Spontaneous Intermolecular Amide Bond Formation between Side Chains for Irreversible Peptide Targeting

Bijan Zakeri and Mark Howarth*

Department of Biochemistry, Oxford University, South Parks Road, Oxford, OX1 3QU, UK.
E-mail: mark.howarth@bioch.ox.ac.uk

Supporting figures

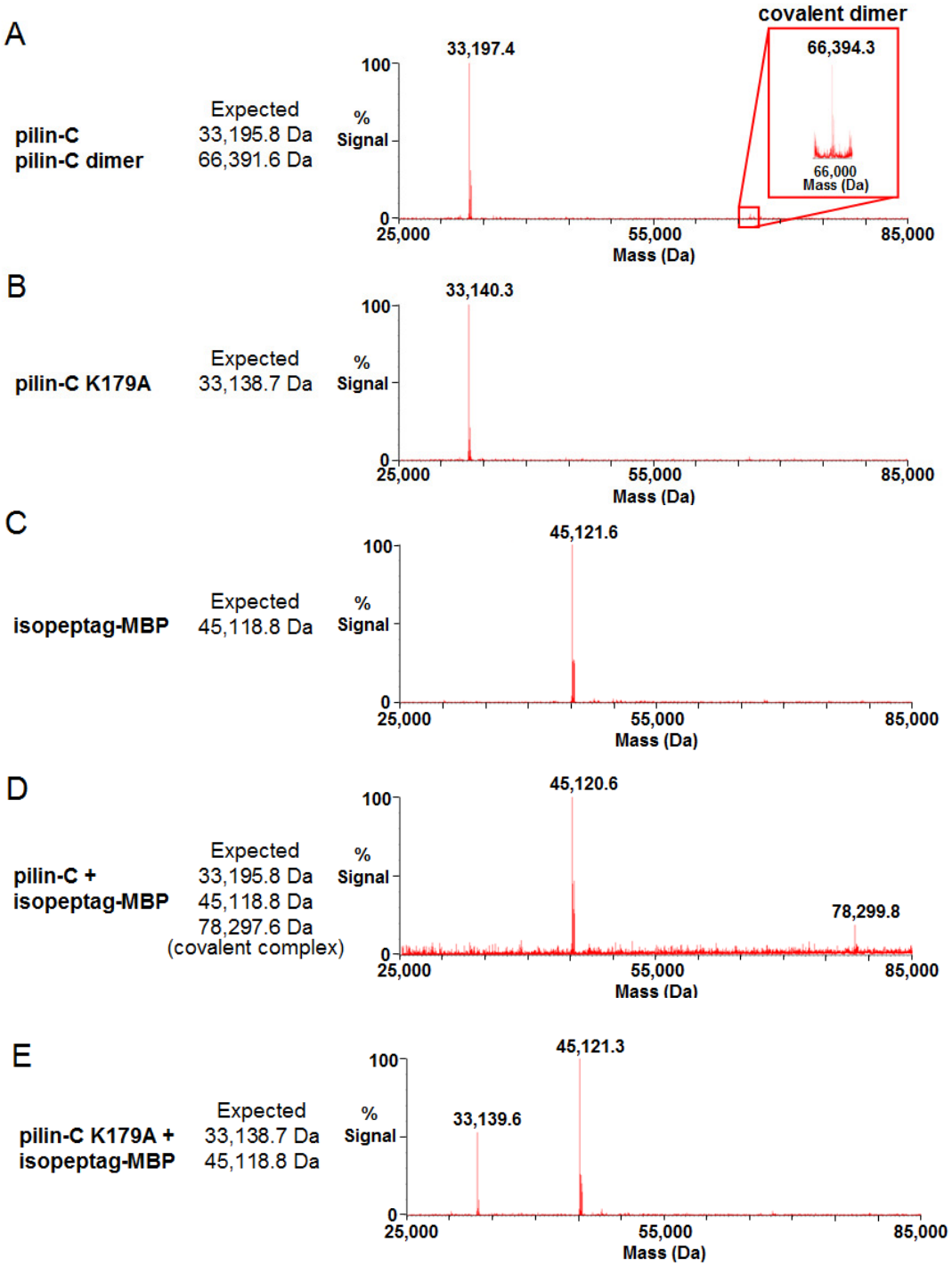

Figure S1. Mass spectrometry of spontaneous intermolecular amide bond formation. (A) Pilin-C alone. Expression of pilin-C also yields a small amount of covalent pilin-C dimer, visible here by MS (inset) and also visible by SDS-PAGE after boiling (Figure 1D). Pilin-C contains one isopeptide in the N-domain. Based on its mass (loss of 17 Da for each isopeptide), the pilin-C dimer contains two isopeptides, one intramolecular and one intermolecular. (B) Pilin-C K179A alone, the negative control for intermolecular amide bond formation. (C) Isopeptag-MBP alone. (D) Pilin-C and isopeptag-MBP were mixed at 100 μM for 16 h at 4 ºC in PBS. The covalent complex was observed, corresponding to loss of NH_3 (17 Da) upon amide bond formation, as well as some residual isopeptag-MBP. (E) A negative control, where pilin-C K179A and isopeptag-MBP were mixed as in (D) but no high molecular weight complex was detected.

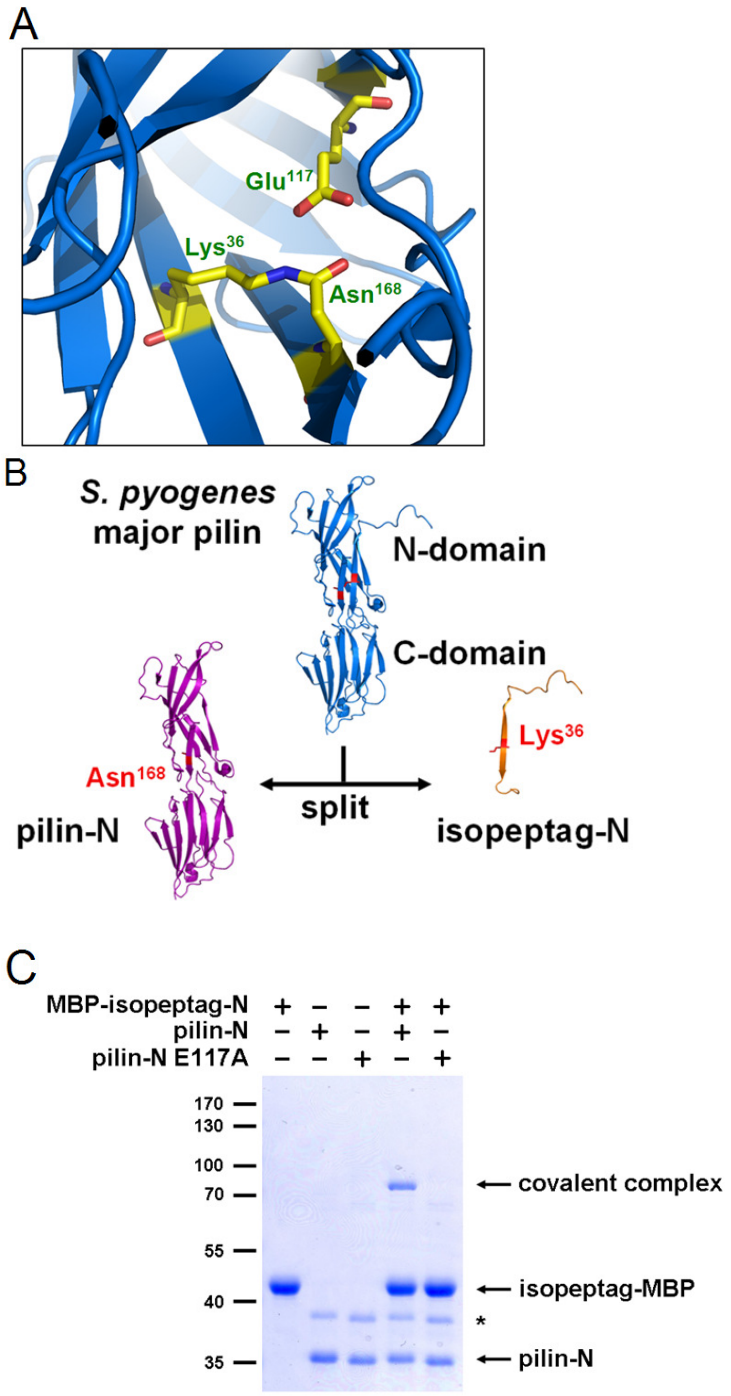

Figure S2. An alternative spontaneous intermolecular amide bond-forming peptide. (A) Key residues for spontaneous amide bond formation in the N-domain of the major pilin protein (Spy0128) of *S. pyogenes*, from Protein Data Bank 3B2M. (B) Isopeptag-N construction. Spy0128 was dissected into a small N-terminal fragment (isopeptag-N in orange) and a large C-terminal fragment (pilin-N in purple). Amide bond forming residues are highlighted in red. (C) Isopeptag-N and pilin-N were able to covalently react *in vitro*. MBP-isopeptag-N and pilin-N, each at 10 μM, were mixed and incubated for 24 h at 25 ˚C, pH 7.0. Pilin-N E117A is a negative control, where the Glu essential for promoting covalent reaction was mutated.[1] Reaction was analyzed by SDS-PAGE and Coomassie staining. *A side-product of pilin-N expression, likely from misformation of the intramolecular isopeptide.
Sequence of isopeptag-N: ATTVHGETVVNGAKLTVTKNLDLVNSNA.

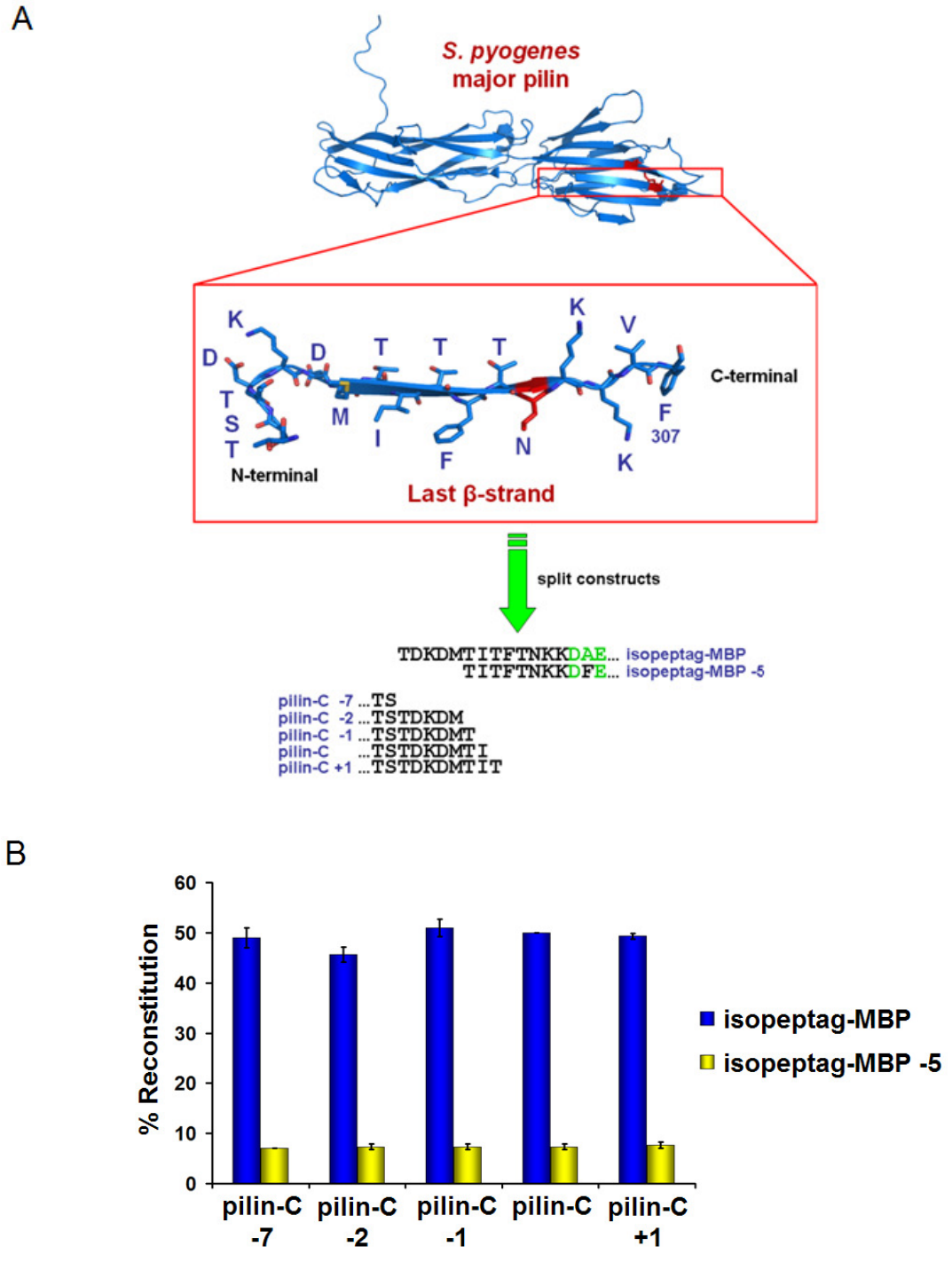

Figure S3. Optimization of split pilin constructs to generate isopeptag and pilin-C. (A) Schematic illustrating the various isopeptag-MBP and pilin-C constructs tested. The reactive Asn is shown in red. The Spy0128 crystal structure (Protein Data Bank 3B2M) stops at residue F307 and has Val at 306 instead of the Asp in *S. pyogenes* M1 GAS (NCBI Entrez Nucleotide accession no. AE004092). Our isopeptag construct had Asp at residue 306 and we also changed the surface-exposed Phe at position 307 to Ala, to increase solubility. Residues different to the crystal structure are shown in green. The numbers depict the change in the number of residues compared to the pilin-C and isopeptag-MBP constructs finally chosen and used in all other figures. Including in the protein partner residues from the final β-strand was evaluated on the basis that these residues could inhibit the other two neighboring β-strands interacting with each other and so occluding the isopeptag binding site. (B) Reconstitution efficiency of the isopeptag-MBP and pilin-C constructs, each at 10 μM, after incubation at 25 ºC for 24 h in PBS pH 7.4, determined by SDS-PAGE and Coomassie staining. The means of triplicate results are shown ± 1 s.d.

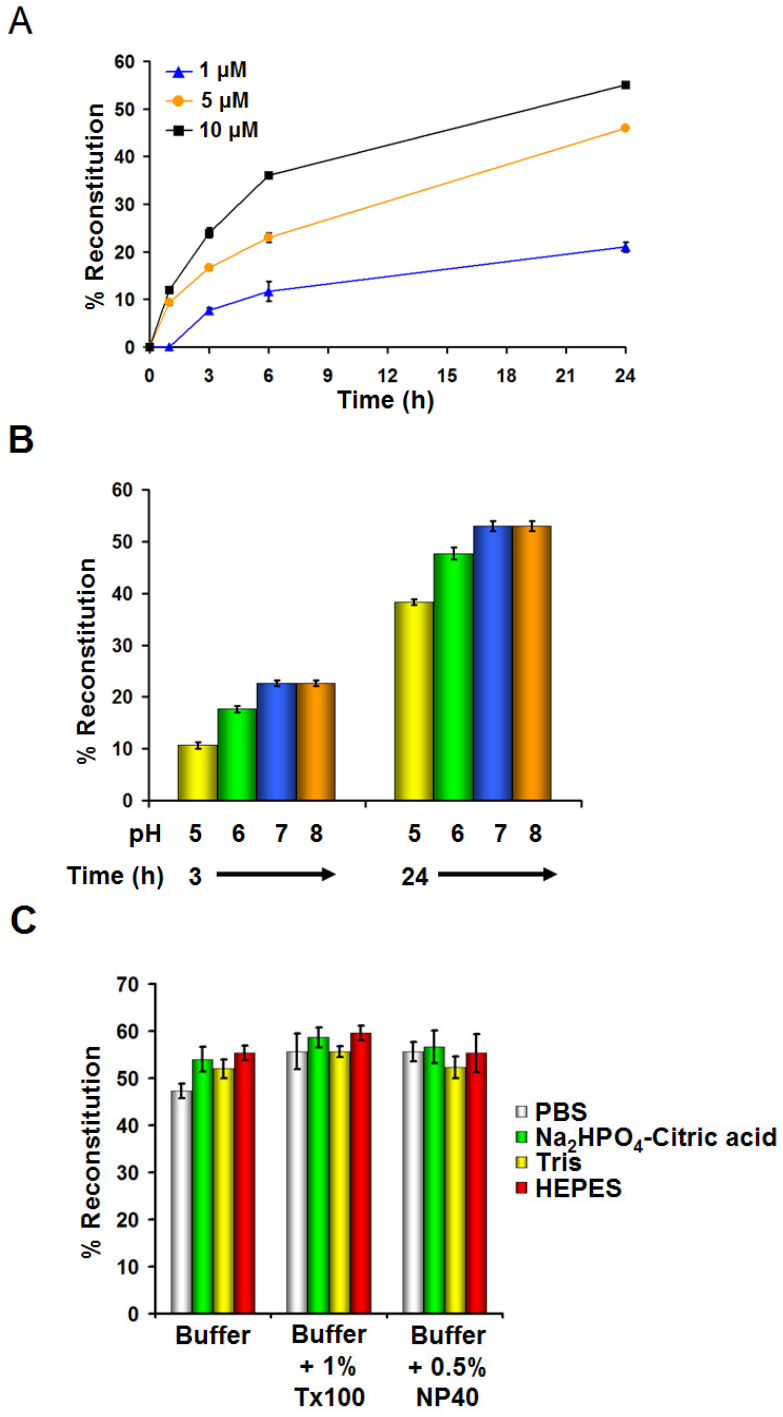

Figure S4. Factors affecting spontaneous amide bond formation. (A) Concentration-dependence of spontaneous amide bond formation. Isopeptag-MBP and pilin-C were mixed each at 1, 5, or 10 µM and incubated at 25 ºC, pH 7.0. The reaction was stopped at varying times and analyzed by SDS-PAGE with Coomassie staining. The means of triplicate results are shown ± 1 s.d. (some errors bars are too small to be visible). (B) pH-dependence of spontaneous amide bond formation. Isopeptag-MBP and pilin-C were mixed at 10 µM in 40 mM Na_2HPO_4 - 20 mM citric acid at the indicated pH and incubated at 25 ºC. The reaction was stopped at varying times and analyzed by SDS-PAGE with Coomassie staining. The means of triplicate results are shown ± 1 s.d. (C) Buffer-dependence of spontaneous amide bond formation. Isopeptag-MBP and pilin-C each at 10 µM were mixed in the indicated buffer with or without the common biological detergents Triton X-100 (Tx100) or Nonidet P-40 (NP40) and incubated for 24 h at 25 ºC, pH 7.0. Detergent might have been expected to change the rate of reconstitution, if amide bond formation depended on exposure of a hydrophobic part of pilin-C. PBS is phosphate buffered saline. Reaction was analyzed by SDS-PAGE with Coomassie staining. The means of triplicate results are shown ± 1 s.d.

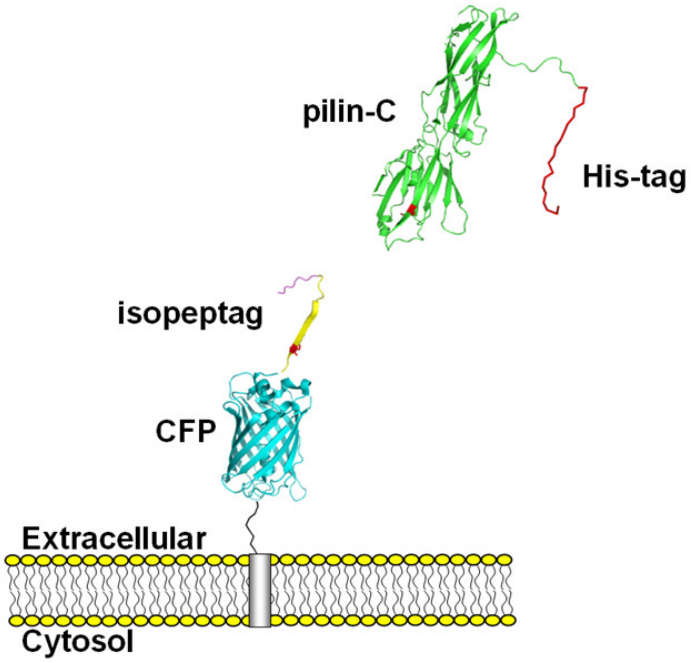

Figure S5. Cartoon of the isopeptag-CFP-TM construct, used to target the isopeptag to the surface of mammalian cells where it could be detected by exogenous pilin-C. CFP = cyan fluorescent protein. TM = transmembrane helix.

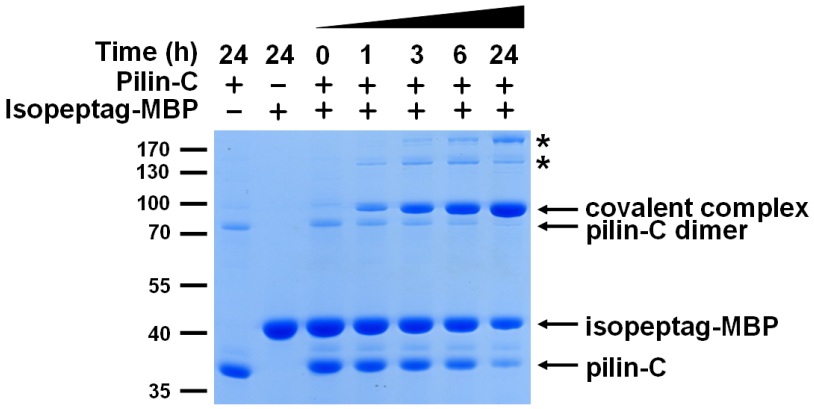

Figure S6. Time-course of spontaneous amide bond formation by SDS-PAGE. Isopeptag-MBP and pilin-C were mixed at 10 μM and incubated at 25 ºC, pH 7.0. The reaction was stopped at varying times and analyzed by SDS-PAGE with Coomassie staining. *This side-product likely relates to the pilin-C dimer reacting with 1 or 2 isopeptag-MBP.

Supporting methods

Plasmids

All residue numbers are based on the Spy0128 protein.[1] PCR was performed with KOD Hot Start DNA polymerase (Roche) and cycling conditions as recommended by the manufacturer. The *Spy0128* gene was PCR-amplified from *Streptococcus pyogenes* M1 GAS strain SF370 genomic DNA (ATCC 700294D-5) using the primers 5'–GGGGCATATGGCTACAACAGTTCACGGG–3' and 5'–GGGGAAGCTTTTATTCAAAGTCTTTTTTATTTG–3', and subsequently cloned in to pET28a (Novagen) using NdeI and HindIII. To construct isopeptag-MBP, isopeptag (Spy0128 residues 293-308) was PCR-amplified from pET28a-Spy0128 using the primers 5'–GGGGCATATGGGAACTGATAAAGATATGACC–3' and 5'–ACCACTTTCACCACTACCTTCAAAGTCTTTTTTATTTG–3', while the maltose binding protein (MBP) was amplified from pMAL (New England Biolabs) using the primers 5'–AAGGTAGTGGTGAAAGTGGTAAAATCGAAGAAGGTAAA–3' and 5'–GGGGAAGCTTTTACGAGCTCGAATTAGTCTG–3'. The two fragments were joined by overlap-extension PCR, and cloned in to pET28a using *NdeI* and *HindIII* sites. This gave a GSGESG linker between isopeptag and MBP. The final isopeptag-MBP construct was generated by a further mutation of F307A (based on the complete *Spy0128* sequence), to remove an exposed hydrophobic residue,[1] using the QuikChange™ (Stratagene) protocol with KOD Hot Start DNA polymerase and the primers 5'– TTTACAAATAAAAAAGACGCTGAAGGTAGTGGTGAAAG–3' and its reverse complement. To generate isopeptag-MBP -5 (Spy0128 residues 298-308), the sequence was PCR-amplified from pET28a-isopeptag-MBP using primers 5'–GGGGCATATGGGAACCATTACTTTTACAAAT–3' and 5'–GGGGAAGCTTTTACGAGCTCGAATTAGTCTG–3' and cloned in to pET28a using *NdeI* and *HindIII* sites. To generate pilin-C (Spy0128 residues 18-299), a stop codon was introduced at residue 300 in pET28a-Spy0128 using the primer 5'–ACAAGAGACATCTACTGATAAAGATATGACCATTTAGTTTACAAATAAAAAAGACTTTGAATAAAAGCTTG–3' and its reverse complement. To generate pilin-C K179A from pET28a-Spy0128, we used the primer 5'–TCTACTACATTAACGGTGAAGGCAAAAGTTTCAGGTACCGGTGG–3' and its reverse complement, and then a stop codon was introduced at residue 300 in pET28a-Spy0128 using the primer 5'–ACAAGAGACATCTACTGATAAAGATATGACCATTTAGTTTACAAATAAAAAAGACTTTGAATAAAAGCTTG–3' and its reverse complement. To generate pilin-C -7 (Spy0128 residues 18-292), a stop codon was introduced at residue 293 in pET28a-Spy0128 using the primer 5'–GCAGGTAATTCAACTGAACAAGAGACATCTTAGGATAAAGATATGACCATTACTTTTACAAAT–3' and its reverse complement. To generate pilin-C -2 (Spy0128 residues 18-297), a stop codon was introduced at residue 298 in pET28a-Spy0128 using the primer 5'–ACTGAACAAGAGACATCTACTGATAAAGATATGTAGATTACTTTTACAAATAAAAAAGACTTTGAAGTG–3' and its reverse complement. To generate pilin-C -1 (Spy0128 residues 18-298), a stop codon was introduced at residue 299 in pET28a-Spy0128 using the primer 5'–GAACAAGAGACATCTACTGATAAAGATATGACCTAGACTTTTACAAATAA

AAAAGACTTTGAATAAAAG–3' and its reverse complement. To generate pilin-C +1 (Spy0128 residues 18-300), a stop codon was introduced at residue 301 in pET28a-Spy0128 using the primer 5'–CAAGAGACATCTACTGATAAAGATATGACCATTACTTAGACAAATAAAAAAGACTTTGAATAAAA–3' and its reverse complement.

To construct MBP-isopeptag-N (Figure S2B), isopeptag-N (Spy0128 residues 18-45) was PCR-amplified from pET28a-Spy0128 using the primers 5'–CTAATTCGAGCTCGGGTTCGGGTGAAAGTGGTGCTACAACAGTTCACGGG–3' and 5'–GGGGAAGCTTTTATGCATTGCTATTAACTAAATC–3', while MBP was amplified from pMAL using primers 5'–CAAGCATATGAAAATCGAAGAAG–3' and 5'–CGAACCCGAGCTCGAATTAGTCTG–3'. The two fragments were joined using overlap-extension PCR, and cloned in to pET28a at *NdeI* and *HindIII* sites. Pilin-N (Spy0128 residues 46-308) was generated from pET28a-Spy0128 using the primers 5'–GGGGCATATGGGATTAATTCCAAATACAGAT–3' and 5'–GGGGAAGCTTCTAGTGATGGTGATGGTGATGTCCTGATCCTTCAAAGTCTTTTTTATTTG–3', and subsequently cloned in to pET28a using *NdeI* and *HindIII* sites. Pilin-N E117A was generated from pET28a-pilin-N using the primer 5'–GTGTTTATTATTACAAAGTAACTGCGGAGAAGATAGATAAAGTTCCTGG–3' and its reverse complement. Isopeptag-CFP-TM was based on pDisplay (Invitrogen). 5'–GACAGATCTGGCGGCACTGATAAAGATATGACCATTACTTTTACAAATAAAAAAGACTTTGAAGGTAGTGGTATGGTGAGCAAGGGCGAG–3' and 5'–ACTCTCGGCATGGACGAGCTATACAAGCGGCCGCGGGGAG–3' were used for PCR from AP-CFP-TM.[2] The PCR product was gel purified and inserted in the *BglII* and *SacII* sites of pDisplay.

To construct a bicistronic vector for simultaneous expression of pilin-C or pilin-C K179A and isopeptag-MBP (Figure 3A), the pilin-C and pilin-C K179A constructs were PCR-amplified from pET28a-pilin-C or pET28a-pilin-C K179A, using primers 5'– GGGGAAGCTTCTCTAGAAATAATTTTGTTTAAC–3' and 5'–GGGGCTCGAGCTAAATGGTCATATCTTTATC–3'. The PCR products were gel purified and inserted in to the *HindIII* and *XhoI* sites of pET28a-isopeptag-MBP, to give pET28a-(isopeptag-MBP + pilin C) and pET28a-(isopeptag-MBP + pilin C K179A). From this plasmid, from a single T7 promoter, an mRNA would be transcribed containing first a sequence that coded for the isopeptag-MBP protein, then a spacer following the isopeptag-MBP stop codon, then a Shine-Dalgarno sequence to promote translation, and finally a sequence coding for the pilin C or pilin-C K179A protein.

We verified all constructs and mutations by sequencing. The nuclear co-transfection marker pECFP-H2B (human histone H2B for nuclear localization fused to enhanced cyan fluorescent protein) was constructed as described.[3]

Protein expression

All proteins were expressed using *E. coli* BL21 DE3 RIPL cells (Stratagene), grown in LB with 0.8 % glucose and 0.05 mg/mL kanamycin. We diluted overnight cultures 100-fold, grew at 37 ºC to A_{600} 0.5, and induced with 0.4 mM IPTG for 4 h at 30 °C. All proteins were purified by nickel affinity chromatography, using standard methods, and dialyzed into PBS. Protein concentration was determined from A_{280}, using the extinction coefficient predicted by ExPASy ProtParam. Typical expression yields per L of culture were 10 mg for pilin-C and 16 mg for isopeptag-MBP.

SDS-PAGE

SDS-PAGE was performed on 10% polyacrylamide gels, using an X-cel SureLock (Invitrogen) at 200 V. Samples were heated at 95 °C for 7 min in SDS loading buffer on a Bio-Rad C1000 thermal cycler before loading. Gels were stained with Coomassie brilliant blue, destained in 60% MilliQ water, 30% methanol and 10% acetic acid, and band intensities were quantified using a ChemiDoc XRS imager and QuantityOne 4.6 software (Bio-Rad).

Reconstitution reactions

Reactions for analyzing speed, temperature-dependence, and pH-dependence of amide bond formation by SDS-PAGE were performed with 10 μM of each protein in 40 mM Na_2HPO_4 - 20 mM citric acid at the indicated pH. For Figure 2C, one reactant was maintained at 10 μM while the concentration of the other reactant was varied. For analyzing amide bond formation in various buffers and detergents (Figure S4C), 10 μM of pilin-C and 10 μM of isopeptag-MBP were mixed in either phosphate buffered saline (PBS) pH 7.4, 40 mM Na_2HPO_4 - 20 mM citric acid pH 7.4, 50 mM Tris (tris-hydroxymethyl aminomethane) pH 7.4, or 50 mM HEPES (4-(2-hydroxyethyl)-1-piperazine ethanesulfonate) pH 7.4, and also in the presence and absence of 1% Triton X-100 or 0.5% Nonidet P-40. pH for all buffers was adjusted with 1 M NaOH or 1 M HCl, except for HEPES which was adjusted with 1 M KOH to establish whether Na^+ was required for reaction. All reactions were incubated for 24 h unless otherwise stated. Reactions were stopped by adding SDS loading buffer and heating at 95 ºC for 7 min. For 4 ºC reactions, samples were incubated in the refrigerator; for 25 ºC and 37 ºC reactions, samples were incubated on a Bio-Rad C1000 thermal cycler with a heated lid to prevent evaporation. The percentage reconstitution between pilin-C and isopeptag-MBP was calculated by dividing the density of the band for the pilin-C:isopeptag-MBP covalent complex by the sum of the density of all the bands in the lane, then multiplying by 100. To calculate the percentage of pilin-C reacted (Figure 2C), the percentage reduction in intensity of the pilin-C band was determined relative to the controls where no reaction partner was added. The percentage of isopeptag-MBP that reacted was calculated by the same approach.

For testing *in situ* reconstitution within living cells (Figure 3A), *E. coli* expressing monocistronic or bicistronic plasmids were induced for 4 h at 30 ºC, as above. Samples of bacterial cultures were mixed with SDS loading buffer and immediately heated at 95 °C for 7 min to prevent further reconstitution, before analysis by SDS-PAGE.

Mass spectrometry (Figure 2A and S1)

We performed mass spectrometry with a Micromass LCT time-of-flight electrospray ionization mass spectrometer (Micromass UK). 100 μM of each protein was mixed in PBS pH 7.4 and incubated at 4 ºC for 16 h. The reactions were then dialyzed against MilliQ H_2O using 0.025 μm VSWP Millipore membrane filters (VSWP 01300) for 1 h at 25 ºC. The reactions were then further de-salted using Millipore ZipTip pipette tips (ZTC04S096). The m/z spectrum was converted to a molecular mass profile using Maximum Entropy processing by the software

MassLynx V4.00.00. Predicted masses were determined from the amino acid sequence by ExPASy ProtParam, subtracting 17 Da for every isopeptide bond.

Cell culture, labeling and microscopy (Figure 3B)

HeLa cells were grown in DMEM with 10% Fetal Calf Serum, 50 U/mL penicillin, and 50 μg/mL streptomycin. Cells were transfected using 0.75 μL Lipofectamine 2000 (Invitrogen) and 0.3 μg of isopeptag-CFP-TM per well of a 48-well plate, according to manufacturer's instructions. One day after transfection, cells were incubated with 5 μM of either pilin-C or pilin-C K179A and incubated at 37 ºC for 4 h. Cells were then stained at 4 ºC as follows: wash 3 times with PBS + 5mM $MgCl_2$ (PBS-Mg), add 100 μL of 11 μg/mL penta-His antibody (Qiagen) in PBS-Mg with 1% BSA and incubate for 10 min, wash 3 times with PBS-Mg, add 100 μL of 20 μg/mL Alexa Fluor 555-anti-mouse antibody (Invitrogen) in PBS-Mg with 1% BSA and incubate for 10 min, wash 3 times with PBS-Mg. Cells were fixed with 100 μL fixing solution (PBS containing 4% formaldehyde, 4% sucrose, and 5 mM $MgCl_2$) for 15 min at 4 ºC and then washed 3 times with PBS-Mg. Cells were imaged using a wide-field DeltaVision Core fluorescent microscope (AppliedPrecision) with a 40× oil-immersion lens. ECFP (436DF20 excitation, 480DF40 emission, Chroma 86002v1 dichroic) and Alexa Fluor 555 (540D420 excitation, 600DF50 emission, Chroma 84100bs polychroic) images were collected and analyzed using softWoRx 3.6.2 software. Typical exposure times were 0.1-1.0 s and fluorescence images were background-corrected. Different samples in the same experiment were prepared, imaged and analyzed under identical conditions.

References

(1) Kang, H. J.; Coulibaly, F.; Clow, F.; Proft, T.; Baker, E. N. *Science* **2007**, *318*, 1625.

(2) Chen, I.; Howarth, M.; Lin, W.; Ting, A. Y. *Nat. Methods* **2005**, *2*, 99.

(3) Platani, M.; Goldberg, I.; Lamond, A. I.; Swedlow, J. R. *Nat. Cell Biol.* **2002**, *4*, 502.