



# **Pseudopilus assembly and substrate recognition factors in the Klebsiella type II secretion system**

Yuanyuan Li

## **► To cite this version:**

Yuanyuan Li. Pseudopilus assembly and substrate recognition factors in the Klebsiella type II secretion system. Microbiology and Parasitology. Université Paris Cité, 2022. English. <NNT : 2022UNIP5192>. <tel-04659781>

**HAL Id: tel-04659781**

**<https://theses.hal.science/tel-04659781v1>**

Submitted on 23 Jul 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

# Université Paris Cité Institut PASTEUR

École doctorale Bio Sorbonne Paris Cité ED562

*Unité de Biochimie des Interactions Macromoléculaires*

## **Pseudopilus assembly and substrate recognition factors in the *Klebsiella* type II secretion system**

Par Yuanyuan LI

Thèse de doctorat de Microbiologie

Dirigée par Olivera Francetic

Présentée et soutenue publiquement le 16 Décembre 2022

Devant un jury composé de :

**Raffaele IEVA**, CR-HDR, Université Paul Sabatier -Toulouse, Rapporteur

**Romé VOULHOUX**, DR, Université Aix Marseille, Rapporteur

**Ombeline ROSSIER**, MCU, Université Paris-Saclay, Examinatrice

**Emmanuelle BOUVERET**, DR, Université Paris Cité, Examinatrice

**Nadia IZADI-PRUNEYRE**, DR, Université Paris Cité, Examinatrice

**Olivera Francetic**, DR, Institut Pasteur, Directrice de thèse

## **Facteurs impliqués dans l'assemblage du pseudopilus et dans la reconnaissance de substrat dans le système de sécrétion de type II chez *Klebsiella***

Les bactéries à Gram-négatif utilisent plusieurs mécanismes pour sécréter des protéines spécifiques à travers leur enveloppe vers la surface cellulaire ou le milieu. Chez *Klebsiella*, notre modèle expérimental, le système de sécrétion de Type II (SST2) permet la sécrétion de la pullulanase (PulA), une enzyme de la famille des amylases. La sécrétion de PulA repose sur l'assemblage de filaments périsplasmiques appelés pseudopili qui facilitent le transport des substrats repliés du périplasma à travers la membrane externe. Les pseudopili, apparentés aux pili de type IV, sont composés de quatre sous-unités mineures (PulHIJK) et d'une sous-unité majeure, PulG. Ils sont polymérisés dans la membrane interne au niveau d'un sous-complexe de SST2 - la plateforme d'assemblage - composé de protéines PulF, PulL et PulM. Le mécanisme d'assemblage du pseudopilus et son rôle éventuel dans la reconnaissance du substrat restent mal compris.

PulA est une lipoprotéine de 117 kDa sécrétée par le SST2 sous forme repliée. Dans ce travail, nous avons cherché à identifier les composants du SST2 qui interagissent directement avec PulA pour permettre sa sécrétion. Dans ce but, nous avons effectué une analyse systématique par l'approche du double hybride bactérien (BACTH), ainsi que par pontage chimique et co-purification. Nous avons également mis en place une approche de pontage *in vivo* après irradiation aux rayons UVs grâce à l'acide aminé non naturel *p*BPA introduit dans les régions de PulA contenant des signaux de sécrétion potentiels. L'analyse par BACTH a identifié les composants de la plate-forme d'assemblage PulF et PulM, ainsi que les pseudopilines PulG, PulL et PulH, comme partenaires d'interaction putatifs de PulA.

Dans la deuxième partie du travail, nous nous sommes intéressés aux composants de la plateforme d'assemblage PulL et PulM impliqués dans l'assemblage des pseudopili et dans la sécrétion de PulA. Ce travail a été réalisé en collaboration avec l'équipe de N. Izadi-Pruneyre, qui a déterminé la structure des domaines C-terminaux (CTD) de PulL et PulM et de leur interface dans le complexe. Nous avons construit des variants contenant des substitutions des résidus d'interface PulL<sub>CTD</sub> - PulM<sub>CTD</sub> et montré qu'ils sont déficients pour la sécrétion de PulA. Nous avons également étudié les interactions PulL - PulM par l'approche du BACTH et utilisé le pontage disulfure pour caractériser les interactions entre leurs segments transmembranaires. Ensemble, ces données nous ont permis de proposer un modèle structural du complexe PulL - PulM dans la membrane. Nos travaux ont révélé un ratio entre PulM et PulL dans la cellule de 20 :1 ce qui suggère un rôle dynamique de PulM dans le transfert de PulG, entre les molécules de PulG libre et celles de PulL associées au site actif d'assemblage du pseudopilus.

Dans un second temps, nous avons étudié le rôle des hélices périsplasmiques et des domaines cytoplasmiques de PulL et PulM dans leur assemblage. Nous avons constaté que la super-hélice formée entre leurs domaines périsplasmiques joue un rôle clé dans la liaison PulL-PulM. Nous avons également constaté que le peptide N-terminal de PulM est nécessaire à la stabilité de PulL. Le même peptide interagit avec PulG, suggérant son rôle dans l'échange dynamique avec le domaine cytoplasmique de PulL. Ces résultats ouvrent la voie à des études plus détaillées sur des interactions et changements conformationnels impliqués dans les différentes étapes de l'extraction membranaire et incorporation de PulG dans le pseudopilus.

**Mots clefs :** *Klebsiella*, sécrétion de protéines, assemblage des pili, interactions protéine-protéine, pontage *in vivo*

## **Pseudopilus assembly and substrate recognition factors in the *Klebsiella* type II secretion system**

Gram-negative bacteria employ several types of complex machineries to secrete specific proteins across their cell envelope to the cell surface or to their milieu. The *Klebsiella* Type II secretion system (T2SS), our experimental model, is dedicated to secretion of pullulanase (PulA), an enzyme of the amylase family. PulA secretion relies on the assembly of periplasmic filaments called pseudopili, a process which is thought to drive the folded T2SS substrates from the periplasm across the outer membrane. Pseudopili, closely related to bacterial type IV pili, are composed of four minor tip subunits (PulHIJK) and a major core subunit PulG. They are polymerized at the inner membrane assembly platform subcomplex of the T2SS, composed of PulF, PulL and PulM inner membrane proteins. The mechanism of pseudopilus assembly and their role in substrate recognition are still poorly understood.

PulA is a 117 kDa lipoprotein secreted in a folded form by the *Klebsiella* T2SS. In this work we aimed to identify T2SS components that interact directly with PulA to promote its transport across the outer membrane. Towards this goal, we performed systematic bacterial two-hybrid (BACTH) analysis, chemical cross-linking and co-purification. We also set up and *in vivo* UV crosslinking approach by introducing the unnatural amino acid *p*BPA in the PulA regions previously identified as putative secretion signals. The BACTH analysis identified the assembly platform components PulF and PulM as putative interacting partners of PulA, as well as the pseudopilins PulG, PulL and PulH.

In the second part of the work, we focused on the assembly platform components PulL and PulM that promote pseudopilus assembly and PulA secretion. This work was done in collaboration with the team of N. Izadi-Pruneyre, who determined the structure of the C-terminal domains (CTDs) of PulL and PulM and of their complex. We tested mutations of residues mapping in the PulL<sub>CTD</sub> –and PulM<sub>CTD</sub> interface and found that they affect PulA secretion. We also investigated the interactions of full-length PulL and PulM by the BACTH approach and used cysteine crosslinking to map the interactions of PulL and PulM transmembrane segments. Together, these data allowed us to propose a structural model of the full-length PulL and PulM complex. The higher abundance of PulM compared to PulL revealed in our work suggested a dynamic role of PulM, which shuttles PulG monomers from their inner membrane pool to PulL at the pseudopilus assembly site.

In a follow-up study, we investigated the role of periplasmic helices and cytoplasmic domains of PulL and PulM in their assembly. We found that the coiled-coil formed between periplasmic domains of PulL and PulM plays a key role in their binding. We also showed a new role of the PulM N-terminal peptide in PulL stability. The same peptide is required for binding to PulG, consistent with its role in a dynamic exchange with the cytoplasmic domain of PulL during pseudopilus assembly. These results open the way for more detailed investigation of interactions and conformational changes involved in different steps of membrane extraction and incorporation of PulG into the growing pseudopilus.

**Keywords:** *Klebsiella*, protein secretion, pilus assembly, protein-protein interactions, *in vivo* crosslinking

## Acknowledgements

First of all, I need to thank my dearest supervisor Dr Olivera Francetic. She paved the way for me to walk into the science world with her great patience and excellent professionalism. Since I came from a very different background – clinical medicine, the beginning of my PhD study was quite a challenge for either me or my supervisor. She taught me every basic experiment hand by hand. I still remembered how shaking my hand was when I had to pipette a 1  $\mu$ L DNA sample on the lens of Nanodrop for the first time, and how excited I was when I successfully made one plasmid for the first time. I even posted that plasmid map on my WeChat's moments (the most popular social media application in China), and said "one small step for science, one giant leap for me". What's more, she taught me the way to think and work in science, how to generate hypotheses based on what is known and how to design experiments to confirm or disprove them step by step. Also, she showed me how to solve problems during the processes. All of these laid a solid foundation for my future career development. At same time, as a non-French speaking person, she helped me so much for my daily life and many administrative procedures. There are too many thanks to Olivera that I cannot list all of them here, but I have to say, without her, I would not have been able to complete my PhD study or make any of these achievements.

I also need to thank everyone in the lab, who helped me in many ways during my whole PhD study. During the 3-month lock down in 2020 because of the COVID-19, Dr Jenny-Lee Thomassin, who was a postdoc in our lab, prepared a reading list for my project, papers about the studies of Type II secretion system and pullulanase from the history to the present, which helped me to build systematically my background knowledge needed for my project. We had regular journal clubs on Teams every week to present one paper that I read. Olivera, Jenny and Ingrid listened patiently and carefully and gave me a lot of suggestions. Also, I learned a lot from listening to them discussing the good sides or shortcomings of those papers. Ingrid Guilvout, who was the engineer in our group, taught me many techniques and helped me to learn good practice habits of working in the lab, as well as Veronique Robbe-Saule, who supported my experiments by maintaining the well-functioning of the lab.

I need to thank my thesis committee, professors Romé Volhous and Raffaele Ieva for accepting to review my thesis manuscript, professor Ombeline Rossier for being the examiner, and Nadia Izadi-Pruneyre and Emmanuelle Bouveret, who also helped me as members of my

Thesis Advisory Committee. Thank you all for being my jury members and attending my thesis defense. I thank the PPU program and my employer—the China national biotech Group (CNBG) for offering me the chance to come to France and to Institut Pasteur, the former Dean of the PPU program professor Susanna Celli and the current Dean Dr Nathalie Pardigon, as well as the director of the Education Department Monica Sala, who helped us all with administrative issues. I thank my TAC members, Dr Nadia Izadi-Pruneyre - my internal expert, Dr Vladimir Shevchik - my external expert, and my tutors, Dr Emmanuelle Bouveret and Dr Louis Lambrechts for attending my TAC meetings and giving me helpful advice and suggestions in every TAC meeting.

My sincere thanks Dr. Daniel Ladant – the head of the Biochemistry of Macromolecular Interactions (BIM) Unit, and Gouzel Karimova, she shared a lot of ideas with me about Russia and China, and taught me how to do BACTH experiment by 96-well plates and microplates reader. I thank every member from our unit and Marie-Bénédicte Benedic, for their support and help in many ways.

Last but not least, I thank my parents and my boyfriend, who supported me in every way during these four years, especially when I had troubles or difficulties, they are the source of my spiritual strength and the driving force behind my fighting for a better life.

感谢妈妈和爸爸给予我生命，抚养我，教育我，尤其是我的妈妈，总是在我最迷茫最虚弱的时候抚慰我帮助我，用她的智慧指引我光明的方向。人生一路走来，从懵懵懂懂蹒跚学步到游刃有余独立行走人间，其中遇到无数挫折和困苦，但无论外面风雨再大，你们永远是我心灵的港湾，是我前进的动力。谁言寸草心，报得三春晖！没有你们就没有我今天的成就，我的每一份收获里都有你们的辛劳付出。博士毕业不是人生的至高时刻，是攀登另一座高山的开始。人生路漫漫，希望你们能陪伴我长长久久。

## Abbreviations

|                           |                                                           |
|---------------------------|-----------------------------------------------------------|
| AP                        | Assembly Platform                                         |
| Ap <sup>R</sup>           | Ampicillin Resistance                                     |
| BACTH                     | Bacterial Adenylate Cyclase Two-Hybrid                    |
| BAM                       | β-Barrel Assembly Machinery                               |
| BlaM                      | β-Lactamase                                               |
| CL                        | Cleared Lysate                                            |
| Cm <sup>R</sup>           | Chloramphenicol Resistance                                |
| CTD                       | C-Terminal Domain                                         |
| DTT                       | Dithiothreitol                                            |
| EM                        | Electron Microscopy                                       |
| FL                        | Flow-Through                                              |
| IM                        | Inner Membrane                                            |
| IMP                       | Inner Membrane Protein                                    |
| LacZ                      | β-Galactosidase                                           |
| LP                        | Lipoprotein                                               |
| LPS                       | Lipopolysaccharide                                        |
| LUCA                      | Last Universal Common Ancestor                            |
| MD                        | Molecular Dynamics                                        |
| OD <sub>600nm/420nm</sub> | Optical Density at A Wavelength of 600nm or 420nm         |
| OM                        | Outer Membrane                                            |
| OMP                       | Outer Membrane Protein                                    |
| <i>p</i> Bpa              | <i>p</i> -Benzoyl-L-Phenylalanine                         |
| PDB                       | Protein Data Bank                                         |
| PG                        | Peptidoglycan                                             |
| PhoA                      | Alkaline Phosphatase                                      |
| POPE                      | 1-Palmitoyl-2-Oleoyl-Sn-Glycero3-Phosphoethanolamine      |
| POTRA                     | Polypeptides Transport-Associated                         |
| PS                        | Periplasmic Space                                         |
| PulA                      | Pullulanase                                               |
| PulA <sub>NA</sub>        | Nonacylated Variant of PulA                               |
| PulL <sub>CTD</sub>       | C-Terminal Domains of PulL                                |
| PulM <sub>CTD</sub>       | C-Terminal Domains of PulM                                |
| SDS-PAGE                  | Sodium Dodecyl Sulfate–Polyacrylamide Gel Electrophoresis |
| SP                        | Signal Peptide                                            |
| SPase                     | Signal Peptidase                                          |
| SPINE                     | Strep-Protein Interaction Experiment                      |
| Sp <sup>R</sup>           | Spectinomycin Resistance                                  |
| SRP                       | Signal Recognition Particle                               |

|       |                                              |
|-------|----------------------------------------------|
| T10SS | Type X Secretion System                      |
| T11SS | Type XI Secretion System                     |
| T1SS  | Type I Secretion System                      |
| T2SS  | Type II Secretion System                     |
| T3SS  | Type III Secretion System                    |
| T4aP  | Type IVa Pili                                |
| T4bP  | Type IVb Pili                                |
| T4F   | Type IV Filament                             |
| T4P   | Type IV Pili                                 |
| T4PM  | Type IV Pilus Machine                        |
| T4SS  | Type IV Secretion System                     |
| T5SS  | Type V Secretion System                      |
| T6SS  | Type VI Secretion System                     |
| T7SS  | Type VII Secretion System                    |
| T8SS  | Type VIII Secretion System                   |
| T9SS  | Type IX Secretion System                     |
| TAT   | Twin-Arginine Translocation Pathway          |
| TBST  | Tris-Buffer Saline Containing 0.05% Tween-20 |
| TMS   | Transmembrane Segment                        |
| TPR   | Tetratricopeptide Repeated                   |

# Table of content

|                                                                                                            |           |
|------------------------------------------------------------------------------------------------------------|-----------|
| <b>Acknowledgements</b>                                                                                    | <b>4</b>  |
| <b>Abbreviations</b>                                                                                       | <b>6</b>  |
| <b>Table of content</b>                                                                                    | <b>8</b>  |
| <b>List of Figures</b>                                                                                     | <b>10</b> |
| <b>List of Tables</b>                                                                                      | <b>12</b> |
| <b>1. Introduction</b>                                                                                     | <b>13</b> |
| <b>1.1 Bacterial cell envelope</b>                                                                         | <b>14</b> |
| <b>1.2 Extra-cytoplasmic and membrane proteins</b>                                                         | <b>15</b> |
| 1.2.1 Signal peptides                                                                                      | 16        |
| 1.2.2 Inner membrane proteins and the SRP pathway                                                          | 17        |
| 1.2.3 Lipoproteins                                                                                         | 18        |
| 1.2.4 Outer membrane proteins                                                                              | 19        |
| 1.2.5 Protein translocation across the IM                                                                  | 20        |
| 1.2.5.1 The Sec pathway                                                                                    | 21        |
| 1.2.5.2 The Tat Pathway                                                                                    | 22        |
| <b>1.3 Cell surface appendages</b>                                                                         | <b>24</b> |
| 1.3.1 Flagella                                                                                             | 24        |
| 1.3.2 Pili                                                                                                 | 26        |
| <b>1.4 Gram-negative bacterial protein secretion systems</b>                                               | <b>29</b> |
| <b>1.5 Type II secretion system</b>                                                                        | <b>34</b> |
| 1.5.1 T2SS as members of the type IV filament superfamily                                                  | 34        |
| 1.5.2 The genetic organization of T2SS and nomenclature of components                                      | 36        |
| 1.5.3 Architecture and biological functions of T2SS                                                        | 37        |
| 1.5.4 The <i>Klebsiella</i> T2SS: function and organization                                                | 38        |
| 1.5.5 The outer membrane complex                                                                           | 40        |
| 1.5.6 The inner membrane assembly platform (AP) and the ATPase                                             | 43        |
| 1.5.7 The pseudopilus                                                                                      | 47        |
| 1.5.7.1 Major pseudopilin                                                                                  | 48        |
| 1.5.7.2 Minor pseudopilins                                                                                 | 50        |
| 1.5.8 PulN                                                                                                 | 51        |
| 1.5.9 PulB                                                                                                 | 51        |
| 1.5.10 PulO                                                                                                | 52        |
| <b>1.6 Substrates and secretion signals of T2SS</b>                                                        | <b>53</b> |
| 1.6.1 Substrates of T2SS                                                                                   | 53        |
| 1.6.2 Structural and functional studies of type II secretion signals                                       | 53        |
| 1.6.3 Secretion signal studies of pullulanase in <i>Klebsiella</i> T2SS                                    | 54        |
| 1.6.4 Interacting partners of substrates in the T2SS                                                       | 58        |
| <b>2. Objectives</b>                                                                                       | <b>62</b> |
| <b>2.1 Identifying PulA interacting partners in the T2SS</b>                                               | <b>63</b> |
| <b>2.2 Structural and functional characterization of PulL and PulM components of the Assembly platform</b> | <b>65</b> |
| <b>3. Results</b>                                                                                          | <b>67</b> |
| <b>3.1 PulA interactions with the T2SS components</b>                                                      | <b>68</b> |
| 3.1.1 The <i>in vivo</i> site-directed photo-cross-linking                                                 | 68        |

|             |                                                                                                                                                       |            |
|-------------|-------------------------------------------------------------------------------------------------------------------------------------------------------|------------|
| 3.1.1.1     | pBpa incorporation and in vivo site-specific photo-cross-linking of PulG                                                                              | 68         |
| 3.1.1.2     | Construction of PulA-pBpa derivatives and cross-linking assays                                                                                        | 76         |
| 3.1.1.3     | Discussion of the pBPA crosslinking results                                                                                                           | 89         |
| 3.1.2       | Chemical crosslinking and copurification of PulA-Strep and variants                                                                                   | 90         |
| 3.1.2.1     | Defining the optimal growth stage for crosslinking                                                                                                    | 91         |
| 3.1.2.2     | Strep-protein interaction experiment (SPINE)                                                                                                          | 92         |
| 3.1.2.3     | Discussion of the SPINE results                                                                                                                       | 96         |
| 3.1.3       | Bacterial adenylate cyclase two-hybrid (BACTH) study of PulA interactions                                                                             | 97         |
| 3.1.3.1     | PulA interactions with the pseudopilins                                                                                                               | 97         |
| 3.1.3.2     | PulA interactions with inner membrane Pul T2SS proteins                                                                                               | 99         |
| 3.1.3.3     | Discussion of the BACTH results                                                                                                                       | 102        |
| <b>3.2</b>  | <b>Analysis of the assembly platform components PulL and PulM</b>                                                                                     | <b>104</b> |
| 3.2.1       | Structure and dynamic association of an assembly platform subcomplex of the bacterial type II secretion system                                        | 106        |
| 3.2.2       | Periplasmic coiled coil formed by assembly platform proteins PulL and PulM is critical for function of the <i>Klebsiella</i> type II secretion system | 161        |
| 3.2.3       | Additional results about PulL and PulM                                                                                                                | 190        |
| <b>4.</b>   | <b>Discussion and Perspectives</b>                                                                                                                    | <b>192</b> |
| <b>4.1</b>  | <b>Crosslinking approaches to identify PulA interacting partners</b>                                                                                  | <b>193</b> |
| <b>4.2</b>  | <b>BACTH studies of PulA interactions with the T2SS</b>                                                                                               | <b>196</b> |
| <b>4.3</b>  | <b>The role of PulL-PulM assembly platform complex</b>                                                                                                | <b>197</b> |
| <b>5.</b>   | <b>Materials and methods</b>                                                                                                                          | <b>201</b> |
| <b>5.1</b>  | <b>Bacterial strains and growth conditions</b>                                                                                                        | <b>202</b> |
| <b>5.2</b>  | <b>Plasmid construction and site-directed mutagenesis</b>                                                                                             | <b>203</b> |
| <b>5.3</b>  | <b>DNA transformation</b>                                                                                                                             | <b>211</b> |
| 5.3.1       | Preparation and transformation to CaCl <sub>2</sub> competent cells and transformation                                                                | 211        |
| 5.3.2       | Preparation of ultra-competent cells and transformation                                                                                               | 212        |
| 5.3.3       | TSS transformation                                                                                                                                    | 212        |
| <b>5.4</b>  | <b>Protein electrophoresis and Western blot analysis</b>                                                                                              | <b>212</b> |
| <b>5.5</b>  | <b>pBpa incorporation</b>                                                                                                                             | <b>213</b> |
| <b>5.6</b>  | <b>The <i>in vivo</i> site-specific photo-cross-linking assay</b>                                                                                     | <b>214</b> |
| <b>5.7</b>  | <b>Protein secretion assays</b>                                                                                                                       | <b>215</b> |
| <b>5.8</b>  | <b>Bacterial adenylate cyclase two-hybrid assay</b>                                                                                                   | <b>215</b> |
| <b>5.9</b>  | <b>Beta-galactosidase assay</b>                                                                                                                       | <b>216</b> |
| <b>5.10</b> | <b>Cysteine crosslinking</b>                                                                                                                          | <b>216</b> |
| <b>6.</b>   | <b>References</b>                                                                                                                                     | <b>218</b> |

## List of Figures

|                                                                                                                                                                               |    |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| Figure 1. The schematic representation of cell envelopes in Gram-positive and Gram-negative bacteria. ....                                                                    | 15 |
| Figure 2. The cleavable signal peptides.....                                                                                                                                  | 17 |
| Figure 3. The BAM machinery and outer membrane protein insertion.....                                                                                                         | 19 |
| Figure 4. Protein translocation through Sec or Tat pathway. ....                                                                                                              | 21 |
| Figure 5. The components of the Sec system.....                                                                                                                               | 22 |
| Figure 6. The components of the Tat translocase. ....                                                                                                                         | 23 |
| Figure 7. Bacterial and archaeal flagella.....                                                                                                                                | 25 |
| Figure 8. Various types of pili in Gram-negative bacteria. ....                                                                                                               | 26 |
| Figure 9. Bacterial type IVa pilus machine (T4PM).....                                                                                                                        | 27 |
| Figure 10. Type IVa pilus assembly machinery from <i>Myxococcus xanthus</i> (PDB:3JC8). ....                                                                                  | 29 |
| Figure 11. Diversity of the bacterial secretion systems. ....                                                                                                                 | 30 |
| Figure 12. The type IV filament superfamily of nanomachines. ....                                                                                                             | 35 |
| Figure 13. Genetic organization of T2SSs from <i>V. cholerae</i> , <i>E. coli</i> , <i>D. dadantii</i> , <i>P. aeruginosa</i> and <i>L. pneumophila</i> .....                 | 37 |
| Figure 14. Schematic representation of the <i>Klebsiella</i> Type II secretion system. ....                                                                                   | 38 |
| Figure 15. Schematic working model of the T2SS. ....                                                                                                                          | 40 |
| Figure 16. The secretin architecture is conserved in different secretion systems. ....                                                                                        | 41 |
| Figure 17. The CryoEM structure of PulD monomer and multimer.....                                                                                                             | 42 |
| Figure 18. The CryoEM structure of PulD Multimer and PulD with PulS and PulC <sub>HR</sub> .....                                                                              | 43 |
| Figure 19. Gallery of T2SS inner membrane components. ....                                                                                                                    | 44 |
| Figure 20. The CryoEM structure of PulD-S-C complex and AlphaFold 2 models of PulC. ...                                                                                       | 45 |
| Figure 21. The structure of PulL <sub>cyto</sub> domain and the AlphaFold 2 models of full-length PulL, PulM and PulF. ....                                                   | 46 |
| Figure 22. The structures of PulE.....                                                                                                                                        | 47 |
| Figure 23. The structures and AF2 models of the <i>Klebsiella</i> pseudopilus subunits. ....                                                                                  | 48 |
| Figure 24. Major pseudopilin assembly.....                                                                                                                                    | 49 |
| Figure 25. The PulG pseudopilus structure. ....                                                                                                                               | 49 |
| Figure 26. Minor pseudopilin periplasmic domains from <i>E. coli</i> . ....                                                                                                   | 50 |
| Figure 27. Pseudopilus assembly initiation and pseudopilus structural model. ....                                                                                             | 51 |
| Figure 28. AlphaFold 2 structural models of PulB, PulN and PulO. ....                                                                                                         | 52 |
| Figure 29. PulA X-ray structure and secretion signal studies.....                                                                                                             | 55 |
| Figure 30. <i>Klebsiella oxytoca</i> T2SS substrate PulA: structure and domain organization. ....                                                                             | 56 |
| Figure 31. PulA secretion signal studies: the roles of Ins and N1 domains. ....                                                                                               | 57 |
| Figure 32. <i>K. oxytoca</i> PulA domain organization and mobility. ....                                                                                                      | 57 |
| Figure 33. Schematic three-dimensional model of the <i>P. aeruginosa</i> Xcp secreton. showing substrate recognition and transport during the type II secretion process. .... | 59 |
| Figure 34. Interactions between the periplasmic domains of the <i>P. aeruginosa</i> Xcp secreton and substrates.....                                                          | 60 |
| Figure 35. The models of type II secretion mechanism.....                                                                                                                     | 61 |
| Figure 36. The principle of pBpa incorporation into proteins in vivo.....                                                                                                     | 64 |

|                                                                                                                                                                                |    |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| Figure 37. The effect of UV irradiation: covalent cross-linking.....                                                                                                           | 69 |
| Figure 38. In silico prediction of interactions between the pseudopilus and PulA. ....                                                                                         | 70 |
| Figure 39. Residues of PulG monomer chosen for site-directed mutagenesis.....                                                                                                  | 70 |
| Figure 40. Selection of <i>E. coli</i> host strains for pBpa incorporation.....                                                                                                | 71 |
| Figure 41. The 3-plasmid system for PulG in vivo site-specific photo-cross-linking in <i>E. coli</i> .<br>.....                                                                | 73 |
| Figure 42. Test of pBpa incorporation in different PulG positions. ....                                                                                                        | 74 |
| Figure 43. The work flow of the in vivo site-specific photo-cross-linking assay. ....                                                                                          | 74 |
| Figure 44. Immunodetection of PulG-pBpa variants after the in vivo site-specific photo-<br>cross-linking assay. ....                                                           | 75 |
| Figure 45. Validation of PulG crosslinks without Pul factors. ....                                                                                                             | 76 |
| Figure 46. Residues in PulA N1 and Ins domains selected for site-direct mutagenesis. ....                                                                                      | 77 |
| Figure 47. The 3-plasmid system for PulA in vivo site-specific photo-cross-linking in <i>E. coli</i> .<br>.....                                                                | 78 |
| Figure 48. PulA <sup>Am</sup> variants pBpa incorporation and in vivo site-specific photo-cross-linking<br>assay. ....                                                         | 79 |
| Figure 49. The assay of PulA-His <sub>6</sub> and PulA <sup>E71Am</sup> -His <sub>6</sub> secretion, pBpa incorporation and in<br>vivo site-specific photo-cross-linking. .... | 80 |
| Figure 50. Coomassie blue staining of purified PulA-His <sub>6</sub> and PulA <sup>E71pBpa</sup> -His <sub>6</sub> proteins.....                                               | 81 |
| Figure 51. PulA-His <sub>6</sub> pBpa incorporation and in vivo site-specific photo-cross-linking assays.<br>.....                                                             | 82 |
| Figure 52. Validation of PulA <sup>Q521Am</sup> -His <sub>6</sub> crosslinked partner by anti-Pul antibodies.....                                                              | 82 |
| Figure 53. Immunodetection of the purified PulA <sup>Q521Am</sup> -His <sub>6</sub> after in vivo site-specific photo-<br>cross-linking assays.....                            | 83 |
| Figure 54. The 2-plasmid system for PulA in vivo site-specific photo-cross-linking in <i>E. coli</i> .<br>.....                                                                | 84 |
| Figure 55. In vivo site-specific photo-cross-linking assays with PulA <sup>Am</sup> -Strep variants in the<br>2-plasmid system. ....                                           | 85 |
| Figure 56. Immunodetection of the purified PulA <sup>Am</sup> -Strep variants after in vivo site-specific<br>photo-cross-linking assays. ....                                  | 86 |
| Figure 57. TCA precipitation of the elution fractions of PulA <sup>Am</sup> -Strep variants after<br>purification. ....                                                        | 87 |
| Figure 58. Analysis of PulA <sup>Q525Am</sup> -Strep crosslinked adducts by different anti-Pul antibodies.<br>.....                                                            | 88 |
| Figure 59. Secretion assay of PulA-Strep and the effect of the Pul factors, PulE and the Ins<br>domain of PulA.....                                                            | 90 |
| Figure 60. PulA-Strep secretion and PulM levels at different growth stages.....                                                                                                | 91 |
| Figure 61. Formaldehyde crosslinking of PulA-Strep + Pul and PulA-Strep + PulΔpule. ....                                                                                       | 93 |
| Figure 62. SPINE with PulA-Strep + Pul (A) and PulA-Strep + PulΔpule (B). ....                                                                                                 | 94 |
| Figure 63. SPINE of PulA-Strep + PulΔE. ....                                                                                                                                   | 95 |
| Figure 64. SPINE of PulA-Strep + PulΔE probed by different anti-Pul antibodies.....                                                                                            | 96 |
| Figure 65. BACTH analysis of PulA-MalF <sup>TMS</sup> with pseudopilins.....                                                                                                   | 98 |
| Figure 66. BACTH analysis of PulG interaction with PulA, N1 domain and PulAΔIns. ....                                                                                          | 99 |

|                                                                                                                 |     |
|-----------------------------------------------------------------------------------------------------------------|-----|
| Figure 67. BACTH analysis of PulA interactions with inner membrane T2SS components.                             | 100 |
| Figure 68. BACTH analysis of PulA (MalF TMS) with PulF, and PulF with PulK/MalF-TMS interactions.....           | 101 |
| Figure 69. BACTH analysis of PulA (MalF TMS) with PulF, PulF <sup>L198Flag</sup> and PulF <sup>A201Flag</sup> . | 102 |
| Figure 70. BACTH analysis of PulA (PulL+PulM).                                                                  | 191 |
| Figure 71. Design of partially overlapping primers .....                                                        | 203 |
| Figure 72. UV LED light equipment. ....                                                                         | 214 |

## List of Tables

|                                                                   |     |
|-------------------------------------------------------------------|-----|
| Table 1. Bacterial strains used in this study. ....               | 202 |
| Table 2. Plasmids used in this study. ....                        | 205 |
| Table 3. Oligonucleotides used for construction of plasmids. .... | 208 |

# **1. Introduction**

## 1.1 Bacterial cell envelope

The majority of bacteria are unicellular organisms that can be found in a wide range of habitats. Bacterial cells are delimited by their envelopes, multi-layered structures that protect the cell and allow for exchange of nutrients and metabolites with the environment. Bacterial envelopes are also the basis to distinguish at least two main groups of bacteria, based on the envelope layer thickness and the number of membranes they contain (Silhavy et al., 2010).

Gram-positive, or monoderm bacteria have a cytoplasmic membrane composed of a phospholipid bilayer and a thick multi-layered peptidoglycan (PG) cell wall that provides the cell with its characteristic shape and rigidity. The cell wall is a large polymer composed of long glycan chains, with alternating units of N-acetyl glucosamine and N-acetyl muramic acid linked via  $\beta$ 1-4 glycosidic bonds. These glycan chains are crosslinked by short peptides, whose number and composition can vary, but which contain D-amino acids. The glycans and peptides create a two-dimensional network with holes of various sizes. In Gram-positive bacteria, multiple PG layers form the external cell wall whose thickness can be up to 80 nm.

Gram-negative bacterial envelope consists of the outer membrane (OM) and the inner membrane (IM), delimiting an aqueous space called the periplasm, which contains a thin layer (about 8 nm) of peptidoglycan (Silhavy et al., 2010). The IM, or the cytoplasmic membrane, is a phospholipid bilayer forming a semi-permeable barrier between the cytoplasm and the periplasm. The OM is asymmetric and composed of phospholipids in the inner leaflet and lipopolysaccharide (LPS) in the outer leaflet (Nikaido, 2003). LPS are macromolecules composed of lipid A and oligosaccharides that form the O-antigens, outer core saccharide (Kamio & Nikaido, 1976). LPS binds  $Mg^{+}$  and  $Ca^{2+}$  ions and reinforces the barrier function of the OM, by making it impermeable to small hydrophobic molecules. The LPS is also called endotoxin because it can cause septic shock associated with the septicemia caused by Gram-negative organisms. It plays an important role at the interface with the host immune system and its mimics can serve as vaccine antigens.

The schematic comparison of Gram-positive and Gram-negative bacterial cell envelopes is shown in Figure 1. However, there are exceptions to these types of envelope compositions: for example, the *Corynebacteria* (Burkovski, 2013) or *Mycobacteria* (Dulberger et al., 2020) are positive for the Gram stain, however, unlike in the typical Gram-positive bacteria, their cell envelope consists of inner and outer membrane layers. Their outer barrier layer is composed of different mycolic acid lipids and is called mycomembrane. The intermembrane space of these

bacteria contains peptidoglycan and arabinogalactan polysaccharides. To include these exceptions, the terms monoderm and diderm are often used to classify the bacteria according to their number of membrane layers.

In addition, many bacteria assemble an outermost coat, the S-layer, which is composed of a single protein that totally encases the organism, and some have the capsule coat which is usually composed of polysaccharides (Silhavy et al., 2010).

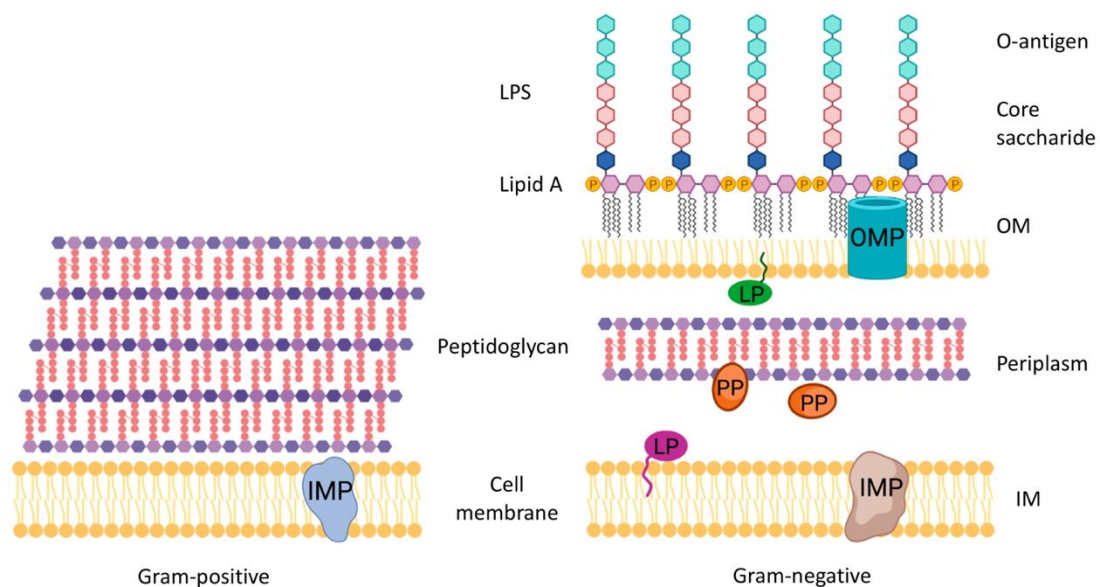


**Figure 1. The schematic representation of cell envelopes in Gram-positive and Gram-negative bacteria.**

IM: inner membrane; OM: outer membrane; IMP: inner membrane protein; OMP: outer membrane protein; LP: lipoprotein, inner membrane lipoprotein (magenta), outer membrane lipoprotein (green); PP: periplasmic protein; LPS: lipopolysaccharide.

## 1.2 Extra-cytoplasmic and membrane proteins

All bacterial proteins are synthesized in the cytoplasm, however about one third of them function in the cell membrane or in extracellular compartments. Proteins can insert in the cytoplasmic membrane or cross this barrier to reach more distant - periplasmic, outer membrane or extracellular compartments. Specific motifs in the primary, secondary or tertiary protein structure promote targeting to these compartments or to diverse transport machineries which allow proteins to reach their correct final localization outside the cytoplasm. The sequence

motifs that allow proteins to cross the cytoplasmic membrane in bacteria are called the signal peptides (Kaushik et al., 2022).

### 1.2.1 Signal peptides

Signal peptides (SP) are short peptides (typically 20-30 residues) located at the N-termini of exported proteins and are the key feature for general protein secretion system in cells of all domains of life-archaeal, eukaryote and bacterial. SP is essential for the secretory proteins to be guided to their target sites, recognized by translocons and translocated across the IM (Owji et al., 2018). The general structural features of SP are: a positively charged region in the N-terminus; an alpha-helical hydrophobic core region in the center; a C-terminus containing residues for recognition and cleavage by a specific signal peptidase (Fig. 2).

The majority of proteins targeted to the periplasm and the outer membrane in Gram-negative bacteria are synthesized with a cleavable Sec-dependent signal peptide (SP). On its C-terminus there is a cleavage site for the type I signal peptidase (SPase) with small side-chains in the -1 and -3 position, often with a consensus Ala-X-Ala (Perlman & Halvorson, 1983). After translocation through the cytoplasmic membrane, the SP cleavage is an essential process for the release of the secreted proteins (Dalbey & Wickner, 1985). Those proteins destined to the periplasm are folded and remain in this compartment as soluble proteins, while the OM proteins are taken up by the specific chaperones and targeted to the Bam complex, which promotes their folding into a beta-barrel and insertion in the OM.

Signal peptidases (SPases) are membrane associated enzymes cleaving off the SP from secretory preproteins. SPase I is the first identified and most studied bacterial SPase, which cleaves off the SP from the majority of secreted proteins that are exported in soluble form by the SecYEG and Tat pathways (Auclair et al., 2012).

Signal peptides directing protein to the Tat pathway are similar to the Sec pathway SPs, however their N-terminal polar region is extended and has a specific consensus sequence.

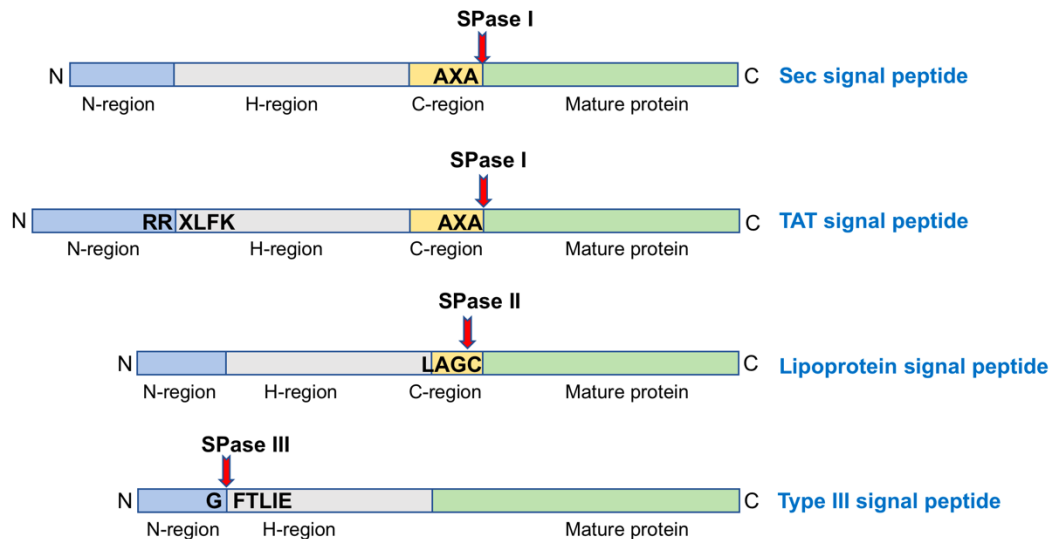


**Figure 2. The cleavable signal peptides.**

Schematic representation of different types of signal peptides with basic, N-regions in blue, the hydrophobic region in grey, the hydrophilic C-region in yellow and the mature proteins in green. The cleavage sites for different peptidases are shown as red arrows. Note that prepilin cleavage by SPase III generates the mature proteins composed of hydrophobic and hydrophilic segment.

### 1.2.2 Inner membrane proteins and the SRP pathway

In bacteria, there are many different types of inner membrane proteins, such as transporters, protein translocases, enzymes, respiratory complexes etc. They insert in the phospholipid membrane *via* hydrophobic transmembrane  $\alpha$ -helices. Based on the number of these transmembrane segments (TMSs), we distinguish bitopic IM proteins (with one TMS) and polytopic proteins with several TMSs. Bitopic proteins can have different orientations in the membrane:  $N_{in} - C_{out}$  or  $N_{out} - C_{in}$ , which depends on their charge distribution and on the presence of anchor sequences or globular domains. Positively charged protein segments tend to localize on the cytoplasmic side which is negatively charged and serve as anchors that prevent protein extraction from the IM.

Their insertion in the membrane is usually co-translational, promoted by the signal recognition particle (SRP) and FtsY receptor. The SRP binds to the hydrophobic segments early during translation and promotes binding of the nascent chain-ribosome complex to the FtsY, and to the SecYEG translocase complex (Lee & Bernstein, 2001). Also, the YidC protein promotes IM protein insertion. YidC forms a groove-like structure with their substrates at an

amphiphilic protein-lipid interface, allowing the substrate's transmembrane segments to slide into the lipid bilayer. YidC can act together with the Sec pathway but is also able to promote insertion of small IM proteins (e.g. phage coat protein Pf3 or membrane ring subunits of the F<sub>1</sub>F<sub>0</sub>ATPase) on its own, by the so-called YidC-only pathway (van der Laan et al., 2004).

### 1.2.3 Lipoproteins

Lipoproteins contain fatty acid- or lipid moieties that are usually attached to an amino terminal cysteine residue. They anchor the protein to the membrane and enhance its stability. In *Escherichia coli*, the best-studied model bacterium, the majority of lipoproteins are typically embedded in the inner leaflet of the OM. The major *E. coli* lipoprotein Lpp connects the OM with the PG. Some lipoproteins can also be secreted and remain anchored in the outer leaflet of the OM. An example of such proteins is the pullulanase from *Klebsiella*, secreted by the type II secretion system (d'Enfert et al., 1987) and many other lipoproteins secreted by this pathway (Rondelet & Condemine, 2013). Some lipoproteins can partially cross the OM through porins while their acyl tail remains anchored in the inner leaflet of the OM, such as the *E. coli* lipoprotein RscF (Konovalova et al., 2014).

Lipoproteins are synthesized in the cytoplasm with a signal peptide at the N-terminus, and cross the inner membrane most often through the Sec pathway. Their maturation in *E. coli* requires three enzymes: Lgt, LspA and Lnt (Buddelmeijer, 2015; Sankaran & Wu, 1994). Lgt attaches two acyl chains onto the cysteine following the signal peptide, after which the lipoprotein signal peptidase LspA removes the signal peptide.

The free N-terminus of the mature protein that results from this cleavage is acylated by Lnt, to give a lipoprotein modified by three fatty acids. Specific sorting signals at the N-terminus determine whether the mature lipoproteins will remain in the IM or end up in the OM. The OM lipoproteins will be recognized by LolCDE complex, extracted from the IM and transferred to LolA, which is a carrier protein in the periplasm. LolA carrying the lipoprotein moves across the periplasm and delivers it to LolB who is located in the outer membrane, under the help of LolB, the lipoprotein will finalize its insertion into the inner leaflet of OM (Narita & Tokuda, 2017). The lipoproteins which have a so-called IM retention signal (an Asp at position +2) will remain in the inner membrane (Terada et al., 2001).

### 1.2.4 Outer membrane proteins

In general, proteins in the OM can be classified into two categories, lipoproteins described above and the  $\beta$ -barrel proteins. While the former are typically peripherally associated with the membrane *via* lipid tails, the  $\beta$ -barrel proteins are fully inserted and form pores in the OM. These pores provide communication gates for passive or active transport of molecules in and out of the cell. Some porins like OmpF and OmpC, function for passive diffusion of small molecules across the OM. Others, allow for specific diffusion of solutes, such as maltose through LamB, or phosphate through PhoE (Delcour, 2009; Nikaido, 2003). In addition, OM contains channels that allow for transport of proteins or polysaccharides, which are often inserted as oligomers where each subunit contributes  $\beta$  strands to form a full  $\beta$ -barrel.

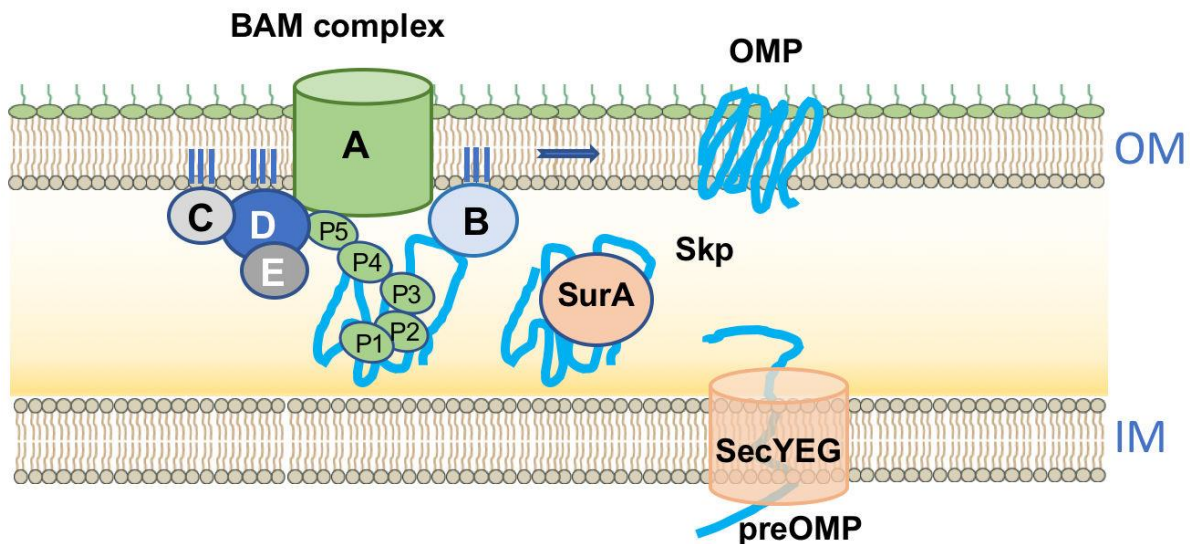


**Figure 3. The BAM machinery and outer membrane protein insertion.**

Schematic representation of outer membrane protein precursor (light blue) exported through the SecYEG, bound to periplasmic chaperons SurA or Skp and transferred to the Potra domains (P1-P5) of BamA. The Bam complex containing lipoproteins BamB, C, D and E promotes folding of the OM into  $\beta$ -barrel and its insertion of the OM.

Porins represent the majority proteins in the outer membrane, they are integral and transmembrane proteins found mainly in  $\beta$ -barrel conformation, wrapped into cylinders. Their size can vary considerably ranging from small 8-strand barrels like OmpA to more than 24-stranded LptDE. The barrels have hydrophobic residues on the outside to interact with the membrane and hydrophilic residues in the lumen. As illustrated in Fig. 3, the OMPs are first transported through the IM to the periplasm by the Sec pathway, then with the help of

chaperones like SurA, the proteins are guided to the outer membrane, where the insertion and folding of these  $\beta$ -barrel porins are promoted by the  $\beta$ -barrel assembly machinery (BAM) (Wu et al., 2020).

BAM is essential for OMP folding into a  $\beta$ -barrel and insertion into the OM and also for bacterial viability. The BAM complex is composed of BamA, itself a  $\beta$ -barrel protein integrated in the OM, and four OM lipoproteins anchored in the periplasmic side of the OM which aid in BamA function (Han et al., 2016). The five components show a relative stoichiometric ratio of 1:1:1:1:1 (Wu et al., 2020). BamA forms a crown on the top of the machinery and embedded in the OM, and its five polypeptides transport-associated (POTRA) domains, named P1-5, interact with the other four lipoproteins components in the periplasm (Fig. 3). Thus, the binding of POTRA domains of BamA to the lipoproteins BamBCDE forms a chamber and promote the processing of OMPs. BamB has an eight-bladed  $\beta$ -propeller fold, and thought to function as scaffold protein to optimize the orientation of the flexible periplasmic domains of BamA for interaction with other BAM components or substrates (Jansen et al., 2015). BamC has a flexible N-terminal domain that interacts with BamD (Kim et al., 2011). BamD consists of 5 tetratricopeptide repeated (TPR) domains, binding to the POTRA-P5 of BamA (Albrecht & Zeth, 2011).

To maintain impermeability, the OM pores are usually plugged by flexible extracellular or periplasmic loops or by extra globular domains that are folded in the pore lumen.

#### 1.2.5 Protein translocation across the IM

The vast majority of bacterial proteins is synthesized on the ribosomes in the cytoplasm, but many functions outside of this compartment. These proteins can use two different pathways to either insert or cross the inner membrane. Below we describe two general export/insertion pathways, the secretory (Sec) pathway and the twin-arginine translocation (TAT) pathway (Pei & Dalbey, 2022), which are illustrated in Fig. 4. The Sec pathway promotes export or membrane insertion of unfolded proteins, during or after their synthesis. The Tat pathway, on the other hand transport already folded proteins, sometimes in complex or bound to metal or other cofactors.

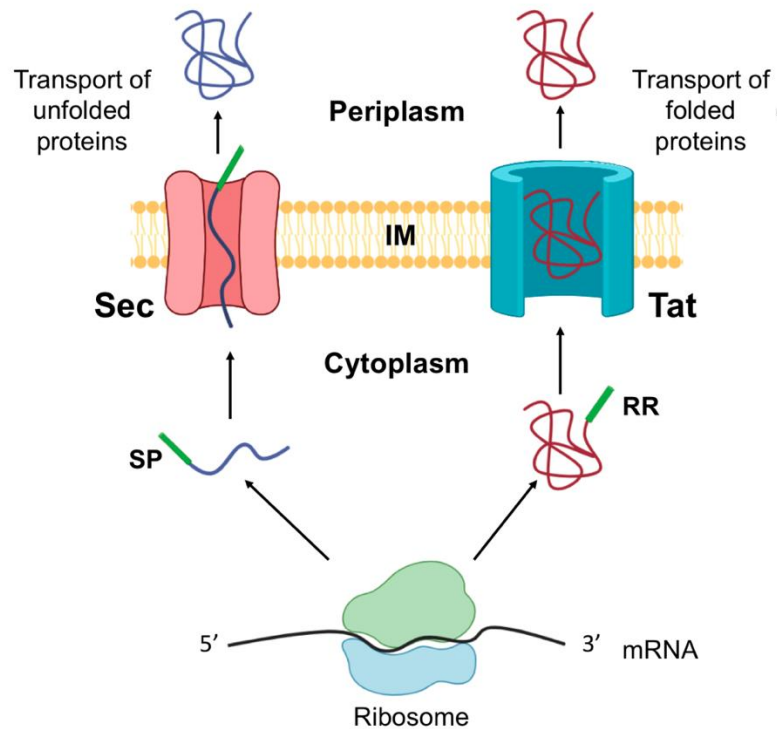


**Figure 4. Protein translocation through Sec or Tat pathway.**

IM: inner membrane; SP: Sec-dependent signal peptide; RR: Tat-dependent signal peptide with a twin-arginine motif.

#### 1.2.5.1 The Sec pathway

The classical secretory (Sec) pathway, the most widely used pathway for protein export from cytoplasm to periplasm, is evolutionarily conserved and is widely present in all domains of life. It operates in the plasma membrane of bacteria, the ER membrane of eukaryotes, and the thylakoid membrane of chloroplasts in plants. The basic Sec system comprises three components, SecY, SecE and SecG, with SecY forming a narrow channel (Cranford-Smith & Huber, 2018) in the inner membrane allowing unfolded proteins to cross or insert into the cytoplasmic membrane.

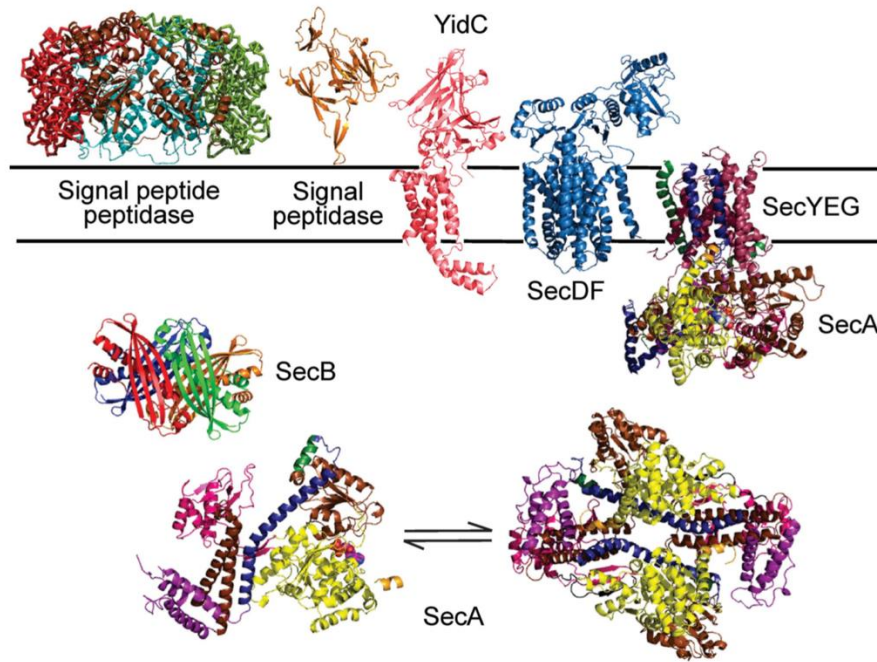


**Figure 5. The components of the Sec system.**

Cytoplasmic chaperon SecB and the ATPase SecA target preproteins to the SecYEG IM complex. SecDE complex couples the proton motive force to protein export. The signal peptidase cleaves the signal peptide, which is further degraded by the signal peptide peptidase. IM: inner membrane. Model taken from (Crane & Randall, 2017).

In bacteria, SecA, the cytoplasmic ATPase motor, associates with SecYEG complex at the IM to recognize and translocate unfolded protein substrates. A subset of exported proteins binds to SecB, which prevents their premature folding and targets them to SecA. Soluble proteins dependent on Sec-pathway are synthesized at the ribosome (Steinberg et al., 2018) as preprotein forms (Tsirigotaki et al., 2017), which have a cleavable N-terminal signal peptides (SP) preceding the mature domain (Chatzi et al., 2012). For the IM proteins, the hydrophobic segments partition into the membrane through the lateral opening formed by the TMS1 and TMS10 segments of SecY. Fig. 5 depicts the structures of all the major components of the Sec pathway.

#### 1.2.5.2 The Tat Pathway

In contrast to the Sec pathway, the Tat system has mostly been implicated in the export of fully folded and/ or cofactor containing proteins (Pei & Dalbey, 2022). Some proteins cannot be secreted in their unfolded state, and must undergo post-translational modifications in the cytoplasm. This is because the molecules required for the modification are only found in the

cytoplasm (Berks et al., 2005). The substrates for Tat pathway are more exclusive. For instance, in *E. coli* and *Salmonella* there are only about 30 different types of substrates for Tat pathway (Cline, 2015; Frain, Robinson, et al., 2019; Frain, van Dijl, et al., 2019). Although the Tat and Sec systems translocate proteins by fundamentally different mechanisms, several common elements can be recognized (Natale et al., 2008), for example, similar to the Sec system, Tat -dependent route is an active, transmembrane proton motive force (PMF)-dependent process. Tat pathway is present in most of the prokaryotes, all chloroplasts and some mitochondria and plays essential roles in photosynthesis, respiratory energy metabolism and virulence (Berks, 2015). In order to be recognized by Tat translocase, those substrates have a N-terminal signal peptide containing a consensus twin arginine motif (RR). Tat apparatus (Fig. 6) comprises small membrane proteins from two structural families, TatA and TatC. In the TatA family, there is a functionally distinct homologous member called TatB, and a third member called TatE which could be a functional partner of *E. coli* Tat pathway (Eimer et al., 2018). In addition, there is another Tat component named TatD whose function is not well characterized (Palmer et al., 2010).

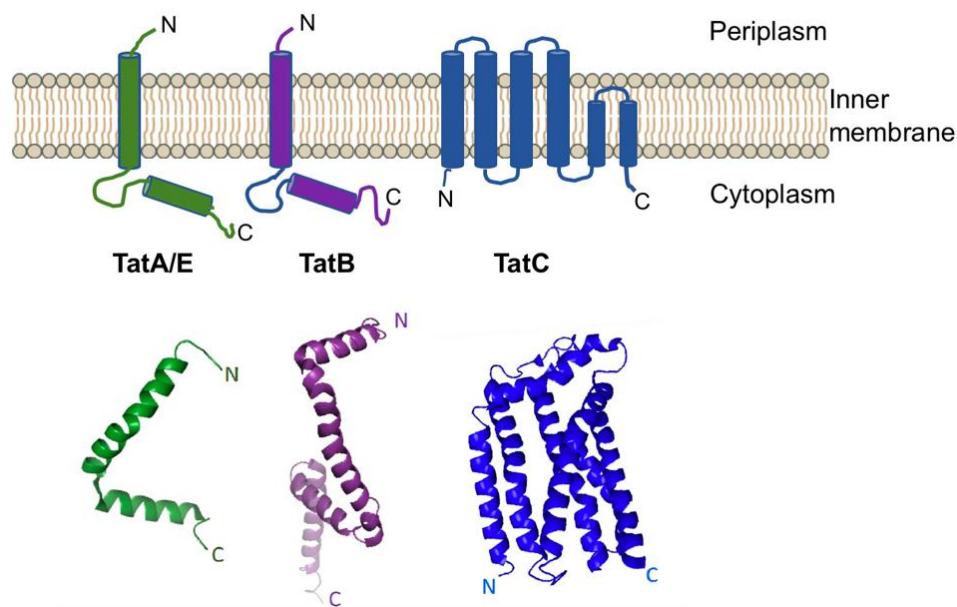


**Figure 6. The components of the Tat translocase.**

*Top: TatA (TatE), TatB and TatC topology in the inner membrane. Bottom: Cartoon models of their structures.*

### 1.3 Cell surface appendages

Bacterial activity is greatly influenced by the physicochemical characteristics and dynamics of the bacterial envelope. There are many different appendages on the bacterial cell surface, like flagella and pili, which play important role in bacterial motility, adhesion, transport of DNA or proteins.

#### 1.3.1 Flagella

Although there are many different mechanisms for bacterial movement, the flagellum is the type of motility organelle that is the best understood (Jarrell & McBride, 2008). Flagella are embedded in the cell envelope and based in bacterial periplasmic space, rotate or undulate the helical cell body for propulsion in liquid media. In some bacteria such as Spirochetes, flagella are in the periplasm serving as a cytoskeleton to define the helicity of the cell body. Flagella are genetically, structurally, and functionally related to Type 3 secretion system (T3SSs) and consist of three major substructures: the filament, the hook, and the basal body (Blocker et al., 2003) (Fig. 7). The long helical flagella filament works as a molecular screw to provide thrust in bacterial species that have flagella at the cell surface. The basal body, or the flagellar motor, which is found at the base of the filament, is a common rotary nanomachine shared by flagellated bacteria. It is made up of a rotor ring complex and several transmembrane stator units. The flagella motor turns the ion flux through an ion channel of each stator unit into the mechanical work necessary for motor rotation. In response to environmental changes, intracellular chemotactic signaling pathways control the direction of flagella-driven motility, enabling bacteria to move toward settings that are more favorable for their survival (Nakamura & Minamino, 2019).

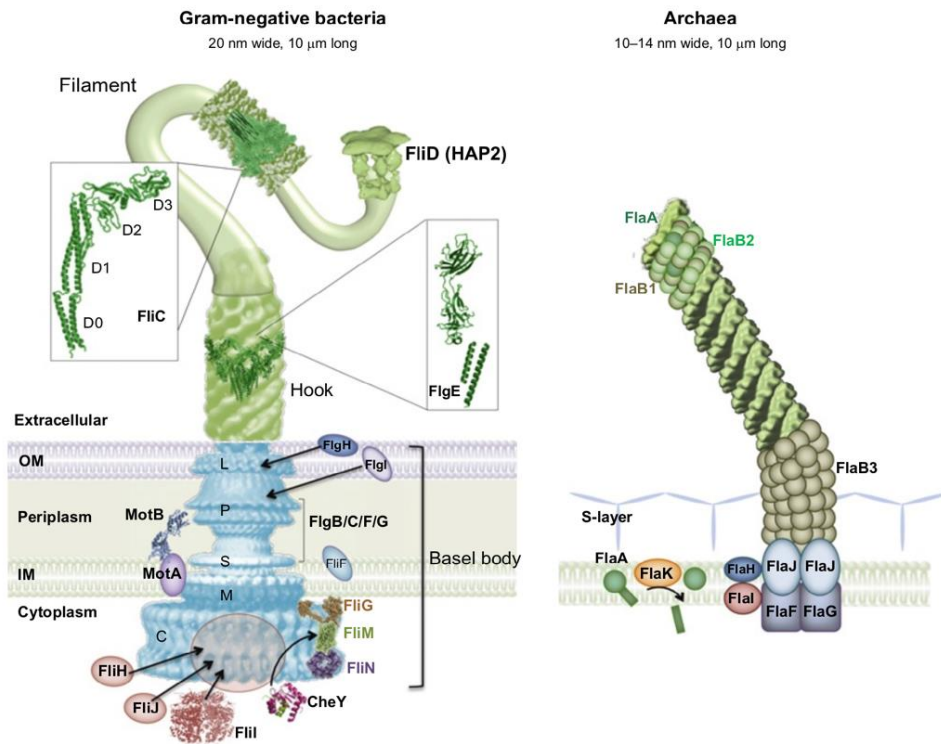


**Figure 7. Bacterial and archaeal flagella.**

Left, the complex machinery of bacterial flagellum with the basal body and motor complex components in the bacterial envelope, extended to extracellular hook and filament. Right, archaeal flagellum, which resembles bacterial type IV pili (Van Gerven et al., 2011).

The cost of flagellar assembly is believed to be significant, and flagella-mediated motility is a conditionally beneficial trait. Over twenty structural proteins, with copies per cell ranging from fewer than ten to more than ten thousand, are used to construct each flagellum. Depending on the species, there might be one to dozens of flagella per cell. Flagellar gene expression is closely controlled to reduce excessive energy consumption (Mukherjee & Kearns, 2014), and the most abundant proteins appear to have evolved to avoid energetically expensive amino acids (Smith & Chapman, 2010). Time is a considerable additional expense in flagellar assembly. A functional flagellum takes roughly an hour to build, and if motility is required for nutrition uptake and growth, the delay between induction and active motility in a flagellate population could be lethal (Guttenplan et al., 2013; Karlinsey et al., 2000). Some bacteria use so-called bet-hedging techniques of bistable/bimodal gene expression to pre-differentiate a motile subpopulation in order to eliminate the time delay and sequester the cost of nutrients (Cummings et al., 2006; Grimbergen et al., 2015; Kearns & Losick, 2005; Norman et al., 2013; Stewart & Cookson, 2014).

### 1.3.2 Pili

Bacteria frequently exhibit long protein extensions on their cell surfaces called pili or fimbriae to interact with their surroundings. These hair-like polymers, which are generally thinner than flagella, are made of repetitive pilin subunits that interact covalently or non-covalently. Depending on their different assembly pathways, pili can be classified as chaperone-usher pili (represented by type I pili), type V pili, type IV pili (T4P) and conjugative pili (represented by type IV secretion pili) (Fig. 8), as well as sortase-dependent pili specific for Gram-positive bacteria. Most pili play crucial roles in bacterial adhesion, but also participate in cell-cell contacts, DNA uptake and protein secretion, biofilm formation, as well as motility (Hospenthal et al., 2017).

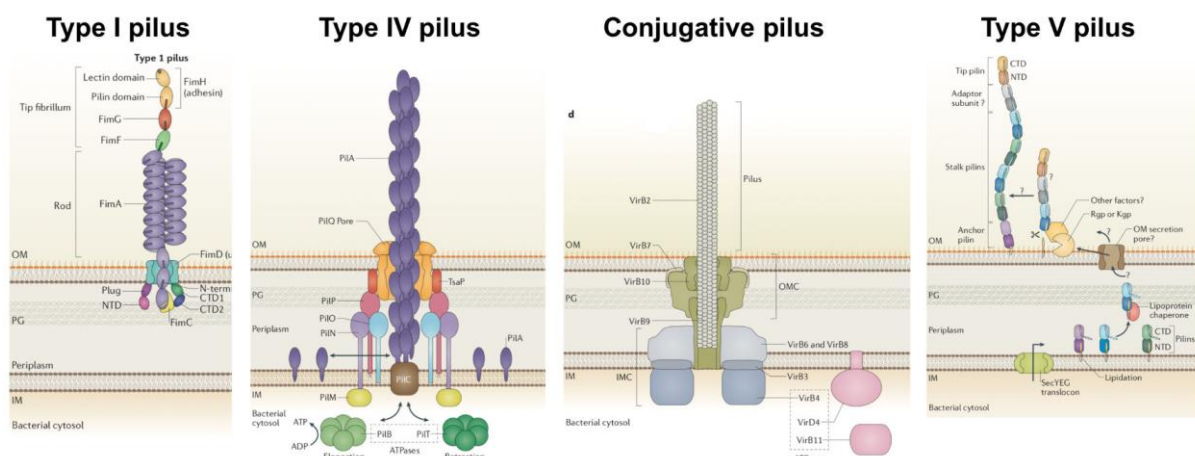


**Figure 8. Various types of pili in Gram-negative bacteria.**

From left to right: type I pilus; type IV pilus; conjugative pilus and type IV pilus (Hospenthal et al., 2017).

The Type IV pili (T4P) will be presented here in more detail, due to their relation with the type II secretion system, the focus of this work. T4P are anchored in the cell envelope and are characteristic dynamic filaments that can rapidly extend and retract while withstanding severe stresses, the dynamic of T4P are drive by type IV pilus machine (T4PM). T4PMs are made of an outer membrane secretin, an inner membrane assembly platform and cytoplasmic ATPase motors (Fig. 9) (Chang et al., 2016). They assemble long flexible filaments on the surface of numerous bacteria. T4PMs belong to type IV filament (T4F) superfamily, which includes Type II secretion system (T2SS) in bacteria as well as flagella (Fig. 7) and pili in archaea. The

similarity and presence of T4F in all domains of life indicates that T4P and T2SS have evolved from the last universal common ancestor (LUCA) (Denise et al., 2019).

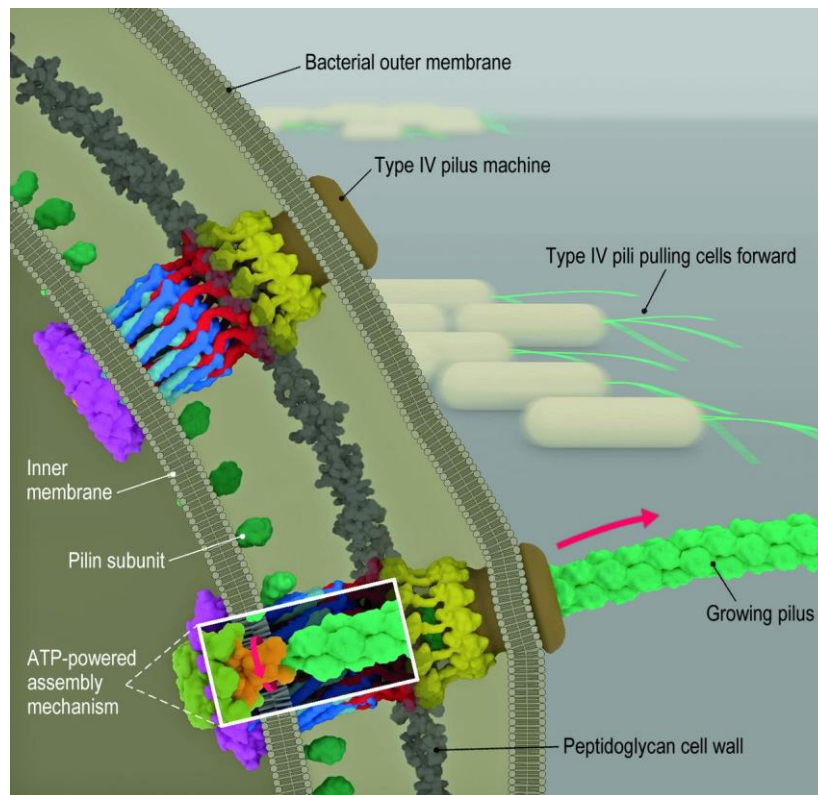


**Figure 9. Bacterial type IVa pilus machine (T4PM).**

Two T4PMs are depicted on the left, spanning the envelope of a Gram-negative bacterial cell. T4PMs extend and retract pili to pull cells forward along the surfaces. The pilin is added to the base of the pilus at the assembly motor powered by that ATP hydrolysis. The process is reversed during retraction (Chang et al., 2016).

There are several classes of T4P: type IVa pili (T4aP), type IVb pili (T4bP) and Tad or type IVc pili. This classification is based on the pilin sequence, the length of the leader peptide, and minor differences in their assembly (Craig et al., 2004). The genes required for pilus formation are organized differently in the T4aP and T4bP systems. T4aP genes are distributed across the genome in several operons, whereas T4bP and Tad genes are concentrated in a single operon (Pelacic, 2008). T4bP is involved in the formation of biofilms, bacterial colonization, and cell adhesion (Roux et al., 2012). T4aP has been identified as the molecular framework underlying bacterial twitching motility as well as other functions like DNA uptake (Piepenbrink, 2019). Their flexibility, elongation, and retraction enable them to perform these various functions.

The stem of T4Ps are constituted of thousands of copies of a single subunit known as the major pilin. Minor pilins, which have a much lower abundance, have been shown to be essential for pilus formation and function (Giltner et al., 2010). Minor pilins are divided into core and non-core minor pilins (Giltner et al., 2012). Despite their scarcity, core minor pilins play an important role in initiating pilus formation. In most systems, there are four core minor pilins, encoded by genes clustered in one operon, which are homologous to the genes encoding four minor pseudopilins in the T2SS (Jacobsen et al., 2020).

The extension or retraction of the pili is made possible by addition or removal of major pilin at the base of pilus; the assembly or disassembly rely on the inner membrane platform protein PilC, and two ATPase motors, PilB and PilT in the cytoplasm. PilC interacts directly with the pilins and the ATPases, which presumably rotate PilC to either elongate or retract the pilus (Takhar et al., 2013; Tsai & Tainer, 2016). PilB and PilT are located in the center of the PilM cytoplasmic ring, close to the inner membrane. PilM's cytoplasmic ring interacts with PilN N-terminal transmembrane tail (Karuppiyah & Derrick, 2011). PilN forms a heterodimer with PilO, which assembles in the periplasm as a ring structure with the same stoichiometry as the cytoplasmic ring. PilP, an inner membrane protein, binds to this complex and connects it to PilQ's periplasmic domain. The large PilQ multimer forms the secretin pore in the outer membrane that allows the pilus to pass through (Berry et al., 2012). As a result, the PilMNOP subcomplex connects the cytoplasmic ATPases to the secretin pore and ensures that the assembly platform motors, PilC and PilB/PilT, are aligned with the secretin (Chang et al., 2016). The pili extend out of the cell through the secretin pore to perform biological functions (Fig. 10).

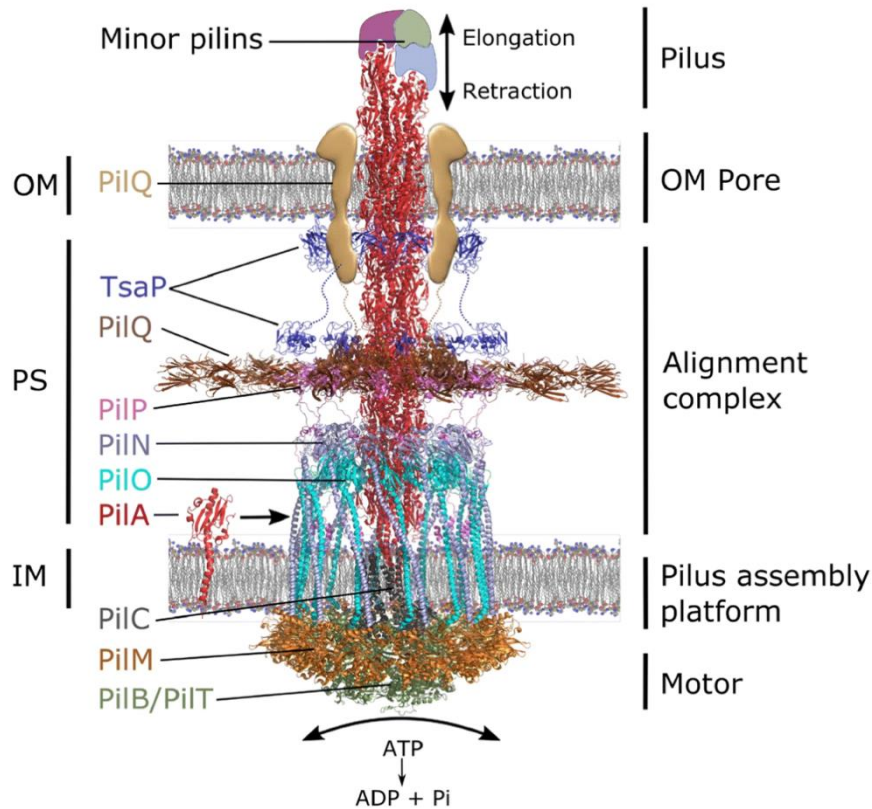


**Figure 10. Type IVa pilus assembly machinery from *Myxococcus xanthus* (PDB:3JC8).**

On the left the outer membrane (OM), the periplasmic space (PS) and the inner membrane (IM) are marked. The individual proteins involved in the formation of T4P are shown and labeled. On the right the spanning of the different sub-complexes (pilus, OM pore, alignment complex, pilus assembly platform and motor) are indicated (Jacobsen et al., 2020).

## 1.4 Gram-negative bacterial protein secretion systems

To adapt to their environment, bacteria have developed different protein secretion systems in their cell envelopes. Size, content, and architecture of these systems vary dramatically. Multiple megadalton sizes can be attained by double-membrane-spanning secretion systems, which are made up of many different protein subunits. These systems transport folded or unfolded substrates and use different energy sources for transport, from ATP to proton or entropy gradients (Costa et al., 2015). The most recent review discusses all secretion systems discovered so far in Gram-negative and diderm bacteria (Filloux, 2022).

Gram-negative bacteria secrete specific proteins across their double-membrane cell envelope to the cell surface or surrounding milieu through several types of complex machineries. The secreted proteins play important roles in nutrient acquisition, adhesion, motility, virulence, respiration, and inter-bacterial competition. In general, secreted proteins allow bacteria to better adapt to their environment. To date, excluding Sec and Tat pathway, 11 types of protein secretion systems have been identified in bacteria.

Gram-positive bacteria use Sec and Tat pathways to secrete proteins to the cell surface. However, some Gram-positive bacteria are diderm and use T7SS to transport specific proteins across their complex envelope.

In general, T2SS, T5SS, T7SS, and T9SS are two-step secretion systems either dependent on Sec or Tat pathway, while the T1SS, T3SS, T4SS, and T6SS span the entire envelope of diderm bacteria promote protein secretion in a single step directly from the cytoplasm (Fig. 11) (Maffei et al., 2017).

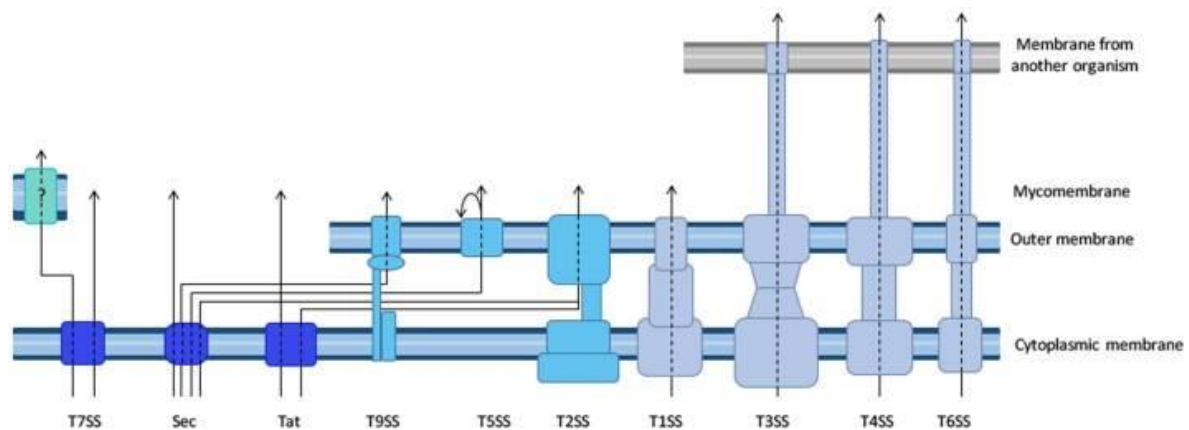


**Figure 11. Diversity of the bacterial secretion systems.**

Schematic representation of the secretion systems identified in bacteria, partially based on structural data. In monoderm bacteria (left), protein export (synonymous for secretion in that case) follows the Sec or Tat pathway, or the signal peptide independent T7SS. In diderm-non-LPS bacteria such as *Mycobacteria* or *Corynebacteria* (far left), it is unknown whether the T7SS system results in protein secretion in one or two steps (question mark). In diderm-LPS bacteria (center and right), secreted proteins can reach the external environment through a one-step process via T1SS, T3SS, T4SS, or T6SS. Other secreted proteins are first exported to the periplasm via the Sec system (T2SS, T5SS, or T9SS) or the Tat system (for T2SS only) (Maffei et al., 2017).

**T1SS** is a one-step secretion system. It secretes unfolded proteins and the substrate will fold after crossing of the OM. The substrate folding can promote its secretion and prevent back-sliding into the secretion apparatus. Three parts make up the canonical T1SS, which is represented by the haemolysin system. On the cytosolic side, the polytopic inner membrane protein HlyB has an ATP-binding and hydrolysis domain. The T1SS substrates feature a C-terminal non-cleavable targeting signal (Filloux, 2010) that directs them to HlyB. The two other accessory proteins are recruited and assembled by HlyB, which leads to the energy-dependent transport across the inner membrane (Koronakis et al., 1991; Thanabalu et al., 1998). By interacting with HlyB and the outer membrane protein TolC, HlyD – originally referred to as a "membrane fusion protein" – creates a continuous channel that runs from the cytosol to the extracellular environment while entirely bypassing the periplasm. TolC is a trimer that forms an exterior membrane-embedded  $\beta$ -barrel with an alpha-helical periplasmic conduit, giving it a remarkable three-dimensional structure (Koronakis et al., 2000).

**T2SS** is a two-step secretion pathway. In the first step, substrate is translocated from the cytoplasm to the periplasm through the Sec or Tat pathways, then the folded protein is secreted by the T2SS across the OM channel – the secretin. The T2SS and the machinery that assembles type IV pili (T4P) at the cell surface, which are involved in twitching motility, are quite similar. The machine that polymerizes the pili and exposes them at the surface for the T4P is almost identical to the T2SS (Ayers et al., 2010). This system will be discussed in detail in Chapter 1.5.

**T3SS** substrate secretion is typically believed to be a one-step process (Lavander et al., 2006). It recruits and transport effectors or toxins directly from cytoplasm into host cells. T3SS consists of three main parts: a base complex or basal body, the needle and the translocon (Abrusci et al., 2014). Due to their construction, T3SSs have been referred to as "injectisomes" and "needle and syringe"-like devices. They span the inner and outer bacterial membranes and secrete a wide range of proteinaceous substrates across the target eukaryotic cell membrane in the same step (Burkinshaw & Strynadka, 2014). The flagellar assembly system and the T3SS are highly similar (Diepold & Armitage, 2015). These similarities between a secretory system and surface structure assembly nanomachines are not unique. As mentioned above, they are also found in T2SS and T4P, or in T4SS described below.

**T4SS** is a one and in some cases a two-step secretion system. It is related to bacterial DNA conjugation and secretes a variety of substrates, monomeric proteins, protein-protein or DNA-protein complexes (Cascales & Christie, 2003). Bacterial type IV secretion systems transfer virulence factors into eukaryotic cells and transport genetic material between bacteria. VirB

system is notably used as a canonical example. The molecular mechanism of type IV secretion has proven challenging to analyze in the absence of structural information for the complete apparatus due to the complicated movements of the substrate through the secretion apparatus. T4SS has significantly distinct architecture from other known bacterial secretion systems, and as a result, mechanism. A sophisticated stoichiometric arrangement of eight proteins results in the formation of a 3 megadalton nanomachine that spans the whole cell envelope. A very large inner membrane complex, with 12 VirB4 ATPase subunits arranged as side-by-side hexameric barrels, is connected by a central stalk to an outer membrane-associated core complex (Low et al., 2014).

**T5SS** is a 2-step and Sec dependent secretion machinery that is one of the more common secretion systems, being found in 62% of Gram-negative bacteria whose genomes have been sequenced, with some species harboring more than 20 different T5SS (Abby et al., 2016). Interestingly, the substrates of T5SS are unique comparing to the other secretion systems, since instead of secreting other proteins or factors from the cell, they secrete themselves. Therefore, they are often called autotransporters (Leyton et al., 2012; van Ulsen et al., 2014, 2018). T5SS are currently classified into six sub-families (T5aSS to T5fSS) based on the three-dimensional structures of their respective translocator and passenger domains, as well as the organization of these domains in precursor polypeptides (Dautin, 2021). These subtypes share 2 crucial common steps: firstly, the Sec complex in IM to move the soluble unfolded substrates from cytoplasm to periplasmic region; and then, the  $\beta$ -barrel assembly machinery (Bam) in the outer membrane promote the folding of outer membrane proteins into a  $\beta$ -barrel conformation and their insertion into the outer membrane (Doyle & Bernstein, 2021; Ryoo et al., 2020).

**T6SS** is a nanomachine that bears resemblance to the bacteriophages. It assembles in the cytoplasm and injects effector proteins into other bacteria or eukaryotic cells in a contact-dependent manner and plays an important role in Gram-negative bacterial communication and competition (Russell et al., 2014). Three major components comprise the T6SS: The membrane complex (TssLMJ) (Bernal et al., 2021), which connects the inner and outer membranes (Rapisarda et al., 2019), has no homologs in the phage system; the remainder of the device is quite similar to contractile phage tails: on top of the membrane complex there is a ring-shaped baseplate complex (TssAEFGK) (Cherrak et al., 2018). The baseplate serves as the site of the contractile sheath assembly (TssBC), which extends in the cytosol and is made up of a helical tube that wraps around a stack of hexameric Hcp rings. A trimeric VgrG in the shape of a cone lies at the tip of this sheath and is known as a puncturing device. The flat side of the cone

interacts with the first Hcp ring and is surrounded by a baseplate. A PAAR protein placed on top can sharpen the conical side that faces the membrane complex. Contraction can begin when the assembling TssBC sheath crosses the entire cytoplasm and reaches the membrane across from the baseplate. The contact with the target cell leads to the baseplate conformational change, the Hcp stacks forcing their way out via the TssJ multimer in the outer membrane, and the VgrG puncturing tool moving outward through the membrane complex (Durand et al., 2015). Various effectors are injected in the target cell to degrade the membrane and the cell wall components, leading to rapid cell death. To protect the producing cell, these effectors are bound by the chaperones that keep them in an inactive state before secretion (Russell et al., 2014).

**T7SS** are found in both Gram-positive monoderm and diderm bacteria across a variety of bacterial phyla. Type VII effectors can influence interactions between bacteria as well as those between bacteria and eukaryotic cells. They exhibit a variety of species-specific biological functions (Unnikrishnan et al., 2017). The secretion of effector proteins by pathogenic and non-pathogenic *Mycobacteria*, including *Mycobacterium tuberculosis*, the primary cause of tuberculosis, depends heavily on Type VII secretion systems (T7SSs). It is still unknown how effector proteins of T7SSs cross the outer membrane of the diderm mycobacterial cell envelope. Although the structures and functions of T7SSs are highly versatile, there are some conserved components: two membrane proteins with (predicted) enzymatic activities: a predicted ATPase (EccC), likely to be required for energy provision of T7SS, and a subtilisin-like protease (MycP) involved in processing of specific substrates (Ates et al., 2016), however, the actual secretion process is poorly understood (Rivera-Calzada et al., 2021).

**T8SS**, also termed as curli biogenesis pathway, is a two-step and Sec-dependent system. Since the secreted proteins build the curli pili, it is debated whether this system should be considered as a protein secretion system strictly speaking. Two divergently transcribed operons, *csgBAC* and *csgDEFG*, control curli biogenesis. The *csgBAC* operon encodes curli structural components, CsgA and CsgB, as well as a chaperone-like protein, CsgC. Accessory proteins needed for effective transcription, secretion, and assembly of curli fibers are encoded by the *csgDEFG* operon. When CsgA and CsgB reach the cell surface, they change from their highly unstructured state as secreted proteins to  $\beta$ -rich structures that assemble into regular threads (Bhoite et al., 2019).

**T9SS** is the recently discovered bacterial secretion system, which only found in some species of the Bacteroidetes phylum (Abby et al., 2016). T9SS also play roles in bacterial movement referred to as gliding motility (McBride, 2019) and in virulence (Hajishengallis,

2015). Protein substrates of the T9SS containing N-terminal signal peptides for transport across the inner membrane by the Sec system and are targeted to the outer membrane translocon via their conserved C-terminal domain signal (CTD) of approximately 80 amino acid residues (Kulkarni et al., 2017). They cross the outer membrane and attach to the cell surface. In the pathogen *Porphyromonas gingivalis* causing periodontal disease at least 19 proteins, including the three transcription regulators PorX, PorY, and SigP, have been discovered part of the T9SS. Although the individual components are determined, little is known about the overall structure and molecular mechanism of the T9SS (Gorasia et al., 2020).

**T10SS** is a holin/peptidoglycan hydrolase-dependent protein secretion system. For many years it has been known that two chitinases are secreted from *Serratia marcescens* independently on the known mechanisms. Recently, Palmer *et al.* showed that this system contains at least four components, including peptidoglycan hydrolase associated and a phage-like holin. This new machinery was designated the Type X secretion system (Palmer et al., 2021). T10SS secretes substrate proteins from the periplasm to the cell surface by utilizing a holin membrane protein in conjunction with cell wall-editing enzymes which vary from different biological systems. According to genomic analysis, this pathway is conserved and involved in the secretion of hydrolytic enzymes and toxins by a variety of bacteria (Palmer et al., 2021).

**T11SS** is a most recently proposed protein secretion system. Recent bioinformatics and other analyses have identified a family of outer membrane transporters capable of secreting lipoproteins such as TbpB from *Neisseria* and soluble proteins in other species (Grossman et al., 2021). T11SS is a two-step pathway dependent on Sec or Tat pathway and can also promote secretion of soluble, nonlipidated proteins across the outer membrane and/or become tethered to the outer membrane, similar to the T2SS (Grossman et al., 2021). This system seems to have few components and its mechanism remains to be explored.

## 1.5 Type II secretion system

### 1.5.1 T2SS as members of the type IV filament superfamily

Among all the bacterial secretion systems, Type II secretion system (T2SS) is one of the most studied and best characterized since its discovery in the 1980s in the genus *Klebsiella* (D'Enfert & Pugsley, 1987; Pugsley, 1992). According to genome sequence analyses, T2SS is

largely specific to the Proteobacteria, and existing in the majority of genera in the Alpha-, Beta-, Gamma- and Delta-proteobacteria classes (Abby et al., 2016). The T2SS belongs to the type IV filament (T4F) superfamily which serve a wide range of functions and share homologous components and similar molecular mechanism (Berry & Pelicic, 2015). In addition to T2SSs (Naskar et al., 2021), other members of this superfamily include the type IV pili (T4P) in Gram-negative and Gram-positive bacteria (Fig. 12). Different subclasses of bacterial T4P include type IVa, and IVb pili (Craig et al., 2019), tight adherence (Tad) pili recently classified as type IVc pili (Tomich et al., 2007), mannose-sensitive hemagglutinin (MSH) pili (Marsh & Taylor, 1999) and competence pili (Piepenbrink, 2019).

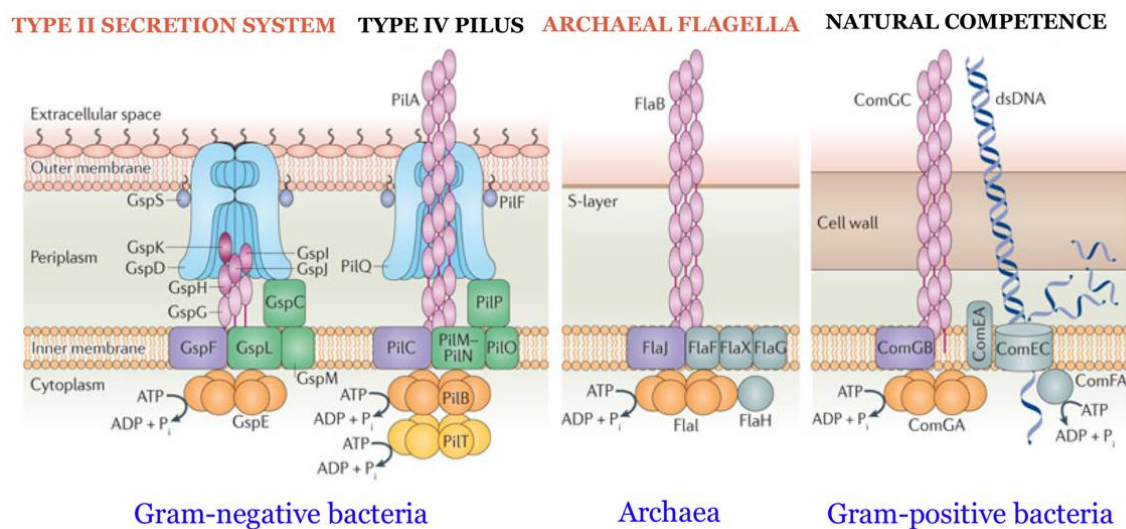


**Figure 12. The type IV filament superfamily of nanomachines.**

Gram-negative bacteria, archaea and Gram-positive bacteria assemble filaments from major (pseudo)pilin/flagellin subunits (in pink). The cytoplasmic ATPase (orange) provides energy for assembly and transfers it to the conserved membrane assembly components (in purple and green). In Gram-negative bacteria, the secretin channel (light blue) ensures the passage of pili and their substrates across the outer membrane (Korotkov et al., 2012).

In archaea, T4F appendages include different types of pili involved in adhesion and aggregation, bindosomes that promote nutrient uptake and flagella (also called archaella) that power the swimming motility (Makarova et al., 2016). All T4F systems share common essential elements and have the capacity to assemble polymers of pilin subunits that are non-covalently linked and dynamic. The pilin subunits initially reside in the IM, to be then extracted and

polymerized by the assembly platform components with the aid of assembly ATPase. Assembled pili have different surface charges and shapes that allow them to bind to specific substrates. Pili are generally thin (6-8 nm) and long, extending for several microns above the cell surface. However, the T2SS assemble short filaments called pseudopili, that reside in the periplasm in physiological conditions. They can be artificially extended and observed on the bacterial surface upon overexpression of the major pseudopilin gene or all the genes encoding the T2SS. It is thought that the pseudopili function to push folded substrates through the secretion channel in the outer membrane.

### 1.5.2 The genetic organization of T2SS and nomenclature of components

In Proteobacteria, the T2SS is usually composed of 12 to 15 different components known as GspA to GspO and GspS, which stands for General Secretion Pathway. T2SS component names are unique to each bacterial system, and share the capital letter designations; for example, OutB to OutO in *Dickeya* and EpsA to EpsO in *Vibrio*. Only the T2SS components of *Pseudomonas* were not included in this nomenclature and have been designated as XcpP to XcpZ, respectively, although this has changed more recently. The genes for the core T2SS components, which range from *gspC* to *gspM*, are generally organized into large operons. The prepilin peptidase gene *gspO* may follow the *gspC-M* operon or be found separately. Additional genes, such as *gspA*, *gspB*, and *gspS*, are independent transcription units that are frequently found near the core *gsp* genes. T2SS gene clusters have a remarkably conserved genetic organization. However, the position of the *gspCD* genes varies between species. In *P. aeruginosa*, for example, they can be found in an inverted order. In fact, the T2SS genes in *L. pneumophila* are divided into five distinct *loci* (Cianciotto, 2009). Some bacteria have a more than one set of T2SS genes that may be organized differently and phylogenetically distant from the "main" T2SS, such as the Hxc system in *P. aeruginosa* or the Stt system in *D. dadantii* (Douzi et al., 2012; Ferrandez & Condemine, 2008; Rondelet & Condemine, 2013) (Fig. 13).

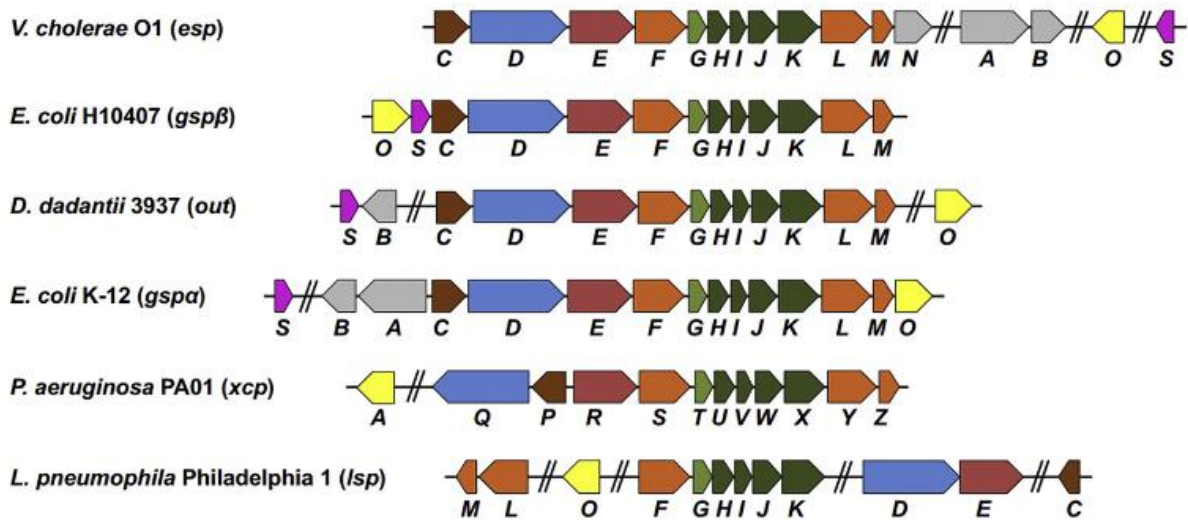


**Figure 13. Genetic organization of T2SSs from *V. cholerae*, *E. coli*, *D. dadantii*, *P. aeruginosa* and *L. pneumophila*.**

Genes are labelled as single letters based on the *Gsp* nomenclature except for *P. aeruginosa* Xcp (XcpP-XcpZ/GspC-GspM; XcpA/GspO). Genes encoding GspC proteins are colored in brown, the secretins are in blue, the inner membrane platform proteins in orange, the pseudopilins in green, pre-pilin peptidases in yellow and accessory components are grey. Pilotins are colored in magenta. Operons are separated by double lines to indicate genetic separation (Gu et al., 2017).

### 1.5.3 Architecture and biological functions of T2SS

All T4F promote various functions including cell adhesion, aggregation, motility and macromolecular transport. The T2SS specializes in secretion of folded proteins from the periplasm to the cell surface. The canonical T2SS architecture contains ~12 to 15 proteins that are more or less conserved in different species of bacteria, named GspC to GspO (Nivaskumar & Francetic, 2014a). These proteins form a macromolecular complex spanning the bacterial cell envelope that promotes secretion of folded protein substrates (exoproteins). T2SS takes up fully folded and even oligomeric substrates from the periplasm, transports them across the outer membrane and releases them into extracellular milieu. Some secreted proteins remain bound to the outer membrane surface (Rondelet & Condemine, 2013). Secretion mechanism involves assembly of a filamentous complex called the pseudopilus. Current models suggest that the pseudopilus promotes translocation of exoproteins by cycles of assembly –disassembly, acting as a piston to push them through the secretin pore (Campos et al., 2013). Other models propose a continuous treadmill-like assembly that lifts the substrates to the level of the OM pore where they are released (Archimedes' screw model) (Nivaskumar & Francetic, 2014b).

The T2SS-secreted exoproteins play important roles in bacterial adhesion, biofilm formation, nutrient acquisition, colonization, and invasion (Cianciotto & White, 2017). For many human pathogens including *Acinetobacter baumannii*, *Vibrio cholerae*, *Klebsiella pneumoniae* and *Pseudomonas aeruginosa*, substrates of their T2SS are key virulence factors (Nivaskumar & Francetic, 2014b). The plant pathogens like *Dickeya dadantii* and *Xanthomonas campestris* also rely on T2SS to secrete enzymes that cause plant disease. In general, T2SS allow bacteria to degrade large and insoluble biopolymers for use as nutrients. These include chitin, cellulose and other polysaccharides, lipids and proteins. The T2SS plays a very important environmental in recycling this organic matter and making it bioavailable.

#### 1.5.4 The *Klebsiella* T2SS: function and organization

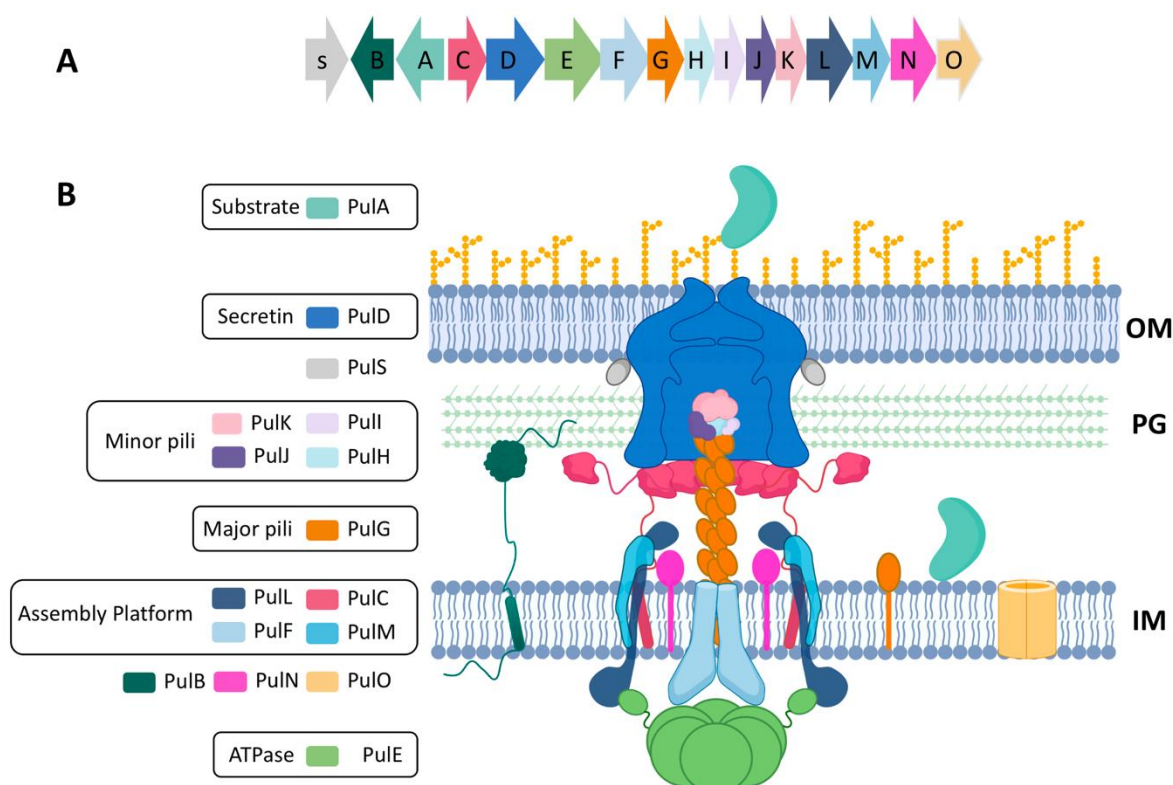


**Figure 14. Schematic representation of the *Klebsiella* Type II secretion system.**

A: The gene organization. B: Architecture of the *Klebsiella* T2SS. Locations of *Klebsiella* T2SS components in the bacterial envelope are illustrated. IM, inner membrane, OM, outer membrane, PG, peptidoglycan.

The study described in this PhD thesis is focused on the working model the *Klebsiella oxytoca* T2SS, the first discovered system of this family. This T2SS is dedicated to the secretion

of pullulanase (D'Enfert & Pugsley, 1987). *Klebsiella* T2SS is a complex of 15 different oligomeric proteins that span the bacterial envelope (Thomassin et al., 2017). It is encoded by 15 genes organized in a large cluster composed of two divergent operons: *pulAB*, *pulCDEFGHIJKLMNO*, and a single transcript *pulS* (Fig. 14A). The two operons are controlled by divergent promoters activated by MalT and by the general catabolic activator protein (CAP) bound to cAMP, while the *pulS* expression is constitutive. The fact that the substrate PulA is encoded by the *pulAB* in the T2SS gene cluster facilitated the reconstitution of the system in *E. coli* (D'Enfert & Pugsley, 1987) and suggested that this is the only substrate of the system. However, it is still possible that other proteins encoded elsewhere in the genome are secreted by the Pul system.

As illustrated in Fig. 14B, the Pul machinery is composed of 4 subcomplexes. The power source of this machinery – the ATPase PulE forms a hexamer that provides energy by hydrolysis of ATP (O. Possot & Pugsley, 1994). PulE is located in cytoplasm at the base of the system. In the inner membrane, PulC (Korotkov et al., 2006), PulF, PulL and PulM comprise an assembly platform (AP) complex (Py et al., 2001), together with the PulE, promoting pseudopilus polymerization in the periplasm. AP might also recognize and recruit the substrate pullulanase (PulA). The AP complex is linked *via* PulC to the outer membrane secretin, multimer of PulD (Guilvout et al., 2006) forming the trans-envelope channel. Each subunit is bound to its lipoprotein pilotin PulS that participates in PulD targeting to the OM and possibly facilitates its insertion (Hardie et al., 1996). The inner membrane components PulC, PulE, PulL, PulM and PulN have a relative stoichiometric ratio of 2:1:1:1:1 (Chernyatina & Low, 2019). The AP assembles the pseudopilus, composed the minor pseudopilins PulH, PulI, PulJ and PulK protein complex at the tip, and a major pseudopilin PulG forming a long homopolymer (Durand et al., 2003, 2005; Sauvonnnet et al., 2000; Vignon et al., 2003). As mentioned above, the pseudopilus is thought to drive secretion, possibly through a direct interaction with the substrate (Nivaskumar & Francetic, 2014a; Thomassin et al., 2017). PulO (Dupuy et al., 1992; Francetić et al., 1998) is a type III prepilin peptidase located in the IM. Its role is to remove the N-terminal signal peptide from the pseudopilin precursors, a cleavage which is essential for fiber assembly (Dupuy et al., 1992; Pugsley & Dupuy, 1992). PulO also adds a methyl group to the N-terminal amine of mature pseudopilins (Strom et al., 1993), which further facilitates their extraction from the membrane (Santos-Moreno et al., 2017). The results of numerous genetic, biochemical and structural analyses in bacterial different T2SS led to a working model as illustrated in Fig. 15.

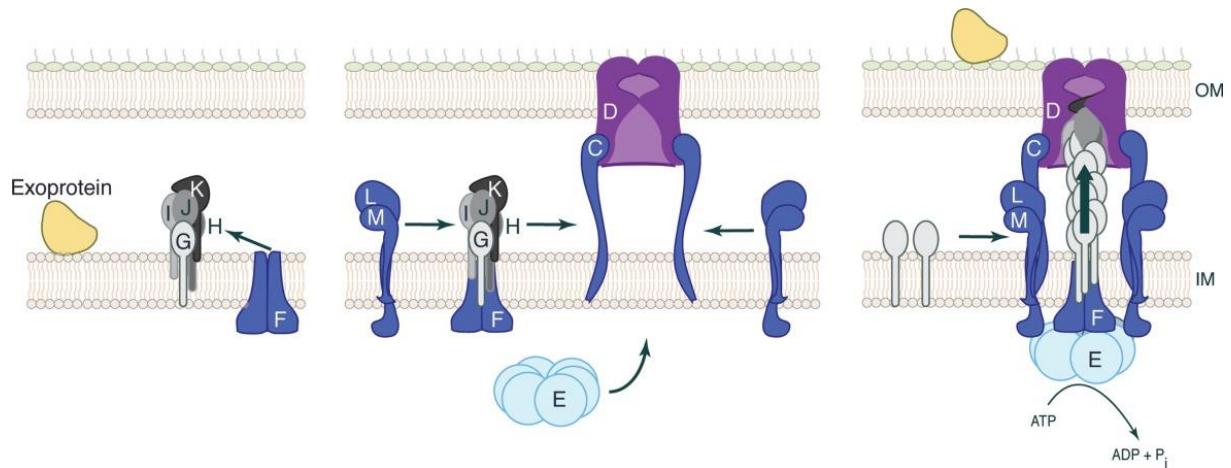


**Figure 15. Schematic working model of the T2SS.**

Shown is a cartoon of the working model for dynamic T2S biogenesis. The four sub-complexes that make up the T2S machinery are highlighted in different colors; the OM secretin sub-complex is purple, the assembly platform (AP) is blue, the ATPase (E) is light blue and the pseudopilin subunits are in different shades of grey (G, H, I, J and K). In this model, a complex formed by the pseudopilins first interacts with AP protein GspF. This sub-complex would be recruited to GspL/M via interactions with the pseudopilin subunits G and H. Interactions between GspL/M and GspC and between GspC and GspD align the IM-embedded components with the pore formed by the secretin (GspD). At this time, the cytoplasmic ATPase would dock to the T2SS and energize pseudopilus assembly. To allow for pseudopilus elongation, new pseudopilin subunits are recruited to the machine through GspM. New pseudopilin subunits and, presumably, exoproteins (wheat), then enter the cage-like compartment formed by the AP and secretin and exoproteins are transported across the OM (Thomassin et al., 2017).

### 1.5.5 The outer membrane complex

Bacterial T2SS, T3SS and T4P share similar outer membrane channels called secretins (Genin & Boucher, 1994; Martin et al., 1993). Secretins from various transport systems and bacteria differ significantly, but they share a common general domain organization and show significant similarity in the C-terminal region. This conserved region, also known as the secretin domain, contributes to the formation of the  $\beta$ -barrel secretin channel in the OM. The N-terminal region of secretins from various transport systems is more variable and may include multiple domains ranging in T2SS from N0 to N3 (Reichow et al., 2010) (Fig. 16). The N-terminal domains are the secretin channel entry point interact with the IM components of T2SS as well as the secreted substrates (Nouwens et al., 1999). The C-terminus of some secretins is terminated by a short S-domain that interacts with a cognate pilotin protein, depending on the transport system and bacterium (Korotkov et al., 2011; Pineau et al., 2014).

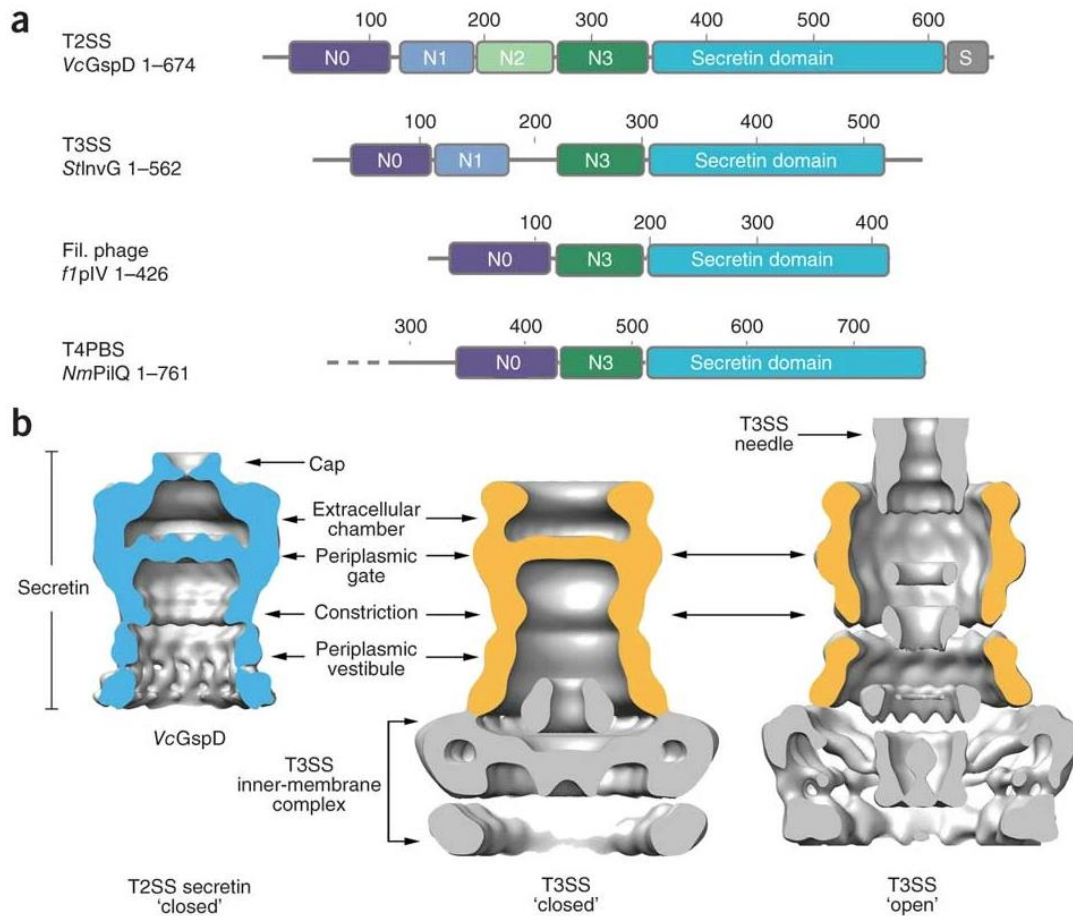


**Figure 16. The secretin architecture is conserved in different secretion systems.**

(a) Domain architecture of secretins from the T2SS, the T3SS, the filamentous phage assembly system and the T4PBS. Members of the secretin superfamily contain a C-terminal secretin core homology domain (cyan). The T2SS secretins generally contain four periplasmic subdomains, termed N0–N3. The N0 subdomain (dark blue) is located at the N terminus and is followed by the three structurally homologous subdomains, N1–N3 (blue, green and dark green, respectively). A T2SS-specific domain, termed the S-domain (gray), is located at the very C terminus. Secretins from other systems share a similar architecture, composed of the secretin domain and at least two periplasmic subdomains that are structurally equivalent to N0 and N3 of VcGspD (EMBD 1763). (b) Structural comparison of the VcGspD density (blue, left) to single-particle reconstructions of the T3SS in its closed state (middle) (EMDB 1224) and to the fully assembled T3SS needle complex in its open state (right) (EMDB 1617). The outer-membrane T3SS secretin sits on top of a large inner-membrane complex (gold, middle and right). VcGspD seems to be in its closed state (compare left panel with middle panel) (Reichow et al., 2010).

The T2SS secretins termed as GspD, and that in *Klebsiella* T2SS is called PulD. PulD forms a symmetric 15-subunit secretin channel that spans the outer membrane and protrudes into the periplasm (Chernyatina & Low, 2019). The PulD monomer is made up of four N-

terminal periplasmic domains known as N0, N1, N2, and N3, as well as a secretin domain and a C-terminal S-domain (Fig. 17). The secretin core is highly ordered, and form a double beta-barrel. Each subunit contributes four beta-strands to the outer barrel, and four shorter strands that make up the periplasmic gate. The N-terminal regions show high flexibility; most cryo-electron microscopy (EM) secretin structures exhibited flexible N2 and N1 domains and totally invisible N0 domain (Hay et al., 2018; Howard et al., 2019; Yin et al., 2018). As a result, the location of N0 in relation to the rest of the secretin was unknown. Purification of a 2.4 MDa T2SS complex from *K. pneumoniae* that included PulD with its pilotin PulS, as well as PulE, PulL, PulM, PulN, and PulC (PulCDELMNS) yielded a complete secretin structure with ordered N0 domains (Chernyatina & Low, 2019) (Fig. 18).

Typically, the assembly and insertion of this multimeric complex into the OM is independent on BAM (Huysmans, 2016). Pilotins, termed PulS in *Klebsiella* T2SS, decorating the outside of the secretin core near the PulD S-domain, are secretin-specific lipoprotein chaperones that help PulD subunits to localize, assemble, and insert correctly into the outer membrane (Hardie et al., 1996; Nickerson et al., 2011). Pilotins are sorted to the outer membrane by the Lol machinery after their signal sequences are cleaved. A conserved cysteine at the PulS N-terminus is lipidated and acts as an anchor within the outer membrane's inner leaflet (Zückert, 2014) (Fig. 18).

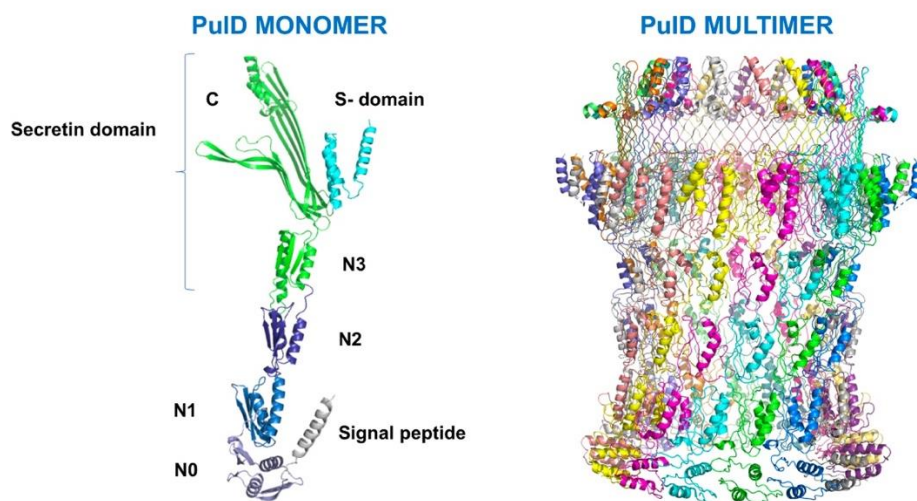


**Figure 17. The CryoEM structure of PulD monomer and multimer.**

*PulD* secretin monomer AF2 model including the N0 to N3 domains in blue, the C domain in green and the S domain in cyan. The signal peptide is removed prior to assembly into multimer (right). The multimer structure PDB: 6HCG.

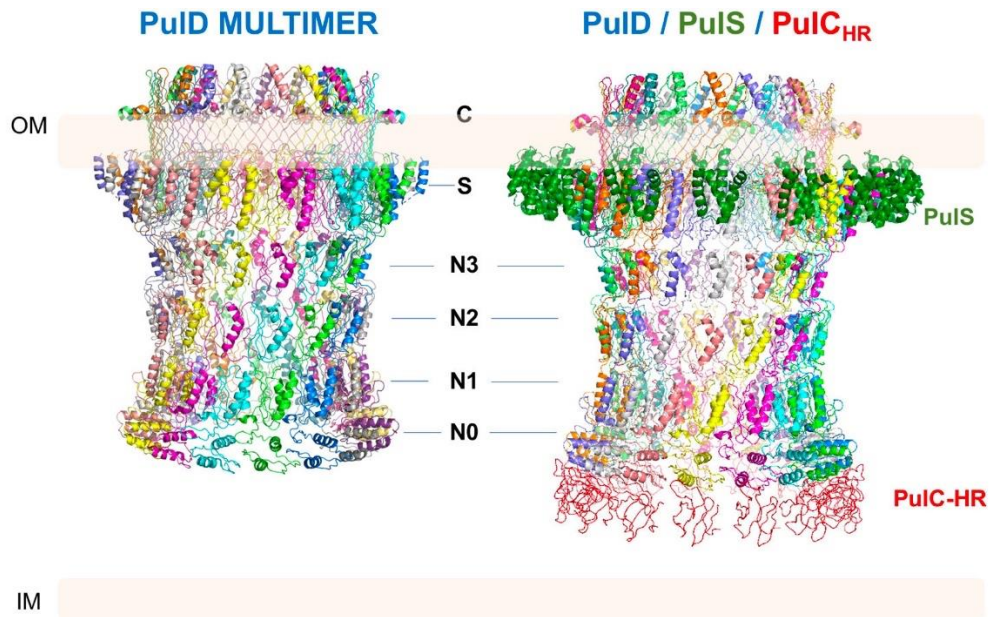


**Figure 18. The CryoEM structure of PulD Multimer and PulD with PulS and PulC<sub>HR</sub>.**

Individual subunits are shown in different colors. IM, OM, inner and outer membrane. PulS is shown in green and PulC HR domain in red. PDB: 6HCG.

#### 1.5.6 The inner membrane assembly platform (AP) and the ATPase

The function of T2SS relies on the assembly of periplasmic pseudopili homologous to bacterial type IV pili. The pseudopilus is assembled from its base in the inner membrane by the assembly platform (AP) complex. AP is composed of GspC, GspF, GspL, GspM, and, in some species, GspN associated with the cytoplasmic ATPase GspE, (Fig. 19).

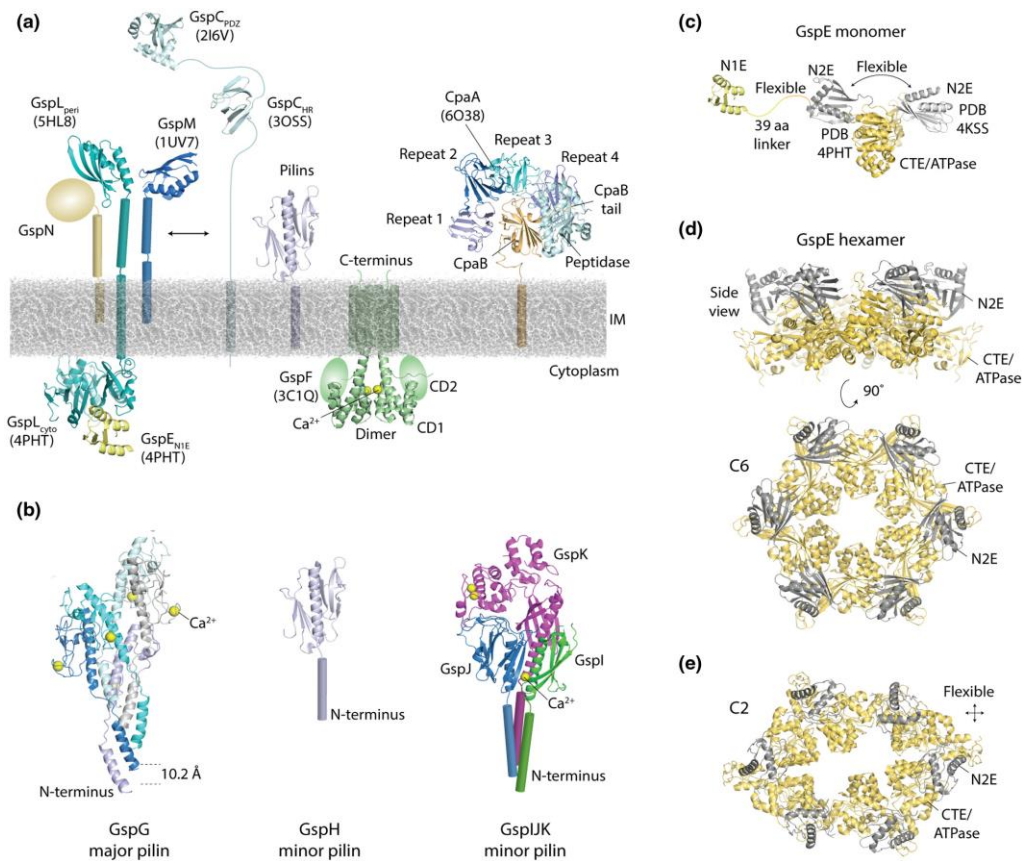


**Figure 19. Gallery of T2SS inner membrane components.**

(a) Gallery of AP and inner membrane components including the substrate CpaA with associated chaperone CpaB. PDB codes are shown in parentheses. GspL, GspM, and GspN form an inner membrane complex. GspC also associates with GspL and GspM. Pseudopilins, as shown in b, are positioned within the inner membrane before pseudopilus assembly. The platform protein GspF may constitute a dimer. CD1 (amino acids 56-171) and CD2 correlate to cytoplasmic domains 1 and 2. The structure of CD2 and GspF transmembrane helices has not been determined. The *Acinetobacter baumannii* protease CpaA complexed with its membrane chaperone CpaB represents a novel chaperone-substrate interaction. (b) Gallery of minor and major pseudopilins that collectively form the pseudopilus. GspG (PDB code 5WDA) forms the bulk of the pseudopilus. GspH (PDB code 2QV8) and GspIJK (PDB code 3CI0) constitute the pseudopilus tip. (c) GspE is a highly flexible and dynamic ATPase. The monomer comprises N1E, N2E and C-terminal CTE domains all interconnected by flexible linkers. Superposition of the CTE domains from *V. vulnificus* and *V. cholerae* (PDB code 4PHT and 4KSS) highlights the flexibility of the N2E domain. (d, e) GspE forms a hexamer when crystallized as a Hcp1 fusion protein. GspE undergoes large scale conformational changes switching between C6 symmetry in d (PDB code 4KSS) and C2 symmetry in e (PDB code 4KSR). In both d and e, the Hcp1 fusion has been removed for clarity (Naskar et al., 2021).

GspC comprises a short cytoplasmic tail at the N-terminus, a transmembrane helix followed by the periplasmic HR (homology region) and a PDZ domain at the C-terminus. The GspC component of the *Dickeya dadantii* T2SS OutC forms dimers via its transmembrane helix (Login & Shevchik, 2006). GspC interacts with the N0 domain of GspD at the secretin base (Fig. 20) and is thought to form a periplasmic cage-like structure that connects the AP with the OM channel. The PDZ crystal structure shows a typical bacteria-specific PDZ fold consisting

of six  $\beta$ -strands bordered by two  $\alpha$  helices, forming a large groove (Login & Shevchik, 2006). The HR crystal structure revealed two stacks of three anti-parallel  $\beta$ -strands (Korotkov et al., 2011), (Fig. 20).

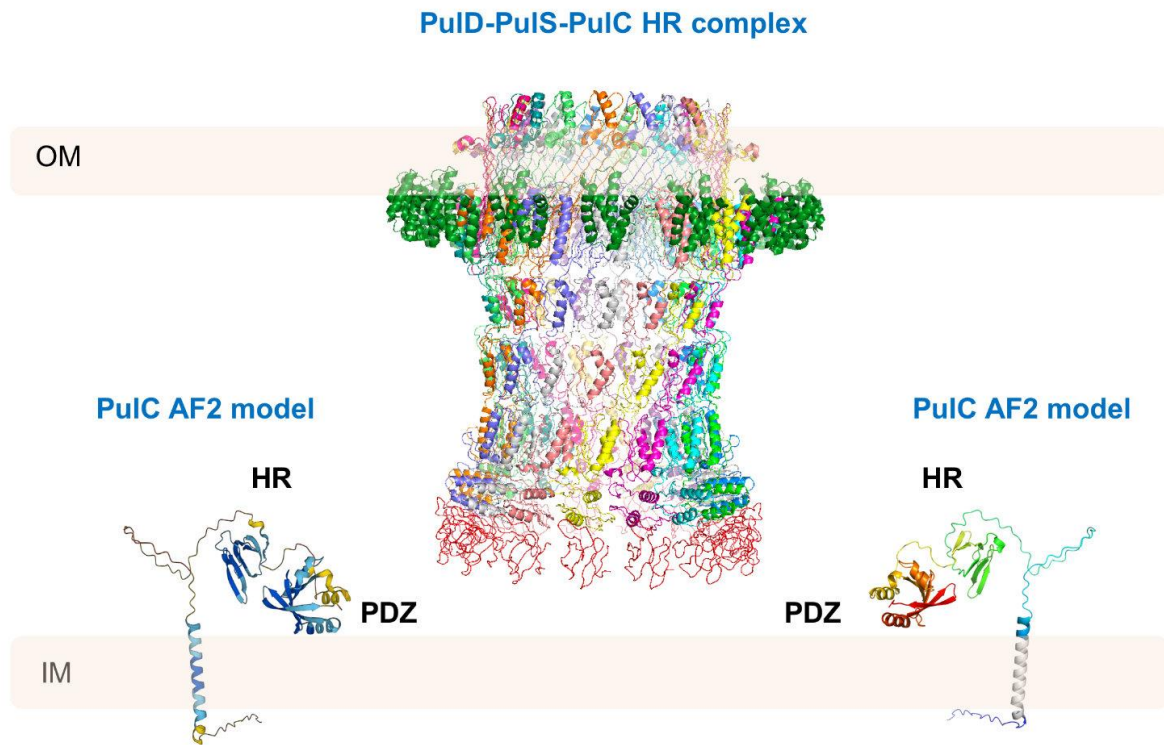


**Figure 20. The CryoEM structure of PulD-S-C complex and AlphaFold 2 models of PulC.**

The PulC model on the left is colored according to the prediction confidence from dark blue (high confidence) through light blue (medium) to yellow (low confidence). The PulC AF2 model on the right is rainbow-colored from the N-terminus (blue) to the C-terminus (red).

The AP components GspL and GspM have a single transmembrane helix followed by a long  $\alpha$ -helical region and the globular C-terminal domains (CTDs) on the periplasmic side. The crystal structures of *Vibrio parahaemolyticus* CTDs of GspL and GspM showed a ferredoxin-like fold (Abendroth, Rice, et al., 2004). It is interesting to note that the T4PS proteins PilN and PilO also have the same fold, which suggests a shared evolutionary history and functional similarity. GspL<sub>CTD</sub> and GspM<sub>CTD</sub> produced homodimers with different interfaces in the crystals. Cross-linking investigations in *Dickeya dadantii* showed that GspL<sub>CTD</sub> and GspM<sub>CTD</sub> form homodimers and heterodimers, suggesting that these interfaces may exist and alternate *in vivo* during the dynamic secretion process (Lallemant et al., 2013). GspL also has a globular N-terminal cytoplasmic domain (GspL<sub>cyto</sub>). The crystal structure of the *V. cholerae* GspL<sub>cyto</sub>

revealed three  $\beta$ -sheet-rich subdomains, I, II, and III (Abendroth, Bagdasarian, et al., 2004) (Fig. 21). This domain interacts with GspE N0 domain and thereby anchors the ATPase to the AP(C. Lu et al., 2014).

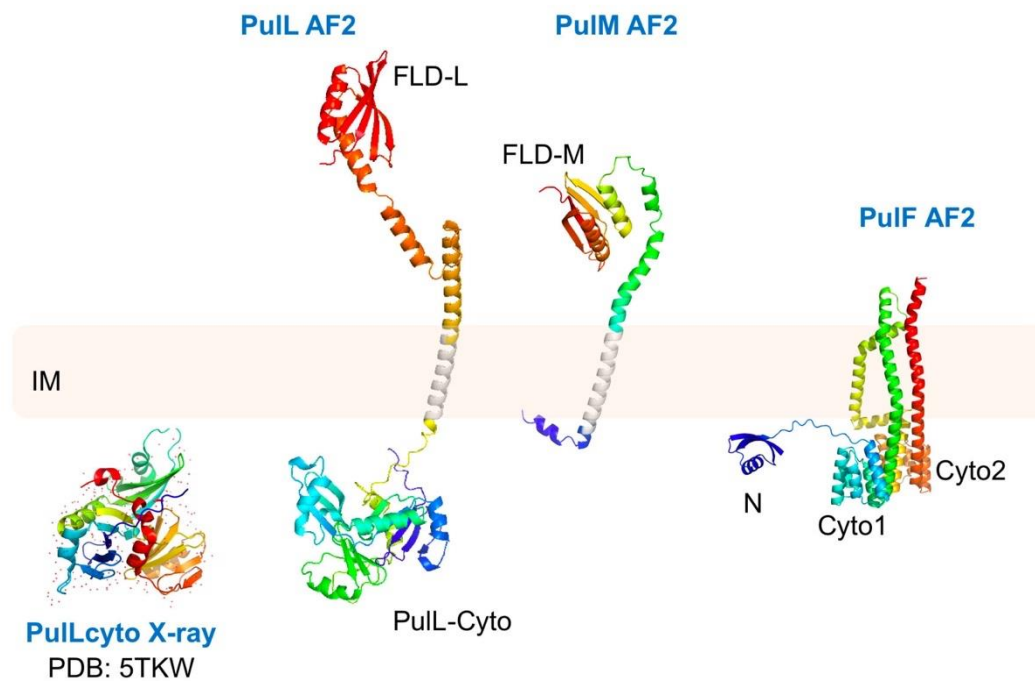


**Figure 21. The structure of PulLcyto domain and the AlphaFold 2 models of full-length PulL, PulM and PulF.**

Proteins are rainbow-colored from N-terminus (dark blue) to C-terminus (red). IM, inner membrane.

GspF is a polytopic membrane protein with three cytoplasmic domains and three transmembrane segments. The alkaline phosphatase fusion experiments supported the N-in – C-out topology membrane architecture (Arts et al., 2007) (Fig. 21). The first cytoplasmic domain of the GspF (EpsF) protein from *V. cholerae* was crystallized, and it revealed a 6-helix bundle fold that dimerized and coordinated calcium through a conserved interface (Abendroth et al., 2009). Since the two cytoplasmic domains share considerable sequence similarity, an analogous fold of Cyto2 was also predicted. Purified GspF (XcpS), according to a recent study in *P. aeruginosa*, adopts a dimeric structure (Putte et al., 2018). Ultimately, because the Cyto1 domains from the T4PS ortholog PilC has been isolated in several systems as a dimer and a tetramer (Bischof et al., 2016), it is still unclear what is the oligomerization state of GspF inside the T2SS.

GspE is a cytoplasmic ATPase that drives substrate extrusion and pseudopilus assembly (Sandkvist et al., 1995). Three domains – the N0, N1, and C – are found in GspE based on

sequence and structural studies (Fig. 22). The flexibility of the linker between the N2 and C domains was underlined by superimposing the GspE structures from *Vibrio cholerae* and *Vibrio vulnificus*. The C domain can be further divided into the C1, C2 and CM subdomains (Lu et al., 2014). Walker A and B motifs that bind nucleotides, as well as additional aspartate and histidine motifs, are all found in the C1 subdomain. It is important to note that the CM contains a unique tetracysteine motif that coordinates zinc (Camberg & Sandkvist, 2005), which is conserved in T4PS assembly ATPases (McCallum et al., 2019). Cycles of ATP binding and hydrolysis produce conformational changes in the ATPase hexamer, which might be transmitted to GspF to drive pseudopilus assembly, possibly by inducing a rotational motion (Nivaskumar et al., 2014).

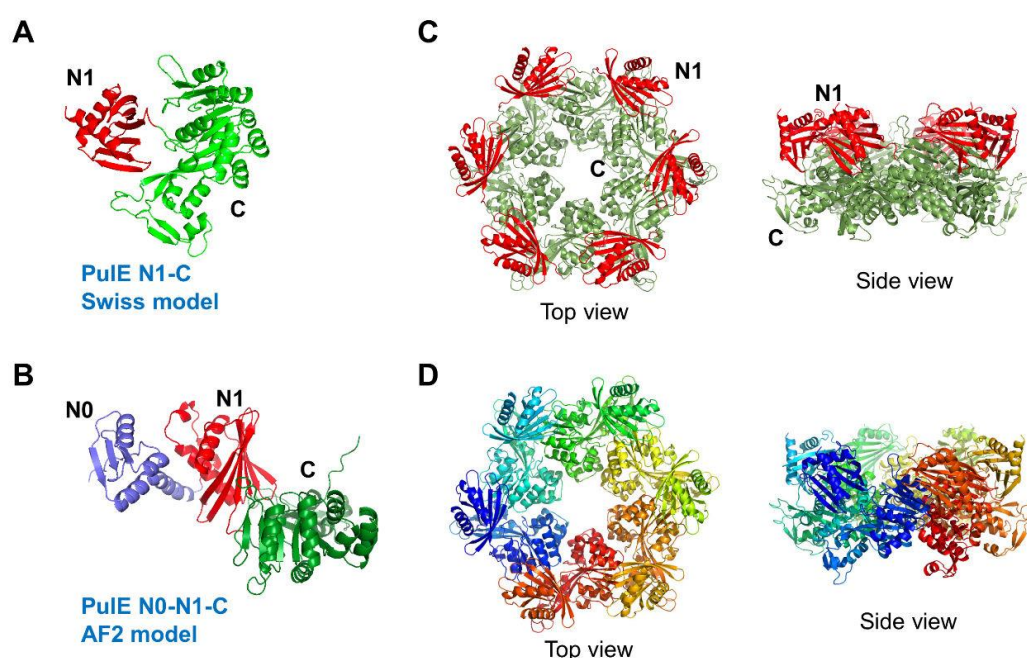


**Figure 22. The structures of PulE.**

A: PulE N1-C structure by Swiss model; B: PulE multimer model, top and side view predicted by Swiss model, with N1 domains in red and C domains in green; C: PulE N0-N1-C AlphaFold 2 model; D: PulE AlphaFold 2 model with monomers shown in different colors.

### 1.5.7 The pseudopilus

The T2SS pseudopili resemble Type IV pili in structural and biosynthesis aspects. The *Klebsiella* T2SS is a filamentous polymer, which consists of major pseudopilin PulG and minor pseudopilins PulH, I, J, K (Fig. 23) and is assembled by the inner membrane AP.

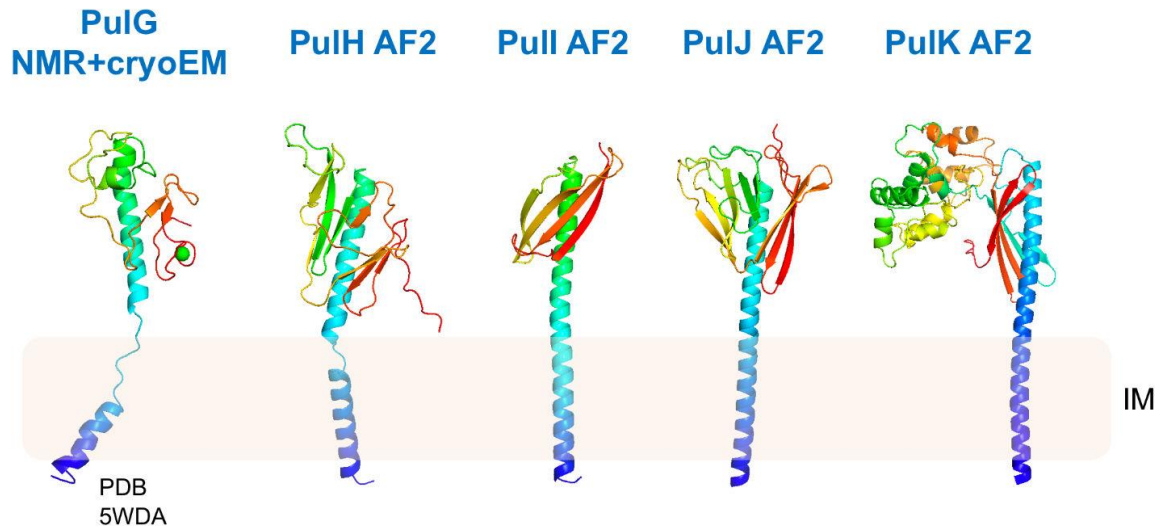


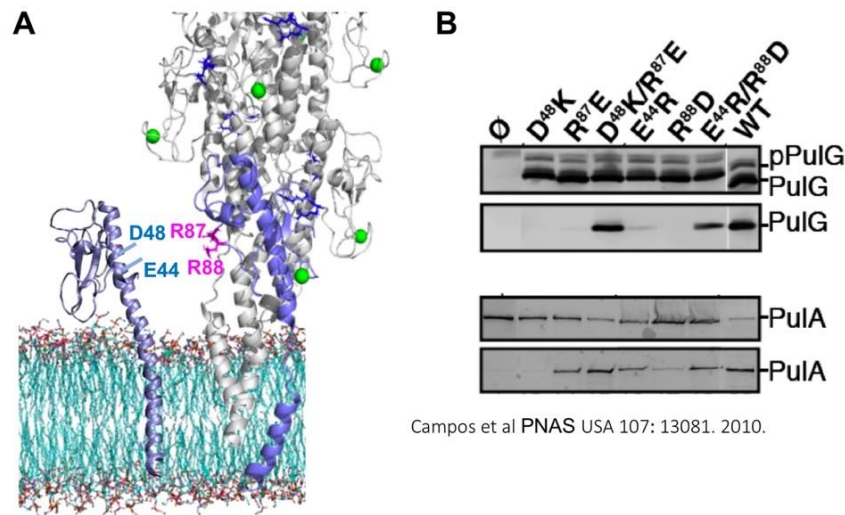
**Figure 23. The structures and AF2 models of the *Klebsiella pseudopilus* subunits.**

The major pseudopilin PulG structure obtained by NMR and cryo-EM and the minor pseudopilin AF2 models as indicated. The cartoon models are rainbow colored from N-terminus (blue) to C-terminus (red). IM, inner membrane.

The mature sequences of GspG, GspH, GspI and GspJ contain a conserved glutamic acid residue at position 5 (Glu5). In the case of GspG and GspH, Glu5 is essential for assembly and for recruitment of these pseudopilins to the AP by direct interaction with GspM (Nivaskumar et al., 2016; Nivaskumar & Francetic, 2014a; Santos-Moreno et al., 2017).

#### 1.5.7.1 Major pseudopilin

The major pseudopilin GspG forms a helical filament that makes up the main stem of the pseudopilus (Campos et al., 2010). Modeling combined with mutational and functional analysis revealed conserved GspG residues that are crucial for pseudopilus assembly and protein secretion (Campos et al., 2010; Nivaskumar et al., 2014). They form two conserved salt bridges between neighboring protomers in the IM (Fig. 24). An N-terminal  $\alpha$ -helix, a short loop, and a four-stranded C-terminal antiparallel  $\beta$ -sheet make up the distinctive fold shared by all pseudopilins. GspG has a conserved calcium-binding site essential for its folding, stability and function (Korotkov et al., 2009; López-Castilla et al., 2017).



Campos et al PNAS USA 107: 13081. 2010.

**Figure 24. Major pseudopilin assembly.**

**A.** PulG monomer (left) docks via the conserved salt bridges to the growing pseudopilus (right). **B.** Single charge inversions of conserved salt bridge residues abolish PulA secretion and PulG pilus assembly. Combined charge inversions restore salt bridges and their function.

Using NMR and cryoEM, the core structure of the PulG pilus has been determined (López-Castilla et al., 2017) (Fig. 25). The structure of the periplasmic PulG domain was determined by NMR. NMR and mutational studies showed an essential role of calcium for PulG folding and stability. The cryoEM structure confirmed essential contacts between subunits characterized previously and revealed unfolding of the N-terminal helix in the pili. This extended region provides flexibility to the pseudopilus (López-Castilla et al., 2017).

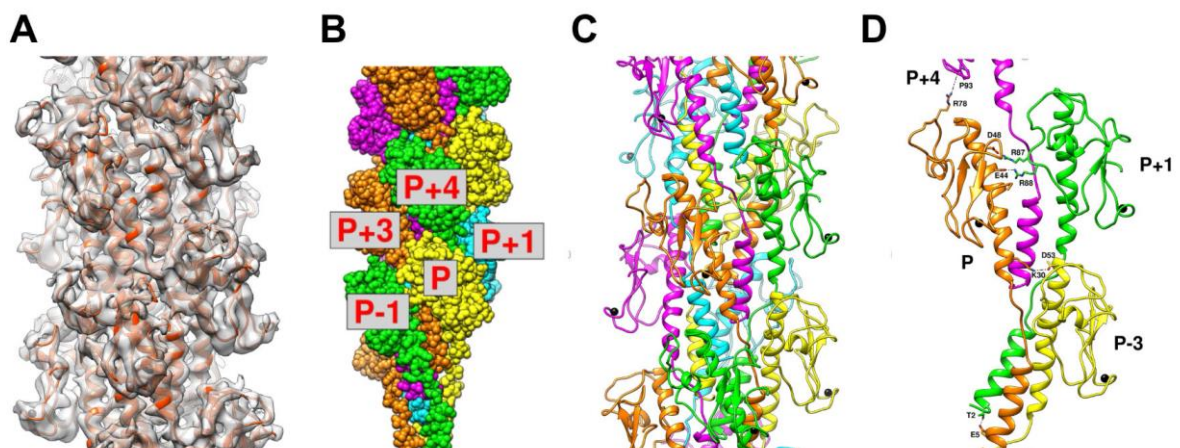


**Figure 25. The PulG pseudopilus structure.**

**A,** the cryo-EM envelope of the PulG pilus with subunits docked in red. **B,** helical organization of PulG protomers; **C,** Cartoon model of the pseudopilus with Ca atoms in black; **D,** contact at different interfaces with the salt bridges highlighted.

#### 1.5.7.2 Minor pseudopilins

The X-ray structure of the GspIJK trimeric complex from *Escherichia coli* (Fig. 26) suggested that this complex is localized at the tip of the filament (Korotkov & Hol, 2008). The periplasmic domains of GspK showed a pilin domain, plus a distinctive extra helical domain of GspK, which also binds calcium ions, is thought to cap the pseudopilus tip and prevent assembly of subunits above (Korotkov & Hol, 2008). GspK also lacks the conserved residue E5, suggesting its position at the very tip.

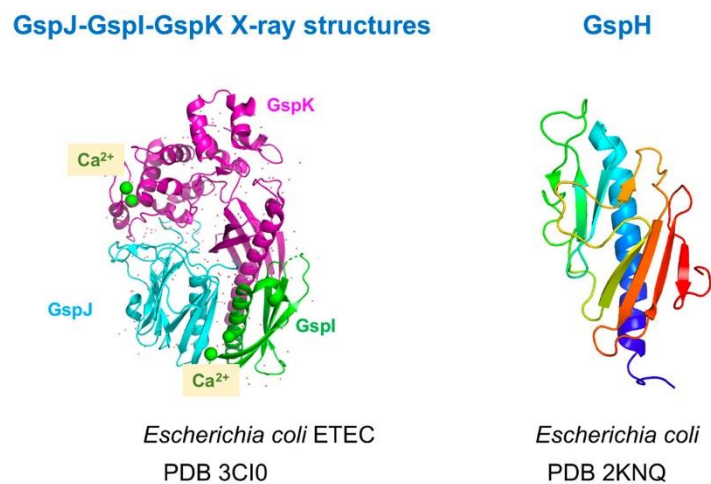


**Figure 26. Minor pseudopilin periplasmic domains from *E. coli*.**

Left: the X-ray structure of the GspJ-GspI-GspK trimer with calcium ions indicated; right: GspH NMR structure.

Minor pseudopilins are essential for protein secretion, which suggests that they may bind the specific substrates (Sauvonnnet et al. 2000). However, they are also required for pseudopilus assembly. Deleting individual genes for PulJ, PulI and PulK reduced the number of assembled pili and extended their length, while removing all minor pseudopilins nearly abolished assembly (Cisneros et al., 2012). Their interactions in the membrane led to the formation of a staggered complex as shown by Cys crosslinking and molecular dynamics simulations. These studies led to propose that minor GspJ-GspI-GspK complex initiates assembly (Cisneros et al., 2012) (Fig. 27).

Using a biochemical approach to study minor pseudopilins Douzi et al. showed that GspI plays a central role in assembly by interacting with GspJ and GspK with different interfaces (Douzi et al., 2009). GspH binds to the complex via GspJ. Recent studies combined crosslinking, mass spectrometry and mutational studies, allowing them to propose a model of the quaternary pseudopilin complex and dock it to the pseudopilus structure (Fig. 27) (Escobar et al., 2021).

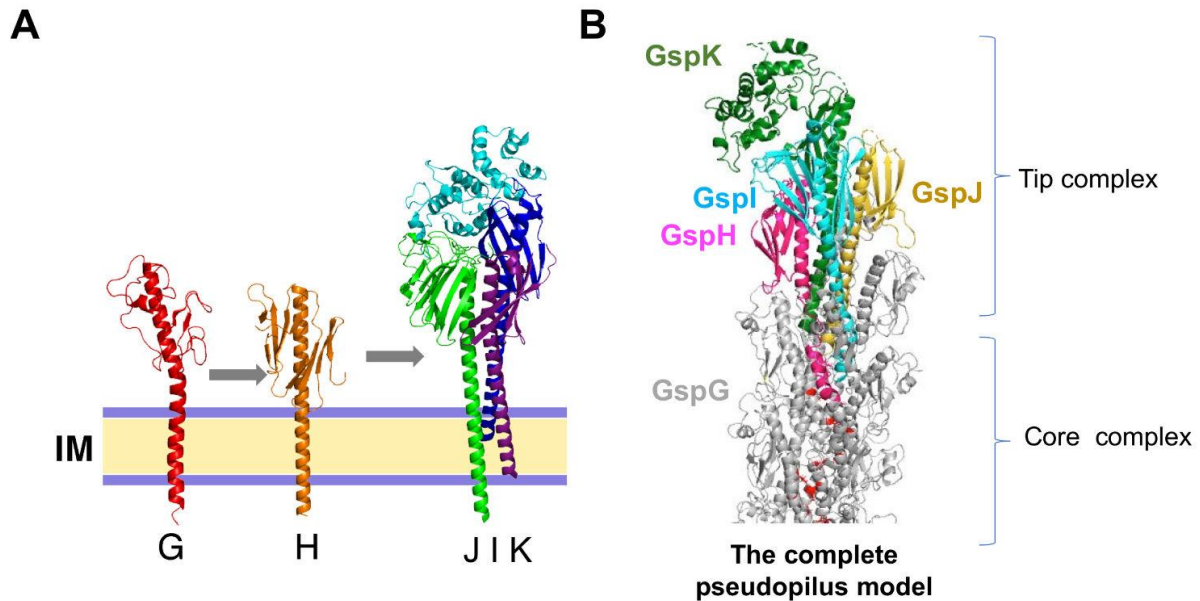


**Figure 27. Pseudopilus assembly initiation and pseudopilus structural model.**

Left: The tip complex of GspJ-GspI and GspK is self-assembled in the membrane upon cleavage by the prepilin peptidase. One GspH and multiple GspG subunits are added via GspM- and GspE dependent mechanism. Right: The structural model of the complete pseudopilus proposed by (Escobar et al., 2021).

#### 1.5.8 PulN

GspN is found only in certain bacteria, including *K. oxytoca*, *V. cholerae*, *P. carotovorum*, *A. hydrophila*, and *A. baumannii*, but not in *D. dadantii*. GspN is a bitopic IM protein with an N-terminal transmembrane helix and a periplasmic globular domain, predicted by the AF2 served to form a  $\beta$ -sheet folded into a funnel-like structure (Fig. 28). The function of GspN is still unknown (Korotkov et al., 2012). Although PulN is not required for pullulanase secretion in *K. oxytoca* (Possot et al., 2000), it has been copurified with the components of the IM assembly platform PulE, PulL, and PulM (Chernyatina & Low, 2019) from *K. pneumoniae*.

#### 1.5.9 PulB

In *Aeromonas* and *Vibrio* a functional *gspAB* operon encodes an additional ATPase GspA and GspB proteins involved in secretin assembly in the outer membrane (Ast et al., 2002; Strozen et al., 2011). The T2SS in *Dickeya*, *Pectobacterium* and *Klebsiella* only have a GspB ortholog. GspB proteins have the same general organization, a short cytoplasmic region, a

single TMS, and a large periplasmic region (Fig. 28). Previous research in *D. dadantii* and *A. hydrophila* revealed that GspB interacts with the secretin N-terminal periplasmic region (Condemine & Shevchik, 2000; Vanderlinde et al., 2014, 2017). Recent studies in *D. dadanti* revealed, by using *in vivo* crosslinking with unnatural amino acids, how OutB binds to the secretin OutD, and also interacts with the peptidoglycan, to promote correct OutD localization and assembly (Zhang et al., 2022). In this regard, the role of OutB is complementary to that of OutS.

#### 1.5.10 PulO

PulO (Dupuy et al., 1992; Francetić et al., 1998) is a type III prepilin peptidase located in the IM (Fig. 28). Its role is to remove the N-terminal signal peptide from the pseudopilin precursors to allow for their membrane extraction and assembly (Dupuy et al., 1992; Pugsley & Dupuy, 1992). PulO also adds a methyl group to the N-terminal amine of mature pseudopilins (Strom et al., 1993), which further facilitates their extraction from the membrane (Santos-Moreno et al., 2017).

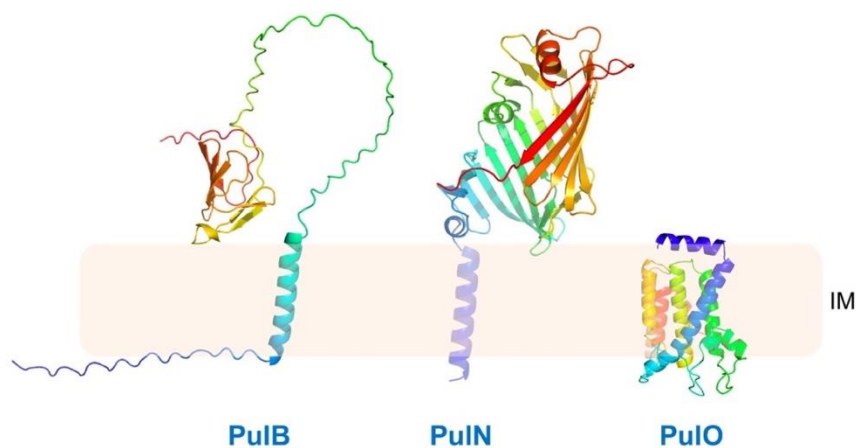


**Figure 28. AlphaFold 2 structural models of PulB, PulN and PulO.**

The cartoon models are depicted in the inner membrane (IM) and colored from the N-terminus (dark blue) to the C-terminus (red).

## 1.6 Substrates and secretion signals of T2SS

### 1.6.1 Substrates of T2SS

The substrates of T2SS are typically large proteins and complexes that reach the periplasm through the Sec or the Tat export systems (Voulhoux et al., 2001). They are recognized and secreted in their native, fully folded state, and in some cases, such as the cholera toxin in *Vibrio*, as oligomeric complexes. The substrates secreted by T2SS in different bacterial species are vary in sizes, shapes, quaternary structures and functions, and the number of substrates secreted by the T2SS in certain bacteria vary from one, like in the *K. oxytoca*, to more than ten in *P. aeruginosa* (Bleves et al., 2010), *V. cholerae* (Sikora et al., 2011) or *L. pneumophila* (Cianciotto, 2009). Nevertheless, there are no sequence or structural similarities found among the T2SS substrates. Their secretion is strictly specific and probably relies on one or more structural motifs. Despite a large amount of structural studies on different bacterial models, the nature of these motifs and their interaction with T2SS components are poorly understood. Some studies suggest that  $\beta$ -sheets are associated with T2SS substrates and that the mechanism of their recognition might be through  $\beta$ -complementation (Korotkov et al., 2012). Below are presented some recent studies that address the molecular nature of the secretion signals.

### 1.6.2 Structural and functional studies of type II secretion signals

Since the discovery of T2SS, the question of secretion signal has been raised and addressed in many studies. In bacteria that secrete several substrates the absence of a common sequence motif suggested that the signal is found in the 3-D structure. In addition, several studies showed that T2SS substrates are folded and have disulfide bonds prior to secretion. This implies that understanding the nature of the secretion signal requires structural information on secreted proteins.

To decipher the mechanisms of the secretion signals of T2SS substrate, a lot of different strategies and techniques – insertion, deletion, protein fusion, orthologous substrates fragments exchange or domain swapping were used in numerous studies. Below we provide a few examples to illustrate these lines of investigation. Studies performed on cellulase EGZ from the plant pathogen *Erwinia chrysanthemi* found that dithiothreitol (DTT) treatment of native EGZ blocks its secretion (Bortoli-German et al., 1994). This result showed that the T2SS substrates are folded and stabilized by disulfide bonds before being secreted. On the other hand, studies

on pullulanase showed that eliminating its disulfide bonds reduced its stability but did not affect its secretion (Sauvonnet & Pugsley, 1998).

Several studies addressed the secretion signal of exotoxin A secreted by the *P. aeruginosa* Xcp system. Lu and Lory used deletions and fusions with  $\beta$ -lactamase to identify 60-aa region containing the secretion signal of Exotoxin A (Lu & Lory, 1996). This region corresponds to a small subdomain of the toxin rich in beta strands. Other studies have identified N-terminal or C-terminal regions of the protein necessary for its secretion (Hamood et al., 1989; McVay & Hamood, 1995). Voulhoux et al. studied the same protein using deletion analysis also involved in binding to a receptor on eukaryotic cells (Voulhoux et al., 2000).

The *Pectobacterium carotovorum* and *Dickeya dadantii* (formerly named *Erwinia chrysanthemi* and *E. carotovora*) are two phytopathogenic proteobacteria that produce almost identical sets of homologous plant cell wall-degrading enzymes, a dozen pectinases and one cellulase, through related T2SSs referred to as Out (Pineau et al., 2014). Although they share 67% sequence identity, the orthologous pectate lyases Pel3 and PelI from these bacteria are only secreted by their cognate T2SS but not in the heterologous T2SS, demonstrating that protein recruitment is precisely controlled. The domain swapping strategy between these two orthologous secretion substrates created a series of fusion proteins containing fragments from each other, and their secretion was tested to determine an essential structural motif specific for Out T2SS (Pineau et al., 2014). In the following work, by combining X-ray crystallography and more precise swapping of loop-regions, these authors revealed that there is not one element with a definite secondary structure but several distant and structurally flexible loop regions that are essential for the secretion of Pel3 and that these loop regions act together as a composite secretion signal (Pineau et al., 2021).

### 1.6.3 Secretion signal studies of pullulanase in *Klebsiella* T2SS

The *K. oxytoca* T2SS, our experimental model, is dedicated to secretion of a 116-kDa surface-anchored lipoprotein pullulanase (PulA), a starch debranching enzyme allowing bacterial growth on branched maltodextrin polymers (Michaelis et al., 1985; Pugsley et al., 1986). Identification and cloning the *pulA* gene from *K. oxytoca* into *Escherichia coli* K-12 initiated a long but exciting process to a final discovery of the Type II secretion system. To identify the factors related to its secretion, d'Enfert and Pugsley used a gene fusion approach. A series of fusions between the gene of pullulanase (*pulA*) and the genes for cytoplasmic  $\beta$ -

galactosidase (*lacZ*) or periplasmic alkaline phosphatase (*phoA*) were created by transposon mutagenesis. These hybrid proteins were expressed in *E. coli* K-12 and their subcellular localization and secretion were determined. The N-terminal cysteine residue of the mature polypeptide is fatty acylated and the lipoprotein PulA uses the Sec pathway to cross the inner membrane (D'Enfert & Pugsley, 1987). Sauvonnnet et al. constructed several PulA- $\beta$ -lactamase (BlaM) hybrid proteins and revealed that PulA lacking the last 256 amino acids was secreted (Sauvonnnet et al., 1995). Based on these findings, they devised a plate screen combined with the deletion strategy and identified two regions, named A and B, that together are essential for PulA secretion (Sauvonnnet & Pugsley, 1996) (Fig. 29). Isolated regions A and B, together with an N-terminal signal sequence, were also sufficient to promote secretion of  $\beta$ -lactamase and a domain from another protein, the *D. dadantii* CelZ.

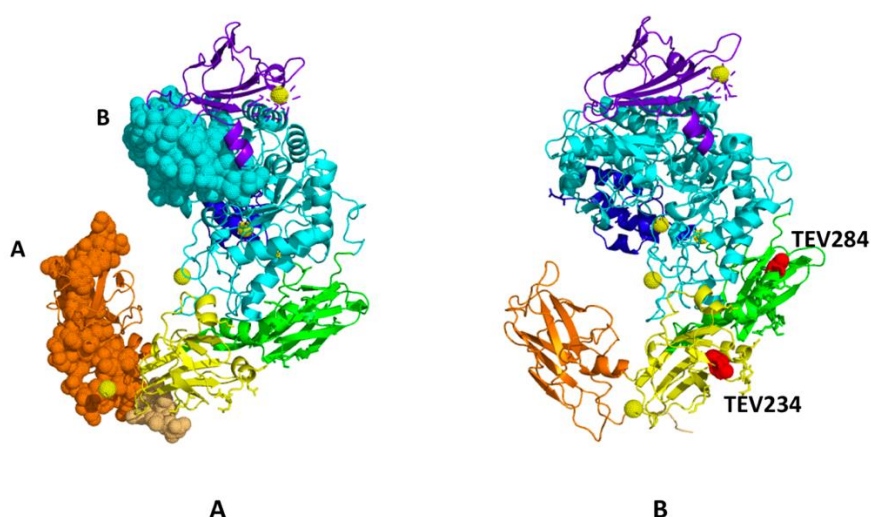


**Figure 29. PulA X-ray structure and secretion signal studies.**

**A:** The X-Ray structure of PulA illustrating the positions of regions A, B and C containing putative secretion signals. Residues 11–78 (region A) are shown in orange with the surface representation and residues 735–814 (region B) in cyan. Deletion of A and B together abolished secretion of lipoPulA. **B:** The TEV peptide insertions shown in red at positions 234 (domain N2 (in yellow)) and 284 (domain N3 (in green)) abolished secretion of PulA<sub>NA</sub>.

Francetić and Pugsley constructed a nonacylated variant of PulA (PulA<sub>NA</sub>), by replacing its native, lipoprotein signal peptide with the signal peptide of pectate lyase PelB from *E. chrysanthemi* (Francetić & Pugsley, 2005). This construct facilitated the study of PulA secretion by its immunodetection in cell and supernatant fractions and showed that PulA<sub>NA</sub> was efficiently secreted into the medium. In this variant linker insertion, fragment deletion and protein fusion approaches were used to identify secretion signal. The results showed that

deletion of region A destabilized PulA and deletion of region B had a small impact on PelBsp-PulA secretion. They defined a third region C between amino acids 234 and 324 in PelBsp-PulA, where four independent linker insertions abolished secretion. Thus, region C of PulA may hold a portion of the PulA secretion signal or be crucial to its proper presentation, without significantly compromising PelBsp-PulA stability. Deletion of region C eliminated its secretion. Taken together, these results suggested that at least three regions of PulA contain the secretion signals (Francetić & Pugsley, 2005) (Fig. 29).

The full-length PulA<sub>NA</sub> structure from *K. oxytoca* was determined X-ray crystallography (East et al., 2016). Atomistic molecular dynamics (MD) simulations of PulA in a 1-palmitoyl-2-oleoyl-sn-glycero3-phosphoethanolamine (POPE) model membrane and biochemical and functional studies provided an unprecedented molecular view of PulA. PulA consists of a N-terminal unstructured linker/tether, domains N1, N2 and N3 and A (comprising the sub-domain Ins) and the C domain (Fig. 30). Two putative secretion motifs, N1 and Ins, have been identified in PulA through a combination of structural, biochemical, and functional analysis.

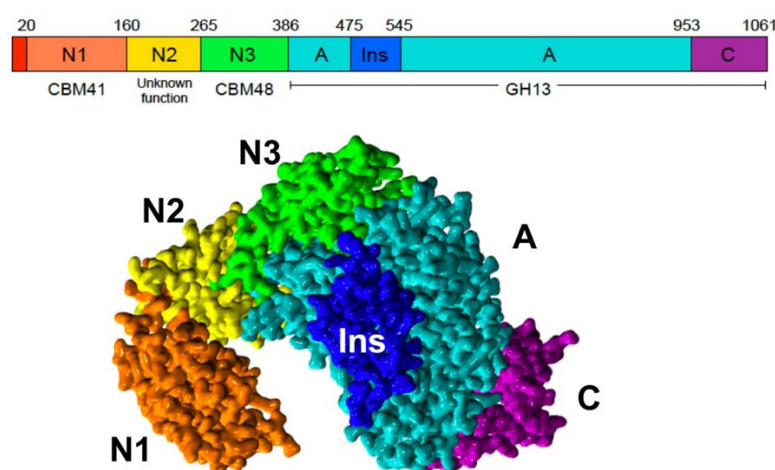


**Figure 30. *Klebsiella oxytoca* T2SS substrate PulA: structure and domain organization.**

Top: Surface representation of the three-dimensional structure of a PulA colored by structural domains (East et al., 2016). Bottom: the positions of PulA domains in the polypeptide chain. Note that the disordered N-terminal peptide of 19 residues shown in red is invisible in the X-ray structure.

East et al. showed that deletion of Ins domain abolished the PulA secretion in low copy number conditions, and deletion of N1 domain almost abolished its secretion and caused degradation (Fig. 31). Domain N1 would correspond to region A identified previously. As revealed by high crystallographic B-factor values, these PulA domains showed also substantial

conformational flexibility (Fig. 32) (East et al., 2016). This suggested that these domains might be related to the PulA secretion function.

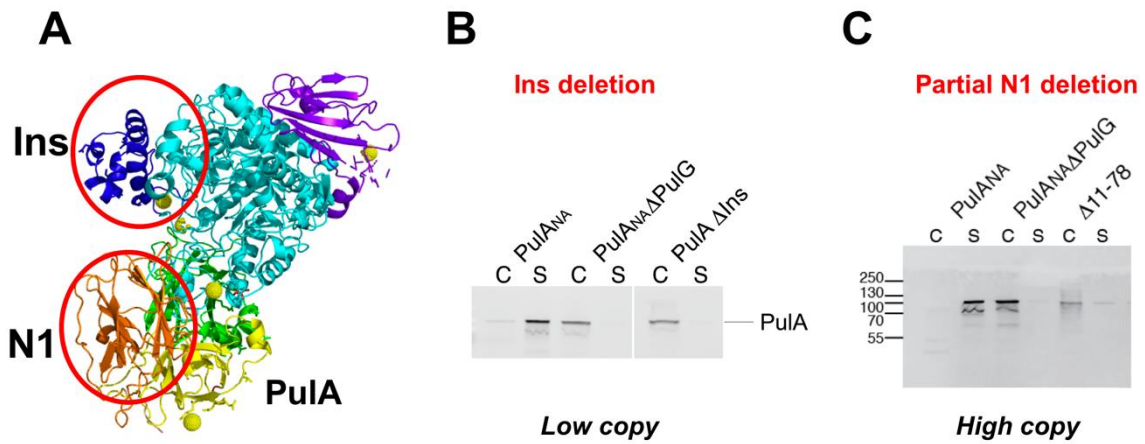


**Figure 31. PulA secretion signal studies: the roles of Ins and N1 domains.**

**A:** The PulA X-ray structure, with Ins and N1 domains indicated in red circles. **B:** Secretion of PulA<sub>NA</sub>ΔIns and controls in strain PAP5207 (low copy). C, cell; S, supernatant. **C:** Secretion of PulA<sub>NA</sub>ΔIns and controls in strain PAP7460 (high copy). B and C are all PulA immunodetection in cell (C) and supernatant (S) fractions from 0.05 OD<sub>600nm</sub> of PAP7460 bacteria containing pCHAP8185 plasmid producing PulA or PulAΔN1 variants and the Pul secretion system (East et al., 2016)

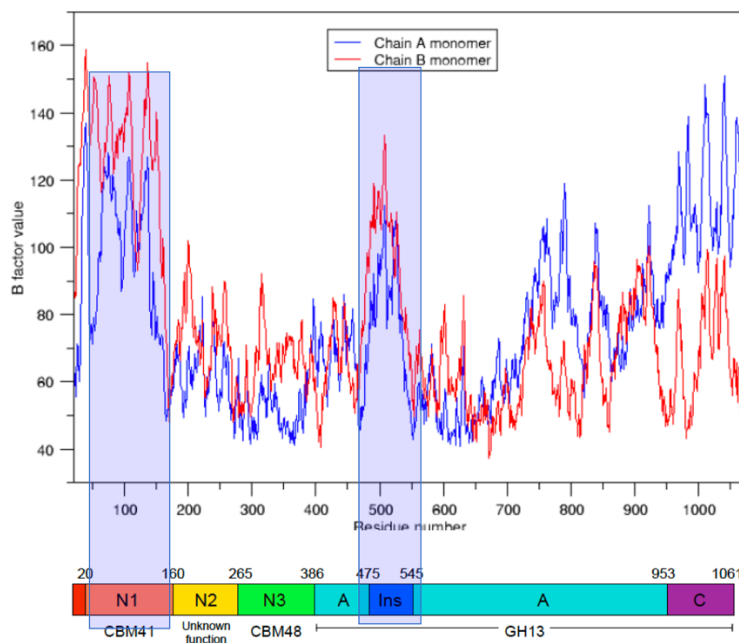


**Figure 32. K. oxytoca PulA domain organization and mobility.**

Top: PulA domain mobility is represented by the crystal B-factor variation of every Ca atom along the primary sequence. The lines correspond to the two PulA copies in the crystallographic unit shown in red and in blue and lined up with the different domains in the primary PulA structure (East et al., 2016).

#### 1.6.4 Interacting partners of substrates in the T2SS

Secretion signal studies raise the question of how the signal works. Secreted proteins (exoproteins) have no distinguishing features other than signal peptides that direct them to the Sec (He et al., 1991; Pugsley et al., 1991) or Tat (Voulhoux et al., 2001) pathways for periplasmic export. The obvious model is that the signal interacts with one or more components of the T2SS. Many studies have addressed this question.

In the Xcp T2SS *P. aeruginosa*, by combining the biophysical and biochemical approaches, biochemical studies defined the interactions between the substrates and three Xcp components, XcpP<sub>C</sub>, the secretin XcpQ<sub>D</sub>, and the pseudopilus tip complex (Fig. 33 and Fig. 34) (Douzi et al., 2011). In the following studies, this team found novel direct and specific interactions between the Xcp effector CbpD and periplasmic domains of XcpY<sub>L</sub> (XcpY<sub>LP</sub>) and XcpZ<sub>M</sub> (XcpZ<sub>MP</sub>) in the inner membrane assembly platform. They also found that there is a competition between XcpY<sub>LP</sub> and XcpZ<sub>MP</sub> for binding to CbpD, and that the CbpD effector might trigger XcpZ<sub>M</sub> dimerization and increase XcpZ<sub>M</sub>/Y<sub>L</sub> interaction (Michel-Souzy et al., 2018). Since all four minor pseudopilins in Xcp T2SS (XcpU<sub>H</sub>, V<sub>I</sub>, W<sub>J</sub>, X<sub>K</sub>) are required for secretion (Escobar et al., 2021), all the results so far suggest that the substrates of *P. aeruginosa* T2SS interact with XcpY<sub>L</sub>, XcpZ<sub>M</sub>, XcpP<sub>C</sub>, XcpQ<sub>D</sub> and four minor pseudopilins (XcpU<sub>H</sub>, V<sub>I</sub>, W<sub>J</sub>, X<sub>K</sub>).

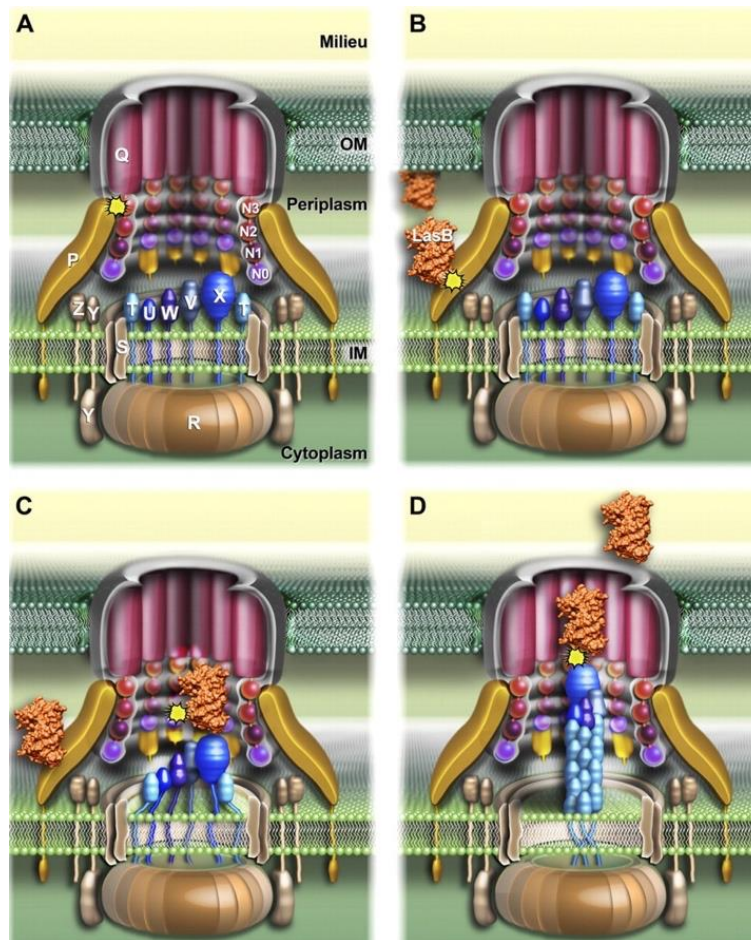


**Figure 33. Schematic three-dimensional model of the *P. aeruginosa* Xcp secreton. showing substrate recognition and transport during the type II secretion process.**

A, the schematic representation of the *P. aeruginosa* Xcp secreton. All Xcp components (labeled following Xcp nomenclature only) are represented according to their cellular localization, topology, and multimerization state. B–D, the different consecutive steps followed by the substrate for its recruitment, transport, and release by the secreton during the type II secretion process. The interactions identified in this study are represented by yellow asterisks (Douzi et al., 2011).

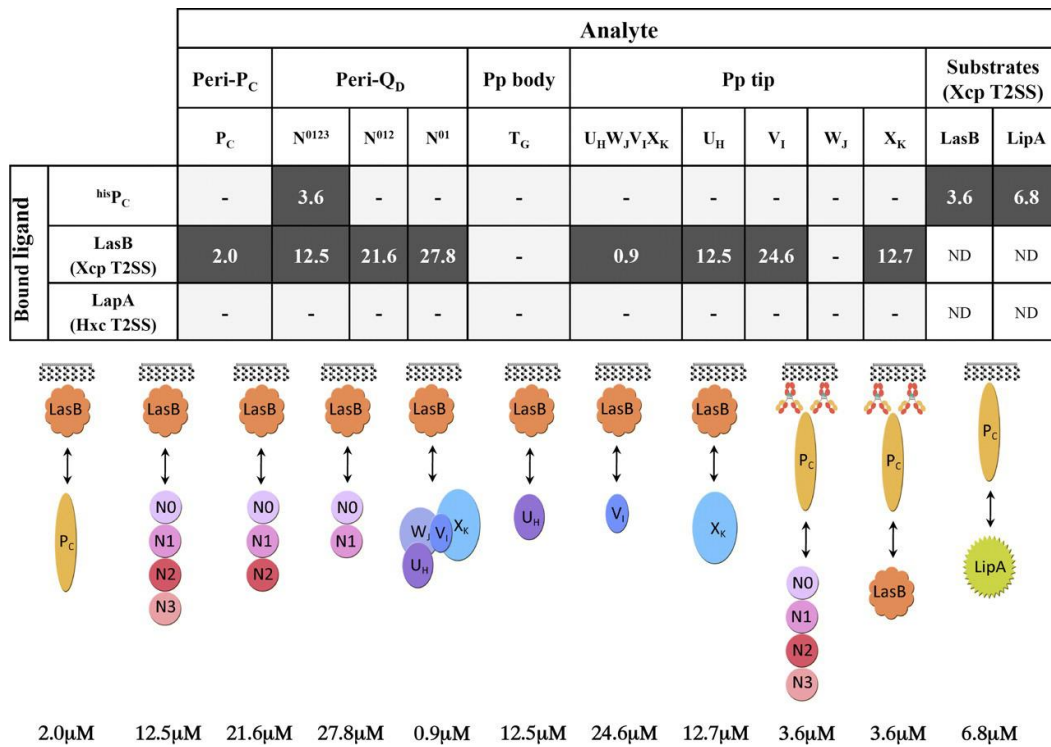


**Figure 34. Interactions between the periplasmic domains of the *P. aeruginosa* Xcp secreton and substrates.**

SPR was used to measure binding. Positive interactions are indicated by black boxes, in which the  $K_d$  values ( $\mu\text{M}$ ) of the interactions are indicated. Gray boxes with dashes indicate that no interaction was detected. The schemes for all positive interactions are represented below the table and follow the color code used in Fig. 33. Pp, pseudopilus; ND, not determined (Douzi et al., 2011).

Previous studies showed that two components of T2SS, GspC and GspD were involved in the recognition of the substrates (Lindeberg et al., 1996; Possot et al., 2000). In *Dickeya dadantii*, studies revealed that the fibronectin-like Fn3 domain of exoprotein Peli is essential for secretion (Pineau et al., 2014). Based on these results, bacterial two-hybrid assay and Surface plasmon resonance (SPR), they tested the interactions between the exoprotein Peli Fn3 domain and the periplasmic region of GspC or the isolated GspC HR or PDZ domain. They found that Fn3 domain interact with PDZ domain of GspC. Similar assays on GspD N0, N1 and N2 domains revealed an interaction between Fn3 domain and GspD N0 and N1 domains (Pineau et al., 2014).

So far, at least two working models for the role of the pseudopilus in T2SS were proposed: the piston model and the Archimedes' screw model (Fig. 35). In the piston model, the pseudopili extend and retract, and the substrate PulA interact with minor pseudopilins (PulH, I, J, K) on

the tip and is pushed out through the secretin across the OM. In the Archimedes' screw model, the pseudopili continuously assemble and extend to the outside of OM and are degraded there. This model predicts that in *K. oxytoca* PulA interacts with the major pseudopilins PulG of pseudopilus core (Fig. 35).

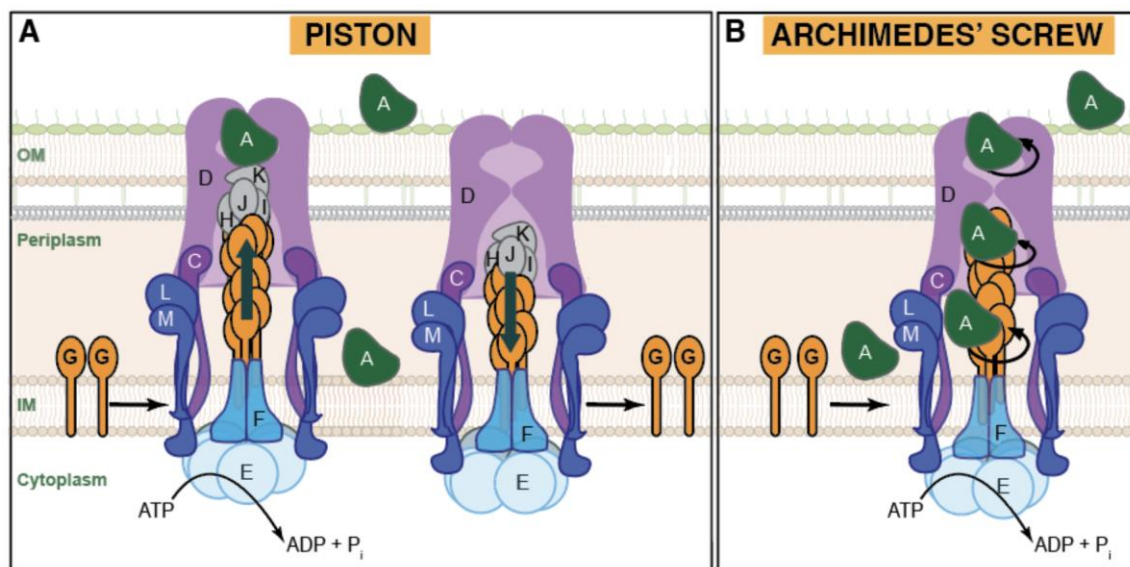


**Figure 35. The models of type II secretion mechanism.**

**A:** The piston model, predicts that substrates interact with the pseudopilus tip complex. Cycles of pseudopilus assembly and retraction drive the exoproteins across the OM channel. **B:** Archimedes' screw model proposes a continuous assembly of the pseudopilus and its degradation at the tip, in addition to substrate binding to the major pseudopilin.

To test which of these models is closer to the reality, or if there is a third model that combines features of both, and to complete the substrate interaction map within the Pul system, we designed and performed different experiments described below.

## **2. Objectives**

## 2.1 Identifying PulA interacting partners in the T2SS

In the previous studies of the PulA secretion, East et al. identified two regions of pullulanase that might contain the secretion signal, designated domain N1 and Ins subdomain. Both regions appeared to be required for PulA secretion, but also to play different roles. Our main goal was to understand the function of these putative secretion signals by identifying T2SS components that interact with these regions.

To study PulA interactions with T2SS proteins, we planned to use several approaches: the bacterial two-hybrid (BACTH) analysis, chemical cross-linking and co-purification, and *in vivo* UV crosslinking with unnatural amino acids.

The BACTH analysis is a powerful method to study interactions of cytoplasmic or membrane proteins in their full-length form and native localization (Karimova et al ref). Proteins fused to the T18 and T25 fragments of CyaA toxin catalytic domain, if they interact with each other, bring T18 and T25 fragments close in space and reconstitute their adenylyl cyclase activity in the cytoplasm. Interaction of periplasmic or membrane domains can therefore result in production of cAMP in the cytoplasm and activate transcription of cAMP-dependent reporter genes, such as *malE* or *lacZ*. The advantage of this approach is that one can probe interactions of full-length proteins in their native environment, cytoplasm or the membrane. The limitation is that soluble periplasmic proteins and OM proteins cannot be studied in these compartments. They can be tethered to a transmembrane segment or produced in the cytoplasm, but in the latter case the risk is that they are not correctly folded because of different physical-chemical conditions, protein and macromolecular content between the cytoplasm and the periplasm.

The *in vivo* site-specific photo-cross-linking approach has the potential to capture transient protein-protein interactions by crosslinking residues in close contact with an UV-activatable, unnatural amino acid residue *p*-benzoyl-L-phenylalanine (*p*Bpa) incorporated in a protein under study (Singh & Wagner, 2019). The genetic encoding and translational incorporation of the synthetic amino acid *p*Bpa at a desired position of a protein of interest is the basis for *in vivo* photo-crosslinking (Farrell et al., 2005). An amber stop codon that is recognized by a suppressor tRNA coupled with *p*Bpa initiates the encoding. A mutant *Methanococcus jannaschii* tRNA synthetase is responsible for charging the tRNA with *p*Bpa (Chin et al., 2002) (Fig. 36). *p*Bpa reacts to UV radiation at 365 nm by forming a covalent bond crosslinking to the peptide

backbone of adjacent proteins. Bacteria like *E. coli* can withstand suppressing amber stop codons. Pineau *et al.* were the first to apply the *in vivo* site-specific photo-cross-linking approach in secretion signal studies of the *D. dadantii* T2SS, and this allowed them to detect interactions between the exoprotein PelI and T2SS components OutC and OutD (Pineau *et al.*, 2014).

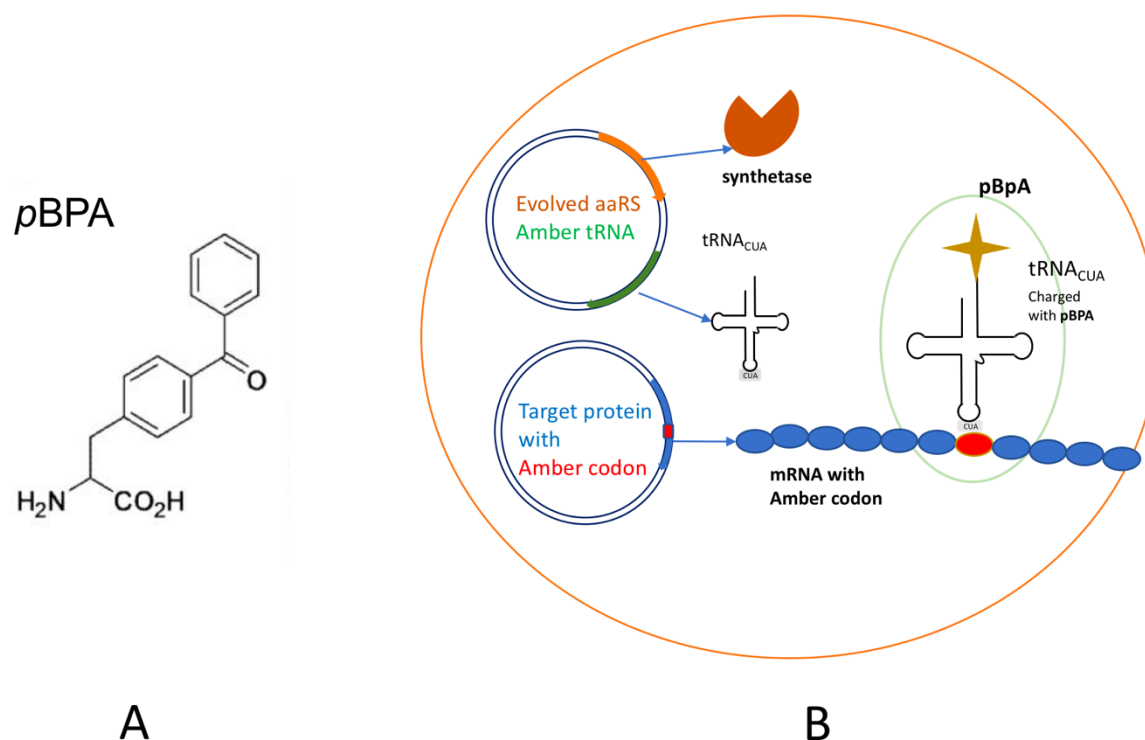


**Figure 36. The principle of pBpa incorporation into proteins in vivo.**

**A.** Schematic representation of the genetic system for targeted pBpa incorporation comprising the target gene with amber codons in specific position, and plasmid encoding amber suppressor tRNAs (several copies) and the pBpa tRNA synthetase evolved to recognize and load pBpa on the suppressor tRNA. **B:** The chemical structure of the artificial amino acid- pBpa used in this study (Chin *et al.*, 2002).

To implement the unnatural amino acid nonsense suppressor approach on the T2SS we proceeded in two steps. Initially I used PulG as a small model protein to implement the method. The major pseudopilin PulG is essential for PulA secretion and we hypothesized that this approach might allow us to test whether it interacts directly with PulA. This interaction would support the Archimedes' screw model of secretion.

In the second stage, we designed strategies to introduce pBpa in PulA domains N1 and Ins. Unpublished results of BACTH analysis and co-purification studies in the host laboratory have

identified several candidate T2SS components that might interact with PulA. These proteins either reside in the inner membrane (PulM, PulC) or form the pseudopilus (PulG). To capture potentially transient interactions of PulA during its transport across the T2SS complex, we planned to employ site-specific UV crosslinking implemented above. Amber codons would be introduced at specific positions of the N1 and Ins domains of PulA. After incorporating the unnatural amino acid *p*Bpa in these positions, we would select PulA-*p*Bpa derivatives that are efficiently secreted. These bacteria would be UV-irradiated to capture the crosslinked PulA adducts during transport. Biochemical analysis and immunodetection would be used to identify crosslinked partners.

The *p*Bpa UV crosslinking approach has its limitations. It is very labor-intensive and risky, since the choice of *p*Bpa insertion site determines whether the interacting partner can be detected. To overcome this limitation, we also envisioned the use of chemical crosslinking on the whole cells followed by co-purification of PulA labeled with an affinity tag – Strep or His<sub>6</sub>.

## 2.2 Structural and functional characterization of PulL and PulM components of the Assembly platform

Toward the end of my thesis, I was involved in a collaborative study with the group of N. Izadi-Pruneyre to decipher the structure and function of two essential components of the assembly platform PulL and PulM. Nadia's group used X-ray crystallography and solution NMR to determine the structures of the C-terminal domains of PulL (PulL<sub>CTD</sub>) and PulM (PulM<sub>CTD</sub>). For the first time, they were able to determine the structure of a heterodimer formed by these T2SS components. To test whether interaction between their CTDs of PulL and PulM in solution corresponds to their interaction *in vivo* in bacterial cells, I used crosslinking and BACTH to study the interactions of full-length proteins. Together these studies allowed us to build a model of the full-length PulL – PulM complex in the inner membrane.

I also participated in the functional analysis of PulL and PulM mutant variants and their interactions. Finally, I determined the ratio of PulL and PulM levels in bacteria, which was important to explain the dynamics of their interaction and to propose a functional model. The manuscript describing these results is under revision in the journal *Structure*.

We extended the study of PulL and PulM to test whether other regions of these proteins contribute to their interaction. We combined deletion analysis of periplasmic and cytoplasmic regions of PulL and PulM with the BACTH and functional analysis. The results showed that the periplasmic helices of PulL and PulM forming a coiled coil play a very important role in their interaction. The cytoplasmic peptide of PulM was also crucial for the stability of PulL and PulM, suggesting a close interaction of the two proteins in the cytoplasm. This cytoplasmic peptide is also crucial for PulM interaction with the major pseudopilin PulG. These results support a model of dynamic interactions of PulM with PulG and PulL involving this N-terminal peptide. These results are presented in Chapter 3.2.2. in a form of a manuscript that is under review in *Research in Microbiology*.

## **3. Results**

### 3.1 PulA interactions with the T2SS components

As emphasized in the Introduction, T2SS plays an important role in bacterial survival and growth on large or complex nutrients. One of the main goals of this project was to characterize the mechanism of substrate specificity in the Pul T2SS by identifying PulA interacting partners during secretion, and their precise contacts during this interaction.

To achieve these goals, we used different approaches established in the host laboratory, such as Bacterial two-hybrid assay and chemical crosslinking coupled to affinity purification. We also planned to introduce a new approach of *in vivo* crosslinking after incorporation of the unnatural amino acid *pBPA* in PulA and UV irradiation. The main axis of my work was aimed at setting up this approach for use in the Pul T2SS reconstituted in *E. coli*.

#### 3.1.1 The *in vivo* site-directed photo-cross-linking

##### 3.1.1.1 *pBpa* incorporation and *in vivo* site-specific photo-cross-linking of PulG

The interactions between the substrate PulA with other T2SS components during secretion, or the interactions between T2SS components during the assembly process, are presumably transient. To capture these transient interactions during these highly dynamic processes, we established *pBpa* incorporation and *in vivo* site-specific photo-cross-linking in our working model, the *K. oxytoca* Pul T2SS.

To implement the unnatural amino acid nonsense suppressor approach on the T2SS we proceeded in two steps. Initially I used PulG as a small model protein to implement the method. The major pseudopilin PulG, which is the most abundant protein in T2SS, is essential for PulA secretion and we hypothesized that this approach might allow us to test whether it interacts directly with PulA (Fig. 37).

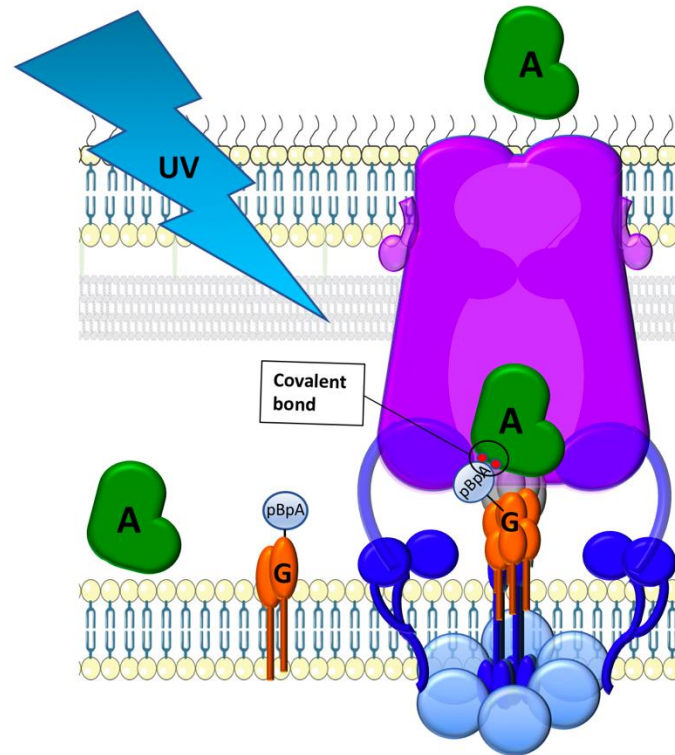


**Figure 37. The effect of UV irradiation: covalent cross-linking.**

The schematic representation of the T2SS with PulG (in orange) modified with pBpa. A covalent bond would be formed between PulG and the protein that comes into close contact (here PulA, in green) when the cells are UV irradiated.

The Archimedes' screw model of type II secretion predicts that the substrate interacts with the pseudopilus core. Our collaborator Benjamin Bardiaux (Structural Bioinformatics Unit, Institut Pasteur) performed docking *in silico* of PulA and the pseudopilus to test this possibility. This docking predicted that the PulA Ins domain containing a putative secretion signal fits into the pseudopilus surface groove. Several residues on PulG and PulA Ins domain were identified as “hot spots” involved in specific contacts (Fig. 38). Based on this model, we focused on PulG surface residues not only in the hot spots but also sites near them, and introduced Amber codons at these specific positions by site-directed mutagenesis.

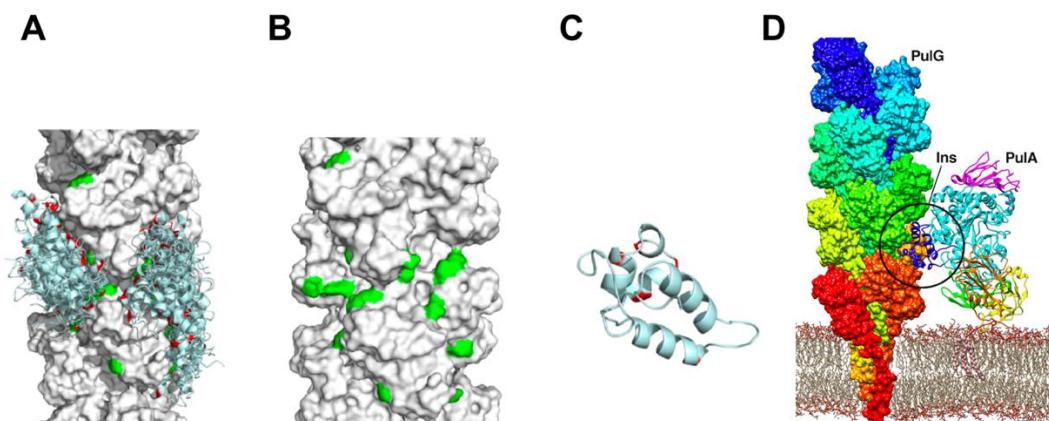


**Figure 38. In silico prediction of interactions between the pseudopilus and Pula.**

The model predicts specific binding of Pula Ins domain to the pseudopilus surface groove, indicated by a black circle. A: Many possible poses. Some redundancy due to the periodicity of the pilus but some residues are more often at the interface (hotspots); B: “hotspots” in PulG pilus in green: 33R, 78R, 79N, 82E, 91Q and 123N; C: “hotspots” in Pula Ins domain, in red: 492E, 498K, 501E, 505Y, 507D and 521Q; D: Docking in silico showing a specific fit of Ins domain into the pseudopilus surface groove. (Figure by Benjamin Bardiaux).

The following PulG residues were substituted: E29, D32, R33, Q65, A70, S72, A73, E74, R78, N79, E82, Q91, D97, L101, Q108, F112, N123. Their positions on the structure of PulG monomer are shown in Fig. 39.

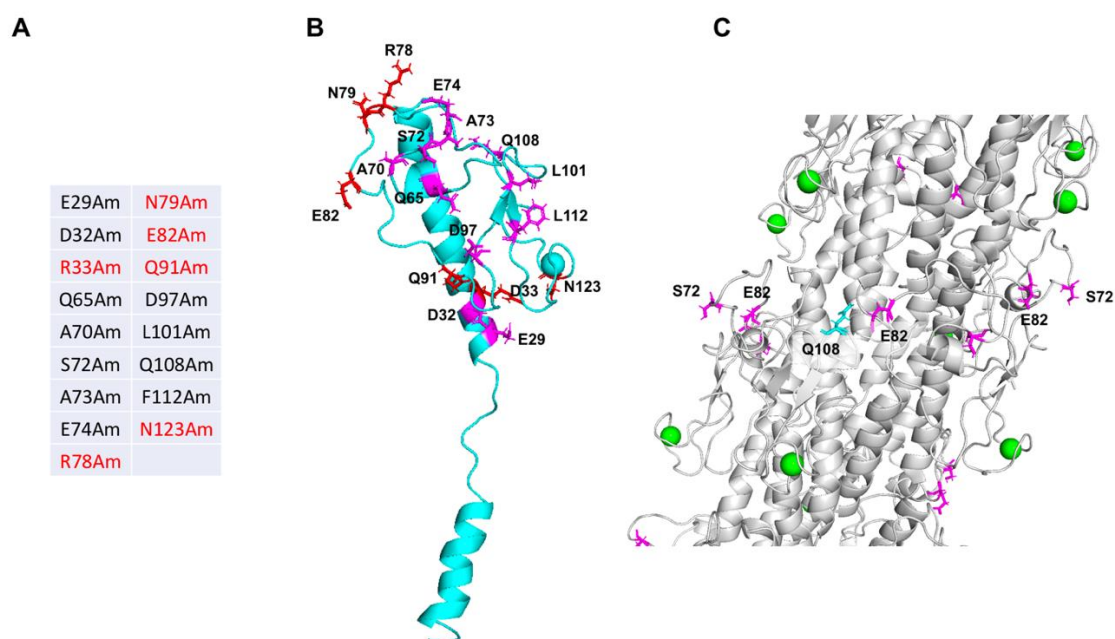


**Figure 39. Residues of PulG monomer chosen for site-directed mutagenesis.**

A. The list of PulG residues selected for pBPA replacement. Residues predicted to bind Pula are shown in red. B. The selected residues to be replaced by pBpa shown in sticks on the cartoon model of PulG: their positions in the PulG sequence are indicated. C. The pseudopilus model showing residues S72, E82 and Q108, exposed on the pseudopilus surface, that might interact with the substrate.

Amber suppressor mutations are present in many *E. coli* strains and might interfere with the incorporation of the unnatural amino acid *p*Bpa in these positions. To select *E. coli* strains without endogenous Amber suppressor that could be used in our experiments, we picked two plasmids encoding PulG Amber variants and transformed them into several different *E. coli* strains, which we tested for the presence of PulG<sup>WT</sup> and PulG<sup>Am</sup> variants by Western blot (Fig. 40). Out of the four strains tested, BL21 showed good expression of *pulG*<sup>WT</sup> and no endogenous Amber suppressors and MG1655 strain was also promising (Fig. 40A). To limit the expression of *pulG* under *placZ* promoter control, we introduced by conjugation an episome overexpressing the *lacI* repressor gene to give MG1655 *F'**lacI*<sup>Q</sup>.

We next compared this strain with strain BW25113 *F'**lacI*<sup>Q</sup> (wild type *E. coli* which was used to construct the Keio collection of *E. coli* gene knockouts (Baba et al., 2006). Strain BW25113 *F'**lacI*<sup>Q</sup> did not have endogenous suppressors, and showed a higher yield of PulG production compared to MG1655 *F'**lacI*<sup>Q</sup> (Fig. 40B), none of them produced full length PulG. According to the results all above, strains without endogenous Amber suppressor that could be used in future experiments are: BL21, MG1655 *F'**lacI*<sup>Q</sup> and BW25113 *F'**lacI*<sup>Q</sup>.

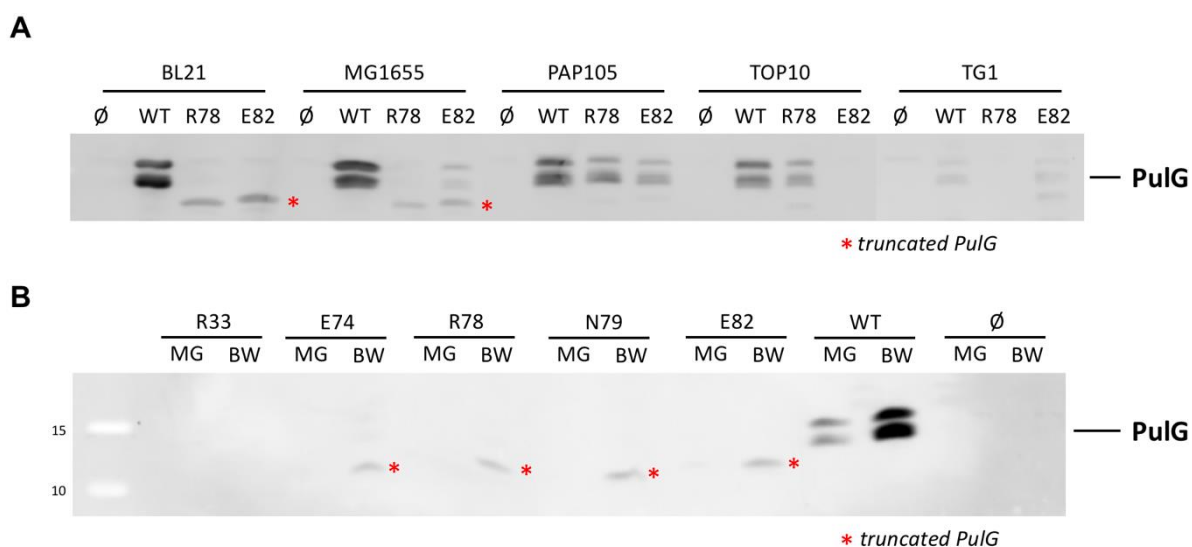


**Figure 40. Selection of *E. coli* host strains for *p*Bpa incorporation.**

**A.** Comparison of PulG levels in total extracts of *E. coli* strains transformed with plasmids *p*SU18 (∅) *p*CHAP8658 encoding PulG (WT), and its derivatives with amber codons at positions 78 (R78) and 82 (E82). The names of *E. coli* strains are indicated on top. The position of full length PulG is indicated on the right. Red stars indicate truncated PulG fragments. **B.** Comparison of strains MG1655 *F'**lacI*<sup>Q</sup> (MG) and BW25113 *F'**lacI*<sup>Q</sup> (BW) transformed with plasmids *p*SU18 (∅), *p*CHAP8658 encoding PulG (WT) and amber codon derivatives replacing residues indicated above the horizontal lines. M, Molecular weight markers (in kDa).

The *K. oxytoca pul* locus contains 16 genes encoding the T2SS, which were cloned on a moderate copy-number plasmid pCHAP231 (D'Enfert & Pugsley, 1987). When introduced into *E. coli*, this system produces a functional T2SS that secretes PulA on the cell surface. It is difficult to introduce Am mutations in this large plasmid by PCR mutagenesis, but such mutations can be easily introduced in plasmids carrying individual *pul* genes. Therefore, we combined pCHAP231 derivatives with nonpolar mutations in individual *pul* genes with plasmids containing the complementing genes on compatible vectors pSU18 or pSU19. To further facilitate biochemical and functional analysis, we used the derivative of plasmid pCHAP231, pCHAP8185, which encodes non-acylated PulA variant.

To make the mutations and study their effect, we combined two compatible plasmids. In the case of *pulG*, plasmid pCHAP8184 (a derivative of plasmid pCHAP8185 carrying ColE1 replication origin and Ap<sup>R</sup> marker) carrying a nonpolar *pulG* deletion was combined with pCHAP8658 carrying *pulG* gene under *placZ* promoter control.

The *pBpa* incorporation at the desired position requires a third plasmid carrying the nonsense suppressor tRNAs and its amino-acyl-tRNA synthetase, both adapted to load *pBpa* on the tRNAs carrying the amber anticodon. The plasmid containing these genes was obtained from the Shultz's team and we made it compatible with our system. We changed its resistance gene and *ori* gene, and constructed a derivative of this plasmid named pMS5005 encoding: an amino-acyl-tRNA synthetase under arabinose promoter control; Amber suppressor tRNAs; pCDF Ori gene; and Spectinomycin resistance (Sp<sup>R</sup>) marker. To accomplish our *in vivo* site-specific photo-cross-linking assay, we transformed these 3 plasmids into *E. coli* and made the so-called 3-plasmid system (Fig. 41).

## *E. coli*

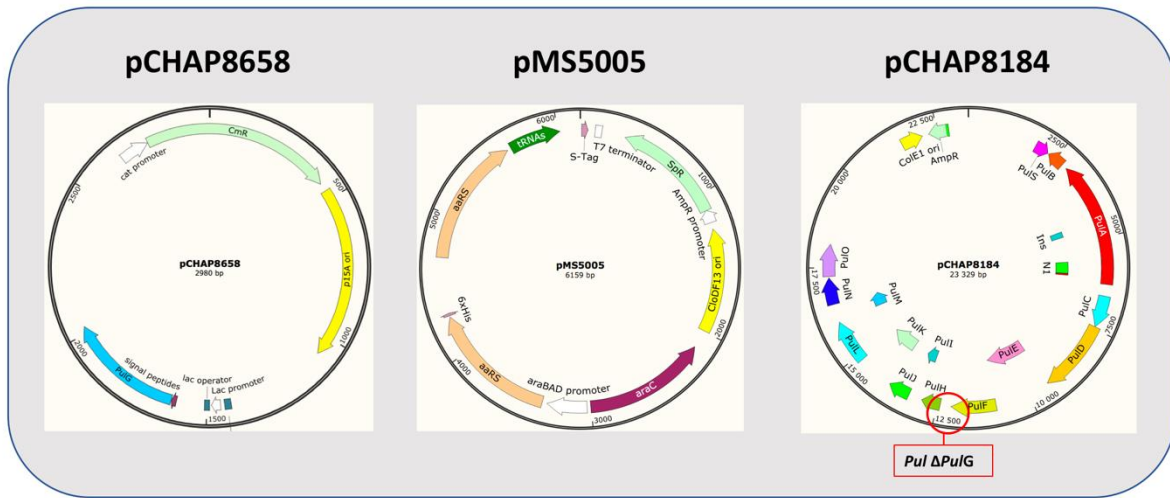


**Figure 41. The 3-plasmid system for *PulG* in vivo site-specific photo-cross-linking in *E. coli*.**

*pCHAP8658*: encoding *PulG*<sup>WT</sup> or *PulG*<sup>Am</sup>, *lacZ* promoter, chloramphenicol resistance (*Cm*<sup>R</sup>); *pMS5005*: encoding tRNAs and *araRS* under *araBAD* promoter, spectinomycin resistance (*Sp*<sup>R</sup>); *pCHAP8184*: encoding every *pul* gene except *pulG*, maltose promoter, ampicillin resistance (*Ap*<sup>R</sup>).

We validated the function of *pMS5005*, using several *PulG*<sup>Am</sup> variants in strain BL21 to test the *pBpa* incorporation efficacy. Bacterial precultures were grown in M63B1 minimal medium containing 0.5% glycerol and antibiotics at 30°C and 1% casamino acids were added to improve growth. Then we performed *pBpa* incorporation assay, following the the *pBpa* incorporation protocol (1) in Materials and Methods, section 4.5.

We collected the bacterial total extracts and analyzed them by SDS-PAGE and Western Blot with anti-*PulG* antibodies. As shown in Fig. 42, *PulG*<sup>E78Am</sup> and *PulG*<sup>E82Am</sup> showed good incorporation of *pBPA*, indicating that the *pMS5005* plasmid was functional and allowed efficient *pBpa* incorporation. However, in other positions, *pBpa* incorporation efficiency was very low. Kensaku Sakamoto's team have studied this phenomenon ([website 1](#)). In their study, they found that it is generally much easier to incorporate *pBPA* in response to UAG just preceding a codon starting with a purine than UAG just preceding a codon starting with a pyrimidine. Our results perfectly match their conclusion.

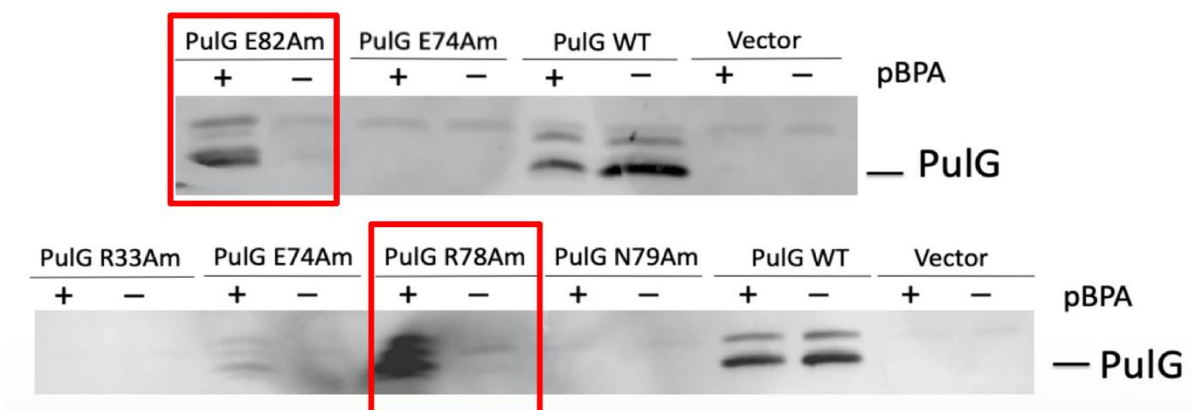


**Figure 42. Test of pBpa incorporation in different PulG positions.**

Samples indicated in red squares showed good and specific incorporation of pBpa compared to the positive PulG<sup>WT</sup> control.

We next tested the 3-plasmid system by adding the plasmid pCHAP8184 encoding the Pul factors. We selected multiple *pulG<sup>Am</sup>* clones showing good pBpa incorporation efficiency for UV-irradiation to capture the crosslinked PulG adducts during secretion. The work flow of the *in vivo* site-specific photo-cross-linking assay is shown in Fig. 43.

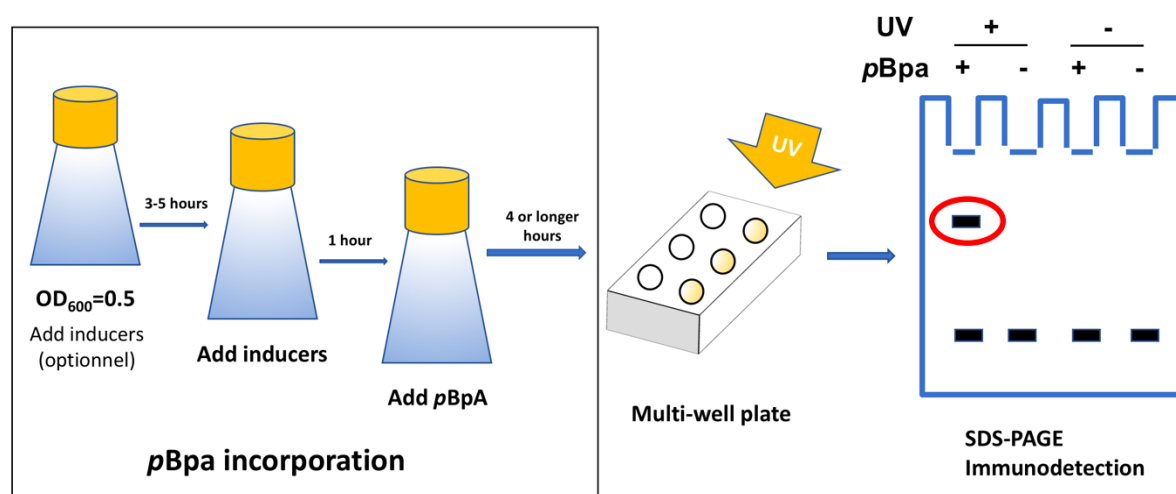


**Figure 43. The work flow of the *in vivo* site-specific photo-cross-linking assay.**

Overnight preculture are normalized in non-inducing minimal M63 medium with 0.5 % glycerol to  $OD_{600}$  of 0.5. For pBpa incorporation protocol (1) or (2) is followed for different experiments (Materials and methods). Half of the cultures are UV irradiated and total bacterial extracts are analyzed by SDS-PAGE and Western blot. The red circle shows the crosslinked band specific for + pBpa and +UV samples.

We tested every *pulG* Amber mutant we made, and we detected two crosslinked bands in PulG S72-*pBpa* and E82-*pBpa* variants. As the Western blot results showed in Fig. 44, the sizes of cross-linked bands are approximately 30 kD and 40 kD, and PulG Mw is 14607 Da, suggesting that the crosslinked partner size should be around 15 kD – 26 kD. Deduced from the size, the crosslinked partners could be PulM and PulH or PulG itself. To verify whether the crosslinked partner(s) is (are) specific to T2SS, we performed the same experiments without *pul* genes in a 2-plasmid system, consisting of PulG S72*pBpa* or E82*pBpa* variants and pMS5005. The same bands at the same positions were detected against anti-PulG antibody (Fig. 45). Thus, we could confirm that the “crosslinked bands” are not T2SS related and probably represent PulG dimers.

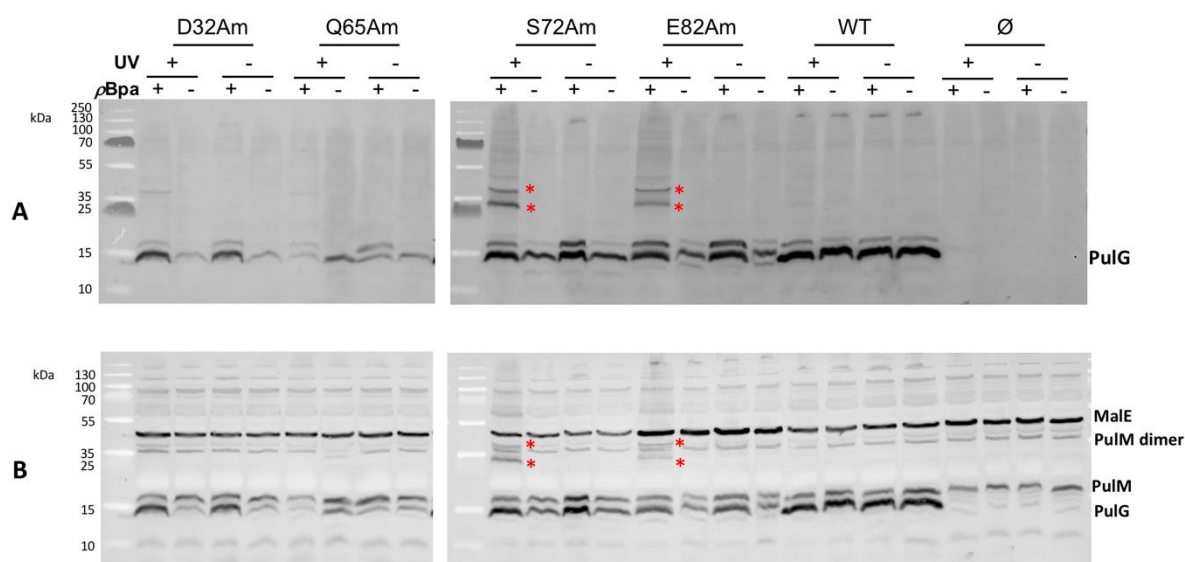


**Figure 44. Immunodetection of PulG-*pBpa* variants after the in vivo site-specific photo-cross-linking assay.** Total bacterial extracts were analyzed on 10% Tricine gels and probed with anti-PulG antibodies (A), followed by anti-PulM antibody (B). The strain used was MG1655 *F'*Tet. UV irradiation time was 30 mins. Red asterisks indicate the crosslinked bands.

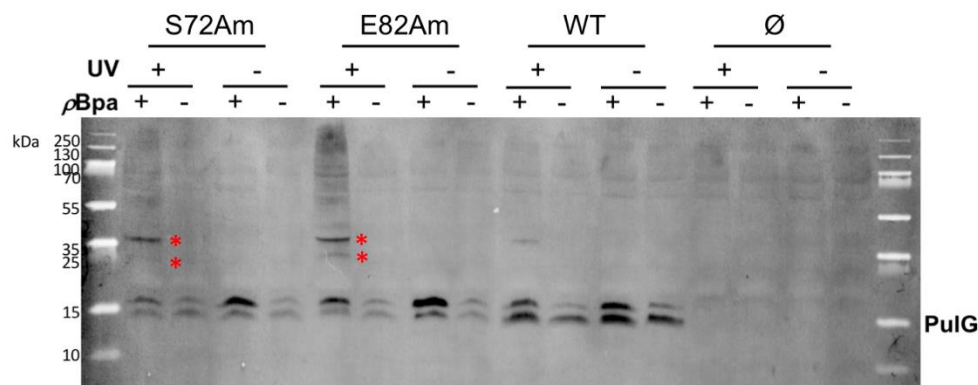


**Figure 45. Validation of PulG crosslinks without Pul factors.**

Samples of *pulG* S72Am, *pulG* E82Am mutants, *pulG*<sup>WT</sup> and empty vector control (Ø) were analyzed by SDS-PAGE on a 10% Tricin gel, followed by Western blot probed by anti-PulG antibodies. Red asterisk indicates the crosslinked bands.

Nevertheless, these preliminary results showed that the 3-plasmid system could promote reasonably good *pBpa* incorporation and that under the conditions of our experiments, UV crosslinking was able to produce adducts. At the same time, we realized that the validation of crosslinked partner(s) by antibody was limited by the detection sensitivity and the ability of the antibodies to recognize the crosslinked proteins. We concluded the sensitivity could be improved by co-purifying PulG complexed *via* tagged variants of PulG and the use Mass Spectrometry (MS) to identify the crosslinked partner.

### 3.1.1.2 Construction of PulA-pBpa derivatives and cross-linking assays

In the next step, we adapted these experimental procedures and conditions for the use for *in vivo* site-specific photo-cross-linking of the nonacylated variant (PulA<sub>NA</sub>). As discussed in the Introduction, previous studies identified two domains in PulA, N1 and Ins that are likely to contain a secretion signal, but play potentially different roles. Deletion of Ins sub-domain abolished the PulA secretion in conditions of physiological expression levels, and deletion of N1 domain almost abolished secretion in overexpression conditions, while also causing PulA degradation. These results, together with the high mobility of these domains, led us to hypothesize that they domains interact with the T2SS factors during secretion.

We therefore focused on N1 and Ins domain to perform the *in vivo* site-specific photo-cross-linking assays. We picked residues on the exposed surface of PulA N1 domain, and the

residues of the Ins domain predicted by *in silico* model to be involved in pseudopilus interaction (Fig. 38). We introduced Amber codon in these positions by site-directed mutagenesis (Fig. 46).

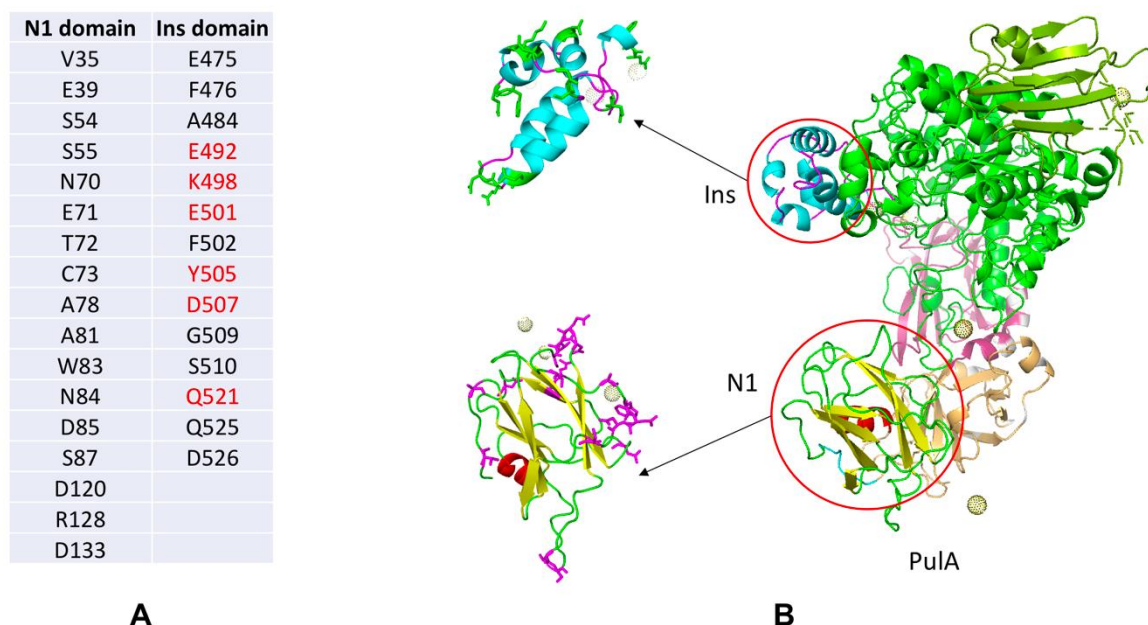


**Figure 46. Residues in PulA N1 and Ins domains selected for site-direct mutagenesis.**

**A:** The list of residues in PulA N1 and Ins domains selected for site-direct mutagenesis; residues in red are the hot spots predicted by *in silico* docking. **B:** X-ray structure of PulA with Ins and N1 domains indicated in red circles. Surface residue positions (shown in sticks) selected for site-direct mutagenesis into Amber codon.

Similar as for PulG, we generated a 3-plasmid system for PulA *in vivo* site-specific photo-cross-linking assay (Fig. 47), comprising the pSU18 derivative pCHAP8255 encoding PulA<sup>WT</sup> or PulA<sup>Am</sup> variants, driven by the *lacZ* promoter, with chloramphenicol resistance (Cm<sup>R</sup>); plasmid pMS5005 encoding tRNAs and *aaRS*, controlled by *araBAD* promoter, carrying spectinomycin resistance (Sp<sup>R</sup>); and the large plasmid pCHAP8861, encoding every *pul* gene except *pulA*, under control of maltose promoters, carrying ampicillin resistance (Ap<sup>R</sup>).

*E. coli*

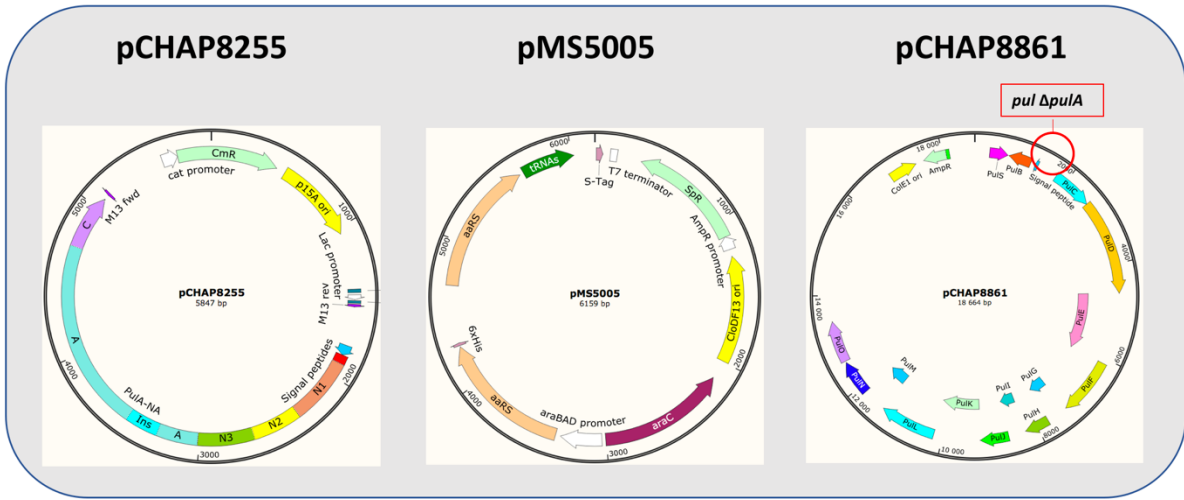


**Figure 47. The 3-plasmid system for PulA in vivo site-specific photo-cross-linking in *E. coli*.**

*pCHAP8255: encoding Pula<sup>WT</sup> or Pula<sup>Am</sup> variants under control of lacZ promoter, chloramphenicol resistance (Cm<sup>R</sup>); pMS5005: encoding tRNAs and aaRS, araBAD promoter, spectinomycin resistance (Sp<sup>R</sup>); pCHAP8861 carrying every pul gene except pula, maltose promoter, ampicillin resistance (Ap<sup>R</sup>).*

To test the system, we first introduced Amber codons into positions of PulA S54 and E71 (domain N1) and S510 or D526 (domains Ins) on plasmid pCHAP8255. We used these constructs to test the *pBpa* incorporation and *in vivo* site-specific photo-cross-linking of these variants in strain BW25113 F'Tet. The protocols used are described in Materials and Methods. These four PulA<sup>Am</sup> variants showed good *pBpa* incorporation, and we detected one crosslinked band in PulA<sup>E71Am</sup> variant upon *pBpa* addition and UV irradiation, which was missing in the negative control (without *pBpa* addition) and in PulA<sup>WT</sup> (Fig. 48). To test if this crosslinked bands are T2SS specific, we performed the same experiments in a 2-plasmid system without Pul factors, however, a band at the same position was detected again, suggesting that the crosslinked band contains an *E. coli* protein that is not T2SS related.

Meanwhile, we also noticed that, due to the high molecular weight of PulA and the bad separation of high molecular weight proteins by SDS-PAGE, it would be difficult to distinguish between the PulA and PulA adducts, if the crosslinked partner is of relatively small molecular weight. Consequently, it was difficult to deduce the size of crosslinked protein.

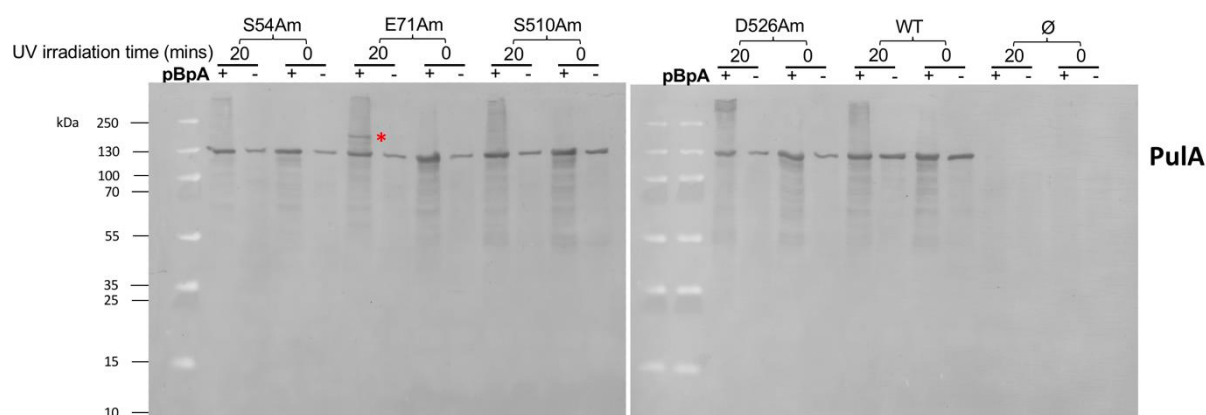


**Figure 48.** *PulA<sup>Am</sup> variants pBpa incorporation and in vivo site-specific photo-cross-linking assay.*

*Strain used is BW25113 F<sup>+</sup>Tet. Samples were analyzed by SDS-PAGE on 8% Tricin gel, followed by Western blot probed with anti-PulA antibodies. Red asterisk indicates the crosslinked band.*

In order to overcome the limitation of immunodetection of the crosslinked adducts, we introduced His<sub>6</sub>-tag at the C terminus of PulA, to be able to purify PulA and PulA complexes after *in vivo* site-specific photo-cross-linking. We contacted the Mass spectrometry (MS) platform to devise an approach to identify the crosslinked proteins. We constructed plasmid pMS5032 encoding PulA-His<sub>6</sub>, under *placZ* promoter, carrying chloramphenicol resistance (Cm<sup>R</sup>), and then introduced Amber codon to obtain plasmid pMS5040 encoding PulA<sup>E71Am</sup>-His<sub>6</sub> variant. We tested the secretion function and performed pBpa incorporation and photo-cross-linking assay and included 3 negative controls: without pBpa addition, PulA<sup>WT</sup>-His<sub>6</sub>, PulA<sup>WT</sup>-His<sub>6</sub> without Pul factors (the corresponding plasmids are listed in Table 2). I transformed plasmids into BW25113 F<sup>+</sup>Tet strain and tested PulA secretion, as well as pBpa incorporation and *in vivo* site-specific photo-cross-linking (see Materials and methods). The results are shown in Fig. 49. In the secretion assay, we noticed that, compared to the PulA-His<sub>6</sub> (indicated as WT) and the negative control ( $\Delta$ Pul), PulA<sup>E71Am</sup>-His<sub>6</sub> variant accumulated in the cells and its secretion was reduced (Fig. 49A). In the pBpa incorporation and *in vivo* site-specific photo-cross-linking assay, a crosslinked band was present specifically for the PulA<sup>E71Am</sup>-His<sub>6</sub> variant, and not in the negative controls (Fig. 49B).

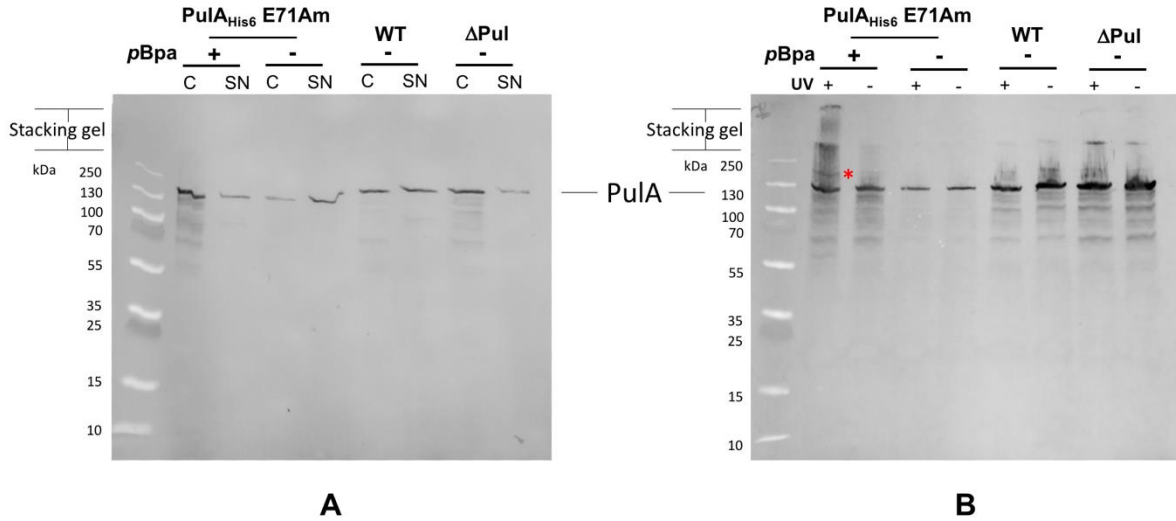


**Figure 49. The assay of PulA-*His*<sub>6</sub> and PulA<sup>E71Am</sup>-*His*<sub>6</sub> secretion, *pBpa* incorporation and *in vivo* site-specific photo-cross-linking.**

**A:** PulA secretion assay with indicated PulA variants and conditions. (C) cell fraction; (SN): supernatant fraction.  
**B:** *pBpa* incorporation and *in vivo* photo-cross-linking assay. Red asterisk indicates the crosslinked band.

To purify the PulA *pBpa*<sup>72</sup> adduct, I performed *pBpa* incorporation and UV crosslinking in a larger scale, using 200 mL of bacterial culture. I grew strains containing plasmids pMS5005, pCHAP8861 encoding Pul factors and either PulA<sup>E71Am</sup>-*His*<sub>6</sub> or PulA-*His*<sub>6</sub>, for protein purification. I performed *pBpa* incorporation and *in vivo* photo-cross-linking. Then, I disrupted the bacteria as described in Materials and methods and purified PulA-*His*<sub>6</sub> variants from the two cultures on a Talon His-tag Affinity Resin. I checked the purification efficiency by SDS-PAGE and Coomassie Blue staining. Both PulA-*His*<sub>6</sub> and PulA<sup>E71Am</sup>-*His*<sub>6</sub> showed good yield and the crosslinked band of PulA<sup>E71Am</sup>-*His*<sub>6</sub> was visible by Coomassie Blue staining, suggesting that the amount of protein should be sufficient for MS determination. The band was absent in the PulA-*His*<sub>6</sub> control (Fig. 50).

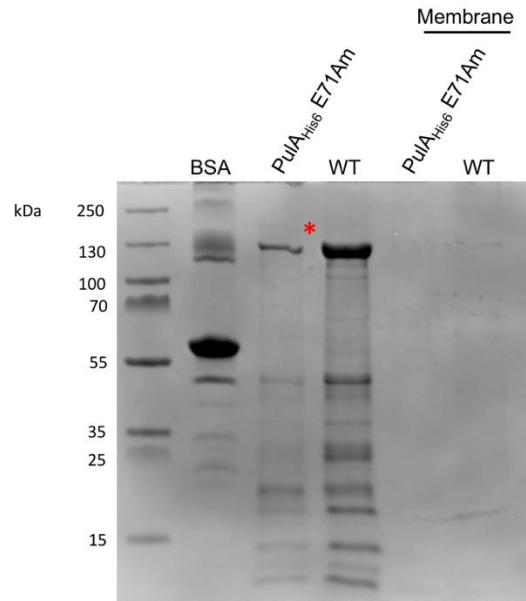


**Figure 50. Coomassie blue staining of purified PulA-His<sub>6</sub> and PulA<sup>E71pBpa</sup>-His<sub>6</sub> proteins.**

The asterisk indicates the co-purified crosslinked adduct. 5  $\mu$ g of bovine serum albumin (BSA) was loaded.

We sent the purified samples of PulA<sup>E71Am</sup>-His<sub>6</sub> and PulA-His<sub>6</sub> (indicated as WT, negative control) to the Institut Pasteur Proteomics platform, with the aim to compare the two samples and identify proteins specifically present only in the PulA<sup>E71Am</sup>-His<sub>6</sub> sample and absent in the negative control. The most abundant protein in the samples was PulA<sup>E71Am</sup>-His<sub>6</sub>. However, hundreds of proteins matched these two criteria. This result suggested that a gel-purification step is essential to enrich for the protein of interest before analysis by MS. In this way, we would probably increase the chance of identifying the crosslinked protein.

After we set up the experimental condition for PulA<sup>Am</sup>-His<sub>6</sub> variant pBpa incorporation and *in vivo* photo-cross-linking assays, we screened all the PulA<sup>Am</sup>-His<sub>6</sub> variants made by performing pBpa incorporation and *in vivo* site-specific photo-cross-linking assays. Among them, we detected a faint crosslinked band in variant PulA<sup>Q521Am</sup>-His<sub>6</sub> (Fig. 51). To test if the crosslinked partner is a T2SS component, we used immunodetection the PulA<sup>Q521Am</sup>-His<sub>6</sub> pBpa incorporated and UV irradiated sample individually by anti-PulC, anti-PulD, anti-PulF, anti-PulG, anti-PulH, anti-PulL, anti-PulM and anti-PulN antibodies, hoping to detect a band that is of higher Mw than these proteins alone, and comigrating with the PulA adduct as the one detected in Fig. 51. However, the result revealed a poor detection of all the Pul factors (Fig. 52). The poor detection of all the Pul factors was probably due to the poor growth of bacteria

carrying 3 plasmids in minimal medium and resulting in poor induction of *pul* genes which are typically induced in early stationary phase.

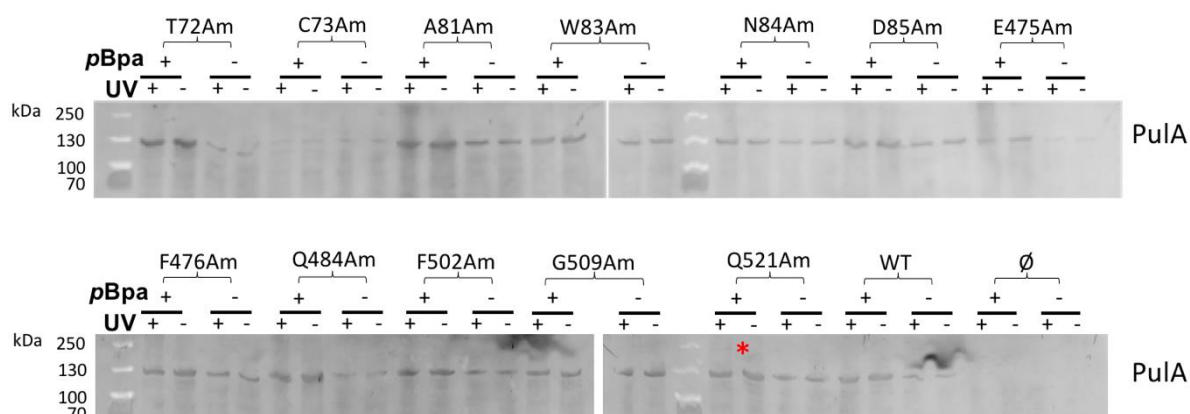


**Figure 51. *PulA*-His<sub>6</sub> *pBpa* incorporation and *in vivo* site-specific photo-cross-linking assays.**

Strain used as host was *BW25113 F'*Tet, UV irradiation time was 25 seconds, LED UV lamp. Red asterisk indicates the *PulA* crosslinked band.

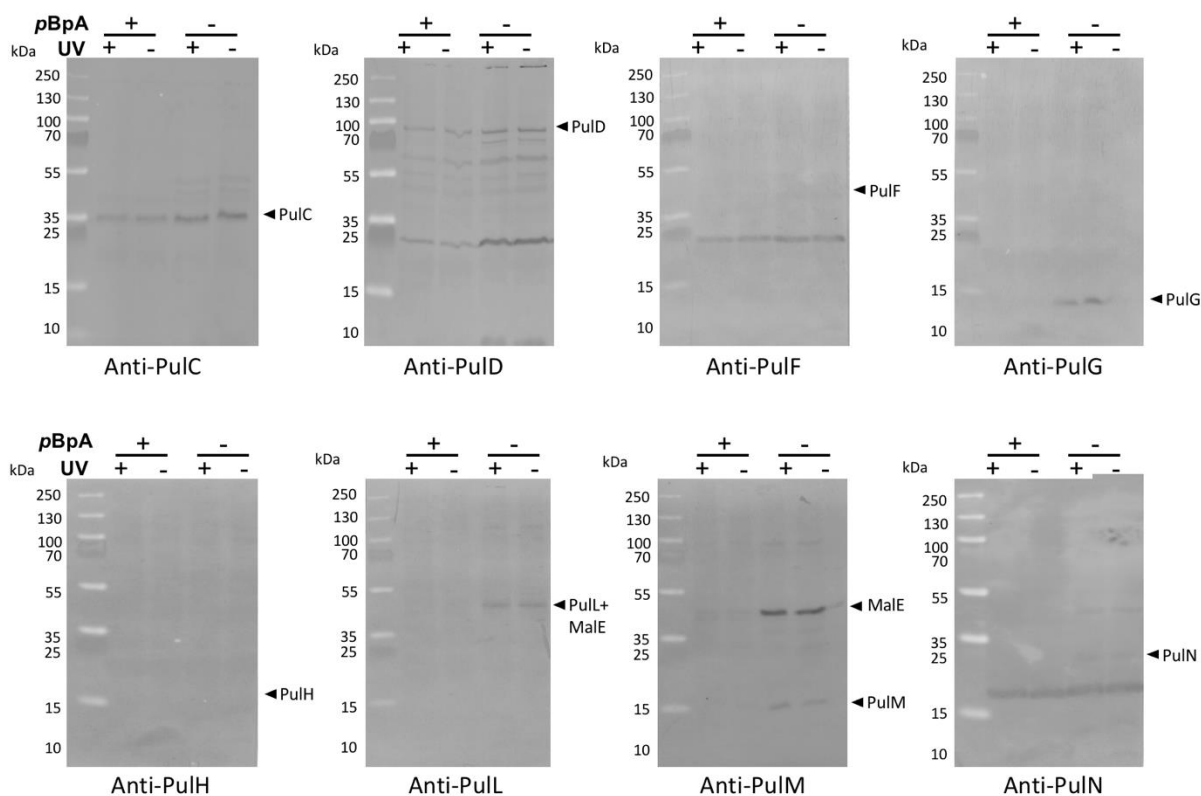


**Figure 52. Validation of *PulA*<sup>Q521Am</sup>-His<sub>6</sub> crosslinked partner by anti-Pul antibodies.**

Strain used as host was *BW25113 F'*Tet, UV irradiation time was 25 seconds with the LED UV lamp.

To concentrate the crosslinked complex from this PulA variant, we grew a 250-mL culture of *E. coli* BW25113 F'Tet harboring the three plasmids and encoding PulA<sup>Q521Am</sup>-His<sub>6</sub>, and a negative control without Pul factors. We performed *pBpa* incorporation assay for these 2 samples. Cultures with or without Pul factors were collected respectively at OD<sub>600</sub> = 1.67 and 1.8, concentrated in 10 mL of the wash buffer used for or His affinity purification (Materials and methods), and UV irradiated for 30 minutes using the Blak-Ray UV lamp that covers a larger surface. We purified PulA<sup>Q521Am</sup>-His<sub>6</sub> and the complex and analyzed them by Western blot. We compared the elution fractions of PulA<sup>Q521Am</sup>-His<sub>6</sub> with and without Pul factors, and unfortunately, we saw no difference between the 2 samples (Fig. 53).

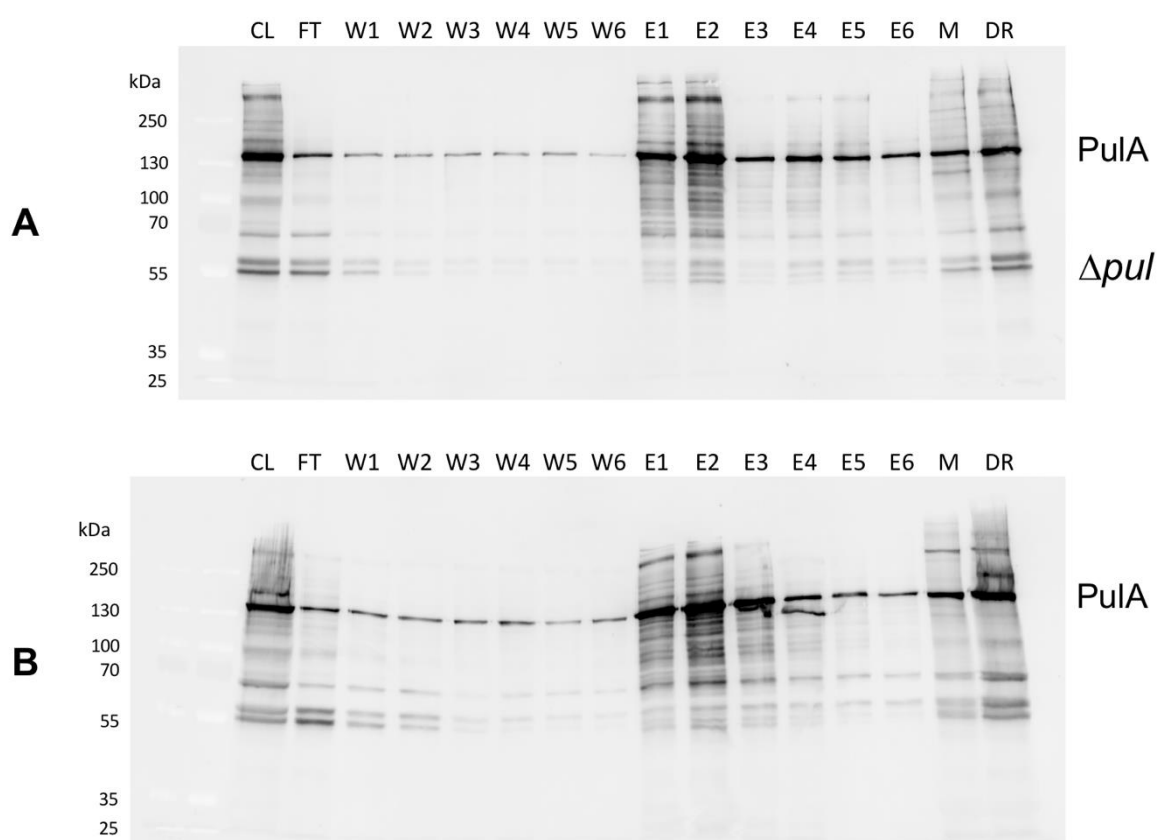


**Figure 53. Immunodetection of the purified PulA<sup>Q521Am</sup>-His<sub>6</sub> after in vivo site-specific photo-cross-linking assays.**

Strain BW25113 F'Tet was used as host. Purification after in vivo site-specific photo-cross-linking of PulA<sup>Q521Am</sup>-His<sub>6</sub> without (panel A) or with (panel B) Pul factors. CL: cleared lysate, the supernatant after ultracentrifugation loaded on the column; FT: flow-through; W1- 6: wash fractions 1- 6; E1-6: elution fractions 1-6; M: membrane pellet after ultracentrifugation; DR: cell debris after cell disruption and low-speed centrifugation. Samples were analyzed by SDS-PAGE on 8% SDS-Glycine gel, and by Western blot with anti-PulA antibodies.

So far, we were facing several problems including low efficiency of *pBPA* incorporation and cross-linking; difficulties in detection of higher Mw species and low levels of Pul proteins. We identified several possible reasons for these issues such as the low copy number of pMS5005, poor bacterial growth with 3 plasmids, potentially also high dynamics of the system and the presence of the His tag that might interfere with PulA secretion and with its interactions with T2SS proteins.

To solve these problems, we replaced the C terminal His tag on PulA by a Strep tag, since PulA-Strep seemed to be more efficiently secreted. It seemed imperative for us to reduce the plasmid load on the cells. We introduced amber codons at all the positions listed in Fig. 46 into PulA<sup>WT</sup> on plasmid pCHAP8255. In plasmid pCHAP8185, we replaced the *pulA*<sup>WT</sup> by *pulA*-Strep to give pMS5073 or its derivatives carrying *pulA*<sup>Am</sup>-Strep mutations. These changes allowed us to set up a 2-plasmid system combining the plasmid encoding the T2SS with pEVOL plasmid as a source of suppressor tRNA and tRNA synthetase (aaRS) (Fig. 54). The higher copy number of pEVOL compared to pMS5005 was expected to improve the *pBpa* incorporation.

## *E. coli*

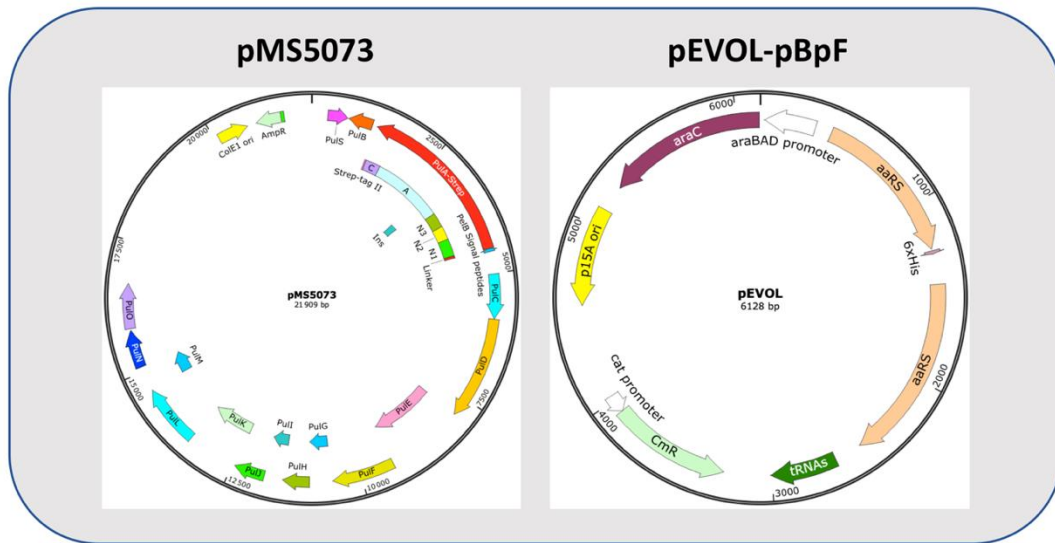


**Figure 54. The 2-plasmid system for *PulA* in vivo site-specific photo-cross-linking in *E. coli*.**

Plasmid pMS5073 carries every *pul* gene and encodes *PulA*<sup>WT</sup>-Strep or *PulA*<sup>Am</sup>-Strep variants, maltose promoters, ampicillin resistance (*Ap*<sup>R</sup>); pEVOL: encoding tRNAs and tRNA synthetase for the amber codon and *pBPA*, *araBAD* promoter, chloramphenicol resistance (*Cm*<sup>R</sup>).

Based on this 2-plasmid system, we also optimized the culture conditions, following the advice of prof. V. Shevchik. Instead of adapting bacteria to minimal medium and then growing

the precultures in M63B1, we grew the precultures in LB medium supplemented with antibiotics at 37°C. The next morning, we inoculated these precultures into M63B1 minimal medium with 0.5% glycerol, 1% casamino acids, and followed the Protocol 2 of *pBpa* incorporation (Materials and methods) followed by *in vivo* site-specific photo-cross-linking assays.

We tested several *PulA*<sup>WT</sup>-Strep variants in the context of 2-plasmid system, UV irradiated 1-mL culture of each in a 24-well plate by the UV led light, and then analyzed the samples by immunodetection with anti-*PulA* antibody. In these experiments, we saw good *pBpa* incorporation of the *PulA*<sup>Am</sup>-Strep variants and what's more, some crosslinked bands in *PulA*<sub>Strep</sub> E492, Q525, Q484, F476, E475 *pBpa* variants were detected (Fig. 55).

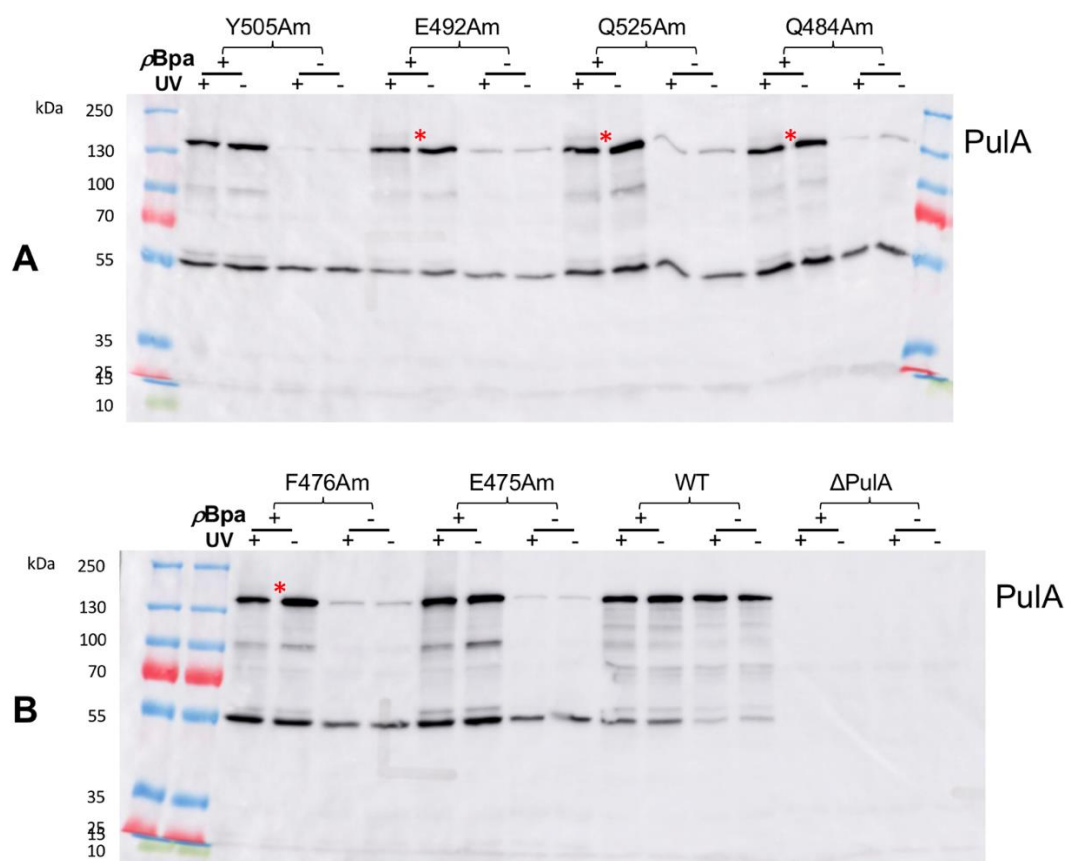


**Figure 55. *In vivo* site-specific photo-cross-linking assays with *PulA*<sup>Am</sup>-Strep variants in the 2-plasmid system.** Strain BW25113 *F'*Tet was used as a host. Total extracts of strains producing the indicated variants or *PulA*<sup>Am</sup>-Strep grown with or without *pBpa*, treated or not with UV light, were analyzed on 8% SDS-Glycine gels. UV irradiation time was 60 sec, LED UV lamp. The potential cross-linked adducts are indicated with red asterisk.

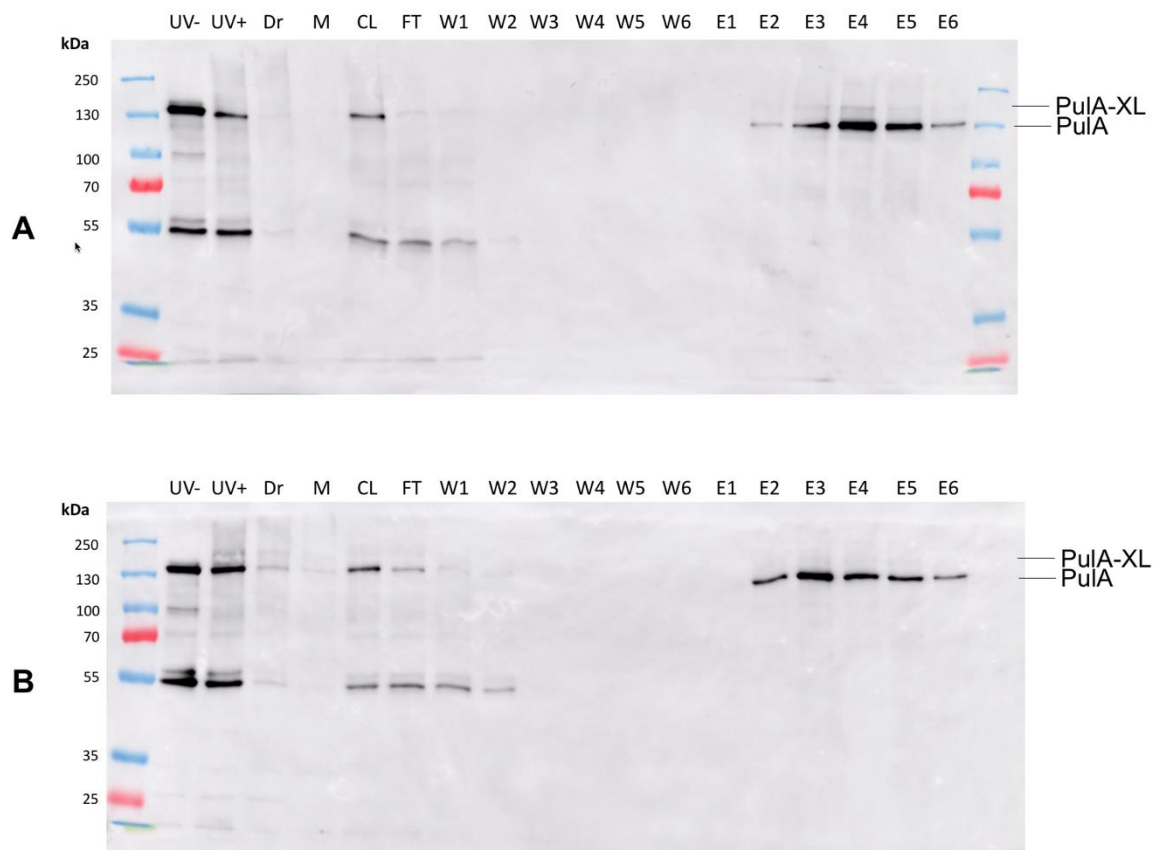


**Figure 56. Immunodetection of the purified  $Pula^{Am}$ -Strep variants after in vivo site-specific photo-cross-linking assays.**

Strain BW25113  $F'Tet$  was used as the host of plasmids  $pEVOL$  and  $pMS5073$  derivatives carrying **A:**  $Pula_{Strep}^{Q525Am}$ . **B:**  $Pula_{Strep}^{E492Am}$ . The samples were analyzed on 8% SDS-Glycine gel. Migration of  $Pula_{Strep}$ , and of the cross-linked  $Pula_{Strep-XL}$  is indicated on the right. UV-/-+: Samples of the total cell extracts before and after UV irradiation. DR: cell debris after cell disruption and centrifugation; M: membranes that were pelleted by the ultracentrifugation; CL: cleared lysate supernatant after ultracentrifugation loaded on the column; FT: flow through; W1- 6: wash fractions; E1-6: elution fractions.

We selected the  $Pula^{E492Am}$ -Strep and  $Pula^{Q525Am}$ -Strep variants to produce cultures of larger volume for purification of the Strep tagged proteins. We grew the bacterial precultures in LB medium at 30°C overnight and used them to inoculate 400 mL of M63B1 minimal medium at an  $OD_{600}$  of 0.5. We followed the modified protocol of  $pBpa$  incorporation (Materials and methods). We harvested the bacteria in 10 mL of wash buffer for Streptactin resin purification and treated them with the Blak Ray UV lamp for 30 minuets on ice for photo-cross-linking. Next, we disrupted the bacteria and purified Pula-Strep on the Strep-Tactin<sup>®</sup> XT 4Flow<sup>®</sup> resin (Materials and methods). The results of the analysis of purification fractions by Western blot with anti-Pula antibodies are shown in Fig. 56. After purification and

concentration of the proteins, we could observe a crosslinked band in the elution fraction 4 for PulA<sup>Q525Am</sup>-Strep, which was enriched after purification (Fig. 56A). However, this was not the case with the PulA<sup>E492Am</sup>-Strep samples. Compared to the total extract before purification (the “UV+” line in Fig. 56B), the elution fractions seemed to have lost the crosslinked complex. Our purification protocol was adapted for purification of soluble PulA-Strep, however this loss suggested that the protocol needs to be adapted for purification of insoluble-membrane-bound complexes. All these bands were not visible by Coomassie Blue staining, which means that the yield of PulA was too low for analysis by MS.

To concentrate the crosslinked adducts, we combined the elution fractions 3 and 4 of each sample, and performed TCA precipitation (see Materials and methods). We analyzed the proteins solubilized in Laemmli sample buffer by Western blot. As shown in Fig. 57, in the sample of PulA<sup>Q525Am</sup>-Strep, the crosslinked band became more intense compare to E4 fraction before TCA precipitation. As for sample of PulA<sup>E492Am</sup>-Strep, the band of PulA<sup>E492Am</sup>-Strep itself become even fainter than before TCA precipitation, which indicated that proteins were lost in the process. Both TCA precipitated samples showed some degradation products and were still invisible by Coomassie Blue staining, which precluded the use of MS.

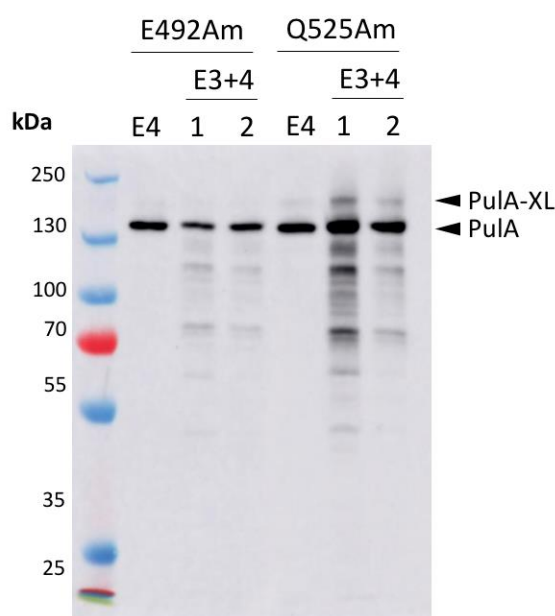


**Figure 57. TCA precipitation of the elution fractions of PulA<sup>Am</sup>-Strep variants after purification.**

The samples were analyzed on 8% SDS-Glycine gel and Western blot with anti-PulA antibodies. The Strep-tag elution fractions from Fig. 56 indicated above. E4: the elution 4 fraction before TCA precipitation, E3+4: mixture of the elution fractions 3 and 4 after TCA precipitation and solubilization.

We attempted to use Western blot to identify the crosslinked partner of PulA<sup>Q525Am</sup>-Strep. We therefore analyzed the purified PulA<sup>WT</sup>-Strep, the cleared lysate (CL), the flow-through (FT) and the wash fractions 1 and 6 from the purification process, as well as the TCA precipitated samples on SDS-PAGE and western blot (Fig 58). We probed the membranes by anti-PulA, anti-PulC, anti-PulD, anti-PulF, anti-PulG, anti-PulJ, anti-PulL, anti-PulM antibodies, however, we failed to detect any bands that comigrated with PulA<sup>Q525Am</sup>-Strep crosslinked band. Also, the BW25113 F'Tet strain contains the *malE* gene, and many of our anti-Pul factors antibodies were developed against the Pul-MalE chimera, including the anti-PulC, anti-PulL and anti-PulM. Therefore, the MalE bands were also revealed affecting the result, especially when the molecular weight of the protein is similar to that of MalE, as in the case of PulL. To overcome this issue, for the next experiments, we used the strain PAP7460 lacking the *malE* gene, which showed good *pBpa* incorporation and lacked an endogenous Amber suppressor (not shown).

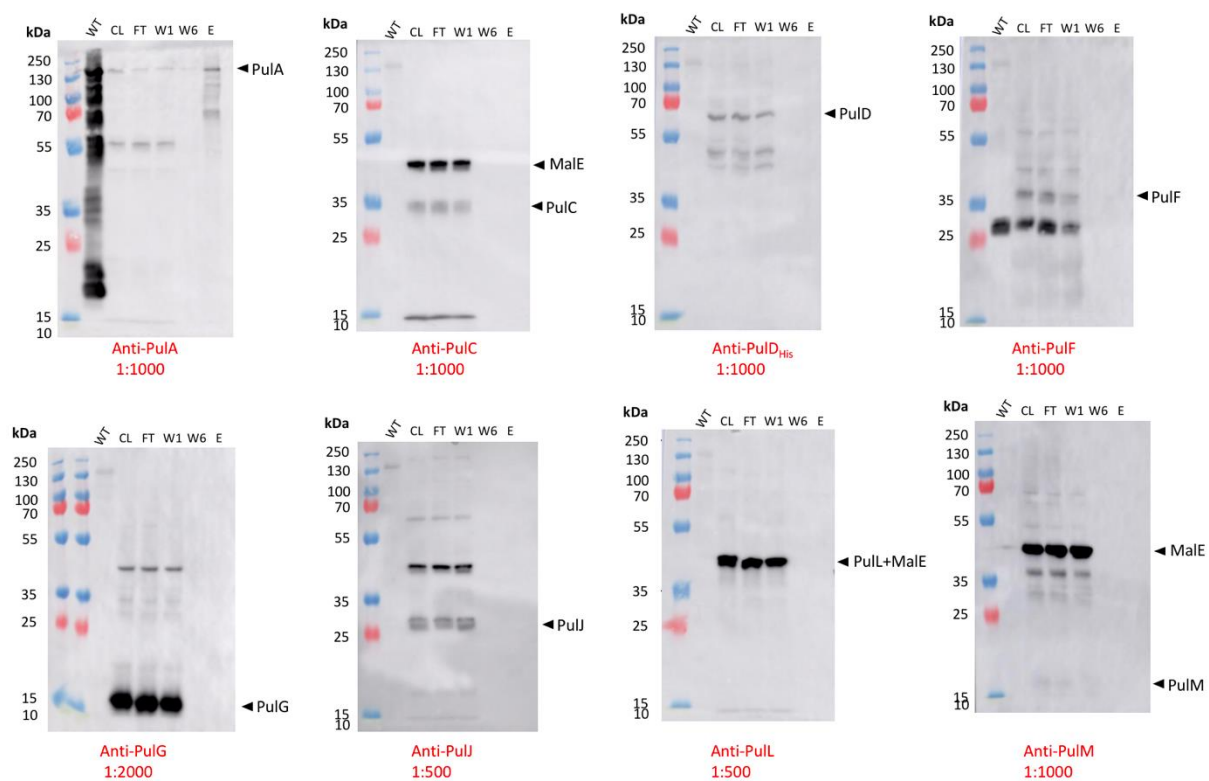


**Figure 58. Analysis of PulA<sup>Q525Am</sup>-Strep crosslinked adducts by different anti-Pul antibodies.**

The samples were analyzed on 10% SDS-Glycine gel. WT: Purified PulA<sub>Strep</sub> sample; CL: cleared lysate, the supernatant after ultracentrifugation used to load on the column; FT: flow-through; W1 and 6: wash fractions 1 and 6; E: TCA precipitated Elution fractions E3+E4.

### 3.1.1.3 Discussion of the *p*BPA crosslinking results

In all the attempts described so far, we failed to identify the interacting partners of PulA using the *p*BPA-UV crosslinking approach. Although we could generate several crosslinked adducts of some PulA-*p*BPA variants, we were unable to complete their characterization by Western blot or by mass spectrometry.

The targeted nature of this approach is one of its limitations: despite multiple *p*BPA positions that we tested, there is a possibility that we missed the optimal sites to capture the interactions by site-specific cross-linking assay. In addition, the highly dynamic secretion process might limit the fraction of PulA that interacts with a given T2SS component. This might result in very low levels of crosslinked adducts, which are below the detection limit by Western blot with our available antibodies. Although we attempted to enrich and concentrate the PulA-Strep labeled adducts, the amounts of protein were too low for analysis by mass spectrometry. In the case that PulA was crosslinked to another cellular protein, the mass spectrometry could have allowed us to identify this complex. However, the analysis of peptides generated by mass spectrometry are further complicated by the presence of *p*BPA in our proteins. The detection of *p*BPA residues by mass spectrometry is not trivial and the databases used in the analysis of spectra would have to be adapted to account for the presence of *p*BPA in the generated peptides.

We also noticed that the secretion of PulA begins as soon as the T2SS complex is formed in the *E. coli* envelope. Even at the early phase of the bacterial growth, most of the PulA has been secreted into the medium before the accumulation of sufficient bacterial mass as well as of the T2SS components. This results in lower abundance of Pul factors and consequently lower abundance of PulA complexes.

To overcome these problems, we performed several complementary approaches including formaldehyde cross-linking and co-purification with PulA-Strep and Bacterial two-hybrid assays.

### 3.1.2 Chemical crosslinking and copurification of PulA-Strep and variants

To overcome some of the limitations of the targeted approach, we turned to the chemical crosslinking followed by purification of the PulA-Strep variant. Based on plasmid pMS5073 described above, encoding PulA-Strep and all the Pul factors, we constructed its derivative, plasmid pMS5078 encoding PulA $\Delta$ Ins-Strep with Pul factors as a control. Another control derived from pMS5078 was plasmid pMS5079 encoding PulA-Strep with a deletion of *pulE* gene encoding the ATPase PulE. Then we transformed these plasmids into strain PAP7460, and performed the PulA secretion assay (Materials and methods) as shown in Fig. 59. As a negative control we included plasmid pCHAP4395 encoding PulA-Strep without the *pul* genes. Four cultures were collected at OD<sub>600nm</sub> = 2.

As the Western blot result shows (Fig. 59), there was no secretion of PulA without the Pul factors, while most of PulA-Strep was in the supernatant fraction when the Pul factors were present. We observed no secretion of PulA-Strep in the absence of PulE. As for the secretion of PulA $\Delta$ Ins-Strep, the secretion was partially defective compared to PulA-Strep (Fig. 59). This is in agreement with previous observations that Ins deletion partially affects secretion when the Pul system is overproduced (East et al., 2016).

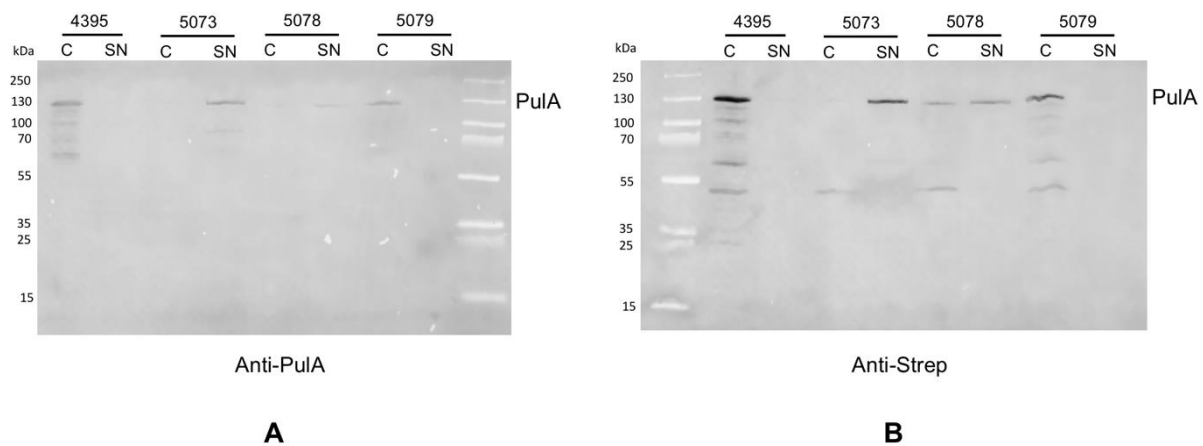


**Figure 59. Secretion assay of PulA-Strep and the effect of the Pul factors, PulE and the Ins domain of PulA.**

The same samples were analyzed by SDS-PAGE and Western blot probed by Anti-PulA or anti-Strep antibodies. 4395: plasmid pCHAP4395 encoding PulA-Strep without the T2SS Pul factors; 5073: plasmid pMS5073 encoding PulA-Strep with the Pul factors; 5078: plasmid pMS5073 encoding PulA $\Delta$ Ins -Strep with the Pul factors; 5079: plasmid pMS5079 encoding PulA-Strep and the Pul T2SS lacking PulE.

### 3.1.2.1 Defining the optimal growth stage for crosslinking

The Pul system is synthesized during late exponential and early stationary phase of bacterial growth. We aimed to capture a stage when the level of the Pul factors was high enough and where PulA was partially secreted. To find the optimal growth window to perform the crosslinking assay, we transformed pMS5073 into PAP7460 strain, grew the bacteria in LB medium with 0.4% maltose, collected cultures at different OD<sub>600</sub>, and performed secretion assays. We probed the same membrane individually by anti-PulA and anti-PulM, to monitor the PulA secretion and the Pul factor accumulation. In the western blot result (Fig. 60), we could see that at the OD<sub>600</sub> = 1.0, there was almost no PulA or PulM synthesized. At OD<sub>600</sub> = 1.4, the PulA and PulM were produced, surprisingly, the majority of PulA was already secreted into the medium at such an early stage. At OD<sub>600</sub> = 1.7, more PulM accumulated in the cell, and only a small portion of PulA remained in the cell fraction. When the OD<sub>600</sub> reached 2.5, the amount of PulM was the highest, but almost all the PulA has been secreted. These results suggested that PulA is secreted rapidly and as soon as the T2SS complex is assembled in the membrane, and that there was not a “perfect” window in which the Pul factors were produced at sufficient levels while the PulA still remained inside the cell. We decided that the most appropriate window to perform crosslinking assays could be around OD<sub>600</sub> = 1.5. Bacteria growing beyond this stage continued to produce PulA and PulM suggesting that the secretion process would continue as well, allowing us to capture at least some intracellular secretion intermediates.

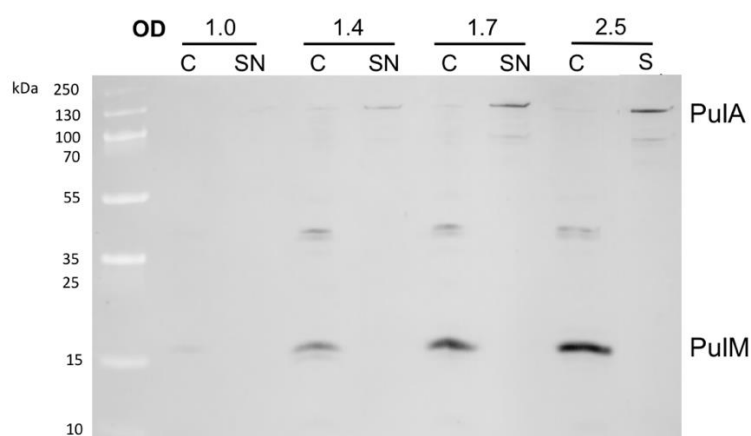


**Figure 60. PulA-Strep secretion and PulM levels at different growth stages.**

Bacteria of strain PAP7460 harboring plasmid pMS5073 were inoculated in LB medium containing 0.4 % D-maltose and Ap and incubated at 30°C with shaking at 200 rpm. Samples were withdrawn when the cultures reached OD<sub>600nm</sub> indicated above, and secretion assays were performed (Materials and methods). The samples were analyzed on 8% Tricine gel and Western blot was performed using anti PulM and anti-PulA antibodies.

### 3.1.2.2 Strep-protein interaction experiment (SPINE)

The method of Strep-protein capture also called SPINE was first reported by Herzberg et al (Herzberg et al., 2007). It combines *in vivo* formaldehyde crosslinking and Strep-protein purification of the target protein and its crosslinked products. The protocol we used here is detailed in Materials and methods. As a crosslinking agent, we used formaldehyde which penetrates the bacteria and cross-links protein in close contact (at a distance of about 2 Å), to allow capturing of direct protein-protein interactions. Formaldehyde crosslinks are also reversible by heating, facilitating the identification of crosslinked proteins by Western blot.

I first optimized the experimental conditions – the concentration of formaldehyde and incubation time, by transforming pMS5073 (PulA-Strep, Pul+) and pMS5079 (PulA-Strep  $\Delta$ pulE) individually into strain PAP7460. I performed SPINE in small volume of cultures grown in minimal medium. I treated samples with 0.6% formaldehyde for different times after adding it directly in the culture medium (Fig. 61A). The treatment was always performed at 30°C, the temperature that was also used for the bacterial growth. As there was no difference between 20, 30, 40 min treatments, we increased the concentration to 0.8% of formaldehyde and treated for 30 min (Fig. 61B). Samples were not boiled before analysis by for SDS-PAGE and western blot, so the crosslinked complexes did not dissociate. We observed two major forms of monomeric PulA, the fully denatured labeled as PulA and the faster migrating PulA', which is probably partially folded. In the formaldehyde treated samples, the PulA bands were fainter than without formaldehyde treatment, and the smear above the PulA bands indicated that some crosslinking complex could have formed. When we used 0.8% formaldehyde and 30 minutes of incubation, the majority of PulA was still monomeric, showing that these conditions were not too harsh – we therefore used them for the future SPINE.

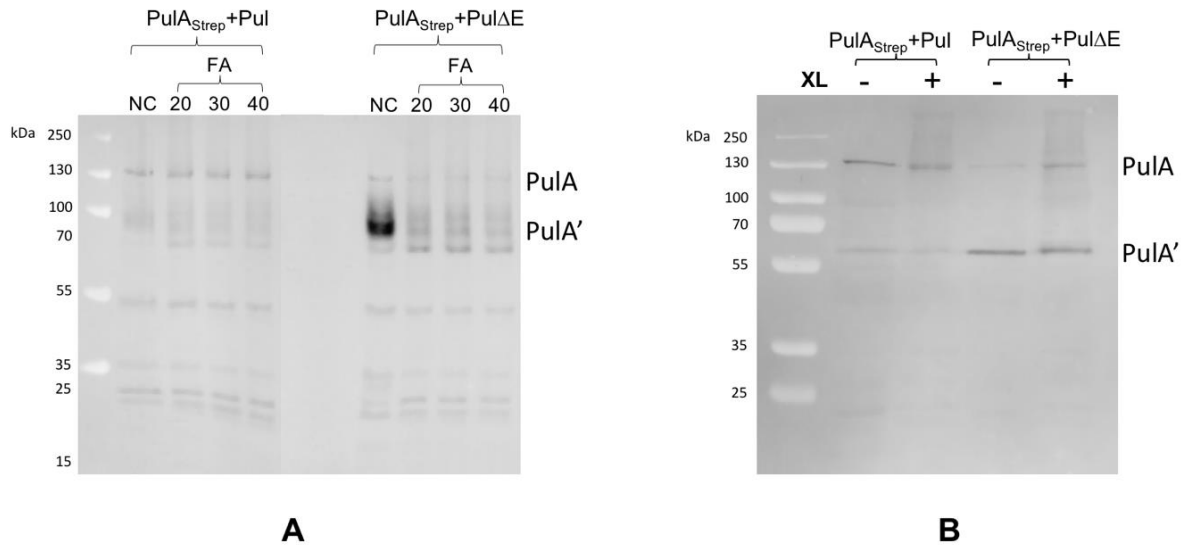


**Figure 61. Formaldehyde crosslinking of PulA-Strep + Pul and PulA-Strep + PulΔE.**

**A.** Treatment with 0.6% formaldehyde, NC: negative control without formaldehyde; 20, 30 and 40 indicate the formaldehyde incubation time (min). The samples were analyzed by SDS-PAGE on 8% Tricine gel and by Western blot with anti-PulA antibodies. **B.** Treatment with 0.8% formaldehyde for 30 min; XL: with (+) or without (-) formaldehyde crosslinking. The samples were analyzed by SDS-PAGE on a 8% SDS-Glycine gel and by Western blot with anti-PulA antibodies.

For the first SPINE, we cultured 200 mL volume of each strain PAP7460 (pMS5073) and PAP7460 (pMS5079). To the culture PAP7460 (pMS5073) (at  $OD_{600}=1.44$ ) and PAP7460 (pMS5079) (at  $OD_{600}=1.25$ ), we added 0.8% formaldehyde and incubated for 30 min. We purified the Strep-proteins by Streptactin resin (IBA). The purified proteins were not visible by Coomassie blue staining, which, together with the Western blot analysis, indicated that the yield of PulA-Strep was quite low (Fig. 62). This was especially the case for the samples PAP7460 (pMS5073), where only a small amount of PulA was detected in the clear lysate fraction (CL). This might be due to the fact that the majority of PulA has been secreted to the culture supernatant. In addition, the presence of PulA in the flow-through (FT) and the wash fractions suggested a low binding of PulA-Strep to the Streptactin resin. Moreover, a large amount of PulA remained in the membrane fractions of each culture (Fig. 62).

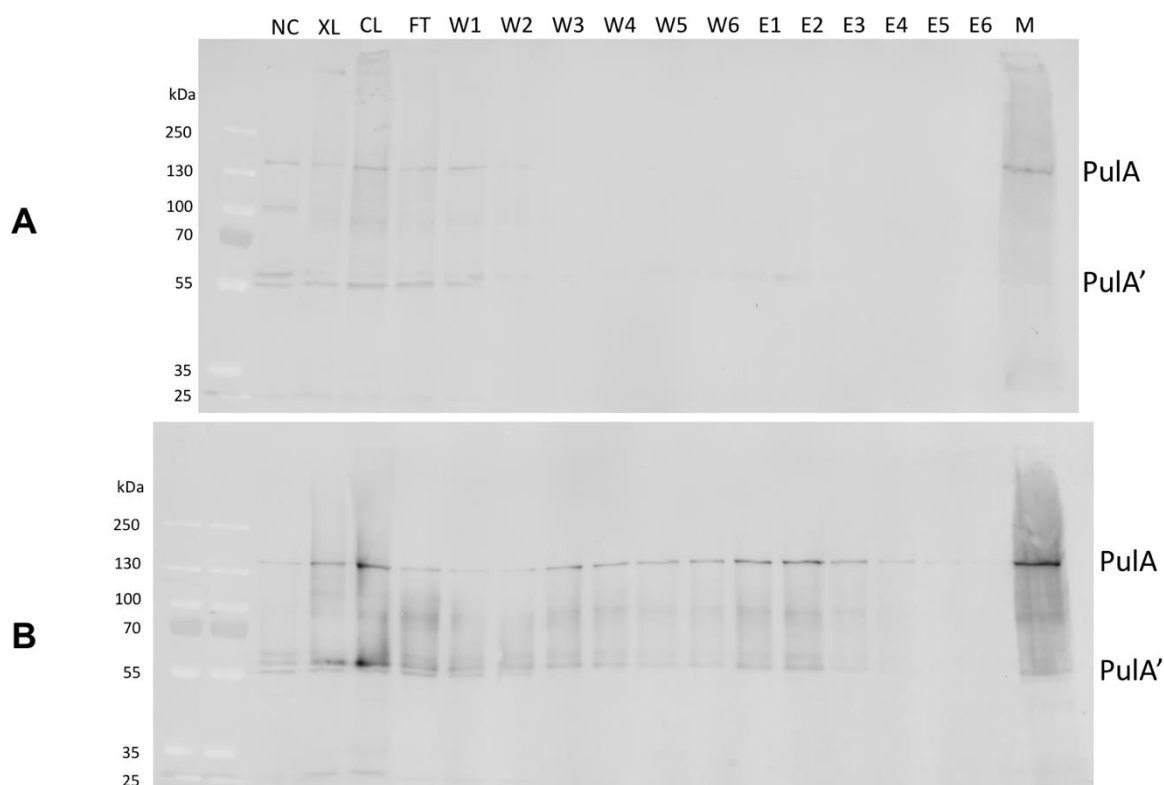


**Figure 62. SPINE with PulA-Strep + Pul (A) and PulA-Strep + PulΔpulE (B).**

Samples were treated with 0.8% formaldehyde for 30 min, followed by purification on Streptactin resin. Samples (without boiling) were analyzed on 8% glycine SDS-PAGE and Western blot with anti-PulA antibodies. NC: negative control that sample without formaldehyde; XL: sample after formaldehyde treatment; CL: cleared lysate, supernatant after cell lysis and ultracentrifugation; FT: flow-through; W1-6: wash fractions 1-6; E1-6: elution fractions 1-6; M: membrane pellet after ultracentrifugation.

To solve the binding affinity issue, we purchased a new Streptactin resin (Strep-Tactin<sup>®</sup>XT 4Flow<sup>®</sup> resin) for PulA-Strep purification. We further increased the formaldehyde concentration to 1%. We cultured 300 mL of PAP7460 (pMS5079) in M63B1 medium, and added 1% formaldehyde to the culture at OD<sub>600</sub> = 1.7 for 30 min at 30°C. The proteins purified on a Strep-Tactin<sup>®</sup>XT 4Flow<sup>®</sup> resin were loaded on 8% SDS-Glycine gel without boiling. On western blot, intense bands corresponding to PulA were present in samples before and after formaldehyde crosslinking (NC and XL), and also in cleared lysate (CL) and elution fractions. Bands corresponding to PulA' were enriched in elution fractions, however. No bands migrating slower than PulA were detected. This might because there were no crosslinked proteins, or the because the crosslinked PulA was present in the PulA' form so that the crosslinked bands were present in the smear between PulA and PulA' (Fig. 63A). By Coomassie blue staining, no bands

corresponding to denatured PulA were visible, and bands co-migrating with PulA' were observed in CL, FT, W1-2 and E2-6 fractions (Fig. 63B).

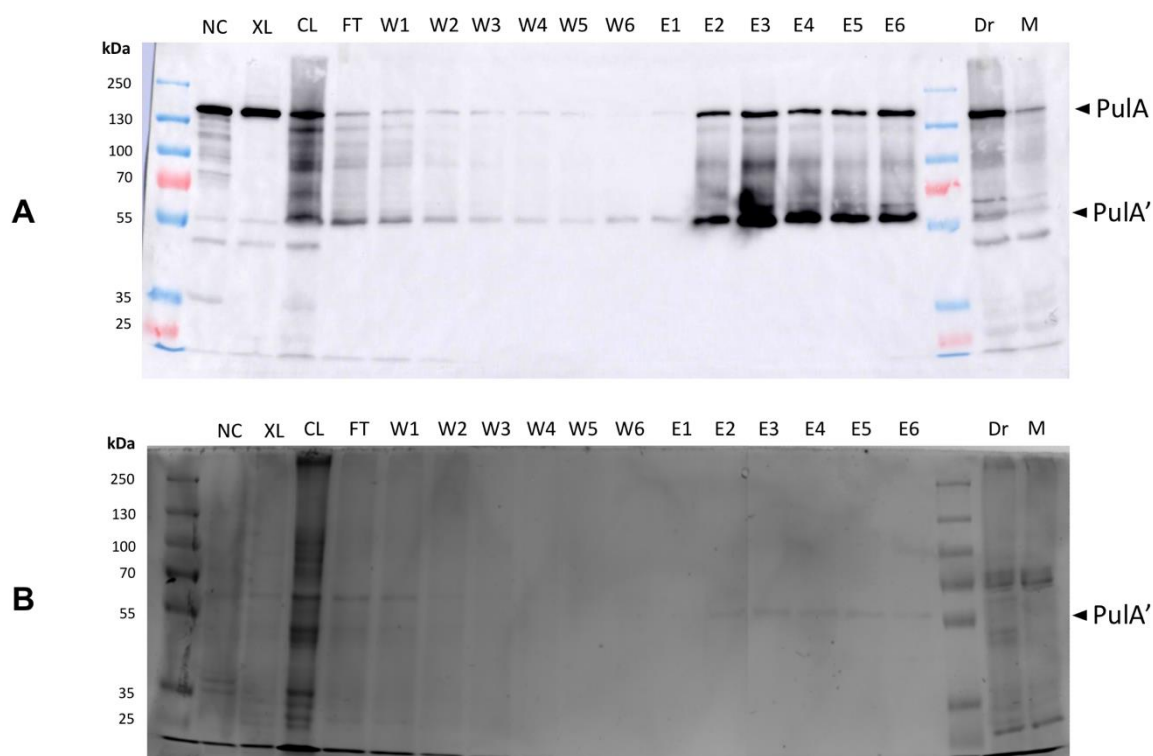


**Figure 63. SPINE of PulA-Strep + Pul $\Delta$ E.**

Samples (without boiling) were analyzed of 8% SDS-Glycine gel **A**: Western blot with anti-PulA antibodies. **B**: Coomassie blue staining of the same samples. NC: negative control without formaldehyde incubation; XL: sample after formaldehyde incubation; CL: cleared lysate after cell lysis and ultracentrifugation; FT: flow-through; W1-6: wash fractions 1-6; E1-6: elution fractions 1-6; Dr: cell debris pellet after cell lysis and centrifugation at 7000 rpm; M: membrane pellet after ultracentrifugation.

Overall, these results showed that PulA-Strep accumulated in the cells due to the absence of PulE could be successfully purified with the improved resin Strep-Tactin®XT 4Flow®. To test if any Pul proteins were crosslinked to PulA-Strep, I tested the purified PulA-Strep (from plasmid pCHAP4395), samples NC, CL, FT, W1, W6 (as negative controls) and E3 fraction (where PulA-Strep was the most abundant), by immunodetection with anti-PulA antibodies, as well as with antibodies against PulC, D, F, G, J, L and M. Samples were boiled for 20 minutes to dissociate the crosslinked complexes and to denature the PulA'. As shown in Fig. 64, PulA was present in every fraction, whereas PulC, D, F, G, J, L and M were only detected in NC, CL, FT and W1 fractions. We did not detect any of the Pul proteins in the elution 3 fraction.

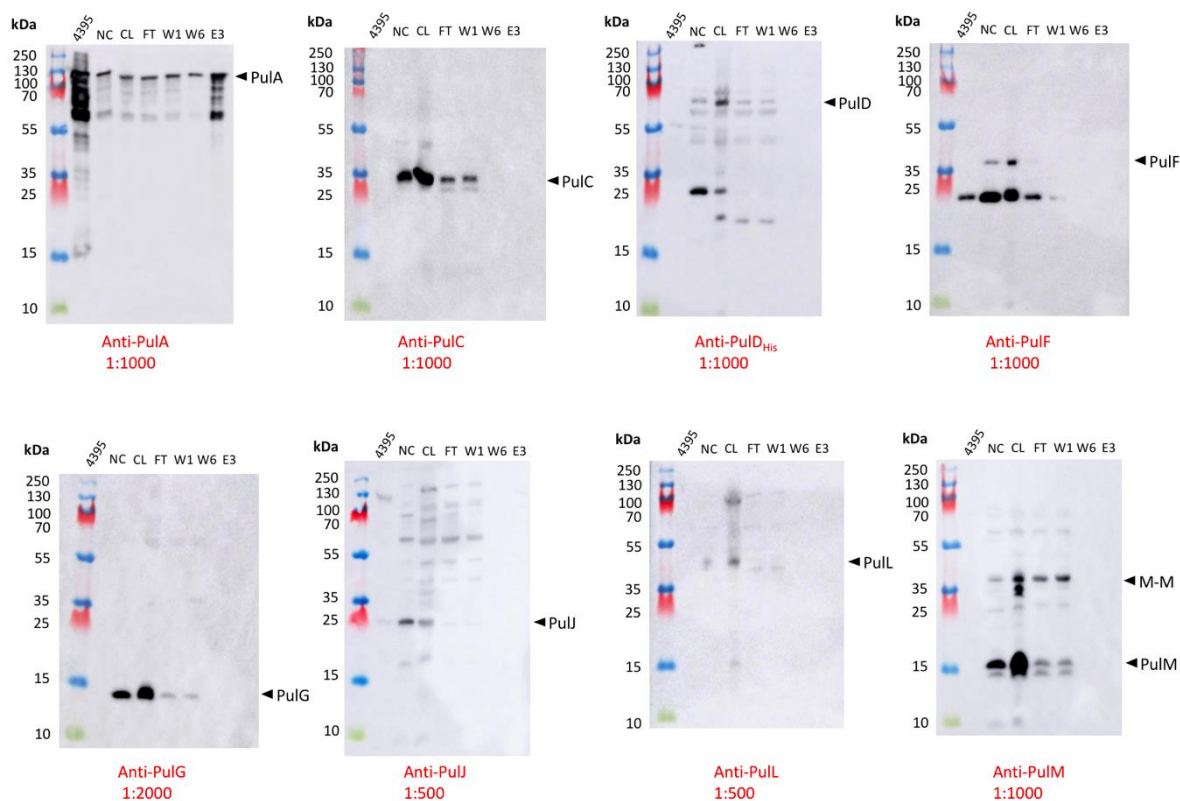


**Figure 64. SPINE of Pula-Strep + Pul $\Delta$ E probed by different anti-Pul antibodies.**

The boiled samples were analyzed by SDS-PAGE on 10% Tricine gel and Western blot with antibodies indicated below each panel. 4395: purified Pula-Strep from plasmid pCHAP4395; NC: negative control without formaldehyde incubation; XL: sample after formaldehyde incubation; CL: cleared lysate, after cell lysis and ultracentrifugation; FT: flow-through; W1/W6: wash fractions 1 and 6; E3: elution fractions 3.

### 3.1.2.3 Discussion of the SPINE results

Unlike the pBPA crosslinking, the SPINE approach has the potential to capture all possible interactions of Pula with the T2SS complex. Here I described preliminary experiments aimed at setting up SPINE for the Pul T2SS. The main issue we observed here was the absence of Pula in the cell fraction of bacteria containing the functional Pul system. Even in the early growth stages, most of the Pula was secreted in the culture supernatant and lost. By using a  $\Delta$ pulE mutant we increased the yield of intracellular Pula. We hypothesized that in this way we would still be able to capture early secretion intermediates and that, theoretically, Pula should bind at least one Pul factor prior to the steps catalyzed by the PulE ATPase. The fact that we failed to detect any complexes with the Pul factors might be due to low levels of such complexes. In addition, the presence of Pula in the membrane fraction (Fig. 62, 63) suggests that this Pula might be bound to some membrane proteins. Since the majority of Pul

components are in the inner or the outer membrane, it is possible that the crosslinked adducts of interest can be found in this fraction. To capture such complexes, it would be important to treat the cleared lysate – the soluble supernatant after cell lysis and before ultracentrifugation - with detergent (e.g. 1% Triton), and perform the purification steps in the presence of detergent.

To improve the yield of intracellular PulA in the cells containing the functional Pul system, it would be important to slow down or block secretion. For example, we could use mutants of the PulE ATPase that are partially functional and show reduced ATPase activity. It would also be interesting to use chemical treatments that slow down or block secretion for example by blocking the proton motive force. Possot et al have shown that adding CCCP to the culture rapidly reduces the amount of secreted PulA (Possot et al., 1997).

### 3.1.3 Bacterial adenylate cyclase two-hybrid (BACTH) study of PulA interactions

As discussed above, we hypothesized that PulA binds to one of the inner membrane components of the Pul T2SS at an early stage of secretion. This interaction might be involved in the recruitment of T2SS substrates, as has been proposed for GspL or GspM in the *Pseudomonas* T2SS (Michel-Souzy et al., 2018) on this assumption, we decided to use BACTH as a complementary approach to study interaction between PulA and T2SS components (Karimova et al., 1998a). Prior to its secretion, the native PulA is an inner membrane bound lipoprotein. This IM localization is necessary for its proper secretion (Pugsley & Kornacker, 1991). The BACTH assay relies on cytoplasmic signaling molecule cAMP, which can be produced in response to interactions of tested proteins in the inner membrane or even in the periplasm. To allow for signaling with the periplasmic PulA, we fused it to a neutral transmembrane segment of MalF protein (Guzman et al., 1997) or to the segment of the PulK minor pseudopilin.

#### 3.1.3.1 PulA interactions with the pseudopilins

We first performed BACTH assay to determine whether PulA fused to the MalF<sup>TMS</sup> interacts with the major pseudopilin PulG or with the minor pseudopilins PulH, PulI, PulJ and PulK. The two models of T2SS secretion mechanism predict interactions of the substrate with the pseudopilus (Nivaskumar & Francetic, 2014a). In the piston model, the tip of the

pseudopilus comprising the minor pseudopilin complex is predicted to interact with the exoprotein substrate during secretion. In the Archimedes' screw model, the specific interaction would rather involve the major pseudopilin.

We fused PulA-MalF<sup>TMS</sup> to the C-terminus of T18 or T25 CyaA fragments and co-transformed strain DHT1 ( $\Delta cya$ ) with the resulting plasmids and with the compatible plasmids encoding T18- and T25 fused to full-length mature pseudopilins. We measured the  $\beta$ -galactosidase activity levels and found that T18-MalF<sup>TMS</sup>-PulA interacts with PulG, PulH, PulJ and PulK significantly compared to the negative control. With the T25-MalF<sup>TMS</sup>-PulA construct, the interactions of PulA with each pseudopilin hybrid was significantly higher than the negative control (Fig. 65).

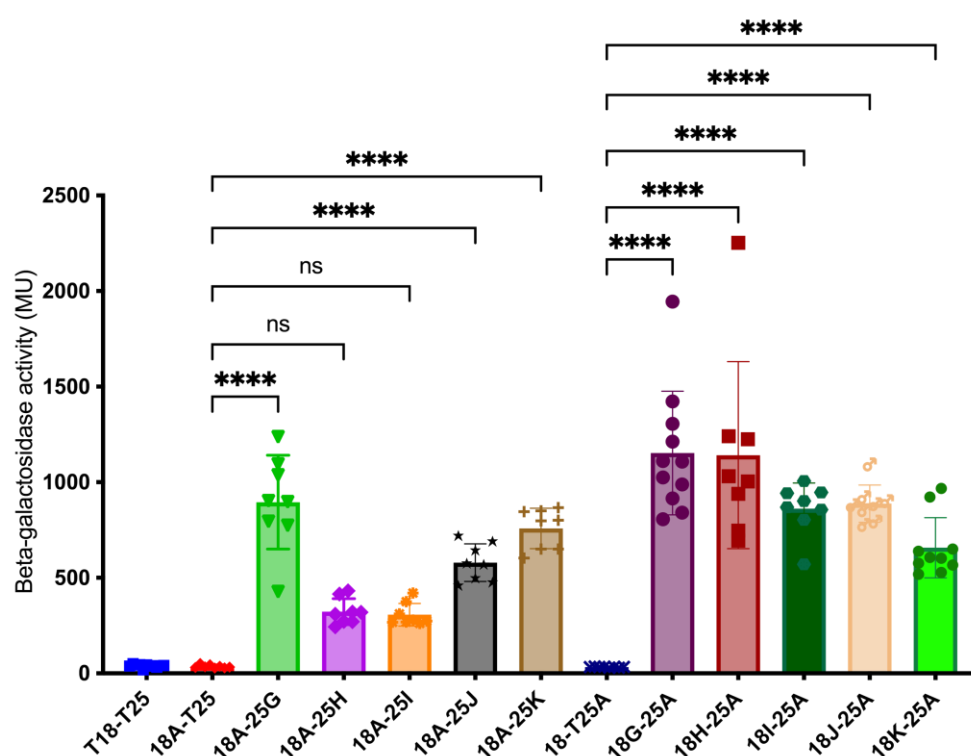


**Figure 65. BACTH analysis of PulA-MalF<sup>TMS</sup> with pseudopilins.**

The  $\beta$ -galactosidase activities measured for 6 or more independent bacterial cultures. Individual values are shown as dots. Bar graphs indicate mean values and error bars standard deviation. Statistical analysis was performed using One-way ANOVA and multiple comparisons test, and plotted using GraphPad Prism 9 software. ns, no significant difference; \*,  $p < 0.05$ ; \*\*,  $p < 0.01$ ; \*\*\*,  $p < 0.001$ ; \*\*\*\*,  $p < 0.0001$ .

The BACTH data showed the strongest interaction signals between PulA and PulG of the highest significance in both directions. To further validate this contact, we compared PulG interaction with PulA, with the PulA domain N1 alone and with PulA $\Delta$ Ins variant encoded by pUT18C or pKT25 constructs (Fig. 66).

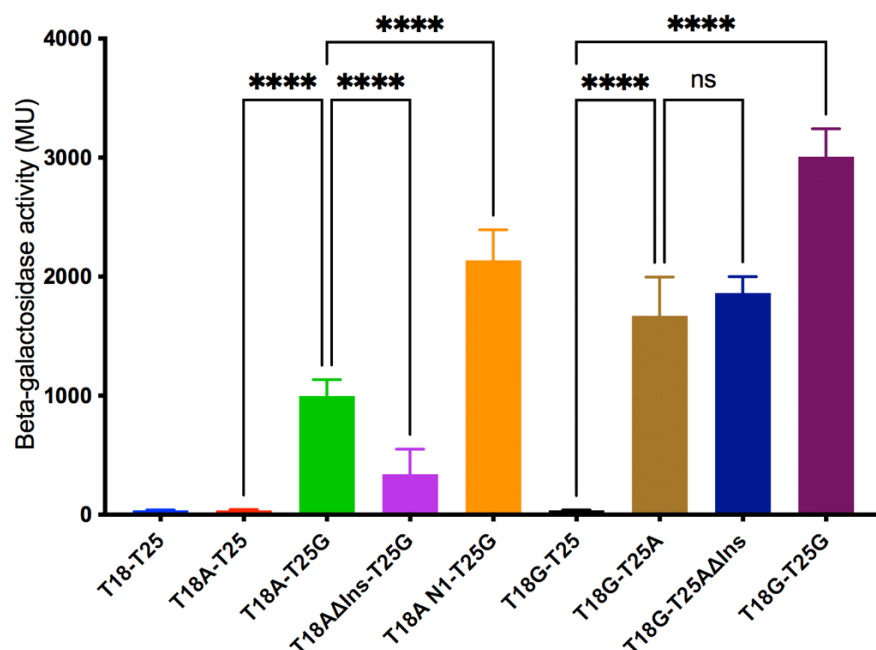


**Figure 66. BACTH analysis of PulG interaction with PulA, N1 domain and PulAΔIns.**

The  $\beta$ -galactosidase activities were measured for at least 4 independent bacterial cultures. The bar graphs show mean values and error bars indicate standard deviation. Statistical analysis was performed using One-way ANOVA and multiple comparisons test, and plotted using GraphPad Prism 9 software. ns, no significant difference; \*\*\*\*,  $p < 0.0001$ .

As shown in Fig. 66, compared to the negative control, interactions between PulA and PulG were highly significant with both combinations of constructs. However, interaction between T18-N1 and T25-PulG was significantly higher compared to that of T18-PulA and T25-PulG, or T18-PulAΔIns with T25-PulG. This is surprising, given that T18-PulA and T18-PulAΔIns also contain the N1 domain. The activity of bacteria producing T18-PulAΔIns with T25-PulG was significantly lower than T18-PulA and T25-PulG, suggesting the importance of the Ins domain for PulA – PulG interaction. However, when the complementary constructs were used, the interaction of T25-PulAΔIns with T18-PulG was as strong as that of T25-PulA with T18-PulG. As shown previously, PulG forms homodimers in this assay (Nivaskumar et al., 2014). Together, these results confirmed the PulA – PulG interaction, however the contributions of N1 and Ins domains remained unclear.

### 3.1.3.2 PulA interactions with inner membrane Pul T2SS proteins

To test whether the transmembrane segment of MalF contributes to PulA interactions in the BACTH assays, we replaced it with the transmembrane segment of PulK. We tested the

interactions of T18/T25-PulK<sup>TMS</sup>-PulA hybrids with the corresponding hybrids of PulB, PulC, PulF, PulL and PulM. The BACTH plasmids (listed in Table 2) were introduced in the *E. coli* strain DHT1 and  $\beta$ -galactosidase activity of at least 6 independent clones was measured. The results are summarized in Fig. 67. The CyaA T18 and T25 fragments, used as a negative control, do not interact and do not restore adenyl cyclase activity. Compared to the negative control, we found that the PulA interaction signals with PulM and PulF were strong and highly significant in both combinations of constructs (Fig. 67).

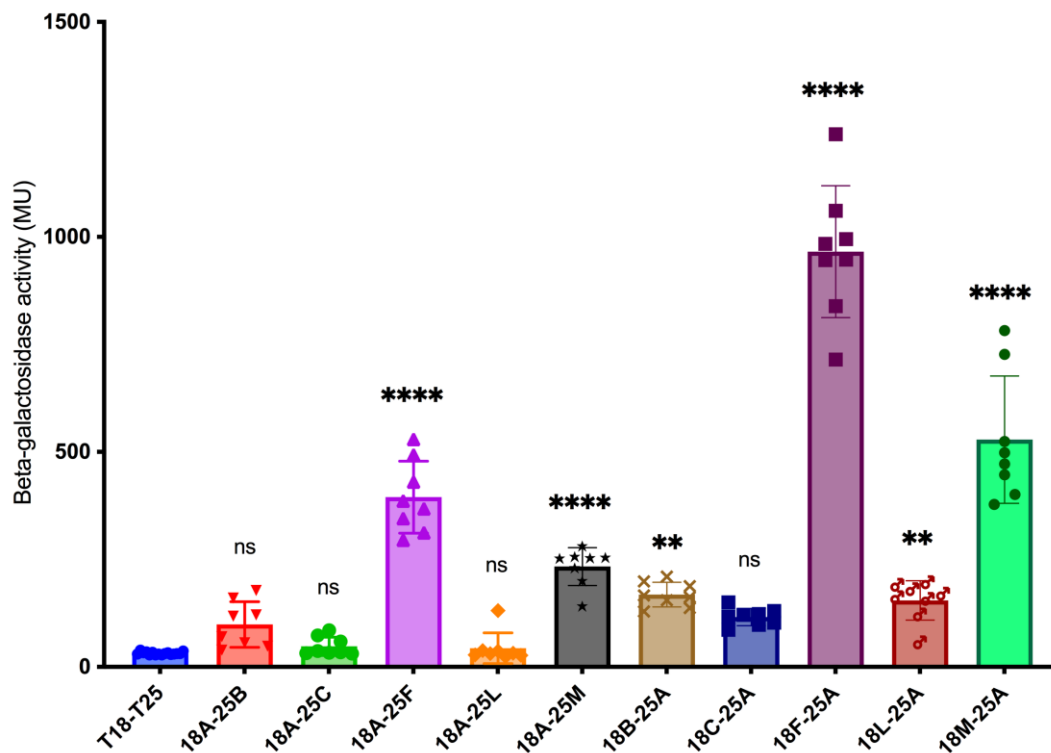


**Figure 67. BACTH analysis of PulA interactions with inner membrane T2SS components.**

The  $\beta$ -galactosidase activities measured for at least 6 independent bacterial cultures. Bar graphs indicate the mean values and error bars show standard deviation. Statistical analysis was performed using One-way ANOVA and multiple comparisons test, and plotted using GraphPad Prism 9 software. All samples were compared with the T18-T25 negative control and the difference is indicated above each column: ns, no significant difference; \*,  $p < 0.05$ ; \*\*,  $p < 0.01$ ; \*\*\*\*,  $p < 0.0001$ .

We were surprised to detect interaction of PulA with PulF, which only has a small portion exposed in the periplasm. We suspected that the TMS of PulK, and not PulA, might be interacting with PulF. Indeed, the BACTH assays showed that T18-PulK<sup>TMS</sup> interacts with T25-PulF, whereas the T18-MalF<sup>TMS</sup> did not (Fig. 68). Therefore, we tested the interactions of T18/T25- MalF<sup>TMS</sup>-PulA hybrids with PulF in the BACTH assay. In the context of MalF<sup>TMS</sup>-

PulA hybrids, the interaction between PulA and PulF was significantly stronger than that of PulF with the negative control MalF<sup>TMS</sup> (Fig. 68).

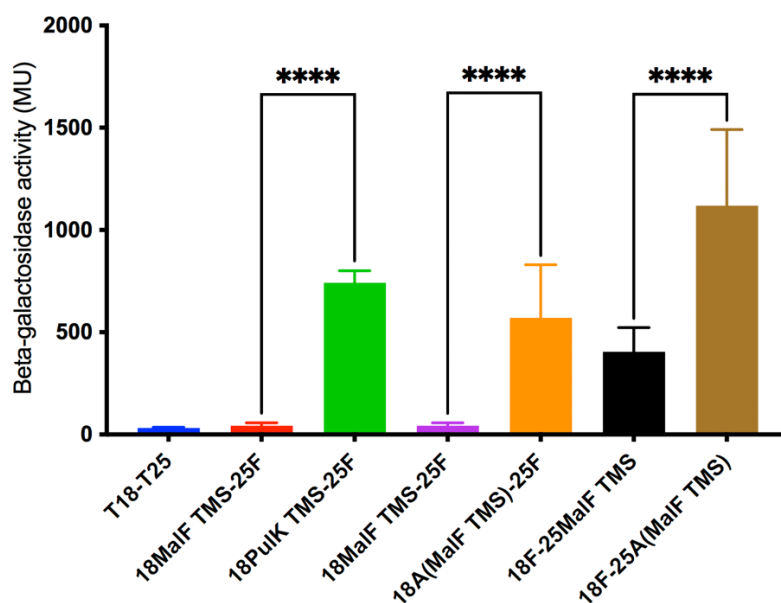


**Figure 68. BACTH analysis of PulA (MalF TMS) with PulF, and PulF with PulK/MalF-TMS interactions.**

The  $\beta$ -galactosidase activities measured for independent bacterial cultures, and the horizontal lines above its corresponding columns show median values. Statistical analysis was performed using One-way ANOVA and multiple comparisons test, and plotted using GraphPad Prism 9 software. ns, no significant difference; \*\*\*\*,  $p < 0.0001$ .

In our lab, Ingrid Guilvout has been focusing on PulF, and she made multiple mutant variants to study PulF function. She isolated two PulF-Flag insertion variants, PulF<sup>L198Flag</sup> and PulF<sup>A201Flag</sup>, which almost abolished the PulA secretion but still promoted piliation (not shown). These Flag insertions map in the periplasmic loop of PulF and we assumed that the PulF L198 and A201 residues might play important roles in interaction with the substrate PulA. Thus, we fused these two variants with T18 and T25 Cya fragments and tested their interactions with PulA (MalF TMS). The interactions between PulA and PulF<sup>L198Flag</sup> or PulF<sup>A201Flag</sup> were all significantly lower compared with that of PulA and PulF<sup>WT</sup>. However, these Flag insertions also affected PulF dimerization (Fig. 69).

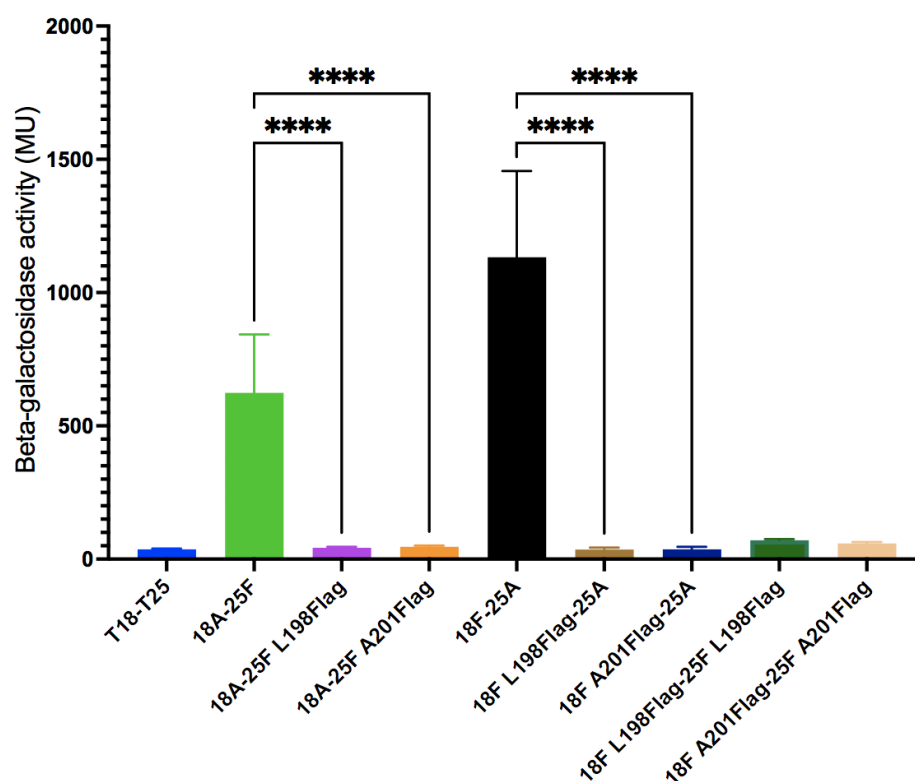


**Figure 69. BACTH analysis of PulA (MalF TMS) with PulF, PulF<sup>L198Flag</sup> and PulF<sup>A201Flag</sup>.**

The  $\beta$ -galactosidase activities measured for at least 4 independent bacterial cultures, and the horizontal lines above its corresponding columns show median values. Statistical analysis was performed using One-way ANOVA and multiple comparisons test, and plotted using GraphPad Prism 9 software. ns, no significant difference; \*\*\*\*,  $p < 0.0001$ .

### 3.1.3.3 Discussion of the BACTH results

The results of BACTH assays identified several strong candidates for PulA interaction partners including the major pseudopilin PulG, and the AP components PulM and PulF. However, other pseudopilins also showed a tendency to bind PulA, depending on the copy number of plasmids encoding the Cya hybrids and protein stability. It is unclear whether the interaction of individual proteins tethered in the membrane can reflect the situation in the functional T2SS. These data need to be confirmed by complementary approaches, e.g. biochemical – copurification or by microscopy – including electron microscopy and fluorescence microscopy. Our collaborators N. Izadi and her team were able to show binding of PulA to pseudopili using biophysical techniques – the biolayer interferometry and co-sedimentation. In the future, chemical crosslinking and mass spectrometry will be combined to identify contacts between PulA and the pseudopili.

Evidence for an interaction of PulA with PulM has also been obtained by Aracelys Lopez Castilla in the team of N. Izadi using surface plasmon resonance. Furthermore, interactions of the *P. aeruginosa* T2SS substrate with XcpL and XcpM have been observed (Michel-Souzy et al., 2018). Further *in vitro* and *in vivo* studies would be required to test whether PulM plays a role in coupling PulA binding with pseudopilus assembly. PulM also interacts with PulL (as detailed in the Chapter 3.2) and with the pseudopilins PulG and PulH (Nivaskumar et al., 2016). These interactions would favor the role of PulL-PulM complex in coupling pseudopilus assembly to substrate loading.

According to the BACTH results, we could see a strong interaction of PulA with PulF. The Flag insertions at positions L198 and A201 are interesting and support the role of periplasmic PulF region in PulA binding. In the future, it would be interesting to study this contact, possibly by the UV crosslinking approach incorporating *pBPA* in this region of PulF.

### 3.2 Analysis of the assembly platform components PulL and PulM

The inner membrane assembly platform composed of GspF, GspL and GspM and their orthologues plays a central role in promoting the function of the T2SS and T4P machineries. So far, the structures of individual domains of most of the AP components have been determined, however, the knowledge on the full-length components and their interactions is still very limited. The co-purification of the entire AP complex has not been accomplished yet, probably due to the weak or transient interactions, notably between their hydrophobic domains. The purification of the *K. pneumoniae* T2SS complex was an important achievement, however PulF and the pseudopilins were not copurified with the rest (Chernyatina & Low, 2019).

Several previous studies of the AP components GspL and GspM in different T2SS models showed that they interact relatively strongly with each other, and act together, which motivated our structure - function studies of this subcomplex. Crystal structures of the periplasmic domains of EpsM and EpsL from the Type II Secretion System of *V. parahaemolyticus* have been determined by Jan Abendroth et al. in 2004 and 2009 (Abendroth, Rice, et al., 2004, p. 2009), as well as the structure of the cytoplasmic domain of EpsL alone (Abendroth, Bagdasarian, et al., 2004) and in complex with the ATPase EpsE N domain (Abendroth et al., 2005). Predictions and topology studies of GspL and GspM suggest that both contain a single transmembrane helix connected to a long helix and periplasmic domains (GspL<sub>peri</sub> and GspM<sub>peri</sub>), which adopt a similar ferredoxin-like fold. The GspL<sub>peri</sub> and GspM<sub>peri</sub> form homodimers in the crystals although different interfaces were involved in dimerization. Meanwhile, biochemical and *in vivo* studies using the bacterial two-hybrid assay (BACTH) showed a strong interaction between periplasmic and transmembrane domains of OutL and that of OutMp, and at same time, a rather weak homo-dimerization for each of them in *D. dadantii* (Lallemant et al., 2013).

Our collaborators, Régine Dazzoni, Aracelys Lopez-Castilla from the group of Nadia Izadi-Pruneyre combined biophysical approaches, NMR and X-ray crystallography to study the *Klebsiella* AP components PulL and PulM. They determined the structure and associations of their periplasmic domains and described the first structure of the heterodimer formed by their ferredoxin-like domains. To relate these results with the *in vivo* situation, I performed the BACTH studies of PulL and PulM and showed that, unlike the isolated CTD domains, the full-length proteins do not form homodimers. Weak homodimerization of PulM was observed

however not statistically significant. I also performed Cys scanning of transmembrane segments of PulL and PulM followed by cross-linking experiments. These studies confirmed that PulM can form some homodimers which could be stabilized by crosslinking. I also contributed to the site-directed mutagenesis and functional assays which aimed at studying the role of the PulL<sub>CTD</sub> - PulM<sub>CTD</sub> interface. Finally, I used quantitative Western blot to evaluate the relative abundance of PulL and PulM in bacterial cells. I showed that PulM was much more abundant than PulL, which agrees with its dynamic role in the T2SS as a factor that binds to PulG and channels these subunits to PulL at the assembly platform. The manuscript by Dazzoni et al. describing this work has been submitted to the journal *Structure* and is presented in its preprint version below.

The Dazzoni et al. study predicted that the long periplasmic helices of PulL and PulM form a coiled coil and the role of these regions had not been tested. In the follow-up study, we investigated their function and their contribution to the PulL-PulM interaction. We found that this region is extremely important for the interaction and that without the coiled coil, the functions of the T2SS are almost abolished. We also investigated the role of the short cytoplasmic peptide of PulM. The deletion of this peptide did not seem to affect the interaction with PulL, but still, it strongly affected PulL stability and led to PulL cleavage at a precise point in the cytoplasm. The same cytoplasmic region of PulM is interesting since it also interacts with PulG. Our studies suggest that this region is crucial during the dynamic process of pseudopilus assembly and that it might interact and stabilize the cytoplasmic domain of PulL. These results are described in the manuscript by Li, Santos-Moreno et al. which is under review in *Research in Microbiology* and is presented below.

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

Article under revision in <Structure>

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

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

<sup>1</sup> Institut Pasteur, Université Paris Cité, CNRS UMR3528, Structural Bioinformatics Unit, F-75015 Paris, France

<sup>2</sup> Institut Pasteur, Université Paris Cité, CNRS UMR3528, Biochemistry of Macromolecular Interactions Unit, F-75015 Paris, France

<sup>3</sup> Institut Pasteur, Université Paris Cité, CNRS UMR3528, Biological NMR and HDX-MS Technological Platform, F-75015 Paris, France

<sup>4</sup> 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**

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 at the inner membrane subcomplex called 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 complex and 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 pseudopilus 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 pseudopilus filament composed of a non-covalent polymer of major pseudopilin GspG and minor pseudopilins 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 (Hobbs and Mattick, 1993, Peabody et al., 2003, Denise et al., 2019) 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).

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 pseudopilus 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  $\alpha$ -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

pseudopilus 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 (Lallemant 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 pseudopilus formation during protein secretion.

## Results

Previous studies provide evidence for a direct interaction between the AP components PulL and PulM in *K. oxytoca* (Possot et al., 2000, Nivaskumar et al., 2016). 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 (Lallemant 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<sub>CTD</sub>, residues 312-398) and PulM (PulM<sub>CTD</sub>, residues 79-161) (See Figure S1 for details on the PulL and PulM domain organization).

## **PuL<sub>CTD</sub> structure**

First, by using NMR and analytical ultracentrifugation (AUC), we determined the oligomerization state of PuL<sub>CTD</sub> at concentrations ranging from 10 to 300  $\mu$ M (Figure S1 and Supplementary data). PuL<sub>CTD</sub> exists in a monomer-dimer equilibrium in solution depending on the protein concentration. At the concentration used for NMR experiments (300-400  $\mu$ M), it is mainly found in dimeric form. Its  $^1\text{H}$ - $^{15}\text{N}$  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  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectrum of PuL<sub>CTD</sub> (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 PuL<sub>CTD</sub> 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 (Brünger et al., 1998, Allain et al., 2020) as detailed in Materials and Methods. The ensemble of PuL<sub>CTD</sub> monomer structures (Figure 1B) presents a backbone RMSD of 0.6 Å for its ordered parts (Table 1). PuL<sub>CTD</sub> displays a ferredoxin-like fold with 2 helices,  $\alpha$ 1 (S319-D332) and  $\alpha$ 2 (S359-R372), and an anti-parallel  $\beta$ -sheet formed by 4  $\beta$ -strands,  $\beta$ 1 (I336-D344),  $\beta$ 2 (N348-A356),  $\beta$ 3 (F373-Q376) and  $\beta$ 4 (I387-G395) (Figure 1B).

To determine the structure of the dimeric form,  $^{13}\text{C}/^{15}\text{N}$  edited-filtered NOESY experiments were performed to collect intermolecular distance restraints between two protomers of PuL<sub>CTD</sub>. 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  $\mu$ s-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 PuL<sub>CTD</sub> monomer structure (Figure 1C), it appears that they are mostly located on the same exposed face of the molecule ( $\alpha$ 1,  $\beta$ 1 and  $\beta$ 2) suggesting that they are at the dimer interface and affected by a conformational exchange between monomeric and dimeric forms.

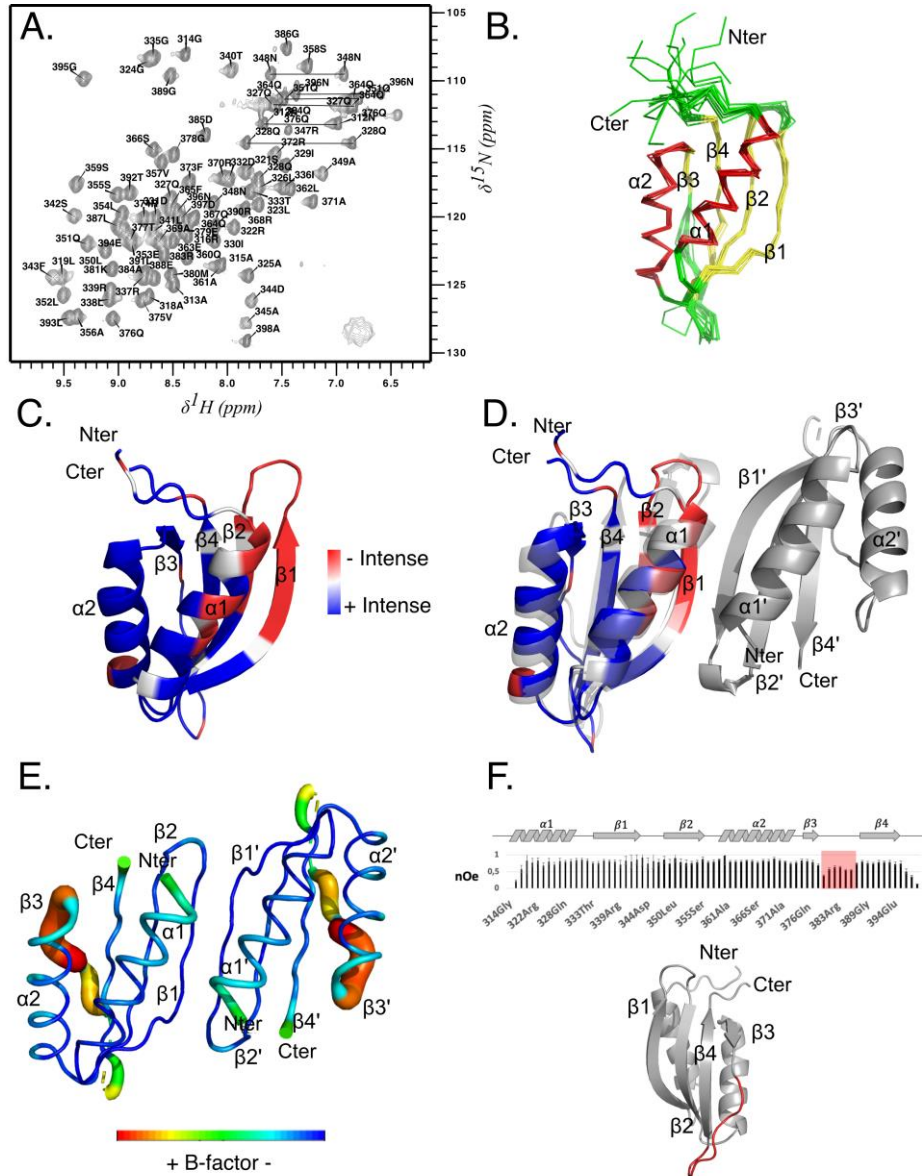


**Figure 1. *PulLCTD* homodimer structure.** **A.**  $^1\text{H}$ - $^{15}\text{N}$  HSQC NMR spectrum of *PulLCTD* (400  $\mu\text{M}$  in 50 mM HEPES buffer pH 6.5, 50 mM NaCl). Backbone resonance assignments are indicated in one-letter amino acid code and side chain  $\text{NH}_2$  peaks of Asn (N) and Gln (Q) are connected by horizontal lines. **B.** NMR structure ensemble of *PulLCTD* exhibiting a ferredoxin-like fold ( $\alpha 1$ - $\beta 1$ - $\beta 2$ - $\alpha 2$ - $\beta 3$ - $\beta 4$ ;  $\alpha$ -helix in red,  $\beta$ -sheet in yellow, loops and turns in green). **C.** Cartoon representation of the lowest-energy NMR structure of the *PulLCTD* monomer. The intensity of the  $^1\text{H}$ - $^{15}\text{N}$  HSQC signals of *PulLCTD* 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 *PulLCTD* NMR structure (coloured as in C) and the dimer structure of *PulLCTD* obtained by X-ray crystallography in grey (PDB ID: 8A9W). **E.** Representation of the B-factor per residue of *PulLCTD* in the X-ray crystallography dimer structure, from blue to red, with increasing B-factor. **F.**  $^1\text{H}$ - $^{15}\text{N}$  heteronuclear NOE values along the *PulLCTD* sequence.

*The red box indicates the lower values corresponding to the most dynamic region (ps-ns time scale), located on the  $\beta 3$ - $\beta 4$  loop shown in red in the cartoon model.*

We then used X-ray crystallography and solved the structure of PulL<sub>CTD</sub> at 1.8 Å resolution. The summary of data collection and refinement statistics is shown in Table 2. The structure of PulL<sub>CTD</sub> monomer exhibits a nearly identical ferredoxin-like fold ( $\alpha 1$ - $\beta 1$ - $\beta 2$ - $\alpha 2$ - $\beta 3$ - $\beta 4$ ) as observed in the NMR structure (Figure 1D). However, the  $\beta 3$  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  $\beta$ -strand. Consistently, higher flexibility on the ps-ns time scale in this strand in solution is evidenced by the low  $^1\text{H}$ - $^{15}\text{N}$  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<sup>2</sup>. This dimeric interface occurs between the  $\alpha 1$  helix and the  $\beta 1$  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  $\beta 1$  and  $\beta 1'$  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  $\alpha 1$  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<sub>CTD</sub> in solution observed by NMR and AUC.

### **PulM<sub>CTD</sub> structure**

PulM<sub>CTD</sub> 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<sub>CTD</sub> molecules (average RMSD 0.28 Å for C $\alpha$  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<sub>CTD</sub> protomer exhibits a ferredoxin-like fold ( $\alpha 1$ - $\beta 1$ - $\beta 2$ - $\alpha 2$ - $\beta 3$ - $\beta 4$ ) (Figure 2). The first helix  $\alpha 1$  extends from P82 to H93, followed by strands  $\beta 1$  (V98-Q103),  $\beta 2$  (R106-V111),  $\alpha 2$  helix (F116-A129),  $\beta 3$  (A134-A140) and  $\beta 4$  (V148-E156).

Among the molecular associations in the crystal, two possible dimer topologies of PulM<sub>CTD</sub> 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  $\alpha$ 2- $\beta$ 3 regions (Figure 2A). Form B was represented by a single homodimer with an  $\alpha$ 1- $\beta$ 1 interface and protomers oriented in a parallel fashion (Figure 2B).

These alternative dimeric forms A and B have, respectively, a buried surface area of 644 Å<sup>2</sup> and 463 Å<sup>2</sup> 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%).

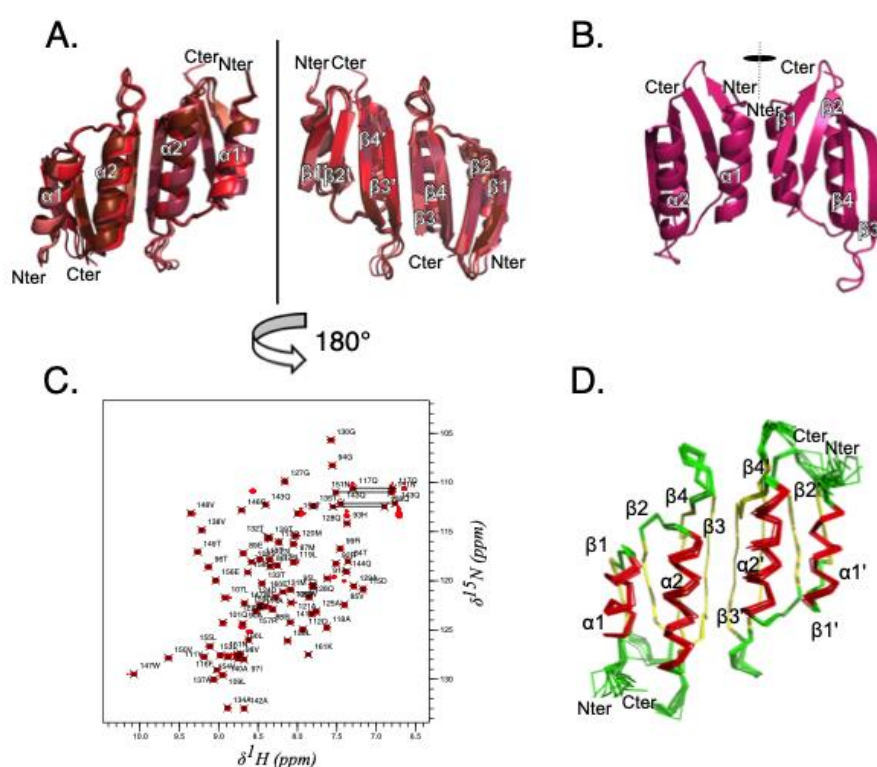


**Figure 2. PulM<sub>CTD</sub> homodimer structure.** **A.** Main topology (form A) found in the PulM<sub>CTD</sub> crystal. The four homodimers observed in the asymmetric unit are superimposed. **B.** Minor topology (form B) found in the PulM<sub>CTD</sub> crystal, represented by a single homodimer. The vertical lines indicate the twofold symmetry axis. **C.** <sup>1</sup>H-<sup>15</sup>N HSQC NMR spectrum of PulM<sub>CTD</sub> 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<sub>2</sub> signals of Asn (N) and Gln (Q) are connected by horizontal lines. **D.** NMR structure ensemble of the PulM<sub>CTD</sub> homodimer. Helices are coloured in red,  $\beta$ -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<sub>CTD</sub> exists in solution. A total of 1623 intramolecular and 25 intermolecular restraints were used to calculate the structure of PulM<sub>CTD</sub> dimer with a backbone RMSD of 0.69 Å for the 15 lowest energy structures (Figure 2D and Table 1). Each PulM<sub>CTD</sub> protomer exhibits a ferredoxin-like fold composed of 2  $\alpha$  helices and one  $\beta$  sheet as follows:  $\alpha$ 1 (T84-H93);  $\beta$ 1 (V98-Q101);  $\beta$ 2 (L107-V111),  $\alpha$ 2 (F116-A129),  $\beta$ 3 (M131-A140) and  $\beta$ 4 (V148-R157). The dimer interface occurs between the  $\alpha$ 2 helix and the  $\beta$ 3 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<sub>CTD</sub>-PulM<sub>CTD</sub> complex

An analysis of the interactions between PulL<sub>CTD</sub> and PulM<sub>CTD</sub> 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<sub>CTD</sub> for 1 PulM<sub>CTD</sub> (Figure S1A and B). For a more accurate characterization of the oligomeric state, we employed native mass spectrometry. Both PulL<sub>CTD</sub> (Figure 3A) and PulM<sub>CTD</sub> (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<sub>CTD</sub>; 17 861.42 Da/17 861.87 Da for PulM<sub>CTD</sub>). The mass spectrum of a mixture of PulL<sub>CTD</sub> and PulM<sub>CTD</sub> 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<sub>CTD</sub>-PulM<sub>CTD</sub> heterodimer (18 465.95 Da). No higher oligomeric states were detected, confirming that PulL<sub>CTD</sub> and PulM<sub>CTD</sub> form heterodimers under our experimental conditions.

To further understand the PulL<sub>CTD</sub>-PulM<sub>CTD</sub> interaction mode, we solved the crystal structure of the PulL<sub>CTD</sub>-PulM<sub>CTD</sub> 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<sub>CTD</sub> molecule flanked on each side by one PulM<sub>CTD</sub> molecule with different orientations (Figure 3D). They form together a continuous anti-parallel  $\beta$ -sheet. This arrangement results in two different interfaces, either PulL<sub>CTD</sub> <sup>$\alpha$ 1 $\beta$ 1</sup> –

$\text{PulM}_{\text{CTD}}^{\alpha 2 \beta 3}$  or  $\text{PulL}_{\text{CTD}}^{\alpha 2 \beta 3} - \text{PulM}_{\text{CTD}}^{\alpha 2 \beta 3}$ . Surface interface analysis (Table S1) indicates that  $\text{PulL}_{\text{CTD}}^{\alpha 1 \beta 1} - \text{PulM}_{\text{CTD}}^{\alpha 2 \beta 3}$  is more likely to represent a biologically relevant arrangement. In addition to the complex form both A and B forms of the  $\text{PulM}_{\text{CTD}}$  dimers, as present in the  $\text{PulM}_{\text{CTD}}$  crystal, are observed (Figure S3A).

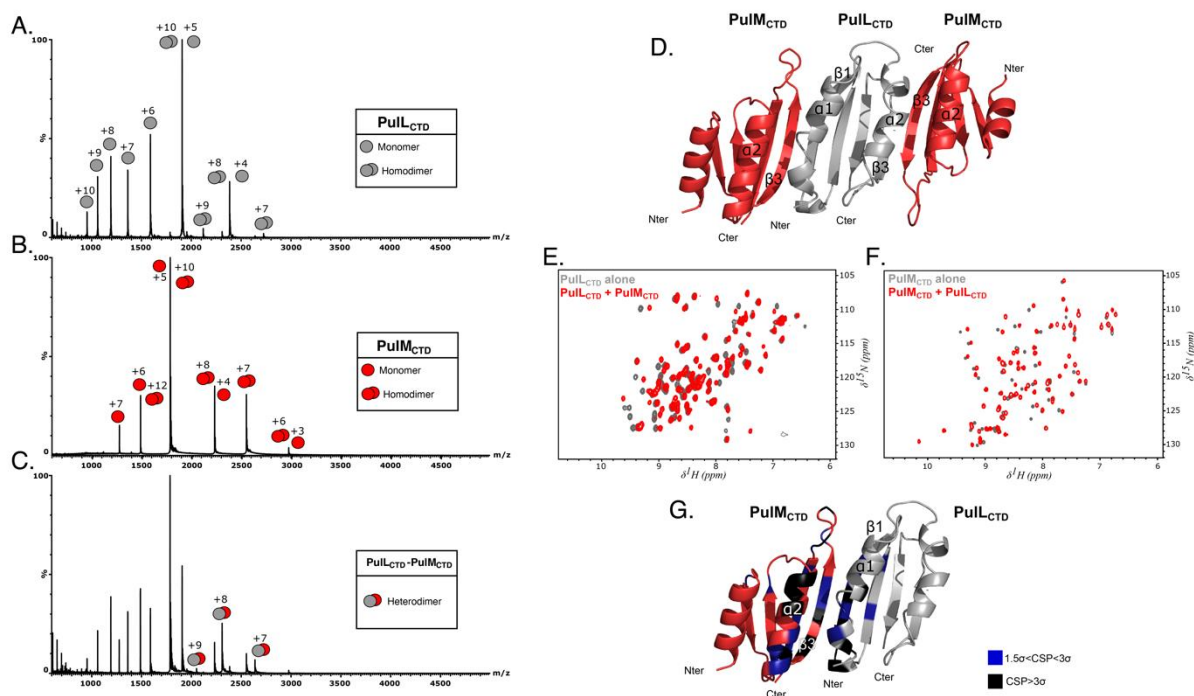


**Figure 3. Oligomeric state and interactions of the  $\text{PulL}_{\text{CTD}}\text{-PulM}_{\text{CTD}}$  complex.** Native mass spectra of  $\text{PulL}_{\text{CTD}}$  (A) and  $\text{PulM}_{\text{CTD}}$  (B) alone and in an equimolar mixture of  $\text{PulL}_{\text{CTD}}$  and  $\text{PulM}_{\text{CTD}}$  (C). From mass spectra we observe  $m/z$  corresponding to the presence of monomeric, homodimeric and heterodimeric species (annotated in grey for  $\text{PulL}_{\text{CTD}}$  and in red for  $\text{PulM}_{\text{CTD}}$ ). Charge states are indicated for all species. D.  $\text{PulL}_{\text{CTD}}\text{-PulM}_{\text{CTD}}$  complexes in the X-ray crystallography asymmetric unit. E. Overlay of  $^1\text{H}\text{-}^{15}\text{N}$  HSQC spectra of  $\text{PulL}_{\text{CTD}}$  in the absence (grey) and presence (red) of unlabeled  $\text{PulM}_{\text{CTD}}$ . F. Overlay of  $^1\text{H}\text{-}^{15}\text{N}$  HSQC spectra of  $\text{PulM}_{\text{CTD}}$  in the absence (grey) and presence (red) of unlabeled  $\text{PulL}_{\text{CTD}}$ . G. X-ray structure of the  $\text{PulL}_{\text{CTD}}^{\alpha 1 \beta 1} - \text{PulM}_{\text{CTD}}^{\alpha 2 \beta 3}$  complex found in solution. The residues at the heterodimer interface displaying the highest CSP ( $>3 \sigma$  in black,  $>1.5 \sigma$  in blue) are colored.

NMR analysis of the  $\text{PulL}_{\text{CTD}}\text{-PulM}_{\text{CTD}}$  interaction in solution, and the spectral comparison between  $^{15}\text{N}\text{-PulM}_{\text{CTD}}$  and  $^{15}\text{N}\text{-PulL}_{\text{CTD}}$  each alone and in the presence of its unlabeled partner show also that this form  $\text{PulL}_{\text{CTD}}^{\alpha 1 \beta 1} - \text{PulM}_{\text{CTD}}^{\alpha 2 \beta 3}$  is the one formed in solution



heterocomplex PulM<sub>CTD</sub>-PulL<sub>CTD</sub> (X-ray structure), from left to right. PulM<sub>CTD</sub> is colored in red and PulL<sub>CTD</sub> in grey. **E, F, G and H.** Hydrogen bonds (yellow dotted lines) in PulM<sub>CTD</sub> (X-ray structure), PulM<sub>CTD</sub> (NMR structure), PulL<sub>CTD</sub> (X-ray structure), and the PulM<sub>CTD</sub>-PulL<sub>CTD</sub> heterodimer complex (X-ray structure), from left to right. In panel H, residues forming hydrogen bonds between  $\beta$  strands are shown in green, residues forming hydrogen bonds between  $\alpha$ -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<sub>CTD</sub>-PulM<sub>CTD</sub> interface

To evaluate the role of the PulL<sub>CTD</sub>-PulM<sub>CTD</sub> 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* carrying the *pul* gene cluster with a non-polar *pulL* gene deletion in plasmid pCHAP8251. While Ala substitutions of Leu residues did not affect the function, replacing polar residue Q327 with alanine led to a small but significant defect in PulA secretion (Figure 5B). This defect 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  $\alpha$ 2 helix, or L136 of the  $\beta$ 3 strand, did not affect the function (Figure 5C). However, the double substituted variant PulM<sup>F116A-L136A</sup> showed a significant secretion defect, suggesting that the interface was weakened (Figure 5C). Replacing the surface exposed residue M120 by charged residues in PulM<sup>M120D</sup> and PulM<sup>M120K</sup> variants also significantly reduced secretion (Figure 5C). In comparison, charge inversions on the surface of the PulM<sub>CTD</sub>  $\alpha$ 1 helix, such as R88E and R92D did not show any effect, consistent with their position distal to the interface with PulL<sub>CTD</sub>. 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<sub>CTD</sub>-PulM<sub>CTD</sub> binding.

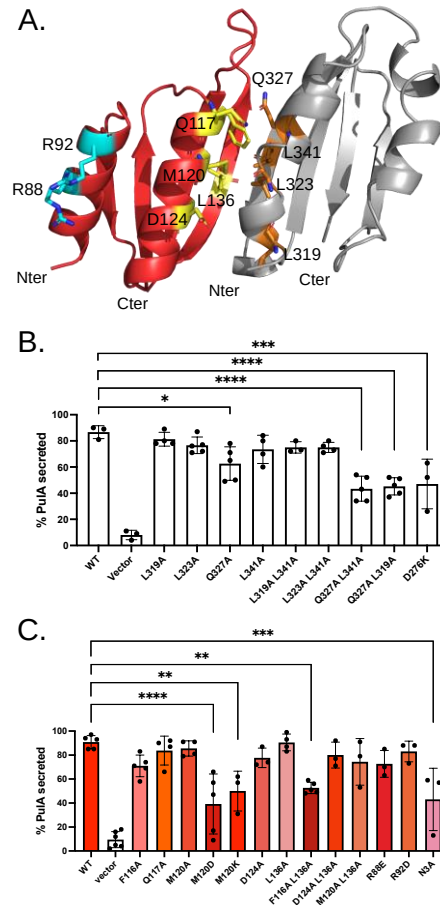


**Figure 5. Functional characterization of the PulL<sub>CTD</sub>-PulM<sub>CTD</sub> interface.** **A.** Cartoon representation of the complex of PulL<sub>CTD</sub> (grey) with PulM<sub>CTD</sub> (red). The substituted residues are highlighted as sticks and marked with a single letter code, those at the PulL-PulM interfaces are colored in orange (PulL<sub>CTD</sub>) or yellow (PulM<sub>CTD</sub>), and those far from the interface are colored in cyan. **B.** Pula secretion in strain PAP7460 carrying plasmid pCHAP8251 ( $\Delta$ pulL) complemented with presence of pulL<sup>WT</sup> (on pCHAP8258) or derivative carrying 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 at least 3 independent experiments. 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 ( $\Delta$ pulM) complemented with pCHAP1353 encoding PulM<sup>WT</sup> 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<sub>CTD</sub>-PulM<sub>CTD</sub> 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<sup>N3A</sup> 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 25) 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*.

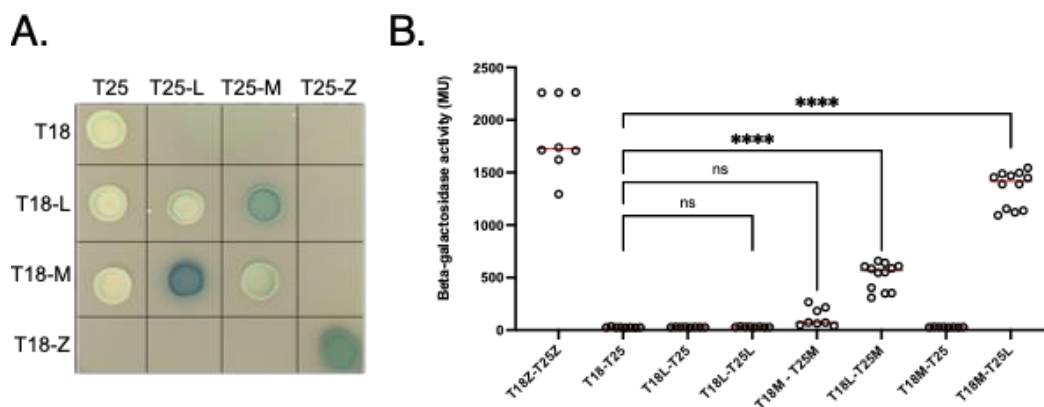


**Figure 6. BACTH analysis of PulL and PulM interactions.** A. Plate assay and B. Beta-galactosidase activity of *E. coli* strains DHT1 producing T18, T25 or their chimera with yeast

*leucin Zipper (Z) (as positive control), PulL (L) or PulM (M). The  $\beta$ -galactosidase activities measured for independent bacterial cultures 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<sup>C17L</sup> 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<sup>C264L</sup>) 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<sub>2</sub> as an oxidant and their total extracts were analyzed by Western blot with anti-PulL or anti-PulM antibodies (Figure 7). The PulL<sup>C264L</sup> 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<sup>C264L</sup> with Cys at positions 250 and 251 produced some homodimers and also heterodimers, notably with PulM with Cys at positions 17 (PulM<sup>WT</sup>) and 18 (PulM<sup>C17L-L18C</sup>). 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<sup>I250C</sup> and PulL<sup>V251C</sup>, 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).

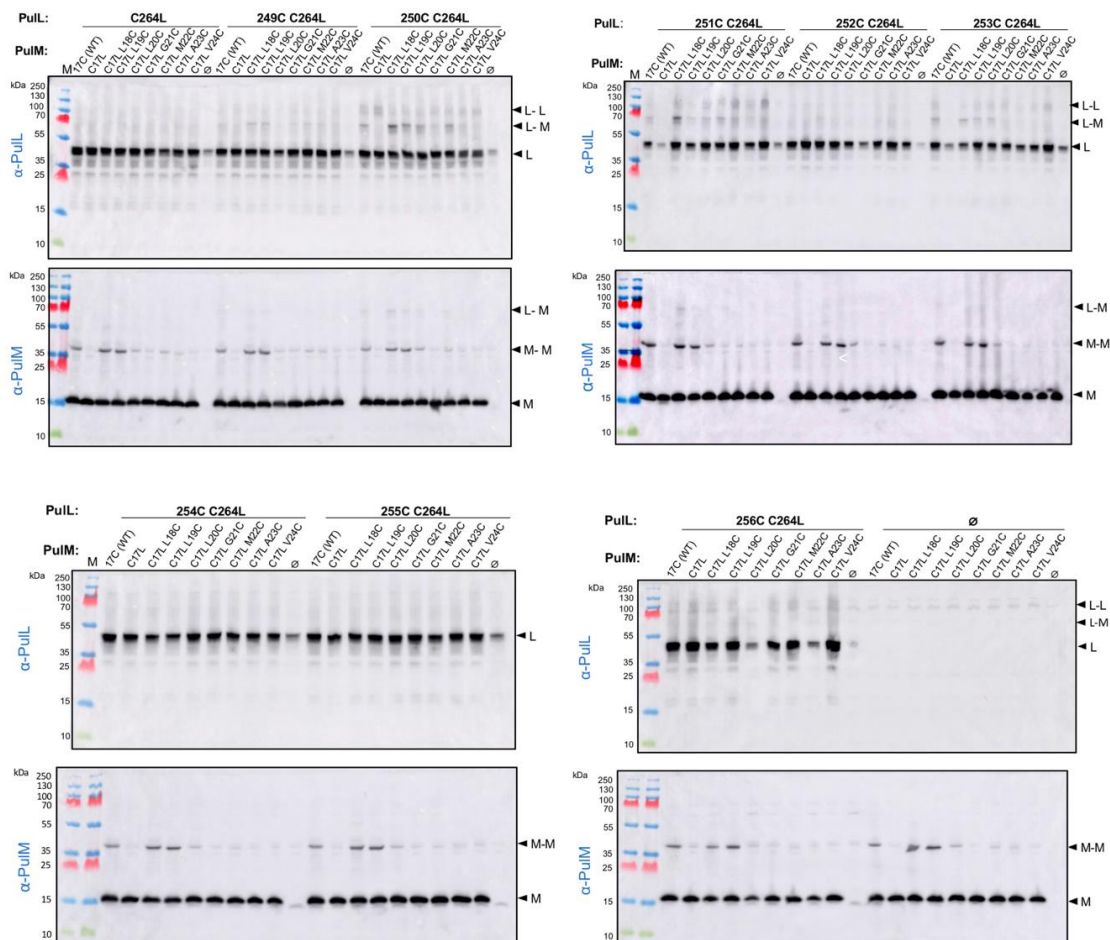


**Figure 7. Cysteine scanning and crosslinking of PulL and PulM transmembrane segments.** Bacteria (strain PAP7460) producing the indicated PulL and PulM variants, or empty vector ( $\emptyset$ ) were oxidized with  $\text{CuCl}_2$  and the total extract from 0.05  $\text{OD}_{600\text{nm}}$  of bacteria 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<sub>CTD</sub> and PulM<sub>CTD</sub> 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<sub>CTD</sub> and PulM<sub>CTD</sub> (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<sub>CTD</sub>-PulM<sub>CTD</sub> 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  $\alpha$ -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  $\alpha$ 1 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<sup>D276K</sup> 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<sub>CTD</sub> and PulM<sub>CTD</sub> each form a homodimer, they assemble as a heterodimer. Functional assays confirm that *in vivo* the dimerization interface and thus the PulL<sub>CTD</sub>-PulM<sub>CTD</sub> 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  $\beta$ -sheets facing the opposite sides in the two protomers compared to PulM (Sampaleanu et al., 2009, Leighton et al., 2018). 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<sub>CTD</sub> and PulM<sub>CTD</sub> 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 (Eck and Dayhoff, 1966, Alva and Lupas, 2018). PulL<sub>CTD</sub> and PulM<sub>CTD</sub> indeed exhibit an internal pseudo-symmetry, where the two faces of the ferredoxin-like fold ( $\alpha$ 1- $\beta$ 1- $\beta$ 2 and  $\alpha$ 2- $\beta$ 3- $\beta$ 4) are structurally superimposable and present high level of sequence similarity (64% for PulM<sub>CTD</sub> and 66% for PulL<sub>CTD</sub>). Interestingly, the dimerization interfaces of PulL<sub>CTD</sub> and PulM<sub>CTD</sub> are also highly similar, at the sequence and structural levels, with overlapping residue positions involved in the interactions (Figure S10).

The PulL<sub>CTD</sub> and PulM<sub>CTD</sub> homodimers associate in an antiparallel fashion, whereas the PulL<sub>CTD</sub>-PulM<sub>CTD</sub> 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 (Lallemant et al., 2013). During the revisions of the manuscript, we became aware of a recent work proposing a similar arrangement of the GspL–GspM ferredoxin-like domains in *E. coli* O18:K1:H7 based on molecular docking (Gutierrez AJ, 2021). Although the antiparallel homodimerization of PulL<sub>CTD</sub> and PulM<sub>CTD</sub> 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.

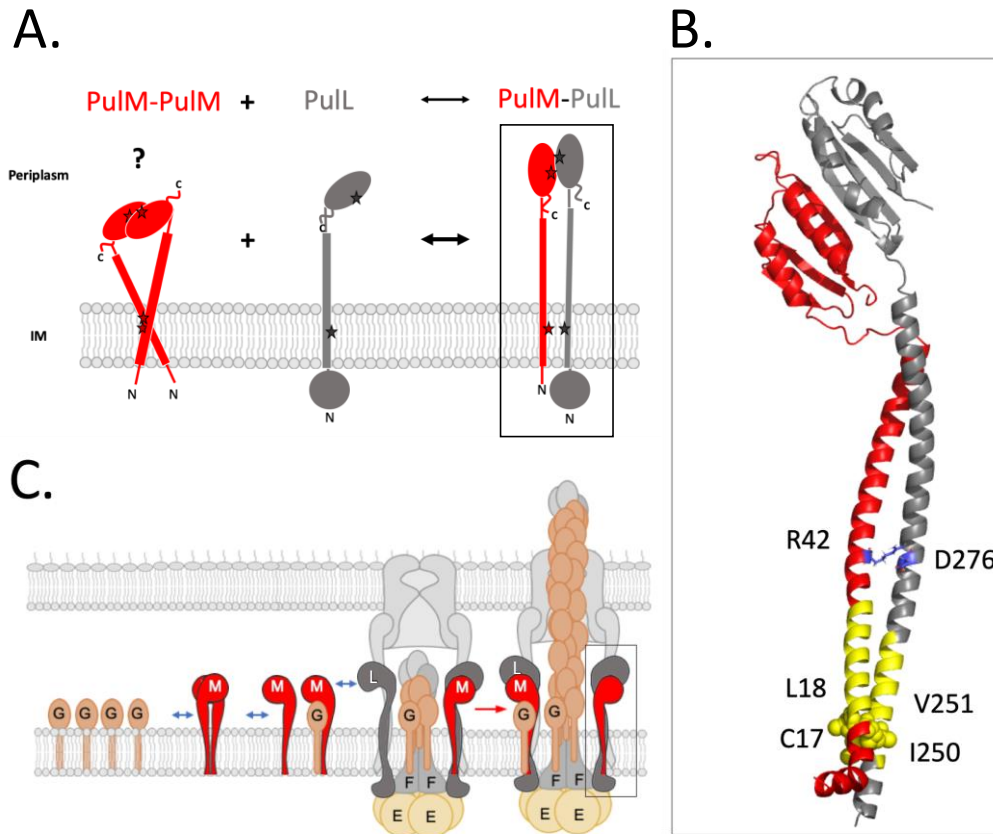


**Figure 8. Full length Pull-PulM interaction model.** **A.** Schematic view of the heterodimerization process of Pull (grey) and PulM (red). PulM can behave as a homodimer (BACTH data) via interfaces at the C-terminal domains in the periplasm (NMR data), and in the inner membrane (IM) via their transmembrane 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<sub>CTD</sub>-PulM<sub>CTD</sub> 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 pseudopilus subunit interacts with PulM. By binding to Pull, PulM targets PulG to the Pull-PulE-PulF complex. PulG, the pseudopilus subunit, is released from PulM and added to the growing pseudopilus.

In addition to the Pull<sub>CTD</sub>-PulM<sub>CTD</sub> 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 (Lallemant 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<sub>CTD</sub> and PulM<sub>CTD</sub> is likely to facilitate this exchange during the highly dynamic pseudopilus 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 pseudopilus (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 (J et al., 2017) to ensure their insertion into the growing pseudopilus. 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 pseudopilins would ensure efficient incorporation into helical filaments, thought to be the driving force for protein secretion. In addition to their role in pseudopilus 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 pseudopilus assembly and protein transport. Future studies are needed to address the influence of the pseudopilin and the secreted substrate on the AP

assembly and to define the sequence of events during pseudopilin and substrate targeting to the secretion complex.

## **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<sub>CTD</sub>), 7ZE0 (PulM<sub>CTD</sub>, NMR), 8A9X (PulM<sub>CTD</sub>, X-ray) and 8AB1 (PulL<sub>CTD</sub>-PulM<sub>CTD</sub>). PulM <sup>13</sup>C/<sup>15</sup>N/<sup>1</sup>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 as *Klebsiella pneumoniae*) described in (d'Enfert et al., 1987).

## **Method details**

### **Plasmid constructions**

*Escherichia coli* K-12 strain DH5α F'*lacI*<sup>Q</sup> 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 *EcoRI* and *KpnI* restriction enzymes (NEB) and ligated with the *EcoRI* - *KpnI* 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 *DpnI* and transformed into DH5 $\alpha$  F'*lacI*<sup>Q</sup> ultracompetent cells. The purified plasmids were verified by DNA sequencing (*Eurofins*). The list of oligonucleotides (*Eurofins*) is shown in Table S3.

Proteins PulL<sub>CTD</sub> and PulM<sub>CTD</sub> 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  $\mu\text{g.mL}^{-1}$ , chloramphenicol (Cm), 25  $\mu\text{g.mL}^{-1}$  and kanamycin (Km) at 25  $\mu\text{g.mL}^{-1}$ . When required, the genes under *placZ* control were induced with 1 mM isopropyl  $\beta$ -D-1-thiogalactopyranoside (IPTG).

### **Production of unlabeled and isotope-labeled proteins**

Derived pMalP2 vectors were used to express either the CTD of PulM (PulM<sub>CTD</sub>: from residues 79 to 161) or the CTD of PulL (PulL<sub>CTD</sub>: 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<sub>6</sub>-tag and a TEV protease cleavage site that is used to remove the MBP and the His<sub>6</sub>-tag. Only an additional N-terminal serine remains after the cleavage.

Uniformly <sup>15</sup>N and <sup>15</sup>N/<sup>13</sup>C labeled PulM<sub>CTD</sub>, PulL<sub>CTD</sub>, were produced in M9 minimal medium using 1 g/L of <sup>15</sup>NH<sub>4</sub>Cl and 4 g/L <sup>13</sup>C glucose, as the sole nitrogen and carbon sources, respectively. Gene expression was induced with 1 mM IPTG (isopropyl  $\beta$ -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 °C.

## 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<sub>6</sub> protease overnight at 10 °C. The mixture was loaded on a HisTrap HP column to remove TEV-His<sub>6</sub> and MBP-His<sub>6</sub>. 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<sub>CTD</sub> protein concentration was determined spectrophotometrically using the absorbance at 280 nm and a calculated extinction coefficient of 11000 M<sup>-1</sup>.cm<sup>-1</sup>. PulL<sub>CTD</sub> 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 <sup>15</sup>N/<sup>13</sup>C labeled proteins in 50 mM HEPES, pH 7.0 (PulM<sub>CTD</sub>) or pH 6.5 (PulL<sub>CTD</sub>), 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 <sup>1</sup>H [<sup>13</sup>C /<sup>15</sup>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 <sup>1</sup>H, <sup>15</sup>N, and <sup>13</sup>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 <sup>15</sup>N-HSQC, <sup>13</sup>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 <sup>13</sup>C-<sup>1</sup>H HBCBCGCDHD and HBCBCGCDCEHE spectra were used (Yamazaki et al., 1993). Side chain resonance assignments were completed by using 3D <sup>13</sup>C and <sup>15</sup>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  $^1\text{H}$ - $^1\text{H}$  NOESY (120 ms mixing time) and TOCSY in  $\text{D}_2\text{O}$ .

2,2-Dimethyl-2-silapentane-5-sulfonate (DSS) signal was taken as 0 ppm for referencing proton chemical shifts and  $^{15}\text{N}$  and  $^{13}\text{C}$  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, H $\alpha$ , C $\alpha$ , C $\beta$ , CO, and N chemical shifts with the TALOS-N prediction server (Shen and Bax, 2013).

The assignment of PulM<sub>CTD</sub>  $^{13}\text{C}/^{15}\text{N}/^1\text{H}$  resonances was deposited to the BMRB under accession number 34719. The assignment of PulL<sub>CTD</sub>  $^{13}\text{C}/^{15}\text{N}/^1\text{H}$  resonances was obtained in our previous work (Dazzoni et al., 2021).

### **PulM<sub>CTD</sub> and PulL<sub>CTD</sub> NMR structure calculation**

The structure of PulM<sub>CTD</sub> and PulL<sub>CTD</sub> 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 (Brünger et al., 1998, Allain et al., 2020), making use of the standard torsion angle simulating annealing protocol. Each cycle consisted of automatic 3D  $^{15}\text{N}$ -NOESY-HSQC and 3D  $^{13}\text{C}$ -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  $> 0.5 \text{ \AA}$  and no dihedral angle violations  $> 5^\circ$  were selected as the final ensemble.

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

PulM<sub>CTD</sub> dimer structure was calculated in two steps, first by using intramolecular distance restraints derived from 3D  $^{13}\text{C}$ - and  $^{15}\text{N}$ -NOESY-HSQC spectra and the PulM<sub>CTD</sub>  $^{13}\text{C}/^{15}\text{N}/^1\text{H}$  resonances assignment to determine the monomer structure. A 3D  $^{13}\text{C}/^{15}\text{N}$  filtered NOESY-HSQC experiment (120 ms mixing time) was performed on a 1:1 double-labeled:unlabeled PulM<sub>CTD</sub> 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<sub>CTD</sub> 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 <sup>1</sup>H-<sup>15</sup>N HSQC spectra of one labeled protein alone and in the presence of its unlabeled partner at 25 °C. <sup>1</sup>H-<sup>15</sup>N HSQC experiments were acquired on 30 μM of either <sup>15</sup>N-PulM<sub>CTD</sub> or <sup>15</sup>N-PulL<sub>CTD</sub> alone or in the presence of 60 μM, 120 μM and 250 μM of unlabeled PulL<sub>CTD</sub> or PulM<sub>CTD</sub>, 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  $CSP = [\Delta\delta H^2 + (\Delta\delta N \cdot 0.159)^2]^{1/2}$ , where  $\Delta\delta H$  and  $\Delta\delta N$  are the observed <sup>1</sup>H and <sup>15</sup>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<sub>CTD</sub>-PulM<sub>CTD</sub> complex was formed before or after buffer exchange by mixing PulM<sub>CTD</sub> with an equimolar or a 2-fold molar excess PulL<sub>CTD</sub>.

Samples (at a final concentration of 5 to 10 μM) were analyzed on a SynaptG2-Si HDMS mass spectrometer (*Waters*) equipped with a nanoelectrospray source. The instrument was calibrated in sensitivity mode using a 2 mg/ml cesium iodide solution prepared in 50% isopropanol, 0.1% formic acid 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<sub>CTD</sub> and PulM<sub>CTD</sub>.

### **Crystallization and diffraction data collection**

For PulL<sub>CTD</sub>-PulM<sub>CTD</sub> 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 Mosquito<sup>TM</sup> 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- $\mu$ l reservoir in multiwell plates (*Greiner Bio-one, GmbH, Frickenhausen, Germany*). The crystallization plates were stored at 18°C in an automated RockImager1000<sup>®</sup> (*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<sub>2</sub> for PulM<sub>CTD</sub>; 0.5 M LiSO<sub>4</sub> and 15% w/v PEG 8K for PulL<sub>CTD</sub>; and 20% w/v PEG 3350, 0.2M KCl for the PulL<sub>CTD</sub>-PulM<sub>CTD</sub> 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 (100K) on the PROXIMA-1 and PROXIMA-2A beamlines at the SOLEIL synchrotron (*St Aubin, France*) and processed with autoPROC (Vonnrhein et al., 2011).

### **X-ray structure determination and model refinement**

The crystal structures of the PulL<sub>CTD</sub>, PulM<sub>CTD</sub> and the PulL<sub>CTD</sub>-PulM<sub>CTD</sub> 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  $\Delta(lac-argF)U169\ araD139\ relA1\ rpsL150\ \Delta malE444\ malG501\ [F' (lacI^Q\ \Delta lacZM15\ pro^+ Tn10)] (Tc^R)$  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 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  $\mu$ L of precultures and grown for 5 hours to OD<sub>600nm</sub> > 1.8. Cultures were normalized to OD<sub>600nm</sub> 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 2 x SDS sample buffer to give the supernatant fraction. The bacterial pellets of the 0.9 mL of normalized cultures were resuspended in 90  $\mu$ L of SDS sample buffer for analysis of PulL and PulM levels (Figure S4). The cell- and supernatant fractions corresponding to the 0.05 OD<sub>600nm</sub> 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<sub>600nm</sub>) 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 was plotted and the data was analysed using GraphPad Prism 9 software. For the qualitative plate tests, 10  $\mu$ l of bacterial cultures was spotted on LB plates containing Ap, Km, X-gal (0.2 mg. mL<sup>-1</sup>) and 100  $\mu$ M 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<sub>600nm</sub> 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<sub>2</sub>, 10% (v/v) glycerol. Bacterial suspensions were prewarmed at 23 °C, CuCl<sub>2</sub> was added to 300  $\mu$ M (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<sub>600nm</sub> mL<sup>-1</sup>. 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<sub>CTD</sub> or PulM<sub>CTD</sub> 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<sub>CTD</sub> and PulM<sub>CTD</sub> 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<sub>CTD</sub> or PulM<sub>CTD</sub> 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<sub>CTD</sub>-PulM<sub>CTD</sub> 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<sub>CTD</sub>-PulM<sub>CTD</sub> heterodimer structure, a full atom model of PulL<sup>239-398</sup>-PulM<sup>1-161</sup> 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 at 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 analysed 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 defined quantities of purified proteins using with ImageJ. The standard curves of signal intensities were plotted as a function of known PulL<sup>CTD</sup> or PulM<sup>CTD</sup> 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<sub>CTD</sub> monomer and PulM<sub>CTD</sub> dimer

|                                             |                               |                               |
|---------------------------------------------|-------------------------------|-------------------------------|
| <b>Number of restraints (per monomer)</b>   |                               |                               |
| NOE Distance restraints                     |                               |                               |
| Intra-residue ( $ i-j  = 0$ )               | 392                           | 501                           |
| Sequential ( $ i-j  = 1$ )                  | 262                           | 324                           |
| Medium-range ( $2 \leq  i-j  < 5$ )         | 116                           | 184                           |
| Long-range ( $ i-j  \geq 5$ )               | 170                           | 455                           |
| Inter-molecular                             | -                             | 25                            |
| Ambiguous                                   | 23                            | 159                           |
| <i>Total</i>                                | 963                           | 1648                          |
| Dihedral angle restraints ( $\phi/\psi$ )   | 144 (72/72)                   | 138 (69/69)                   |
| Hydrogen bonds restraints                   | 28                            | 12                            |
| <b>Restraints statistics<sup>a</sup></b>    |                               |                               |
| Average no. of violations per structure     |                               |                               |
| NOE restraints $>0.5 \text{ \AA}$           | $0.7 \pm 0.8$                 | 0                             |
| H-bond restraints $>0.5 \text{ \AA}$        | 0                             | 0                             |
| Dihedral restraints $>5^\circ$              | $1.1 \pm 1.4$                 | 0                             |
| RMS of distance violations                  |                               |                               |
| NOE restraints                              | $0.053 \pm 0.013 \text{ \AA}$ | $0.046 \pm 0.001 \text{ \AA}$ |
| H-bond restraints                           | $0.034 \pm 0.004 \text{ \AA}$ | $0.021 \pm 0.001 \text{ \AA}$ |
| RMS of dihedral violations                  | $1.009 \pm 0.408^\circ$       | $0.869 \pm 0.083^\circ$       |
| <b>RMS from idealized covalent geometry</b> |                               |                               |
| bonds                                       | $0.005 \pm 0.001 \text{ \AA}$ | $0.007 \pm 0.001 \text{ \AA}$ |

|                                                          |                             |                             |
|----------------------------------------------------------|-----------------------------|-----------------------------|
| angles                                                   | $0.582 \pm 0.035^\circ$     | $0.793 \pm 0.009^\circ$     |
| impropers                                                | $1.639 \pm 0.150^\circ$     | $1.948 \pm 0.039^\circ$     |
| <b>Structural quality<sup>a</sup></b>                    |                             |                             |
| <i>Ramachandran statistics<sup>b</sup></i>               |                             |                             |
| Most favoured regions                                    | $89.4 \pm 2.8\%$            | $93.5 \pm 1.3\%$            |
| Allowed regions                                          | $10.1 \pm 3.2\%$            | $5.9 \pm 1.2\%$             |
| Generously allowed regions                               | $0.4 \pm 0.6\%$             | $0 \pm 0\%$                 |
| Disallowed regions                                       | $0.1 \pm 0.4\%$             | $0.6 \pm 0.7\%$             |
| <i>Global quality scores (Raw / Z-score)<sup>c</sup></i> |                             |                             |
| 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                |
| <b>Coordinates precision<sup>d</sup></b>                 |                             |                             |
| All residues                                             | (311-398)                   | (79-161)                    |
| Backbone atoms                                           | $1.44 \pm 0.40 \text{ \AA}$ | $0.69 \pm 0.14 \text{ \AA}$ |
| Heavy atoms (311-398)                                    | $1.74 \pm 0.38 \text{ \AA}$ | $1.10 \pm 0.16 \text{ \AA}$ |
| Ordered residues <sup>e</sup>                            | (317-344, 346-385, 387-395) | (84-101, 106-160)           |
| Backbone atoms                                           | $0.60 \pm 0.21 \text{ \AA}$ | $0.35 \pm 0.09 \text{ \AA}$ |
| Heavy atoms                                              | $1.13 \pm 0.30 \text{ \AA}$ | $0.69 \pm 0.07 \text{ \AA}$ |

<sup>a</sup> Average values and standard deviations over the final conformer ensemble

<sup>b</sup> Percentage of residues in the Ramachandran plot regions determined by PROCHECK (Laskowski et al., 1996)

<sup>c</sup> Calculated using PSVS ver. 1.5 (Bhattacharya et al., 2007)

<sup>d</sup> Average root mean square deviation (RMSD) over the ensemble atomic coordinates with respect to the average structure.

<sup>e</sup> Ordered residues [S(phi) + S(psi) > 1.8].

**Table 2:** Crystallography data collection and refinement statistics.

| Parameters           | PuLCTD                           | PuLCTD-PuMCTD                        | PuMCTD                         |
|----------------------|----------------------------------|--------------------------------------|--------------------------------|
| Beamline             | Proxima 2A                       | Proxima 1                            | Proxima 1                      |
| Resolution range (Å) | 40.15 - 1.895<br>(1.963 - 1.895) | 58.95 - 2.771<br>(2.87 - 2.771)      | 40.4 - 1.52<br>(1.575 - 1.52)  |
| Space group          | I 21 3                           | P 63 2 2                             | C 2 2 21                       |
| Unit cell (Å, °)     | 80.3 80.3 80.3<br>90 90 90       | 117.899 117.899 110.002<br>90 90 120 | 80.81 135.48 109.4<br>90 90 90 |
| Total reflections    | 277247 (26424)                   | 464974 (47811)                       | 1239705 (112086)               |
| Unique reflections   | 7012 (687)                       | 11980 (1165)                         | 91879 (8852)                   |
| Multiplicity         | 39.5 (38.5)                      | 38.8 (41.0)                          | 13.5 (12.7)                    |
| Completeness (%)     | 99.97 (100.00)                   | 99.92 (100.00)                       | 99.60 (96.68)                  |
| Mean I/sigma(I)      | 33.73 (3.03)                     | 29.12 (2.73)                         | 18.95 (1.52)                   |
| Wilson B-factor      | 41.28                            | 95.24                                | 27.69                          |
| R-merge              | 0.06935 (1.739)                  | 0.09584 (1.471)                      | 0.06888 (1.059)                |
| R-meas               | 0.07029 (1.762)                  | 0.09716 (1.49)                       | 0.07162 (1.102)                |
| R-pim                | 0.01132 (0.2832)                 | 0.01573 (0.2311)                     | 0.01943 (0.3013)               |
| CC1/2                | 1 (0.86)                         | 1 (0.846)                            | 0.999 (0.837)                  |
| CC*                  | 1 (0.962)                        | 1 (0.958)                            | 1 (0.955)                      |

|                                |                 |                 |                 |
|--------------------------------|-----------------|-----------------|-----------------|
| Reflections used in refinement | 7012 (687)      | 11980 (1165)    | 91879 (8830)    |
| Reflections used for R-free    | 339 (32)        | 581 (50)        | 4593 (442)      |
| R-work                         | 0.2408 (0.3641) | 0.2569 (0.3558) | 0.2082 (0.4667) |
| R-free                         | 0.2765 (0.2755) | 0.2980 (0.4413) | 0.2321 (0.4923) |
| CC(work)                       | 0.946 (0.725)   | 0.907 (0.681)   | 0.955 (0.801)   |
| CC(free)                       | 0.915 (0.713)   | 0.952 (0.552)   | 0.949 (0.779)   |
| Number of non-hydrogen atoms   | 606             | 2365            | 4696            |
| Macromolecules                 | 579             | 2326            | 4141            |
| Ligands                        | 10              | 0               | 0               |
| Solvent                        | 17              | 39              | 555             |
| Protein residues               | 75              | 308             | 556             |
| RMS (bonds) (Å)                | 0.010           | 0.012           | 0.024           |
| RMS (angles) (°)               | 1.41            | 1.62            | 3.61            |
| Ramachandran plot              |                 |                 |                 |
| favored (%)                    | 94.37           | 90.00           | 95.74           |
| allowed (%)                    | 2.82            | 8.00            | 2.78            |
| outliers (%)                   | 2.82            | 2.00            | 1.48            |
| Rotamer outliers (%)           | 10.00           | 10.29           | 2.07            |
| Clashscore                     | 11.83           | 9.52            | 14.79           |
| Average B-factor               | 64.98           | 122.89          | 35.56           |
| Macromolecules                 | 65.09           | 123.41          | 35.13           |
| Ligands                        | 79.76           |                 |                 |
| Solvent                        | 52.56           | 92.02           | 38.74           |
| Number of TLS groups           |                 |                 | 39              |

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., Vornrhein, 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/s09074444998003254.

Buddelmeijer, N., Francetic, O., and Pugsley, A. (2006). Green fluorescent chimeras indicate nonpolar localization of pullulanase secretion 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., Sogaard-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). <sup>1</sup>H, <sup>15</sup>N and <sup>13</sup>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.

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

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

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

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

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

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

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

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

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

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

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

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

- Lallemant, M., Login, F.H., Guschinskaya, N., Pineau, C., Effantin, G., Robert, X., and Shevchik, V.E. (2013). Dynamic interplay between the periplasmic and transmembrane domains of GspL and GspM in the type II secretion system. *PLoS One* 8, e79562. 10.1371/journal.pone.0079562.
- Laskowski, R., Rullmann, J., MacArthur, M., Kaptein, R., and Thornton, J. (1996). AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. *Journal of biomolecular NMR* 8, 477-486. 10.1007/bf00228148.
- Leighton, T., Mok, M., Junop, M., Howell, P., and Burrows, L. (2018). Conserved, unstructured regions in *Pseudomonas aeruginosa* PilO are important for type IVa pilus function. *Scientific reports* 8, 2600. 10.1038/s41598-018-20925-w.
- Linge, J., Williams, M., Spronk, C., Bonvin, A., and Nilges, M. (2003). Refinement of protein structures in explicit solvent. *Proteins* 50, 496-506. 10.1002/prot.10299.
- Lopez-Castilla, A., Thomassin, J.L., Bardiaux, B., Zheng, W., Nivaskumar, M., Yu, X., Nilges, M., Egelman, E.H., Izadi-Pruneyre, N., and Francetic, O. (2017). Structure of the calcium-dependent type 2 secretion pseudopilus. *Nat Microbiol* 2, 1686-1695. 10.1038/s41564-017-0041-2.
- 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).
- Peabody, C., Chung, Y., Yen, M., Vidal-Ingigliardi, D., Pugsley, A., and Saier, M. (2003). Type II protein secretion and its relationship to bacterial type IV pili and archaeal flagella. *Microbiology (Reading, England)* 149, 3051-3072. 10.1099/mic.0.26364-0.
- Pettersen, E., Goddard, T., Huang, C., Couch, G., Greenblatt, D., Meng, E., and Ferrin, T. (2004). UCSF Chimera--a visualization system for exploratory research and analysis. *Journal of computational chemistry* 25, 1605-1612. 10.1002/jcc.20084.
- Possot, O.M., Vignon, G., Bomchil, N., Ebel, F., and Pugsley, A.P. (2000). Multiple interactions between pullulanase secretion components involved in stabilization and cytoplasmic membrane association of PulE. *J Bacteriol* 182, 2142-2152. 10.1128/jb.182.8.2142-2152.2000.
- Pugsley, A.P. (1993). The complete general secretory pathway in gram-negative bacteria. *Microbiological reviews* 57, 50-108. 10.1128/mr.57.1.50-108.1993.
- Py, B., Loiseau, L., and Barras, F. (2001). An inner membrane platform in the type II secretion machinery of Gram-negative bacteria. *EMBO reports* 2, 244-248. 10.1093/embo-reports/kve042.
- Rieping, W., Habeck, M., Bardiaux, B., Bernard, A., Malliavin, T., and Nilges, M. (2007). ARIA2: automated NOE assignment and data integration in NMR structure calculation. *Bioinformatics* 23, 381-382. 10.1093/bioinformatics/btl589.
- Sali, A., and Blundell, T. (1993). Comparative protein modelling by satisfaction of spatial restraints. *Journal of molecular biology* 234, 779-815. 10.1006/jmbi.1993.1626.
- Sampaleanu, L., Bonanno, J., Ayers, M., Koo, J., Tammam, S., Burley, S., Almo, S., Burrows, L., and Howell, P. (2009). Periplasmic domains of *Pseudomonas aeruginosa* PilN and PilO form a stable heterodimeric complex. *Journal of molecular biology* 394, 143-159. 10.1016/j.jmb.2009.09.037.
- Sandkvist, M., Hough, L.P., Bagdasarian, M.M., and Bagdasarian, M. (1999). Direct interaction of the EpsL and EpsM proteins of the general secretion apparatus in *Vibrio cholerae*. *J Bacteriol* 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). <sup>1</sup>H, <sup>13</sup>C and <sup>15</sup>N 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.

**Supplementary data:** Oligomeric state of PulL<sub>CTD</sub> and PulM<sub>CTD</sub>.

**Supplementary Figure S1:** Oligomeric state of PulL<sub>CTD</sub>, PulM<sub>CTD</sub> and PulL<sub>CTD</sub>-PulM<sub>CTD</sub> complex analysis.

**Supplementary Figure S2:** Superposition of PulM<sub>CTD</sub> X-ray structure with PulM<sub>CTD</sub> NMR structure.

**Supplementary Figure S3:** X-ray crystallographic structure and NMR chemical shift perturbation mapping of the PulL<sub>CTD</sub>-PulM<sub>CTD</sub> complex.

**Supplementary Figure S4:** Levels of PulL and PulM variants in bacterial extracts.

**Supplementary Figure S5:** Estimating the PulL and PulM molar ratio in the bacteria.

**Supplementary Figure S6:** Sequence and structure analysis used for PulL-PulM complex modeling.

**Supplementary Figure S7:** Homologs of PulL

**Supplementary Figure S8:** Homologs of PulM

**Supplementary Figure S9:** Conserved residues in PulL and PulM homologs shown on the PulL<sub>CTD</sub>-PulM<sub>CTD</sub> hetero-dimer structure.

**Supplementary Figure S10:** Structural alignments of PulM<sub>CTD</sub> and PulL<sub>CTD</sub> dimerization regions.

**Supplementary Table S1:** Scoring biological interfaces in the X-ray crystallographic structure of the PulL<sub>CTD</sub>-PulM<sub>CTD</sub> heterodimer complex

**Supplementary Table S2:** Plasmids used in this study

**Supplementary Table S3:** Oligonucleotides used in this study

**Supplementary Materials and Methods**

**Supplementary References**

### **Supplementary data: Oligomeric state of PulL<sub>CTD</sub> and PulM<sub>CTD</sub>**

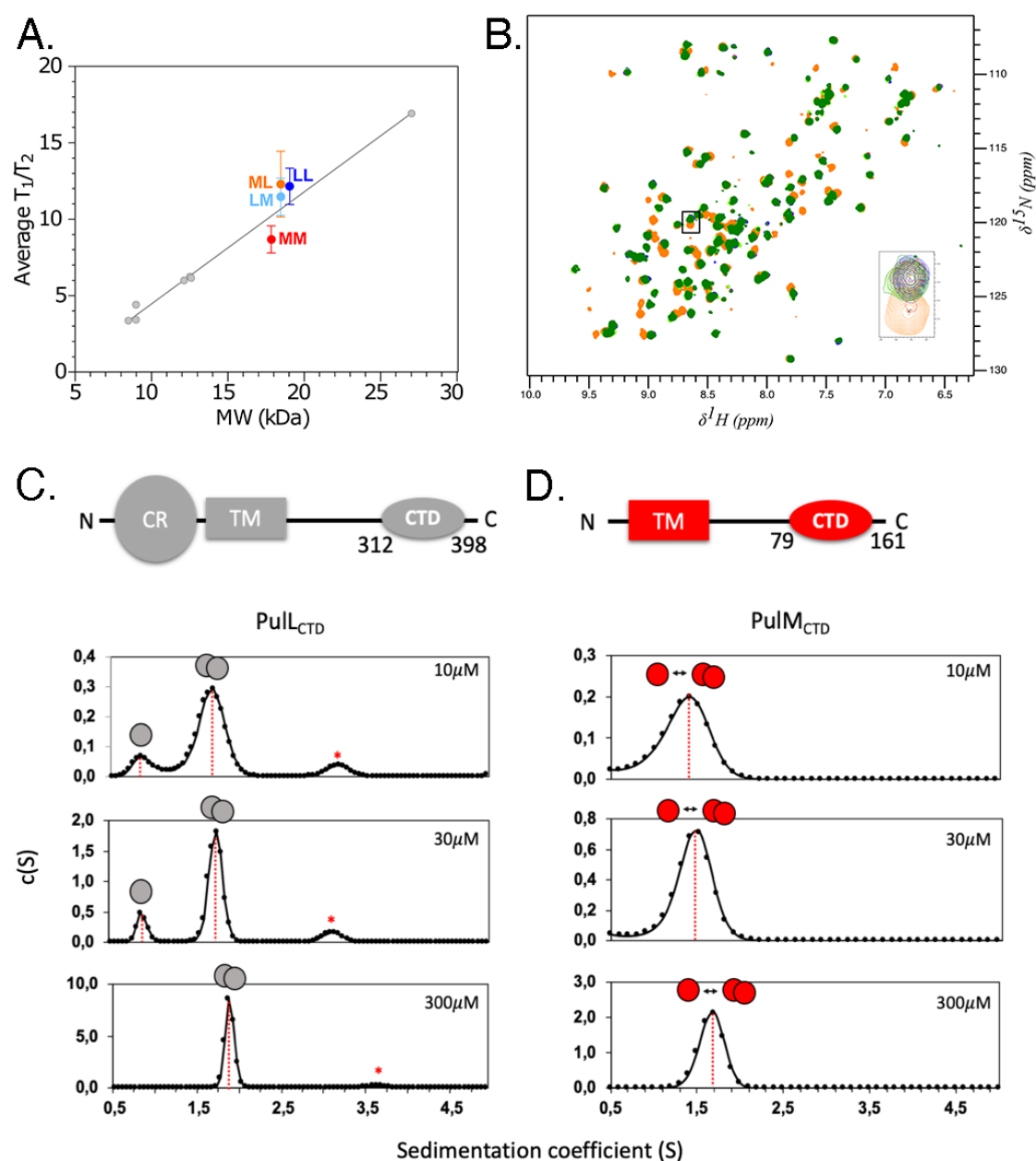
Oligomeric forms of both PulL<sub>CTD</sub> and PulM<sub>CTD</sub> have been first suspected by NMR analysis. The rotational correlation time ( $\tau_c$ ) was estimated as 10.8 ns for PulL<sub>CTD</sub> and 8.90 ns for PulM<sub>CTD</sub> by using  $T_1$  and  $T_2$  relaxation times measurements NMR experiments. For both proteins, these values are consistent with a dimeric form with a molecular weight of 19.07 kDa for PulL<sub>CTD</sub> and 17.86 kDa PulM<sub>CTD</sub>, rather than a monomeric form (Figure S1A). To determine the oligomeric states more accurately, we performed analytical ultracentrifugation (AUC). As the NMR experimental conditions require high protein concentrations (300-400  $\mu$ M in our experiments), we performed AUC with samples at concentrations ranging from 10 to 300  $\mu$ M.

For PulL<sub>CTD</sub>, we observed two main peaks with sedimentation coefficient of 1 and 1.81 S (Figure S1C). These two peaks correspond to a monomeric and a dimeric form. We observed an additional peak with higher sedimentation coefficient, compatible with a tetrameric form that evolves to a larger oligomeric form at high concentration. Nevertheless, the dimeric form represents the major form in solution ( $\cong 75\%$  of total protein) even at low concentrations, compared to the monomeric species. A monomer-dimer equilibrium can thus exist depending

on the protein concentration. Evidence for this exchange at the NMR time scale ( $\mu$ s-ms) is observed on the PulL<sub>CTD</sub>  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectrum (Figure 1A) where the line widths were broader than would be expected for a monomeric protein.

Regarding PulM<sub>CTD</sub>, we obtained a sedimentation coefficient of 1.6 S, (Figure S1D) for all tested concentrations, including that of the NMR sample (300  $\mu$ M). These values are compatible with a dimeric form of PulM<sub>CTD</sub>. Nonetheless, the PulM<sub>CTD</sub> peaks are quite broad at the tested concentrations (Figure S1D) and are getting narrower when increasing concentrations, suggesting a monomer-dimer equilibrium.

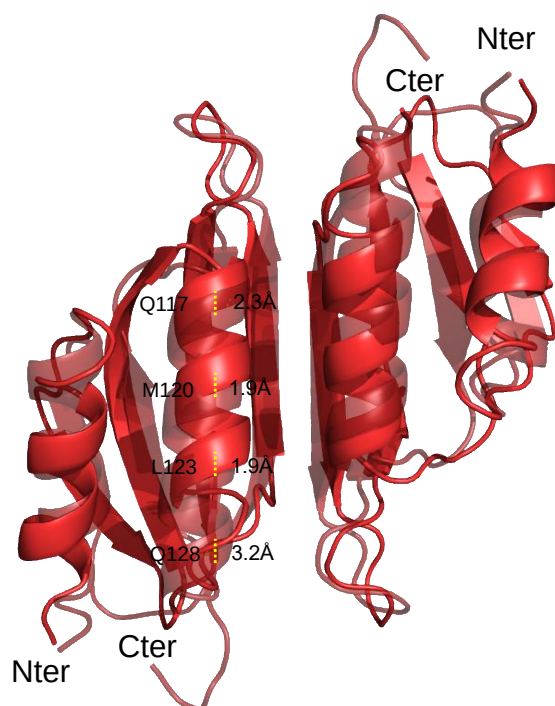
**Supplementary Figure S1:** Oligomeric state of PulL<sub>CTD</sub>, PulM<sub>CTD</sub> and PulL<sub>CTD</sub>-PulM<sub>CTD</sub> complex analysis.



**A.** Average  $T_1/T_2$  relaxation times ratios plotted as a function of the molecular weight (MW) for the dimeric states of PulM<sub>CTD</sub> (red, MM) and PulL<sub>CTD</sub> (dark blue, LL). For the mixture PulM<sub>CTD</sub>-PulL<sub>CTD</sub> (orange, ML) and PulL<sub>CTD</sub>-PulM<sub>CTD</sub> (light blue, LM) the measurements were performed on  $^{15}N$ -labeled PulM<sub>CTD</sub> (with unlabeled PulL<sub>CTD</sub>) or on  $^{15}N$ -labeled PulL<sub>CTD</sub> (with unlabeled PulM<sub>CTD</sub>), respectively (see Material and Methods). For comparison, grey circles highlight data recorded under similar conditions, at 25 °C and 600 MHz, on other in-house protein samples such as human ubiquitin (8.5 kDa, PDB id 1UBQ); *Klebsiella oxytoca* PulG

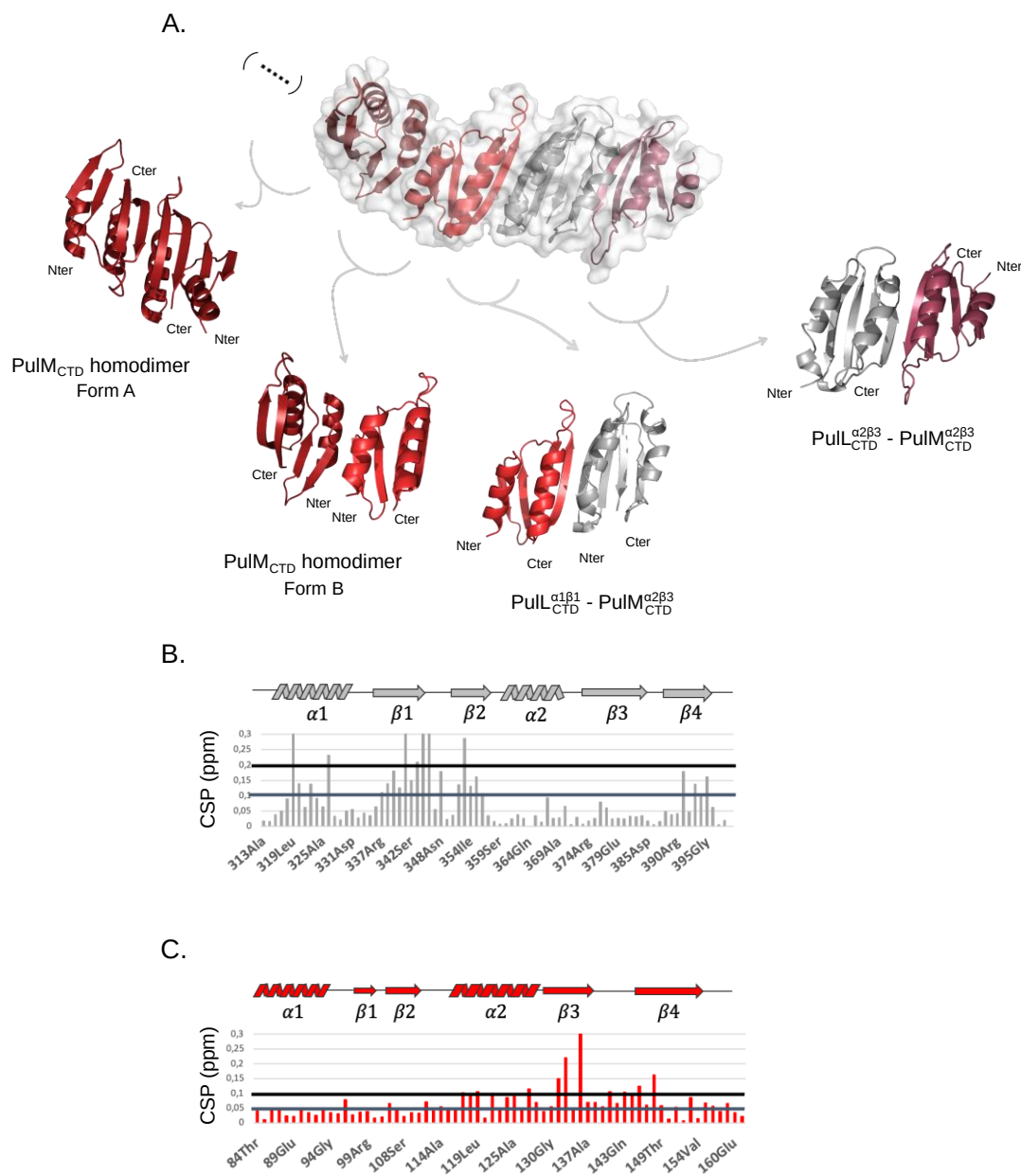
(12.1 kDa PDB id 5O2Y), etc. **B.** PulL<sub>CTD</sub>-PulM<sub>CTD</sub> titration by NMR. Left: Superimposition of <sup>1</sup>H-<sup>15</sup>N HSQC of <sup>15</sup>N-PulL<sub>CTD</sub> at 100 μM in 50 mM HEPES pH 6.5, 50 mM NaCl (orange), in the presence of PulM<sub>CTD</sub> at 50 μM (red), 100 μM (purple), 150 μM (blue), 200 μM PulM<sub>CTD</sub> (dark green), and 300 μM PulM<sub>CTD</sub> (light green). All spectra were acquired on a 600 MHz spectrometer with the same parameters. The insert shows a zoom of a representative signal (Asp 331). **C.** Determination of PulL<sub>CTD</sub> oligomeric state by analytical ultracentrifugation (AUC). Top: Schematic representation of PulL protein with the cytoplasmic region (CR), transmembrane segment (TM) and the C-terminal domain (CTD). Bottom: Sedimentation coefficient distribution of PulL<sub>CTD</sub> obtained by AUC at the indicated concentrations. Red dashed lines highlight the main oligomeric forms, schematized with single (monomer) or double (dimer) grey circles. The red star indicates an additional minor peak corresponding to a higher oligomeric state. **D.** Determination of PulM<sub>CTD</sub> oligomeric state by AUC. Top: Schematic representation of PulM with the transmembrane segment (TM) and the C-terminal domain (CTD). Bottom: Sedimentation coefficient distribution of PulM<sub>CTD</sub> forms obtained by AUC at the indicated concentrations. Red dashed lines highlight the main oligomeric forms. The different forms are schematized with single (monomer) or double (dimer) red circles.

**Supplementary Figure S2:** Superposition of PulM<sub>CTD</sub> X-ray structure with PulM<sub>CTD</sub> NMR structure



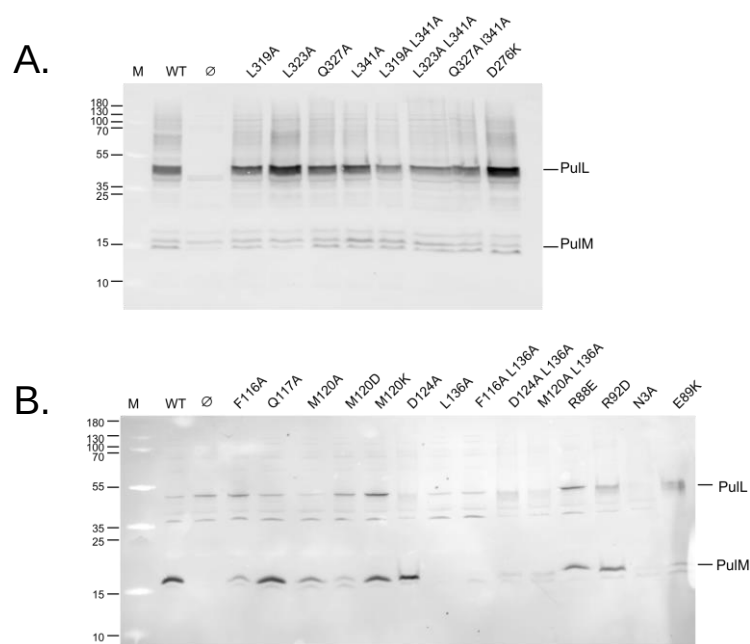
Superposition of PulM<sub>CTD</sub> X-ray structure (opaque red) with PulM<sub>CTD</sub> NMR structure (transparent red). The average backbone RMSD is 1.6 Å when the secondary structure elements of both chains are considered. The most noticeable difference between the NMR and X-ray crystallographic structures is a slight displacement of the helix  $\alpha_2$ , 2.3 Å on average along the helix axis. In dashed yellow line, measurement of the position of C $\alpha$  of the helix  $\alpha_2$  residues between the X-ray structure and the NMR structure of PulM<sub>CTD</sub>.

**Supplementary Figure S3:** X-ray crystallographic structure and NMR chemical shift perturbation (CSP) mapping of the PulL<sub>CTD</sub>-PulM<sub>CTD</sub> complex.



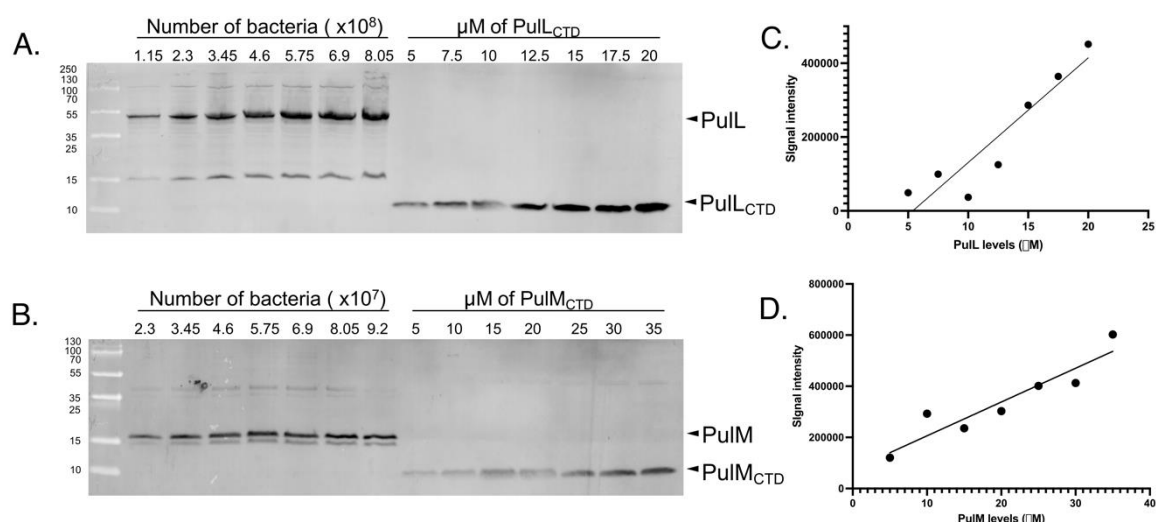
**A.** The crystallographic asymmetric unit is composed of PulM<sub>CTD</sub>/PulM<sub>CTD</sub>/PulL<sub>CTD</sub>/PulM<sub>CTD</sub> molecules, from left to right (PulM<sub>CTD</sub> are colored in shades of red and PulL<sub>CTD</sub> in grey). Below, the different homo/heterodimeric forms found in the asymmetric unit and crystal lattice are shown. **B.** Chemical shift perturbation (CSP) of PulL<sub>CTD</sub> backbone amide signals as a function of residue number. The grey lines indicate CSP value higher than 3 or 1.5 standard deviations ( $\sigma$ ) from the mean CSP. **C.** CSP of PulM<sub>CTD</sub> backbone amide signals as a function of residue numbers. The grey lines indicate CSP value higher than 3  $\sigma$  or 1.5  $\sigma$  from the mean CSP.

**Supplementary Figure S4: Levels of PulL and PulM variants in bacterial extracts.**



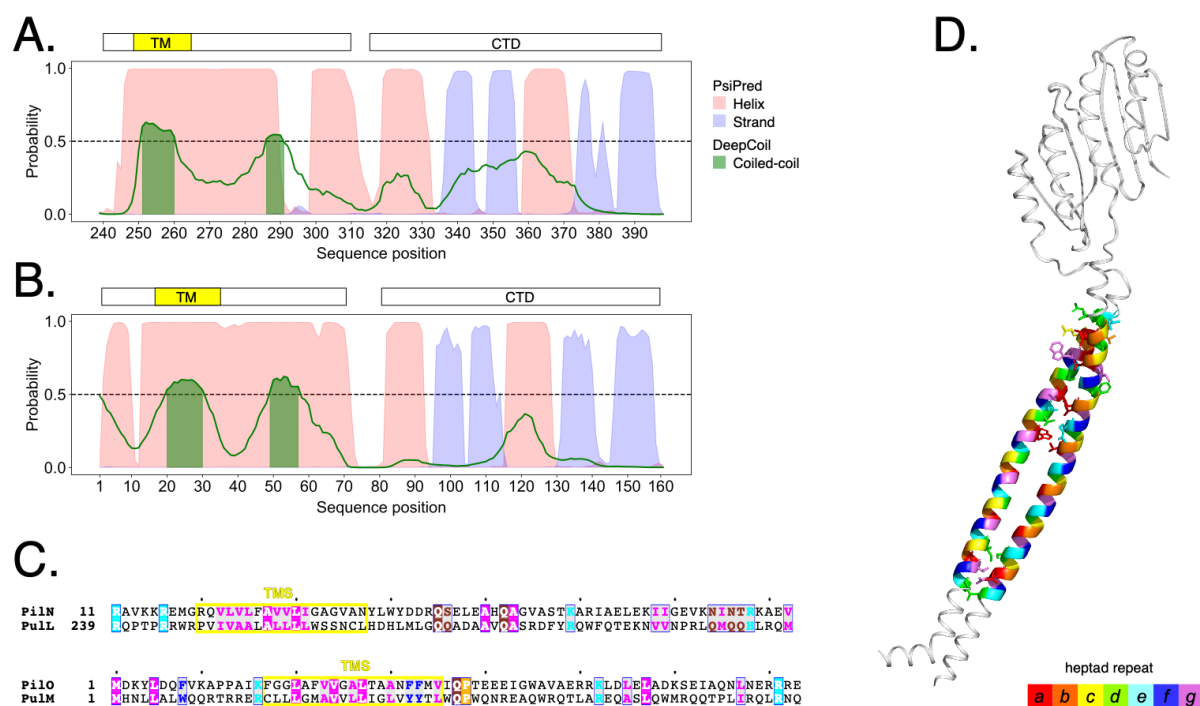
**A.** Total extracts of strain PAP7460 harboring pCHAP8251 (lacking the *pulL* gene) complemented with plasmids encoding PulL or its mutant variants. Total extracts from 0.1 OD<sub>600nm</sub> of cultures were analyzed on 10% Tris-Tricine SDS-PAGE followed by Western blot with anti-PulM and anti-PulL antibodies. Note that the band migrating slightly faster than PulM is a PulL degradation product. **B.** Total extracts of strain PAP7460 harboring pCHAP8496 (lacking the *pulM* gene) complemented with plasmids encoding PulM or its mutant variants. Total extracts from 0.1 OD<sub>600nm</sub> of cultures were analyzed on 10% Tris-Tricine SDS-PAGE followed by Western blot with anti-PulM and anti-PulL antibodies. Note that the L136A substitution in PulM affects the epitope recognized by anti-PulM antibodies.

**Supplementary Figure S5: Estimation of the PulL and PulM molar ratio in the bacteria.**



Bacteria of strain PAP7460 carrying plasmid pCHAP8185 were grown to early exponential phase under conditions inducing the *pul* gene expression. Colony forming units were enumerated and total bacterial extracts were prepared from the same cultures at concentrations of 1 and 10 OD<sub>600nm</sub>.mL<sup>-1</sup>. Bacterial counts per 1 OD<sub>600nm</sub>.mL<sup>-1</sup> were 2.3 x10<sup>9</sup> colony forming units. **A.** Western blots of total bacterial extracts from the numbers of bacteria indicated above each lane and known concentrations of purified PulL<sub>CTD</sub>, revealed with anti-PulL antibodies. Positions of PulL and PulL<sub>CTD</sub> are indicated on the right. The Mw standards are shown on the left (in kDa). **B.** Western blots of total bacterial extracts from the number of bacteria indicated above each lane and known concentrations of purified PulM<sub>CTD</sub>, revealed with anti-PulM antibodies. **C.** Standard curve quantifying the signal intensity of known amounts of purified PulL<sub>CTD</sub> (in  $\mu\text{M}$ ) analyzed in panel (A). **D.** Standard curve quantifying the signal intensity of known amounts of purified PulM<sub>CTD</sub> (in  $\mu\text{M}$ ) analyzed in panel (B). Based on these measurements, we estimated that the PulM signal from 1 mOD<sub>600nm</sub> of bacteria corresponds to 2  $\mu\text{M}$  of PulM<sub>CTD</sub>, while for PulL it corresponds to 0.1  $\mu\text{M}$  of PulL<sub>CTD</sub>, suggesting an approximate PulL/PulM ratio of 1/20 per cell.

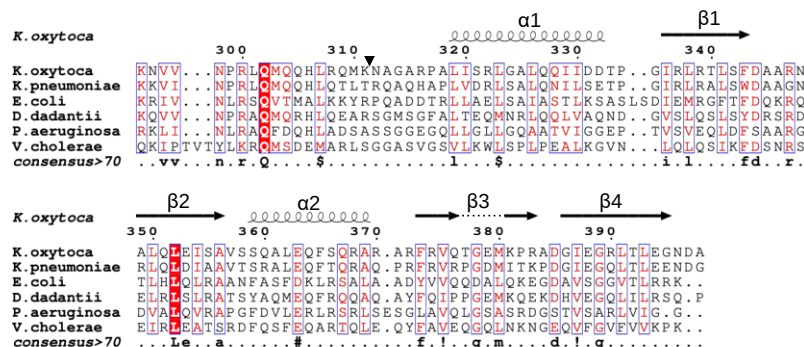
**Supplementary Figure S6:** Sequence and structure analysis used for PulL-PulM complex modeling.



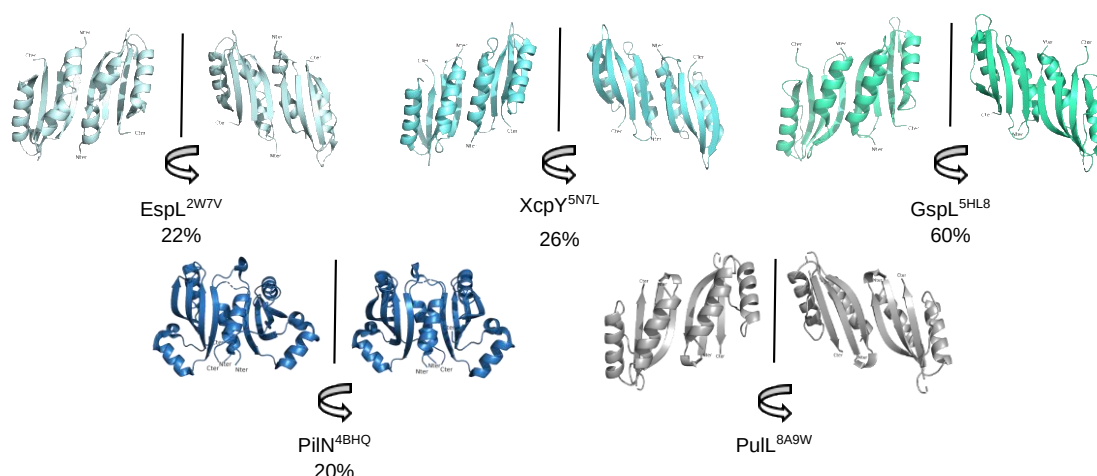
Probabilities for secondary structures [PsiPred] (Jones, 1999) and coiled-coil [DeepCoil] (Ludwiczak *et al.*, 2019) predictions for PulL (A) and PulM (B) sequences. Helical probabilities are shown in red, beta-strand in blue and coiled-coil in green. Position of the transmembrane segments (TM) is shown in yellow. Regions with coiled-coil probability above 0.5 are highlighted in green. C. Sequence alignments of the N-terminal regions of *M. xanthus* PilN<sub>11-81</sub>/PilO<sub>1-73</sub> with PulL<sub>239-310</sub>/PulM<sub>1-73</sub> as used for modeling. Positions with similar amino acids are boxed and colored by amino acid properties: hydrophobic in pink, polar in maroon, positively charged in cyan, aromatics in blue, and glycine and proline in orange. The predicted TMS are highlighted in yellow. D. Knobs-into-Holes and heptad repeats found in the model of the PulL-PulM complex using the SOCKET2 program (Kumar and Woolfson, 2021). Residues are colored according to their positions in the heptad repeat [a-g] of the coiled-coil. Side-chains of residues detected as knobs or forming holes are shown as sticks.

## Supplementary Figure S7: Homologs of Pull

A.



B.

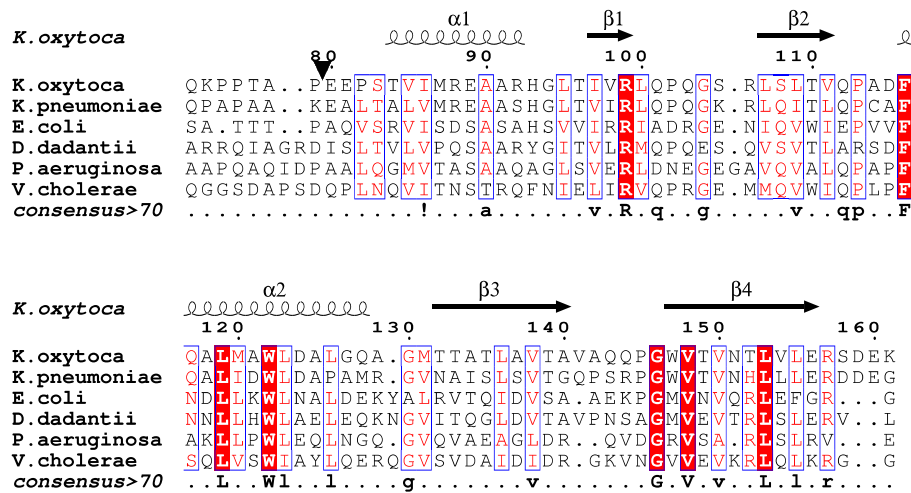


**A.** Sequence alignment of representative examples of the C-terminal domain of Pull homologs. The residue numbering is for Pull, and the triangle indicates the beginning of the C-terminal domain of Pull used in this study. The secondary structure elements correspond to the NMR structure of the Pull. The alignment was performed by using Escript (Robert and Gouet, 2014).

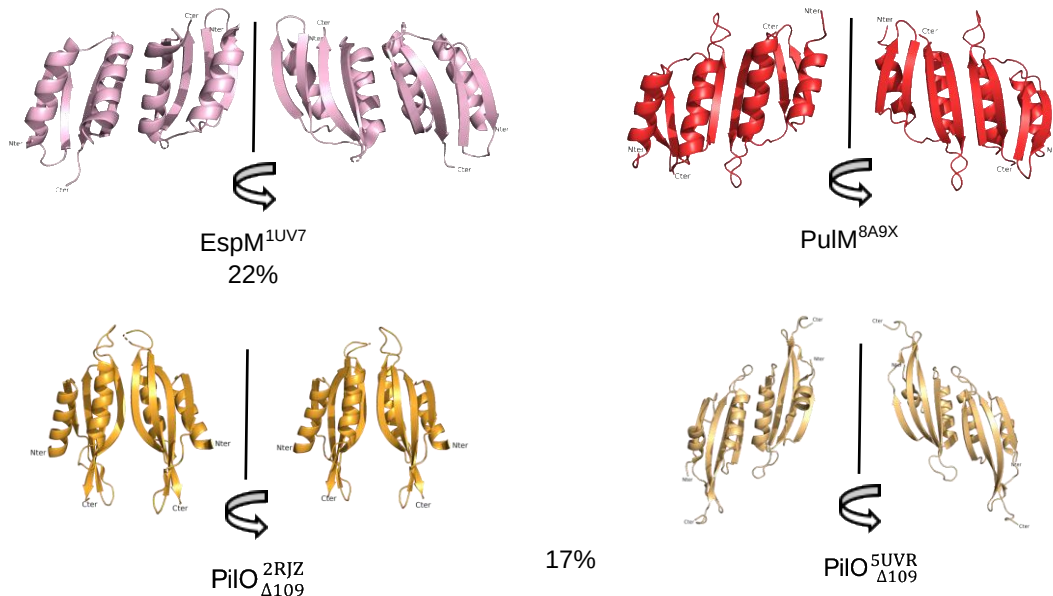
**B.** Structures of the Pull<sub>CTD</sub> homologs obtained by X-ray crystallography. The percentage of sequence identity with Pull<sub>CTD</sub> is indicated below each structure.

## Supplementary Figure S8: Homologs of PulM

**A.**

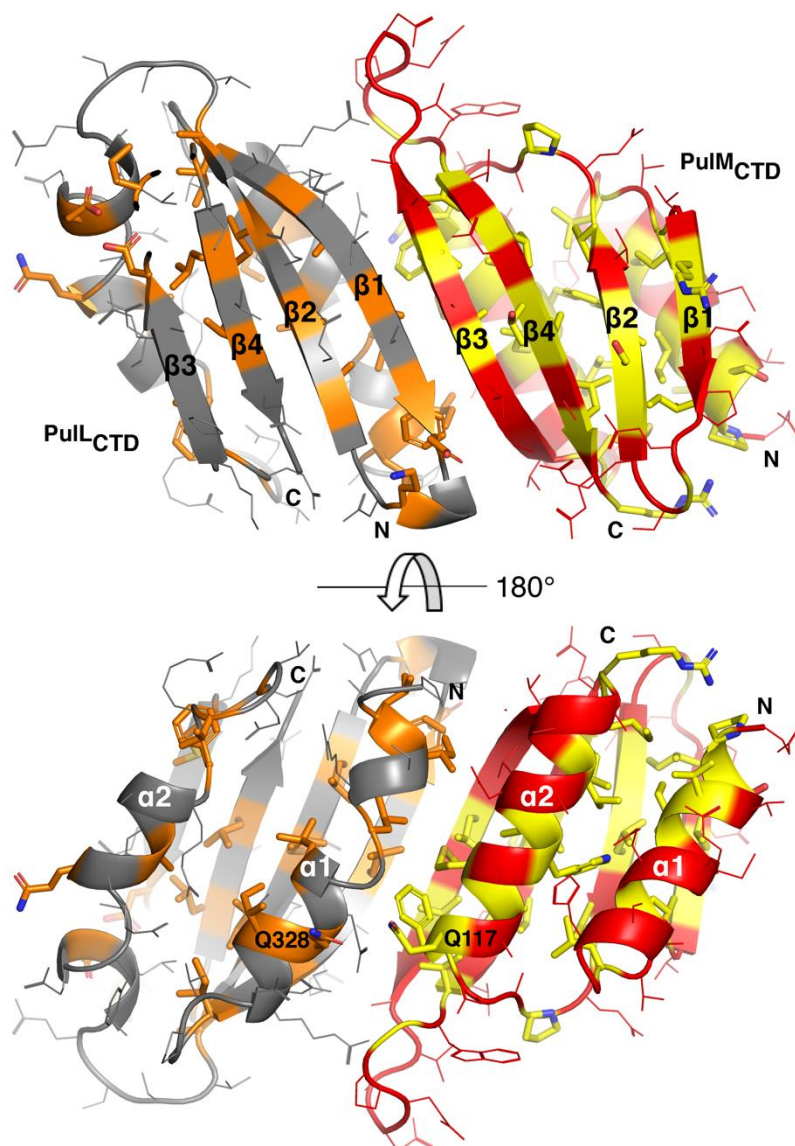


**B.**



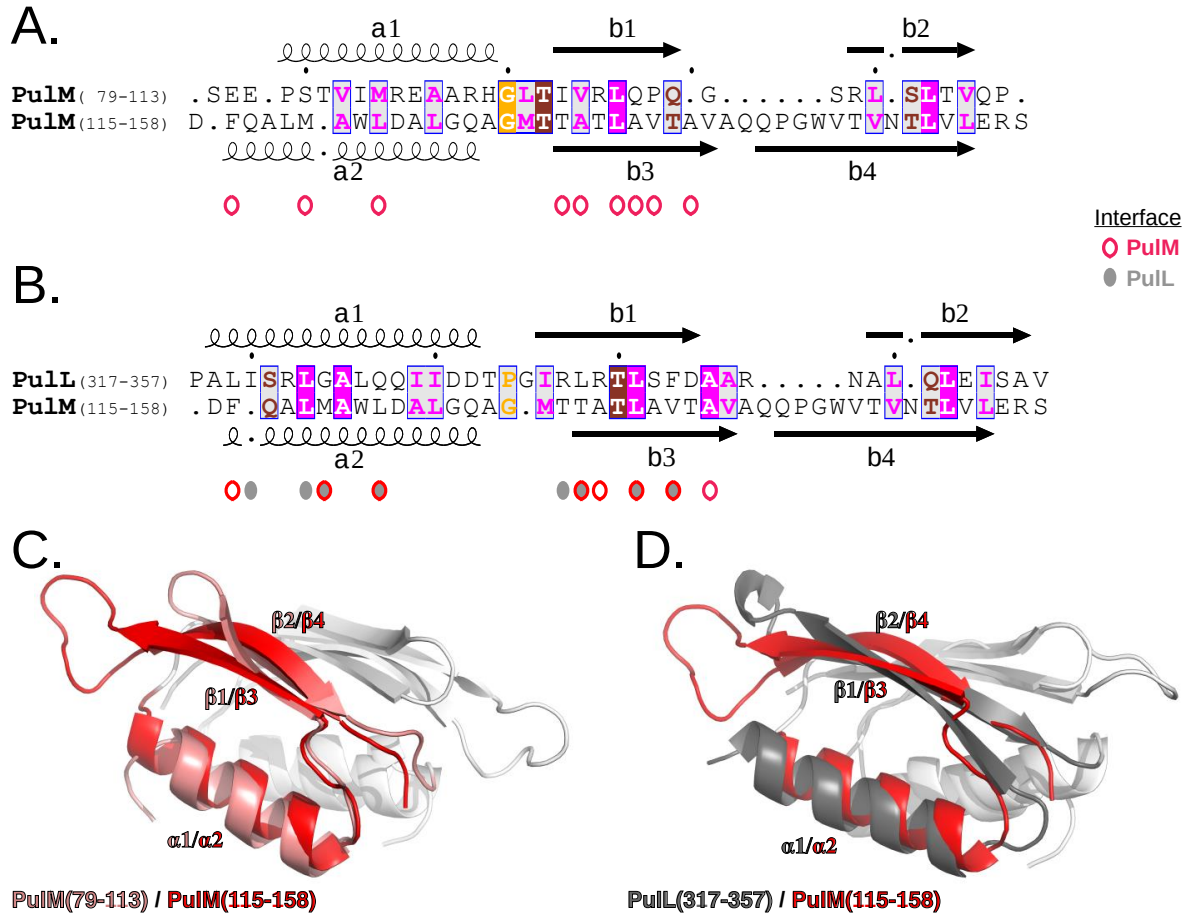
**A.** Sequence alignment of representative examples of the C-terminal domain of PulM homologs. The residue numbering is for PulM; the triangle corresponds to the beginning of the CTD of PulM used in this study. The secondary structure elements correspond to the NMR structure of PulM<sub>CTD</sub>. The alignment was performed by using Escript (Robert and Gouet, 2014). **B.** Structures of PulM<sub>CTD</sub> homologs determined by X-ray crystallography. The percentage of sequence identity with PulM<sub>CTD</sub> is indicated for each protein.

**Supplementary Figure S9:** Conserved residues in PulL and PulM homologs shown on the PulL<sub>CTD</sub>-PulM<sub>CTD</sub> heterodimer structure.



Representation of the X-ray structure of the PulL<sub>CTD</sub>-PulM<sub>CTD</sub> heterodimer complex where PulL is colored in grey and PulM in red. Conserved residues in sequence alignments of homologs (Figures S7 for PulL and S8 for PulM) are colored in orange for PulL and in yellow for PulM with side chains as sticks. Secondary structures elements are labeled along with the positions of N and C termini. The bottom representation is a rotated view by 180°. Most of the conserved residues are involved in the hydrophobic core of each ferredoxin-like domain, in addition to some exposed residues at the hetero-dimer interface such as Q328 (PulL) and Q117 (PulM). In addition, only a single residue (Q367) appears to be conserved and exposed on the  $\beta 3/\alpha 2$  side of PulL which is neither involved in homo-dimerization nor hetero-dimerization in our data.

**Supplementary Figure S10: Structural alignments of PulM<sub>CTD</sub> and PulL<sub>CTD</sub> dimer interfaces.**



**A.** Sequence alignment based on structural superimposition of PulM(79-113) and PulM(115-158) with TM-align (Zhang and Skolnick, 2005). Positions with similar amino acids are boxed and colored according to amino acid properties (hydrophobic in pink, polar in maroon, glycine and proline in orange). Secondary structure elements are labelled on top for PulM(79-113) and bottom for PulM(115-158). Positions involved in PulM<sub>CTD</sub> homodimerization are highlighted with red open circles. **B.** Structural alignment of PulL(317-358) and PulM(115-158), with amino acids colored as in (A). Secondary structure elements are labelled on top for PulL(317-358) and bottom for PulM (115-158). Positions involved in PulL<sub>CTD</sub>-PulM<sub>CTD</sub> heterodimerization are highlighted with filled grey circles for PulL and red open circles for PulM. **C.** Superimposition of PulM(79-113) in pink and PulM(115-158) in red. **D.** Superimposition of PulL(317-357) in grey and PulM(115-158) in red. In C and D, secondary structure elements are labeled in the respective colors. Parts of the proteins not used for superimposition are colored in white for sake of clarity.

**Supplementary Table S1:** Scoring biological interfaces in the X-ray crystallographic structure of the PulL<sub>CTD</sub>-PulM<sub>CTD</sub> heterodimer complex

|                              | PulL <sub>CTD</sub> <sup>α2β3</sup> – PulM <sub>CTD</sub> <sup>α2β3</sup> | PulL <sub>CTD</sub> <sup>α1β1</sup> – PulM <sub>CTD</sub> <sup>α2β3</sup> |
|------------------------------|---------------------------------------------------------------------------|---------------------------------------------------------------------------|
| <b>Buried interface area</b> | 544 Å <sup>2</sup>                                                        | 639 Å <sup>2</sup>                                                        |
| <b>Prodigy-Crystal</b>       | 7%                                                                        | 19%                                                                       |
| <b>ClusPro-DC</b>            | 11%                                                                       | 58%                                                                       |

We tested the probability of biological relevance of heterodimers formed by PulL<sub>CTD</sub> and two PulM<sub>CTD</sub> molecules within the asymmetric unit. Both Prodigy-Crystal (Jiménez-García *et al.*, 2019) and ClusPro-DC (Yueh *et al.*, 2017) predicted a higher probability for PulL<sub>CTD</sub><sup>α1β1</sup> – PulM<sub>CTD</sub><sup>α2β3</sup> to be a biologically relevant interface than PulL<sub>CTD</sub><sup>α2β3</sup> – PulM<sub>CTD</sub><sup>α2β3</sup>. These results agree with PISA analysis of the buried interface for each heterodimer and, more importantly, with the results of the NMR CSP analysis of heterodimer interface in solution.

**Supplementary Table S2:** Plasmids used in this study

| Plasmid              | Markers                                                                                   | Source/Reference                     |
|----------------------|-------------------------------------------------------------------------------------------|--------------------------------------|
| pSU18                | <i>placZ</i> , p15A <i>ori</i> , Cm <sup>R</sup>                                          | (Bartolomé <i>et al.</i> , 1991)     |
| pMalP2               | <i>placZ</i> , <i>malE</i> , ColE1 <i>ori</i> , Ap <sup>R</sup>                           | New England Biolabs                  |
| pPulL <sub>CTD</sub> | pMalP2-His-TEV-PulL <sub>CTD</sub>                                                        | (Dazzoni <i>et al.</i> , 2021)       |
| pPulM <sub>CTD</sub> | pMalP2-His-TEV-PulM <sub>CTD</sub>                                                        | <i>Proteogenix</i>                   |
| pCHAP8185            | <i>pulS</i> , <i>pulAB</i> , <i>pulCDEFGHIJKLMNO</i> , ColE1 <i>ori</i> , Ap <sup>R</sup> | (Cisneros <i>et al.</i> , 2012)      |
| pCHAP8496            | pCHAP8185 <i>ΔpulM</i> , Ap <sup>R</sup>                                                  | (Santos-Moreno <i>et al.</i> , 2017) |
| pCHAP8251            | pCHAP8185 <i>ΔpulL</i> , Ap <sup>R</sup>                                                  | (Santos-Moreno <i>et al.</i> , 2017) |
| pCHAP8258            | pSU18- <i>pulL</i> , Cm <sup>R</sup>                                                      | (Santos-Moreno <i>et al.</i> , 2017) |
| pCHAP1353            | pSU18- <i>pulM</i> , Cm <sup>R</sup>                                                      | (Possot <i>et al.</i> , 2000)        |
| pMS1185              | pSU18- <i>pulL</i> L319A, Cm <sup>R</sup>                                                 | This study                           |
| pMS1186              | pSU18- <i>pulL</i> L323A, Cm <sup>R</sup>                                                 | This study                           |
| pMS1187              | pSU18- <i>pulL</i> Q327A, Cm <sup>R</sup>                                                 | This study                           |
| pMS1188              | pSU18- <i>pulL</i> L341A, Cm <sup>R</sup>                                                 | This study                           |
| pMS1235              | pSU18- <i>pulL</i> L319A L341A, Cm <sup>R</sup>                                           | This study                           |
| pMS1236              | pSU18- <i>pulL</i> L323A L341A, Cm <sup>R</sup>                                           | This study                           |
| pMS1237              | pSU18- <i>pulL</i> Q327A L341A, Cm <sup>R</sup>                                           | This study                           |
| pMS1239              | pSU18- <i>pulL</i> Q327A L319A, Cm <sup>R</sup>                                           | This study                           |
| pMS1189              | pSU18- <i>pulM</i> F116A, Cm <sup>R</sup>                                                 | This study                           |
| pMS1190              | pSU18- <i>pulM</i> M120A, Cm <sup>R</sup>                                                 | This study                           |
| pMS1166              | pSU18- <i>pulM</i> M120D, Cm <sup>R</sup>                                                 | This study                           |
| pCHAP8705            | pSU18- <i>pulM</i> M120K, Cm <sup>R</sup>                                                 | This study                           |

|            |                                                              |                                   |
|------------|--------------------------------------------------------------|-----------------------------------|
| pMS1192    | pSU18- <i>pulM</i> D124A, Cm <sup>R</sup>                    | This study                        |
| pMS1201    | pSU18- <i>pulM</i> L136A, Cm <sup>R</sup>                    | This study                        |
| pMS1223    | pSU18- <i>pulM</i> D124A L136A, Cm <sup>R</sup>              | This study                        |
| pMS1224    | pSU18- <i>pulM</i> F116A L136A, Cm <sup>R</sup>              | This study                        |
| pMS1225    | pSU18- <i>pulM</i> M120A L136A, Cm <sup>R</sup>              | This study                        |
| pCHAP8468  | pSU18- <i>pulM</i> R88E, Cm <sup>R</sup>                     | This study                        |
| pCHAP8525  | pSU18- <i>pulM</i> R92D, Cm <sup>R</sup>                     | This study                        |
| pCHAP8728  | pSU18- <i>pulM</i> N3A, Cm <sup>R</sup>                      | This study                        |
| pMS1297    | pSU18 <i>pulL</i> C264L Cm <sup>R</sup>                      | This study                        |
| pMS1296    | pSU18 <i>pulL</i> 249C C264L, Cm <sup>R</sup>                | This study                        |
| pMS1290    | pSU18 <i>pulL</i> 250C C264L, Cm <sup>R</sup>                | This study                        |
| pMS1291    | pSU18 <i>pulL</i> 251C C264L, Cm <sup>R</sup>                | This study                        |
| pMS1292    | pSU18 <i>pulL</i> 252C C264L, Cm <sup>R</sup>                | This study                        |
| pMS1293    | pSU18 <i>pulL</i> 253C C264L, Cm <sup>R</sup>                | This study                        |
| pMS1294    | pSU18 <i>pulL</i> 254C C264L, Cm <sup>R</sup>                | This study                        |
| pMS1295    | pSU18 <i>pulL</i> 255C C264L, Cm <sup>R</sup>                | This study                        |
| pMS1300    | pSU18 <i>pulL</i> 256C C264L, Cm <sup>R</sup>                | This study                        |
| pCHAP8813  | pUC18 <i>pulM</i> (17C), Ap <sup>R</sup>                     | This study                        |
| pCHAP8893  | pUC18 <i>pulM</i> C17L, Ap <sup>R</sup>                      | This study                        |
| pCHAP8894  | pUC18 <i>pulM</i> C17L L18C, Ap <sup>R</sup>                 | This study                        |
| pCHAP8817  | pUC18 <i>pulM</i> C17L L19C, Ap <sup>R</sup>                 | This study                        |
| pCHAP8818  | pUC18 <i>pulM</i> C17L L20C, Ap <sup>R</sup>                 | This study                        |
| pCHAP8819  | pUC18 <i>pulM</i> C17L G21C, Ap <sup>R</sup>                 | This study                        |
| pCHAP8820  | pUC18 <i>pulM</i> C17L M22C, Ap <sup>R</sup>                 | This study                        |
| pCHAP8821  | pUC18 <i>pulM</i> C17L A23C, Ap <sup>R</sup>                 | This study                        |
| pCHAP8822  | pUC18 <i>pulM</i> C17L V24C, Ap <sup>R</sup>                 | This study                        |
| pUT18C     | T18 <i>cyaA</i> fragment, ColE1 <i>ori</i> , Ap <sup>R</sup> | (Karimova G <i>et al.</i> , 1998) |
| pKT25      | T25 <i>cyaA</i> fragment, p15A <i>ori</i> , Km <sup>R</sup>  | (Karimova G <i>et al.</i> , 1998) |
| pUT18C-Zip | pUT18C fused to yeast Leu zipper motif, Ap <sup>R</sup>      | (Karimova G <i>et al.</i> , 1998) |
| pKT25-Zip  | pKT25 with yeast Leu zipper Km <sup>R</sup>                  | (Karimova G <i>et al.</i> , 1998) |
| pMS1222    | pUT18C- <i>pulL</i> , Ap <sup>R</sup>                        | This study                        |
| pMS1229    | pKT25- <i>pulL</i> , Km <sup>R</sup>                         | This study                        |
| pCHAP8154  | pUT18C- <i>pulM</i> , Ap <sup>R</sup>                        | (Nivaskumar <i>et al.</i> , 2016) |
| pCHAP8155  | pKT25- <i>pulM</i> , Km <sup>R</sup>                         | (Nivaskumar <i>et al.</i> , 2016) |

**Supplementary Table S3:** Oligonucleotides used in this study

| Oligonucleotide name | DNA sequence (5'-3')                 |
|----------------------|--------------------------------------|
| PulL Kpn 5           | GCAGGTACCTATGAATAACCACCATACC         |
| PulL Eco 3           | CACGAATTCCTGTTGCCATAAGGCGAG          |
| PulL L319A For*      | GGCGCCCGGCCGGCGGCTATCTCGCGGCTCGGCGC  |
| PulL L323A For       | CGCTTATCTCGCGGGCCGGCGCGCTGCAGCAAATC  |
| PulL Q327A For       | CGGCTCGGCGCGCTGGCGCAAATCATCGACGACAC  |
| PulL L341A For       | CATCCGCCTGCGGACGGCGAGCTTTGACGCCGCGC  |
| PulM F116A For       | GTGCAGCCCCGCCGATGCCCAGGCGTTGATGGCATG |
| PulM M120A For       | CGATTTCCAGGCGTTGGCGGCATGGCTGGACGCGC  |
| PulM M120D For       | GATTTCCAGGCGTTGGACGCATGGCTGGACGCG    |
| PulM M120K For       | CGATTTCCAGGCGTTGAAGGCATGGCTGGAC      |
| PulM D124A For       | CGTTGATGGCATGGCTGGCCGCGCTGGGGCAGGC   |
| PulM L136A For       | GATGACCACCGCCACCGCGGCGGTGACCGCCGTAG  |
| PulM R88E For        | GACGGTGATTATGGACGAAGCCGCCCGTC        |
| PulM R92D For        | CGCGAAGCCGCCGATCACGGTCTCACCATCG      |
| PulM N3A For         | GGAACGATGCATGCCCTGCTCGCCTTATG        |
| PulL C264L For       | CTGTGGAGCAGTAATCTCC TCCACGATCATC     |
| PulL V249C For       | CGACGCTGGCGCCCGTGTATCGTCGCGGCGCTG    |
| PulL I250C For       | CTGGCGCCCGGTATGCGTCGCGGCGCTG         |
| PulL V251C For       | CTGGCGCCCGGTAATCTGCGCGGCGCTGGCGCTAC  |
| PulL A252C For       | CGCCCGGTAATCGTCTGCGCGCTGGCGCTAC      |
| PulL A253C For       | CGGTAATCGTCGCGTGCCTGGCGCTACTGCTGC    |
| PulL L254C For       | GTAATCGTCGCGGCGTGTGCGCTACTGCTGC      |
| PulL A255C For       | CGTCGCGGCGCTGTGCCTACTGCTGCTGTG       |
| PulL L256C For       | GTCGCGGCGCTGGCGTGTCTGCTGCTGTGGAG     |
| PulM C17L For        | GCCGCGAACGGCTTCTGCTGCTGG             |
| PulM C17L L18C For   | CCGCGAACGGCTTTGTCTGCTGGGTATGG        |
| PulM C17L L19C For   | CCGCGAACGGCTTCTGTGTCTGGGTATGGCC      |
| PulM C17L L20C For   | CGAACGGCTTCTGCTGTGTGGTATGGCCGTGG     |
| PulM C17L G21C For   | CGGCTTCTGCTGCTGTGTATGGCCGTGGTAC      |
| PulM C17L M22C For   | CTTCTGCTGCTGGGTGTGCGGTGGTACTGC       |
| PulM C17L A23C For   | GCTGCTGGGTATGTGCGTGGTACTGCTC         |
| PulM C17L V24C For   | CTGCTGGGTATGGCCTGTGTACTGCTCATCG      |

\* Only the Forward primers used for site-directed mutagenesis are shown. They were used in combination with Reverse primers with a fully overlapping antiparallel sequence.

## Supplementary Materials and Methods

### NMR relaxation experiments

The  $^{15}\text{N}$  relaxation times ( $T_1$  and  $T_2$ ) were estimated at 25 °C on the 600 MHz spectrometer on samples of PulL<sub>CTD</sub> (120  $\mu\text{M}$  of  $^{15}\text{N}$ -PulL<sub>CTD</sub>, pH 6.5), PulM<sub>CTD</sub> (30  $\mu\text{M}$  of  $^{15}\text{N}$ -PulM<sub>CTD</sub>, pH 7.0), PulL<sub>CTD</sub>-PulM<sub>CTD</sub> complex (30  $\mu\text{M}$  of  $^{15}\text{N}$ -PulL<sub>CTD</sub> and 120  $\mu\text{M}$  of unlabeled-PulM<sub>CTD</sub>, pH 6.5) and PulM<sub>CTD</sub>-PulL<sub>CTD</sub> complex (30  $\mu\text{M}$  of  $^{15}\text{N}$ -PulM<sub>CTD</sub> and 120  $\mu\text{M}$  of unlabeled-PulL<sub>CTD</sub>, pH 7.0). Experiments were recorded by standard methods (Barbato *et al.*, 1992) in an interleaved manner with a recycling time of 4.5 s. Two relaxation delays for

$T_1$  (20 and 500 ms) and  $T_2$  (17 and 51 ms) were used with 184 scans. For each residue, the  $T_1$  and  $T_2$  were estimated from the peak intensities ( $Int_1$  and  $Int_2$ ) for the two delays ( $d_{relax1}$  and  $d_{relax2}$ ) by the formula  $(d_{relax1} - d_{relax2})/\ln(Int_2/Int_1)$ . The overall  $T_1$  and  $T_2$  values (and errors) for each sample were calculated as the averages (and standard deviations) over all non-flexible residues.

### **Analytical ultracentrifugation experiments**

Sedimentation velocity experiments were carried out at 20 °C using a Beckman Coulter Optima AUC centrifuge equipped with an AN60-Ti rotor. Protein samples at different concentrations (10  $\mu$ M, 30  $\mu$ M, 300  $\mu$ M) were loaded in a 1.2 cm or 3 mm centerpiece and centrifuged overnight at 42000 rpm. Data were analyzed with SEDFIT 15.1 using a continuous size distribution  $c(S)$  model. The partial specific volume, the viscosity and the density of the samples were calculated with SEDNTERP software. The processed data were used to obtain sedimentation coefficients values at null concentration in our experimental conditions ( $S_0$ ), and to get the standard sedimentation coefficients in buffer (50 mM HEPES pH 7.0, 50 mM NaCl).

### **Supplementary References:**

Barbato, G., Ikura, M., Kay, L., Pastor, R., and Bax, A. (1992). Backbone dynamics of calmodulin studied by  $^{15}\text{N}$  relaxation using inverse detected two-dimensional NMR spectroscopy: the central helix is flexible. *Biochemistry* **31**, 5269-5278. 10.1021/bi00138a005.

Bartolomé, B., Jubete, Y., Martínez, E., and de la Cruz, F. (1991). Construction and properties of a family of pACYC184-derived cloning vectors compatible with pBR322 and its derivatives. *Gene* **102**, 75-78. 10.1016/0378-1119(91)90541-i.

Cisneros, D., Bond, P., Pugsley, A., Campos, M., and Francetic, O. (2012). Minor pseudopilin self-assembly primes type II secretion pseudopilus elongation. *The EMBO journal* **31**, 1041-1053. 10.1038/emboj.2011.454.

Dazzoni, R., López-Castilla, A., Cordier, F., Bardiaux, B., Nilges, M., Francetic, O., and Izadi-Pruneyre, N. (2021).  $^1\text{H}$ ,  $^{15}\text{N}$  and  $^{13}\text{C}$  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.

Jiménez-García, B., Elez, K., Koukos, P., Bonvin, A., and Vangone, A. (2019). PRODIGY-crystal: a web-tool for classification of biological interfaces in protein complexes. *Bioinformatics* **35**, 4821-4823. 10.1093/bioinformatics/btz437.

Jones, D.T. (1999). Protein secondary structure prediction based on position-specific scoring matrices. *J Mol Biol* 292, 195-202. 10.1006/jmbi.1999.3091.

Karimova G, Pidoux J, Ullmann A, and Ladant D (1998). A bacterial two-hybrid system based on a reconstituted signal transduction pathway. *Proc Natl Acad Sci U S A* 95, 5752-5756. 10.1073/pnas.95.10.5752.

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

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.

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.

Possot, O., Vignon, G., Bomchil, N., Ebel, F., and Pugsley, A. (2000). Multiple interactions between pullulanase secreton components involved in stabilization and cytoplasmic membrane association of PulE. *Journal of bacteriology* 182, 2142-2152. 10.1128/JB.182.8.2142-2152.2000.

Robert, X., and Gouet, P. (2014). Deciphering key features in protein structures with the new ENDscript server. *Nucleic acids research* 42, W320-324. 10.1093/nar/gku316.

Santos-Moreno, J., East, A., Guilvout, I., Nadeau, N., Bond, P., Tran Van, N.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.

Zhang, Y., and Skolnick, J. (2005). TM-align: a protein structure alignment algorithm based on the TM-score. *Nucleic Acids Res* 33, 2302-2309. 10.1093/nar/gki524.

3.2.2 Periplasmic coiled coil formed by assembly platform proteins PulL and PulM is critical for function of the *Klebsiella* type II secretion system

Article under review in <Research in Microbiology>

**Periplasmic coiled coil formed by assembly platform proteins PulL and PulM is critical for function of the *Klebsiella* type II secretion system**

Yuanyuan Li<sup>a</sup>, Javier Santos-Moreno<sup>a,#</sup>, Olivera Francetic<sup>a\*</sup>

<sup>a</sup>Institut Pasteur, Université Paris Cité, CNRS UMR3528, Biochemistry of Macromolecular Interactions Unit, F-75015 Paris, France

<sup>#</sup>Present address: University Pompeu Fabra, Department of Medicine and Life Sciences, Barcelona, Spain

Authors' e-mails:

Yuanyuan Li: [yuanyuan.li@pasteur.fr](mailto:yuanyuan.li@pasteur.fr)

Javier Santos-Moreno: [santosmoreno.j@gmail.com](mailto:santosmoreno.j@gmail.com)

\*Correspondence and reprints: Olivera Francetic: [olivera.francetic@pasteur.fr](mailto:olivera.francetic@pasteur.fr)

## ABSTRACT

Bacteria use type II secretion systems (T2SS) to secrete to their surface folded proteins that confer diverse functions, from nutrient acquisition to virulence. In the *Klebsiella* species, T2SS-mediated secretion of pullulanase (PulA) requires assembly of a dynamic filament called pseudopilus. The inner membrane assembly platform (AP) subcomplex is essential for pseudopilus assembly and PulA secretion. AP components PulL and PulM interact with each other through their C-terminal globular domains and transmembrane segments. Here, we investigated the roles of their periplasmic helices, predicted to form a coiled coil, in assembly and function of the PulL–PulM complex. PulL and PulM variants lacking these periplasmic helices were defective for interaction in the bacterial two-hybrid (BACTH) assay. Their functions in PulA secretion and assembly of PulG subunits into pseudopilus filaments were strongly reduced. Interestingly, deleting the cytoplasmic peptide of PulM nearly abolished function of variant PulM $\Delta$ N and its interaction with PulG, but not with PulL, in the BACTH assay. Nevertheless, PulL was specifically proteolyzed in the presence of the PulM $\Delta$ N variant, suggesting that PulM N-terminal peptide binds PulL in the cytoplasm and plays a stabilizing role. We discuss the implication of these results for the T2S pseudopilus and type IV pilus assembly mechanisms.

Keywords:

Pseudopilus; pullulanase; protein secretion; type IV pili; assembly platform

## 1. Introduction

Bacteria have an astonishing capacity to adapt to diverse environments, growth conditions and nutrient sources. Understanding the molecular basis of this adaptation is fundamentally important and might lead to applications in biotechnology and bioremediation. Among the numerous nanomachines that contribute to bacterial adaptation, type II secretion system (T2SS) plays a prominent role (Cianciotto & White, 2017).

T2SS is typically composed of 15 proteins forming a megadalton complex that spans the envelope of Gram-negative bacteria. Recent reviews summarize our current knowledge of T2SS architecture and function (Gu, Shevchik et al., 2017; Korotkov & Sandkvist, 2019; Naskar, Hohl et al., 2021). In the outer membrane, 15 subunits of the GspD protein form the secretin channel that allows the passage of large folded proteins or their complexes to the cell surface. According to the current models, protein secretion is driven by the polymerization of periplasmic filaments called pseudopili in the inner membrane. The core of the pseudopilus is a helical homopolymer composed of the major subunit GspG (Sauvonnnet, Gounon et al., 2000; Durand, Bernadac et al., 2003; Kohler, Schafer et al., 2004). The minor pseudopilins GspH, I, J and K form a complex at the pseudopilus tip (Korotkov & Hol, 2008; Douzi, Ball et al., 2011) and initiate pseudopilus assembly (Cisneros, Bond et al., 2012). The minor pseudopilins are essential for protein secretion and have also been implicated in substrate recognition (Douzi et al., 2011). Pseudopili are continuously polymerized by the inner membrane assembly platform (AP) subcomplex composed of GspF, GspL and GspM proteins (Py, Loiseau et al., 2001). GspL forms a stable complex with the cytoplasmic ATPase GspE (Abendroth, Murphy et al., 2005) and is itself stabilized by GspM (Sandkvist, Hough et al., 1999; Possot, Vignon et al., 2000; Lallemand, Login et al.,

2013). During pseudopilus elongation, GspM promotes targeting of pseudopilins GspG and GspH to the AP (Nivaskumar, Santos-Moreno et al., 2016; Santos-Moreno, East et al., 2017). In the cytoplasm, cycles of ATP binding and hydrolysis promote conformational changes of the hexameric GspE ATPase complex (Patrick, Korotkov et al., 2011; Robien, Krumm et al., 2003). These changes are proposed to induce rotational movements of the central AP component GspF that orchestrates pseudopilus assembly (Nivaskumar, Bouvier et al., 2014).

Deciphering the structural basis of pseudopilus assembly is important to understand how T2SS functions at the molecular level. Structure-function analyses have provided molecular insight into the majority of individual T2SS components. Extending this knowledge towards larger subcomplexes will bring clues on the connectivity between the different nanomachine components and their roles. So far, cryo-electron microscopy (cryo-EM) has been essential to gain insight into the secretin channel structure (Yan, Yin et al., 2017; Hay, Belousoff et al., 2018; Chernyatina & Low, 2019), a large pentadecameric complex with a periplasmic gate. Combining cryo-EM, NMR, modeling and mass spectrometry provided a detailed view of the pseudopilus core (Campos, Nilges et al., 2010; López-Castilla, Thomassin et al., 2017) and tip complexes (Escobar, Douzi et al., 2021).

Recent studies of the *Klebsiella oxytoca* T2SS that secretes pullulanase (PulA) combined X-ray crystallography and NMR to gain information on the structure of a complex formed between C-terminal domains (CTDs) of GspL (PulL) and GspM (PulM) (Dazzoni, Li et al., 2022). Mutations at the interface of these CTDs affect the stability of the two proteins and their function in protein secretion. The model of the PulL – PulM complex (lacking the cytoplasmic domain of PulL) predicts additional interactions *via* periplasmic helices of the two partners. Here we investigated the importance of this

interface for the formation and function of the PulL – PulM complex. In addition, we studied the role of the short cytoplasmic region of PulM in the stability and function of the AP complex.

## **2. Materials and methods**

### *2.1. Bacterial strains and culture conditions*

*Escherichia coli* strain DH5 $\alpha$ F'*lacI*<sup>Q</sup> was used for cloning and transformation experiments. Functional assays were performed in *E. coli* PAP7460 (Possot et al., 2000) and PAP5378 strains. PAP5378 was constructed by P1 transduction by introducing the *ompT::kan* gene from the Keio collection strain BW2511 *ompT::kan* into strain PAP5299 (Cisneros et al., 2012). The *kan* cassette was then deleted thanks to the helper plasmid pCP20 as described (Datsenko & Wanner, 2000) to give strain PAP5299  $\Delta$ *ompT*. The bacterial two-hybrid assays were performed in strain DHT1 (Dautin, Karimova et al., 2000). Bacteria were grown at 30°C or 37°C in LB medium (Miller, 1972) supplemented with antibiotics as required: ampicillin (Ap) at 100  $\mu$ g.mL<sup>-1</sup>, chloramphenicol (Cm) at 25  $\mu$ g.mL<sup>-1</sup>, Kanamycin (Km) at 25  $\mu$ g.mL<sup>-1</sup>. Solid LB media contained 1.5% agar (Difco). Expression of genes under *lacZ* promoter control was induced with 1 mM isopropyl- $\beta$ -1-D-thiogalactopyranoside (IPTG). The *pul* gene promoters were induced with 0.4% D-maltose. For secretion assays, media were buffered with 0.1 volume of M63 salt solution (Miller, 1972).

### *2.2. Plasmid construction and site-directed mutagenesis*

Plasmids used in this study are listed in Table 1. Site-directed mutagenesis was performed using a modified QuickChange protocol, with fully overlapping oligonucleotide primers listed in Table S1 (Eurofins). Template DNA was amplified with

high fidelity DNA polymerase Q5 (New England Biolabs) in two-steps: first, 6 cycles of amplification were performed with forward and reverse primers alone. The reactions were mixed and another 15 cycles of amplification were performed. The DNA was treated with *DpnI* enzyme (Fermentas) and transformed into DH5 $\alpha$ F' ultracompetent cells. Isolated colonies were inoculated into 5 mL of LB with appropriate antibiotics and plasmids were purified using the alkaline lysis protocol and Qiaquick miniprep kit (Qiagen). Plasmids were verified by sequencing (GATC, Eurofins). For the *pulL* bacterial two hybrid constructs, vectors pUT18C and pKT25 were digested with *EcoRI* and *KpnI* and the purified vector fragments were ligated with the PCR amplified *pulL*  $\Delta$ 270-302 digested the same enzymes. For the *pulM* cloning, the BACTH vectors and PCR fragments were digested with *EcoRI* and *BamHI*. Deletion of *pulM* residues

**Table 1. Plasmids used in this study**

| Name      | Relevant characteristics                                                                          | Source/Reference                |
|-----------|---------------------------------------------------------------------------------------------------|---------------------------------|
| pCHAP8185 | <i>pulC-O, pulAB, pulS, Ap<sup>R</sup>, ColE1ori</i>                                              | (Cisneros et al., 2012)         |
| pCHAP8496 | pCHAP8185 $\Delta$ <i>pulM</i>                                                                    | (Santos-Moreno et al., 2017)    |
| pCHAP8251 | pCHAP8185 $\Delta$ <i>pulL</i>                                                                    | (Santos-Moreno et al., 2017)    |
| pCHAP8258 | <i>placZ-pulL, Cm<sup>R</sup>, p15A ori</i>                                                       | (Santos-Moreno et al., 2017)    |
| pMS1349   | pCHAP8258 <i>pulL</i> $\Delta$ 270-302 ( <i>pulL</i> <sup>ACC</sup> ), Cm <sup>R</sup> , p15A ori | This study                      |
| pCHAP1353 | <i>placZ-pulM, Cm<sup>R</sup>, p15A ori</i>                                                       | (Possot et al., 2000)           |
| pMS1350   | pCHAP8258 <i>pulM</i> $\Delta$ 37-68 ( <i>pulM</i> <sup>ACC</sup> ), Cm <sup>R</sup> , p15A ori   | This study<br>This study        |
| pCHAP8882 | <i>pulM</i> $\Delta$ 2-15 ( <i>pulM</i> $\Delta$ N) Cm <sup>R</sup> , p15A ori                    | (Karimova, Pidoux et al., 1998) |
| pUT18C    | <i>cyaT18 Ap<sup>R</sup>, ColE1ori</i>                                                            | (Karimova et al., 1998)         |
| pKT25     | <i>cyaT25 Km<sup>R</sup>, p15A ori</i>                                                            |                                 |
| pMS1222   | <i>T18-pulL Ap<sup>R</sup>, ColE1ori</i>                                                          | (Dazzoni et al., 2022)          |
| pMS1229   | <i>T25-pulL, Km<sup>R</sup>, p15A ori</i>                                                         | (Dazzoni et al., 2022)          |
| pCHAP8154 | <i>T18-pulM, Ap<sup>R</sup>, ColE1ori</i>                                                         | (Nivaskumar et al., 2016)       |
| pCHAP8155 | <i>T25-pulM, Km<sup>R</sup>, p15A ori</i>                                                         | (Nivaskumar et al., 2016)       |
| pMS1355   | <i>T18-pulL</i> $\Delta$ 270-302, Ap <sup>R</sup> , ColE1ori                                      | This study                      |
| pMS1356   | <i>T25-pulL</i> $\Delta$ 270-302, Km <sup>R</sup> , p15A ori                                      | This study                      |
| pMS1357   | <i>T18-pulM</i> $\Delta$ 37-68 Ap <sup>R</sup> , ColE1ori                                         | This study                      |
| pMS1358   | <i>T25-pulM</i> $\Delta$ 37-68 Km <sup>R</sup> , p15A ori                                         | This study                      |
| pCHAP7330 | <i>T18-pulG, Ap<sup>R</sup>, ColE1ori</i>                                                         | (Nivaskumar et al., 2014)       |
| pCHAP7332 | <i>T25-pulG, Km<sup>R</sup>, p15A ori</i>                                                         | (Nivaskumar et al., 2014)       |
| pMS1147   | <i>T18-pulM</i> $\Delta$ 2-15 Ap <sup>R</sup> , ColE1ori                                          | This study                      |
| pMS1148   | <i>T25-pulM</i> $\Delta$ 2-15 Km <sup>R</sup> , p15A ori                                          | This study                      |

2-15 in plasmid pCHAP8882 was performed by QuickChange mutagenesis using the primers PulM delN5 and PulM delN3 (Table S1). The same approach was used to introduce PulM2-15 deletion in the BACTH constructs pCHAP8154 and pCHAP8155 with primers PulM delN For and PulM delN Rev2 (Table S1).

### 2.3. *Protein electrophoresis and Western blot analysis*

Total bacterial extracts were analyzed on denaturing sodium dodecyl-sulphate polyacrylamide gel electrophoresis (SDS-PAGE) using the Tris-tricine gels (Schägger & von Jagow, 1987). Proteins were transferred on nitrocellulose membranes using the Fast Blot semi-dry transfer system and one-step transfer buffer (Thermo). Membranes were blocked with 5% skim milk in Tris-buffer saline containing 0.05% Tween-20 (TBST) for one hour and probed with primary antibodies diluted in TBST. Anti-PulM and anti-PulL sera were used at 1:1000 dilution and anti-PulA and anti PulG sera were diluted at 1:2000. After 4 washes for 10 minutes in TBST, membranes were incubated 1 hour in the secondary goat anti-rabbit antibodies coupled to horse-radish peroxidase (Amersham) (1:10000). Fluorescence signals were developed with ECL2 (Thermo) and recorded on Typhoon FLA9000 (GE, Cytiva).

### 2.4. *PulA secretion assays*

Secretion of the non-acylated variant of PulA was assessed by cell fractionation in strain PAP7460 harboring plasmid pCHAP8251 complemented with pCHAP8258 derivatives carrying *pulL* variants or pCHAP8496 complemented with pCHAP1353 derivatives carrying *pulM* variants. Bacteria were cultured overnight at 30°C in LB containing Ap and Cm, then inoculated in inducing medium containing in addition 0.4 % D-maltose and 0.1 vol of M63 salts for another 4 hours. Cultures were normalized to

OD<sub>600nm</sub> of 1. Bacterial pellets were collected after 5-min centrifugation at 16000 x g in a table-top Eppendorf centrifuge at 4°C and resuspended in an equal volume of Laemmli sample buffer. Supernatants were subjected to another 10-min round of centrifugation and 0.1 mL was taken off top, then mixed with 0.1 mL of 2 x Laemmli sample buffer. Cell and supernatant fractions from the same amount of cultures were analyzed by SDS-PAGE on 10% Tris-glycine gels followed by Western blot with anti-PulA antibodies. PulA bands were quantified with ImageJ and the fraction of PulA in the supernatant was calculated. Data were plotted and analyzed with GraphPad Prism9 software.

## 2.5. *PulG pilus assembly assays*

To quantify the PulG pilus assembly, bacteria of strain PAP7460 harboring the *pul* operon on plasmid pCHAP8185 derivatives were cultured for 48 or 72 hours at 30°C on LB plates containing 1.5 % agar (Difco), appropriate antibiotics and 0.2% D-maltose. Bacteria were resuspended in 1 mL of LB and the cell density was normalized to OD<sub>600nm</sub> of 1. The suspensions were vigorously vortexed for 1 min to detach surface pili. Bacterial pellets were then collected by 5-min centrifugation at 4°C and 16000xg and resuspended in Laemmli sample buffer at the concentration of 10 OD<sub>600nm</sub>.mL<sup>-1</sup>. The pili-containing supernatants were cleared from remaining bacteria by a 10-min centrifugation at 16000xg and 0.7 mL was mixed with tri-chloro-acetic acid (TCA) at a final concentration of 10%. After a 30-min precipitation on ice, the samples were centrifuged for 30 min at 4°C and 16000xg. Pellets were washed twice with acetone and air dried, then resuspended in 70 µL of Laemmli sample buffer. Equivalent volumes were analyzed by SDS-PAGE on 10 % Tris-tricine gels followed by Western

blot with anti-PulG antibodies. PulG bands were quantified using ImageJ and the results were analyzed and plotted with Prism 9 software (GraphPad).

## 2.6. *Bacterial two-hybrid assays*

Bacterial two-hybrid (BACTH) assay was used to study interactions between proteins fused to T18 and T25 fragments of the CyaA catalytic domain (Karimova et al., 1998). Plasmids were co-transformed into DHT1 competent cells and grown on LB Ap Km plates for 48-60 hours at 30°C. Independent single co-transformants were picked at random, inoculated into 1 mL of LB Ap Km and grown overnight at 30°C. The overnight precultures were used to inoculate 1-mL cultures of LB Ap Km containing 1 mM IPTG, which were incubated 4-6 hours at 30°C with shaking. The cultures were used to measure beta-galactosidase activity as described by (Miller, 1972). The data were plotted and analyzed using the Prism 9 software (GraphPad).

### 3. Results

#### 3.1. The periplasmic helices of PulL and PulM are important for T2SS function

PulL and PulM share a similar architecture of their transmembrane and periplasmic domains (Fig. 1A). Their hydrophobic transmembrane helices extend into the periplasm and are followed by the flexible linkers connected to the globular ferredoxin-like C-terminal domains (CTDs). The two periplasmic helices are predicted to form a coiled coil (Dazzoni et al., 2022). To test their functional importance, we generated

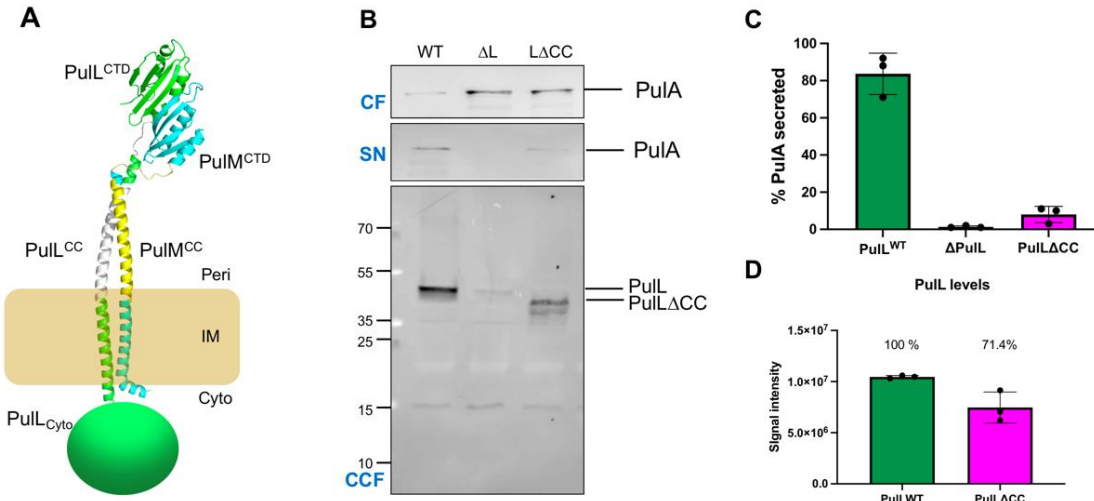


**Figure 1.** Deletion of the PulL periplasmic helix affects PulA secretion. **A.** Schematic view of the PulL-PulM dimer in the inner membrane (IM). PulL is shown in green with its periplasmic helix in white; PulM is shown in cyan and its periplasmic helix in yellow. **B.** Secretion of PulA in strains producing native PulL (WT), no PulL ( $\Delta L$ ) and PulL $\Delta CC$ . PulA secretion assay was performed (Materials and Methods) and 0.005 OD<sub>600nm</sub> equivalents of cell- (CF) and supernatant fractions (SN) were analyzed by Western blot with anti-PulA antibodies. In the lower panel, the concentrated cell fractions (CCF) (from 0.05 OD<sub>600nm</sub> of bacteria) were analyzed with anti-PulL antibodies. **C.** PulA secretion efficiency quantified from 3 independent assays as in (B). The column height shows the mean values and dots represent values from independent experiments. **D.** Bar graphs indicate mean signal intensities of bands detected with anti-PulL antibodies in strains producing PulL<sup>WT</sup> and PulL $\Delta CC$ . Mean levels of PulL $\Delta CC$  signal relative to PulL<sup>WT</sup> are indicated above the bar. Black dots indicate signal levels from 3 independent measurements and error bars show standard deviation.

PulL and PulM variants lacking these periplasmic helices. In PulL, we removed the region comprising residues 270 to 302 (Fig. 1A) to yield variant PulL<sup>ΔCC</sup>, encoded by plasmid pMS1349 (Table 1). We first tested the ability of PulL<sup>ΔCC</sup> to promote secretion of the nonacylated variant of pullulanase (PulA), the substrate of the *K. oxytoca* T2SS (Michaelis, Chapon et al., 1985). Whereas the native PulL (PulL<sup>WT</sup>) promoted efficient PulA secretion, the majority of the PulA pool accumulated in the cell fraction in the presence of PulL<sup>ΔCC</sup> (Fig. 1B). The intracellular accumulation of PulA was complete in the negative control lacking PulL. We used antibodies directed against PulL<sup>CTD</sup> to assess PulL stability in these strains (Fig. 1B, lower panel labeled CCF). Quantification of PulL<sup>ΔCC</sup> signals showed that protein levels were reduced by about 30% compared to the native PulL (Fig. 1D).

To study the role of the periplasmic helix of PulM, we removed the region comprising residues 37 to 68 (Fig. 1A) resulting in variant PulM<sup>ΔCC</sup> encoded by plasmid pMS1350 (Table 1). Similar to the PulL case, this deletion resulted in a strong PulA secretion defect (Fig. 2A, 2B). PulA accumulated in the cell fraction in the presence PulM<sup>ΔCC</sup> at levels similar to the negative control (Fig. 2A). PulM<sup>ΔCC</sup> stability was reduced by 10% on average compared to PulM (Fig. 2C).

### 3.2. *Pilus assembly defects in pulL<sup>ΔCC</sup> and pulM<sup>ΔCC</sup> mutants*

The AP components PulL and PulM directly participate in pseudopilus assembly, which is thought to drive PulA secretion. We therefore asked whether the periplasmic helices of PulL and PulM are required for PulG pilus assembly. Under conditions where bacteria are cultured on solid media, *E. coli* harboring a moderate copy-number plasmid encoding a complete set of *pul* genes produces surface pili composed of PulG,

the major pseudopilin subunit (Sauvonnet, Vignon et al., 2000). We tested the ability of PulL<sup>ΔCC</sup>

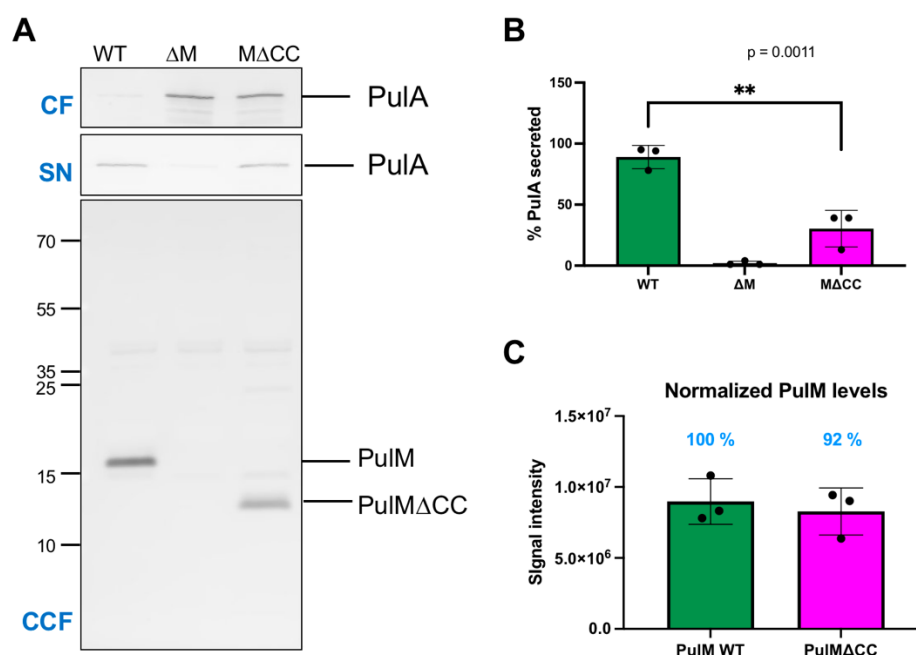


**Figure 2.** PulM periplasmic helix region is required for full T2SS function. **A.** Secretion of PulA in strains producing native PulM (WT), no PulM ( $\Delta M$ ) and PulM<sup>ΔCC</sup>. PulA secretion assay was performed (Materials and Methods) and 0.005 OD<sub>600nm</sub> equivalents of cell- (CF) and supernatant fractions (SN) were analyzed by Western blot with anti-PulA antibodies. Lower panel, the concentrated cell fractions (CCF) (0.05 OD<sub>600nm</sub>) analyzed with anti-PulM antibodies. **B.** PulA secretion efficiency quantified from 3 independent assays in the presence of PulM (WT) or PulM<sup>ΔCC</sup> and without PulM ( $\Delta M$ ). The column height shows the mean value and dots represent values from independent experiments. **C.** Normalized PulM levels. Bar heights indicate mean signal intensity of bands detected with anti-PulM antibodies in strains producing PulM<sup>WT</sup> and PulM<sup>ΔCC</sup>. The mean signal intensity values of PulM<sup>ΔCC</sup> relative to PulM<sup>WT</sup> are indicated above the bars. Black dots show signal intensities from 3 independent experiments and error bars show standard deviations.

and PulM<sup>ΔCC</sup> to complement this function in mutants lacking plasmid-encoded *pulL* (pCHAP8251) or *pulM* (pCHAP8496). After 2 days of growth in the presence of maltose to induce the *pul* gene expression, bacteria complemented with wild type *pulL* and *pulM* genes produced pili that could be sheared from their surface and separated from the cell-associated PulG pool (Fig. 3). In the presence of PulL<sup>ΔCC</sup> and PulM<sup>ΔCC</sup> variants,

the PulG pilus assembly was nearly abolished. When bacteria were cultured for a standard period of 48 hrs for this assay, no PulG could be detected in the pilus fractions (Fig. S1). Only upon extending the growth period for another 24 hours, we could detect some PulG pili on the surface of the coiled coil deletion mutants (Fig. 3). In addition, we noticed that cellular PulG levels were reduced in the absence of PulL or PulM (Fig. 3B).

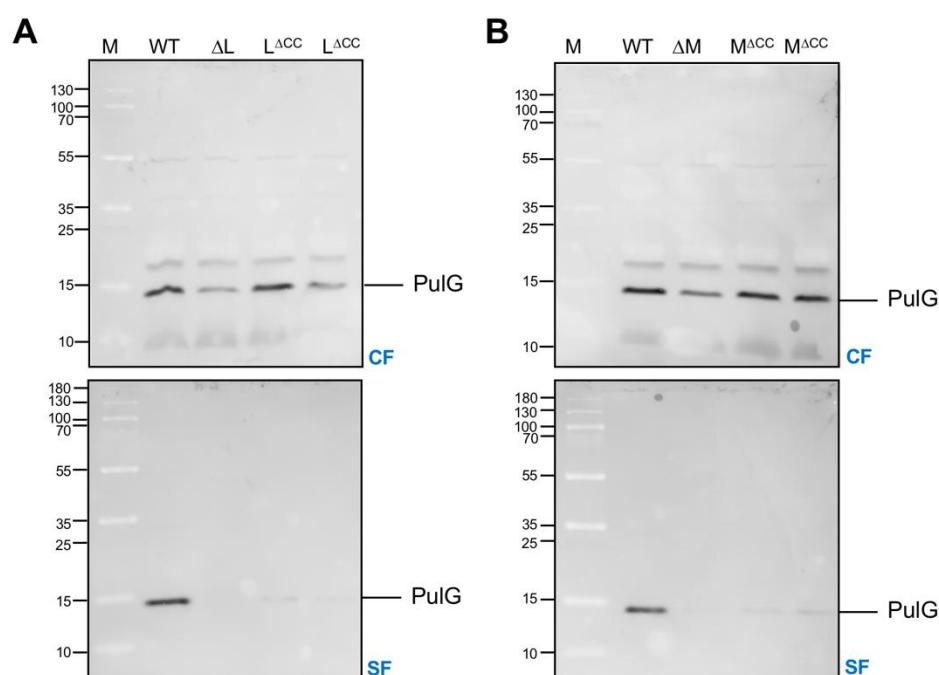


**Figure 3.** Piliation defect of PulL and PulM lacking periplasmic helices. **A.** PulG pilus assembly on the surface of PAP7460 bacteria containing plasmid pCHAP8251 complemented with pCHAP8252 (PulL<sup>WT</sup>), pMS1348 (PulL<sup>ΔCC</sup>) or vector pSU19 (ΔL). Migration of Mw markers (M) (in kDa) is indicated on the left. **B.** PulG pilus assembly on the surface of PAP7460 bacteria containing plasmid pCHAP8496 complemented with pCHAP1353 (PulM<sup>WT</sup>), pMS1350 (PulM<sup>ΔCC</sup>) or vector pSU18 (ΔM). Cell- (CF) and sheared fractions (SF) were prepared from bacteria cultured for 72 h at 30°C as described in Materials and Methods and 0.05 OD<sub>600nm</sub> equivalents were analyzed on 10 % Tris-Tricine SDS PAGE and western blot using anti-PulG antibodies.

### 3.3. The periplasmic coiled coil is required for assembly of the PulL–PulM complex

To determine whether the deletion of periplasmic helices affects the PulL–PulM interaction, we used the bacterial two-hybrid (BACTH) assay based on the reconstitution of CyaA adenylyl cyclase signaling (Karimova et al., 1998). Full-length

PuL, PuM and their variants were fused to the C-termini of T18 and T25 CyaA fragments (Table 1). Constructs were co-transformed in *E. coli* cya mutant strain DHT1 to evaluate the reconstitution of adenylyl cyclase activity. By measuring the beta-galactosidase activity, we compared interactions of full-length PuL and PuM constructs described previously (Dazzoni et al., 2022) with constructs harboring the PuL<sup>ΔCC</sup> and PuM<sup>ΔCC</sup> variants. For each tested couple, activity was measured for 8 independent co-transformants picked at random (Fig. 4).

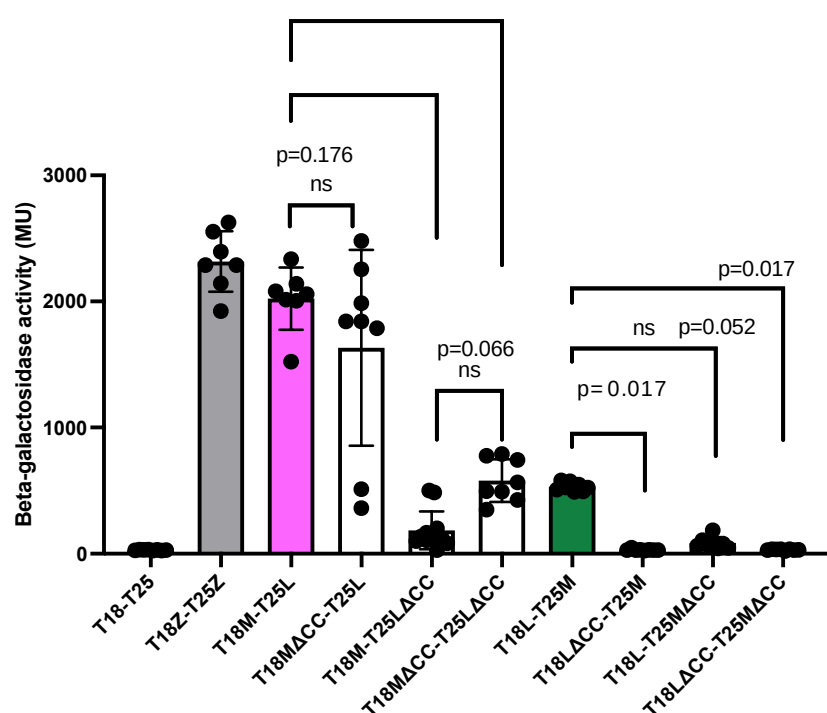


**Figure 4.** Periplasmic helices are required for PuL – PuM interaction. Beta-galactosidase activity (Miller units) of DHT1 bacteria co-transformed with plasmids producing T18 and T25 or their hybrids as indicated: Z, yeast leucine zipper; M (PuM), L (PuL); M<sup>ΔCC</sup> (PuM<sup>ΔCC</sup>) and L<sup>ΔCC</sup> (PuL<sup>ΔCC</sup>). The bar graph heights indicate mean values and error bars show standard deviation. Black dots show β-gal. activities of individual cultures. Statistical analysis was performed with GraphPad Prism 9 using One-way ANOVA test with multiple comparisons. \*\*\*\*, p<0.0001. Exact p values are indicated above the bars.

Bacteria containing plasmids encoding T18-PuM and T25-PuL showed high beta-galactosidase activity, comparable to the yeast leucine zipper positive control,

indicating efficient heterodimer formation. In strains harboring T18-PulM<sup>ACC</sup> and T25-PulL, the mean activity remained high, suggesting that PulM<sup>ACC</sup> interaction with PulL is similar to that of PulM<sup>WT</sup>. However, in several isolates the activity was strongly reduced, indicating that removing the PulM periplasmic helix affected the efficiency of binding to PulL. Deleting the PulL helix in T25-PulL<sup>ACC</sup> had a much stronger effect and nearly abolished the interaction with T18-PulM. In comparison, combining the two deletions in T18-PulM<sup>ACC</sup> and T25-PulL<sup>ACC</sup> increased the mean activity, presumably by placing the respective CTDs at the same level relative to the membrane to facilitate their contacts. However, this increase was weak and not statistically significant in this experimental context.

Strains harboring the T18-PulL and T25-PulM hybrids showed weaker  $\beta$ -galactosidase activity overall. This difference is probably due to the lower stability of T18-PulL compared to T18-PulM hybrids. Beta-galactosidase activity was reduced to background levels upon removal of the periplasmic helix residues in T18-PulL<sup>ACC</sup> or T25-PulM<sup>ACC</sup>. Combining the two deletions in this context did not measurably improve the interaction. Overall, these data suggest that the periplasmic coiled coil - forming regions play a major role in the PulL-PulM assembly.

#### *3.4. Periplasmic helix of PulM is not required for binding to PulG*

PulM has been characterized as a targeting factor for PulH and PulG subunits during pseudopilus elongation (Nivaskumar et al., 2016; Santos-Moreno et al., 2017). Since the deletion of the PulM periplasmic helix affected pseudopilus assembly, we asked whether it also affected the interactions with PulG. As shown in the BACTH assay results (Fig. 5), although deletion of the PulM periplasmic helix reduced the interaction of the pair T18-PulM – T25-PulG, the PulM<sup>ACC</sup> variant could still interact with

PulG strongly and significantly relative to the negative control. In addition, there was no effect on the interaction with T18-PulG. Previous studies have shown that Pull does not interact with PulG in the BACTH assay (Nivaskumar et al., 2016). We confirmed this result here, by showing that  $\beta$ -galactosidase activities of the bacteria co-expressing *pulG* and *pull* in BACTH constructs were at the level of the negative control. Like the  $\text{Pull}^{\text{WT}}$ , the  $\text{Pull}^{\Delta\text{CC}}$  variant did not interact with PulG (Fig. 5). Based on these data, we concluded that it is unlikely that the piliation defect caused by  $\text{Pull}^{\Delta\text{CC}}$  is due to a defect in pseudopilin binding.

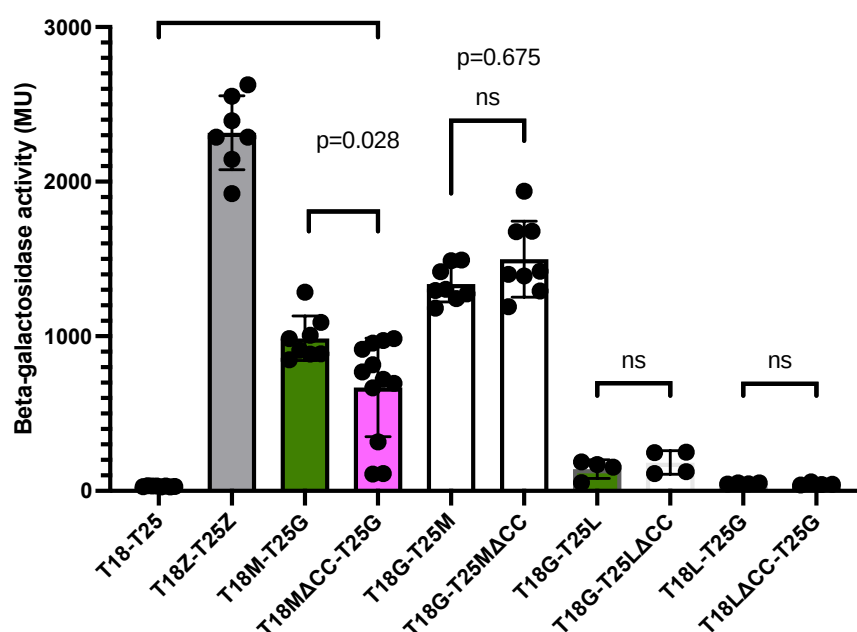


**Figure 5.** PulM periplasmic  $\alpha$ -helix is not essential for interaction with PulG. Beta galactosidase activity measured for indicated pairs of T18- and T25- hybrids. Z, yeast leucine zipper; M (PulM), L (Pull); M $\Delta$ CC (PulM $\Delta$ CC), L $\Delta$ CC (Pull $\Delta$ CC) and G (PulG). Bar graphs indicate mean values and error bars standard deviation. Black dots show activity values of independent cultures. The data were plotted and analysed with GraphPad Prism 9 software.

### 3.5. The role of the cytoplasmic PulM peptide

The cytoplasmic domain of Pull interacts stably with the ATPase PulE, based on the study of their homologues in *Vibrio cholerae* (Abendroth et al., 2005). The

cytoplasmic region of PulM is a 16-residue peptide, presumably functioning as a positively charged membrane anchor. We hypothesized that this region interacts with the polar N-terminal residues of PulG, as their substitutions in variants PulG<sup>T2A</sup> and PulG<sup>E5A</sup> strongly reduce interactions with PulM (Nivaskumar et al., 2016; Santos-Moreno et al., 2017). To test this hypothesis, we deleted the cytoplasmic peptide of PulM, and studied interactions of this variant, designated PulM $\Delta$ N, with PulL and PulG. In the BACTH assay, PulM $\Delta$ N variant showed a strong and significant defect in binding to PulG, confirming our predictions (Fig. 6A). This defect was comparable to that of

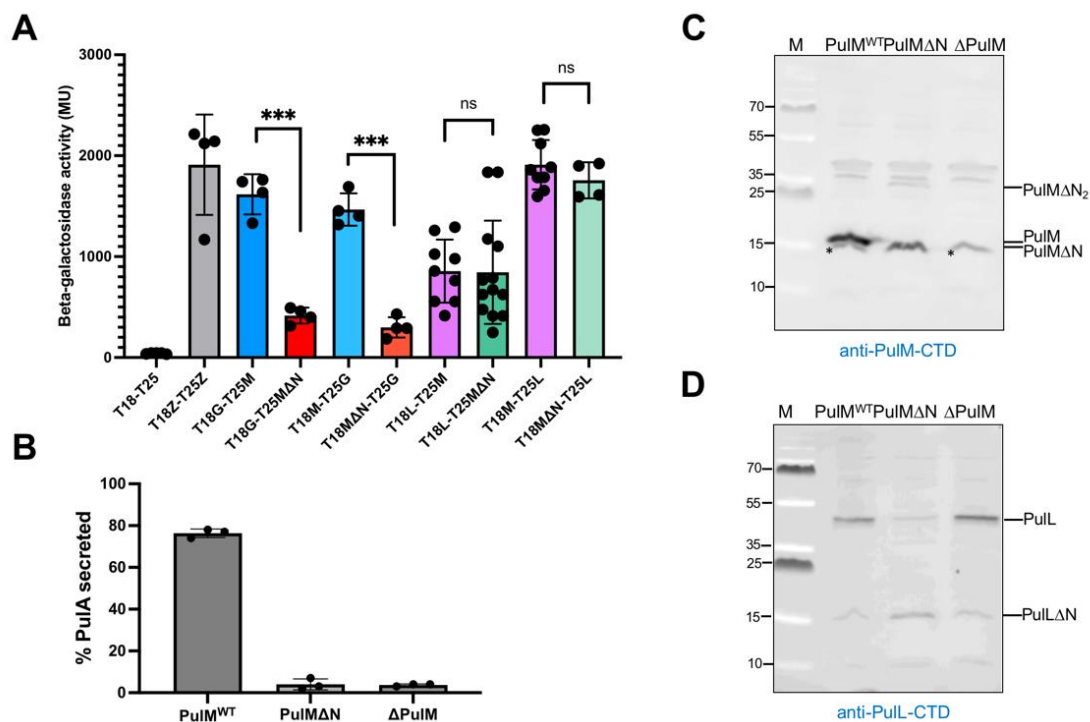


**Figure 6.** PulM cytoplasmic peptide is essential for interaction with PulG and for the PulL stability. **A.** BACTH analysis of PulM and PulM $\Delta$ N interactions with PulG and PulL. **B.** PulA secretion in the presence of PulM $\Delta$ N is fully defective. Each data point represents % of PulA secretion in strain PAP5378 containing plasmid pCHAP8496 ( $\Delta$ pulM) complemented with *pulM*<sup>WT</sup> (on pCHAP1353), *pulM* $\Delta$ N (on pCHAP8882) and empty vector (pSU18). Secretion assay was performed as described in Materials and Methods. **(C)** PulM (top) and **(D)** PulL (bottom) levels in strain PAP7460 carrying pCHAP8496 ( $\Delta$ pulM) complemented with *pulM*<sup>WT</sup> (on pCHAP1353), empty vector (pSU18) and *pulM* $\Delta$ N (on pCHAP8882). The asterisk in **(C)** marks a nonspecific band cross-reacting with the antibody.

PulM binding to the PulG<sup>E5A</sup> variant, characterized previously (Santos-Moreno et al., 2017). At the same time, deleting the PulM cytoplasmic peptide did not affect its interactions with PulL in this assay.

We tested the ability of PulM $\Delta$ N to promote PulA secretion and found that nearly all PulA remained cell bound in the secretion assay, similar to the negative control lacking PulM (Fig. 6B). This phenotype is probably caused by the defective interaction of PulM $\Delta$ N with PulG. Although the levels of PulM $\Delta$ N variant were reduced compared to PulM<sup>WT</sup>, the protein was still well produced and migrated on SDS-PAGE according to its expected Mw (~16.5 kDa) (Fig. 6C). Interestingly, this variant seemed to form some SDS-resistant dimers that were not detectable with PulM<sup>WT</sup>. We also tested the levels of PulL in these strains. Surprisingly, although the N-peptide deletion of PulM did not affect its interaction with PulL in the BACTH assay, we found that PulL was strongly destabilized in the presence of PulM $\Delta$ N, resulting in a proteolytic product of a size similar to that of PulM (Fig. 6D). Antibodies directed against the PulL<sup>CTD</sup> detected this product, suggesting that it comprises the inner membrane and periplasmic regions of PulL. As it appears to be generated by a specific cleavage that removes the PulL cytoplasmic domain, we designated this protein PulL $\Delta$ N. These results suggest that the PulM cytoplasmic peptide not only ensures PulG targeting to the assembly site through direct interaction, but also protects PulL from specific proteolysis.

#### **4. Discussion**

Structural information on the PulL<sup>CTD</sup>–PulM<sup>CTD</sup> heterodimer, together with the BACTH and cysteine crosslinking data, allowed Dazzoni and collaborators to propose a structural model of membrane-embedded PulL–PulM complex (Fig. 1A) (Dazzoni et

al., 2022). Sequence-based predictions (Ludwiczak, Winski et al., 2019) and biochemical data (Dazzoni et al., 2022) suggest that periplasmic helices of PulL and PulM interact together to form a coiled coil. Our mutational, interaction and functional studies described here support this model and show that this coiled coil plays a major role in PulL–PulM assembly. Deleting residues 270-302 of PulL abolished interaction with PulM in the BACTH assay, and removal of PulM residues 37-68 reduced or abolished PulL binding, depending on the plasmid construct and expression levels. In structural models of PulL and PulM, these deletions would shorten their periplasmic helices by 4.8 nm. This might preclude interaction of ferredoxin-like CTDs of the deletion variant and the native, full-length partner. Consistent with this model, the BACTH interaction signal was somewhat improved when the two deletions were combined in T18-PulM<sup>ΔCC</sup> and T25-PulL<sup>ΔCC</sup>. However, this interaction remained weak overall and was not restored to original levels, suggesting that deleted regions play more than just an alignment role. The binding of PulL<sub>CTD</sub> to PulM<sub>CTD</sub> is weak and dynamic in solution (Dazzoni et al., 2022). Our results show that this interaction is also weak even in the context of PulL<sup>ΔCC</sup> or PulM<sup>ΔCC</sup> variants, where the two CTDs are anchored in the membrane *via* transmembrane segments followed by their flexible linkers. This suggests that the main role of the coiled coil is to stimulate PulM binding to PulL.

When isolated and in solution, PulL<sub>CTD</sub> and PulM<sub>CTD</sub> can form different types of parallel and anti-parallel homodimers, as shown by X-ray crystallography (Dazzoni et al., 2022). Similar behavior was found for the CTDs of EpsM, the PulM homologue from *V. cholerae* (Abendroth, Rice et al., 2004) and for XcpL, the PulL homologue from *P. aeruginosa* (Fulara, Vandenberghe et al., 2018). Likewise, homo-oligomerization was observed in the ferredoxin-like domains of T4P assembly proteins PilN and PilO,

considered as orthologues of PulL transmembrane and periplasmic regions and PulM, respectively. The CTD of PilN from *Thermus thermophilus* (Karuppiah, Collins et al., 2013) and soluble periplasmic regions of *P. aeruginosa* PilO and PilN formed homodimers in crystals (Sampaleanu, Bonanno et al., 2009). This is in contrast to their poor oligomerization *in vivo* when they are anchored in the bacterial membrane *via* their transmembrane segments, a constraint which reduces the degrees of freedom for CTDs binding.

The recent BACTH data indicate that PulL and PulM preferentially form heterodimers, with a parallel orientation compatible with their membrane insertion (Dazzoni et al., 2022). Similarly, full-length *P. aeruginosa* PilN and PilO form heterodimers, with no homodimers observed in the BACTH assay (Leighton, Dayalani et al., 2015). The long periplasmic helices thus seem to provide a compatible interacting surface to promote efficient formation of the heterodimer coiled coil. Studies in other systems also support the role precise and dynamic interactions between AP components. Mutations in the coiled coil regions of PilN and PilO in the *P. aeruginosa* T4P assembly system cause defects in pilus retraction (Leighton, Yong et al., 2016). In the *Dickeya dadanti* T2SS, the periplasmic regions of OutL and OutM comprise the major interacting interfaces revealed by BACTH and copurification studies, although the results of these studies suggested a less prominent role for the coiled coil regions (Lallemand et al., 2013).

In the Pul T2SS, deletions of the PulL and PulM coiled coil regions strongly affected the function. Although variant PulM<sup>ACC</sup> was more stable and retained the capacity to interact with PulL in the BACTH assay, it was as defective as PulL<sup>ACC</sup> for both PulA secretion and PulG pilus assembly, suggesting an impaired structure and function of the PulL–PulM complex. The coiled coil might provide structural rigidity to the complex

compared to the individual proteins. Shortening of PulL and PulM helices by 4.8 nm could also result in defective interactions with other T2SS components or cellular structures, such as the peptidoglycan layer. In PulL<sup>ΔCC</sup> and PulM<sup>ΔCC</sup> variants, the globular ferredoxin-like domains (CTDs) would be placed proximal to the membrane, connected *via* the flexible linkers to the transmembrane helices. This might affect the overall architecture of the T2SS and interactions with its substrates.

While we showed here that PulM<sup>ΔCC</sup> retained its ability to bind PulG, the next step in pseudopilus assembly, possibly involving binding to PulL, might be impaired. Like other pilins and pseudopilins, the TM segment of PulG is alpha-helical in its membrane-embedded state, but undergoes stretching upon incorporation into the pseudopilus (López-Castilla et al., 2017). This results in the loss of secondary structure in the region between Gly14 and Pro22 of PulG, which adopts an extended coil conformation (López-Castilla et al., 2017). The force that causes this stretching is likely to be generated during PulG extraction from the membrane, a key step in pseudopilus assembly which involves PulG binding to PulM (Santos-Moreno et al., 2017). The PulL–PulM coiled coil might provide an anchoring point during this step. The rigidity or the length of PulL–PulM dimer might be essential to exert force on PulG and operate its stretching during membrane extraction. The Pule ATPase in complex with the PulL cytoplasmic domain might be directly involved in this process. A defect in membrane extraction could be the cause of the major piliation defect observed in *pulL*<sup>ΔCC</sup> and *pulM*<sup>ΔCC</sup> mutants.

Our BACTH data suggest that the PulM cytoplasmic region is not required for interaction with PulL, and that the periplasmic regions play a dominant role. The specific cleavage of PulL cytoplasmic domain in the *pulM*ΔN mutant was therefore surprising. We hypothesize that the N-terminal peptide of PulM protects PulL from

proteolysis by binding to a disordered linker that connects PulL<sub>cyto</sub> domain with the transmembrane and periplasmic regions. This linker is present in the Pull model retrieved from the AF2 database (Varadi, Anyango et al., 2022) and predicted by the Alpha Fold 2 algorithm (Jumper, Evans et al., 2021) (Fig. S2).

Binding to the disordered Pull linker might require the free N-terminus of PulM, and thus might occur in the native context, but not in the BACTH constructs, where PulM N-terminus is fused to the T18 and T25 CyaA fragments. A similar interaction is observed in T4P assembly systems, between the conserved N-terminal peptide of PilN and PilM, which is a paralogue of the PulL<sub>cyto</sub> domain (Karupiah & Derrick, 2011; McCallum, Tammam et al., 2016). This contact is highly important for pilus assembly (Georgiadou, Castagnini et al., 2012) and involves a conserved sequence ΦNLLP (where Φ is a V, I or L) at the N-terminus of PilN homologues. Structural studies have shown that this N-terminal peptide fits into a groove in PilM, allowing them to form a stable PilM–PilN complex – a functional equivalent of Pull. The conserved Asn residue (N5 in PilN of *T. thermophilus*) plays a major role in PilM binding by forming a network of hydrogen bonds with several residues in the PilM groove (Karupiah & Derrick, 2011). Functional data support the crucial role of its equivalent Asn8 in the *Neisseria meningitidis* PilN, reflected by the strong piliation defect of the PilN<sup>N8A</sup> variant (Georgiadou et al., 2012). The presence of a similar N-terminal sequence in PulM, **MHNLL** (identical residues shown in bold) led us to test the effect of N3A substitution in PulM. The PulM<sup>N3A</sup> variant showed a strong secretion defect and affected the stability of both PulM and Pull (Dazzoni et al., 2022). Similar effects were observed in the PulM $\Delta$ N variant described here. The *pulM* $\Delta$ N mutants show reduced PulM and Pull stability, generating a distinct Pull fragment corresponding in size and likely in domain organization to the PilN component of T4P assembly systems. Of note, a

similar PulL degradation product is also observed in the presence of native PulM (Dazzoni et al., 2022), suggesting that the cytoplasmic linker region of PulL is dynamic and protease sensitive, even in the presence of its interacting partner(s).

The above results argue that, although PulM is often considered as a paralogue of PilO in T4P systems, its role may be more similar to that of PilN. Importantly, direct interaction of PulM with the major pseudopilin PulG parallels the binding of the PilN homologue HofN to the major pilin PpdD in the T4P assembly system of Enterohaemorrhagic *E. coli* (Luna Rico, Zheng et al., 2019).

The N-terminal regions of PulM and PilN might therefore play analogous roles in pilin targeting to the assembly site. The phenotype of the PulM $\Delta$ N variant, strongly defective in PulG binding and affecting PulL stability, suggests a role of the cytoplasmic PulM peptide in passing the pseudopilin subunits to PulL for incorporation into the growing pilus. PulM is also more abundant in the cell relative to PulL (Dazzoni et al., 2022) consistent with its role as a shuttle between the large membrane pool of free PulG and the PulL protein stably bound to the ATPase PulE at the pseudopilus assembly site. Crosslinking studies of *V. cholerae* T2SS provided evidence for a transient complex of EpsL with EpsG (Gray, Bagdasarian et al., 2011), supporting at least a transient formation of G-M-L tripartite complex. Clearly, further studies are needed to identify their precise interactions within this complex and conformational changes of partners that promote this key step during pseudopilus assembly.

### **Conflict of interest statement**

The authors declare that they have no conflict of interest.

## Acknowledgments

We thank N. Izadi-Pruneyre, I. Guilvout, E. Bouveret and members of the BIM unit for helpful discussions. YL was funded by the Pasteur Paris University (PPU) international PhD program. JSM was funded by a fellowship from the Basque Government. This work was funded by Institut Pasteur, CNRS and ANR grant Synergy-T2SS ANR-19-CE11-0020-01.

## References

- Abendroth J, Murphy P, Sandkvist M, Bagdasarian M, Hol WG (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*. J Mol Biol 348: 845-55.
- Abendroth J, Rice AE, McLuskey K, Bagdasarian M, Hol WG (2004) The crystal structure of the periplasmic domain of the type II secretion system protein EpsM from *Vibrio cholerae*: the simplest version of the ferredoxin fold. J Mol Biol 338: 585-96.
- Campos M, Nilges M, Cisneros M, Francetic O (2010) Detailed structure and assembly model of the type II secretion pilus from sparse data. Proc Natl Acad Sci USA 107: 13081-6.
- Chernyatina AA, Low HH (2019) Core architecture of a bacterial type II secretion system. Nat Commun 10: 5437.
- Cianciotto NP, White RC (2017) Expanding Role of Type II Secretion in Bacterial Pathogenesis and Beyond. Infect Immun 85: e00014-17
- Cisneros DA, Bond PJ, Pugsley AP, Campos M, Francetic O (2012) Minor pseudopilin self-assembly primes type II secretion pseudopilus elongation. The EMBO J 31: 1041-1053.
- Datsenko KA, Wanner BL (2000) One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. Proc Natl Acad Sci U S A 97: 6640-6645.
- Dautin N, Karimova G, Ullmann A, Ladant D (2000) Sensitive Genetic Screen for Protease Activity Based on a Cyclic AMP Signaling Cascade in *Escherichia coli*. J Bacteriol 182: 7060-7066.
- Dazzoni R, Li Y, Lopez-Castilla A, Brier S, Mechaly A, Cordier F, Haouz A, Nilges M, Francetic O, Bardiaux B, Izadi-Pruneyre N (2022) Structure and dynamic association of an assembly

platform subcomplex of the bacterial type II secretion system. BioRxiv  
doi: <https://doi.org/10.1101/2022.07.16.500195>

Douzi B, Ball G, Cambillau C, Tegoni M, Voulhoux R (2011) Deciphering the Xcp *Pseudomonas aeruginosa* Type II Secretion Machinery through Multiple Interactions with Substrates. J Biol Chem 286: 40792-40801.

Durand E, Bernadac A, Ball G, Lazdunski A, Sturgis JN, Filloux A (2003) Type II protein secretion in *Pseudomonas aeruginosa*: the pseudopilus is a multifibrillar and adhesive structure. J Bacteriol 185: 2749-58.

Escobar CA, Douzi B, Ball G, Barbat B, Alphonse S, Quinton L, Voulhoux R, Forest KT (2021) Structural interactions define assembly adapter function of a type II secretion system pseudopilin. Structure 29: 1116-1127.e8.

Fulara A, Vandenberghe I, Read RJ, Devreese B, Savvides SN (2018) Structure and oligomerization of the periplasmic domain of GspL from the type II secretion system of *Pseudomonas aeruginosa*. Sci Rep 8: 16760

Georgiadou M, Castagnini M, Karimova G, Ladant D, Pelicic V (2012) Large-scale study of the interactions between proteins involved in type IV pilus biology in *Neisseria meningitidis*: characterization of a subcomplex involved in pilus assembly. Mol Microbiol 84: 857-873.

Gray MD, Bagdasarian M, Hol WGJ, Sandkvist M (2011) In vivo cross-linking of EpsG to EpsL suggests a role for EpsL as an ATPase-pseudopilin coupling protein in the Type II secretion system of *Vibrio cholerae*. Mol Microbiol 79: 786-798.

Gu S, Shevchik VE, Shaw R, Pickersgill RW, Garnett JA (2017) The role of intrinsic disorder and dynamics in the assembly and function of the type II secretion system. Biochim Biophys Acta (BBA) - Proteins and Proteomics 1865: 1255-1266.

Hay ID, Belousoff MJ, Dunstan RA, Bamert RS, Lithgow T (2018) Structure and membrane topography of the vibrio-type secretin complex from the type 2 secretion system of enteropathogenic *Escherichia coli*. J Bacteriol 200: e00521-17.

Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, Tunyasuvunakool K, Bates R, Žídek A, Potapenko A, Bridgland A, Meyer C, Kohl SAA, Ballard AJ, Cowie A, Romera-Paredes B, Nikolov S, Jain R, Adler J, Back T et al. (2021) Highly accurate protein structure prediction with AlphaFold. Nature 596: 583-589.

Karimova G, Pidoux J, Ullmann A, Ladant D (1998) A bacterial two-hybrid system based on a reconstituted signal transduction pathway. Proc Natl Acad Sci U S A 95: 5752-5756.

Karuppiiah V, Collins RF, Thistlethwaite A, Gao Y, Derrick JP (2013) Structure and assembly of an inner membrane platform for initiation of type IV pilus biogenesis. Proc Natl Acad Sci USA 110: E4638-E4647.

Karuppiah V, Derrick JP (2011) Structure of the PilM-PilN Inner Membrane Type IV Pilus Biogenesis Complex from *Thermus thermophilus*. J Biol Chem 286: 24434-24442.

Kohler R, Schafer K, Muller S, Vignon G, Diederichs K, Philippsen A, Ringler P, Pugsley AP, Engel A, Welte W (2004) Structure and assembly of the pseudopilin PulG. Mol Microbiol 54: 647-64.

Korotkov K, Sandkvist M (2019) Architecture, Function, and Substrates of the Type II Secretion System. EcoSal Plus 8: 10.1128/ecosalplus.ESP-0034-2018.

Korotkov KV, Hol WG (2008) Structure of the GspK-GspI-GspJ complex from the enterotoxigenic *Escherichia coli* type 2 secretion system. Nature structural & molecular biology 15: 462-8.

Lallemant M, Login FH, Guschinskaya N, Pineau C, Effantin G, Robert X, Shevchik VE (2013) Dynamic Interplay between the Periplasmic and Transmembrane Domains of GspL and GspM in the Type II Secretion System. PLoS One 8: e79562

Leighton TL, Dayalani N, L.M. S, Howell PL, LL. B (2015) Novel Role for PilNO in Type IV Pilus Retraction Revealed by Alignment Subcomplex Mutations. J Bacteriol 197: 2229-2238.

Leighton TL, Yong DH, Howell PL, Burrows LL (2016) Type IV Pilus Alignment Subcomplex Proteins PilN and PilO Form Homo- and Heterodimers *n Vivo*. J Biol Chem 291: 19923-19938.

López-Castilla A, Thomassin J-L, Bardiaux B, Zheng W, Nivaskumar M, Yu X, Nilges M, Egelman EH, Izadi-Pruneyre N, Francetic O (2017) Structure of the calcium-dependent type 2 secretion pseudopilus. Nat Microbiol 2: 1686-1695.

Ludwiczak J, Winski A, Szczepaniak K, Alva V, Dunin-Horkawicz S (2019) DeepCoil—a fast and accurate prediction of coiled-coil domains in protein sequences. Bioinformatics (Oxford, England) 35: 2790-2795.

Luna Rico A, Zheng W, Petiot N, Egelman EH, Francetic O (2019) Functional reconstitution of the type IVa pilus assembly system from enterohaemorrhagic *Escherichia coli*. Mol Microbiol 111: 732-749.

McCallum M, Tammam S, Little DJ, Robinson H, Koo J, Shah M, Calmettes C, Moraes TF, Burrows LL, Howell PL (2016) PilN Binding Modulates the Structure and Binding Partners of the *Pseudomonas aeruginosa* Type IVa Pilus Protein PilM\*. J Biol Chem 291: 11003-11015.

Michaelis S, Chapon C, d'Enfert C, Pugsley AP, Schwartz M (1985) Characterization and expression of the structural gene for pullulanase, a maltose-inducible secreted protein of *Klebsiella pneumoniae*. J Bacteriol 164: 633-638.

Miller JH (1972) Experiments in Molecular Genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY. USA,

Naskar S, Hohl M, Tassinari M, Low HH (2021) The structure and mechanism of the bacterial type II secretion system. *Mol Microbiol* 115: 412-424.

Nivaskumar M, Bouvier G, Campos M, Nadeau N, Yu X, Egelman EH, Nilges M, Francetic O (2014) Distinct Docking and Stabilization Steps of the Pseudopilus Conformational Transition Path Suggest Rotational Assembly of Type IV Pilus-like Fibers. *Structure* 22: 685-696.

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

Patrick M, Korotkov KV, Hol WGJ, Sandkvist M (2011) Oligomerization of EpsE Coordinates Residues from Multiple Subunits to Facilitate ATPase Activity. *J Biol Chem* 286: 10378-10386.

Possot O, Vignon G, Bomchil N, Ebel F, Pugsley AP (2000) Multiple interactions between pullulanase secretion components involved in stabilization and cytoplasmic membrane association of PulE. *J Bacteriol* 182: 2142-2152.

Py B, Loiseau F, Barras F (2001) An inner membrane platform in the type II secretion machinery of Gram-negative bacteria. *EMBO Rep* 2: 244-248.

Robien MA, Krumm BE, Sandkvist M, Hol WG (2003) Crystal structure of the extracellular protein secretion NTPase EpsE of *Vibrio cholerae*. *J Mol Biol* 333: 657-74.

Sampaleanu LM, Bonanno JB, Ayers M, Koo J, Tammam S, Burley SK, Almo SC, Burrows LL, Howell PL (2009) Periplasmic Domains of *Pseudomonas aeruginosa* PilN and PilO Form a Stable Heterodimeric Complex. *J Mol Biol* 394: 143-159.

Sandkvist M, Hough LP, Bagdasarian MM, Bagdasarian M (1999) Direct interaction of the EpsL and EpsM proteins of the general secretion apparatus in *Vibrio cholerae*. *J Bacteriol* 181: 3129-35.

Santos-Moreno J, East A, Guilvout I, Nadeau N, Bond PJ, Tran Van Nhieu G, Francetic O (2017) Polar N-terminal Residues Conserved in Type 2 Secretion Pseudopilins Determine Subunit Targeting and Membrane Extraction Steps during Fibre Assembly. *J Mol Biol* 429: 1746-1765.

Sauvonnet N, Gounon P, Pugsley AP (2000) PpdD type IV pilin of *Escherichia coli* K-12 can be assembled into pili in *Pseudomonas aeruginosa*. *J Bacteriol* 182: 848-854.

Sauvonnet N, Vignon G, Pugsley AP, Gounon P (2000) Pilus formation and protein secretion by the same machinery in *Escherichia coli*. *EMBO J* 19: 2221-2228.

Schägger H, von Jagow G (1987) Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. *Anal Biochem* 166: 368-379.

Varadi M, Anyango S, Deshpande M, Nair S, Natassia C, Yordanova G, Yuan D, Stroe O, Wood G, Laydon A, Židek A, Green T, Tunyasuvunakool K, Petersen S, Jumper J, Clancy E, Green R, Vora A, Lutfi M, Figurnov M et al. (2022) AlphaFold Protein Structure Database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. *Nucl Acids Res* 50: D439-D444.

Yan Z, Yin M, Xu D, Zhu Y, Li X (2017) Structural insights into the secretin translocation channel in the type II secretion system. . *Nat Struct Mol Biol* 24: 177–183.

## Supplementary Information

### Contains:

Supplementary Table S1

Supplementary Fig. 1

Supplementary Fig. 2

**Table S1 Oligonucleotides used in this study**

| Primer name    | DNA sequence (5'-3')                        |
|----------------|---------------------------------------------|
| PuIL dCC For   | CCTCCACGATCATCTCATGCAGCAGCATCTGCGGCAG       |
| PuIL dCC Rev   | GCTTCTGATTGCGAAGCTGCCATAACGTGTAATAG         |
| PuIM dCC For   | CTATTACACGTTATGGCAGCTTCGCAATCAGAAGC         |
| PuIM dCC Rev   | GCTTCTGATTGCGAAGCTGCCATAACGTGTAATAG         |
| PuIM del N5    | CTGACGCTGGAGGGGAACGATGCGGTGTCTGCTGCTGGGTATG |
| PuIM del N3    | CATACCCAGCAGCAGACACCGCATCGTTCCCCTCCAGCGTCAG |
| PuIM delN For  | CGGTGTCTGCTGCTGGGTATG                       |
| PuIM delN Rev2 | ATGCATCGGATCCTCTAGAGTC                      |

### Supplementary figures.

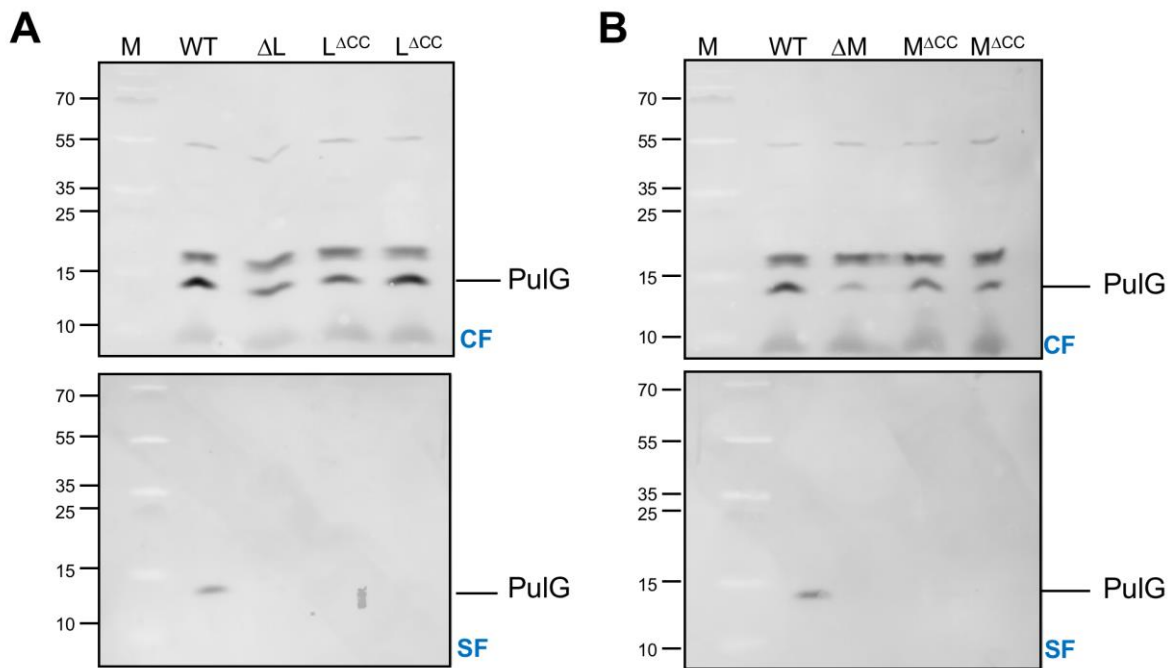


**Figure S1.** Piliation defect of PulL and PulM lacking periplasmic helices. **A.** PulG pilus assembly on the surface of PAP7460 bacteria containing plasmid pCHAP8251 complemented with pCHAP8252 (PulL<sup>WT</sup>), pMS1348 (PulL <sup>$\Delta$ CC</sup>) or vector pSU19 ( $\Delta$ L). Migration of Mw markers (M) (in kDa) is indicated on the left. **B.** PulG pilus assembly on the surface of PAP7460 bacteria containing plasmid pCHAP8496 complemented with pCHAP1353 (PulM<sup>WT</sup>), pMS1350 (PulM <sup>$\Delta$ CC</sup>) or vector pSU18 ( $\Delta$ M). Cell- (CF) and sheared fractions (SF) were prepared as described in Materials and Methods from bacteria cultured for 48 h on LB agar at 30°C. For each sample, 0.05 OD<sub>600nm</sub> equivalents were analyzed on 10% Tris-Tricine SDS PAGE and western blot using anti-PulG antibodies.

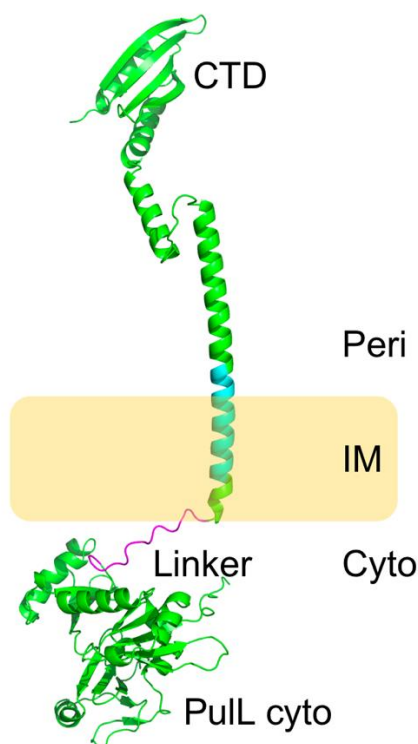


**Figure S2.** The AlphaFold 2 model of PulL in green represented in the bacterial inner membrane (IM), shown in yellow. The PulL cytoplasmic (PulL<sub>cyto</sub>) and C-terminal domains (CTD) are labeled. The disordered linker is shown in magenta. The periplasmic (peri) and cytoplasmic (cyto) compartments are indicated.

### 3.2.3 Additional results about PulL and PulM

As the BACTH studies suggest that PulA interacts with PulM, and that PulL and PulM form heterodimers. We were wondering whether the presence of PulA would affect the interaction of PulL and PulM. We designed a “triple-hybrid assay”, by co-transforming PulL and PulM in the DHT1 strain harboring plasmid encoding PulA or the empty vector (pSU18). As result shown in Fig. 70, the presence of PulA or the vector in DHT1 strain not affect the  $\beta$ -galactosidase levels in the T18-T25 negative or T18Zip-T25Zip positive control. When we compared the activity of PulL-PulM pairs, the presence of PulA significantly affected the PulL and PulM interaction in one combination of constructs, but not the other. The presence of PulA reduced the interaction of T18-PulM with T25-PulL. This could mean that PulA could cause

the dissociation of the complex by binding to one of the partners. So far, BACTH and biochemical studies indicate that this partner could be PulM.

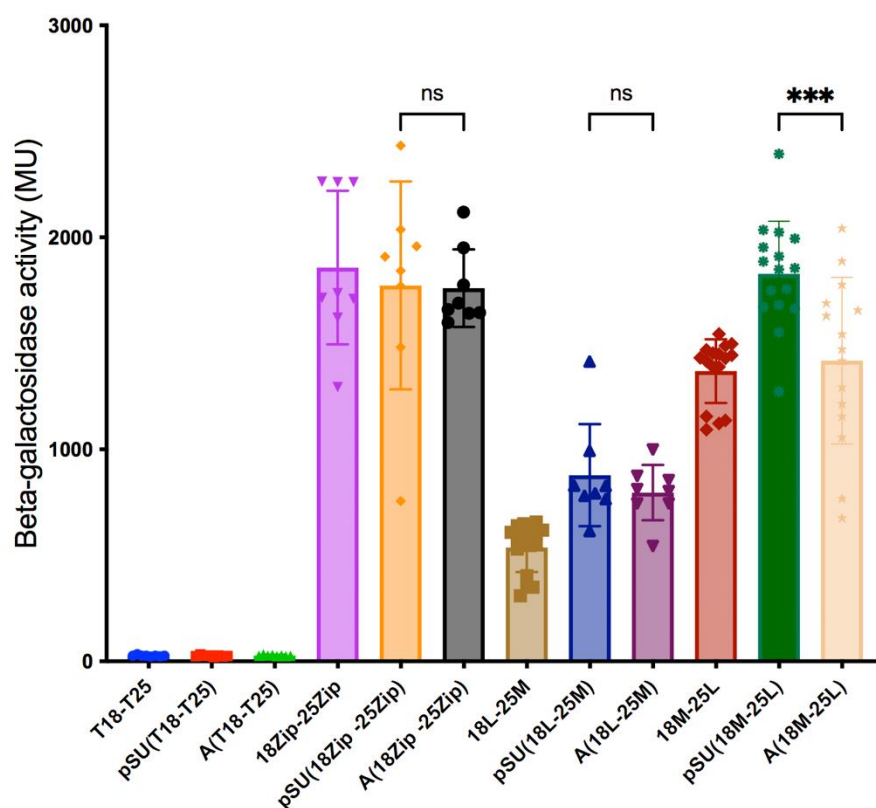


**Figure 70. BACTH analysis of PulA (PulL+PulM).**

The  $\beta$ -galactosidase activities measured for 8 or more independent bacterial cultures. Individual values are shown as dots. Bar graphs indicate mean values and error bars standard deviation. Statistical analysis was performed using One-way ANOVA and multiple comparisons test, and plotted using GraphPad Prism 9 software. ns, no significant difference; \*\*\*,  $p < 0.001$ .

We further verified by Western blot that PulA was produced in the BACTH samples under the assay conditions. Although the effect of PulA was observed only in one of the PulL-PulM pairs, this result suggests mean that PulA binding might dissociate the PulL-PulM complex at their site of interaction in the T2SS. This possibility should be investigated further *in vitro* and *in vivo*.

## **4. Discussion and Perspectives**

## 4.1 Crosslinking approaches to identify PulA interacting partners

Type II secretion systems are complex bacterial nanomachines embedded in the envelope of Gram-negative bacteria. The Pul system from *K. oxytoca* studied in this thesis has been discovered in the 80s. Since then T2SS was found in many species across multiple bacterial classes (Denise et al., 2019). The ability of this system to secrete folded proteins may have played a role in the bacterial ancestors of mitochondria, and might have contributed to the evolution of eukaryotic cells (Horváthová et al., 2021). Folded proteins are stabilized by disulfide bonds, calcium or other cofactors and are secreted by T2SS ready to be active and to resist to the environmental conditions and proteolysis. The mechanism of substrate recognition remains the main open question in the T2SS field, together with the question of secretion mechanism.

In this thesis we attempted to address these questions by focusing on PulA interactions with the T2SS. The main goal of my work was to set up the *pBPA* UV cross-linking technique in the lab, to address this question. I also extended the use of the bacterial two-hybrid (BACTH) system to identify PulA interactions with T2SS components. Previous interaction studies in the lab showed that PulG interacts with PulF, PulG, PulH, PulI, PulJ, PulK, PulL and PulM (Nivaskumar et al., 2016).

Previously, unpublished results of co-purification studies in the host laboratory have identified several candidate T2SS components that might interact with PulA. These proteins either reside in the inner membrane (PulM, PulC) or form the pseudopilus (PulG). Based on these results, we employed the *in vivo* site-directed photo-crosslinking assay, by implementing the unnatural amino acid nonsense suppressor approach on the T2SS model protein PulG and PulA. We hoped to capture some transient protein-protein interactions by crosslinking residues in close contact with an UV-activatable residue *pBPA* incorporated in a protein under study. By using PulG-*pBPA* derivatives, we expected to identify known PulG partners such as PulM and minor pseudopilins, and potentially PulA itself.

However, in all the attempts described so far, we failed to identify the interacting partners of PulA using the *pBPA*-UV crosslinking approach. Although we could generate several crosslinked adducts of some PulA-*pBPA* variants, we were unable to complete their characterization by Western blot or by mass spectrometry. Although our results were negative

so far regarding PulA, we identified possible ways to improve the application of this method on our system and this approach that are worth testing.

We were facing several limitations of the *pBPA* approach. The site-targeted nature of this approach results in its limitations: despite multiple *pBPA* positions that we tested, there is a possibility that we missed the optimal sites to capture the interactions by site-specific cross-linking assay. In addition, the highly dynamic secretion process might limit the fraction of PulA that interacts with a given T2SS component. This might result in very low levels of crosslinked adducts, which are below the detection limit by Western blot with our available antibodies. Although we attempted to enrich and concentrate the PulA-Strep labeled adducts, the amounts of protein were too low for analysis by mass spectrometry. In the case that PulA was crosslinked to another cellular protein, the mass spectrometry could have allowed us to identify this complex. However, the analysis of peptides generated by mass spectrometry is further complicated by the presence of *pBPA* in the proteins. In our initial trial, we failed to identify peptides containing *pBPA* residues by mass spectrometry. Unlike *pBPA*, more recently other unnatural amino acids have been developed that are better suited for analysis by MS, however they are not commercially available.

Another issue we faced was low efficiency of crosslinking. Initially we grew our bacteria in minimal medium, but the precultures in LB suggested by prof. Shevchik seemed to improve the subsequent *pBPA* incorporation yield and crosslinking. It would be worth reanalyzing the PulG-*pBPA* variants with this modified protocol, since it was surprising that with this small protein which interacts with many T2SS components we only identified homodimers. The reason for this could be also the relative abundance of components. PulG and PulA are the most abundant substrates of the system and all the other proteins are less abundant. Therefore, the proportion of PulA in contact with the partner protein would be very small and perhaps not detectable by Western blot. Still, the sensitivity of mass spectrometry should allow us to detect such proteins after copurification and band extraction. This approach should be pursued further in the future. Also, it may be a better strategy to modify less abundant components with *pBPA* and increase the chance to detect the crosslinked adducts with more abundant substrates.

We also noticed that the secretion of PulA begins as soon as the T2SS complex is formed in the *E. coli* envelope. Even at the early phase of the bacterial growth, most of the PulA is secreted in the medium before the accumulation of sufficient bacterial mass and buildup of T2SS components. This results in lower abundance of Pul factors and consequently lower abundance of PulA complexes.

To overcome these problems, we performed several complementary approaches including formaldehyde cross-linking and co-purification with PulA-Strep and Bacterial two-hybrid assays.

Unlike the targeted pBPA crosslinking, the SPINE approach has the potential to capture all possible interactions of PulA with the T2SS complex. I performed preliminary experiments aimed at setting up SPINE for the Pul T2SS. The main issue we observed here was the absence of PulA in the cell fraction of bacteria containing the functional Pul system. This phenomenon might be caused by two main reasons: firstly, even in the early growth stages, most of the PulA was secreted in the culture supernatant and lost. By using a  $\Delta pulE$  mutant we increased the yield of intracellular PulA and hoped that we would still be able to capture early secretion intermediates. We assumed that PulA should bind at least one Pul factor prior to the steps catalyzed by the PulE ATPase. The fact that we failed to detect any complexes with the Pul factors might be due to low levels of such complexes. Secondly, the presence of PulA in the membrane fraction (Fig. 62, 63) suggests that this PulA might be bound to some membrane proteins. Since the majority of Pul components are in the inner or the outer membrane, the crosslinked adducts of interest should be found in this fraction. To capture such complexes, the cell extract should be treated with detergent to solubilize the membrane-bound T2SS components and purification should be done in the presence of detergent. The caveat in this case is that we probably would not be able to solubilize the secretin PulD, which is highly stable and detergent resistant. Also, the loss of the crosslinked protein might be due to the limitation of the approach itself. In a study of the utility of formaldehyde cross-linking and mass spectrometry, the authors identified a great loss (90%) of the crosslinked complex as well as the global proteins of the bacteria when treating the culture by 1% formaldehyde for 30 min, which is exactly my incubation condition. The significant loss of all the proteins was probably due to lysis of the cell by the formaldehyde treatment (Sutherland et al., 2008). If this is the case, that means we may have lost most of the PulA as well as its complexes with Pul proteins in the supernatant since we centrifuge the whole culture after formaldehyde crosslinking, and discard all the supernatant. To improve the yield, we would need to optimize the experimental conditions and use 0.3% - 0.5% formaldehyde and shorter incubation times or lower temperature (Sutherland et al., 2008). In addition, a “glycine quenching” step could be added after formaldehyde incubation to prevent over-reaction (Vasilescu et al., 2004).

In view of all the problems we faced when attempting to capture transient interactions in the dynamic secretion process *in vivo*, the first and foremost issue is the design of a good

experiment to monitor and generate a quantification curve of the secretion process, to localize the best window with a good balance of the amount of PulA and other Pul proteins present in the cell, to perform the *in vivo* crosslinking assays. If there does not exist such moment, the next is to find a way to slow down the process. To achieve this, we could use mutants of the PulE ATPase that are partially functional and show reduced ATPase activity. It would also be interesting to use chemical treatments that slow down or block secretion for example by blocking the proton motive force, which is essential for PulA secretion, by CCCP or other ionophores (Possot et al., 1997).

## 4.2 BACTH studies of PulA interactions with the T2SS

As a complementary approach, we used the BACTH analysis to T2SS components that interact with PulA at the level of the inner membrane. Although substrate interactions with the T2SS components are often considered as transient, this is probably the case in the active system which is powered by the ATPase activity. However, it is possible that stable intermediates are formed in the absence of ATPase. If so, we should be able to identify such early intermediates by the BACTH assay. As described in this work, the BACTH assays identified several strong candidates for PulA interaction partners including the major pseudopilin PulG, and the AP components PulM and PulF. In addition, other pseudopilins also showed a tendency to bind PulA, depending on the copy number of plasmids encoding the Cya hybrids and protein stability. To test whether these interactions reflect the situation in the functional T2SS, the results need to be corroborated by complementary approaches, e.g. biochemical – copurification, binding studies, or by microscopy – including electron microscopy and fluorescence microscopy. Our collaborator N. Izadi-Pruneyre and her team were able to show binding of PulA to the pseudopili using biophysical techniques – SPR, the biolayer interferometry and co-sedimentation. In collaboration within an ITN training network, they will use chemical crosslinking and mass spectrometry to characterize contacts between PulA and the pseudopili.

Evidence for an interaction of PulA with PulM has also been obtained by Aracelys Lopez Castilla in the team of N. Izadi using surface plasmon resonance. This result agrees with the data in *Pseudomonas* T2SS, where interactions of the substrate with XcpL and XcpM have been observed (Michel-Souzy et al., 2018). As discussed above, PulM might play a role in coupling

PulA binding with pseudopilus assembly. PulM also interacts with PulL and with the pseudopilins PulG and PulH (Nivaskumar et al., 2016). We can propose that the PulM-PulG complex might be targeted to PulL to form a tripartite complex. It would be interesting to test PulA binding and compare its affinities to PulG, PulM, PulG-PulM and PulL-PulM complexes. This might provide information on the possible mechanism of substrate recruitment and transfer at the site of pseudopilus assembly.

According to the BACTH results, we could see an unexpected and strong interaction of PulA with PulF. The Flag peptide insertions at positions L198 and A201 of the periplasmic loop of PulF abolished this interaction and affected specifically PulA, secretion but not pseudopilus assembly. These results support a direct role of PulF in secretion but need further investigation. In the future, it would be interesting to study the PulA-PulF contacts and identify the interface more precisely, possibly by using the UV crosslinking upon incorporating *pBPA* in the periplasmic regions of PulF. Site-directed mutagenesis of this region of PulF could also bring information on its function.

### 4.3 The role of PulL-PulM assembly platform complex

An important part of our work has been focused on the AP components PulL and PulM. PulM is required for PulL stability, and PulL is essential for pseudopilus assembly and PulA secretion (Possot et al., 2000), which is presumably due to its role in anchoring the ATPase PulE to the base of the complex. It has been proposed that GspL and GspM signal the presence of the substrate across the IM to the ATPase GspE (Lallemant et al., 2013; Michel-Souzy et al., 2018). However, the absence of the substrate in the Pul T2SS does not prevent pseudopilus assembly, which argues against the signaling model. The Pul T2SS lacking PulM is partially functional, as a very small number of long pseudopili can be assembled and a small amount of PulA is secreted (Santos-Moreno et al., 2017). Extensive biophysical and structural analysis of the PulL<sub>CTD</sub>-PulM<sub>CTD</sub> complex performed by the group of N. Izadi-Pruneyre provided a novel insight into these domains and their interface. Despite numerous structural studies, the structure of a L-M heterodimer had not been determined before and could only be modeled. Here, Dazzoni *et al.* combined NMR to characterize PulL binding to PulM and obtained crystals with parallel heterodimers, which are highly likely to reflect the native state of this complex. Our

BACTH data support this parallel heterodimer state, and provide no evidence for homodimerization. The Cysteine crosslinking of PulL and PulM transmembrane segments allowed us to identify specific residues in the TMS of PulL and PulM that made specific crosslinks under oxidizing conditions. Define these contacts were helpful in building the final model of the L-M complex comprising their membrane and cytoplasmic domains. The flexible linkers, which connect the ferredoxin like domains of L and M to their periplasmic helices remain difficult to predict and might provide some freedom to the movement of the CTDs. Given the monomer-dimer equilibrium in solution, it is possible that *in vivo* CTDs also bind to each other and dissociate in different stages of pseudopilus assembly and secretion process. It would be important to understand whether and how their substrates – PulA and PulG- or other factors influence the dynamics.

Purification of the *Klebsiella* T2SS subcomplex containing the secretin and parts of the assembly platform supports the 1:1 stoichiometry of PulL and PulM (Chernyatina & Low, 2019). However, we noticed that PulM was much more abundant in bacteria producing the T2SS system from the *pul* locus. Thanks to the purified PulM<sub>CTD</sub> and PulL<sub>CTD</sub> which are recognized by the available PulL and PulM antibodies, we were able to determine the molar ratio of PulM and PulL in the bacteria and found that PulM is 20-fold more abundant. This agrees with the model of a dynamic role of PulM, which is predicted to shuttle between the membrane pool of PulG and the assembly platform-associated PulL. To test this model, it would be interesting to perform live fluorescence microscopy localization studies and FRAP experiments with PulL and PulM fused to GFP or other fluorescent proteins.

The same regions of PulL<sub>CTD</sub> and PulM<sub>CTD</sub> are involved in both homo- and heterodimer formation in solution. This interface is hydrophobic and probably forces the dimerization to provide stability to the complex in solution. In the native mass spectrometry studies, PulL and PulM formed only dimers, but not higher order multimers, which fits with the presence of a single binding determinant in each protein. However, *in vivo* in the context of full-length proteins, these C-terminal domains are restricted by the presence of other domains/regions and by their membrane embedded conformation. The study described in the second manuscript, Li et al., suggests that when the CTDs of L and M are connected to the transmembrane segments directly *via* their native flexible linkers, the interaction between them is rather weak in the BACTH system. This suggests that the coiled coil formed by the periplasmic helices of PulL and PulM plays a dominant role in driving the heterodimer formation. As discussed in that manuscript, it is also possible that this coiled coil ensures a distance between the CTDs and the

membrane and prevents an extensive interaction of these domains with the membrane surface. In addition, the coiled coil might provide rigidity to the complex and might also allow it to interact with other proteins or peptidoglycan, although there is so far no evidence for these possibilities. In the *D. dadantii* Out T2SS, L and M proteins interact with OutC, but when we used the BACTH approach to test these interactions in the Pul system, we did not detect any interactions signals.

We used site-directed mutagenesis to target the Pul<sub>L</sub><sub>CTD</sub>-Pul<sub>M</sub><sub>CTD</sub> interface. We found that was difficult to disrupt this interface without affecting the stability of the CTDs. Initially we used charge inversions, but they strongly affected the domain stability, and we had to introduce double alanine substitutions to minimize this effect. These mutations produced a mild but reproducible and significant effect on PulA secretion. It is unclear whether Pul<sub>L</sub><sub>CTD</sub>-Pul<sub>M</sub><sub>CTD</sub> complex binds to PulA, as is the case in *Pseudomonas* system (Michel-Souzy et al., 2018). While in this study the addition of substrate CbpD stimulated the interaction of L and M in the BACTH assay, we observed an opposite effect in a similar experiment (Fig. 70). In that experiment, PulA showed a tendency to reduce PulL-PulM interaction, which might indicate that PulA binding to one of the interacting partners competes rather than stimulates the interaction of the T18/T25-PulL/PulM hybrid proteins.

We found that deleting the cytoplasmic peptide of PulM affected PulL stability and generated a PulL fragment that is an equivalent of its orthologue in T4P, PilN. This fragment, which remains to be characterized by N-terminal sequencing or mass spectrometry, probably comprises the membrane and periplasmic domains of PulL. Interestingly, this fragment could also be detected in the presence of the complete T2SS (in experiments that served to quantify the PulL/PulM ratio, Fig. S5, Dazzoni *et al.* 2022), which suggests that the protease-sensitive site in PulL is exposed during secretion. This fragment was generated upon deleting the cytoplasmic peptide of PulM much more frequently than in the absence of PulM or in the presence of native PulM. This is puzzling and might suggest that the PulL-PulM interactions in the periplasm promote conformational changes of PulL that expose the protease-sensitive region of PulL. *In vitro* studies could be used to study binding of PulM peptide to the PulL cytoplasmic domain. In T4P the orthologue of Pul<sub>L</sub><sub>cyto</sub> domain PilM was co-crystallized with the N-terminal peptide of PilN, which might be the orthologue of PulM in the T2SS.

Overall, the structural, biochemical and biophysical studies of PulL and PulM opened new avenues of research. While it remains difficult to study the function of dynamic membrane-

embedded proteins, combining complementary *in vitro* and *in vivo* approaches might provide insights into the mechanisms of PulG assembly and PulA secretion.

## **5. Materials and methods**

## 5.1 Bacterial strains and growth conditions

The plasmids used in the work are listed in Table 1.

*Escherichia coli* K-12 strain DH5 $\alpha$  F' *lacI*<sup>Q</sup> was used as a host for cloning purposes. Strains BL21, BW25113 F' *lacI*<sup>Q</sup>, MG1655 F' *lacI*<sup>Q</sup>, PAP105, PAP7460, TG1 and TOP10 are used for selection of host strains for *pBpa* incorporation. BW25113 F' *lacI*<sup>Q</sup>, MG1655 F' *lacI*<sup>Q</sup> and PAP7460 was used as a *pul* gene expression host. Bacterial two-hybrid assays were performed in strain DHT1 (Dautin et al., 2000). Bacteria were grown at 30°C or 37°C in LB medium (Miller, 1972) or in minimal glycerol medium (M63 medium supplemented with 0.5% glycerol, 10 mM MgSO<sub>4</sub> and 1 mM CaCl<sub>2</sub>, 0,5% B1 (thiamine) and antibiotics, added as required at following concentration: ampicillin (Ap) 100  $\mu$ g. mL<sup>-1</sup>, chloramphenicol (Cm) 25  $\mu$ g. mL<sup>-1</sup>, kanamycin (Km) at 25 $\mu$ g. mL<sup>-1</sup> and spectinomycin (Sp) 50  $\mu$ g. mL<sup>-1</sup>. Expression of the *pul* genes was induced with 0.4 % of D-maltose. The genes under *placZ* control were induced with Isopropyl  $\beta$ - D-1-thiogalactopyranoside (IPTG).

**Table 1. Bacterial strains used in this study.**

| Strain                                   | Genotype and/or relevant markers                                                                                                                                                                                                                                                      | Source                    |
|------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------|
| BL21( $\lambda$ DE3)                     | F <sup>-</sup> <i>ompT hsdS<sub>B</sub> (r<sub>B</sub><sup>-</sup> m<sub>B</sub><sup>-</sup>) gal dcm</i> ( $\lambda$ DE3)                                                                                                                                                            | Bill Studier              |
| BW25113 F' <i>lacI</i> <sup>Q</sup>      | $\Delta$ ( <i>araD-araB</i> )567, $\Delta$ <i>lacZ</i> 4787(::rrnB-3), $\lambda^-$ , <i>rph-1</i> , $\Delta$ ( <i>rhaD-rhaB</i> )568, <i>hsdR</i> 514                                                                                                                                 | (Datsenko & Wanner, 2000) |
| DH5 $\alpha$ F' <i>lacI</i> <sup>Q</sup> | F <sup>-</sup> $\Phi$ 80 <i>lacZ</i> $\Delta$ M15 $\Delta$ ( <i>lacZYA-argF</i> ) U169 <i>recA1 endA1 hsdR</i> 17 (rK <sup>-</sup> , mK <sup>+</sup> ) <i>phoA supE</i> 44 $\lambda^-$ <i>thi-1 gyrA</i> 96 <i>relA</i> 1                                                             | Laboratory collection     |
| DHT1                                     | $\Delta$ <i>cya</i>                                                                                                                                                                                                                                                                   | (Dautin et al., 2000)     |
| MG1655 F' <i>lacI</i> <sup>Q</sup>       | <i>rph1 ilvG rfb-50</i> $\lambda^-$ F                                                                                                                                                                                                                                                 | (Guyer et al., 1981)      |
| PAP105                                   | $\Delta$ ( <i>lac-pro</i> ) F'( <i>lacI</i> <sup>Q</sup> $\Delta$ <i>lacZ</i> M15 <i>pro</i> <sup>+</sup> <i>Tn10</i> ).                                                                                                                                                              | Laboratory collection     |
| PAP7460                                  | $\Delta$ ( <i>lac-argF</i> )U169 <i>araD</i> 139 <i>relA1 rpsL</i> 150 $\Delta$ <i>malE</i> 444 <i>malG</i> 501 [F' ( <i>lacI</i> <sup>Q</sup> $\Delta$ <i>lacZ</i> M15 <i>pro</i> <sup>+</sup> <i>Tn10</i> )] (Tc <sup>R</sup> )                                                     | (Possot et al., 2000)     |
| TG1                                      | F <sup>-</sup> <i>traD</i> 36 <i>lacI</i> <sup>Q</sup> $\Delta$ <i>lacZ</i> M15 <i>proA+B supE D</i> ( <i>HsdM-mcrB</i> 5) <i>r<sub>k</sub>-m<sub>k</sub><sup>-</sup> thi</i> $\Delta$ ( <i>lac-proAB</i> )                                                                           | Laboratory collection     |
| TOP10                                    | <i>mcrA</i> , $\Delta$ ( <i>mrr-hsdRMS-mcrBC</i> ), $\Phi$ 80 <i>lacZ</i> ( $\Delta$ M15, $\Delta$ <i>lacX</i> 74, <i>deoR</i> , <i>recA1</i> , <i>araD</i> 139, $\Delta$ ( <i>ara-leu</i> ) 7697, <i>galU</i> , <i>galK</i> , <i>rpsL</i> ( <i>SmR</i> ), <i>endA1</i> , <i>nupG</i> | Invitrogen                |

## 5.2 Plasmid construction and site-directed mutagenesis

All the plasmids used in the work are listed in Table 2. All plasmids were verified by DNA sequencing (Eurofins).

Site-directed mutagenesis was performed using a modification of the QuikChange™ method (Agilent) using Q5 DNA polymerase (New England Biolabs) or Platinum SuperFi DNA Polymerase (Invitrogen). All the mutagenic oligonucleotides (Eurofins) used in this study are listed in Table 3.

Point mutation primers were designed either by fully overlapping antiparallel sequence, or following the “overlapping mutagenic primers with 3’-overhangs” (Fig. 41). The mutagenic primers comprised two major parts: ① a 5’-complementary region of 15–20 nt and ② at least 8 non-overlapping bases at the 3’-terminus. The targeted mutation was in the middle of the complementary region in both primers. At least one G or C presented at the 5’ and 3’ termini of the primer will increase the success rate (Thermo Fisher Scientific protocol).

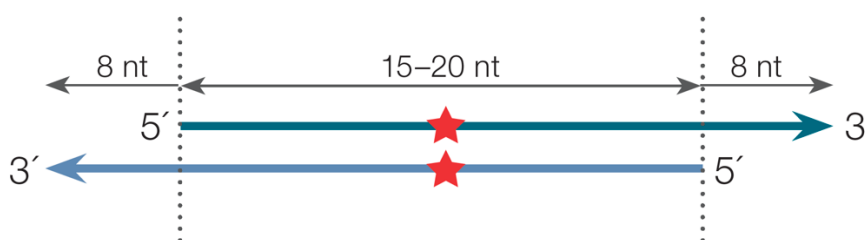


**Figure 71. Design of partially overlapping primers**  
The principle of primer designing (Thermo Fisher Scientific protocol).

All the PulG<sup>Am</sup> variants were generated from template pMS5008 (PulG<sup>Am135Oc</sup> in pCHAP8658, PCR mutation to replace the Amber stop codon by Ochre stop codon using primers PulG Stop (Am135Oc) For/Rev) by site-directed mutagenesis.

To construct plasmids pMS5032 (*pulA*-His<sub>6</sub>), I digested pCHAP4486 by EcoRI and HindIII, then separated the digested bands on 1% agarose gel after adding the CyberGreen dye. After visualizing the bands under the blue light trans-illuminator (Dark Reader), I cut and purified the A band, and extracted the DNA using the Qiaquick DNA extraction kit (Qiagen). The vector pSU18 was digested by EcoRI and HindIII, and the A band was extracted as above. I ligated these two EcoRI-HindIII fragments by T4 ligase (NEB) in T4 ligase buffer (NEB), incubate for 2 hours at room temperature. The reaction mixture was transformed into DH5α

F' *lacI*<sup>Q</sup> ultracompetent cells. The resulting plasmid pMS5032 was purified and tested by DNA sequencing (Eurofins).

All the PulA<sup>Am</sup>-His<sub>6</sub> variants were constructed by site-directed mutagenesis on pMS5032. Site-directed mutagenesis was performed using a modification of the QuikChange™ method (Agilent) using Q5 DNA polymerase or Platinum SuperFi DNA Polymerase (Invitrogen). The PCR products were treated with DpnI and transformed into ultracompetent cells of strain F' *lacI*<sup>Q</sup>. All plasmids were verified by DNA sequencing (Eurofins).

To construct pMS5036 (*pulAΔIns*), I used PCR to delete the Ins domain by primers PulAΔIns For/Rev on pCHAP8255. To construct pMS5072 (*pulB-S + pulA<sub>NA</sub>-His<sub>6</sub>*), I digested pCHAP8861 by restriction enzyme XbaI and NcoI, and purified the A band as vector. To prepare the insert, I used primers PulA NcoI For(Rev) and pulA-His XbaI Rev (Table 4) for PCR on pMS5032 to generate a NcoI-*pulA*-strep-XbaI fragment. I digested the PCR products by XbaI and NcoI, purified the A fragment and ligated the insert and vector DNA with T4 ligase (NEB).

To construct pMS5073 (*pulB-S + pulA<sub>NA</sub>-Strep*), I digested pMS5072 by restriction enzymes AscI and XbaI, and purified the A fragment as vector. To generate the insert, I used primers PulA NcoI For(Rev) and pulA-Strep XbaI For to PCR-amplify NcoI-*pulA*-strep-XbaI fragment on pCHAP4395, I digested the PCR products by AscI and XbaI, purified the A fragment and ligated the insert and vector fragments by T4 DNA ligase.

To construct pMS5078 (*pulB-S + pulA<sub>NA</sub>ΔIns-Strep*), I digested pMS5073 by enzymes AscI and NcoI, purified the A fragment as vector. I digested pMS5036 by AscI and NcoI, purified the A fragment and ligated with the vector fragment by T4 DNA ligase.

To construct pMS5079 (*pulΔE + pulA<sub>NA</sub>-Strep*), I digested pCHAP8200 by restriction enzymes AscI and EcoRI, and purified the A fragment as vector and ligated it with the purified B fragment of pMS5073 digested with AscI and EcoRI.

Plasmid pMS5037 (pUT18c-*pulK*-TM-*pulA*) was constructed by PCR-amplification of template pCHAP8123 using primers PulK-*pulA* IV For and pUT18c IV Rev and PCR-amplification on template pCHAP8247, using primers pUT18c IV For and PulK-*pulA* IV Rev. These PCR products were digested with *Dpn*I, column-purified, then co-transformed in to DH5α F' *lacI*<sup>Q</sup> ultracompetent cells for *in vivo* cloning.

To construct pMS5038 (pKT25-*pulK*-TM-*pulA*), insert fragment was PCR-amplified on template pHCAP8124, using primers “PulK-*pulA* IV For” and “pKT25 IV Rev”; vector fragment was PCR-amplified on template pHCAP8250, using primers “pKT25 IV For” and “PulK-*pulA* IV Rev”. The PCR products were digested with *DpnI*, purified, then co-transform in to DH5 $\alpha$  F’*lacI*<sup>Q</sup> ultracompetent cell for *in vivo* cloning.

To construct pMS5064 (pUT18c-*pulK*-TM) and pMS5065 (pKT25-*pulK*-TM) I used PCR deletion on pMS5037 or pMS5038, respectively by primers T18/25 *EcoRI* For and PulK-TM Och-*EcoRI* Rev. The PCR products were digested with *DpnI*, purified, then digested with *EcoRI*, ligated with T4 DNA ligase and transformed into DH5 $\alpha$  F’*lacI*<sup>Q</sup> ultracompetent cells.

**Table 2. Plasmids used in this study.**

| Plasmid    | Description                                                                     | Origin, Resistance                 | Source/Reference             |
|------------|---------------------------------------------------------------------------------|------------------------------------|------------------------------|
| pSU18      | <i>placZ</i> ,                                                                  | p15A <i>ori</i> , Cm <sup>R</sup>  | (Bartolomé et al., 1991)     |
| pSU19      | <i>placZ</i> , <i>lacZ'</i> ,                                                   | p15A <i>ori</i> , Cm <sup>R</sup>  | (Bartolomé et al., 1991)     |
| pCHAP231   | <i>pulS</i> , <i>pulAB-pulCDEFGHIJKLMNOP</i> ,                                  | ColE1 <i>ori</i> , Ap <sup>R</sup> | (D’Enfert & Pugsley, 1987)   |
| pCHAP8185  | <i>pulS</i> , <i>pulA</i> <sub>NA</sub> , <i>B</i> , <i>pulCDEFGHIJKLMNOP</i> , | ColE1 <i>ori</i> , Ap <sup>R</sup> | (Cisneros et al., 2012)      |
| pCHAP8184  | <i>ApulG</i> in <i>pCHAP8185</i> ,                                              | ColE1 <i>ori</i> , Ap <sup>R</sup> | (Campos et al., 2010)        |
| pEVOL-pBpF | M.j. p-benzoylphenylalanine RS (2 copies +tRNA)                                 | p15A <i>ori</i> , Cm <sup>R</sup>  | Addgene                      |
| pCDFDuet-1 | 2 MCSs, T7lac promoter, <i>lacI</i>                                             | CDF <i>ori</i> , Sm <sup>R</sup>   | Novagen                      |
| pCHAP8658  | pSU19 containing <i>pulG</i>                                                    | p15A <i>ori</i> , Cm <sup>R</sup>  | (Nivaskumar et al., 2014)    |
| pCHAP8255  | <i>pulA</i> <sub>NA</sub> in pSU19                                              | p15A <i>ori</i> , Cm <sup>R</sup>  | (Cisneros et al., 2012)      |
| pCHAP4486  | <i>pelBsp-pulA</i> <sub>NA</sub> -Thr-His <sub>x6</sub>                         | ColE1 <i>ori</i> , Ap <sup>R</sup> | (East et al., 2016)          |
| pCHAP8861  | <i>ApulA</i> in pCHAP8185                                                       | ColE1 <i>ori</i> , Ap <sup>R</sup> | Olivera Francetic            |
| pCHAP8200  | all <i>pul</i> genes <i>ApulE</i> in pCHAP8185                                  | ColE1 <i>ori</i> , Ap <sup>R</sup> | (Santos-Moreno et al., 2017) |
| pUT18c     | <i>lacZp-cyaAT18</i>                                                            | ColE1 <i>ori</i> , Ap <sup>R</sup> | (Karimova et al., 1998b)     |
| pKT25      | <i>lacZp-cyaAT25</i>                                                            | p15A <i>ori</i> , Km <sup>R</sup>  | (Karimova et al., 1998b)     |
| pCHAP7330  | pUT18c- <i>pulG</i>                                                             | ColE1 <i>ori</i> , Ap <sup>R</sup> | (Nivaskumar et al., 2014)    |
| pCHAP7332  | pKT25- <i>pulG</i>                                                              | p15A <i>ori</i> , Km <sup>R</sup>  | (Nivaskumar et al., 2014)    |
| pCHAP8116  | pUT18c- <i>malF</i> -TMS                                                        | ColE1 <i>ori</i> , Ap <sup>R</sup> | Olivera Francetic            |
| pCHAP8117  | pKT25- <i>malF</i> -TMS                                                         | p15A <i>ori</i> , Km <sup>R</sup>  | Olivera Francetic            |
| pCHAP8123  | pUT18c- <i>malF</i> -TMS- <i>pulA</i>                                           | ColE1 <i>ori</i> , Ap <sup>R</sup> | Olivera Francetic            |
| pCHAP8124  | pKT25- <i>malF</i> -TMS- <i>pulA</i>                                            | p15A <i>ori</i> , Km <sup>R</sup>  | Olivera Francetic            |
| pCHAP8905  | pUT18c- <i>malF</i> -TMS- <i>pulA</i> $\Delta$ <i>Ins</i>                       | ColE1 <i>ori</i> , Ap <sup>R</sup> | Olivera Francetic            |
| pCHAP8906  | pKT25- <i>malF</i> -TMS- <i>pulA</i> $\Delta$ <i>Ins</i>                        | p15A <i>ori</i> , Km <sup>R</sup>  | Olivera Francetic            |

|           |                                                               |                                    |                           |
|-----------|---------------------------------------------------------------|------------------------------------|---------------------------|
| pCHAP8588 | pUT18c- <i>malF</i> -TMS- <i>pulA</i> N1                      | ColE1 <i>ori</i> , Ap <sup>R</sup> | Olivera Francetic         |
| pCHAP8841 | pKT25- <i>malF</i> -TMS- <i>pulA</i> N1                       | p15A <i>ori</i> , Km <sup>R</sup>  | Olivera Francetic         |
| pCHAP8266 | pUT18c- <i>pulB</i>                                           | ColE1 <i>ori</i> , Ap <sup>R</sup> | Olivera Francetic         |
| pCHAP8267 | pKT25- <i>pulB</i>                                            | p15A <i>ori</i> , Km <sup>R</sup>  | Olivera Francetic         |
| pCHAP8113 | pUT18c- <i>pulC</i>                                           | ColE1 <i>ori</i> , Ap <sup>R</sup> | (Nivaskumar et al., 2016) |
| pCHAP8119 | pKT25- <i>pulC</i>                                            | p15A <i>ori</i> , Km <sup>R</sup>  | (Nivaskumar et al., 2016) |
| pCHAP8364 | pUT18c- <i>pulF</i>                                           | ColE1 <i>ori</i> , Ap <sup>R</sup> | (Nivaskumar et al., 2016) |
| pCHAP8365 | pKT25- <i>pulF</i>                                            | p15A <i>ori</i> , Km <sup>R</sup>  | (Nivaskumar et al., 2016) |
| pCHAP8154 | pUT18c- <i>pulM</i>                                           | ColE1 <i>ori</i> , Ap <sup>R</sup> | (Nivaskumar et al., 2016) |
| pCHAP8155 | pKT25- <i>pulM</i>                                            | p15A <i>ori</i> , Km <sup>R</sup>  | (Nivaskumar et al., 2016) |
| pCHAP8256 | pUT18c- <i>pulH</i>                                           | ColE1 <i>ori</i> , Ap <sup>R</sup> | (Nivaskumar et al., 2016) |
| pCHAP8257 | pKT25- <i>pulH</i>                                            | p15A <i>ori</i> , Km <sup>R</sup>  | (Cisneros et al., 2012)   |
| pCHAP8245 | pUT18c- <i>pulI</i>                                           | ColE1 <i>ori</i> , Ap <sup>R</sup> | (Cisneros et al., 2012)   |
| pCHAP8248 | pKT25- <i>pulI</i>                                            | p15A <i>ori</i> , Km <sup>R</sup>  | (Cisneros et al., 2012)   |
| pCHAP8249 | pUT18c- <i>pulJ</i>                                           | ColE1 <i>ori</i> , Ap <sup>R</sup> | (Cisneros et al., 2012)   |
| pCHAP8246 | pKT25- <i>pulJ</i>                                            | p15A <i>ori</i> , Km <sup>R</sup>  | (Nivaskumar et al., 2016) |
| pCHAP8247 | pUT18c- <i>pulK</i>                                           | ColE1 <i>ori</i> , Ap <sup>R</sup> | (Nivaskumar et al., 2016) |
| pCHAP8250 | pKT25- <i>pulK</i>                                            | p15A <i>ori</i> , Km <sup>R</sup>  | (Cisneros et al., 2012)   |
| pMS1222   | pUT18c- <i>pulL</i>                                           | ColE1 <i>ori</i> , Ap <sup>R</sup> | This study                |
| pMS1229   | pKT25- <i>pulL</i>                                            | p15A <i>ori</i> , Km <sup>R</sup>  | This study                |
| pMS5005   | M.j. p-benzoylphenylalanine RS (2 copies +tRNA) in pCDFDuet-1 | CDF <i>Ori</i> , Sm <sup>R</sup>   | This study                |
| pMS5008   | PulG stop (Am135Oc) in pCHAP8658                              | p15A <i>ori</i> , Cm <sup>R</sup>  | This study                |
| pMS5009   | PulA in pCHAP8255 S54Am                                       | p15A <i>ori</i> , Cm <sup>R</sup>  | This study                |
| pMS5011   | PulA in pCHAP8255 E71Am                                       | p15A <i>ori</i> , Cm <sup>R</sup>  | This study                |
| pMS5012   | PulA in pCHAP8255 S510Am                                      | p15A <i>ori</i> , Cm <sup>R</sup>  | This study                |
| pMS5013   | PulA in pCHAP8255 D526Am,                                     | p15A <i>ori</i> , Cm <sup>R</sup>  | This study                |
| pMS5014   | PulG in pMS5008 E29Am                                         | p15A <i>ori</i> , Cm <sup>R</sup>  | This study                |
| pMS5015   | PulG in pMS5008 D32Am                                         | p15A <i>ori</i> , Cm <sup>R</sup>  | This study                |
| pMS5016   | PulG in pMS5008 R33Am                                         | p15A <i>ori</i> , Cm <sup>R</sup>  | This study                |
| pMS5017   | PulG in pMS5008 Q65Am                                         | p15A <i>ori</i> , Cm <sup>R</sup>  | This study                |
| pMS5018   | PulG in pMS5008 A70Am                                         | p15A <i>ori</i> , Cm <sup>R</sup>  | This study                |
| pMS5019   | PulG in pMS5008 S72Am                                         | p15A <i>ori</i> , Cm <sup>R</sup>  | This study                |
| pMS5020   | PulG in pMS5008 A73Am                                         | p15A <i>ori</i> , Cm <sup>R</sup>  | This study                |
| pMS5021   | PulG in pMS5008 E74Am                                         | p15A <i>ori</i> , Cm <sup>R</sup>  | This study                |
| pMS5022   | PulG in pMS5008 N79Am                                         | p15A <i>ori</i> , Cm <sup>R</sup>  | This study                |
| pMS5023   | PulG in pMS5008 E82Am                                         | p15A <i>ori</i> , Cm <sup>R</sup>  | This study                |
| pMS5024   | PulG in pMS5008 D97Am                                         | p15A <i>ori</i> , Cm <sup>R</sup>  | This study                |

|         |                                      |                                    |            |
|---------|--------------------------------------|------------------------------------|------------|
| pMS5025 | PulG in pMS5008 L101Am               | p15A <i>ori</i> , Cm <sup>R</sup>  | This study |
| pMS5026 | PulG in pMS5008 Q108Am               | p15A <i>ori</i> , Cm <sup>R</sup>  | This study |
| pMS5027 | PulG in pMS5008 F112Am               | p15A <i>ori</i> , Cm <sup>R</sup>  | This study |
| pMS5028 | PulG in pMS5008 R78Am                | p15A <i>ori</i> , Cm <sup>R</sup>  | This study |
| pMS5029 | PulAΔN1 in pCHAP8255                 | p15A <i>ori</i> , Cm <sup>R</sup>  | This study |
| pMS5030 | pUT18c-PulAΔN1 in pCHAP8123          | ColE1 <i>ori</i> , Ap <sup>R</sup> | This study |
| pMS5031 | pKT25-pulAΔN1 in pCHAP8124           | p15A <i>ori</i> , Km <sup>R</sup>  | This study |
| pMS5032 | PulA-His6 in pSU18                   | p15A <i>ori</i> , Cm <sup>R</sup>  | This study |
| pMS5033 | PulA-His in pMS5032 V35Am            | p15A <i>ori</i> , Cm <sup>R</sup>  | This study |
| pMS5034 | PulA-His in pMS5032 E39Am            | p15A <i>ori</i> , Cm <sup>R</sup>  | This study |
| pMS5035 | PulA-His in pMS5032 N70Am            | p15A <i>ori</i> , Cm <sup>R</sup>  | This study |
| pMS5036 | PulAΔIns in pCHAP8255                | p15A <i>ori</i> , Cm <sup>R</sup>  | This study |
| pMS5037 | pUT18c- <i>pulK</i> -TM- <i>pulA</i> | ColE1 <i>ori</i> , Ap <sup>R</sup> | This study |
| pMS5038 | pKT25- <i>pulK</i> -TM- <i>pulA</i>  | p15A <i>ori</i> , Km <sup>R</sup>  | This study |
| pMS5040 | PulA-His in pMS5032 E71Am            | p15A <i>ori</i> , Cm <sup>R</sup>  | This study |
| pMS5041 | PulA-His in pMS5032 T72Am            | p15A <i>ori</i> , Cm <sup>R</sup>  | This study |
| pMS5042 | PulA-His in pMS5032 C73Am            | p15A <i>ori</i> , Cm <sup>R</sup>  | This study |
| pMS5043 | PulA-His in pMS5032 A78Am            | p15A <i>ori</i> , Cm <sup>R</sup>  | This study |
| pMS5044 | PulA-His in pMS5032 A81A             | p15A <i>ori</i> , Cm <sup>R</sup>  | This study |
| pMS5045 | PulA-His in pMS5032 W83Am            | p15A <i>ori</i> , Cm <sup>R</sup>  | This study |
| pMS5046 | PulA-His in pMS5032 N84Am            | p15A <i>ori</i> , Cm <sup>R</sup>  | This study |
| pMS5047 | PulA-His in pMS5032 D85Am            | p15A <i>ori</i> , Cm <sup>R</sup>  | This study |
| pMS5048 | PulA-His in pMS5032 S87Am            | p15A <i>ori</i> , Cm <sup>R</sup>  | This study |
| pMS5049 | PulA-His in pMS5032 D120Am           | p15A <i>ori</i> , Cm <sup>R</sup>  | This study |
| pMS5050 | PulA-His in pMS5032 R128Am           | p15A <i>ori</i> , Cm <sup>R</sup>  | This study |
| pMS5051 | PulA-His in pMS5032 D133Am           | p15A <i>ori</i> , Cm <sup>R</sup>  | This study |
| pMS5052 | PulA-His in pMS5032 E475Am           | p15A <i>ori</i> , Cm <sup>R</sup>  | This study |
| pMS5053 | PulA-His in pMS5032 F476Am           | p15A <i>ori</i> , Cm <sup>R</sup>  | This study |
| pMS5054 | PulA-His in pMS5032 A484Am           | p15A <i>ori</i> , Cm <sup>R</sup>  | This study |
| pMS5055 | PulA-His in pMS5032 F502Am           | p15A <i>ori</i> , Cm <sup>R</sup>  | This study |
| pMS5056 | PulA-His in pMS5032 G509Am           | p15A <i>ori</i> , Cm <sup>R</sup>  | This study |
| pMS5057 | PulA-His in pMS5032 Q521Am           | p15A <i>ori</i> , Cm <sup>R</sup>  | This study |
| pMS5058 | PulA-His in pMS5032 Q525Am           | p15A <i>ori</i> , Cm <sup>R</sup>  | This study |
| pMS5059 | PulA-His in pMS5032 E492Am           | p15A <i>ori</i> , Cm <sup>R</sup>  | This study |
| pMS5060 | PulA-His in pMS5032 K498Am           | p15A <i>ori</i> , Cm <sup>R</sup>  | This study |
| pMS5061 | PulA-His in pMS5032 E501Am           | p15A <i>ori</i> , Cm <sup>R</sup>  | This study |
| pMS5062 | PulA-His in pMS5032 Y505Am           | p15A <i>ori</i> , Cm <sup>R</sup>  | This study |

|         |                                                                         |                                    |                   |
|---------|-------------------------------------------------------------------------|------------------------------------|-------------------|
| pMS5063 | PulA-His in pMS5032 D507Am                                              | p15A <i>ori</i> , Cm <sup>R</sup>  | This study        |
| pMS5064 | pUT18c- <i>pulK</i> -TMS                                                | ColE1 <i>ori</i> , Ap <sup>R</sup> | This study        |
| pMS5065 | pKT25- <i>pulK</i> -TMS                                                 | p15A <i>ori</i> , Km <sup>R</sup>  | This study        |
| pMS5072 | All <i>pul</i> genes with <i>pulA</i> <sub>NA</sub> -His6               | ColE1 <i>ori</i> , Ap <sup>R</sup> | This study        |
| pMS5073 | All <i>pul</i> genes with <i>pulA</i> <sub>NA</sub> -Strep              | ColE1 <i>ori</i> , Ap <sup>R</sup> | This study        |
| pSM5076 | PulG in pMS5008 Q91Am                                                   | p15A <i>ori</i> , Cm <sup>R</sup>  | This study        |
| pSM5077 | PulG in pMS5008 N123Am                                                  | p15A <i>ori</i> , Cm <sup>R</sup>  | This study        |
| pMS5078 | All <i>pul</i> genes with <i>pulA</i> <sub>NA</sub> Δ <i>Ins</i> -strep | ColE1 <i>ori</i> , Ap <sup>R</sup> | This study        |
| pMS5079 | <i>pulΔE</i> genes with <i>pulA</i> <sub>NA</sub> -Strep                | ColE1 <i>ori</i> , Ap <sup>R</sup> | This study        |
| pMS1196 | pCHAP8185 <i>pulA</i> Y505Am Strep                                      | ColE1 <i>ori</i> , Ap <sup>R</sup> | Olivera Francetic |
| pMS1203 | pCHAP8185 <i>pulA</i> E492Am Strep                                      | ColE1 <i>ori</i> , Ap <sup>R</sup> | Olivera Francetic |
| pMS1204 | pCHAP8185 <i>pulA</i> Q525Am Strep                                      | ColE1 <i>ori</i> , Ap <sup>R</sup> | Olivera Francetic |
| pMS1207 | pCHAP8185 <i>pulA</i> Q484Am Strep                                      | ColE1 <i>ori</i> , Ap <sup>R</sup> | Olivera Francetic |
| pMS1208 | pCHAP8185 <i>pulA</i> F476Am Strep                                      | ColE1 <i>ori</i> , Ap <sup>R</sup> | Olivera Francetic |
| pMS1209 | pCHAP8185 <i>pulA</i> E475Am Strep                                      | ColE1 <i>ori</i> , Ap <sup>R</sup> | Olivera Francetic |

**Table 3. Oligonucleotides used for construction of plasmids.**

| Primer Name    | Sequence (5'--3')               |
|----------------|---------------------------------|
| PulG E29Am For | GCAACAAGTAGAAGGCCGACCGGCAAA     |
| PulG E29Am Rev | TCGGCCTTCTACTTGTGCCCATCAGG      |
| PulG D32Am For | AAAAGGCCTAGCGGCAAAAAGTGGTCA     |
| PulG D32Am Rev | TTTTGCCGCTAGGCCTTTTCCTTGTTG     |
| PulG R33Am For | AGGAAAAGGCCGACTAGCAAAAAGTGGTCAG |
| PulG R33Am Rev | CTGACCACTTTTTGCTAGTCGGCCTTTTCCT |
| PulG Q65Am For | GAGCAGGGGCTGTAGGCGCTGGTCAGCG    |
| PulG Q65Am Rev | CGCTGACCAGCGCCTACAGCCCCTGCTC    |
| PulG S69Am For | TGCAGGCGCTGGTCTAGGCGCCATCCGCCGA |
| PulG S69Am Rev | TCGGCGGATGGCGCCTAGACCAGCGCCTGCA |
| PulG A70Am For | GCGCTGGTCAGCTAGCCATCCGCCGAG     |
| PulG A70Am Rev | CTCGGCGGATGGCTAGCTGACCAGCGC     |
| PulG S72Am For | GTCAGCGCGCCATAGGCCGAGCCGCAC     |
| PulG S72Am Rev | GTGCGGCTCGGCCTATGGCGCGCTGAC     |
| PulG A73Am For | AGCGCGCCATCCTAGGAGCCGCACGCGCGCA |
| PulG A73Am Rev | TGCGCGCGTGGCGCTCCTAGGATGGCGCGCT |
| PulG E74Am For | GCCATCCGCCTAGCCGCACGCGCGCAAC    |
| PulG E74Am Rev | GTTGCGCGCGTGGCGCTAGGCGGATGGC    |
| PulG H76Am For | ATCCGCCGAGCCGTAGGCGCGCAACTAT    |
| PulG H76Am Rev | ATAGTTGCGCGCCTACGGCTCGGCGGAT    |

|                        |                                     |
|------------------------|-------------------------------------|
| PulG A77Am For         | TCCGCCGAGCCGCACTAGCGCAACTATCCCGA    |
| PulG A77Am Rev         | TCGGGATAGTTGCGCTAGTGCGGCTCGGCGGA    |
| PulG R78Am For         | CCGAGCCGCACGCGTAGAACTATCCCGAAGG     |
| PulG R78Am Rev         | CCTTCGGGATAGTTCTACGCGTGCGGCTCGG     |
| PulG N79Am For         | AGCCGCACGCGCGCTAGTATCCCGAAGGCGG     |
| PulG N79Am Rev         | CCGCCTTCGGGATACTAGCGCGCGTGCGGCT     |
| PulG E82Am For         | GCAACTATCCCTAGGGCGGCTATATCC         |
| PulG E82Am Rev         | GGATATAGCCGCCCTAGGGATAGTTGC         |
| PulG Q91Am For         | TCCGCCGTCTGCCGTAGGACCCCTGGGGCAG     |
| PulG Q91Am Rev         | CTGCCCCAGGGGTCCTACGGCAGACGGCGGA     |
| PulG S96Am For         | AGGACCCCTGGGGCTAGGATTACCAACTGCT     |
| PulG S96Am Rev         | AGCAGTTGGTAATCCTAGCCCCAGGGGTCTT     |
| PulG D97Am For         | ACCCCTGGGGCAGCTAGTACCAACTGCTGA      |
| PulG D97Am Rev         | TCAGCAGTTGGTACTAGCTGCCCCAGGGGT      |
| PulG L101Am For        | GATTACCAACTGTAGAGCCCCGGCCAG         |
| PulG L101Am Rev        | CTGGCCGGGGCTCTACAGTTGGTAATC         |
| PulG Q105Am For        | CTGAGCCCCGGCTAGCATGGTCAGGTG         |
| PulG Q105Am Rev        | CACCTGACCATGCTAGCCGGGGCTCAG         |
| PulG Q108Am For        | CGGCCAGCATGGTTAGGTGGATATTTTCA       |
| PulG Q108Am Rev        | TGAAAATATCCACCTAACCATGCTGGCCG       |
| PulG F112Am For        | GTCAGGTGGATATTTAGAGCCTCGGGCCGGA     |
| PulG F112Am Rev        | TCCGGCCCCGAGGCTCTAAATATCCACCTGAC    |
| PulG L114Am For        | TGGATATTTTCAGCTAGGGGCCGGACGGCGT     |
| PulG L114Am Rev        | ACGCCGTCCGGCCCCTAGCTGAAAATATCCA     |
| PulG N123Am For        | GCGTGCCGGAGAGCTAGGACGATATCGGCAA     |
| PulG N123Am Rev        | TTGCCGATATCGTCCTAGCTCTCCGGCACGC     |
| PulG Stop(Am135Oc) For | GATCGGGAAGAAATAATGCGACAGCGCGGCTTTAC |
| PulG Stop(Am135Oc) Rev | GTAAAGCCGCGCTGTGCGATTATTTCTTCCCGATC |
| PulG K133Am For        | CGATCGGGTAGAAATAATGCGACAGCG         |
| PulG K133Am Rev        | CATTATTTCTACCCGATCGTCCAGTTG         |
| PulG R78Am For II      | CGCACGCGTAGAACTATCCCGAAGGCG         |
| PulG R78Am Rev II      | GGATAGTTCTACGCGTGCGGCTCGGCG         |
| PulG Q91Am For II      | GTCTGCCGTAGGACCCCTGGGGCAG           |
| PulG Q91Am Rev II      | CAGGGGTCTACGGCAGACGGCGGATATAG       |
| PulG N123Am For II     | CGGAGAGCTAGGACGATATCGGCAACT         |
| PulG N123Am Rev II     | GATATCGTCCTAGCTCTCCGGCACGCCGTC      |
| PulG R87Am For         | GCTATATCTAGCGTCTGCCGCAGGACC         |
| PulG R87Am Rev         | GGCAGACGCTAGATATAGCCGCCTTCG         |
| PulG R88Am For         | ATATCCGCTAGCTGCCGCAGGACCCCTG        |
| PulG R88Am Rev         | CTGCGGCAGCTAGCGGATATAGCCGCCT        |
| PulG D97Am For         | GGGGCAGCTAGTACCAACTGCTGAGCC         |
| PulG D97Am Rev         | CAGTTGGTACTAGCTGCCCCAGGGGTCC        |

|                    |                                  |
|--------------------|----------------------------------|
| PulA S54/S55Am For | TACAGCAGCAAAAACCTCTATCTCTGG      |
| PulA S55Am Rev     | GTCCGCCGCGCTCTAGCTGGTGATGCC      |
| PulA S54Am Rev     | GTCCGCCGCGCTACTCTAGGTGATGCCGGAAT |
| PulA E71Am For     | GGAACAACCTAGACCTGCGACGCGCTGA     |
| PulA E71Am Rev     | TCGCAGGTCTAGTTGTTCCAGAGATAG      |
| PulA S510Am For    | ACAGCGGCTAGACGGTGAAGAGGTAC       |
| PulA S510Am Rev    | TCCACCGTCTAGCCGCTGTCGCAGTAT      |
| PulA D526Am For    | ATAGCCAGTAGAATCCGCAGGTGCAGG      |
| PulA D526Am Rev    | TGCGGATTCTACTGGCTATCGCTCTGC      |
| PulA V35Am For     | GCGAAGCGTAGACGGCGGTAGAAAACC      |
| PulA V35Am Rev     | ACCGCCGTCTACGCTTCGCCGGGCACC      |
| PulA E39Am For     | CGGCGGTATAGAACCAGGCGGTCATCC      |
| PulA E39Am Rev     | GCCTGGTTCTATAACCGCCGTGACCGCTTC   |
| PulA N70Am For     | GGAACAACCTAGACCTGCGACGCGCTGAG    |
| PulA N70Am Rev     | CAGGTTTCCTAGTTCCAGAGATAGAGG      |
| PulA T72Am For     | CAACGAATAGTGCGACGCGCTGAGC        |
| PulA T72Am Rev     | CGTCGCACTATTCGTTGTTCCAGAG        |
| PulA C73Am For     | GAAACCTAGGACGCGCTGAGCGCGC        |
| PulA C73Am Rev     | GCGCGTCCTAGGTTTCGTTGTTCC         |
| PulA A78Am For     | CGCTGAGCTAGCCGGTGGCCGACTGG       |
| PulA A78Am Rev     | CCACCGGCTAGCTCAGCGCGTCGCAG       |
| PulA A81Am For     | CCGGTGTAGGACTGGAATGACGTC         |
| PulA A81Am Rev     | CCAGTCCTACACCGGCGCGCTCAG         |
| PulA W83Am For     | GGCCGACTAGAATGACGTCAGCACC        |
| PulA W83Am Rev     | CGTCATTCTAGTCGGCCACCGGCGC        |
| PulA N84Am For     | CCGACTGGTAGGACGTCAGCACCACGC      |
| PulA N84Am Rev     | CTGACGTCCTACCAGTCGGCCACCGGC      |
| PulA D85Am For     | CTGGAATTAGGTCAGCACCACGCCGAG      |
| PulA D85Am Rev     | TGCTGACCTAATTCCAGTCGGCCAC        |
| PulA S87Am For     | GACGTCTAGACCACGCCGAGCGG          |
| PulA S87Am Rev     | CGTGGTCTAGACGTCATTCCAGTC         |
| PulA D120Am For    | GGCACCTAGAAGCTGATCGATAG          |
| PulA D120Am Rev    | CAGCTTCTAGGTGCCGTGCGGAC          |
| PulA R128Am For    | CGACCTGTAGGTGGCGTTTGCGGAC        |
| PulA R128Am Rev    | CGCCACCTACAGGTCGCTATCGATC        |
| PulA D133Am For    | CGTTTGGCTAGTTCACCGATCGTACGG      |
| PulA D133Am Rev    | TCGGTGAACCTAGCCAAACGCCACCCGC     |
| PulA E475Am For    | CGGTGAACTAGTTCAGCGACAAGGTC       |
| PulA E475Am Rev    | GCTGAACTAGTTCACCGTCGCCAGATCGA    |
| PulA F476Am For    | GAACGAGTAGAGCGACAAGGTCGCCG       |
| PulA F476Am Rev    | TTGTGCTCTACTCGTTCACCGTCGCCAGA    |
| PulA Q484Am For    | CGATATTTAGCAACCGTTCAGTCG         |

|                       |                                  |
|-----------------------|----------------------------------|
| PulA Q484Am Rev       | GGTTGCTAAATATCGGCGACCTTG         |
| PulA F502Am For       | CAGCGAGTAGGCCGGATACTGCGACAG      |
| PulA F502Am Rev       | TATCCGGCCTACTCGCTGCTTTTCAC       |
| PulA G509Am For       | CGACAGCTAGTCCACGGTGGAAGAG        |
| PulA G509Am Rev       | CCGTGGACTAGCTGTGCGAGTATCC        |
| PulA Q521Am For       | GCTTAAGTAGAGCGATAGCCAGG          |
| PulA Q521Am Rev       | TCGCTCTACTTAAGCTGATTCAG          |
| PulA Q525Am For       | CGATAGCTAGGATAATCCGCAGG          |
| PulA Q525Am Rev       | TTATCCTAGCTATCGCTCTGCTTAAG       |
| PulA NcoI For(Rev)    | GATGGCCATGGATAACGGATC            |
| PulA AscI For         | GACTATAACGGCGCGCCTGG             |
| PulA-His XbaI Rev     | CAGGTCGACTCTAGAATCAGTG           |
| PulA-Strep XbaI For   | CACTCTAGATTATTTTCGAACTGCGGG      |
| PulA E492Am For       | GCCTGTGTTAGGTGAACAGCGCGGTGA      |
| PulA E492Am Rev       | CTGTTACCTAACACAGGCGACTGAAC       |
| PulA K498Am For       | GCGCGGTGTAGAGCAGCGAGTTCGCCG      |
| PulA K498Am Rev       | CGCTGCTCTACACCGCGCTGTTAC         |
| PulA E501Am For       | GAAAAGCAGCTAGTTCGCCGGATACTGCG    |
| PulA E501Am Rev       | CCGGCGAACTAGCTGCTTTTCACCGCGCT    |
| PulA Y505Am For       | CGCCGGATAGTGCACAGCGGCTCCAC       |
| PulA Y505Am Rev       | CTGTCGCACTATCCGGCGAACTCGC        |
| PulA D507Am For       | GATACTGCTAGAGCGGCTCCACGGTGG      |
| PulA D507Am Rev       | GAGCCGCTCTAGCAGTATCCGGCGAAC      |
| PulAΔN1 Rev           | CCACTCGAGCTGATTATCAGGAGTGTCG     |
| PulAΔN1 For           | CACCTCGAGGGCGTCGCGCTGGCCGAAG     |
| PulAΔIns For          | CACCTCGAGGGCTACGATCCGTTCCAC      |
| PulAΔIns Rev          | CACCTCGAGCTCGTTACCGTCGCCAGAT     |
| T18/25 EcoRI For      | CACGAATTTCGTCGACCTGCAGGCATGC     |
| PulK-TM Och-EcoRI Rev | CACGAATTCTTAGAGATCCGCGTACATGGCG  |
| PulA-N1 Och-EcoRI Rev | CACGAATTCTTAAAAAGCGGCGCGGAAGGCGT |

## 5.3 DNA transformation

### 5.3.1 Preparation and transformation to CaCl<sub>2</sub> competent cells and transformation

To prepare chemically competent bacteria, the cells were grown in LB at 37°C to an early log phase, pelleted by centrifugation for 5 min at 7000 g and cooled on ice for 10 min. Bacteria were washed with cold solution of 0.1 M CaCl<sub>2</sub> and incubated on ice for 20 min in 0.1 M

CaCl<sub>2</sub>. The bacteria were pelleted by centrifugation as above, resuspended in 0.1 M CaCl<sub>2</sub> containing 15% glycerol and frozen in aliquots at -80°C. For transformation, the competent cells were thawed on ice, DNA was added and incubated 30-60 min on ice, heat shocked for 1 minute in a 42°C water bath. The tubes were placed back on ice for 2 min, and the bacteria were plated on LB plate with appropriate antibiotics.

### 5.3.2 Preparation of ultra-competent cells and transformation

To prepare ultracompetent cells, the bacteria of strain DH5α *F'lacI<sup>Q</sup>* were grown in 500 mL of SOB medium at 18°C to OD<sub>600nm</sub> of 0.6. Bacteria were pelleted by centrifugation at 4°C and resuspended in 80 mL of ice-cold transformation buffer (TB) containing 10 mM PIPES, 15 mM CaCl<sub>2</sub>, 250 mM KCl and 55 mM MnCl<sub>2</sub>. After 20-min incubation on ice, bacteria were pelleted by centrifugation and resuspended in 20 mL of ice-cold TB, to which 1.4 mL of DMSO was added. The bacteria were aliquoted on ice and frozen at -80°C.

For transformation, ultracompetent bacteria were thawed on ice and DNA was added. After incubation for 10-30 min on ice bacteria were plated directly on pre-warmed selective plates.

### 5.3.3 TSS transformation

Bacteria were grown in LB medium (with antibiotics if required) to mid-log phase (typically for 2 hours). The culture was transferred in an Eppendorf tube and centrifuged at 10 000 rpm for 5 minutes. The supernatant was discarded and the bacteria were resuspended in 100 µL of cold TSS buffer containing LB, 5 % DMSO, 10 % PEG 3000 and 50 mM MgCl<sub>2</sub>. The bacteria were mixed with DNA, incubated on ice for 30 minutes, and plated on LB with antibiotics.

## 5.4 Protein electrophoresis and Western blot analysis

Total bacterial extracts were analyzed on denaturing sodium dodecyl-sulphate polyacrylamide gel electrophoresis (SDS-PAGE) using the Tris-tricine gels (Schägger & von Jagow, 1987) or Tris-Glycine gels. Proteins were transferred on nitrocellulose membranes using the Fast Blot semi-dry transfer system and one-step transfer buffer (Thermo). Membranes were blocked with 5% skim milk in Tris-buffer saline containing 0.05% Tween-20 (TBST) for one hour and probed with primary antibodies diluted in TBST. Anti-PulA, anti-PulC, anti-PulD,

anti-PulF and anti-PulM sera were used at 1:1000 dilution, anti-PulJ and anti-PulL sera were diluted at 1:500, and anti PulG sera were diluted at 1:2000. After 4 washes for 10 minutes in TBST, membranes were incubated 1 hour in the secondary goat anti-rabbit antibodies coupled to horse-radish peroxidase (Amersham) (1:10000). Fluorescence signals were developed with ECL2 (Thermo) and recorded on Typhoon FLA9000 (GE, Cytiva), or with ECL chemiluminescence kit (*Thermo*) and recorded on Amersham 680 imager.

## 5.5 *pBpa* incorporation

Protocol 1: Transformed bacteria were selected on LB + corresponding antibiotics, then streaked on minimal glycerol plates (M63 minimal medium supplemented with 0.5% glycerol, 10 mM MgSO<sub>4</sub> and 1 mM CaCl<sub>2</sub>, 0.5% B1 and 1.5% agar). Antibiotics were added at following concentrations: 25 µg/mL chloramphenicol (Cm), 100 µg.mL<sup>-1</sup> spectinomycin (Sp) or 100 µg.mL<sup>-1</sup> ampicillin (Ap). Overnight cultures were grown in minimal glycerol medium supplemented with 1 % casamino acid and antibiotics, with maximum aeration at 30°C. The next morning, the OD<sub>600nm</sub> of 1:10 dilutions were measured, and bacteria were diluted to OD<sub>600nm</sub> = 0.5 in fresh minimal glycerol medium with 1% casamino acids. Two cultures were inoculated for each strain (one to add *pBpa* and a negative control). Maltose was added to 0.4% to induce the *pul* genes on pCHAP8185 and derivatives in this step. The bacteria were grown with vigorous shaking for 3-5 hours at 30°C. Then 0.2% L-arabinose was added to induce the tRNA and tRNA synthetase genes.

Protocol 2: Transformed bacteria were selected on LB medium, antibiotics were added at following concentrations: 25 µg.mL<sup>-1</sup> chloramphenicol (Cm), 100 µg.mL<sup>-1</sup> spectinomycin (Sp) or 100 µg.mL<sup>-1</sup> ampicillin (Ap). Overnight cultures were grown in LB medium with antibiotics, with maximum aeration at 30°C. The next morning, the OD<sub>600nm</sub> of 1:10 dilutions were measured, and bacteria were diluted to OD<sub>600nm</sub> = 0.5 in fresh minimal glycerol medium with 1% casamino acids. Two cultures were inoculated for each strain (one to add *pBpa* and a negative control). The bacteria were grown with vigorous shaking for 3-5 hours at 30°C. Then 0.2% L-arabinose was added to induce the tRNA and tRNA synthetase genes, and 0.4% maltose was added to induce the *pul* genes on pCHAP8185 and derivatives.

Then, cultures from either protocol 1 or 2 were grown with shaking for 1 hour at 30°C. Then, 1 mM of freshly prepared *pBpa* was added to one of the two cultures and bacteria were grown for another 3 hours or longer until the desired OD<sub>600nm</sub>. Then I measured the OD<sub>600nm</sub> of

a 1: 10 dilution, and after this step, 1 mL of culture was prepared for analysis by SDS-PAGE or western blot by resuspending the bacteria in SDS sample buffer at a concentration of 10 OD<sub>600nm</sub> .mL<sup>-1</sup>. Alternatively, the cultures were used for *in vivo* site-specific photo-cross-linking assay.

## 5.6 The *in vivo* site-specific photo-cross-linking assay

After *pBpa* incorporation, we measured the OD<sub>600nm</sub> of cultures diluted 1:10, prepared 1 mL of each culture as non-UV irradiated negative control. For UV-irradiation, we used 2 different parameters in this study: ① for the purpose of batch screening, 1 mL of each culture was transferred into 24-well plates on ice, and on aliquot of each strain was UV irradiated using the UV LED lamp (HAMAMATSU, Lightningcure LC-L1V5) ([website 2](#)) by 365nm UV for 10-30s at distance of 1cm (Fig. 71). Then, 1 mL of bacteria were centrifuged, the supernatant was discarded and bacteria were resuspended in SDS sample buffer at a concentration of 10 OD<sub>600nm</sub>.mL<sup>-1</sup>. The samples were kept at -20°C before analyzing on SDS-PAGE; ② for the large volume cultures for purification of the crosslinked proteins, the cultures were transferred to a 30 mL petri dish on ice and UV-irradiated using the UV lamp (Krackeler, Blak-Ray® B-100AP High Intensity UV Lamp) ([website 3](#)) by 365nm UV for 20-40 minutes at distance of 5cm, before harvesting bacteria for the cell lysis and protein purification protocols.

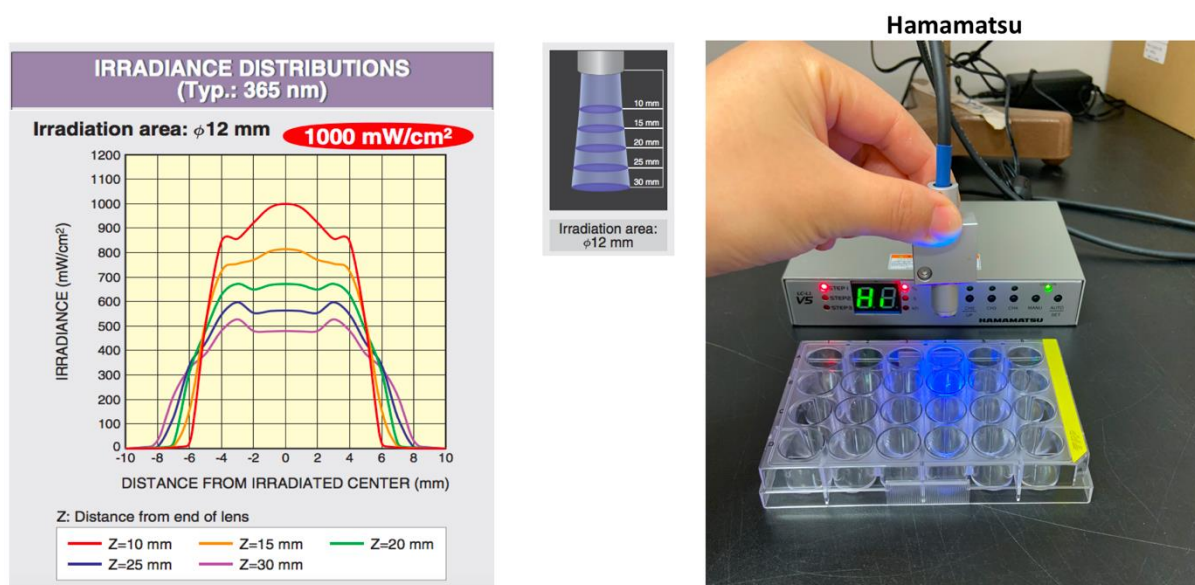


**Figure 72. UV LED light equipment.**

*Lightningcure LC-L1V5* parameters and the demonstration of its use. In real experiments of UV irradiation of bacteria, the plates were always placed on ice, to avoid heating of bacteria.

## 5.7 Protein secretion assays

Bacteria harboring plasmids encoding *pul* genes including *pulA* were cultured overnight in LB containing antibiotics at 30°C. The next day, we measured the OD<sub>600nm</sub> of each preculture, normalized before inoculating into fresh LB supplemented with antibiotics, 0.4 % D-maltose and M63 salts (0.1 vol). Bacteria were cultured at 30°C for several hours until an OD<sub>600nm</sub> >2. Cultures were normalized to OD<sub>600nm</sub> 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 at 4°C, 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 Cell and Supernatant fractions of bacteria were analyzed on 8% or 10% Tris-glycine or Tris-Tricine SDS-PAGE and transferred to the nitrocellulose membranes. The membranes were incubated anti-PulA and then with secondary anti-rabbit antibodies coupled to HRP (Cytiva) and revealed with ECL-2 (Thermo). The fluorescence signals were recorded using Typhoon SLA-9000. Secretion efficiency was calculated as the fraction of total PulA signal detected in the supernatant.

## 5.8 Bacterial adenylate cyclase two-hybrid assay

Bacterial adenylate cyclase two-hybrid (BACTH) assays (Karimova et al., 1998a) 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 (100 µg.mL<sup>-1</sup>) and Km (25 µg.mL<sup>-1</sup>). 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 was plotted and the data was analysed using GraphPad Prism 9 software.

## 5.9 Beta-galactosidase assay

For beta-galactosidase assays (Miller, 1972), the OD<sub>600nm</sub> 1:10 of each culture was measured. We mixed 0.1 mL of bacteria with 0.9 mL of reaction buffer (0.003% SDS, 0.8 mM DTT in M63 salts) in clean Kahn tubes, and added 20 µL chloroform before vortexing each tube for 10 seconds. The reaction tubes were pre-warmed in a 28°C water bath for 5 minutes. To start the reaction, 200 µL of ONPG was added from a stock solution at 4 mg.mL<sup>-1</sup> in M63 medium. The reaction was incubated at 28°C, mixing occasionally; when the yellow color developed, the reactions were stopped by adding 0.5 mL of 1M Na<sub>2</sub>CO<sub>3</sub>. The time was recorded in minutes. The cultures were centrifuged for 5 min at 13000 rpm in 2-mL Eppendorf tubes and 1mL of supernatant was transferred into semi-micro cuvettes to measure the OD<sub>420nm</sub>. The activity (in Miller units) was calculated according to the formula:  $A_{420nm} \times 1000 / \text{reaction time (min)} \times \text{Vol (mL)} \times A_{600nm}$ . The mean and standard deviation was calculated using Prism 9 software (GraphPad).

## 5.10 Cysteine crosslinking

Plasmids encoding PulL and PulM variants were co-transformed in 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, 5 mL of fresh LB Ap Cm medium were inoculated with 0.2 mL of the precultures, and bacteria were incubated for 5 hours at 30°C with shaking. Cultures were normalized to OD<sub>600nm</sub> of 1 and 1 mL of bacteria was centrifuged for 3 min at 16000 x g in the table-top Eppendorf centrifuge. Bacterial pellets were washed once with phosphate buffer saline (PBS) and resuspended in 1 mL of the crosslinking buffer composed of 50 mM MOPS pH 7.0, 5 mM MgCl<sub>2</sub>, 10% glycerol. Bacterial suspensions were prewarmed at 23°C, CuCl<sub>2</sub> was added to 300 µM (final concentration) and incubated for 23 min at 23°C in the Thermomixer with shaking at 650 rpm. To stop the reaction, EDTA pH 8.0 was added to a final concentration of 22.5 mM. The reaction mixtures were pelleted by centrifugation and bacteria were resuspended in the SDS sample buffer at 10 OD<sub>600nm</sub> mL<sup>-1</sup>. Total extracts were boiled for 10 min and analyzed by sodium dodecyl sulfate gel electrophoresis (SDS-PAGE) on 10% Tris-Tricine gels. Proteins were transferred to nitrocellulose membrane (Cytiva) using a semi-dry fast blotting system powerblot (Invitrogen) in 1-Step transfer buffer (Thermo). The membranes were blocked in 5% skim milk in Tris buffer saline solution containing 0.05% Tween-20 (TBST) and probed with anti-PulL

or anti-PulM antibodies. The chemiluminescence ECL kit (Thermo) was used to reveal the signals, which were recorded on Amersham 680 imager.

## **6. References**

- Abby, S. S., Cury, J., Guglielmini, J., Néron, B., Touchon, M., & Rocha, E. P. C. (2016). Identification of protein secretion systems in bacterial genomes. *Scientific Reports*, 6(1), Article 1. <https://doi.org/10.1038/srep23080>
- Abendroth, J., Bagdasarian, M., Sandkvist, M., & Hol, W. G. J. (2004). The Structure of the Cytoplasmic Domain of EpsL, An Inner Membrane Component of the Type II Secretion System of *Vibrio cholerae*: An Unusual Member of the Actin-like ATPase Superfamily. *Journal of Molecular Biology*, 344(3), 619–633. <https://doi.org/10.1016/j.jmb.2004.09.062>
- Abendroth, J., Mitchell, D. D., Korotkov, K. V., Johnson, T. L., Kreger, A., Sandkvist, M., & Hol, W. G. J. (2009). The three-dimensional structure of the cytoplasmic domains of EpsF from the type 2 secretion system of *Vibrio cholerae*. *Journal of Structural Biology*, 166(3), 303–315. <https://doi.org/10.1016/j.jsb.2009.03.009>
- Abendroth, J., Murphy, P., Sandkvist, M., Bagdasarian, M., & Hol, W. G. J. (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(4), 845–855. <https://doi.org/10.1016/j.jmb.2005.02.061>
- Abendroth, J., Rice, A. E., McLuskey, K., Bagdasarian, M., & Hol, W. G. J. (2004). The Crystal Structure of the Periplasmic Domain of the Type II Secretion System Protein EpsM From *Vibrio cholerae*: The Simplest Version of the Ferredoxin Fold. *Journal of Molecular Biology*, 338(3), 585–596. <https://doi.org/10.1016/j.jmb.2004.01.064>
- Abrusci, P., McDowell, M. A., Lea, S. M., & Johnson, S. (2014). Building a secreting nanomachine: A structural overview of the T3SS. *Current Opinion in Structural Biology*, 25, 111–117. <https://doi.org/10.1016/j.sbi.2013.11.001>

- Albrecht, R., & Zeth, K. (2011). Structural Basis of Outer Membrane Protein Biogenesis in Bacteria. *Journal of Biological Chemistry*, 286(31), 27792–27803.  
<https://doi.org/10.1074/jbc.M111.238931>
- Arts, J., de Groot, A., Ball, G., Durand, E., Khattabi, M. E., Filloux, A., Tommassen, J., & Koster, M. (2007). Interaction domains in the *Pseudomonas aeruginosa* type II secretory apparatus component XcpS (GspF). *Microbiology (Reading, England)*, 153(Pt 5), 1582–1592. <https://doi.org/10.1099/mic.0.2006/002840-0>
- Ast, V. M., Schoenhofen, I. C., Langen, G. R., Stratilo, C. W., Chamberlain, M. D., & Howard, S. P. (2002). Expression of the ExeAB complex of *Aeromonas hydrophila* is required for the localization and assembly of the ExeD secretion port multimer. *Molecular Microbiology*, 44(1), 217–231. <https://doi.org/10.1046/j.1365-2958.2002.02870.x>
- Ates, L. S., Houben, E. N. G., & Bitter, W. (2016). Type VII Secretion: A Highly Versatile Secretion System. *Microbiology Spectrum*, 4(1).  
<https://doi.org/10.1128/microbiolspec.VMBF-0011-2015>
- Auclair, S. M., Bhanu, M. K., & Kendall, D. A. (2012). Signal peptidase I: Cleaving the way to mature proteins. *Protein Science*, 21(1), 13–25. <https://doi.org/10.1002/pro.757>
- Ayers, M., Howell, P. L., & Burrows, L. L. (2010). Architecture of the type II secretion and type IV pilus machineries. *Future Microbiology*, 5(8), 1203–1218.  
<https://doi.org/10.2217/fmb.10.76>
- Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: The Keio collection. *Molecular Systems Biology*, 2(1), 2006.0008. <https://doi.org/10.1038/msb4100050>

- Bartolomé, B., Jubete, Y., Martínez, E., & de la Cruz, F. (1991). Construction and properties of a family of pACYC184-derived cloning vectors compatible with pBR322 and its derivatives. *Gene*, 102(1), 75–78. [https://doi.org/10.1016/0378-1119\(91\)90541-I](https://doi.org/10.1016/0378-1119(91)90541-I)
- Berks, B. C. (2015). The Twin-Arginine Protein Translocation Pathway. *Annual Review of Biochemistry*, 84(1), 843–864. <https://doi.org/10.1146/annurev-biochem-060614-034251>
- Berks, B. C., Palmer, T., & Sargent, F. (2005). Protein targeting by the bacterial twin-arginine translocation (Tat) pathway. *Current Opinion in Microbiology*, 8(2), 174–181. <https://doi.org/10.1016/j.mib.2005.02.010>
- Bernal, P., Furniss, R. C. D., Fecht, S., Leung, R. C. Y., Spiga, L., Mavridou, D. A. I., & Filloux, A. (2021). A novel stabilization mechanism for the type VI secretion system sheath. *Proceedings of the National Academy of Sciences of the United States of America*, 118(7), e2008500118. <https://doi.org/10.1073/pnas.2008500118>
- Berry, J.-L., & Pelicic, V. (2015). Exceptionally widespread nanomachines composed of type IV pilins: The prokaryotic Swiss Army knives. *FEMS Microbiology Reviews*, 39(1), 134–154. <https://doi.org/10.1093/femsre/fuu001>
- Berry, J.-L., Phelan, M. M., Collins, R. F., Adomavicius, T., Tønjum, T., Frye, S. A., Bird, L., Owens, R., Ford, R. C., Lian, L.-Y., & Derrick, J. P. (2012). Structure and assembly of a trans-periplasmic channel for type IV pili in *Neisseria meningitidis*. *PLoS Pathogens*, 8(9), e1002923. <https://doi.org/10.1371/journal.ppat.1002923>
- Bhoite, S., van Gerven, N., Chapman, M. R., & Remaut, H. (2019). Curli Biogenesis: Bacterial Amyloid Assembly by the Type VIII Secretion Pathway. *EcoSal Plus*, 8(2). <https://doi.org/10.1128/ecosalplus.ESP-0037-2018>

- Bischof, L. F., Friedrich, C., Harms, A., S gaard-Andersen, L., & van der Does, C. (2016). The Type IV Pilus Assembly ATPase PilB of *Myxococcus xanthus* Interacts with the Inner Membrane Platform Protein PilC and the Nucleotide-binding Protein PilM. *The Journal of Biological Chemistry*, 291(13), 6946–6957.  
<https://doi.org/10.1074/jbc.M115.701284>
- Bleves, S., Viarre, V., Salacha, R., Michel, G. P. F., Filloux, A., & Voulhoux, R. (2010). Protein secretion systems in *Pseudomonas aeruginosa*: A wealth of pathogenic weapons. *International Journal of Medical Microbiology*, 300(8), 534–543.  
<https://doi.org/10.1016/j.ijmm.2010.08.005>
- Blocker, A., Komoriya, K., & Aizawa, S.-I. (2003). Type III secretion systems and bacterial flagella: Insights into their function from structural similarities. *Proceedings of the National Academy of Sciences of the United States of America*, 100(6), 3027–3030.  
<https://doi.org/10.1073/pnas.0535335100>
- Bortoli-German, I., Brun, E., Py, B., Chippaux, M., & Barras, F. (1994). Periplasmic disulphide bond formation is essential for cellulase secretion by the plant pathogen *Erwinia chrysanthemi*. *Molecular Microbiology*, 11(3), 545–553.  
<https://doi.org/10.1111/j.1365-2958.1994.tb00335.x>
- Buddelmeijer, N. (2015). The molecular mechanism of bacterial lipoprotein modification—How, when and why? *FEMS Microbiology Reviews*, 39(2), 246–261.  
<https://doi.org/10.1093/femsre/fuu006>
- Burkinshaw, B. J., & Strynadka, N. C. J. (2014). Assembly and structure of the T3SS. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research*, 1843(8), 1649–1663.  
<https://doi.org/10.1016/j.bbamcr.2014.01.035>

- Burkovski, A. (2013). Cell Envelope of Corynebacteria: Structure and Influence on Pathogenicity. *ISRN Microbiology*, 2013, 935736.  
<https://doi.org/10.1155/2013/935736>
- Camberg, J. L., & Sandkvist, M. (2005). Molecular Analysis of the Vibrio cholerae Type II Secretion ATPase EpsE. *Journal of Bacteriology*, 187(1), 249–256.  
<https://doi.org/10.1128/JB.187.1.249-256.2005>
- Campos, M., Cisneros, D. A., Nivaskumar, M., & Francetic, O. (2013). The type II secretion system – a dynamic fiber assembly nanomachine. *Research in Microbiology*, 164(6), 545–555. <https://doi.org/10.1016/j.resmic.2013.03.013>
- Campos, M., Nilges, M., Cisneros, D. A., & Francetic, O. (2010). Detailed structural and assembly model of the type II secretion pilus from sparse data. *Proceedings of the National Academy of Sciences of the United States of America*, 107(29), 13081–13086. <https://doi.org/10.1073/pnas.1001703107>
- Cascales, E., & Christie, P. J. (2003). The versatile bacterial type IV secretion systems. *Nature Reviews. Microbiology*, 1(2), 137–149. <https://doi.org/10.1038/nrmicro753>
- Chang, Y.-W., Rettberg, L. A., Treuner-Lange, A., Iwasa, J., Søgaaard-Andersen, L., & Jensen, G. J. (2016). Architecture of the type IVa pilus machine. *Science (New York, N.Y.)*, 351(6278), aad2001. <https://doi.org/10.1126/science.aad2001>
- Chatzi, K. E., Sardis, M. F., Karamanou, S., & Economou, A. (2012). Breaking on through to the other side: Protein export through the bacterial Sec system. *Biochemical Journal*, 449(1), 25–37. <https://doi.org/10.1042/BJ20121227>
- Chernyatina, A. A., & Low, H. H. (2019). Core architecture of a bacterial type II secretion system. *Nature Communications*, 10(1), 5437. <https://doi.org/10.1038/s41467-019-13301-3>

- Cherrak, Y., Rapisarda, C., Pellarin, R., Bouvier, G., Bardiaux, B., Allain, F., Malosse, C., Rey, M., Chamot-Rooke, J., Cascales, E., Fronzes, R., & Durand, E. (2018). Biogenesis and structure of a type VI secretion baseplate. *Nature Microbiology*, 3(12), Article 12. <https://doi.org/10.1038/s41564-018-0260-1>
- Chin, J. W., Martin, A. B., King, D. S., Wang, L., & Schultz, P. G. (2002). Addition of a photocrosslinking amino acid to the genetic code of *Escherichia coli*. *Proceedings of the National Academy of Sciences*, 99(17), 11020–11024. <https://doi.org/10.1073/pnas.172226299>
- Cianciotto, N. P. (2009). Many substrates and functions of type II secretion: Lessons learned from *Legionella pneumophila*. *Future Microbiology*, 4, 797–805. <https://doi.org/10.2217/FMB.09.53>
- Cianciotto, N. P., & White, R. C. (2017). Expanding Role of Type II Secretion in Bacterial Pathogenesis and Beyond. *Infection and Immunity*, 85(5), e00014-17. <https://doi.org/10.1128/IAI.00014-17>
- Cisneros, D. A., Bond, P. J., Pugsley, A. P., Campos, M., & Francetic, O. (2012). Minor pseudopilin self-assembly primes type II secretion pseudopilus elongation. *The EMBO Journal*, 31(4), 1041–1053. <https://doi.org/10.1038/emboj.2011.454>
- Cline, K. (2015). Mechanistic Aspects of Folded Protein Transport by the Twin Arginine Translocase (Tat). *The Journal of Biological Chemistry*, 290(27), 16530–16538. <https://doi.org/10.1074/jbc.R114.626820>
- Condemine, G., & Shevchik, V. E. (2000). Overproduction of the secretin OutD suppresses the secretion defect of an *Erwinia chrysanthemi* outB mutant. *Microbiology (Reading, England)*, 146 ( Pt 3), 639–647. <https://doi.org/10.1099/00221287-146-3-639>

- Costa, T. R. D., Felisberto-Rodrigues, C., Meir, A., Prevost, M. S., Redzej, A., Trokter, M., & Waksman, G. (2015). Secretion systems in Gram-negative bacteria: Structural and mechanistic insights. *Nature Reviews. Microbiology*, 13(6), 343–359.  
<https://doi.org/10.1038/nrmicro3456>
- Craig, L., Forest, K. T., & Maier, B. (2019). Type IV pili: Dynamics, biophysics and functional consequences. *Nature Reviews. Microbiology*, 17(7), 429–440.  
<https://doi.org/10.1038/s41579-019-0195-4>
- Craig, L., Pique, M. E., & Tainer, J. A. (2004). Type IV pilus structure and bacterial pathogenicity. *Nature Reviews Microbiology*, 2(5), Article 5.  
<https://doi.org/10.1038/nrmicro885>
- Crane, J. M., & Randall, L. L. (2017). The Sec System: Protein Export in Escherichia coli. *EcoSal/Plus*, 7(2). <https://doi.org/10.1128/ecosalplus.ESP-0002-2017>
- Cranford-Smith, T., & Huber, D. (2018). The way is the goal: How SecA transports proteins across the cytoplasmic membrane in bacteria. *FEMS Microbiology Letters*, 365(11).  
<https://doi.org/10.1093/femsle/fny093>
- Cummings, L. A., Wilkerson, W. D., Bergsbaken, T., & Cookson, B. T. (2006). In vivo, fliC expression by Salmonella enterica serovar Typhimurium is heterogeneous, regulated by ClpX, and anatomically restricted. *Molecular Microbiology*, 61(3), 795–809.  
<https://doi.org/10.1111/j.1365-2958.2006.05271.x>
- d’Enfert, C., Ryter, A., & 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(11), 3531–3538.
- Dalbey, R. E., & Wickner, W. (1985). Leader peptidase catalyzes the release of exported proteins from the outer surface of the Escherichia coli plasma membrane. *Journal of*

- Biological Chemistry*, 260(29), 15925–15931. [https://doi.org/10.1016/S0021-9258\(17\)36347-0](https://doi.org/10.1016/S0021-9258(17)36347-0)
- Datsenko, K. A., & Wanner, B. L. (2000). One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proceedings of the National Academy of Sciences*, 97(12), 6640–6645. <https://doi.org/10.1073/pnas.120163297>
- Dautin, N. (2021). Folding Control in the Path of Type 5 Secretion. *Toxins*, 13(5), 341. <https://doi.org/10.3390/toxins13050341>
- Dautin, N., Karimova, G., Ullmann, A., & Ladant, D. (2000). Sensitive Genetic Screen for Protease Activity Based on a Cyclic AMP Signaling Cascade in *Escherichia coli*. *Journal of Bacteriology*, 182(24), 7060–7066. <https://doi.org/10.1128/JB.182.24.7060-7066.2000>
- Delcour, A. H. (2009). Outer Membrane Permeability and Antibiotic Resistance. *Biochimica et Biophysica Acta*, 1794(5), 808–816. <https://doi.org/10.1016/j.bbapap.2008.11.005>
- D’Enfert, C., & Pugsley, A. p. (1987). A gene fusion approach to the study of pullulanase export and secretion in *Eschenchia coli*. *Molecular Microbiology*, 1(2), 159–168. <https://doi.org/10.1111/j.1365-2958.1987.tb00508.x>
- Denise, R., Abby, S. S., & Rocha, E. P. C. (2019). Diversification of the type IV filament superfamily into machines for adhesion, protein secretion, DNA uptake, and motility. *PLoS Biology*, 17(7), e3000390. <https://doi.org/10.1371/journal.pbio.3000390>
- Diepold, A., & Armitage, J. P. (2015). Type III secretion systems: The bacterial flagellum and the injectisome. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 370(1679), 20150020. <https://doi.org/10.1098/rstb.2015.0020>
- Douzi, B., Ball, G., Cambillau, C., Tegoni, M., & Voulhoux, R. (2011). Deciphering the Xcp *Pseudomonas aeruginosa* Type II Secretion Machinery through Multiple Interactions

- with Substrates\* ♦. *Journal of Biological Chemistry*, 286(47), 40792–40801.  
<https://doi.org/10.1074/jbc.M111.294843>
- Douzi, B., Durand, E., Bernard, C., Alphonse, S., Cambillau, C., Filloux, A., Tegoni, M., & Voulhoux, R. (2009). The XcpV/GspI Pseudopilin Has a Central Role in the Assembly of a Quaternary Complex within the T2SS Pseudopilus\*. *Journal of Biological Chemistry*, 284(50), 34580–34589. <https://doi.org/10.1074/jbc.M109.042366>
- Douzi, B., Filloux, A., & Voulhoux, R. (2012). On the path to uncover the bacterial type II secretion system. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 367(1592), 1059–1072. <https://doi.org/10.1098/rstb.2011.0204>
- Doyle, M. T., & Bernstein, H. D. (2021). BamA forms a translocation channel for polypeptide export across the bacterial outer membrane. *Molecular Cell*, 81(9), 2000-2012.e3. <https://doi.org/10.1016/j.molcel.2021.02.023>
- Dulberger, C. L., Rubin, E. J., & Boutte, C. C. (2020). The mycobacterial cell envelope—A moving target. *Nature Reviews Microbiology*, 18(1), Article 1. <https://doi.org/10.1038/s41579-019-0273-7>
- Dupuy, B., Taha, M. K., Possot, O., Marchal, C., & Pugsley, A. P. (1992). PulO, a component of the pullulanase secretion pathway of *Klebsiella oxytoca*, correctly and efficiently processes gonococcal type IV prepilin in *Escherichia coli*. *Molecular Microbiology*, 6(14), 1887–1894. <https://doi.org/10.1111/j.1365-2958.1992.tb01361.x>
- Durand, E., Bernadac, A., Ball, G., Lazdunski, A., Sturgis, J. N., & Filloux, A. (2003). Type II protein secretion in *Pseudomonas aeruginosa*: The pseudopilus is a multifibrillar and adhesive structure. *Journal of Bacteriology*, 185(9), 2749–2758. <https://doi.org/10.1128/JB.185.9.2749-2758.2003>

- Durand, E., Michel, G., Voulhoux, R., Kürner, J., Bernadac, A., & Filloux, A. (2005). XcpX controls biogenesis of the *Pseudomonas aeruginosa* XcpT-containing pseudopilus. *The Journal of Biological Chemistry*, 280(36), 31378–31389. <https://doi.org/10.1074/jbc.M505812200>
- Durand, E., Nguyen, V. S., Zoued, A., Logger, L., Péhau-Arnaudet, G., Aschtgen, M.-S., Spinelli, S., Desmyter, A., Bardiaux, B., Dujeancourt, A., Roussel, A., Cambillau, C., Cascales, E., & Fronzes, R. (2015). Biogenesis and structure of a type VI secretion membrane core complex. *Nature*, 523(7562), Article 7562. <https://doi.org/10.1038/nature14667>
- East, A., Mechaly, A. E., Huysmans, G. H. M., Bernarde, C., Tello-Manigne, D., Nadeau, N., Pugsley, A. P., Buschiazio, A., Alzari, P. M., Bond, P. J., & Francetic, O. (2016). Structural Basis of Pullulanase Membrane Binding and Secretion Revealed by X-Ray Crystallography, Molecular Dynamics and Biochemical Analysis. *Structure (London, England: 1993)*, 24(1), 92–104. <https://doi.org/10.1016/j.str.2015.10.023>
- Eimer, E., Kao, W.-C., Fröbel, J., Blümmel, A.-S., Hunte, C., & Müller, M. (2018). Unanticipated functional diversity among the TatA-type components of the Tat protein translocase. *Scientific Reports*, 8(1), 1326. <https://doi.org/10.1038/s41598-018-19640-3>
- Escobar, C. A., Douzi, B., Ball, G., Barbat, B., Alphonse, S., Quinton, L., Voulhoux, R., & Forest, K. T. (2021). Structural interactions define assembly adapter function of a type II secretion system pseudopilin. *Structure*, 29(10), 1116-1127.e8. <https://doi.org/10.1016/j.str.2021.05.015>
- Farrell, I. S., Toroney, R., Hazen, J. L., Mehl, R. A., & Chin, J. W. (2005). Photo-cross-linking interacting proteins with a genetically encoded benzophenone. *Nature Methods*, 2(5), Article 5. <https://doi.org/10.1038/nmeth0505-377>

- Ferrandez, Y., & Condemine, G. (2008). Novel mechanism of outer membrane targeting of proteins in Gram-negative bacteria. *Molecular Microbiology*, 69(6), 1349–1357.  
<https://doi.org/10.1111/j.1365-2958.2008.06366.x>
- Filloux, A. (2010). Secretion Signal and Protein Targeting in Bacteria: A Biological Puzzle. *Journal of Bacteriology*, 192(15), 3847–3849. <https://doi.org/10.1128/JB.00565-10>
- Filloux, A. (2022). Bacterial protein secretion systems: Game of types. *Microbiology (Reading, England)*, 168(5). <https://doi.org/10.1099/mic.0.001193>
- Frain, K. M., Robinson, C., & van Dijl, J. M. (2019). Transport of Folded Proteins by the Tat System. *The Protein Journal*, 38(4), 377–388. <https://doi.org/10.1007/s10930-019-09859-y>
- Frain, K. M., van Dijl, J. M., & Robinson, C. (2019). The Twin-Arginine Pathway for Protein Secretion. *EcoSal Plus*, 8(2). <https://doi.org/10.1128/ecosalplus.ESP-0040-2018>
- Francetić, O., Lory, S., & Pugsley, A. P. (1998). A second prepilin peptidase gene in *Escherichia coli* K-12. *Molecular Microbiology*, 27(4), 763–775.  
<https://doi.org/10.1046/j.1365-2958.1998.00723.x>
- Francetić, O., & Pugsley, A. P. (2005). Towards the Identification of Type II Secretion Signals in a Nonacylated Variant of Pullulanase from *Klebsiella oxytoca*. *Journal of Bacteriology*, 187(20), 7045–7055. <https://doi.org/10.1128/JB.187.20.7045-7055.2005>
- Genin, S., & Boucher, C. A. (1994). A superfamily of proteins involved in different secretion pathways in gram-negative bacteria: Modular structure and specificity of the N-terminal domain. *Molecular and General Genetics MGG*, 243(1), 112–118.  
<https://doi.org/10.1007/BF00283883>

- Giltner, C. L., Habash, M., & Burrows, L. L. (2010). *Pseudomonas aeruginosa* Minor Pilins Are Incorporated into Type IV Pili. *Journal of Molecular Biology*, 398(3), 444–461.  
<https://doi.org/10.1016/j.jmb.2010.03.028>
- Giltner, C. L., Nguyen, Y., & Burrows, L. L. (2012). Type IV pilin proteins: Versatile molecular modules. *Microbiology and Molecular Biology Reviews: MMBR*, 76(4), 740–772.  
<https://doi.org/10.1128/MMBR.00035-12>
- Gorasia, D. G., Veith, P. D., & Reynolds, E. C. (2020). The Type IX Secretion System: Advances in Structure, Function and Organisation. *Microorganisms*, 8(8), 1173.  
<https://doi.org/10.3390/microorganisms8081173>
- Grimbergen, A. J., Siebring, J., Solopova, A., & Kuipers, O. P. (2015). Microbial bet-hedging: The power of being different. *Current Opinion in Microbiology*, 25, 67–72.  
<https://doi.org/10.1016/j.mib.2015.04.008>
- Grossman, A. S., Mauer, T. J., Forest, K. T., & Goodrich-Blair, H. (2021). *Description of a widespread bacterial secretion system with chemically diverse substrates* (p. 2020.01.20.912956). bioRxiv. <https://doi.org/10.1101/2020.01.20.912956>
- Gu, S., Shevchik, V. E., Shaw, R., Pickersgill, R. W., & Garnett, J. A. (2017). The role of intrinsic disorder and dynamics in the assembly and function of the type II secretion system. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics*, 1865(10), 1255–1266.  
<https://doi.org/10.1016/j.bbapap.2017.07.006>
- Guilvout, I., Chami, M., Engel, A., Pugsley, A. P., & Bayan, N. (2006). Bacterial outer membrane secretin PulD assembles and inserts into the inner membrane in the absence of its pilotin. *The EMBO Journal*, 25(22), 5241–5249.  
<https://doi.org/10.1038/sj.emboj.7601402>

- Guttenplan, S. B., Shaw, S., & Kearns, D. B. (2013). The cell biology of peritrichous flagella in *Bacillus subtilis*. *Molecular Microbiology*, 87(1), 211–229.  
<https://doi.org/10.1111/mmi.12103>
- Guyer, M. S., Reed, R. R., Steitz, J. A., & Low, K. B. (1981). Identification of a Sex-factor-affinity Site in *E. coli* as  $\gamma\delta$ . *Cold Spring Harbor Symposia on Quantitative Biology*, 45, 135–140. <https://doi.org/10.1101/SQB.1981.045.01.022>
- Guzman, L. M., Weiss, D. S., & Beckwith, J. (1997). Domain-swapping analysis of FtsI, FtsL, and FtsQ, bitopic membrane proteins essential for cell division in *Escherichia coli*. *Journal of Bacteriology*, 179(16), 5094–5103.  
<https://doi.org/10.1128/jb.179.16.5094-5103.1997>
- Hajishengallis, G. (2015). Periodontitis: From microbial immune subversion to systemic inflammation. *Nature Reviews. Immunology*, 15(1), 30–44.  
<https://doi.org/10.1038/nri3785>
- Hamood, A. N., Olson, J. C., Vincent, T. S., & Iglewski, B. H. (1989). Regions of toxin A involved in toxin A excretion in *Pseudomonas aeruginosa*. *Journal of Bacteriology*, 171(4), 1817–1824. <https://doi.org/10.1128/jb.171.4.1817-1824.1989>
- Han, L., Zheng, J., Wang, Y., Yang, X., Liu, Y., Sun, C., Cao, B., Zhou, H., Ni, D., Lou, J., Zhao, Y., & Huang, Y. (2016). Structure of the BAM complex and its implications for biogenesis of outer-membrane proteins. *Nature Structural & Molecular Biology*, 23(3), Article 3.  
<https://doi.org/10.1038/nsmb.3181>
- Hardie, K. R., Lory, S., & Pugsley, A. P. (1996). Insertion of an outer membrane protein in *Escherichia coli* requires a chaperone-like protein. *The EMBO Journal*, 15(5), 978–988.
- Hay, I. D., Belousoff, M. J., Dunstan, R. A., Bamert, R. S., & Lithgow, T. (2018). Structure and Membrane Topography of the Vibrio-Type Secretin Complex from the Type 2

- Secretion System of Enteropathogenic *Escherichia coli*. *Journal of Bacteriology*, 200(5), e00521-17. <https://doi.org/10.1128/JB.00521-17>
- He, S. Y., Schoedel, C., Chatterjee, A. K., & Collmer, A. (1991). Extracellular secretion of pectate lyase by the *Erwinia chrysanthemi* out pathway is dependent upon Sec-mediated export across the inner membrane. *Journal of Bacteriology*, 173(14), 4310–4317. <https://doi.org/10.1128/jb.173.14.4310-4317.1991>
- Herzberg, C., Weidinger, L. A. F., Dörrbecker, B., Hübner, S., Stülke, J., & Commichau, F. M. (2007). SPINE: A method for the rapid detection and analysis of protein–protein interactions in vivo. *PROTEOMICS*, 7(22), 4032–4035. <https://doi.org/10.1002/pmic.200700491>
- Horváthová, L., Žárský, V., Pánek, T., Derelle, R., Pyrih, J., Motyčková, A., Klápšťová, V., Vinopalová, M., Marková, L., Voleman, L., Klimeš, V., Petrů, M., Vaitová, Z., Čepička, I., Hryzáková, K., Harant, K., Gray, M. W., Chami, M., Guilvout, I., ... Doležal, P. (2021). Analysis of diverse eukaryotes suggests the existence of an ancestral mitochondrial apparatus derived from the bacterial type II secretion system. *Nature Communications*, 12(1), Article 1. <https://doi.org/10.1038/s41467-021-23046-7>
- Hospenthal, M. K., Costa, T. R. D., & Waksman, G. (2017). A comprehensive guide to pilus biogenesis in Gram-negative bacteria. *Nature Reviews Microbiology*, 15(6), Article 6. <https://doi.org/10.1038/nrmicro.2017.40>
- Howard, S. P., Estrozi, L. F., Bertrand, Q., Contreras-Martel, C., Strozen, T., Job, V., Martins, A., Fenel, D., Schoehn, G., & Dessen, A. (2019). Structure and assembly of pilotin-dependent and -independent secretins of the type II secretion system. *PLoS Pathogens*, 15(5), e1007731. <https://doi.org/10.1371/journal.ppat.1007731>

- Huysmans, G. H. M. (2016). Folding outer membrane proteins independently of the  $\beta$ -barrel assembly machinery: An assembly pathway for multimeric complexes? *Biochemical Society Transactions*, 44(3), 845–850. <https://doi.org/10.1042/BST20160003>
- Jacobsen, T., Bardiaux, B., Francetic, O., Izadi-Pruneyre, N., & Nilges, M. (2020). Structure and function of minor pilins of type IV pili. *Medical Microbiology and Immunology*, 209(3), 301–308. <https://doi.org/10.1007/s00430-019-00642-5>
- Jansen, K. B., Baker, S. L., & Sousa, M. C. (2015). Crystal Structure of BamB Bound to a Periplasmic Domain Fragment of BamA, the Central Component of the  $\beta$ -Barrel Assembly Machine \*. *Journal of Biological Chemistry*, 290(4), 2126–2136. <https://doi.org/10.1074/jbc.M114.584524>
- Jarrell, K. F., & McBride, M. J. (2008). The surprisingly diverse ways that prokaryotes move. *Nature Reviews Microbiology*, 6(6), Article 6. <https://doi.org/10.1038/nrmicro1900>
- Kamio, Y., & Nikaido, H. (1976). Outer membrane of *Salmonella typhimurium*: Accessibility of phospholipid head groups to phospholipase c and cyanogen bromide activated dextran in the external medium. *Biochemistry*, 15(12), 2561–2570. <https://doi.org/10.1021/bi00657a012>
- Karimova, G., Pidoux, J., Ullmann, A., & Ladant, D. (1998a). A bacterial two-hybrid system based on a reconstituted signal transduction pathway. *Proceedings of the National Academy of Sciences of the United States of America*, 95(10), 5752–5756. <https://doi.org/10.1073/pnas.95.10.5752>
- Karimova, G., Pidoux, J., Ullmann, A., & Ladant, D. (1998b). A bacterial two-hybrid system based on a reconstituted signal transduction pathway. *Proceedings of the National Academy of Sciences*, 95(10), 5752–5756. <https://doi.org/10.1073/pnas.95.10.5752>

- Karlinsey, J. E., Tanaka, S., Bettenworth, V., Yamaguchi, S., Boos, W., Aizawa, S. I., & Hughes, K. T. (2000). Completion of the hook-basal body complex of the *Salmonella typhimurium* flagellum is coupled to FlgM secretion and fliC transcription. *Molecular Microbiology*, 37(5), 1220–1231. <https://doi.org/10.1046/j.1365-2958.2000.02081.x>
- Karuppiah, V., & Derrick, J. P. (2011). Structure of the PilM-PilN Inner Membrane Type IV Pilus Biogenesis Complex from *Thermus thermophilus*. *Journal of Biological Chemistry*, 286(27), 24434–24442. <https://doi.org/10.1074/jbc.M111.243535>
- Kaushik, S., He, H., & Dalbey, R. E. (2022). Bacterial Signal Peptides- Navigating the Journey of Proteins. *Frontiers in Physiology*, 13. <https://www.frontiersin.org/articles/10.3389/fphys.2022.933153>
- Kearns, D. B., & Losick, R. (2005). Cell population heterogeneity during growth of *Bacillus subtilis*. *Genes & Development*, 19(24), 3083–3094. <https://doi.org/10.1101/gad.1373905>
- Kim, K. H., Aulakh, S., & Paetzel, M. (2011). Crystal Structure of  $\beta$ -Barrel Assembly Machinery BamCD Protein Complex\*. *Journal of Biological Chemistry*, 286(45), 39116–39121. <https://doi.org/10.1074/jbc.M111.298166>
- Konovalova, A., Perlman, D. H., Cowles, C. E., & Silhavy, T. J. (2014). Transmembrane domain of surface-exposed outer membrane lipoprotein RcsF is threaded through the lumen of  $\beta$ -barrel proteins. *Proceedings of the National Academy of Sciences*, 111(41), E4350–E4358. <https://doi.org/10.1073/pnas.1417138111>
- Koronakis, V., Hughes, C., & Koronakis, E. (1991). Energetically distinct early and late stages of HlyB/HlyD-dependent secretion across both *Escherichia coli* membranes. *The EMBO Journal*, 10(11), 3263–3272.

- Koronakis, V., Sharff, A., Koronakis, E., Luisi, B., & Hughes, C. (2000). Crystal structure of the bacterial membrane protein TolC central to multidrug efflux and protein export. *Nature*, 405(6789), Article 6789. <https://doi.org/10.1038/35016007>
- Korotkov, K. V., Gray, M. D., Kreger, A., Turley, S., Sandkvist, M., & Hol, W. G. J. (2009). Calcium is essential for the major pseudopilin in the type 2 secretion system. *The Journal of Biological Chemistry*, 284(38), 25466–25470. <https://doi.org/10.1074/jbc.C109.037655>
- Korotkov, K. V., & Hol, W. G. J. (2008). Structure of the GspK–GspI–GspJ complex from the enterotoxigenic Escherichia coli type 2 secretion system. *Nature Structural & Molecular Biology*, 15(5), Article 5. <https://doi.org/10.1038/nsmb.1426>
- Korotkov, K. V., Johnson, T. L., Jobling, M. G., Pruneda, J., Pardon, E., Héroux, A., Turley, S., Steyaert, J., Holmes, R. K., Sandkvist, M., & Hol, W. G. J. (2011). Structural and functional studies on the interaction of GspC and GspD in the type II secretion system. *PLoS Pathogens*, 7(9), e1002228. <https://doi.org/10.1371/journal.ppat.1002228>
- Korotkov, K. V., Krumm, B., Bagdasarian, M., & Hol, W. G. J. (2006). Structural and functional studies of EpsC, a crucial component of the type 2 secretion system from Vibrio cholerae. *Journal of Molecular Biology*, 363(2), 311–321. <https://doi.org/10.1016/j.jmb.2006.08.037>
- Korotkov, K. V., Sandkvist, M., & Hol, W. G. J. (2012). The type II secretion system: Biogenesis, molecular architecture and mechanism. *Nature Reviews. Microbiology*, 10(5), 336–351. <https://doi.org/10.1038/nrmicro2762>

- Kulkarni, S. S., Zhu, Y., Brendel, C. J., & McBride, M. J. (2017). Diverse C-Terminal Sequences Involved in *Flavobacterium johnsoniae* Protein Secretion. *Journal of Bacteriology*, 199(12), e00884-16. <https://doi.org/10.1128/JB.00884-16>
- Lallemant, M., Login, F. H., Guschinskaya, N., Pineau, C., Effantin, G., Robert, X., & Shevchik, V. E. (2013). Dynamic Interplay between the Periplasmic and Transmembrane Domains of GspL and GspM in the Type II Secretion System. *PLoS ONE*, 8(11), e79562. <https://doi.org/10.1371/journal.pone.0079562>
- Lavander, M., Ericsson, S. K., Bröms, J. E., & Forsberg, A. (2006). The twin arginine translocation system is essential for virulence of *Yersinia pseudotuberculosis*. *Infection and Immunity*, 74(3), 1768–1776. <https://doi.org/10.1128/IAI.74.3.1768-1776.2006>
- Lee, H. C., & Bernstein, H. D. (2001). The targeting pathway of *Escherichia coli* presecretory and integral membrane proteins is specified by the hydrophobicity of the targeting signal. *Proceedings of the National Academy of Sciences*, 98(6), 3471–3476. <https://doi.org/10.1073/pnas.051484198>
- Leyton, D. L., Rossiter, A. E., & Henderson, I. R. (2012). From self sufficiency to dependence: Mechanisms and factors important for autotransporter biogenesis. *Nature Reviews Microbiology*, 10(3), Article 3. <https://doi.org/10.1038/nrmicro2733>
- Lindeberg, M., Salmond, G. P. C., & Collmer, A. (1996). Complementation of deletion mutations in a cloned functional cluster of *Erwinia chrysanthemi* out genes with *Erwinia carotovora* out homologues reveals OutC and OutD as candidate gatekeepers of species-specific secretion of proteins via the type II pathway. *Molecular Microbiology*, 20(1), 175–190. <https://doi.org/10.1111/j.1365-2958.1996.tb02499.x>

- Login, F. H., & Shevchik, V. E. (2006). The Single Transmembrane Segment Drives Self-assembly of OutC and the Formation of a Functional Type II Secretion System in *Erwinia chrysanthemi*\*. *Journal of Biological Chemistry*, 281(44), 33152–33162.  
<https://doi.org/10.1074/jbc.M606245200>
- López-Castilla, A., Thomassin, J.-L., Bardiaux, B., Zheng, W., Nivaskumar, M., Yu, X., Nilges, M., Egelman, E. H., Izadi-Pruneyre, N., & Francetic, O. (2017). Structure of the calcium-dependent type 2 secretion pseudopilus. *Nature Microbiology*, 2(12), Article 12. <https://doi.org/10.1038/s41564-017-0041-2>
- Low, H. H., Gubellini, F., Rivera-Calzada, A., Braun, N., Connery, S., Dujeancourt, A., Lu, F., Redzej, A., Fronzes, R., Orlova, E. V., & Waksman, G. (2014). Structure of a type IV secretion system. *Nature*, 508(7497), 550–553. <https://doi.org/10.1038/nature13081>
- Lu, C., Korotkov, K. V., & Hol, W. G. J. (2014). Crystal structure of the full-length ATPase GspE from the *Vibrio vulnificus* type II secretion system in complex with the cytoplasmic domain of GspL. *Journal of Structural Biology*, 187(3), 223–235.  
<https://doi.org/10.1016/j.jsb.2014.07.006>
- Lu, H. M., & Lory, S. (1996). A specific targeting domain in mature exotoxin A is required for its extracellular secretion from *Pseudomonas aeruginosa*. *The EMBO Journal*, 15(2), 429–436. <https://doi.org/10.1002/j.1460-2075.1996.tb00373.x>
- Maffei, B., Francetic, O., & Subtil, A. (2017). Tracking Proteins Secreted by Bacteria: What's in the Toolbox? *Frontiers in Cellular and Infection Microbiology*, 7, 221.  
<https://doi.org/10.3389/fcimb.2017.00221>
- Makarova, K. S., Koonin, E. V., & Albers, S.-V. (2016). Diversity and Evolution of Type IV pili Systems in Archaea. *Frontiers in Microbiology*, 7, 667.  
<https://doi.org/10.3389/fmicb.2016.00667>

- Marsh, J. W., & Taylor, R. K. (1999). Genetic and transcriptional analyses of the *Vibrio cholerae* mannose-sensitive hemagglutinin type 4 pilus gene locus. *Journal of Bacteriology*, 181(4), 1110–1117. <https://doi.org/10.1128/JB.181.4.1110-1117.1999>
- Martin, P. R., Hobbs, M., Free, P. D., Jeske, Y., & Mattick, J. S. (1993). Characterization of pilQ, a new gene required for the biogenesis of type 4 fimbriae in *Pseudomonas aeruginosa*. *Molecular Microbiology*, 9(4), 857–868. <https://doi.org/10.1111/j.1365-2958.1993.tb01744.x>
- McBride, M. J. (2019). Bacteroidetes Gliding Motility and the Type IX Secretion System. *Microbiology Spectrum*, 7(1), 7.1.15. <https://doi.org/10.1128/microbiolspec.PSIB-0002-2018>
- McCallum, M., Benlekber, S., Nguyen, S., Tammam, S., Rubinstein, J. L., Burrows, L. L., & Howell, P. L. (2019). Multiple conformations facilitate PilT function in the type IV pilus. *Nature Communications*, 10(1), 5198. <https://doi.org/10.1038/s41467-019-13070-z>
- McVay, C. S., & Hamood, A. N. (1995). Toxin A secretion in *Pseudomonas aeruginosa*: The role of the first 30 amino acids of the mature toxin. *Molecular and General Genetics MGG*, 249(5), 515–525. <https://doi.org/10.1007/BF00290577>
- Michaelis, S., Chapon, C., D’Enfert, C., Pugsley, A. P., & Schwartz, M. (1985). Characterization and expression of the structural gene for pullulanase, a maltose-inducible secreted protein of *Klebsiella pneumoniae*. *Journal of Bacteriology*, 164(2), 633–638.
- Michel-Souzy, S., Douzi, B., Cadoret, F., Raynaud, C., Quinton, L., Ball, G., & 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. *Journal of*

- Biological Chemistry*, 293(50), 19441–19450.
- <https://doi.org/10.1074/jbc.RA117.001127>
- Miller, J. H. (1972). *Experiments in molecular genetics*. Cold Spring Harbor Laboratory.
- Mukherjee, S., & Kearns, D. B. (2014). The Structure and Regulation of Flagella in *Bacillus subtilis*. *Annual Review of Genetics*, 48(1), 319–340.
- <https://doi.org/10.1146/annurev-genet-120213-092406>
- Nakamura, S., & Minamino, T. (2019). Flagella-Driven Motility of Bacteria. *Biomolecules*, 9(7), E279. <https://doi.org/10.3390/biom9070279>
- Narita, S., & Tokuda, H. (2017). Bacterial lipoproteins; biogenesis, sorting and quality control. *Biochimica et Biophysica Acta (BBA) - Molecular and Cell Biology of Lipids*, 1862(11), 1414–1423. <https://doi.org/10.1016/j.bbalip.2016.11.009>
- Naskar, S., Hohl, M., Tassinari, M., & Low, H. H. (2021). The structure and mechanism of the bacterial type II secretion system. *Molecular Microbiology*, 115(3), 412–424.
- <https://doi.org/10.1111/mmi.14664>
- Natale, P., Brüser, T., & Driessen, A. J. M. (2008). Sec- and Tat-mediated protein secretion across the bacterial cytoplasmic membrane—Distinct translocases and mechanisms. *Biochimica Et Biophysica Acta*, 1778(9), 1735–1756.
- <https://doi.org/10.1016/j.bbamem.2007.07.015>
- Nickerson, N. N., Tosi, T., Dessen, A., Baron, B., Raynal, B., England, P., & Pugsley, A. P. (2011). Outer membrane targeting of secretin PulD protein relies on disordered domain recognition by a dedicated chaperone. *The Journal of Biological Chemistry*, 286(45), 38833–38843. <https://doi.org/10.1074/jbc.M111.279851>

Nikaido, H. (2003). Molecular basis of bacterial outer membrane permeability revisited.

*Microbiology and Molecular Biology Reviews: MMBR*, 67(4), 593–656.

<https://doi.org/10.1128/MMBR.67.4.593-656.2003>

Nivaskumar, M., Bouvier, G., Campos, M., Nadeau, N., Yu, X., Egelman, E. H., Nilges, M., &

Francetic, O. (2014). Distinct docking and stabilization steps of the Pseudopilus conformational transition path suggest rotational assembly of type IV pilus-like fibers. *Structure (London, England: 1993)*, 22(5), 685–696.

<https://doi.org/10.1016/j.str.2014.03.001>

Nivaskumar, M., & Francetic, O. (2014a). Type II secretion system: A magic beanstalk or a

protein escalator. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research*, 1843(8), 1568–1577. <https://doi.org/10.1016/j.bbamcr.2013.12.020>

Nivaskumar, M., & Francetic, O. (2014b). Type II secretion system: A magic beanstalk or a

protein escalator. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research*, 1843(8), 1568–1577. <https://doi.org/10.1016/j.bbamcr.2013.12.020>

Nivaskumar, M., Santos-Moreno, J., Malosse, C., Nadeau, N., Chamot-Rooke, J., Tran Van

Nhieu, G., & Francetic, O. (2016). Pseudopilin residue E5 is essential for recruitment by the type 2 secretion system assembly platform. *Molecular Microbiology*, 101(6), 924–941. <https://doi.org/10.1111/mmi.13432>

Norman, T. M., Lord, N. D., Paulsson, J., & Losick, R. (2013). Memory and modularity in cell-fate decision making. *Nature*, 503(7477), 481–486.

<https://doi.org/10.1038/nature12804>

Nouwen, N., Ranson, N., Saibil, H., Wolpensinger, B., Engel, A., Ghazi, A., & Pugsley, A. P.

(1999). Secretin PulD: Association with pilot PulS, structure, and ion-conducting

- channel formation. *Proceedings of the National Academy of Sciences*, 96(14), 8173–8177. <https://doi.org/10.1073/pnas.96.14.8173>
- Owji, H., Nezafat, N., Negahdaripour, M., Hajiebrahimi, A., & Ghasemi, Y. (2018). A comprehensive review of signal peptides: Structure, roles, and applications. *European Journal of Cell Biology*, 97(6), 422–441. <https://doi.org/10.1016/j.ejcb.2018.06.003>
- Palmer, T., Finney, A. J., Saha, C. K., Atkinson, G. C., & Sargent, F. (2021). A holin/peptidoglycan hydrolase-dependent protein secretion system. *Molecular Microbiology*, 115(3), 345–355. <https://doi.org/10.1111/mmi.14599>
- Palmer, T., Sargent, F., & Berks, B. C. (2010). The Tat Protein Export Pathway. *EcoSal Plus*, 4(1). <https://doi.org/10.1128/ecosalplus.4.3.2>
- Pei, D., & Dalbey, R. E. (2022). Membrane Translocation of Folded Proteins. *The Journal of Biological Chemistry*, 102107. <https://doi.org/10.1016/j.jbc.2022.102107>
- Pellicic, V. (2008). Type IV pili: E pluribus unum? *Molecular Microbiology*, 68(4), 827–837. <https://doi.org/10.1111/j.1365-2958.2008.06197.x>
- Perlman, D., & Halvorson, H. O. (1983). A putative signal peptidase recognition site and sequence in eukaryotic and prokaryotic signal peptides. *Journal of Molecular Biology*, 167(2), 391–409. [https://doi.org/10.1016/S0022-2836\(83\)80341-6](https://doi.org/10.1016/S0022-2836(83)80341-6)
- Piepenbrink, K. H. (2019). DNA Uptake by Type IV Filaments. *Frontiers in Molecular Biosciences*, 6, 1. <https://doi.org/10.3389/fmolb.2019.00001>
- Pineau, C., Guschinskaya, N., Gonçalves, I. R., Ruaudel, F., Robert, X., Gouet, P., Ballut, L., & Shevchik, V. E. (2021). Structure–function analysis of pectate lyase Pel3 reveals essential facets of protein recognition by the bacterial type 2 secretion system.

*Journal of Biological Chemistry*, 296, 100305.

<https://doi.org/10.1016/j.jbc.2021.100305>

Pineau, C., Guschinskaya, N., Robert, X., Gouet, P., Ballut, L., & Shevchik, V. E. (2014).

Substrate recognition by the bacterial type II secretion system: More than a simple interaction. *Molecular Microbiology*, 94(1), 126–140.

<https://doi.org/10.1111/mmi.12744>

Possot, O. M., Letellier, L., & Pugsley, A. P. (1997). Energy requirement for pullulanase

secretion by the main terminal branch of the general secretory pathway. *Molecular Microbiology*, 24(3), 457–464. <https://doi.org/10.1046/j.1365-2958.1997.3451726.x>

Possot, O. M., Vignon, G., Bomchil, N., Ebel, F., & Pugsley, A. P. (2000). Multiple Interactions between Pullulanase Secretion Components Involved in Stabilization and Cytoplasmic Membrane Association of PulE. *Journal of Bacteriology*, 182(8), 2142–2152.

Possot, O., & Pugsley, A. P. (1994). Molecular characterization of PulE, a protein required for pullulanase secretion. *Molecular Microbiology*, 12(2), 287–299.

<https://doi.org/10.1111/j.1365-2958.1994.tb01017.x>

Pugsley, A. P. (1992). Translocation of a folded protein across the outer membrane in

*Escherichia coli*. *Proceedings of the National Academy of Sciences of the United States of America*, 89(24), 12058–12062. <https://doi.org/10.1073/pnas.89.24.12058>

Pugsley, A. P., Chapon, C., & Schwartz, M. (1986). Extracellular pullulanase of *Klebsiella pneumoniae* is a lipoprotein. *Journal of Bacteriology*, 166(3), 1083–1088.

Pugsley, A. P., & Dupuy, B. (1992). An enzyme with type IV prepilin peptidase activity is required to process components of the general extracellular protein secretion pathway of *Klebsiella oxytoca*. *Molecular Microbiology*, 6(6), 751–760.

<https://doi.org/10.1111/j.1365-2958.1992.tb01525.x>

- Pugsley, A. P., & Kornacker, M. G. (1991). Secretion of the cell surface lipoprotein pullulanase in *Escherichia coli*. Cooperation or competition between the specific secretion pathway and the lipoprotein sorting pathway. *Journal of Biological Chemistry*, 266(21), 13640–13645. [https://doi.org/10.1016/S0021-9258\(18\)92748-1](https://doi.org/10.1016/S0021-9258(18)92748-1)
- Pugsley, A. P., Poquet, I., & Kornacker, M. G. (1991). Two distinct steps in pullulanase secretion by *Escherichia coli* K12. *Molecular Microbiology*, 5(4), 865–873. <https://doi.org/10.1111/j.1365-2958.1991.tb00760.x>
- Putte, W. V., Vos, T. D., Broeck, W. V. D., Stahlberg, H., Kudryashev, M., & Savvides, S. N. (2018). *Inner-membrane GspF of the bacterial type II secretion system is a dimeric adaptor mediating pseudopilus biogenesis* (p. 435982). bioRxiv. <https://doi.org/10.1101/435982>
- Py, B., Loiseau, L., & Barras, F. (2001). An inner membrane platform in the type II secretion machinery of Gram-negative bacteria. *EMBO Reports*, 2(3), 244–248. <https://doi.org/10.1093/embo-reports/kve042>
- Rapisarda, C., Cherrak, Y., Kooger, R., Schmidt, V., Pellarin, R., Logger, L., Cascales, E., Pilhofer, M., Durand, E., & Fronzes, R. (2019). In situ and high-resolution cryo-EM structure of a bacterial type VI secretion system membrane complex. *The EMBO Journal*, 38(10), e100886. <https://doi.org/10.15252/emboj.2018100886>
- Reichow, S. L., Korotkov, K. V., Hol, W. G. J., & Gonen, T. (2010). Structure of the cholera toxin secretion channel in its closed state. *Nature Structural & Molecular Biology*, 17(10), Article 10. <https://doi.org/10.1038/nsmb.1910>
- Rivera-Calzada, A., Famelis, N., Llorca, O., & Geibel, S. (2021). Type VII secretion systems: Structure, functions and transport models. *Nature Reviews Microbiology*, 19(9), Article 9. <https://doi.org/10.1038/s41579-021-00560-5>

- Rondelet, A., & Condemine, G. (2013). Type II secretion: The substrates that won't go away. *Research in Microbiology*, 164(6), 556–561.  
<https://doi.org/10.1016/j.resmic.2013.03.005>
- Roux, N., Spagnolo, J., & de Bentzmann, S. (2012). Neglected but amazingly diverse type IVb pili. *Research in Microbiology*, 163(9), 659–673.  
<https://doi.org/10.1016/j.resmic.2012.10.015>
- Russell, A. B., Peterson, S. B., & Mougous, J. D. (2014). Type VI secretion system effectors: Poisons with a purpose. *Nature Reviews. Microbiology*, 12(2), 137–148.  
<https://doi.org/10.1038/nrmicro3185>
- Ryoo, D., Rydmark, M. O., Pang, Y. T., Lundquist, K. P., Linke, D., & Gumbart, J. C. (2020). BamA is required for autotransporter secretion. *Biochimica Et Biophysica Acta. General Subjects*, 1864(7), 129581. <https://doi.org/10.1016/j.bbagen.2020.129581>
- Sandkvist, M., Bagdasarian, M., Howard, S. P., & DiRita, V. J. (1995). Interaction between the autokinase EpsE and EpsL in the cytoplasmic membrane is required for extracellular secretion in *Vibrio cholerae*. *The EMBO Journal*, 14(8), 1664–1673.
- Sankaran, K., & Wu, H. C. (1994). Lipid modification of bacterial prolipoprotein. Transfer of diacylglycerol moiety from phosphatidylglycerol. *Journal of Biological Chemistry*, 269(31), 19701–19706. [https://doi.org/10.1016/S0021-9258\(17\)32077-X](https://doi.org/10.1016/S0021-9258(17)32077-X)
- Santos-Moreno, J., East, A., Guilvout, I., Nadeau, N., Bond, P. J., Tran Van Nhieu, G., & 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(11), 1746–1765.  
<https://doi.org/10.1016/j.jmb.2017.04.005>

- Sauvonnet, N., Poquet, I., & Pugsley, A. P. (1995). Extracellular secretion of pullulanase is unaffected by minor sequence changes but is usually prevented by adding reporter proteins to its N- or C-terminal end. *Journal of Bacteriology*, 177(18), 5238–5246.
- Sauvonnet, N., & Pugsley, A. P. (1996). Identification of two regions of *Klebsiella oxytoca* pullulanase that together are capable of promoting  $\beta$ -lactamase secretion by the general secretory pathway. *Molecular Microbiology*, 22(1), 1–7.  
<https://doi.org/10.1111/j.1365-2958.1996.tb02650.x>
- Sauvonnet, N., & Pugsley, A. P. (1998). The requirement for DsbA in pullulanase secretion is independent of disulphide bond formation in the enzyme. *Molecular Microbiology*, 27(3), 661–667. <https://doi.org/10.1046/j.1365-2958.1998.00722.x>
- Sauvonnet, N., Vignon, G., Pugsley, A. P., & Gounon, P. (2000). Pilus formation and protein secretion by the same machinery in *Escherichia coli*. *The EMBO Journal*, 19(10), 2221–2228. <https://doi.org/10.1093/emboj/19.10.2221>
- Schägger, H., & von Jagow, G. (1987). Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. *Analytical Biochemistry*, 166(2), 368–379. [https://doi.org/10.1016/0003-2697\(87\)90587-2](https://doi.org/10.1016/0003-2697(87)90587-2)
- Sikora, A. E., Zielke, R. A., Lawrence, D. A., Andrews, P. C., & Sandkvist, M. (2011). Proteomic Analysis of the *Vibrio cholerae* Type II Secretome Reveals New Proteins, Including Three Related Serine Proteases. *The Journal of Biological Chemistry*, 286(19), 16555–16566. <https://doi.org/10.1074/jbc.M110.211078>
- Silhavy, T. J., Kahne, D., & Walker, S. (2010). The Bacterial Cell Envelope. *Cold Spring Harbor Perspectives in Biology*, 2(5), a000414. <https://doi.org/10.1101/cshperspect.a000414>

- Singh, N., & Wagner, S. (2019). Investigating the assembly of the bacterial type III secretion system injectisome by in vivo photocrosslinking. *International Journal of Medical Microbiology*, 309(6), 151331. <https://doi.org/10.1016/j.ijmm.2019.151331>
- Smith, D. R., & Chapman, M. R. (2010). Economical evolution: Microbes reduce the synthetic cost of extracellular proteins. *MBio*, 1(3), e00131-10. <https://doi.org/10.1128/mBio.00131-10>
- Steinberg, R., Knüpffer, L., Origi, A., Asti, R., & Koch, H.-G. (2018). Co-translational protein targeting in bacteria. *FEMS Microbiology Letters*, 365(11). <https://doi.org/10.1093/femsle/fny095>
- Stewart, M. K., & Cookson, B. T. (2014). Mutually repressing repressor functions and multi-layered cellular heterogeneity regulate the bistable Salmonella fliC census. *Molecular Microbiology*, 94(6), 1272–1284. <https://doi.org/10.1111/mmi.12828>
- Strom, M. S., Nunn, D. N., & Lory, S. (1993). A single bifunctional enzyme, PilD, catalyzes cleavage and N-methylation of proteins belonging to the type IV pilin family. *Proceedings of the National Academy of Sciences*, 90(6), 2404–2408. <https://doi.org/10.1073/pnas.90.6.2404>
- Strozen, T. G., Stanley, H., Gu, Y., Boyd, J., Bagdasarian, M., Sandkvist, M., & Howard, S. P. (2011). Involvement of the GspAB Complex in Assembly of the Type II Secretion System Secretin of Aeromonas and Vibrio Species. *Journal of Bacteriology*, 193(9), 2322–2331. <https://doi.org/10.1128/JB.01413-10>
- Sutherland, B. W., Toews, J., & Kast, J. (2008). Utility of formaldehyde cross-linking and mass spectrometry in the study of protein–protein interactions. *Journal of Mass Spectrometry*, 43(6), 699–715. <https://doi.org/10.1002/jms.1415>

- Takhar, H. K., Kemp, K., Kim, M., Howell, P. L., & Burrows, L. L. (2013). The platform protein is essential for type IV pilus biogenesis. *The Journal of Biological Chemistry*, 288(14), 9721–9728. <https://doi.org/10.1074/jbc.M113.453506>
- Terada, M., Kuroda, T., Matsuyama, S., & Tokuda, H. (2001). Lipoprotein Sorting Signals Evaluated as the LolA-dependent Release of Lipoproteins from the Cytoplasmic Membrane of *Escherichia coli* \*. *Journal of Biological Chemistry*, 276(50), 47690–47694. <https://doi.org/10.1074/jbc.M109307200>
- Thanabalu, T., Koronakis, E., Hughes, C., & Koronakis, V. (1998). Substrate-induced assembly of a contiguous channel for protein export from *E. coli*: Reversible bridging of an inner-membrane translocase to an outer membrane exit pore. *The EMBO Journal*, 17(22), 6487–6496. <https://doi.org/10.1093/emboj/17.22.6487>
- Thomassin, J.-L., Santos Moreno, J., Guilvout, I., Tran Van Nhieu, G., & Francetic, O. (2017). The trans-envelope architecture and function of the type 2 secretion system: New insights raising new questions. *Molecular Microbiology*, 105(2), 211–226. <https://doi.org/10.1111/mmi.13704>
- Tomich, M., Planet, P. J., & Figurski, D. H. (2007). The tad locus: Postcards from the widespread colonization island. *Nature Reviews. Microbiology*, 5(5), 363–375. <https://doi.org/10.1038/nrmicro1636>
- Tsai, C.-L., & Tainer, J. A. (2016). The ATPase Motor Turns for Type IV Pilus Assembly. *Structure*, 24(11), 1857–1859. <https://doi.org/10.1016/j.str.2016.10.002>
- Tsirigotaki, A., De Geyter, J., Šoštarić, N., Economou, A., & Karamanou, S. (2017). Protein export through the bacterial Sec pathway. *Nature Reviews Microbiology*, 15(1), Article 1. <https://doi.org/10.1038/nrmicro.2016.161>

- Unnikrishnan, M., Constantinidou, C., Palmer, T., & Pallen, M. J. (2017). The Enigmatic Esx Proteins: Looking Beyond Mycobacteria. *Trends in Microbiology*, 25(3), 192–204.  
<https://doi.org/10.1016/j.tim.2016.11.004>
- van der Laan, M., Bechtluft, P., Kol, S., Nouwen, N., & Driessen, A. J. M. (2004). F1F0 ATP synthase subunit c is a substrate of the novel YidC pathway for membrane protein biogenesis. *Journal of Cell Biology*, 165(2), 213–222.  
<https://doi.org/10.1083/jcb.200402100>
- Van Gerven, N., Waksman, G., & Remaut, H. (2011). Chapter 2 - Pili and Flagella: Biology, Structure, and Biotechnological Applications. In S. Howorka (Ed.), *Progress in Molecular Biology and Translational Science* (Vol. 103, pp. 21–72). Academic Press.  
<https://doi.org/10.1016/B978-0-12-415906-8.00005-4>
- van Ulsen, P., Rahman, S. ur, Jong, W. S. P., Daleke-Schermerhorn, M. H., & Luirink, J. (2014). Type V secretion: From biogenesis to biotechnology. *Biochimica Et Biophysica Acta*, 1843(8), 1592–1611. <https://doi.org/10.1016/j.bbamcr.2013.11.006>
- van Ulsen, P., Zinner, K. M., Jong, W. S. P., & Luirink, J. (2018). On display: Autotransporter secretion and application. *FEMS Microbiology Letters*, 365(18).  
<https://doi.org/10.1093/femsle/fny165>
- Vanderlinde, E. M., Strozen, T. G., Hernández, S. B., Cava, F., & Howard, S. P. (2017). Alterations in Peptidoglycan Cross-Linking Suppress the Secretin Assembly Defect Caused by Mutation of GspA in the Type II Secretion System. *Journal of Bacteriology*, 199(8), e00617-16. <https://doi.org/10.1128/JB.00617-16>
- Vanderlinde, E. M., Zhong, S., Li, G., Martynowski, D., Grochulski, P., & Howard, S. P. (2014). Assembly of the type two secretion system in *Aeromonas hydrophila* involves direct interaction between the periplasmic domains of the assembly factor ExeB and the

- secretin ExeD. *PLoS One*, 9(7), e102038.  
<https://doi.org/10.1371/journal.pone.0102038>
- Vasilescu, J., Guo, X., & Kast, J. (2004). Identification of protein-protein interactions using in vivo cross-linking and mass spectrometry. *PROTEOMICS*, 4(12), 3845–3854.  
<https://doi.org/10.1002/pmic.200400856>
- Vignon, G., Köhler, R., Larquet, E., Giroux, S., Prévost, M.-C., Roux, P., & Pugsley, A. P. (2003). Type IV-like pili formed by the type II secretin: Specificity, composition, bundling, polar localization, and surface presentation of peptides. *Journal of Bacteriology*, 185(11), 3416–3428. <https://doi.org/10.1128/JB.185.11.3416-3428.2003>
- Voulhoux, R., Ball, G., Ize, B., Vasil, M. L., Lazdunski, A., Wu, L. F., & Filloux, A. (2001). Involvement of the twin-arginine translocation system in protein secretion via the type II pathway. *The EMBO Journal*, 20(23), 6735–6741.  
<https://doi.org/10.1093/emboj/20.23.6735>
- Voulhoux, R., Taupiac, M.-P., Czjzek, M., Beaumelle, B., & Filloux, A. (2000). Influence of Deletions within Domain II of Exotoxin A on Its Extracellular Secretion from *Pseudomonas aeruginosa*. *Journal of Bacteriology*, 182(14), 4051–4058.  
<https://doi.org/10.1128/JB.182.14.4051-4058.2000>
- Wu, R., Stephenson, R., Gichaba, A., & Noinaj, N. (2020). The big BAM theory: An open and closed case? *Biochimica et Biophysica Acta (BBA) - Biomembranes*, 1862(1), 183062.  
<https://doi.org/10.1016/j.bbamem.2019.183062>
- Yin, M., Yan, Z., & Li, X. (2018). Structural insight into the assembly of the type II secretion system pilotin-secretin complex from enterotoxigenic *Escherichia coli*. *Nature Microbiology*, 3(5), 581–587. <https://doi.org/10.1038/s41564-018-0148-0>

- Zhang, S., Gu, S., Rycroft, P., Ruaudel, F., Delolme, F., Robert, X., Ballut, L., Pickersgill, R. W., & Shevchik, V. E. (2022). Scaffolding Protein GspB/OutB Facilitates Assembly of the *Dickeya dadantii* Type 2 Secretion System by Anchoring the Outer Membrane Secretin Pore to the Inner Membrane and to the Peptidoglycan Cell Wall. *MBio*, 13(3), e00253-22. <https://doi.org/10.1128/mbio.00253-22>
- Zückert, W. R. (2014). Secretion of bacterial lipoproteins: Through the cytoplasmic membrane, the periplasm and beyond. *Biochimica Et Biophysica Acta*, 1843(8), 1509–1516. <https://doi.org/10.1016/j.bbamcr.2014.04.022>

# Résumé substantiel en français

## Introduction

La plupart des bactéries sont des organismes unicellulaires qui se trouvent dans un large éventail d'habitats. Les cellules bactériennes sont délimitées par leurs enveloppes, structures multicouches qui protègent la cellule et permettent l'échange de nutriments et de métabolites avec l'environnement. Pour s'adapter à leur environnement, les bactéries ont développé différents systèmes de sécrétion de protéines dans leurs enveloppes cellulaires. Les bactéries à Gram-négatif font appel à plusieurs types de machineries complexes afin de sécréter des protéines spécifiques à travers leur enveloppe à double membrane. Les protéines sécrétées jouent des rôles importants dans l'acquisition des nutriments, l'adhésion, motilité, virulence, respiration ou encore la compétition inter-bactérienne. En général, les protéines sécrétées permettent aux bactéries de mieux s'adapter à leur environnement. À ce jour, en excluant les voies Sec et Tat, 11 types de systèmes de sécrétion de protéines ont été identifiés chez les bactéries.

Parmi eux, le système de sécrétion de type II (SST2) est l'un des plus étudiés et des mieux caractérisés depuis sa découverte chez *Klebsiella* dans les années 1980. Selon les analyses de séquences du génome, le SST2 est largement spécifique aux protéobactéries et existe dans la majorité des genres des classes de protéobactéries Alpha, Beta, Gamma et Delta.

Le sujet d'étude dans cette thèse est l'SST2 de *Klebsiella oxytoca*, le premier système de cette famille mis en évidence. Ce système, dédié à la sécrétion de pullulanase, est un complexe de 15 protéines différentes oligomériques pour la plupart, qui traverse l'enveloppe de la bactérie. Le système Pul est codé par 15 gènes organisés en un large cluster composé de deux opérons divergents : *pulAB*, *pulCDEFGHIJKLMNO*, et d'un gène monocystronique *pulS*. Les deux opérons sont contrôlés par des promoteurs divergents activés par MalT et par la protéine activatrice catabolique générale (CAP) liée à l'AMPc, tandis que l'expression de *pulS* est constitutive. Le fait que le substrat PulA soit codé par le *pulAB* dans le cluster de gènes SST2 a facilité la reconstitution du système chez *Escherichia coli* et a laissé supposer que PulA soit le seul substrat du système. Cependant, il est toujours possible que d'autres protéines, codées ailleurs dans le génome, soient sécrétées par le système Pul.

La machinerie Pul est composée de 4 sous-complexes: ① l'ATPase (PulE). En tant que source d'énergie de cette machinerie, PulE forme un hexamère qui permet hydrolyse de l'ATP,

situé dans le cytoplasme à la base du système ; cette énergie est nécessaire pour l'assemblage du pseudopilus et sécrétion de substrat ; ② la plate-forme d'assemblage de la membrane interne (AP) un complexe composé de PulF, PulL, PulM et PulN. Avec PulE, ce complexe favorise la polymérisation du pseudopilus dans le périplasma. ③ la sécrétine, un pentadécamère de PulD formant un canal dans la membrane externe. Le complexe AP est lié *via* PulC à la sécrétine dans la membranaire externe. Chaque sous-unité de PulD est liée à la lipoprotéine pilotine PulS qui permet l'adressage de PulD vers la membrane externe et facilite son insertion ; ④ le pseudopilus dans le périplasma. L'AP permet l'assemblage du pseudopilus, composé des pseudopilines mineures PulH, PulI, PulJ et PulK au sommet, et de la pseudopiline majeure PulG formant un long homopolymère. Selon le modèle actuel, le pseudopilus permettrait la sécrétion en interagissant avec le substrat. La prépiline peptidase PulO est localisée dans la membrane interne et permet la maturation des pseudopilines, par clivage de leur peptide signal N-terminal afin de permettre leur extraction de la membrane lors de l'assemblage. PulO modifie également l'extrémité N-terminale des pseudopilines matures en ajoutant un groupement méthyl, ce qui facilite aussi leur extraction de la membrane.

La sécrétion se déroule en deux étapes : dans un premier temps, les substrats du SST2 sont exportés vers le périplasma *via* les systèmes Sec ou Tat, pour être pris en charge dans le périplasma et sécrétés dans leur état natif, entièrement replié. Dans certains cas, comme la toxine cholérique dans *Vibrio*, les substrats sont sécrétés sous forme de complexes oligomériques. Chez différentes espèces bactériennes les substrats varient en taille, forme, structure quaternaire et fonction. De plus, le nombre de substrats sécrétés chez les bactéries varie d'un, comme chez *K. oxytoca*, à plus de dix chez *Pseudomonas aeruginosa*, *Vibrio cholerae* ou *Legionella pneumophila*. Il n'y a pas similitudes de séquence ou de structure parmi les substrats chez une même espèce ou entre espèces. La sécrétion est strictement spécifique et repose probablement sur un ou plusieurs motifs structuraux. Malgré un grand nombre d'études structurales sur différents modèles bactériens, la nature de ces motifs et leur interaction avec les composants du SST2 sont mal comprises. Certaines études suggèrent la  $\beta$ -complémentation comme mécanisme de reconnaissance en vue de la présence des feuilletts  $\beta$  associés aux substrats d'SST2. Ci-dessous sont présentées quelques études récentes qui traitent de la nature moléculaire des signaux de sécrétion. Depuis la découverte du SST2, la question du signal de sécrétion a été soulevée et abordée dans de nombreuses études. Chez les bactéries qui sécrètent plusieurs substrats, l'absence d'un motif de séquence commun suggère que le signal se trouve dans leur structure tridimensionnelle. De plus, plusieurs études ont montré que les substrats

d'SST2 sont repliés et stabilisés par des ponts disulfures avant la sécrétion. Cela implique que la compréhension de la nature du signal de sécrétion nécessite des informations structurales sur les protéines sécrétées.

Le *K. oxytoca* SST2, notre modèle expérimental, est dédié à la sécrétion d'une lipoprotéine de 116 kDa, la pullulanase (PulA), ancrée à la surface de la bactérie. PulA est une enzyme qui dégrade les polymères branchés de maltotriose comme le pullulan, permettant la croissance bactérienne sur ces polymères comme seule source de carbone. L'identification et le clonage du gène *pulA* de *K. oxytoca* dans *Escherichia coli* K-12 ont marqué le début d'un long et passionnant processus qui a abouti à la découverte du système de sécrétion de type II. Afin d'identifier les facteurs de la sécrétion de PulA, d'Enfert et Pugsley ont utilisé une approche de fusion de gènes. Une série de fusions entre le gène de la pullulanase (*pulA*) et les gènes de la  $\beta$ -galactosidase cytoplasmique (*lacZ*) ou de la phosphatase alcaline périplasmique (*phoA*) ont été créées par la mutagenèse par des transposons. Ces fusions ont été exprimées chez *E. coli* K-12 afin de déterminer la localisation subcellulaire des protéines hybrides. Le résidu cystéine N-terminal du polypeptide mature est acylé et la lipoprotéine PulA utilise la voie Sec pour traverser la membrane interne. Sauvonnnet *et al.* ont construit plusieurs protéines hybrides PulA- $\beta$ -lactamase (BlaM) et ont montré que la PulA dépourvue des 256 derniers acides aminés était sécrétée. Sur cette base, ils ont conçu un crible phénotypique, permettant à identifier, suite à des délétions aléatoires, deux régions, nommées A et B, qui sont ensemble essentielles à la sécrétion de PulA. Les régions isolées A et B, fusionnées à une séquence signal N-terminale, étaient également suffisantes pour permettre la sécrétion de la  $\beta$ -lactamase, ainsi que d'un domaine de la protéine CelZ de *Dickeya dadantii*.

Francetić et Pugsley ont construit un variant non acylée de PulA (PulA<sub>NA</sub>), en remplaçant son peptide signal natif par le peptide signal de la pectate lyase PelB de *D. dadantii*. Cette construction a facilité l'étude de la sécrétion de PulA par son immuno-détection dans les fractions cellulaires et de surnageant de culture grâce à la sécrétion efficace de PulA<sub>NA</sub> dans le milieu. Des études du signal de sécrétion de PulA<sub>NA</sub> ont été réalisées par des approches d'insertion de peptides, délétion et fusion de protéines. Les résultats ont montré que la délétion de la région A déstabilisait PulA et que la délétion de la région B avait un faible impact sur la sécrétion de PulA<sub>NA</sub>. Une troisième région C importante pour la sécrétion a été identifiée, entre les acides aminés 234 et 324 de PulA<sub>NA</sub>. Quatre insertions de linker indépendantes dans cette région ont aboli la sécrétion, suggérant que la région C contiendrait une partie du signal de sécrétion. De

plus, la délétion de la région C avait abolit la sécrétion. Pris ensemble, ces résultats suggèrent qu'au moins trois régions de PulA contiennent les signaux de sécrétion.

La structure PulA<sub>NA</sub> de *K. oxytoca* a été déterminée par la cristallographie aux rayons X (East et al, 2016). Les simulations de dynamique moléculaire atomistique (MD) de PulA dans une membrane modèle de 1-palmitoyl-2-oléoyl-sn-glycéro3-phosphoéthanolamine (POPE) et des études biochimiques et fonctionnelles ont apporté une vue moléculaire sans précédent de PulA. PulA consiste en un linker N-terminal non structuré, les domaines N1, N2 et N3 et A (comprenant le sous-domaine Ins) et le domaine C. Deux motifs de sécrétion putatifs, N1 et Ins, ont été identifiés dans PulA grâce à une combinaison d'approches structurales, biochimiques et fonctionnelles.

East et al. ont montré que la délétion du domaine Ins abolit la sécrétion de PulA dans des conditions physiologiques lorsque le système *pul* est exprimé sur un plasmide en faible nombre de copies, et que la délétion du domaine N1 presque abolit sécrétion et conduit à la dégradation de PulA. Le domaine N1 correspondrait à la région A identifiée précédemment. Comme le révèlent les valeurs élevées du facteur B cristallographique, ces domaines PulA ont également montré une flexibilité conformationnelle importante. Cela suggère que ces domaines pourraient être impliquées dans la sécrétion de PulA.

Les études du signal de sécrétion soulèvent également la question de son mécanisme. Les protéines sécrétées (exoprotéines) n'ont pas d'autres caractéristiques distinctives que des peptides signal qui les dirigent vers les voies Sec ou Tat pour le transport dans le périplasma. Le modèle évident est que le signal interagit avec un ou plusieurs composants du SST2. De nombreuses études se sont penchées sur cette question.

Au moins deux modèles concernant le rôle du pseudopilus dans la sécrétion ont été proposés : le modèle de piston et le modèle de vis d'Archimède. Dans le modèle de piston, les pseudopili s'étendent et se rétractent, et le substrat PulA interagit avec des pseudopilines mineures (PulH, I, J, K) au sommet du filament. Le substrat est expulsé à travers la sécrétine dans la membrane externe par ces mouvements de piston. Dans le modèle de vis d'Archimède, les pseudopili s'assemblent et s'étendent continuellement à vers l'extérieur mais la longueur du filament est constante grâce à son désassemblage et dégradation. Ce modèle prédit que chez *K. oxytoca*, PulA interagit avec la pseudopiline majeure PulG dans le corps du pseudopilus.

Pour tester lequel de ces modèles est le plus proche de la réalité, ou s'il existe un troisième modèle qui combine les caractéristiques des deux, et pour compléter la cartographie des

interactions du substrat avec le système Pul, nous avons conçu et réalisé différentes expériences décrites ci-dessous.

## Objectifs

### Identification des partenaires d'interaction PulA dans le SST2

Dans les études antérieures de la sécrétion de PulA, East *et al.* ont identifié deux régions de pullulanase qui pourraient contenir le signal de sécrétion, désignées domaine N1 et le sous-domaine Ins. Ces deux régions semblaient nécessaires à la sécrétion de PulA, mais pourraient aussi jouer des rôles différents. Notre objectif principal était de comprendre la fonction de ces signaux de sécrétion putatifs en identifiant les composants SST2 qui interagissent avec ces régions.

Pour étudier les interactions de PulA avec les protéines du SST2, nous avons prévu d'utiliser plusieurs approches : l'analyse par le système de double hybride bactérien (BACTH), le pontage chimique et co-purification, et le pontage par des rayons UV *in vivo* avec des acides aminés non naturels.

L'analyse BACTH est une méthode puissante permettant d'étudier les interactions des protéines cytoplasmiques ou membranaires. Les protéines sont fusionnées aux fragments T18 et T25 du domaine catalytique de la toxine CyaA, et, si elles interagissent entre elles, rapprochent dans l'espace les fragments T18 et T25 et reconstituent leur activité adénylyl cyclase dans le cytoplasme. L'interaction des domaines périplasmiques ou membranaires peut donc entraîner la production d'AMPC dans le cytoplasme et activer la transcription de gènes rapporteurs dépendant de l'AMPC, tels que *malE* ou *lacZ*. L'avantage de cette approche est que l'on peut sonder les interactions des protéines entières dans leur environnement natif, le cytoplasme ou la membrane. La limitation est que les protéines périplasmiques solubles et les protéines de la membrane externe ou sécrétées ne peuvent pas être étudiées dans ces compartiments. Elles peuvent être attachées à un segment transmembranaire ou produites dans le cytoplasme, bien que dans ce dernier cas, il existe un qu'elles ne soient pas correctement repliées en raison de conditions physico-chimiques, ou du contenu protéique et macromoléculaire différent entre le cytoplasme et le périplasma.

L'approche de photo-pontage site-spécifique *in vivo* a le potentiel de capturer les interactions protéine-protéine transitoires en liant les résidus en contact étroit avec un résidu d'acide aminé non naturel p-benzoyl-L-phénylalanine (pBpa) incorporé dans la protéine d'intérêt. Le codage génétique et l'incorporation traductionnelle de l'acide aminé synthétique

*pBpa* à une position souhaitée d'une protéine d'intérêt constituent la base du photo-pontage *in vivo*. Un codon stop TAG (ambre) reconnu par un ARNt suppresseur couplé au *pBpa* initie le codage. Une synthétase d'ARNt de *Methanococcus jannaschii* mutante est responsable de la charge de l'ARNt avec du *pBpa*. Le *pBpa* réagit au rayonnement UV à 365 nm en formant une liaison covalente avec la chaîne peptidique des protéines adjacentes. Les bactéries comme *E. coli* peuvent résister à la suppression des codons stop TAG.

Pour mettre en œuvre l'approche de suppresseur ambre par des acides aminés non naturels dans le SST2, nous avons procédé en deux étapes. Au départ, j'ai utilisé PulG comme petite protéine modèle pour mettre en œuvre la méthode. La pseudopiline majeure PulG est essentielle à la sécrétion de PulA et nous avons émis l'hypothèse que cette approche pourrait nous permettre de tester si elle interagit directement avec PulA. Une telle interaction conforterait le modèle de sécrétion par le mécanisme de vis d'Archimède.

Dans un second temps, nous avons conçu des stratégies pour introduire *pBpa* dans les domaines PulA N1 et Ins de PulA. Les résultats non publiés de l'analyse BACTH et des études de co-purification au laboratoire ont identifié plusieurs composants qui pourraient interagir avec PulA. Ces protéines résident soit dans la membrane interne (PulM, PulC), soit elles forment le pseudopilus (PulG). Pour capturer les interactions potentiellement transitoires de PulA lors de son transport à travers le complexe SST2, nous avons prévu d'utiliser le pontage UV site-spécifique. Des codons TAG seraient introduits à des positions spécifiques des domaines N1 et Ins de PulA. Après avoir incorporé l'acide aminé non naturel *pBpa* dans ces positions, nous avons prévu de sélectionner des dérivés PulA-*pBpa* qui sont efficacement sécrétés. Ces bactéries seraient irradiées aux UV pour capturer les dérivés de PulA pontés à leurs partenaires pendant le transport. L'analyse biochimique et l'immunodétection seraient utilisées pour identifier ces partenaires.

L'approche de pontage de *pBpa* aux UV a ses limites. Elle est très laborieuse et risquée, car le choix du site d'insertion du *pBpa* détermine si le partenaire d'interaction sera détecté. Pour surmonter cette limitation, nous avons également envisagé l'utilisation d'un pontage chimique sur les cellules entières suivie d'une co-purification de PulA marqué avec une étiquette d'affinité - Strep ou His<sub>6</sub>.

## Caractérisation structurale et fonctionnelle des composants PulL et PulM de la plateforme d'assemblage

Vers la fin de ma thèse, j'ai participé à une étude collaborative avec le groupe de N. Izadi-Pruneyre pour décrypter la structure et le fonctionnement de deux composants essentiels de la plateforme d'assemblage, PulL et PulM. Le groupe de Nadia a utilisé la cristallographie aux rayons X et la RMN en solution pour déterminer les structures des domaines C-terminaux de PulL (PulL<sub>CTD</sub>) et PulM (PulM<sub>CTD</sub>). Pour la première fois, ils ont pu déterminer la structure d'un hétérodimère formé par ces composants du SST2. Pour tester si l'interaction entre le CTD de PulL et PulM en solution correspond à leur interaction *in vivo* dans les cellules bactériennes, j'ai utilisé le pontage disulfure et le BACTH pour étudier les interactions des protéines entières localisées dans la membrane. Ensemble, ces études nous ont permis de proposer un modèle du complexe PulL – PulM.

J'ai également participé à l'analyse fonctionnelle des variants mutants de PulL et PulM et de leurs interactions. Enfin, j'ai déterminé le rapport des niveaux de PulL et PulM dans les bactéries, ce qui était important pour expliquer la dynamique de leur interaction et pour proposer un modèle fonctionnel. Le manuscrit décrivant ces résultats est en cours de révision dans la revue Structure.

Nous avons étendu l'étude de PulL et PulM pour tester si d'autres régions de ces protéines contribuent à leur interaction. Nous avons délété des régions périplasmiques et cytoplasmiques de PulL et PulM et analysé leurs interactions par le BACTH et études fonctionnelles. Les résultats ont montré que les hélices périplasmiques de PulL et PulM formant un motif coiled coil (« bobine enroulée ») jouent un rôle très important dans leur interaction. Le peptide cytoplasmique de PulM était également crucial pour la stabilité de PulL et PulM, suggérant une interaction étroite des deux protéines dans le cytoplasme. Ce même peptide cytoplasmique de PulM est également crucial pour son interaction avec la pseudopiline majeure PulG. Ces résultats sont en accord avec un modèle d'interactions dynamiques de PulM avec PulG et PulL impliquant ce peptide N-terminal.

## Résultats

### Interactions PulA avec les composants d'SST2

Comme souligné dans l'Introduction, le SST2 joue un rôle important dans la survie et la croissance des bactéries sur des polymères larges ou complexes. L'un des principaux objectifs de ce projet était de caractériser le mécanisme de spécificité de substrat dans le Pul SST2 en identifiant les partenaires d'interaction PulA lors de la sécrétion, et leurs contacts précis lors de cette interaction.

Pour atteindre ces objectifs, nous avons voulu mettre en place l'approche de pontage *in vivo* après incorporation de l'acide aminé non naturel *pBpa* dans PulA et irradiation UV. L'axe principal de mon travail visait à mettre en place cette approche dans le système Pul reconstitué chez *E. coli*.

### Incorporation de *pBpa* et photo-pontage site-spécifique *in vivo* de PulG

Pour mettre en œuvre l'approche de pontage par le *pBPA*, au départ j'ai utilisé PulG comme petite protéine modèle. La pseudopiline majeure PulG est la protéine la plus abondante dans le SST2 et essentielle à la sécrétion de PulA. Nous avons émis l'hypothèse que cette approche pourrait nous permettre de tester si elle interagit directement avec PulA.

Le modèle de vis d'Archimède de sécrétion de type II prédit une interaction de PulA avec PulG. Notre collaborateur Benjamin Bardiaux (Unité de Bioinformatique Structurale, Institut Pasteur) a modélisé *in silico* l'interaction de PulA et du pseudopilus pour tester cette possibilité. Il a prédit que le domaine PulA Ins contenant un signal de sécrétion putatif s'insère dans le sillon de surface du pseudopilus. Plusieurs résidus sur le domaine PulG et Ins de PulA ont été identifiés comme impliqués dans ces contacts spécifiques. Sur la base de ce modèle, nous nous sommes concentrés sur les résidus de surface PulG non seulement dans ces résidus prédits de contact, mais également sur les sites proches. Nous avons introduit des codons Amber à ces positions spécifiques par mutagenèse dirigée. Les résidus PulG suivants ont été remplacés : E29, D32, R33, Q65, A70, S72, A73, E74, R78, N79, E82, Q91, D97, L101, Q108, F112, N123.

Nous avons sélectionné les souches d'*E. coli* qui ne possèdent pas des supprimeurs endogènes des codons Amber tel que BL21, BW25113 ou MG1655, et nous avons introduit un épisode contenant le gène *lacI* afin de limiter l'expression du gène *pulG*. Le locus *K. oxytoca pul* contient 16 gènes codant pour le SST2, qui ont été clonés sur un plasmide pCHAP8185 à nombre de copies modéré. Lorsqu'il est introduit dans *E. coli*, ce plasmide produit un SST2

fonctionnel qui sécrète PulA à la surface de la cellule. Il est difficile d'introduire des mutations Amber dans ce grand plasmide par mutagenèse par PCR, mais de telles mutations peuvent être facilement introduites dans des plasmides portant des gènes *pul* individuels. Par conséquent, nous avons combiné des dérivés de pCHAP8185 avec des mutations non polaires dans des gènes *pul* individuels avec des plasmides contenant les gènes manquants clonés dans les vecteurs compatibles, comme pSU18 ou pSU19. Dans le cas de *pulG*, le plasmide pCHAP8184 (un dérivé du plasmide pCHAP8185 portant l'origine de répllication ColEI et le marqueur Ap<sup>R</sup>) portant une délétion non polaire de *pulG* a été combiné avec pCHAP8658 portant le gène *pulG* sous le contrôle du promoteur *placZ*.

L'incorporation de *pBpa* à la position souhaitée nécessite un troisième plasmide portant les ARNt supprimeurs et son amino-acyl-ARNt synthétase, tous deux adaptés pour charger *pBpa* sur les ARNt portant l'anticodon ambre. Le plasmide contenant ces gènes a été obtenu auprès de l'équipe de Shultz et nous l'avons rendu compatible avec notre système. Nous avons modifié son gène de résistance et son gène *ori*, et construit un dérivé de ce plasmide nommé pMS5005 codant pour : une amino-acyl-ARNt synthétase sous contrôle du promoteur arabinose ; ARNt supprimeur de TAG ; gène pCDF *ori*; et le marqueur de résistance à la spectinomycine (Sp<sup>R</sup>). Pour accomplir notre test de photo-pontage *in vivo*, nous avons transformé *E. coli* avec ces 3 plasmides.

Afin de tester nous avons sélectionné plusieurs clones *pulGAm* montrant une bonne efficacité d'incorporation de *pBpa* pour l'irradiation aux UV afin de capturer les dérivés de PulG pontés pendant la sécrétion. Nous avons testé chaque mutant *pulG* Amber que nous avons produit et nous avons détecté deux nouvelles bandes dérivées des variants PulG S72-*pBpa* et E82-*pBpa*. Les tailles de ces bandes d'environ 30 kD et 40 kD, en vue du poids moléculaire de de PulG de 14607 Da, suggèrent que la taille du partenaire ponté serait entre 15 et 26 kDa. Les candidate pour ces partenaires seraient PulM et PulH ou PulG elle-même. Pour vérifier si le(s) partenaire(s) sont spécifique(s) du SST2, nous avons réalisé les mêmes expériences sans les gènes *pul* dans un système à 2 plasmides, constitué des variants PulG S72-*pBpa* ou E82-*pBpa* et pMS5005. Les mêmes bandes aux mêmes positions ont été détectées par des anticorps anti-PulG. Ainsi, nous avons pu confirmer que ces bandes représentent probablement des dimères PulG.

Néanmoins, ces résultats préliminaires ont montré que le système à 3 plasmides permet une bonne incorporation de *pBpa* et que dans nos conditions expérimentales et permet le pontage aux UVs. Parallèlement, nous nous sommes rendu compte que la validation du ou des

partenaires par les anticorps était limitée par la sensibilité de détection et la capacité des anticorps à reconnaître les protéines pontées. Nous avons conclu que la sensibilité pouvait être améliorée en co-purifiant PulG complexé *via* des variantes étiquetées de PulG et en utilisant la spectrométrie de masse (MS) pour identifier le partenaire ponté.

#### Construction de dérivés PulA-*pBpa* et tests de pontage

Dans l'étape suivante, nous avons adapté ces procédures et conditions expérimentales pour le de PulA<sub>NA</sub> et domaines N1 et Ins. Nous avons choisi des résidus sur la surface exposée du domaine PulA N1 et les résidus du domaine Ins prédits par le modèle *in silico* comme étant impliqués dans l'interaction avec le pseudopilus. Nous avons introduit le codon Amber dans ces positions par mutagenèse dirigée.

Nous avons généré un système à 3 plasmides avec le dérivé de pSU18 pCHAP8255 codant pour PulA ou PulA<sub>Am</sub>, sous contrôle du promoteur *lacZ* avec résistance au chloramphénicol (Cm<sup>R</sup>) ; le plasmide pMS5005 codant pour les ARNt et aaRS, contrôlé par le promoteur *araBAD*, portant la résistance à la spectinomycine (Sp<sup>R</sup>) ; et le grand plasmide pCHAP8861, codant pour chaque gène *pul* sauf *pulA*, sous contrôle des promoteurs maltose, portant la résistance à l'ampicilline (Ap<sup>R</sup>).

Pour tester le système, nous avons d'abord introduit des codons Amber dans les positions de PulA S54 et E71 (domaine N1) et S510 ou D526 (domaines Ins) sur le plasmide pCHAP8255. Nous avons utilisé ces constructions pour tester l'incorporation de *pBpa* et la photo-pontage *in vivo* dans la souche BW25113 F<sup>Tet</sup>. Ces quatre variantes PulA<sub>Am</sub> ont montré une bonne incorporation de *pBpa*, et nous avons détecté une bande dérivée dans le variant PulAE71Am lors de l'ajout de *pBpa* et de l'irradiation UV, absente dans le contrôle négatif (sans addition de *pBpa*) et dans PulA<sup>WT</sup>. Pour tester si ces bandes pontées sont spécifiques au SST2, nous avons effectué les mêmes expériences dans un système à 2 plasmides sans facteurs Pul, cependant, une bande à la même position a été détectée à nouveau, suggérant que la bande pontée contient une protéine d'*E. coli* qui n'est pas lié au SST2.

Parallèlement, nous avons également remarqué qu'en raison du poids moléculaire élevé de PulA et de la mauvaise séparation des protéines de haut poids moléculaire par SDS-PAGE, il serait difficile de faire la distinction entre PulA et ses dérivés PulA, notamment si le partenaire est de relativement petite masse moléculaire. Par conséquent, il était difficile de déduire la taille de la protéine pontée.

Afin de surmonter les limitations de l'immunodétection, nous avons introduit l'étiquette His<sub>6</sub> à l'extrémité C-terminale de PulA, afin de purifier PulA et ses complexes après une photo-pontage. Nous avons contacté la plateforme de spectrométrie de masse (MS) pour concevoir une approche d'identification des protéines pontées. Nous avons construit le plasmide pMS5032 codant pour PulA-His<sub>6</sub>, sous le promoteur *placZ*, portant la résistance au chloramphénicol (Cm<sup>R</sup>), puis avons introduit le codon Amber pour obtenir le plasmide pMS5040 codant pour le variant PulAE71Am-His<sub>6</sub>. Nous avons testé la sécrétion et effectué un test d'incorporation de *pBpa* et de photo-pontage en incluant 3 contrôles négatifs : sans addition de *pBpa*, PulA<sup>WT</sup>-His<sub>6</sub>, PulA<sup>WT</sup>-His<sub>6</sub> sans facteurs Pul. J'ai transformé ces plasmides dans la souche BW25113 F<sup>-</sup>Tet et testé la sécrétion de PulA, ainsi que l'incorporation de *pBpa* et le photo-pontage. Dans le test de sécrétion, nous avons remarqué que, par rapport au PulA-His<sub>6</sub> et au contrôle négatif, le variant PulAE71Am-His<sub>6</sub> s'accumule dans les cellules et sa sécrétion est réduite. Dans le test d'incorporation de *pBpa* et pontage *in vivo*, une bande pontée était présente spécifiquement pour le variant PulAE71Am-His<sub>6</sub>, et non dans les témoins négatifs.

J'ai purifié des échantillons de PulAE71Am-His<sub>6</sub> et PulA-His<sub>6</sub> et les ai envoyés à la plateforme protéomique de l'Institut Pasteur, dans le but de comparer les deux échantillons et d'identifier des protéines spécifiquement présentes uniquement dans l'échantillon PulAE71Am-His<sub>6</sub> et absentes du contrôle négatif. La protéine la plus abondante dans les échantillons était PulAE71Am-His<sub>6</sub>. Cependant, des centaines de protéines correspondaient à ces deux critères. Ce résultat suggère qu'une étape de purification sur gel est essentielle pour enrichir la protéine d'intérêt avant l'analyse par MS. De cette façon, nous augmenterions probablement les chances d'identifier la protéine pontée.

Après avoir mis en place les conditions expérimentales, nous avons criblé toutes les variantes PulA<sup>Am</sup>-His<sub>6</sub> construites en effectuant des essais d'incorporation de *pBpa* et de photo-pontage *in vivo*. Parmi eux, nous avons détecté une faible bande pontée dans le variant PulAQ521Am-His<sub>6</sub>. Pour tester si le partenaire potentiel est un composant du SST2, nous avons utilisé l'immunodétection du PulAQ521Am-His<sub>6</sub> *pBpa* irradié aux UV par anticorps anti-PulC, anti-PulD, anti-PulF, anti-PulG, anti-PulH, anti-PulL, anti-PulM et anti-PulN, en espérant détecter une bande de Mw supérieur à ces protéines seules, et co-migrant avec le dérivé PulA. Cependant, le résultat a révélé une mauvaise détection de tous les facteurs Pul. La faible présence des facteurs Pul était probablement due à la mauvaise croissance des bactéries portant les 3 plasmides en milieu minimal et résultant en une mauvaise induction des gènes *pul* qui sont typiquement induits en phase stationnaire précoce.

Pour concentrer le complexe ponté de ce variant PulA, nous avons cultivé 250 ml d'*E. coli* BW25113 F<sup>+</sup>Tet contenant les trois plasmides et codant pour PulAQ521Am-His6, et un contrôle négatif sans facteurs Pul. Nous avons effectué un test d'incorporation de *pBpa* pour ces 2 échantillons. Les cultures avec ou sans facteurs Pul ont été collectées respectivement à DO<sub>600</sub> de 1,67 et 1,8, concentrées dans 10 mL du tampon de lavage utilisé pour la purification par affinité ou His, et irradiées aux UV pendant 30 minutes à l'aide de la lampe UV Blak-Ray qui couvre une surface de diamètre de 10 à 15 cm. Nous avons purifié PulAQ521Am-His6 et le complexe et les avons analysés par Western blot. En comparant les fractions d'élution de PulAQ521Am-His6 avec et sans facteurs Pul, nous n'avons vu aucune différence entre les 2 échantillons.

Jusqu'à présent, nous étions confrontés à plusieurs problèmes, notamment la faible efficacité de l'incorporation du *pBpa* et de pontage, difficultés de détection des espèces à Mw plus élevé et faibles niveaux de protéines Pul. Nous avons identifié plusieurs raisons possibles à ces problèmes, telles que le faible nombre de copies de pMS5005, une faible croissance bactérienne avec 3 plasmides, une dynamique potentiellement élevée du système et la présence de l'étiquette His qui pourrait interférer avec la sécrétion de PulA et avec ses interactions avec les protéines Pul.

Pour résoudre ces problèmes, nous avons remplacé le tag His terminal C sur PulA par un tag Strep, puisque PulA-Strep semblait être sécrété plus efficacement. Il nous a semblé impératif de réduire la charge plasmidique sur les cellules. Nous avons introduit des codons TAG dans toutes les positions dans PulA<sup>WT</sup> sur le plasmide pCHAP8255. Dans le plasmide pCHAP8185, nous avons remplacé le *pulA*<sup>WT</sup> par *pulA-Strep* pour donner pMS5073 ou ses dérivés portant des mutations *pulA*Am-Strep. Ces changements nous ont permis de mettre en place un système à 2 plasmides combinant le plasmide codant pour le SST2 avec le plasmide pEVOL comme source d'ARNt suppresseur et d'ARNt synthétase (aaRS). Le nombre de copies plus élevé de pEVOL par rapport à pMS5005 devait améliorer l'incorporation de *pBpa*.

Sur la base de ce système à 2 plasmides, nous avons également optimisé les conditions de culture, suivant les conseils du prof. V. Shevchik. Au lieu d'adapter les bactéries au milieu minimal, puis de cultiver les précultures dans M63B1, nous avons cultivé les précultures dans du milieu LB additionné d'antibiotiques à 37°C. Le lendemain, nous avonsensemencé ces précultures dans du milieu minimal M63B1 avec 0, 5% de glycérol, 1% de casamino acides et nous avons suivi le protocole 2 d'incorporation de *pBpa* suivi d'essais de photo-pontage *in vivo*.

Nous avons testé plusieurs variants de PulA-Strep dans le contexte d'un système à 2 plasmides. Les cultures de 1 ml ont été irradiées aux UV dans une plaque à 24 puits par la lumière LED UV, puis analysés par immunodétection avec un anticorps anti-PulA. Dans ces expériences, nous avons vu une bonne incorporation de *pBpa* chez les variants PulA<sup>Am</sup>-Strep et de plus, certaines bandes pontées pour les variants PulA<sup>Strep</sup> E492, Q525, Q484, F476, E475 *pBpa*.

Nous avons sélectionné les variants PulA<sup>E492Am</sup>-Strep et PulA<sup>Q525Am</sup>-Strep pour produire des cultures de plus grand volume pour la purification des protéines tagées Strep. Nous avons cultivé les précultures bactériennes dans du milieu LB à 30°C sur la nuit et ensemencé 400 ml de milieu minimal M63B1 à une DO<sub>600</sub> de 0,5. Nous avons suivi le protocole modifié d'incorporation de *pBpa*. Nous avons récolté les bactéries dans 10 ml de tampon de lavage pour la purification de la résine Streptactin et les avons traitées avec la lampe UV Blak Ray pendant 30 minutes sur glace pour le photo-pontage. Ensuite, nous avons cassé les bactéries et purifié PulA-Strep sur la résine Strep-Tactin®XT 4Flow®. Après purification et concentration des protéines, nous avons pu observer une bande pontée dans la fraction d'élution 4 pour PulA<sup>Q525Am</sup>-Strep, qui a été enrichie après purification. Cependant, ce n'était pas le cas avec les échantillons PulA<sup>E492Am</sup>-Strep. Par rapport à l'extrait total avant purification, les fractions d'élution semblaient avoir perdu le complexe ponté. Notre protocole de purification a été adapté pour la purification de PulA-Strep soluble, mais cette perte a suggéré que le protocole doit être adapté pour la purification de complexes liés à la membrane. Toutes ces bandes n'étaient pas visibles par coloration au bleu de Coomassie, ce qui signifie que le rendement en PulA était trop faible pour une analyse par MS.

Pour concentrer les dérivés pontés, nous avons combiné les fractions d'élution 3 et 4 de chaque échantillon et effectué une précipitation au TCA. Nous avons analysé les protéines solubilisées dans un tampon d'échantillon de Laemmli par Western blot. Dans l'échantillon de PulA<sup>Q525Am</sup>-Strep, la bande pontée est devenue plus intense par rapport à la fraction E4 avant la précipitation du TCA. En ce qui concerne l'échantillon de PulA<sup>E492Am</sup>-Strep, la bande de PulA<sup>E492Am</sup>-Strep elle-même est devenue encore plus faible qu'avant la précipitation au TCA, ce qui indique que des protéines ont été perdues au cours du processus. Les deux échantillons précipités au TCA présentaient des produits de dégradation et étaient toujours invisibles par la coloration au bleu de Coomassie, ce qui excluait l'utilisation de la spectrométrie de masse.

Nous avons tenté d'utiliser Western blot pour identifier le partenaire ponté de PulA<sup>Q525Am</sup>-Strep. Nous avons donc analysé le PulA<sup>WT</sup>-Strep purifié, le lysat clarifié (CL),

le flux traversant (FT) et les fractions de lavage 1 et 6 du processus de purification, ainsi que les échantillons précipités au TCA sur SDS-PAGE et western blot. Nous avons sondé les membranes avec des anticorps anti-PulA, anti-PulC, anti-PulD, anti-PulF, anti-PulG, anti-PulJ, anti-PulL, anti-PulM, cependant, nous n'avons détecté aucune bande qui a migré avec PulAQ525Am Strep ponté. De plus, la souche BW25113 F<sup>Tet</sup> contient le gène *malE*, et nombre de nos anticorps anti-Pul ont été développés contre la chimère Pul-MalE, notamment les anti-PulC, anti-PulL et anti-PulM. Par conséquent, les bandes MalE ont également été révélées affectant le résultat, en particulier lorsque le poids moléculaire de la protéine est similaire à celui de MalE, comme dans le cas de PulL. Pour surmonter ce problème, pour les expériences suivantes, nous avons utilisé la souche PAP7460 dépourvue du gène *malE*, qui présentait une bonne incorporation de *pBpa* et ne semblait pas avoir de suppresseur Amber endogène.

#### Pontage chimique et co-purification de PulA-Strep et variants

Pour surmonter certaines des limites de l'approche ciblée, nous nous sommes tournés vers le pontage chimique suivie d'une purification du variant PulA-Strep. Sur la base du plasmide pMS5073 décrit ci-dessus, codant pour PulA-Strep et tous les facteurs Pul, nous avons construit son dérivé, le plasmide pMS5078 codant pour PulAΔIns-Strep avec des facteurs Pul comme contrôle. Un autre témoin dérivé de pMS5078 était le plasmide pMS5079 codant pour PulA-Strep avec une délétion du gène codant pour l'ATPase PulE. Ensuite, nous avons transformé ces plasmides dans la souche PAP7460 et effectué le test de sécrétion de PulA. Comme témoin négatif, nous avons inclus le plasmide pCHAP4395 codant pour PulA-Strep sans les gènes *pul*. Quatre cultures ont été collectées à OD<sub>600nm</sub> = 2.

Comme le montre le résultat d'immunodétection, il n'y avait pas de sécrétion de PulA sans les facteurs Pul, alors que la majeure partie de PulA-Strep se trouvait dans la fraction surnageant lorsque les facteurs Pul étaient présents. Nous n'avons observé aucune sécrétion de PulA-Strep en l'absence de PulE. Quant à la sécrétion de PulAΔIns-Strep, elle était partiellement défectueuse par rapport à PulA-Strep. Ceci est en accord avec les observations précédentes selon lesquelles la délétion de Ins affecte partiellement la sécrétion lorsque le système Pul est surproduit.

#### Définir le stade de croissance optimal pour la réticulation

Le système Pul est synthétisé au cours de la phase exponentielle tardive et stationnaire précoce de la croissance bactérienne. Nous avons cherché à capturer une étape où le niveau des facteurs Pul était suffisamment élevé et où PulA était partiellement sécrété. Pour trouver la

fenêtre de croissance optimale pour effectuer le test de pontage, nous avons transformé pMS5073 en souche PAP7460, cultivé les bactéries dans un milieu LB avec 0,4% de maltose, collecté des cultures à différentes  $DO_{600}$  et effectué des tests de sécrétion. Nous avons sondé la même membrane individuellement par anti-PulA et anti-PulM, pour surveiller la sécrétion de PulA et l'accumulation des facteurs Pul. Dans le résultat du western blot, nous avons pu voir qu'à la  $DO_{600} = 1,0$ , il n'y avait presque pas de PulA ou de PulM synthétisés. À  $OD_{600} = 1,4$ , la PulA et la PulM ont été produites, et étonnamment, la majorité de la PulA était déjà sécrétée dans le milieu à un stade aussi précoce. À  $OD_{600} = 1,7$ , plus de PulM s'est accumulé dans la cellule et seule une petite partie de PulA est restée dans la fraction cellulaire. Lorsque l' $OD_{600}$  a atteint 2,5, la quantité de PulM était la plus élevée, mais presque tout le PulA a été sécrété. Ces résultats suggèrent que PulA est sécrété rapidement et dès que le complexe SST2 est assemblé dans la membrane, et qu'il n'y avait pas de fenêtre "parfaite" dans laquelle les facteurs Pul étaient produits à des niveaux suffisants alors que PulA restait encore à l'intérieur de la cellule. Nous avons décidé que la fenêtre la plus appropriée pour effectuer des tests de réticulation pourrait être d'environ  $OD_{600} = 1,5$ . Les bactéries au-delà de ce stade ont continué à produire PulA et PulM, ce qui suggère que le processus de sécrétion se poursuivrait également, nous permettant de capturer au moins certains intermédiaires intracellulaires de sécrétion.

#### Capture de Strep-protéine (SPINE)

La méthode de capture de Strep-protéines également appelée SPINE a été rapportée pour la première fois par Herzberg et al. Elle associe le pontage *in vivo* au formaldéhyde et la purification par l'étiquette Strep de la protéine cible et de ses dérivés de pontage. Comme agent de pontage, nous avons utilisé du formaldéhyde qui pénètre dans la bactérie et lie les protéines en contact étroit (à une distance d'environ 2 Å), pour permettre la capture des interactions directes protéine-protéine. Les pontages par formaldéhyde sont également réversibles par chauffage, facilitant l'identification des protéines pontées par Western blot.

J'ai d'abord optimisé les conditions expérimentales - la concentration de formaldéhyde et le temps d'incubation, en transformant pMS5073 (PulA-Strep, Pul+) et pMS5079 (PulA-Strep  $\Delta$ PulE) individuellement dans la souche PAP7460. J'ai réalisé SPINE dans un petit volume de cultures en milieu minimal. J'ai traité des échantillons avec du formaldéhyde à 0,6 % à différents temps après l'avoir ajouté directement dans le milieu de culture. Le traitement a toujours été effectué à 30°C, la température qui a également été utilisée pour la croissance bactérienne. Comme il n'y avait pas de différence entre les traitements de 20, 30, 40 min, nous avons augmenté la concentration à 0,8 % de formaldéhyde et traité pendant 30 min. Les échantillons

n'ont pas été bouillis avant l'analyse par SDS-PAGE et western blot, de sorte que les complexes pontés ne soient pas dissociés. Nous avons observé deux formes principales de PulA monomérique, la PulA entièrement dénaturée et la PulA à migration plus rapide, qui est probablement partiellement repliée. Dans les échantillons traités au formaldéhyde, les bandes PulA étaient plus faibles que sans traitement au formaldéhyde, et les trainées au-dessus des bandes PulA indiquait que des complexes auraient pu se former. Lorsque nous avons utilisé 0,8% de formaldéhyde et 30 minutes d'incubation, la majorité de PulA était encore monomérique, ce qui montre que ces conditions n'étaient pas trop fortes - nous les avons donc utilisées pour la future SPINE.

Pour la première SPINE, nous avons cultivé un volume de 200 ml de chaque souche PAP7460 (pMS5073) et PAP7460 (pMS5079). Aux cultures PAP7460 (pMS5073) (à  $OD_{600} = 1,44$ ) et PAP7460 (pMS5079) (à  $OD_{600} = 1,25$ ), nous avons ajouté 0,8 % de formaldéhyde et incubé pendant 30 min. Nous avons purifié les protéines Strep avec la résine Streptactin (IBA). Les protéines purifiées n'étaient pas visibles par la coloration au bleu de Coomassie, qui, avec l'analyse Western blot, indiquait que le rendement de PulA-Strep était assez faible. C'était particulièrement le cas pour les échantillons PAP7460 (pMS5073), où seule une petite quantité de PulA a été détectée dans la fraction de lysat clair (CL). Cela pourrait être dû au fait que la majorité de PulA a été sécrétée dans le surnageant de culture. De plus, la présence de PulA dans l'écoulement (FT) et les fractions de lavage suggèrent une faible liaison de PulA-Strep à la résine Streptactin. De plus, une grande quantité de PulA restait dans les fractions membranaires de chaque culture.

Pour résoudre le problème d'affinité, nous avons commandé une nouvelle résine Streptactin (Strep-Tactin®XT 4Flow®) pour la purification PulA-Strep. Nous avons encore augmenté la concentration de formaldéhyde à 1 %. Nous avons cultivé 300 mL de PAP7460 (pMS5079) dans du milieu M63B1 et ajouté 1 % de formaldéhyde à la culture à  $OD_{600} = 1,7$  pendant 30 min à 30°C. Les protéines purifiées sur résine Strep-Tactin®XT 4Flow® ont été chargées sur gel SDS-Glycine 8% sans ébullition. Sur Western blot, des bandes intenses correspondant à PulA étaient présentes dans les échantillons avant et après pontage au formaldéhyde (NC et XL), ainsi que dans les fractions de lysat (CL) et d'élution clarifiées. Cependant, les bandes correspondant à PulA' étaient enrichies dans les fractions d'élution. Aucune bande migrant plus lentement que PulA n'a été détectée. Cela pourrait être dû au fait qu'il n'y avait pas de protéines pontées, ou parce que le PulA pontée était présente sous la forme rapide PulA', de sorte que les bandes pontée étaient présentes dans les trainées entre PulA et PulA'. Par coloration au bleu de

Coomassie, aucune bande correspondant à la PulA dénaturée n'était visible et des bandes co-migrant avec la PulA' ont été observées dans les fractions CL, FT, W1-2 et E2-6.

Dans l'ensemble, ces résultats ont montré que PulA-Strep accumulé dans les cellules en raison de l'absence de PulE pouvait être purifié avec succès avec la résine améliorée Strep-Tactin®XT 4Flow®. Pour tester si des protéines Pul étaient pontée à PulA-Strep, j'ai testé le PulA-Strep purifié (à partir du plasmide pCHAP4395), les échantillons NC, CL, FT, W1, W6 (comme témoins négatifs) et la fraction E3 (où PulA-Strep était le plus abondant), par immunodétection avec des anticorps anti-PulA, ainsi qu'avec des anticorps contre PulC, D, F, G, J, L et M. Les échantillons ont été bouillis pendant 20 minutes pour dissocier les complexes pontés et dénaturer le PulA'. PulA était présent dans chaque fraction, tandis que PulC, D, F, G, J, L et M n'étaient détectés que dans les fractions NC, CL, FT et W1. Nous n'avons détecté aucune des protéines Pul dans la fraction d'élution 3.

### Étude des interactions PulA par double hybride bactérien (BACTH)

Comme discuté ci-dessus, nous avons émis l'hypothèse que PulA interagit avec un des composants de la membrane interne du Pul SST2 à un stade précoce de la sécrétion. Cette interaction pourrait être impliquée dans le recrutement de substrats SST2, comme cela a été proposé pour GspL ou GspM dans le SST2 chez *Pseudomonas*. Nous avons décidé à utiliser BACTH comme approche complémentaire pour étudier ces interactions. Avant sa sécrétion, la PulA native est une lipoprotéine liée à la membrane interne. Cette localisation IM est nécessaire à sa bonne sécrétion. Le test BACTH repose sur la signalisation cytoplasmique par AMPc, qui peut être produite en réponse aux interactions des protéines dans la membrane interne ou même dans le périplasme. Pour permettre la signalisation avec le PulA périplasmique, nous l'avons fusionné à un segment transmembranaire neutre de la protéine MalF ou au segment de la pseudopiline mineure PulK.

### Interactions de PulA avec les pseudopilines

Nous avons d'abord effectué un test BACTH pour déterminer si PulA fusionné au TMS de MalF interagit avec la pseudopiline majeure PulG ou avec les pseudopilines mineures PulH, PulI, PulJ et PulK. Les deux modèles de mécanisme de sécrétion SST2 prédisent les interactions du substrat avec le pseudopilus. Dans le modèle de piston, la sommet du pseudopilus comprenant le complexe des pseudopilines mineures devrait interagir avec le substrat pendant

la sécrétion. Dans le modèle de la vis d'Archimède, l'interaction spécifique impliquerait plutôt la pseudopiline majeure.

Nous avons fusionné PulA-MalF<sup>TMS</sup> à l'extrémité C-terminale des fragments T18 ou T25 CyaA et co-transformé la souche DHT1 ( $\Delta cya$ ) avec les plasmides résultants et avec les plasmides compatibles codant pour T18 et T25 fusionnés à des pseudopilines matures entières. Nous avons mesuré les niveaux d'activité de la  $\beta$ -galactosidase et constaté que T18-MalF<sup>TMS</sup>-PulA interagit de manière significative avec PulG, PulH, PulJ et PulK comparé au contrôle négatif. Avec la construction T25-MalF<sup>TMS</sup>-PulA, les interactions de PulA avec toutes les pseudopilines étaient significativement plus élevées que le contrôle négatif.

Les données BACTH ont montré les signaux d'interaction les plus forts entre PulA et PulG dans les deux sens. Pour valider ce contact, nous avons comparé l'interaction entre PulG et PulA, avec le domaine PulA N1 seul et avec le variant PulA $\Delta$ Ins codé par les constructions pUT18C ou pKT25. Par rapport au contrôle négatif, les interactions entre PulA et PulG étaient hautement significatives avec les deux combinaisons de constructions. Cependant, l'interaction entre T18-N1 et T25-PulG était significativement plus élevée par rapport à celle de T18-PulA et T25-PulG, ou T18-PulA $\Delta$ Ins avec T25-PulG. Ceci est surprenant, étant donné que T18-PulA et T18-PulA $\Delta$ Ins contiennent également le domaine N1. L'activité des bactéries produisant T18-PulA $\Delta$ Ins avec T25-PulG était significativement plus faible qu'entre T18-PulA et T25-PulG, suggérant l'importance du domaine Ins pour l'interaction PulA - PulG. Cependant, lorsque les constructions complémentaires ont été utilisées, l'interaction entre T25-PulA $\Delta$ Ins et T18-PulG était aussi forte que celle de T25-PulA avec T18-PulG. Comme indiqué précédemment, PulG forme des homodimères dans ce test. Ensemble, ces résultats ont confirmé l'interaction PulA - PulG, mais les contributions des domaines N1 et Ins sont restées floues.

#### Interactions PulA avec les protéines Pul SST2 de la membrane interne

Pour tester si le segment transmembranaire de MalF contribue aux interactions de PulA dans BACTH, nous l'avons remplacé par le segment transmembranaire de PulK. Nous avons testé les interactions des hybrides T18/T25-PulK<sup>TMS</sup>-PulA avec les hybrides correspondants de PulB, PulC, PulF, PulL et PulM. Les plasmides BACTH ont été introduits dans la souche *E. coli* DHT1 et l'activité  $\beta$ -galactosidase d'au moins 6 clones indépendants a été mesurée. Les fragments CyaA T18 et T25, utilisés comme contrôle négatif, n'interagissent pas et ne restaurent pas l'activité adénylyl cyclase. Par rapport au contrôle négatif, nous avons constaté que les

signaux d'interaction PulA avec PulM et PulF étaient forts et hautement significatifs dans les deux combinaisons de constructions.

Nous avons été surpris de détecter l'interaction de PulA avec PulF, qui n'a qu'une petite partie exposée dans le périplasme. Nous soupçonnions que le TMS de PulK, et non PulA, pourrait interagir avec PulF. En effet, les tests BACTH ont montré que T18-PulK<sup>TMS</sup> interagit avec T25-PulF, contrairement à T18-MalF<sup>TMS</sup>. Par conséquent, nous avons testé les interactions des hybrides T18/T25- MalF<sup>TMS</sup>-PulA avec PulF par BACTH. Dans le contexte des hybrides MalF<sup>TMS</sup>-PulA, l'interaction entre PulA et PulF était significativement plus forte que celle de PulF avec le contrôle négatif MalF<sup>TMS</sup>.

Dans notre laboratoire, Ingrid Guilvout a créé plusieurs variants mutants pour étudier la fonction de PulF. Elle a isolé deux variants d'insertion de PulF-Flag, PulFL198Flag et PulFA201Flag, qui ont presque aboli la sécrétion de PulA mais qui favorisent la piliation. Les peptides Flag sont insérés dans la boucle périplasmique de PulF suggérant que cette région de PulF pourrait jouer un rôle important dans l'interaction avec PulA. Ainsi, nous avons fusionné ces deux variants avec des fragments T18 et T25 Cya et testé leurs interactions avec PulA (MalF<sup>TMS</sup>). Les interactions entre PulA et PulFL198Flag ou PulFA201Flag étaient toutes significativement inférieures à celles de PulA et PulF<sup>WT</sup>. Cependant, ces insertions de Flag ont également affecté la dimérisation et le niveau de PulF.

#### Analyse des composants de la plate-forme d'assemblage Pull et PulM

La plate-forme d'assemblage de la membrane interne composée de GspF, GspL et GspM et de leurs orthologues joue un rôle central dans le fonctionnement des machineries SST2 et T4P. Jusqu'à présent, les structures des domaines individuels de la plupart des composants AP ont été déterminées, cependant, les connaissances sur les protéines entières et leurs interactions sont encore très limitées. La co-purification de l'ensemble du complexe AP n'a pas encore été accomplie, probablement en raison des interactions faibles ou transitoires, notamment entre leurs domaines hydrophobes. La purification du complexe SST2 de *K. pneumoniae* a été une avancée importante, mais PulF et les pseudopilines n'ont pas été co-purifiés avec le reste.

Plusieurs études antérieures des composants AP GspL et GspM dans différents modèles SST2 ont montré leurs interactions relativement fortes, ce qui a motivé nos études structure-fonction de ce sous-complexe. Les structures cristallines des domaines périplasmiques d'EpsM et d'EpsL du système de sécrétion de type II de *V. parahaemolyticus* ont été déterminées par Abendroth et al. en 2004 et 2009, ainsi que la structure du domaine cytoplasmique d'EpsL seul

et en complexe avec le domaine ATPase EpsE<sub>N</sub>. Les prédictions et les études de topologie de GspL et GspM suggèrent que les deux contiennent une seule hélice transmembranaire connectée à une longue hélice périplasmique et domaines C-terminaux (GspL<sup>CTD</sup> et GspM<sup>CTD</sup>), qui adoptent un pli similaire à la ferrédoxine. Les GspL<sup>CTD</sup> et GspM<sup>CTD</sup> forment des homodimères dans les cristaux, bien que différentes interfaces aient été impliquées dans la dimérisation. Parallèlement, des études biochimiques et *in vivo* utilisant le BACTH ont montré une forte interaction entre les domaines périplasmiques et transmembranaires de OutL et celui de OutMp, et en même temps, une homo-dimérisation assez faible pour chacun chez *D. dadantii*.

Nos collaboratrices, Régine Dazzoni, Aracelys Lopez-Castilla du groupe de Nadia Izadi-Pruneyre ont combiné des approches biophysiques, RMN et cristallographie aux rayons X pour étudier les composants de l'AP chez *Klebsiella* - PulL et PulM. Elles ont déterminé la structure et les associations de leurs domaines périplasmiques et décrit la première structure de l'hétérodimère formé par leurs domaines de type ferrédoxine. Pour relier ces résultats à la situation *in vivo*, j'ai réalisé les études BACTH de PulL et PulM et montré que, contrairement aux domaines CTD isolés, les protéines entières ne forment pas d'homodimères. Une faible homodimérisation de PulM a été observée mais statistiquement non significative. J'ai également effectué une analyses de segments transmembranaires de PulL et PulM par pontage Cys suivi d'oxydation et pontage disulfure. Ces études ont confirmé que PulM peut former des homodimères qui pourraient être stabilisés par oxydation. J'ai également contribué à la mutagenèse dirigée et aux essais fonctionnels qui visaient à étudier le rôle de l'interface PulL<sup>CTD</sup> - PulM<sup>CTD</sup>. Enfin, j'ai utilisé le western blot quantitatif pour évaluer les quantités relatives de PulL et PulM dans les cellules bactériennes. J'ai montré que PulM était environ 20 fois plus abondant que PulL, ce qui est en accord avec son rôle dynamique dans le SST2 en tant que facteur qui se lie à PulG et canalise ces sous-unités vers PulL au niveau de la plateforme d'assemblage. Le manuscrit de Dazzoni et al. décrivant ce travail a été soumis à la revue Structure et est présenté ici dans sa version préimprimée.

L'étude de Dazzoni et al. a prédit que les longues hélices périplasmiques de PulL et PulM forment une bobine enroulée et le rôle de ces régions n'a pas été testé. Dans l'étude suivante, nous avons étudié leur fonction et leur contribution à l'interaction PulL-PulM. Nous avons constaté que cette région est extrêmement importante pour l'interaction et que sans elles, les fonctions du SST2 sont presque abolies. Nous avons également étudié le rôle du court peptide cytoplasmique de PulM. La délétion de ce peptide ne semble pas affecter l'interaction avec PulL, mais elle affecte néanmoins fortement la stabilité de PulL et conduit à un clivage de PulL en un

point précis du cytoplasme. La même région cytoplasmique de PulM est intéressante car elle interagit également avec PulG. Nos études suggèrent que cette région est cruciale au cours du processus dynamique d'assemblage du pseudopilus et qu'elle pourrait interagir et stabiliser le domaine cytoplasmique de PulL. Ces résultats sont décrits dans le manuscrit de Li, Santos-Moreno et al. qui est en cours d'examen dans la revue *Research in Microbiology* présenté ci-dessous.

#### Validation fonctionnelle de l'interface PulL<sup>CTD</sup>-PulM<sup>CTD</sup>

Pour évaluer le rôle de l'association PulL<sup>CTD</sup>-PulM<sup>CTD</sup> dans la sécrétion de PulA, nous avons muté les résidus impliqués dans l'interface. Nous avons introduit des substitutions d'alanine simples aux positions L319, L323, L341 et Q327 de PulL et testé la capacité de ces variants à restaurer la sécrétion de pullulanase (PulA) chez *Escherichia coli* portant les gènes *pul* avec une délétion non polaire du gène *pulL* dans le plasmide pCHAP8251. Alors que les substitutions Ala des résidus Leu n'affectaient pas la fonction, le remplacement du résidu polaire Q327 par l'alanine a conduit à un défaut, faible mais significatif, de la sécrétion de PulA. Ce défaut était encore exacerbé en présence de substitutions L319A ou L341A. Du côté PulM, des substitutions Ala simples de F116, Q117, M120 et D124 de l'hélice  $\alpha 2$ , ou L136 du brin  $\beta 3$ , n'ont pas affecté la fonction. Cependant, le variant PulM<sup>F116A-L136A</sup> a montré un défaut de sécrétion important, suggérant que l'interface était affaiblie. Le remplacement du résidu M120 exposé en surface par des résidus chargés dans les variantes PulM<sup>M120D</sup> et PulM<sup>M120K</sup> a également réduit de manière significative la sécrétion. En comparaison, les inversions de charge à la surface de l'hélice PulM<sup>CTD</sup>  $\alpha 1$ , telles que R88E et R92D, n'ont montré aucun effet, conformément à leur position éloignée par rapport à l'interface avec PulL<sup>CTD</sup>. Bien que nous ne puissions pas exclure que ces mutations affectent l'interaction avec d'autres composants du SST2, leur localisation à l'interface et la corrélation entre leurs effets avec les modifications plus ou moins conservatrices des acides aminés plaident en faveur de leur effet direct sur la liaison PulL<sup>CTD</sup>-PulM<sup>CTD</sup>.

Ensemble, ces données montrent que l'interface intacte PulL<sup>CTD</sup>-PulM<sup>CTD</sup>, et donc le complexe PulL-PulM, jouent un rôle important dans la sécrétion. Il est possible que d'autres régions PulL et PulM contribuent également à cette interaction *in vivo*. Une indication est le variant PulM<sup>N3A</sup> dans l'extrémité N-terminale cytoplasmique de la protéine, qui a considérablement affecté la stabilité des deux protéines et la sécrétion.

Pour mieux comprendre l'interaction entre PulL et PulM entières dans le contexte cellulaire et dans la membrane, nous avons combiné des approches de BACTH et de pontage par cystéines. L'approche BACTH permet d'analyser les interactions entre protéines membranaires dans leur environnement natif. Pour étudier comment les PulL et PulM entières interagissent, nous les avons fusionnés aux extrémités C-terminales de T18 et T25, fragments du domaine catalytique de l'adenylyl cyclase (CyaA). Les plasmides ont été introduits dans la souche *E. coli* DHT1 portant une délétion du gène *cya* endogène. Les fragments CyaA T18 et T25 n'interagissent pas et ne restaurent pas l'activité de l'adenylyl cyclase comme indiqué par les colonies blanches sur les boîtes X-gal et les faibles niveaux d'expression du gène *lacZ*. Les bactéries contenant des hybrides T18-PulL et T25-PulL ont également montré des colonies blanches et une faible activité, comparable aux témoins négatifs, ce qui indique que le PulL ne forme pas d'homodimères. Les bactéries portant les chimères T18-PulM et T25-PulM étaient bleu pâle sur l'indicateur X-gal et présentaient une activité un peu plus élevée, indiquant une tendance de PulM à s'homodimériser. Enfin, les bactéries co-produisant les chimères PulL et PulM ont montré des signaux d'interaction forts et hautement significatifs. Ces résultats montrent que les PulL et PulM entières interagissent fortement les unes avec les autres et forment préférentiellement des hétérodimères *in vivo*.

PulL et PulM sont des protéines membranaires de type I avec une orientation N-in C-out, chacune contenant un seul segment transmembranaire (TM). Pour étudier l'implication de ces segments dans l'interaction PulL-PulM, nous avons utilisé le pontage par cystéines. PulM natif possède un résidu Cys en position 17, dans son segment TM. Nous avons d'abord remplacé ce résidu par un Leu et utilisé PulM<sup>C17L</sup> comme base pour générer une série de variants PulM avec une seule substitution Cys aux positions entre 18 et 24. De la même manière, nous avons remplacé Cys264, localisé dans la partie périplasmique de PulL (PulL<sup>C264L</sup>) et utilisé ce variant pour générer des substitutions Cys uniques aux positions 249 à 256 du segment TM. Les bactéries produisant toutes les combinaisons de variants Cys simples de PulL et PulM ont été traitées avec CuCl<sub>2</sub> comme oxydant et leurs extraits totaux ont été analysés par Western blot avec des anticorps anti-PulL ou anti-PulM. Le variant PulL<sup>C264L</sup> n'a produit aucun pontage significatif. Le résultat similaire a été observé en présence de résidus Cys uniques aux positions 249, 252, 253, 254, 255 et 256. Cependant, PulL<sup>C264L</sup> avec Cys aux positions 250 et 251 a produit certains homodimères ainsi que des hétérodimères avec PulM contenant Cys aux positions 17 (PulM<sup>WT</sup>) et 18 (PulM<sup>C17L-L18C</sup>). Ce dernier variant a donné les bandes

d'hétérodimères les plus intenses, qui ont également été détectées avec des anticorps anti-PulM. Les variants PulM avec Cys aux positions 17, 18 et 19 ont également produit des bandes homodimères. Ces résultats montrent que PulL et PulM peuvent interagir *via* leurs segments transmembranaires et que dans cette région ils utilisent également la même interface pour l'homo- et l'hétérodimérisation. Les niveaux d'homodimères PulM n'étaient pas affectés par la présence de PulL<sup>I250C</sup> et de PulL<sup>V251C</sup>, peut-être parce que PulM est plus abondant que PulL dans ces souches. Conformément à une fraction plus élevée de PulL ponté avec PulM, les hétérodimères LM étaient plus facilement détectables dans les membranes révélées par anti-PulL que ceux avec anti-PulM. Confirmant le rôle stabilisateur de PulM, les niveaux de tous les variants de PulL ont été significativement réduits en l'absence de PulM (Ø). En revanche, les niveaux de PulM n'ont pas changé en l'absence de PulL.

Les expériences ci-dessus ont été réalisées dans des bactéries produisant PulL et PulM en l'absence d'autres composants SST2 et à des niveaux qui peuvent différer de ceux du système fonctionnel. Dans l'opéron *pulC-O* codant pour l'SST2 de *Klebsiella*, les gènes *pulL* et *pulM* sont adjacents et transcrits à partir du même promoteur. Pour déterminer leur abondance relative, nous avons quantifié PulL et PulM dans les bactéries produisant le SST2 fonctionnel à partir d'un seul plasmide pCHAP8185 à nombre de copies modéré. Nous avons utilisé une analyse Western blot avec des anticorps dirigés contre PulL<sup>CTD</sup> et PulM<sup>CTD</sup> pour quantifier leurs niveaux dans des extraits bactériens exprimant les gènes SST2, en comparant ces signaux avec ceux des quantités connues de PulL<sup>CTD</sup> et PulM<sup>CTD</sup> purifiés. Ces mesures ont indiqué que PulM est présent en grand excès par rapport à PulL, avec un rapport molaire d'environ 20 : 1.

La combinaison des données structurales sur l'hétérodimère PulL<sup>CTD</sup>-PulM<sup>CTD</sup> avec des expériences de pontage cystéine nous a permis de construire un modèle 3D du complexe PulL-PulM. L'analyse de séquence a révélé la présence de longues hélices  $\alpha$  périplasmiques à la fois dans PulL et PulM, englobant les segments transmembranaires et en amont des CTD. De plus, deux régions impliquées dans la formation de coiled-coil dans les deux protéines sont prédites, la première chevauchant les segments transmembranaires. Par conséquent, les hélices périplasmiques de PulL et PulM ont été modélisées comme une bobine enroulée, dans laquelle l'emballage des boutons dans les trous a été confirmé par l'analyse SOCKET2. Dans le modèle, l'alignement relatif des segments transmembranaires prédits est compatible avec une orientation orthogonale des hélices périplasmiques par rapport à la membrane. La bobine enroulée formée par les hélices périplasmiques de PulL et PulM est suivie de courtes régions linker désordonnées reliant les hélices  $\alpha$ 1 de l'hétérodimère CTD. Dans le modèle du complexe, un pont salin

potentiel se trouve entre D276 de PulL et R42 de PulM. Fait intéressant, la variante PulM<sup>D276K</sup> a affiché une sécrétion considérablement réduite, ce qui pourrait indiquer que la bobine enroulée contribue également à la surface d'interaction globale dans le complexe PulL-PulM.

Les hélices périplasmiques de PulL et PulM sont importantes pour la fonction SST2

PulL et PulM partagent une architecture similaire de leurs domaines transmembranaires et périplasmiques. Leurs hélices transmembranaires hydrophobes s'étendent dans le périplasme et sont suivies par les lieurs flexibles connectés aux domaines C-terminaux globulaires de type ferredoxine (CTD). Les deux hélices périplasmiques devraient former une bobine enroulée. Pour tester leur importance fonctionnelle, nous avons généré des variants PulL et PulM dépourvus de ces hélices périplasmiques. Dans PulL, nous avons éliminé la région comprenant les résidus 270 à 302 pour donner le variant PulL<sup>ACC</sup>. Nous avons d'abord testé la capacité de PulL<sup>ACC</sup> à favoriser la sécrétion de la pullulanase (PulA), le substrat du *K. oxytoca* SST2. Alors que le PulL natif (PulL<sup>WT</sup>) permet une sécrétion efficace de PulA, la majorité du pool de PulA s'accumule dans la fraction cellulaire en présence de PulL<sup>ACC</sup>. L'accumulation intracellulaire de PulA était complète dans le contrôle négatif dépourvu de PulL. Nous avons utilisé des anticorps dirigés contre PulL<sup>CTD</sup> pour évaluer la stabilité de PulL dans ces souches. La quantification des signaux PulL<sup>ACC</sup> a montré que les niveaux de protéines étaient réduits d'environ 30 % par rapport au PulL natif.

Pour étudier le rôle de l'hélice périplasmique de PulM, nous avons retiré la région comprenant les résidus 37 à 68 résultant en variant PulM<sup>ACC</sup> codée par le plasmide pMS1350. Comme dans le cas de PulL, cette délétion a entraîné un fort défaut de sécrétion de PulA. PulA s'accumule dans la fraction cellulaire en présence de PulM<sup>ACC</sup> à des niveaux similaires au témoin négatif. La stabilité de PulM<sup>ACC</sup> a été réduite de 10 % en moyenne par rapport à PulM.

Défauts d'assemblage des pilus chez les mutants pulL<sup>ACC</sup> et pulM<sup>ACC</sup>

Les composants AP PulL et PulM participent directement à l'assemblage du pseudopilus, qui devrait entraîner la sécrétion de PulA. Nous avons donc demandé si les hélices périplasmiques de PulL et PulM sont nécessaires pour l'assemblage du pilus PulG. Dans des conditions où les bactéries sont cultivées sur milieu solide, *E. coli* avec un plasmide codant pour un ensemble de gènes *pul* produit des pili de surface composés de PulG, pseudopiline majeure. Nous avons testé la capacité de PulL<sup>ACC</sup> et PulM<sup>ACC</sup> à compléter cette fonction chez des mutants dépourvus de *pulL* ou *pulM*. Après 2 jours de croissance en présence de maltose pour induire l'expression du gène *pul*, les bactéries complémentées avec les gènes *pulL* et *pulM* de

type sauvage ont produit des pili qui ont pu être détachés et séparés du pool PulG associé aux cellules. En présence des variants PulL<sup>ACC</sup> et PulM<sup>ACC</sup>, l'assemblage du pilus PulG a été presque aboli. Lorsque les bactéries ont été cultivées pendant une période standard de 48 heures pour ce test, aucune PulG n'a pu être détectée dans les fractions des pili. Ce n'est qu'en prolongeant la période de croissance pendant 24 heures supplémentaires que nous avons pu détecter des pili PulG à la surface des mutants de délétés de bobine enroulée. De plus, nous avons remarqué que les niveaux de PulG cellulaire étaient réduits en l'absence de PulL ou de PulM.

#### Les domaines périplasmiques sont nécessaires pour l'assemblage du complexe PulL-PulM

Pour déterminer si la délétion des hélices périplasmiques affecte l'interaction PulL – PulM, nous avons utilisé le double hybrides bactérien (BACTH). PulL, PulM entières et leurs variantes ont été fusionnés aux extrémités C-terminales des fragments T18 et T25 CyaA. Les constructions ont été co-transformées dans la souche mutante *E. coli cya* DHT1 pour évaluer la reconstitution de l'activité de l'adénylyl cyclase. En mesurant l'activité bêta-galactosidase, nous avons comparé les interactions des constructions PulL et PulM entières avec les constructions hébergeant les variants PulL<sup>ACC</sup> et PulM<sup>ACC</sup>. Pour chaque couple testé, l'activité a été mesurée pour 8 co-transformants indépendants pris au hasard.

Les bactéries contenant des plasmides codant pour T18-PulM et T25-PulL ont montré une activité bêta-galactosidase élevée, comparable au contrôle positif, indiquant une formation efficace d'hétérodimères. Dans les souches hébergeant T18-PulM<sup>ACC</sup> et T25-PulL, l'activité moyenne est restée élevée, ce qui suggère que l'interaction de PulM<sup>ACC</sup> avec PulL est similaire à celle de PulM<sup>WT</sup>. Cependant, dans plusieurs isolats, l'activité était fortement réduite, indiquant que la suppression de l'hélice périplasmique PulM affectait l'efficacité de la liaison à PulL. La délétion de l'hélice PulL dans T25-PulL<sup>ACC</sup> a eu un effet beaucoup plus fort et a presque aboli l'interaction avec T18-PulM. En comparaison, la combinaison des deux délétions dans T18-PulM<sup>ACC</sup> et T25-PulL<sup>ACC</sup> a augmenté l'activité moyenne, vraisemblablement en plaçant les CTD respectifs au même niveau par rapport à la membrane pour faciliter leurs contacts. Cependant, cette augmentation était faible et non statistiquement significative dans ce contexte expérimental.

Les souches hébergeant les hybrides T18-PulL et T25-PulM ont montré une activité  $\beta$ -galactosidase globalement plus faible. Cette différence est probablement due à la plus faible stabilité des hybrides T18-PulL par rapport aux hybrides T18-PulM. L'activité de la bêta-galactosidase a été réduite à des niveaux de base lors de l'élimination des résidus d'hélice

périplasmique dans T18-PulL<sup>ACC</sup> ou T25-PulM<sup>ACC</sup>. La combinaison des deux suppressions dans ce contexte n'a pas amélioré l'interaction. Dans l'ensemble, ces données suggèrent que les régions formant des bobines enroulées périplasmiques jouent un rôle majeur dans l'assemblage PulL-PulM.

#### L'hélice périplasmique de PulM n'est pas nécessaire pour la liaison à PulG

PulM a été caractérisé comme un facteur d'adressage pour les sous-unités PulH et PulG lors de l'élongation du pseudopilus. Étant donné que la suppression de l'hélice périplasmique PulM affectait l'assemblage du pseudopilus, nous avons demandé si cela affectait également les interactions avec PulG. Comme le montrent les résultats du test BACTH, bien que la délétion de l'hélice périplasmique de PulM réduise l'interaction de la paire T18-PulM - T25-PulG, le variant PulM<sup>ACC</sup> pouvait toujours interagir fortement et significativement avec PulG par rapport au contrôle négatif. De plus, il n'y avait aucun effet sur l'interaction avec T18-PulG. Des études antérieures ont montré que PulL n'interagit pas avec PulG dans le test BACTH. Nous avons ici confirmé ce résultat, en montrant que les activités  $\beta$ -galactosidase des bactéries co-exprimant *pulG* et *pulL* dans les constructions BACTH se situaient au niveau du contrôle négatif. Comme le PulL<sup>WT</sup>, le variant PulL<sup>ACC</sup> n'a pas interagi avec PulG. Sur la base de ces données, nous avons conclu qu'il est peu probable que le défaut de pilation causé par PulM<sup>ACC</sup> soit dû à un défaut de liaison à la pseudopiline.

#### Le rôle du peptide cytoplasmique PulM

Le domaine cytoplasmique de PulL interagit de manière stable avec l'ATPase PulE, d'après l'étude de leurs homologues chez *Vibrio cholerae*. La région cytoplasmique de PulM est un peptide à 16 résidus, fonctionnant vraisemblablement comme une ancre membranaire chargée positivement. Nous avons émis l'hypothèse que cette région interagit avec les résidus polaires N-terminaux de PulG, car leurs substitutions dans les variants PulG<sup>T2A</sup> et PulG<sup>E5A</sup> réduisent fortement l'interaction avec PulM. Pour tester cette hypothèse, nous avons supprimé le peptide cytoplasmique de PulM, et étudié les interactions de ce variant, désigné PulM $\Delta$ N, avec PulL et PulG. Dans le test BACTH, variant PulM $\Delta$ N a montré un défaut fort et significatif d'interaction avec PulG, confirmant nos prédictions. Ce défaut était comparable à celui de la liaison de PulM au variant PulG<sup>E5A</sup>, caractérisé précédemment. Dans le même temps, la délétion du peptide cytoplasmique PulM n'a pas affecté ses interactions avec PulL dans ce test.

Nous avons testé la capacité de PulM $\Delta$ N à favoriser la sécrétion de PulA et avons constaté que presque tout PulA restait dans les cellules lors des tests de sécrétion, comme le contrôle

négatif dépourvu de PulM. Ce phénotype est probablement causé par l'interaction défectueuse de PulM $\Delta$ N avec PulG. Bien que les niveaux de variant PulM $\Delta$ N aient été réduits par rapport à PulM<sup>WT</sup>, la protéine était toujours bien produite et migrait sur SDS-PAGE selon son Mw attendu (~ 16,5 kDa). Fait intéressant, ce variant semblait former des dimères résistants au SDS qui n'étaient pas détectables avec PulM<sup>WT</sup>. Nous avons également testé les niveaux de PulL dans ces souches. Étonnamment, bien que la délétion du N-peptide de PulM n'ait pas affecté son interaction avec PulL dans le test BACTH, nous avons constaté que PulL était fortement déstabilisé en présence de PulM $\Delta$ N, résultant en un produit protéolytique d'une taille similaire à celle de PulM. Des anticorps dirigés contre le PulL<sup>CTD</sup> ont détecté ce produit, suggérant qu'il comprend les régions de la membrane interne et périplasmiques de PulL. Comme elle semble être générée par un clivage spécifique qui enlève le domaine cytoplasmique PulL, nous avons désigné cette protéine PulL $\Delta$ N. Ces résultats suggèrent que le peptide cytoplasmique PulM assure non seulement l'adressage de PulG vers le site d'assemblage par interaction directe, mais qu'il protège également PulL d'une protéolyse spécifique.

#### Résultats complémentaires sur PulL et PulM

Les études BACTH suggèrent que PulA interagit avec PulM, et que PulL et PulM forment des hétérodimères. Nous nous demandions si la présence de PulA affecterait l'interaction de PulL et PulM. Nous avons conçu un « test triple-hybride », en co-transformant PulL et PulM dans la souche DHT1 hébergeant le plasmide codant pour PulA ou le vecteur vide (pSU18). Les résultats ont montré que la présence de PulA ou du vecteur dans la souche DHT1 n'affecte pas les niveaux de  $\beta$ -galactosidase dans le contrôle négatif T18-T25 ou positif T18Zip-T25Zip. Lorsque nous avons comparé l'activité des paires PulL-PulM, la présence de PulA a affecté de manière significative l'interaction PulL - PulM dans une combinaison de constructions, mais pas dans l'autre. La présence de PulA réduit l'interaction de T18-PulM avec T25-PulL. Cela pourrait signifier que PulA induirait la dissociation du complexe en se liant à l'un des partenaires. Jusqu'à présent, les études BACTH et biochimiques indiquent que ce partenaire pourrait être PulM.

Nous avons en outre vérifié par Western blot que PulA était produite dans les échantillons BACTH dans les conditions de dosage. Bien que l'effet de PulA n'ait été observé que dans l'une des paires PulL-PulM, ce résultat suggère que PulA pourrait dissocier le complexe PulL-PulM au niveau de leur site d'interaction dans le SST2. Cette possibilité devrait être étudiée plus avant *in vitro* et *in vivo*.

## Discussion and Perspectives

### Approches de pontage pour identifier les partenaires d'interaction de PulA

Les systèmes de sécrétion de type II sont des nanomachines bactériennes complexes intégrées dans l'enveloppe de bactéries Gram-négatives. Le système Pul de *K. oxytoca* étudié dans cette thèse a été découvert dans les années 80. Depuis lors, le SST2 a été trouvé chez de nombreuses espèces dans plusieurs classes bactériennes. La capacité de ce système à sécréter des protéines repliées peut avoir joué un rôle chez les ancêtres bactériens des mitochondries et pourrait avoir contribué à l'évolution des cellules eucaryotes. Les protéines repliées sont stabilisées par des ponts disulfure, du calcium ou d'autres cofacteurs et sont sécrétées par SST2 prêtes à être actives et à résister aux conditions environnementales et à la protéolyse. Le mécanisme de reconnaissance du substrat reste la principale question ouverte dans le domaine du SST2, avec la question du mécanisme de sécrétion.

Dans cette thèse, nous avons tenté de répondre à ces questions en nous concentrant sur les interactions de PulA avec le SST2. L'objectif principal de mon travail était de mettre en place la technique de réticulation UV du *pBpa* au laboratoire, pour répondre à cette question. J'ai également étendu l'utilisation du système double hybride bactérien (BACTH) pour identifier les interactions PulA avec les composants du SST2. Des études d'interaction antérieures en laboratoire ont montré que PulG interagit avec PulF, PulG, PulH, PulI, PulJ, PulK, PulL et PulM.

Auparavant, des résultats non publiés d'études de co-purification dans le laboratoire hôte ont identifié plusieurs composants candidats SST2 qui pourraient interagir avec PulA. Ces protéines résident soit dans la membrane interne (PulM, PulC), soit forment le pseudopilus (PulG). Sur la base de ces résultats, nous avons utilisé le test de photo-pontage *in vivo*, en mettant en œuvre l'approche de suppression du non-sens sur les protéines modèle PulG et PulA. Nous espérons capturer certaines interactions protéine-protéine transitoires en pontant des résidus en contact étroit avec un résidu *pBpa*. En utilisant des dérivés de PulG-*pBpa*, nous nous attendions à identifier des partenaires PulG connus tels que PulM et des pseudopilines mineures, et potentiellement PulA lui-même.

Cependant, dans toutes les tentatives décrites jusqu'à présent, nous n'avons pas réussi à identifier les partenaires d'interaction de PulA en utilisant l'approche de pontage *pBpa*-UV. Bien que nous ayons pu générer plusieurs dérivés pontés de certains variants PulA-*pBpa*, nous n'avons pas pu compléter leur caractérisation par Western blot ou par spectrométrie de masse.

Bien que nos résultats aient été négatifs jusqu'à présent concernant PulA, nous avons identifié des moyens possibles d'améliorer l'application de cette méthode sur notre système et cette approche qui méritent d'être testés.

Nous étions confrontés à plusieurs limites de l'approche *pBpa*. La nature ciblée de cette approche entraîne ses limites : malgré les multiples positions de *pBpa* que nous avons testées, il est possible que nous ayons manqué les sites optimaux pour capturer les interactions par un test de pontage site-spécifique. De plus, le processus de sécrétion hautement dynamique pourrait limiter la fraction de PulA qui interagit avec un composant SST2 donné. Cela pourrait entraîner de très faibles niveaux dérivés pontés, qui seraient inférieurs à la limite de détection par Western blot avec les anticorps dont nous disposons. Bien que nous ayons tenté d'enrichir et de concentrer les dérivés marqués de PulA-Strep, les quantités de protéines étaient trop faibles pour une analyse par spectrométrie de masse. Dans le cas où PulA était ponté à une autre protéine cellulaire, la spectrométrie de masse aurait pu nous permettre d'identifier ce complexe. Cependant, l'analyse des peptides générés par spectrométrie de masse est encore compliquée par la présence de *pBpa* dans les protéines. Dans notre essai initial, nous n'avons pas réussi à identifier les peptides contenant des résidus de *pBpa* par spectrométrie de masse. Contrairement au *pBpa*, plus récemment, d'autres acides aminés non naturels ont été développés qui sont mieux adaptés à l'analyse par MS, mais ils ne sont pas disponibles dans le commerce.

Un autre problème auquel nous avons été confrontés était la faible efficacité de pontage. Initialement nous cultivions nos bactéries en milieu minimal, mais les précultures en LB proposées par le prof. Shevchik ont semblé améliorer le rendement d'incorporation de *pBpa* et le pontage ultérieur. Il serait intéressant de réanalyser les variants PulG-*pBpa* avec ce protocole modifié, car il était surprenant qu'avec cette petite protéine qui interagit avec de nombreux composants du SST2, nous n'ayons identifié que des homodimères. La raison en est peut-être aussi l'abondance relative des composants. PulG et PulA sont les substrats les plus abondants du système et toutes les autres protéines sont moins abondantes. Par conséquent, la proportion de PulA en contact avec la protéine partenaire serait très faible et peut-être non détectable par Western blot. Pourtant, la sensibilité de la spectrométrie de masse devrait nous permettre de détecter de telles protéines après copurification et extraction de bande. Cette approche devrait être poursuivie à l'avenir. En outre, il peut être préférable de modifier les composants moins abondants avec du *pBpa* et d'augmenter les chances de détecter les dérivés pontés avec des substrats plus abondants.

Nous avons également remarqué que la sécrétion de PulA commence dès que le complexe SST2 est formé dans l'enveloppe d'*E. coli*. Même au début de la croissance bactérienne, la majeure partie du PulA est sécrétée dans le milieu avant l'accumulation d'une masse bactérienne suffisante et l'accumulation de composants SST2. Cela se traduit par une plus faible abondance de facteurs Pul et par conséquent une plus faible abondance de complexes PulA.

Pour surmonter ces problèmes, nous avons effectué plusieurs approches complémentaires, y compris le pontage au formaldéhyde et co-purification avec PulA-Strep en plus de double hybride bactérien.

Contrairement au pontage ciblé avec *pBpa*, l'approche SPINE a le potentiel de capturer toutes les interactions possibles de PulA avec le complexe SST2. J'ai effectué des expériences préliminaires visant à mettre en place SPINE pour le SST2 Pul. Le principal problème que nous avons observé ici était l'absence de PulA dans la fraction cellulaire des bactéries contenant le système Pul fonctionnel. Ce phénomène pourrait être causé par deux raisons principales : premièrement, même dans les premiers stades de croissance, la majeure partie de la PulA était sécrétée dans le surnageant de culture et perdue. En utilisant un mutant *ApulE*, nous avons augmenté le rendement de PulA intracellulaire et nous espérons pouvoir encore capturer les intermédiaires de sécrétion précoces. Nous avons supposé que PulA devrait lier au moins un facteur Pul avant les étapes catalysées par l'ATPase PulE. Le fait que nous n'ayons pas détecté de complexes avec les facteurs Pul pourrait être dû aux faibles niveaux de ces complexes. Deuxièmement, la présence de PulA dans la fraction membranaire suggère que ce PulA pourrait être lié à certaines protéines membranaires. Étant donné que la majorité des composants Pul se trouvent dans la membrane interne ou externe, les dérivés pontés d'intérêt devraient se trouver dans cette fraction. Pour capturer de tels complexes, l'extrait cellulaire doit être traité avec un détergent pour solubiliser les composants SST2 liés à la membrane et la purification doit être effectuée en présence de détergent. La mise en garde dans ce cas est que nous ne serions probablement pas en mesure de solubiliser la sécrétine PulD, qui est très stable et résistante aux détergents. En outre, la perte de la protéine pontée pourrait être due à la limitation de l'approche elle-même. Dans une étude sur l'utilité du pontage par formaldéhyde et de la spectrométrie de masse, les auteurs ont identifié une grande perte (90%) du complexe ponté, ainsi que des protéines globales de la bactérie lors du traitement de la culture par 1% de formaldéhyde pendant 30 min, qui est exactement ma condition d'incubation. La perte significative de toutes les protéines était probablement due à la lyse de la cellule par le traitement au formaldéhyde. Si tel est le cas, cela signifie que nous avons peut-être perdu la plupart des PulA ainsi que ses

complexes avec les protéines Pul dans le surnageant puisque nous centrifugions toute la culture après pontage au formaldéhyde et éliminons tout le surnageant. Pour améliorer le rendement, nous aurions besoin d'optimiser les conditions expérimentales et d'utiliser 0,3 % à 0,5 % de formaldéhyde et des temps d'incubation plus courts ou une température plus basse. De plus, une étape de "trempage à la glycine" pourrait être ajoutée après l'incubation au formaldéhyde pour éviter une réaction excessive.

Compte tenu de tous les problèmes auxquels nous avons été confrontés lorsque nous avons tenté de capturer les interactions transitoires dans le processus de sécrétion dynamique *in vivo*, la première et principale question est la conception d'une bonne expérience pour surveiller et générer une courbe de quantification du processus de sécrétion, pour localiser la meilleure fenêtre avec un bon équilibre de la quantité de PulA et d'autres protéines Pul présentes dans la cellule, avant d'effectuer les essais de pontage *in vivo*. S'il n'existe pas un tel moment, la prochaine possibilité est de trouver un moyen de ralentir le processus. Pour ce faire, nous pourrions utiliser des mutants de PulE partiellement fonctionnels et présentant une activité ATPase réduite. Il serait également intéressant d'utiliser des traitements chimiques qui ralentissent ou bloquent la sécrétion par exemple en bloquant la force protomotrice, essentielle à la sécrétion de PulA, par CCCP ou d'autres ionophores.

#### Etudes BACTH des interactions PulA avec le SST2

Comme approche complémentaire, nous avons utilisé l'analyse BACTH pour identifier les composants du SST2 qui interagissent avec PulA au niveau de la membrane interne. Bien que les interactions du substrat avec les composants du SST2 soient souvent considérées comme transitoires, c'est probablement le cas dans le système actif qui est alimenté par l'activité ATPase. Cependant, il est possible que des intermédiaires stables soient formés en l'absence d'ATPase. Si tel est le cas, nous devrions être en mesure d'identifier ces intermédiaires précoces par le test BACTH. Comme décrit dans ce travail, les tests BACTH ont identifié plusieurs candidats forts pour les partenaires d'interaction PulA, notamment la pseudopiline majeure PulG et les composants de l'AP PulM et PulF. De plus, d'autres pseudopilines ont également montré une tendance à se lier à PulA, en fonction du nombre de copies de plasmides codant pour les hybrides Cya et de la stabilité des protéines. Pour tester si ces interactions reflètent la situation dans le SST2 fonctionnel, les résultats doivent être corroborés par des approches complémentaires, par ex. biochimiques - copurification, études de liaison ou par microscopie - y compris la microscopie électronique et la microscopie à fluorescence. Notre collaboratrice N. Izadi-Pruneyere et son équipe ont pu montrer la liaison de PulA aux pseudopili en utilisant des

techniques biophysiques – SPR, l'interférométrie de la biocouche et la co-sédimentation. En collaboration au sein d'un réseau de formation ITN, des approches de pontage chimique et spectrométrie de masse sont employées pour caractériser les contacts entre PulA et les pseudopili.

La preuve d'une interaction de PulA avec PulM a également été obtenue par Aracelys Lopez Castilla dans l'équipe de N. Izadi-Pruneyre en utilisant la résonance plasmonique de surface. Ce résultat est en accord avec les données chez *Pseudomonas*, où des interactions du substrat avec XcpL et XcpM ont été observées. Comme discuté ci-dessus, PulM pourrait jouer un rôle dans le couplage de sécrétion de PulA avec l'assemblage de pseudopilus. PulM interagit également avec PulL et avec les pseudopilines PulG et PulH. Nous pouvons proposer que le complexe PulM-PulG soit adressé à PulL pour former un complexe tripartite. Il serait intéressant de tester la liaison de PulA et de comparer ses affinités de liaison au PulG, PulM, et aux complexes PulG-PulM et PulL-PulM. Cela pourrait fournir des informations sur le mécanisme possible de recrutement et de transfert du substrat sur le site d'assemblage du pseudopilus.

Selon les résultats de BACTH, nous avons pu voir une interaction inattendue et forte de PulA avec PulF. Les insertions du peptide Flag aux positions L198 et A201 de la boucle périplasmique de PulF ont aboli cette interaction et affecté spécifiquement la sécrétion de PulA, mais pas l'assemblage du pseudopilus. Ces résultats soutiennent un rôle direct de PulF dans la sécrétion, mais nécessitent une analyse plus approfondie. Dans le futur, il serait intéressant d'étudier les contacts PulA-PulF et d'identifier plus précisément l'interface, éventuellement en utilisant le pontage UV après incorporation de *pBpa* dans les régions périplasmiques de PulF. La mutagenèse dirigée de cette région de PulF pourrait également apporter des informations sur sa fonction.

#### Le rôle du complexe de plateforme d'assemblage PulL-PulM

Une partie importante de notre travail s'est concentrée sur les composants AP PulL et PulM. PulM est nécessaire pour la stabilité de PulL, et PulL est essentiel pour l'assemblage du pseudopilus et la sécrétion de PulA, ce qui est probablement dû à son rôle dans l'ancrage de l'ATPase PulE à la base du complexe. Il a été proposé que GspL et GspM signalent la présence du substrat à travers l'IM à l'ATPase GspE. Cependant, l'absence du substrat dans le Pul SST2 n'empêche pas l'assemblage du pseudopilus, ce qui va à l'encontre du modèle de signalisation. Le Pul SST2 dépourvu de PulM est partiellement fonctionnel, car un très petit nombre de

pseudopili longs peuvent être assemblés et une petite quantité de PulA est sécrétée. Une analyse biophysique et structurale approfondie du complexe PulL<sup>CTD</sup>-PulM<sup>CTD</sup> réalisée par le groupe de N. Izadi-Pruneyre a fourni un nouvel aperçu de ces domaines et de leur interface. Malgré de nombreuses études structurales, la structure d'un hétérodimère L-M n'avait pas été déterminée auparavant et ne pouvait être que modélisée. Ici, Dazzoni et al. ont combiné RMN et cristallographie pour caractériser la liaison de PulL à PulM pour obtenir des cristaux avec des hétérodimères parallèles, qui sont très susceptibles de refléter l'état natif de ce complexe. Nos données BACTH soutiennent cet état d'hétérodimère parallèle et ne fournissent aucune preuve d'homodimérisation. Le pontage disulfure des segments transmembranaires de PulL et PulM nous a permis d'identifier des résidus spécifiques dans le TMS de PulL et PulM qui font des contacts et pontages spécifiques dans des conditions oxydantes. Définir ces contacts a été utile pour construire le modèle final du complexe L-M comprenant leurs domaines membranaires et cytoplasmiques. Les linkers flexibles, qui relient les domaines de type ferrédoxine de L et M à leurs hélices périplasmiques restent difficiles à prédire et pourraient fournir une certaine liberté au mouvement des CTD. Compte tenu de l'équilibre monomère-dimère en solution, il est possible que les CTD *in vivo* se lient également les uns aux autres et se dissocient à différentes étapes du processus d'assemblage du pseudopilus et de sécrétion. Il serait important de comprendre si et comment les substrats – PulA et PulG – ou d'autres facteurs influencent cette dynamique.

La purification du sous-complexe SST2 de *Klebsiella* contenant la sécrétine et des parties de la plate-forme d'assemblage indique la stoechiométrie 1: 1 de PulL et PulM. Cependant, nous avons remarqué que PulM était beaucoup plus abondant que PulL chez les bactéries produisant le système SST2 à partir du locus *pul*. Grâce aux PulM<sup>CTD</sup> et PulL<sup>CTD</sup> purifiés qui sont reconnus par les anticorps PulL et PulM disponibles, nous avons pu déterminer le rapport molaire de PulM et PulL dans les bactéries et constater que PulM est 20 fois plus abondant. Cela concorde avec le modèle d'un rôle dynamique de PulM, qui devrait faire la navette entre le pool membranaire de PulG et le PulL associé à la plate-forme d'assemblage. Pour tester ce modèle, il serait intéressant d'effectuer des études de localisation par microscopie à fluorescence sur des bactéries vivantes et des expériences FRAP avec PulL et PulM fusionnés à la GFP ou à d'autres sondes fluorescentes.

Les mêmes régions de PulL<sup>CTD</sup> et PulM<sup>CTD</sup> sont impliquées dans la formation d'homo- et d'hétérodimères en solution. Cette interface est hydrophobe et force probablement la dimérisation afin d'assurer la stabilité du complexe en solution. Dans les études de

spectrométrie de masse native, PulL et PulM n'ont formé que des dimères, mais pas des multimères d'ordre supérieur, ce qui correspond à la présence d'un seul déterminant de liaison dans chaque protéine. Cependant, *in vivo* dans le contexte des protéines entières, la conformation des domaines C-terminaux serait restreinte par la présence d'autres domaines/régions et par leur conformation enchâssée dans la membrane. L'étude décrite dans le deuxième manuscrit, Li et al., suggère que lorsque les CTD de L et M sont connectés aux segments transmembranaires directement *via* leurs linkers flexibles natifs, leur interaction serait plutôt faible, comme indiqué dans le système BACTH. Cela suggère que la bobine enroulée formée par les hélices périplasmiques de PulL et PulM joue un rôle dominant dans la formation d'hétérodimères. Comme discuté dans ce manuscrit, il est également possible que cette bobine enroulée assure une distance entre les CTD et la membrane et empêche une interaction de ces domaines avec la surface de la membrane. De plus, les régions périplasmiques pourraient fournir de la rigidité au complexe et pourrait également lui permettre d'interagir avec d'autres protéines ou avec le peptidoglycane, bien qu'il n'y ait jusqu'à présent aucune évidence pour ces possibilités. Dans l'STT2 Out chez *D. dadantii*, les protéines L et M interagissent avec OutC, mais lorsque nous avons utilisé l'approche BACTH pour tester ces interactions dans le système Pul, nous n'avons détecté aucun signal d'interaction.

Nous avons utilisé la mutagenèse dirigée pour cibler l'interface PulL<sup>CTD</sup>-PulM<sup>CTD</sup>. Nous avons constaté qu'il était difficile de perturber cette interface sans affecter la stabilité des CTD. Initialement, nous avons utilisé des inversions de charge, mais elles ont fortement affecté la stabilité de CTD, et nous avons dû introduire des simples et doubles substitutions d'alanine pour minimiser cet effet. Ces mutations ont produit un effet léger mais reproductible et significatif sur la sécrétion de PulA. On ne sait pas si le complexe PulL<sup>CTD</sup>-PulM<sup>CTD</sup> se lie à PulA, comme c'est le cas dans le système *Pseudomonas*. Alors que dans cette étude, l'ajout de substrat CbpD a stimulé l'interaction de L et M dans le test BACTH, nous avons observé un effet opposé dans une expérience similaire. Dans cette expérience, PulA a montré une tendance à réduire l'interaction PulL-PulM, ce qui pourrait indiquer que la liaison de PulA à l'un des partenaires d'interaction entre en compétition plutôt qu'elle ne stimule l'interaction des protéines hybrides T18/T25-PulL/PulM.

Nous avons constaté que la délétion du peptide cytoplasmique de PulM affectait la stabilité de PulL et générerait un fragment de PulL équivalent à son orthologue dans T4P, PilN. Ce fragment, qui reste à caractériser par séquençage N-terminal ou spectrométrie de masse, comprend probablement les domaines membranaire et périplasmique de PulL. Fait intéressant,

ce fragment pourrait également être détecté en présence du SST2 complet (dans des expériences qui ont servi à quantifier le rapport PulL/PulM), ce qui suggère que le site sensible aux protéases dans PulL est exposé pendant la sécrétion. Ce fragment a été généré après délétion du peptide cytoplasmique de PulM beaucoup plus fréquemment qu'en l'absence de PulM ou en présence de PulM natif. Ceci est déroutant et pourrait suggérer que les interactions Pull-PulM dans le périplasme favorisent les changements conformationnels de PulL qui exposent la région sensible aux protéases de PulL. Des études *in vitro* pourraient être utilisées pour étudier la liaison du peptide PulM au domaine cytoplasmique PulL. Dans T4P, l'orthologue du domaine PulLcyto, PilM a été co-cristallisé avec le peptide N-terminal de PilN, qui pourrait être l'orthologue de PulM dans le SST2.

Globalement, les études structurales, biochimiques et biophysiques de PulL et PulM ont ouvert de nouvelles voies de recherche. Bien qu'il reste difficile à étudier la fonction des protéines dynamiques dans la membrane, la combinaison d'approches complémentaires *in vitro* et *in vivo* pourrait fournir des informations sur leur rôle dans les mécanismes d'assemblage de PulG et de sécrétion de PulA.