The molecular basis of FimT-mediated DNA uptake during bacterial natural transformation

Sebastian A. G. Braus1, Francesca L. Short2, Stefanie Holz1,3, Matthew J. M. Stedman1,3, Alvar D. Gossert1 & Manuela K. Hospenthal1

Naturally competent bacteria encode sophisticated protein machinery for the uptake and translocation of exogenous DNA into the cell. If this DNA is integrated into the bacterial genome, the bacterium is said to be naturally transformed. Most competent bacterial species utilise type IV pili for the initial DNA uptake step. These proteinaceous cell-surface structures are composed of thousands of pilus subunits (pilins), designated as major or minor according to their relative abundance in the pilus. Here, we show that the minor pilin FimT plays an important role in the natural transformation of Legionella pneumophila. We use NMR spectroscopy, in vitro DNA binding assays and in vivo transformation assays to understand the molecular basis of FimT’s role in this process. FimT binds to DNA via an electropositive patch, rich in arginines, several of which are well-conserved and located in a conformationally flexible C-terminal tail. FimT orthologues from other Gammaproteobacteria share the ability to bind to DNA. Our results suggest that FimT plays an important role in DNA uptake in a wide range of competent species.
C ompetent bacteria can take up exogenous DNA, present in their environment, and integrate it into their genomes by the process of natural transformation. This is an important avenue of horizontal gene transfer (HGT), which has widespread consequences for bacterial evolution and the spread of antibiotic resistance and other pathogenicity traits. In contrast to other modes of HGT, namely transduction and conjugation, natural transformation is entirely controlled by the recipient cell that encodes all the required machinery for DNA uptake, translocation and integration. More than 80 bacterial species, including Gram-negative and Gram-positive organisms, have been shown to be naturally competent, yet the true prevalence of this mechanism amongst bacteria likely remains under-appreciated. The Gram-negative bacterium Legionella pneumophila is naturally competent, consistent with the observation that its genome bears evidence of frequent HGT and recombination events. Although L. pneumophila could be described as an accidental human pathogen, it is the aetiological agent of Legionnaire’s disease, a serious and life-threatening form of pneumonia, that results from an infection of alveolar macrophages by contaminated aerosols.

Legionella, like most Gram-negative bacteria, are thought to utilise type IV pili (T4P) for DNA uptake, which is defined as the movement of DNA across the outer membrane (OM) and into the periplasmic space. However, the molecular mechanisms involved in this step remain poorly defined. T4P are extracellular proteinaceous appendages composed of thousands of individual pilus subunits (pilins), designated as major or minor depending on their relative abundance in the pilus. A prevailing model suggests that T4P can bind to DNA and transport it into the cell via pilus retraction, which is powered by the retraction ATPase PilT. Pilus retraction is thought to bring the DNA into proximity with the OM and be taken up across the OM-embedded secretin channel PilQ, which is the same pore traversed by the T4P themselves. Once in the periplasm, ComEA binds to incoming DNA to prevent its back-diffusion by acting like a Brownian ratchet. Subsequently, DNA is converted into single-stranded DNA (ssDNA) and transported across the inner membrane (IM) by a putative channel called ComEC. In the cytoplasm, ssDNA is protected by single-stranded DNA-binding protein (Ssb) and DNA processing protein A (DpaA), before being integrated into the genome by homologous recombination in a RecA- and ComM-dependent manner.

In recent years, studies of several competent bacteria have shown that their T4P (or their pilins) can directly interact with DNA and form DNA-binding protein complexes. This function was attributed to specialised minor pilins or pilin-like proteins in Neisseria species (ComP), Vibrio cholerae (VC0858 and VC0859), and Thermus thermophilus (ComZ), although a major pilin (PilA4) has also been suggested to contribute in the latter. Of these, ComP found in Neisseria species, is the best-characterised DNA-binding minor pilin to date. ComP displays a sequence preference for neisserial DNA containing so-called DNA uptake sequences (DUS) and binds to DNA through an electropositive surface patch. Co-localisation of ComP with the pilus tip has been suggested to either be incorporated throughout the pilus fibre or at the pilus tip. In addition to these proteins, the minor pilin FimT has also been implicated in natural transformation, as its loss leads to a reduction in transformation efficiency in Acinetobacter baumannii. However, this phenotype was never followed up with further DNA-binding studies.

We set out to study DNA uptake during natural transformation in Legionella pneumophila. It is not known whether Legionella’s T4P can interact with DNA, and if so, which pilins are responsible. Here we show that FimT efficiently interacts with DNA in vitro and in vivo, and that loss of binding, just like fimT deletion, results in almost complete abrogation of natural transformation. We present the structure of FimT and show that a conserved electropositive surface patch rich in arginines is required for DNA binding. Finally, we show that FimT is not only important for natural transformation in L. pneumophila, but that it likely plays a role in many other bacterial species, as suggested by DNA binding studies and bioinformatic analyses. Together, our work provides the molecular basis of FimT’s role in natural transformation.

**Results**

FimT is critical for natural transformation in L. pneumophila and interacts with DNA. L. pneumophila encodes type IV pilins in several genetic loci. The largest of these operons contains fimU, pilV, pilW, pilX, pilY1 and pilE, and is shared, in a similar arrangement, by many other T4P-containing organisms. In addition, the L. pneumophila Philadelphia-1 genome encodes several other type IV pilins and putative type IV pilins in locations seemingly far removed from the fimU-containing operon, or indeed from other genes encoding components of the T4P machinery. Amongst these are two consecutively encoded PilA homologues, PilA1 and PilA2, three further putative pilin genes, lpg1995, lpg1996 and lpg1997, arranged in an operonic structure, and fimT, which appears entirely on its own. Recent work has suggested that PilA2 is the major pilin in L. pneumophila and therefore required for natural transformation, while PilA1 appears dispensable.

FimT and FimU are minor type IV pilins that belong to the GspH/FimT family of proteins (Pfam: PF12019; InterPro: IPR022346), which also includes the type II secretion system (T2SS) pseudopilin GspH/XcpU. All three genes are encoded in the L. pneumophila genome and share an overall amino acid sequence identity of ~15–25%. L. pneumophila FimT (FimTlp) and FimU (FimULp) possess all the features of typical type IV pilins, including an N-terminal signal sequence required for their targeting to the inner membrane (IM), followed by a hydrophobic transmembrane helix required for IM insertion prior to pilus assembly and proper packing into the filament structure post-assembly. To test whether FimTlp or FimULp play a role in natural transformation in L. pneumophila, we performed transformation assays comparing the fimT and fimU deletion strains with the parental strain and strains harbouring deletions in genes known to be important for natural transformation (Fig. 1a). Deletion of comEC, encoding the putative IM DNA channel, pilQ, encoding the OM secretin, and pilT, encoding the retraction ATPase, resulted in undetectable levels of natural transformation in our assay. These observations are in close agreement with previous studies in L. pneumophila, as well as other competent Gram-negative organisms such as V. cholerae, where deletion of these genes resulted in a complete loss or severe reduction of natural transformation. Deletion of fimU did not produce a phenotype, whereas natural transformation was undetectable in the fimT deletion strain, as observed previously in A. baylyi. Expression of FimTlp in trans from an IPTG-inducible promoter restored the transformation efficiency of our L. pneumophila strain to wild-type levels, showing that the transformation defect is specific to FimTlp. Although the absence of FimT does not appear to affect pilus biogenesis or other T4P functions such as twitching motility in other organisms, this possibility cannot be ruled out.

We reasoned that FimT contributes to the OM DNA uptake step of natural transformation by forming a constituent part of type IV pili (T4P) able to directly bind to DNA. Therefore, we performed electrophoretic mobility shift assays (EMSA) to test
Fig. 1 FimT is critical for the transformation of \( L. \) pneumophila and binds to DNA. a Natural transformation efficiencies of the parental \( L. \) pneumophila Lp02 strain and Lp02 strains harbouring deletions of genes known to play a role in transformation compared to the fimU and fimT deletion strains. The \( \Delta \)fimT strain was complemented by ectopic expression of wild-type FimT, under the control of an IPTG-inducible promoter. The mean transformation efficiencies of three independent biological replicates is shown (error bars represent standard deviation [SD]). < d.l., below detection limit (d.l.) (average d.l. = \( 2.0 \times 10^{-8} \pm 8.2 \times 10^{-9} \)). Statistical significances of transformation differences were determined on log-transformed data using an unpaired two-sided t-test with Welch’s correction. All strains were compared to the parental strain. n.s., not statistically significant, \( p > 0.05 \) (\( p_{\Delta \text{fimU}} = 0.39; p_{\Delta \text{fimT Ptac-fimT}} = 0.89 \)). b In vitro DNA binding of purified \( L. \) pneumophila PilA1, PilA2, FimU and FimT assessed by an EMSA. A 30 bp dsDNA fragment (1\( \mu \)M) was incubated with increasing concentrations of purified pilins (0–100 \( \mu \)M) and resolved by agarose gel electrophoresis. This experiment was independently performed three times with reproducible results. c ITC binding studies of wild-type FimT binding to 12meric dsDNA (left) and ssDNA (right). In both cases, DNA (syringe) was injected into FimT (cell). Data were fitted using the "one set" of sites model, assuming that both binding sites on the dsDNA are of equal affinity. Source data are provided as a Source Data file.
whether FimTLP is able to bind to DNA in vitro (Fig. 1b). In order to produce soluble protein samples, all pilins were expressed as truncations lacking the N-terminal transmembrane helix (Supplementary Fig. 1a). Indeed, purified FimTLP interacted with all DNA probes tested, including ssDNA, dsDNA, linear and circular DNA molecules, whereas neither FimULP nor the putative major pilin subunits (PilA1 and PilA2) showed any interaction (Fig. 1b and Supplementary Fig. 1b). These experiments suggest that the dissociation constant (K_D) of the interaction between FimTLP and 30meric DNA is in the low µM range. In order to determine the K_D more precisely and to learn about the binding stoichiometry of this interaction, we performed isothermal titration calorimetry (ITC) utilising shorter 12meric ssDNA or dsDNA fragments (Fig. 1c). We determined a K_D of 7.0 µM and 2.8 µM for 12meric ssDNA and dsDNA, respectively. Interestingly, these experiments revealed that a single FimTLP binds to 12meric ssDNA, whereas two molecules can bind to the dsDNA ligand, suggesting two binding sites on opposite sides of the double helix.

The solution structure of FimTLP. We determined the solution structure of the soluble N-terminally truncated (residues 28–152, mature pilin sequence number) FimTLP by nuclear magnetic resonance (NMR) spectroscopy (Fig. 2a, Supplementary Table 1). The structure consists of an N-terminal α-helix (α1C) (the transmembrane portion of this helix, α1 N, has been removed in the construct), two β-sheets that complete the C-terminal globular pilin domain, and a C-terminal tail, which exhibits conformational flexibility. Both β-sheets are composed of antiparallel strands: β-sheet I is formed by β1, β2, β3 and β5, and β-sheet II by β4, β6 and β7. The closest structural homologue is FimU from Pseudomonas aeruginosa (FimU) (PDB ID: 4IPU, 4IPV) (Fig. 2b). While the two structures share a common fold, there are some key differences. In the FimU structure, the loop between β2 and β3 in β-sheet I forms an additional β-hairpin (β2′ and β2″). It is possible, however, that this additional β-hairpin of FimULP simply represents the conformation captured in the crystal structure, as the length of the β2–β3 loop is similar in both proteins. In addition, the β7 strand of β-sheet II is longer in FimU, and FimULP contains an additional strand (β8)36. Furthermore, FimULP contains a disulphide bond connecting Cys127 of β6 to the penultimate residue, Cys158, effectively stapling the C-terminal tail in place on top of β-sheet II. Such a disulphide bond is found in various major and minor pilins and the

Fig. 2 The structure of FimTLP. a The solution structure of FimTLP 28–152 (state 18) in ribbon representation (left) and the corresponding topology diagram (right). Secondary structure elements are indicated: truncated N-terminal α-helix (α1C) (blue), β-sheet I formed by β1, β2, β3, and β5 (yellow), and β-sheet II formed by β4, β6, and β7 (magenta). A vertical arrow indicates the pilus axis from the cell surface towards the pilus tip. b Structure alignment of FimTLP (blue) and FimU (grey; PDB ID: 4IPV) (left) and the topology diagram of FimU (right). The disulphide bond of FimU is indicated in stick representation with sulphur atoms in yellow. c Superimposed 20 lowest energy structures calculated by NMR spectroscopy. An arrow indicates the conformational flexibility of the C-terminal tail (140–152). The pairwise backbone root-mean-square deviation (RMSD) for the structured region (residues 32–62, 70–139) is 1.13 Å. N- and C-termini are indicated in each panel. d ξ α chemical shift values (top) and T2(1H) transverse relaxation data (bottom), encompassing the last 27 residues of FimTLP. Secondary structural elements are indicated and error bars represent the fitting errors of the respective exponential decay curves. Source data are provided as a Source Data file.
The C-terminal tail (residues 140–152) of FimTLp is unique amongst the currently determined structures of GspH/FimT family members. Different pieces of NMR data suggest significant conformational exchange, but not an entirely flexibly disordered tail. The amide resonances of residues 140–149 are very weak and those of residues 142–144 are not visible at all. We could not observe any intense long-range nuclear Overhauser effects (NOEs) for those of residues 142–152, which would be expected for a well-defined β-sheet conformation. T2 relaxation measurements indicated conformational exchange on the millisecond timescale, as the T2 values for backbone amide 1H and 15N nuclei for the C-terminal tail were approximately half the value of the structured part of the protein (Fig. 2d, Supplementary Fig. 3). A fully disordered C-terminal tail could however be excluded by 1H-15N heteronuclear NOE measurements, as the NOE intensity for the amides in the tail was close to the theoretical value of 0.78, which is expected for amides on globular particles. Finally, the deviations of Ca chemical shifts from random coil values clearly indicated a β-strand propensity (Fig. 2d, Supplementary Fig. 3). The data therefore suggest that the C-terminal amino acids have a β-strand-like backbone conformation but sample different states in the micro- to millisecond timescale. These findings are further supported by low amide proton temperature coefficients43 and increased proteolytic susceptibility of this region, compared to the rest of the structure, witnessed by disappearance of the NMR resonances of the tail after prolonged storage of samples.

**FimTLp interacts with DNA via a conserved C-terminal region rich in arginines.** Next, we characterised the residues of FimTLp involved in DNA binding using NMR spectroscopy (Fig. 3a–c). We performed binding experiments titrating increasing amounts of 12 bp dsDNA into 15N-labelled FimTLp and recorded 15N, 1H heteronuclear single-quantum correlation (2D [15N,1H]-HSQC) spectra. Most FimTLp resonances remained unperturbed (Fig. 3a), which suggests that no global conformational change occurs upon DNA binding. However, a subset of resonances exhibit marked chemical shift perturbations (CSPs) (Fig. 3a), indicating changes in the local chemical environment resulting from direct contact with DNA or other indirect conformational changes. A plot of CSPs against the amino acid sequence is shown in Fig. 3b, and we mapped CSPs greater than a threshold (Δppm > 1σ) onto the FimTLp surface (Fig. 3c, Supplementary Fig. 4). The largest CSPs correspond to residues located in three adjacent loop regions in the C-terminal globular domain of the protein, the β4-β5 loop (residues 103–106), the β5-β6 loop (118–126) and the C-terminal tail (140–152) (Fig. 3b). These shifts predominantly map to an elongated surface patch connecting the C-terminal tail with the globular C-terminal domain of FimTLp (Fig. 3c). Most of these residues are predicted to be accessible in the context of the assembled pilus, particularly when considering the flexibility of this region (Fig. 2d, Supplementary Fig. 3). CSPs corresponding to residues outside this contiguous surface patch can be explained by indirect conformational changes. We attempted to further structurally characterise the DNA-bound state, with special emphasis on possible changes in the structure or dynamics of the C-terminus. However, the FimTLp-DNA complex was not stable long-term and NMR signals were generally strongly weakened upon DNA binding, such that relaxation or triple resonance experiments did not yield spectra of sufficient quality. An analysis of evolutionary conservation of the FimTLp surface revealed that many of the interacting residues are also well conserved (Fig. 3c). In particular, residues of the C-terminal tail show marked sequence conservation and include a number of positively charged arginines, which are often involved in protein-DNA contacts through binding to the negatively charged DNA backbone via electrostatic interactions44.

**Interface mutations inhibit DNA binding and natural transformation in vivo.** We conducted microscale thermophoresis/temperature-related intensity change (MST/TRIC) experiments to measure the binding of labelled 12 bp dsDNA to purified FimTLp variants (Supplementary Fig. 1a), in order to further understand the nature of the FimTLp-DNA interaction and the importance of specific interface residues. First, we conducted experiments under different buffer conditions to test whether the affinity of the interaction between wild-type FimTLp and DNA is dependent on ionic strength. Indeed, when we increased the NaCl concentration from 50 mM to 150 mM, thereby raising the ionic strength, the KD increased from ~6.3 µM to ~70.1 µM (Fig. 4a). This is consistent with a non-sequence specific protein-DNA interaction, which is electrostatically driven. Furthermore, the KD determined at a NaCl concentration of 50 mM agrees very well with the affinities determined from the ITC experiments (KD ~ 2.8 µM) (Fig. 1c), as well as our NMR binding studies (KD ~ 81.1 µM) (Supplementary Fig. 5), which were all conducted in the same buffer. Next, we used MST/TRIC to test the importance of several charged residues at the DNA binding interface identified by our NMR analyses (Fig. 4b). We substituted arginine or lysine residues for glutamine in the three loop regions we identified to be important for binding. As expected, the loss of a single charged residue (e.g., K103 in the β4-β5 loop; R119 in the β5-β6 loop; R143, R146 or R148 in the C-terminal tail) only led to a small reduction in the affinity (~1.4–4 fold). However, the combined loss of two (R146/R148) or three (R143/R146/R148) charged residues next to each other on the FimTLp surface was more detrimental to binding, resulting in ~10 fold or greater reduction in affinity.

In the absence of FimT, Lp02 cells are not transformable (Fig. 1a). However, we wanted to rule out that this phenotype could result as a consequence of a pilus biogenesis defect by testing what effect our DNA binding mutations have on natural transformation in vivo (Fig. 4c). We reasoned that our substitution mutations are unlikely to affect pilus biogenesis given that they are located in a portion of the pilin not known to play a role in pilus assembly or packing. Additionally, we generated control mutations in residues of FimT not involved in DNA binding, within the same three regions (β4-β5 loop, β5-β6 loop and the C-terminal tail) previously identified to contribute to the DNA binding patch of FimT (Fig. 4d). These data show that mutations of the DNA binding residues of FimT lead to reduction or loss of transformation, whereas our FimT control mutations in residues not involved in DNA binding support wild-type levels of transformation in vivo. Mutation of a single charged residue involved in DNA binding reduces Legionella’s transformability by ~30–600 fold, whereas the double and triple mutants completely abrogate DNA uptake in our assay and thus phenocopy the effect observed upon fimT deletion (Fig. 1a). These results suggest that our FimT mutants are incorporated into the pilus fibre and that the reduction or loss of transformation efficiency is due to the diminished ability of the FimT mutants to bind to DNA in the context of a DNA uptake pilus.

**FimT of other Gram-negative bacteria also interacts with DNA.** Given that FimT, and the residues involved in DNA binding...
identified in FimT<sub>Lp</sub>, appear to be conserved, we wondered whether FimT orthologues from other bacteria are also capable of binding DNA. We expressed and purified FimT and FimU from the human pathogen <i>P. aeruginosa</i> and the plant pathogen <i>Xanthomonas campestris</i> (both γ-Proteobacteria) and performed EMSAs to assess DNA binding in vitro (Fig. 5a). Interestingly, FimT from both species binds to DNA and the affinity appears to be within the same order of magnitude as <i>L. pneumophila</i> FimT. On the other hand, FimU does not interact with DNA, except for the <i>X. campestris</i> homologue, which shows very weak binding at very high FimU concentrations. This may be due to FimUXc's slightly more basic C-terminal tail region, and/or other the presence of other residues elsewhere in the protein that allow FimUXc to weakly interact with DNA.

Since both FimT orthologues (FimTPa and FimTXc) likely share structural similarities to FimT<sub>Lp</sub>, we tested whether they are capable of restoring natural transformability in a <i>L. pneumophila</i> fimT<sub>Lp</sub> deletion strain. First, we ectopically expressed full-length FimTPa and FimTXc orthologues and performed a transformation assay. In this set-up, natural transformation was not restored in the Legionella system (Fig. 5b). Possible reasons for this lack of complementation include that the signal sequences of the orthologues are not compatible with the Legionella system and do not lead to IM localisation, that the orthologues do not interact appropriately with the Legionella T4P machinery during assembly, that the orthologues’ N-terminal helices do not allow for proper incorporation into or packing with the Legionella T4P filament, and that the orthologues do not promote DNA binding.
in vivo. To circumvent possible limitations of targeting and assembly inherent with introducing a non-native protein into another system, we ectopically expressed chimeric proteins, where we replaced only the flexible C-terminal tail region (lacking a disulphide bond) of FimTLp with the bona fide D-region of the FimT orthologues. Intriguingly, the chimeric construct, append-aged according to two binding sites with equal affinity. Error bars represent the mean ± SD (n = 3). c, Natural transformation efficiencies of parental Lp02, Lp02 ΔfimT, and the Lp02 ΔfimT strain complemented by ectopic expression of wild-type and FimT ΔmT mutants that disrupt DNA binding (c) or do not contribute to DNA binding according to Fig. 3 (d). The mean transformation efficiencies of three independent biological replicates are plotted with error bars representing the SD. < d.l., below detection limit (d.l.) (average d.l. used to calculate the mean transformation efficiency). The assay in panel c was performed in parallel to those displayed in Fig. 1a, and statistical differences were determined on log-transformed data using an unpaired two-sided t-test with Welch’s correction. The Lp02 ΔfimT strain complemented with mutants was compared to the Lp02 ΔfimT strain complemented with wild-type FimT, which was in turn compared to the parental strain. **, p < 0.01 (P<0.003; P<0.002; P<0.004); n.s., not statistically significant, p > 0.05 (P<0.38; P=0.08; P<0.89). Source data are provided as a Source Data file.
we have identified to be important for DNA binding and thus natural transformation (Fig. 3c, d). Indeed, it appears as though these DNA binding residues can be identified in proteins with as little as 18% overall amino acid sequence identity with FimT_{Lp}. Taken together, FimT homologues share an overall fold and a conserved DNA-binding motif near the C-terminus of the protein, and can be found in diverse genomic locations within diverse proteobacterial species.

**Discussion**

Natural transformation is an important mode of horizontal gene transfer with widespread consequences for bacterial evolution. Furthermore, the spread of pathogenicity traits and antibiotic resistance genes leads to the emergence of increasingly virulent and difficult to treat bacterial strains. The first step of this process involves DNA uptake mediated by T4P, which has only been studied in a handful of competent species. The minor type IV pilin FimT, but not the closely related FimU, from *A. baylyi* was previously implicated in natural transformation yet its mechanism remained obscure. Here, we characterised FimT from the naturally competent human pathogen *L. pneumophila* (*L. pneumophila*) and revealed the molecular mechanisms underlying its role in natural transformation.

We hypothesised that FimT_{Lp} is involved in DNA uptake by binding to extracellular DNA in the context of T4P and showed that *Legionella* strains lacking fimT display a marked reduction in transformation efficiency (Fig. 1a). Indeed, purified FimT_{Lp}
interacted with DNA in vitro, regardless of the nature of DNA probe tested (Fig. 1b, Supplementary Fig. 1b). Furthermore, we determined the structure of FimT<sub>Lp</sub> by NMR spectroscopy (Fig. 2) and mapped its DNA interaction surface by chemical shift perturbation experiments (Fig. 3). This binding surface consists of several positively charged residues, some of which are highly conserved, located primarily in two loop regions (the β4-β5 and β5-β6 loops) and the C-terminal tail (Fig. 3b, c). The importance of key residues for DNA binding and natural transformation was confirmed by in vitro DNA binding assays and in vivo transformation assays (Fig. 4b, c). Although our ITC experiments (Fig. 1c) indicate a 2:1 (FimT<sub>Lp</sub>:dsDNA) binding mode in which two FimT molecules engage the DNA double helix from opposite sides, we do not think this is physiologically relevant in the context of the T4P. This is because at least two FimT moieties would need to be located near each other in the pilus and oriented such that DNA could be sandwiched between them, which would be difficult to achieve given that FimT’s N-terminal helices would need to pack into the centre of the filament. Although it remains a possibility that more than one FimT, or indeed more than one DNA-binding pilin, are present within the T4P, producing an avidity effect.

Our structure of FimT<sub>Lp</sub> shares the same overall fold as the closely related T4P minor pilin FimU<sub>Pa</sub>, and the T2SS minor pseudopilins GspH<sub>Ec</sub> and EpsH<sub>VC</sub>, albeit with some key differences (Fig. 2, Supplementary Fig. 2). In place of the last β-strand (β8), part of β-sheet II in all other currently determined FimT/ GspH family structures, FimT<sub>Lp</sub> contains a conformationally flexible C-terminal tail (Supplementary Fig. 3). In our NMR studies, the heteronuclear [1,13]<sup>1</sup>H-15<Nsup>1</Nsup>NQ data and Ca chemical shifts for the C-terminal residues are indicative of a β-strand conformation, while the T<sub>2</sub> transverse relaxation times for backbone amide 1H and 15N nuclei, increased line broadening and the absence of H-bonds indicate a less well-structured conformation. A plausible interpretation of these results is that this region can exchange between a β-strand and a less-structured conformation on a millisecond timescale. The flexibility of this region is further supported by its increased proteolytic susceptibility. FimT<sub>Lp</sub>, as well as all FimT homologues from the order Legionellales, lack the D-region defining disulphide bond present in many major and minor pilins, including other FimT and FimU homologues (Fig. 5d). Therefore, it is likely that disulphide bond-containing FimT orthologues do not possess a conformationally flexible C-terminal tail. It remains to be seen whether or not the C-terminal tail of FimT<sub>Lp</sub> adopts a more structured conformation in complex with DNA and/or in the context of a fully assembled pilus. The uniqueness of FimT<sub>Lp</sub>’s C-terminal tail is further highlighted by the observation that the structure of GspH<sub>Ec</sub> was also determined in solution by NMR spectroscopy, yet it possesses a clearly defined and complete four-stranded β-sheet II region. This suggests that this region, also shared by FimU<sub>Pa</sub> and EpsH<sub>VC</sub>, is not simply a result of crystal lattice effects (Fig. 2, Supplementary Fig. 2).

FimU and GspH have been suggested to serve as adaptors linking the tip subunits to the remainder of the filament structure. Whereas minor pilins in general have been suggested to play a role in pilus priming/pilus biogenesis, the deletion of FimU, but not FimT affected pilus biogenesis in <i>P. aeruginosa</i> and <i>Pseudomonas syringae</i>. Furthermore, FimU, but not FimT of <i>P. aeruginosa</i> has been shown to play a role in bacterial twitching motility. In <i>A. baylyi</i> on the other hand, loss of either protein showed near wild-type levels of twitching, but FimT appeared to play a role in natural transformation. Orthologues of the GspH pseudopilin are critical components of the T2SS and may play a role in binding to T2SS protein substrates. To this end, the crystal structure of the <i>V. cholerae</i> orthologue EpsH revealed an extended and disordered β4-β5 loop (Supplementary Fig. 2d), which has been proposed to play a role in substrate binding. Interestingly, we have identified this same loop to contribute to FimT<sub>Lp</sub>-DNA binding (Fig. 3b). Therefore it appears that, although sharing a common evolutionary origin, FimT/GspH family proteins have become functionally diverged and specialised for the binding of different macromolecular substrates. In the case of FimT<sub>Lp</sub>, a surface patch rich in arginines enables it to function in DNA uptake during natural transformation.

The currently best-characterised DNA binding minor pilin is ComP<sup>27,28,34</sup>. While ComP homologues seem to be restricted to species of the family Neisseriaceae<sup>27</sup>, FimT homologues are present in diverse γ-Proteobacteria and some Hydrophilales (Fig. 5c). Both proteins share a conserved type IV pilin core structure, including the N-terminal helix and a four-stranded antiparallel β-sheet, but differ substantially in their C-terminal regions. In the case of ComP, this region is characterised by its so-called DD-region containing two disulphide bonds<sup>27</sup> (Supplementary Fig. 7). By contrast, FimT contains a second three-stranded antiparallel β-sheet followed by its conformationally flexible C-terminal tail and contains no disulphide bonds. In both proteins, important DNA binding residues are located near the C-terminus, which would be exposed to the solvent in the context of a fully assembled pilus. Interestingly, competent Neisseriaceae species preferentially take up DNA sequences from related species<sup>31-33</sup>. This has been attributed to ComP’s increased binding affinity towards DUS-sequences, which are DNA sequences that are highly enriched in their own genomes<sup>27</sup>. It was proposed that ComP engages DNA via an initial electrostatic attraction, followed by ComP’s α1-β1, β1-β2, DD-region binding to successive grooves of the dsDNA to achieve specificity. In contrast, no sequence selectivity has been reported for <i>L. pneumophila</i>, which is consistent with the electrostatic binding mode of FimT<sub>Lp</sub>. In addition to FimT<sub>Lp</sub> and ComP, other type IV pilins or pilin-like proteins that contribute to T4P DNA binding include ComZ and PilA4 from <i>T. thermophila</i> and <i>V. cholerae</i>. Once again, positively charged lysine and/or arginine residues likely contribute to DNA binding in all these proteins. Whereas VC0858, VC0859 and ComZ have been proposed to be located at the pilus tip<sup>15,29</sup>, ComP was initially modelled within the pilus fibre<sup>28</sup>, and later also suggested to potentially occupy a position within the pilus tip complex<sup>28</sup>. In the case of FimT, the precise localisation in the pilus remains to be determined, yet it seems plausible that it may be located within the pilus tip complex, given that its closest structural homologues (FimU, GspH and EpsH) are also proposed to occupy tip complex positions<sup>35-47</sup>. Lastly, we showed that other FimT orthologues, including FimT of the human pathogen <i>P. aeruginosa</i> and the plant pathogen <i>X. campestris</i>, are also capable of DNA binding (Fig. 5a). These experiments showed that FimT orthologues, whether they contain or lack the D-region defining disulphide bond, are capable of DNA binding. This was demonstrated even more strikingly by the FimT chimera, where the fusion of FimT<sub>Lp</sub> with FimT<sub>Pa</sub> introduced a non-native disulphide bond into the Legionella system, yet resulted in a functional protein in vivo capable of supporting natural transformation (Fig. 5b). This observation further supports the notion that perturbation of FimT’s C-terminal region does not affect pilus biogenesis. In addition, our bioinformatic analyses showed that FimT is present across a wide range of γ-Proteobacteria and that the DNA-binding C-terminal region is well-conserved on a sequence level (Fig. 5d). In particular, our alignments of high-confidence FimTs revealed a conserved GRxR motif (where x is often, but not always, a hydrophobic residue) at their C-terminus (Fig. 5d).
FimTLp, these two arginines correspond to R146 and R148, which are derived from L. pneumophila Philadelphia-1) was cultured in ACES [N-(2-acet-amido)-2-aminoethanesulfonic acid]-buffered yeast extract (AYE) liquid medium or on ACES-buffered charcoal yeast extract (CYE) solid medium, supplemented with 100 μg/mL streptomycin and 100 μg/mL thymidine. When appropriate, chloramphenicol and kanamycin were added at 5 μg/mL and 15 μg/mL, respectively. For the construction of knockout FpT2 strains, the relevant genes and 1000 bp of upstream and downstream regions were first cloned into the plsR475 suicide plasmid (derivative of pSR475). Following deletion of the gene of interest from the plasmid, the modification was introduced onto the Fp2 chromosome by triparental conjugation and subsequent selection as described previously26.

All strains were verified by colony PCR and DNA sequencing (Microsynth) and are listed in Supplementary Table 2.

Plasmids. All protein expression constructs were generated using the pOPINS or pOPINB vectors27,28, carrying an N-terminal His6-SUMO or His6 tag, respectively. Constructs for in vivo studies were generated using pMMB207C, by cloning the relevant genes downstream of the Ttac promoter. DNA fragments were amplified from L. pneumophila (RefSeq NC_002942.2) genomic DNA by PCR using CloneAmp HiFi PCR premix (Takara) and the relevant primers. For FimT and FimU orthologues from P. aeruginosa PAO1 (RefSeq NC_003512.6) and X. campesiris pv. campesiris str. ATCC 33913 (RefSeq NC_003902.1), template DNA was first synthesised (Twist Bioscience). In-Fusion cloning and site-directed mutagenesis was carried out according to the manufacturer’s guidelines (Takara). All plasmids and constructs used in this study can be found in Supplementary Table 2 and Supplementary Table 3, respectively. A summary of the gene locus tags of genes mentioned in this study from their respective genomes can be found in Supplementary Tables 5 and 6.

Protein production. Recombinant His6-SUMO-tagged proteins (FimTLP, FimUlp, FimUlpFm, FimTXc, FimUXc, FimTXc, FimUXc, and His6-tagged proteins (PilA1, PilA2) were expressed in BL21 (DE3) or Shuffle T7 E. coli cells (NEB). All constructs were N-terminally truncated to remove the transmembrane helix (a1 N) (Supplementary Table 3). Cultures were grown in Luria-Bertani (LB) media to an optical density at 600 nm (OD_{600}) of 0.6–0.8, induced using 0.5 mM IPTG and further incubated at 16 °C for 12–18 h while shaking. Cells were lysed in 50 mM Tris-HCl pH 7.2, 2 mM NaCl, 20 mM imidazole, 0.1 mg/mL lysozyme, 1 mg/mL DNAse, and one complete mini EDTA-free protease inhibitor cocktail tablet (Roche), by passing the sample three times through a pressurised cell disruptor (M110-L, Microfluidics) at 12000 psi. The clarified lyate was applied to a 5 mL HisTrap HP column (Cytiva) and HisSUMO or His-tagged plums were eluted with a linear 20–500 mM imidazole gradient. The HisSUMO- or His tag was cleaved using the catalytic domain of the human SENP1 protease or PreScission protease, respectively, while the sample was dialysed against 50 mM Tris-HCl pH 7.2, 50 mM NaCl. Protein samples were further purified by cation exchange chromatography using a 5 mL HisTrap SP HP column (Cytiva), from which plums were eluted using a linear 50–1000 mM NaCl gradient. Lastly, the pilin samples were purified by one exclusion chromatography in 50 mM Tris-HCl pH 7.2, 50 mM NaCl using a HiLoad 16/600 Superdex 75 pg column (Cytiva). Protein samples were concentrated using Amicon Ultra-15 centrifugal filters (3 kDa molecular weight cut-off, Millipore). Reducing agent (2 mM DTT) was included in the buffers for those plums with free cysteines. All purification steps were performed at 4 °C.

NMR spectroscopy
Production of isotope-labelled FimTLP. To produce uniformly labelled FimTLP, cells were grown in M9 minimal medium containing 1 g/L ^15N,15C_l-glucose and further supplemented with 3 g/L glucose (or ^13C_l-glucose for double labelled FimTLP). 2 mM MgSO4, trace elements, vitamin mix and appropriate antibiotics for selection. Each strain was grown at an OD600 of 0.5 with 15 mM IPTG and cells were harvested after 20 h at 16 °C. FimTLP was purified as described previously29.

Data acquisition and structure determination. For resonance assignments and structure determination the following spectra were recorded with a 500 MHz sample of uniformly ^15N-labelled FimTLP-28–152 in 25 mM NaCl, pH 7.2, 150 mM NaCl, and 10% D2O at 298 K in a 3 mm diameter NMR tube: 3D HNCA, 3D CBCA(CO)NH, 3D HN(C)CA, 3D HNCACB, 3D HNCBHCA, 3D HN(C)COHA, 3D HNCANHCA, and 3D HNCAHC(C)OHA spectra were recorded on a 700 MHz AVIIIHD spectrometer equipped with a TCI cryoprobe (Bruker) using the software package Topspin (versions 3.6.2 and 4.0.7). The spectra consisted of 2048 × 50 × 100 complex points in the ^H, ^13N, and ^15C dimensions with respective spectral widths of 16, 34, and 64 ppm, and were recorded with 8 scans per increment resulting in 2 and 1.5 days of measurement time, respectively. A 3D He(Calalato)-H-TOSY35 was recorded on a 600 MHz AVIIIHD spectrometer equipped with a TCI cryoprobe (Bruker). The spectrum consisted of 1536 × 100 × 150 complex points in the ^H, ^13H, and ^13C dimensions with respective spectral widths of 16, 12, and 140 ppm and was recorded with 2 scans per increment in 3 days using a recycle delay of 2 s. A time shared 3D [15N,1H,F,N]-HSQC NOESY (modified from36) was recorded on a 900 MHz AVIIIHD spectrometer equipped with a TCI cryoprobe (Bruker). The spectrum consisted of 1536 × 100 × 256 complex points in the ^H, ^15H, and ^13C dimensions with respective spectral widths of 16, 12, and 140 ppm and was recorded with 2 scans per increment in 3 days using a recycle delay of 2 s.

Resonance assignments were determined with the program CARA release 1.9.2 (www.cara.mrr.de, Keller R (2005), ETH Zurich) to 98.2% completeness. Signals in the N and H dimensions were subsequently automatically picked in the program the CcpNmr v2.5.1 software package25. Peaklists and assignments were used as input for a structure calculation with CYANA v3.9813.49 where angle constraints were automatically generated from Ca chemical shifts. Manual inspection of the automatically picked peak lists resulted in a set of 4595 picked NOE peaks of which 4220 were assigned in the CYANA calculation which yielded an average target function value of 0.21. The structures were subjectively minimised in the program Amber2065. Statistics for the resulting bundle of 20 conformers can be found in Supplementary Table 1. Additional analysis of the structural bundle after the CYANA calculation revealed 42 hydrogen bonds (each present in more than 6 structures) and the following Ramachandran statistics: 91.3%, 5.4% of residues in allowed, and additionally allowed regions, respectively. All structural figures were generated using PyMOL v2.4.1 (https://www.pymol.org).

DNA binding studies by NMR. To map the surface patch of FimTLP involved in DNA binding, chemical shift perturbation experiments were performed using 12 bp or 36 bp dsDNA fragments (Supplementary Table 4). The DNA fragments were dialysed into 50 mM Tris-HCl pH 7.2, 50 mM NaCl, 0.1% SDS, 20 mM imidazole, 1 mg/mL lysozyme, 1 mg/mL DNAase, and one complete mini EDTA-free protease inhibitor cocktail tablet (Roche), by passing the sample three times through a pressurised cell disruptor (M110-L, Microfluidics) at 12000 psi. The clarified lyate was applied to a 5 mL HisTrap HP column (Cytiva) and HisSUMO- or His-tagged plums were eluted with a linear 20–500 mM imidazole gradient. The HisSUMO- or His tag was cleaved using the catalytic domain of the human SENP1 protease or PreScission protease, respectively, while the sample was dialysed against 50 mM Tris-HCl pH 7.2, 50 mM NaCl. Protein samples were further purified by cation exchange chromatography using a 5 mL HiTrap SP HP column (Cytiva), from which plums were eluted using a linear 50–1000 mM NaCl gradient. Lastly, the pilin samples were purified by one exclusion chromatography in 50 mM Tris-HCl pH 7.2, 50 mM NaCl using a HiLoad 16/600 Superdex 75 pg column (Cytiva). Protein samples were concentrated using Amicon Ultra-15 centrifugal filters (3 kDa molecular weight cut-off, Millipore). Reducing agent (2 mM DTT) was included in the buffers for those plums with free cysteines. All purification steps were performed at 4 °C.

Electrophoretic mobility shift assay. Various DNA probes were tested for interaction with purified pilin samples using an agarose gel-based electrophoretic mobility shift assay (EMSA). Short 30 bp dsDNA fragments were generated by annealing complementary 30-mer DNA oligonucleotides. All oligonucleotides were obtained from Microsynth and are listed in Supplementary Table 4. The TpCR99a-lp12953-2958-Kan (9074 bp) plasmid, left intact or linearised by a 6.5% agarose gels containing SYBR Safe DNA stain (Invitrogen). DNA was visualised using UV illumination in a gel imaging system (Carestream).

Isothermal titration calorimetry. Isothermal titration calorimetry (ITC) experiments were carried out in duplicate on a VP-ITC microcalorimeter (MicroCal). All relevant genes downstream of the Pσ70 promoter (from 9074 bp) plasmid, left intact or linearised by a 6.5% agarose gels containing SYBR Safe DNA stain (Invitrogen). DNA was visualised using UV illumination in a gel imaging system (Carestream).
measurements were performed in 50 mM Tris-HCl pH 7.2, 50 mM NaCl buffer at 30°C. Following a pre-injection of 1 μL, titrations consisted of 19 consecutive 15 μL injections of 12 mM (0.06% final concentration) or 35 μM (0.6% final concentration) protein at 180 s or 240 s intervals, respectively. The heat of dilution, obtained by injecting DNA into buffer, was subtracted from the reaction heat, and curve fitting was performed in Origin v.7.0.383 (OriginLab) using a model assuming two binding sites of equal affinity or “one set” of binding sites.

Microscale thermophoresis/temperature-related intensity change measurements. Microscale thermophoresis (MST) experiments were conducted measuring the temperature-related intensity change (TRIC) of the fluorescence signal66. A 12 μL fluorescently labelled dsDNA probe was generated by annealing a 5’ fluorescein (FAM)-labelled and an unlabelled strand (Microsynth; Supplementary Table 4) and used in all MST/TRIC experiments. Equilibrium binding assays were performed in 50 mM Tris-HCl pH 7.2, 50 mM NaCl buffer at 30°C. Injections were repeated at 190 mM NaCl. 20, 40 and 60 nM of purified wild-type or mutant FimT samples were incubated with 100 nM of FAM-labelled 12 bp dsDNA probe at 25°C for 30 min prior to measurement. MST/TRIC measurements were performed at 20°C using a Monolight NT.115 instrument (NanoTemper) at 25% LED power and 20% MST laser power. Curve fitting was performed using the Affinity Analysis (MST) v2.0.2. For the experiment conducted with wild-type FimT1p measured at 50 mM NaCl, the data appeared slightly biphasic in nature. This suggested the presence of two binding sites with similar, yet non-identical binding affinities. When these data were fitted with a binding model assuming two non-identical binding sites (Kd1 = 100 nM, Kd2 = 1 μM) the very similar dissociation constants obtained when fitting to two identical sites (~2.9 vs 6.3 μM). All other binding experiments using other methods (ITC and NMR), as well as MST/TRIC experiments conducted with FimTpe mutants, did not reveal an obvious biphasic binding signature, which could be explained by insufficient resolution. Therefore, we chose to fit all data in the same manner, assuming two identical binding sites, to allow for their comparison. All MST/ TRIC measurements were performed at least three times. In addition, all samples were measured twice, 30 min apart, resulting in very similar binding curves and derived dissociation constants, indicating that the binding equilibrium had been attained at the time of measurement.

Transformation assay. All transformation assays were performed with the L. pneumophila Lp2953-2958 strain in liquid medium at 30°C, similar to transformation assays performed by others11,26. Strains were streaked onto CYE solid medium from frozen stocks and incubated at 37°C for 3–4 days. From this plate, bacteria were resuspended in a liquid starter culture (5 mL of AYE medium) and incubated at 37°C overnight while shaking at 200 rpm. The starter culture was diluted into a fresh 10 mL AYE medium. The culture was counted on non-selective medium. The minimum counting threshold was set at 10 colonies per plate. Transformation assays to test complementation of knockout strains of FimT were conducted against the refseq_select protein database which, for prokaryotes, contains only sequences from representative and reference genomes. The FimT and FimU gene neighbourhoods were examined using custom Biopython scripts and NCBI resources as follows 1) source proteins for each protein target were identified from the Identiﬁed Protein Groups (IPG) resource (this was necessary because many of the blast results were non-redundant entries comprising multiple identical proteins), 2) the genome region corresponding to the gene of interest and 5000 bp up- and downstream was down-loaded from the nucleotide database for one representative of each IPG (<5000 bp flanking up- and downstream sequence was available the entry was excluded from further analysis), and 3) coding sequences in the neighbouring region were extracted as multifasta and searched against the Pfam database of domain pro files using HMMER3.3.172 (hmm scan function, e-value threshold 0.0001). FimT or FimU gene sequences were classiﬁed as orphans or minor pilin locus components based on the presence of one or more of the Pfam domain domains PBC, PDX, PDX_1, and PIV in the domain region. The presence of just one of these domains was deﬁned as indicating a minor pilin locus, to account for the possibility that proteins only weakly matching the relevant Pfam domain would be missed, or that relevant proteins may be found >5000 bp away. NCBI scripts used in this study are available at Zenodo (https://doi.org/10.5281/zenodo.5920973)3. Generation of high-confidence FimT set and phylogenetic analysis. Because FimT is a GspH domain protein and shares overall structural similarity with the type IV minor pilin FimU and the TSS2 protein GspH, putative homologues from the diverse FimT/U set were filtered based on their gene neighbourhood to exclude likely fimT/U genes and generate a subset of high-confidence putative fimT/U genes for further analyses. At 100% of genes in the FimT set were located in minor pilin operons, orphan genes within the diverse FimT/U set were presumed to encode genuine FimT proteins, and these sequences were aligned along with FimT1p, FimTp, FimTAb and FimTxc, and used to generate a FimT HMM profile using HMMER72 (hmm build function). A FimT HMM profile was generated from sequence set 2 (FimU homologues), following alignment with MUSCLE v3.8.3172 and removal of entries showing >80% amino acid identity to another entry. Each sequence from the diverse FimT/U set was scanned against both the FimT and FimU sequence HMMs and reported as a likely FimT if its match score to the FimT profile was >20 points greater than its match to the FimU profile. In this way, a set of 196 putative FimT protein sequences was obtained. FimT protein sequences were aligned using MUSCLE with default (high-accuracy) settings, and the alignment was visualised and manually improved using JalView v2.11.73. The FimT alignment was processed using TrimAL v1.4.176 to remove low-quality positions and uninformative sequences (parameters: -strictplus -resoverlap 0.8 -seqoverlap 75). A maximum-likelihood phylogenetic tree of the FimT homologue was constructed using IQTree v1.6.12.77 with the substitution model LG + F + R878 and ultrafast bootstrapping79. The phylogenetic tree and associated metadata was viewed using ToIto v590. The tree was midpoint-rooted and branches with less than 50% bootstrap support removed. Gene neighborhood diagrams for selected FimT homologues were generated using Clinker v0.20.01. The FimT motif diagram was generated using WebLogo v2.8.2. Reporting summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability. NMR spectra and corresponding model coordinates have been deposited in the BioMag Resonance Data Bank (BMRB: 34704) and Protein Data Bank (PDB ID: 7QTY), respectively. Any other data that support the findings of this study are available from the corresponding author upon request. Source data are provided with this paper.

Code availability. The code used in this study has been deposited in the GitHub repository (https://github.com/francesca-short/NCBI_scripts)3. Received: 16 July 2021; Accepted: 8 February 2022; Published online: 04 March 2022

References.

1. Johnborg, O., Eldholm, V. & Hävarstein, L. S. Natural genetic transformation: prevalence, mechanisms and function. Res. Microbiol. 158, 767–778 (2007).
2. Johnston, C., Martin, B., Fichant, G., Polard, P. & Claverys, J.-P. Bacterial transformation: distribution, shared mechanisms and divergent control. Nat. Rev. Microbiol. 12, 181–196 (2014).
3. Stone, B. J. & Kwaik, Y. A. Natural competence for DNA transformation by Legionella pneumophila and its association with triph expression of type IV pil. J. Bacteriol. 181, 1395–1402 (1999).
4. Gomez-Valero, L. et al. Extensive recombination events and horizontal gene transfer shaped the Legionella pneumophila genomes. BMC Genomics 12, 536 (2011).
5. Sanchez-Busó, L., Comas, I., Jorques, G. & González-Candelas, F. Recombination drives genome evolution in outbreak-related Legionella pneumophila isolates. Nat. Genet. 46, 1205–1211 (2014).
6. Davis, S. et al. Multiple major disease-associated clones of Legionella pneumophila have emerged recently and independently. Genome Res. 26, 1555–1564 (2016).
7. Newton, H. J., Ang, D. K. Y., Driel, I. Ryan & Hartland, E. L. Molecular pathogenesis of infections caused by Legionella pneumophila. Clin. Microbiol. Rev. 23, 274–298 (2010).

8. Cunha, P. B. A., Burillo, A. & Bouza, P. E. Legionnaires’ disease. Lancet 387, 376–385 (2016).

9. Piepenbrink, K. H. DNA uptake by type IV filaments. Front. Mol. Biosci. 6, 1441–13 (2019).

10. Hepp, C. & Maier, B. Kinetics of DNA uptake during transformation provide a basis for natural competence and switching motility in pilinated Neisseria gonorrhoeae. Mol. Microbiol. 209, 301–308 (2020).

11. Jacobsen, T., Bardiaux, B., Francetic, O., Izadi-Pruneyre, N. & Nilges, M. Structural and function of minor pilins of type IV pili. Med. Microbiol. Immunol. 209, 301–308 (2020).

12. Vessey, J.-L. & Pelicic, V. Exceptionally widespread nanomachines composed of type IV pilins: the prokaryotic swiss army knives. Structure and function of minor pilins of type IV pili. Med. Microbiol. Immunol. 209, 301–308 (2020).

13. Dubnau, D. & Blokesch, M. Mechanisms of DNA uptake by naturally competent bacteria. Annu. Rev. Genet. 53, 217–237 (2019).

14. Bergé, M., Mortier-Barrière, I., Martin, B. & Claverys, J.-P. Transformation of periplasm in naturally transformable Vibrio cholerae. J. Bacteriol. 195, 301–308 (2013).

15. Shaikh, E. J., van et al. DNA Binding: a Novel Function of Type II Secretome Reveals Unique Exoproteins and a Chitinase That for DNA Binding During Transformation of Competent Myxococcus xanthus. J. Bacteriol. 195, 55–62 (2013).

16. Seitz, P. & Blokesch, M. DNA-uptake machinery of naturally competent Vibrio cholerae. Proc. Natl Acad. Sci. USA 110, 17987–17993 (2013).

17. Weaver, S. J. et al. CryoEM structure of the type IVa pilus secretin required for Type II Secretion System. Nat. Struct. Mol. Biol. 24, 385–392 (2018).

18. Seitz, P. et al. ComEA is essential for the transfer of external DNA into the periplasm in naturally transformable Vibrio cholerae cells. PLoS Genet. 10, e1004066–14 (2014).

19. Hepp, C. & Maier, B. Kinetics of DNA uptake during transformation provide evidence for a translocation ratchet mechanism. Proc. Natl Acad. Sci. USA 113, 12467–12472 (2016).

20. Draskovic, I. & Dubnau, D. Biogenesis of a putative channel protein, ComEC, required for DNA uptake: membrane topology, oligomerization and formation of disulfide bonds. Mol. Microbiol. 55, 881–896 (2004).

21. Attiaich, L. et al. Role of the single-stranded DNA-binding protein SsbB in Pneumococcal transformation: maintenance of a reservoir for genetic plasticity. PLoS Genet. 7, e1002156 (2011).

22. Bergé, M., Mortier-Barrière, I., Martin, B. & Claverys, J.-P. Transformation of Periplasm in Naturally Transformable Vibrio cholerae Reveals a Long Flexible Loop. Biochimica et. Biophysica Acta (BBA) - Proteins Proteom. 1844, 406–415 (2015).

23. Corina, R. I. & Gou, J. J. Synthetic analysis of structural determinants for protein-DNA-binding specificity. Proteins 84, 1147–1161 (2016).

24. Douzi, B. et al. The XcpV/GspL Pseudopilin has a central role in the assembly of a quaternary complex within the T2SS Pseudopilus*. J. Biol. Chem. 284, 34580–34589 (2009).

25. Konoto, K. V. & Sandkvist, M. Architecture, Function, and Substrates of the Type II Secretion System. Ecosal 8, (2019).

26. Debruyne, J. et al. The Pili of Myxococcus xanthus. Nat. Commun. 11, 5054 (2020).

27. Belete, B., Hu, L. & Wozniak, D. J. Pseudomonas aeruginosa AlgR regulates type IV Pilus biosynthesis by activating transcription of the fimU-fimVWWXY12E operon. J. Bacteriol. 190, 2032–2030 (2008).

28. Douzi, B., Ball, G., Cambillau, C., Tegoni, M. & Voulhoux, R. Deciphering the Xcp Pseudopilus as a Type II secretion machinery through multiple interactions with substrates*. J. Biol. Chem. 286, 40792–40801 (2011).

29. Denise, R., A. & Rocha, E. P. C. The evolution of protein secretion systems by co-option and tinkering of cellular machineries. Trends Microbiol. 28, 372–386 (2020).

30. Debruyne, J. et al. Pilin Assembly in Type IVa Pili are Highly Dynamic Structures that Retract to Promote DNA uptake. Microbiol. 2016 (2020) https://doi.org/10.1111/mmi.14718.

31. Chung, Y. S. & Dubnau, D. All Seven comG open reading frames are required for DNA binding during transformation of competent Bacillus subtilis. J. Bacteriol. 180, 41–45 (1998).

32. Merriam, J. J., Mathur, M., Maxfield-Boumill, R. & Isberg, R. R. Analysis of the Legionella pneumophila filA Gene: Intracellular Growth of a Defined Mutant Defective for Flagellum Biosynthesis. Infect. Immun. 65, 2497–2501 (1997).

33. Zuckman, D. M., Hung, J. B. & Roy, C. R. Pore-forming activity is not sufficient for Legionella pneumophila phagosome trafficking and intracellular growth. Mol. Microbiol. 32, 990–1001 (1999).

34. Roy, C. R. & Isberg, R. R. Topology of Legionella pneumophila DotA: an inner membrane protein required for replication in macrophages. Infect. Immun. 65, 571–578 (1997).

35. Berrow, N. J. et al. A versatile ligation-independent cloning method suitable for high-throughput expression screening applications. Nucleic Acids Res. 35, 645–645 (2007).

36. Schau, G. et al. Expression, purification and crystallization of a lyssavirus matrix (Mr) protein. Acta Crystalllography Sect. F. Struct. Biol. Cryst. Commun. 64, 258–262 (2008).

37. Chen, J. et al. Legionella effectors that promote nonlytic release from protozoa. Science 303, 1358–1361 (2004).

38. Muhandiram, D. & Kay, L. E. Gradient-enhanced triple-resonance three-dimensional nmr experiments with improved sensitivity. J. Magn. Reson. Ser. B 103, 203–216 (1994).

39. Kovacs, H. & Gossert, A. Improved NMR experiments with 13C-isotopic mixing for assignment of aromatic and aliphatic side chains in labeled proteins. J. Biomol. Nmr 58, 101–112 (2014).
Acknowledgements

This work was funded by an SNF PRIMA grant PR00P3_179728 to M.K.H. F.L.S. is supported by an Australian Research Council Discovery Early Career Research Award DE200101524. We would like to thank G. Waksman and A. Meir for the Lp02, CR019 and DH5a apr strains, and the pSR47S plasmid. We would also like to thank H. Hilbi for the pMMB207C plasmid. We are grateful to R. Glockshuber, E. Weber-Ban, and M. Pihler for helpful discussions and sharing of reagents and instruments, and to J. Scheuermann for the use of the VP-ITC instrument.

Author contributions

SAGB cloned constructs, created Legionella strains, purified proteins, performed DNA binding studies, transformation assays, Western blots, and analysed results. FLS designed and performed all bioinformatic analyses. SH constructed FimT chimera constructs and performed the corresponding transformation assays. MJMS purified proteins and performed ITC experiments. ADG performed and analysed all NMR-related experiments with help from SAGB. MKH designed and supervised the study, made figures and wrote the manuscript with help from all authors.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information

The online version contains supplementary material available at https://doi.org/10.1038/s41467-022-28690-1.

Correspondence and requests for materials should be addressed to Manuela K. Hospenthal.

Peer review information

Nature Communications thanks Jeremy Derrick, Kurt Piepenbrink and the other, anonymous, reviewer for their contribution to the peer review of this work. Peer reviewer reports are available.

Reprints and permission information

is available at http://www.nature.com/reprints

Publisher’s note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s) 2022

Open Access

This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2022