Reconstitution of a minimal machinery capable of assembling type IV pili

Vivianne J. Goosens[a,1], Andreas Busch[b,1], Michaella Georgiadou[a], Marta Castagnini[a], Katrina T. Forest[c], Gabriel Waksman[b] and Vladimir Pelicic[a,2]

[a]MRC Centre for Molecular Bacteriology and Infection, Imperial College London, London, United Kingdom

[b]Institute of Structural and Molecular Biology, University College London and Birkbeck, London, United Kingdom

[c]Department of Bacteriology, University of Wisconsin-Madison, Madison, USA

[1]These authors contributed equally

[2]To whom correspondence should be addressed. Email: v.pelicic@imperial.ac.uk

BIOLOGICAL SCIENCES (Microbiology)

Abstract

Type IV pili (Tfp), which are key virulence factors in many bacterial pathogens, define a large group of multi-purpose filamentous nanomachines, widespread in Bacteria and Archaea. Tfp biogenesis is a complex multi-step process, which relies on macromolecular assemblies composed of 15 conserved proteins in model Gram-negative species. To improve our limited understanding of the molecular mechanisms of filament assembly, we have used a synthetic biology approach to reconstitute, in a non-native heterologous host, a minimal machinery capable of building Tfp. Here, we show that eight synthetic genes are sufficient to promote Tfp assembly and that the corresponding proteins form a macromolecular complex at the cytoplasmic membrane, which we have purified and characterised biochemically. Our results contribute to a better mechanistic understanding of the assembly of remarkable dynamic filaments, nearly ubiquitous in prokaryotes.

\body

Introduction

Evolution has provided prokaryotes with sophisticated surface nanomachines that endow them with many functions instrumental to their ability to colonise most niches on Earth. Among these engineering marvels, type IV filamentous (Tff) nanomachines (1), of which Tfp are the paradigm, are unique for two reasons. They are exceptionally (i) widespread, with genes encoding distinctive proteins found in virtually every prokaryotic genome, and (ii) multi-purpose, associated with functions as diverse as adhesion, motility, protein secretion, DNA uptake, electric conductance *etc.* (1). Much of this broad distribution and multi-functionality is due to Tfp (1).

All Tff nanomachines share multiple components and are thought to use common basic operating principles. They have at their core a filament, which can be long or short (depending on the studied Tff) and is a polymeric assembly of a protein named pilin, PilE in our model Tfp-expressing species *Neisseria meningitidis* (the meningococcal nomenclature will be used here). Type IV pilins are produced as prepilins with a distinctive N-terminal class III signal peptide (2), consisting of a short hydrophilic leader peptide followed by a stretch of 21 hydrophobic residues, always forming an extended α-helix (3). This signal peptide is first recognised by the Sec machinery (4, 5), which translocates prepilins across the cytoplasmic membrane where they remain embedded as bitopic proteins. The leader peptide is then cleaved by an integral membrane aspartic protease (6, 7), the prepilin peptidase PilD. This processing, which does not require other Pil proteins (8), is a pre-requisite for polymerisation of pilins into filaments. Filaments are helical polymers in which the pilins' extended N-terminal α-helices are buried within the filament core, almost parallel to its long axis (9). Finally, in Gram-negative Tfp-expressing bacteria, filaments cross the outer membrane through a pore formed by the secretin PilQ (10).

The molecular mechanisms of filament assembly remain poorly understood. However, there is consensus that assembly occurs at the cytoplasmic membrane

and requires energy, which is generated by PilF, a cytoplasmic ATPase (11-13). This energy is transmitted via an ill-defined membrane-embedded assembly sub-complex to the processed pilins, which are thereby extruded from the lipid bilayer and polymerised into filaments. Filament assembly has been best studied in Gram-negative species expressing retractable Tfp, where piliation relies on 15 highly conserved proteins (1) (PilC, PilD, PilE, PilF, PilG, PilH, PilI, PilJ, PilK, PilM, PilN, PilO, PilP, PilQ and PilW). Genetic studies have shown that seven of these proteins are not involved in filament assembly *per se* since piliation can be restored in the corresponding mutants by a second mutation in *pilT*, which encodes an ATPase powering pilus retraction/disassembly (14). As confirmed in different species, these seven proteins are the outer membrane component PilC (15, 16), the four pilin-like proteins (PilH, PilI, PilJ and PilK) (16-18), the secretin PilQ (16, 19), and the secretin-associated lipoprotein PilW (20, 21). Interestingly, in the *pilQpilT* double mutant, filaments remain trapped in the periplasm (16, 19), showing that filament assembly can be genetically uncoupled from their emergence on the cell surface. As a corollary, when piliation was not restored in a double mutant, this was viewed as indirect evidence that the corresponding Pil protein might be involved in filament assembly. While different studies agree that PilD, PilE and PilF fall in this class (16, 19), conflicting results have been obtained for PilG, PilM, PilN, PilO and PilP. In *N. meningitidis*, PilM, PilN, PilO and PilP were deemed essential for filament assembly while the integral membrane protein (PilG) was not (16), while in *P. aeruginosa* it was the opposite scenario (22). As a result, the exact role of these five proteins is unclear but there is ample evidence that they establish multiple binary/ternary interactions at the cytoplasmic membrane (23-33). Moreover, in a recent study in *Myxococcus xanthus*, in which the entire Tfp machinery was visualised by cryo-electron tomography (34), it was shown that these five proteins form a series of interconnected layers spanning the cytoplasmic membrane, which is *a priori* compatible with a role in filament assembly.

Although the above mutational studies defining Pil components essential for Tfp assembly have provided a useful blueprint for subsequent experiments, they are inherently limited by their negative readout (absence of piliation in a *pilT* mutant background) and the contrasting findings in two closely related systems (*N. meningitidis* and *P. aeruginosa*). Here, we have directly defined the proteins required for Tfp assembly by using a previously unexplored synthetic biology approach. By identifying the minimal set of Pil proteins capable of assembling Tfp in a heterologous host in which they are not natively produced, and characterising biochemically the macromolecular complexes these proteins form, we provide novel insights into a fundamental but poorly understood phenomenon.

Results

Engineering large synthetic operons encoding proteins involved in Tfp assembly

Reconstituting a minimal machinery capable of assembling Tfp is challenging because of (i) the large number of genes required, and (ii) the fact that these genes are scattered over many genomic loci. To overcome these challenges, *pil* genes from the sequenced *N. meningitidis* 8013 strain (35), codon-optimised for expression in *E. coli*, were synthesised for each meningococcal protein potentially involved in Tfp assembly (PilD, PilE, PilF, PilG, PilM, PilN, PilO and PilP). To engineer large artificial operons with these synthetic genes, we used an iterative cloning approach (36). Genes were combined into operons of increasing size, where each gene was preceded by a ribosome-binding site (RBS) and the expression of the entire operon was driven by a T7 promoter (Fig. S1). First, to test experimentally the two contrasting models for Tfp assembly, we engineered *pilDFGE* and *pilDFMNOPE* operons (abbreviated as DFGE or DFMNOP) (Fig. 1A). However, since toxicity and plasmid instability were observed in a variety of BL21-based expression strains, we sub-cloned these operons into pBAD18, under a tighter arabinose-inducible promoter (37). These pBAD18-derived plasmids were stable and did not significantly affect bacterial growth. All the Pil components included in these operons were expressed, as tested by immunoblotting using specific antibodies (Fig. 1B). Since we have been unable to generate a good anti-PilD antibody, we confirmed the presence of a functional prepilin peptidase by showing that PilE was processed only when *pilD* was present. In the absence of PilD (first lane), PilE has a slightly larger molecular weight than in bacteria where PilD is present (lanes 2-7) (Fig. 1B). Our attempts to construct operons encoding all of the above eight Pil components were thwarted by plasmid instability. We therefore used an alternative cloning strategy to create pBAD18 derivatives expressing these eight genes from two different promoters (Fig. 1A). The

DFGE and DFMNOPE operons were sub-cloned in pBAD18 under a constitutive σ70 promoter, which was mapped using 5' RACE (Fig. S2). The resulting plasmids were then used to sub-clone the remaining *pil* gene(s) under the arabinose-inducible promoter, yielding MNOP[DFGE] and G[DFMNOPE] constructs (genes within brackets are those whose expression is driven by σ70). The final plasmids were stable, did not significantly affect bacterial growth and led to the expression of all the Pil components as tested by immunoblotting (Fig. 1B).

Pil proteins form membrane-embedded macromolecular assemblies, which can be purified to homogeneity

In order to promote Tfp assembly, the Pil proteins expressed in *E. coli* must interact to form a macromolecular complex at the cytoplasmic membrane. Therefore, to test complex formation/stability and unravel protein-protein interactions between PilF, PilG, PilM, PilN, PilO and PilP components, we added a *Strep*-tag to either PilO or PilP (indicated as P_{Strep} or O_{Strep}) and purified under native conditions the complexes formed by various protein combinations. Notably, when the *pilMNOP*$_{Strep}$ operon was expressed, we could purify a native PilMNOP$_{Strep}$ complex solubilised in *n*-Dodecyl β-D-maltoside (β-DDM) using a combination of affinity and size exclusion (SEC) chromatographies. As shown in Fig. 2, PilMNOP$_{Strep}$, which is stable throughout the purification process, eluted as a single, symmetric peak during SEC. The purified complex consisted of the four PilM, PilN, PilO and PilP$_{Strep}$ components as assessed by Coomassie after SDS-PAGE (Fig. 2) and immunoblotting (Fig. S3). A MALDI-TOF MS analysis of the purified complex in solution confirmed that the four Pil components were intact (Fig. S4). Using SEC coupled with in-line multi-angle light scattering (SEC-MALS), PilMNOP$_{Strep}$ was found to be a homogeneous and monodisperse sample with an estimated molecular weight of 132.7 ± 1.6 kDa (Table S1). This value is reasonably close to the theoretical mass for a hetero-tetramer (108 kDa) suggesting that the purified PilMNOP$_{Strep}$ complex consists of one copy of each

protein (1:1:1:1 stoichiometry). Next, we tested whether these four components are all necessary for complex formation and/or stability by generating alternative constructs with *pilMNO*$_{Strep}$ and *pilNOP*$_{Strep}$ operons. While PilNOP$_{Strep}$ could be purified as a stable and homogeneous complex (Fig. 3A), indicating that PilM is dispensable for complex formation/stability, the PilMNO$_{Strep}$ complex eluted in several peaks and the proteins in the different fractions tended to aggregate after purification (Fig. S5), indicating that PilP is important for PilMNOP$_{Strep}$ stability. Strikingly, a *Strep*-tagged 9.8 kDa truncated version of PilP (named PilP$_{NT\text{-}Strep}$), consisting only of the N-terminal 77 residues previously shown to interact with PilNO (29), was sufficient to restore stability and homogeneity to the PilMNOP complex (Fig. 3B). Using SEC-MALS, PilNOP$_{Strep}$ was found to be a homogeneous and monodisperse sample (Table S1). Intriguingly, the estimated molecular weight of PilNOP$_{Strep}$, 214.9 ± 9.2 kDa, was much larger than the theoretical mass of a hetero-trimer (66.7 kDa) and even larger than the molecular weight measured for PilMNOP$_{Strep}$. This finding, which is consistent with the lower retention volumes for PilNOP$_{Strep}$ when compared to PilMNOP$_{Strep}$, suggests that PilNOP$_{Strep}$ adopts a different 3:3:3 stoichiometry in the absence of PilM.

We next assessed whether the other two putative filament assembly components (PilF and PilG) form a complex with PilMNOP. While co-expression of the six genes failed to yield amounts of proteins sufficient for analysis, *pilF*$_{His}$*MNOP*$_{Strep}$ and *pilG*$_{His}$*MNOP*$_{Strep}$, where PilF and PilG have C-terminally fused His-tags, produced complexes consisting of all five Pil components. Both PilF$_{His}$MNOP$_{Strep}$ (Fig. 4A) and PilG$_{His}$MNOP$_{Strep}$ (Fig. 4B) are stable and homogeneous complexes, which eluted during SEC as single, symmetric peaks. The presence of all complex components was confirmed by Coomassie and verified by immunoblotting after SDS-PAGE (Fig. S6). In order to determine whether PilF and PilG could interact with the stable PilNOP complex, we co-expressed *pilF*$_{His}$ or *pilG*$_{His}$ with *pilNOP*$_{Strep}$. We found that PilF$_{His}$ could not be pulled-down with the PilNOP$_{Strep}$

complex (Fig. S7A), suggesting that PilM is the likely interaction partner of PilF. Similarly, $PilG_{His}$ did not form a stable complex with $PilNOP_{Strep}$. The protein complexes that were pulled-down eluted in several peaks in which no $PilG_{His}$ was detectable by Coomassie (Fig. S7B), although the protein (probably minute amounts) could be detected by immunoblotting in the higher molecular weight peaks. These findings show that PilM is important for the stability of PilG within the PilGMNOP complex.

Taken together, these results show that the Pil components predicted to play a role in Tfp assembly form stable membrane macromolecular complexes, which could be purified in native fashion for the first time.

Addressing the PilG "paradox" in *N. meningitidis*

It has been reported in *N. meningitidis* 8013 that piliation is restored in a *pilGpilT* mutant (16), which is in stark contrast with subsequent results in *P. aeruginosa* (22) and the central position of PilG in subtomograms of the *M. xanthus* Tfp machinery (34). The biochemical evidence above that PilG forms a complex with PilMNOP prompted us to revisit PilG's role in *N. meningitidis*. First, because the anti-PilG antibody was not available during our original study, we determined whether the *pilG* transposon insertion (Tn) mutant that was used then (called *pilG1* hereafter), which has a *mariner* mini-transposon inserted close to the beginning of this gene (Fig. 5A), disrupts protein production. As could be seen in Fig. 5B, this is indeed the case since no PilG could be detected by immunoblotting in whole cell protein extracts from the *pilG1* mutant. Then, as assessed by immunofluorescence (IF) microscopy, using the 20D9 monoclonal antibody that is highly sensitive and specific for strain 8013 Tfp (38), we confirmed earlier findings (16) that the *pilG1* mutant is non-piliated and that piliation is restored in a *pilG1pilT* double mutant, which is heavily piliated (Fig. 5C). To determine whether different *pilG* mutations would yield similar results, or not, we constructed two additional double mutants using either a *pilG2* Tn mutant with an

insertion closer to the middle of the gene, or a *ΔpilG* mutant in which the gene was cleanly deleted (27) (Fig. 5A). Strikingly, while both these mutations abolish PilG production (Fig. 5B), the *pilG2pilT* and *ΔpilGΔpilT* mutants behaved differently from *pilG1pilT* since they were non-piliated (Fig. 5C). Not a single filament could be detected by IF microscopy in the *pilG2pilT* and *ΔpilGΔpilT* mutants. Together, these results show that although the non-piliated phenotype in a *N. meningitidis pilG1* mutant can indeed be suppressed by a second mutation in *pilT*, this is dependent on the nature of this *pilG* mutation. Therefore, since no filaments are restored when *pilG* is cleanly deleted, PilG is likely to be involved in pilus assembly in *N. meningitidis*.

Eight proteins are sufficient to assemble Tfp

Using our monoclonal 20D9 anti-Tfp antibody, which specifically and efficiently recognises filaments from strain 8013 (38), we assessed by IF microscopy whether the expression of any of the above *pil* operons would promote filament assembly in *E. coli*. An important caveat is that since no *pilQ* was included in our constructs, potential Tfp were expected to be trapped in the periplasm, much like the filaments in a meningococcal double *pilQpilT* mutant (16). Bacteria were therefore submitted to an osmotic shock treatment prior to IF microscopy as previously done for *N. meningitidis pilQpilT* (16). We first tested the two models for Tfp assembly pre-existing to this study, PilDEFG vs. PilDEFMNOP (16, 22). No filaments were detected when the [DFGE] or [DFMNOPE] operons were expressed in *E. coli* (Fig. 6A and 6B). This shows that (i) none of these two operons promotes Tfp assembly, and (ii) filament assembly does not occur spontaneously, confirming previous findings that the anti-Tfp monoclonal antibody shows specificity for assembled Tfp (38). Instead, with both of the above gene combinations, we saw green spots/foci localised on the bacterial cells (Fig. 6A and 6B). Since we showed above that all the corresponding proteins are expressed and form membrane-embedded macromolecular complexes, the absence of piliation suggests that none of the

PilDEFG and PilDEFMNOP subsets of proteins is sufficient to promote Tfp assembly. Therefore, since we found in this study that PilG is essential for filament assembly in *N. meningitidis* and that it interacts with PilMNOP to form a PilGMNOP complex, we tested whether *E. coli* strains transformed with the MNOP[DFGE] (Fig. 6C and 6E) and G[DFMNOPE] (Fig. 6D and 6F) constructs would be capable of producing Tfp. Strikingly, µm long filaments (a length similar to native meningococcal Tfp) were readily and reproducibly detected. Filaments were seen with both operons, in which the genes are in different orders and expressed from different promoters, confirming that the eight proteins are sufficient to promote filament assembly. Filaments were not detected when the bacteria were not osmotically shocked, confirming that they were originally trapped in the periplasm (Fig. 6G and 6H). Taken together, these results show that eight proteins (PilD, PilE, PilF, PilG, PilM, PilN, PilO and PilP) are sufficient to promote Tfp assembly, indicating that these proteins form the minimal machinery capable of polymerising PilE into filaments.

Discussion

Tff nanomachines are nearly ubiquitous in prokaryotes and have been studied for decades. However, our understanding of the molecular mechanism(s) leading to the assembly of filaments composed of type IV pilins remains limited. This is in part due to the complexity of the protein machinery involved with as many as 15 highly conserved proteins involved in Gram-negative model species (1). In addition, the integral membrane nature of these protein complexes has hindered biochemical and structural studies. Here, we have used a synthetic biology approach to reconstitute in *E. coli* a minimal machinery capable of assembling Tfp, which we characterised in depth biochemically. This led to the notable findings discussed below.

Our new genetic evidence that the "platform" protein PilG is involved in filament assembly is important as it solves the *Neisseria* PilG paradox. This finding is now consistent with PilG's (i) presence in all Tff nanomachines (1), (ii) central position in the Tfp machinery of *M. xanthus* (34) and (iii) role in Tfp assembly in *P. aeruginosa* (22). Nevertheless, the piliated phenotype of the *pilG1pilT* mutant is intriguing and unique to the *pilG1* mutation, where a Tn is inserted early in the gene after the first 100 bp. Although speculative, the most likely scenario is that a truncated PilG protein is still produced in this mutant. Although this putative truncated protein, which we could not detect by immunoblotting, is unable to promote piliation in an otherwise WT genetic background, it might be partially active and capable of promoting filament assembly in a *pilT* mutant background. The N-terminus of PilG is therefore likely to be dispensable for filament assembly, and its role could be to promote piliation by controlling PilT-mediated pilus retraction. This scenario is consistent with the observation that the N-terminus is the least conserved portion in PilG orthologs.

A key finding in this study is that a minimal machinery capable of assembling Tfp can be reconstituted in *E. coli* by co-expressing only eight of the 15 highly conserved Pil proteins in species expressing retractable Tfp. Our results indicate that the seven Pil proteins acting after pilus assembly (PilC, PilH, PilI, PilJ, PilK, PilQ and

PilW) are dispensable *en bloc* for filament assembly. Since PilD, PilE and PilF roles are known, our results suggest that five components (PilG, PilM, PilN, PilO and PilP) form the macromolecular complex involved in filament assembly. Therefore, both pre-existing models for Tfp assembly (16, 22) were partially correct since both PilG and PilMNOP are required. However, in the presence of PilFDGE or PilFDMNOPE only, foci but not filaments were detected, which suggests an abortive filament assembly process. The PilMNOP sub-complex is therefore not merely an "alignment sub-complex" responsible for aligning the filament assembly machinery with the secretin pore in the outer membrane (29). Instead, it is an integral and key part of the assembly machinery itself. This is supported by the presence of PilM and PilN orthologs in Tfp-expressing Gram-positive species which lack an outer membrane (39). Nevertheless, in Gram-negative species, PilMNOP also plays an aligning role via the interaction of PilP with the secretin (29, 40). Our findings illustrate the following series of events leading to filament assembly (Fig. 7). PilE subunits are first processed by the prepilin peptidase PilD (8), and accumulate in the cytoplasmic membrane. Importantly, processing in this study was only seen in the presence of PilD, indicating that there was no interference of the prepilin peptidase activity previously reported in a lab strain of *E. coli*, which was due to the product of the cryptic *pppA* gene (41). Mature PilE is then "loaded" on the membrane-embedded assembly sub-complex, composed of PilGMNOP, which is powered by the filament extension motor PilF. The role of PilGMNOP would thus be to translate the mechanical energy generated by motion of domains within PilF (13, 42) to pilins, which are simultaneously extruded from the membrane and polymerised into the base of a growing filament (Fig. 7). The wide conservation of the above components (1) suggests that this scenario for filament assembly is broadly applicable.

The other major finding in this study is that native membrane-embedded macromolecular complexes of Pil proteins can be purified, which provides a clearer topological picture of the Tfp assembly machinery (Fig. 7). Purification of PilMNOP as

a homogeneous species shows that these four proteins can form a stable sub-complex in the absence of other Pil proteins, with a probable 1:1:1:1 stoichiometry. The findings that PilM, which is the only cytoplasmic component, is dispensable for complex assembly/stability (PilNOP is very stable), while PilP is essential (PilMNO is unstable) are in agreement with previous genetic studies characterising binary interactions between these proteins (23, 27, 30, 31). The finding that only a small N-terminal portion of PilP, which was shown to interact with PilNO (29), is sufficient for PilNOP stability suggests that PilP plays an indirect role in filament assembly by stabilising the PilNO hetero-dimer (24, 26). Curiously, unlike PilMNOP, PilNOP adopts a likely 3:3:3 stoichiometry, which suggests that it is a highly dynamic macromolecular assembly. PilM, which interacts with the cytoplasmic N-terminal portion of PilN (25, 27, 29) and forms a cytoplasmic ring in Tfp subtomograms (34), is thus a peripheral component of the PilMNOP complex, probably recruited to the cytoplasmic membrane once PilNOP is pre-assembled. Purification of stable PilGMNOP and PilFMNOP complexes confirms that the extension ATPase and platform protein are integral components of the assembly machinery (Fig. 7). PilG is likely to become a stable part of the machinery once PilM has been recruited to the PilNOP complex (PilG does not form a stable complex with PilNOP), which is consistent with Tfp subtomograms showing that the PilG "dome" structure requires the presence of the PilM ring (34). Similarly, the ATPase PilF is likely to be recruited to the complex via interactions with PilM (PilF does not co-purify with PilNOP) (33, 43) and/or PilG (32), which is consistent with Tfp subtomograms where the PilF "disc" requires both the PilM ring and PilG dome (34). Finally, PilE from a pool of processed subunits awaiting in the membrane would diffuse to the PilFGMNOP complex (28), which would scoop them out of the lipid bilayer and into the base of a growing filament (Fig. 7).

In conclusion, we provide here the first integrated molecular view of the functioning of the machinery involved in the assembly of a filamentous polymer

composed of type IV pilins. This provides a layout for the understanding of current and past findings in the field and paves the way for structural analysis of the macromolecular complex involved, which suggests that an atomic level understanding of Tfp assembly is achievable.

Materials and methods

Bacterial strains and plasmids

The *N. meningitidis* strains used in this study are all derived from the sequenced serogroup C clinical isolate 8013 (35). *N. meningitidis* was grown on GCB agar plates (Difco) containing Kellogg's supplements and, when required, 100 µg/ml kanamycin and 3 µg/ml erythromycin. Plates were incubated in a moist atmosphere containing 5% CO_2. The *pil* mutants used were described in earlier studies (21), or constructed by splicing PCR as described elsewhere (27) using the primers listed in Table S2.

E. coli DH5α was used for cloning experiments in pET-29b (Novagen), while cloning in pBAD18 was performed in *E. coli* TOP10 (Invitrogen). Cells were grown in liquid or solid Lysogenic Broth (LB) (Difco), or LBG (LB supplemented with 1% glucose) either at 37°C or at 30°C. When appropriate, the following antibiotics were used: ampicillin 100 µg/ml, chloramphenicol 34 µg/ml and kanamycin 50 µg/ml. Chemically competent cells were prepared using standard molecular biology techniques. For filament detection and immunoblot analyses, TOP10 cells transformed with pBAD18-borne *pil* operons were grown overnight in LBG at 30ºC in the presence of the relevant antibiotic. Bacteria were re-inoculated (1/100) in Terrific Broth supplemented with 1% glucose and without antibiotics, and grown at 30ºC to late exponential phase. Gene expression was then induced with 0.5% arabinose for one hour before bacteria were placed on ice. Aliquots of each sample were taken for immunoblots and/or IF.

The *pil* operons were constructed as follows using synthetic genes codon-optimised for expression in *E. coli* generated by GeneArt. The plasmids used in this study are listed in Table S3. Genes *pilD*, *pilE*, *pilF* and *pilG* were synthesised separately, while *pilM*, *pilN*, *pilO* and *pilP* were synthesised as an operon where the last three genes were preceded by canonical RBS. Each synthetic gene/group of genes was preceded by a unique *Nde*I site (CATATG, in which the ATG is the start

368 codon of the gene) and followed by consecutive and unique *Nhe*I and *Xho*I sites
369 (right after the stop codon of the last gene). To construct the various operons, each
370 gene/group of genes extracted as an *Nde*I and *Xho*I fragment was sub-cloned in
371 pET-29b cut with the same enzymes. Then, genes were combined into operons of
372 increasing size using an iterative cloning approach (36). In brief, gene B is extracted
373 from the pET-derived plasmid together with its RBS on a *Xba*I-*Xho*I fragment, and
374 cloned into *Nhe*I-*Xho*I immediately downstream of gene A (*Xba*I and *Nhe*I generate
375 compatible cohesive ends). This effectively creates an artficial AB operon whose
376 expression is driven by the T7 promoter. Since the *Nhe*I site downstream gene A is
377 destroyed during this cloning step, the strategy can be employed iteratively to create
378 operons of increasing size. Using this methodology, multiple variations of pET29-
379 based operons were generated, numbering up to seven *pil* genes. Toxicity/plasmid
380 instability with pET29-based plasmids prompted us to sub-clone the above operons
381 into the arabinose-inducible pBAD18 (32). Genes or group of genes were extracted
382 from pET-29 derivatives on a *Xba*I-*Xma*I fragment and sub-cloned in pBAD18 cut
383 with *Nhe*I and *Xma*I. This effectively placed these genes/operons under the control of
384 the arabinose-inducible promoter in pBAD18. Since it was impossible to combine all
385 eight *pil* genes in a single operon, we used an alternative cloning strategy to co-
386 express the corresponding proteins in *E. coli*. We noticed, serendipitously, that *pil*
387 operons sub-cloned in pBAD18 in the reverse direction of the arabinose-inducible
388 promoter were efficiently expressed from an endogenous σ70 promoter, which we
389 have mapped (Fig. S2). We therefore cloned the *pilFDGE* and *pilFDMNOPE* operons
390 under the control of that σ70 promoter by extracting them from pET-29 derivatives on
391 a *Xba*I-*Xma*I fragment and sub-cloning in pBAD18 cut with the same enzymes.
392 Finally, the missing assembly *pil* gene(s) were sub-cloned into these plasmids, which
393 were cut by *Nhe*I, as *Xba*I-*Nhe*I fragments extracted from pET-29-derivatives. This
394 placed them under the arabinose-inducible promoter, in the reverse direction of the
395 *pil* genes under the σ70 promoter.

E. coli BL21 Star (DE3) (ThermoFisher Scientific) was used for heterologous expression of full-length native complexes of Pil proteins. Transformed cells were grown in 3-12 l of LB at 37ºC, under 180 rpm agitation, until OD_{660} reached 0.6 and then cooled down to 16ºC for 30 min, before inducing overnight with 200 µg/l of anhydrotetracycline (ATc) (iba). Cells where harvested the next day by centrifuging at 5,000 *g* at 4ºC for 30 min, and resuspended in buffer (20 mM Tris-HCl pH 7.5, 100 mM NaCl, 1 mM EDTA) before being flash-frozen in liquid nitrogen and stored at -80ºC.

Plasmids for co-expression and purification of full-length native Pil proteins complexes were constructed by sub-cloning the above *pil* operons into pASK-IBA3C (iba) vector, which puts them under an ATc-inducible promoter and fuses a *Strep*-tag to the C-terminus of the last protein encoded. Initially, the *pilMNOP* operon was cloned into the two *Bsa*I sites in pASK-IBA3C. Addition of *pil* genes with a His-Tag, as well as deletion or truncation of *pil* genes was performed using the pASK-*pilMNOP*$_{Strep}$ construct with the In-Fusion cloning kit (Clonech).

Purification of native membrane-embedded macromolecular complexes

Deep-frozen cell pellets were thawed on ice and incubated for 30 min at 4ºC, under constant agitation (100 rpm), upon addition of lysozyme (1 mg/ml), 1 mM $MgCl_2$ and 5 U of Benzonase (EMD Millipore). Cell lysis was performed with the EmulsiFlex-C5 homogeniser (Avestin), used at 750-1,000 psi. Cell debris was pelleted by centrifugation at 34,000 *g* for 30 min at 4ºC, before the whole membrane fraction was pelleted at 112,000 *g* for 90 min at 4ºC. Cell membranes were resuspended in 20 mM Tris-HCl, pH 7.5 buffer, containing 1 mM EDTA and 100-250 mM NaCl (depending on the Pil macromolecular complex studied). Membrane proteins were solubilised by adding 1% (w/v) of β-DDM (Anatrace) detergent to this suspension and stirring at 100 rpm for 1 h at 4ºC. Remaining cell debris was pelleted by centrifugation for 20 min at 112,000 *g* at 4ºC. The solubilised membrane protein

extract was then loaded onto a 5 ml StrepTrap HP affinity column (GE Healthcare) and the Pil multi-protein complexes pulled-down using the C-terminal Strep-tag on one of the proteins. In some cases, a second pull-down purification was done using an additional C-terminal His-Tag on a different protein. Finally, complexes were purified by SEC using HiLoad Superose 6 XK 16/70 PG or 16/600 Superdex 200 PG columns (both from GE Healthcare), depending on the molecular weight of the complex. When the protein yields obtained after the initial affinity purification step were lower, analytical Superose 6 Increase 10/300 GL or Superdex 200 Increase 10/300 GL columns (both from GE Healthcare) were used instead. Throughout the entire purification, performed at 4ºC, identical buffer conditions were used (20 mM Tris-HCl pH 7.5, 100-250 mM NaCl, 1 mM EDTA, 0.05% β-DDM).

MALDI-TOF and SEC-MALS analysis of macromolecular complexes

Absolute molar masses of purified $PilMOP_{Strep}$ and $PilNOP_{Strep}$ complexes were determined by SEC-MALS as follows. Protein samples (100 µl at 1 mg/ml, in 20 mM Tris-HCl pH 7.5, 150 mM NaCl, 1 mM EDTA, 0.05% β-DDM) were loaded onto a Superose 6 Increase 10/300 GL column at 0.5 ml/min using an Agilent 1100 series HPLC (Agilent). The column output was fed into a DAWN HELEOS II MALS detector (Wyatt Technology) followed by an Optilab T-rEX differential refractometer (Wyatt Technology), which measures absolute and differential refractive indexes. Data were collected and analysed using the Astra 6.1.2 software (Wyatt Technology). Molecular masses were calculated across eluted protein peaks through extrapolation from Zimm plots using dn/dc values of 0.185 ml/g for the protein fraction and 0.1435 ml/g for β-DDM. Quoted molecular weights of protein or β-DDM and estimated errors relate to the overall mass calculation across a single peak. Three repeat runs were performed for both $PilNOP_{Strep}$ and $PilMNOP_{Strep}$ complexes under identical experimental conditions.

MALDI-TOF analysis of the PilMOP$_{Strep}$ complex in solution was performed at the Max Planck Institute of Biophysics.

SDS-PAGE and immunoblotting

Whole-cell protein extracts were prepared as previously described for *N. meningitidis* (44), or by resuspending *E. coli* cells directly in Laemmli sample buffer (Bio-Rad) and heating 5-10 min at 95ºC. When needed, proteins were quantified using the Bio-Rad Protein Assay as suggested by the manufacturer. Separation of the proteins by SDS-PAGE and subsequent blotting to Amersham Hybond ECL membranes (GE Healthcare) was carried out using standard molecular biology techniques. Blocking overnight (in PBS with 0.5% milk), incubation with primary and/or secondary antibodies (60 min each) and detection using Amersham ECL Plus (GE Healthcare) were carried out following the manufacturer's instructions. Primary antisera were used at 1/100,000 (anti-PilP) or 1/10,000 (anti-PilE, anti-PilF, anti-PilG, anti-PilM, anti-PilN and anti-PilO). Amersham ECL HRP-linked secondary antibody (GE Healthcare) was used at a 1/10,000 dilution. Blots were imaged with a Bio-Rad Chemidoc Touch imaging system.

Tfp immuno-detection

Bacteria were spotted and dried into the wells of a microscope glass slide, fixed with 2.5% paraformaldehyde (in PBS) for 20 min, and quenched with 0.1 M glycine (in PBS) for 5 min. After blocking with 5 % milk (in PBS) for 30 min, the monoclonal anti-Tfp 20D9 mouse antibody (1/2,000 in blocking solution) was added and incubated for 30 min. After washing slides with PBS, cells were stained with DAPI (ThermoFisher Scientific) while Tfp were labelled with a goat anti-mouse coupled to Alexa Fluor 488 (ThermoFisher Scientific), both added at 1/1,000 dilution (in PBS). After 30 min incubation and a PBS wash, a coverslip was mounted using Aqua Poly/Mount (Polysciences). After overnight incubation at 4ºC, samples were viewed and

photographed using an Axio Imager A2 microscope (Zeiss). When indicated, cells were submitted to a cold osmotic shock treatment prior spotting on slides. Four ml of cultures were centrifuged at 1,200 *g* for 10 min at 4ºC, resuspended in 300 µl osmotic buffer (0.1 M Tris acetate pH 8.2, 0.5 M sucrose and 5 mM EDTA). Lysozyme at 0.1 mg/ml was then added, and samples were left on ice for 5 min. Cells were osmotically shocked and filaments released by adding 18 mM $MgSO_4$.

RNA extraction and 5' RACE

RNA was extracted from *E. coli* cultures grown in LBG to late exponential phase. RNA was extracted using a PureLink RNA mini kit (Ambion Life Technologies) and stabilised with RNAprotect cell reagent (Qiagen). Transcription start site mapping was done using the 5’ RACE system for rapid amplification of cDNA ends (Invitrogen), according to the manufacturer instructions.

Acknowledgements

This work was funded by MRC grants to VP (MR/L008408/1) and GW (MR/K018434/1), and an NIH grant (RO1GM59721) to KTF. AB was supported by a long-term fellowship from the European Molecular Biology Organization and a Marie Curie IEF. We acknowledge the support of employees and the use of experimental resources of Instruct, a Landmark ESFRI project. We are grateful to Sophie Helaine (Imperial College London), Christoph Tang (University of Oxford) and Romé Voulhoux (CNRS, Marseille) for critical reading of this manuscript.

References

1. Berry JL & Pelicic V (2015) Exceptionally widespread nano-machines composed of type IV pilins: the prokaryotic Swiss Army knives. *FEMS Microbiol. Rev.* 39:134-154.

2. Szabó Z*, et al.* (2007) Identification of diverse archaeal proteins with class III signal peptides cleaved by distinct archaeal prepilin peptidases. *J. Bacteriol.* 189:772-778.

3. Giltner CL, Nguyen Y, & Burrows LL (2012) Type IV pilin proteins: versatile molecular modules. *Microbiol. Mol. Biol. Rev.* 76:740-772.

4. Francetic O, Buddelmeijer N, Lewenza S, Kumamoto CA, & Pugsley AP (2007) Signal recognition particle-dependent inner membrane targeting of the PulG pseudopilin component of a type II secretion system. *J. Bacteriol.* 189:1783-1793.

5. Arts J, van Boxtel R, Filloux A, Tommassen J, & Koster M (2007) Export of the pseudopilin XcpT of the *Pseudomonas aeruginosa* type II secretion system via the signal recognition particle-Sec pathway. *J. Bacteriol.* 189:2069-2076.

6. LaPointe CF & Taylor RK (2000) The type 4 prepilin peptidases comprise a novel family of aspartic acid proteases. *J. Biol. Chem.* 275:1502-1510.

7. Hu J, Xue Y, Lee S, & Ha Y (2011) The crystal structure of GXGD membrane protease FlaK. *Nature* 475:528-531.

8. Aly KA*, et al.* (2013) Cell-free production of integral membrane aspartic acid proteases reveals zinc-dependent methyltransferase activity of the *Pseudomonas aeruginosa* prepilin peptidase PilD. *MicrobiologyOpen* 2:94-104.

9. Kolappan S*, et al.* (2016) Structure of the *Neisseria meningitidis* type IV pilus. *Nature Communications* 7:13015.

10. Korotkov KV, Gonen T, & Hol WG (2011) Secretins: dynamic channels for protein transport across membranes. *Trends Biochem. Sci.* 36:433-443.

526 11. Yamagata A & Tainer JA (2007) Hexameric structures of the archaeal
527 secretion ATPase GspE and implications for a universal secretion mechanism.
528 *EMBO J.* 26:878-890.
529 12. Lu C, Korotkov KV, & Hol WG (2014) Crystal structure of the full-Length
530 ATPase GspE from the *Vibrio vulnificus* type II secretion system in complex with the
531 cytoplasmic domain of GspL. *J. Struct. Biol.* 187:223-235.
532 13. Reindl S*, et al.* (2013) Insights into FlaI functions in archaeal motor assembly
533 and motility from structures, conformations, and genetics. *Mol. Cell* 49:1069-1082.
534 14. Merz AJ, So M, & Sheetz MP (2000) Pilus retraction powers bacterial
535 twitching motility. *Nature* 407:98-102.
536 15. Wolfgang M, Park HS, Hayes SF, van Putten JP, & Koomey M (1998)
537 Suppression of an absolute defect in type IV pilus biogenesis by loss-of-function
538 mutations in *pilT*, a twitching motility gene in *Neisseria gonorrhoeae*. *Proc. Natl.*
539 *Acad. Sci. USA* 95:14973-14978.
540 16. Carbonnelle E, Helaine S, Nassif X, & Pelicic V (2006) A systematic genetic
541 analysis in *Neisseria meningitidis* defines the Pil proteins required for assembly,
542 functionality, stabilization and export of type IV pili. *Mol. Microbiol.* 61:1510-1522.
543 17. Giltner CL, Habash M, & Burrows LL (2010) *Pseudomonas aeruginosa* minor
544 pilins are incorporated into type IV pili. *J. Mol. Biol.* 398:444-461.
545 18. Winther-Larsen HC*, et al.* (2005) A conserved set of pilin-like molecules
546 controls type IV pilus dynamics and organelle-associated functions in *Neisseria*
547 *gonorrhoeae*. *Mol. Microbiol.* 56:903-917.
548 19. Wolfgang M, van Putten JP, Hayes SF, Dorward D, & Koomey M (2000)
549 Components and dynamics of fiber formation define a ubiquitous biogenesis pathway
550 for bacterial pili. *EMBO J.* 19:6408-6418.
551 20. Koo J*, et al.* (2008) PilF is an outer membrane lipoprotein required for
552 multimerization and localization of the *Pseudomonas aeruginosa* type IV pilus
553 secretin. *J. Bacteriol.* 190:6961-6969.

554 21. Carbonnelle E, Helaine S, Prouvensier L, Nassif X, & Pelicic V (2005) Type IV
555 pilus biogenesis in *Neisseria meningitidis*: PilW is involved in a step occuring after
556 pilus assembly, essential for fiber stability and function. *Mol. Microbiol.* 55:54-64.
557 22. Takhar HK, Kemp K, Kim M, Howell PL, & Burrows LL (2013) The platform
558 protein is essential for type IV pilus biogenesis. *J. Biol. Chem.* 288:9721-9728.
559 23. Ayers M*, et al.* (2009) PilM/N/O/P proteins form an inner membrane complex
560 that affects the stability of the *Pseudomonas aeruginosa* type IV pilus secretin. *J.*
561 *Mol. Biol.* 394:128-142.
562 24. Sampaleanu LM*, et al.* (2009) Periplasmic domains of *Pseudomonas*
563 *aeruginosa* PilN and PilO form a stable heterodimeric complex. *J. Mol. Biol.* 394:143-
564 159.
565 25. Karuppiah V & Derrick JP (2011) Structure of the PilM-PilN inner membrane
566 type IV pilus biogenesis complex from *Thermus thermophilus*. *J. Biol. Chem.*
567 286:24434-24442.
568 26. Tammam S*, et al.* (2011) Characterization of the PilN, PilO and PilP type IVa
569 pilus subcomplex. *Mol. Microbiol.* 82:1496-1514.
570 27. Georgiadou M, Castagnini M, Karimova G, Ladant D, & Pelicic V (2012)
571 Large-scale study of the interactions between proteins involved in type IV pilus
572 biology in *Neisseria meningitidis*: characterization of a subcomplex involved in pilus
573 assembly. *Mol. Microbiol.* 84:857-873.
574 28. Karuppiah V, Collins RF, Thistlethwaite A, Gao Y, & Derrick JP (2013)
575 Structure and assembly of an inner membrane platform for initiation of type IV pilus
576 biogenesis. *Proc. Natl. Acad. Sci. USA* 110:E4638-4647.
577 29. Tammam S*, et al.* (2013) PilMNOPQ from the *Pseudomonas aeruginosa* type
578 IV pilus system form a transenvelope protein interaction network that interacts with
579 PilA. *J. Bacteriol.* 195:2126-2135.

30. Li C, Wallace RA, Black WP, Li YZ, & Yang Z (2013) Type IV pilus proteins form an integrated structure extending from the cytoplasm to the outer membrane. *PLoS One* 8:e70144.

31. Friedrich C, Bulyha I, & Sogaard-Andersen L (2014) Outside-in assembly pathway of the type IV pilus system in *Myxococcus xanthus*. *J. Bacteriol.* 196:378-390.

32. Bischof LF, Friedrich C, Harms A, Sogaard-Andersen L, & van der Does C (2016) The type IV pilus assembly ATPase PilB of *Myxococcus xanthus* interacts with the inner membrane platform protein PilC and the nucleotide-binding protein PilM. *J. Biol. Chem.* 291:6946-6957.

33. McCallum M*, et al.* (2016) PilN binding modulates the structure and binding partners of the *Pseudomonas aeruginosa* type IVa pilus protein PilM. *J. Biol. Chem.* 291:11003-11015.

34. Chang YW*, et al.* (2016) Architecture of the type IVa pilus machine. *Science* 351:aad2001.

35. Rusniok C*, et al.* (2009) NeMeSys: a resource for narrowing the gap between sequence and function in the human pathogen *Neisseria meningitidis*. *Genome Biol.* 10:R110.

36. McLaughlin LS (2013) Heterologous expression of the type IV pilus assembly complex. PhD (University of Wisconsin-Madison).

37. Guzman LM, Belin D, Carson MJ, & Beckwith J (1995) Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter. *J. Bacteriol.* 177:4121-4130.

38. Pujol C, Eugène E, de Saint Martin L, & Nassif X (1997) Interaction of *Neisseria meningitidis* with a polarized monolayer of epithelial cells. *Infect. Immun.* 65:4836-4842.

39. Gurung I*, et al.* (2016) Functional analysis of an unusual type IV pilus in the Gram-positive *Streptococcus sanguinis*. *Mol. Microbiol.* 99:380-392.

608 40. Balasingham SV*, et al.* (2007) Interactions between the lipoprotein PilP and
609 the secretin PilQ in *Neisseria meningitidis*. *J. Bacteriol.* 189:5716-5727.
610 41. Francetic O, Lory S, & Pugsley AP (1998) A second prepilin peptidase gene
611 in *Escherichia coli* K-12. *Mol. Microbiol.* 27:763-775.
612 42. Mancl JM, Black WP, Robinson H, Yang Z, & Schubot FD (2016) Crystal
613 structure of a type IV pilus assembly ATPase: insights into the molecular mechanism
614 of PilB from *Thermus thermophilus*. *Structure*.
615 43. Sandkvist M, Bagdasarian M, Howard SP, & DiRita VJ (1995) Interaction
616 between the autokinase EpsE and EpsL in the cytoplasmic membrane is required for
617 extracellular secretion in *Vibrio cholerae*. *EMBO J.* 14:1664-1673.
618 44. Helaine S*, et al.* (2005) PilX, a pilus-associated protein essential for bacterial
619 aggregation, is a key to pilus-facilitated attachment of *Neisseria meningitidis* to
620 human cells. *Mol. Microbiol.* 55:65-77.

Legends to Figures

Fig. 1. Engineering large operons composed of synthetic meningococcal genes optimised for expression in *E. coli*, which encode proteins involved in Tfp assembly. (**A**) Gene organisation of *pil* operons generated in pBAD18 in this study. Expression of genes in black is driven by an arabinose-inducible promoter (Ara). Expression of genes in white, indicated within brackets, is driven by a constitutive σ70 promoter (σ70). All the genes are drawn to scale, with the scale bar representing 1 kb. (**B**) Immunoblot analysis of the production of Pil proteins from various constructs. Whole-cell protein extracts of *E. coli* TOP 10 transformed with the various constructs (indicated above each lane) were successively probed using specific anti-Pil antibodies (indicated on the right of each immunoblot). Since no antibody is available against PilD, we confirmed the presence of a functional prepilin peptidase by showing that the pilin detected (top immunoblot) in a strain expressing only PilE (lane 1) has a slightly larger molecular weight than the pilin detected in the bacteria where PilD is present (lanes 2-7). Bacterial cultures were equalised based on OD_{600} readings and equivalent amounts of cells were loaded in each lane.

Fig. 2. PilMNOP proteins form a stable membrane complex when expressed in *E. coli*, which can be purified to homogeneity. SEC profile of $PilMNOP_{Strep}$ on a Superose 6 XK 16/70 column. The inset shows the SDS-PAGE/Coomassie analysis of the purified complex. A molecular weight marker was run in the first lane. Molecular weights are indicated in kDa. The molecular weights for the individual proteins are: PilM (41.4 kDa), PilN (22.2 kDa), PilO (23.3 kDa) and $PilP_{Strep}$ (21.3 kDa).

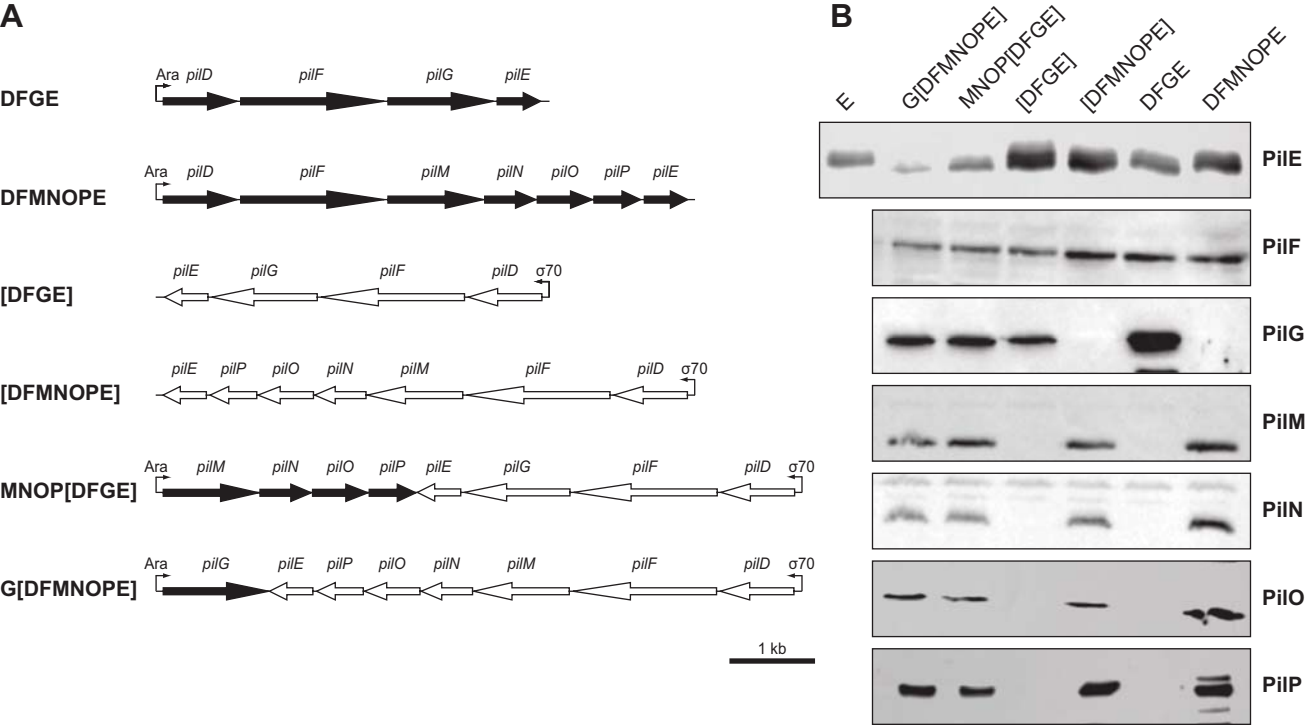
Fig. 3. PilM is dispensable for the stability of the PilMNOP complex, while PilP is essential via its unstructured N-terminal domain. (**A**) $PilNOP_{Strep}$ is a stable

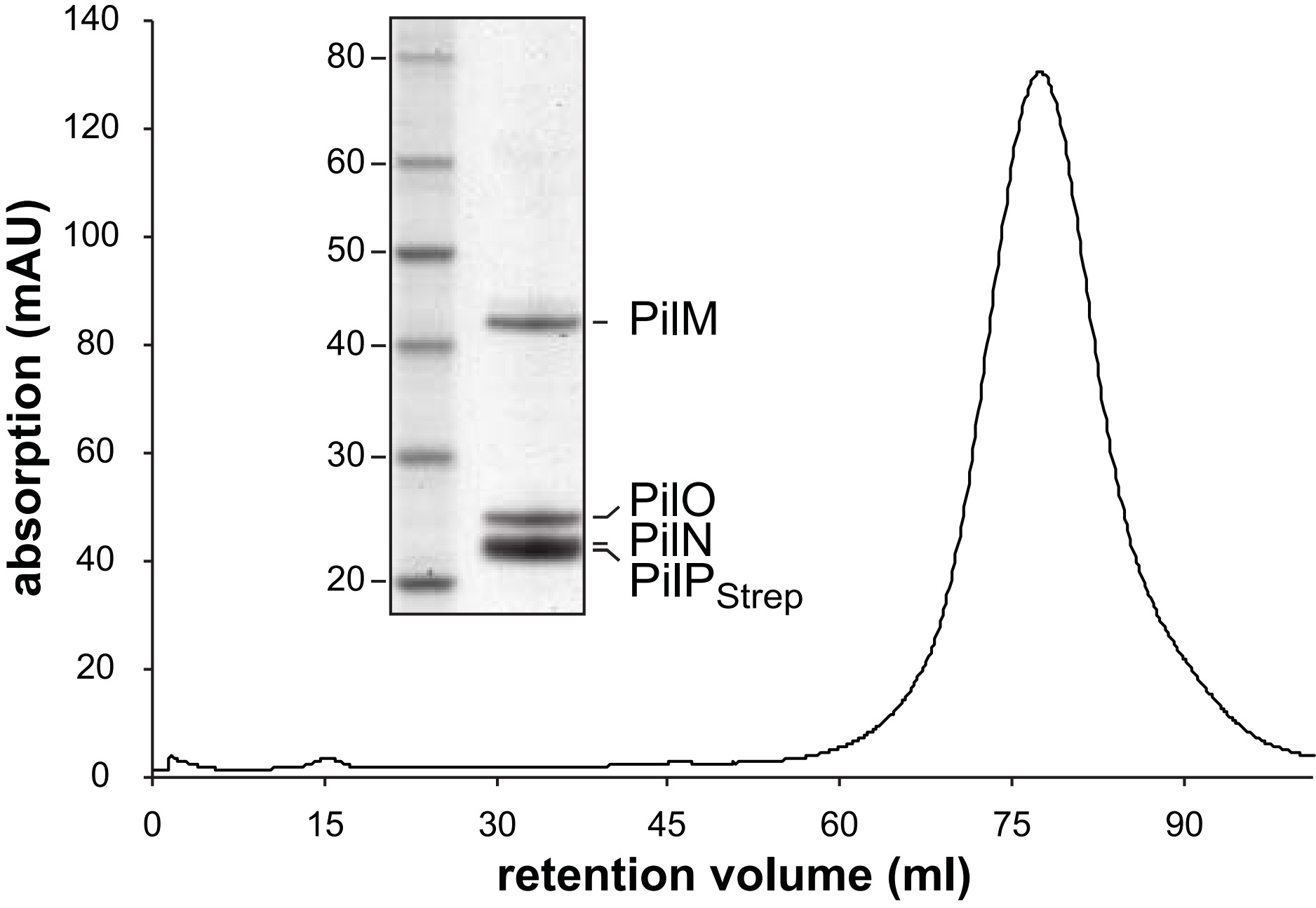
membrane complex, which can be purified to homogeneity. SEC profile of $PilNOP_{Strep}$ on a Superose 6 XK 16/70 column. The inset shows the SDS-PAGE/Coomassie analysis of the purified complex. (**B**) $PilMNOP_{NT\text{-}Strep}$, in which PilP has been truncated down to its predicted unstructured N-terminal domain is a stable membrane complex, which can be purified to homogeneity. SEC profile of $PilMNOP_{NT\text{-}Strep}$ on a Superose 6 10/300 GL column. The inset shows the SDS-PAGE/Coomassie analysis of the purified complex. The molecular weight for $PilP_{NT\text{-}Strep}$ is 9.8 kDa.

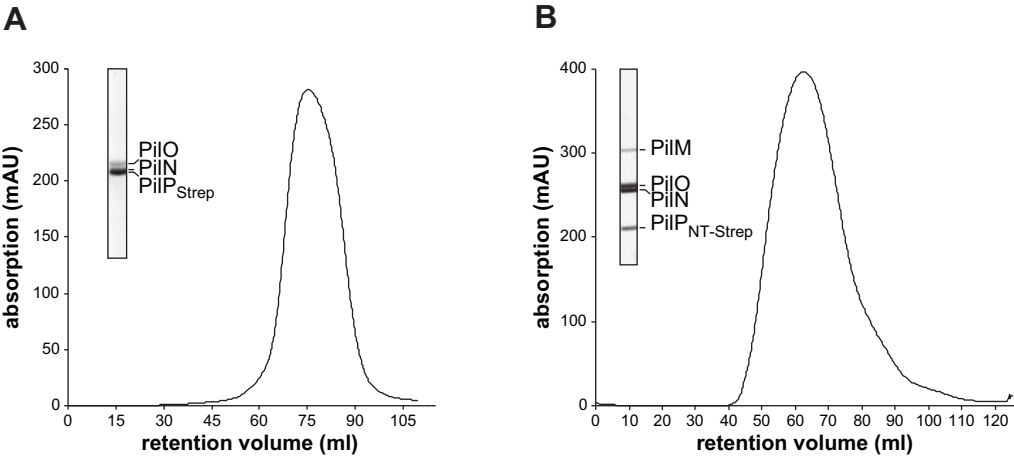
Fig. 4. Tfp assembly proteins PilG and PilF interact with PilMNOP, forming stable membrane complexes, which can be purified to homogeneity. (**A**) SEC profile of $PilG_{His}MNOP_{Strep}$ on Superose 6 XK 16/70 column. The inset shows the SDS-PAGE/Coomassie analysis of the purified complex. The molecular weight for $PilG_{His}$ is 46.5 kDa. (**B**) SEC profile of $PilF_{His}MNOP_{Strep}$ on a Superose 6 10/300 GL column. The inset shows the SDS-PAGE/Coomassie analysis of the purified complex. The molecular weight for $PilF_{His}$ is 63.2 kDa.

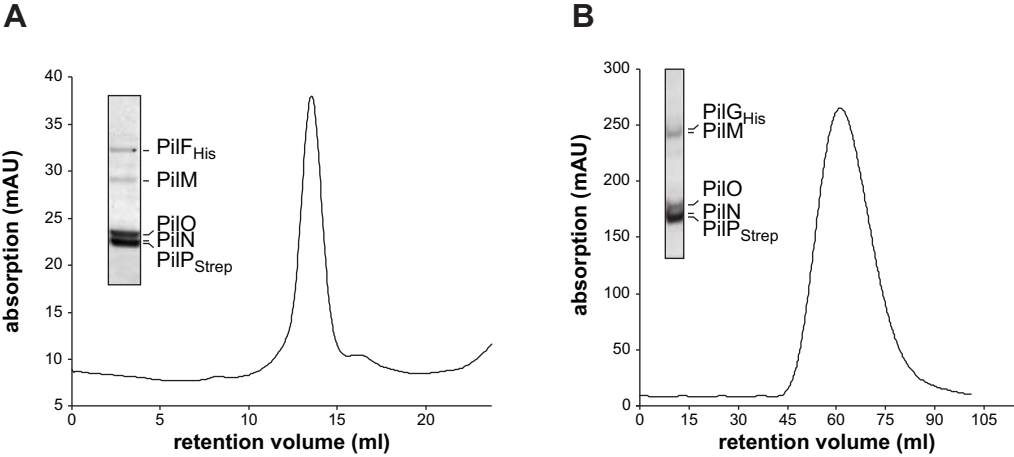
Fig. 5. PilG is important for Tfp assembly in *N. meningitidis*. (**A**) Schematic representation of *pilG* from *N. meningitidis* 8013 and the different mutations analysed in this study. *pilG1* and *pilG2* are Tn insertion mutants, while *ΔpilG* is a mutant in which the gene has been deleted and replaced with a cassette encoding kanamycin resistance. Picture is drawn to scale, with the scale bar representing 1 kb. (**B**) Immunoblot analysis of PilG production in the different mutants. Double mutants include either a *pilT* mutant with an inserted cassette encoding erythromycin resistance, or *ΔpilT* deletion mutant in which the gene has been deleted and replaced by a cassette encoding erythromycin resistance. Whole-cell meningococcal protein extracts were probed using anti-PilG and anti-PilE (as a positive control) antibodies. Protein extracts were quantified, equalised and equivalent amounts of total proteins were loaded in each lane. Molecular weights are indicated in kDa. (**C**) Piliation as

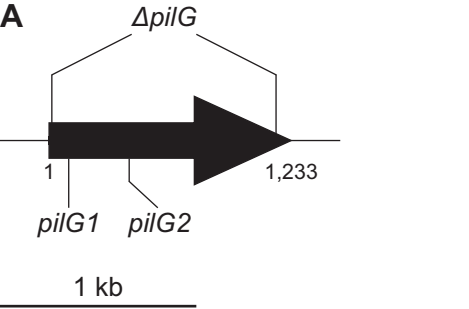
677 assessed by IF microscopy in the various *N. meningitidis pilG* mutants. The WT
678 strain was included as a positive control. Tfp (green) were labelled with a monoclonal
679 antibody specific for strain 8013 filaments and a secondary antibody coupled to Alexa
680 fluor 488, while the bacteria (red) were stained with DAPI. All the pictures were taken
681 at the same magnification. Scale bars represent 10 µm.
682
683 **Fig. 6. Eight Pil proteins (PilD, PilE, PilF, PilG, PilM, PilN, PilO and PilP) are**
684 **necessary and sufficient to promote Tfp assembly.** The presence of filaments in
685 *E. coli* TOP 10 transformed with various *pil* constructs (indicated in the upper left
686 corner of each panel) was assessed by IF microscopy. Tfp (green) were labelled with
687 a monoclonal antibody highly specific for strain 8013 filaments and a secondary
688 antibody coupled to Alexa fluor 488, while the bacteria (red) were stained with DAPI.
689 Except where indicated (panels **G** and **H**), the presence of filaments was assessed
690 by IF microscopy after the bacteria were submitted to an osmotic shock treatment to
691 release their periplasmic content (panels **A**-**F**). For those *pil* combination that lead to
692 Tfp assembly, MNOP[DFGE] (panels **C** and **E**) and G[DFMNOPE] (panels **D** and **F**),
693 two different experiments are shown. Long filaments are indicated by white
694 arrowheads. All the pictures were taken at the same magnification. Scale bars
695 represent 10 µm.
696
697 **Fig. 7. Working model for the functioning of the minimal machinery capable of**
698 **building Tfp. 1.** The PilMNOP complex is formed. PilM is likely to be recruited to the
699 cytoplasmic membrane upon PilNOP pre-assembly. **2.** The extension ATPase (PilF)
700 and platform protein (PilG) are recruited to the PilMNOP complex, forming the
701 filament assembly machinery. **3.** PilE subunits awaiting in the membrane, from a pool
702 of pilins processed by PilD, diffuse to the PilFGMNOP complex, which scoops
703 subunits out of the lipid bilayer and into the base of a growing filament.

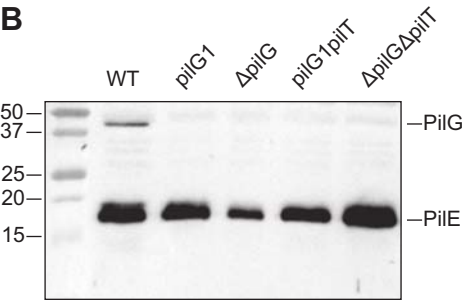

A
DFGE
Ara pilD pilF pilG pilE
DFMNOPE
Ara pilD pilF pilM pilN pilO pilP pilE
[DFGE]
pilE pilG pilF pilD σ70
[DFMNOPE]
pilE pilP pilO pilN pilM pilF pilD σ70
MNOP[DFGE]
Ara pilM pilN pilO pilP pilE pilG pilF pilD σ70
G[DFMNOPE]
Ara pilG pilE pilP pilO pilN pilM pilF pilD σ70
1 kb
B
E
G[DFMNOPE]
MNOP[DFGE]
[DFGE]
[DFMNOPE]
DFGE
DFMNOPE
PilE
PilF
PilG
PilM
PilN
PilO
PilP

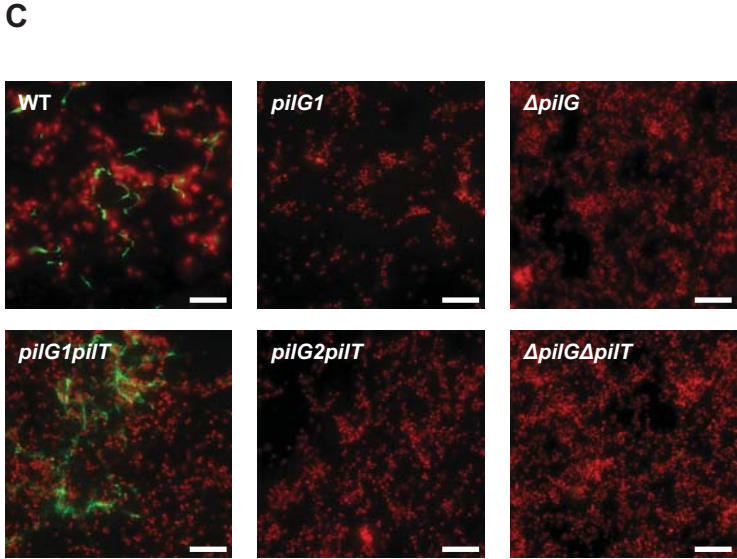

absorption (mAU)
140
120
100
80
60
40
20
0
0
15
30
45
60
75
90
retention volume (ml)
80
60
50
40
30
20
PilM
PilO
PilN
PilP Strep

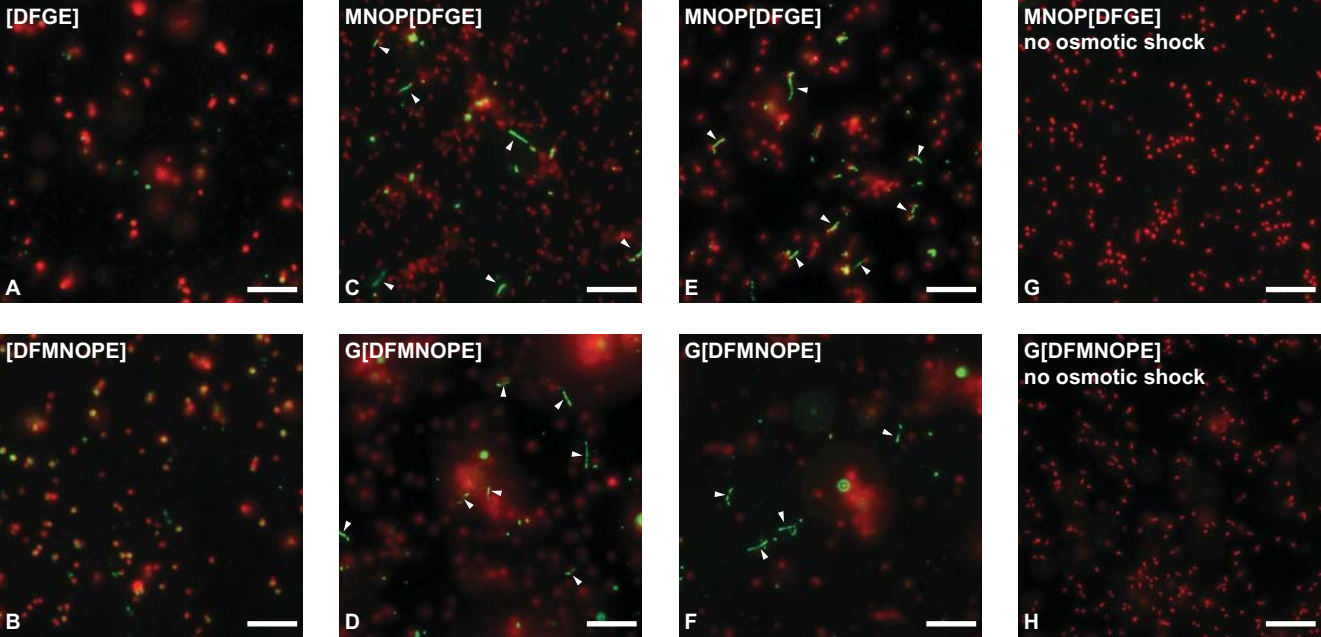

A
absorption (mAU)
300
250
200
150
100
50
0
0
15
30
45
60
75
90
105
retention volume (ml)
PilO
PilN
PilP Strep
B
absorption (mAU)
400
300
200
100
0
0
10
20
30
40
50
60
70
80
90
100
110
120
retention volume (ml)
PilM
PilO
PilN
PilP NT-Strep

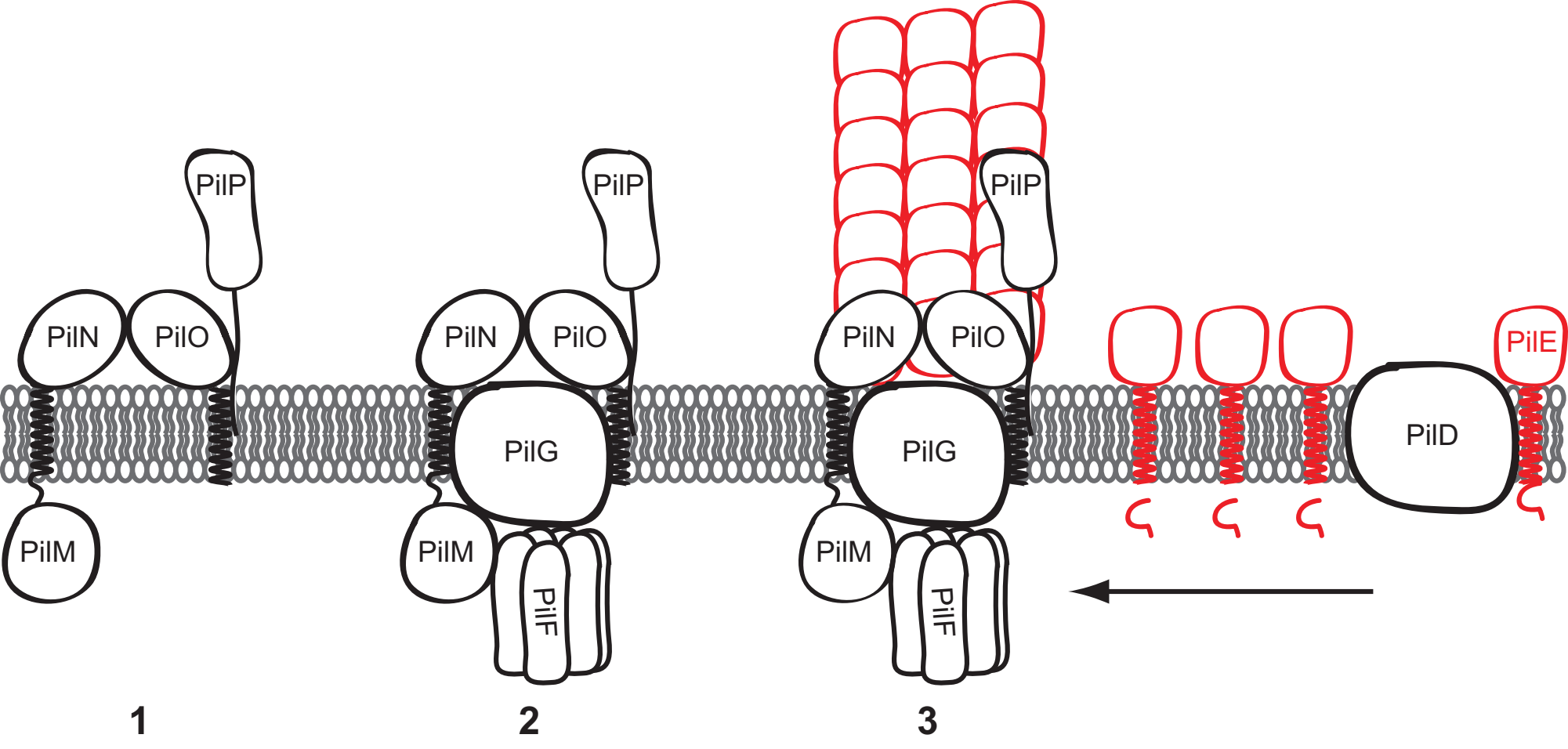

A
absorption (mAU)
retention volume (ml)
PilF$_{His}$
PilM
PilO
PilN
PilP$_{Strep}$
B
absorption (mAU)
retention volume (ml)
PilG$_{His}$
PilM
PilO
PilN
PilP$_{Strep}$


A
ΔpilG
1
1,233
pilG1
pilG2
1 kb


B
WT
pilG1
ΔpilG
pilG1pilT
ΔpilGΔpilT
50
37
25
20
15
PilG
PilE


C
WT
pilG1
ΔpilG
pilG1pilT
pilG2pilT
ΔpilGΔpilT


[DFGE]
A
MNOP[DFGE]
C
MNOP[DFGE]
E
MNOP[DFGE]
no osmotic shock
G
[DFMNOPE]
B
G[DFMNOPE]
D
G[DFMNOPE]
F
G[DFMNOPE]
no osmotic shock
H


PilP
PilN
PilO
PilM
PilG
PilF
PilD
PilE
1
2
3

Supplementary information

Reconstitution of a minimal machinery capable of assembling type IV pili

Vivianne J. Goosens[a,1], Andreas Busch[b,1], Michaella Georgiadou[a], Marta Castagnini[a], Katrina T. Forest[c], Gabriel Waksman[b] and Vladimir Pelicic[a,2]

[a]MRC Centre for Molecular Bacteriology and Infection, Imperial College London, London, United Kingdom

[b]Institute of Structural and Molecular Biology, University College London and Birkbeck, London, United Kingdom

[c]Department of Bacteriology, University of Wisconsin-Madison, Madison, USA

[1]These authors contributed equally

[2]To whom correspondence should be addressed. Email: v.pelicic@imperial.ac.uk

BIOLOGICAL SCIENCES (Microbiology)

Legends to supplementary figures

Fig. S1. Iterative cloning strategy used to construct the various *pil* operons. Pil genes/group of genes, initially cloned in pET-29b (not drawn to scale) were combined as follows. Gene B is extracted together with its RBS from the corresponding pET-derivative on a *Xba*I and *Xho*I fragment (sites underlined), and cloned into *Nhe*I-*Xho*I immediately downstream of gene A (*Xba*I and *Nhe*I generate compatible cohesive ends). This effectively creates an artificial AB operon whose expression is driven by the T7 promoter and where each gene is preceded by a canonical RBS. Since the *Nhe*I site downstream gene A is destroyed during this cloning step, the strategy can be employed iteratively to create operons of increasing size. Abbreviations are X for *Xba*I, N for *Nhe*I, and Xh for *Xho*I.

Fig. S2. Promoters driving the expression of *pil* genes/operons in pBAD18-derivatives. (**A**) Schematic representation and sequence of the pBAD18 region centered on the multiple cloning site (MCS). Drawn to scale, with the scale bar representing 50 bp. The two transcriptional start sites used to express *pil* genes/operons are represented in bold are indicated by an arrow. (**B**) Sequence of the promoter region. The two transcriptional start sites are highlighted in bold. Restriction sites are in italics.

Fig. S3. Immunoblot analysis of the purified PilMNOP$_{Strep}$ complex. Molecular weights are indicated in kDa.

Fig. S4. MALDI-MS analysis of the purified PilMNOP$_{Strep}$ complex in solution.

Fig. S5. A PilMNO complex is unstable. The PilMNO$_{Strep}$ complex elutes in several peaks and is unstable since the proteins in the different fractions aggregate after

purification. SEC profile of PilMNO$_{Strep}$ on a Superose 6 10/300 GL column. The left panel shows the SDS-PAGE/Coomassie analysis of the different fractions. A molecular weight marker was run in the last lane. The molecular weight for PilO$_{Strep}$ is 24.5 kDa.

Fig. S6. Immunoblot analysis of the purified complexes PilF$_{His}$MNOP$_{Strep}$ (panel A) and PilG$_{His}$MNOP$_{Strep}$ (panel B). Molecular weights are indicated in kDa.

Fig. S7. PilF and PilG do not form stable complexes with PilNOP. (**A**) SEC profile of the complex purified on a Superose 6 XK 16/70 column upon expression of *pilF$_{His}$NOP$_{Strep}$*. * void volume. The inset shows the SDS-PAGE/Coomassie analysis of the purified complex, which corresponds to PilNOP$_{Strep}$. A molecular weight marker was run in the first lane. (**B**) SEC profile of the complex purified on a Superose 6 XK 16/70 column upon expression of *pilG$_{His}$NOP$_{Strep}$*. The left panels show the SDS-PAGE/Coomassie and anti-PilG immunoblot analysis of the different peaks. A molecular weight marker was run in the first lane of each gel. In each peak we detected a lower molecular weight species, indicated with *, which is either a truncated form of PilG$_{His}$ or a cross-reacting species. This band was also present in the immunoblot analysis of the PilG$_{His}$MNOP$_{Strep}$ complex shown in Fig. S6B.

Table S1. SEC-MALS data.

Complex	Mp (kDa)	Mw (kDa)	Polydispersity (Mw/Mn)	Mw protein (kDa)	Mw β-DDM (kDa)
$PilMNOP_{Strep}$	198. 6 ± 3.51	200.3 ± 3.24	1.003 ± 0.001	132.67 ± 1.62	67.6 ± 2.68
$PilNOP_{Strep}$	257.47 ± 2.05	258.47 ± 4.5	1.001 ± 0.001	214.87 ± 9.21	43.57 ± 4.96

Values are average ± standard deviation (*n* = 3).

Table S2. Primers used in this study.

Name	Sequence
Cloning in pET-29b	
Doperon-F	GGCATATGAGCGATCTGAGCGTGCTGA
Doperon-R	CTCGAGTGGCTAGCTTACAGAACCGGATGTGTCAGCC
Eoperon-F	GGCATATGAACACCCTGCAGAAAG
Eoperon-R	CTCGAGTGGCTAGCTTAGCTTGCGCTGCTATCATCAC
Foperon-F	GGCATATGAGCGTTGGTCTGCTGCGTAT
Foperon-R	CTCGAGTGGCTAGCTTAATCATTGGTATTTGCGGTAAC
MNOPoperon-F	GGCATATGCGCCTGTTTAAAAGCCTGA
MNOPoperon-R	CTCGAGTGGCTAGCTTAATTCTGTTCTGCTGCCGGTG
Goperon-F	GGCATATGGCTAAAAATGGTGGCTTTAG
Goperon-R	CTCGAGTGGCTAGCTCATGCAACAACATTACCCAG
Cloning in pASK-IBA3C	
$MNOP_{Strep}$-F	ATGGTAGGTCTCAAATGCGCCTGTTTAAAAGCCTG
$MNOP_{Strep}$-R	ATGGTAGGTCTCAGCGCTATTCTGTTCTGCTGCCGGTGC
MNO_{Strep}-F	AACAGAAAAGCGCTTGGAGCCACCCG
MNO_{Strep}-R	AAGCGCTTTTCTGTTCTGCATTCTGTGCTG
NOP_{Strep}-F	TAACGAGGGCAAAAAATGAACAACCTGATCAAAATC
NOP_{Strep}-R	TTTTTGCCCTCGTTATCTAGATTTTTG
$MNOP_{NT\text{-}Strep}$-F	AAAATGCAAGCGCTTGGAGCCACCC
$MNOP_{NT\text{-}Strep}$-R	AAGCGCTTGCATTTTCACCTTTTTTGTCGG
F_{His}-F	TAACGAGGGCAAAAAATGAGCGTTGGTCTG
$F_{His}MNOP_{Strep}$-R	TTTAAACAGGCGCATGGTATATCTCCTTCTTAATCATTGGTATT
$F_{His}NOP_{Strep}$-R	GATCAGGTTGTTCATGGTATATCTCCTTCTTAATCATTGGTATT
$G_{His}MNOP_{Strep}$-F	TAACGAGGGCAAAAAATGGCTAAAAATGGTGGCTTTAGCC
$G_{His}MNOP_{Strep}$-R	TTTAAACAGGCGCATTGTATATCTCCTTCTCATGCAACAAC
linearise $MNOP_{Strep}$-F	ATGCGCCTGTTTAAAAGCCTGAAAA
linearise $MNOP_{Strep}$-R	TTTTTGCCCTCGTTATCTAGATTTT
linearise $MNOP_{Strep}$/NOP_{Strep}-R	TTTTTGCCCTCGTTATCTAGATTTT
5' RACE	
GSP1-PilD	GCATAACAATTGCGGCAACC
GSP2-PilD1	GCCCATACCAATTTCGCCTT
GSP2-PilD2	AGAGGCAGGGTCATGCTATC
GSP3-PilD	TTCCGGTGATGATGGAACGT
Design of *N. meningitidis* deletion mutants	
aph1	**ATGGCTAAAATGAGAATATCACC**
aph2	**CTAAAACAATTCATCCAGTAAAA**
ΔpilG-F1	GCCCGGGTAAAAGCATTCAT
ΔpilG-R1	**GGTGATATTCTCATTTTAGCCAT**ATTCTAGCCCCGTAAAGTAATTGG
ΔpilG-F2	**TTTTACTGGATGAATTGTTTTAG**TTTTAGGCTTGGTCATCGGC
ΔpilG-R2	TTCAGACGGCATGGAAATCAGAAACGCGGTCA
erm1	**ATGAACAAAAATATAAAATATTC**
erm2	**TTATTTCCTCCCGTTAAATAATA**
ΔpilT-F1	ACCTCGATACACACCTGCAA
ΔpilT-R1	**GAATATTTTATATTTTTGTTCAT**AATGCGGCTCTGTTTAGTATAATG
ΔpilT-F2	**TATTATTTAACGGGAGGAAATAA**CGGAAGTCGCACGCAGAC
ΔpilT-R2	TTCAGACGGCATCGGAGTAGTTGATTTCGCCG

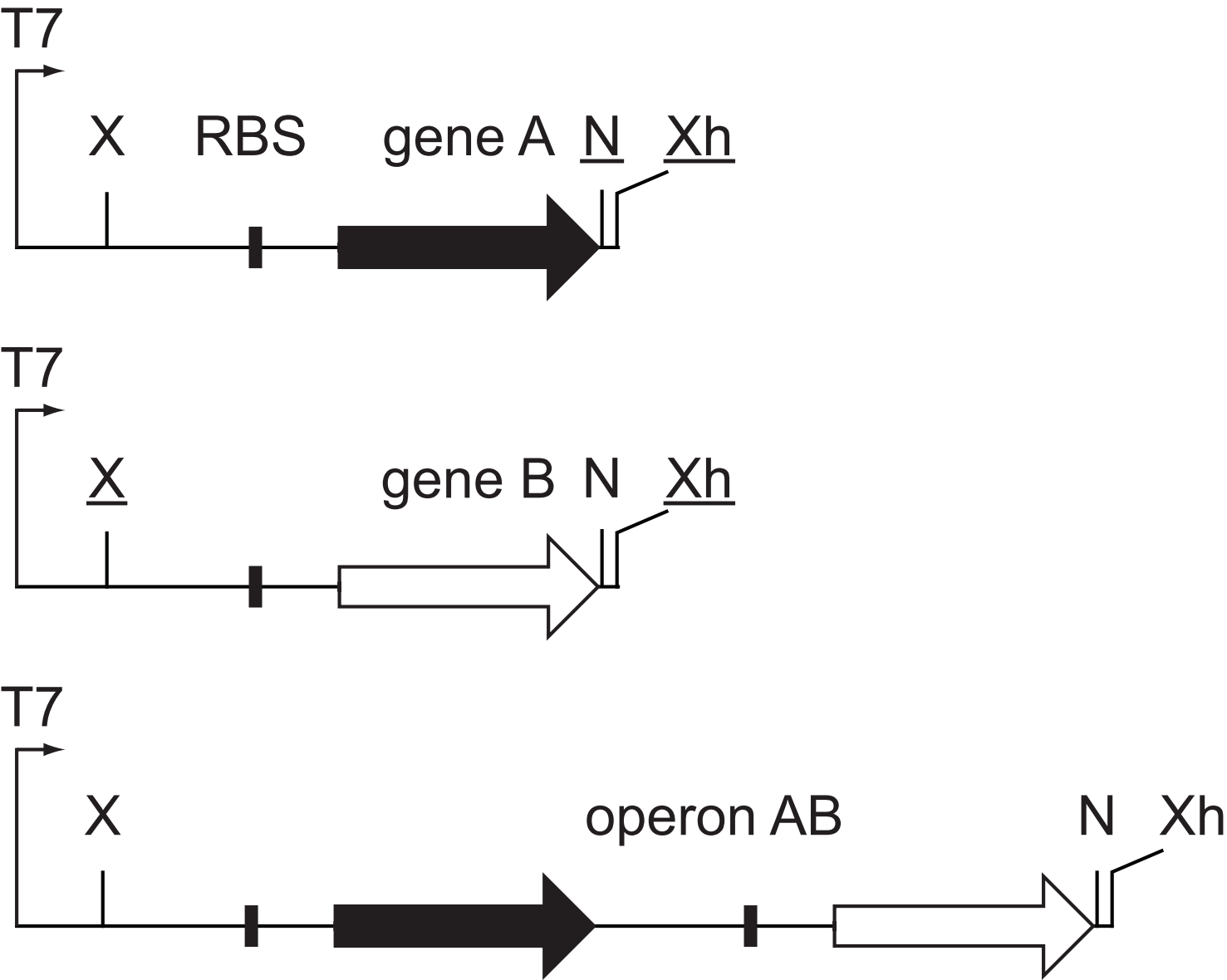
Restriction sites are underlined. Regions of complementarity for splicing PCR are in bold.

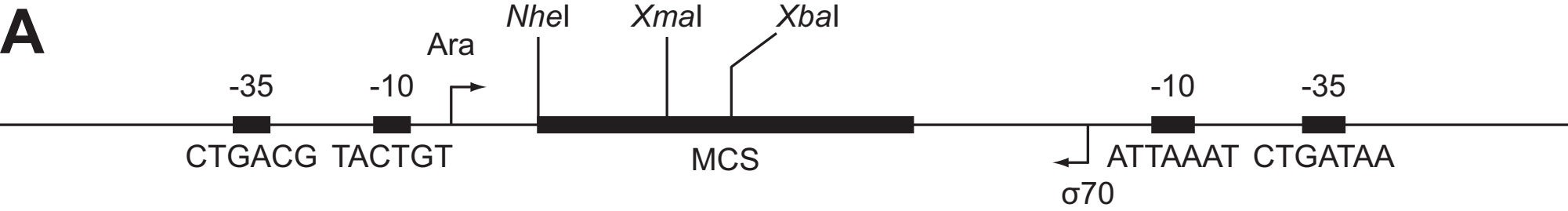
Table S3. Plasmids used in this study.

Name[a]	Details[b]	Source/Reference
pET-29b	T7-based expression vector	Novagen
pBAD18	arabinose-inducible expression vector	(32)
pASK-IBA3C	ATc-inducible expression vector, *Strep*-tag fused to the C-terminus of target protein	iba
pET-*pilD*	pET-derivative with *pilD*	this study
pET-*pilE*	pET-derivative with *pilE*	this study
pET-*pilF*	pET-derivative with *pilF*	this study
pET-*pilG*	pET-derivative with *pilG*	this study
pET-*pilMNOP*	pET-derivative with *pilMNOP*	this study
pET-*pilDF*	pET-derivative with *pilDF*	this study
pET-*pilDFG*	pET-derivative with *pilDFG*	this study
pET-*pilDFMNOP*	pET-derivative with *pilDFMNOP*	this study
pET-*pilDFGE*	pET-derivative with *pilDFGE*	this study
pET-*pilDFMNOPE*	pET-derivative with *pilDFMNOPE*	this study
pBAD-*pilE*	pBAD-derivative with *pilE*	this study
pBAD-*pilDFGE*	pBAD-derivative with *pilDFGE*	this study
pBAD-*[pilDFGE]*	pBAD-derivative with *[pilDFGE]*	this study
pBAD-*pilDFMNOPE*	pBAD-derivative with *pilDFMNOPE*	this study
pBAD-*[pilDFMNOPE]*	pBAD-derivative with *[pilDFMNOPE]*	this study
pBAD-*pilG[pilDFMNOPE]*	pBAD-derivative with *pilG[pilDFMNOPE]*	this study
pBAD-*pilMNOP[pilDFGE]*	pBAD-derivative with *pilMNOP[pilDFGE]*	this study
pASK-$pilMNOP_{Strep}$	pASK derivative with $pilMNOP_{Strep}$	this study
pASK-$pilMNO_{Strep}$	pASK derivative with $pilMNO_{Strep}$	this study
pASK-$pilNOP_{Strep}$	pASK derivative with $pilNOP_{Strep}$	this study
pASK-$pilMNOP_{NT\text{-}Strep}$	pASK derivative with $pilMNOP_{NT\text{-}Strep}$	this study
pASK-$pilF_{His}MNOP_{Strep}$	pASK derivative with $pilF_{His}MNOP_{Strep}$	this study
pASK-$pilG_{His}MNOP_{Strep}$	pASK derivative with $pilG_{His}MNOP_{Strep}$	this study
pASK-$pilF_{His}NOP_{Strep}$	pASK derivative with $pilF_{His}NOP_{Strep}$	this study
pASK-$pilG_{His}NOP_{Strep}$	pASK derivative with $pilG_{His}NOP_{Strep}$	this study
pASK-$pilFG_{His}MNOP_{Strep}$	pASK derivative with $pilFG_{His}MNOP_{Strep}$	this study

[a]In pBAD18, genes within brackets *[pil]* are under the control of a σ70 promoter, if not they are under the control of the arabinose-inducible promoter.

[b]All the *pil* genes are synthetic *N. meningitidis* genes codon-optimised for expression in *E. coli*.


T7
X
RBS
gene A
N
Xh
T7
X
gene B
N
Xh
T7
X
operon AB
N
Xh



B

-35

atgccatagcatttttatccataagattagcggatcctacctgacgctttttatcgcaactctctactgtttctccatacccgttttttttgggctagcga

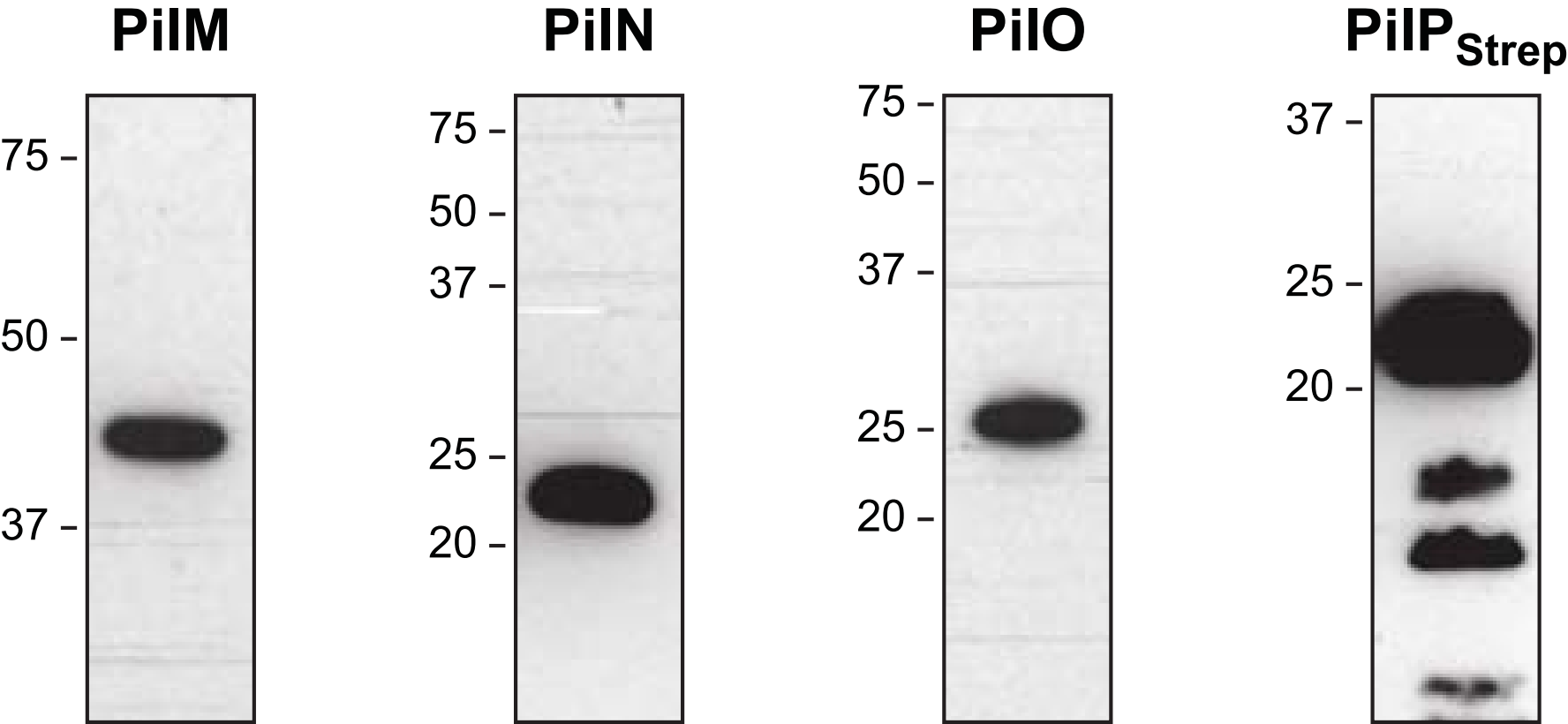
-10 NheI

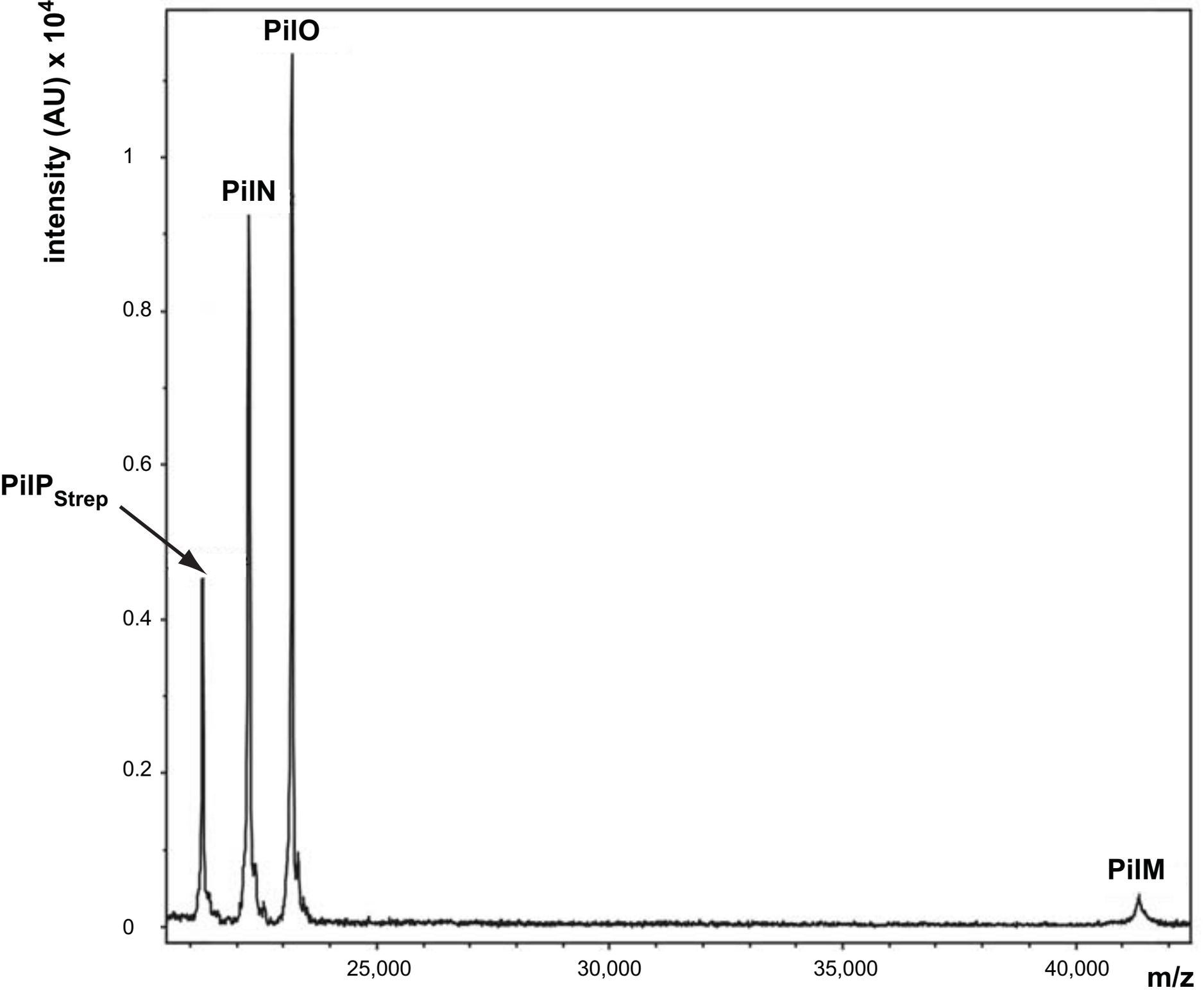
attcgagctcggtacccggggatcctctagagtcgacctgcaggcatgcaagcttggctgttttggcggatgagagaagattttcagcctgatacagatt

XmaI XbaI

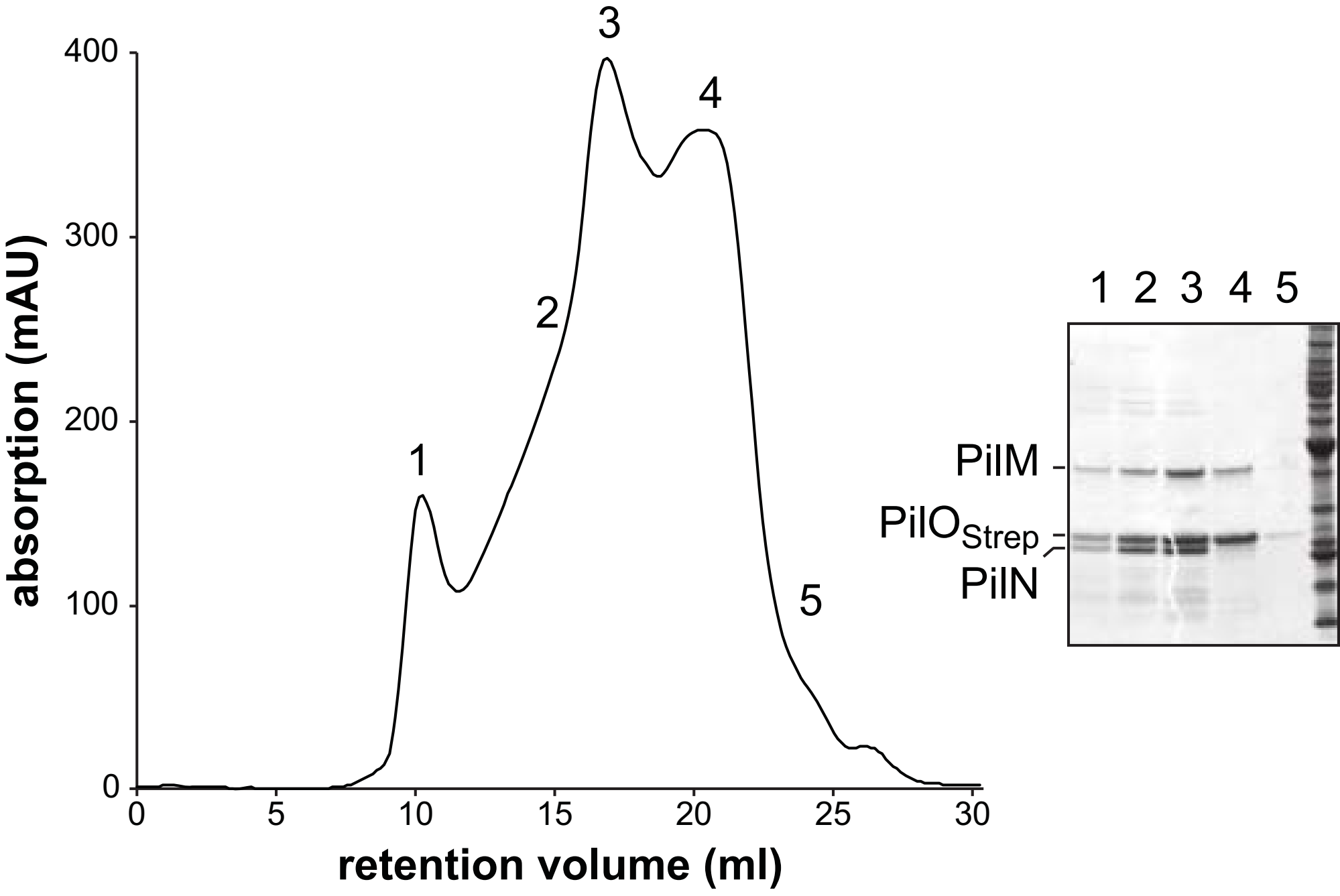
-10 -35

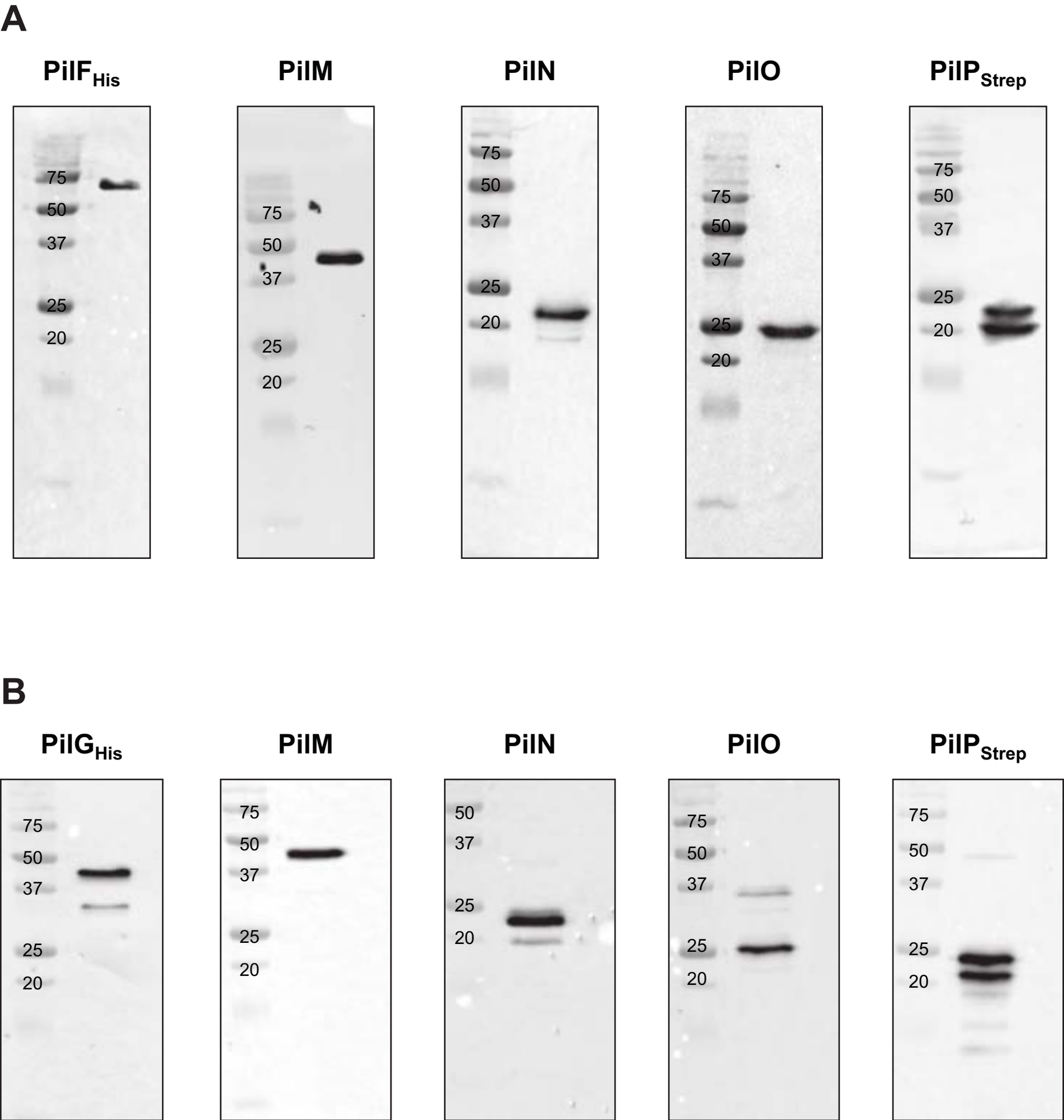
aaatcagaacgcagaagcggtctgataaaacagaatttgcctggcggcagtagcgcggtggtcccacc

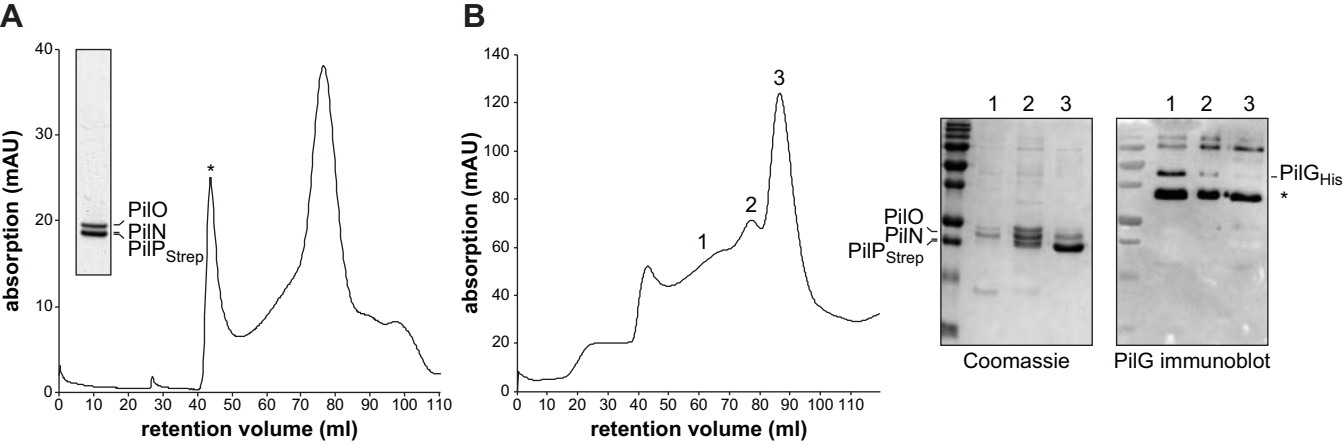

PilM
75
50
37
PilN
75
50
37
25
20
PilO
75
50
37
25
20
PilP Strep
37
25
20



mass (Da)	PilM	PilN	PilO	$PilP_{Strep}$
theoretical mass	41,366.7	22,257.4	23,315.4	21,265.8
measured mass	41,363.7	22,261.3	23,185.8	21,269.3


absorption (mAU)
400
300
200
100
0
0
5
10
15
20
25
30
retention volume (ml)
1
2
3
4
5
1 2 3 4 5
PilM
PilO Strep
PilN


A
PilF$_{His}$
PilM
PilN
PilO
PilP$_{Strep}$
75
50
37
25
20
B
PilG$_{His}$
PilM
PilN
PilO
PilP$_{Strep}$


A
absorption (mAU)
retention volume (ml)
PilO
PilN
PilP Strep
*
B
absorption (mAU)
retention volume (ml)
1
2
3
PilO
PilN
PilP Strep
1 2 3
Coomassie
1 2 3
PilG His
*
PilG immunoblot