Pili are essential for conjugation also in many Gram-positive bacteria

Josy ter Beek[1,2], Dennis Svedberg[3], Kieran Deane-Alder[1,2], André Mateus[3,4] and Ronnie P.-A. Berntsson[1,2]

[1] Department of Medical Biochemistry and Biophysics, Umeå University, SE-90187 Umeå, Sweden
[2] Wallenberg Centre for Molecular Medicine & Umeå Centre for Microbial Research, Umeå University, Umeå, Sweden
[3] Department of Chemistry, Umeå University, Umeå, Sweden.
[4] The Laboratory for Molecular Infection Medicine Sweden (MIMS) & Umeå Centre for Microbial Research, Umeå

Correspondence should be addressed to J.t.B (email: josy.beek@umu.se) or R.P-A.B. (email: ronnie.berntsson@umu.se)

Abstract

Type IV secretion systems (T4SS) enable the spread of antibiotic resistance and other virulence factors. In Gram-positive bacteria, T4SSs have long been thought to lack VirB2-like proteins that form conjugative pili and instead rely on adhesins for cell-cell contacts. Yet, it has remained unclear how subsequent DNA transfer (conjugation) is physically mediated. Here we identify a VirB2-like protein, $PrgF_{B2}$, from the clinically isolated conjugative plasmid pCF10 in *Enteroccocus faecalis* and show that it is essential for conjugation. Structural modeling confidently predicts a pilus-like assembly for $PrgF_{B2}$. We validate this prediction through mutagenesis, conjugation assays, and targeted chemical labeling. By combining various machine learning bioinformatic techniques, we analysed >1000 Gram-positive conjugative plasmids from diverse species, including major pathogens, and identified pili forming VirB2-like proteins in almost all of them. Our findings overturn the prevailing view that Gram-positive T4SSs lack pili. This discovery provides a new framework for understanding horizontal gene transfer and highlights critical targets for combating antimicrobial resistance and virulence in Gram-positive bacteria.

Main

Type 4 Secretion Systems (T4SS) are megadalton-sized membrane protein complexes that facilitate the transfer of proteins and/or DNA from bacterial donor cells to recipient cells. Conjugative T4SSs are main drivers for horizontal DNA transfer in bacteria and are of major interest as they spread antibiotic resistance and other virulence factors. T4SSs have been studied for decades, and multiple labs have provided structural and functional models for these systems, but the focus has been on a few model systems from Gram-negative (G-) bacteria[1].

In conjugative T4SSs, the mobile DNA, often encoded on a plasmid, is recognized by its *oriT* sequence. A specific relaxase then nicks the DNA, makes it single stranded and becomes covalently attached to the 5' end. With the help of the coupling protein (T4CP), this protein:ssDNA substrate is then translocated through the T4SS channel. In G- T4SSs the transport channel comprises an inner membrane complex, an outer membrane complex, a connecting stalk and a conjugative pilus that extends from the cell. Pili of G- T4SSs are helical polymers of VirB2-like proteins, often capped by a VirB5 pentamer that facilitates adhesion to recipient cells. Although it has been debated whether these pili form the transport channel for the protein:ssDNA substrate, this has been shown for the conjugative T4SSs from the F and pED208 plasmids from *E. coli* [2–4] .

Pili forming VirB2-like proteins are essential in G- T4SSs [2,5]. Despite high sequence divergence, all characterized mature VirB2 homologs are small membrane proteins with two transmembrane helices, connected by a short and charged cytoplasmic loop [6–8]. Pili assembly is presumed to occur from a pool of these membrane-inserted VirB2 proteins that oligomerize onto the central VirB6 pentamer, from the T4SS inner membrane complex, and are thereby extruded from the membrane (Fig. 1A) [1,9]. In all available structures of assembled conjugative pili, the charged loop is facing the lumen, while the N and C-terminal ends are on the outside of the tubular structure. The charge of the lumen is often modulated by specifically bound phospholipids [6,10–12].

We know much less of Gram-positive (G+) T4SSs, but they are consistently described to lack conjugative pili [13–18]. It is known that some G+ T4SSs instead rely on cell-wall anchored adhesins for cell-to-cell contacts that enhance mating pair formation [19–21]. These adhesins are important for conjugation to occur but they are not essential and they do not form a channel for substrate transfer [19,21,22]. Instead, different studies have suggested that such a channel could be formed either by the VirB1-like cell-wall hydrolases or the VirB8-like proteins in G+ T4SSs [14,16,23]. However, all these proteins are membrane anchored and it is unclear how they could extend far enough to form a transfer pathway [24]. How G+ T4SSs facilitate the transfer of the protein:ssDNA substrate through the cell-walls of both donor and recipient cells, which in *e.g.* enterococci equals roughly 80 nm, and then through the membrane of the recipient cell has therefore remained an unanswered question.

Despite all reports that G+ T4SS lack pili, we wondered whether they could have been overlooked, since identification of VirB2-like proteins is difficult as *virB2* genes are small and poorly conserved [1].

Model of a Gram-positive conjugative pilus

We previously modelled each predicted protein of the T4SS encoding *prgQ* operon from the pCF10 plasmid in *Enterococcus faecalis* [25]. One of these modelled proteins was PrgF, a small uncharacterized protein that is predicted to have two transmembrane helices connected by a short, charged loop. Its predicted structure bears some resemblance to the determined structures of mature VirB2-like proteins from G- bacteria (Fig. 1B). AlphaFold modelling of oligomeric structures of up to 50 copies of PrgF only yielded low-confidence structures (Table S1). However, when modelling 5 or more PrgF subunits together with the pentameric VirB6-like protein from pCF10, $PrgH_{B6}$, AlphaFold yielded high-confidence predictions (Fig. 1C and Fig. S1). These models have ipTM scores of up to 0.87 (Table S1) and resemble pilus-like structures with a helical architecture analogous to that observed in G- T4SSs (Fig. 1D). Based on these findings we hypothesized that PrgF might function as a VirB2-like protein and will therefore name it $PrgF_{B2}$.

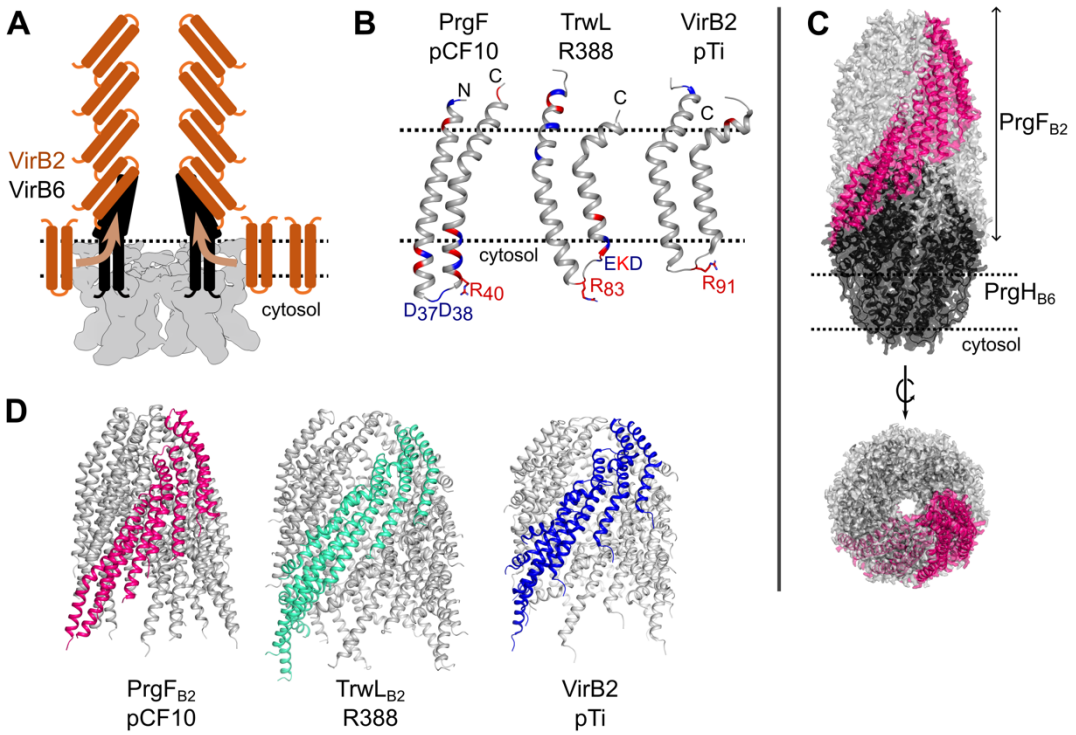


Figure 1. Overview of VirB2-like pili forming proteins. A) Cartoon of VirB2-like proteins visualizing their membrane insertion and assembly into a pilus (slice-through) on top of VirB6, a central protein of the inner-membrane complex. Figure adapted from Paillard *et al,* 2025 [1]. B) AlphaFold model of PrgF alongside characterized VirB2-like proteins from R388 (PDB ID: 8S6H) and pTi (PDB ID: 8CUE). Dotted lines indicate the approximate location of the membrane. C) Overview of the $PrgF_{B2}$-$PrgH_{B6}$ AlphaFold model, with 5 $PrgH_{B6}$ in black, 4 $PrgF_{B2}$ from one helical rise in pink and the remaining 16 $PrgF_{B2}$ in grey. The dotted lines indicate the approximate location of the membrane. D) Comparison of the $PrgF_{B2}$ pili model (as in C) with the determined structures of the pili from R388 and pTi (PDB codes as in B).

$PrgF_{B2}$ is essential for conjugation

To determine whether $PrgF_{B2}$ is important for T4SS function, we deleted this gene from pCF10 in *E. faecalis* OG1RF, which resulted in a complete loss of conjugation efficiency (Fig. 2A and Fig. S2). Notably, conjugation was fully restored to wild-type levels upon complementation with heterologous expressed $PrgF_{B2}$ (Fig. 2A, blue bar). These findings demonstrate that, consistent with its role as a VirB2-like protein, $PrgF_{B2}$ is essential for conjugation.

The AlphaFold model of the pilus suggests that D37, D38 and R40 from the charged loop of $PrgF_{B2}$ line the lumen (Fig. 2B). In the structures from G- conjugative pili, the charged loop also lines the lumen with a bound phospholipid that modulates the overall surface charge of the pili cavity [6]. The surface charge of the $PrgF_{B2}$-$PrgH_{B6}$ AlphaFold model indicates that the initial cavity, formed by $PrgH_{B6}$, has an overall positive surface charge, while the lumen of the $PrgF_{B2}$ pilus is mainly negatively charged (Fig. 2C). In line with having an important functional role, we found that all modifications of charged residues in the $PrgF_{B2}$ loop, except for D38C, strongly affected conjugation (Fig. 2A). $PrgF_{B2}$ variants where the charge of the R40 was reversed, or where both aspartates were exchanged for asparagine, could no longer support conjugation.

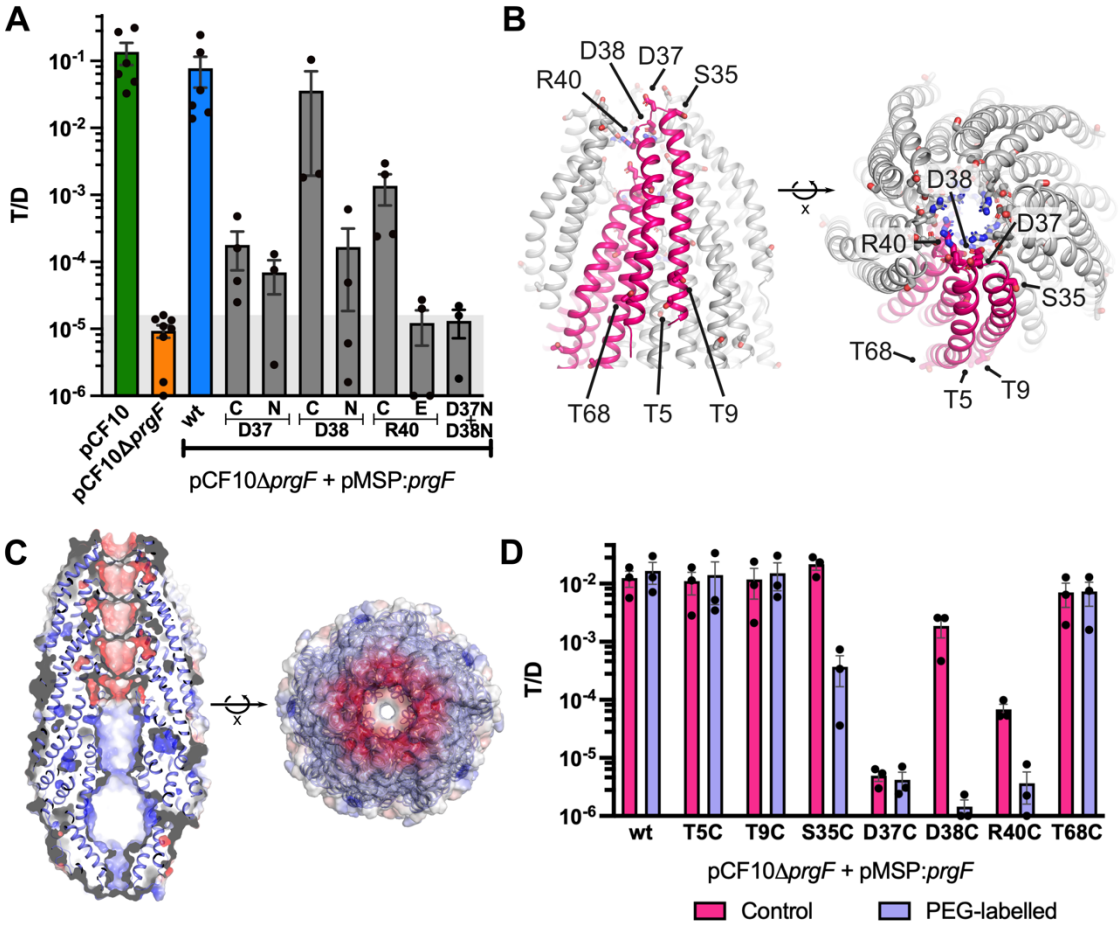


Figure 2. $PrgF_{B2}$ is essential for conjugation and as in other VirB2-like proteins its charged loop is important for function. A) The conjugation efficiency of wild-type pCF10 (first bar, in green) is completely reduced to background (indicated by grey area, see Fig. S2) when *prgF* is knocked-out (pCF10Δ*prgF*, second bar, in orange). Conjugation is restored when wild-type PrgF is expressed from a pMSP vector (third bar, in blue). The grey bars show pCF10Δ*prgF* with various PrgF variants with substituted loop residues expressed from the same pMSP vector. B) AlphaFold model of the $PrgF_{B2}$ pili (same colours as in Fig 1C-D) with substituted residues shown as sticks and for the upper pink subunit their locations are indicated. C) Surface charge of the $PrgF_{B2}$-$PrgH_{B6}$ model (same as in Fig. 1C), showing a slice-through to visualize the lumen of the pili and a top (extracellular) view, coloured from red (-5 kT/e) to blue (+5 kT/e). D) Conjugation efficiency of $PrgF_{B2}$ variants or wild type (wt) expressed from the pMSP vector in the pCF10Δ*prgF* background without (control) or with maleimide-PEG_{11}-biotin labelling.

PEG-labelling of residues predicted to point towards the pili lumen blocks conjugation

To further verify the AlphaFold model, we identified residues on the inside and outside of the proposed pili-structure (Fig. 2B), exchanged them individually for cysteines and used them for PEG-maleimide labelling after T4SS assembly [26]. We reasoned that the conjugation efficiency should be decreased if steric hindrances were introduced inside the lumen of the proposed pili. To only probe transport through labelled T4SS complexes, T4SS expression was artificially induced before labelling and mating time was reduced from 3.5 hours to just 15 minutes, leading to ca. 10-fold lower conjugation efficiency (compare Fig. 2A & D).

Residues that are predicted to be on the outside of the pili in the AlphaFold model, T5C, T9C, and T68C, were unaffected by maleimide-PEG_{11}-biotin labelling. S35C, predicted to be in a protein-protein interface area of the pilus was affected by labelling, but conjugation remained above background. In contrast, the residues situated in the charged loop that lines the pilus lumen in the model, D38C and R40C, showed a complete loss of conjugation efficiency after labelling (Fig. 2D). The D37C variant has a very low conjugation efficiency from the start (Fig. 2A &D), and it is therefore difficult to determine if it is affected by labelling. As in G- VirB2-like proteins, the charged loop is predicted to face the cytoplasm when $PrgF_{B2}$ is inserted in the membrane on its own (Fig. 1B) and is therefore predicted to only become accessible for labelling if the $PrgF_{B2}$ is indeed extruded from the membrane (Fig. 1A). Therefore, these results strongly support the hypothesis that $PrgF_{B2}$ forms a pilus with the charged loop residues lining the pilus lumen.

VirB2-like proteins are prevalent in many clinically relevant G+ T4SSs

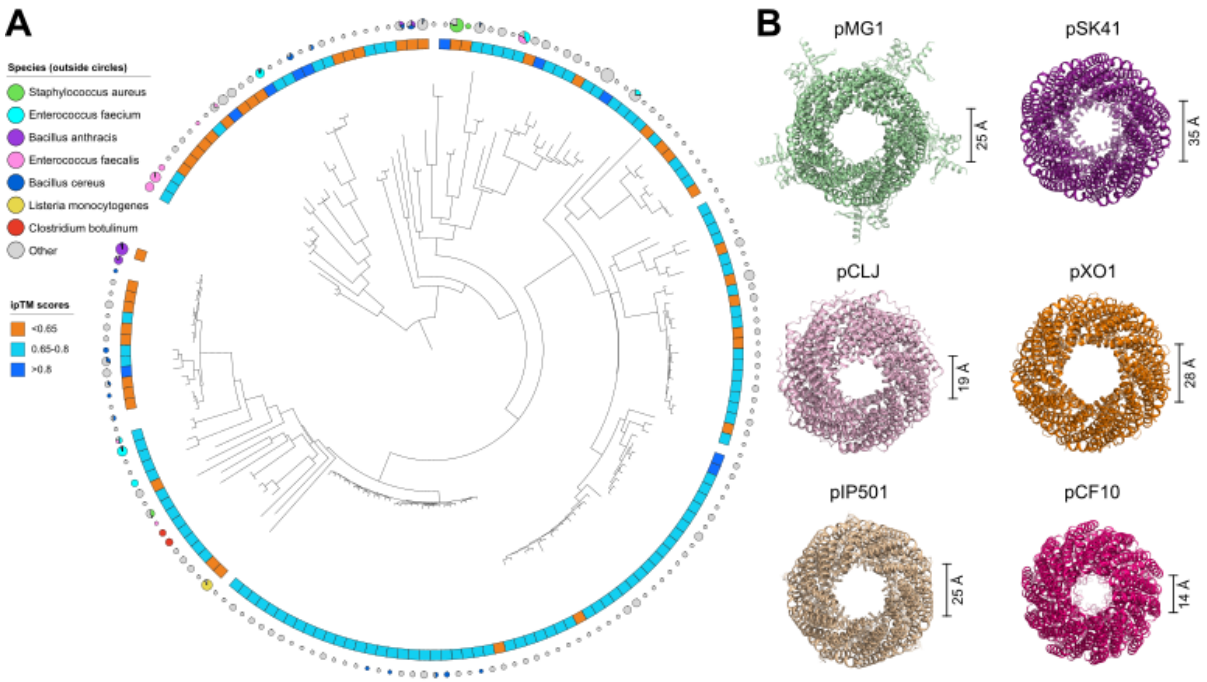
While assessing the function of $PrgF_{B2}$, we wondered how widespread VirB2-like proteins are in G+ T4SSs. To assess this, we retrieved all annotated G+ conjugative plasmids from the plasmid database (PLSDB) and clustered all of them (1103 plasmids) based on their predicted VirB6-like protein. This resulted in 174 clusters and for each a representative plasmid was selected, originating from a wide variety of G+ bacteria, to limit computation time (Fig. 3A). We then searched all open reading frames of the 174 representative plasmids, to identify small proteins (50-150 amino acids) with 2 predicted transmembrane helices and a topology similar to $PrgF_{B2}$ from pCF10. Potential VirB2-like proteins were found for 95% of the plasmids and these were encoded in close genetic proximity to the corresponding *virB6* gene (Extended data file 1, Fig. S3). Each VirB2 candidate was analysed by AlphaFold modelling, using 10 copies of the VirB2-like protein together with 5 copies of the respective VirB6-like protein. In 91 plasmids (ca 52%) we identified multiple VirB2-like candidates. However, upon visual inspection of 16 example

177 plasmids only one of each clearly modelled as a pilus and these had ipTM scores clearly
178 separated from models with other small transmembrane proteins (Fig. S4). For 71% of
179 our reference plasmids (123/174), this resulted in reliable (ipTM >0.65) pilus-like models
180 with the VirB2-like proteins assembled with their pentameric VirB6-like protein (Fig. 3A,
181 Extended data file 1).

182 Noteworthy examples of G+ T4SSs for which AlphaFold could confidently model a pilus-
183 like structure include plasmids from the important clinical pathogens *Enterococcus*
184 *faecium* (pMG1), *Staphylococcus aureus* (pSK41), *Clostridium botulinum* (in the
185 neurotoxin gene cluster on pCLJ) and *Bacillus anthracis* (on the anthrax capsule
186 plasmids pXO1/pXO2), as well as pIP501, one of the few other characterized G+ T4SSs
187 [27,28] (Fig. 3B). We measured the diameter of the pili lumen formed by their VirB2-like
188 protein in these models and although the measurements should be regarded as
189 approximates, they range from ca 14 Å in pCF10 up to ca 35 Å in pSK41(Fig. 3B). The
190 modelled surface charge of the pili lumen in these models vary, with pMG1, pCLJ and
191 pXO1 being predominantly positive, while pSK41 and pIP501 have an overall negative
192 surface charge (Fig. S5).

193 Overall, this bioinformatics analysis shows that VirB2-like proteins are present in the vast
194 majority of G+ T4SSs and strongly indicate that they form a pilus-like structure together
195 with their corresponding VirB6.

196



197

198 **Figure 3**. Bioinformatics analysis of VirB2-like proteins in G+ T4SSs. A) Phylogenetic tree
199 of the VirB6-like proteins of G+ T4SSs, showing the 174 reference classes. Each coloured

box indicates the ipTM score (orange <0.65, light blue 0.65-0.8, dark blue >0.8) when 10 VirB2- and 5 VirB6-like proteins from the reference plasmid are modelled using AlphaFold. The size of the outermost circles indicates how many plasmids are in each representative class and the colour highlights the proportion of a selection of clinically relevant strains. B) Top views of selected VirB2-VirB6-like protein complexes from *E. faecium* (pMG1, ipTM: 0.74), *S. aureus* (pSK41, ipTM: 0.77), *C. botulinum* (pCLJ, ipTM:0.75), *B. anthracis* (pXO1, ipTM: 0.75), *E. faecalis* pIP501 (ipTM: 0.74) and pCF10 (ipTM: 0.85).

Discussion

Gram-positive T4SSs have long been described to lack conjugative pili. Competence pili for the uptake of DNA do exist in G+ bacteria[29,30], but these are not found in T4SS operons and differ greatly from conjugative VirB2-like pili. Since no previously described G+ T4SS component could form a long enough scaffold to cross the cell-wall of even just the donor cells, we wondered if VirB2-like pili forming proteins could have been overlooked. AlphaFold provided high-confidence models of $PrgF_{B2}$, from pCF10, forming a pili-like structure together with $PrgH_{B6}$. We confirmed this pilus model by combining mutagenesis and *in vivo* assays and showed that $PrgF_{B2}$ is essential for conjugation. As far as we are aware this is the first VirB2-like protein identified in a Gram-positive T4SS, but by using a combination of recently developed bioinformatics software, we identified VirB2-like proteins in the majority of G+ T4SSs. We propose that they provide the missing substrate passageway through the thick cell-wall of G+ bacteria, by assembly with their VirB6-like protein, as has been postulated for G- T4SS, such as the characterized R388 system [1].

VirB2-like proteins in conjugative pili from G- systems have been shown to bind specific lipids, usually phosphatidyl glycerol [5,31]. These lipids line the lumen and modulate the generally positive surface charge to moderately negative to facilitate efficient DNA transfer through the pili. In our modelled G+ $PrgF_{B2}$ pilus the lumen is already negatively charged (Fig. 2C), as is the lumen of the modelled pIP501 and pSK41 (Fig. S5). This suggests that no lipid binding is necessary for these examples, similar to the G- RP4 pilus [32]. However, other modelled pili have an overall positive lumen, so they might bind a lipid, but this remains to be determined.

In our AlphaFold models, the diameters of the modelled pili lumen varies from 14 up to 35 Å (Fig. 3B). Although these numbers should be seen as approximations, they are in line with pili structures from G- conjugative plasmids, many of which have a diameter of around 25 Å [33]. In all cases, the pore sizes of the channels are large enough to support ssDNA transfer, which has been shown to need a diameter of at least 10 Å [34].

We searched but could not identify any VirB5-like protein to act as a pilus tip in pCF10. This does not conclusively prove that no VirB5-like proteins exist, but it is worth noting that there also are G- systems without a pili tip, such as the T4SSs from the F-plasmid from *E. coli,* and from Ptl from *B. pertussis*, GGI from *N. gonorrhoeae* and ComB from *H. pylori* ComB [35–37].

To our knowledge, no conjugative pilus has been visualized using electron microscopy on T4SS expressing G+ bacteria. Possibly, the thick and heavily crosslinked cell wall obscures pili-like features and/or pili extension occurs only transiently, when a mating pair has formed and does not extend far outside the cell. This would be reminiscent of G- systems with dynamic pili, such as the F-plasmid, when it is in its retracted state [38].

T4SSs and VirB2-like proteins are present in numerous plasmids from a wide range of species, including those prominent in nosocomial infections, such as *E. faecium*. Moreover, we also identified them in plasmids carrying genes for the anthrax toxin and botulinum neurotoxin, both of which are of significant concern for healthcare.

To conclude, we have here shown that G+ conjugative T4SSs depend on VirB2-like proteins that form pili. Due to almost non-existing sequence homology, these proteins were overlooked until now but using newly developed bioinformatics tools we could identify them. There are still many questions that remain unanswered, including how pilus assembly works, but our findings open completely new ways to study the spread of both virulence factors and antibiotic resistance in Gram-positive bacteria.

Materials and methods

In silico analysis of PrgF and PrgH

AlphaFold 3 modelling for PrgH and PrgF was done on the AlphaFold server [39]. The topology of PrgF was predicted using DeepTMHMM (v1.0.24) [40]. Surface charge calculations of the $PrgF_{B2}$-$PrgH_{B6}$ pili model were done in PyMOL (v3.1.5.1), using the APBS plugin (v3.4.1) [41].

Construction of OG1RF pCF10ΔprgF

The construction of *E. faecalis* OG1RF:pCF10Δ*prgF* strain was carried out with allelic exchange and counterselection aided by a pCJK218-derived plasmid [42]. The first 5 and last 8 amino acids of PrgF were kept intact. 900 bp upstream and downstream regions of *prgF* were PCR amplified from pCF10 with primer pairs of KOprgF_UF_fwd_BamHI, KOprgF_UF_rev_XbaI and KOprgF_DF_fwd_XbaI, KOprgF_DF_rev_NcoI respectively (Table S2). The PCR products were digested with either BamHI/XbaI or XbaI/NcoI and ligated into BamHI/NcoI/XbaI digested pCJK218. For the rest of the procedure the protocol from Jäger et al. [43] was followed.

Construction of pMSP:*prgF* (wild type and mutants)

prgF was PCR amplified from pCF10 with primer pairs pMSP_prgF_fwd_NcoI and pMSP_prgF_rev_XbaI (Table S2). The PCR product was digested with NcoI/XbaI and ligated into the NcoI/XbaI digested pMSP3545S:MCS vector [22]. After transformation and selection in *E. coli* Top10 cells, the obtained pMSP:*prgF* vector was sequenced and used as a template for iPCR reactions to construct pMSP:prgF_D37C, pMSP:prgF_D37N, pMSP:prgF_D38C, pMSP:prgF_D38N, pMSP:prgF_R40C, pMSP:prgF_R40E, pMSP:prgF_D37NandD38N (see Table S2 for used primers).prgF_T5C, prgF_T9C, prgF_S35C and prgF_T68C were instead ordered synthetically with a NcoI site around the start codon and a BamHI site after the stop codon and digested and ligated into the pMSP3545S:MCS (see Table S2 for further details).

All obtained constructs were sequenced and transformed to OG1RF:pCF10Δ*prgF* according to the procedure described in Sun et al [22]. Cysteine mutations, T5C, T9C, S35C, T68C were designed to be mild mutations accessible for labelling, based on the AlphaFold model and NetSurP 3.0 [44] scores as described in Ellison et al. [26].

Conjugation assays

Donor (OG1RF pCF10 and derivatives thereof) and recipient (OG1ES) cells were inoculated from glycerol stock and grown overnight (~16 hrs) at 37°C and 180 rpm in BHI with 10 µg/mL tetracycline (to select for pCF10) and 100 µg/mL erythromycin for donors

292 with pMSP or just with 20 µg/mL erythromycin for the recipient cells (see strains and
293 plasmid list Table S2). The following day the cultures were refreshed 1:1000 (donors) or
294 1:500 (recipient) in BHI without antibiotics and incubated for 2.5 hrs at 37°C and 180 rpm.
295 Afterwards, 1 mL of each donor culturewas mixed with 1 mL of the recipient cells. For
296 conjugation assays with pMSP containing donors, nisin (Sigma) was added at a final
297 concentration of 50 ng/ml at this mixing step. Conjugation mixtures were then incubated
298 for 2.5 hrs at 37°C and 105 rpm. After this incubation the mixtures were serially diluted in
299 PBS and 10 µL drops of undiluted to 10^5 times diluted cultures were plated out in
300 triplicates on BHI agar plates supplemented with 10 µg/mL tetracycline and 25 µg/mL
301 fusidic acid (to select for donor cells) or 10 µg/mL tetracycline, 20 µg/mL erythromycin
302 and 1 mg/mL streptomycin (to select for transconjugants). Plates were incubated at 37°C
303 for 48 hr, counted and enumerated for colony-forming units (CFU). The plasmid transfer
304 rate was determined as CFU of transconjugant over CFU of donor (T_c/ D). To determine
305 the background level of our OG1RF and OG1ES strains under the tested conditions,
306 OG1RF without pCF10 was also used as a "donor", which CFU of "donor" were based on
307 BHI agar plates with only fusidic acid.

308 To be able to label PrgF before conjugation the protocol was adjusted slightly. The
309 overnight cultures of donors and recipient were refreshed 1:100 to get enough cells for
310 labelling and incubated for 2 hrs at 37°C and 180 rpm. Then both cCF10 and nisin were
311 added to the donor cultures at a final concentration of 10 ng/mL and 50 ng/mL
312 respectively and the agitation was reduced to 105 rpm for all cultures. After 1 hour of
313 induction, 500 µL aliquots for each donor culture were spun down at 16,000 rcf for 1
314 minute. Cell pellets were washed in 1 mL PBS, spun down again and resuspended in 250
315 µL PBS. Then two 100 uL aliquots were taken out to which either 1 µL 125 mM maleimide-
316 PEG11-biotin (EZ-link) in DMSO was added (for labelling) or 1 µL DMSO (for control). All
317 aliquots were mixed by vortexing and incubated for 10 minutes at 25°C. Afterwards 200
318 µL of the recipient cell culture (in BHI) was added to each of the (labelled or control)
319 donor fractions and mixed by vortexing. These conjugation mixtures were incubated
320 statically at 37°C for just 15 minutes to only assay mating via the preformed channels.
321 The rest of the procedure was done as described above.

322 All conjugation data shown in figure 1 is done with strains with a pMSP plasmid, either
323 empty (first two bars of Fig. 1A) or with *prgF* (all other bars as indicated) and with nisin
324 induction. The conjugation data shown in figure S2 is from strains without pMSP vector
325 and without the addition of nisin. All conjugation results are shown as bars of the mean,
326 with error bars showing the standard error of the mean of at least three independent
327 experiments, that are also shown as individual datapoints (black dots). To include data
328 with no observed transconjugants in the analysis, these were given a very small score:
329 10^{-6} (Fig. 1A and Fig. S2) or 10^{-7} for labelling experiments (Fig. 1D). Grey area indicates the
330 background Tc/D value as calculated from conjugation experiments of OG1RF without

pCF10 (Fig. S2). Data visualization and analysis was done with GraphPad Prism (version 10.6, GraphPad software).

Bioinformatics analysis of G+ plasmids

Plasmid sequences from 72,360 plasmids were retrieved from PLSDB (v2024_05_31_v2) [45]. Plasmids annotated as conjugative and from G+ origin, as identified by PLSDB, were selected creating a set of 1,787 plasmids for further analysis. All proteins from these plasmids were clustered at 90% identity and 90% coverage using MMseqs2 (v18-8cc5c) [46] resulting in 28,172 protein clusters. Protein structures for these protein representatives were predicted using ESMfold (v1.0.3, ESM-2-3B), and transmembrane domains and signal peptides were predicted using DeepTMHMM (v1.0.24) [40,47].

Proteins with predicted structures similar to $PrgH_{B6}$ were identified using Foldseek (v10-941cd33) [48], resulting in 1,113 proteins originating from 174 protein clusters. VirB6-like protein candidates were then modelled using AlphaFold 3 (v3.0.1) [39]. For identifying potential VirB2-like proteins we searched for proteins with a length between 50-150 amino acids and with a predicted DeepTMHMM topology of (S)OMIMO:with/without Signal peptide, Outside, Membrane, Inside, Membrane, Outside. Proteins with a signal peptide were further analysed without this part of the sequence. From each of the 174 representative plasmids, all VirB2-like candidates were predicted using AF3 with their corresponding VirB6-like protein, all modelled with 10 VirB2- and 5 VirB6-like proteins, to find high confidence predictions.

For tree construction, VirB6-like proteins were aligned using Muscle5 (v5.3), sparse columns were removed using trimAl (v1.5.0) and unrooted trees were constructed using IQ-TREE3 (v3.0.1) [49–51]. Trees were then rooted at midpoint and visualised using iTOL (v7) [52]. All sequences and structure models used for the bioinformatics have been deposited to Zenodo, and can be found using the following DOI: 0.5281/zenodo.17964500

Acknowledgments

This work was supported by grants from the Swedish Research Council (2023-02423), Knut and Alice Wallenberg Foundation, and Kempestiftelserna (SMK-2059) to R.P-A.B. and from Kempestiftelserna (JCK-3126) to A.M. The bioinformatics predictions made in this study were enabled by the Berzelius resource provided by the Knut and Alice Wallenberg Foundation at the National Supercomputer Centre.

Author contributions

Josy ter Beek: Conceptualization, Methodology, Investigation, Writing – Original Draft, Writing – Review & Editing. Dennis Svedberg: Methodology, Investigation, Data curation,

Writing – Review & Editing. Kieran Deane-Alder: Investigation, Writing – Review & Editing. André Mateus: Validation, Writing – Review & Editing, Supervision, Funding acquisition. Ronnie P-A Berntsson: Conceptualization, Writing – Original Draft, Writing – Review & Editing, Supervision, Funding acquisition.

References

1. Paillard, P., Rouger, Q., Thomet, M. & Macé, K. Type IV secretion systems: from structures to mechanisms. *EMBO J* https://doi.org/10.1038/s44318-025-00584-0 (2025) doi:10.1038/s44318-025-00584-0.

2. Goldlust, K. *et al.* The F pilus serves as a conduit for the DNA during conjugation between physically distant bacteria. *Proc Natl Acad Sci U S A* 120, (2023).

3. Beltrán, L. *et al.* The mating pilus of E. coli pED208 acts as a conduit for ssDNA during horizontal gene transfer. *mBio* 15, (2024).

4. Babić, A., Lindner, A. B., Vulić, M., Stewart, E. J. & Radman, M. Direct visualization of horizontal gene transfer. *Science (1979)* 319, (2008).

5. Vadakkepat, A. K. *et al.* Cryo-EM structure of the R388 plasmid conjugative pilus reveals a helical polymer characterized by an unusual pilin/phospholipid binary complex. *Structure* 32, (2024).

6. Vadakkepat, A. K. *et al.* Cryo-EM structure of the R388 plasmid conjugative pilus reveals a helical polymer characterized by an unusual pilin/phospholipid binary complex. *Structure* 32, 1335–1347.e5 (2024).

7. Paiva, W. D., Grossman, T. & Silverman, P. M. Characterization of F-pilin as an inner membrane component of Escherichia coli K12. *Journal of Biological Chemistry* 267, 26191–26197 (1992).

8. Beijersbergen, A., Smith, S. J. & Hooykaas, P. J. J. Localization and Topology of VirB Proteins of Agrobacterium tumefaciens. *Plasmid* 32, 212–218 (1994).

9. Macé, K. *et al.* Cryo-EM structure of a type IV secretion system. *Nature* 607, 191–196 (2022).

10. Tiago Costa *et al.* Structure of the Bacterial Sex F Pilus Reveals an Assembly of a Stoichiometric Protein-Phospholipid Complex. *Cell* 166, 1436–1444 (2016).

11. Amro, J. *et al.* Cryo-EM structure of the Agrobacterium tumefaciens T-pilus reveals the importance of positive charges in the lumen. *Structure* 31, 375–384.e4 (2023).

12. Ishimoto, N., Frankel, G. & Beis, K. *Phospholipid-Independent Biogenesis of a Functional RP4 Conjugation Pilus*. (2025) doi:10.1101/2025.06.27.661960.

400 13. Guglielmini, J. *et al*. Key components of the eight classes of type IV secretion
401 systems involved in bacterial conjugation or protein secretion. *Nucleic Acids Res*
402 42, 5715–5727 (2014).

403 14. Minny Bhatty *et al*. The expanding bacterial type IV secretion lexicon. *Res Microbiol*
404 164, 620–639 (2013).

405 15. Elisabeth Grohmann *et al*. Conjugative Plasmid Transfer in Gram-Positive Bacteria.
406 *Microbiology and Molecular Biology Reviews* 67, 277–301 (2003).

407 16. Maffo-Woulefack, R. *et al*. Elucidating assembly and function of VirB8 cell wall
408 subunits refines the DNA translocation model in Gram-positive T4SSs. *Science*
409 *Advances* 11, 5975 (2025).

410 17. Ryan, M. E., Damke, P. P. & Shaffer, C. L. DNA Transport through the Dynamic Type
411 IV Secretion System. *Infect Immun* 91, e00436–22 (2023).

412 18. Grohmann, E., Christie, P. J., Waksman, G. & Backert, S. Type IV secretion in Gram-
413 negative and Gram-positive bacteria. *Mol Microbiol* 107, 455–471 (2018).

414 19. Bhatty, M. *et al*. Enterococcus faecalispCF10-encoded surface proteins PrgA,
415 PrgB (aggregation substance) and PrgC contribute to plasmid transfer, biofilm
416 formation and virulence. *Mol Microbiol* 95, 660–677 (2015).

417 20. Järvå, M. A., Hirt, H., Dunny, G. M. & Berntsson, R. P. A. Polymer Adhesin Domains
418 in Gram-Positive Cell Surface Proteins. *Frontiers in Microbiology* vol. 11 Preprint at
419 https://doi.org/10.3389/fmicb.2020.599899 (2020).

420 21. Revitt-Mills, S. A., Watts, T. D., Lyras, D., Adams, V. & Rood, J. I. The ever-expanding
421 tcp conjugation locus of pCW3 from Clostridium perfringens. *Plasmid* 113, 102516
422 (2021).

423 22. Sun, W.-S. *et al*. Structural foundation for the role of enterococcal PrgB in
424 conjugation, biofilm formation, and virulence. *Elife* 12, RP84427 (2023).

425 23. Goessweiner-Mohr, N., Arends, K., Keller, W. & Grohmann, E. Conjugation in
426 Gram-Positive Bacteria. *Microbiol Spectr* 2, 2.4.19 (2014).

427 24. Sun, W.-S., Torrens, G., ter Beek, J., Cava, F. & Berntsson, R. P.-A. Breaking
428 barriers: pCF10 type 4 secretion system relies on a self-regulating muramidase to
429 modulate the cell wall. *mBio* 0, e00488–24 (2024).

430 25. Breidenstein, A., Svedberg, D., Ter Beek, J. & Berntsson, R. P.-A. Advances in
431 Protein Structure Prediction Highlight Unexpected Commonalities Between Gram-
432 positive and Gram-negative Conjugative T4SSs. *J Mol Biol* 437, 168924 (2025).

433 26. Ellison, C. K., Dalia, T. N., Dalia, A. B. & Brun, Y. V. Real-time microscopy and
434 physical perturbation of bacterial pili using maleimide-conjugated molecules. *Nat*
435 *Protoc* 14, 1803–1819 (2019).

436 27. Thompson, J. K. & Collins, M. A. Completed sequence of plasmid pIP501 and origin
437 of spontaneous deletion derivatives. *Plasmid* 50, 28–35 (2003).

438 28. Kurenbach, B. *et al.* Intergeneric transfer of the Enterococcus faecalis plasmid
439 pIP501 to Escherichia coli and Streptomyces lividans and sequence analysis of its
440 tra region. *Plasmid* 50, 86–93 (2003).

441 29. Chen, I., Provvedi, R. & Dubnau, D. A macromolecular complex formed by a pilin-
442 like protein in competent Bacillus subtilis. *Journal of Biological Chemistry* 281,
443 (2006).

444 30. Lam, T. *et al.* Competence pili in Streptococcus pneumoniae are highly dynamic
445 structures that retract to promote DNA uptake. *Mol Microbiol* 116, (2021).

446 31. Costa, T. R. D. *et al.* Structure of the Bacterial Sex F Pilus Reveals an Assembly of
447 a Stoichiometric Protein-Phospholipid Complex. *Cell* 166, 1436-1444.e10 (2016).

448 32. Ishimoto, N., Frankel, G. & Beis, K. Phospholipid-independent biogenesis of a
449 functional RP4 conjugation pilus. *bioRxiv* (2025).

450 33. Beltran, L. C. *et al.* Archaeal DNA-import apparatus is homologous to bacterial
451 conjugation machinery. *Nat Commun* 14, (2023).

452 34. Heng, J. B. *et al.* The electromechanics of DNA in a synthetic nanopore. *Biophys J*
453 90, (2006).

454 35. Fernandez-Gonzalez, E. & Backert, S. DNA transfer in the gastric pathogen
455 Helicobacter pylori. *J Gastroenterol* 49, 594–604 (2014).

456 36. Hamilton, H. L., Domínguez, N. M., Schwartz, K. J., Hackett, K. T. & Dillard, J. P.
457 Neisseria gonorrhoeae secretes chromosomal DNA via a novel type IV secretion
458 system. *Mol Microbiol* 55, (2005).

459 37. Alvarez-Martinez, C. E. & Christie, P. J. Biological Diversity of Prokaryotic Type IV
460 Secretion Systems. *Microbiology and Molecular Biology Reviews* 73, 775–808
461 (2009).

462 38. Margaret Clarke *et al.* F-pili dynamics by live-cell imaging. *Proc Natl Acad Sci U S*
463 *A* 105, 17978–17981 (2008).

464 39. Abramson, J. *et al.* Accurate structure prediction of biomolecular interactions with
465 AlphaFold 3. *Nature* 630, 493–500 (2024).

40. Hallgren, J. *et al.* DeepTMHMM predicts alpha and beta transmembrane proteins using deep neural networks. Preprint at https://doi.org/10.1101/2022.04.08.487609 (2022).

41. Jurrus, E. *et al.* Improvements to the APBS biomolecular solvation software suite. *Protein Science* 27, 112–128 (2018).

42. Vesić, D. & Kristich, C. J. A rex family transcriptional repressor influences H2O2 accumulation by Enterococcus faecalis. *J Bacteriol* 195, 1815–1824 (2013).

43. Jäger, F., Lamy, A., Sun, W.-S., Guerini, N. & Berntsson, R. P.-A. Structure of the enterococcal T4SS protein PrgL reveals unique dimerization interface in the VirB8 protein family. *Structure* 30, 876-885.e5 (2022).

44. Høie, M. H. *et al.* NetSurfP-3.0: accurate and fast prediction of protein structural features by protein language models and deep learning. *Nucleic Acids Res* 50, (2022).

45. Molano, L.-A. G., Hirsch, P., Hannig, M., Müller, R. & Keller, A. The PLSDB 2025 update: enhanced annotations and improved functionality for comprehensive plasmid research. *Nucleic Acids Res* 53, D189–D196 (2025).

46. Steinegger, M. & Söding, J. MMseqs2 enables sensitive protein sequence searching for the analysis of massive data sets. *Nat Biotechnol* 35, 1026–1028 (2017).

47. Lin, Z. *et al.* Evolutionary-scale prediction of atomic-level protein structure with a language model. *Science (1979)* 379, 1123–1130 (2023).

48. van Kempen, M. *et al.* Fast and accurate protein structure search with Foldseek. *Nat Biotechnol* 42, 243–246 (2024).

49. Edgar, R. C. Muscle5: High-accuracy alignment ensembles enable unbiased assessments of sequence homology and phylogeny. *Nat Commun* 13, 6968 (2022).

50. Capella-Gutiérrez, S., Silla-Martínez, J. M. & Gabaldón, T. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* 25, 1972–1973 (2009).

51. Wong, T. *et al.* IQ-TREE 3: Phylogenomic Inference Software using Complex Evolutionary Models. Preprint at https://doi.org/10.32942/X2P62N (2025).

52. Letunic, I. & Bork, P. Interactive Tree of Life (iTOL) v6: recent updates to the phylogenetic tree display and annotation tool. *Nucleic Acids Res* 52, W78–W82 (2024).

53. Dunny, G., Funk, C. & Adsit, J. Direct stimulation of the transfer of antibiotic resistance by sex pheromones in Streptococcus faecalis. *Plasmid* 6, 270–278 (1981).

54. Vesić, D. & Kristich, C. J. A rex family transcriptional repressor influences H2O2 accumulation by Enterococcus faecalis. *J Bacteriol* 195, 1815–24 (2013).

55. Sun, W.-S. *et al*. Structural foundation for the role of enterococcal PrgB in conjugation, biofilm formation and virulence. *Elife* 12, RP84427 (2023).

56. Laverde Gomez, J. A., Bhatty, M. & Christie, P. J. PrgK, a Multidomain Peptidoglycan Hydrolase, Is Essential for Conjugative Transfer of the Pheromone-Responsive Plasmid pCF10. *J Appl Microbiol* 196, 527–539 (2014).

Supplementary Tables and Figures

Table S1: Scores and short descriptions of various AlphaFold 3 models.

$PrgF_{B2}$	$PrgH_{B6}$	ipTM	pTM	Comment
1	-	-	0.77	See Fig.1B
5	-	0.11	0.25	Low confidence pentameric ring
20	-	0.19	0.23	Forms a third of a circular structure, with a much larger diameter
50	-	0.34	0.35	Resembles a pilus, but with very low confidence scores
5	5	0.75	0.77	Highly similar to model shown in Fig. 1C
10	5	0.87	0.88	Highly similar to model shown in Fig. 1C
15	5	0.80	0.81	Highly similar to model shown in Fig. 1C
20	5	0.73	0.74	Highly similar to model shown in Fig. 1C
21	5	0.72	0.74	Used for visualisation in Fig. 1C-D, with one PrgF removed
22	5	0.69	0.71	This is the highest number of PrgF copies that can be modelled reliably together with PrgH

519 Table S2: Strains, plasmids and primers used in this study.

Strains	**Relevant feature(s)**	**Reference**
E. feacalis OG1RF	Parental strain, RifR, FusR	[53]
E. feacalis OG1ES	Used as recipient strain, EryR, StrepR	
Plasmids	**Relevant feature(s)**	**Reference**
pCF10	TetR, clinical isolate with T4SS encoded in pQ operon	[53]
pCJK218	for homologous recombination in *E. faecalis*, used to make pCF10Δ*prgF*	[54]
pCF10ΔprgF	pCF10 plasmid with *prgF* deletion, TetR	this study
pMSP:MCS	plasmid with nisin inducible promotor, EryR (derivitive of pMSP3545S)	[55]
pMSP:prgF	*prgF* cloned into pMSP:MCS via NcoI/XbaI, under nisin inducible promotor, EryR	this study
pMSP:prgF_D37C	pMSP:*prgF* with D37C mutation (GAT --> TGT), derived from pMSP:prgF by iPCR	this study
pMSP:prgF_D37N	pMSP:*prgF* with D37N mutation (GAT --> AAT), derived from pMSP:prgF by iPCR	this study
pMSP:prgF_D38C	pMSP:*prgF* with D38C mutation (GAT --> TGT), derived from pMSP:prgF by iPCR	this study
pMSP:prgF_D38N	pMSP:*prgF* with D38N mutation (GAT --> AAT), derived from pMSP:prgF by iPCR	this study
pMSP:prgF_R40C	pMSP:*prgF* with R40C mutation (CGT --> TGT), derived from pMSP:prgF by iPCR	this study
pMSP:prgF_R40E	pMSP:*prgF* with R40E mutation (CGT --> GAA), derived from pMSP:prgF by iPCR	this study
pMSP:prgF_D37NandD38N	pMSP:*prgF* with D37N and D38N mutations (GAT GAT--> AATAAT), derived from pMSP:prgF by iPCR	this study
pMSP:prgF_T5C	pMSP:*prgF* with T5C mutation (ACA--> TGC), ordered from Eurofins and cloned via NcoI/BamHI	this study
pMSP:prgF_T9C	pMSP:*prgF* with T9C mutation (ACA--> TGC), ordered from Eurofins and cloned via NcoI/BamHI	this study
pMSP:prgF_S35C	pMSP:*prgF* with S35C mutation (TCG--> TGC), ordered from Eurofins and cloned via NcoI/BamHI	this study
pMSP:prgF_T68C	pMSP:*prgF* with T68C mutation (ACA--> TGC), ordered from Eurofins and cloned via NcoI/BamHI	this study
Primers	**Sequence 5'-3'**	**Reference**
KOprgF_UF_fwd_BamHI	ataGGATCCgattctaagttagaggaacgattaacac	this study
KOprgF_UF_rev_XbaI	ataTCTAGAtgtgactatatccatagttttgtctctcc	this study
KOprgF_DF_fwd_XbaI	ataTCTAGAacacaggtaggagctattcg	this study
KOprgF_DF_rev_NcoI	catgtaccccatggcaatcgaacgtaattc	this study
pMSP_prgF_fwd_NcoI	ataCCatggatatagtcacaaaaattattacacta	this study
pMSP_prgF_rev_XbaI	ataTCTAGAttaaaaacgaatagctcctacctg	this study

prgF_D37C_fwd	gtcgaatTGtgatccacgtacgcttgataaaggaatag	this study
prgF_D37C_Rev	gtacgtggatcaCAattcgacatacctgaacgaatttc	this study
prgF_D37N_fwd	gtatgtcgaatAatgatccacgtacgcttgataaaggaatag	this study
prgF_D37N_rev	gtacgtggatcatTattcgacatacctgaacgaatttc	this study
prgF_D38C_fwd	gtcgaatgatTGTccacgtacgcttgataaaggaatag	this study
prgF_D38C_rev	gtacgtggACAatcattcgacatacctgaacgaatttc	this study
prgF_D38N_fwd	gtcgaatgatAATccacgtacgcttgataaaggaatag	this study
prgF_D38N_rev	gtacgtggATTatcattcgacatacctgaacgaatttc	this study
prgF_R40C_fwd	gaatgatgatccaTGTacgcttgataaaggaatagaaaaagtagttg	this study
prgF_R40C_rev	caagcgtACAtggatcatcattcgacatacctgaacg	this study
prgF_R40E_fwd	gaatgatgatccaGAAacgcttgataaaggaatagaaaaagtagttg	this study
prgF_R40E_rev	caagcgtTTCtggatcatcattcgacatacctgaacg	this study
prgF_D37NandD38N_fwd	gtatgtcgaatAATAATccacgtacgcttgataaaggaatag	this study
prgF_D37NandD38N_rev	gtacgtggATTATTattcgacatacctgaacgaatttc	this study

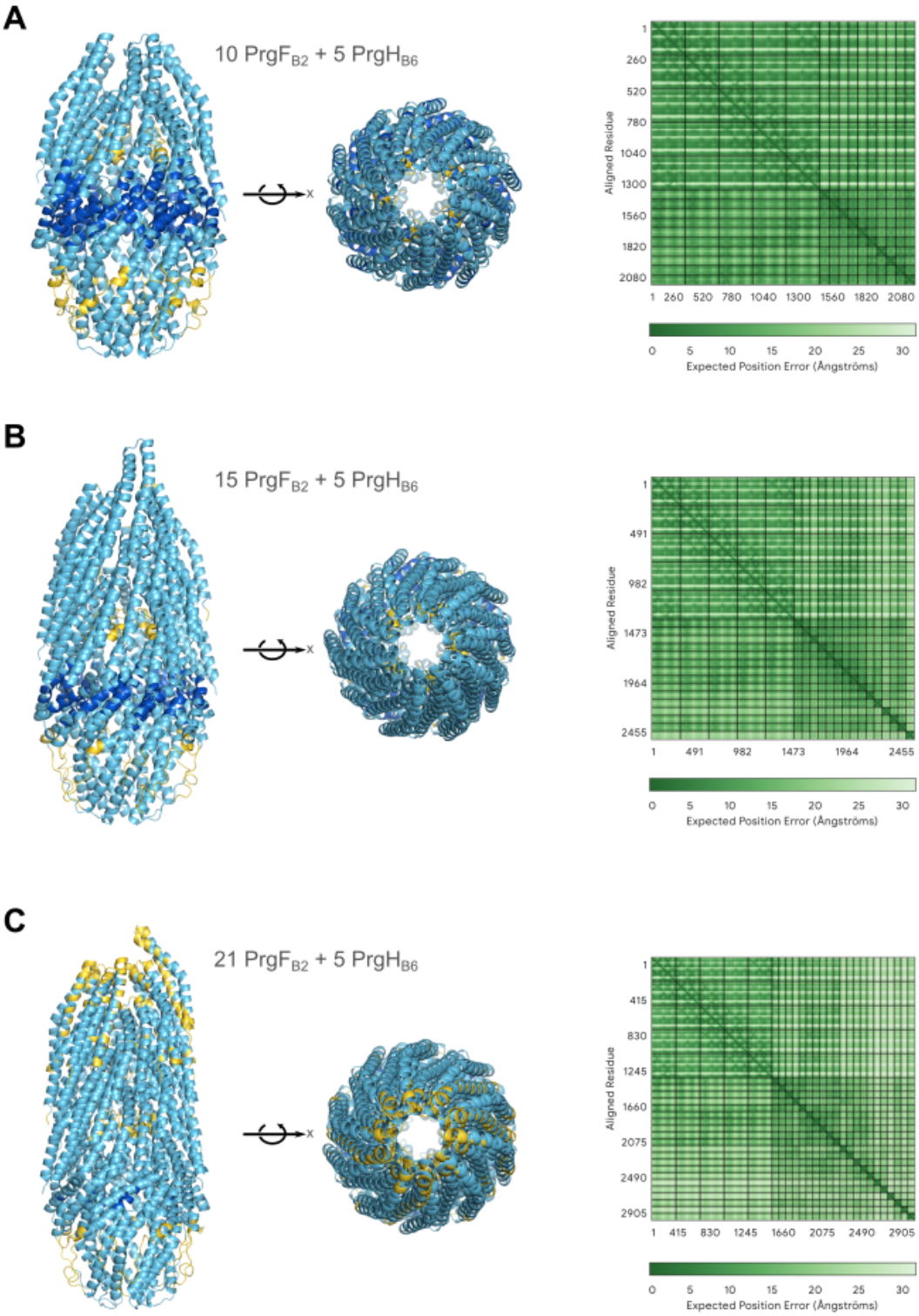
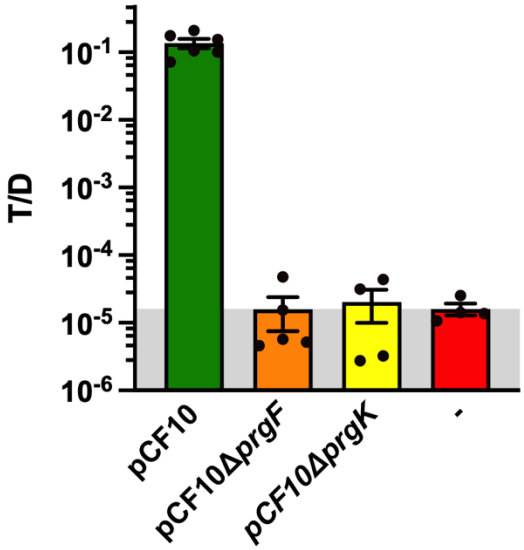


Figure S1. AlphaFold models of PrgF$_{B2}$ with PrgH$_{B6}$. All panels have 5 copies of PrgH$_{B6}$, with increasing numbers of PrgF$_{B2}$ (10 in panel A, 15 in panel B and 21 in panel C). Models are coloured according to their pLDDT scores, and the PAE plot of each model is shown on the right (starting with 5 copies of PrgH$_{B6}$ from the upper left corner).

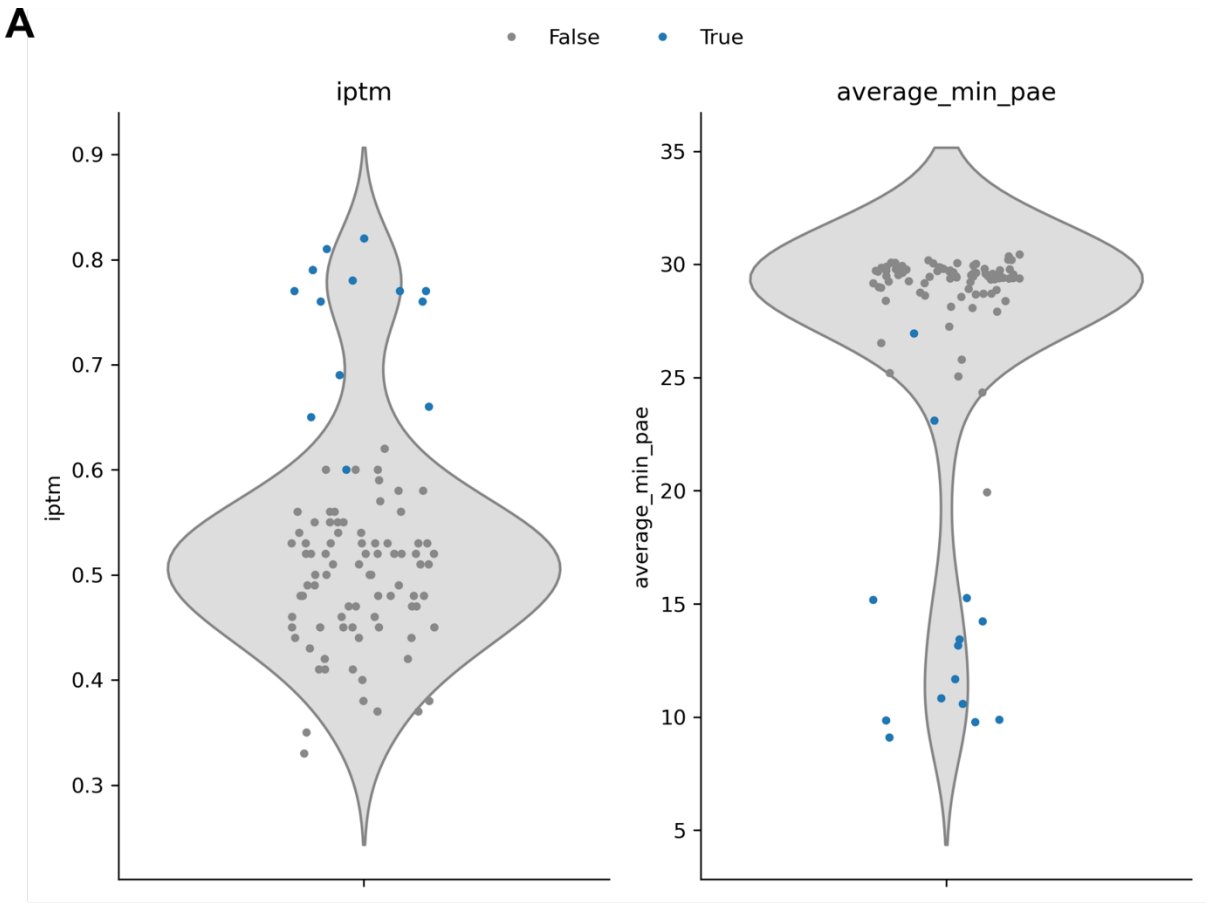


528

529 **Figure S2**: $PrgF_{B2}$ is essential for conjugation. The pCF10Δ*prgF* strain has the same
530 conjugation background level, calculated as the number of transconjugants over donors
531 (T/D), as the pCF10Δ*prgK* strain (PrgK is known to be essential [56]) and the same strain
532 without pCF10 (-). For OG1RF without pCF10 (-) the number of ”donors” were counted
533 on a plate with just fusidic acid. All strains in this figure are OG1RF without pMSP vector
534 and with or without pCF10 (as indicated).

535

Figure S3. Genetic organisation of T4SSs in example G+ conjugative plasmids.


A
False
True
iptm
average_min_pae
iptm
average_min_pae
0.9
0.8
0.7
0.6
0.5
0.4
0.3
35
30
25
20
15
10
5

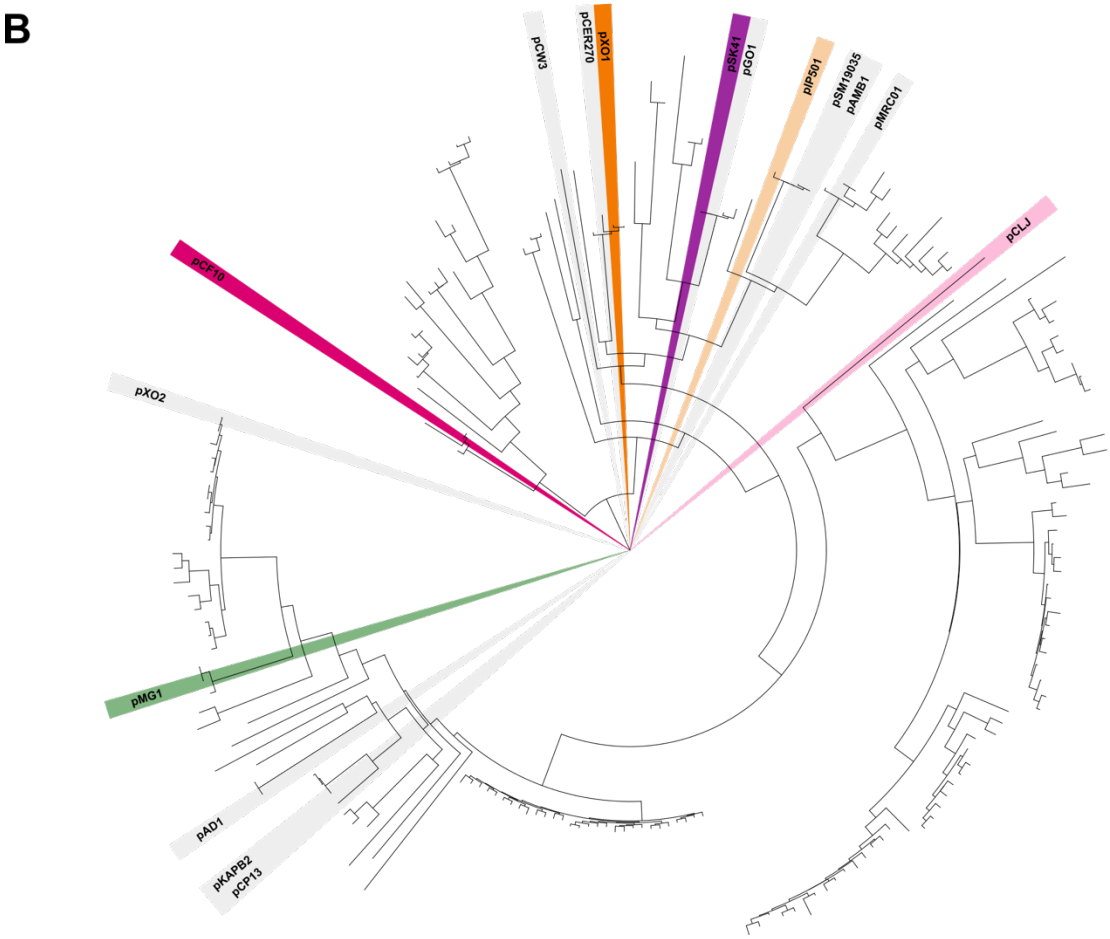
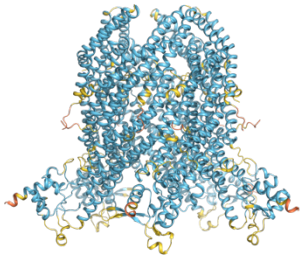
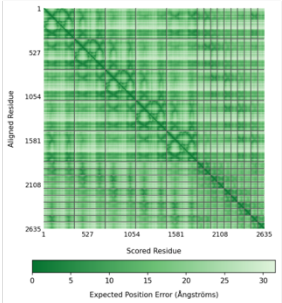
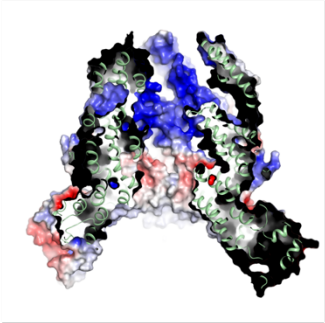

B
pCW3
pCER270
pXO1
pSK41
pGO1
pIP501
pSM19035
pAMβ1
pMRC01
pCLJ
pCF10
pXO2
pMG1
pAD1
pKAPB2
pCP13

Figure S4. Distribution of modelling scores of 16 reference plasmids. A) Distribution plots of ipTM scores (left) and average PAE scores (right) of all VirB2-VirB6 complexes modelled from 16 selected reference plasmids. By inspection of the models and their corresponding values, we set an ipTM score of >0.65 as the cut-off value for a reliable model. This is a conservative value to avoid false negatives. B) Phylogenetic tree of the VirB6-like proteins of G+ T4SSs, showing the 174 reference classes (as in Fig. 3A) as well as the 16 reference plasmids that were used in panel A. The plasmids used for the models showed in Fig. 3B are highlighted with colour, the remaining reference plasmids in gray.

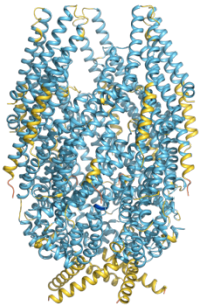
pMG1

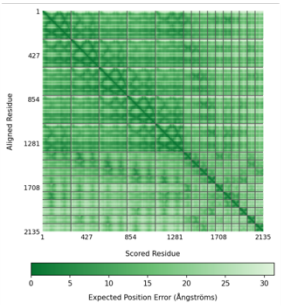


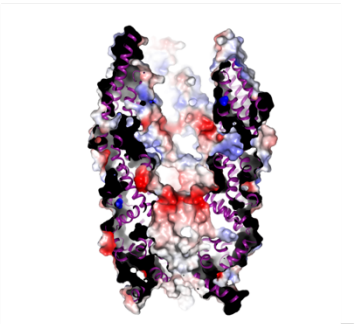




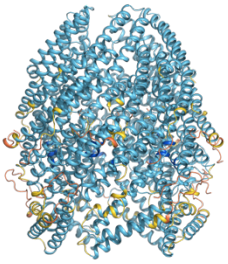
pSK41

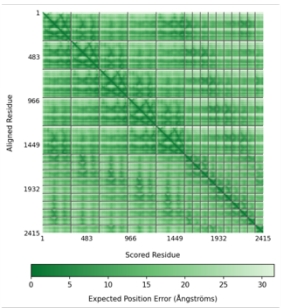


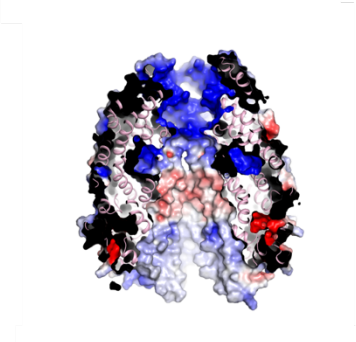




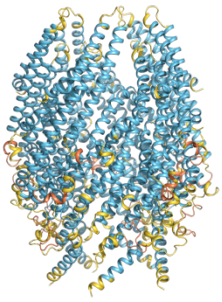
pCLJ

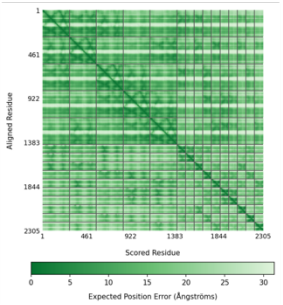


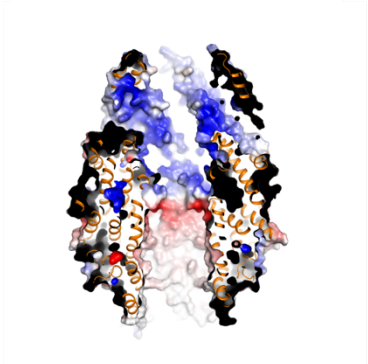




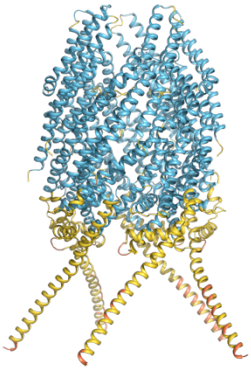
pXO1

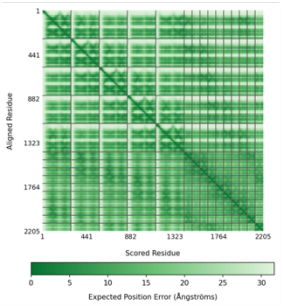






pIP501





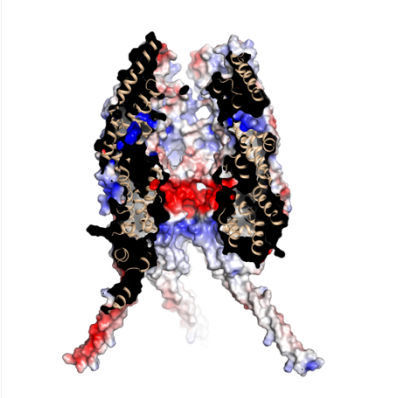


Figure S5. AlphaFold models of VirB2-VirB6 complexes of the T4SSs from Fig. 4B. All models have 10 VirB2- and 5 VirB6-like proteins. In the left panel models are coloured according to their pLDDT scores, and the corresponding PAE plot of each model is shown on the middle panel. The right panels show the surface charge of the models, coloured from red (-5 kT/e) to blue (+5 kT/e), with the cartoon model having the same colours as in Fig 3B.