Structure and dynamic association of an assembly platform subcomplex of the bacterial type II secretion system

Régine Dazzoni, Yuanyuan Li, Aracelys López-Castilla, Sébastien Brier, Ariel Mechaly, Florence Cordier, Ahmed Haouz, Michael Nilges, Olivera Francetic, Benjamin Bardiaux, et al.

Structure and dynamic association of an assembly platform subcomplex of the bacterial type II secretion system

Régine Dazzoni[1], Yuanyuan Li[2], Aracelys López-Castilla[1], Sébastien Brier[3], Ariel Mechaly[4], Florence Cordier[1,3], Ahmed Haouz[4], Michael Nilges[1], Olivera Francetic[2], Benjamin Bardiaux[1*], Nadia Izadi-Pruneyre[1*]

[1] Institut Pasteur, Université Paris Cité, CNRS UMR3528, Structural Bioinformatics Unit, F-75015 Paris, France
[2] Institut Pasteur, Université Paris Cité, CNRS UMR3528, Biochemistry of Macromolecular Interactions Unit, F-75015 Paris, France
[3] Institut Pasteur, Université Paris Cité, CNRS UMR3528, Biological NMR and HDX-MS Technological Platform, F-75015 Paris, France
[4] Institut Pasteur, Université Paris Cité, CNRS UMR3528, Crystallography Platform, F-75015 Paris, France

* corresponding author: benjamin.bardiaux@pasteur.fr, nadia.izadi@pasteur.fr

Keywords: Type II Secretion System, Assembly platform, Type IV pili, protein-protein interaction, Ferredoxin-like fold.

Abstract (148 words)

Type II secretion systems (T2SS) allow diderm bacteria to secrete hydrolytic enzymes, adhesins or toxins important for growth and virulence. To promote secretion of folded proteins, T2SSs assemble periplasmic filaments called pseudopili or endopili at an inner membrane subcomplex or the assembly platform (AP). Here, we combined biophysical approaches, NMR and X-ray crystallography to study the *Klebsiella* AP components PulL and PulM. We determined the structure and associations of their periplasmic domains and describe the first structure of the heterodimer formed by their ferredoxin-like domains. We show how structural complementarity and plasticity favor their association during the secretion process. Cysteine scanning and cross-linking data provided additional constraints to build a structural model of the PulL–PulM assembly in the cellular context. Our structural and functional insights, together with the relative cellular abundance of its components, support the role of AP as a dynamic hub that orchestrates pilus polymerization.

Introduction

Gram-negative bacteria have developed multiple protein secretion systems that are important for their survival and pathogenesis (Maffei *et al*., 2017). Among these, the type II secretion system (T2SS), discovered in the late 1980s (d'Enfert *et al*., 1987), is one of the most widespread and at the same time most relevant, from biomedical and environmental standpoints. It allows the bacterium to secrete fully folded proteins with a wide range of functions – toxins, adhesins, cytochromes and hydrolytic enzymes (Cianciotto and White, 2017).

The T2SS is a transmembrane nano-machine composed of 12 to 14 proteins designated here by the Gsp (General secretory pathway) nomenclature (Pugsley, 1993). It is organized into four sub-complexes (reviewed in (Naskar *et al*., 2021)): (1) an outer membrane secretin channel (GspD); (2) a periplasmic filament called endopilus (previously called pseudopilus) composed of a non-covalent polymer of major pilin GspG and minor pilins GspH, I, J, K; (3) an assembly platform (AP) composed of GspL, M, C and F (Py *et al*., 2001) ; and (4) a cytosolic ATPase GspE forming a hexamer. The T2SS has a common evolutionary origin (Denise *et al*., 2019; Hobbs and Mattick, 1993; Peabody *et al*., 2003) with nanomachines involved in assembly of type IV pili (T4P) and archaeal flagella and pili and shares a similar architecture. Together they form the superfamily of type IV filament assembly systems (Berry and Pelicic, 2015). Since the T2SS pili are not extracellular, they have been called pseudopilus. However, we prefer to call them endopili.

Although the general architecture of the T2SS has been extensively investigated, the biogenesis of this nanomachine and the secretion mechanism are still largely unknown. Current models propose that the substrate secretion is coupled with the polymerization of the endopilus in the periplasm, which is driven by the assembly platform (AP) complex by an unknown mechanism (Naskar *et al*., 2021). Information on the interaction mode of AP components and their structure is therefore essential to understand this key step of the secretion process.

GspL and GspM are the peripheral AP components that share a similar domain organization, suggesting their common evolutionary origin. Both are predicted to insert into the inner membrane *via* a single hydrophobic helix that is followed by a periplasmic α-helical region and a C-terminal globular domain. X-ray crystallography has provided structural information for the periplasmic domains of GspM (Abendroth *et al*., 2005) and GspL (Abendroth *et al*., 2009; Fulara *et al*., 2018). Additionally, GspL has an N-terminal cytoplasmic domain, which forms a stable complex with the GspE ATPase motor, as demonstrated in the

Vibrio cholerae T2SS (Abendroth *et al*., 2005). This interaction anchors the GspE hexamer to the cytoplasmic base of the AP complex, and couples the energy of the ATPase motor to the endopilus polymerization. *In vivo*, GspM forms a complex with GspL, protecting it from degradation, as shown in *V. cholerae* (Sandkvist *et al.*, 1999) and studies in *Dickeya dadantii* highlight a dynamic nature of this complex (Lallemand *et al*., 2013). An electron-microscopy study of a purified T2SS subcomplex from *Klebsiella pneumoniae* suggested a C6 symmetry of the GspE-GspL-GspM complex (Chernyatina and Low, 2019). However, neither that study, nor the *in situ* analysis of the *Legionella pneumophila* T2SS by cryo-tomography (Ghosal *et al.*, 2019) were able to provide a clear view of the L and M complex architecture required to understand their molecular function.

Here we focused on two essential AP components GspL and GspM from *Klebsiella oxytoca* T2SS called respectively PulL and PulM, with reference to the lipoprotein pullulanase (PulA), the only identified exoprotein secreted by this system. This enzyme degrades branched maltotriose polymers, allowing them to be taken up and used as nutrient (d'Enfert *et al*., 1987). To understand the assembly of PulL and PulM, we first analyzed their periplasmic globular C-terminal domains (CTDs). We solved their structures by NMR and X-ray crystallography, and combined these approaches to determine the first structure of their heterodimer complex available to date. Mutational and functional studies based on this structure showed the key role of the heterodimer interface in the secretion process. To reach a comprehensive view of PulL-PulM assembly in the cellular and membrane context, we used bacterial two-hybrid (BACTH) and cysteine crosslinking. These data and the information on the relative abundance of PulL and PulM in bacterial cells, allowed us to propose a model of the dynamic association of AP proteins and of the way this drives the endopilus formation during protein secretion.

Results

Previous studies provide evidence for a direct interaction between the AP components PulL and PulM in *K. oxytoca* (Nivaskumar *et al*., 2016; Possot *et al*., 2000). In the *Dickeya dadantii* T2SS, pulldown, bacterial two-hybrid and cysteine crosslinking assays suggest that GspL and GspM interact *via* their periplasmic C-terminal domains (Lallemand *et al*., 2013). However, their assembly mode is still not understood at the molecular and atomic levels. Here, to study their structure and interaction, we produced and purified the predicted globular C-

terminal domains of PulL ($PulL_{CTD}$, residues 312-398) and PulM ($PulM_{CTD}$, residues 79-161) (See Figure S1 for details on the PulL and PulM domain organization).

$PulL_{CTD}$ structure

First, by using NMR and analytical ultracentrifugation (AUC), we determined the oligomerization state of $PulL_{CTD}$ at concentrations ranging from 10 to 300 µM (Figure S1 and Supplementary data). $PulL_{CTD}$ exists in a monomer-dimer equilibrium in solution depending on the protein concentration. At the concentration used for NMR experiments (300-400 µM), it is mainly found in dimeric form. Its ^{1}H-^{15}N HSQC spectrum was used for an initial structural analysis. In this spectrum, each signal arises from the backbone amide group of one residue. The number of observed signals in the ^{1}H-^{15}N HSQC spectrum of $PulL_{CTD}$ (82) is consistent with the number of its backbone amide groups (83), and thus with a symmetrical homodimer (Figure 1A). To solve the structure of the $PulL_{CTD}$ dimeric form, we first determined the structure of its monomer by NMR. An ensemble of 15 monomeric structures was calculated based on chemical shift assignments obtained previously (Dazzoni *et al.*, 2021) and NOESY experiments by using the program ARIA and CNS (Allain *et al.*, 2020; Brünger *et al.*, 1998) as detailed in Materials and Methods. The ensemble of $PulL_{CTD}$ monomer structures (Figure 1B) presents a backbone RMSD of 0.6 Å for its ordered parts (Table 1). $PulL_{CTD}$ displays a ferredoxin-like fold with 2 helices, α1 (S319-D332) and α2 (S359-R372), and an anti-parallel β-sheet formed by 4 β-strands, β1 (I336-D344), β2 (N348-A356), β3 (F373-Q376) and β4 (I387-G395) (Figure 1B).

To determine the structure of the dimeric form, $^{13}C/^{15}N$ edited-filtered NOESY experiments were performed to collect intermolecular distance restraints between two protomers of $PulL_{CTD}$. Unfortunately, no intermolecular cross-peaks could be observed in these experiments, although various conditions (temperature, magnetic field and pulse sequence) were tested. The absence of signal is most likely due to conformational exchange at the µs-ms NMR time scale. Such dynamic behavior has been already suspected for several residues displaying very low NMR signal intensity (Dazzoni *et al.*, 2021). When mapping these residues on the $PulL_{CTD}$ monomer structure (Figure 1C), it appears that they are mostly located on the same exposed face of the molecule (α1, β1 and β2) suggesting that they are at the dimer interface and affected by a conformational exchange between monomeric and dimeric forms.

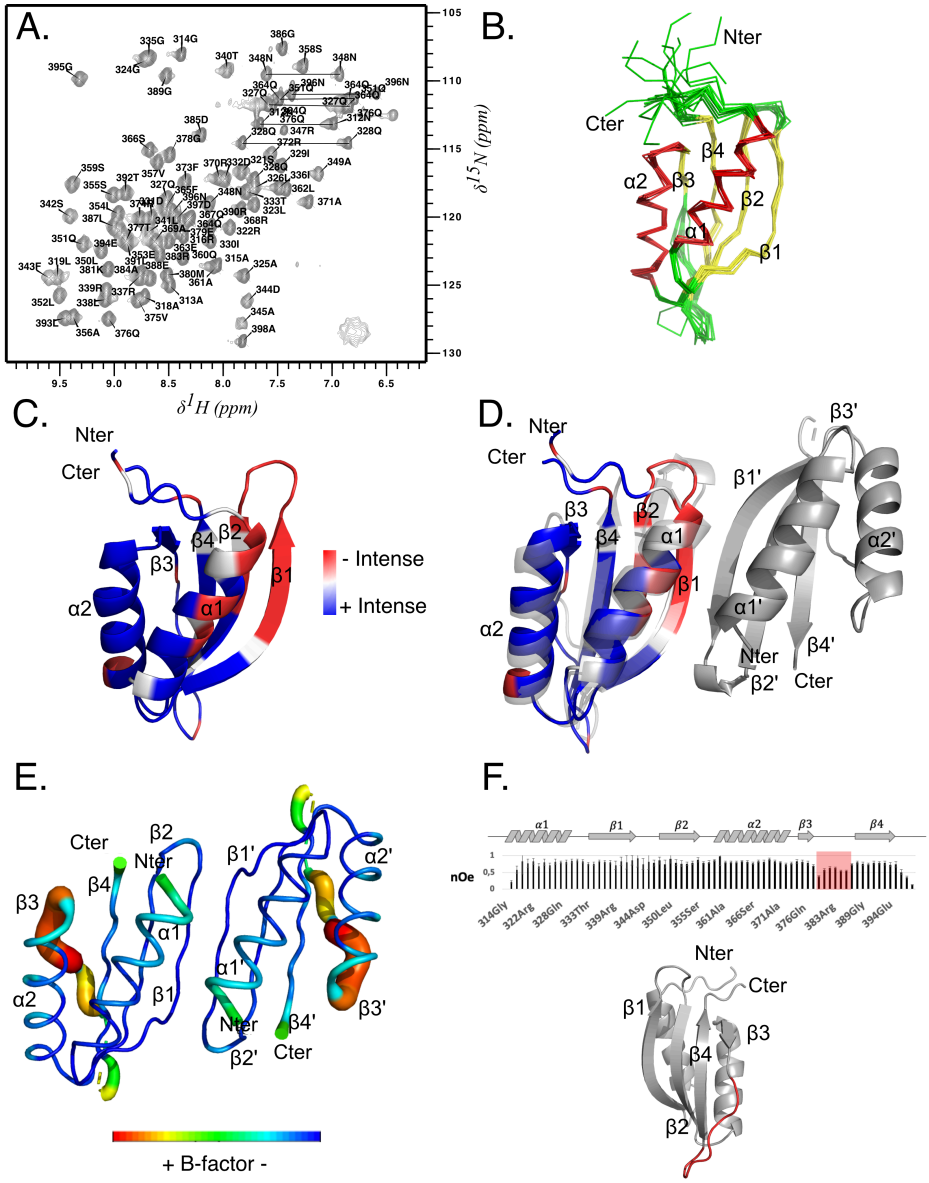

Figure 1. PulL$_{CTD}$ homodimer structure. ***A***. *^{1}H-^{15}N HSQC NMR spectrum of PulL$_{CTD}$ (400 µM in 50 mM HEPES buffer pH 6.5, 50 mM NaCl). Backbone resonance assignments are indicated in one-letter amino acid code and side chain NH_2 peaks of Asn (N) and Gln (Q) are connected by horizontal lines.* ***B.*** *NMR structure ensemble of PulL$_{CTD}$ exhibiting a ferredoxin-like fold (α1-β1-β2-α2-β3-β4; α-helix in red, β-sheet in yellow, loops and turns in green).* ***C.*** *Cartoon representation of the lowest-energy NMR structure of the PulL$_{CTD}$ monomer. The intensity of the ^{1}H-^{15}N HSQC signals of PulL$_{CTD}$ is reported with a specific colour code on the structure from blue to red for decaying peak intensity, outlining the likely homodimeric interface.* ***D.*** *Superposition of the PulL$_{CTD}$ NMR structure (coloured as in C) and the dimer structure of PulL$_{CTD}$ obtained by X-ray crystallography in grey (PDB ID: 8A9W).* ***E.*** *Representation of the B-factor per residue of PulL$_{CTD}$ in the X-ray crystallography dimer structure, from blue to red, with increasing B-factor.* ***F.*** *^{1}H–^{15}N heteronuclear NOE values along the PulL$_{CTD}$ sequence.*

The red box indicates the lower values corresponding to the most dynamic region (ps-ns time scale), located on the β3-β4 loop shown in red in the cartoon model.

We then used X-ray crystallography and solved the structure of $PulL_{CTD}$ at 1.8 Å resolution. The summary of data collection and refinement statistics is shown in Table 2. The structure of $PulL_{CTD}$ monomer exhibits a nearly identical ferredoxin-like fold (α1-β1-β2-α2-β3-β4) as observed in the NMR structure (Figure 1D). However, the β3 strand (residues F373-Q376) is not fully formed in the X-ray structure. As shown in Figure 1E, higher B-factor values are observed in this region compared to the rest of the protein, reflecting a disordered β-strand. Consistently, higher flexibility on the ps-ns time scale in this strand in solution is evidenced by the low $^{1}H–^{15}N$ heteronuclear NOE values (Figure 1F).

Only one possible dimer interface with C2 symmetry exists in the crystal, across two asymmetric units in the crystal lattice and with a buried surface area of 526 $Å^2$. This dimeric interface occurs between the α1 helix and the β1 strand oriented in an anti-parallel fashion (Figure 1D). Although the interaction surface is quite large relative to the size of the domain, only 6 hydrogen bonds are observed at the dimerization interface between residues of β1 and β1' belonging the other protomer: one between A345 and R339 and two between F343 and L341 (see also Figure 4G below). There are also a few interactions involving α1 helices, notably hydrophobic contacts between the I320 side-chains of each protomer. This small number of interprotein contacts explains that the assemblies are short-lived (Janin *et al.*, 2007) and hence the monomer/dimer exchange of $PulL_{CTD}$ in solution observed by NMR and AUC.

$PulM_{CTD}$ structure

$PulM_{CTD}$ behaves as a stable dimer in solution at low and high concentrations (Figure S1 and Supplementary data). To determine its structure, we performed NMR experiments and crystallization trials in parallel. We obtained crystals in two days, with an asymmetric unit containing seven virtually identical $PulM_{CTD}$ molecules (average RMSD 0.28 Å for Cα atoms). The final structure was refined at 1.5 Å of resolution (Figure 2A, B). The summary of data collection and refinement statistics is shown in Table 2. The structure of the $PulM_{CTD}$ protomer exhibits a ferredoxin-like fold (α1-β1-β2-α2-β3-β4) (Figure 2). The first helix α1 extends from P82 to H93, followed by strands β1 (V98-Q103), β2 (R106-V111), α2 helix (F116-A129), β3 (A134-A140) and β4 (V148-E156).

Among the molecular associations in the crystal, two possible dimer topologies of $PulM_{CTD}$ were found, named form A and form B. Form A was represented by four virtually superimposable anti-parallel dimers (average RMSD= 0.7 Å), with an interface involving the α2-β3 regions (Figure 2A). Form B was represented by a single homodimer with an α1-β1 interface and protomers oriented in a parallel fashion (Figure 2B).

These alternative dimeric forms A and B have, respectively, a buried surface area of 635 $Å^2$ and 453 $Å^2$ across the dimerization interface, as calculated by using the PISA server (Krissinel and Henrick, 2007). The A form is thus more favorable from a structural aspect, considering its larger buried interface and the fact that it is also the most abundant form in the crystal (80%).

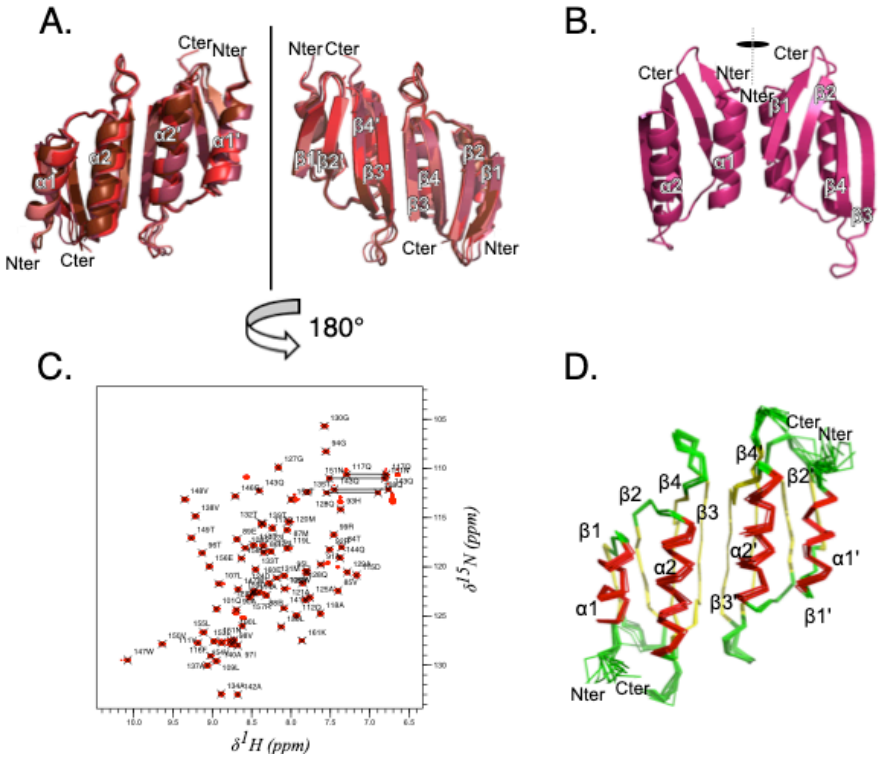

Figure 2. $PulM_{CTD}$ homodimer structure. ***A.*** *Main topology (form A) found in the $PulM_{CTD}$ crystal. The four homodimers observed in the asymmetric unit are superimposed.* ***B.*** *Minor topology (form B) found in the $PulM_{CTD}$ crystal, represented by a single homodimer. The vertical lines indicate the twofold symmetry axis.* ***C.*** *^{1}H-^{15}N HSQC NMR spectrum of $PulM_{CTD}$ in 50 mM HEPES pH 7.0, 50 mM NaCl. The backbone resonance assignments are indicated in one-letter amino acid code. Side chain NH_2 signals of Asn (N) and Gln (Q) are connected by horizontal lines.* ***D.*** *NMR structure ensemble of the $PulM_{CTD}$ homodimer. Helices are coloured in red, β-strands in yellow and turns/loops in green.*

Since crystallization conditions and crystal packing might bias these results, we used NMR spectroscopy to determine which dimeric form of $PulM_{CTD}$ exists in solution. A total of 1623 intramolecular and 25 intermolecular restraints were used to calculate the structure of $PulM_{CTD}$ dimer with a backbone RMSD of 0.69 Å for the 15 lowest energy structures (Figure 2D and Table 1). Each $PulM_{CTD}$ protomer exhibits a ferredoxin-like fold composed of 2 α helices and one β sheet as follows: α1 (T84-H93); β1 (V98-Q101); β2 (L107-V111), α2 (F116-A129), β3 (M131-A140) and β4 (V148-R157). The dimer interface occurs between the α2 helix and the β3 strand of the two protomers, oriented in an antiparallel fashion (Figure 2D). This dimer in solution is therefore the form A obtained by X-ray crystallography (RMSD of backbone atoms 1.6 Å between the NMR dimer structure and the crystallographic form A, Figure S2).

Structure of the $PulL_{CTD}$-$PulM_{CTD}$ complex

An analysis of the interactions between $PulL_{CTD}$ and $PulM_{CTD}$ gave us insight into their assembly and mode of association. Titration experiments of each protein with its unlabelled partner up to three times the molar ratio along with the rotational correlation time estimate highly suggested a stoichiometry of 1 $PulL_{CTD}$ for 1 $PulM_{CTD}$ (Figure S1A and B). For a more accurate characterization of the oligomeric state, we employed native mass spectrometry. Both $PulL_{CTD}$ (Figure 3A) and $PulM_{CTD}$ (Figure 3B) alone are detected as homodimers in accordance with their expected molecular weight (expected/measured MW: 19 069.48 Da/19 069.15 Da for $PulL_{CTD}$; 17 861.42 Da/17 861.87 Da for $PulM_{CTD}$). The mass spectrum of a mixture of $PulL_{CTD}$ and $PulM_{CTD}$ showed the presence of additional species with a well-resolved charge state series, from +7 (*m/z* 2638.9) to +9 (*m/z* 2052.7) (Figure 3C). The measured molecular weight (18 465.63 Da) fits perfectly with the expected mass of a $PulL_{CTD}$-$PulM_{CTD}$ heterodimer (18 465.95 Da). No higher oligomeric states were detected, confirming that $PulL_{CTD}$ and $PulM_{CTD}$ form heterodimers under our experimental conditions.

To further understand the $PulL_{CTD}$-$PulM_{CTD}$ interaction mode, we solved the crystal structure of the $PulL_{CTD}$-$PulM_{CTD}$ complex. A summary of the crystallographic parameters and data, as well as refinement statistics, is shown in Table 2. The asymmetric unit is organized in an arc-like arrangement where we found one $PulL_{CTD}$ molecule flanked on each side by one $PulM_{CTD}$ molecule with different orientations (Figure 3D). They form together a continuous anti-parallel β-sheet. This arrangement results in two different interfaces, either $PulL_{CTD}^{\alpha1\beta1} -$

$PulM_{CTD}^{\alpha2\beta3}$ or $PulL_{CTD}^{\alpha2\beta3} - PulM_{CTD}^{\alpha2\beta3}$. Surface interface analysis (Table S1) indicates that $PulL_{CTD}^{\alpha1\beta1} - PulM_{CTD}^{\alpha2\beta3}$ is more likely to represent a biologically relevant arrangement. In addition to the complex form both A and B forms of the $PulM_{CTD}$ dimers, as present in the $PulM_{CTD}$ crystal, are observed (Figure S3A).

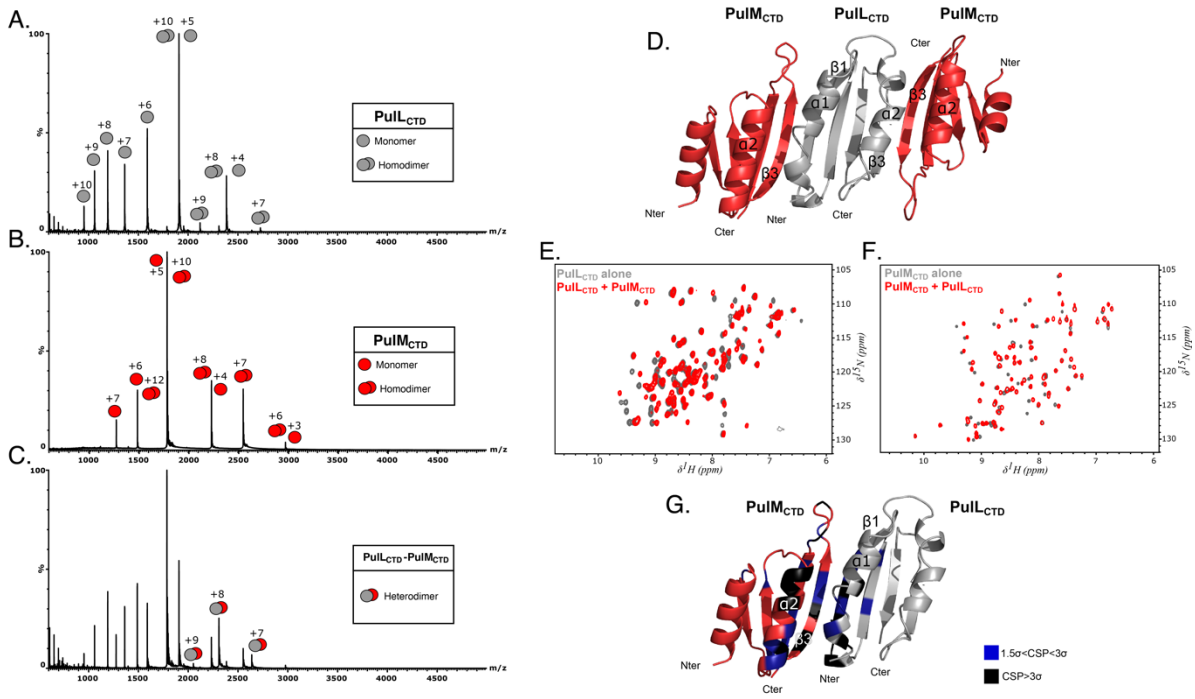

Figure 3. Oligomeric state and interactions of the $PulL_{CTD}$-$PulM_{CTD}$ complex. *Native mass spectra of $PulL_{CTD}$ (**A**) and $PulM_{CTD}$ (**B**) alone and in an equimolar mixture of $PulL_{CTD}$ and $PulM_{CTD}$ (**C**). From mass spectra we observe m/z corresponding to the presence of monomeric, homodimeric and heterodimeric species (annotated in grey for $PulL_{CTD}$ and in red for $PulM_{CTD}$). Charge states are indicated for all species. **D.** $PulL_{CTD}$-$PulM_{CTD}$ complexes in the X-ray crystallography asymmetric unit. **E.** Overlay of ^{1}H-^{15}N HSQC spectra of $PulL_{CTD}$ in the absence (grey) and presence (red) of unlabeled $PulM_{CTD}$. **F**. Overlay of ^{1}H-^{15}N HSQC spectra of $PulM_{CTD}$ in the absence (grey) and presence (red) of unlabeled $PulL_{CTD}$.. **G.** X-ray structure of the $PulL_{CTD}^{\alpha1\beta1} - PulM_{CTD}^{\alpha2\beta3}$ complex found in solution. The residues at the heterodimer interface displaying the highest CSP (>3 σ in black, >1.5 σ in blue) are colored.*

NMR analysis of the $PulL_{CTD}$-$PulM_{CTD}$ interaction in solution, and the spectral comparison between ^{15}N-$PulM_{CTD}$ and ^{15}N-$PulL_{CTD}$ each alone and in the presence of its unlabeled partner show also that this form $PulL_{CTD}^{\alpha1\beta1} - PulM_{CTD}^{\alpha2\beta3}$ is the one formed in solution

(Figure 3E-G). For $PulL_{CTD}$, residues displaying the highest chemical shift perturbations (CSP) (>1.5 σ above the mean CSP) belong to the α1 helix (L319, I320, S321 and L326); the β1-β2 strands (R337-A356) and the middle of β4 at the C-terminus (L391-G395) (Figure 3G and S3B). On $PulM_{CTD}$, the highest CSPs (>1.5 σ) are observed within the α2 helix (Q117-Q128), the β3 strand (G130-V138) and the β3-β4 loop (V141 to V148) (Figure 3G and S3C). As for the $PulL_{CTD}$ homodimer, due to the dynamics of this assembly and of the residues at the interface, no intermolecular NOE cross-peaks could be detected for $PulL_{CTD}$-$PulM_{CTD}$ mixtures.

Interestingly, the same structural elements and the same residues are involved in the homodimerization and heterodimerization interfaces either through hydrophobic contacts (Figure 4A-D) or hydrogen bonds (Figure 4E-H). While $PulM_{CTD}$ β3 and $PulL_{CTD}$ β1 are each involved in an antiparallel β sheet formation within the homodimers (Figure 4E-G), they are hydrogen bonded to form the heterodimer interface (Figure 4H, in green). Additional hydrogen bonds involving $PulM_{CTD}^{V141,P145,Q144}$-$PulL_{CTD}^{R337}$ (in blue in Figure 4H) and $PulM_{CTD}^{Q117}$-$PulL_{CTD}^{G324,Q327}$ (in magenta in Figure 4H) further stabilize the heterodimer complex.

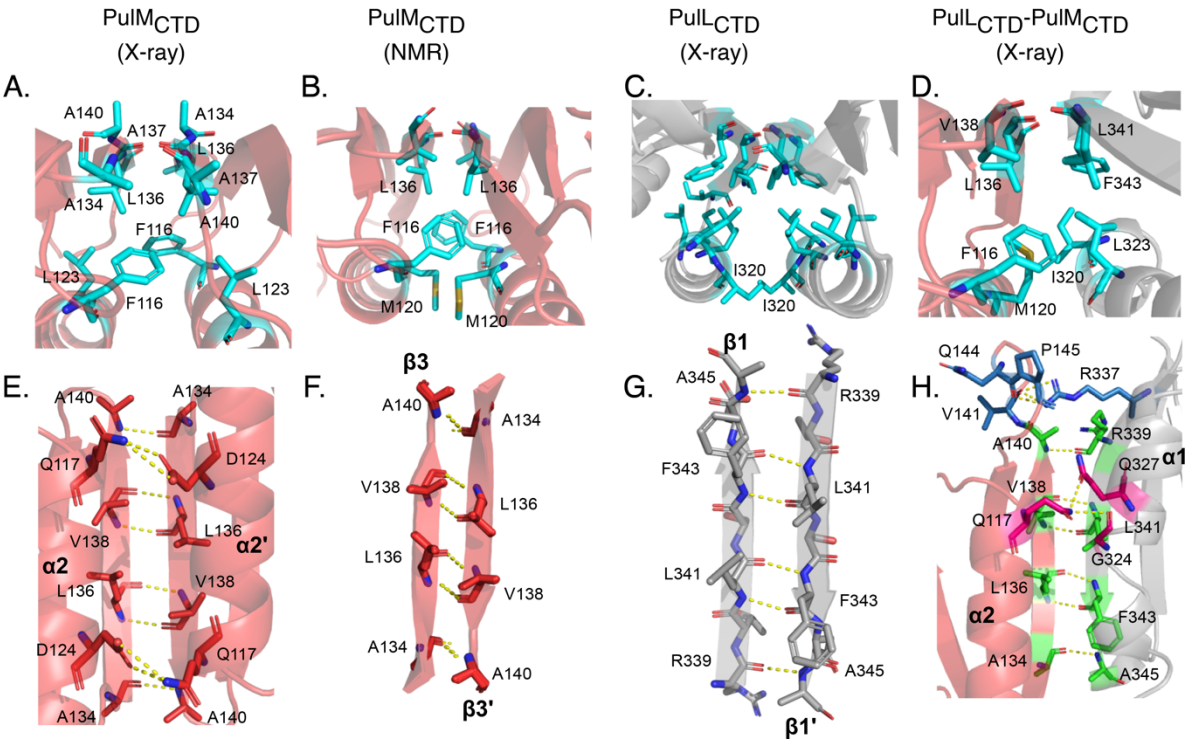

Figure 4. Comparison of the interfaces in homodimers of $PulL_{CTD}$ or $PulM_{CTD}$ and the $PulL_{CTD}$-$PulM_{CTD}$ heterodimer complex. A, B, C and D*. Hydrophobic interactions (cyan) in $PulM_{CTD}$ (X-ray structure), $PulM_{CTD}$ (NMR structure), $PulL_{CTD}$ (X-ray structure), and the*

heterocomplex $PulM_{CTD}$-$PulL_{CTD}$ (X-ray structure), from left to right. $PulM_{CTD}$ is colored in red and $PulL_{CTD}$ in grey. ***E, F, G and H.*** *Hydrogen bonds (yellow dotted lines) in $PulM_{CTD}$ (X-ray structure), $PulM_{CTD}$ (NMR structure), $PulL_{CTD}$ (X-ray structure), and the $PulM_{CTD}$-$PulL_{CTD}$ heterodimer complex (X-ray structure), from left to right. In panel H, residues forming hydrogen bonds between β strands are shown in green, residues forming hydrogen bonds between α-helices in magenta and residues hydrogen bonding in the loops in blue. PulL and PulM protomers are colored in grey and red, respectively.*

Functional validation of the $PulL_{CTD}$-$PulM_{CTD}$ interface

To evaluate the role of the $PulL_{CTD}$-$PulM_{CTD}$ association in protein secretion, we mutated the residues involved in the interface. We introduced single alanine substitutions in positions L319, L323, L341 and Q327 of PulL (Figure 5A, side chains shown as orange sticks) and tested the ability of these variants to restore pullulanase (PulA) secretion in *Escherichia coli* strain PAP7460 carrying the *pul* gene cluster with a non-polar *pulL* gene deletion in plasmid pCHAP8251. In this strain, wherein the *pul* genes are moderately overexpressed, complementation with $pulL^{WT}$ allele in pCHAP8258 (WT) restored PulA secretion to native levels (> 90%). The single Leu to Ala substitutions in PulL did not affect the function, replacing polar residue Q327 with alanine in $PulL^{Q327A}$ led to a small but significant defect in PulA secretion (Figure 5B), which was further exacerbated in the presence of L319A or L341A substitutions (Figure 5B). On the PulM side, single Ala substitutions of F116, Q117, M120 and D124 of the α2 helix, or L136 of the β3 strand, did not affect the function (Figure 5C). However, the double substituted variant $PulM^{F116A\text{-}L136A}$ showed a significant secretion defect, suggesting that the interface was weakened (Figure 5C). Replacing the surface exposed M120 by charged residues in $PulM^{M120D}$ and $PulM^{M120K}$ variants also significantly reduced secretion (Figure 5C). In comparison, charge inversions on the surface of the $PulM_{CTD}$ α1 helix, such as R88E and R92D did not show any effect, consistent with their position distal to the interface with $PulL_{CTD}$. Although we cannot exclude that these mutations affect interaction with other IM components of T2SS, their localization at the interface and the correlation between their effects with the more or less conservative amino acid changes argue for their direct effect on $PulL_{CTD}$-$PulM_{CTD}$ binding.

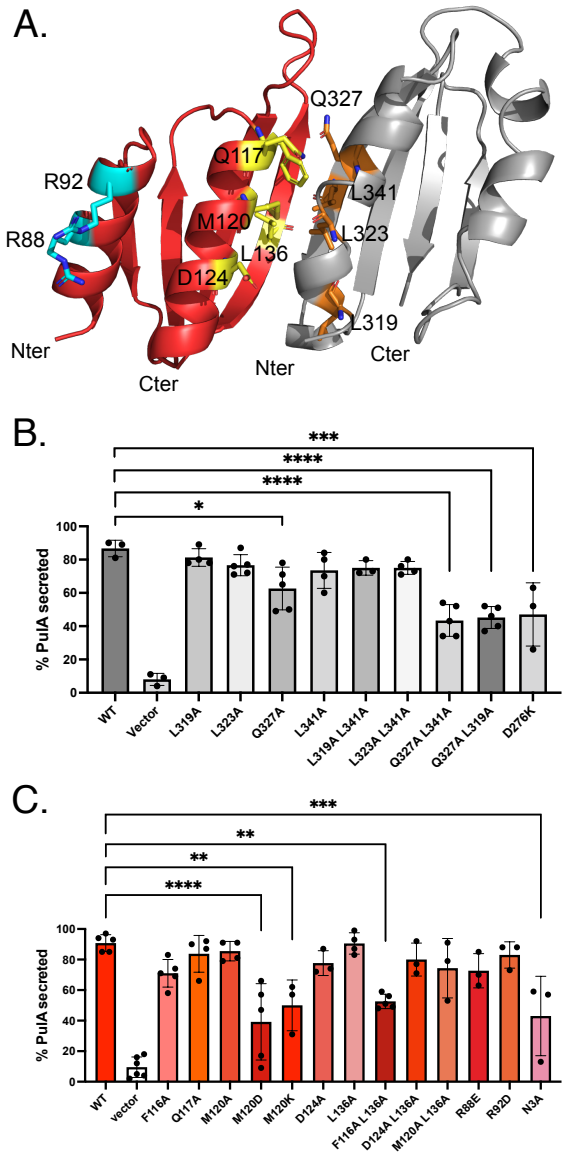

Figure 5. Functional characterization of the PulL$_{CTD}$-PulM$_{CTD}$ interface. ***A***. *Cartoon representation of the complex of PulL$_{CTD}$ (grey) with PulM$_{CTD}$ (red). The substituted residues are highlighted as sticks and marked with a single letter code, those at the PulL-PulM interfaces are colored in orange (PulL$_{CTD}$) or yellow (PulM$_{CTD}$), and those far from the interface are colored in cyan.* ***B***. *PulA secretion in strain PAP7460 carrying plasmid pCHAP8251 (*ΔpulL*) complemented with pCHAP8258 encoding PulL (WT) or its derivatives encoding mutant PulL variants (Table S2). Details of the secretion assays are provided in Materials and Methods. The bar graph heights represent the mean values and dots represent the percentage of secreted PulA from independent experiments (n ≧ 3). One-way ANOVA and multiple comparisons were done with GraphPad PRISM 9. Statistically significant differences relative to the wild-type PulL are indicated.* ***C.*** *PulA secretion in strain PAP7460 carrying plasmid pCHAP8496* (ΔpulM) *complemented with pCHAP1353 encoding WT PulM or its mutant derivatives*

encoding indicated PulM variants (Table S2). Data representation and analysis were performed as in (B).

Together, these data show that the intact $PulL_{CTD}$-$PulM_{CTD}$ interface, and thus the PulL-PulM assembly, play an important role in the secretion. Possibly, other PulL and PulM regions might also contribute to this interaction *in vivo*. One indication is the $PulM^{N3A}$ variant mapping in the cytoplasmic N-terminus of the protein, which dramatically affected the stability of both proteins (Figure S4) and was defective in secretion (Figure 5C).

Interactions of full-length membrane-anchored PulL and PulM

To gain further insights into the interaction between full-length PulL and PulM in the cellular context and in the presence of the membrane, we combined bacterial adenylate cyclase two-hybrid system (BACTH) and cysteine cross-linking approaches. The BACTH approach allows us to analyze interactions between membrane proteins in their native environment (Karimova *et al.*, 1998). To study how the full-length PulL and PulM interact, we fused them to the C-terminal ends of T18 and T25, fragments of the *Bordetella pertussis* adenylyl cyclase (CyaA) catalytic domain. The resulting plasmids (listed in Table S2) were introduced in the *E. coli* strain DHT1 (Dautin *et al.*, 2000) carrying a deletion of the endogenous *cya* gene. The CyaA T18 and T25 fragments do not interact and do not restore adenylyl cyclase activity as indicated by white colonies on indicator plates (Figure 6A) and low expression levels of the chromosomal *lacZ* gene (Figure 6B). Bacteria containing T18-PulL and T25-PulL hybrids also showed white colonies and low activity, comparable to the negative controls, indicating that the full-length membrane-embedded PulL does not form homodimers. Bacteria carrying the T18-PulM and T25-PulM chimera were pale blue on X-gal indicator and showed somewhat higher activity, indicating a tendency of PulM to homodimerize (Figure 6). Finally, bacteria co-producing PulL and PulM chimera showed strong and highly significant interaction signals. These results show that full-length PulL and PulM interact strongly with each other and preferentially form heterodimers *in vivo*.

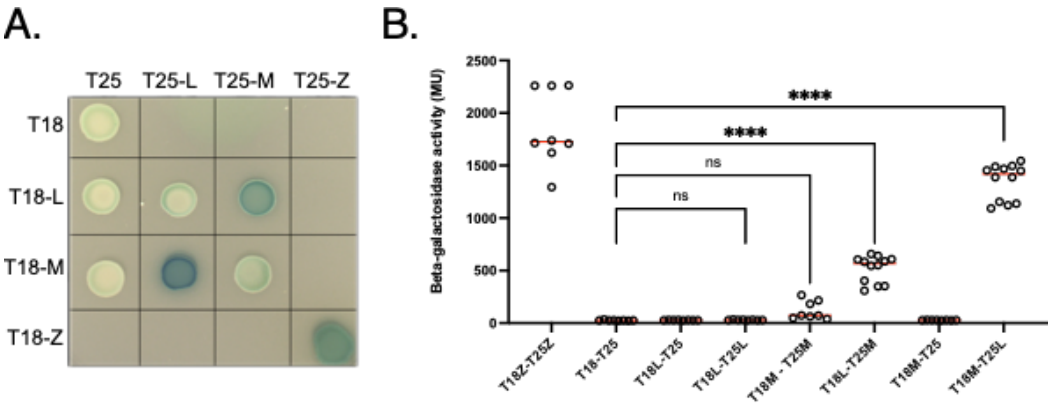

Figure 6. BACTH analysis of PulL and PulM interactions. ***A.*** *Plate assay and* ***B***. *Beta-galactosidase activity of E. coli strains DHT1 producing T18, T25 or their chimera with yeast leucin Zipper (Z) (as positive control), PulL (L) or PulM (M). The β-galactosidase activities measured for independent bacterial cultures (n ≧ 8) are plotted as dots and horizontal red lines show median values. Statistical analysis was performed with One-way ANOVA and multiple comparisons test, and plotted with GraphPad Prism 9 software.*

PulL and PulM are type I membrane proteins with an N-in C-out orientation, each containing a single transmembrane (TM) segment. To investigate the involvement of these segments in the PulL-PulM interaction, we used cysteine crosslinking. Native PulM has one Cys residue in position 17, within its TM segment. We first replaced this residue by a Leu and used PulM^{C17L} as a starting construct to generate a series of PulM variants with single Cys substitution at positions 18 through 24. In a similar manner, we substituted Cys264, localized in the periplasmic part of PulL (PulL^{C264L}) and used it to generate single Cys substitutions at positions 249 to 256 of the TM segment. Bacteria producing all the combinations of single Cys variants of PulL and PulM single Cys variants were treated with $CuCl_2$ as an oxidant and their total extracts were analyzed by Western blot with anti-PulL or anti-PulM antibodies (Figure 7). The PulL^{C264L} variant did not produce any significant level of crosslinked species and a similar pattern was observed in the presence of single Cys residues at positions 249, 252, 253, 254, 255 and 256 (Figure 7). However, PulL^{C264L} with Cys at positions 250 and 251 produced some homodimers and also heterodimers, notably with PulM with Cys at positions 17 (PulM^{WT}) and 18 (PulM$^{C17L-L18C}$). The latter variant gave the most intense heterodimer bands, which were also detected with anti-PulM antibodies (Figure 7). The PulM variants with Cys at positions 17, 18 and 19 also produced homodimer bands. These results show that PulL and PulM can interact *via* their transmembrane segments and that in this region they also use the same interface for homo- and heterodimerization. The levels of PulM homodimers were not affected by the

presence of PulL^{I250C} and PulL^{V251C}, possibly because PulM was more abundant than PulL in these strains. Consistent with a higher fraction of PulL crosslinked with PulM, the L-M heterodimers were more readily detectable in anti-PulL Western blots compared to those with anti-PulM. Confirming the stabilizing role of PulM, the levels of all PulL variants were significantly reduced in the absence of PulM (Ø). In contrast, the levels of PulM did not change in the absence of PulL (Figure 7, last panel, Ø).

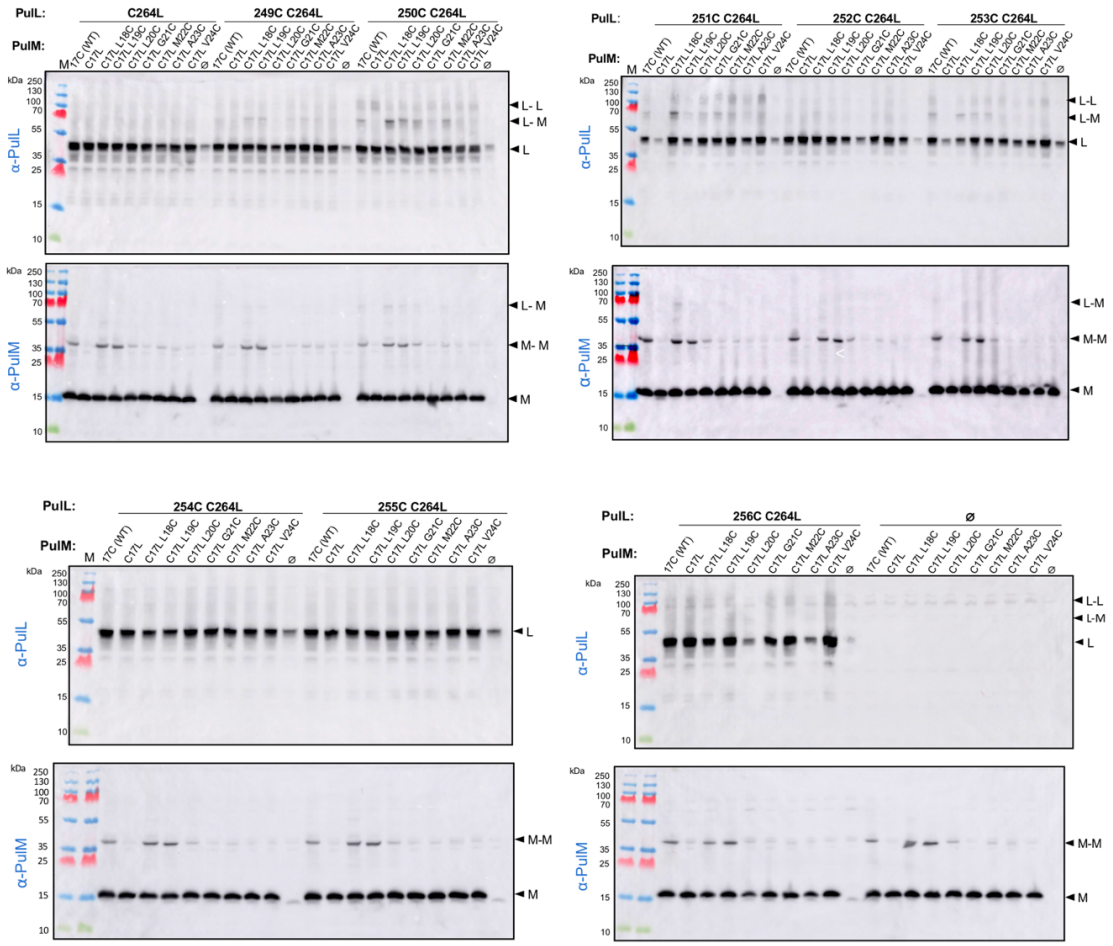

Figure 7. Cysteine scanning and crosslinking of PulL and PulM transmembrane segments. *Bacteria of strain PAP7460 producing the indicated PulL and PulM variants, or empty vector (Ø) were oxidized with $CuCl_2$ and the total extract from 0.05 OD_{600nm} of bacteria was analyzed by SDS-PAGE and Western blot. Molecular weight markers are shown on the left, and migration of PulL (L) and PulM (M) monomers, homodimers (L-L, M-M) and heterodimers (L-M) are indicated on the right. The same total fractions were analyzed with anti-PulL antibody (top) and anti-PulM antibody (bottom).*

The above experiments were performed in bacteria producing PulL and PulM in the absence of other T2SS components and at levels which may differ from those in the functional system. In the *pulC-O* operon encoding the *Klebsiella* T2SS, the *pulL* and *pulM* genes are adjacent and transcribed from the same promoter. To determine their relative abundance, we quantified the PulL and PulM amounts in bacteria producing the functional T2SS from a single, moderate copy-number plasmid pCHAP8185 (Table S2). We used Western blot analysis with antibodies raised against $PulL_{CTD}$ and $PulM_{CTD}$ to quantify the level of these proteins in bacterial extracts expressing the T2SS genes and compared these signals with those known amounts of purified $PulL_{CTD}$ and $PulM_{CTD}$ (Figure S5). These measurements indicated that PulM is present in large excess over PulL, with a molar ratio of about 20:1.

Combining the structural data on the $PulL_{CTD}$-$PulM_{CTD}$ heterodimer with cysteine crosslinking experiments allowed us to build a 3D model of the membrane-anchored and periplasmic region of the PulL-PulM complex (Figure 8A, B). Sequence analysis revealed the presence of long periplasmic α-helices in both PulL and PulM, encompassing the transmembrane segments and upstream of the CTDs. Additionally, two regions involved in the formation of coiled-coil in both proteins are predicted, the first one overlapping with the transmembrane segments (Figure S6A-C). Consequently, periplasmic helices of PulL and PulM were modeled as a coiled-coil, in which knobs-into-holes packing was confirmed by SOCKET2 analysis (Kumar and Woolfson, 2021) (Figure S6D). In the model, the relative alignment of the predicted transmembrane segments is compatible with an orthogonal orientation of the connecting periplasmic helices with respect to the membrane. The coiled-coil formed by the periplasmic helices of PulL and PulM is followed by short disordered linkers connecting the α1 helices of the CTDs heterodimer. In the model of the complex, a potential salt-bridge is found between D276 of PulL and R42 of PulM (Figure 8B). Interestingly, the $PulM^{D276K}$ variant displayed significantly reduced secretion (Figure 5B) which could indicate that the coiled-coil also contributes to the overall interaction surface in the PulL-PulM complex.

Discussion

Here we report the first high-resolution structure of the PulL-PulM AP subcomplex from a bacterial T2SS and investigate its assembly in solution and *in vivo*. The structural data obtained here could be relevant for the PulL and PulM orthologues PilN and PilO in T4P assembly systems, for which no complex structures are not yet available. Due to the dynamics of PulL, PulM and their association, their structural study was only possible by an integrative

approach. By combining native mass spectrometry, NMR and X-ray crystallography we show that while $PulL_{CTD}$ and $PulM_{CTD}$ each form a homodimer, they assemble as a heterodimer. The complementarity between NMR and X-ray crystallography was crucial for the structural study of such dynamic assemblies. While multiple molecules in the crystal asymmetric unit displayed several protein-protein interfaces, NMR gave us unambiguous insights on the dimerization interfaces in solution for both the isolated proteins and their assembly. Functional assays confirm that *in vivo* the dimerization interface and thus the $PulL_{CTD}$-$PulM_{CTD}$ association play an important role in the secretion process.

Intriguingly, the same interfaces and residues are involved in both homo- and heterodimerization, however resulting in different orientations of the subunits. These different parallel and antiparallel topologies were also observed in homodimers of other GspL, GspM and their orthologues, indicating the high plasticity of their interfaces and their need for stabilization in dimeric form (Figure S7, S8). For instance, in T4P assembly systems, the CTD of the PulM orthologue PilO forms homodimers, however with β-sheets facing the opposite sides in the two protomers compared to PulM (Leighton *et al.*, 2018; Sampaleanu *et al.*, 2009). The distribution of conserved residues in PulL and PulM homologs is remarkably consistent with the respective dimerization interfaces of their C-terminal domains (Figure S9). The ferredoxin-like fold of $PulL_{CTD}$ and $PulM_{CTD}$ and their orthologues explains well their structural complementarity and their dynamic assembly mode, both as homodimer and as heterodimer. The ferredoxin-like fold is highly favorable for molecular association. Its symmetrical repetition has been proposed by Eck and Dayhoff (known as Dayhoff's theory) as the result of the repetition of simple peptides to form a protein and would represent the protein ancestor (Alva and Lupas, 2018; Eck and Dayhoff, 1966). $PulL_{CTD}$ and $PulM_{CTD}$ indeed exhibit an internal pseudo-symmetry, where the two faces of the ferredoxin-like fold (α1-β1-β2 and α2-β3-β4) are structurally superimposable and present high level of sequence similarity (64% for $PulM_{CTD}$ and 66% for $PulL_{CTD}$). Interestingly, the dimerization interfaces of $PulL_{CTD}$ and $PulM_{CTD}$ are also highly similar, at the sequence and structural levels, with overlapping residue positions involved in the interactions (Figure S10).

The $PulL_{CTD}$ and $PulM_{CTD}$ homodimers associate in an antiparallel fashion, whereas the $PulL_{CTD}$-$PulM_{CTD}$ complex is parallel and thus more compatible with the membrane insertion of the proteins where both N-termini are found on the same side of the dimer (Figure 8). This interaction mode is consistent with the BACTH data in which full-length membrane-inserted proteins form heterodimers in parallel arrangement, and with previous data showing that the formation of GspL–GspM heterodimers in *D. dadantii* is favored over their homodimerization.

During the revision of this manuscript, we became aware of a recent work using molecular docking and proposing a similar arrangement of the GspL–GspM ferredoxin-like domains from *E. coli* O18:K1:H7, but as a tetramer (Pacheco Gutierrez, 2021). Although the antiparallel homodimerization of $PulL_{CTD}$ and $PulM_{CTD}$ would require a tangled arrangement with the respect to the membrane (Figure 8A), we cannot exclude that it might occur during the process. The long coiled-coil regions and flexible linkers that connect them to the CTDs could provide additional plasticity in the periplasmic regions of the proteins allowing this tangled association.

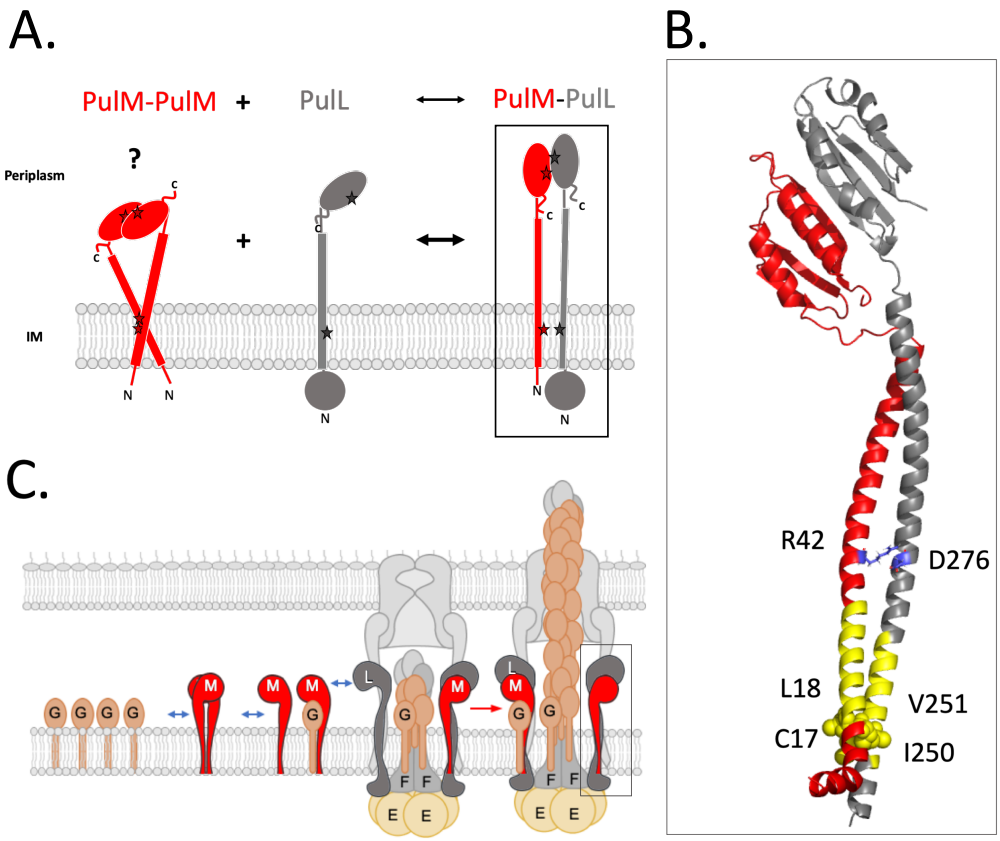

Figure 8. Full length PulL-PulM interaction model. ***A****. Schematic view of the heterodimerization process of PulL (grey) and PulM (red). PulM can behave as a homodimer (BACTH data) via interfaces at the C-terminal domains in the periplasm (NMR data), and in the inner membrane (IM) via their transmembrane domains (crosslinking data). PulL is mainly in monomeric form (BACTH data) and promotes PulL-PulM heterodimerization via the same interface (NMR and crosslinking data). Stars indicate the regions of interaction of PulL and PulM, in homo- and heterodimeric forms.* ***B.*** *3D model of the PulL-PulM complex based on the* $PulL_{CTD}$*-*$PulM_{CTD}$ *heterodimer structure (PDB ID: 8AB1) and crosslinking data, both obtained in this study and on the structure of the orthologue PilN-PilO complex (PDB ID: 3JC8). PulL*

is colored in grey, PulM in red, transmembrane segments in yellow. Residues close in space as revealed by crosslinking are annotated on the structural model. ***C.*** *Model of full-length PulL and PulM in the context of the T2SS. The highly abundant PulG endopilus subunit interacts with PulM. By binding to PulL, PulM targets PulG to the PulL-PulE-PulF complex. PulG, the endopilus subunit, is released from PulM and added to the growing endopilus.*

In addition to the $PulL_{CTD}$-$PulM_{CTD}$ interface, native full-length forms of PulL and PulM also interact *via* their transmembrane segments, as shown by Cys scanning experiments. Like for the CTDs, we found an overlap between homo- and heterodimer interfaces involving the transmembrane segments (Figure 7). This is consistent with the previously observed competition between L-M homo- and heterodimerization (Lallemand *et al.*, 2013). The relative abundance of these proteins *in vivo* is likely to determine whether homo- and heterodimers coexist and exchange. In most of the T2SS models, GspL and GspM are shown with a 1:1 stoichiometry, which finds confirmation in the purified T2SS subcomplex (Chernyatina and Low, 2019). However, in this study we showed that in the cellular context, when the whole system is expressed from endogenous promoters, the amount of PulM per cell is about 20 times higher than that of PulL (Figure S5). Higher cellular levels of PulM compared to PulL have also been observed by fluorescence microscopy using the GFP (green fluorescent protein) chimera fused to their N-terminal ends (Buddelmeijer *et al.*, 2006). Considering the static role of PulL, which serves as an anchor for the PulE ATPase, the higher abundance of PulM would ensure its dynamic role in PulG targeting (Nivaskumar *et al.*, 2016; Santos-Moreno *et al.*, 2017), which requires rapid exchange of its binding partners. The structural flexibility and complementarity of $PulL_{CTD}$ and $PulM_{CTD}$ is likely to facilitate this exchange during the highly dynamic endopilus assembly and turnover. The homodimerization would be more rapid than the heterodimerization, which is more stable, similarly to Jun-Fos proteins where the homodimers are formed with a favorable kinetics, while the heterodimer is more stable, both forms are found *in vivo* and *in vitro* (O'Shea *et al.*, 1992).

We have previously shown that PulG proteins that are assembled to form the endopilus (Lopez-Castilla *et al.*, 2017) interact with PulM *via* their transmembrane regions (Nivaskumar *et al.*, 2016; Santos-Moreno *et al.*, 2017). By binding to PulL, PulM might therefore channel PulG to the assembly site defined by the PulE-PulL complex (Figure 8B). The higher abundance of PulM is consistent with its role in targeting of the even more abundant PulG subunits (Santos-Moreno *et al.*, 2017) to ensure their insertion into the growing endopilus. While PulG does not directly interact with PulL (Nivaskumar *et al.*, 2016), crosslinking studies in *V. cholerae*

showed their physical proximity (Gray *et al.*, 2011) supporting the existence of either a ternary complex or their rapid exchange of partners. Rapid influx of pilins would ensure efficient incorporation into helical filaments, thought to be the driving force for protein secretion. In addition to their role in endopilus assembly, the PulL and PulM homologs have been implicated in interactions with the secreted substrate in the *P. aeruginosa* T2SS (Michel-Souzy *et al.*, 2018). Together, these data suggest that the PulL-PulM complex play a major role in the dynamic coupling of endopilus assembly and protein transport. Future studies are needed to address the influence of the pilin and the secreted substrate on the AP assembly and to define the sequence of events during pilin and substrate targeting to the secretion complex.

STAR Methods

Key resources table

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Anti-MalE-PulL rabbit polyclonal antibodies	(Possot *et al.*, 2000)	N/A
Anti-MalE-PulM rabbit polyclonal antibodies	(Possot *et al.*, 2000)	N/A
Anti-PulA rabbit polyclonal antibodies	(d'Enfert *et al.*, 1987)	N/A
Goat anti-Rabbit antibody coupled to HRP	Amersham	
Bacterial Strains		
Escherichia coli DH5α F'*lacI*^{Q}	Laboratory collection	N/A
Escherichia coli BL21(DE3) Star	Novagen	Cat# 69450
Escherichia coli PAP7460	(Possot *et al.*, 2000)	N/A
Escherichia coli DHT1	(Dautin *et al.*, 2000)	N/A
Biological Samples		
Escherichia coli strain PAP7460 carrying plasmids pCHAP8251 and pCHAP8258 or derivatives listed in Table S2	This study	N/A
Escherichia coli strain PAP7460 carrying plasmids pCHAP8496 and pCHAP1353 or derivatives listed in Table S2	This study	N/A
Escherichia coli strain DHT1 carrying plasmids pKT25 and pUT18C or derivatives listed in Table S2	This study	N/A
Chemicals, Peptides, and Recombinant Proteins		
T4 DNA ligase	New England Biolabs	
pMalE-His-TEV-PulM	This study	N/A
pMalE-His-TEV-PulL	(Dazzoni *et al.*, 2021)	N/A
Q5 High Fidelity Polymerase	New England Biolabs	Cat# M0491
*Kpn*I-HF restriction enzyme	New England Biolabs	R3142L
*Eco*RI-HF restriction enzyme	New England Biolabs	R3010L
Tricine	Sigma-Aldrich	Cat# T5816
$CuCl_2$	Sigma-Aldrich	Cat# C3279
Isopropyl β-D-1-thiogalactopyranoside (IPTG)	Fischer Scientific	Cat# BP1755
o-Nitrophenyl- β -D-thiogalactoside (ONPG)	Sigma-Aldrich	Cat# N1127
Carbenicillin (Ap)	Sigma-Aldrich	Cat# C1389

Kanamycine (Km)	Sigma-Aldrich	Cat# K-4000
Chloramphenicol (Cm)	Sigma-Aldrich	Cat# C-0378
Sodium-dodecyl sulfate (SDS) 20% solution	Interchim	Cat# UP896826
D2O	Eurisotop	D214F
Critical Commercial Assays		
Miniprep kit	Qiagen	Cat#: 27104
PCR purification kit	Qiagen	Cat#: 28104
Pierce®ECL 2 Western Blotting Substrate	Fisher Scientific	Cat# : 993PT
Pierce®ECL Western Blotting Substrate	Thermo Scientific	Cat# : 32109
Deposited Data		
$PulL_{CTD}$ NMR resonance assignments	(Dazzoni *et al.*, 2021)	BMRB: 50966
$PulM_{CTD}$ NMR resonance assignments	This study	BMRB: 34719
NMR structure ensemble of dimeric $PulM_{CTD}$	This study	PDB: 7ZE0
X-ray crystallographic structure of $PulM_{CTD}$	This study	PDB: 8A9X
X-ray crystallographic structure of $PulL_{CTD}$	This study	PDB: 8A9W
X-ray crystallographic structure of $PulL_{CTD}$-$PulM_{CTD}$	This study	PDB: 8AB1
Pullulanase secretion assay dataset shown in Figure 5	This study	doi: 10.5281/zenodo.7178734
Oligonucleotides		
See Table S3 for list of primers used in this study	N/A	NA/
Recombinant DNA		
See Table S2 for list of plasmids used in this study	N/A	N/A
Software and Algorithms		
PRISM version 9	GraphPad	https://www.graphpad.com/scientific-software/prism/
Image J	(Abramoff *et al.*, 2004)	https://imagej.nih.gov/ij/
TOPSPIN 3.6.1	Bruker Biospin.	N/A
CcpNmr Analysis 2.4	(Vranken *et al.*, 2005)	https://www.ccpn.ac.uk/
ARIA 2.3	(Rieping *et al.*, 2007)	http://aria.pasteur.fr
ARIAweb	(Allain *et al.*, 2020)	http://ariaweb.pasteur.fr
CNS 1.2	(Brünger *et al.*, 1998)	http://cns-online.org
TALOS-N	(Shen and Bax, 2013)	http://spin.niddk.nih.gov/bax/software/TALOS-N
Modeller 9.24	(Sali and Blundell, 1993)	https://salilab.org/modeller/
Pymol 2	Schrödinger, LLC	https://pymol.org/
UCSF Chimera 1.13	(Pettersen *et al.*, 2004)	https://www.cgl.ucsf.edu/chimera/
PROCHECK-NMR	(Laskowski *et al.*, 1996)	https://www.ebi.ac.uk/thornton-srv/software/PROCHECK/
PSVS 1.5	(Bhattacharya *et al.*, 2007)	https://montelionelab.chem.rpi.edu/PSVS/PSVS/
MassLynx 4.1	Waters	N/A
Phaser	(McCoy *et al.*, 2007)	https://www.phaser.cimr.cam.ac.uk/

Coot	(Emsley and Cowtan, 2004)	https://www2.mrc-lmb.cam.ac.uk/personal/pemsley/coot/
Buster	(Bricogne *et al.*, 2011)	http://www.globalphasing.com/buster/
DeepCoil	(Ludwiczak *et al.*, 2019)	https://github.com/labstructbioinf/DeepCoil
PSIPRED	(Jones, 1999)	http://bioinf.cs.ucl.ac.uk/psipred/
SOCKET2	(Kumar and Woolfson, 2021)	http://coiledcoils.chm.bris.ac.uk/socket2/home.html
PISA server	(Krissinel and Henrick, 2007)	https://www.ebi.ac.uk/pdbe/pisa/

Resource availability

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Nadia Izadi-Pruneyre (nadia.izadi@pasteur.fr).

Materials availability

Plasmids and strains produced in this study are available from the authors and detailed in Table S2.

Data and code availability

Atomic coordinates and structure factors have been deposited in the RCSB Protein Data Bank under the accession codes 8A9W ($PulL_{CTD}$), 7ZE0 ($PulM_{CTD}$, NMR), 8A9X ($PulM_{CTD}$, X-ray) and 8AB1 ($PulL_{CTD}$-$PulM_{CTD}$). PulM $^{13}C/^{15}N/^{1}H$ resonance assignments were deposited to the BMRB under accession number 34719.

Experimental model and subject details

Escherichia coli strains used in this study are listed in Key resources table. They were cultured in M9 or LB medium (Miller, 1972) at temperatures indicated in the corresponding Methods sections. The genes encoding the type II secretion system components, including PulL and PulM proteins studied here originate from the *Klebsiella oxytoca* strain UNF5023 (formerly classified and *Klebsiella pneumoniae*) described in (d'Enfert *et al*., 1987).

Method details

Plasmid constructions

Escherichia coli K-12 strain DH5α F'*lacI*^{Q} was used as a host for cloning purposes. The plasmids used in this study are listed in Table S2. To construct plasmids pMS1222 and pMS1229, the *pulL* gene was PCR-amplified from plasmid pCHAP8258 as template using primers PulL Kpn 5 and PulL Eco 3 with the high-fidelity Q5 DNA polymerase (*New England Biolabs*). The PCR products were purified on a Qiaquick spin column, digested with *Eco*RI and *Kpn*I restriction enzymes (NEB) and ligated with the *Eco*RI - *Kpn*I digested plasmids pUT18C and pKT25. Site-directed mutagenesis to generate derivatives of pCHAP8258 and pCHAP1353 was performed using a two-step amplification with perfectly overlapping mutagenic primers. Two separate reactions with single mutagenic primers were performed for 6 cycles using the Q5 DNA polymerase under conditions recommended by the manufacturer. The reactions were mixed and additional 15 cycles of amplification were performed. The PCR reactions were treated with *Dpn*I and transformed into DH5α F'*lacI*^{Q} ultracompetent cells. The purified plasmids were verified by DNA sequencing *(Eurofins*). The list of oligonucleotides (*Eurofins*) is shown in Table S3.

Proteins $PulL_{CTD}$ and $PulM_{CTD}$ were produced in *E. coli* BL21 (DE3) under control of the T7 promoter. Bacteria were grown at indicated temperatures in LB medium or in minimal M9 medium (Miller, 1972). Antibiotics were added as required at following concentrations: ampicillin (Ap) 100 μg.mL^{-1}, chloramphenicol (Cm), 25 μg.mL^{-1} and kanamycin (Km) at 25 μg.mL^{-1}. When required, the genes under p*lacZ* control were induced with 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG).

Production of unlabeled and isotope-labeled proteins

Derived pMalP2 vectors were used to express either the CTD of PulM ($PulM_{CTD}$: from residues 79 to 161) or the CTD of PulL ($PulL_{CTD}$: from residues 312 to 398) in *E. coli* BL21(DE3) cells (Table S2). Each protein was expressed in the periplasm fused to maltose binding protein (MBP) followed by a His_6-tag and a TEV protease cleavage site that is used to remove the MBP and the His_6-tag. Only an additional N-terminal serine remains after the cleavage.

Uniformly ^{15}N and $^{15}N/^{13}C$ labeled $PulM_{CTD}$, $PulL_{CTD}$, were produced in M9 minimal medium using 1 g/L of $^{15}NH_4Cl$ and 4 g/L ^{13}C glucose, as the sole nitrogen and carbon sources, respectively. Gene expression was induced with 1 mM IPTG (isopropyl β-D-1-thiogalactopyranoside) overnight at 18 °C in *E. coli* BL21 (DE3) cells. Unlabeled protein

samples were prepared from *E. coli* BL21 (DE3) cell cultures in LB medium and expression was induced with 1 mM IPTG for 4 hours at 30 ºC.

Protein purification

Proteins were purified from the supernatant after the sonication of bacterial cells and centrifugation at 16000 g, during 1 hour at 4 °C. The supernatant of the cell lysis was filtered on a 0.22 µm filter, then loaded onto a HisTrap HP column (*Cytiva*) equilibrated with 50 mM Tris-HCl, pH 8.0, 100 mM NaCl, 10 mM imidazole. Bound proteins were eluted with a linear imidazole gradient going from 10 to 500 mM. The eluted fractions were then incubated with the TEV-His_6 protease overnight at 10 ºC. The mixture was loaded on a HisTrap HP column to remove TEV-His_6 and MBP-His_6. Pul proteins from the unbound fractions were collected, concentrated on a 3 kDa cutoff centricon device (*Cytiva*) and applied on a Sephacryl S-100 column (*Cytiva)* equilibrated with 50 mM HEPES, pH 7.0, 50 mM NaCl. The fractions containing the purified Pul proteins were pooled and concentrated using centricon devices (*Cytiva*). Protease inhibitor cocktail, EDTA-free (*Roche*) was added to all buffers used during the purification. SDS-PAGE was used for analyzing the purity and protein contents of the fractions at each step of the purification.

$PulM_{CTD}$ protein concentration was determined spectrophotometrically using the absorbance at 280 nm and a calculated extinction coefficient of 11000 $M^{-1}.cm^{-1}$. $PulL_{CTD}$ concentration was obtained by BCA (BiCinchoninic acid Assay) method (Simpson, 2008), since this protein lacks tryptophan.

NMR experiments for assignment

NMR spectra were acquired with a range of 0.3 to 0.4 mM $^{15}N/^{13}C$ labeled proteins in 50 mM HEPES, pH 7.0 ($PulM_{CTD}$) or pH 6.5 ($PulL_{CTD}$), 50 mM NaCl at 25°C on a 600 MHz Avance III HD and a 800 MHz Avance NEO spectrometers (*Bruker Biospin*) both equipped with a cryogenically cooled triple resonance ^{1}H [^{13}C /^{15}N] probe (*Bruker Biospin*). The pulse sequences were used as implemented in the TOPSPIN 3.6.1 (*Bruker Biospin*) and IBS libraries (Favier and Brutscher, 2019). TOPSPIN 3.6.1 (*Bruker Biospin*) was used for NMR data acquisition and processing. The ^{1}H, ^{15}N, and ^{13}C backbone and side chain resonance assignments were carried out as previously described (Dazzoni *et al.*, 2021). Briefly standard experiments (Cavanagh *et al.*, 1996) were used: 2D ^{15}N-HSQC, ^{13}C-HSQC, and 3D HNCA/HN(CO)CA, HNCACB/ HN(CO)CACB pair, HNCO/HN(CA)CO pair, HCCH-TOCSY, C(CO)NH-TOCSY and H(CCO)NH-TOCSY. For the assignment of CδHδ and CεHε

of phenylalanines and tyrosines, the 2D ^{13}C-^{1}H HBCBCGCDHD and HBCBCGCDCEHE spectra were used (Yamazaki *et al.*, 1993). Side chain resonance assignments were completed by using 3D ^{13}C and ^{15}N NOESY-HSQC each with a mixing time of 120 ms and optimized for the detection of either aliphatic or aromatic carbon (Iwahara *et al.*, 2001), together with 2D ^{1}H-^{1}H NOESY (120 ms mixing time) and TOCSY in D_2O.

2,2-Dimethyl-2-silapentane-5-sulfonate (DSS) signal was taken as 0 ppm for referencing proton chemical shifts and ^{15}N and ^{13}C chemical shifts were indirectly referenced to DSS (Wishart *et al.*, 1995). CcpNmr Analysis (Vranken *et al.*, 2005) was used for NMR data analysis. Secondary structure analysis was performed by using HN, Hα, Cα, Cβ, CO, and N chemical shifts with the TALOS-N prediction server (Shen and Bax, 2013).

The assignment of $PulM_{CTD}$ $^{13}C/^{15}N/^{1}H$ resonances was deposited to the BMRB under accession number 34719. The assignment of $PulL_{CTD}$ $^{13}C/^{15}N/^{1}H$ resonances was obtained in our previous work (Dazzoni *et al.*, 2021).

$PulM_{CTD}$ and $PulL_{CTD}$ NMR structure calculation

The structure of $PulM_{CTD}$ and $PulL_{CTD}$ in their dimeric and monomeric form respectively, were determined by performing several cycles of calculation with ARIA 2.3 software (Rieping *et al.*, 2007) coupled to CNS 1.2 software and ARIAweb (Allain *et al.*, 2020; Brünger *et al.*, 1998), making use of the standard torsion angle simulating annealing protocol. Each cycle consisted of automatic 3D ^{15}N-NOESY-HSQC and 3D ^{13}C-NOESY-HSQC spectra assignment and structure calculations with 9 or 8 iterations with default parameters. In the last iteration 200 or 50 structures were calculated and further refined in an explicit water box (Linge *et al.*, 2003). Some corrections to the NOE assignment were made manually. The 15 lowest energy structures exhibiting no NOE restraint violations > 0.5 Å and no dihedral angle violations > 5 ° were selected as the final ensemble.

For $PulL_{CTD}$, an ensemble of 15 monomer structures was calculated. During the calculation process, most of the NOESs restraints from 3D ^{15}N-NOESY-HSQC and 3D ^{13}C-NOESY-HSQC spectra were automatically assigned by ARIA based on the chemical shifts previously obtained (Dazzoni *et al.*, 2021).

$PulM_{CTD}$ dimer structure was calculated in two steps, first by using intramolecular distance restraints derived from 3D ^{13}C- and ^{15}N-NOESY-HSQC spectra and the $PulM_{CTD}$ $^{13}C/^{15}N/^{1}H$ resonances assignment to determine the monomer structure. A 3D $^{13}C/^{15}N$ filtered NOESY-HSQC experiment (120 ms mixing time) was performed on a 1:1 double-labeled:unlabeled $PulM_{CTD}$ mixture to obtain intermolecular distance restraints to calculate the

dimer structure with ARIA. Chemical shift tolerances were set to 0.045 ppm for protons and 0.4 ppm for the bound heteroatoms. Phi and psi dihedral angles were predicted with TALOS-N (Shen and Bax, 2013), and predictions classified as "strong" or "good" were incorporated as dihedral angle restraints. The structure ensemble was visualized and inspected with PyMOL (The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC), their quality was evaluated with PROCHECK-NMR (Laskowski *et al.*, 1996) and the PSVS server (Bhattacharya *et al.*, 2007). The atomic coordinates of $PulM_{CTD}$ dimer and restraints used in the calculation were deposited in the Protein Data Bank (PDB ID: 7ZE0).

Protein-protein interaction analysis by NMR

The protein-protein interactions were monitored by comparison of the ^{1}H-^{15}N HSQC spectra of one labeled protein alone and in the presence of its unlabeled partner at 25 °C. ^{1}H-^{15}N HSQC experiments were acquired on 30 µM of either ^{15}N-$PulM_{CTD}$ or ^{15}N-$PulL_{CTD}$ alone or in the presence of 60 µM, 120 µM and 250 µM of unlabeled $PulL_{CTD}$ or $PulM_{CTD}$, respectively. To avoid dilution, we lyophilized the unlabeled protein in the same buffer and added it to the labeled protein sample in solution. Chemical shift perturbations (CSP) of backbone amide cross-peaks were quantified by using the equation $CSP = [\Delta\delta H^2 + (\Delta\delta N*0.159)^2]^{1/2}$, where $\Delta\delta H$ and $\Delta\delta N$ are the observed ^{1}H and ^{15}N chemical shift changes between the two experimental conditions. CSP higher than 1.5, 2 or 3 standard deviations (σ) from the mean were considered for the analysis.

Native mass spectrometry

Prior to performing native mass analysis, the quality of each protein was assessed by intact mass measurement under denaturing conditions. For native mass spectrometry, protein samples were buffer exchanged against 250 mM ammonium acetate (pH 7.0) using Zeba spin desalting columns with a 7 kDa cutoff (*Thermo Fisher Scientific, Waltham, MA, USA*). The $PulL_{CTD}$-$PulM_{CTD}$ complex was formed before or after buffer exchange by mixing $PulM_{CTD}$ with an equimolar or a 2-fold molar excess $PulL_{CTD}$.

Samples (at a final concentration of 5 to 10 µM) were analyzed on a SynaptG2-Si HDMS mass spectrometer (*Waters*) equipped with a nano-electrospray source. The instrument was calibrated in sensitivity mode using a 2 mg/ml cesium iodide solution prepared in 50% isopropanol (v/v), 0.1% formic acid (v/v) in the 50 to 5000 *m/z* range, and the quadrupole profile was adjusted to ensure the best transmission in the selected mass range. To preserve the integrity of non-covalent complexes in the gas phase, the instrument settings were carefully adjusted to

the following values: capillary voltage, 1.5-2.0 kV; sampling cone, 150 V; source offset, 150 V; trap gas flow: between 5 and 7 mL/min, trap collision energy, 4 V; cone gas, 20 L/h; source temperature, 30 °C. Spectra were acquired in positive mode for 5 to 10 minutes to obtain a good signal-to-noise ratio and processed with MassLynx 4.1 software (*Waters*) with minimal smoothing. Validation of the settings of the above-mentioned instrument was performed using equine skeletal muscle myoglobin (*Sigma Aldrich*) prepared in 250 mM ammonium acetate buffer, pH 7.0 and using the same experimental procedure as for $PulL_{CTD}$ and $PulM_{CTD}$.

Crystallization and diffraction data collection

For $PulL_{CTD}$-$PulM_{CTD}$ complex analysis, an equimolar mixture of two proteins was prepared at a final concentration of 1 mM and co-eluted on a Sephacryl S-100 column (*Cytiva)* equilibrated with 50 mM HEPES, pH 7.0, 50 mM NaCl. Initial screening of crystallization conditions was carried out using the vapor diffusion method with a MosquitoTM nanoliter-dispensing system (*TTP Labtech, Melbourn, United Kingdom*) following the established protocols (Weber *et al.*, 2019). Briefly, sitting drops were set up using 400 nl of a 1:1 mixture of each sample protein and crystallization solutions (672 different commercially available conditions) equilibrated against a 150-µl reservoir in multiwell plates (*Greiner Bio-one, GmbH, Frichenhausen, Germany*). The crystallization plates were stored at 18°C in an automated RockImager1000® (*Formulatrix, Bedford, MA, United States*) imager to monitor crystal growth. The best crystals were obtained in crystallization conditions containing 30% (w/v) PEG 4K, 0.1 M HEPES pH 7.5, 0.2 M $CaCl_2$ for $PulM_{CTD}$; 0.5 M $LiSO_4$ and 15% (w/v) PEG 8K for $PulL_{CTD}$; and 20% (w/v) PEG 3350, 0.2M KCl for the $PulL_{CTD}$-$PulM_{CTD}$ complex. Crystals were then flash cooled in liquid nitrogen using the crystallization condition supplemented with 30% (V/V) of glycerol as cryoprotectant.

Diffraction data were collected at cryogenic temperatures (100 K) on the PROXIMA-1 and PROXIMA-2A beamlines at the SOLEIL synchrotron (*St Aubin, France*) and processed with autoPROC (Vonrhein *et al.*, 2011).

X-ray structure determination and model refinement

The crystal structures of the $PulL_{CTD}$, $PulM_{CTD}$ and the $PulL_{CTD}$-$PulM_{CTD}$ complex were solved by the molecular replacement method with Phaser (McCoy *et al.*, 2007), using NMR models as search probes. Final models were obtained through interactive cycles of manual model building with Coot (Emsley and Cowtan, 2004) and reciprocal space refinement with Buster (Bricogne *et al.*, 2011) and REFMAC (Murshudov *et al.*, 2011). X-ray diffraction data

collection and model refinement statistics are summarized in Table 2. All structure figures were generated with Chimera (version 1.13rc) (Pettersen *et al.*, 2004) or with PyMOL (version 2.5.2 The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC).

Protein secretion assays

Strain PAP7460 Δ*(lac-argF)U169 araD139 relA1 rpsL150* Δ*malE444 malG501 [F' (lacI^{Q}* Δ*lacZM15 pro+ Tn10)] (Tc^{R})*) was used as a *pul* gene expression host (Possot *et al.*, 2000). PAP7460 bacteria harboring plasmids pCHAP8251 or pCHAP8496 (Table S2) were transformed with compatible plasmids encoding PulL and PulM variants (Table S2). Bacteria were cultured overnight in LB containing Ap and Cm, 0.4% D-maltose (w/v) and 0.1 vol of M63 salts (Miller, 1972) at 30°C. The next day, 6 mL of the same fresh medium was inoculated with 300 µL of precultures and grown for 5 hours to OD_{600nm} > 1.8. Cultures were normalized to OD_{600nm} of 1 in a total volume of 1 mL and fractionated as follows: 0.1 mL of cultures were centrifuged for 10 min at 16000 x g in an Eppendorf centrifuge, the supernatant was aspirated off and the bacterial pellets were resuspended in 0.1 mL of SDS sample buffer to give the cell fraction. The remaining 0.9 mL was centrifuged for 5 min at 16000 x g. The supernatant was transferred to a fresh tube and centrifuged for another 10 min at 16000 x g. A fraction (0.1 mL) of the supernatant was mixed with 0.1 mL of 2x SDS sample buffer to give the supernatant fraction. The bacterial pellets of the 0.9 mL of normalized cultures were resuspended in 90 µL of SDS sample buffer for analysis of PulL and PulM levels (Figure S4). The cell- and supernatant fractions corresponding to the 0.05 OD_{600nm} of bacteria were analyzed on 10% Tris-glycine SDS-PAGE and transferred to the nitrocellulose membranes. The membranes were probed by Western blot with anti-PulA antibodies and revealed by fluorescence using ECL-2 and Typhoon SLA-9000. Concentrated cell pellets (0.1 OD_{600nm}) were analyzed on 10% Tris-tricine SDS-PAGE and Western blot using anti-PulL and anti-PulM antibodies. Membranes were incubated with secondary anti-rabbit antibodies coupled to HRP (*Cytiva*) and revealed with ECL-2 (*Thermo*). The fluorescence signals were recorded using Typhoon SLA-9000. The signal intensity of PulA bands in the cell- and supernatant fractions was quantified using ImageJ to calculate the fraction of PulA in the supernatant. Data were analyzed with Prism GraphPad software, using Ordinary one-way ANOVA test with multiple comparisons.

Bacterial two-hybrid assays

Bacterial two-hybrid assays (Karimova *et al.*, 1998) were performed in strain DHT1 (Dautin *et al.*, 2000). Plasmids encoding compatible pUT18C and pKT25 vectors and their

derivatives were co-transformed in calcium-competent DHT1 bacteria and selected on LB plates containing Ap and Km. After 47-72 hours of growth at 30 °C, single colonies were picked at random and inoculated in LB containing the same antibiotics. Bacteria were grown overnight at 30°C and precultures were used to inoculate the same medium supplemented with 1 mM IPTG. After 4 hours, cultures were placed on ice and beta-galactosidase assays were performed as described previously (Miller, 1972). Beta-galactosidase activity from at least 8 independent cultures (indicated by the number of dots on the plot) was plotted and the data was analysed using GraphPad Prism 9 software. For the qualitative plate tests, 10 µl of bacterial cultures was spotted on LB plates containing Ap, Km, X-gal (0.2 mg. mL^{-1}) and 100 µM IPTG. The plates were cultured for 24-36 hours at 30 °C and images were recorded with a digital camera.

Cysteine crosslinking assays

Plasmids encoding different PulL and PulM variants or the empty vectors were co-transformed into strain PAP7460 and selected on LB Ap Cm plates at 30 °C. Single colonies were inoculated into 5 mL of LB containing Ap and Cm and grown overnight at 30 °C with shaking at 200 rpm. The next day, 0.2 mL of the precultures was diluted into 5 mL of fresh LB Ap Cm medium and bacteria were incubated for 5 hours at 30 °C with vigorous shaking. Cultures were normalized to OD_{600nm} of 1 and 1 mL of bacteria was centrifuged for 3 min at 16000 x g in a table-top Eppendorf centrifuge. Bacterial pellets were washed once with phosphate buffer saline (PBS) and resuspended in 1 mL of buffer containing 50 mM MOPS pH 7.0, 5 mM $MgCl_2$, 10% (v/v) glycerol. Bacterial suspensions were prewarmed at 23 °C, $CuCl_2$ was added to 300 µM (final concentration) and incubated for 23 min at 23 °C in a Thermomixer shaken at 650 rpm. To stop the reaction, EDTA pH 8.0 was added to a final concentration of 22.5 mM. Reaction mixtures were pelleted by centrifugation and bacteria were resuspended in the SDS sample buffer at 10 OD_{600nm} mL^{-1}. Total extracts were analyzed by sodium dodecyl sulfate gel electrophoresis (SDS-PAGE) on 10% Tris-Tricine gels. Proteins were transferred to nitrocellulose membranes (*Cytiva*) using the fast blotting system PowerBlot (Invitrogen) and a 1-Step transfer buffer (*Thermo*). The membranes were blocked in 5% skim milk (w/v) in Tris buffer saline solution containing 0.05% Tween-20 (TBST) and probed with anti-PulL or anti-PulM antibodies. The signals were revealed with ECL chemiluminescence kit (*Thermo*) and recorded on *Amersham* 680 imager.

Quantification of cellular PulL and PulM levels

To estimate the ratio between the cellular levels of PulL and PulM, we used semi-quantitative Western blot analysis. Different amounts of total bacterial extracts in parallel with the purified $PulL_{CTD}$ or $PulM_{CTD}$ were analyzed on 10% Tris-Tricine SDS-PAGE. Total bacterial extracts were prepared from cultures of strains PAP7460 harboring plasmid pCHAP8185 containing the *pul* operons. Bacteria were grown in LB supplemented with Ap, 0.4% D-maltose (w/v) and 0.1 vol of M63 salts. Colony forming units from these cultures were counted by plating bacterial serial dilutions in triplicate. The proteins were transferred on nitrocellulose membranes, probed with antibodies directed against $PulL_{CTD}$ and $PulM_{CTD}$ domains followed by secondary goat anti-rabbit antibodies coupled with HRP. The blots were developed using the ECL-2 kit and quantified on Typhoon FLA-9000 imager. The bands were quantified with ImageJ. The standard curves of signal intensities were plotted as a function of known $PulL_{CTD}$ or $PulM_{CTD}$ molar concentrations and linear regression analysis was used to determine the amounts of PulL and PulM in total bacterial extracts and their molar ratio, using the GraphPad Prism software.

Full-length PulL-PulM heterocomplex modeling

To model the structure of the membrane embedded and periplasmic regions of the PulL/PulM complex, we used the structure of the PilN/PilO complex from the piliated state of the type IV pilus machine from *Myxococcus xanthus* (PDB ID: 3JC8) (Chang *et al.*, 2016). First, the structure of $PulL_{CTD}$-$PulM_{CTD}$ heterodimer was superimposed onto the C-terminal domains (ferredoxin like-domains) of PilN and PilO. In both proteins, DeepCoil (Ludwiczak *et al.*, 2019) predicts two regions potentially involved in coiled-coil formation. Thus, the N-terminal helical region of PilN/PilO was used as template for comparative modeling of the homologous PulL/PulM region as a coiled-coil. The sequences of PulL and PulM encompassing the transmembrane segments (TMS) and the long helices between the TMS and the C-terminal domains, as predicted by PSIPRED (Jones, 1999), were aligned on the corresponding regions of PilN and PilO, respectively. In addition to sequence similarity, we consider the following to guide the alignment: i) the positions of the hydrophobic TMS, ii) the length of the predicted helices and iii) the expected spatial proximity of PulL residues 250/251 and PulM residues 17/18 from the cysteine cross-linking experiments. Using the resulting alignments (Figure S7C) and the $PulL_{CTD}$-$PulM_{CTD}$ heterodimer structure, a full atom model of $PulL^{239-398}$-$PulM^{1-161}$ was built by using Modeller (Sali and Blundell, 1993). SOCKET2 (Kumar and Woolfson, 2021) identified a coiled-coil interface involving helical segments from PulL (L254-R300) and PulM

(G21-I67). Regions with knobs-into-holes packing between PulL and PulM correspond to the ones predicted as coiled-coils (Figure S6D).

Quantification and statistical analysis

For the PulA secretion assays, the fluorescence signals revealed by Western blot and ECL2 were quantified using Image J from 3 or more independent experiments (indicated by the number of dots on the graphs representing individual values). For each sample, the fraction of secreted PulA was calculated as the signal intensity in the supernatant fraction divided by the sum of signal intensities of PulA in the cell and supernatant fractions. Data were analyzed with Prism GraphPad software, using Ordinary one-way ANOVA test with multiple comparisons.

For the BACTH experiments, beta-galactosidase activity measured from at least 8 independent cultures as described in Methods (Miller, 1972). The mean values were plotted and the data was analyzed using GraphPad prism 9 software using Ordinary one-way ANOVA test with multiple comparisons.

Quantification of cellular protein levels of PulL and PulM was performed by comparing Western blot ECL signals from defined number of bacteria with the signals from define quantities of purified proteins using with ImageJ. The standard curves of signal intensities were plotted as a function of known $PulL_{CTD}$ or $PulM_{CTD}$ molar concentrations and linear regression analysis was used to determine the amounts of PulL and PulM in total bacterial extracts and their molar ratio, using the GraphPad Prism software.

Table 1: NMR structure statistics and restraints of $PulL_{CTD}$ monomer and $PulM_{CTD}$ dimer

Number of restraints (per monomer)	**$PulL_{CTD}$**	**$PulM_{CTD}$**
NOE Distance restraints		
Intra-residue (|i-j| = 0)	392	501
Sequential (|i-j| = 1)	262	324
Medium-range (2 ≤ |i-j| < 5)	116	184
Long-range (|i-j| ≥ 5)	170	455
Inter-molecular	-	25
Ambiguous	23	159
Total	963	1648
Dihedral angle restraints (φ/ψ)	144 (72/72)	138 (69/69)
Hydrogen bonds restraints	28	12
Restraints statistics[a]		
Average no. of violations per structure		
NOE restraints >0.5 Å	0.7 ± 0.8	0
H-bond restraints >0.5 Å	0	0
Dihedral restraints >5 °	1.1 ± 1.4	0
RMS of distance violations		
NOE restraints	0.053 ± 0.013 Å	0.046 ± 0.001 Å
H-bond restraints	0.034 ± 0.004 Å	0.021 ± 0.001 Å
RMS of dihedral violations	1.009 ± 0.408 °	0.869 ± 0.083 °
RMS from idealized covalent geometry		
bonds	0.005 ± 0.001 Å	0.007 ± 0.001 Å
angles	0.582 ± 0.035 °	0.793 ± 0.009 °
impropers	1.639 ± 0.150 °	1.948 ± 0.039 °
Structural quality[a]		
Ramachandran statistics[b]		
Most favoured regions	89.4 ± 2.8 %	93.5 ± 1.3 %
Allowed regions	10.1 ± 3.2 %	5.9 ± 1.2 %
Generously allowed regions	0.4 ± 0.6 %	0 ± 0 %
Disallowed regions	0.1 ± 0.4 %	0.6 ± 0.7 %
Global quality scores (Raw / Z-score)[c]		
Verify3D	0.20 / -4.17	0.15 / -4.98
ProsaII	0.52 / -0.54	0.85 / 0.83
ProCheck (all)	-0.44 / -2.60	-0.34 / -2.01
MolProbity clashscore	38.6 / -5.10	48.8 / -6.85
Coordinates precision[d]		
All residues	(311-398)	(79-161)
Backbone atoms	1.44 ± 0.40 Å	0.69 ± 0.14 Å
Heavy atoms (311-398)	1.74 ± 0.38 Å	1.10 ± 0.16 Å
Ordered residues[e]	(317-344, 346-385,3 87-395)	(84-101,106-160)
Backbone atoms	0.60 ± 0.21 Å	0.35 ± 0.09 Å
Heavy atoms	1.13 ± 0.30 Å	0.69 ± 0.07 Å

[a] Average values and standard deviations over the final conformer ensemble
[b] Percentage of residues in the Ramachandran plot regions determined by PROCHECK (Laskowski *et al.*, 1996)
[c] Calculated using PSVS ver. 1.5 (Bhattacharya *et al.*, 2007)
[d] Average root mean square deviation (RMSD) over the ensemble atomic coordinates with respect to the average structure.
[e] Ordered residues [S(phi) + S(psi) > 1.8].

Table 2: Crystallography data collection and refinement statistics.

Parameters	PulL$_{CTD}$	PulL$_{CTD}$-PulM$_{CTD}$	PulM$_{CTD}$
Beamline	Proxima 2A	Proxima 1	Proxima 1
Resolution range (Å)	40.15 - 1.895 (1.963 - 1.895)	58.95 - 2.771 (2.87 - 2.771)	67.765 - 1.523 (1.549 - 1.523)
Space group	I 21 3	P 63 2 2	C 2 2 21
Unit cell (Å, °)	80.3 80.3 80.3 90 90 90	117.899 117.899 110.002 90 90 120	80.85 135.529 109.408 90 90 90
Total reflections	277247 (26424)	464974 (47811)	1238306 (62080)
Unique reflections	7012 (687)	11980 (1165)	91712 (4530)
Multiplicity	39.5 (38.5)	38.8 (41.0)	13.5 (13.7)
Completeness (%)	99.97 (100.00)	99.92 (100.00)	100.0 (100.0)
Mean I/sigma(I)	33.73 (3.03)	29.12 (2.73)	17.6 (2.2)
Wilson B-factor	41.28	95.24	25.02
R-merge	0.06935 (1.739)	0.09584 (1.471)	0.070 (0.951)
R-meas	0.07029 (1.762)	0.09716 (1.49)	0.073 (0.987)
R-pim	0.01132 (0.2832)	0.01573 (0.2311)	0.020 (0.264)
CC1/2	1 (0.86)	1 (0.846)	0.999 (0.900)
CC*	1 (0.962)	1 (0.958)	1 (0.975)
Reflections used in refinement	7012 (687)	11980 (1165)	91712 (4530)
Reflections used for R-free	339 (32)	581 (50)	4683 (494)
R-work	0.2408 (0.3641)	0.2569 (0.3558)	0.2167 (0.2828)
R-free	0.2765 (0.2755)	0.2980 (0.4413)	0.2305 (0.2933)
CC(work)	0.946 (0.725)	0.907 (0.681)	0.954 (0.874)
CC(free)	0.915 (0.713)	0.952 (0.552)	0.954 (0.820)
Number of non-hydrogen atoms	606	2365	4680
Macromolecules	579	2326	4124
Ligands	10	0	0
Solvent	17	39	556
Protein residues	75	308	549
RMS (bonds) (Å)	0.010	0.012	0.016
RMS (angles) (°)	1.41	1.62	1.53
Ramachandran plot			
favored (%)	94.37	90.00	98.88
allowed (%)	2.82	8.00	1.12
outliers (%)	2.82	2.00	0.00
Rotamer outliers (%)	10.00	10.29	1.37
Clashscore	11.83	9.52	2.04
Average B-factor	64.98	122.89	32.90
Macromolecules	65.09	123.41	32.01
Ligands	79.76		
Solvent	52.56	92.02	39.48

Statistics for the highest-resolution shell are shown in parentheses.

Acknowledgements:

This work was supported by the French Agence Nationale de la Recherche (ANR Synergy-T2SS ANR-19-CE11-0020-01) and the Fondation pour la Recherche Médicale (Equipe FRM 2017M.DEQ20170839114). YL was funded by the Pasteur Paris University (PPU) international PhD program. We thank Ingrid Guilvout and Maylis Lejeune for their constant help. We acknowledge Iñaki Guijarro, Rémy Le Meur, Bertrand Raynal, Sébastien Brûlé and Christophe Thomas of C2RT for their help and assistance. The 800-MHz NMR spectrometer and the optima AUC of the Institut Pasteur were partially funded by the Région Ile de France (SESAME 2014 NMRCHR grant no 4014526) and DIM One Health, respectively. The authors are grateful to the staff of the Institut Pasteur Crystallography platform for robot-driven crystallization screening. We acknowledge the Synchrotron SOLEIL (St Aubin, France) staff for assistance and advice during data collection on PROXIMA-1 and PROXIMA-2A beamlines. This work used the computational and storage service (TARS cluster) provided by the IT Department at Institut Pasteur, Paris.

Author contributions:

RD, ALC, SB, AH, OF, BB and NIP conceived and designed the experiments. RD, YL, ALC, SB, AM, AH, FC, OF, BB performed the experiments. RD, YL, ALC, SB, AM, AH, FC, OF, BB and NIP analyzed the data. All authors wrote the manuscript. RD, MN, OF, BB and NIP edited the manuscript.

Conflict of interest:

The authors declare that they have no conflict of interest.

References

Abendroth, J., Kreger, A., and Hol, W. (2009). The dimer formed by the periplasmic domain of EpsL from the Type 2 Secretion System of Vibrio parahaemolyticus. Journal of structural biology *168*, 313-322. 10.1016/j.jsb.2009.07.022.

Abendroth, J., Murphy, P., Sandkvist, M., Bagdasarian, M., and Hol, W. (2005). The X-ray structure of the type II secretion system complex formed by the N-terminal domain of EpsE and the cytoplasmic domain of EpsL of Vibrio cholerae. Journal of molecular biology *348*, 845-855. 10.1016/j.jmb.2005.02.061.

Abramoff, M.D., Magalhaes, P.J., and Ram, S.J. (2004). Image Processing with ImageJ. Biophotonics International *11*, 36-42.

Allain, F., Mareuil, F., Ménager, H., Nilges, M., and Bardiaux, B. (2020). ARIAweb: a server for automated NMR structure calculation. Nucleic acids research *48*, W41-W47. 10.1093/nar/gkaa362.

Alva, V., and Lupas, A. (2018). From ancestral peptides to designed proteins. Current opinion in structural biology *48*, 103-109. 10.1016/j.sbi.2017.11.006.

Berry, J., and Pelicic, V. (2015). Exceptionally widespread nanomachines composed of type IV pilins: the prokaryotic Swiss Army knives. FEMS microbiology reviews *39*, 134-154. 10.1093/femsre/fuu001.

Bhattacharya, A., Tejero, R., and Montelione, G. (2007). Evaluating protein structures determined by structural genomics consortia. Proteins *66*, 778-795. 10.1002/prot.21165.

Bricogne, G., Blanc, E., Brandl, M., Flensburg, C., Keller, P., Paciorek, W., Roversi, P., Sharff, A., Smart, O.S., Vonrhein, C., and Womack, T.O. (2011). BUSTER version 2.11.1. Cambridge, UK: Global Phasing Ltd.

Brünger, A.T., Adams, P.D., Clore, G.M., DeLano, W.L., Gros, P., Grosse-Kunstleve, R.W., Jiang, J.S., Kuszewski, J., Nilges, M., Pannu, N.S*., et al.* (1998). Crystallography & NMR system: A new software suite for macromolecular structure determination. Acta crystallographica. Section D, Biological crystallography *54*, 905-921. 10.1107/s0907444998003254.

Buddelmeijer, N., Francetic, O., and Pugsley, A. (2006). Green fluorescent chimeras indicate nonpolar localization of pullulanase secreton components PulL and PulM. Journal of bacteriology *188*, 2928-2935. 10.1128/JB.188.8.2928-2935.2006.

Cavanagh, J., Fairbrother, W.J., Palmer III, A.G., and Skelton, N.J. (1996). Protein NMR Spectroscopy: Principles and Practice (Academic press).

Chang, Y., Rettberg, L., Treuner-Lange, A., Iwasa, J., Søgaard-Andersen, L., and Jensen, G. (2016). Architecture of the type IVa pilus machine. Science *351*, aad2001. 10.1126/science.aad2001.

Chernyatina, A., and Low, H. (2019). Core architecture of a bacterial type II secretion system. Nature communications *10*, 5437. 10.1038/s41467-019-13301-3.

Cianciotto, N.P., and White, R.C. (2017). Expanding Role of Type II Secretion in Bacterial Pathogenesis and Beyond. Infection and immunity *85*, e00014-00017. 10.1128/iai.00014-17.

d'Enfert, C., Ryter, A., and Pugsley, A.P. (1987). Cloning and expression in Escherichia coli of the Klebsiella pneumoniae genes for production, surface localization and secretion of the lipoprotein pullulanase. The EMBO journal *6*, 3531-3538. 10.1002/j.1460-2075.1987.tb02679.x.

Dautin, N., Karimova, G., Ullmann, A., and Ladant, D. (2000). Sensitive genetic screen for protease activity based on a cyclic AMP signaling cascade in Escherichia coli. Journal of bacteriology *182*, 7060-7066. 10.1128/jb.182.24.7060-7066.2000.

Dazzoni, R., López-Castilla, A., Cordier, F., Bardiaux, B., Nilges, M., Francetic, O., and Izadi-Pruneyre, N. (2021). 1 H, 15 N and 13 C resonance assignments of the C-terminal domain of PulL, a component of the Klebsiella oxytoca type II secretion system. Biomolecular NMR assignments *15*, 455-459. 10.1007/s12104-021-10045-4.

Denise, R., Abby, S., and Rocha, E. (2019). Diversification of the type IV filament superfamily into machines for adhesion, protein secretion, DNA uptake, and motility. PLoS biology *17*, e3000390. 10.1371/journal.pbio.3000390.

Eck, R., and Dayhoff, M. (1966). Evolution of the structure of ferredoxin based on living relics of primitive amino Acid sequences. Science *152*, 363-366. 10.1126/science.152.3720.363.

929 Emsley, P., and Cowtan, K. (2004). Coot: model-building tools for molecular graphics. Acta
930 crystallographica. Section D, Biological crystallography *60*, 2126-2132.
931 10.1107/s0907444904019158.

932 Favier, A., and Brutscher, B. (2019). NMRlib: user-friendly pulse sequence tools for Bruker
933 NMR spectrometers. Journal of biomolecular NMR *73*, 199-211. 10.1007/s10858-019-00249-
934 1.

935 Fulara, A., Vandenberghe, I., Read, R.J., Devreese, B., and Savvides, S.N. (2018). Structure
936 and oligomerization of the periplasmic domain of GspL from the type II secretion system of
937 Pseudomonas aeruginosa. Sci Rep *8*, 16760. 10.1038/s41598-018-34956-w.

938 Ghosal, D., Kim, K.W., Zheng, H., Kaplan, M., Truchan, H.K., Lopez, A.E., McIntire, I.E.,
939 Vogel, J.P., Cianciotto, N.P., and Jensen, G.J. (2019). In vivo structure of the Legionella type
940 II secretion system by electron cryotomography. Nat Microbiol *4*, 2101-2108. 10.1038/s41564-
941 019-0603-6.

942 Gray, M.D., Bagdasarian, M., Hol, W.G.J., and Sandkvist, M. (2011). In vivo cross-linking of
943 EpsG to EpsL suggests a role for EpsL as an ATPase-pseudopilin coupling protein in the Type
944 II secretion system of Vibrio cholerae. Mol Microbiol. *79*, 786–798. 10.1111/j.1365-
945 2958.2010.07487.x.

946 Hobbs, M., and Mattick, J. (1993). Common components in the assembly of type 4 fimbriae,
947 DNA transfer systems, filamentous phage and protein-secretion apparatus: a general system for
948 the formation of surface-associated protein complexes. Molecular microbiology *10*, 233-243.
949 10.1111/j.1365-2958.1993.tb01949.x.

950 Iwahara, J., Wojciak, J.M., and Clubb, R.T. (2001). Improved NMR spectra of a protein-DNA
951 complex through rational mutagenesis and the application of a sensitivity optimized isotope-
952 filtered NOESY experiment. Journal of biomolecular NMR *19*, 231-241.
953 10.1023/a:1011296112710.

954 Janin, J., Rodier, F., Chakrabarti, P., and Bahadur, R. (2007). Macromolecular recognition in
955 the Protein Data Bank. Acta crystallographica. Section D, Biological crystallography *63*, 1-8.
956 10.1107/s090744490603575x.

957 Jones, D. (1999). Protein secondary structure prediction based on position-specific scoring
958 matrices. Journal of molecular biology *292*, 195-202. 10.1006/jmbi.1999.3091.

959 Karimova, G., Pidoux, J., Ullmann, A., and Ladant, D. (1998). A bacterial two-hybrid system
960 based on a reconstituted signal transduction pathway. Proceedings of the National Academy of
961 Sciences of the United States of America *95*, 5752-5756. 10.1073/pnas.95.10.5752.

962 Krissinel, E., and Henrick, K. (2007). Inference of macromolecular assemblies from crystalline
963 state. J. Mol. Biol. *372*, 774-797. 10.1016/j.jmb.2007.05.022.

964 Kumar, P., and Woolfson, D. (2021). Socket2: A Program for Locating, Visualising, and
965 Analysing Coiled-coil Interfaces in Protein Structures. Bioinformatics *37*, 4575-4577.
966 10.1093/bioinformatics/btab631.

967 Lallemand, M., Login, F.H., Guschinskaya, N., Pineau, C., Effantin, G., Robert, X., and
968 Shevchik, V.E. (2013). Dynamic interplay between the periplasmic and transmembrane
969 domains of GspL and GspM in the type II secretion system. PLoS One *8*, e79562.
970 10.1371/journal.pone.0079562.

971 Laskowski, R., Rullmannn, J., MacArthur, M., Kaptein, R., and Thornton, J. (1996). AQUA
972 and PROCHECK-NMR: programs for checking the quality of protein structures solved by
973 NMR. Journal of biomolecular NMR *8*, 477-486. 10.1007/bf00228148.

974 Leighton, T., Mok, M., Junop, M., Howell, P., and Burrows, L. (2018). Conserved, unstructured
975 regions in Pseudomonas aeruginosa PilO are important for type IVa pilus function. Scientific
976 reports *8*, 2600. 10.1038/s41598-018-20925-w.

977 Linge, J., Williams, M., Spronk, C., Bonvin, A., and Nilges, M. (2003). Refinement of protein
978 structures in explicit solvent. Proteins *50*, 496-506. 10.1002/prot.10299.

979 Lopez-Castilla, A., Thomassin, J.L., Bardiaux, B., Zheng, W., Nivaskumar, M., Yu, X., Nilges,
980 M., Egelman, E.H., Izadi-Pruneyre, N., and Francetic, O. (2017). Structure of the calcium-
981 dependent type 2 secretion pseudopilus. Nat Microbiol *2*, 1686-1695. 10.1038/s41564-017-
982 0041-2.

Ludwiczak, J., Winski, A., Szczepaniak, K., Alva, V., and Dunin-Horkawicz, S. (2019). DeepCoil-a fast and accurate prediction of coiled-coil domains in protein sequences. Bioinformatics *35*, 2790-2795. 10.1093/bioinformatics/bty1062.

Maffei, B., Francetic, O., and Subtil, A. (2017). Tracking Proteins Secreted by Bacteria: What's in the Toolbox? Frontiers in cellular and infection microbiology *7*, 221. 10.3389/fcimb.2017.00221.

McCoy, A., Grosse-Kunstleve, R., Adams, P., Winn, M., Storoni, L., and Read, R. (2007). Phaser crystallographic software. Journal of applied crystallography *40*, 658-674. 10.1107/s0021889807021206.

Michel-Souzy, S., Douzi, B., Cadoret, F., Raynaud, C., Quinton, L., Ball, G., and Voulhoux, R. (2018). Direct interactions between the secreted effector and the T2SS components GspL and GspM reveal a new effector-sensing step during type 2 secretion. J Biol Chem *293*, 19441-19450. 10.1074/jbc.RA117.001127.

Miller, J.H. (1972). Experiments in Molecular Genetics (Cold Spring Harbor).

Murshudov, G., Skubák, P., Lebedev, A., Pannu, N., Steiner, R., Nicholls, R., Winn, M., Long, F., and Vagin, A. (2011). REFMAC5 for the refinement of macromolecular crystal structures. Acta crystallographica. Section D, Biological crystallography *67*, 355-367. 10.1107/s0907444911001314.

Naskar, S., Hohl, M., Tassinari, M., and Low, H. (2021). The structure and mechanism of the bacterial type II secretion system. Molecular microbiology *115*, 412-424. 10.1111/mmi.14664.

Nivaskumar, M., Santos-Moreno, J., Malosse, C., Nadeau, N., Chamot-Rooke, J., Tran Van Nhieu, G., and Francetic, O. (2016). Pseudopilin residue E5 is essential for recruitment by the type 2 secretion system assembly platform. Mol Microbiol *101*, 924-941. 10.1111/mmi.13432.

O'Shea, E., Rutkowski, R., and Kim, P. (1992). Mechanism of specificity in the Fos-Jun oncoprotein heterodimer. Cell *68*, 699-708. 10.1016/0092-8674(92)90145-3.

Pacheco Gutierrez, A.J. (2021). Structural studies of the bacterial Type II secretion system. PhD thesis (Queen Mary University of London).

1010 Peabody, C., Chung, Y., Yen, M., Vidal-Ingigliardi, D., Pugsley, A., and Saier, M. (2003). Type
1011 II protein secretion and its relationship to bacterial type IV pili and archaeal flagella.
1012 Microbiology (Reading, England) *149*, 3051-3072. 10.1099/mic.0.26364-0.

1013 Pettersen, E., Goddard, T., Huang, C., Couch, G., Greenblatt, D., Meng, E., and Ferrin, T.
1014 (2004). UCSF Chimera--a visualization system for exploratory research and analysis. Journal
1015 of computational chemistry *25*, 1605-1612. 10.1002/jcc.20084.

1016 Possot, O.M., Vignon, G., Bomchil, N., Ebel, F., and Pugsley, A.P. (2000). Multiple
1017 interactions between pullulanase secreton components involved in stabilization and
1018 cytoplasmic membrane association of PulE. J Bacteriol *182*, 2142-2152.
1019 10.1128/jb.182.8.2142-2152.2000.

1020 Pugsley, A.P. (1993). The complete general secretory pathway in gram-negative bacteria.
1021 Microbiological reviews *57*, 50-108. 10.1128/mr.57.1.50-108.1993.

1022 Py, B., Loiseau, L., and Barras, F. (2001). An inner membrane platform in the type II secretion
1023 machinery of Gram-negative bacteria. EMBO reports *2*, 244-248. 10.1093/embo-
1024 reports/kve042.

1025 Rieping, W., Habeck, M., Bardiaux, B., Bernard, A., Malliavin, T., and Nilges, M. (2007).
1026 ARIA2: automated NOE assignment and data integration in NMR structure calculation.
1027 Bioinformatics *23*, 381-382. 10.1093/bioinformatics/btl589.

1028 Sali, A., and Blundell, T. (1993). Comparative protein modelling by satisfaction of spatial
1029 restraints. Journal of molecular biology *234*, 779-815. 10.1006/jmbi.1993.1626.

1030 Sampaleanu, L., Bonanno, J., Ayers, M., Koo, J., Tammam, S., Burley, S., Almo, S., Burrows,
1031 L., and Howell, P. (2009). Periplasmic domains of Pseudomonas aeruginosa PilN and PilO
1032 form a stable heterodimeric complex. Journal of molecular biology *394*, 143-159.
1033 10.1016/j.jmb.2009.09.037.

1034 Sandkvist, M., Hough, L.P., Bagdasarian, M.M., and Bagdasarian, M. (1999). Direct interaction
1035 of the EpsL and EpsM proteins of the general secretion apparatus in Vibrio cholerae. J Bacteriol
1036 *181*, 3129-3135. 10.1128/JB.181.10.3129-3135.1999.

Santos-Moreno, J., East, A., Guilvout, I., Nadeau, N., Bond, P.J., Tran Van Nhieu, G., and Francetic, O. (2017). Polar N-terminal Residues Conserved in Type 2 Secretion Pseudopilins Determine Subunit Targeting and Membrane Extraction Steps during Fibre Assembly. Journal of molecular biology *429*, 1746-1765. 10.1016/j.jmb.2017.04.005.

Shen, Y., and Bax, A. (2013). Protein backbone and sidechain torsion angles predicted from NMR chemical shifts using artificial neural networks. Journal of biomolecular NMR *56*, 227-241. 10.1007/s10858-013-9741-y.

Simpson, R. (2008). Quantifying protein by bicinchoninic Acid. CSH protocols *2008*, pdb.prot4722. 10.1101/pdb.prot4722.

Vonrhein, C., Flensburg, C., Keller, P., Sharff, A., Smart, O., Paciorek, W., Womack, T., and Bricogne, G. (2011). Data processing and analysis with the autoPROC toolbox. Acta crystallographica. Section D, Biological crystallography *67*, 293-302. 10.1107/s0907444911007773.

Vranken, W.F., Boucher, W., Stevens, T.J., Fogh, R.H., Pajon, A., Llinas, M., Ulrich, E.L., Markley, J.L., Ionides, J., and Laue, E.D. (2005). The CCPN data model for NMR spectroscopy: development of a software pipeline. Proteins *59*, 687-696. 10.1002/prot.20449.

Weber, P., Pissis, C., Navaza, R., Mechaly, A., Saul, F., Alzari, P., and Haouz, A. (2019). High-Throughput Crystallization Pipeline at the Crystallography Core Facility of the Institut Pasteur. Molecules *24*, 4451. 10.3390/molecules24244451.

Wishart, D., Bigam, C., Yao, J., Abildgaard, F., Dyson, H., Oldfield, E., Markley, J., and Sykes, B. (1995). 1H, 13C and 15N chemical shift referencing in biomolecular NMR. Journal of biomolecular NMR *6*, 135-140. 10.1007/bf00211777.

Yamazaki, T., Yoshida, M., and Nagayama, K. (1993). Complete assignments of magnetic resonances of ribonuclease H from Escherichia coli by double- and triple-resonance 2D and 3D NMR spectroscopies. Biochemistry *32*, 5656-5669. 10.1021/bi00072a023.