


MicroReview

Monoderm bacteria: the new frontier for type IV pilus biology

Vladimir Pelicic *

MRC Centre for Molecular Bacteriology and Infection, Imperial College London, London, UK.

Summary

In the diverse world of bacterial pili, type IV pili (Tfp) are unique for two reasons: their multifunctionality and ubiquity. This latter feature offers an extraordinary possibility, that is, to perform comparative studies in evolutionarily distant species in order to improve our fragmentary understanding of Tfp biology. Regrettably, such potential has remained largely untapped, because, for 20 years, Tfp have only been characterised in diderm bacteria. However, recent studies of Tfp in monoderms have started closing the gap, revealing many interesting commonalities and a few significant differences, extending the frontiers of knowledge of Tfp biology. Here, I review the current state of the art of the Tfp field in monoderm bacteria and discuss resulting implications for our general understanding of the assembly and function of these widespread filamentous nanomachines.

Tfp at a glance

Pili are hair-like filamentous polymers composed of proteins generically named pilins, found at the surface of most prokaryotes (Bacteria and Archaea). Multiple unrelated pilus types have been defined, which are composed of different types of pilins and are assembled by distinct protein machineries (Hospenthal *et al.*, 2017). As revealed by three decades of studies, Tfp display a set of unique and highly distinctive features (Berry and Pelicic, 2015).

First, Tfp share both morphological and molecular characteristics, which allow their easy identification. In brief, Tfp are (i) flexible filaments, 4–8 nm in diameter and several

µm long (Fig. 1A), with often a propensity to form bundles, and (ii) polymers of predominantly one major type IV pilin, assembled by conserved multiprotein machineries. Type IV pilins (hereafter simply named pilins) share an N-terminal sequence motif known as class III signal peptide (Fig. 1B) and a typical 'lollipop' 3D structure (Giltner *et al.*, 2012). The N-terminal motif (IPR012902 entry in the InterPro database) (Jones *et al.*, 2014) consists of a hydrophilic leader peptide (6–51 residues long) ending with a Gly, followed by a tract of 21 mostly hydrophobic residues, except for a negatively charged Glu₅ (Fig. 1B). This hydrophobic tract, which is part of an extended N-terminal α -helix (α 1N) of about 50 residues, represents the 'stick' of the lollipop protruding from a globular head (Giltner *et al.*, 2012). The globular head consists of the C-terminal half of α 1 packed against a β -sheet made of several antiparallel β -strands (Fig. 1C). Besides pilins, many other proteins involved in Tfp biology are also very distinctive (Berry and Pelicic, 2015) including an aspartic acid protease named prepilin peptidase (IPR000045), a platform protein at the base of the assembly machinery (IPR003004), a hexameric filament assembly (extension) ATPase (IPR007831), *etc.*

The second defining feature of Tfp is that they are exquisitely multifunctional and have been associated with an array of seemingly unrelated functions, which is likely to expand even further with the characterisation of Tfp from seldom studied species of Bacteria and Archaea. Properties associated with Tfp include (i) adhesion to a wide variety of biotic and abiotic surfaces, (ii) the formation of cell aggregates, (iii) a form of surface-associated motility known as twitching motility, (iv) DNA uptake during natural transformation *etc.* It should be noted, however, that no Tfp-expressing species display all these properties, some of which are more common than others.

The third defining feature is that although Tfp have been studied only in a handful of diderm bacteria for 20 years, mostly Proteobacteria, genes encoding pilus assembly proteins are actually ubiquitous in bacterial and archaeal genomes (Imam *et al.*, 2011; Berry and Pelicic, 2015). However, not all the corresponding proteins are involved in the assembly of *bona fide* Tfp. Rather, they

Accepted 23 September, 2019. *For correspondence. E-mail: v.pelicic@imperial.ac.uk; Tel: +44 20 7594 2080.

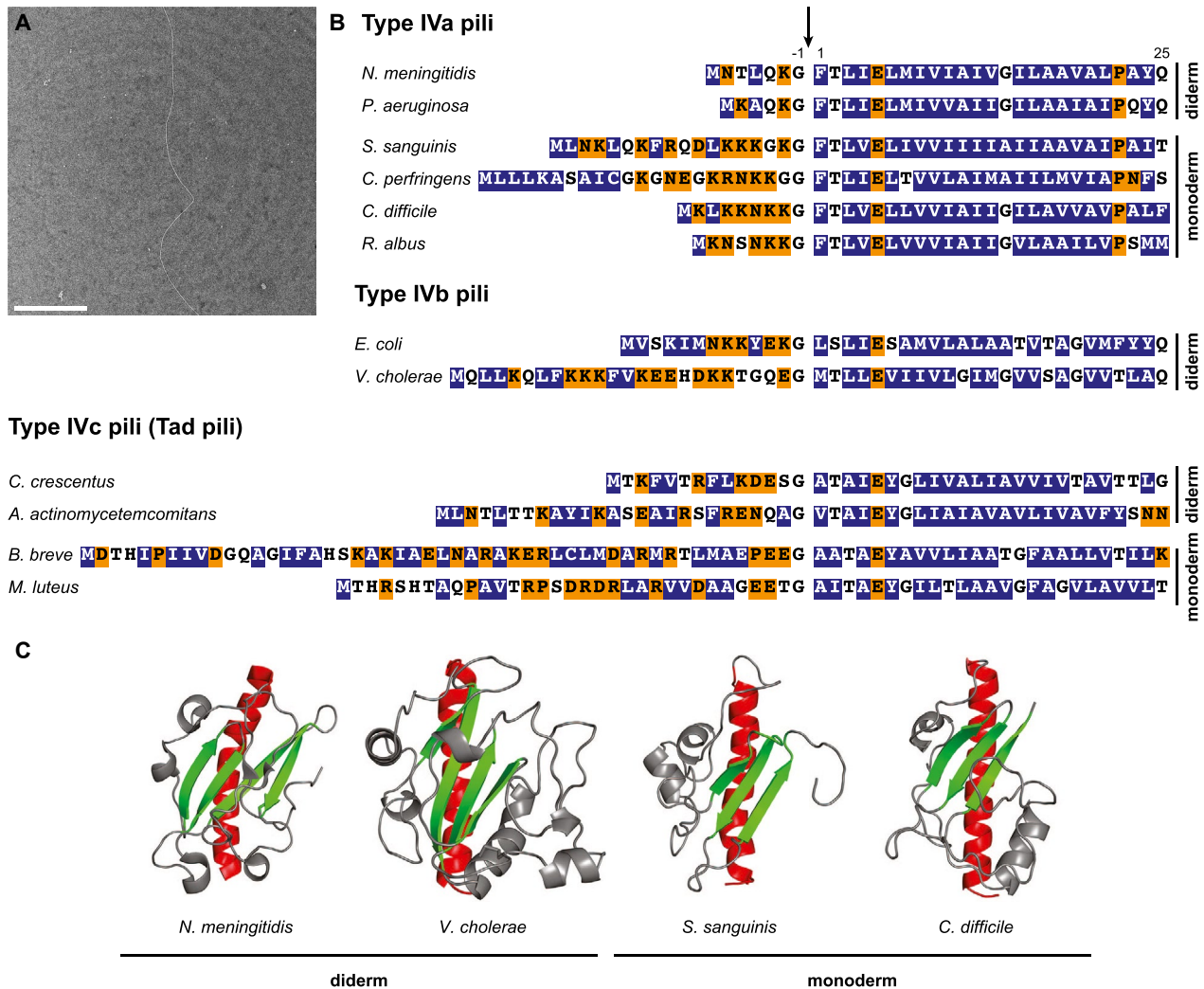


Fig. 1. Tfp and their major pilin subunits in monoderms.

A. Tfp morphology in *S. sanguinis* as assessed by transmission EM after negative staining. Scale bar represents 500 nm.

B. Sequence alignment of class III signal peptides in major pilins in selected monoderm and diderm bacteria. Hydrophilic residues are shaded in orange, hydrophobic residues in blue, and neutral residues are unshaded. The site of processing by the prepilin peptidase is indicated by a vertical arrow.

C. 3D structures of major pilins in monoderm and diderm bacteria. Conserved structural features have been highlighted including the C-terminal part of the $\alpha 1$ helix in red and the anti-parallel β -sheet in green. *Nota bene*: the protruding N-terminus of the $\alpha 1$ helix ($\alpha 1N$), which corresponds to the long stretch of hydrophobic residues in the class III signal peptide and gives full-length pilins their typical 'lollipop' 3D structure, is not represented here because it is often truncated in structural studies to improve protein solubility. The PDB entries that were used to generate this figure are: 5JW8 (*N. meningitidis*), 1OQV (*V. cholerae*), 6I2O (*S. sanguinis*) and 4PE2 (*C. difficile*).

assemble a variety of filamentous nanomachines – type 2 secretion systems (T2SS), competence pseudopili, and archaella – which are evolutionarily related to Tfp, hence their grouping in a super-family named type IV filaments (Tff) (Berry and Pelicic, 2015). However, although there is often some overlap, these nanomachines can be readily distinguished from Tfp because they have different morphological characteristics and/or mediate different properties. For example, T2SS form short pseudopili mediating protein secretion in diderm bacteria (Korotkov and Sandkvist, 2019).

Monoderm bacteria produce Tfp

In this review, I will focus exclusively on *bona fide* Tfp in monoderm bacteria (hereafter simply referred to as monoderms), the first description of which was in *Ruminococcus albus*, a cellulolytic species of Firmicutes from the order of Clostridiales. The ability of *R. albus* to adhere to cellulose was found to be mediated by Tfp (Rakotoarivonina *et al.*, 2002). High-resolution negative staining electron microscopy (EM) micrographs revealed thin, long, flexible and moderately bundling pili on the

surface of *R. albus*, which were shown by immuno-EM to be composed of GP25 – a pilin with a canonical class III signal peptide (Fig. 1B) (Rakotoarivonina *et al.*, 2002).

Tfp were subsequently found to be widespread in other species of Clostridiales, including in important human pathogens *Clostridium perfringens* and *Clostridioides difficile* (formerly known as *Clostridium difficile*). *C. perfringens* was found to produce short filaments mediating twitching motility (although it was originally referred to as gliding motility), which were absent in a mutant in the platform protein (Varga *et al.*, 2006). In *C. difficile*, high-resolution EM and immuno-EM micrographs confirmed that filaments in Clostridiaceae display classical Tfp morphology and are primarily composed of PilA1, a pilin with canonical class III signal peptide (Fig. 1B) and 3D structure (Fig. 1C) (Piepenbrink *et al.*, 2014; 2015).

The last Firmicutes in which Tfp have been described is *Streptococcus sanguinis*, a common cause of infective endocarditis in humans. Tfp are not found in other *Streptococcus* species as they are encoded by a locus found only in *S. sanguinis* (Gurung *et al.*, 2016), probably acquired by horizontal transfer from an unknown donor. *S. sanguinis* Tfp, which have a Tfp morphology (Fig. 1A), were purified to homogeneity and shown to be heteropolymers of two similar pilins (PilE1 and PilE2) in a 4:3 ratio (Berry *et al.*, 2019). These pilins exhibit canonical class III signal peptide (Fig. 1B) and 3D structure (Fig. 1C) (Berry *et al.*, 2019). It remains to be seen whether such a heteropolymeric structure might be a more general property of Tfp in monoderms, which awaits purification and biochemical characterisation of filaments in other species. However, it is now clear that the classical definition that Tfp are always homopolymers composed of one major pilin must be moderated.

Firmicutes are not the only phylum of monoderms in which Tfp have been studied. *Bifidobacterium breve*, a probiotic species of the phylum Actinobacteria, was found to express filaments composed of a canonical pilin (Fig. 1B), as shown by immuno-EM (O'Connell Motherway *et al.*, 2011). In addition, piliation was abolished in a mutant in the extension of ATPase. Tfp have also been identified in another Actinobacteria, *Micrococcus luteus*, where they are involved in natural transformation as shown by a systematic genetic analysis (Angelov *et al.*, 2015). Visualisation of *M. luteus* filaments was difficult, but they could be enriched after shearing off the bacterial surface and were shown to be composed primarily of the major pilin Flp (Fig. 1B).

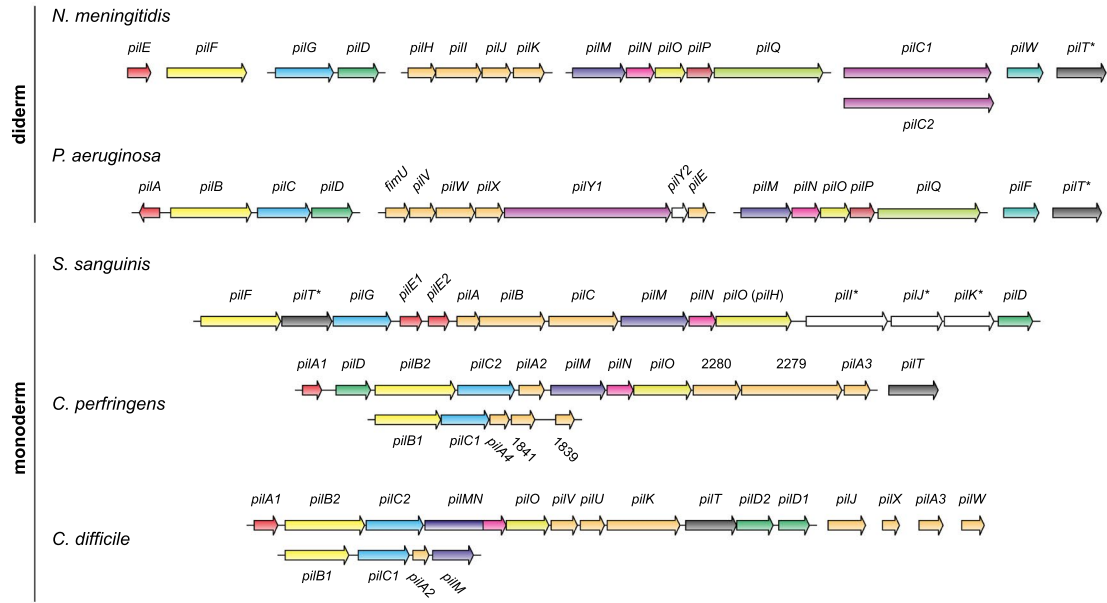
Tfp biogenesis machineries in monoderm bacteria: the beauty of simplicity

Early studies have identified two different sub-types of Tfp, named Tfpa and Tfpb, based on sequence features of their

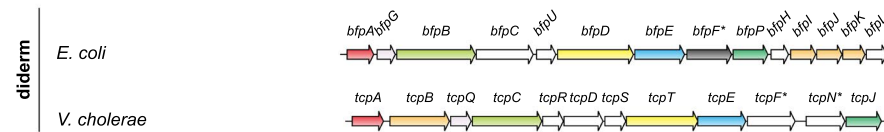
major pilin and genetic organisation of the genes involved in filament assembly (Pelicic, 2008; Giltner *et al.*, 2012). In brief, Tfpa were defined as consisting of smaller pilins with shorter leader peptides, with Tfp assembly genes scattered around the genome. In contrast, Tfpb consist of larger pilins with longer leader peptides, with assembly genes clustering together. While it is now clear that these defining features are obsolete (see below), phylogenetic studies have confirmed the validity of this sub-division by showing that Tfpa and Tfpb are phylogenetically distinct systems (Kachlany *et al.*, 2001; Denise *et al.*, 2019). Moreover, this revealed that Tad pili, which were previously thought to be a sub-family of Tfpb (Tomich *et al.*, 2007), actually constitute a distinct clade and should therefore be renamed Tfpc. A global survey (Denise *et al.*, 2019) revealed that Tfpa and Tfpc display a wider taxonomic distribution in Bacteria than Tfpb, which are found only in Proteobacteria. Interestingly, Tfpa and Tfpc are likely to be expressed by hundreds of bacterial species belonging to all major phyla of monoderms (*e.g.* Actinobacteria, Firmicutes, Thermotogae, Chloroflexi, *etc.*) (Denise *et al.*, 2019). Consistent with this, Tfp studied so far in monoderms belong to the Tfpa and Tfpc sub-types.

Tfp studied so far in Firmicutes are clearly Tfpa, although their major pilins have long leader peptides (Fig. 1B) and their Tfp assembly genes cluster together (Fig. 2). Indeed, these Tfp are (i) composed of pilins with a α 1N highly similar to that of major pilins in diderm Tfpa models (Fig. 1B), (ii) assembled by machineries that contain the PilM (IPR005883), PilN (IPR007813) and PilO (IPR007445) proteins (Fig. 2) specific of Tfpa (Denise *et al.*, 2019) and (iii) retractile filaments whose retraction is powered by the retraction ATPase PilT (IPR006321). A detailed comparison of Tfpa biogenesis genes in monoderms and diderms reveals striking overall conservation (Fig. 2), but also a few interesting differences. In the diderm models, *Neisseria meningitidis* and *Pseudomonas aeruginosa*, Tfpa biogenesis machineries are virtually identical. Systematic mutagenesis studies have shown that filament biogenesis requires 15-17 proteins (Pelicic, 2008): a pilin, several minor (low abundance) pilins (four in *N. meningitidis*, five in *P. aeruginosa*), a prepilin peptidase, a platform protein, an extension ATPase, an assembly sub-complex (PilM, PilN and PilO), a secretin pore (PilQ) in the outer membrane (OM), a linker (PilP) between the assembly sub-complex and secretin, an OM lipoprotein stabilising PilQ (PilW in *N. meningitidis*) and an OM protein involved in the late stages of Tfp biogenesis (PilC1/PilC2 in *N. meningitidis*). Filament retraction is powered by PilT, which is dispensable for piliation (Craig *et al.*, 2019). In *S. sanguinis*, the only monoderm in which a systematic genetic analysis of piliation has been performed, filament biogenesis relies on a very similar machinery, even more than we previously appreciated (Gurung *et al.*, 2016).

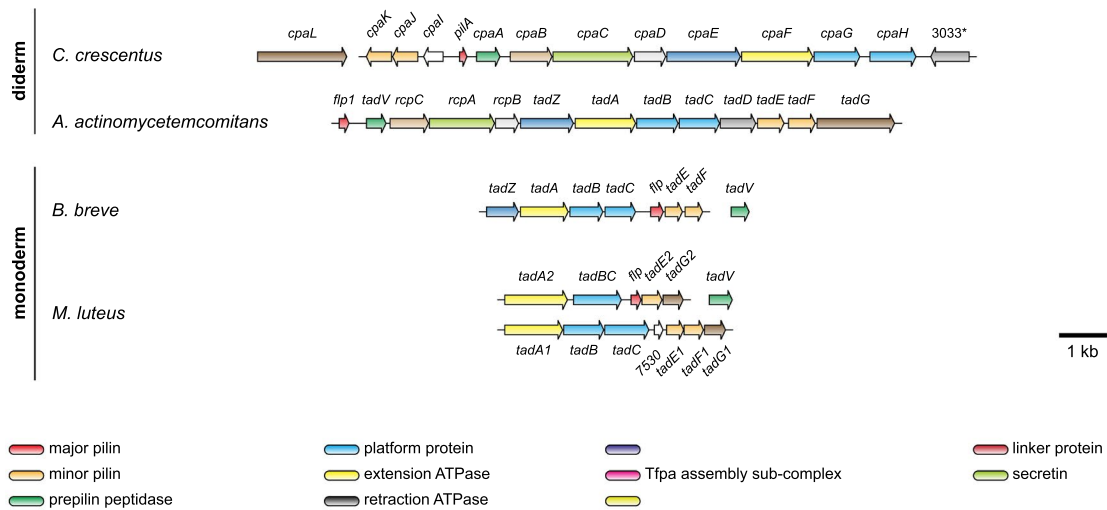
Type IVa pili



Type IVb pili



Type IVc pili (Tad pili)



█ major pilin █ platform protein █ Tfp assembly sub-complex █ linker protein
█ minor pilin █ extension ATPase █ retraction ATPase █ secretin
█ prepilin peptidase

Fig. 2. Genes involved in Tfp biogenesis in monoderms. All the genes are drawn to scale, with the scale bar representing 1 kb. Genes shaded in the same colour are likely to encode orthologous proteins based on sequence similarity, functional domain conservation and/or conserved synteny. Names of genes often mentioned in this review are listed at the bottom. No colour coding indicates genes with no clear homologs in other species. Genes with no name are identified by their ORF number in the corresponding genome, genes highlighted with an * have been shown to encode proteins dispensable for Tfp biogenesis but playing key roles in Tfp biology. The former nomenclature for the *pilO* gene in *S. sanguinis* is indicated within parentheses. The major pilin in *C. perfringens* has not been experimentally determined and has been attributed to *pilA1* based on findings in *C. difficile*. The NCBI entries that were used to generate this figure are: PRJEA34687 (*N. meningitidis*), PRJNA331 (*P. aeruginosa*), PRJEB7884 (*S. sanguinis*), PRJNA79 (*C. perfringens*), PRJNA78 (*C. difficile*), PRJNA15630 (*E. coli*), PRJNA36 (*V. cholerae*), PRJNA32027 (*C. crescentus*), PRJNA40107 (*A. actinomycetemcomitans*), PRJNA13487 (*B. breve*) and PRJNA238160 (*M. luteus*).

Indeed, the use of a sensitive protein homology detection tool (HHpred) capable of identifying remote homology (Zimmermann *et al.*, 2018) reveals that *pilH* actually encodes a protein showing N-terminal similarity to PilO. I therefore propose to rename this gene *pilO* to reflect this important finding, and to be consistent with the general nomenclature. Filament biogenesis in *S. sanguinis* requires 'only' 10 proteins (Gurung *et al.*, 2016): one major pilin (PilE1 or PilE2), three minor pilins (PilA, PilB and PilC), a prepilin peptidase, a platform protein, an extension ATPase and an assembly sub-complex (PilM, PilN and PilO). Strikingly, all the components involved in late stages of Tfp_a biogenesis in diderms (such as crossing of the OM), that most often localise to the OM, are missing in monoderms which do not have a second membrane. As in diderms, filament retraction is powered by the retraction of ATPase PilT, which is dispensable for piliation (Gurung *et al.*, 2016). It is important to underline that the term 'minor pilin' is just generic, only indicating that the corresponding proteins are cleaved by the prepilin peptidase to become low abundance components of the filaments (Berry *et al.*, 2019). This does not imply that these proteins play identical or even similar functions in each system, which is supported by the fact that minor pilins in different systems share no sequence homology. For example, PilA, PilB and PilC in *S. sanguinis* are unrelated to PilH, PilI, PilJ and PilK in *N. meningitidis*. However, findings in *S. sanguinis* are expected to be broadly applicable to Tfp_a in other monoderms, such as Clostridiaceae, where all the above genes are conserved (Melville and Craig, 2013; Maldarelli *et al.*, 2014), with only minor differences (Fig. 2). Another important conclusion is that the molecular mechanisms of Tfp_a assembly and retraction are likely to be the same in monoderms and diderms.

Tfp studied in Actinobacteria are clearly Tfp_c. Indeed, they are composed of pilins (i) with the signature Flp motif (Kachlany *et al.*, 2001) identified in the α 1N of Tfp_c major pilins (Fig. 1B), (ii) with a very small size (44 residues in *B. breve*, 68 residues in *M. luteus*) and (iii) that are assembled by machineries that contain two paralogs of the platform protein (Fig. 2) specific of Tfp_c (Denise *et al.*, 2019). A detailed comparison of Tfp_c biogenesis genes in monoderms and diderms reveals some conservation (Fig. 2) and significant differences. In the diderm models, *Caulobacter crescentus* and *Aggregatibacter actinomycetemcomitans*, Tfp_c machineries are virtually identical (Tomich *et al.*, 2007). Systematic mutagenesis studies have shown that filament biogenesis requires 13 proteins: a pilin, two minor pilins, a prepilin peptidase, two paralogous platform proteins, an extension ATPase, a secretin and several Tfp_c-specific proteins of unknown function (RcpB, RcpC, TadD, TadG and TadZ) (the *A. actinomycetemcomitans* nomenclature is used here for the sake of clarity). In monoderms, the simplification

trend discussed above for Tfp_a is also obvious for Tfp_c and is even more extreme. In *B. breve*, the Tfp_c machinery apparently consists only of a pilin, two minor pilins, a prepilin peptidase (encoded in a different locus), two paralogous platform proteins, an extension ATPase and TadZ (O'Connell Motherway *et al.*, 2011). Most of these genes are conserved in *M. luteus*, where there is extensive gene duplication, except for *tadZ* that is curiously 'replaced' by *tadG* (Fig. 2). A systematic genetic analysis of Tfp_c biogenesis in *M. luteus* revealed that all the genes from both loci are essential for piliation suggesting that paralogs have non-redundant roles (Angelov *et al.*, 2015). As in the case of Tfp_a, the striking simplicity when compared to diderm models can be explained by the fact that all the components in diderms likely to be involved at late stages of Tfp_c biogenesis (*e.g.* the secretin) and/or localising to the OM (RcpB, RcpC and TadD) are 'missing' in monoderms.

Taken together, the simplicity of Tfp_a and Tfp_c biogenesis machineries in monoderms suggests that bacteria with one membrane could be advantageous to unravel how Tfp are assembled, which remains one of the most important open questions in the field.

Tfp assembly: lessons from monoderm bacteria

What have studies of Tfp in monoderms revealed about filament assembly? In the case of Tfp_a, which have been extensively studied in diderm models, the lessons are many and potentially far-reaching.

First, since PilM, PilN and PilO proteins are found in monoderms and are essential for Tfp_a biogenesis, they do not merely constitute an 'alignment complex' as they are often referred to in the literature (Craig *et al.*, 2019; McCallum *et al.*, 2019), whose role would be to connect the pilus assembly sub-complex to the secretin in the OM, simply because there is neither secretin, nor OM in monoderms. Instead, these three proteins are involved in Tfp_a assembly and are part of the assembly sub-complex together with the platform protein and extension ATPase.

Second, unlike in diderms where PilN and PilO are similar in size and structure (their globular domains exhibiting a ferredoxin-like fold) (McCallum *et al.*, 2019), PilO in monoderms is actually much larger than PilN (56.6 vs. 22.3 kDa in *S. sanguinis*) and shows similarity to PilO only in its N-terminal portion. This suggests that it is a bifunctional protein with an unknown second role unrelated to filament assembly. A fascinating possibility, although entirely speculative at this stage, is that PilO in monoderms might be involved in the translocation of the filaments across the thick layer of peptidoglycan (PG), and/or anchoring of the pilus assembly machinery to the cell wall.

Third, the set of proteins essential for Tfp biogenesis in *S. sanguinis* strikingly overlaps with the set of *N.*

meningitidis proteins found to be involved in filament assembly *per se* (major pilin, prepilin peptidase, platform protein, extension ATPase, assembly sub-complex PilM, PilN, PilO and PilP) (Carbonnelle *et al.*, 2006; Goosens *et al.*, 2017). The only apparent differences concern PilP, which is absent in *S. sanguinis*, and the fact that minor pilins are not part of the minimal synthetic machinery necessary for Tfp assembly in a non-native heterologous host (Goosens *et al.*, 2017). There are suitable explanations for both these differences. Since PilP is the linker between the assembly sub-complex and the secretin (Tammam *et al.*, 2013), its absence in monoderms where there is no secretin is logical. PilP requirement in a minimal synthetic Tfp assembly machinery (Goosens *et al.*, 2017) is therefore likely to be only indirect, because PilP is essential for the stability of the PilN-PilO complex in diderms (Georgiadou *et al.*, 2012). As for the minor pilins, it has been shown in diderms that they become dispensable for piliation when pilus retraction is abolished by a second mutation in *pilT* (Winther-Larsen *et al.*, 2005; Carbonnelle *et al.*, 2006), which means that they are not involved in filament assembly *per se*. It remains to be seen, however, if such a genetic suppression strategy would yield similar findings in monoderms.

Altogether, observations in monoderms and diderms suggest a hypothetical scenario for Tfp assembly in monoderms (Fig. 3). Prepilins (major and minor) are first translocated via the SRP-Sec pathway across the cytoplasmic membrane (CM) (Arts *et al.*, 2007; Francetic *et al.*, 2007), where they remain initially as bitopic proteins with

their $\alpha 1N$ behaving as a transmembrane domain. Their leader peptide is then processed after the conserved Gly by the prepilin peptidase (Aly *et al.*, 2013), which also methylates the first residue of mature pilins as confirmed in *S. sanguinis* (Berry *et al.*, 2019). It is unclear at this stage, if the minor pilins are assembled first and cap the filaments as proposed for other Tff (Korotkov and Hol, 2008). Next, filament assembly involves extrusion of the pilins from the CM and their polymerisation at the base of a growing filament, which is performed by a sub-complex consisting of the platform protein, extension ATPase and PilM-PilN-PilO sub-complex. Multiple findings, including cryo-electron tomography (ET) of the entire machinery (Chang *et al.*, 2016), indicate that these components form a structure at the CM with several interconnected layers. The protein-protein interactions are likely to be the same in all Tfp systems (Fig. 3). In brief, PilM interacts with the N-terminus of PilN (Georgiadou *et al.*, 2012), forming a ring at cytoplasmic side of the CM (Chang *et al.*, 2016). In monoderms, this interaction is further supported by the fact that PilM and PilN are fused in *C. difficile* (Fig. 2) (Melville and Craig, 2013), and it might involve a conserved C-terminal RD(MILF)N(FYW)FS motif in PilN, which is different from the motif involved in the PilM-PilN interaction in diderms (Georgiadou *et al.*, 2012). PilN then interacts with PilO (Georgiadou *et al.*, 2012), forming a ring on the periplasmic side of the CM (Chang *et al.*, 2016). As speculated above, PilO in monoderms might be a bifunctional protein involved in the translocation of the filaments across the PG and/or anchoring of the assembly

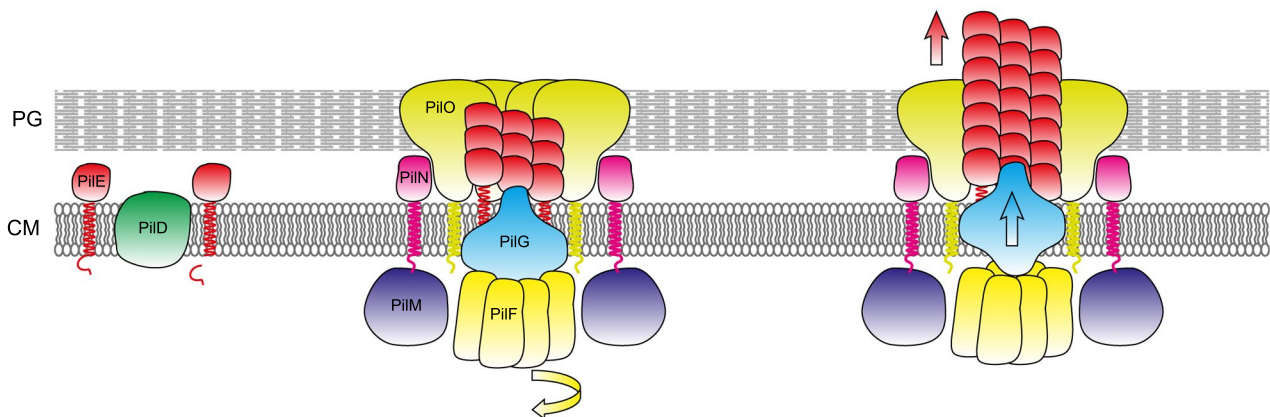


Fig. 3. Molecular mechanisms of Tfp assembly in monoderms: a working model. The model involves three steps, from left to right. The *S. sanguinis* nomenclature has been used here for the sake of clarity, and the same colour code than in Fig. 2 has been used. (1) Prepilins, which are translocated across the CM via the SRP-Sec pathway, are processed by the prepilin peptidase PilD and constitute a reservoir of mature subunits, ready to be polymerised into filaments. (2) Filament assembly involves extrusion of the pilins from the CM and their polymerisation at the base of a growing filament, which is performed by a sub-complex consisting of PilF-PilG-PilM-PilN-PilO. The topology of this sub-complex is based on existing literature. ATP binding, hydrolysis and release drive a clockwise rotation of central sub-pores in the PilF ATPase (indicated by the yellow arrow). (3) This rotatory movement thrusts the interacting PilG platform protein upwards (cyan arrow) and rotates it clockwise, before it falls back into the CM. This rotational motion pushes pilins out of the CM (red arrow) and allow their polymerisation in a right-handed helical filament. Although the PilT retraction motor is not represented on this figure, it is thought that ATP binding, hydrolysis and release drive a counterclockwise rotation of PilT (hence reverse to PilF), which wrenches PilG downwards and facilitates PilE depolymerisation from the filament. CM, cytoplasmic membrane; PG, peptidoglycan.

machinery to the cell wall (Fig. 3). The two rings spanning the CM accommodate the platform protein (Chang *et al.*, 2016), which interacts with the hexameric extension ATPase (Goosens *et al.*, 2017). A recent structural study has suggested that cycles of ATP binding, hydrolysis and release drive a clockwise rotation of central sub-pores in the ATPase hexamer, which thrusts the interacting platform protein upwards and rotates it clockwise, before it falls back in the CM (McCallum *et al.*, 2017). This rotational motion pushes pilins out of the CM for their subsequent polymerisation in a right-handed helical filament (Fig. 3). Recent cryo-EM reconstructions of several Tfp revealed that the filaments are right-handed helical polymers of pilins held together by extensive hydrophobic interactions between their $\alpha 1$ helices, which run approximately parallel to each other within the filament core (Kolappan *et al.*, 2016; Wang *et al.*, 2017). Extensive sequence conservation in $\alpha 1N$ (Fig. 1B) suggests that Tfp_a in monoderms have a similar structure. Furthermore, conservation of the helix-breaking Pro₂₂ suggests that $\alpha 1N$ helices will be partially melted between Ala₁₄/Gly₁₄ and Pro₂₂ as in diderm filaments, which is thought to make Tfp_a flexible and elastic (Kolappan *et al.*, 2016; Wang *et al.*, 2017).

Since Tfp_c assembly is less understood (Tomich *et al.*, 2007), fewer lessons can be learned from monoderms but they are interesting nonetheless. It appears that Tfp_c assembly in monoderms relies on a machinery much simpler than Tfp_a. Actually, it might be one of the smallest, if not the smallest machinery capable of assembling canonical Tfp. Findings in monoderms suggest that the minimal machinery capable of assembling Tfp_c consists of a major pilin and two minor pilins, which are processed by a prepilin peptidase, and polymerised by two paralogue platform proteins powered by an extension ATPase (O'Connell Motherway *et al.*, 2011; Angelov *et al.*, 2015). A CM protein is also involved, which is curiously different in *B. breve* (TadZ) and *M. luteus* (TadG). As a corollary, all the other Tfp_c biogenesis proteins found in diderm models (Tomich *et al.*, 2007) are likely to act after filament assembly, possibly at the step during which filaments cross the periplasm and OM in order to emerge on the cell surface. An important question is how the assembly sub-complex is held together in Tfp_c systems, when compared to Tfp_a. Can the Tfp_c-specific CM component (TadZ/TadG) 'substitute' for the PilM-PilN-PilO complex? Or has the Tfp_c platform extension ATPase complex evolved to be stable and functional on its own, and if so how?

Monoderm bacteria too display a wide array of Tfp-mediated functions

Our mechanistic understanding of the many Tfp-mediated functions remains incomplete, at best. Since many of

these functions have also been reported for Tfp in monoderms, the study of these simpler filaments might illuminate some poorly understood mechanistic aspects.

Perhaps unsurprisingly, one of the first role attributed to Tfp in monoderms is adhesion to biotic/abiotic surfaces, a very common Tfp property and possibly the main reason why these filaments have been studied for decades (Berry and Pelicic, 2015). In contrast to other types of pili that usually harbour a minor pilin with intrinsic adhesive properties at their tip (Hospenthal *et al.*, 2017), it is not well understood how Tfp mediate adhesion. This could be clarified by further studying Tfp_a and Tfp_c in monoderms as there is growing evidence that both are 'sticky' filaments. Tfp_c were found to be essential for efficient *in vivo* colonisation of the murine gut by *B. breve*, since a non-piliated mutant was severely impaired in colonisation and persistence (O'Connell Motherway *et al.*, 2011). Strikingly, a minor pilin (TadE) was found to promote epithelial cell proliferation *in vivo*, suggesting that *B. breve* Tfp contribute to the maturation of the human gut in early life by stimulating growth of the epithelial mucosa (O'Connell Motherway *et al.*, 2019). For Tfp_a, there is also ample evidence of their role in adhesion in a variety of monoderms. In *R. albus*, a cellulolytic species playing an important role in fibre breakdown in the rumen of herbivores, Tfp bind cellulose but it is not clear whether the major pilin GP25 has intrinsic binding activity or not (Pegden *et al.*, 1998; Rakotoarivonina *et al.*, 2002). In monoderms pathogenic for humans, Tfp_a have been implicated in binding to host cells. Heterologous expression of *pilA2* from *C. perfringens* in a *pilE* non-piliated mutant of *N. gonorrhoeae* was sufficient to restore piliation and confer attachment to mouse or rat muscle cell lines, to which the gonococcus does not normally adhere (Rodgers *et al.*, 2011). These findings directly implicate *C. perfringens* Tfp_a, more specifically PilA2, in the ability of this species to adhere to muscle fibres, which is key to its pathogenic potential. Similarly, Tfp_a were found to promote (i) the adherence of *C. difficile* to human intestinal epithelial cells and persistence in the intestine (McKee *et al.*, 2018), and (ii) the binding of *S. sanguinis* to human cells (Chen *et al.*, 2019). It is not known which pilus component mediates binding, but the presence of very large (>50 kDa) minor pilins with a 'modular' architecture, that is, with bulky C-terminal functional domains, is certainly an intriguing possibility. In *S. sanguinis*, PilB and PilC contain C-terminal domains often involved in adhesion in proteins unrelated to Tfp, a von Willebrand factor type A domain (IPR002035) and a concanavalin A-like lectin/glucanase structural domain (SSF49899), respectively (Berry *et al.*, 2019). This suggests that *S. sanguinis* Tfp harbour two minor pilins with intrinsic adhesive properties, maybe at their tip, allowing the bacterium to bind host protein and carbohydrate ligands. In addition, in the above

species, colonisation of host cells and surfaces is further enhanced by the ability of Tfp_a to mediate the formation of cell aggregates (Bordeleau *et al.*, 2015), which eventually results in the formation of biofilms (Varga *et al.*, 2008; Maldarelli *et al.*, 2016; Chen *et al.*, 2019). It remains to be seen whether pilus-pilus contacts leading to aggregation involve major or minor pilins.

Twitching motility, a distinctive property of Tfp_a (Berry and Pelicic, 2015), was observed and studied in several monoderm Tfp_a-expressing species: *C. perfringens*, *C. difficile* and *S. sanguinis* (Varga *et al.*, 2006; Mendez *et al.*, 2008; Gurung *et al.*, 2016; Purcell *et al.*, 2016). Actually, studies in the 1970's, showing that most *S. sanguinis* isolates were capable of surface-associated motility with the hallmarks of twitching motility, were the first evidence for the presence of Tfp in monoderms (Henriksen and Henriksen, 1975). Twitching motility is a PilT-powered process during which bacteria use Tfp as 'grappling hooks' to attach to a surface and then pull themselves upon filament retraction towards the site where the pilus is attached. Tensile forces generated upon PilT-mediated pilus retraction were measured in *S. sanguinis* and found to be remarkably similar to those in Gram-negative Tfp_a-expressing species (*i.e.* hundreds of pN), despite these bacteria having widely different surface architectures. A systematic genetic analysis identified four *S. sanguinis* proteins required for twitching motility but dispensable for pilus biogenesis (PilI, PilJ, PilK and PilT), the first three of which are not found in diderm Tfp_a-expressing species (Gurung *et al.*, 2016). This offers the unprecedented possibility to characterise twitching motility at an atomic level without interference of an OM.

The last property that was associated with Tfp in monoderms is DNA uptake during natural transformation. In most diderm competent species such as *N. meningitidis*, Tfp_a bind-free extracellular DNA and promote its import (or uptake) across the OM, upon PilT-powered pilus retraction (Ellison *et al.*, 2018). DNA is then transferred to a second (Com) machinery that further translocates it across the PG and CM. In competent monoderm species, this early step in natural transformation is not mediated by Tfp but by a closely related Tff known as competence pseudopilus. This filamentous structure, which can be long or not, has been shown to bind DNA (Laurenceau *et al.*, 2013; Muschiol *et al.*, 2017). However, an unexpected twist comes from the analysis of DNA uptake in *M. luteus* that was found to be mediated by Tfp_c (Angelov *et al.*, 2015), which were not previously known to be capable of mediating this property. This finding has important functional implications, especially since it suggests that Tfp_c in *M. luteus* are retractile filaments capable of generating significant tensile force, which has recently been confirmed in diderms (Ellison *et al.*, 2017). Owing to its simple pilus biogenesis machinery (Fig. 2), *M. luteus* is

therefore a promising model for unravelling the molecular mechanisms of PilT-independent Tfp_c retraction, which remain mysterious.

Concluding remarks and future directions

Although the study of Tfp in monoderms is still in its infancy, it is now obvious that these filaments are not mere curiosities distantly related to Tfp in diderm models but *bona fide* Tfp, which might represent the long sought new research avenue capable of filling important gaps in our understanding of Tfp biology. The key advantage of monoderms is the simplicity of their Tfp biogenesis machineries, which stems from their simpler surface architecture. These new Tfp models, especially those with developed genetics like *S. sanguinis* and *M. luteus*, could help us to understand how filaments are assembled and function by performing in-depth studies using a combination of genetics, two-hybrid assays, cryo-ET, structural biology, synthetic biology (the compact *S. sanguinis* system is a handy source of genes for reconstituting Tfp_a into a non-native monoderm host) *etc.*

Important advances (although this is obviously subjective) concerning the molecular mechanisms of filament biogenesis would be to (i) precisely define the roles of PilM, PilN and PilO proteins in Tfp_a assembly, (ii) determine whether the minor pilins sit at the filament tip and prime filament assembly (Cisneros *et al.*, 2012; Nguyen *et al.*, 2015), (iii) understand how Tfp cross the mesh-like PG layer (could the non-canonical PilO protein be involved) and (iv) understand how the more rudimentary Tfp_c machinery works. As for Tfp-mediated functions, monoderms could prove key for understanding how (i) Tfp mediate adhesion by studying PilB and PilC modular pilins (this would have implications reaching far beyond adhesion since modular pilins with hundreds of different domain architectures are widespread), (ii) Tfp_a retraction/extension dynamics orchestrate motility, possibly by using elegant labelling or label-free strategies that allow dynamic filament visualisation (Ellison *et al.*, 2017; Tala *et al.*, 2019) and (iii) Tfp_c retraction occurs in the absence of PilT.

In conclusion, recent advances in Tfp biology in monoderms – including Archaea that have been covered elsewhere (Chaudhury *et al.*, 2018) – outline a very exciting prospect for the years to come, that is, that monoderms might finally contribute to our mechanistic understanding of the assembly/function of Tfp and related filamentous nanomachines, which are ubiquitous in prokaryotes.

Acknowledgements

I would like to thank the present and former members of my group who have contributed to the *S. sanguinis* Tfp project, especially Ishwori Gurung who initiated it. Work in my

laboratory is currently supported by the Medical Research Council. I would like to thank Nicolas Biais (City University of New York), Angelika Gründling (Imperial College London) and Romé Voulhoux (CNRS Marseille) for critical reading of this manuscript.

References

- Aly, K.A., Beebe, E.T., Chan, C.H., Goren, M.A., Sepulveda, C., Makino, S., *et al.* (2013) Cell-free production of integral membrane aspartic acid proteases reveals zinc-dependent methyltransferase activity of the *Pseudomonas aeruginosa* prepilin peptidase PilD. *MicrobiologyOpen*, **2**, 94–104.
- Angelov, A., Bergen, P., Nadler, F., Hornburg, P., Lichev, A., Ubelacker, M., *et al.* (2015) Novel Flp pilus biogenesis-dependent natural transformation. *Frontiers in Microbiology*, **6**, 84.
- Arts, J., van Bortel, R., Filloux, A., Tommassen, J. and Koster, M. (2007) Export of the pseudopilin XcpT of the *Pseudomonas aeruginosa* type II secretion system via the signal recognition particle-Sec pathway. *Journal of Bacteriology*, **189**, 2069–2076.
- Berry, J.L., Gurung, I., Anonsen, J.H., Spielman, I., Harper, E., Hall, A.M.J., *et al.* (2019) Global biochemical and structural analysis of the type IV pilus from the Gram-positive bacterium *Streptococcus sanguinis*. *Journal of Biological Chemistry*, **294**, 6796–6808.
- Berry, J.L. and Pelicic, V. (2015) Exceptionally widespread nano-machines composed of type IV pilins: the prokaryotic Swiss Army knives. *FEMS Microbiology Reviews*, **39**, 134–154.
- Bordeleau, E., Purcell, E.B., Lafontaine, D.A., Fortier, L.C., Tamayo, R. and Burrus, V. (2015) C-di-GMP riboswitch-regulated type IV pili contribute to aggregation of *Clostridium difficile*. *Journal of Bacteriology*, **197**, 819–832.
- Carbonnelle, E., Helaine, S., Nassif, X. and Pelicic, V. (2006) A systematic genetic analysis in *Neisseria meningitidis* defines the Pil proteins required for assembly, functionality, stabilization and export of type IV pili. *Molecular Microbiology*, **61**, 1510–1522.
- Chang, Y.W., Rettberg, L.A., Treuner-Lange, A., Iwasa, J., Sogaard-Andersen, L. and Jensen, G.J. (2016) Architecture of the type IVa pilus machine. *Science*, **351**, aad2001.
- Chaudhury, P., Quax, T.E.F. and Albers, S.V. (2018) Versatile cell surface structures of archaea. *Molecular Microbiology*, **107**, 298–311.
- Chen, Y.M., Chiang, Y.C., Tseng, T.Y., Wu, H.Y., Chen, Y.Y., Wu, C.H., *et al.* (2019) Molecular and functional analysis of the type IV pilus gene cluster in *Streptococcus sanguinis* SK36. *Applied and Environmental Microbiology*, **85**, e02788-18.
- Cisneros, D.A., Bond, P.J., Pugsley, A.P., Campos, M. and Francetic, O. (2012) Minor pseudopilin self-assembly primes type II secretion pseudopilus elongation. *The EMBO Journal*, **31**, 1041–1053.
- Craig, L., Forest, K.T. and Maier, B. (2019) Type IV pili: dynamics, biophysics and functional consequences. *Nature Reviews in Microbiology*, **17**, 429–440.
- Denise, R., Abby, S.S. and Rocha, E.P.C. (2019) Diversification of the type IV filament superfamily into machines for adhesion, protein secretion, DNA uptake, and motility. *PLoS Biology*, **17**, e3000390.
- Ellison, C.K., Dalia, T.N., Vidal Ceballos, A., Wang, J.C., Biais, N., Brun, Y.V., *et al.* (2018) Retraction of DNA-bound type IV competence pili initiates DNA uptake during natural transformation in *Vibrio cholerae*. *Nature Microbiology*, **3**, 773–780.
- Ellison, C.K., Kan, J., Dillard, R.S., Kysela, D.T., Ducret, A., Berne, C., *et al.* (2017) Obstruction of pilus retraction stimulates bacterial surface sensing. *Science*, **358**, 535–538.
- Francetic, O., Buddelmeijer, N., Lewenza, S., Kumamoto, C.A. and Pugsley, A.P. (2007) Signal recognition particle-dependent inner membrane targeting of the PulG pseudopilin component of a type II secretion system. *Journal of Bacteriology*, **189**, 1783–1793.
- Georgiadou, M., Castagnini, M., Karimova, G., Ladant, D. and Pelicic, V. (2012) Large-scale study of the interactions between proteins involved in type IV pilus biology in *Neisseria meningitidis*: characterization of a subcomplex involved in pilus assembly. *Molecular Microbiology*, **84**, 857–873.
- Giltner, C.L., Nguyen, Y. and Burrows, L.L. (2012) Type IV pilin proteins: versatile molecular modules. *Microbiology and Molecular Biology Reviews*, **76**, 740–772.
- Goosens, V.J., Busch, A., Georgiadou, M., Castagnini, M., Forest, K.T., Waksman, G., *et al.* (2017) Reconstitution of a minimal machinery capable of assembling periplasmic type IV pili. *Proceedings of the National Academy of Sciences of the United States of America*, **114**, E4978–E4986.
- Gurung, I., Spielman, I., Davies, M.R., Lala, R., Gaustad, P., Biais, N., *et al.* (2016) Functional analysis of an unusual type IV pilus in the Gram-positive *Streptococcus sanguinis*. *Molecular Microbiology*, **99**, 380–392.
- Henriksen, S.D. and Henriksen, J. (1975) Twitching motility and possession of polar fimbriae in spreading *Streptococcus sanguis* isolates from the human throat. *Acta Pathologica e Microbiologica Scandinavica [B]*, **83**, 133–140.
- Hospenthal, M.K., Costa, T.R.D. and Waksman, G. (2017) A comprehensive guide to pilus biogenesis in Gram-negative bacteria. *Nature Reviews in Microbiology*, **15**, 365–379.
- Imam, S., Chen, Z., Roos, D.S. and Pohlschröder, M. (2011) Identification of surprisingly diverse type IV pili, across a broad range of Gram-positive bacteria. *PLoS ONE*, **6**, e28919.
- Jones, P., Binns, D., Chang, H.Y., Fraser, M., Li, W., McAnulla, C., *et al.* (2014) InterProScan 5: genome-scale protein function classification. *Bioinformatics*, **30**, 1236–1240.
- Kachlany, S.C., Planet, P.J., DeSalle, R., Fine, D.H., Figurski, D.H. and Kaplan, J.B. (2001) *flp-1*, the first representative of a new pilin gene subfamily, is required for non-specific adherence of *Actinobacillus actinomycetemcomitans*. *Molecular Microbiology*, **40**, 542–554.
- Kolappan, S., Coureuil, M., Yu, X., Nassif, X., Egelman, E.H. and Craig, L. (2016) Structure of the *Neisseria meningitidis* type IV pilus. *Nature Communications*, **7**, 13015.
- Korotkov, K.V. and Hol, W.G. (2008) Structure of the GspK-GspI-GspJ complex from the enterotoxigenic *Escherichia coli* type 2 secretion system. *Nature Structural and Molecular Biology*, **15**, 462–468.

- Korotkov, K.V. and Sandkvist, M. (2019) Architecture, function, and substrates of the type II secretion system. *EcoSal Plus*, **8**. <https://doi.org/10.1128/ecosalplus.ESP-0034-2018>.
- Laurenceau, R., Pehau-Arnudet, G., Baconnais, S., Gault, J., Malosse, C., Dujeancourt, A., *et al.* (2013) A type IV pilus mediates DNA binding during natural transformation in *Streptococcus pneumoniae*. *PLoS Path*, **9**, e1003473.
- Maldarelli, G.A., De Masi, L., von Rosenvinge, E.C., Carter, M. and Sonnenberg, M.S. (2014) Identification, immunogenicity, and cross-reactivity of type IV pilin and pilin-like proteins from *Clostridium difficile*. *Pathogens and Disease*, **71**, 302–314.
- Maldarelli, G.A., Piepenbrink, K.H., Scott, A.J., Freiberg, J.A., Song, Y., Achermann, Y., *et al.* (2016) Type IV pili promote early biofilm formation by *Clostridium difficile*. *Pathogens and Disease*, **74**, ftw061.
- McCallum, M., Burrows, L.L. and Howell, P.L. (2019) The dynamic structures of the type IV pilus. In: Sandkvist, M., Cascales, E. and Christie, P.J. (Eds.) *Protein secretion in bacteria*. Washington DC: ASM Press, pp. 113–128.
- McCallum, M., Tammam, S., Khan, A., Burrows, L.L. and Howell, P.L. (2017) The molecular mechanism of the type IV pilus motors. *Nature Communications*, **8**, 15091.
- McKee, R.W., Aleksanyan, N., Garrett, E.M. and Tamayo, R. (2018) Type IV pili promote *Clostridium difficile* adherence and persistence in a mouse model of infection. *Infection and Immunity*, **86**, e00943-17.
- Melville, S. and Craig, L. (2013) Type IV pili in Gram-positive bacteria. *Microbiology and Molecular Biology Reviews*, **77**, 323–341.
- Mendez, M., Huang, I.H., Ohtani, K., Grau, R., Shimizu, T. and Sarker, M.R. (2008) Carbon catabolite repression of type IV pilus-dependent gliding motility in the anaerobic pathogen *Clostridium perfringens*. *Journal of Bacteriology*, **190**, 48–60.
- Muschiol, S., Erlendsson, S., Aschtgen, M.S., Oliveira, V., Schmieder, P., de Lichtenberg, C., *et al.* (2017) Structure of the competence pilus major pilin ComGC in *Streptococcus pneumoniae*. *Journal of Biological Chemistry*, **292**, 14134–14146.
- Nguyen, Y., Sugiman-Marangos, S., Harvey, H., Bell, S.D., Charlton, C.L., Junop, M.S., *et al.* (2015) *Pseudomonas aeruginosa* minor pilins pime type IVa plus assembly and promote surface display of the PilY1 adhesin. *Journal of Biological Chemistry*, **290**, 601–611.
- O'Connell Motherway, M., Houston, A., O'Callaghan, G., Reunanen, J., O'Brien, F., O'Driscoll, T., *et al.* (2019) A bifidobacterial pilus-associated protein promotes colonic epithelial proliferation. *Molecular Microbiology*, **111**, 287–301.
- O'Connell Motherway, M., Zomer, A., Leahy, S.C., Reunanen, J., Bottacini, F., Claesson, M.J., *et al.* (2011) Functional genome analysis of *Bifidobacterium breve* UCC2003 reveals type IVb tight adherence (Tad) pili as an essential and conserved host-colonization factor. *Proceedings of the National Academy of Sciences of the United States of America*, **108**, 11217–11222.
- Pegden, R.S., Larson, M.A., Grant, R.J. and Morrison, M. (1998) Adherence of the Gram-positive bacterium *Ruminococcus albus* to cellulose involves a cellulose-binding protein which belongs to the Pil family of proteins. *Journal of Bacteriology*, **180**, 5921–5927.
- Pellicic, V. (2008) Type IV pili: *e pluribus unum?* *Molecular Microbiology*, **68**, 827–837.
- Piepenbrink, K.H., Maldarelli, G.A., de la Pena, C.F., Mulvey, G.L., Snyder, G.A., De Masi, L., *et al.* (2014) Structure of *Clostridium difficile* PilJ exhibits unprecedented divergence from known type IV pilins. *Journal of Biological Chemistry*, **289**, 4334–4345.
- Piepenbrink, K.H., Maldarelli, G.A., Martinez de la Pena, C.F., Dingle, T.C., Mulvey, G.L., Lee, A., *et al.* (2015) Structural and evolutionary analyses show unique stabilization strategies in the type IV pili of *Clostridium difficile*. *Structure*, **23**, 385–396.
- Purcell, E.B., McKee, R.W., Bordeleau, E., Burrus, V. and Tamayo, R. (2016) Regulation of type IV pili contributes to surface behaviors of historical and epidemic strains of *Clostridium difficile*. *Journal of Bacteriology*, **198**, 565–577.
- Rakotoarivonina, H., Jubelin, G., Hebraud, M., Gaillard-Martinie, B., Forano, E. and Mosoni, P. (2002) Adhesion to cellulose of the Gram-positive bacterium *Ruminococcus albus* involves type IV pili. *Microbiology*, **148**, 1871–1880.
- Rodgers, K., Arvidson, C.G. and Melville, S. (2011) Expression of a *Clostridium perfringens* type IV pilin by *Neisseria gonorrhoeae* mediates adherence to muscle cells. *Infection and Immunity*, **79**, 3096–3105.
- Tala, L., Fineberg, A., Kukura, P. and Persat, A. (2019) *Pseudomonas aeruginosa* orchestrates twitching motility by sequential control of type IV pili movements. *Nature Microbiology*, **4**, 774–780.
- Tammam, S., Sampaleanu, L.M., Koo, J., Manoharan, K., Daubaras, M., Burrows, L.L., *et al.* (2013) PilMNOPQ from the *Pseudomonas aeruginosa* type IV pilus system form a transenvelope protein interaction network that interacts with PilA. *Journal of Bacteriology*, **195**, 2126–2135.
- Tomich, M., Planet, P.J. and Figurski, D.H. (2007) The *tad* locus: postcards from the widespread colonization island. *Nature Reviews in Microbiology*, **5**, 363–375.
- Varga, J.J., Nguyen, V., O'Brien, D.K., Rodgers, K., Walker, R.A. and Melville, S.B. (2006) Type IV pili-dependent gliding motility in the Gram-positive pathogen *Clostridium perfringens* and other *Clostridia*. *Molecular Microbiology*, **62**, 680–694.
- Varga, J.J., Therit, B. and Melville, S.B. (2008) Type IV pili and the CcpA protein are needed for maximal biofilm formation by the Gram-positive anaerobic pathogen *Clostridium perfringens*. *Infection and Immunity*, **76**, 4944–4951.
- Wang, F., Coureuil, M., Osinski, T., Orlova, A., Altindal, T., Gesbert, G., *et al.* (2017) Cryoelectron microscopy reconstructions of the *Pseudomonas aeruginosa* and *Neisseria gonorrhoeae* type IV pili at sub-nanometer resolution. *Structure*, **25**, 1423–1435.
- Winther-Larsen, H.C., Wolfgang, M., Dunham, S., van Putten, J.P., Dorward, D., Lovold, C., *et al.* (2005) A conserved set of pilin-like molecules controls type IV pilus dynamics and organelle-associated functions in *Neisseria gonorrhoeae*. *Molecular Microbiology*, **56**, 903–917.
- Zimmermann, L., Stephens, A., Nam, S.Z., Rau, D., Kubler, J., Lozajic, M., *et al.* (2018) A completely reimplemented MPI bioinformatics toolkit with a new HHpred server at its core. *Journal of Molecular Biology*, **430**, 2237–2243.