SUPPLEMENTARY ONLINE MATERIAL

Supplementary Methods

Plasmids. The Δ*papH pap* operon plasmid (pEMM13) is a pMMB91 based vector expressing a *papH* deleted *pap* operon under the control of the P_{tac} promoter. The operon, which carries an in-frame deletion in the *papH* gene, was cut from pPap207 (Baga et al., 1987) using EcoRI-BamHI and ligated into pMMB91. Plasmid pYML113 is a pACYC184 based vector containing *papC* inserted using EcoRI-SalI restriction sites.

For PapD/PapH and PapD/PapA complex production, the PapD gene was cloned by PCR from *E. coli* UTI89 genomic DNA using oligonucleotides flanked by BamHI and XbaI restriction sites (Supporting Table 1). The PapA and PapH genes were extracted from plasmid pMMB66 using EcoRI and BamHI (Lee et al., 2004). The PapA and PapD genes or the PapH and PapD genes were introduced into a single pTrc99A plasmid by a 3 point ligation (pTRC-DH and pTRC-DA). In order to design the Ntd1 mutants of PapA and PapH (missing residues 2 to 8 for PapA and 2 to 22 for PapH), a PCR of the above pTrc99A plasmid containing either the PapA and PapD genes or the PapH and PapD genes was carried out using phosphorylated primers flanking the region to be deleted (residues 2 to 8 for $PapA_{Ntd1}$ or residues 2 to 22 for $PapH_{Ntd1}$) followed by ligation of the amplified product (Supporting Table 1), producing pTRC-DHNtd1 and pTRC-DANtd1. The $PapD/PapA_{Ntd1G15N}$ mutant was made from the $PapD/PapA_{Ntd1}$ plasmid, using the Strategene QuickChange Mutagenesis Kit (Supporting Table 1).

In order to design a $PapD_{His}/PapA_{Ntd1G15N}$ and $PapD_{His}/PapH_{Ntd1}$ constructs, the above $PapD/PapA_{Ntd1G15N}$ and $PapD/PapH_{Ntd1}$ constructs were modified using 5'-phosphorylated primers containing the coding sequence for 6 histidines plus a short linker region (Ser-Gly; Supporting Table 1).

An HA (influenza A virus hemagglutin) epitope tag was incorporated into PapH using a two-step PCR protocol that has been described previously (Morrison and Desrosiers, 1993). The plasmid pDH-HA was constructed from pTRC-DH by replacing residues Trp14 through Asp22 in PapH with the HA tag sequence, YPYDVPDYA. Primers were used to replace the existing sequence with the HA tag and introduce EcoRI and BamHI restriction enzyme sites at the 5' and 3' end of the gene (PapD/PapH$_{EcoRI/BamHI}$). The *papH* gene was cleaved from pTRC-DH with EcoRI-BamHI; and the PCR product, also digested with EcoRI-BamHI, was ligated into the linearized vector, producing pDH-HA. Proper incorporation of the HA tag was confirmed by sequencing. Supporting Table 1 provides information on oligonucleotide design for all plasmids constructed in this section.

Peptides. The Nte$_{PapA}$ peptide (sequence: Ac-APTIPQGQGKVTFNGTVVD-NH_2) and Nte2$_{PapH}$ peptide (sequence: Ac-GRAAFHGEVVR-NH_2; the entire Nte of PapH was not stable) were purchased from CSS-Albachem (Gladsmuir, UK), resuspended in 50 mM TrisHCl pH 7.5 and 50 mM NaCl, and dialyzed in that same buffer. Quantitative amino acid analysis (University of Cambridge, UK) was performed to obtain accurate extinction coefficients.

Purification of PapD$_{His}$/PapH$_{Ntd1}$, PapD$_{His}$/PapK, PapD$_{His}$/PapE$_{Ntd}$ or PapD$_{His}$/PapA$_{Ntd1G15N}$. The PapD$_{His}$/PapK (Sauer et al., 1999), PapD$_{His}$/PapE$_{Ntd}$ (Sauer et al., 2002), PapD$_{His}$/PapH$_{Ntd1}$, or PapD$_{His}$/PapA$_{Ntd1G15N}$ mutants were purified from periplasmic extracts, using Cobalt- affinity chromatography (Talon; Clontech, Montain View, USA) followed by hydrophobic interaction chromatography (phenyl source (GE Healtcare, Amersham, UK)). Pure His-tagged proteins were then dialysed overnight against a solution containing 50mM TrisHCl pH 7.5 and 50mM NaCl.

Purification of PapD/PapH$_{Ntd1}$. The PapD/PapH$_{Ntd1}$ complex was purified from periplasmic extracts using cation-exchange chromatography (SP HiTrap HP column (GE

Healthcare, Amersham, UK)) followed by hydrophobic interaction chromatography (Phenyl Source column (GE Healthcare, Amershaw, UK)). Purification is then completed by a gel filtration step in 20 mM MES pH6.0 and 50 mM NaCl.

PapD/PapH$_{Ntd1}$ crystallization, data collection, structure determination and refinement. PapD/PapH$_{Ntd1}$ complex was crystallized at room temperature in a hanging drop equilibrated against a reservoir solution containing 0.01M Cobalt Chloride, 0.1M MES pH 6.5 and 1.6 to 1.8M Ammonium Sulphate. Crystals belonged to space group $C222_1$, with one complex per asymmetric unit (67% solvent). The cell dimensions are a = 103 Å, b = 149 Å and c = 83 Å. The data was processed to a resolution of 2.3 Å (see Supplementary Table 2). The structure was solved by molecular replacement using PapD/PapE$_{Ntd}$ as a search model (PDB entry code 1N0L) using the programme AMoRe (Navaza and Saludjian, 1997). The first 2.3 Å map enabled Arp/Warp (Perrakis et al., 1999) to rebuild most of the model along to the PapH$_{Ntd1}$ sequence. Then successive cycles of manual rebuilding with O (Jones et al., 1991) and conjugate gradient minimization using CNS (Brunger et al., 1998) were performed. B factors were refined individually. The refinement converged to the final values of R = 20.9% and Rfree = 24.0% (20 to 2.3 Å range; $F/\sigma F \geq 0.0$) with good stereochemistry (see Supplementary Table 2 for additional statistics). The final model includes all PapD main chain atoms except the last two residues and all PapH$_{Ntd1}$ main chain atoms except the first 2 (residues 1 and 23). The final model also misses one side chain of PapD and 13 side chains of PapH$_{Ntd1}$; the electron density was absent for these exposed residues thus they were modelled as alanines. The final model also includes one cobalt atom, coordinating N-terminal extension residue His28 of PapH$_{Ntd1}$, 11 sulphate ions, and 274 water molecules.

Negative-Stain Electron Microscopy. *E. coli* K12 strain MC1061 transformed with pPap5 (plasmid expressing the entire *pap* operon; (Lund et al., 1985)), pEMM13, pEMM13/pTRC-DH, pEMM13/pTRC-DHNtd1, or pEMM13/pDH-HA were grown on tryptic soy agar (TSA) plates with isopropyl-β-D-thiogalactopyranoside (IPTG) and the appropriate antibiotics for 36 hours at 37°C. All plasmids, except pPap5, were induced with 0.1μM IPTG; and the strain containing pEMM13/pTRC-DH was also induced with 1.0μM IPTG to represent an overexpressed *papH* system. Bacterial cells were harvested, washed once in PBS, and resuspended in PBS to an OD_{600} of 1.0. Bacteria were allowed to absorb onto formvar/carbon-coated grids for 1 min. Grids were washed in dH_2O and stained with 1% aqueous uranyl acetate (Ted Pella Inc., Redding CA) for 1 min. Excess liquid was gently wicked off and grids were allowed to air dry. Samples were viewed on a JEOL 1200EX transmission electron microscope (JEOL USA, Peabody, MA) at an accelerating voltage of 80kV.

References

Baga M, Norgren M, and Normark S (1987) Biogenesis of E. coli Pap pili: papH, a minor pilin subunit involved in cell anchoring and length modulation. *Cell* **49**: 241-251.

Brunger AT, Adams PD, Clore GM, DeLano WL, Gros P, Grosse-Kunstleve RW, Jiang JS, Kuszewski J, Nilges M, Pannu NS*, et al.* (1998) Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr D Biol Crystallogr* **54 (Pt 5)**: 905-921.

Jones TA, Zou JY, Cowan SW, and Kjeldgaard (1991) Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr A* **47 (Pt 2)**: 110-119.

Lee YM, DiGiuseppe PA, Silhavy TJ, and Hultgren SJ (2004) P pilus assembly motif necessary for activation of the CpxRA pathway by PapE in Escherichia coli. *J Bacteriol* **186**: 4326-4337.

Lund B, Lindberg FP, Baga M, and Normark S (1985) Globoside-specific adhesins of uropathogenic Escherichia coli are encoded by similar trans-complementable gene clusters. *J Bacteriol* **162**: 1293-1301.

Morrison HG, and Desrosiers RC (1993) A PCR-based strategy for extensive mutagenesis of a target DNA sequence. *Biotechniques* **14**: 454-457.

Navaza J, and Saludjian P (1997) AMoRe: An automated molecular replacement program package. *Macromolecular Crystallography, Pt A* **276**: 581-594.

Perrakis A, Morris R, and Lamzin VS (1999) Automated protein model building combined with iterative structure refinement. *Nat Struct Biol* **6**: 458-463.

Sauer FG, Futterer K, Pinkner JS, Dodson KW, Hultgren SJ, and Waksman G (1999) Structural basis of chaperone function and pilus biogenesis. *Science* **285**: 1058-1061.

Sauer FG, Pinkner JS, Waksman G, and Hultgren SJ (2002) Chaperone priming of pilus subunits facilitates a topological transition that drives fiber formation. *Cell* **111**: 543-551.

SUPPLEMENTARY TABLES

Supplementary Table 1. Oligonucleotides used in this study. For the $PapD_{His}PapH_{Ntd1}$, $PapD_{His}PapA_{Ntd1G15N}$, PapD/$PapH_{Ntd1}$, and PapD/$PapA_{Ntd1}$ constructs, primers 5' ends were phosphorylated.

Protein	Forward primer	Reverse primer
PapD	CGC GGA TCC CGG CTG ATG AAT GCG GTT ATG CC	CTA GTC TAG ACA CAA CCA TAA TAT CCC ACC CGC
PapD/$PapH_{Ntd1}$	GGC AGG GCT GCT TTT CAT GGT GAG G	ACC GGC AAA AAC ACC ATG AAC AAA CAC ACA GC
PapD/$PapA_{Ntd1}$	GGT AAA GTA ACT TTT AAC GGA ACT GTT G	AGC AGC ATT ATT TAC ACC AAA AGA CAC
PapD/$PapA_{Ntd1G15N}$	GGT AAA GTA ACT TTT AAC AAC ACT GTT GTT GAT GCT CC	GGA GCA TCA ACA ACA GTG TTG TTA AAA GTT ACT TTA CC
$PapD_{His}PapH_{Ntd1}$	CAT CAC CAT CAC TAA TGT ACC GCA ATA ACG GTT AAA TGC GGG	GTG ATG CGA TCC TTT CTC TTT TTT CAC AGA GCA ACG GCT GCC
$PapD_{His}PapA_{Ntd1G15N}$	CAT CAC CAT CAC TAA TGT ACC GCA ATA ACG GTT AAA TGC GGG	GTG ATG CGA TCC TTT CTC TTT TTT CAC AGA GCA ACG GCT GCC
PapD/$PapH_{EcoRI/BamHI}$	GCC GGA ATT CGC GGA AAT ATT CCG CTG	AGC GGG ATC CTC ACT CAT AAT CGA CCC GG
PapD/$PapH_{HA}$	GTC CCT TCC TGA ATA CTA TCC GTA TGA TGT GCC GGA TTA TGC GGG CAG GGC TGC TTT TC	GAA AAG CAG CCC TGC CCG CAT AAT CCG GCA CAT CAT ACG GAT AGT ATT CAG GAA GGG AC

Supplementary Table 2. Data Collection and Refinement Statistics

Data collection		
Data set	Native PapD/$PapH_{Ntd1}$	
(SRS Beamline 14.2)		
Resolution	20-2.3 Å	
Total/Unique	258,065/28,676	
Reflections		
Completeness [1](%)	99.9 (99.5)	
R_{sym}(%)[2]	0.071 (0.268)	
$I/\sigma(I)$[3]	8.7 (2.3)	
Refinement		
Resolution (Å)	20-2.3	
Number of reflections[4]	27,212 (1456)	
(20 – 2.3Å)		
Total number of atoms[5]	3141	
R/R_{free} (%)[6]	20.9/24.0	
r.m.s. deviations[7]	Bonds (Å)	
	0.006	
	Angles (°)	
	1.3	
	B-values ($Å^2$)	
	main chain	side chain
	1.2	2.0

[1]Completeness for $I/\sigma(I) > 0$, high resolution shell (2.42-2.3Å) in parentheses.

[2]$R_{sym} = \sum|I-\langle I\rangle|\sum I$, where I=observed intensity and $\langle I\rangle$=average intensity from multiple observations of symmetry related reflections; high resolution shell in parentheses.

[3]$I/\sigma(I)$ in high resolution shell (2.42-2.3Å) in parentheses.

[4]Number of reflections ($F/\sigma(F) > 0.0$) in the "working" set; number of reflections ($F/\sigma(F) > 0.0$) in the "test" set in parentheses (5% of the total reflections).

[5]Includes 274 water molecules, 11 sulfate ions and 1 cobalt ion.

[6]R_{free} was calculated on the basis of 5% of the total number of reflections randomly omitted from the refinement.

[7]Deviations from ideal bond lengths and angles and in B factors of bonded atoms.