STRUCTURE-FUNCTION RELATIONSHIPS IN EXPANDED TYPE IV SECRETION SYSTEMS AND THE CONJUGATIVE F-PILUS

Thesis by

JONASZ B. PATKOWSKI

in fulfilment of the degree of

DOCTOR OF PHILOSOPHY (PhD)

CENTRE FOR BACTERIAL RESISTANCE BIOLOGY

DEPARTMENT OF LIFE SCIENCES

IMPERIAL COLLEGE LONDON

APRIL 2024

Copyright

The copyright of this thesis rests with the author.

Unless otherwise indicated, its contents are licensed under the

Creative Commons Attribution-Non-Commercial 4.0 International Licence (CC BY-NC).

Under this licence, you may copy and redistribute the material in any medium or format.

You may also create and distribute modified versions of the work. This is on the condition that: you credit the author and do not use it, or any derivative works, for a commercial purpose.

When reusing or sharing this work, ensure you make the licence terms clear to others by naming the licence and linking to the licence text.

Where a work has been adapted, you should indicate that the work has been changed and describe those changes.

Please seek permission from the copyright holder for uses of this work that are not included in this licence or permitted under UK Copyright Law.

Originality

I declare that the work presented in this thesis is my own,
except where otherwise stated.

All work here was conducted under the supervision of Dr Tiago Costa.

All cited work is appropriately referenced, and any figure reproduction was done so with permission and is properly annotated.

© Jonasz B. Patkowski 2024
ORCID: 0000-0001-7301-1569

Rodzicom,
za wszystko.

Acknowledgements

I would like to express my gratitude to Tiago Costa, for your mentorship and the freedom you gave me to explore the incredible world of bacterial secretion systems, the enthusiasm and passion you always had during our discussions and for the hours spent together on reiterating the hundredth (and counting) version of our manuscripts. Thank you to the fantastic Crumble Gang: Himani Amin and Ambre Bexter. I will be forever grateful for the way we always supported each other throughout the days, nights, and weekends in the lab together.

Thank you to Morgan Beeby and Thomas Meier, my progress review panel, and to Alain Filloux, Aravindan Ilangovan, Wael Elhenawy, José Penadés, Alfred Fillol, Lingchen He, Gerald Larrouy-Maumus, Leticia Beltran, Edward Egelman, Joseph Baker, Magnus Andersson, Dharmender Kumar, Joseph Vogel, and Luke Allsopp for fruitful collaborations. To Paul Simpson, for training me on electron microscopes.

To my Warsaw family: Alicja, Patrycja, and my best friend, Alicja, thank you for all your encouragement, friendship, advice, all the calls that kept me sane throughout this madness. You inspired me to keep going no matter how hard it was.

Rodzinie - Mamie, Tacie i moim Braciom - którzy zawsze pielęgnowali moją ciekawość, gdy dorastałem, i pomogli przekształcić ją w żarzące się pragnienie wiedzy, które mam teraz. Dziękuję za wasze przewodnictwo, które poprowadziło mnie do moich marzeń, i za to, że pozwoliliście mi spakować walizkę i wyjechać, aby je realizować. Przepraszam, że nie mogę być z wami częściej. Moim ukochanym dziadkom: Babci Irence, Babci Jadzi, Dziadkowi Ryśkowi i Dziadkowi Waldkowi; to osiągnięcie jest świadectwem wartości, które we mnie zaszczepiliście.

And finally, to my dearest Jordi, thank you for being my calm harbour where I can always recharge and refocus, for your enthusiasm that charted me out of moments of doubt, and for the weeks together in the beautiful Colònia de Sant Pere, where this thesis bloomed. I would not have done this without you.

Abstract

Antimicrobial resistance (AMR), a major threat to public health, is mediated by bacterial conjugation – a process whereby bacteria exchange antibiotic resistance genes. On the molecular level, it relies on the Type IV Secretion Systems (T4SS), a protein nanomachine localised in bacterial membranes, responsible for two operations: processing of the substrate DNA in the donor cell and its subsequent delivery into the neighbouring recipient cell. Despite the clinical relevance of bacterial conjugation, the details of how T4SS recognises DNA, energises its translocation, and directs it to the recipient cell, are unknown. This thesis aims to provide insight into the sequence of events that occur between DNA processing and its release in the recipient cell during bacterial conjugation.

Here, I unveil the structure of TraD, a T4SS-associated DNA receptor, and document its ability to bind DNA and form oligomers upon ATP hydrolysis, suggesting an FtsK-like mechanism of DNA transfer. I demonstrate the biomechanical properties of the F-pilus - an extracellular filament responsible for creating conjugative contacts - that enhance conjugation efficiency in severe mechanical and thermochemical environments, and attribute these properties to the presence of phospholipid molecules in the F-pilus atomic structure, thereby documenting their structural role for the first time. Furthermore, I generate a fluorescent variant of the F-pilus that allowed for the visualisation of DNA journey through its lumen – the first direct evidence of this phenomenon. Lastly, I document that T4SS also play a role in the colonisation of epithelial cells by adherent-invasive *E. coli* during Leśniowski-Crohn's disease development.

Taken together, my findings present novel structure-function relationships between proteins that establish T4SS, providing additional detail to our understanding of the molecular mechanisms of bacterial conjugation and the role of T4SS as a virulence factor during infection.

Published works related to this study

Costa, T. R. D., **Patkowski, J. B**., Macé, K., Christie, P. J. and Waksman, G. (2023). "Structural and functional diversity of type IV secretion systems." Nature Reviews Microbiology.

Patkowski, J. B., Dahlberg, T., Amin, H., Gahlot, D. K., Vijayrajratnam, S., Vogel, J. P., Francis, M. S., Baker, J. L., Andersson, M. and Costa, T. R. D. (2023). "The F-pilus biomechanical adaptability accelerates conjugative dissemination of antimicrobial resistance and biofilm formation." Nature Communications 14(1): 1879.

Beltrán, L., Torsilieri, H., **Patkowski, J. B**., Yang, J. E., Casanova, J., Costa, T. R. D., Wright, E. R. and Egelman, E. H. (2023). "The mating pilus of *E. coli* pED208 acts as a conduit for ssDNA during horizontal gene transfer." mBio 15(1): e02857-02823.

Beltran, L. C., Cvirkaite-Krupovic, V., Miller, J., Wang, F., Kreutzberger, M. A. B., **Patkowski, J. B.**, Costa, T. R. D., Schouten, S., Levental, I., Conticello, V. P., Egelman, E. H. and Krupovic, M. (2023). "Archaeal DNA-import apparatus is homologous to bacterial conjugation machinery." Nature Communications 14(1): 666.

List of Contents

Acknowledgements **4**

Abstract **5**

Published works related to this study **6**

List of Contents **7**

List of Figures **10**

List of Tables **12**

Abbreviations **13**

Chapter 1 General Introduction **15**

1.1 Antimicrobial resistance (AMR) 15
1.1.1 Antimicrobial resistance is a public health threat *15*
1.1.2 Sources and mobility of AMR *16*

1.2 Bacterial Conjugation 18
1.2.1 Incompatibility of conjugative plasmids and the IncF plasmids *21*
1.2.2 The F-plasmid *22*
1.2.3 Mechanism of conjugation: mating pair formation *25*
1.2.4 Mechanism of conjugation: DNA processing and transfer *28*

1.3 Type IV Secretion Systems (T4SS) 32
1.3.1 T4SS diversity *32*
1.3.2 Architectures of the minimised T4SS *35*
1.3.3 Architectures of the expanded T4SS *40*
1.3.4 The F-pilus biogenesis *47*
1.3.5 The F-pilus structure *49*
1.3.6 The Type IV Coupling Proteins (T4CP) *54*

1.4 Cryo-electron microscopy (cryo-EM) 60
1.4.1 Principles of cryo-EM *60*
1.4.2 Image generation *61*
1.4.3 Sample preparation *64*
1.4.4 Data processing *66*
1.4.5 Model building and validation *70*

1.5 Aims of this thesis 72

Chapter 2 Structural and biochemical analysis of the F-plasmid T4CP, TraD **74**

2.1 Introduction 74

2.2 Experimental procedures 76
2.2.1 Materials *76*
2.2.2 Purification of $TraD^{130-717}$ and Hcp-$TraD^{130-717}$ *76*
2.2.3 Purification of TraM *78*
2.2.4 $TraD^{130-717}$ ATPase assay *78*
2.2.5 DNA binding assay *79*
2.2.6 Trypsin digestion assay *79*
2.2.7 TraD-TraM pulldown procedure *79*
2.2.8 Chemical crosslinking of $TraD^{130-717}$ *80*
2.2.9 Negative stain EM (NS-EM) procedures *80*
2.2.10 Cryo-EM procedures *80*

2.3 Results 81
2.3.1 TraD does not form hexamers in solution *81*

2.3.2 TraD exhibits ATPase activity ... *83*
2.3.3 TraD interacts with DNA non-specifically ... *85*
2.3.4 TraD interacts with TraM in vitro ... *85*
2.3.5 Structure of TraD ... *87*
2.3.6 The lumen of TraD is fit for interaction with DNA ... *90*
2.3.7 Interaction interface between adjacent TraD subunits ... *92*
2.3.8 Nucleotide-binding domain of TraD ... *93*
2.3.9 ATP induces oligomerisation of TraD ... *97*
2.3.10 Cryo-EM of TraD$^{130-717}$ reveals dimer and trimer formation ... *100*
2.3.11 Scaffold protein is insufficient to induce hexamerisation of TraD ... *102*

2.4 Discussion ... 106

Chapter 3 Biomechanics of the F-pilus and the role of its phospholipids in conjugation ... 111

3.1 Introduction ... 111

3.2 Experimental procedures ... 113
3.2.1 Materials ... *113*
3.2.2 Conjugation efficiency assay ... *114*
3.2.3 Biofilm formation assay ... *115*
3.2.4 Purification of F-pili ... *116*
3.2.5 Microscopy procedures ... *117*
3.2.6 Cryo-EM imaging of F-pili ... *117*
3.2.7 Thermochemical stability assessment of F-pili ... *117*
3.2.8 Bacterial strain and growth conditions for force measurements ... *118*
3.2.9 Force measurements and sample preparation ... *118*
3.2.10 Molecular dynamics simulations ... *119*
3.2.11 f1 phage infection ... *121*

3.3 Results ... 121
3.3.1 Conjugative transfer is enhanced under hydrodynamic stress ... *121*
3.3.2 f1 phage infection is also enhanced in hydrodynamic conditions ... *124*
3.3.3 F-pilus accelerates biofilm formation ... *125*
3.3.4 The F-pilus is robust, flexible, and thermostable ... *127*
3.3.5 The F-pilus is an elastic filament ... *129*
3.3.6 Presence of phospholipids confers structural robustness to the F-pilus ... *132*

3.4 Discussion ... 134

Chapter 4 Direct evidence of ssDNA transport through the F-pilus ... 137

4.1 Introduction ... 137

4.2 Experimental procedures ... 139
4.2.1 Materials ... *139*
4.2.2 Generation of F-pili (pED208) for maleimide staining ... *140*
4.2.3 Complementation of cysteine-derivative of F-pilus ... *140*
4.2.4 Cell growth and Alexa Fluor 568 labelling ... *141*
4.2.5 Fluorescence microscopy ... *141*

4.3 Results ... 142
4.3.1 Generation of F-pili for maleimide staining ... *142*
4.3.2 Co-localisation of F-pili with the travelling ssDNA ... *144*

4.4 Discussion ... 146

Chapter 5 Identification of a novel T4SS, essential for AIEC colonisation in L-CD ... 148

5.1 Introduction ... 148

5.2 Experimental Procedures ... 151

5.2.1 Materials ... *151*
5.2.2 Genomic search for T4SS-associated loci in NRG857c ... *152*
5.2.3 Classification of the genomic T4SS locus ... *153*
5.2.4 Protein structure prediction of the OMCC components ... *153*
5.2.5 Generation of pASK_pO83-OMCC and pBAD_ICE-OMCC ... *153*
5.2.6 Purification of the pO83- and ICE-encoded OMCC ... *154*
5.2.7 Western Blotting ... *156*
5.2.8 NS-EM imaging ... *156*
5.2.9 Imaging of NRG857c infection on Caco-2 cells ... *156*
5.3 Results ... 157
5.3.1 The pO83-T4SS does not form a typical OMCC ... *157*
5.3.2 Discovery of a novel T4SS in NRG857c AIEC ... *161*
5.3.3 Comparison of OMCC subunits between pO83-T4SS and ICE-T4SS ... *165*
5.3.4 The ICE-T4SS forms a typical OMCC ... *168*
5.3.5 The novel ICE-T4SS is essential for biofilm formation ... *170*
5.4 Discussion ... 171
Chapter 6 Concluding remarks and outlook ... 175
6.1 Overview of the main findings ... 175
6.2 Future directions ... 177
References ... 180
Supplementary Figures ... 202
Supplementary Tables ... 204

List of Figures

Figure 1. Routes of horizontal gene transfer between bacteria 17
Figure 2. *E. coli* during bacterial conjugation 19
Figure 3. Genetic organisation of the F-plasmid 24
Figure 4. Two possible routes of bacterial conjugation 26
Figure 5. The mechanism of bacterial conjugation 30
Figure 6. Diversity of T4SS 33
Figure 7. Cryo-EM structure of a minimised T4SS encoded on the R388 plasmid 37
Figure 8. Cryo-EM structures of OMCC and cryo-ET maps of the expanded T4SS 42
Figure 9. Structural transitions of F-T4SS during different stages of bacterial conjugation 44
Figure 10. The F-pilus biogenesis 48
Figure 11. Comparison of conjugative pili species between minimised and expanded T4SS 51
Figure 12. Structures of type IV coupling proteins 56
Figure 13. Principles of electron microscopy 62
Figure 14. Typical cryo-EM workflow overview 65
Figure 15. Motion correction of a cryo-EM dataset 67
Figure 16. Principle of CTF correction 68
Figure 17. 3D reconstruction based on 2D classes 69
Figure 18. Building of the atomic model and its evaluation 71
Figure 19. Purification of TraD$^{130\text{-}717}$ 82
Figure 20. NS-EM micrograph of TraD$^{130\text{-}717}$ 83
Figure 21. Biochemical analysis of TraD$^{130\text{-}717}$ 84
Figure 22. TraD$^{130\text{-}717}$-TraM double pull-down assay 86
Figure 23. AlphaFold-2 prediction model of TraD structure 89
Figure 24. Positive charge of TraD lumen suggests involvement of lysine residues in DNA transfer 91
Figure 25. The interaction interface between TraD subunits 92
Figure 26. Comparison of the model of TraD with TrwB and Cagβ 94
Figure 27. The nucleotide-binding site of TraD 96
Figure 28. Glutaraldehyde crosslinking reaction between two proteins 97
Figure 29. Glutaraldehyde crosslinking assay of TraD 98
Figure 30. Large-scale crosslinking of TraD$^{130\text{-}717}$ 100
Figure 31. Cryo-EM of crosslinked TraD$^{130\text{-}717}$ 101
Figure 32. Principle of using Hcp as scaffold 103
Figure 33. Purification of Hcp-TraD$^{130\text{-}717}$ 104
Figure 34. Cryo-EM map of Hcp-TraD$^{130\text{-}717}$ 105
Figure 35. Functional transitions of TraD during bacterial conjugation 110
Figure 36. Effect of a dynamic environment on bacterial conjugation 123
Figure 37. Effect of a dynamic environment on biofilm formation 126
Figure 38. The F-pilus exhibits visible flexibility 128

Figure 39. Testing thermochemical limits of the F-pilus 129
Figure 40. Force extension measurements of the F-pilus 131
Figure 41. Steered molecular dynamic simulations of the F-pilus 133
Figure 42. Equilibrium simulations of the F-pilus 134
Figure 43. Hydrodynamic model of conjugation and biofilm formation 136
Figure 44. Construction of an F-pilus suitable for maleimide labelling 143
Figure 45. F-pilus mediates distant transfer during conjugation 145
Figure 46. Homology between pO83-T4SS and the F-T4SS 158
Figure 47. pO83-OMCC does not assemble large rings 160
Figure 48. NS-EM imaging of the pO83-OMCC SEC fractions via NS-EM 161
Figure 49. Organisation and classification of ICE-T4SS 164
Figure 50. Comparison of OMCC between ICE-T4SS, pO83-T4SS and F-T4SS 167
Figure 51. The ICE-OMCC forms typical large ring structures 169
Figure 52. The ICE-T4SS is involved in biofilm formation on Caco-2 cells 171

List of Tables

Table 1. Bacterial strains 76
Table 2. Plasmids and genetic constructs used in this study 76
Table 3. Bacterial strains 113
Table 4. Plasmids and genetic constructs used in this study 114
Table 5. Bacterial strains 139
Table 6. Plasmids and genetic constructs used in this study 139
Table 7. Primers used in this study 139
Table 8. Bacterial strains 151
Table 9. Plasmids and genetic constructs used in this study 151
Table 10. Primers used in this study 152
Table 11. Identification of a chromosomal T4SS in AIEC strain NRG857c 163

Abbreviations

AAA+	ATPases Associated with Diverse Cellular Activities
AAD	All-alpha Domain
AIEC	Adherent-Invasive *Escherichia coli*
AMR	Antimicrobial Resistance
ARG	Antibiotic Resistance Gene
CFU	Colony-Forming Unit
Cryo-EM	Cryo-Electron Microscopy
Cryo-ET	Cryo-Electron Tomography
CTD	C-terminal Domain
CTF	Contrast Transfer Function
DDM	n-Dodecyl β-D-Maltoside
DED	Direct Electron Detector
dsDNA	Double-stranded DNA
EHEC	Enterohemorrhagic *Escherichia coli*
EMDB	Electron Microscopy Data Bank
GA	Glutaraldehyde
HGT	Horizontal Gene Transfer
IBD	Inflammatory Bowel Disease
ICE	Integrative and Conjugative Element
IEC	Intestinal Epithelial Cells
IMC	Inner Membrane Complex
Inc	Incompatibility Group
IPTG	Isopropyl-ß-D-1-thiogalactopyranoside
IR	Inner Ring
kbp	Kilobase Pair
L-CD	Leśniowski-Crohn's Disease
LB	Luria-Bertani
LPS	Lipopolysaccharide
MGE	Mobile Genetic Element
NBD	Nucleotide-Binding Domain
NHS	National Health Service
NS-EM	Negative-stain Electron Microscopy
OMCC	Outer Membrane Core Complex

OR	Outer Ring
PDB	Protein Data Bank
PFU	Plaque-forming Unit
PR	Periplasmic Ring
RCR	Rolling Circle Replication
RMSD	Root Mean Square Deviation
SDS-PAGE	Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis
SEC	Size-Exclusion Chromatography
SPA	Single Particle Analysis
ssDNA	Single-stranded DNA
STA	Subtomogram Averaging
T1SS-T6SS	Type I Secretion System - Type VI Secretion System
T4CP	Type IV Coupling Protein
TA	Toxin-Antitoxin
TBS-T	Tris-Buffered Saline with Tween
TIS	Transposon-Insertion Sequencing
UPEC	Uropathogenic *Escherichia coli*
WHO	World Health Organisation

Chapter 1 General Introduction

1.1 Antimicrobial resistance (AMR)

1.1.1 Antimicrobial resistance is a public health threat

The era of penicillin as the "wonder drug" has come to an end. Persistent and reckless overconsumption of antibiotics has triggered the emergence of bacteria that no longer respond to antimicrobial treatment. This phenomenon, known as antimicrobial resistance (AMR), has become one of the most pressing public health threats worldwide, resulting in 5 million deaths in 2019 alone (Murray *et al.* 2022). According to the United Nations, AMR infections are projected to surpass cancer as the leading cause of annual deaths by 2050 (UN 2023). Seven highly virulent biofilm-forming pathogens are responsible for 70% of drug-resistance-related fatalities. They are collectively acronymised as ESKAPEE pathogens: *Enterococcus faecium, Staphylococcus aureus, Klebsiella pneumoniae, Acinetobacter baumannii, Pseudomonas aeruginosa, Enterobacter spp*., and *Escherichia coli* (Kurbatfinski *et al.* 2023). ESKAPEE strains often exhibit multi-drug resistance and hence can "escape" antibiotic treatment.

In recognition of the severity of the AMR crisis, the World Health Organisation (WHO) published a list in 2017 that identifies the pathogens of greatest threat to humanity, with the highest priority given to *E. coli, K. pneumoniae* and *Enterobacter spp.* (WHO 2017). Subsequently, in 2021 the WHO launched the "Call to Action on Antimicrobial Resistance" as a desperate call for improved antimicrobial stewardship, increased financial support towards AMR actions, and intensified research on the AMR (Tacconelli *et al.* 2018, Mulani *et al.* 2019, WHO 2021).

Since then, numerous initiatives have been devised to address the development of antibiotic resistance, including the innovative "Subscription Model" introduced by the National Health Service (NHS) England in July 2023. Under this pioneering policy, the UK government will

procure antibiotics based on their value to the NHS rather than market demand (NHS 2023). Manufacturers will thus receive a fixed annual fee for their antibiotics, regardless of the volumes used, which incentivises the development of novel antibiotics. The recently formulated Cefiderocol and Ceftazidime-avibactam are the world-first subscription-contracted antibiotics (NHS 2023).

The Subscription Model approach represents a paradigm shift in global policymaking, emphasising the urgency of curbing antibiotic consumption and uniting support from all sectors - government, industry, and academia - to combat AMR. However, despite the combined efforts, our understanding of the molecular mechanisms underlying antibiotic resistance development and spread remains incomplete, highlighting the urgent need for accelerated research in this area to safeguard the future of public health.

1.1.2 Sources and mobility of AMR

The mechanisms of AMR emergence and dissemination are highly diverse (Blair *et al.* 2015, Darby *et al.* 2023). The global repository of antibiotic resistance genes (ARG) is termed the "resistome" which can be categorised into intrinsic and acquired resistance (Darby *et al.* 2023). Intrinsic resistance is inherent to bacteria and arises through genetic information already present on their chromosome and is often shared by all members of the genus (Darby *et al.* 2023). An example of intrinsic resistance is the rich assortment of protein efflux systems, which play a role in conferring various resistances, particularly in Gram-negative bacteria, including resistance to glycopeptides, β-lactams, tetracyclines, and macrolides (Reygaert 2018). Intrinsic resistance constitutes the "ancient resistome" (Olaitan and Rolain 2016), which developed ahead of the post-1940s antibiotic era through Darwinian mechanisms, such as the accumulation of random mutations over evolutionary timescales.

In contrast, the "modern resistome" evolves rapidly through the acquired resistance mechanisms, often involving the acquisition of resistance genes via horizontal gene transfer

(HGT) (Datta and Hughes 1983, Brito 2021). HGT facilitates the transfer of mobile genetic elements (MGE) between bacterial cells through four primary mechanisms: transformation (Figure 1A), bacterial conjugation (Figure 1B), transduction (Figure 1C) and membrane vesicle release and uptake (Figure 1D) (Brito 2021). MGE act as carriers of genetic information, accumulating and disseminating various genes that are collectively referred to as the mobilome.

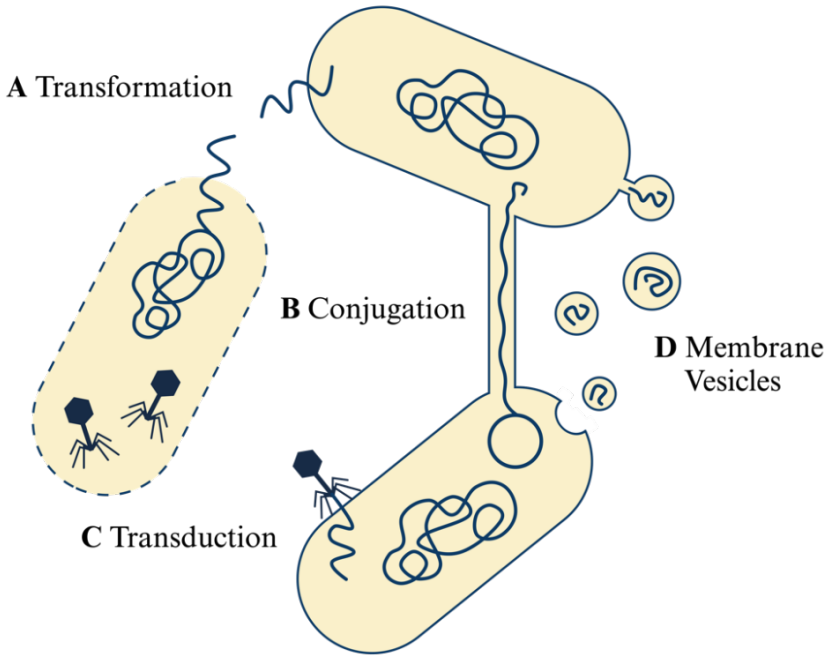


Figure 1. Routes of horizontal gene transfer between bacteria

A| Transformation is a process in which bacteria uptake free DNA from the environment, present after lysis of neighbouring bacteria. **B|** Bacterial conjugation is the most prominent form of HGT and involves a targeted transfer of DNA between two bacterial cells, which are connected via a conjugative pilus. **C|** Transduction encompasses the spread of genes mediated by phage infection. **D|** MGE can also travel between bacteria through membrane vesicles.

The existence of MGE has profound evolutionary implications as it enables homologous recombination between bacterial cells – a process typically associated with eukaryotes (Thomas and Nielsen 2005, González-Torres *et al.* 2019). MGE therefore increase the genomic plasticity of bacteria, offering in return a rapid adaptation to environments and niche establishment, achieved through MGE-encoded blueprints for novel metabolic strategies.

These include fundamental pathways of primary metabolism such as carbohydrate digestion (Hehemann *et al.* 2010) or specific catabolic pathways (Ikuma and Gunsch 2012), as well as pathogenic traits of secondary metabolism, such as virulence factors (Böhm *et al.* 2015) or mercury resistance (Hemme *et al.* 2016). However, the most alarming aspect of MGE existence is their capacity to encode and disseminate ARG among bacterial populations.

1.2 Bacterial Conjugation

Bacterial conjugation describes the process whereby two bacterial cells establish a physical connection to exchange genetic information. It occurs through the targeted, unidirectional transfer of single-stranded DNA (ssDNA) from a donor to a recipient bacterial cell, which remain connected through a conjugative pilus (Figure 2). This process has been observed in both Gram-negative and Gram-positive bacteria and is a widespread phenomenon occurring within microbial communities across diverse environments – such as soil, water sources, sewage, biofilms or in bacterial communities associated with plant or animal hosts (Davison 1999, Virolle *et al.* 2020). In these niches, conjugation serves as a mechanism for bacteria to adapt to the environment by disseminating MGE-encoded advantageous metabolic traits, such as virulence factors or ARG. In contrast to transduction or transformation, which occur as a side effect of phage lifecycle or nutrient acquisition, respectively, conjugation evolved as a specialised mechanism for deliberate and efficient movement of MGE between bacteria (Frost *et al.* 2005, Redfield *et al.* 2005). Consequently, bacterial conjugation became the major contributor to HGT thus shaping the evolution of bacterial genomes and the development of AMR (Lawrence 1999, Dolejska and Papagiannitsis 2018, Ott and Mellata 2022).

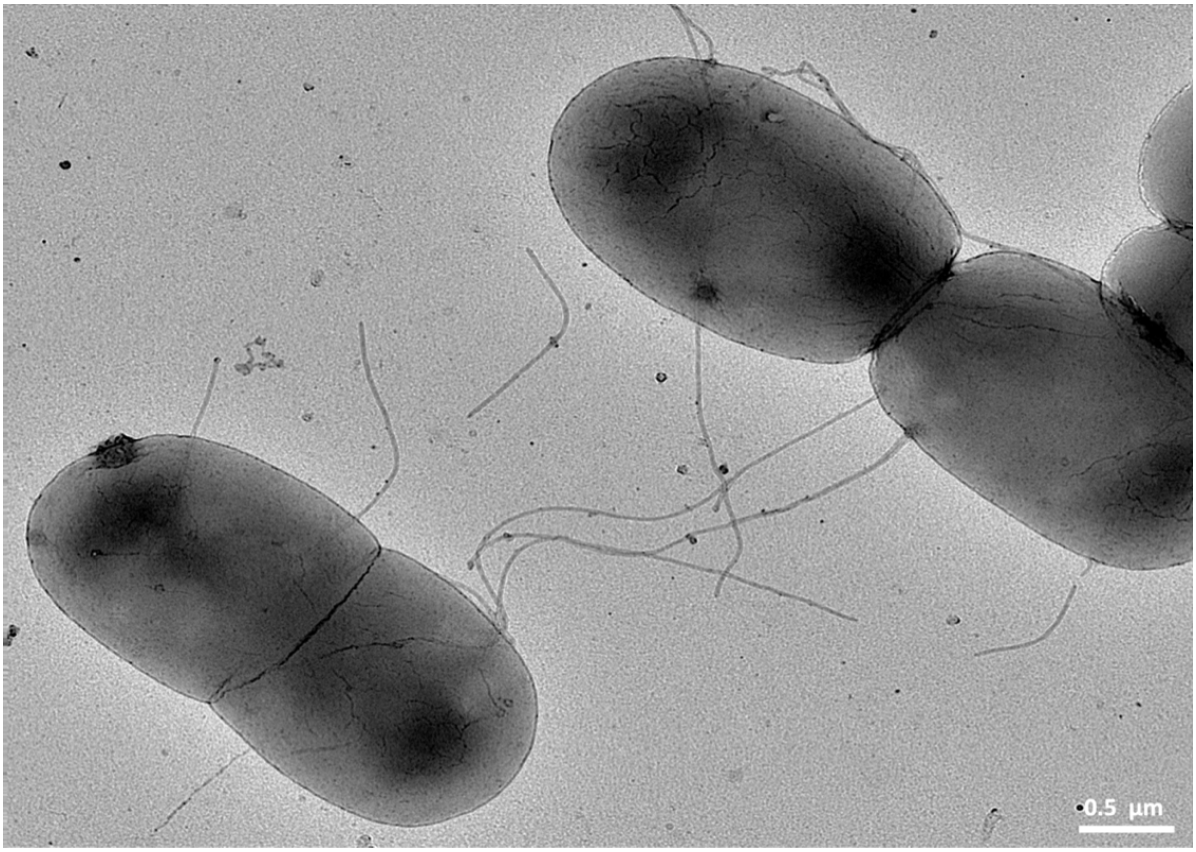


Figure 2. *E. coli* during bacterial conjugation

The negative stain electron microscopy (NS-EM) micrograph represents two bacterial cells undergoing conjugation. Both cells remain connected by the conjugative F-pilus, an extracellular appendage responsible for establishing physical interbacterial contacts during conjugation.

Edward Tatum and Joshua Lederberg, who first observed the phenomenon of bacterial conjugation, described the travelling genetic information as the bacterial “fertility factor”, or shortly - the F-factor (Lederberg and Tatum 1946). Subsequently, Lederberg realised that the F-factor is an extrachromosomal replicative element and introduced the term “plasmid”, thus the F-factor became the F-plasmid (Lederberg 1952, Helinski 2022).

Plasmids are functional assemblies of genetic modules, organised into stable and self-replicating entities that do not usually contain genes essential for cellular processes. Typically, plasmids are smaller than chromosomes and harbour a set of indispensable genes associated with their maintenance and replication alongside a diverse array of fluctuating accessory genes (Frost *et al.* 2005). For their survival, plasmids must ensure their transmission during cell

division via a process termed plasmid partitioning or be mobilised via conjugation (Frost *et al.* 2005). Therefore, half of the currently circulating plasmids have been documented as mobilisable, meaning they can serve as substrates for bacterial conjugation (Smillie *et al.* 2010, Coluzzi *et al.* 2022). A further half of mobilisable plasmids are categorised as self-transmissible, also known as "conjugative" plasmids - these not only serve as substrates for conjugation but also possess a full set of genes (referred to as transfer genes or simply – *tra* genes) necessary to construct the machinery responsible for bacterial conjugation – a type IV secretion system (T4SS) (Smillie *et al.* 2010). The division into "mobilisable" and "conjugative" plasmids arose because plasmid complexity correlates with their size, and thus bigger plasmids can encode multiple transmission strategies, such as several replicons or a conjugative machinery, whereas smaller plasmids will rely on the co-residing conjugative plasmids for their interbacterial mobility (Frost *et al.* 2005).

A noteworthy study of the era of plasmid discovery and classification shed light on the human-induced impact on the incorporation of AGRs into conjugative plasmids. N. Datta and V. M. Hughes conducted an analysis of the plasmid collection accumulated by E.D.G. Murray between 1917 and 1954, that is from the pre-antibiotic era, which revealed that none of the plasmids in the collection carried ARG, despite belonging to the same plasmid classes that are currently known to carry various antibiotic resistances (Datta and Hughes 1983). This study underscores the circulation of conjugative plasmids even before the incorporation of ARG conferred benefits to their bacterial hosts and shows that the evolutionary pressure to mobilise ARG on such a scale arose only after the penicillin discovery, that is in the post-1940s world, making it sadly, entirely man-made. Nowadays, conjugative plasmids harbour some of the most clinically concerning ARG, such as the colistin-resistance-associated gene *mcr-1*, and carbapenem-resistance-associated genes bla_{KPC} and bla_{NDM} (Wang et al. 2018, Wu et al. 2019, Yang et al. 2021, Shen et al. 2022), all three commonly found on conjugative plasmids of *E.*

coli and *K. pneumoniae* – two of the ESKAPEE strains (Quan et al. 2017, Wu et al. 2019, Yang et al. 2021).

Notably, while plasmids are typically the primary conjugative substrates, chromosome-borne MGE can also be mobilised through conjugation and are referred to as integrative conjugative elements (ICE) (Carraro *et al.* 2015). The mobility of ICE is poorly documented but is assumed to resemble that of plasmid, with the additional initial step of DNA excision and a final step of genomic integration. Therefore, ICE often encode additional genes similar to those found in temperate phages, such as integrases and transposases (Caparon and Scott 1989, Ravatn *et al.* 1998). In contrast to plasmids, there is limited evidence of autonomous replication of ICE, therefore they must rely on chromosome replication and cannot be maintained extrachromosomally (Wozniak and Waldor 2010). Similarly to the mobilisable and conjugative plasmids, ICE can encode a conjugative machinery or rely on a coexisting ICE/conjugative plasmid for mobility (Guglielmini *et al.* 2011). Many of them have been found to encode a wide range of antibiotic and heavy metal resistance genes (Whittle *et al.* 2002, Davies *et al.* 2009, Wozniak and Waldor 2010), thus ICE also play a role in ARG dissemination.

1.2.1 Incompatibility of conjugative plasmids and the IncF plasmids

It is impossible for two plasmids with the same replication or partitioning regions to coexist in a single cell, hence they are categorised into incompatibility (Inc) groups (Novick 1987, Frost *et al.* 2005, Garcillán-Barcia *et al.* 2023). Initially, due to the lack of proficient methods of gene sequencing, plasmid incompatibility was assessed by introducing two plasmids into a single cell by conjugation and examining their persistence after growth in an antibiotic-free medium (Datta and Hedges 1971). Currently, 26 incompatibility groups have been identified for *Enterobacteriaceae,* 14 for *Pseudomonas spp.* and 18 for the Gram-positive *Staphylococcus spp.* (Frost *et al.* 2005, Garcillán-Barcia *et al.* 2023). Interestingly, the number of the discovered incompatibility groups appears to be stabilising, suggesting a finite number of successful

replication mechanisms (Frost *et al.* 2005). Aside from their incompatibility, conjugative plasmids can be further classified based on several properties, such as the ability to conjugate in liquids (Dennison and Baumberg 1975, Bradley *et al.* 1980); repression or derepression of their *tra* genes (Baumberg and Dennison 1975, Bradley *et al.* 1980, Bradley 1984); thermosensitivity (Silver and Cohen 1972, Odakura *et al.* 1977, Bradley *et al.* 1980); susceptibility to bacteriophages (Bradley 1979, Bradley *et al.* 1981, Coetzee *et al.* 1983); and lastly by their interactions with plasmids of other Inc groups (Chabbert *et al.* 1972, Olsen and Shipley 1975).

Among different Inc groups, the IncF family is the main contributor to the dissemination of ARG among humans and animals, as it is the most prevalent type of conjugative plasmids in isolates coming from those two sources (Moran *et al.* 2015, Rozwandowicz *et al.* 2018, Stephens *et al.* 2020). IncF plasmids are exclusive to *Enterobacteriaceae* and incorporate a wide variety of ARG, including extended-spectrum beta-lactams, aminoglycosides, quinolones, sulphonamides, tetracyclines, and carbapenems, making them extremely prolific spreaders of ARG (Liao *et al.* 2013, Yang *et al.* 2015, Rozwandowicz *et al.* 2018). Dissemination of IncF-associated ARG is observed worldwide (Tamang *et al.* 2008, Woodford *et al.* 2009, Ho *et al.* 2012, Li *et al.* 2012, Ruiz *et al.* 2012, Dahmen *et al.* 2013), and nearly all plasmids encoding virulence genes, that is genes involved in bacterial pathogenesis, belong to the IncF family (Johnson and Nolan 2009). Taken together, understanding IncF plasmids and their modes of dissemination is incredibly pressing in the battle against AMR.

1.2.2 The F-plasmid

The archetypal representative of the IncF family is the iconic F-plasmid of *E. coli*, with more than 200 different plasmids currently termed F- or F-like plasmids (Koraimann 2018). The F-plasmid host range is restricted to closely related genera of *Escherichia spp.*, *Salmonella spp.*, *Klebsiella spp.*, *Shigella spp.*, *Enterobacter spp.*, and *Citrobacter spp.* (Koraimann 2018),

including commensal and pathogenic strains found predominantly in the gastrointestinal tracts of humans and animals.

Genetically, the F-plasmid is made up of four functional sections: replication region, cargo, mobility region and leading region (Figure 3A) (Koraimann 2018). The replication region comprises the IncF replicon, usually FII or FIC, responsible for plasmid replication, plasmid-partitioning systems, and multiple toxin-antitoxin (TA) systems that prevent the loss of the plasmid during segregation (Koraimann 2018). The cargo region is highly variable and includes all genes that provide a phenotypic advantage to the host cell (Koraimann 2018). It typically encodes a wide range of ARG, but also virulence genes and enterotoxins, often flanked by transposons and integrons, incorporation of which is an efficient way to maximise mobility and capturing of beneficial genes into the integron region, and hence into the plasmid (Wedel *et al.* 2023). The flagship integrons of F-like plasmids include the integron Tn21 which conveys mercury resistance and Tn3 with resistance to β-lactams (Nicolas *et al.* 2015, Partridge *et al.* 2018).

The last two regions of the F-plasmid – the mobility and leading regions – are both involved in bacterial conjugation. The leading region is the first one to enter the recipient cell during conjugative transfer and encodes a variety of genes responsible for a rapid settlement of the transferring plasmid within the new host (Manwaring *et al.* 1999). Specific DNA protection mechanisms of the incoming F-plasmid are necessary as foreign DNA triggers the deployment of antiforeign DNA weapons by the recipient cell, such as restriction enzymes, CRISPR-Cas systems, and SOS response (Koraimann 2018). The leading region must therefore counteract these defences, for example, by encoding copies of genes such as ArdB, which possesses antirestriction activity (Serfiotis-Mitsa *et al.* 2010) or PsiB, which interferes with induction of the SOS response (Bagdasarian *et al.* 1992, Petrova *et al.* 2009).

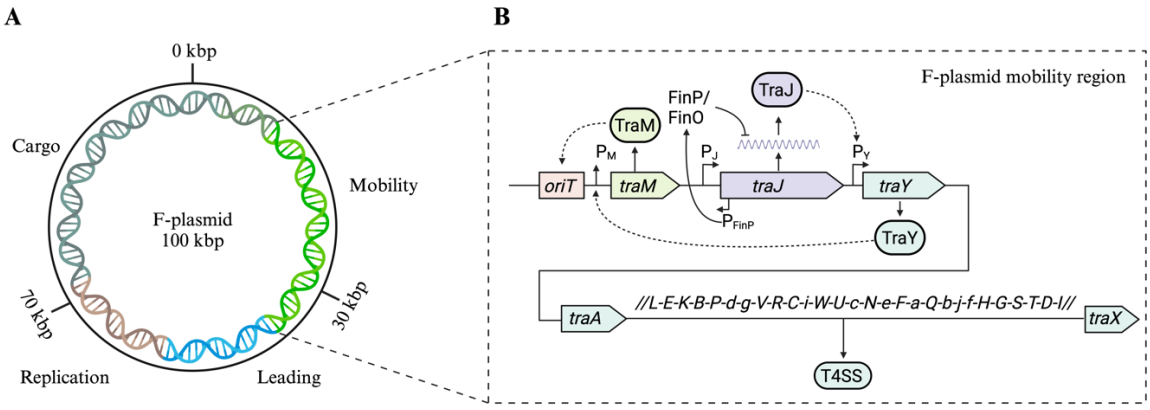


Figure 3. Genetic organisation of the F-plasmid

A| Genetic organisation of the F-plasmid includes four main regions: replication, cargo, mobility, and leading regions. These are involved in plasmid maintenance, encode AGRs, facilitate plasmid transfer, and aid plasmid settlement in the recipient cell, respectively. **B|** Genetic organisation and regulation of the mobility region. In summary, three promoters are involved in expression of mobility genes: P_M, P_J, and P_Y. The cascade is initiated by the P_J promoter that drives the transcription of *traJ*. The translated protein, TraJ binds to the P_Y promoter, triggering expression of all transfer proteins that assemble a type IV secretion system (T4SS), the transfer machinery. One of them is the transcriptional regulator TraY, which activates the P_M promoter, causing production of TraM – a relaxosome protein that initiates the processing of the plasmid DNA before conjugation. Regulation of this cascade is modulated by FinP/FinO complex that represses the translation of *traJ* mRNA. Arrows indicate protein transcription/translation. Dashed arrows indicate activation. Blunt arrow indicates inhibition. Between *traA* and *traX,* the genes denoted with upper-case letters indicate *tra* nomenclature (for example *traL)*, whereas lower-case letters indicate *trb* nomenclature (for example, *trbD*).

The mobility region of F-plasmids accounts for the ubiquity of this plasmid family as it enables the transfer of the plasmid between bacterial cells during bacterial conjugation. The mobility region includes two important sequences – the *oriT* site and the *tra* operon (Figure 3B).

OriT, or origin of transfer, contains specific DNA sequences needed for the processing of the F-plasmid DNA, potentiating it as a substrate for bacterial conjugation. These sequences are recognised by specific products of genes encoded in the *tra* operon. Their expression is however stringently controlled and is typically repressed, which minimises the metabolic cost of conjugation, being triggered only under specific (largely unknown) circumstances (Wong *et*

al. 2012b). A notable exception to this regulation is the F-like plasmid pED208 of *S. enterica*, which has an IS2 element inserted in its regulation region, resulting in constitutive expression of the *tra* region, making it a great candidate for structural and functional studies of bacterial conjugation due to the overproduction of all the proteins involved in conjugation.

The regulation of the *tra* genes expression is notably complex and relies on three promoters P_M, P_J, and P_Y (Figure 3B). The key regulatory element is the FinOP repressor system that controls the TraJ transcriptional activator protein (Figure 3B) (Biesen and Frost 1994, Wong *et al.* 2012b). In the repressed state, FinP directly binds the mRNA of TraJ, masking its ribosome binding sites (Virolle *et al.* 2020). To trigger the expression of transfer genes, FinO associates with FinP, which releases translation of the *traJ* gene. Subsequently, TraJ induces transcription of the *tra* region (Figure 3B), which lies under a single P_Y promoter and is nearly 35 kbp in length, notably representing the longest transcript found in *E. coli* (Fernandez-Lopez et al. 2016). The first expressed protein – TraY – acts as a transcriptional activator for the expression of TraM, a protein required for DNA processing in the early stages of conjugation (Figure 3B). The remainder of the transcript comprises over 25 genes that, upon expression, will assemble a large double-membrane-spanning protein complex that orchestrates the transfer of the F-plasmid DNA – the type IV secretion system (Costa *et al.* 2021, Costa *et al.* 2023).

1.2.3 Mechanism of conjugation: mating pair formation

Once all genes in the mobility region are expressed, bacterial conjugation can commence. This process relies on two concomitant events: the establishment of interbacterial connections and DNA processing and transfer. During bacterial conjugation, cell-to-cell contacts are formed by the conjugative F-pilus – an extracellular appendage elaborated by the cell harbouring a conjugative plasmid (henceforth referred to as the donor cell). The F-pilus is responsible for exploring the environment for a potential mating opportunity, that is a bacterial cell devoid of a conjugative plasmid of the same type (the recipient cell). For this, the F-pilus has the capacity

for repeated and highly dynamic extension and retraction cycles and rotation along its central axis (Figure 4A) (Clarke *et al.* 2008, Goldlust *et al.* 2023). Upon establishing a connection, the F-pilus adheres to the recipient cell, establishing a firm linkage between the two cells known as a mating pair (Figure 4A).

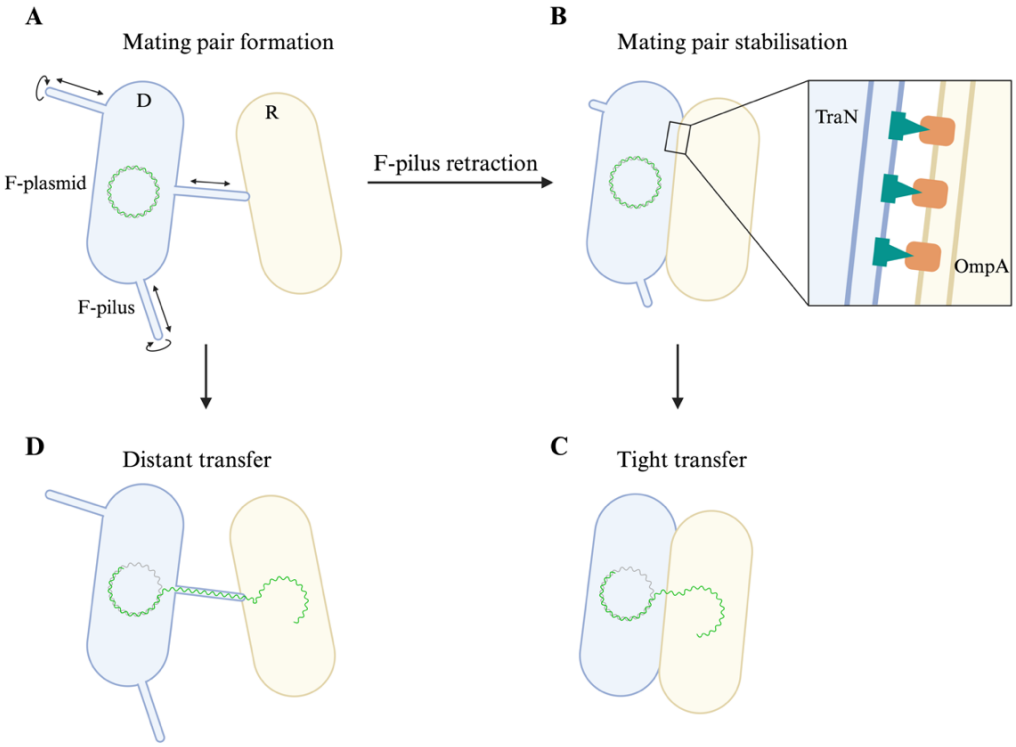


Figure 4. Two possible routes of bacterial conjugation

A| During the conjugation of the F-plasmid, the initial contact between donor and recipient cells occurs through the attachment of the donor's conjugative F-pilus to the surface of the recipient cell, known as mating pair formation. **B|** Pilus retraction results in establishment of membrane-to-membrane contacts, which are fortified by TraN–OmpA interactions, a process known as mating pair stabilisation. **C|** Subsequently, the close connection between cells leads to DNA transfer, named as tight transfer. **D|** Alternatively, DNA can be also transferred from the donor to the recipient cell through the lumen of the F-pilus. This is called the distant transfer. "D" denotes the donor cell, while "R" denotes the recipient cell.

The subsequent events following the formation of a mating pair have been subject to debate. Two potential pathways have been proposed: first, the retraction of the F-pilus draws the cells closer together, followed by the reinforcement of the connection known as the stabilisation of

the mating pair (Figure 4B). This process occurs through the interaction of the outer membrane proteins between the donor and the recipient, namely: F-plasmid encoded TraN in the donor cell and chromosomally-encoded and abundantly-expressed OmpA in the recipient cell (Figure 4B). Notably, the TraN-OmpA interactions were demonstrated to confer a degree of species specificity, with specific pairs being more prevalent during conjugation within the same species, while intraspecies conjugation may exhibit TraN-OmpA incompatibility (Low *et al.* 2022). Following stabilisation, the mating pair triggers DNA transfer, wherein the cell envelopes of both mating cells establish direct membrane-to-membrane contact (Figure 4C). This is known as "tight" transfer.

The second proposed pathway involves a distant transfer of DNA between mating cells, where no stabilisation of the mating pair occurs (Figure 4D). In this model, pilus retraction is not complete, hence membrane-to-membrane contacts cannot be established. DNA transfer must proceed through the lumen of the F-pilus – a mechanism that has been long hypothesised but only recently has been directly shown (Babić *et al.* 2008, Beltrán *et al.* 2023, Goldlust *et al.* 2023). It is currently not known what triggers distant transfer; however, it is hypothesised that while tight transfer is the preferred route due to the added stabilisation mechanism, it is only possible in conditions where bacterial cells have unrestrained mobility – such as in loosely populated aquatic environments. Distant transfer might therefore be an alternative route triggered in densely populated bacterial communities where pilus retraction is likely to be obstructed (Christie 2023).

The connections established by conjugative pili have implications beyond bacterial conjugation as they are also involved in the formation of bacterial biofilms (Ghigo 2001). These densely populated syntropic communities of bacteria formed through attachment either to surfaces or to one another pose a clear populational advantage to bacteria due to the secretion and establishment of an extracellular matrix. It is composed of polymeric substances like proteins,

polysaccharides, or DNA, forming a shield that protects them from antimicrobials or immune factors of the host organism (Vestby *et al.* 2020). This protective matrix accounts for 90% of the biofilm biomass and is a key virulence factor for many human pathogens, including uropathogenic *E. coli* or *K. pneumoniae* and medical-device-associated infections of *E. faecalis* (Kostakioti *et al.* 2013). Biofilms are also responsible for the recurrence of chronic infections, such as those seen in cystic fibrosis patients (Vestby *et al.* 2020).

Strains carrying IncF plasmids and elaborating conjugative F-pili form robust biofilms, a trait that was lost upon plasmid removal or when the gene encoding the F-pilin, *traA*, is deleted (Ghigo 2001, Reisner *et al.* 2003, Reisner *et al.* 2006). The mechanism of F-pili-mediated biofilm formation is unclear, however, it has been proposed that the F-pili are directly responsible for the initial cell-to-cell connections established between bacteria, after which the biofilms begin to mature and adhere to biotic or abiotic surfaces via elaboration of curli (May and Okabe 2008, May *et al.* 2011). Altogether, the elaboration of F-pili benefits the bacterial host on two levels that facilitate niche establishment and their virulence: it provides increased genome plasticity through the involvement of F-pili in horizontal gene transfer, and protection from environmental stressors via F-pilus-mediated biofilm formation. This underscores the importance of understanding the adaptations of the F-pilus to successfully establish cell-to-cell contacts.

1.2.4 Mechanism of conjugation: DNA processing and transfer

Concomitantly to the establishment of an F-pilus-mediated connection between the donor and the recipient cell, the conjugative F-plasmid must undergo a sophisticated network of processing to potentiate it for transfer. These processes can be broadly divided into two stages – those occurring within the donor cell and those within the recipient cell.

On the molecular level, bacterial conjugation begins with the expression of transfer genes, clustered into the *tra* operon within the transfer region of the F-plasmid (Figure 5, step 1)

(Virolle *et al.* 2020). These genes encode all the protein factors necessary for the formation of the T4SS - a multi-megadalton complex responsible for the establishment of the mating pair between cells (Figure 5, step 1a) (Virolle *et al.* 2020, Couturier *et al.* 2023). The T4SS is required for the elaboration of the conjugative pilus and creates a channel through both bacterial membranes of the donor cell, facilitating DNA passage (Lawley *et al.* 2003). Alongside the T4SS, the components of another protein complex - the relaxosome - are also expressed (Figure 5, step 1b) (Wong *et al.* 2012a). The relaxosome comprises four proteins: TraI, TraM, TraY, and host-encoded integration host factor (IHF), all of which assemble on the circular double-stranded DNA (dsDNA) of the F-plasmid and are involved in the processing of plasmid DNA ahead of its transfer (Figure 5, step 2) (Virolle *et al.* 2020, Couturier *et al.* 2023). Subsequently, TraI, the relaxase, introduces a single-stranded cleavage in the *oriT* site of the F-plasmid and remains covalently attached to its 5' end (Ilangovan *et al.* 2017). The relaxosome-DNA complex is then directed to the T4SS by its affinity to TraD – the type IV coupling protein (T4CP) of the F-plasmid, which acts as a substrate receptor and couples DNA processing and DNA transfer during conjugation (Figure 5, step 3) (Lu *et al.* 2008b, Virolle *et al.* 2020). TraD directs the travelling DNA towards the lumen of the T4SS, facilitating DNA translocation. As the transfer continues, the remaining 3' end of the plasmid DNA serves as a primer for the host DNA polymerase III, initiating rolling circle replication (RCR) to synthesise the complementary strand of F-plasmid DNA in the donor cell, making it proficient to undergo another round of transfer (Figure 5, step 4) (Mok and Marians 1987, Kim *et al.* 1996).

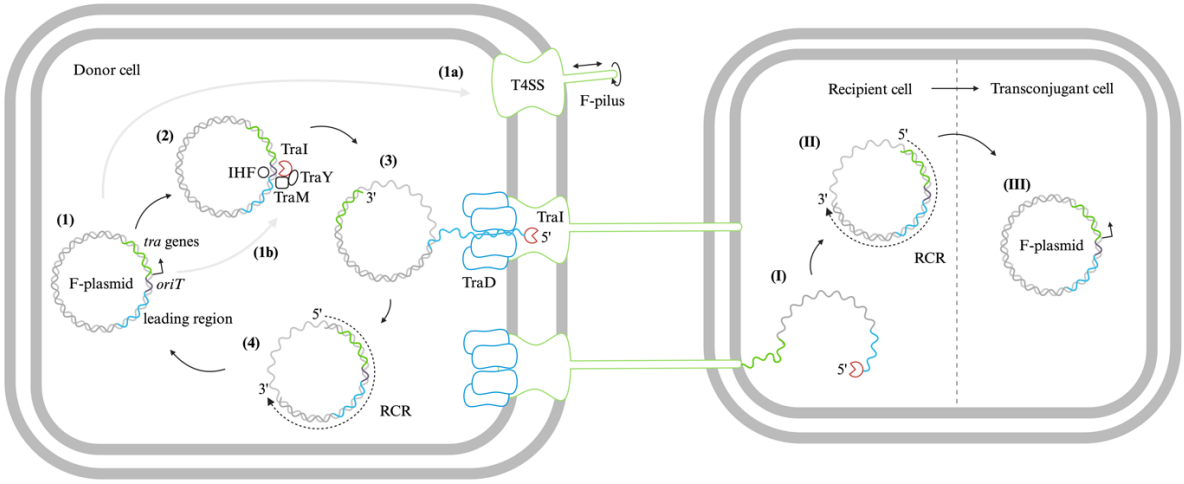


Figure 5. The mechanism of bacterial conjugation

(1) Two regions of the F-plasmid are crucial for bacterial conjugation, that is the leading region (blue) and the transfer region (green), with the *oriT* motif shown in purple. As *tra* genes are expressed, some assemble into the T4SS involved in the elaboration of the F-pilus **(1a),** whereas others, namely TraM, TraY and TraI form the relaxosome complex along with the chromosomally-encoded integration host factor (IHF) **(1b)**. Subsequently, the relaxosome assembles upon the F-plasmid DNA within the *oriT* locus. **(2)** The relaxosome prepares the F-plasmid for transfer by introducing a single stranded cleavage and TraI-DNA complex remains covalently bound throughout transfer. **(3)** The relaxosome-DNA complex is specifically recognised by TraD, a substrate receptor found at the base of the T4SS. Then, the F-plasmid DNA transferred through the lumen of T4SS to the recipient cell. **(4)** The remaining strand of the F-plasmid serves as a template for synthesis of the complementary strand, causing conversion from ssDNA to dsDNA by rolling circle replication (RCR). **(I)** In the recipient cell, TraI facilitates recircularisation of the F-plasmid DNA. **(II)** The complementary strand is synthesised through RCR, resulting in a double stranded copy of the F-plasmid. **(III)** The newly acquired genes of the transfer region can now be expressed by the recipient cell, which causes phenotypic conversion of the recipient cell into a transconjugant cell, which can now propagate the F-plasmid.

Within the recipient cell, the helicase activity of TraI performs 5' to 3' tracking of the incoming DNA in search for the other end of *oriT,* which will be recognised as TraI termination site (Figure 5, step I) (Garcillán-Barcia *et al.* 2007, De La Cruz *et al.* 2010). Once both ends are brought together, TraI joins them, thereby recircularising the F-plasmid (Becker and Meyer 2012, Chandler *et al.* 2013). As single-stranded DNA (ssDNA) accumulates in the recipient cell, its conversion to dsDNA commences at a rapid rate to prevent degradation (Couturier *et*

al. 2023). The leading region of the F-plasmid contains a specific *Frpo* region (F-plasmid RNA polymerase) that forms a stem-loop secondary structure, which reconstitutes the canonical -10 and -35 boxes (Masai and Arai 1997). The existence of *Frpo* facilitates recognition by the host RNA polymerase, thereby initiating primer synthesis (Masai and Arai 1997). Subsequently, DNA polymerase III begins complementary strand synthesis, which results in a fully synthesised F-plasmid (Figure 5, step II) (Yoichi *et al.* 1988).

Expression of genes present in the newly established F-plasmid by the recipient cell causes its phenotypic conversion into a transconjugant cell (Figure 5, step III). Functionally, transconjugants can serve as donors with competence for plasmid transfer, accounting for the exponential rate of conjugative plasmid dissemination across bacterial populations. Notably, the F-plasmid also encodes an exclusion system that prevents transconjugants from serving as recipients for the same plasmid, which limits redundant self-mating. This exclusion is conveyed by two proteins, TraT and TraS, outer- and inner-membrane proteins, respectively (Achtman *et al.* 1977, Thomas and Nielsen 2005, Marrero and Waldor 2007). The former provides surface exclusion, inhibiting the formation of another mating pair, whereas the latter achieves entry exclusion, preventing repetitive DNA transfer where a mating pair has already been established (Jalajakumari *et al.* 1987, Virolle *et al.* 2020).

1.3 Type IV Secretion Systems (T4SS)

1.3.1 T4SS diversity

The metabolic exchange between bacteria and their environment is pivotal for their survival, driving not only the primary cellular metabolism but also virulence, intercellular communication and developing adaptations to the ever-changing circumstances of the environment. To facilitate this exchange, bacteria have evolved a range of secretion machineries - large protein complexes localised within their membranes - that provide transport of macromolecules across the cell envelope, called secretion systems. In Gram-negative bacteria, the secretion systems range from type I secretion systems (T1SS) to type XI secretion systems (T11SS) (Costa *et al.* 2015, Veith *et al.* 2017, Filloux 2022), except for type VII secretion system (T7SS) characteristic of Gram-positive bacteria (Bunduc *et al.* 2021). Each system performs a highly specialised role in bacterial survival and fitness, including in processes such as pathogenesis (Kanonenberg *et al.* 2013), adhesion to surfaces (Aly and Baron 2007), cell-to-cell contact, biofilm formation (Leo *et al.* 2012), modulating host metabolism (Büttner 2012), competition killing (Ho *et al.* 2014), toxin secretion (Nivaskumar and Francetic 2014), antimicrobial resistance, and horizontal gene transfer (Christie *et al.* 2014).

Early secretion systems have been evolving alongside bacterial plasmids and eventually attained the ability to recognise their rolling-cycle replicases as transfer substrates (Grohmann *et al.* 2018). Subsequently, secretion systems adapted to efficiently shuttle whole plasmids across the cell envelope, thereby conveying what is now known as bacterial conjugation. Over time, conjugative systems began to recognise additional protein effectors as a response to developing relationships between bacteria and eukaryotes. Presently, all secretion systems ancestrally related to conjugation, despite the kind of substrates they carry, are classified into the T4SS family, making it incredibly functionally diverse.

T4SS are broadly categorised into three super-groups: contact-dependent interkingdom T4SS (Figure 6A), contact-dependent interbacterial T4SS (Figure 6B) and contact-independent DNA uptake or release T4SS (Figure 6C).

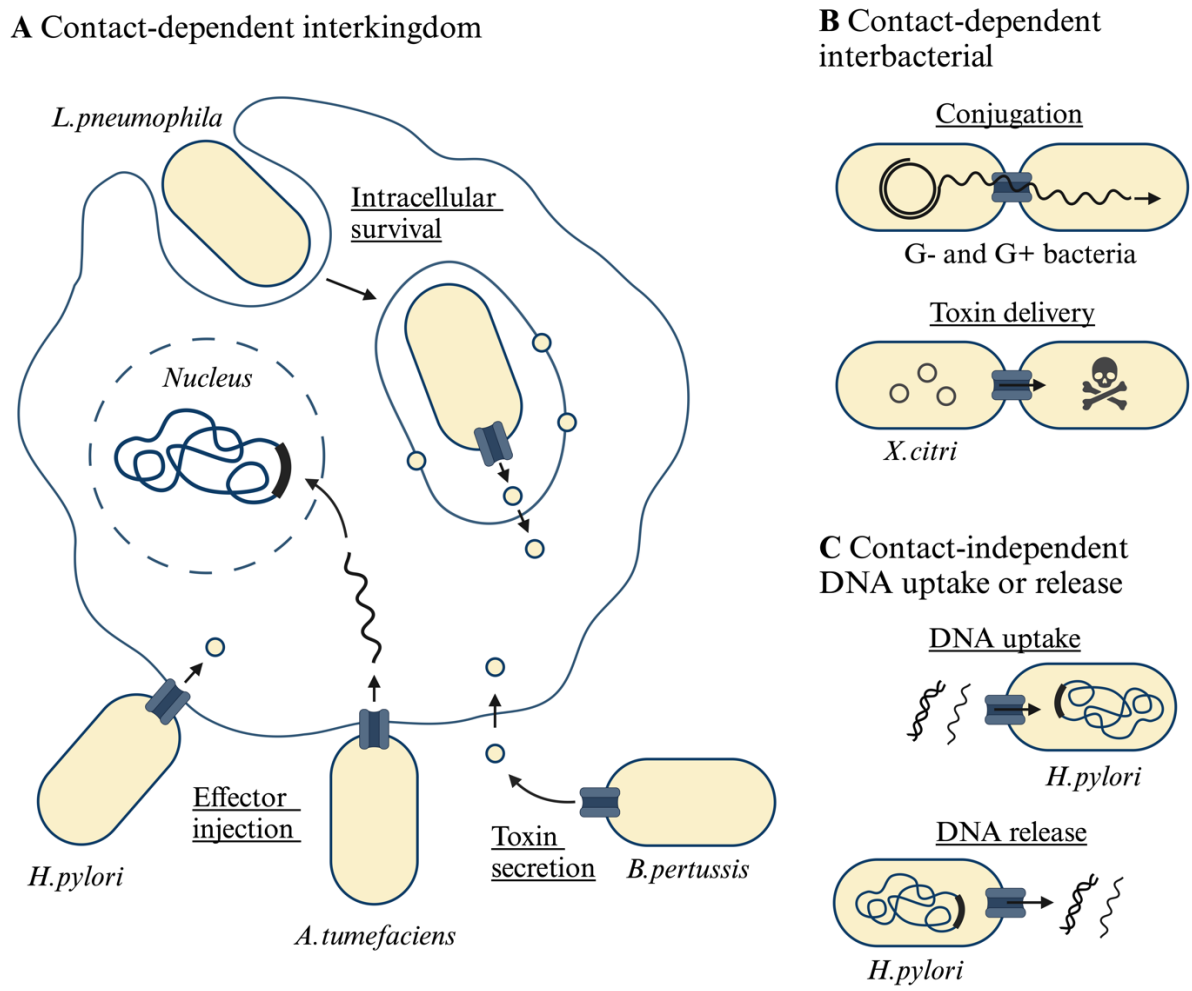


Figure 6. Diversity of T4SS

A| Contact-dependent interkingdom T4SS can be divided into T4SS involved in intracellular survival, such as *L. pneumophila* that uses its Dot/Icm-T4SS to establish a protective vacuole inside a macrophage; *H. pylori* utilises its T4SS for injection of effector oncoprotein CagA into human epithelial cells causing ulcers, or *A. tumefaciens* that injects effector DNA into plant cells inducing tumour growth; and T4SS that are involved in toxin secretion such as *B. pertussis* that secretes pertussis toxin which causes whooping cough in humans. **B|** Contact-dependent interbacterial T4SS can be involved in bacterial conjugation – the most widespread function of T4SS. In the case of *X. citri*, T4SS can be also involved in direct toxin delivery to neighbouring bacteria. **C|** Contact-independent DNA uptake or release T4SS are involved in the exchange of DNA with the environment.

These can provide DNA import such as in the case of *H. pylori* or DNA secretion such as in the case of *N. gonorrhoeae*.

T4SS that interact with eukaryotes are prevalent in clinically relevant pathogens, including *Legionella pneumophila, Helicobacter pylori*, or *Bordetella pertussis* (Christie and Cascales 2005, Backert and Meyer 2006). Despite losing their ability to mobilise DNA, they employ a wide variety of protein effectors to establish infection. For instance, the Dot/Icm-T4SS of *L. pneumophila* mobilises over 300 distinct protein effectors (Nagai and Kubori 2011, Lockwood *et al.* 2022), which orchestrate the formation of *Legionella*-containing vacuoles that promote survival within the host macrophages, thereby causing Legionnaire's disease (Figure 6A) (Meir *et al.* 2020, Lockwood *et al.* 2022, Viasus *et al.* 2022). *H. pylori* employs its Cag-T4SS to inject CagA, an oncoprotein, into epithelial cells lining human intestines, which leads to the development of ulcers and gastric cancer (Figure 6A). Similarly, *B. pertussis* secretes pertussis toxin, which, upon endocytosis, intoxicates mammalian cells, resulting in whooping cough (Figure 6A) (Burns 2021). T4SS can also mediate the transfer of DNA to eukaryotic cells, such as in the case of *Agrobacterium tumefaciens* which utilises its T4SS to translocate effector ssDNA into plant cells to induce tumour growth (Figure 6A) (Pitzschke and Hirt 2010).

The group of contact-dependent T4SS includes the most widespread and thus archetypal group of T4SS – the conjugative T4SS. These systems are ubiquitous across Gram-negative, Gram-positive, and Archaeal species (Figure 6B) (Costa *et al.* 2023). They rely on establishing direct physical contact between two bacterial cells to enable the transfer of MGE such as conjugative plasmids, transposons, and ICE, therefore they are paramount for genomic plasticity of bacterial populations, including the spread of MGE-associated ARG. Another contact-dependent interbacterial function of T4SS involves toxin delivery, such as the T4SS found in *Xanthomonas citri*, which is deployed to deliver lysis-inducing protein effectors to neighbouring Gram-positive and Gram-negative bacteria, resulting in interbacterial killing

(Figure 6B) (Souza *et al.* 2015, Sgro *et al.* 2018, Sgro *et al.* 2019). Functionally similar systems have also been recently documented in *Stenotrophomonas maltophila* (Bayer-Santos et al. 2019) and *Pseudomonas putida* (Purtschert-Montenegro et al. 2022).

In contrast to conjugative systems, T4SS involved in DNA uptake and release facilitate the exchange of genetic material between bacterial cells and their immediate environment without the need for direct contact with neighbouring cells (Figure 6C). An example of DNA release T4SS is encoded on the gonococcal genetic island of *N. gonorrhoeae*, which can release ssDNA directly into the environment for subsequent uptake by naturally competent *Neisseria* species (Callaghan *et al.* 2017). Additionally, DNA uptake, that is the transfer from the environment into the cell, can be deployed by the ComB-T4SS of *H. pylori*, providing natural competence to these bacteria (Figure 6C) (Hofreuter *et al.* 2001).

1.3.2 Architectures of the minimised T4SS

Operons encoding T4SS exhibit a diverse organisational pattern, however, their genomic arrangement displays substantial variability even among homologous T4SS. The order of genes within these operons, along with their lengths and the sizes of the corresponding genes and their encoded proteins, show considerable diversity. Remarkably, T4SS may extend beyond their designated operons, with emerging evidence highlighting the involvement of various cytoskeletal proteins or chromosome/plasmid partitioning ATPases in facilitating T4SS function.

In Gram-negative bacteria, T4SS are composed of a minimum of 12 core subunits designated as VirB1-VirB11 and VirD4 (Costa *et al.* 2021, Costa *et al.* 2023). These systems are categorised as "minimised" and primarily include the family of conjugative T4SS (Costa *et al.* 2021). Throughout evolution, minimised systems have adapted functions beyond conjugation and thus have integrated additional subunits beyond the canonical VirB/D4, hence “expanded” systems have branched out. These are made up of greater numbers of subunits – even over 35

as documented in certain *Neisseria* species (Ramsey *et al.* 2011). They can convey a variety of functionalities, including toxin and effector secretion and in many cases have lost their ancestral ability to transfer DNA (Gonzalez-Rivera *et al.* 2016, Sgro *et al.* 2019). Presently, the division into "minimised" and "expanded" T4SS is widely accepted as the leading architecture-based categorisation (Costa *et al.* 2021, Sheedlo *et al.* 2022, Costa *et al.* 2023).

Minimised T4SS have been the subject of structural investigation for over 20 years, with the first X-ray crystallography structures elucidating smaller, highly conserved subunits of T4SS, such as that of the VirB11-like ATPase from the Cag-T4SS of *H. pylori* or the VirD4-like ATPase TrwB of the R388 conjugative plasmid (Yeo *et al.* 2000, Gomis-Rüth and Coll 2001) However, larger T4SS subcomplexes were unattainable through X-ray crystallography alone, because of their large size (often encompassing multi-megadalton assemblies), and presence of multiple hydrophobic transmembrane domains – the two factors that hinder crystallisation. Consequently, the understanding of T4SS architecture has greatly benefited from the development and implementation of cryo-electron microscopy (cryo-EM) and cryo-electron tomography (cryo-ET) coupled with fluorescence microscopy techniques.

Early studies of minimised systems have shown that subunits VirB7, VirB9, and VirB10 are vital for assembling a T4SS, forming the first stable subcomplex within T4SS. This complex is known as the outer membrane core complex (OMCC) and serves as the foundation of the T4SS, upon which the remainder of the T4SS apparatus assembles (Christie *et al.* 2005, Khara *et al.* 2021). Consequently, the OMCC was the first large subunit of T4SS to be structurally characterised, with high-resolution structures obtained for the minimised OMCC of T4SS encoded by the conjugative pKM101 plasmid (Chandran *et al.* 2009, Fronzes *et al.* 2009), followed by OMCC of other minimised systems, namely from the R388 conjugative plasmid, the *A. tumefaciens* Ti-plasmid and of the *X. citri* bacterial-killing T4SS (Rivera-Calzada *et al.* 2013, Low *et al.* 2014, Gordon *et al.* 2017, Sgro *et al.* 2018).

Excitingly, a detailed insight into the T4SS organisation came in 2022, providing the first high-resolution cryo-EM structure of an intact T4SS encoded by the conjugative R388 plasmid (Figure 7A) (Macé *et al.* 2022). This structure can be divided into three subcomplexes, including the OMCC, a periplasmic stalk and a distinctive inner membrane complex (IMC) (Figure 7A). The periplasmic stalk and the IMC are surrounded by "arches" - these are sometimes described as either a standalone fourth complex, or as forming a periplasmic complex (PC) in conjunction with the stalk (Figure 7A).

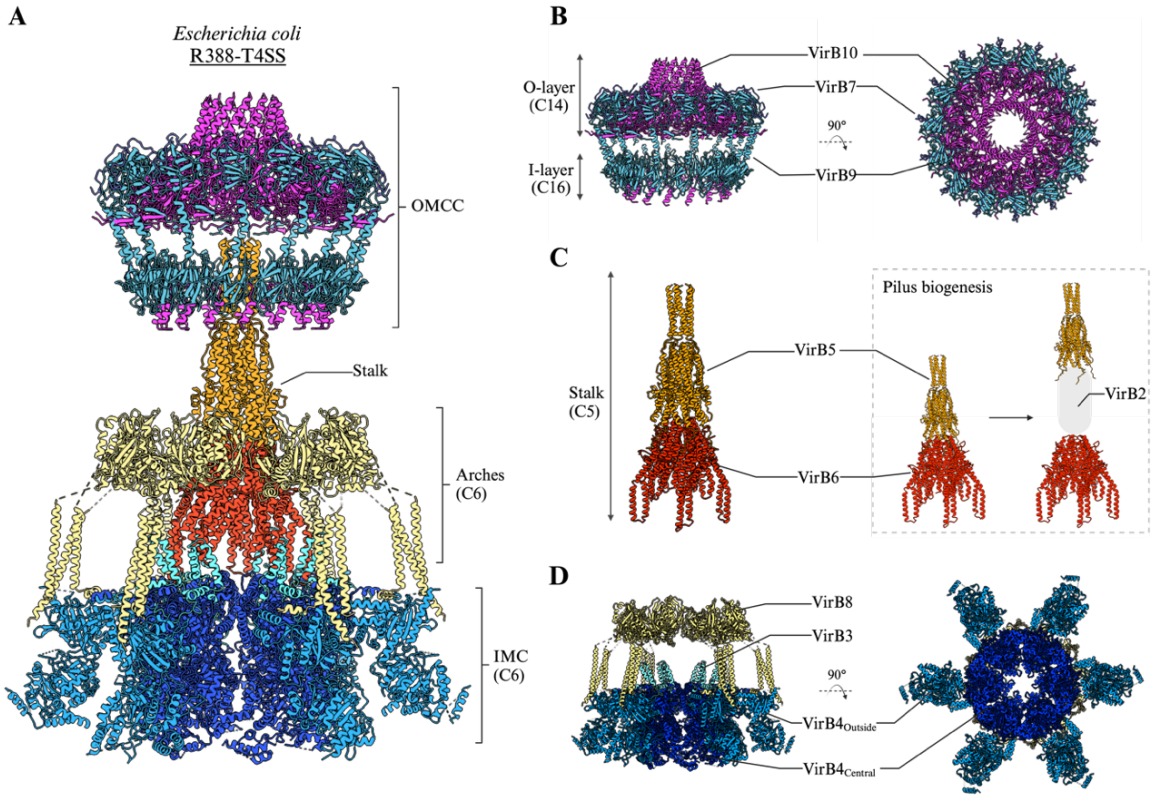


Figure 7. Cryo-EM structure of a minimised T4SS encoded on the R388 plasmid

A| The R388-T4SS is made up of four major assemblies: the outer membrane core complex (OMCC), the periplasm-residing stalk and arches, and the inner membrane complex (IMC). **B|** The OMCC consists of the outer (O-) layer with 14-fold symmetry (C14) and the inner (I-) layer with 16-fold symmetry (C16). The OMCC is made up of VirB10, VirB7 and VirB9 subunits. **C|** The periplasmic stalk is made up of VirB5 and VirB6 subunits with 5-fold symmetry (C5). Pilus biogenesis occurs by VirB2 subunits assembling at the interface between VirB6 and VirB5 (as depicted in frame). **D|** The

IMC is surrounded by arches that are made up of six trimers of VirB8, assuming a 6-fold symmetry (C6). These connect to the outside hexamer of VirB4, which form a concentric ring around the central hexamer of VirB4. The central hexamer also interacts with inner-membrane-embedded VirB3. The double VirB4 hexamer and VirB3 comprise the main components of the IMC.

The minimised OMCC structure assumes a barrel-like shape of 130 Å in height and 185 Å in diameter, consisting of proteins VirB7, VirB9, and VirB10 (Figure 7B). It comprises two distinctive layers that are symmetrically mismatched: the outer (O-) layer of 14-fold symmetry and the inner (I-) layer of 16-fold symmetry (Figure 7B). This C14:C16 mismatch is possible via the unique assembly mechanism in which the two diametrically opposite VirB9 and VirB10 heterodimers insert their N-termini into the I-layer, but their C-termini do not integrate into the O-layer. The β-barrel domains of VirB10 bundle up into a hollow opening in the outer membrane, connecting the periplasm with the extracellular milieu (Figure 7B). This opening is predicted to exist in two conformations: open and closed, which are interchangeable via a hinge-like mechanism along α-helices that protrude towards the centre of the pore (Cascales and Christie 2004a, Banta *et al.* 2011, Cascales *et al.* 2013, Darbari *et al.* 2020, Macé *et al.* 2022). Interestingly, there is a high heterogeneity in the structures of minimised OMCC. For instance, the constituents of the OMCC of the minimised T4SS of *X. citri* involved in bacterial killing are highly homologous to those of the T4SS encoded on the R388 plasmid; however, they do not present the O- and I-layer symmetry mismatch (Sgro *et al.* 2018, Sgro *et al.* 2019). Currently, the functional significance of this mismatch remains unclear.

Connecting the OMCC and the IMC is the periplasmic stalk, a cone-shaped assembly spanning 216 Å, composed of a VirB5 pentamer located proximally to the OMCC atop a VirB6 pentamer anchored in the inner membrane (Figure 7C). The topology of the stalk provides an exciting insight into the steps of pilus biogenesis. By structural-complementarity simulations, the VirB2 pilin subunit of the conjugative pilus was successfully docked onto the VirB6 platform, which yielded a novel model for the assembly of conjugative pili, considering the previous evidence

that VirB5 locates at the tip of the T-pilus elaborated by the Ti-plasmid encoded T4SS of *A. tumefaciens* (Aly and Baron 2007) together with the fact that the bulk of VirB6 resides in the periplasm (Figure 7C). As the symmetry of the pilus matches that of both VirB5 and VirB6 (C5), one can envision that VirB2 subunits are extracted from the inner membrane directly onto the interphase between VirB5 and VirB6 (Macé *et al.* 2022, Costa *et al.* 2023). Through repetitive recruitment of VirB2 subunits, the pilus elongates and displaces VirB5 pentamer outward through the OMCC and into the extracellular environment (Figure 7C) (Macé *et al.* 2022). Subsequently, the VirB5 pentamer remains at the tip of the pilus and is presumably involved in interaction with specific receptors on the recipient's cell surface (Costa *et al.* 2023).

The IMC of the R388-encoded minimised T4SS was shown to adopt a two-side-by-side barrel structure (Low *et al.* 2014, Redzej *et al.* 2017). However, the recent advances in cryo-ET and cryo-EM have shown a different organisation of the IMC of minimised systems. The newer cryo-ET map of T4SS from the pKM101 plasmid (Khara *et al.* 2021), together with the high-resolution cryo-EM structure of the R388-encoded T4SS (Macé *et al.* 2022) both unveil a more intricate arrangement, consisting of a hexamer of homodimers of VirB4 forming two concentric rings—an inner and outer ring—projecting into the cytoplasm from the inner membrane (Figure 7D). VirB4 is an ATPase that functions in orchestrating and energising pilus biogenesis (Kerr and Christie 2010) and remains permanently associated with the T4SS. Specifically, the central VirB4 ring interacts with VirB3, anchoring it to the inner membrane (Figure 7D), while the outer VirB4 ring associates with the N-terminal segment of VirB8 in a 1:3 stoichiometry, resulting in 18 VirB8 subunits per T4SS (Figure 7D). VirB8 further forms periplasmic arches, organised as hexamers of trimers (Figure 7D).

Functionally, two additional ATPases interact with the T4SS, that is VirB11 and VirD4, which associate only transiently with the IMC (Atmakuri *et al.* 2004, Cascales *et al.* 2013, Ripoll-Rozada *et al.* 2013). Because of this, the precise localisation of VirB11 and VirD4 relative to

the T4SS remains ambiguous. VirB11 self-assembles as a homohexamer and localises to the inner face of the cytoplasmic membrane independently of the remainder of the T4SS, despite not having a canonical transmembrane domain (Stephens *et al.* 1995, Rashkova *et al.* 1997, Savvides *et al.* 2003, Hare *et al.* 2006). VirB11 is expected to interact directly with VirB4 to stabilise its hexameric assembly (Macé *et al.* 2022), and as such the VirB4-VirB11 complex will function in orchestrating the elaboration of the conjugative pilus (Kerr and Christie 2010).

The other ATPase, VirD4 is the T4CP of the R388 system, which acts as the substrate receptor by specifically interacting with protein or nucleoprotein substrates, and subsequently directing them towards the T4SS for translocation (Llosa and Alkorta 2017, Redzej *et al.* 2017). VirD4 transverses the inner membrane via its N-terminal transmembrane domain, which interacts transiently with VirB10 – the OMCC protein (Segura *et al.* 2014), suggesting a potential role of VirD4 in signal transduction from the IMC to the OMCC (Christie *et al.* 2014). In solution, VirD4 displays a range of oligomeric states – including dimers and monomers (Redzej *et al.* 2017) or hexamers upon crystallisation (Gomis-Rüth *et al.* 2001). Functionally, only dimeric VirD4 was captured attached to the VirB4 at the base of the T4SS using chemical crosslinking (Redzej *et al.* 2017), suggesting a specific role of the dimeric VirD4 state in T4SS operation.

1.3.3 Architectures of the expanded T4SS

Expanded T4SS consist of over 25 distinct proteins, including the essential 12 core components described in the minimised systems, along with additional protein subunits (Costa *et al.* 2021, Sheedlo *et al.* 2022). Among the extensively studied expanded systems are the conjugative T4SS encoded by the F-plasmid (F-T4SS), the Dot/Icm-T4SS of *L. pneumophila* and Cag-T4SS of *H. pylori* – last two are involved in effector translocation and facilitate pathogenicity of their hosts (Lawley *et al.* 2003, Gonzalez-Rivera *et al.* 2016, Guzmán-Herrador *et al.* 2017) (Figure 6). While detailed structures of the IMC present in these systems are yet to be

determined, near-atomic-resolution structures have been obtained for the OMCC in all three of them.

The F-T4SS stands as the prime model for investigating conjugative T4SS even though the genetic organisation of its operon presents a more intricate assembly when compared to minimised conjugative machineries, such as those encoded on the R388 or pKM101 plasmids. The subunits involved in F-T4SS assembly are encoded on an operon made up of 30 genes instead of the conventional 12 genes known from R388 and pKM101. While nine of these proteins exhibit homology between minimised and F-T4SS, the remainder is unique to the F-T4SS (Bragagnolo *et al.* 2020).

The OMCC of the F-T4SS is composed of VirB7-like TraV, VirB9-like TraK, and VirB10-like TraB subunits (Figure 8A) (Amin *et al.* 2021, Liu *et al.* 2022). It assumes a distinct "flying saucer" shape with a diameter of approximately 268 Å and a width of 115 Å and forms two concentric rings termed the outer and inner rings (OR and IR, respectively) (Figure 8A). The IR is characterised by a 17-fold symmetry and consists of 17 C-terminal domains (CTDs) of TraB and N-terminal domains (NTDs) of TraV. In contrast, the OR, featuring a 13-fold symmetry, comprises 26 CTDs of TraV and 26 copies of TraK, assembling as 13 homodimers. Each dimer contributes its CTDs to form 13 "knobs" that extend radially from the core of the OMCC (Amin *et al.* 2021, Costa *et al.* 2023). The symmetry mismatch between the two rings is accommodated by flexible linkers of TraV and TraB.

Cryo-ET studies of the F-T4SS have revealed that the thickness of the stalk is much greater compared to the slender periplasmic stalk observed in the minimised conjugative T4SS (Figure 8A). This increase in periplasmic volume can attributed to a group of F-exclusive periplasmic proteins, that is —TraH, -U, -F, -W, TrbC, and TrbI, which have been shown to form a periplasmic complex associated with the F-T4SS OMCC (Koch *et al.* 2020). Functionally,

these proteins are expected to be involved specifically in the F-pilus extension and retraction, however their precise functions remain unknown (Bragagnolo *et al.* 2020).

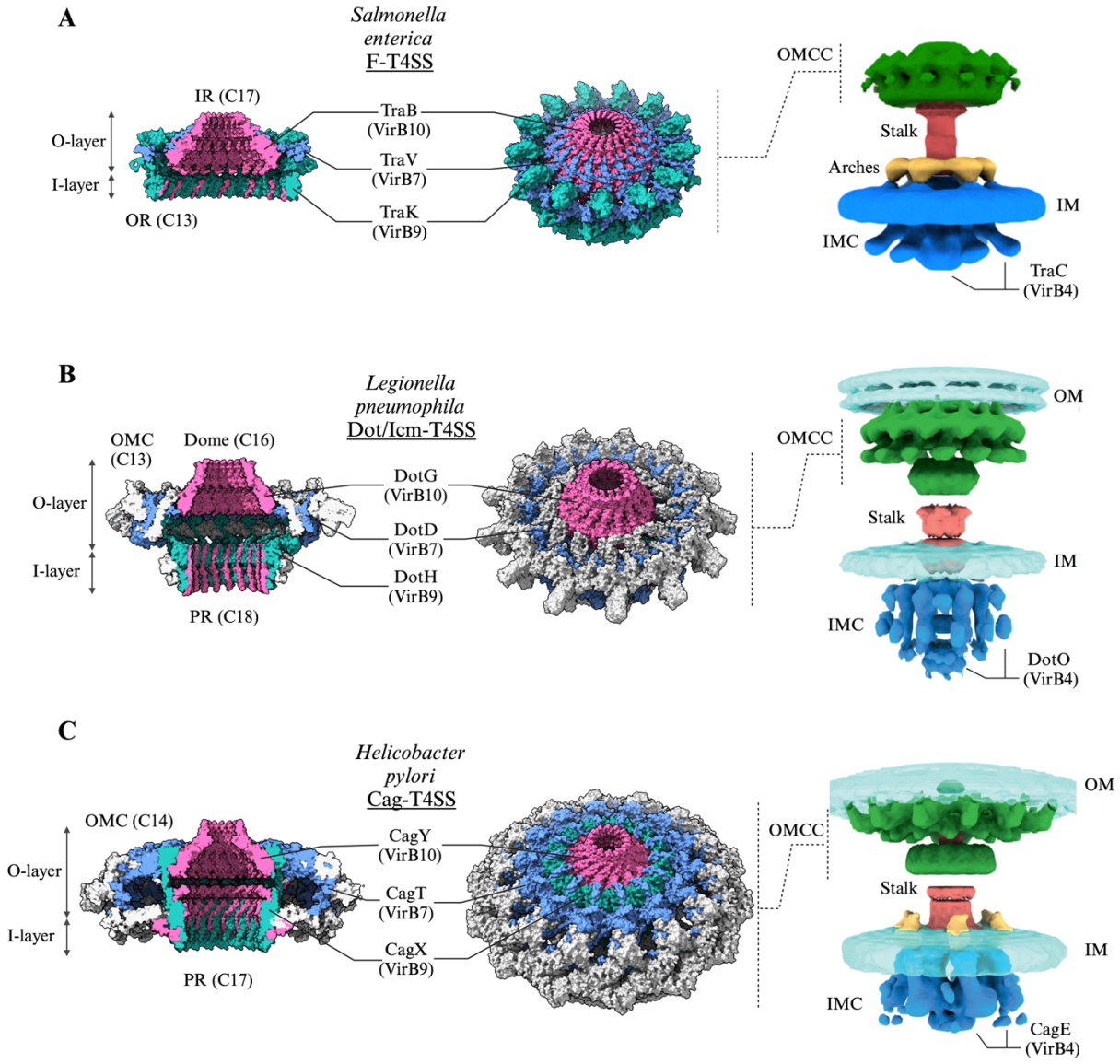


Figure 8. Cryo-EM structures of OMCC and cryo-ET maps of the expanded T4SS

Protein constituents homologous to the VirB7, VirB9 and VirB10 are coloured in blue, green and pink, respectively. **A|** Left: the OMCC encoded on an F-like plasmid pED208 from *S. enterica* is the smallest of the described OMCC and consists of the O- and I- layers. It assumes a pore-like structure of two concentric rings, the inner ring (IR) of 17C symmetry, and the outer ring (OR) of 13C symmetry with 17C and 13C symmetries. The F-T4SS OMCC is made up of TraB, TraV and TraK – which are homologous to VirB10, VirB7 and VirB9, respectively. Right: cryo-ET structure of the whole F-T4SS (EMDB ID: 9344), coloured by subcomplex: green for the OMCC, red for the stalk, yellow for the

arches and blue for the IMC. **B|** Left: atomic structure of *L. pneumophila* OMCC, part of the Dot/Icm-T4SS. It consists of a dome that protrudes towards the extracellular milieu and assumes 16-fold symmetry, an Outer Membrane Cap (OMC) with 13C symmetry, and a periplasmic ring (PR) with 18-fold symmetry. (PDB ID: 7MUS) Right: cryo-ET structure of the whole Dot/Icm-T4SS (EMDB ID: 7611 and 7612), coloured as in A. **C|** Left: atomic structure of the Cag-T4SS present in *H. pylori*. The OMC has 14C symmetry, whereas the PR – 17C. This OMCC is currently the biggest known OMCC, due to larger size of its constituents – CagY, CagT and CagX as well as additional proteins that make it up. (PDB ID: 6X6S and 6X6J). Right: cryo-ET structure of the Cag-T4SS, coloured as in A. (EMDB ID: 0634 and 0635). IM – inner membrane, OM- outer membrane.

The IMC of the F-T4SS comprises a central hexamer-of-dimers assembly made up of the VirB4-like TraC ATPase (Figure 8A), resembling that presented in the cryo-EM structure of the fully assembled T4SS (Figure 7) (Macé *et al.* 2022), however, additional F-exclusive proteins have been identified as inner membrane proteins of the F-T4SS that are absent in the minimised system, including TraG, TrbI, or TraP (Bragagnolo *et al.* 2020). The localisation of these proteins in relation to the entire T4SS apparatus is currently dubious as they could not be easily distinguished in the cryo-ET map. A high-resolution structure of the fully assembled F-T4SS is urgently needed to understand where the F-exclusive proteins localise within the apparatus.

The cryo-ET imaging of the F-T4SS also provided insight into its functional "life cycle", which involves a series of assembly and disassembly events (Hu *et al.* 2019a). Different conformational states of the apparatus could be observed across the cell envelopes of the imaged bacteria, expected to represent functional transitions of the F-T4SS that occur during bacterial conjugation (Hu *et al.* 2019a). The authors propose a model whereby the fully assembled F-T4SS – that is the structure named the "F1-Channel" represents the quiescent F-T4SS that corresponds to 30% of all F-T4SS observed during imaging (Figure 9A). Subsequently, in the early stages of bacterial conjugation, the F-T4SS is involved in mate-seeking, thus it elaborates the F-pilus, giving rise to the F2-Channel/Pilus assembly which

builds and retracts pili (Figure 9B). After the F-pilus establishes contact with a potential mating partner, conjugation begins, and the apparatus is activated by interaction with TraD, the VirD4-like ATPase, which recruits the relaxosome – that is protein complex TraIYM and IHF - carrying the F-plasmid DNA (Figure 9C).

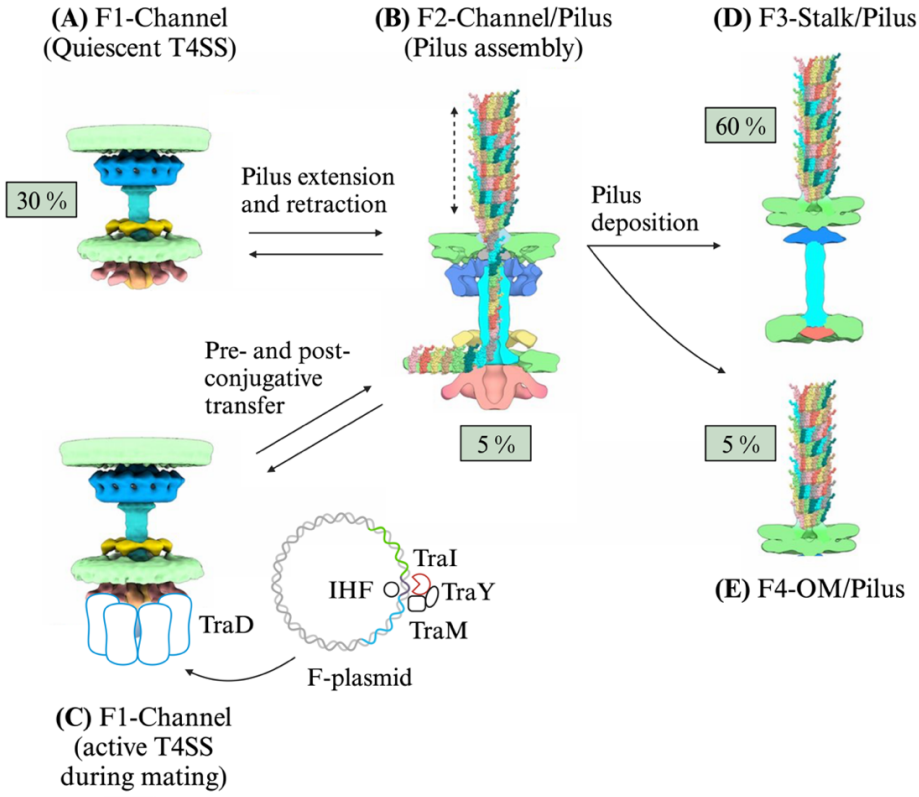


Figure 9. Structural transitions of F-T4SS during different stages of bacterial conjugation

A| Initially, the F-T4SS subunits assemble into a dormant F1-Channel complex. **B|** For mate-seeking, the F1 complex undergoes transition into the dynamic F2-Channel/Pilus conformation, representing extension and retraction of the F-pilus by the F-T4SS. This process promotes the establishment of interbacterial contacts that will facilitate conjugation. When contact with the recipient cell is established, the F-pilus retracts, returning to the conformation A. **C|** Bacterial conjugation begins and the F1 structure is activated, triggering the recruitment of TraD, which acts as a receptor for the relaxosome complex (made up of TraIYM and IHF subunits) together with the F-plasmid DNA. In some cases, F-pili can be deposited on alternative platforms along the cell outer membrane. These platforms are represented by structures **D|** and **E|** which can provide non-conjugative contacts and function in development of biofilms. In green boxes, the proportional representation of each state as observed on *E. coli* minicells via *in situ* cryo-ET is given (Hu *et al.* 2019a).

Curiously, conjugative pili were also observed to be deposited in the membrane onto alternative basal platforms which are devoid of some subunits of the F-T4SS. A large share of the observed F-T4SS (60%) represented the F3-Stalk/Pilus assembly that lacks the IMC and parts of the OMCC (Figure 9D). Strikingly, some F-pili were observed to be deposited directly in the OM without any distinguishable T4SS subunits, comprising the F4-OM/Pilus structure (Figure 9E). As the F3-Stalk./Pilus and F4-OM/Pilus structures are devoid of significant parts of the functional T4SS subassemblies, they are likely incapable of mediating bacterial conjugation. The significance of these deposited F-pili is unknown, but they are hypothesised to function in providing biofilm-forming interbacterial contacts rather than their original function in conjugation (Hu *et al.* 2019a, Patkowski *et al.* 2023).

The second well-studied expanded T4SS is the Dot/Icm-T4SS, which plays a crucial role in the invasion of alveolar macrophages by *L. pneumophila*, deploying a complex network of over 300 protein effectors that constitute an astonishing 10% of the bacterium's proteome (Gomez-Valero *et al.* 2019, Kitao *et al.* 2022, Lockwood *et al.* 2022). Through this mechanism, *L. pneumophila* manipulates the phagosomes into protective compartments called *Legionella*-containing vacuoles, enabling the bacterium to survive and replicate inside macrophages and evade the immune defences of its host, thereby facilitating infection (Asrat *et al.* 2014, Lockwood *et al.* 2022).

The OMCC of Dot/Icm-T4SS has a diameter of approximately 420 Å and is primarily composed of VirB7-like DotD, VirB9-like DotH, and VirB10-like DotG proteins, with additional integration of DotC, Dis1, Dis2, and Dis3, which are Dot/Icm-exclusive proteins (Figure 8B) (Durie *et al.* 2020, Sheedlo *et al.* 2021). This OMCC comprises three distinct layers: a disk-shaped outer membrane cap (OMC) exhibiting 13-fold symmetry, a proximal dome with 16-fold symmetry, and a smaller periplasmic ring with 18-fold symmetry (Figure

8B). Notably, the OMC and periplasmic ring correspond to distinct O- and I-layers observed in minimised OMCC (Figure 7B), respectively.

In situ visualisation of the Dot/Icm-T4SS by cryo-ET structure reveals that its OMCC is linked to a cylinder with an outer diameter of approximately 20 nm, featuring a central lumen measuring around 6 nm, which narrows down with proximity to the inner membrane (Figure 8B) (Chetrit *et al.* 2018). Meanwhile, the IMC exhibits a configuration much like that of minimised systems, with a hexamer of dimers topology attributed to the VirB4-like DotO (Figure 8B). Notably, the second ATPase of this system, VirB11-like DotB, forms a hexamer that docks at the inner hexamer of DotO, a feature yet to be observed in the minimised systems (Chetrit *et al.* 2018).

The third well-characterised T4SS is the Cag-T4SS, essential for the pathogenesis of *H. pylori*. These bacteria stand out as a major bacterial pathogen, infecting half of the world's population (Shirani *et al.* 2023). They adhere to human gastric epithelial cells, contributing to the development of ulcers, gastric adenocarcinomas, and lymphomas of the mucosa (Coppens *et al.* 2018, Reyes 2023). During virulence, *H. pylori* employs its Cag-T4SS to translocate various effectors into the host's gastric cells, including LPS metabolites or DNA, but most importantly the only known protein effector - the CagA oncoprotein, which triggers a cascade of morphological alterations of host cells and induces their epithelial-mesenchymal transition that leads to the development of tumours (Cover *et al.* 2020).

The mushroom-shaped OMCC of the Cag system spans 400 Å in width and 250 Å in height, comprised of VirB7-like CagT, VirB9-like CagX, and VirB10-like CagY proteins, complemented by two Cag-exclusive proteins, Cag3 and CagM (Figure 8C) (Frick-Cheng *et al.* 2016). Structurally, the Cag-OMCC forms an OMC with separate O- and I-layers, constituted by CagT, CagX, CagY, Cag3, and CagM, while the periplasmic ring is formed by

CagY and CagX. Both the O- and I-layers of the OMC exhibit 14-fold symmetry, while the periplasmic ring adopts 17-fold symmetry (Figure 8C).

In situ cryo-ET studies of the Cag-T4SS reveal that its OMCC is linked to a hollow cylinder traversing the periplasm, establishing a connection between the OMCC and the IMC (Figure 8C) (Hu *et al.* 2019b). The IMC exhibits an unusual architectural configuration with three concentric rings instead of the conventional two (Figure 8C). The inner and middle rings of the IMC align with the central and outer hexamers of VirB4-like CagE (Figure 8C), while the assembly of the outer ring depends on the production of VirD4-like Cagβ, documenting the second direct observation of a T4CP density associated with its cognate T4SS (Hu *et al.* 2019a), after the previously observed binding of the dimeric VirD4-like TrwB to the IMC of the T4SS encoded by the conjugative plasmid R388 (Redzej *et al.* 2017). VirB11-like Cagα associates at the base of the CagE central hexamer. Unlike the homologues present in the Dot/Icm system, the association of Cagα with CagE appears stable rather than dynamic.

1.3.4 The F-pilus biogenesis

To establish physical cell-to-cell contacts for DNA transfer, the conjugative F-T4SS elaborates the conjugative F-pilus in a complex and multistep process. Initially, the F-pilin subunits (TraA) are synthesised as pro-pilins with an unusually long leader peptide of about 50 residues (Costa *et al.* 2016, Hospenthal *et al.* 2017). These pro-pilins are inserted into the inner membrane utilising an ATP-dependent and proton motive force (PMF)-dependent mechanism chaperoned by TraQ (Figure 10, step 1) (Majdalani and Ippen-Ihler 1996, Hospenthal *et al.* 2017). Upon insertion, the leader peptide is cleaved by the periplasmic peptidase LepB, which results in mature pilin with both N- and C-termini exposed to the periplasm (Figure 10, step 2) (Majdalani *et al.* 1996). Subsequently, TraX acetylates the N-terminus of TraA (Figure 10, step 3) (Moore *et al.* 1993), finalising maturation. Mature TraA subunits form a pool of inner membrane-associated subunits known as TraA-pilins presumably located in proximity to the

F-T4SS and are readily available for pilus assembly. When extracted from the inner membrane and directed to F-pilus nucleation, each TraA molecule specifically binds one molecule of phospholipid (Figure 10, step 4).

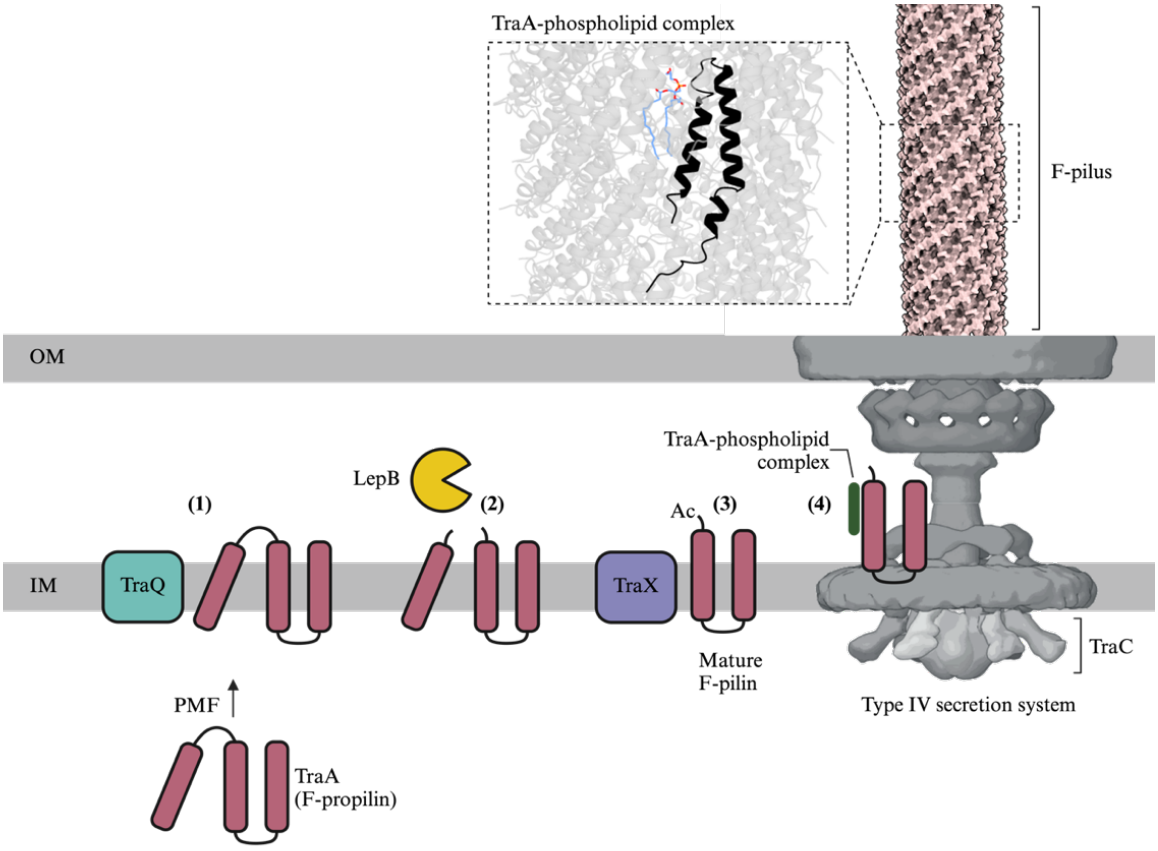


Figure 10. The F-pilus biogenesis

Before assembling into the F-pilus, F-pilin subunits must undergo four steps of maturation. **(1)** Initially, immediate products of the gene *traA* - the F-propilins - are transported to the inner membrane through proton motive force (PMF) mechanism that is facilitated by protein TraQ. **(2)** The leader peptide is cleaved off by LepB peptidase present in the periplasm, followed by **(3)** acetylation of F-pilins at the N-terminus by TraX. Lastly, **(4)** the F-pilins associate with phospholipid molecules and as such assemble into a helical filament by the T4SS, energised by TraC ATPase localised at its base.

The F-pilus assembly is orchestrated specifically by the VirB4-like ATPase TraC, the main constituent of the F-4SS IMC, however, the precise mechanism of that is not understood (Hospenthal *et al.* 2017, Bragagnolo *et al.* 2020). Interestingly, in minimised conjugative systems, such as those encoded on the pKM101 or R388, a VirB11 ATPase is also implicated

in pilus assembly, but there is no homologue for VirB11 in the F-T4SS (Sagulenko *et al.* 2001, Hospenthal *et al.* 2017). This suggests that the F-pilus elaboration must rely solely on VirB4-like TraC.

TraC remains static throughout pilus outgrowth, but the F-pilus filament rotates along its central axis during the incorporation of new subunits (Clarke *et al.* 2008). This has been evidenced by immobilising the distal end of the F-pilus and the observation that as the pilus grows, the cell on the other end rotates (Clarke *et al.* 2008). The F-pilus extends through the F-OMCC, which presents a highly dynamic structure that is perfectly fit to mediate the F-pilus extension and retraction (Amin *et al.* 2021). The mature F-pili of up to 20 μm in length have been observed, which is approximately 10-20 times the size of an average *E. coli* cell (Shiomi *et al.* 2009, Wang *et al.* 2009), however on average, they rarely exceed 4 μm (Goldlust *et al.* 2023).

It is unclear whether the recent findings on the biogenesis of the minimised R388 pilus, whereby the platform for pilus elongation is established between VirB5 and VirB6 pentamers (Figure 7C), can be extrapolated to the pili of expanded systems such as the F-pilus. The only obvious homologue of either of these proteins is the inner membrane protein TraG, which combines structural properties of both VirB5 and VirB6 and could be a potential fusion protein, but it is unclear if such fusion could complement both functions of VirB5 and VirB6 (Christie 2016, Bragagnolo *et al.* 2020). Perhaps, TraG requires maturation steps, such as proteolytic cleavage, that would separate the VirB5 and VirB6-like domains, allowing for pilus outgrowth between these domains.

1.3.5 The F-pilus structure

Early structural studies of the F-pilus have brought about the first structural understanding of its helical topology (Wang *et al.* 2009), however with the development of cryo-EM, higher resolution structures of three members of the F-pilus family were recently determined, namely

those of the *E. coli* pOX38, *S. enterica* pED208 and *K. pneumoniae* pKpQIL plasmids (Figure 11) (Costa *et al.* 2016, Zheng *et al.* 2020).

Generally, all F-pili share structural features, presenting around 85 Å in diameter and a hollow lumen of approximately 25 Å (Figure 11), which is suitable to physically fit the travelling DNA. Based on the highest resolution structure (pED208), F-pili architecture can be described as an elongated filament of 5-start helical assembly, easier envisioned as layers of F-pilin pentamers stacked on top of each other, each layer related to the adjacent one by a specific rotational angle and raise, with slight variations in those values seen between all F-pili (Costa *et al.* 2021). In the case of pED208 F-pili, the angle between subunits is 28.2° and the rise is 12.1 Å (Costa *et al.* 2016).

The pED208 TraA F-pilin is a 64-residue protein that folds into three α-helices (α_1, α_2, α_3) (Figure 11). An elongated N-terminus precedes the α_1, which is a short α-helix, followed by the longer α_2 α-helix. What connects α-helices α_2 and α_3 is a loop that extends outwards before folding back into the core of the pilin structure. TraA orientation situates that loop within the pilus's central cavity, while the N- and C-terminal regions lie externally (Costa *et al.* 2016). This arrangement aligns with earlier findings indicating that the N- and C-terminal regions of the F TraA pilin are exposed for phage attachment, necessitating their external positioning (Frost and Paranchych 1988, Costa *et al.* 2016). Following the loop, the α_3 forms a dual-helix structure ending with the pilin's C-terminus.

Notably, the year 2023 finally brought about the first structures of pili associated with minimised systems, namely those produced by the *A. tumefaciens* Ti plasmid and that of *E. coli* pKM101 plasmids, named, respectively, T- and N-pili (Figure 11) (Amro *et al.* 2023, Beltran *et al.* 2023, Kreida *et al.* 2023). Excitingly, conjugative pili of Archea were also documented and present architectural homology to conjugative pili present in Gram-negative bacteria (Beltran *et al.* 2023).

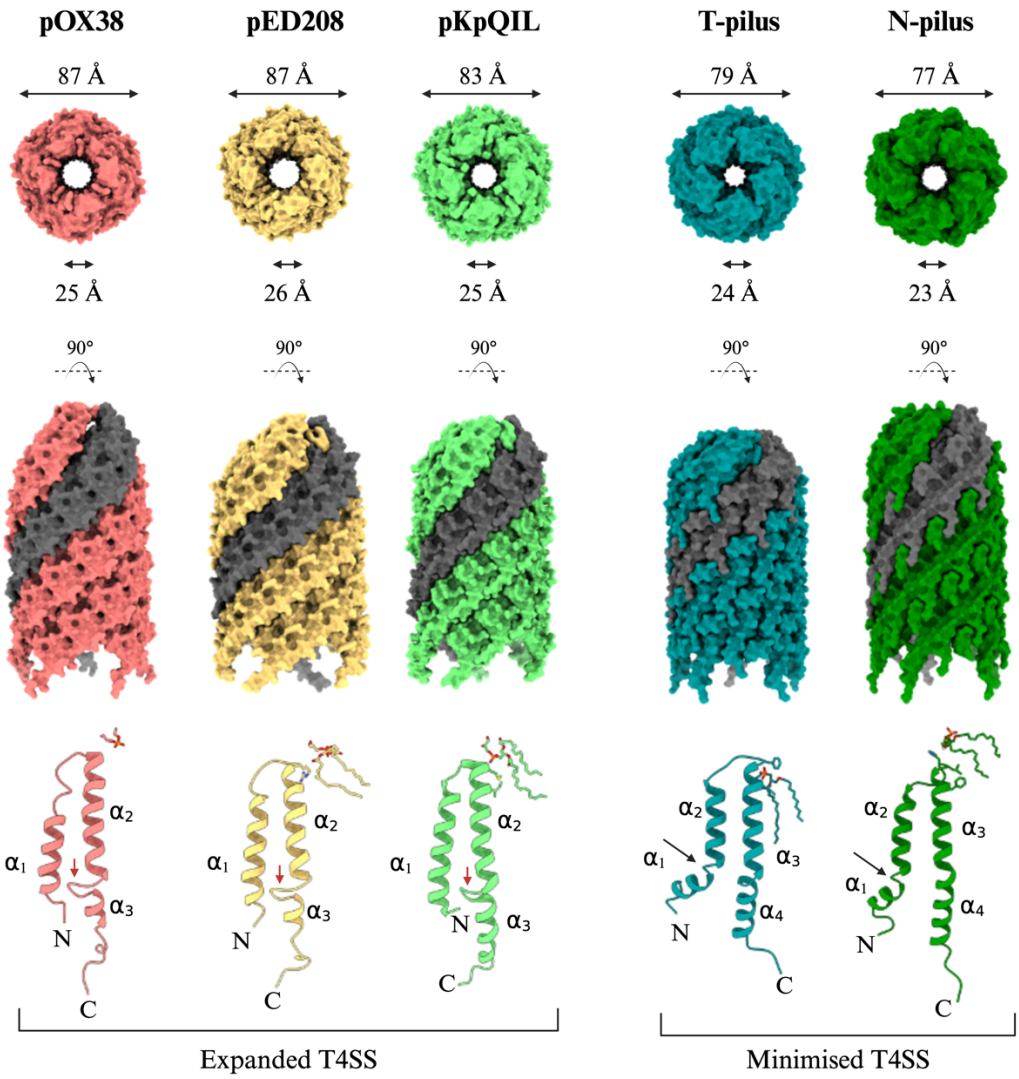


Figure 11. Comparison of conjugative pili species between minimised and expanded T4SS

The top row displays top view of each pilus with indication the entire pilus and its lumen diameters. The middle row shows a side view of the pili. The bottom row includes the atomic structure of cognate pilin subunits from each of the pili, with the associated phospholipids. Pili from expanded T4SS include the F-pili from plasmids pOX38 of *E. coli* (PDB ID: 5LER), pED208 of *Salmonella typhimurium* (PDB ID: 5LEG) and pKpQIL of *K. pneumoniae* (PDB ID: 7JSV). Pili from the minimised systems include: the T-pilus of *A. tumefaciens* (PDB ID: 8EXH) and the N-pilus of the pKM101-encoded T4SS of *E. coli* (PDB ID: 8CW4). Red arrows indicate the loops between α_2 and α_3 in pili of the expanded systems. Black arrows indicate presence of bends between α_1 and α_2 in pili of the minimised systems.

The T- and N-pili share some features described for the F-pili, presenting helical assemblies of 79 and 77 Å in diameter, with lumens of 24 and 23 Å in diameter, respectively (Figure 11). The notable difference is the presence of a bend between α_1 and α_2 helices in T- and N-pilin subunits

(Figure 11), resulting in asymmetry of N- and T-pilin subunits. Its significance to the structure of the entire filament is elusive. It is plausible that the alterations in pilin symmetry contribute to their biophysical properties, tailoring different pili for specialised functions under different environmental conditions. Recent biophysical investigations have revealed that F-pili, which are made up of symmetrical and slender F-pilins, exhibit remarkable flexibility and possess spring-like properties along with notable structural and thermochemical resilience (Patkowski *et al.* 2023). These attributes were shown to be important for enhancing the rate of conjugation and the formation of biofilms by cells carrying F-plasmids, thereby enabling efficient operation of the F-pilus even in turbulent, liquid environments such as the human gastrointestinal tract (Patkowski *et al.* 2023). In contrast to F-pili, conjugative pili of minimised T4SS tend to be shorter and stiffer and are not capable of conjugating in liquid environments (Bradley 1980, Arutyunov and Frost 2013, Virolle *et al.* 2020) or require additional supporting pili that enable liquid matings, as in the case of pili encoded on IncI plasmids (Horiuchi and Komano 1998). Therefore, pili of minimised systems are hypothesised to easily break from the cell surface and thus not be capable of mediating conjugation in turbulent environments, a property that might arise from the asymmetry of pilins imposed by the presence of a bend between α_1 and α_2 helices.

All available structures of pili revealed an unconnected electron density neighbouring each pilin subunit, which was attributed to various types of phospholipid molecules (Figure 11). The F-pilins and N-pilins are associated with phosphatidylglycerol molecules, whereas a different phospholipid has been identified in each study that documented the T-pilus structure, that is: phosphatidyl-ethanolamine (Beltran *et al.* 2023), phosphatidyl-choline (Amro *et al.* 2023) and phosphatidylglycerol (Kreida *et al.* 2023). It is unclear what is the significance of these differences, and perhaps, the discrepancy might come as a limitation of the methods used in each study, as only one of those three studies used phospholipase treatment of purified pili before subjecting samples to mass spectrometry to establish the type of the phospholipid

present (Beltran *et al.* 2023). This step is crucial, as pili fractions are often contaminated with leftovers of the biological membranes, which might overwhelm the signal and thus give a false readout of the phospholipid type. Therefore, phosphatidyl-ethanolamine will be shown as the most probable phospholipid associated with the T-pilus (Figure 11).

The presence of phospholipid groups has not been observed in the molecular structures of other extracellular filaments, such as type IV pili, flagella, or fimbriae, making it an exclusive trait of conjugative pili. Their structural significance remained unidentified until 2023 when by steered molecular dynamics of the F-pilus, it was shown that the presence of phospholipid molecules in the molecular structure is essential to maintaining the integrity of the F-pilus filament by acting as "molecular glue" that preserves F-pilins together (Patkowski *et al.* 2023). The simulations of F-pili devoid of phospholipids exhibited very high vulnerability of the filament and collapsed much more easily, especially upon applying force that imitates the environmental turbulence (Patkowski *et al.* 2023). Therefore, the presence of phospholipids has powerful implications in both - distant and tight routes of conjugation (Figure 4). In distant conjugation, a robust filament will better maintain distant interbacterial connections in a more turbulent filament and protect the travelling DNA. On the other hand, in tight conjugation, the establishment of membrane-to-membrane connections by bacteria requires pulling of the recipient cell via the F-pilus retraction to bring it closer to the donor cell, which undoubtedly exerts significant forces on the filament. The presence of phospholipids allows the F-pilus to withstand these forces.

Why are there different types of phospholipids incorporated by different conjugative pili? Perhaps their specific properties will have implications on the type of substrate that will be transferred through the pili. As all phospholipids in the atomic structures of conjugative pili are directed toward the lumen, they impart changes in the electrostatic charge of their lining. In the case of F- and N-pili, the phospholipids impart an overall negative charge of the lumen,

whereas in the case of the T-pili, the phospholipids impart an overall positive charge. Such charge discrepancies could have evolved as an adaptation of the conjugative pili F- and N-pili to translocate DNA through their lumens, whereby the negative charge would help "lubricate" the lining of the lumen to prevent electrostatic "sticking" of negatively charged DNA to its walls (Costa *et al.* 2016). On the other hand, the T-pilus is responsible for the translocation of a variety of protein effectors, hence the positive charge of its lumen might better reflect the electrostatic interactions between them and the T-pilus lumen (Li and Christie 2018).

1.3.6 The Type IV Coupling Proteins (T4CP)

The type IV coupling proteins (T4CP) are an indispensable component of type IV secretion, found in all conjugative systems and most effector-translocating systems (Gonzalez-Rivera *et al.* 2016, Llosa and Alkorta 2017). Their name reflects their role as the coupling factor between the substrate reception and processing (either DNA or effector protein) and translocation by the T4SS. Consequently, T4CP present a complex and dynamic set of interactions with both, the substrates and the T4SS.

Functionally, T4CP play multifaceted roles in type IV secretion, including substrate recognition, translocation (Moncalián *et al.* 1999, Gunton *et al.* 2007) and signal transduction (Larrea *et al.* 2017). Due to their pivotal role in conjugation, they are being widely explored as potential drug targets aimed at inhibiting bacterial conjugation and the spread of associated AMR genes (Álvarez-Rodríguez *et al.* 2020). Moreover, advances in genetic engineering have introduced exciting applications of T4CP in biotechnology, exemplified by successful DNA delivery into mammalian cells through the utilisation of a T4SS with its cognate T4CP (Guzmán-Herrador *et al.* 2017).

In the cell, T4CP are localised mostly at the cell poles, which was shown for various T4CP (Kumar and Das 2002, Chen *et al.* 2008, Segura *et al.* 2014). In conjugative systems, the cell pole localisation might be a result of the interaction of their T4CP with plasmid partitioning

proteins ParM and ParR, thereby imparting a similar localisation pattern (Gruber *et al.* 2016). That however is contrary to the recent discovery that the preferential exit point of F-plasmid, and hence the localisation of the T4SS apparatus was at the side of the cell, not at its pole (Couturier *et al.* 2023). This might indicate the involvement of conjugative T4CP in a handover between plasmid maintenance and partitioning (cell poles) to plasmid transfer (cell sides).

Structurally, T4CP belong to the superfamily of ATPases associated with various activities (AAA+) that also include the GspE ATPases from T2SS and type IV pilus assembly systems (Christie *et al.* 2014) or the DNA-transporting SpoIIIE/FtsK proteins (Cabezon and de la Cruz 2006). Currently, the atomic structures of three T4CP are available (Figure 12), the first of which is the X-ray crystallography structure of the soluble domains of TrwB, the T4CP from the R388 plasmid (Gomis-Rüth and Coll 2001) (Figure 12A), which became the paradigm for studying T4CP. The second structure is that of the soluble domain Cagβ, the T4CP from the *H. pylori* Cag-T4SS (Figure 12B), which was documented very recently (Wu *et al.* 2023). Both crystal structures showed a homohexameric topology of these T4CP, which does not agree with their oligomeric state when they associate with their cognate T4SS, as shown by NS-EM imaging of the IMC of R388-T4SS associated with the TrwB T4CP (R388-IMC/T4CP) and cryo-ET imaging of the IMC of Cag-T4SS associated with Cagβ T4CP (Cag-IMC/T4CP) (Figure 12C). In the former, TrwB was shown to form dimers sandwiched between VirB4 densities (Figure 12C) (Redzej *et al.* 2017), whereas, in the latter, Cagβ was shown to form a hexamer of dimers assembling an additional ring around the VirB4-like ATPase, CagE (Figure 12C) (Hu *et al.* 2019b).

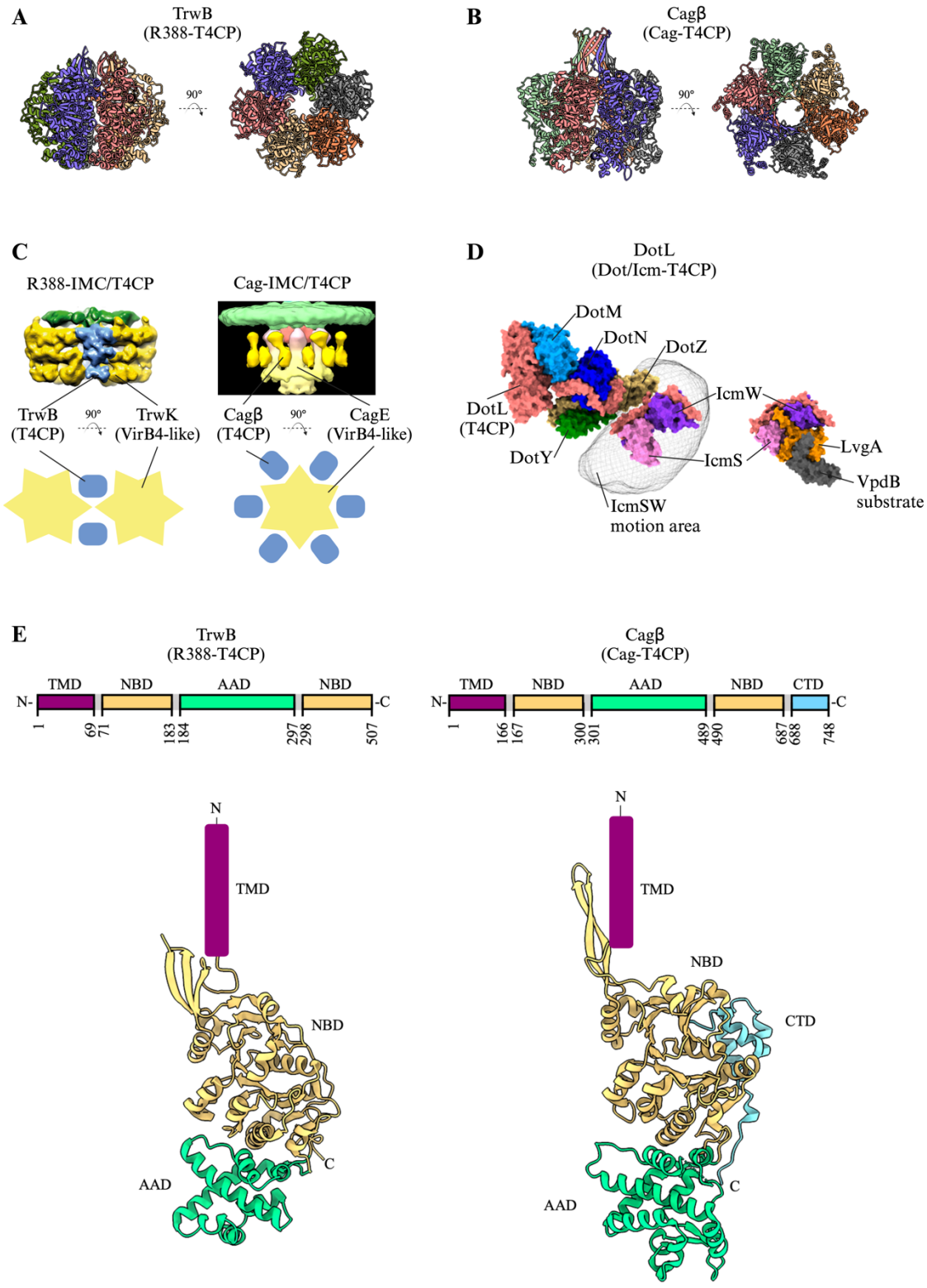


Figure 12. Structures of type IV coupling proteins

A| Crystal structure of hexameric TrwB, the T4CP encoded on the R388 plasmid of *E. coli* (PDB ID: 1GL7). **B|** Crystal structure of hexameric Cagβ, the T4CP of *H. pylori* Cag-T4SS (PDB ID: 8DOL). **C|** Left: NS-EM structure of the R88-T4SS IMC (EMDB ID: EMD-3585) with the associated TrwB

T4CP dimers (blue) sandwiched between VirB4-like ATPase TrwK (yellow). Right: cryo-ET map of Cag-T4SS IMC (EMDB ID: EMD-0634) with the associated hexameric ring of Cagβ T4CP dimers (blue in the diagram) formed around the VirB4-like CagE (yellow). **D|** Cryo-EM structure of DotL, the T4CP of *L. pneumophila* Dot/Icm-T4SS, purified as monomer with an assay of dynamic adaptor proteins that form a type IV coupling complex and an interacting substrate, VpdB. (PDB ID of complex DotLMNYZ-IcmSW structure: 6SZ9, PDB ID of DotL-IcmSW-LvgA-VpdB: 7BWK). **E|** Domain organisation of TrwB and Cagβ. TMD – transmembrane domain, NBD - nucleotide-binding domain, AAD - all-alpha domain, CTD – C-terminal domain.

The third structure of a T4CP was obtained via a cryo-EM study of DotL, the T4CP from *L. pneumophila* Dot/Icm-T4SS (Figure 12D). It reveals the DotL subunit bound to an array of adaptor proteins that altogether facilitate the specificity of DotL to a wide array of substrates carried by the Dot/Icm-T4SS, forming a type IV coupling complex of DotLMNYZ-IcmSW-LvgA, which exhibits a degree of flexibility necessary for efficient substrate binding (Figure 12D). The observed complex was additionally bound to the substrate VpdB (Figure 12D) (Meir *et al.* 2020). Strikingly, in the deposited structure, DotL exists as a monomer, and up to date, no additional structural information is available on the association of DotL with the T4SS, therefore the presence of other functional oligomeric arrangements of DotL cannot be excluded.

The oligomerisation discrepancy observed among T4CP, that is the presence of hexameric (upon crystallisation) and dimeric (upon association with cognate T4SS) TrwB and Cagβ together with a monomeric DotL (in solution) is common for the AAA+ family that all T4CP belong to. It has been shown that AAA+ ATPases may exist in a range of oligomeric states in solution, however, upon receiving a specific activator – such as binding of a substrate or a cofactor - they transiently assume more intricate arrangements to convey their specific functions (Erzberger and Berger 2006). This is true for T4CP, which undergo dynamic transitions during their assembly and/or substrate translocation. Further study is needed to

better understand the mechanism and significance of oligomeric transitions of T4CP to the operation of the T4SS.

On the atomic level, obtaining both TrwB and Cagβ structures gave valuable insight into the anatomy of T4CP subunits, which share an overall similar three-domain topology, consisting of the transmembrane domain (TMD), nucleotide-binding domain (NBD) and an all-alpha domain (AAD) (Figure 12E).

The TMD of T4CP is made up of multiple membrane-spanning helices connected by the periplasmic loop(s) and stands out as the most heterogeneous domain within T4CP. It is responsible for docking of the T4CP to the cognate T4SS apparatus and any modifications result in changes to the subcellular localisation of T4CP (Kumar and Das 2002, Chen *et al.* 2008, Segura *et al.* 2014). Notably, it is a common strategy to truncate TMD during the structural determination of T4CP, thereby designating them "soluble domains", hence TrwB and Cagβ structures lack densities corresponding to the TMD (Figure 12E). That is because for structural determination, membrane-associated proteins must be extracted from the lipid phase, which adds many complications to the purification workflow, such as the need for using detergents.

The NBD of T4CP shows the highest evolutionary conservation out of all T4CP domains. It forms a twisted β-sheet conformation surrounded by α-helices (Figure 12E), resembling that of the FtsK DNA motor (Jean *et al.* 2020), which is involved in the translocation of chromosomal DNA during cellular division. The NBD consists of two motifs, the Walker A and Walker B motifs, which have the capacity to hydrolyse ATP, thereby providing energy for substrate translocation by the T4CP (Ramakrishnan *et al.* 2002, delToro *et al.* 2016, Llosa and Alkorta 2017). The Walker A motif is rich in glycine and is mostly involved in capturing ATP, whereas Walker B plays a direct role in ATP hydrolysis (Ramakrishnan *et al.* 2002, delToro *et al.* 2016). Mutations in either motif have catastrophic consequences not only to the T4CP, but

also to the entire T4SS, as Walker A mutants abolished conjugation whatsoever (Atmakuri *et al.* 2004).

The AAD domain of T4CP reveals an all-helix bundle (Figure 12E) that protrudes into the cytoplasm and resembles the N-terminal domain of recombinase XerD (Gomis-Rüth *et al.* 2001, Whitaker *et al.* 2015). In XerD, this domain is involved in interaction with DNA (Cao *et al.* 1997), hence suggesting a similar role in T4CP. In effector translocating systems, AADs were shown to play a role in substrate recognition, therefore the size and organisation of will AADs vary between T4CP, enabling them to capture different substrates (Whitaker *et al.* 2015, Wu *et al.* 2023).

Additional C-terminal domains (CTDs) are present in some T4CP, characteristic but not exclusive to conjugative systems (Alvarez-Martinez and Christie 2009). These CTDs form tails that function in capturing substrates such as relaxosome components via an extensive network of highly specific inter-residue contacts, making them the major determinant of substrate/DNA specificity to the T4SS (Lu *et al.* 2008a).

1.4 Cryo-electron microscopy (cryo-EM)

1.4.1 Principles of cryo-EM

Structural biology offers a high-resolution insight into bacterial molecular mechanisms by providing three-dimensional atomic maps of proteins. For decades, the dominant technique used for this purpose was X-ray crystallography, which relies on directing the X-ray radiation at crystallised proteins. This method is burdened with many limitations – the main of them being the crystallisation of the protein sample. The ideal protein for this purpose is of smaller size (<100 kDa), can be purified in enormous quantities, maintained at high concentrations and, most importantly, can form crystals. Most of these characteristics disqualify membrane proteins, such as secretion systems (Sgro and Costa 2018).

An alternative arose thanks to recent improvements of transmission electron microscopy (TEM). TEM aims to magnify objects in an analogous way to light microscopy, however a beam of electrons is used instead of photons. This property enables visualisation of biological macromolecules, with resolution proportional to the wavelength of radiation use. In TEM, the wavelength is inversely proportional to the square root of the acceleration voltage – for instance, an electron accelerated by a standard 200 kV voltage yields a theoretical resolution of approximately 1 Å, which is sufficient to visualise atoms (Penczek 2010, Ke *et al.* 2015). The image arises when electrons are scattered off the specimen, which is either dried and negatively stained (negative stain-electron microscopy, or NS-EM) or cryogenically frozen to keep them in the solvent-solubilised and hydrated state (cryo-EM), which adds a layer of protection from the radiation damage (Baker and Rubinstein 2010). In the cryogenic environment, the specimen decays much slower as the breakage of bonds is reduced at lower temperatures (Bock and Grubmüller 2022). Then, thousands of movies of the sample are taken, their frames are aligned and averaged, thereby generating a 2D image, and thousands of those 2D images can be used to reconstruct the final 3D model.

Central to TEM is the electron microscope, which resides in a column with a radiation source on one end and an electron detector at the other, separated by multiple sets of lenses (Figure 13A). The column must be maintained under vacuum to remove potential gas molecules that would scatter electrons. The beam of electrons is generated by an electron gun, located at the top of the column (Figure 13A). As voltage is applied to the needle-like tip of the gun, it behaves as a cathode, from which electrons are drawn and travel further down the column. The resulting beam passes through a set of condenser lenses, which create a magnetic field that alters the trajectory of electrons, focusing them on the specimen (Figure 13A). The condenser system also includes condenser apertures, which ensure that the beam is parallel by removing a fraction of the electrons. Subsequently, the focused beam interacts with the specimen, scattering electrons (Figure 13A). The scattered electrons pass through the objective lens with its aperture, refocusing them and providing the initial magnification, further increased by a set of projector lenses. Finally, the beam reaches the detector camera to generate the image. Modern TEM microscopes are equipped with direct electron detectors (DED), which do not collect images, but rather short movies that consist of multiple frames.

1.4.2 Image generation

The main challenge of cryo-EM imaging is the low signal-to-noise ratio of the imaged specimen because the dose of electrons per surface must be kept low to avoid radiation damage (usually at 40-50 electrons/$Å^2$). This results in particle images that have low contrast, which is defined as the difference in the intensity of the particle versus the intensity of its environment. Biological samples are organic compounds, made up of carbons and hydrogens, which are submerged in an aqueous, hydrogen- and oxygen-rich environment. All of these are light atoms and hence not very electron-dense, thus the probability of an electron scattering event is low. Therefore, the total number of electrons that entered the specimen will be the same as the number of electrons that exit the sample, resulting in no intensity differences between the

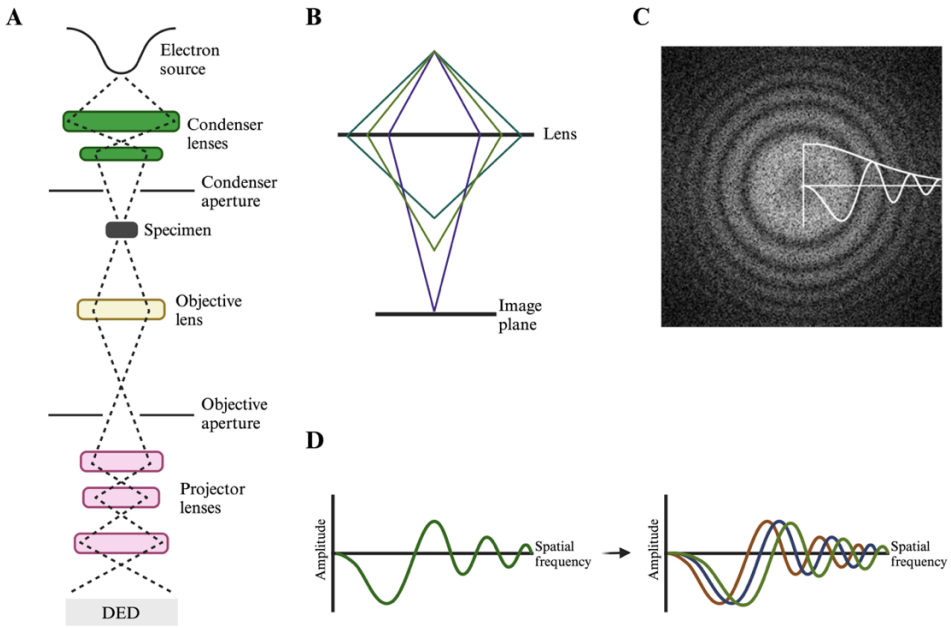


Figure 13. Principles of electron microscopy

A| Electron microscope anatomy. The electrons are accelerated from an electron source and focused onto the specimen by a set of condenser lenses, followed by focusing by the objective lens. The magnification is acquired by a set of projector lenses. Two apertures – the condenser aperture and the objective aperture – help remove signal from high-angle electrons. Eventually, the scattered electrons reach the direct electron detector (DED), which collects the magnified movies of the specimen. **B|** Representation of the spherical aberration suffered due to imperfection of the EM lenses, electrons passing further away from the lens centre, will be focused at sharper angles. **C|** Example Fourier transform of an image, also known as a power spectrum. Bright areas indicate frequencies where image details are well-preserved, known as Thon rings. By analysing the power spectrum, the Contrast Transfer Function (CTF) can be estimated, revealing how different detail sizes (spatial frequencies) are affected by factors like defocus and aberrations. **D|** To recover information that is lost at point of zero amplitude (where CTF crosses the X axis), images are taken at different values of defocus, which alters their CTF and hence information is recovered when images are averaged.

particle and the environment. The change between electrons entering and leaving the specimen is called the amplitude contrast, which is negligible during the imaging of biological samples. Instead, what facilitates cryo-EM imaging of such samples is the phase contrast, which occurs when the specimen scatters the entering electrons at different angles, resulting in different path lengths through the sample. Phase contrast is generally poor and needs to be enhanced by a

collective effect of two mechanisms – imaging out of focus and the spherical aberration of the lenses.

While imaging closer or in focus, the scattered information will be recombined with the non-scattered beam – that is because biological samples scatter electrons very weakly. However, when imaging slightly under focus, the scattered and non-scattered electrons will be phase-shifted, with the scattered electrons having a delayed phase, depending on the angle of scattering (Franken *et al.* 2020). While this provides better contrast, imaging at a defocus causes a loss of information, limiting attainable resolution.

Spherical aberration is one of the aberrations that magnetic lenses suffer from. It arises from the non-uniform magnetic field within the lens: when an electron passes closer to the centre of the lens, it experiences a weaker electromagnetic field strength than when passing further away from it (Figure 13B). Consequently, these central electrons converge at a point closer to the lens compared to the intended focal plane, resulting in overfocusing (Franken *et al.* 2020). Similar to underfocusing, overfocusing introduces phase shifts in the electron wavefunction, leading to an increased contrast.

Defocus and spherical aberration can be analysed using the contrast transfer function (CTF). This function describes how details of varying sizes (called spatial frequencies) in the image are affected by these aberrations. At certain frequencies, the CTF goes to zero, indicating a complete loss of information for that specific detail size. In the Fourier transform of an image also known as the power spectrum, this appears as dark gaps between the characteristic Thon rings (Figure 13C). Therefore, the CTF can be estimated from the power spectrum because the high contrast areas (details preserved) correspond to bright rings in the Fourier transform, which signify peaks in the CTF (Figure 13C). Conversely, black gaps, where detail is lost, represent zero-crossings in the CTF. To overcome this information loss, images should be acquired at a range of defocus levels (Figure 13D). Since each defocus setting affects different

spatial frequencies, averaging these images helps mitigate the overall information loss (Figure 13D).

1.4.3 Sample preparation

Cryo-EM offers two complementary approaches for structural elucidation: single-particle analysis (SPA) and cryo-ET with subtomogram averaging (STA). STA excels at studying large protein complexes within their native cellular context (*in situ*), enabling visualisation of their overall organisation, albeit at a lower resolution. This technique employs a "tilt series" approach, where numerous two-dimensional images, called subtomograms, are captured from different orientations of the target protein within intact cells. Subsequent computational averaging of these subtomograms allows for the reconstruction of a three-dimensional model, or tomogram, revealing the general architecture of the protein complex. In contrast, SPA offers high-resolution structural analysis but necessitates meticulous sample preparation (Figure 14). Ideal SPA samples require homogeneity, symmetry, and a size that facilitates a favourable signal-to-noise ratio during imaging (Wentinck *et al.* 2022). To achieve this, a multi-step purification process is required to isolate the protein of interest from cellular lysates, followed by optimisation of the buffer composition to ensure a uniform biochemical state. This sample preparation stage can often be the most time-consuming aspect of the SPA workflow.

For successful data acquisition in cryo-EM, the protein concentration in the sample needs to be optimised. The ideal concentration maximises the number of particles visible in each image while preventing particle overlap. Before applying the sample, the grid must be "glow discharged" to create a hydrophilic surface that promotes sample adherence within the grid holes. This involves placing the grid in a chamber and subjecting it to ionisation. Once the sample is applied, the excess solution is blotted away. The blotting process needs careful control to achieve the desired ice thickness on the grid. This ice layer serves a critical purpose: it must be thick enough to fully encapsulate the particles and prevent them from interacting

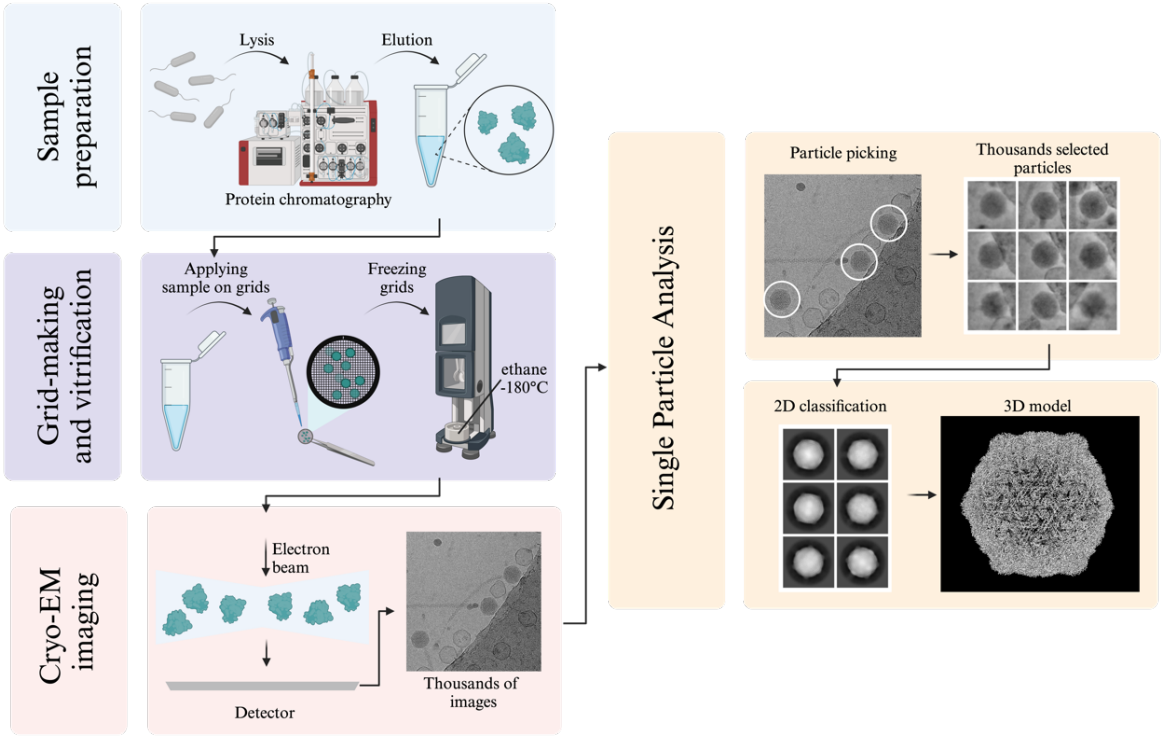


Figure 14. Typical cryo-EM workflow overview

The sample preparation (blue) includes the extensive process of protein purification, biochemical characterisation and preparation for structural determination. The sample needs to be applied on grids and vitrified in liquid ethane (purple). Resulting grids can then be imaged (red) by passing an electron beam through ice-suspended particles. Thousands of images are normally collected during one collection. The Single Particle Analysis (yellow) is a computational process, in which particles are picked from the micrographs, extracted and then subjected to 2D classification. These 2D classes can then be used to generate a 3D model.

with the air-water interface, yet thin enough to avoid particle stacking and multi-layer formation (Figure 14). Furthermore, for highly asymmetric particles, the ice needs to be thick enough to allow all possible particle orientations, ensuring a random distribution of particle views in the final reconstruction. The blotting time and force are manually adjusted to achieve this optimal ice thickness.

The grid with the sample is then plunge-frozen in a liquid that was cooled down to cryogenic temperatures (Figure 14). This rapid freezing process is crucial to preserve the sample in its

native state, as slow freezing allows water molecules to form ice crystals that would damage the specimen. Liquid ethane, with its high heat capacity, is preferred over liquid nitrogen because it achieves ultra-fast freezing times (less than 200 μs) (Dubochet *et al.* 1988). This rapid cooling vitrifies the sample, essentially trapping the molecules in a glassy, non-crystalline state that mimics their original, hydrated environment (Dubochet *et al.* 1988). This technique also prevents conformational changes that might occur during slower cooling. If freezing is too slow, molecules rearrange to adapt to the changing temperature, leading to a wider range of conformational states and ultimately lower resolution in the final image (Bock and Grubmüller 2022).

Despite being a conceptually straightforward process, freezing particles for cryo-EM includes a complex network of factors that affect how particles are distributed and oriented within the ice. In recent years, multiple strategies have arisen to overcome some common difficulties of sample preparation, such as chemical crosslinking, buffer optimisation, adding protein scaffolds (Yeates *et al.* 2020) or megabodies (Uchański *et al.* 2021), altering grid surface, etc. All these novel technologies aim at removing limitations of cryo-EM SPA, for example, by increasing the protein size or enforcing protein stability.

1.4.4 Data processing

The most optimal frozen grids can now be used for the collection of thousands of movies using automated collection on a designated cryo-EM microscope (Figure 14). Once such a dataset is ready, one can advance to data processing. As DED collect movies, the frames capture slight movements of particles that occur during imaging (Figure 15A), which can be tracked throughout the frames (Figure 15B), and then aligned computationally to generate a single image devoid of movement-induced blur of particles (Figure 15C). This process is called motion correction this is usually the first step of cryo-EM data processing.

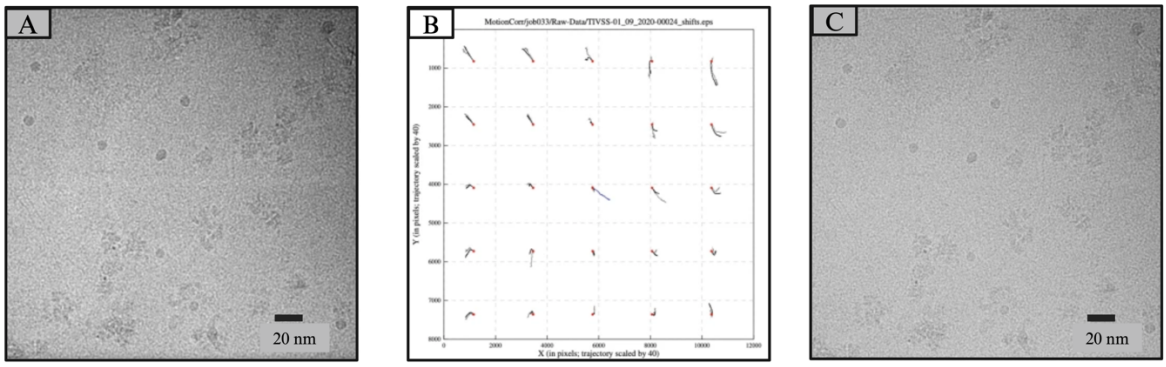


Figure 15. Motion correction of a cryo-EM dataset

A| Representative uncorrected image from a Cryo-EM data collection**. B|** Motion patches between frames in X and Y directions. **C|** Motion corrected image. Adapted from Ignatiou *et al.* (2024).

Following motion correction, comes the CTF correction - a crucial step that tackles distortions introduced by the CTF (Scheres 2012, Sheth *et al.* 2015). Uncorrected CTF shows up as alternating bands of positive and negative contrast in the image's Fourier transform (Figure 13C). Therefore, CTF correction involves manipulating the image in Fourier space. Once the CTF is estimated based on the power spectrum, the correction is carried out by a phase flip. A square wave function, with a value of -1 for negative contrast regions and +1 for positive regions, is multiplied across the image in the Fourier space (Scheres 2012). This effectively flips the negative phases of the CTF (Figure 16), essentially reversing the distortions and restoring the accurate phase information in the original image (Scheres 2012). With CTF correction, high-resolution details of the biological molecule become more prominent, facilitating a more precise 3D reconstruction in the later stages of cryo-EM data processing.

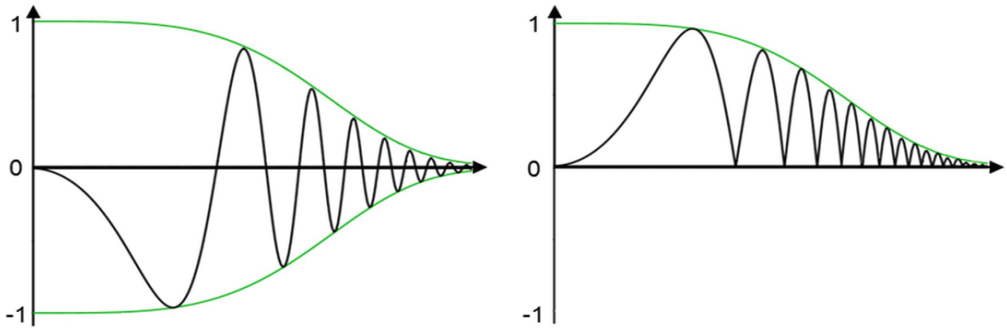


Figure 16. Principle of CTF correction

The uncorrected CTF (left) alternates between negative and positive contrasts based on spatial frequencies. Information loss occurs only when the CTF intersects the zero line. The negative segments of the uncorrected CTF (right) are inverted to positive through phase flipping. Adapted from Costa *et al.* (2017).

Processed micrographs can then be used to train the processing software to recognise and pick particles from each micrograph – that is accomplished by providing an estimated size of the particle and manually picking a pool of representatives. Automated picking can then be applied to the entire dataset, and then individual particles can be extracted from micrographs to generate particle stacks. These are now subjected to 2D classification, where particles are aligned together and grouped into classes by overall similarity. A resulting 2D class is an average of a distinctive group of particles that share similarities. Ideally, different views of the particle will give separate 2D classes.

Because the 2D classes will represent projections of a 3D object, all of them will be translated into the original object by the concept of the central slice theorem. It describes how each of the Fourier transforms of each 2D projection represents a different plane passing through the 3D Fourier transform of the object (Figure 17). Since an ideal cryo-EM sample contains particles in random orientations, all slices of the 3D Fourier transform are present in the dataset (Figure 17). Projections will have a relative "common line" – a line of identical pixels that represent their intersection in the 3D Fourier transform of the object. Based on this, the relative angles between planes can be calculated, and the initial 3D map can be generated (Figure 17). This

process is called 3D reconstruction. An initial 3D map can then be used as the reference for the following round of matching different projections and refining relative angles between them. The resulting refined structure provides an average of hundreds of thousands of particles, which reveals high-resolution features of the map (Figure 17).

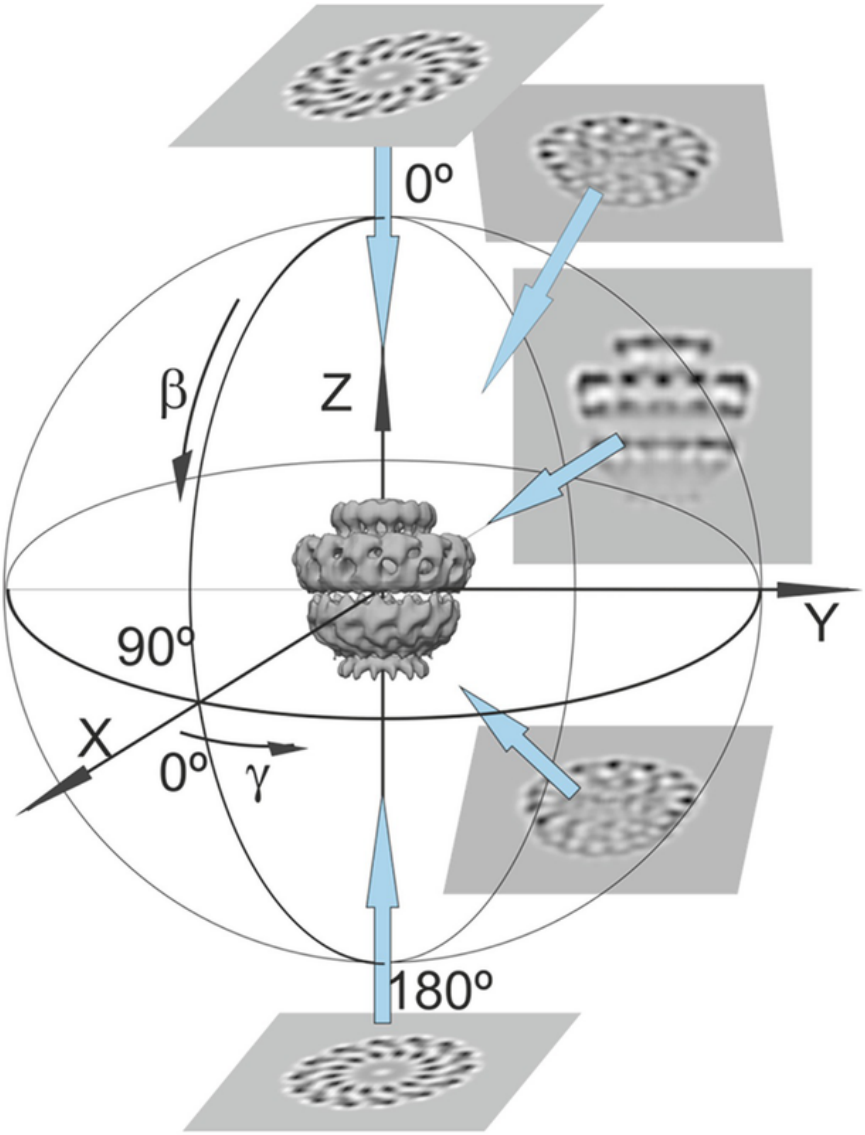


Figure 17. 3D reconstruction based on 2D classes

Each 2D will relate to the final 3D object by a certain angle. Thus, a set of 2D projection that encompasses the full 3D object can be therefore back-projected to generate the 3D model. Figure adapted from Costa *et al.* (2017).

1.4.5 Model building and validation

The atomic model of a protein can be built based on the generated 3D map and the amino acid sequence. At resolutions higher than 3.7 Å, domains of α-helices and β-sheets are clearly discernible together with side chains of large amino acids, which allow for *de novo* tracing of polypeptide chains (Figure 18A) (Ignatiou *et al.* 2024). This can be done through softwares such as Coot, Phenix or Isolde, which had been developed for X-ray crystallography (Emsley *et al.* 2010, Afonine *et al.* 2018, Croll 2018). The resulting models need to be evaluated for correctness and consistency, through parameters such as Ramachandran plots and steric clashes.

Ramachandran plots are based on dihedral torsion angles across the protein, that is φ and ψ angles in the N-C_α-C bonds, as shown in Figure 18B (Sasisekharan 1963). In a Ramachandran plot, φ and ψ angles are plotted for all residues in the protein model (Figure 18C). Certain regions of the plot indicate angle combinations that are more favourable than others, with specific regions indicating α-helices and β-sheets (Figure 18C) (Sasisekharan 1963). Residues falling outside these allowed regions are unlikely due to steric clashes, which occur when atoms are too close together, causing repulsion and hindering the protein from folding properly (Figure 18D). The final assessment of the atomic model is generated using MolProbity (Williams *et al.* 2018), which takes into consideration torsion angles and steric clashes. Once those factors are satisfied, the structure can be analysed and deposited in the Protein Data Bank (PDB).

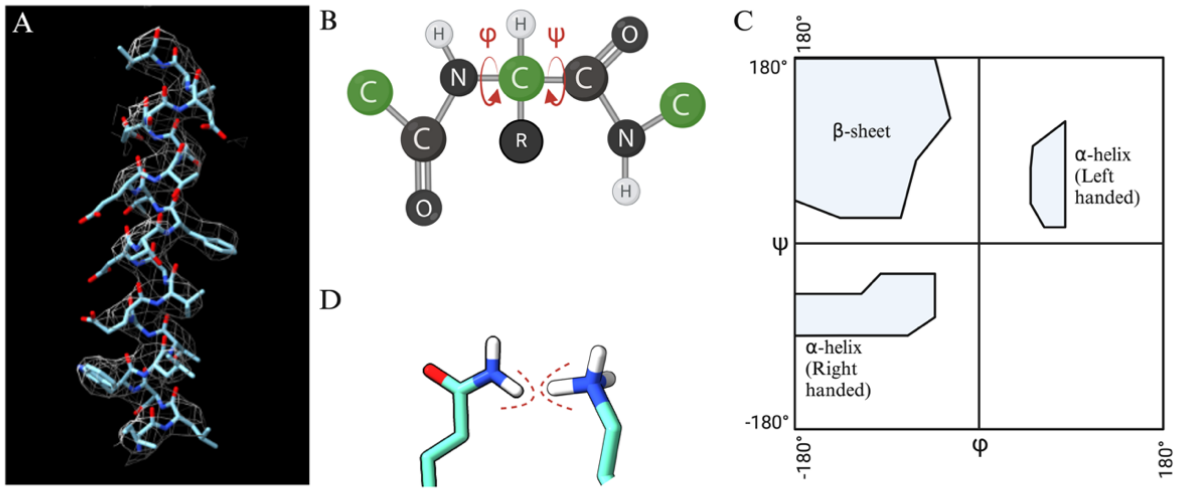


Figure 18. Building of the atomic model and its evaluation

A| Building a model of an α-helix involves fitting residues (coloured in cyan) in the experimentally obtained 3D map (white mesh). The map shows clearly discernible densities for larger side chains. **B|** The N-C_α and C_α-C bonds can freely rotate, being only limited by the size and character of the R-groups. The torsion angles of these bonds are called φ and ψ angles (shown in red). **C|** Example of a Ramachandran plot showing favourable φ and ψ combinations for α-helices and β-sheets (regions in blue). Regions in white indicate steric collisions**. D|** An example of steric collision, where atoms were placed too close together, indicating insufficient optimisation of the atomic model.

1.5 Aims of this thesis

This thesis aims to understand the molecular mechanisms underlying bacterial conjugation, specifically how it contributes to the spread of AMR. Chapters 2 to 4 aim to describe different events during conjugation, including substrate DNA recognition by TraD (Chapter 2); adaptations of the F-pilus to successfully form mating pairs in the human microbiome (Chapter 3); and the mechanism of distant conjugative transfer (Chapter 4). Chapter 5 aims to explore the existence of pathogenic T4SS in AIEC associated with L-CD. Detailed goals are described below.

Chapter 2: Structural and biochemical analysis of the F-plasmid T4CP, TraD

This chapter focuses on TraD, the F-plasmid coupling protein to understand its structural adaptations and mechanisms for substrate capture and transport. The specific objectives were:

- determine the structure of TraD and analyse its implications in TraD function in the recognition and transfer of substrate DNA;
- assess TraD ability to interact with DNA and hydrolyse ATP;
- investigate TraD capacity to form higher-order oligomers and the factors influencing this process.

Chapter 3: Biomechanics of the F-pilus and the role of its phospholipids in conjugation

This chapter investigates how the environment affects mating pair formation by the F-pilus and aims to understand its structural adaptations that allow for high conjugative efficiency in turbulent conditions. The specific objectives were:

- evaluate the effect of the environment on the efficiency of F-pilus-mediated conjugation and biofilm formation;
- understand the dynamic and mechanical properties of the F-pilus;

- asses the structural adaptations of the F-pilus that convey its biomechanical properties;
- define the role of phospholipids in the atomic structures of conjugative pili.

Chapter 4: Direct evidence of ssDNA transport through the F-pilus

This chapter explores the distant transfer route of conjugation to explore the capacity of the F-pilus to accommodate the travelling DNA. The specific objectives were:

- develop a fluorescently labelled F-pilus to aid its visualisation;
- record the colocalisation of the travelling DNA within the F-pilus during distant transfer.

Chapter 5: Identification of a novel T4SS, essential for AIEC colonisation in L-CD

The final chapter investigates the T4SS found in the NGR857 AIEC strain associated with L-CD patients. The specific objectives were:

- characterise the structure and function of the recently discovered plasmid-encoded T4SS in the NRG857c;
- understand the mechanism by which this T4SS aids AIEC pathogenesis despite presenting an unusual operon;
- identify other virulence factors within the strain, such as other T4SS.

Chapter 2 Structural and biochemical analysis of the F-plasmid T4CP, TraD

2.1 Introduction

Bacterial conjugation relies on delivering the substrate nucleoprotein complex (comprising the relaxosome and plasmid DNA) to the cytoplasmic entrance of the T4SS apparatus for DNA translocation to the recipient cell (Figure 5). Initially, the relaxosome complex must recognise the substrate DNA and introduce a single-stranded cleavage, potentiating it for translocation through the lumen of T4SS (Figure 5). Subsequently, the type IV coupling proteins (T4CP) guide the DNA-relaxosome complex to the translocation machinery, thereby coupling two main processes of bacterial conjugation: DNA processing by the relaxosome and translocation by the T4SS.

In F-plasmid conjugation, the F-plasmid encoded T4CP - TraD, recruits the DNA-bound TraM, a relaxosome component that specifically recognises the *oriT* locus within the F-plasmid (Figure 5) (Lu *et al.* 2008a). The subsequent steps of TraD operation are unknown, however early studies show the TraD affinity for DNA-paved cellulose columns, suggesting a potential role in DNA binding (Panicker and Minkley 1992). This can indicate that TraD serves as a DNA-translocating ATPase, potentially by a mechanism similar to that of FtsK – a hexameric AAA+ ATPase involved in the translocation of DNA during cell division (Jean *et al.* 2020). The ATPase activity of TraD was never documented, however, T4CP of other T4SS facilitate substrate recognition and translocation by ATP hydrolysis (Cabezon *et al.* 1997, Christie and Vogel 2000, Cascales and Christie 2004b, Cabezón *et al.* 2015), implying a similar role of TraD.

The prominent feature of T4CP is their ability for dynamic transitions between oligomeric states. Atomic structures of TrwB (T4CP from R388-T4SS) and Cagβ (T4CP from the *H. pylori*

Cag-T4SS) exhibit homohexameric topologies when crystallised (Figure 12A&B) (Gomis-Rüth and Coll 2001, Wu *et al.* 2023), however in the solution they may exist as monomers or dimers (Redzej *et al.* 2017, Wu *et al.* 2023). This stands in contrast with their topology upon association with their cognate T4SS, where TrwB was shown to form dimers sandwiched between VirB4-like subunits of TrwK, whereas Cagβ forms a ring around the VirB4-like CagE as a hexamer of dimers (Figure 12C) (Redzej *et al.* 2017, Hu *et al.* 2019b). On the other hand, the cryo-EM study of DotL (T4CP from *L. pneumophila* Dot/Icm-T4SS) reveals a monomeric structure (Figure 12D) (Meir *et al.* 2020). This oligomerisation variability, typical in the AAA+ family, implies dynamic transitions during T4CP assembly and/or substrate translocation, regulated by specific triggers like substrate recognition or signal transduction (Erzberger and Berger 2006). Further studies to understand the mechanisms underlying T4CP oligomerisation transitions are needed.

Little structural information on TraD is available. It is limited to the seven C-terminal residues of TraD ($TraD^{711-717}$), which were co-crystallised with TraM (Lu *et al.* 2008b), providing insight into the high specificity between TraD-TraM interactions during substrate DNA capture by TraD. Early topological studies evidenced the localisation of TraD to the inner membrane, with both N- and C-termini facing the cytoplasm (Lee *et al.* 1999), and the subsequent *in vivo* crosslinking studies of TraD have identified three residues essential for oligomerisation of TraD, that is K273, A392 and E358 (Haft *et al.* 2007). Nevertheless, the structural details of TraD, including the topology of its domains, the oligomeric states, potential transition triggers and biochemical properties, such as DNA binding capacity or ATPase activity remain unknown.

Here, I use an AlphaFold2-powered cryo-EM analysis of TraD to reveal its high structural similarity to TrwB and a dynamic oligomerisation pattern. I identify the ability of TraD to bind DNA and hydrolyse ATP, and I furthermore attribute these functions to specific structural domains of TraD, mapping the Walker A and B motifs responsible for ATP hydrolysis, and the

positively charged lumen potentially involved in interaction with DNA. I show the evident interaction between TraD and TraM, and, lastly, I reveal that ATP hydrolysis is a trigger for oligomerisation of TraD. Altogether, my findings provide structural and biochemical detail to the operation of TraD as the T4CP during conjugation of the F-plasmid, which I culminated here into an updated model of conjugation.

2.2 Experimental procedures

2.2.1 Materials

Bacterial strains (Table 1) and plasmids (Table 2) used in this study.

Table 1. Bacterial strains

Bacterial strain	Information	Source
BL21 (DE3)	Common expression strain of *E. coli.*	Costa Lab

Table 2. Plasmids and genetic constructs used in this study

Plasmid	Information	Source
pASK_TraD$^{130\text{-}717}$	Used for overexpression of the soluble domain of TraD, that is ΔN130-TraD with an N-terminal Strep-Tag II. Contains a tetracycline-inducible promoter. Amp^{R}	Ilangovan Lab
pASK_Hcp-TraD$^{130\text{-}717}$	Used for overexpression a fusion between hexameric scaffold Hcp and ΔN130-TraD with an N-terminal Strep-tag II. Contains a tetracycline-inducible promoter. Amp^{R}	Ilangovan Lab
pET28a_TraM	Used for overexpression of TraM protein with an N-terminal 6x his-tag under IPTG-inducible promoter. Kan^{R}	Ilangovan Lab
pASK_TraD	Used for overexpression full length TraD with an N-terminal Strep-tag II. Contains a tetracycline-inducible promoter. Amp^{R}	Ilangovan Lab

2.2.2 Purification of TraD$^{130\text{-}717}$ and Hcp-TraD$^{130\text{-}717}$

For purification, the TMD of TraD was removed to prevent TraD insertion in the membrane, so only the soluble domain of TraD was used, henceforth referred to as TraD$^{130\text{-}717}$. Purification of both constructs followed the protocol outlined below with slight changes where indicated.

pASK_TraD$^{130-717}$ or pASK_Hcp-TraD$^{130-717}$ were transformed into *E. coli* BL21 and a single resulting colony was inoculated into LB supplemented with ampicillin (100 µg/ml) and incubated with shaking at 37°C overnight. On the next day, 30 ml of the overnight culture was transferred to 3L of LB with ampicillin (100 µg/ml) and grown at 37°C until the cultures reached an OD_{600} of 0.6-0.8. The expression was induced by addition of anhydrotetracycline (200 µg/l) and maintained overnight with shaking at 18°C. The cells were then harvested through centrifugation at 7,000 rpm for 7 minutes. The cell pellet was resuspended in 50 mM HEPES pH 7.2 300 mM NaCl 2 mM $MgCl_2$ buffer with a protease inhibitor tablet (Roche). Subsequently, cell lysis was achieved through double cell disruption at 27 kpsi. To eliminate cell debris, the lysed mixture underwent centrifugation at 35,000 g for 30 minutes at 4°C. For purification, the clarified lysate was applied on the Strep Trap XT column (Cytiva) and subjected to washing with 50 mM HEPES (pH 7.2) 300 mM NaCl 2 mM $MgCl_2$. Then, the proteins bound to the column were eluted using the same buffer supplemented with 10 mM desthiobiotin at a flow rate of 1 ml/min. Elution fractions were collected at 0.5 ml intervals and visualised using SDS-PAGE. The elution fractions obtained from affinity chromatography were pooled and 500 µl was applied on size-exclusion Superose 6 Increase 10/300 GL column (Cytiva) in a 50 mM HEPES (pH 7.2) 300mM NaCl 2mM $MgCl_2$. Elution fractions were collected at 0.5 mL intervals and visualised using SDS-PAGE.

SDS-PAGE was conducted using Novex WedgeWell SDS-polyacrylamide gels (10-20%) (Thermofisher). SDS-PAGE loading buffer (Thermo-Fisher) was added, and the samples were heat-treated at 95°C for 3 minutes. Gel electrophoresis was conducted at 200 V for 50 minutes in SDS-PAGE running buffer (Tris, glycine, SDS), and visualisation was accomplished through overnight staining with Coomassie blue stain (Thermo-Fisher).

2.2.3 Purification of TraM

BL21 cells were transformed with pET28a_TraM, and a single colony harbouring was inoculated into LB supplemented with kanamycin (20 µg/mL) at 37°C overnight. On the next day, 30 ml of the overnight culture was transferred to 3L of LB with kanamycin (20 µg/ml) and grown at 37°C until the cultures reached an OD_{600} of 0.6-0.8. The expression of TraM was induced by addition of IPTG (1mM) and maintained overnight at 18°C. The cells were then harvested through centrifugation at 7,000 rpm for 7 minutes. The cell pellet was resuspended in buffer HEPES pH 7.2 100 mM NaCl supplemented with a protease inhibitor tablet (Roche). Subsequently, cell lysis was achieved through double cell disruption at 27 kpsi. To eliminate cell debris, the lysed mixture underwent centrifugation at 35,000 g for 30 minutes at 4°C. The clarified lysate was applied on His-trap HP column (Cytiva) and washed with HEPES pH 7.2 100 mM NaCl buffer supplemented with 10 mM imidazole. The proteins bound to the column were then eluted using a linear imidazole gradient ranging from 10 mM to 1000 mM over 50 ml at a flow rate of 1 ml/min. Elution fractions were collected at 0.5 ml intervals. Fractions corresponding to TraM elution were dialysed overnight at 4°C into HEPES pH 7.2 50 mM NaCl overnight, and then bound to anion-exchange HiTrap Q column (Cytvia) and eluted in a gradient of 50 mM to 1000 mM NaCl over 30ml at a flow rate of 1 ml/min. Fractions corresponding to the second peak were combined and concentrated using 1,000 MWCO concentrator (Merck) to 500µl and applied on Superose 6 Increase 10/300 GL column (Cytiva) equilibrated with HEPES pH 7.2 100 mM NaCl buffer. The peak corresponding to TraM was collected and visualised by SDS-PAGE.

2.2.4 $TraD^{130-717}$ ATPase assay

ATPase assays were conducted using Malachite Green Phosphate Assay Kit (Sigma) following the manufacturer's instructions. The initial concentration of ATP was 0.25 mM ATP, and

$TraD^{130-717}$ concentration was 1µM. The base reaction buffer was HEPES pH 7.2 100mM NaCl, altered by additives as indicated in the figure.

2.2.5 DNA binding assay

An increasing amount of $TraD^{130-717}$ (from 0 nmol to 3 nmol) was incubated with 200 fmol of 3.2 kbp dsDNA in 20 µl reaction volume containing HEPES pH 7.0 100 mM NaCl buffer for 15 minutes on ice, followed by resolving on 0.5% agarose in 1xTBE buffer for 2 hours at 90V. The resulting gel was stained in 1x Sybr Safe for 30 minutes and imaged using UV light.

2.2.6 Trypsin digestion assay

To observe changes in trypsin digestion patterns upon addition of dsDNA to $TraD^{130-717}$, 1 mg/ml of $TraD^{130-717}$ was incubated with or without 50 ng of 3.2 kbp dsDNA in buffer containing HEPES pH 7.0 100 mM NaCl 2 mM $MgCl_2$. After 20 min incubation on the ice of $TraD^{130-717}$ with or without dsDNA, the solution was transferred to room temperature and 0.2 mg/ml of trypsin was added. At timepoints 0, 15, 30, 45, 60 min after addition of trypsin, aliquots of the reactions were taken out and denatured by boiling for 5 minutes at 95°C to stop digestion. Such prepared samples were subjected to SDS-PAGE.

2.2.7 TraD-TraM pulldown procedure

To assess the interaction between TraD and TraM, equal amounts of lysates of both proteins were prepared as indicated above for each protein. The clarified lysates were combined and applied on a Strep Trap XT column (Cytiva) column, washed extensively with buffer HEPES pH 6.8 100 mM NaCl and eluted with the same buffer containing 10 mM desthiobiotin. Then, the eluted fractions were applied on the his-trap column, washed extensively with buffer HEPES pH 6.8 100 mM NaCl and eluted with the same buffer supplemented with 500 mM imidazole. Such prepared sample was dialysed into buffer HEPES pH 6.8 100 mM NaCl, diluted to 0.05 mg/ml total protein concentration and applied on EM grids as indicated below.

2.2.8 Chemical crosslinking of TraD$^{130-717}$

For small-scale assays, purified TraD$^{130-717}$ in buffer HEPES pH 6.8 100 mM NaCl 2 mM $MgCl_2$ was supplemented with varying concentrations of glutaraldehyde and crosslinked for 30 minutes at room-temperature. The reaction was quenched by addition of 50 mM final concentration of Tris pH 6.8. The reaction was supplemented with different additives as indicated in the assay, and the formation of crosslinked populations was assessed via SDS-PAGE.

For large-scale crosslinking of TraD$^{130-717}$, purified TraD$^{130-717}$ after affinity chromatography in buffer HEPES pH 6.8 100 mM NaCl was supplemented with 2 mM $MgCl_2$ and 10 mM ATP and incubated at room temperature for 10 minutes. Then, 0.1% of glutaraldehyde was added, and crosslinking was allowed to proceed for 30 minutes, after which it was quenched with 50 mM Tris pH 6.8. The crosslinked TraD$^{130-717}$ was then applied on size-exclusion chromatography with Superose 6 Increase 10/300 GL column (Cytiva) in buffer HEPES pH 6.8 100 mM NaCl. Such prepared sample was diluted EM grids.

2.2.9 Negative stain EM (NS-EM) procedures

For NS-EM procedures, fractions of TraD or TraD with additives were prepared as indicated above. 10 µl of such fraction was diluted to ~0.04 mg/ml in buffer HEPES pH 6.8 100 mM NaCl or buffer HEPES pH 6.8 100 mM NaCl 0.03% DDM in the case of full-length TraD and applied on glow discharged 300 mesh carbon-coated grids (Agar Scientific). Following two subsequent washes with 10 µl water, the grid was stained with 10 µl 0.2% uranyl acetate for 60s. The grids were imaged on a Technai12 Spirit transmission electron microscope (FEI) operating at 120 kV and equipped with a 2K Eagle camera.

2.2.10 Cryo-EM procedures

For cryo-EM imaging of crosslinked TraD$^{130-717}$, fractions were prepared as in *Chemical crosslinking of TraD$^{130-717}$*. Following size-exclusion chromatography, fractions were diluted

to 0.5 mg/ml and 4 µl was applied on copper Quantifoil 1.2/1.3 grids. Grids were vitrified on vitrobot (FEI) operating at 95 % humidity at 4°C with a wait time of 1 min and blot time of 4 s before freezing in liquid ethane. The images were collected on Glacios microscope operating at 200 kV with a Falcon 4 direct electron detector with a nominal defocus range of −0.7 to −3.0 µm. Image processing was done with CryoSPARC (Punjani *et al.* 2017) with no symmetry applied.

For cryo-EM imaging of Hcp-TraD$^{130\text{-}717}$, 4 µl of 0.3 mg/ml TraD was applied on copper Quantifoil 1.2/1.3 grids with ultrathin carbon film (300 mesh, Agar Scientific. Grids were vitrified on vitrobot (FEI) operating at 95 % humidity at 4°C with a wait time of 1 min and blot time of 3.5 s before freezing in liquid ethane. The images were collected on Glacios microscope operating at 200 kV with a Falcon 4 direct electron detector with a nominal defocus range of −0.7 to −3.0 µm. Image processing was done with CryoSPARC (Punjani *et al.* 2017) with C6 symmetry applied.

2.3 Results

2.3.1 TraD does not form hexamers in solution

For protein purification, the N-terminal transmembrane domain (TMD) of TraD was removed to prevent TraD insertion in the membrane. The removal of the TMD is common during structural and biochemical studies of various proteins, including T4CP (Gomis-Rüth and Coll 2001, Wu *et al.* 2023), that is because lack of TMD facilitates expression, improves protein yield, and does not rely on detergents for proteins to exist in solution, which can have polar effect on protein functions.

Based on early topological studies, the TMD of TraD consists of 130 first residues (Lee *et al.* 1999), hence TraD$^{130\text{-}717}$ was used for subsequent studies. TraD$^{130\text{-}717}$ was expressed using pASK_TraD$^{130\text{-}717}$ construct in *E. coli* BL21 cells and purified by two rounds of

chromatography: affinity chromatography using the Strep Trap XT column (Cytiva) (Figure 19A) and size-exclusion chromatography (SEC) using Superose 6 Increase 10/300 GL column (Cytiva) (Figure 19B). The presence of purified TraD$^{130-717}$ was controlled with SDS-PAGE after consecutive rounds of purification, and the size of the protein corresponding to a molecular weight marker represents the calculated theoretical molecular weight of TraD$^{130-717}$ (approximately 68,5 kDa) (Figure 19).

The elution volume of TraD$^{130-717}$ during SEC (~15 ml) corresponds to a population of particles smaller than a hexamer (which would have a calculated theoretical molecular weight of 410,5 kDa) based on the manufacturers' handbook (Figure 19B). To observe the morphology of TraD$^{130-717}$, the fraction of the highest concentration was used for NS-EM imaging, revealing particles that do not resemble ring-like shapes, which would be expected of a hexameric TraD$^{130-717}$ (Figure 20). These results indicate that TraD$^{130-717}$ remains monomeric in the solution similarly to DotL, the T4CP of *L. pneumophila* Dot/Icm-T4SS.

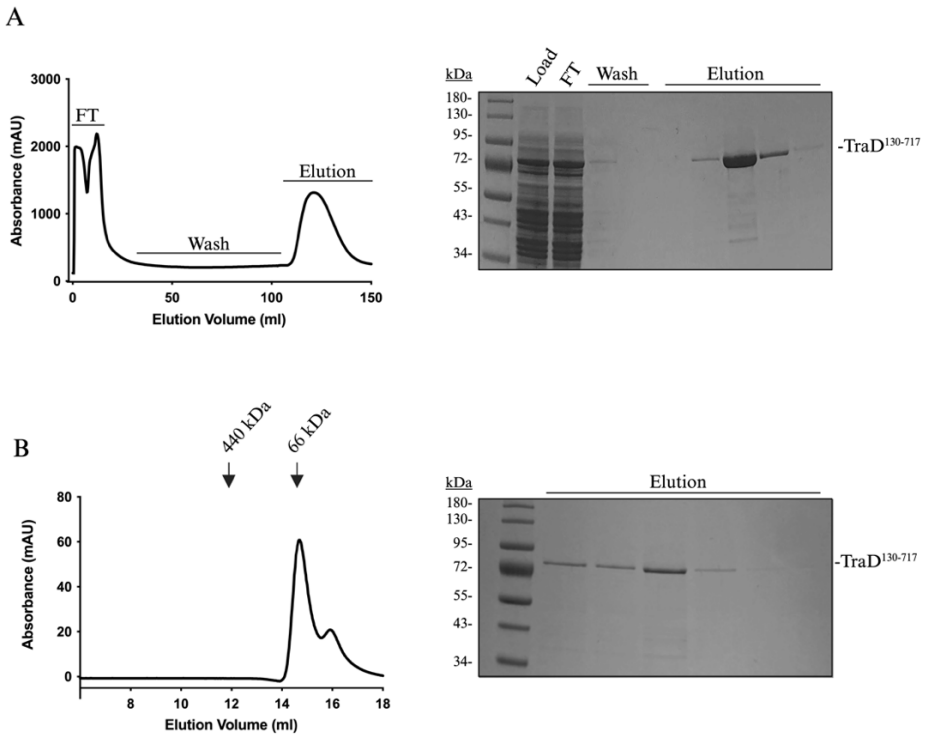


Figure 19. Purification of TraD$^{130-717}$

A| A representative chromatogram of affinity chromatography of TraD$^{130-717}$ and the corresponding SDS-PAGE gel, showing the purified TraD$^{130-717}$. The molecular weight of TraD$^{130-717}$ as shown on the SDS-PAGE gel corresponds to the calculated theoretical molecular weight of TraD$^{130-717}$ (approximately 68,5 kDa). **B|** Representative chromatogram of SEC using the Superose 6 Increase 10/300 GL column (Cytiva) and the corresponding SDS-PAGE gel of the chromatography. The elution volume of TraD$^{130-717}$ (15 ml) corresponds to a population of proteins much smaller than hexamers (calculated theoretical molecular weight 410,5 kDa).

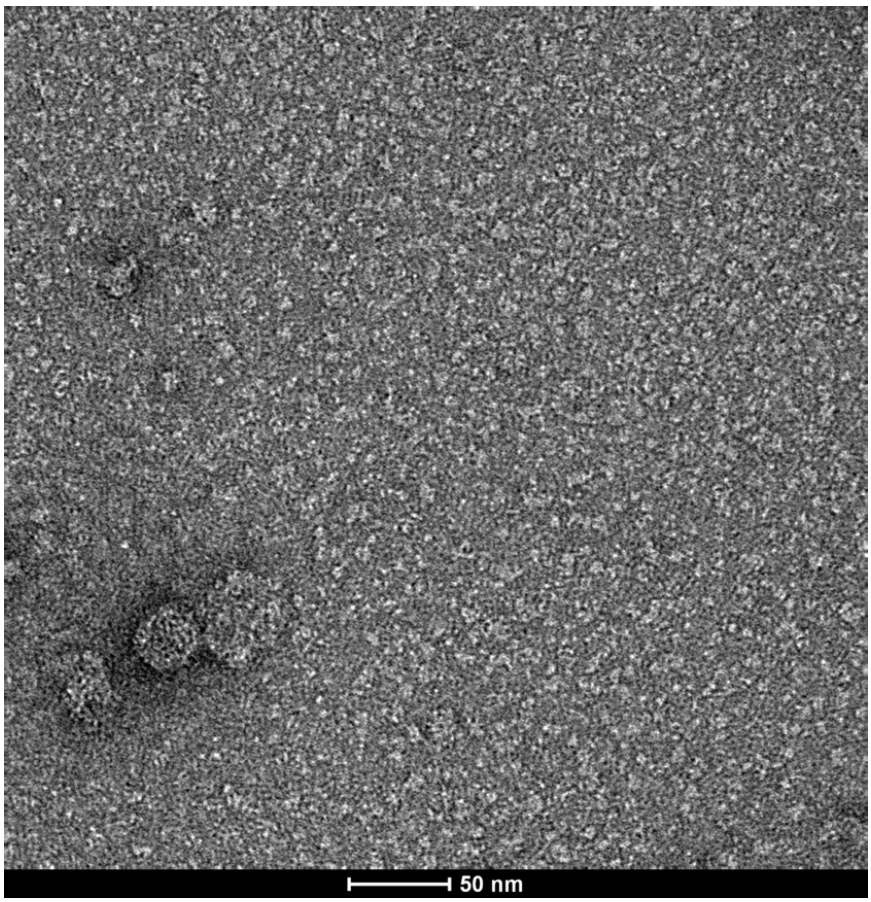


Figure 20. NS-EM micrograph of TraD$^{130-717}$

Following purification by SEC, NS-EM imaging was conducted. The visualisation did not reveal any regular ring-like shapes, indicating no hexameric TraD.

2.3.2 TraD exhibits ATPase activity

No prior research documented ATPase activity in TraD. Therefore, I investigated this using a Malachite green assay (New England Biolabs), which is a colourimetric assay that measures the amount of released phosphate as a product of ATP hydrolysis reaction. Here, it revealed that TraD$^{130-717}$ exhibits proficient ATPase activity. To define the optimal conditions for this

enzymatic function, I conducted a series of experiments focusing on buffer composition (Figure 21A). These experiments revealed the highest activity of TraD$^{130\text{-}717}$ in buffers containing sodium-based salts, neutral pH, and critically requires bivalent cations like Mn^{2+} or Mg^{2+}.

The obtained optimal buffer conditions were subsequently used for the downstream structural and biochemical investigation of TraD$^{130\text{-}717}$.

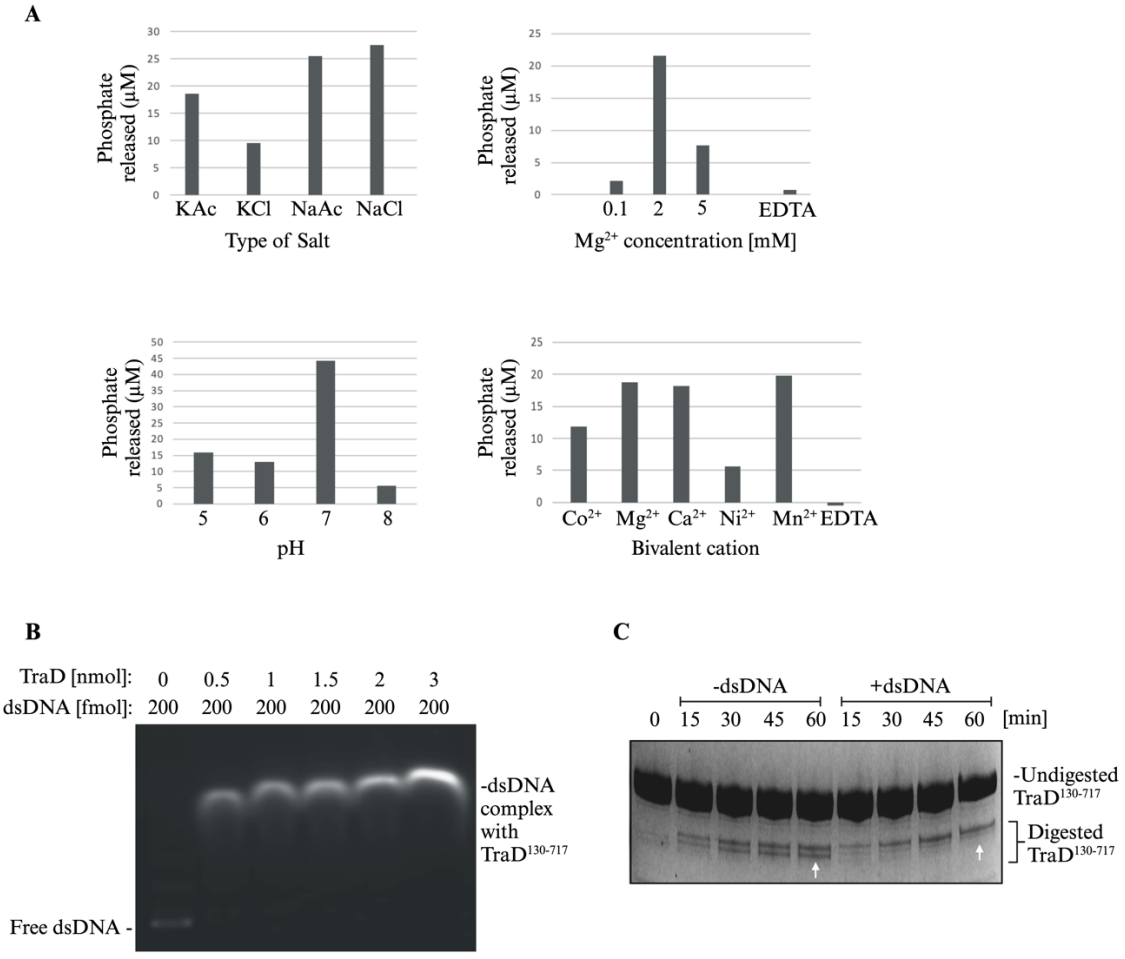


Figure 21. Biochemical analysis of TraD$^{130\text{-}717}$

A| ATPase assay of TraD$^{130\text{-}717}$ exemplifying the optimal buffer conditions for its ATPase activity. **B|** Electromobility shift assay (EMSA) showing interaction of TraD$^{130\text{-}717}$ with non-specific dsDNA. Upon addition of TraD$^{130\text{-}717}$, dsDNA changes its migration pattern, signifying interaction. **C|** Trypsin digestion of TraD$^{130\text{-}717}$. Upon addition of dsDNA, differences in digestion pattern of TraD$^{130\text{-}717}$ can be seen, which signify conformational changes of protein upon DNA binding, protecting the binding site from the protease (white arrows).

2.3.3 TraD interacts with DNA non-specifically

To assess the TraD interaction with DNA, I employed an electrophoretic mobility shift assay (EMSA) (Figure 21B). The purified $TraD^{130-717}$ at various concentrations (0.5-3 nmol) was incubated with a fixed amount (200 fmol) of dsDNA from the empty pASK plasmid (3.2 kbp) (Figure 21B). Following incubation, the samples were subjected to agarose gel electrophoresis. The results revealed a characteristic super-shift, indicating a change in the DNA's mobility due to protein binding (Figure 21B). This confirms the TraD ability to bind DNA non-specifically.

To investigate potential DNA-induced conformational changes in $TraD^{130-717}$, I performed a trypsin digestion assay (Figure 21C). This method relies on the hypothesis that DNA binding alters the protein conformation, rendering specific regions inaccessible to trypsin cleavage (Ilangovan *et al.* 2017). Therefore, DNA binding can offer a protective effect against proteolysis. This technique has been successfully used to study conformational changes in the F-plasmid relaxase TraI (Ilangovan *et al.* 2017). Here, $TraD^{130-717}$ was subjected to mild trypsin digestion in the presence or absence of non-specific 3.2 kbp dsDNA (the same as used during EMSA). Samples were collected at different time points (15-60 minutes) after trypsin addition, and digestion was stopped by boiling. SDS-PAGE analysis revealed an additional band only in the samples digested without DNA (Figure 21C). This suggests that DNA binding induces conformational changes of TraD, making certain sites inaccessible for trypsin.

2.3.4 TraD interacts with TraM *in vitro*

Since the function of TraD relies on specific recognition of the nucleoprotein complex (DNA-relaxosome), I investigated its interaction with TraM (a relaxosome subunit), using a double pull-down assay. The principle of this assay can be visualised in the schematic in Figure 22 and relies on the presence of two different affinity tags on the potentially interacting proteins and monitoring whether both proteins are captured after two rounds of affinity purifications. If no interaction occurs, both proteins would be lost through the lack of specificity to the affinity

column specific to the other potential interaction partner. However, if the proteins interact, the complex will be preserved throughout all purification steps.

To assess TraD-TraM interactions I combined equal volumes of a bacterial lysate containing strep-tagged TraD$^{130-717}$ and a lysate containing his-tagged TraM. This mixture was loaded onto the Strep Trap XT column (Cytiva), washed extensively, and eluted with desthiobiotin. At this stage, only proteins with a Strep-tag II or those interacting with them would elute. The eluate was then applied onto a His Trap HP column, washed with a low imidazole concentration buffer (10 mM) to remove non-specifically bound proteins, and eluted with high imidazole buffer (500 mM). Both TraD$^{130-717}$ and TraM were retained throughout the double pull-down, indicating a strong interaction (Figure 22).

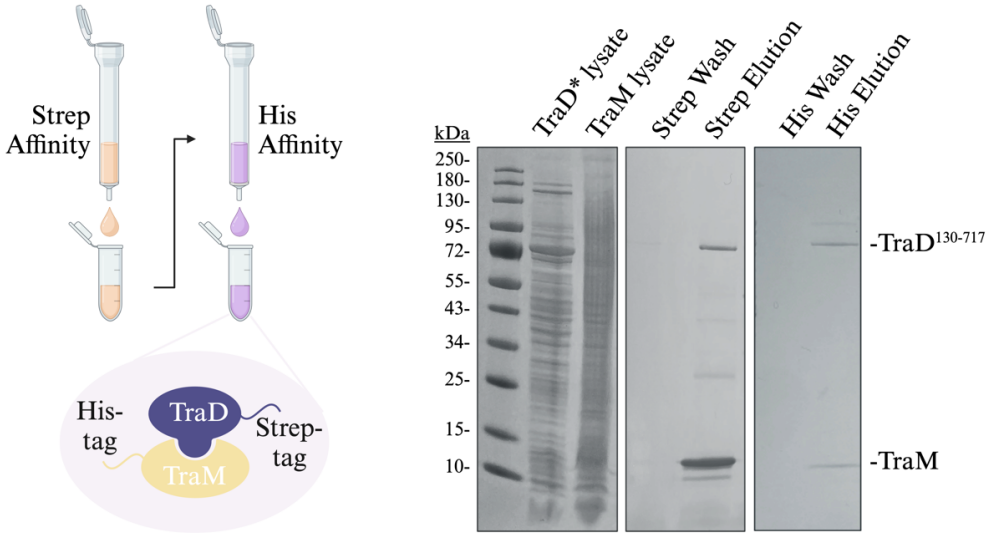


Figure 22. TraD$^{130-717}$-TraM double pull-down assay

Left: a schematic graph displaying the principle of double-pulldown. Right: the results of the double-pulldown between TraD$^{130-717}$ (theoretical molecular weight 68,5 kDa) and TraM (theoretical molecular weight 14,5 kDa), showing that interaction between these proteins was maintained throughout purification and rounds of extensive washing.

As a control, non-specific binding of TraM to the Strep Trap XT column, and TraD$^{130-717}$ to the His Trap HP column was tested and proved to be negative.

2.3.5 Structure of TraD

The development of AlphaFold-2 has revolutionised protein structure prediction, particularly for highly conserved domains and proteins within large families, such as AAA+ ATPases or T4CP, which share architectural features and tend to adopt more uniform conformations, making them easier to predict. AlphaFold-2 was recently experimentally tested for accuracy in predicting T4CP structures in the study that documented the structure of Cagβ (T4CP of Cag-T4SS) (Wu *et al.* 2023). The authors of this study reported that their AlphaFold-2 prediction of Cagβ achieved a root mean square deviation (RMSD) of 1.2 Å compared to the experimentally solved structure. The RMSD value indicates how far apart atoms of two different structures are on average, with a smaller number meaning a closer match in 3D shape. Hence, the obtained value indicates an almost identical conformation between the experimentally derived structure and the predicted model of Cagβ (Wu *et al.* 2023). This observation reinforces the high confidence in AlphaFold-2 predictions for T4CPs.

Here, I used AlphaFold-2 to generate a highly confident model of TraD atomic structure, specifically focusing on the functional, conserved region. The model achieved an impressive confidence score exceeding 90%, with low scores only in the unstructured C-terminal tail of TraD (Supplementary Figure 1). AlphaFold-2 has the ability to incorporate a specific oligomeric state in its prediction, which allowed me to impose hexameric symmetry onto the predicted model of TraD, facilitating a direct comparison with the experimentally determined structures of the hexameric TrwB and Cagβ (T4CP of R388-T4SS and the Cag-T4SS of *H. pylori,* respectively).

The obtained prediction revealed that the full-length TraD consists of four separate functional domains, the transmembrane domain (TMD), the nucleotide-binding domain (NBD), the all-alpha domain (AAD) and the C-terminal domain (CTD) (Figure 23A). This domain

organisation resembles other T4CP, such as TrwB and Cagβ, the T4CP of R388-T4SS and Cag-T4SS of *H. pylori*, respectively (Figure 12E).

Structurally, a hexameric TraD resembles a ring with 110 Å in diameter and 80 Å in height of the globular domain (Figure 23B). A single subunit of TraD assumes the shape of a letter "b", with the N-terminal TMD of TraD consisting of three α-helices, separated by a periplasmic β-hairpin (Figure 23C). Such organisation of TMD is essential for anchoring in the inner membrane, as the hydrophobic character of α-helices will direct the protein towards the lipids within bacterial membranes. The NBD is involved in ATP binding and hydrolysis and assumes an α/β topology made up of nine β-sheets flanked by α-helices on each side (Figure 23C). Immediately underneath the NBD, a bundle of α-helices forms the AAD, a domain known among T4CP to be responsible for the interaction of T4CP with substrates (Llosa and Alkorta 2017). The AAD of TraD is made up of eight α-helices (Figure 23C), one more than observed in TrwB, potentially conveying specificity to the recognition of the F-encoded relaxosome (Gomis-Rüth and Coll 2001). The additional α-helix of AAD is provided by the last domain of TraD, the CTD, which forms an unstructured, flexible tail, flanked by two α-helices, one of which (N-terminal) is incorporated into AAD, and the other (C-terminal) plays a role in recognition of TraM and interaction with the relaxosome (Figure 23C) (Lu *et al.* 2008a).

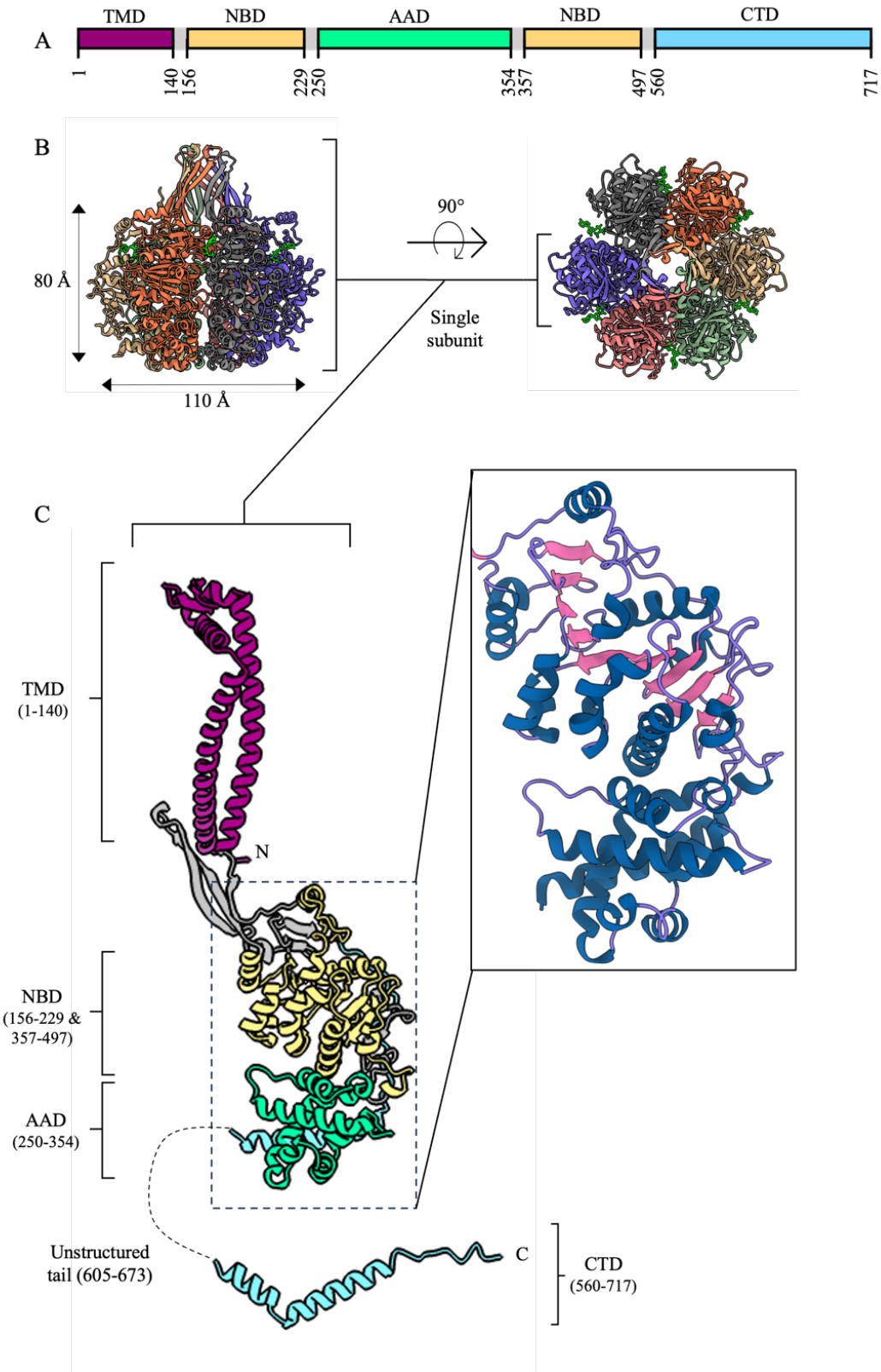


Figure 23. AlphaFold-2 prediction model of TraD structure

A| Domain organisation of TraD, reveals four distinctive domains – a transmembrane domain (TMD); a nucleotide-binding domain (NBD) which is interrupted by all-alpha domain (AAD) and a

pronounced C-terminal domain (CTD). **B|** The hexameric AlphaFold-2 prediction model of TraD reveals a ring-like shape particle reminiscent of TrwB. **C|** A single subunit of full-length TraD with clear segmented topology into functional domains. Blowout of the soluble domain shows organisation of α-helices (deep blue) and β-sheets (pink).

2.3.6 The lumen of TraD is fit for interaction with DNA

It is hypothesised that as conjugative T4CP oligomerise, their lumens act as channels for translocating ssDNA substrates, energised by ATP hydrolysis. Examining a cross-section through two *vis-à-vis* subunits reveals that the lumen of TraD is lined with predominantly positively charged residues (Figure 24A). These include lysines 273, 335, 467, and 488, alongside a negatively charged glutamate 466 (Figure 24A&B). This positively charged lumen, rich in lysines, resembles that of FtsK, a hexameric ATPase involved in DNA translocation during cell division (Jean *et al.* 2020). The arginine and lysine residues that line the lumen of FtsK directly interact with and accelerate DNA movement. Moreover, this characteristic is also present in TrwB (T4CP of the R388-T4SS), which contains three specific lysine residues lining its lumen that are important for the transport of DNA (Larrea *et al.* 2017). This similarity suggests that the lysine-rich lumen of TraD might play a role in DNA translocation, consistent with the previous positive results for $TraD^{130\text{-}717}$-dsDNA binding (Figure 21B&C).

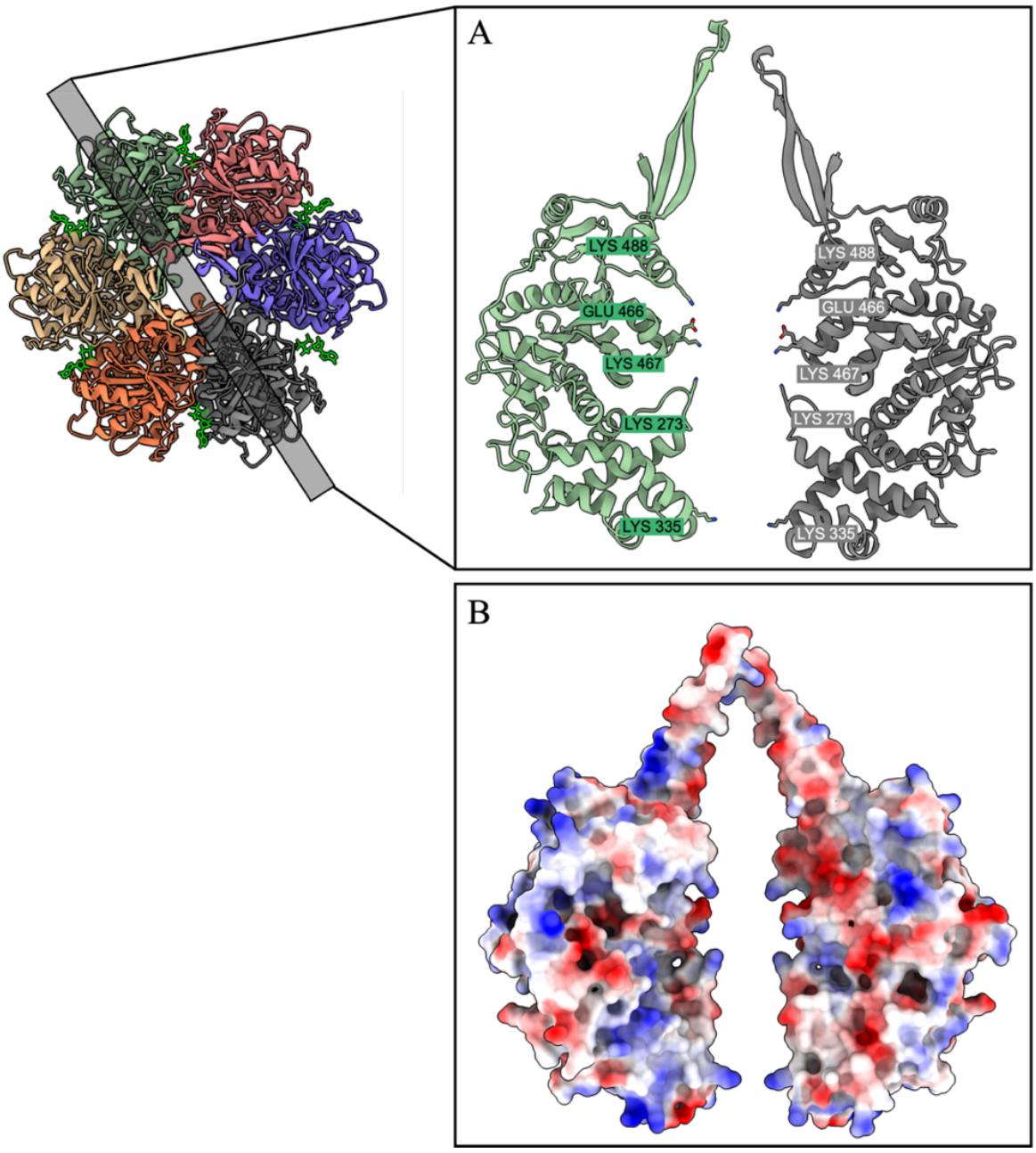


Figure 24. Positive charge of TraD lumen suggests involvement of lysine residues in DNA transfer

A| A slice through two *vis-à-vis* subunits of TraD hexamer reveals five main residues that protrude to its lumen, that is positively charged lysines 273, 335,467 and 488, as well as positively charged glutamate 466. **B|** Electrostatic potential of TraD and its lumen reveals a positively charged lumen, possibly functioning in attracting negatively charged DNA. The colour coding symbolised, a local negative, neutral, or positive charge represented by colours red, white, and blue, respectively.

2.3.7 Interaction interface between adjacent TraD subunits

The oligomeric state of T4CP proves to be dynamic – a feature characteristic to AAA+ ATPases (Erzberger and Berger 2006). Microscopic analyses of TrwB and Cagβ (T4CP of R388-T4SS and Cag-T4SS of *H. pylori*) suggest that T4CP associate with their cognate T4SS as oligomers – dimers or hexamer of dimers, respectively (Figure 12C). This suggests that the ability to form higher-order oligomers is crucial to their function, however, the details about TraD oligomerisation remain unclear. The generated structural model of TraD allowed to identify the precise residues that form the interface between neighbouring TraD subunits (Figure 25). The detailed interaction network is broad and is provided in Supplementary Table 1.

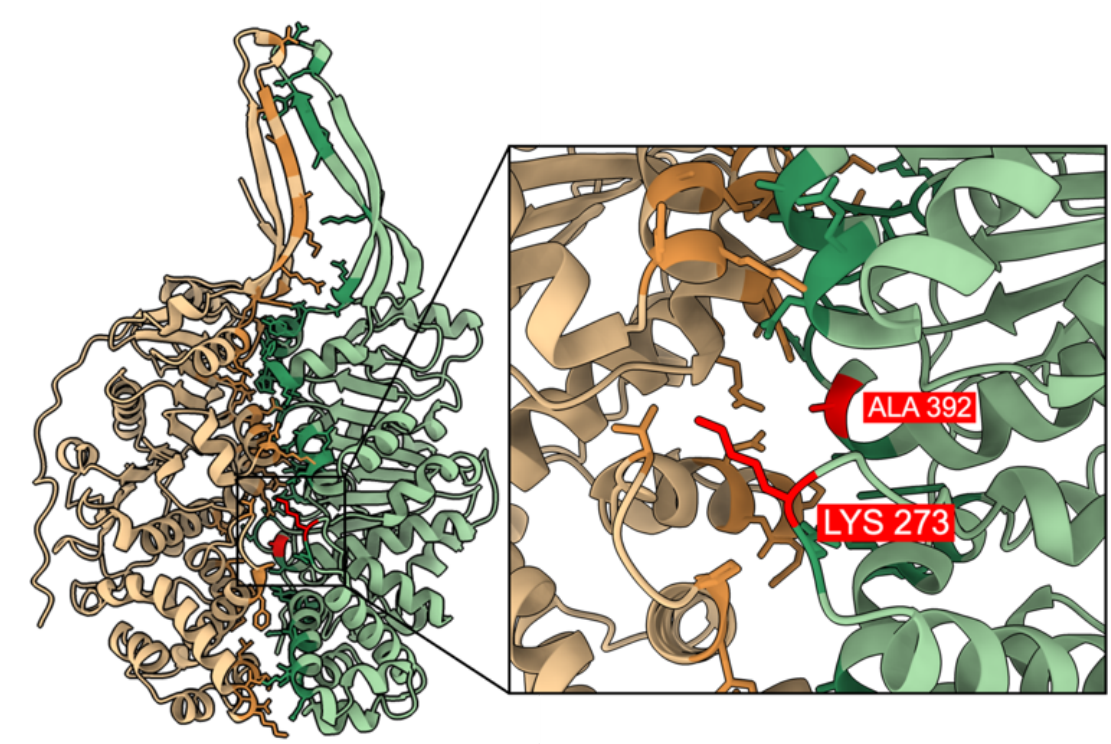


Figure 25. The interaction interface between TraD subunits

Two TraD subunits (tan and light green) oligomerise through an extensive network of residues (coloured in orange and dark green, respectively) which are place across the entire length of the subunits, encompassing all domains. The precise list of interacting residues is provided as Supplementary Table 1. Residues Lys273 and Ala392 (red) correspond to residues that were experimentally confirmed to be involved in *in vivo* oligomerisation of TraD (Haft et al. 2007).

Two key residues were mapped, that is lysine 273 and alanine 392. In 2007, a mutational analysis of these residues showed an abolishment of TraD oligomer formation evidenced by i*n vivo* crosslinking studies (Haft *et al.* 2007). Here, the predicted structural model of TraD positions both residues at the interface between subunits, hence mutations in those residues are likely to abolish interactions between them (Figure 25). Therefore, the model presented here explains the observed abrogation of homodimer formation (Haft *et al.* 2007).

2.3.8 Nucleotide-binding domain of TraD

Earlier, I demonstrated the $TraD^{130-717}$ ability to hydrolyse ATP (Figure 21A). To pinpoint the specific location of ATP hydrolysis within TraD structure, I aligned the predicted TraD model with the X-ray crystal structures of TrwB (T4CP of R388-T4SS, PDB ID: 1GL7) and Cagβ (T4CP of Cag-T4SS, PDB ID: 8DOL).

The alignment revealed a near match between TraD and TrwB, with RMSD of 3 Å (Figure 26A). This is unsurprising since TrwB, like TraD, functions as a T4CP in a conjugative T4SS, thus both proteins share DNA as their primary substrate. In contrast, TraD displayed a lower degree of structural similarity with Cagβ, with an RMSD of 27.8 Å (Figure 26B). This might be because Cagβ functions in an effector translocating Cag-T4SS, and thus its primary substrate is the oncoprotein CagA, not DNA. The high degree of similarity between the predicted TraD model and the TrwB crystal structure becomes particularly evident when overlaying the conserved NBD of both proteins (Figure 26C).

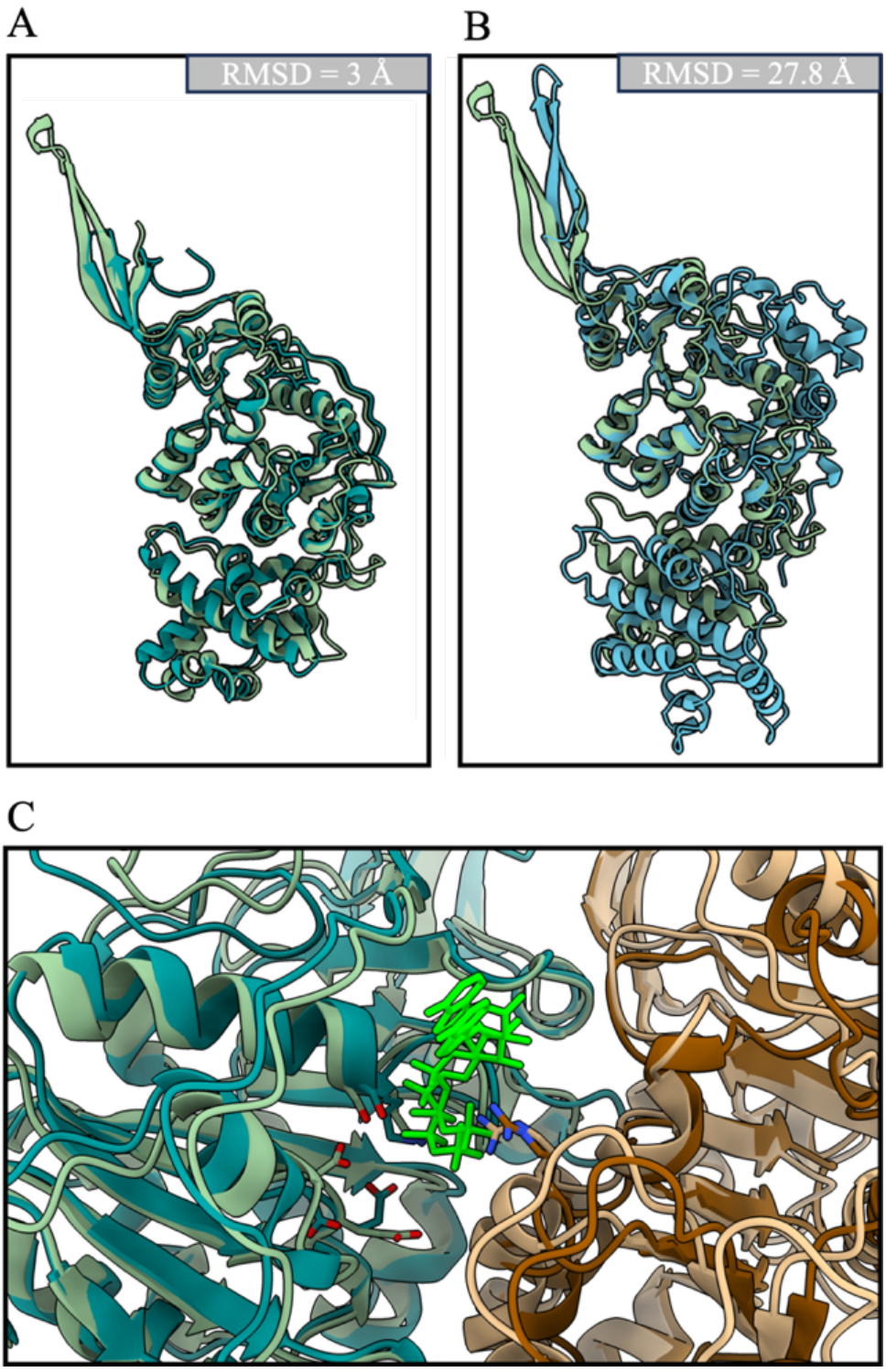


Figure 26. Comparison of the model of TraD with TrwB and Cagβ

A| Overlay between TraD (light green) and TrwB (dark green, PDB ID: 1GL7) with calculated RMSD of 3 Å overall. **B|** Overlay between TraD (light green) and Cagβ (light blue, PDB ID: 8DOL) with calculated RMSD of 27.8 Å overall. **C|** Overlay of the nucleotide-binding sites between TraD (light green) and TrwB (dark green) showing a nearly perfect overlay between the secondary structures as well as the location of residues that interact with an ATP-analogue nucleotide ADPNP (lime).

The specific amino acid sequences forming the motifs within NBD that are directly responsible for ATP hydrolysis are well-known and designated as Walker A and Walker B motifs, corresponding to sequences: [G-xxxx-GK-T/S] (x-any residue) and [hh-D-hh] (h-hydrophobic residue), respectively. The search for these conserved motifs in TraD revealed that the Walker A motif consists of the sequence [192-GTVGAGKS-199], whereas the two nearby residues on the same TraD subunit, aspartic acid 245 and glutamic acid 246, form the Walker B motif (Figure 27A).

Interestingly, the Walker A and B motifs in both TraD and TrwB are composed of the same residues, and when mapped onto the structures and aligned, they show a near-identical arrangement. Specifically, the TraD glycine 197, lysine 198, serine 199, aspartic acid 425, and glutamic acid 426 occupy positions similar to glycine 135, lysine 136, serine 137, aspartic acid 356, and glutamic acid 357 in TrwB, respectively (Figure 27B).

Aside from Walker A and B motifs, AAA+ ATPases have an additional characteristic motif essential for ATP hydrolysis, which is the presence of an arginine finger that stabilises the negative charge of the γ-phosphate group, acting as a catalyst of the reaction (Miller and Enemark 2016). The arginine residue is always provided *in trans* by the neighbouring subunit, not by the same subunit that contains Walker A and B motifs reaction (Miller and Enemark 2016). Consequently, ATP hydrolysis occurs at the interface between two subunits (Gomis-Rüth *et al.* 2001, Ogura *et al.* 2004, Wu *et al.* 2023). In the model, arginine 444 of the adjacent TraD subunit is mapped in close proximity to Walker A and B motifs, suggesting the role as the essential arginine finger (Figure 27B).

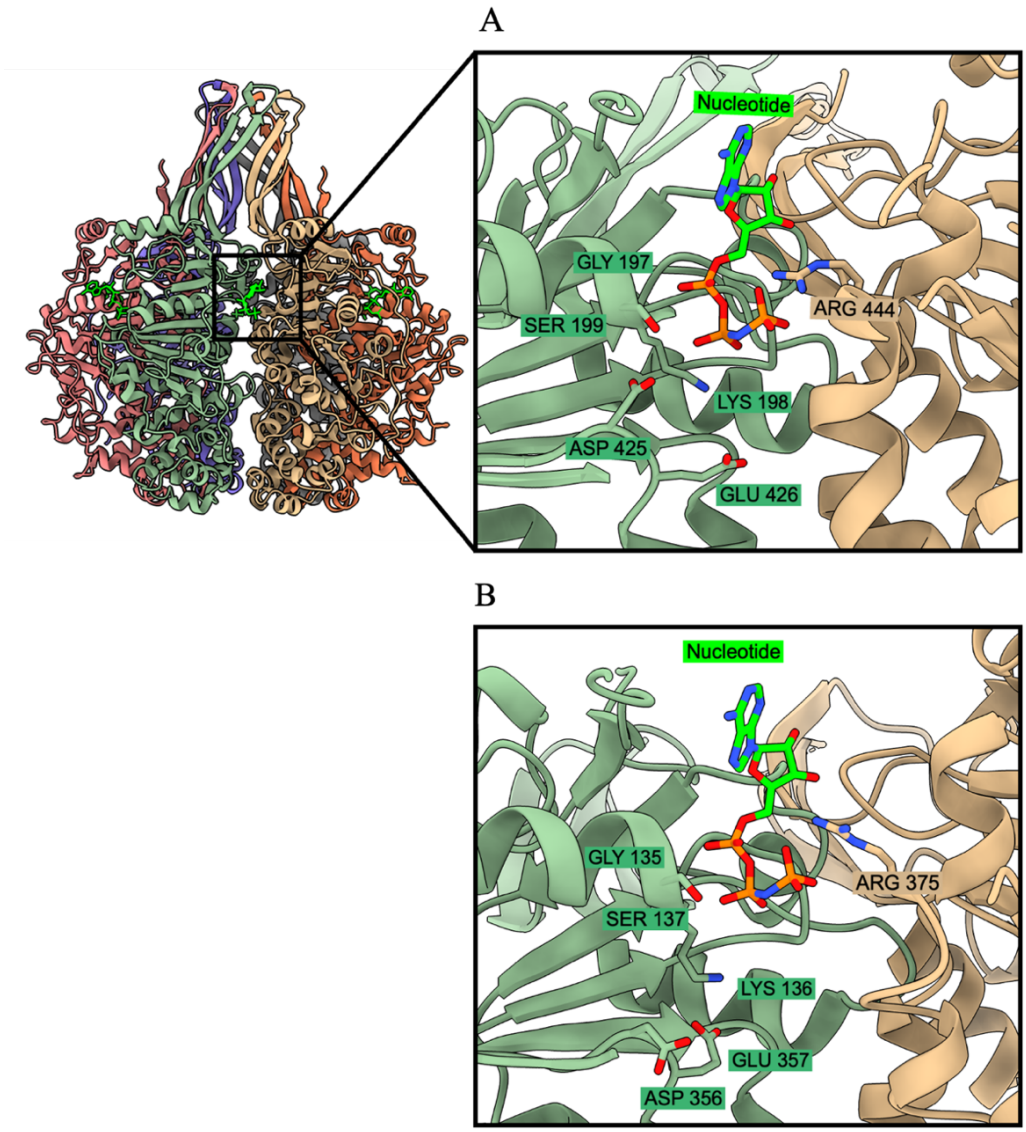


Figure 27. The nucleotide-binding site of TraD

ATPase activity is conveyed by two characteristic motifs – Walker A motif known as P-loop and Walker B motifs. **A|** Representation of the NBD of TraD. Walker A motif in TraD is comprised of Gly197, Lys198 and Ser199, whereas the Walker B motif of Asp425 and Glu426, all found on one subunit of TraD (green). A distinctive feature of AAA+ ATPases is an arginine finger provided *in trans* by an adjacent subunit, in this case arg444 provided by the neighbouring subunit of TraD (tan). ATP analogue, ADPNP is also fitted based on homology to TrwB (lime). **B|** Representation of the NBD in the structure of TrwB (PDB ID: 1GL7), showing the corresponding Walker A and B motifs, made up by: Gly135, Lys136, Ser137 and Asp356, Glu357, respectively, on the same subunit of TrwB (green). The arginine finger Arg375 is provided by the adjacent subunit (tan). ATP analogue, ADPNP is shown in lime.

Altogether, by mapping all three motifs – that is Walker A and B and the arginine finger, I could precisely fit an ATP-analogue nucleotide, ADPNP, into the predicted model, signifying the exact binding site of ATP. This docking was based on the orientation of the same nucleotide within the TrwB crystal structure with ADPNP (PDB ID: 1GL7).

2.3.9 ATP induces oligomerisation of TraD

It is established that T4CP can oligomerise dynamically and transiently oligomerise, forming multisubunit complexes for optimal functionality. Such oligomerisation is often induced by functional regulation *in vivo*, such as the binding of substrates or specific activators (Erzberger and Berger 2006). Here, I explored the factors that may influence the assembly of TraD *in vitro*. To observe the formation of higher oligomeric states of TraD, I used chemical crosslinking with glutaraldehyde (GA) – a chemical compound that introduces covalent bonds between -NH_2 moieties of polypeptide chains, following the reaction presented in Figure 28. Once crosslinked, the covalently bound proteins become resistant to denaturing conditions, such as those present in SDS-PAGE. Hence, such crosslinked interacting proteins will run at the higher molecular weight that corresponds to the sum of their masses, forming just one distinctive band on SDS-PAGE. I applied this principle to develop a quick method for understanding the oligomeric state of TraD upon addition of different compounds.

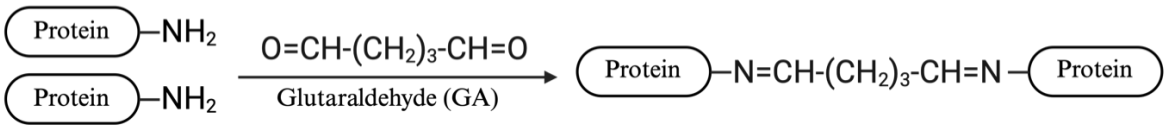


Figure 28. Glutaraldehyde crosslinking reaction between two proteins

Glutaraldehyde introduces covalent bonds between -NH2 moieties of polypeptide chains.

Purified $TraD^{130-717}$ was incubated with various cofactors (ATP, dsDNA, ssDNA), substrates (ATP and non-hydrolysable ATP analogue, ATP-PNP) or interaction partner (TraM) and subjected to chemical cross-linking with a varying concentration of GA (0.001%-0.1%). After

incubation, the resulting crosslinked samples were subjected to SDS-PAGE to assess the formation of higher order oligomers by TraD$^{130-717}$.

The results of the crosslinking assay revealed that neither dsDNA nor ssDNA significantly influenced TraD$^{130-717}$ oligomerisation compared to the freely crosslinked TraD control (Figure 29). Similarly, the presence of the known interaction partner, TraM, did not induce a change in the oligomeric state.

However, a distinct pattern emerged upon supplementation with ATP. The addition of both low (0.001%) and high (0.1%) concentrations of GA resulted in the formation of distinct bands corresponding to dimeric and hexameric TraD$^{130-717}$, respectively (Figure 29). Interestingly, the non-hydrolysable ATP analogue - ATP-PNP - did not induce a similar effect. These observations collectively suggest that TraD oligomerisation strictly depends on ATP hydrolysis and the sole binding of a nucleotide is insufficient for oligomerisation, as evidenced by the lack of response with ATP-PNP.

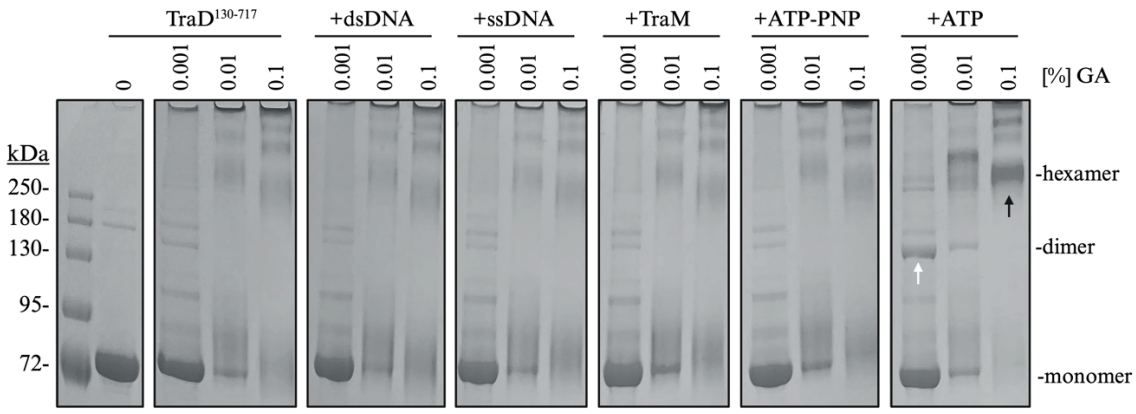


Figure 29. Glutaraldehyde crosslinking assay of TraD

Crosslinking with an increasing concentration of GA (from 0.001% to 0.1%) was used to understand whether addition of an excess of TraD substrates or interaction partners (namely, dsDNA, ssDNA, TraM, non-hydrolysable ATP analogue ATP-PNP, ATP) in 1:2 molar ratio (TraD monomer: substrate). TraD formed clear bands corresponding to dimers (white arrow) at low GA concentration (0.001%) and hexamers (black arrow) at higher GA concentration (0.1%).

Furthermore, the formation of dimers at lower GA concentrations suggests a potential two-step oligomerisation mechanism. TraD may initially form dimers, followed by the subsequent assembly of these dimers into trimers of dimers, ultimately resulting in a hexameric complex. This proposed mechanism aligns well with the structural predictions of TraD, which localise the nucleotide-binding site at the interface between subunits, as presented earlier (Figure 27).

Building upon the observation that ATP stimulates TraD oligomerisation, large-scale crosslinking was performed using purified $TraD^{130\text{-}717}$ in the presence of excess ATP (10 mM) and 0.1% GA. The formation of hexamers was subsequently confirmed by size-exclusion chromatography (SEC) using a Superose 6 Increase 10/300 GL column (Cytiva). As depicted in Figure 30, a distinct shift in elution volume was observed, with the peak corresponding to uncrosslinked monomeric $TraD^{130\text{-}717}$ (calculated theoretical molecular weight: 68.5 kDa) migrating towards a higher molecular weight fraction when crosslinked in the presence of ATP (10 mM), corresponding to the theoretical molecular weight of hexameric $TraD^{130\text{-}717}$ (410.5 kDa) according to the manufacturer's handbook.

SDS-PAGE analysis confirmed the presence of crosslinked hexameric $TraD^{130\text{-}717}$ in Peak 2 (Figure 30), together with some crosslinking-induced aggregation in Peak 1, which was anticipated as the crosslinking reaction is not specific. The size of the crosslinked hexameric TraD surpasses the capabilities of standard molecular ladders, but clearly exceeds 250 kDa.

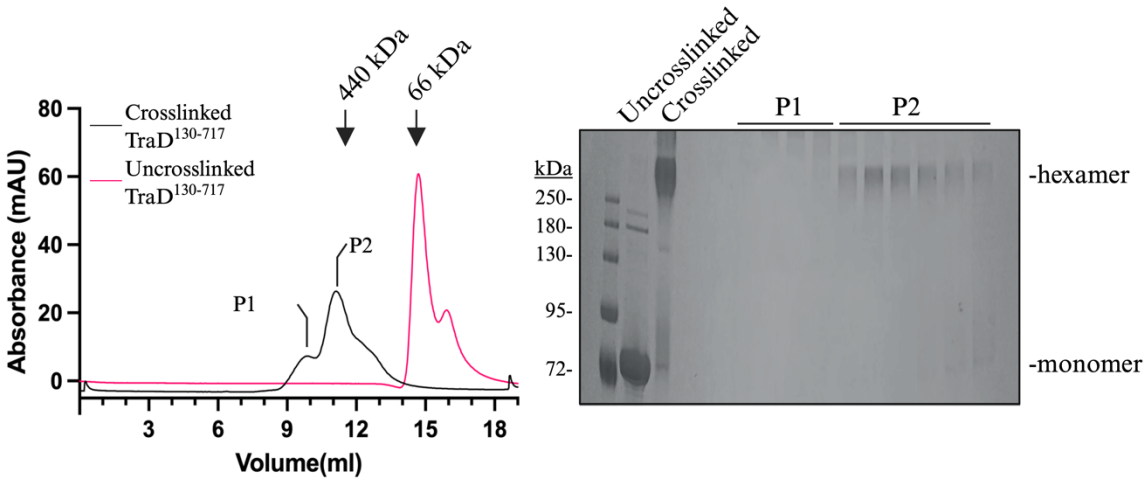


Figure 30. Large-scale crosslinking of TraD$^{130-717}$

Pure fraction of TraD$^{130-717}$ was crosslinked with 0.1% GA in presence of excess of ATP (10 mM). Then, such prepared fraction was applied on size-exclusion chromatography Superose 6 Increase 10/300 GL column (Cytiva) to assess the size of the resulting complex. The crosslinked fraction showed a visible shift to the higher molecular weight that roughly corresponds to that of hexameric TraD$^{130-717}$ (left). The corresponding SDS-PAGE also confirms presence of TraD$^{130-717}$ at the higher molecular weight, corresponding to hexamers (right). P1, P2 -Peak 1 and Peak 2.

2.3.10 Cryo-EM of TraD$^{130-717}$ reveals dimer and trimer formation

To further investigate the oligomeric state of TraD, the crosslinked fraction (Peak 2) obtained from size-exclusion chromatography was subjected to NS-EM analysis. As depicted in Figure 31A, NS-EM revealed the presence of regular ring-like structures, potentially corresponding to hexameric TraD, in agreement with previous structural predictions.

The same sample was subsequently employed for cryo-EM analysis. The crosslinked TraD fraction was applied to copper Quantifoil 1.2/1.3 grids, followed by optimisation of vitrification conditions. The grids were then visualisation using an in-house Glacios XFEG microscope (ThermoFisher) equipped with a Falcon 4 direct electron detector. After overnight data collection, the curation of exposures yielded a final dataset comprising 581 high-quality, motion-corrected micrographs. Analysis was performed using CryoSPARC (Punjani *et al.* 2017).

The dataset exhibited significant heterogeneity, likely a consequence of the non-specific nature of GA crosslinking. However, 2D classification revealed that the most prominent particle population consisted of dimers (Figure 31B), with 3,089 particles identified. Additionally, 950 trimeric particles were detected (Figure 31C). These particles were utilised for independent 3D reconstructions without imposing symmetry, resulting in their respective 3D volume maps.

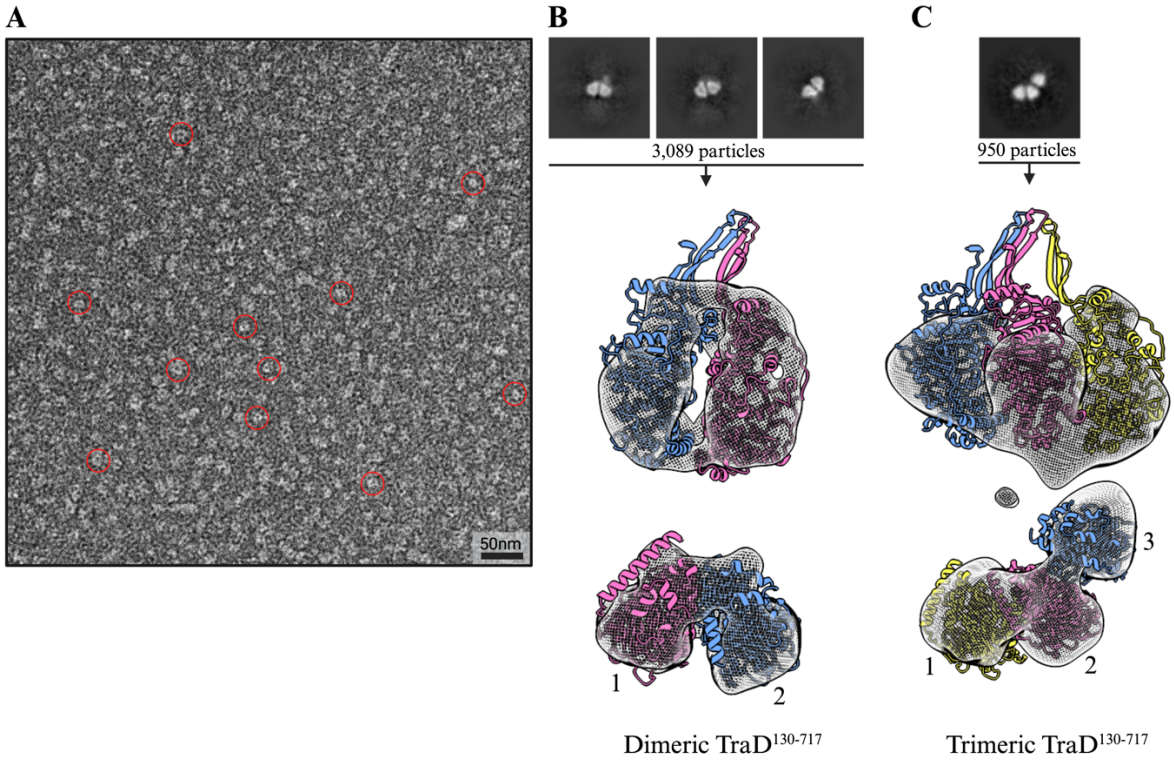


Figure 31. Cryo-EM of crosslinked TraD$^{130\text{-}717}$

A| Representative NS-EM micrograph of crosslinked TraD$^{130\text{-}717}$ showing a representation of the crosslinked particles. Ring-like shapes are visible and are circled in red. **B|** The most prominent populations were TraD$^{130\text{-}717}$ dimers. Top: 2D-classification of particles showing particles with homodimeric symmetry. Bottom: 3D-reconstruction of the obtained 2D-classes with AlphaFold-2 model of a dimeric TraD$^{130\text{-}717}$ fitted. **C|** The second prominent population of particles was that corresponding to trimeric TraD$^{130\text{-}717}$. Top: 2D-classification of particles showing formation of a half-ring structure made up of three subunits of TraD$^{130\text{-}717}$. Bottom: the 3D-reconstruction of the obtained 2D-classes with AlphaFold-2 model of trimeric TraD$^{130\text{-}717}$.

The map presented in Figure 31B corresponds to two side-by-side TraD$^{130\text{-}717}$ subunits, indicative of homodimers. Conversely, Figure 31C displays a distinctive half-ring-shaped map,

consistent with a TraD$^{130-717}$ trimer. Notably, the predicted dimeric and trimeric TraD models exhibited good coverage when fitted into the obtained volumes (Figures 31B & 31C), further supporting the validity of the prediction model.

Despite the successful demonstration of dimer and trimer formation, the resolution of the obtained volumes was insufficient to discern high-resolution structural features. This limitation likely stems from the high level of heterogeneity within the dataset. Nonetheless, the obtained volumes showcase the inherent flexibility of TraD oligomeric state. Even crosslinking was not sufficient to enforce a single oligomeric configuration, suggesting that *in vivo* oligomerisation of T4CPs requires multiple functional cues, ATP hydrolysis being only one of them.

2.3.11 Scaffold protein is insufficient to induce hexamerisation of TraD

A compelling approach to increase the homogeneity of protein samples for structural studies involves the creation of fusions with strategically selected scaffold proteins. These scaffolds exploit their intrinsic properties to impart specific features onto the protein of interest. A prominent example of such a scaffold is the highly stable hexameric protein, Hcp, which plays a crucial role in constructing the contractile sheath of the T6SS secretion system. Notably, Hcp has been previously reported for its successful application in the structural determination of other AAA+ ATPases (Macé *et al.* 2022). Figure 32 provides a visual representation of the expected outcome of generating an Hcp fusion with TraD.

To promote homogeneity, an N-terminal fusion protein, designated Hcp-TraD$^{130-717}$, was created by linking TraD$^{130-717}$ to the highly stable hexameric protein Hcp. The purification of Hcp-TraD$^{130-717}$ followed the established protocol for TraD$^{130-717}$ alone, employing a combination of affinity chromatography with a Strep Trap XT column (Cytiva) and size-exclusion chromatography with a Superose 6 Increase 10/300 GL column (Cytiva). Notably, the chromatography buffer was supplemented with 10 mM ATP, following the positive cross-linking results presented earlier (Figure 29).

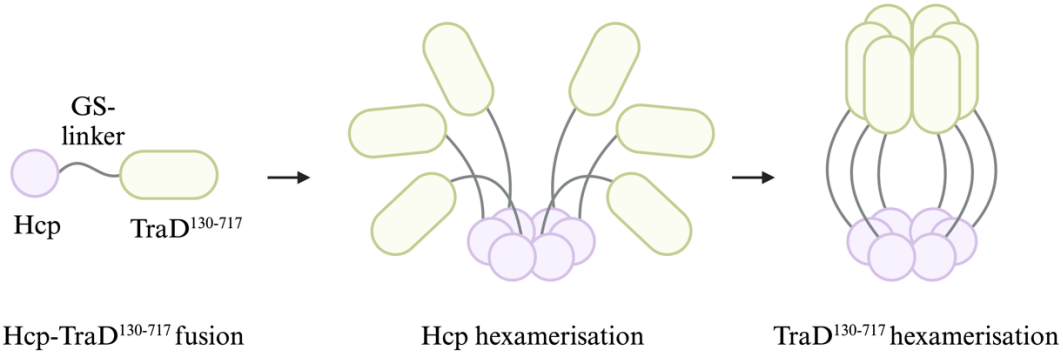


Figure 32. Principle of using Hcp as scaffold

Hcp is a highly stable hexameric protein, here a Hcp-TraD$^{130-717}$ fusion was generated as an attempt to impose a hexameric symmetry onto TraD$^{130-717}$.

The resulting chromatogram for Hcp-TraD$^{130-717}$ displayed a significant shift in elution volume compared to TraD$^{130-717}$ alone, indicating a higher molecular weight fraction (Figure 33). This shift corresponds to the theoretical molecular weight of a hexameric Hcp-TraD$^{130-717}$ complex (534 kDa) compared to the TraD$^{130-717}$ monomer (68.5 kDa). Fractions corresponding to the peak (designated P) were analysed by NS-EM, revealing the presence of hexameric rings (Figure 33).

The apparent homogeneity of the peak on NS-EM and the symmetric profile of the SEC peak (Figure 33) justified further analysis using cryo-EM. The sample was prepared for cryo-EM using vitrified grids on copper Quantifoil 1.2/1.3 grids. Following optimisation of vitrification conditions, the grids were visualisation using the in-house Glacios XFEG microscope

(ThermoFisher) equipped with a Falcon 4 direct electron detector. After overnight data collection, the final dataset was curated and motion-corrected.

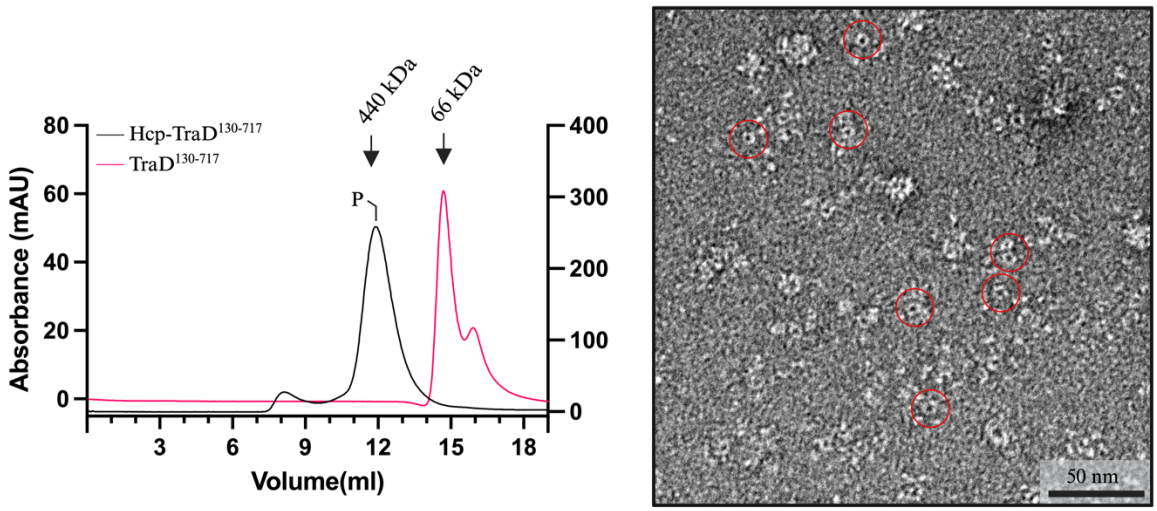


Figure 33. Purification of Hcp-TraD$^{130-717}$

Left: Comparison of size-exclusion chromatographs using column Superose 6 Increase 10/300 GL column (Cytiva) of Hcp-TraD$^{130-717}$ supplemented with 2 mM ATP (black curve) and TraD$^{130-717}$ (red curve) showing a shift in elution volume towards higher molecular weight, indicating presence of hexamers (calculated theoretical molecular weight of Hcp-TraD$^{130-717}$ hexamer 534 kDa; of TraD$^{130-717}$ 68.5 kDa) P-peak. Left Y axis corresponds to the black curve, whereas right Y axis corresponds to the red curve. Right: A representative NS-EM micrograph of fractions corresponding to peak P of Hcp-TraD$^{130-717}$. Formation of rings is visible (circled in red).

2D classification of the cryo-EM dataset revealed successful hexamerisation of Hcp, but not of TraD$^{130-717}$ itself (Figure 34). Nonetheless, the resulting 2D classes were used for 3D reconstruction with imposed C6 symmetry, generating a map of the complex. Notably, both the crystal structure of Hcp (PDB ID: 5XEU) and the generated AlphaFold-2 predicted model of TraD$^{130-717}$ fit well within the obtained density map. Data analysis was performed using CryoSPARC (Punjani *et al.* 2017).

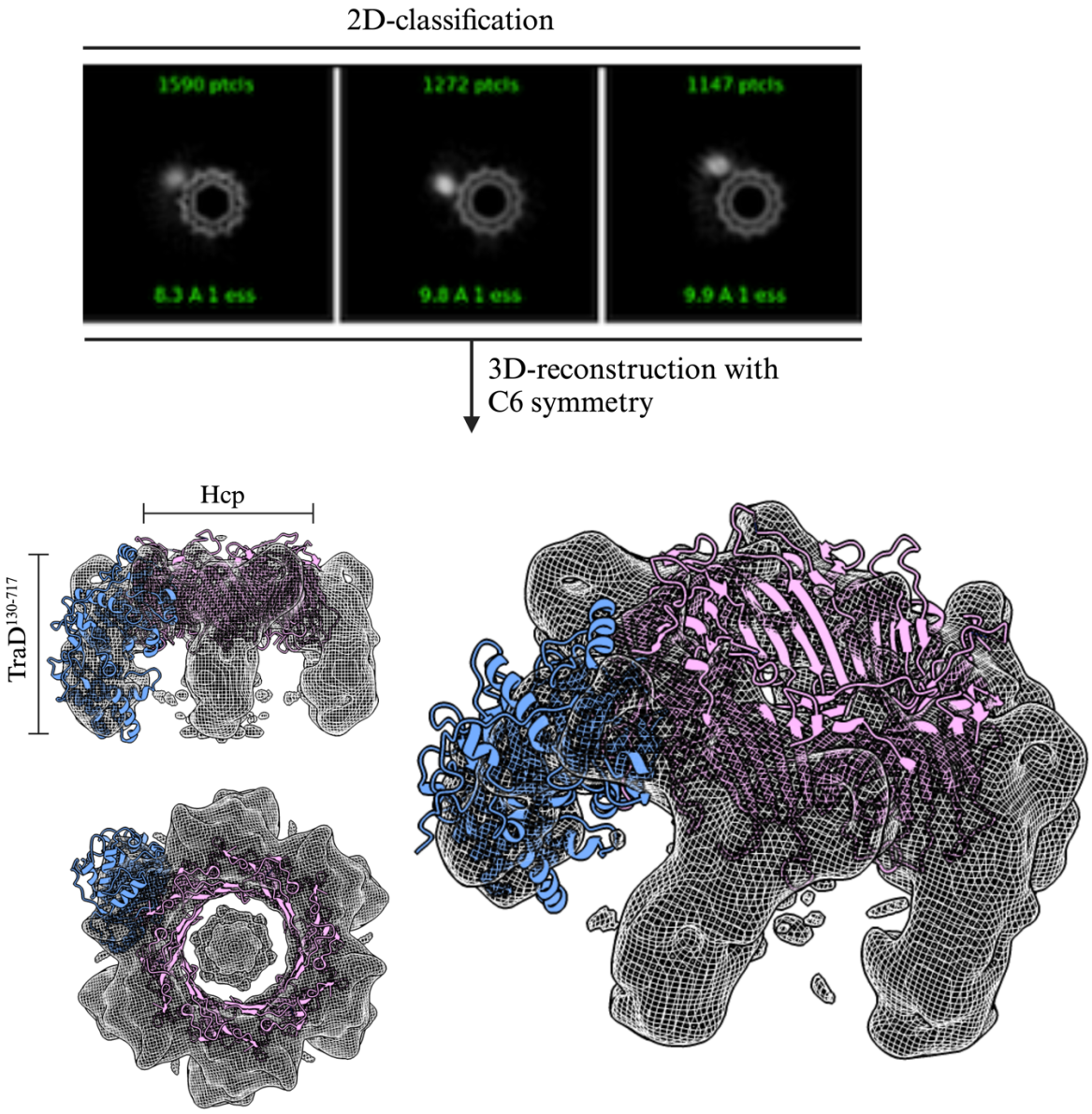


Figure 34. Cryo-EM map of Hcp-TraD$^{130\text{-}717}$

Top: following cryo-EM data collection, the most prominent 2D-classes from the dataset of Hcp-TraD$^{130\text{-}717}$ show proficient formation of the hexameric Hcp ring but no associated oligomerisation of TraD$^{130\text{-}717}$ was observed. Bottom: the 3D-map following 3D-reonstruction of the resulting 2D-classes with applied C6 symmetry. The structure of hexameric Hcp (PDB ID: 5XEU) together with AlphaFold-2 prediction of TraD$^{130\text{-}717}$ was fitted within the generated map.

2.4 Discussion

The rise of high-confidence structural predictions by AlphaFold-2 in 2021 (Jumper *et al.* 2021) has revolutionised structural biology (Baek *et al.* 2021, Thornton *et al.* 2021, Lin *et al.* 2023), thanks to the high accuracy of its models. Their utilisation marks a ground-breaking paradigm shift in structural biology and drug discovery (Subramaniam and Kleywegt 2022) with some claiming that the entire field of structural biology is now "already solved" (Ourmazd *et al.* 2022), as nearly all known proteins have their predictions (Callaway 2022). While such models must be taken with caution, and they will never entirely replace experimentally derived models (Terwilliger *et al.* 2024), they represent an invaluable tool, especially reliable for proteins with many conserved domains, such as ATPases. The prediction structures are slowly finding their place between experimental structures, with the first AlphaFold-2-driven biological stories being published as standalone studies (Goulet and Cambillau 2022). AlphaFold-2 structures have also recently entered the T4SS field with the model of the entire Dot/Icm-T4SS of *L. pneumophila*, which relies heavily on the obtained structural predictions complemented by validation of the model, rather than experimentally obtained structures (Dutka *et al.* 2023). Here, I employ a similar approach, where an AlphaFold-2 prediction model of TraD is used, validated experimentally, and then serves as the basis for structural analysis.

The T4CP of the iconic F-plasmid encoded T4SS, TraD mediates reception of the relaxosome-processed F-plasmid DNA, and then directs it towards T4SS lumen for the conjugative transfer across the cell envelopes of the donor and the recipient cell. Thus, TraD is indispensable during conjugation, coupling two important events: substrate DNA processing and transfer, marking it a very attractive drug target against bacterial conjugation and the associated spread of ARG. Nevertheless, the structural details about TraD: its oligomeric state, mechanism of oligomeric transition or functional properties, such as the ability to hydrolyse ATP or bind DNA, have been understudied.

Here, I obtain a high-confidence prediction model of TraD covering the functional region of the protein at 90% confidence score (Jumper *et al.* 2021), with low scores only in the unstructured C-terminal tail of TraD (Figure 23, Supplementary Figure 1). Subsequently, I validated the predicted model on three levels. Firstly, I identify that residues Lys273 and Ala392 form the interaction interface between TraD subunits in the prediction model (Figure 25), which is in line with their experimentally evidenced role in the oligomerisation of TraD through *in vivo* crosslinking (Haft *et al.* 2007). Secondly, the predicted model of TraD shows an almost ideal match of the ATP-binding socket to that of TrwB, where all residues characteristic for ATP hydrolysis were present in the same location as in the experimental structure of TrwB, which confirms the functional plausibility of the model (Figure 27). Lastly, the experimentally obtained cryo-EM maps of dimeric and trimeric TraD as well as hcp-TraD both fit the prediction model, (Figures 31&34). All the mentioned factors increase my trust in the model, despite it already having a very reliable confidence score, and allow for subsequent structural analysis of the obtained model of TraD.

The overall topology of TraD is nearly identical to that of TrwB, suggesting that despite being distantly related T4SS, albeit performing the same conjugative function, the organisation of T4CP of conjugative systems remains the same (Figure 26). The same could not be said about alignment between TraD and Cagβ, an effector translocating T4CP from the Cag-T4SS of *H. pylori* not involved in bacterial conjugation (Figure 26). This implies that conjugative T4CP assume a specific architecture exclusive to conjugative systems, as opposed to effector translocating T4CP which exhibit different architectures to facilitate interactions with their respective types of substrates. Presumably, that is due to the role of conjugative T4CP in translocating DNA, a substrate that will always be of similar structure - single-stranded nucleic acid with a negative charge. On the other hand, the substrate effector proteins will differ in their shapes and charges, and thus effector translocating T4CP will represent a higher level of

structural variability to match the diversity of their substrates. This is also in line with the fact that DotL, the T4CP of *L. pneumophila* Dot/Icm-T4SS must employ a range of adaptors that will facilitate its interactions with different substrates (Figure 12D) (Meir *et al.* 2020). No such adaptors have been observed for TrwB or TraD, indicating that their structures have inherent adaptations for the interaction with substrate DNA.

The prediction model displays the details of the nucleotide-binding cassette, which resembles characteristics typical of AAA+ ATPases: the presence of Walker A and Walker B motifs and an arginine finger provided *in trans* by an adjacent subunit. In TraD, these correspond specifically to Gly197, Lys198, Ser199 as Walker A motif; Asp425 and Glu426 as Walker B motif, and Arg44 of the neighbouring TraD subunit as the arginine finger (Figure 27). Moreover, I biochemically confirm the ability of TraD to hydrolyse ATP (Figure 21A) and provide evidence that ATP hydrolysis induces TraD oligomerisation (Figure 29). This is in line with the localisation of the ATP binding site at the interface between adjacent TraD subunits (Figure 27), thus the affinity of TraD subunits to ATP might aid its dimerisation and further oligomerisation. Functionally, such a mechanism of oligomerisation might be important to prevent maintaining an unnecessary pore within the inner membrane when DNA is not actively transferred - that is no ATP hydrolysis occurs. Conversely, the pore will only form upon a functional trigger – that is ATP hydrolysis – and will disassemble shortly after, when it is no longer needed.

The slice through the TraD hexamer reveals positively charged lysines (Lys273, Lys335, Lys467, and Lys488), which form the lining of the central lumen of TraD (Figure 24), reminiscent of FtsK – an AAA+ ATPase involved in DNA translocation during cell division, which also has a positively charged lumen made up of lysine and arginine residues (Jean *et al.* 2020). Similar topology of their lumens might suggest that TraD and FtsK can share a similar mechanism of substrate translocation, whereby substrate DNA interacts with the positively

charged residues, which are then reiteratively displaced in one direction by conformational changes induced by the ATPase activity, resulting in the transfer of DNA. This is in line with my results evidencing DNA binding and DNA-induced conformational changes by TraD (Figure 21B&C).

Lastly, I propose a model of TraD operation during bacterial conjugation (Figure 35), which adds molecular detail to its role as the coupling protein based on the presented results. Initially, TraD exists as a monomer and forms a pool of readily available subunits in proximity to its cognate T4SS apparatus (Figure 35, step 1), reminiscent of its monomeric state in solution (Figure 19). In the cytosol, the relaxosome complex (particularly, proteins TraM and TraI) begins to assemble upon the o*riT* region of the F-plasmid DNA. Then, through its pronounced flexible and unstructured CTD, TraD interacts with TraM and hence captures the processed relaxosome-DNA complex (Figure 35, step 2). Subsequently, the oligomerisation of TraD occurs through ATP hydrolysis, forming a channel in the inner membrane, which will serve for DNA translocation (Figure 35, step 3). Interaction between TraD and negatively charged DNA is mediated by the positively charged lysine residues that protrude towards the lumen of the TraD channel and the transfer is energised by ATP hydrolysis. Once all DNA is transferred, and no ATP hydrolysis occurs, TraD instantly disassembles, preventing maintaining a pore in the inner membrane that is no longer needed (Figure 35, step 4). Now, the available TraD subunits can be recycled for another round of conjugation.

Overall, several properties of TraD were documented here, including ATP hydrolysis, DNA binding, and mechanism of oligomerisation, providing molecular detail to our understanding of TraD operation in bacterial conjugation of the F-plasmid. In return, the multifaceted function of TraD makes it a great drug target against conjugative transfer.

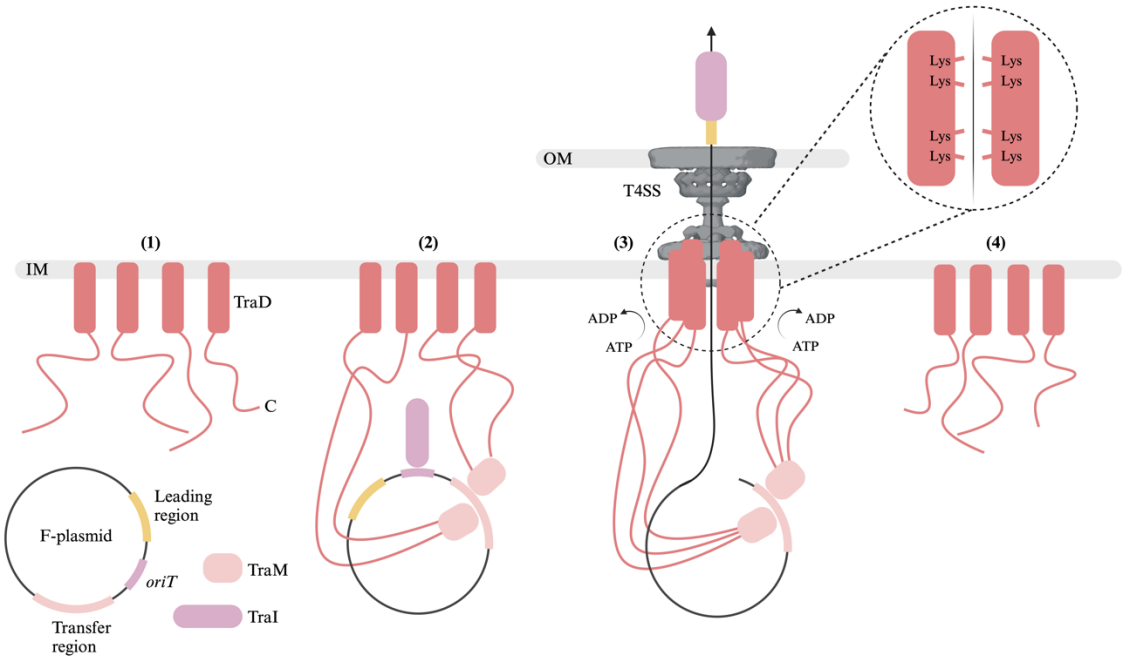


Figure 35. Functional transitions of TraD during bacterial conjugation

Operation of TraD can be divided into four main steps. **(1)** TraD forms a pool of monomeric subunits, that reside in the inner membrane. **(2)** As the relaxosome (TraM and TraI) begins assembling on the F-plasmid DNA, TraD interacts with TraM via its CTD, hence bringing substrate DNA towards TraD. **(3)** TraD oligomerises upon ATP hydrolysis, which energises translocation of DNA, that is orchestrated by the positively charged lysine residues (blowout). **(4)** Once transfer is finished, TraD returns to its non-oligomerised form to prevent leakage through its lumen across the inner membrane.

Chapter 3 Biomechanics of the F-pilus and the role of its phospholipids in conjugation

3.1 Introduction

Bacterial conjugation is the major contributor to the rapid spread of AMR genes (Lederberg and Tatum 1946, Thomas and Nielsen 2005, Murray *et al.* 2022) and relies on bacterial surface filaments generated by T4SS, known as conjugative pili. These elongated, helical appendages scan the environment to find a mating opportunity, and once such is found, they are critical to establishing and maintaining a physical connection between the donor cell and the recipient cell to facilitate DNA passage between them (Clarke *et al.* 2008, Arutyunov and Frost 2013, Costa *et al.* 2015, Waksman 2019, Costa *et al.* 2021)

Various conjugative plasmids circulate in nature, however the IncF plasmids, with their conjugative F-pili have become the predominant plasmid family found in bacteria isolated from both human and animal sources and are commonly found in human enteropathogenic species, such as *E. coli*, *S. enterica*, *K. pneumoniae*, *Shigella spp.* (Rozwandowicz *et al.* 2018). IncF plasmids are prevalent mainly in bacterial isolates from urine (Bengtsson *et al.* 2011, Narciso *et al.* 2012, Mukherjee and Mukherjee 2019, Decano *et al.* 2021) and blood samples (Mshana *et al.* 2013, Usein *et al.* 2016, Hastak *et al.* 2020, Shankar *et al.* 2022) and were shown to mobilise in mouse gut microbiome at similar rates to those obtained *in vitro* (Licht et al. 1999), documenting their incredibly high conjugation efficiencies. It is not known what contributes to the ubiquity of IncF plasmids and their high dissemination efficiency in human and animal microbiomes; however, it suggests a specific adaptational advantage of IncF plasmids which other types of plasmids lack.

The most fundamental factor that can affect the ubiquity of conjugative plasmids in a specific environment is their ability to successfully establish and maintain mating pairs in that particular

environment. This function is attributed to the conjugative pili, which are morphologically diverse between different incompatibility groups, for instance, the P-pilus of the IncP plasmid encoded VirB/D4 system of *A. tumefaciens* has a short and rigid morphology, shared with the N-pili of IncN plasmids such as R388 and pKM101; on the other hand, the F-pili of IncF plasmids are significantly less rigid and much more flexible (Bradley 1980). Differences in the flexible versus rigid morphology of pili correlate with their ability to conjugate either on solid surfaces or liquid media, as rigid pili are not proficient in conjugation in liquids (del Campo *et al.* 2012, Arutyunov and Frost 2013, Virolle *et al.* 2020). In some cases, such as the I-pili of IncI plasmids, rigid pili require additional supporting flexible pili that enable liquid matings, and their removal abrogates liquid mating (Horiuchi and Komano 1998). This contrasts with the flexible F-pili, which inherently conjugate well in liquid environments (Bradley 1980, Virolle *et al.* 2020). Moreover, the F-pili also play a role in biofilm formation and help establish early cell-to-cell contacts that in turn lead to biofilm formation (Ghigo 2001, Reisner *et al.* 2003, May *et al.* 2011) and knockout mutation of the F-pilus was also shown to be detrimental to the ability of adherent-invasive *E. coli* to adhere to epithelial cells during mouse gut infection (Elhenawy *et al.* 2021). It is not known what allows the F-pili to operate in liquids and provide such a variety of functions.

At the atomic level, all currently known conjugative pili are made up of pilin subunits and the associated phospholipid molecules in 1:1 stoichiometry (Costa *et al.* 2016, Zheng *et al.* 2020, Amro *et al.* 2023, Beltran *et al.* 2023, Costa *et al.* 2023, Kreida *et al.* 2023). The pilin subunits assume similar folds, with minimal differences between minimised, expanded, and archaeal systems (Figure 11) (Beltran *et al.* 2023, Costa *et al.* 2023). Incorporation of phospholipids into the atomic structure is a unique trait of conjugative pili that separates them from other bacterial surface filaments, such as fimbriae, flagella, or type IV pili, where the presence of

phospholipids has not been documented (Hospenthal *et al.* 2017). Nevertheless, the function of phospholipids in the conjugative pili also has been unknown.

I hypothesised that the environmental success of IncF plasmids, exhibited by their dominance over other conjugative plasmids in human and animal isolates (Rozwandowicz *et al.* 2018), is ensured by the unique adaptations present in the atomic structure of the F-pilus to form flexible but highly robust connections. Here, I present that hydrodynamic stress accelerates F-pilus-mediated conjugation and biofilm formation; I document the biomechanical properties of the F-pilus, such as its resistance to bending and stretching, which prevent snapping and ensure successful operation in those turbulent environments. I furthermore attribute these properties to the presence of phospholipids in the F-pilus molecular structure, thereby revealing for the first time the role of phospholipids as an essential adaptation that grants the robustness of conjugative pili. All this provides a better understanding of the evolutionary success of IncF plasmids in mobilising at high frequencies in humans and animals, attributed to a single molecular detail of the F-pilus structure - the presence of the phospholipids.

3.2 Experimental procedures

3.2.1 Materials

Bacterial strains (Table 3) and plasmids (Table 4) used in this study.

Table 3. Bacterial strains

Bacterial strain	Information	Source
JE2571	Harbours the plasmid pED208 encoding the F-T4SS	(Costa *et al.* 2016)
EDL933 F+	Pathogenic *E. coli* O157:H7, an enterohemorrhagic strain transformed with pOX38:Cm^R plasmid, making it F+.	This study
EDL933 F-	Pathogenic *E. coli* O157:H7, an enterohemorrhagic strain.	Penadés Lab
DH5α F+	Common laboratory strain, transomed with pOX38:Cm^R plasmid, making it F+.	(Costa *et al.* 2016)
DH5α F-	Common laboratory strain.	Costa Lab
DH10β	Common laboratory strain used as recipient for conjugation. Str^R	Costa Lab

Table 4. Plasmids and genetic constructs used in this study

Plasmid	Information	Source
pED208	Derepressed IncF plasmid	Costa Lab
pOX38	Repressed IncF plasmid, a direct derivative of the original F plasmid generated by HindIII digestion and recircularisation of the largest fragment that retained conjugative ability. Cm^R	Costa Lab

3.2.2 Conjugation efficiency assay

E. coli strains DH5α and EDL933 harbouring the pOX38 plasmid (donor) were revived from glycerol stocks by streaking on LB agar plates with 30 µg/ml chloramphenicol and the *E. coli* DH10β strain (recipient) was streaked on LB agar plates with 100 µg/ml streptomycin. On the next day, single colonies were picked to inoculate an overnight culture at 37 °C with shaking. The following day, 500 µl of each culture was used to inoculate independent 25 ml cultures with the appropriate antibiotics, and growth was monitored until OD_{600} reached 0.6 in both cultures. Subsequently, cells were washed twice in 25 ml of chilled LB to eliminate antibiotics. The donor cell pellet was resuspended in 10 ml of chilled LB, while the recipient cells were resuspended in 1 ml of chilled LB. In a separate pre-chilled tube, 1 ml of each resuspension was combined and thoroughly mixed by pipetting, generating a suspension with a 1:10 donor to recipient ratio. 500 µl of the conjugation suspension was used to inoculate four different tubes filled with 4.5 ml of LB. Each tube was placed in the respective tested condition: (1) steady (no shaking); (2) shaking at 180 rpm; (3) shaking at 180 rpm and supplemented with 2.5 g of sterile glass beads (2-3 mm in diameter); (4) shaking at 180 rpm but interrupted with one-minute high-speed vortexing every four minutes. All conjugation reactions were allowed to proceed for 30 minutes at 37 °C. The conjugation reaction of plasmid-cured DH5α and EDL933 (F−) followed a similar procedure, except for the addition of chloramphenicol to

culture the cells. To evaluate the number of transconjugants in each condition, serial dilutions of the resulting conjugation reactions were prepared, ranging from 10^{-1} to 10^{-6}, and 7 µl of each dilution was plated on LB agar supplemented with 30 µg/ml chloramphenicol and 100 µg/ml streptomycin. Additionally, to count the total viable transconjugants and recipients, 7 µl of each dilution was plated on LB agar supplemented with 100 µg/ml streptomycin. Colony counts for each condition were statistically compared using two-way ANOVA with Tukey's adjustment for multiple comparisons - performed and visualised using GraphPad Software (www.graphpad.com). The experiments were conducted in six biological replicates (n = 6).

3.2.3 Biofilm formation assay

Overnight bacterial cultures (20 ml each) were cultivated in LB medium (DH5α and EDL933), LB supplemented with 30 µg/ml chloramphenicol (DH5α and EDL933 with pOX38 plasmid), or LB supplemented with 100 µg/ml streptomycin (DH10β). The following day, the cultures were centrifuged at 2000 x g for 20 min and then washed with 50 ml fresh LB to eliminate antibiotics. After an additional centrifugation round, the resulting pellet was resuspended in 5 ml fresh LB. For the assessment of biofilm formation in dynamic environments, 100 µl of each donor strain (F+) or plasmid-cured strain (F−) was mixed with 1 ml of DH10β recipient strain and thoroughly pipetted. The resulting inocula were evenly distributed into three glass tubes containing 3 ml LB for each tested condition. Under steady conditions, the culture was incubated in a stationary incubator, while conditions involving shaking were incubated at 180 rpm. In the latter case, glass beads (1.5 g of sterile glass beads with a diameter of 2-3 mm) were added. Biofilms were allowed to develop for 24 h at 37 °C with continuous shaking. For quantification, resulting biofilms underwent careful washing five times with 10 ml of water to remove unattached planktonic cells. Subsequently, 1 ml of crystal violet suspension was added and incubated on a roller for 10 min to ensure staining of biofilms. After removing excess

crystal violet, additional five rounds of 10 ml washes with water were performed. Stained biofilms were solubilised by adding 1 ml of 30% acetic acid and vortexing until no crystal violet remained on the glass tube walls. Absorbance was measured on 10-fold dilutions of resulting crystal violet suspensions using a 96-well plate reader at a wavelength of 595 nm. The experiments were conducted in three biological replicates, each followed by three replicates of absorbance measurements (n = 9). The resulting values were plotted using GraphPad Prism. Statistical significance was determined by a two-way ANOVA test with Tukey's adjustment for multiple comparisons. To visualise bacteria within the biofilm matrix through negative-stain electron microscopy, 10 µl of bacterial cell culture prior to crystal violet staining was saved and imaged as explained in *Microscopy procedures.*

3.2.4 Purification of F-pili

E. coli JE2571, harbouring the F-plasmid pED208, were cultivated on large LB agar plates. The bacterial lawn was collected from the plate surface using SSC buffer (150 mM NaCl pH 7.2 15 mM sodium citrate). The cells were left to resuspend for 2 h at 4 °C with gentle agitation, followed by two centrifugations at 10,800 × g for 20 min each. The supernatant obtained was precipitated using 500 mM NaCl and 5% PEG 6,000, and the mixture was allowed to incubate for 2 h at 4 °C. The resulting precipitate was separated by centrifugation at 15,000 × g for 10 min after resuspending in sterile water. The supernatant was subjected to a second precipitation following the same procedure, followed by centrifugation at 15,000 × g for 20 min. The resulting F-pili pellet was resuspended in PBS at pH 7.5 and applied onto a pre-made CsCl step gradient (1.0–1.3 g/cm3). After centrifugation at 192,000 × g for 17 h at 4 °C, the fraction containing the F-pili was carefully collected and extensively dialysed against PBS at pH 7.4.

3.2.5 Microscopy procedures

To conduct negative staining of whole bacteria producing F-pili, a 10 µl culture of bacterial cells (grown in LB until OD_{600} reached 1.5) was applied onto glow-discharged 300 mesh carbon-coated grids (Agar Scientific) and incubated for 2 min. After two subsequent washes with 10 µl of water, the grids were stained with 0.2% (w/v) phosphotungstic acid for 7 s before blotting. Subsequently, the grids were imaged using a Technai12 Spirit transmission electron microscope (FEI) operating at 120 kV and equipped with a 2 K Eagle camera (FEI). For purified F-pili, they were applied onto glow-discharged 400 mesh carbon-coated grids (Agar Scientific) and incubated for 2 min, followed by two washes with 10 µl of water. The grids were then stained with 2% uranyl acetate for 1 min, blotted with filter paper, and imaged using a Technai12 Spirit transmission electron microscope (FEI) operating at 120 kV and equipped with a 2 K Eagle camera (FEI).

3.2.6 Cryo-EM imaging of F-pili

A 4 µl aliquot of the purified F-pili sample was applied to glow-discharged 400 mesh copper Lacey grids coated with a 2 nm layer of carbon. Grids were vitrified using a Vitrobot (FEI) operating at 100% humidity and 25 °C, with a wait time of 1 min and a blot time of 3.5 s before the grid was plunged into liquid ethane for freezing. Imaging was conducted on a FEI Technai G2 Polara operating at 300 kV, equipped with a Gatan K2 Summit direct electron detector, a Quantum energy filter, and an energy-selecting slit width of 20 eV. The images were captured with a pixel size of 1.1 Å/pixel and a total dose of 50 $e^-/Å^2$, split over 60 frames. Subsequently, the movie frames were aligned and summed using the MotionCor2 software.

3.2.7 Thermochemical stability assessment of F-pili

The purified F-pili underwent a thermochemical stability test wherein they were incubated for 1 h in water, 1 M NaOH, or 8 M Urea, followed by boiling at 100 °C. Subsequently, the samples were imaged using TEM, employing the same procedure as described previously.

3.2.8 Bacterial strain and growth conditions for force measurements

The E. coli JE2571 strain carrying the F-plasmid, pED208, was revived from frozen storage by streaking for well-isolated single colonies on Lysogeny broth (LB) agar, supplemented with 96 μ/mL 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-gal). After overnight incubation at 37 °C, a well-isolated single dark blue pigmented colony was selected and inoculated into 10 ml fresh LB, then incubated at 37 °C for 18 h with shaking at 160 rpm. The following day, bacteria were subcultured into 10 ml of fresh LB and further grown at 37 °C until reaching an OD_{600} of 0.6 to 0.8. The bacteria were maintained at 37 °C without shaking before conducting force-extension measurements. Prior to the force-extension measurements, 10 μl of the liquid culture was suspended in 1000 μl of ice-cold PBS. This step was taken to halt and slow down potential polymerisation or depolymerisation during the measurements. The force-extension response of F-pili was measured for 30 cells from 2 biological replicates.

3.2.9 Force measurements and sample preparation

Force spectroscopy measurements were undertaken using a custom-built optical tweezers setup centred around an inverted microscope (Olympus IX71, Olympus, Japan) equipped with a high numerical aperture water immersion objective (model: UplanSApo60xWIR 60X N.A. = 1.2; Olympus, Japan). Trapping of cells and microspheres was accomplished using a 2.0-watt DPSS laser (Rumba, Cobolt AB, Solna, Sweden) operating at 1064 nm in continuous wave (CW) mode. The position of the trapped object was tracked by directing the beam of a low-power fiber-coupled HeNe laser (operating at 632.8 nm, 117 A, Spectra Physics/Newport, Irvine, CA) onto a position-sensitive detector (L20 SU9, Sitek Electro Optics, Partille, Sweden), enabling sub-nanometer spatial resolution of a trapped object. Prior to constructing a flow chamber, a 1:500 suspension of 9.5 μm carboxylate-modified latex beads (product no.2-10000, Interfacial Dynamics, Portland, OR) was prepared in Milli-Q water. These larger-sized beads were employed as platforms for bacteria to keep them distant from the coverslips, thus preventing

interactions with the surface. Ten µl of the bead-water suspension were deposited onto 24 × 60 mm coverslips (no.1, Paul Marienfeld GmbH & Co, Lauda-Königshofen, Germany) and incubated in an oven at 60 °C for 60 minutes to immobilise the beads on the surface. To enhance bacterial adhesion to the beads, a solution of 20 µl of 0.01% poly-L-lysine (catalogue no. P4832, Sigma-Aldrich, St. Louis, MO) was applied to the coverslips, followed by incubation in the oven at 60 °C for 45 minutes. The coverslips were then gently rinsed to remove excess poly-L-lysine. A sample was prepared by encircling the area containing the poly-L-lysine-coated beads with a ring of vacuum grease (Dow Corning, Midland, MI). Next, 5 µl of bacterial suspension and a 5 µl suspension of probe beads (surfactant-free 2 µm white amidine polystyrene latex beads, product no. 3-2600, Invitrogen, Carlsbad, CA) were delicately placed onto the area and sealed by positioning a 20 × 20 mm coverslip (no.1, Paul Marienfeld GmbH & Co, Lauda-Königshofen, Germany) on top. A single bacterium was trapped and securely affixed onto a poly-L-lysine-coated 9.5 µm latex bead. Subsequently, a 2 µm amidine bead was trapped, and the trap stiffness was calibrated using the power spectrum. To evaluate the stability of the setup and determine the optimal calibration time, an Allan variance assay was conducted, followed by attaching the trapped bead to an F-pilus by manoeuvring the bead in proximity to the mounted bacterium. By zooming in optically on the live video feed, it became feasible to observe when a bead interacted with a pilus. Once the pilus was attached to the bead, the tensile force was exerted by translating the piezo-stage (PI-P5613CD, Physik Instrumente GmbH & Co, Karlsruhe, Germany) utilising a custom LabView program.

The eWLC fitting was conducted as in Patkowski *et al.* (2023)

3.2.10 Molecular dynamics simulations

Molecular dynamics simulations were carried out utilising the initial structure of the F-pilus acquired from PDB ID: 5LEG, examining both the F-pilus with and without phospholipids. To ensure a manageable initial structure for simulation, 11 layers of the fundamental F-pilus

pentamer and 10 layers of interacting lipid chains were incorporated into the simulation, resulting in 55 protein subunits and 50 lipid molecules within the simulated systems. The simulations and system preparation were conducted using Amber20/AmberTools21 software. System parameterisation was achieved with tLeap, employing force field parameters from FF14SB for the protein (Maier *et al.* 2015), TIP3P for water (Jorgensen *et al.* 1983), and Lipid21 for phospholipids (Dickson *et al.* 2022). Monatomic counterions were introduced for system charge neutralisation, following the parameters outlined by Joung and Cheatham (Joung and Cheatham III 2008). The protein was solvated in a rectangular water box measuring 12 Å x 12 Å x 100 Å, oriented along the z-axis, with a substantial water buffer in the z direction allowing for system extension during steered molecular dynamics simulations. Energy minimisation consisted of 5000 steps, with restraints of 10 kcal mol^{-1} $Å^{-2}$ on alpha carbon atoms. Gradual heating occurred in two stages, reaching 100 K over 20 ps and then 310 K over another 20 ps. Equilibration comprised six stages, progressively reducing restraints, with the final coordinates from equilibration initiating steered molecular dynamics simulations. Steered molecular dynamics simulations were conducted in the NPT ensemble at 310 K, with the z-distance between alpha carbons of the bottommost and topmost protein pentamers serving as the collective variable. A restraint of 0.5 kcal mol^{-1} $Å^{-2}$ was applied to prevent translations and rotations, utilising pulling speeds of 1 Å/ns. The integration timestep for all phases was 2 fs, with the SHAKE algorithm utilisation for hydrogen bond constraints (Ryckaert *et al.* 1977). Temperature and pressure were maintained using the Langevin thermostat and Monte Carlo barostat, respectively. Long-range electrostatics were managed using the particle mesh Ewald method, with a real space interaction cut-off of 8 Å (Salomon-Ferrer *et al.* 2013). All analyses were conducted using Visual Molecular Dynamics (VMD) and AmberTools21, with visualisation, including image and movie creation, accomplished through VMD (Humphrey *et al.* 1996).

3.2.11 f1 phage infection

The assay for filamentous f1 phage infection began with the utilisation of an overnight culture (500 µl) of *E. coli* containing the pOX38 plasmid (donor), which was used to inoculate 20 ml of fresh LB supplemented with 30 µg/ml chloramphenicol. The culture was allowed to grow at 37 °C with shaking until the OD_{600} reached 0.6. Subsequently, cells were collected by centrifugation at 2,000 x g for 10 minutes and resuspended in 5 ml of fresh LB. To initiate the infection assay, 2 µl of a 10^{-8} dilution from a clarified f1 phage lysate was added to the cell suspension, and the solution was thoroughly mixed by pipetting. The resulting mixture was then divided into two separate tubes, each containing 2.5 ml. One tube was kept in a static incubator, while the other was subjected to shaking at 180 rpm. Both infection reactions were allowed to proceed at 37 °C for 1 hour. To evaluate the infection rates at specific time intervals (5 min, 30 min, 1 h), 400 µl samples were withdrawn from each condition. These samples were mixed with 3 ml of LB-agarose (prepared as LB with 0.4% w/v agarose) and evenly spread on "bottom agar," which consisted of LB agar supplemented with 30 µg/ml chloramphenicol. This process followed the standard double-layer plating procedure (Stachurska *et al.* 2021). The plates were left to dry for 20 minutes and then incubated overnight at 37 °C, resulting in the formation of well-defined plaques on the following day.

3.3 Results

3.3.1 Conjugative transfer is enhanced under hydrodynamic stress

Conjugation has traditionally been viewed as delicate, requiring stable conditions to preserve bacterial mating. However, this does not accurately represent the conjugation resilience in the dynamic environments of the human gut or natural habitats. In nature, bacterial conjugation occurs in a variety of environments, and thus conjugative pili adapted to formation of mating pairs in varying levels of hydrodynamic stress, including in highly turbulent environments. To test how the dynamics of the environment affect bacterial conjugation via the F-pilus, I

conducted an array of conjugation reactions with increasing mechanical stress as a way to imitate the varying conditions that bacteria experience in nature. Two different strains were tested as donors, a laboratory strain *E. coli* DH5α and an enterohaemorrhagic *E. coli* (EHEC) O157:H7 clinical strain EDL933, both harbouring the conjugative pOX38 plasmid and elaborate F-pili (designated F+ strains). The conjugative pOX38 plasmid is a direct derivative of the original F-plasmid generated by HindIII digestion and recircularisation of the largest fragment that retained conjugative ability with the addition of a chloramphenicol resistance cassette, and it has been widely used as the representative of the IncF family. Both donors, DH5α F+ and EDL933 F+, were combined with a recipient strain, DH10β *E. coli*, chromosomally streptomycin resistant. As a control, the same reaction was conducted with plasmid-cured donors without conjugative F-pili (DH5α F- and EDL933 F-). Such prepared conjugation reactions were then split into four conditions represented in Figure 36A: conjugation in steady media without shaking; shaken at 180 rpm, shaken at 180 rpm with 50% (w/v) glass beads, and in media vortexed for 1 min every 4-minute interval, and conjugation was allowed to proceed for 30 minutes at 37°C in each condition. When conjugation is finished, cells were counted by plating on chloramphenicol and streptomycin plates to select for transconjugants (which will be chromosomally streptomycin resistant and chloramphenicol resistant from the pOX38 plasmid), as well as on plates with streptomycin to count the total viable transconjugant and recipient population – representative results are shown in Figure 36B. Conjugation efficiency was calculated as a ratio between colony forming units (CFU) of transconjugants (T) over transconjugants and recipients (T+R).

The obtained transconjugant numbers revealed a significant rise in conjugation efficiency when conjugation proceeded in agitated media compared to static conditions in both DH5α and EDL933 strains (Figure 36C), specifically a three-log increase in the conjugation efficiency was observed in shaking conditions as opposed to the steady condition (Figure 36C).

Intriguingly, the inclusion of glass beads in the mating media, which introduced additional hydrodynamic and mechanical stress, did not alter conjugation efficiency when compared to matings without glass beads (Figure 36C). Furthermore, in a highly unexpected outcome, subjecting mating pairs to cycles of intense turbulence—characterised by 1 minute of vortexing followed by 4 minutes of steady incubation—did not completely abolish conjugation efficiency.

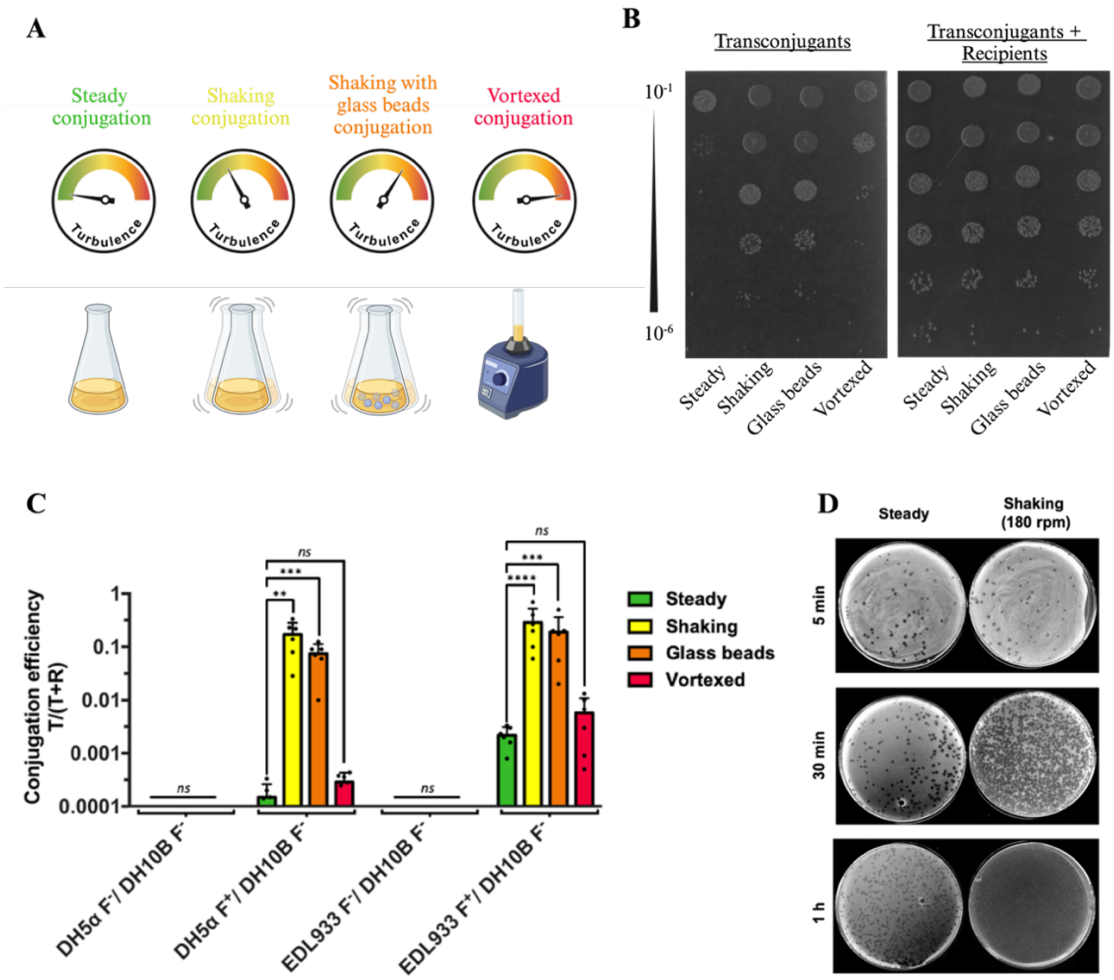


Figure 36. Effect of a dynamic environment on bacterial conjugation

A| Representation of tested conjugation conditions. From left: conjugation in steady media without any agitation; conjugation with agitation imposed by shaking at 180 rpm; conjugation with shaking at 180 rpm in media with glass beads; conjugation in steady media interrupted by vigorous vortexing for 1 minute every 4 minutes. **B|** Representative result of conjugation efficiency tested in different conditions. Left: plate with growth on media selective for transconjugants (conjugation efficiency).

Right: plate with grows on media selective for transconjugants and recipients (growth control). **C**| Share of transconjugants (T) over transconjugants + recipients (T + R) after 30 min of conjugation reaction. Two donor strains were used: DH5α (F+) and EDL 933 (F+) and two plasmid-cured strains DH5α (F−) and EDL 933 (F−) with the recipient strain DH10β (F−) in all four conditions described above—steady (green), shaking at 180 rpm (yellow), supplemented with Glass Beads (orange) and vortexed (red). All p values were derived from comparisons of each condition via two-way ANOVA with Tukey's correction for multiple comparisons. DH5α F+/DH10B F− steady vs shaking ***$p = 0.0008$ and steady vs glass beads *$p = 0.0243$; EDL933 F+/DH10B F− steady vs shaking ****$p = 0.0001$, steady vs glass beads ***$p = 0.0002$. $n = 6$ biological replicates. Bars in graphs represent mean values, whiskers represent standard deviation, and black circles represent each replicate value. ****$p < 0.0001$; ***$p < 0.001$; **$p < 0.01$; *$p < 0.05$; ns not significant. D| f1 filamentous phage infection in environments of different turbulence. Left column: infection in steady media without agitation at timepoints 5 min, 30 min, 1 hours. Right column: infection in media with shaking at 180 rpm in the same timepoints.

3.3.2 f1 phage infection is also enhanced in hydrodynamic conditions

It is not obvious what could be causing the increase in conjugation efficiency in turbulent environments. The most likely hypothesis is rather simple – the enhanced hydrodynamic flow of liquids provides a higher probability of encountering a recipient. As a proof-of-concept experiment, the f1 phage infection rate was tested in steady and shaking conditions, similar to those tested in the conjugation experiments. The f1 phage attaches specifically to the tip of the F-pilus and hence can represent a recipient cell albeit without triggering other polar effects that might influence conjugation efficiency, such as mating pair formation and stabilisation. If turbulent environments enhance the probability of encountering a recipient, the infection rates should increase alike. Therefore, a DH5α (F+) cell culture was infected with the f1 phage and split into two vials incubated with or without shaking for 30 min at 37°C. Aliquots were taken at different timepoints of reaction and plated on soft agar to allow for plaque-formation visualisation. The number of the plaque forming units (PFU) in the shaking conditions, and hence the attachment of f1 phage to the F-pilus, was clearly much greater in the shaking reaction as compared to the steady reaction (Figure 36D), confirming the previous

observations. This result also allows to conclude that the enhanced conjugation efficiency in turbulent environments is attributed solely to the operation of the F-pilus rather than other processes, such as mating pair stabilisation.

3.3.3 F-pilus accelerates biofilm formation

Cells that elaborate conjugative pili were shown to exhibit increased capacity to form biofilms (Ghigo 2001). However, the effect of the environmental turbulence in biofilm formation by the F-pilus remains unclear. Given the F-pilus role in facilitating bacterial contacts, it is conceivable that F-pili evolved not only for DNA delivery but also as "sterile" or "non-conjugative" adhesion appendages, aiding the establishment of donor-recipient biofilm communities, which would be in line with the presence of free F-pili deposited in the cell membranes as observed in cryo-ET studies of the F-T4SS lifecycle (Figure 9). Using light microscopy and confocal autofluorescence microscopy, biofilm formation ability of *E. coli* strain EDL933, with (F+) and without (F-) the F-plasmid pOX38 was assessed, showing prominent formation of aggregates, known as microcolonies, in the piliated (F+) population (Figure 37A). To investigate the impact of hydrodynamic stress on biofilm formation*, E. coli* strains DH5α and EDL933, with (F+) and without (F-) the F-plasmid pOX38, were co-incubated with an excess of recipient strain *E. coli* DH10β under three conditions: steady, shaking, and shaking with glass beads (50% w/v), mimicking increasingly dynamic environments as described previously for conjugation efficiency experiments (Figure 36A). After 24 hours of incubation at 37°C, planktonic cells were removed, revealing visible biofilms at the air-liquid interface in F+ populations with shaking. These biofilms were stained with crystal violet, and representative results are shown in Figure 37B. The absorbance measurements at 595 nm of the solubilised stain allowed for quantification of the results, confirming higher biofilm production under conditions of elevated hydrodynamic stress (Figure 37C), consistent with the enhanced conjugation efficiency observed (Figure 36C).

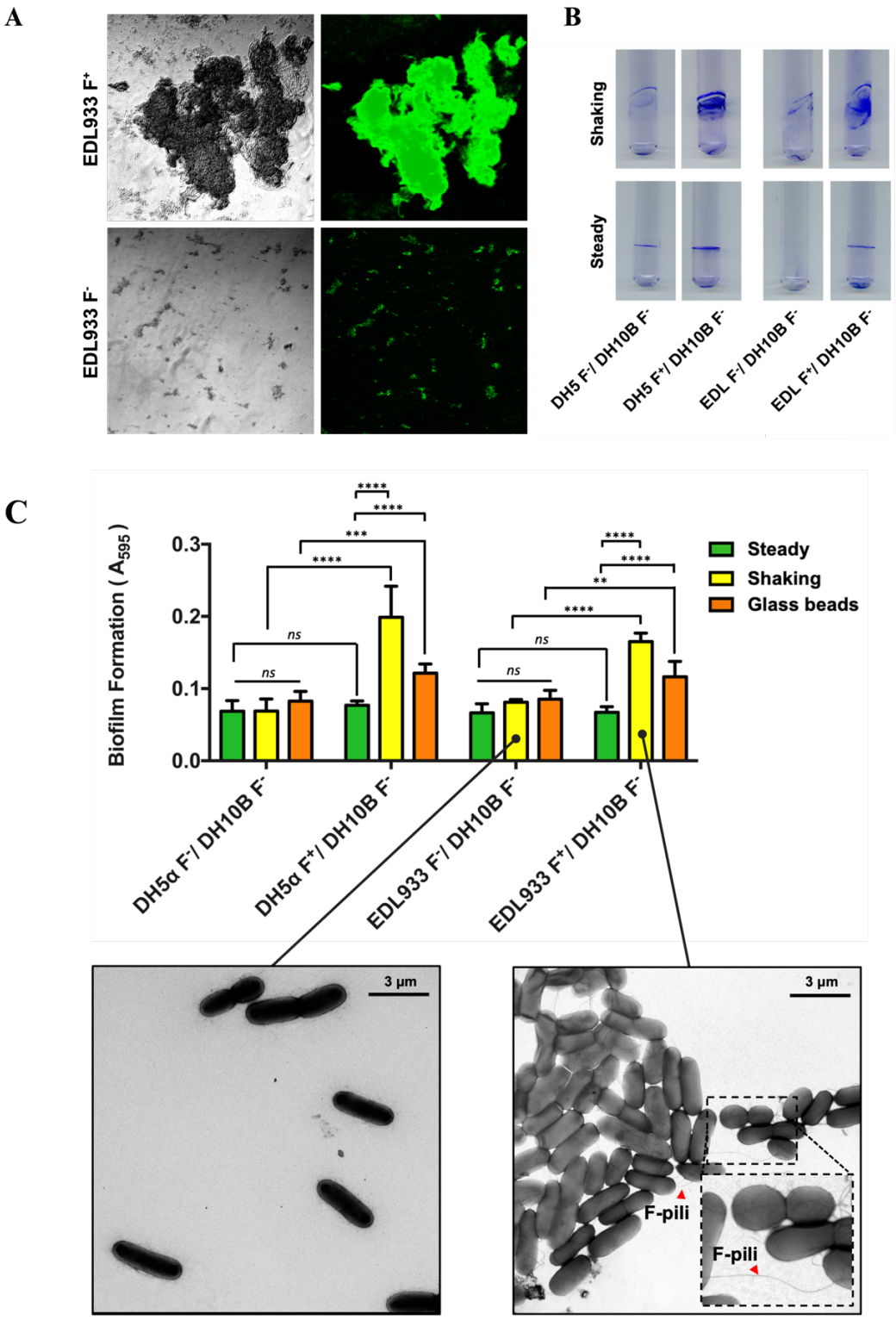


Figure 37. Effect of a dynamic environment on biofilm formation

A| Representation of F-pilus dependent biofilm formation by *E. coli* EDL933 using transmitted light confocal microscopy and autofluorescence confocal microscopy. Formation of biofilm was tested in F-plasmid containing (F+) strain and F-plasmid cured strain (F-). **B|** Representative result of F-pilus mediated biofilm formation in conditions of different turbulence. Two donors were tested, DH5α (F+)

and EDL 933 (F+) and two plasmid-cured strains DH5α (F−) and EDL 933 (F−) with the recipient strain DH10β (F−). Biofilms were allowed to form for 24 hours with or without shaking, after which planktonic cells were decanted and remaining biofilms were stained. **C|** Biofilm formation ability obtained after 24 h growth of two donor strains DH5α (F+) and EDL 933 (F+) and two plasmid-cured strains DH5α (F−) and EDL 933 (F−) with recipient strain DH10β (F−). All p-values were obtained via two-way ANOVA with Tukey's correction for multiple comparisons. DH5α F−/DH10B F− glass beads vs DH5α F+/DH10B F− glass beads $p = 0.0003$; EDL933 F−/DH10B F−glass beads vs EDL933 F+/DH10B F− glass beads $*p = 0.0106$. All p values denoted as **** were $p < 0.0001$. Data were obtained from $n = 3$ biological replicates measured in triplicate. Negative stain images for chosen conditions were presented below as an example of the F-pilus mediated cell aggregation. ****$p < 0.0001$; ***$p < 0.001$; **$p < 0.01$; *$p < 0.05$; ns not significant. Bars in all graphs represent mean values, whiskers represent standard deviation, and black circles represent each replicate value.

F-plasmid-cured DH5α (F-) and EDL933 (F-) cells formed minimal biofilms, whereas DH5α (F+) and EDL933 (F+) displayed substantial biofilm development under the shaking condition (Figure 37C). To visualise the formed biofilms with high magnification, NS-EM imaging was employed and revealed an intricate network of F-pili connecting bacterial cells, likely contributing to the biofilm matrix formation (Figure 37C). These findings suggest that the F-pilus plays a dual role in mediating enhanced conjugation efficiency as well as biofilm formation in dynamic environments.

3.3.4 The F-pilus is robust, flexible, and thermostable

I demonstrated that F-pilus mediated conjugation and biofilm formation thrive in dynamic environments (Figures 36&37), suggesting that the structure of the F-pilus filament may have evolved unique biomechanical properties beneficial to its operation in such conditions. Therefore, the dynamic properties of the F-pilus will be observed here.

Firstly, to elucidate the morphology of F-pili attached to bacterial cells, I conducted whole-cell NS-EM imaging of intact *E. coli* DH5α (F+) harbouring the pED208 plasmid, a member of the IncF family isolated from *S. typhimurium*, which constitutively expresses all *tra* genes, hence resulting in a higher number of pili per cell. The imaging revealed that most pili project straight

out of the cell (Figure 38A), while some exhibited pronounced curvature of nearly 180 degrees without signs of breakage (Figure 38A). For high resolution imaging, F-pili were sheared off from the bacterial surface and purified by caesium chloride density gradient, with purity confirmed by SDS-PAGE and further validation of TraA F-pilin presence via mass spectrometry. Cryo-EM imaging of vitrified F-pili affirmed their ability to endure significant bending without breakage (Figure 38B). Notably, the purified F-pili were more likely to carry portions of the membrane throughout purification steps rather than breaking apart, as signified by presence of bacterial membranes at the base of F-pili (Figure 38C). This further highlights the resilience of the filament.

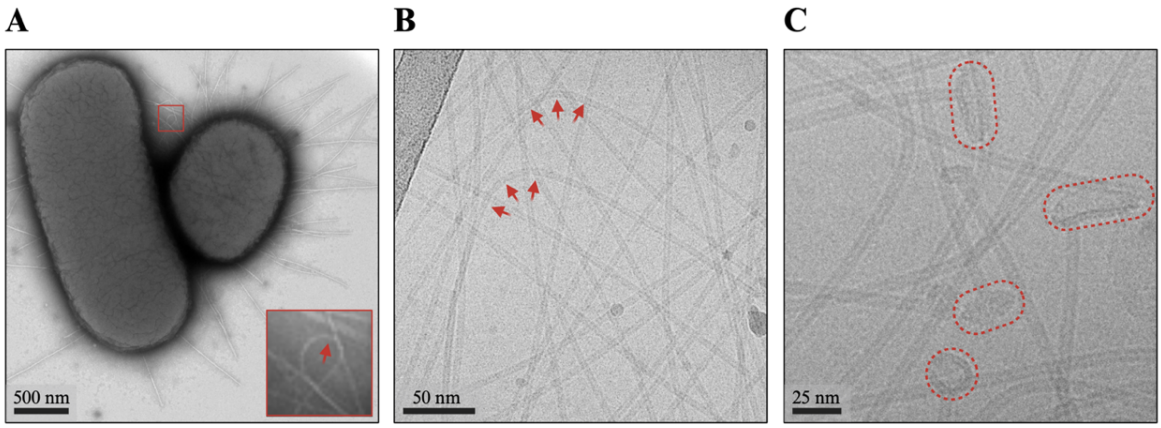


Figure 38. The F-pilus exhibits visible flexibility

A| NS-EM micrograph of *E. coli* DH5α harbouring the pED208 F-plasmid, the elaborated F-pili show a significant ability to form high degree bends, highlighted by a red arrow. **B|** cryo-EM micrograph of purified F-pili, showing F-pilus flexibility at higher resolution (red arrows). **C|** Purified F-pili are likely to crave portions of bacterial membranes (circled in dashed red lines) during purification, displaying significant robustness.

Secondly, the IncF plasmids have been reported in extremophilic species (Flores-Ríos *et al.* 2019), suggesting adaptations to withstand extreme environmental conditions, therefore the thermochemical resistance of F-pili was also assessed. The purified fractions of F-pili were subjected to extreme conditions, followed by NS-EM to monitor any morphology changes

compared to the morphology of F-pili in their native state (Figure 39A). F-pili exhibit resistance to treatments with 1 M NaOH (Figure 39B) or boiling in 8 M urea (Figure 39C), maintaining an intact morphology resembling that in water. Both conditions otherwise would trigger denaturation of protein complexes and their disassembly into smaller subunits, however, this was not the case for the F-pilus.

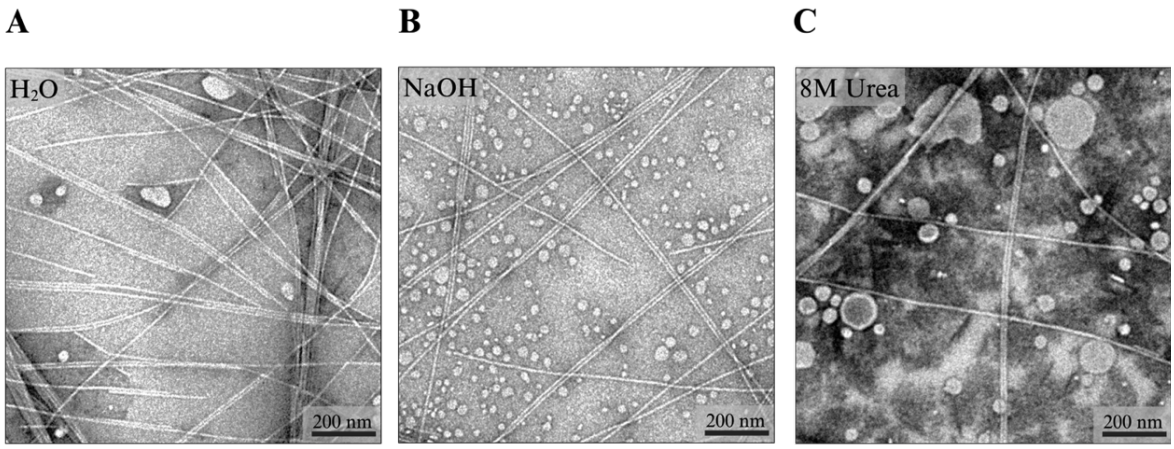


Figure 39. Testing thermochemical limits of the F-pilus

A| NS-EM micrograph of a control fraction of purified F-pili showing structured elongated filaments of F-pili. **B|** F-pili after treatment with 1M NaOH **C|** F-pili after treatment with 8M Urea, followed by boiling in 100°C for 30 min.

Overall, both the remarkable mechanical and thermochemical resistance of the F-pilus suggest evolution under selective pressure to ensure stability of the mating pair formed between bacterial cells during tight conjugative transfer, and potentially also protection of DNA from environmental stressors during DNA passage through the F-pilus during distant conjugative transfer.

3.3.5 The F-pilus is an elastic filament

Force measurements in this section were conducted by Tobias Dahlberg from the Andersson Lab (University of Umeå).

To evaluate the flexibility and resilience of F-pili, force measurements with optical tweezers were employed, a technique well-established for analysing the biomechanical properties of various pilus structures such as Type-1, Pap, and CFA/I pili (Jass *et al.* 2004, Andersson *et al.* 2012, Mortezaei *et al.* 2015). Utilising this setup, a controlled tension was applied to individual F-pili and their response was monitored with sub-pN precision. An illustrative setup is depicted in Figure 40A, with the corresponding graphical representation of the pilus behaviour during the experiment in Figure 40B. Initially, the F-pilus remains relaxed in the absence of external force, upon application of tension its quaternary structure extends (Figure 40B) and when the force is subsequently removed, the F-pilus retracts to its relaxed state (Figure 40B). The force response of the F-pilus is shown as a function of its extension and demonstrated by two curves, corresponding to F-pilus extension (black) and retraction (blue) (Figure 40C). Despite cycles of retraction and extension, no irreversible conformational alterations were noted in the pili at forces up to 100 pN, which exhibits purely elastic behaviour under those conditions, indicating spring-like properties of the F-pilus filament (Figure 40C). Any alterations, if present, would manifest as a sigmoidal (plateauing) force-extension response curve.

The biomechanical properties of flexible filaments are often described by two properties: persistence length and stretch modulus. The former is used to express how much force is needed to bend the filament, whereas the latter indicates how much force is needed to stretch it. To obtain these fractions, the experimentally obtained force response data needs to be fitted to an elastic worm-like-chain model (eWLC), which allows for a more accurate representation of how filaments respond to different forces.

The response curve of the retraction phase of the F-pilus (blue curve) was fitted to the corresponding eWLC model (black dashed curve), as depicted in Figure 40D. Consequently, the fit yields persistence length and stretch modulus, which for the 20 bacterial cells tested, assumed the average values of 104 ± 36 nm (mean ± SE) and 2330 ± 469 pN (mean ± SE),

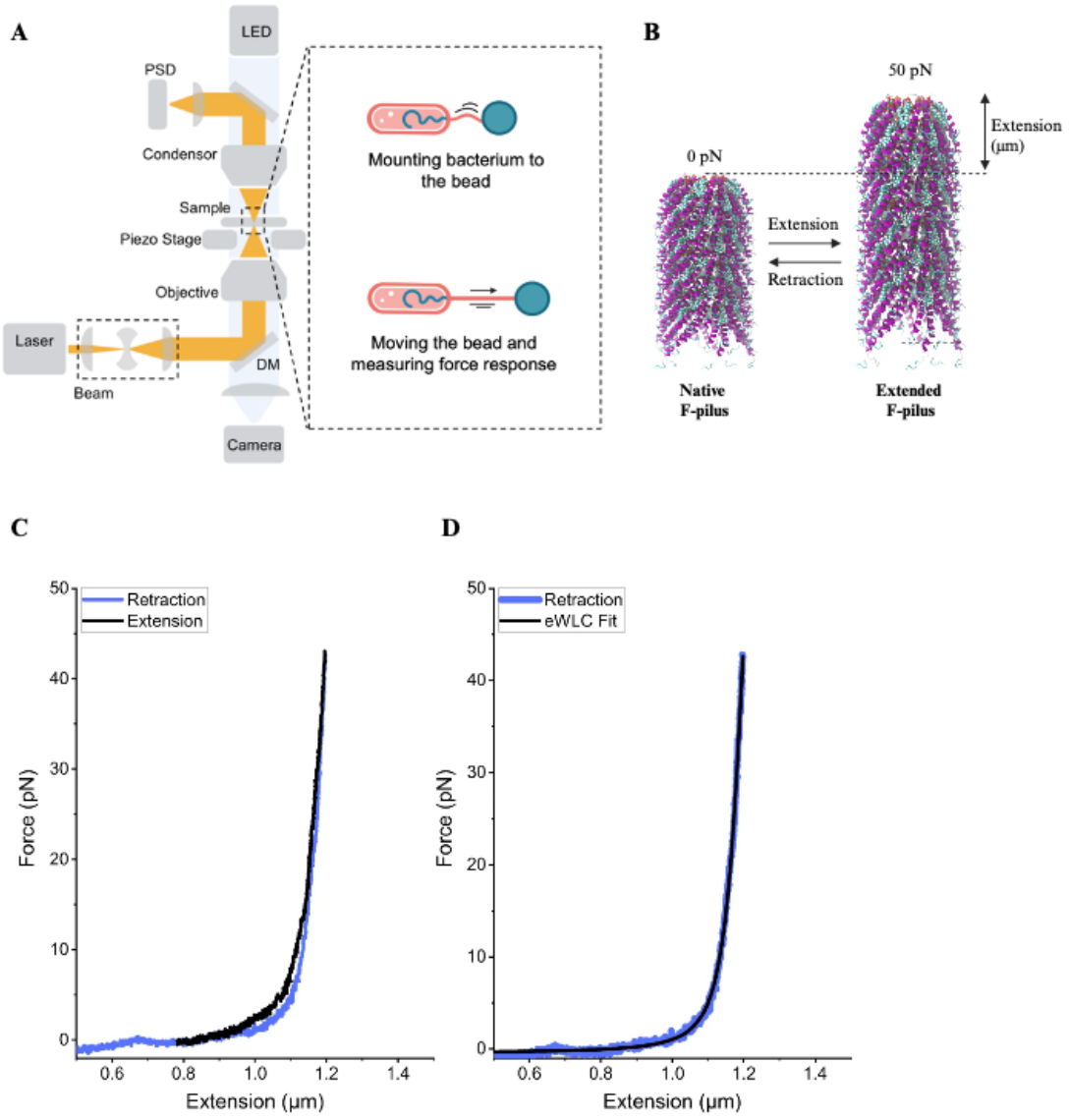


Figure 40. Force extension measurements of the F-pilus

A| The experimental setup of "molecular tweezers." A bacterium was immobilised on the stage and its F-pilus was attached to a bead. The bead was then moved, imposing extension of the F-pilus, after which it was relaxed, causing retraction of the F-pilus. Force response was measured. **B|** Graphical representation of the results. The native F-pilus represents the F-pilus before force is applied. After applying force, the F-pilus extends, and when the force is reduced, the F-pilus retracts, behaving like a molecular spring. Extension is the difference in filament length between these two states. **C|** Force response of the F-pilus as a function of F-pilus extension. The force response was measured during extension (black) as well as in the opposite direction – retraction (blue), and both curves overlap. **D|** Force response curve (blue) fitted with eWLC model (black).

respectively. Variability in these parameters, including outliers may arise from the potential attachment of two or more pili to the probe bead or bundling of a number of F-pili.

Altogether, the elasticity of the F-pilus is evident in Figure 40C, where a linear force increase is observed following a 1.2 µm extension until reaching approximately 45 pN, at which point extension ceases. Due to the absence of a covalent pilus-bead attachment, higher forces could not be applied, however, the obtained results suggest that the F-pilus exhausted the possibilities of the used models and can accommodate forces of at least 50-100 pN, akin to Type IV pili (Merz *et al.* 2000). Notably, the persistence length of F-pili (104 ± 36 nm) surpasses that of Type IV pili (1-50 nm), suggesting a more robust filament (Lu *et al.* 2015).

3.3.6 Presence of phospholipids confers structural robustness to the F-pilus

The molecular dynamics simulations in this section were conducted by Joseph Baker (The College of New Jersey).

The F-pilus is structurally made up of F-pilin subunits and phosphatidylglycerol phospholipid molecules in a 1:1 stoichiometry (Figure 11). Due to the essential role of phospholipids in bacterial viability, directly assessing the F-pilus structure lacking phospholipids is not possible *in vivo* (Costa et al. 2016). Hence, steered molecular dynamics (sMD) simulations were employed here to compare the properties of the native F-pilus structure (with phospholipids) to the F-pilus structure without phospholipids, thereby aiming at understanding the specific impact of PG phospholipids on the structural integrity of the F-pilus filament.

In the sMD simulations, the F-pilus tip was subjected to external pulling forces while restraining the base pentamer, mimicking experimental conditions. Notably, simulations with phospholipids depict elongation of the F-pilus without structural rupture, and illustrate the force applied to the native F-pilus beyond 2000 pN (Figure 41A) while those without phospholipids reveal severe structural failure under force application at much lower forces, of approximately

1000 pN (Figure 41B). The distinctive force peaks present in time-force curve of the native F-pilus (Figure 41A) correspond to extraction of single protein subunits from the pilus tip, while simulations lacking phospholipids display no discernible force peaks (Figure 41B), consistent with an overall structure failed of the F-pilus filament.

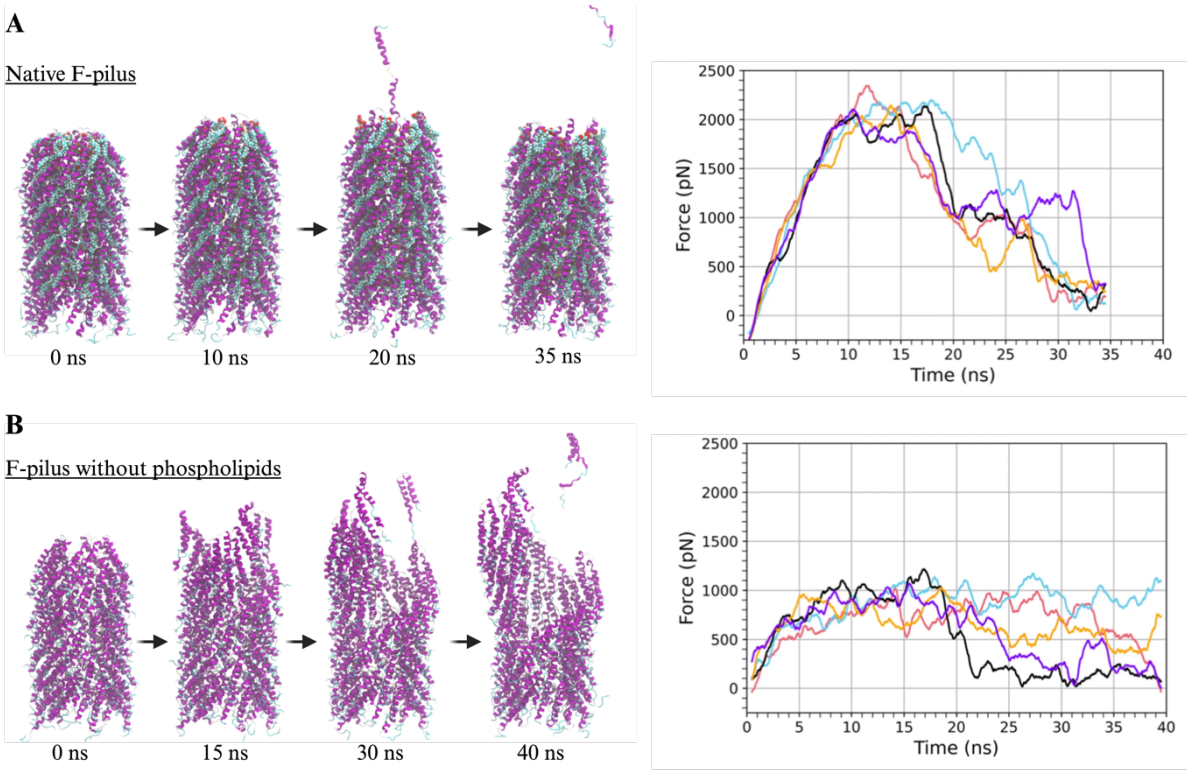


Figure 41. Steered molecular dynamic simulations of the F-pilus

Molecular dynamics simulations of the F-pilus with (A) and without (B) phospholipids upon application of a pulling force. **A|** Left: Four representative snapshots from the simulation show high resilience of the F-pilus despite the pulling force. Right: force versus time curve shown for five separate runs (each curve colour represents a different run). **B|** Left: Four representative snapshots from the simulation show low resilience of the F-pilus upon application of a pulling force. The filament falls apart easily. Right: force versus time curve shown for five separate runs (each curve colour represents a different run).

Equilibrium simulations, that is simulations in which F-pilus behaviour is monitored over time, however without applying pulling forces, similarly highlight significant differences between native F-pilus and the F-pilus lacking phospholipids, as evidenced by root mean square deviation (RMSD) calculations before and after 100 ns simulation. Divergence from the initial

structure is represented by higher RMSD values (Figure 42 A) and is clear in the absence of phospholipids (orange curve), contrasting with the relatively stable conformation of the filament of the native F-pilus (blue curve). The initial and final structures are shown in Figure 42 B, with a clear deterioration of the structure of the F-pilus lacking phospholipid.

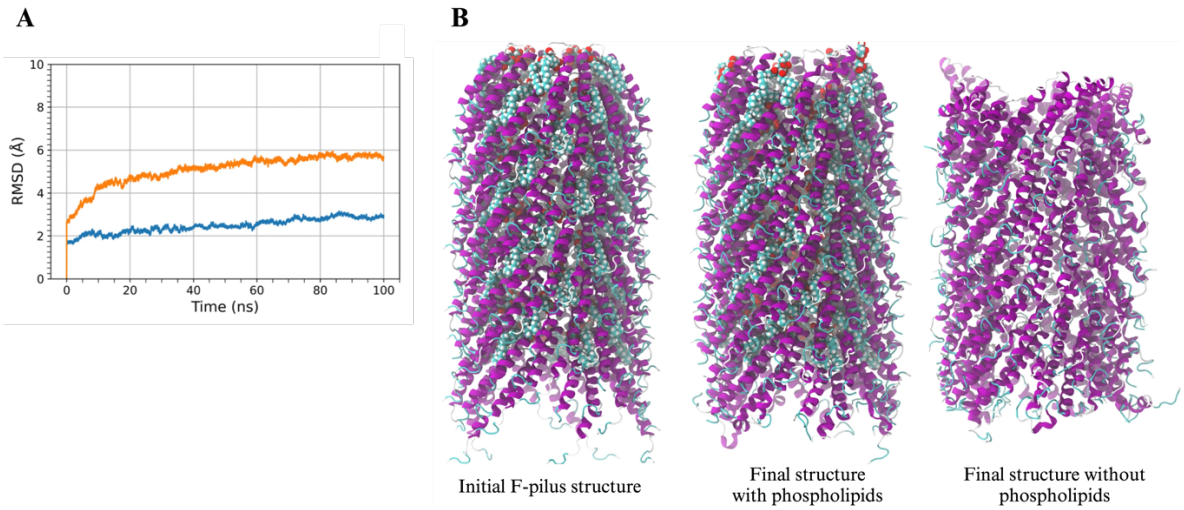


Figure 42. Equilibrium simulations of the F-pilus

A| Equilibrium simulation of the F-pilus with (blue curve) and without (orange curve) phospholipids as the root mean square deviation of the alpha carbon coordinates of the final structure were calculated with respect to the initial filament structure. Alpha carbons belonging to the five subunits at the base and the five subunits at the tip of the filament are excluded. **B|** From the left: the initial F-pilus structure including phospholipid; structure of the F-pilus with phospholipids after 100 ns equilibrium simulation; structure of the F-pilus without phospholipids after 100 ns equilibrium simulation.

3.4 Discussion

The most prevalent plasmid family found in human and animal isolates is based on the IncF replicon (Rozwandowicz *et al.* 2018). These plasmids encode the highly efficient conjugative F-T4SS that have evolved to facilitate successful DNA transfer even in challenging environments, such as those present in human and animal guts. The aim of this study was to explore the structural adaptations of the F-pilus and their implications for maintaining stable mating pairs between donor and recipient bacteria, crucial for the efficient propagation of drug-resistant plasmids and the formation of biofilms in challenging environments.

To simulate the conditions experienced by bacteria in the human or animal host, two *E. coli* strains (laboratory strain DH5α and EHEC strain EDL933) harbouring the IncF plasmid pOX38 were subjected to a range of hydrodynamic and mechanical environments, and their ability to conjugate and form biofilm communities under these conditions was assessed. The results revealed that the interbacterial connections established by the F-pilus are extremely flexible (Figure 38) and can withstand highly turbulent conditions (Figure 36&37), which adds to the previous findings that the F-plasmids have the unique ability to conjugate in liquid environments (Bradley 1980). This suggests that bacteria leverage their well-adapted F-pili to exploit environmental flow forces to scan larger volumes of their surroundings, resulting in a higher chance of finding a mating partner.

The attempts to disrupt mating connections and abolish conjugation through extreme mechanical forces (vortexing) were unsuccessful (Figure 36), indicating that once the F-pilus from the donor cell contacts the surface of the recipient cell, a highly stable connection is established, locking the two cells together. This strong interbacterial attachment mediated by the F-pilus provides an ecological advantage beyond gene transfer, serving as the early ignition point for aggregation of bacterial cells, leading to F-pilus-mediated biofilm formation (Figure 37) (Ghigo 2001). Given that most bacterial infections are associated with biofilms, the dual strategy of the F-pilus, that is involvement in horizontal gene transfer and in biofilm formation, presents a clear benefit to bacteria, enhancing genome plasticity and virulence, respectively.

Here, the exceptional biomechanical properties of the F-pilus were characterised, including high flexibility and spring-like properties (Figure 40). These properties allow for establishing robust intercellular connections by the F-pilus despite environmental turbulence, which in turn helps exploit the dynamic flow forces for enhanced conjugation and biofilm formation. The F-pilus resistance to harsh thermochemical conditions was also demonstrated (Figure 39), suggesting adaptations to mediating conjugation in extremophilic bacterial communities.

Moreover, the unique robustness of the F-pilus is specifically attributed to a single detail in its atomic structure - that is the presence of phospholipids in the F-pilus structure (Figure 41&42). This finding documents for the first time the role of phospholipids in the structure of conjugative pili as fundamental in maintaining filament integrity.

Lastly, I propose a model where the inherent flexibility and robustness of the F-pilus presented here allow for exploiting the enhanced mixing of donor and recipient bacteria in dynamic environments, leading to faster formation of mating pairs (Figure 43). When compared to rigid, non-flexible pili, which are ineffective in turbulent conditions and incapable of conjugation in liquid environments, the F-pili have a biomechanical advantage that prevents snapping and allows for conjugation in dynamic environments, resulting in the accelerated rates of conjugation of IncF plasmids to the cells harbouring them, which could explain their ubiquity in human and animal isolates.

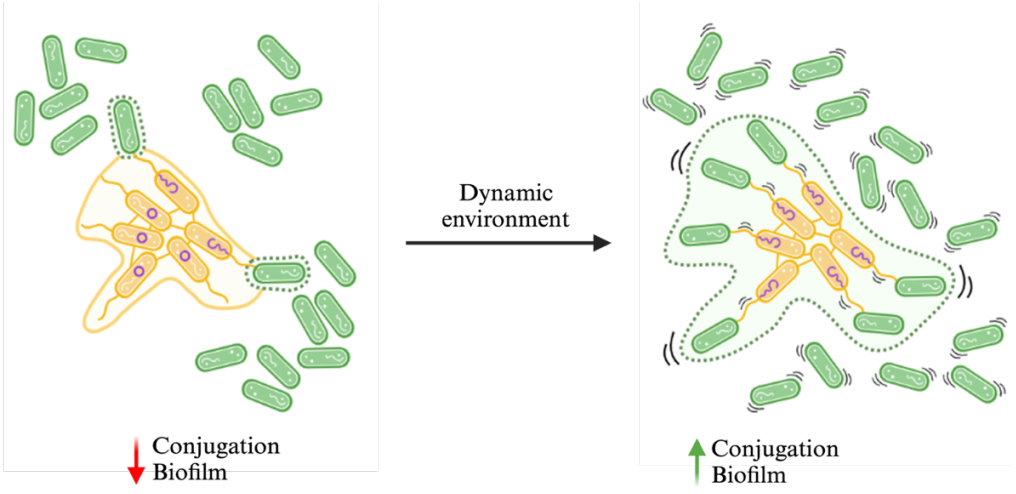


Figure 43. Hydrodynamic model of conjugation and biofilm formation

Left: In conditions without cell being mixed, contacts between donor cells (yellow) and recipient cells (green) are sparce. This is due to uneven distribution of recipient and donor cells, resulting in low biofilm formation and low HGT. Right: Increased hydrodynamic stress and presence of flow forces result in better mixing of donor and recipient populations, stimulating and expanding contacts between cells, leading to an improved conjugation efficiency and biofilm formation.

Chapter 4 Direct evidence of ssDNA transport through the F-pilus

4.1 Introduction

Bacterial conjugation has evolved as a specialised and deliberate movement of MGE between bacteria, becoming the primary contributor to the transfer of ARG between bacteria, thus driving the AMR crisis (Lawrence 1999, Frost *et al.* 2005, Dolejska and Papagiannitsis 2018, Ott and Mellata 2022). It relies on the ability of conjugative T4SS to elaborate long conjugative pili – extracellular polymers that are responsible for creating physical connections between the donor and recipient cell, a process called mating pair formation (Figure 4A) (Cabezón *et al.* 2015, Christie 2016, Brito 2021, Coluzzi *et al.* 2022).

Conjugative transfer of the F-plasmid requires elaboration of the F-pilus, which undergoes cycles of extension and retraction to probe the environment for a mating opportunity (Daehnel *et al.* 2005, Clarke *et al.* 2008). Once such is found, a complete F-pilus retraction is triggered (in conjunction with the recipient cell attached at the distal end of the pilus), which brings the cells close together to initiate the transfer of DNA in a process known as tight transfer (Figure 4C). The membrane-to-membrane contacts are reinforced by a network of interactions between the outer membrane proteins, that is TraN present in the donor cell and OmpA in the recipient cell, a process called the mating pair stabilisation (Figure 4B) (Achtman 1975, Klimke and Frost 1998, Klimke *et al.* 2005, Low *et al.* 2022). This process additionally provides a level of species-specificity as some TraN-OmpA pairs might be incompatible, thereby hindering conjugation between distantly related species (Low *et al.* 2022).

The ability of the F-pilus to retract, together with the existence of mating pair stabilisation strategies resulted in a notion that the establishment of membrane-to-membrane contacts is a requirement for bacterial conjugation, and it became widely accepted that DNA transfer occurs

only when the F-pilus is fully depolymerised (Llosa *et al.* 2002). This model implies that the conjugative journey of DNA occurs solely through the T4SS apparatus, and the F-pilus has a role merely in drawing the cells close together.

However, an alternative model has been considered by Harrington and Rogerson (1990) - who documented successful conjugation events occurring through semipermeable membranes that prevent membrane-to-membrane contacts between donor and recipient cells. This model implies that conjugation can occur at a distance, further called the distant transfer (Figure 4D), and was backed by live-cell microscopy experiments conducted by Babić *et al.* (2008), who showed ssDNA acquisition by recipient cells not in direct contact with any donors. What forms the conduit for the distant transfer of DNA has been a matter of debate. The capacity of the F-pilus to mediate the transfer of DNA through its lumen has been hypothesised, further backed by the first high-resolution structure of the F-pilus, which showed that its lumen diameter and electrostatic charge are suitable to accommodate the travelling ssDNA (Costa *et al.* 2016). So far, distant transfer through the F-pilus has not been directly observed, resulting in many opponents of this model.

This study aims to settle the distant vs tight transfer debate and repeat the experiment conducted by Babić *et al.* (2008), but improve the F-pilus visibility by concomitant fluorescent labelling of the F-pilus, thus directly showing the DNA journey through its lumen. Here, through site-directed mutagenesis of F-pilin subunits – the main "building blocks" of the F-pilus (Figure 10), I generate a functional F-pilus mutant that conforms with fluorescent labelling with Alexa Fluor 568 maleimide, which served as a basis to co-localise the conjugative F-pilus with the simultaneous transfer of conjugative ssDNA, thereby providing the first direct evidence of the F-pilus-mediated distant transfer route of conjugation.

4.2 Experimental procedures

4.2.1 Materials

Table 5. Bacterial strains

Bacterial strain	Information	Source
JE2571	*E. coli* strain. arbours the plasmid pED208 encoding the F-T4SS	(Costa *et al.* 2016)
JE2571 *ΔtraA*	Harbours the plasmid pED208 encoding the F-T4SS with deletion mutation of F-pilin gene, *ΔtraA.*	(Costa *et al.* 2016)
MG1655 *seqA-YFP*	*E. coli strain.* A *Dam*- mutant deficient in DNA methylation. Harbours *seqA-YFP* fusion gene as reporter of hemimethylated DNA.	Egelman Lab
TOP10	Common cloning and expression strain of *E. coli.*	Costa Lab

Table 6. Plasmids and genetic constructs used in this study

Plasmid	Information	Source
pBAD_TraA	pBAD24 cloned with *traA* gene from pED208 plasmid under arabinose-inducible promoter. Amp^R	(Costa *et al.* 2016)
pBAD_TraA:T69C	T69C mutant of pBAD_TraA	This study
pBAD_TraA:T111C	T111C mutant of pBAD_TraA	This study
pBAD_TraA:T112C	T112C mutant of pBAD_TraA	This study
pBAD_TraA:T116C	T116C mutant of pBAD_TraA	This study

Table 7. Primers used in this study

Name	DNA sequence 5' to 3'	Purpose
T69C F	GAAAGCCTGCTTCGGTGCCGACTCATTCG	Generation of pBAD_TraA:T69C
T69C R	CCGAAGCAGGCTTTCACATCATCCTTGCCC	
T116C F	CGGTCTTTGCTTCATCAAATGATAAGAGAGCTTGG	Generation of pBAD_TraA:T116C
T116C R	ATGAAGCAAAGACCGACGGTAGTGAAGAC	

T112C F	CTTCACTTGCGTCGGTCTTACCTTCATCAAATG	Generation of pBAD_TraA:T112C
T112C R	CCGACGCAAGTGAAGACGATAACCACAACCAG	
T111C F	CGTCTTCTGCACCGTCGGTCTTACCTTCATC	Generation of pBAD_TraA:T111C
T111C R	ACGGTGCAGAAGACGATAACCACAACCAGGC	

4.2.2 Generation of F-pili (pED208) for maleimide staining

The native F-pilus structure does not possess surface-exposed thiol groups and hence is not conforming with maleimide staining. It was necessary to artificially increase the number of thiol groups via site-directed mutagenesis. Drawing from the 3D structure of the F-pilus (PDB ID: 5LEG), four threonine residues—T69, T111, T112, and T116—were selected as candidates for substitution with cysteine residues. Consequently, F-pilin (TraA) mutants T69C, T111C, T112C, and T116C were created using site-directed mutagenesis via inverse PCR following the Takara In-Fusion Snap Assembly protocol. The primer pairs T69C F&R, T111C F&R, T112C F&R and T116C F&R were employed with pBAD_TraA construct (Costa *et al.* 2016) as the template to yield the constructs pBAD_TraA:T69C, pBAD_TraA:T111C, pBAD_TraA:T112C, and pBAD_TraA:T116C, respectively.

4.2.3 Complementation of cysteine-derivative of F-pilus

To assess the functional complementation of the generated constructs, a pilus-specific phage f1 spot assay was employed. The plasmid pBAD_TraA encoding the wild-type F-pilin (TraA WT) along with various variants (T69, T111, T112, and T116), was transformed into *E. coli* strain JE2571 containing the pilin-deficient pED208:*ΔtraA* plasmid. The cultures were grown until reaching OD_{600} = 0.5, at which point expression was induced with 0.2% L-arabinose and allowed to proceed for 2 hours at 37°C. Subsequently, the cultures were plated on LB-agar supplemented with 0.2% L-arabinose. Following this, 2 µL of the F-pilus-specific phage f1

was spotted on each of the freshly dispersed bacterial inocula, and the plates were then incubated at 37°C overnight. The next day, an examination of bacteriophage plaques on the plates was conducted. Positive complementation, indicative of the presence of functional F-pilus, was confirmed by the observation of f1 phage plaques. Notably, cells harbouring the pBAD_TraA:T116C variant exhibited plaques identical to those with the wild-type pBAD_TraA. Consequently, this construct was selected for further maleimide labelling and imaging.

4.2.4 Cell growth and Alexa Fluor 568 labelling

The *E. coli* DH5α dam+ strain carrying the pED208:*ΔtraA* plasmid and the complementing pBAD_TraA:T116C, a culture was grown at 37°C and 180 RPM until it reached an OD_{600} of approximately 1.1 in 100 mL of LB media with 50 µg/ml kanamycin and 100 µg/ml ampicillin. The expression of the TraA variant induced by adding 0.2% L-arabinose. Cells were then treated with Alexa-Fluor 568 maleimide (AF568) and incubated for 15 minutes.

For the *E. coli* MG1655 *dam-* strain with the *seqA-YFP* fusion gene (Babić *et al.* 2008), a culture was grown at 37°C and 180 RPM until it reached an OD_{600} of about 0.6 in 100 mL of liquid LB with 25 µg/mL chloramphenicol.

4.2.5 Fluorescence microscopy

A 5 µL of cells containing Alexa-Fluor 568 labelled *E. coli* DH5α (F+) with pBAD_TraA:T116C donor cells mixed with the recipient *E. coli dam- seqA-YFP* strain was applied to a 25 × 75 mm 1.0-mm-thick glass microslide, with or without an agar pad, and covered with a 22 × 22 mm micro cover glass slide. Imaging was performed using a Nikon Ti2-E inverted microscope equipped with an AX-R confocal microscopy system, 25 mm FOV resonant scanner, a 25 mm FOV Galvano scanner, and an NSPARC detector unit. Photomultipliers and the Nikon NIS Elements C software were used to capture images. Live cell imaging took place at 23°C, and the samples were observed using a 100×/NA1.52 oil

objective with perfect focus during live image acquisition. Captured images and movies underwent processing using Elements Denoise.ai, followed by compilation in FIJI software.

4.3 Results

4.3.1 Generation of F-pili for maleimide staining

One way to label extracellular appendages such as the conjugative pili is through a direct fluorophore attachment to the surface of the filament. Fluorescent labelling of type IV pilus has been used in the past for spatiotemporal studies of this appendage (Lam *et al.* 2021). This method relies on labelling the type IV pili with thiol-reactive dyes, which have affinity to surface-exposed thiol groups such as those present in cysteine residues. An example of such a dye is Alexa Fluor 568 C5 maleimide - a bright, water-soluble, orange-fluorescent dye, which is active at neutral pH (~7). Importantly, it does not permeate the bacterial membrane, making it suitable for specifically targeting extracellular appendages.

To label conjugative F-pili, the pED208 plasmid is the ideal candidate as it is a non-repressed type of F-plasmid, resulting in higher number of T4SS and F-pili per cell, facilitating visualisation and downstream spatiotemporal studies. However, the native pED208 encoded F-pilus presents a challenge for labelling with thiol-reactive dyes. This is because the mature F-pilin subunit possesses only a single cysteine residue, which locates within the lumen of the F-pilus (Costa *et al.* 2016), rendering it inaccessible to the dye. To overcome this limitation, I designed a set of cysteine-derivatives of the F-pilus by introducing strategic mutations (T to C substitutions) at four surface-exposed threonine residues: T69C, T111C, T112C, and T116C (Figure 44A).

Each variant was constructed within an arabinose-inducible pBAD expression vector, resulting in independent constructs expressing TraA:T69C, T111C, T112C, and T116C. Additionally, the wild-type pBAD_TraA construct was included as a control. These constructs were then

transformed into *E. coli* harbouring a pED208:*ΔtraA* background (Costa et al., 2016) to assess pilus functionality. To determine which TraA variant supported functional pilus formation, an f1 phage infection assay was employed. It specifically targets the F-pilus as its receptor, hence only a functional F-pilus, that is one that has unaffected structure and can retract, will support f1 phage infection. Any alterations in the pilus function would prevent infection. Therefore, each TraA variant was tested for complementation of f1 phage before and after arabinose induction, the principle of which is described in figure 44B. Only the T116C variant displayed positive complementation, making it the ideal candidate for further live-cell imaging experiments. (Figure 44B).

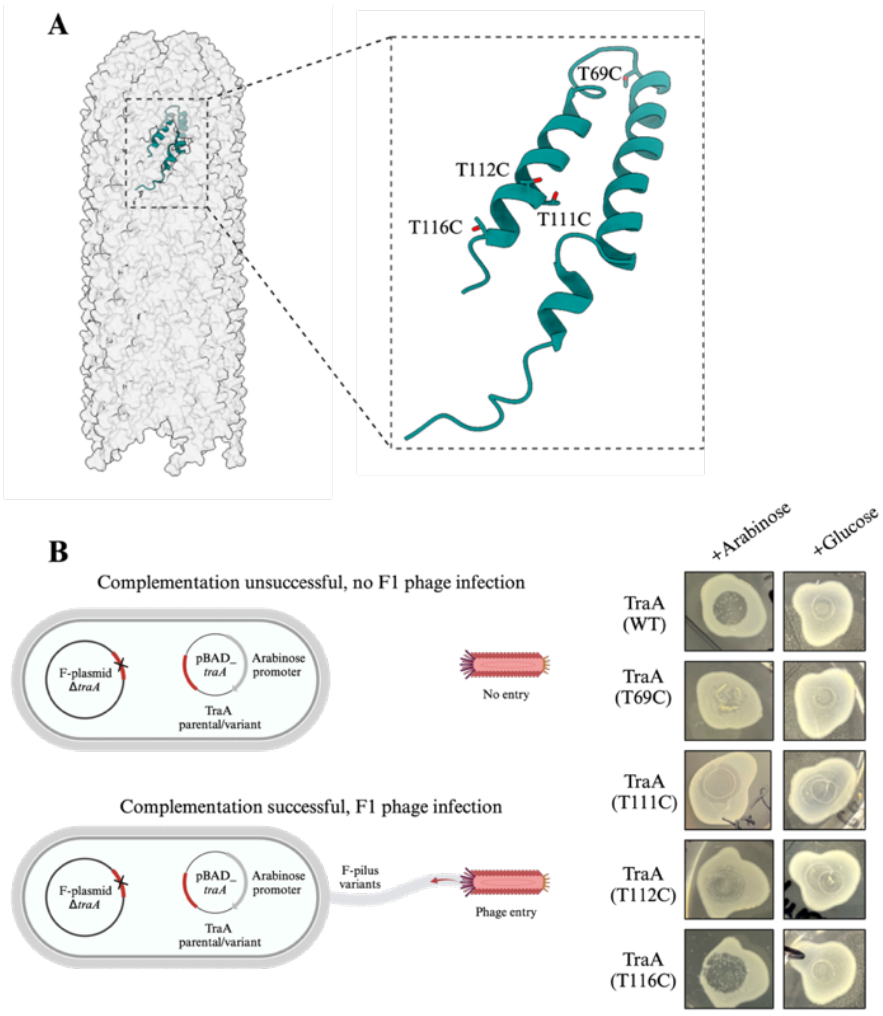


Figure 44. Construction of an F-pilus suitable for maleimide labelling

A| F-pilus (PDB ID: 5LEG) atomic structure, with one F-pilin subunit zoomed in to display potential mutation sites to introduce cysteines. **B|** f1 phage assay was used to indicate formation of a functional F-pilus. Left: Graphical representation of the mutagenesis: a *ΔtraA* mutation of F-plasmid (pED208) was complemented with pBAD_*traA* construct including different mutation variants. If mutation is obstructing F-pilus formation, the f1 phage will not infect the cell (top), but when the mutation is functional, the f1 phage will infect successfully (bottom). Right: plaque formation was observed with different TraA variants upon induction with arabinose and repression of TraA expression with glucose. Clear plaques upon induction indicate a functional mutation.

4.3.2 Co-localisation of F-pili with the travelling ssDNA

Experiments in this section were conducted by Leticia Beltran from the Egelman Lab (University of Virginia).

To co-localise the fluorescent pili together with the travelling ssDNA, the AF568-stained T116C-derivative of F-pilus was combined with the previously documented SeqA-YFP system (Babić *et al.* 2008). The principle of the system is that methylation-deficient recipient cells (MG1655 *dam*-) contain the SeqA-YFP reporter protein, which binds with high affinity to hemimethylated DNA. As the donor is methylation-proficient, the only hemimethylated DNA in the recipient cell would be that of the travelling DNA from the *dam*+ donor to the *dam*- recipient.

Live-cell imaging of the conjugative mixture of *E. coli* DH5α (*dam*+) cells containing the AF568-stained T116C-derivative of F-pilus together with the recipient *E. coli* MG1655 (dam-*seqA-YFP*) revealed multiple intense fluorescent foci where conjugative DNA (green) was transferred through the F-pilus (red) (Figure 45A). Some cells undergoing conjugation were separated even by micrometres. Moreover, the entire process of conjugation was also captured as a movie, frames of which are shown in (Figure 45B).

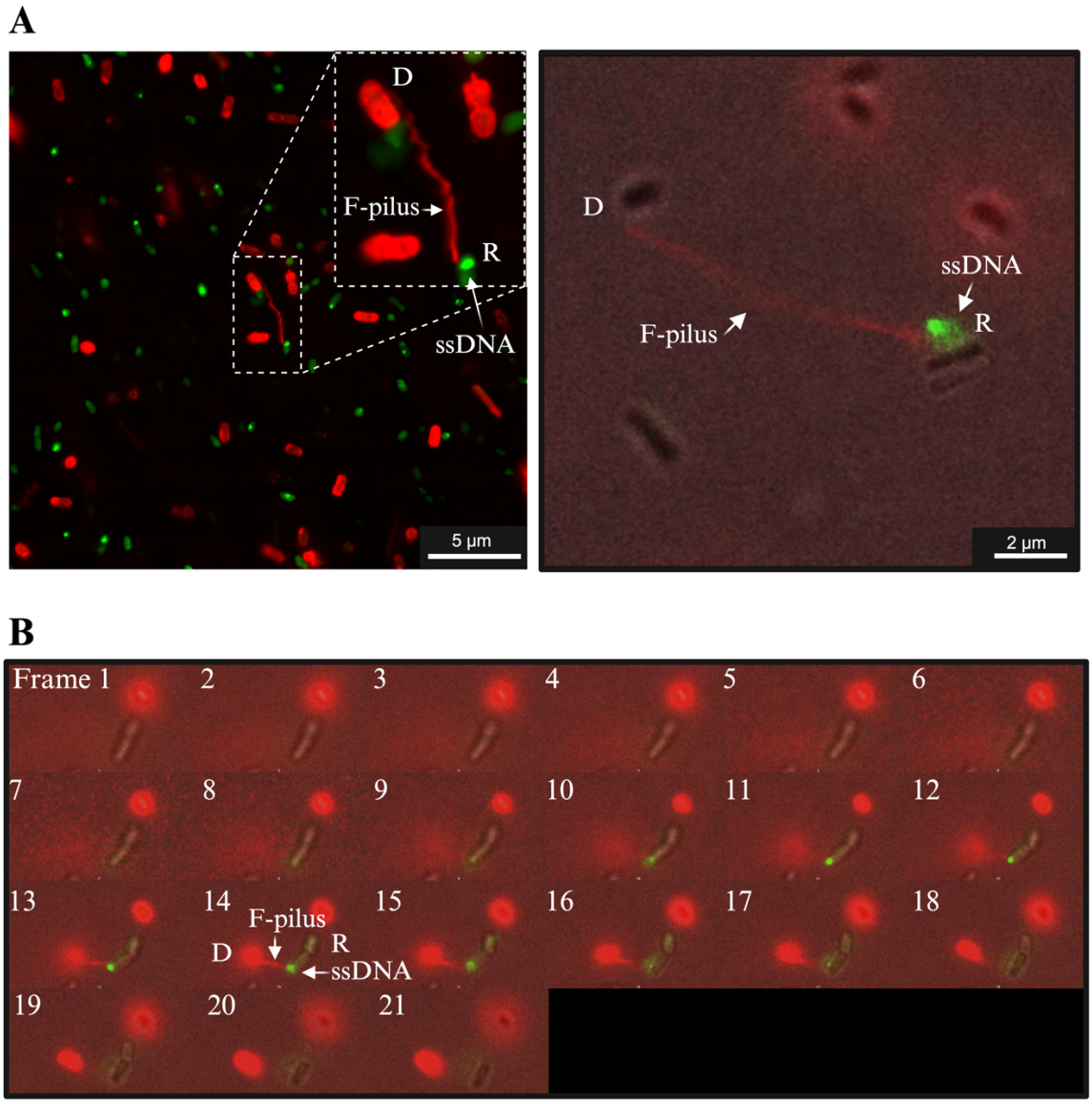


Figure 45. F-pilus mediates distant transfer during conjugation

A| Fluorescence microscopy images of conjugation at different distances. The F-pilus (red) is elaborated by the donor cell (D) and mediates transfer of ssDNA (green) to the recipient cell (R) without both being in direct membrane-to-membrane contact. **B|** Frame-by-frame representation of conjugation at a distance. Frames 10-17 show green puncta in the non-piliated cell (recipient) that display hemi-methylated DNA – the travelling F-plasmid. In the same frames, an elongated red shape stemming from a red-coloured donor cell shows the F-pilus. Frame 14 best displays co-localisation of the F-pilus that forms bridge between bacterial cells and established point of plasmid DNA entry into the recipient cell.

4.4 Discussion

There has been an ongoing debate that polarised the T4SS community into supporters of two distinct routes of bacterial conjugation, one that occurs either through membrane-to-membrane contacts between bacterial cells – the so-called "tight transfer", and the other that occurs at a distance, where DNA travels via the conjugative F-pilus, called the "distant transfer" (Figure 4). Here, I generated a fluorescent F-pilus through the site-directed mutagenesis of the F-pilin subunits (Figure 44), which was subsequently used to provide direct visualisation of the transfer of ssDNA through the F-pilus (Figure 45), thereby concluding that both routes of conjugation are now sufficiently experimentally supported.

In nature, the retraction of the F-pilus and hence the tight transfer route of conjugation might be the preferred route, as it includes the step of mating pair stabilisation which fortifies the interbacterial conjugative connections (Figure 4B). This mode however is only possible in liquid, unrestricted environments, where mating cells can be brought close together without obstruction. On the other hand, in a more densely populated setting, such as in bacterial biofilms, the movement of cells might not be possible, therefore distant transfer of DNA must be triggered. This dual strategy has advantages on multiple levels. Firstly, the ability to conjugate at a distance minimises the overall metabolic cost of bacterial conjugation – that is because in an event where the pilus retraction is not possible, an alternative route can be triggered to mitigate wasting resources that had already been used to generate the conjugative pilus and establish a mating pair. Secondly, the F-pilus ability to retract will always minimise the distance between the mating cells, which ensures the shortest possible path for DNA transfer, thereby lowering the probability of its damage through external stressors. Whether the retraction is finalised at complete depolymerisation of the F-pilus or at partial depolymerisation, will be determined solely by the presence of obstructions on the path of retraction. This dual strategy ensures the shortest path for DNA translocation in sparsely

populated environments, where establishment of a tight connection is possible, but also in densely populated biofilms, where pilus retraction is impossible, and DNA must be transmitted at a distance.

This study reveals the additional roles of the F-pilus during conjugation, serving as a specialised conduit for DNA translocation, where tight conjugative transfer is not possible. This finding complements the documented incredible mechanical and thermochemical resilience of F-pili mentioned in Chapter 3 of this thesis. Now, one could hypothesise that the robustness of the F-pilus filament plays an important role in ensuring the protection of the travelling DNA against environmental stressors during its journey through the pilus during distant transfer, but also fortifies the F-pilus structure to be able to withstand the immense force exerted on the F-pilus during its retraction with the recipient cell attached at the distal end. The combined effect of those strategies has likely contributed to the ubiquity of IncF plasmids in human and animal microbiomes (Rozwandowicz *et al.* 2018).

It is important to acknowledge that shortly before the publication of these results (Beltran *et al.* 2023), a very elegant work by Goldlust *et al.* (2023) was released, also documenting distant conjugation via the F-pilus. They employed a similar approach – by generating a cysteine-derivative, they were able to use Alexa-Fluor 488 maleimide to fluorescently label the F-pilus, and as such combined it with a fluorescent fusion of plasmid-partitioning protein ParB-GFP instead of SeqA-YFP to visualise DNA. Excitingly, their results reach beyond those presented here, additionally documenting the piliation patterns of donor cells and dynamics of pilus extension and retraction. Nevertheless, documentation of such an important mechanism by two independent groups adds to the validity of the conclusions presented here and will undoubtedly help shift the current opinions about distant transfer in the T4SS community.

Chapter 5 Identification of a novel T4SS, essential for AIEC colonisation in L-CD

5.1 Introduction

A peculiar type of chronic inflammatory bowel disease (IBD) was first described by a Polish surgeon Prof. Antoni Leśniowski of the University of Warsaw, Poland, who noticed a high inflammatory activity with omental adhesion in a patient's ileocecum without characteristics of any of the known diseases back in that time. In 1903, he presented his findings to the Warsaw Medical Society and published them in the Polish journals "Medycyna" (Leśniowski 1903, Gibiński 1989) terming the disease *ileitis terminalis.* Due to the occupation of Polish lands by foreign forces at the time, his scientific works could not be translated into other European languages (Eder *et al.* 2019, Van Hootegem and Travis 2020), and therefore it took until 1932 for the disease to be recognised internationally – it was when Dr Burrill B. Crohn presented *ileitis terminalis* to the American Gastroenterological Association and published his findings (Crohn *et al.* 1932). From that point on, the disease is often simply termed Crohn's disease (CD), or Leśniowski-Crohn's (L-CD) disease. The latter will be used in this thesis.

L-CD is characterised by an excessive immune response towards commensal gut microbiota, resulting in tissue injury and recurrent symptoms such as fever, abdominal pains, and diarrhoea (Binder and Orholm 1996). Even in healthy individuals, the human gut is naturally prone to inflammation, displaying the so-called "physiological inflammation", which occurs as a normal response to antigens coming from dietary and microbial sources (Fiocchi 2005). Some individuals, however, develop a much greater immune response to harmless commensals, resulting in unrestrained inflammation, in turn causing L-CD (Eckburg and Relman 2007).

There is increasing evidence that the genetic makeup of the microbiome plays a crucial role in the development of L-CD, however definitive genetic markers of L-CD-associated bacterial

species remain unknown (Lloyd-Price *et al.* 2019). Gut metagenome studies show that L-CD patients display a much poorer microbial diversity, with a concomitant increase in the prevalence of *Enterobacteriaceae,* such as *E. coli* (Darfeuille-Michaud *et al.* 1998). The proliferation of adherent-invasive *E. coli* (AIEC) is particularly enhanced as the likelihood of isolating an AIEC strain from L-CD patients is six times greater than from healthy individuals (Kamali Dolatabadi *et al.* 2021, Mansour *et al.* 2023), with additional evidence showing that AIEC colonisation results in chronic gut inflammation (Palmela *et al.* 2018, Elhenawy *et al.* 2019, Buisson *et al.* 2023). Altogether, this suggests an important role of AIEC colonisation in L-CD pathogenesis.

AIEC harbour a variety of traits that separate them from commensal *E. coli*, including the ability to adhere to intestinal epithelial cells (IEC) (Darfeuille-Michaud *et al.* 1998); the ability to invade IEC via host cell actin polymerisation and microtubule mobilisation (Boudeau *et al.* 1999); and the ability to replicate in macrophages (Glasser *et al.* 2001). Several potential virulence factors have been identified in AIEC, such as type I fimbrial adhesins involved in adhesion to IEC (Boudeau *et al.* 2001, Barnich *et al.* 2007), long polar fimbriae that facilitate interaction with Payer's patches (Chassaing *et al.* 2011, Coppens *et al.* 2015) or mucin-degradation proteases that help with crossing the mucous layer (Gibold *et al.* 2016). Nevertheless, AIEC lack some other prominent virulence factors, such as T3SS, which are known to promote host invasion (Mansour *et al.* 2023).

A recent study has evidenced that a T4SS present in NRG857c AIEC plays a role in the formation of bacterial biofilms and the persistence of this strain in the mouse gut (Elhenawy *et al.* 2021). This T4SS is encoded on a conjugative plasmid pO83 which contains a *tra* operon almost identical to that encoded on the F-plasmid (over 80% nucleotide sequence identity between the shared genes). Outside of the shared genes, however, the pO83 T4SS *tra* operon contains a severe deletion when compared to the *tra* operon of the canonical F-plasmid

(Elhenawy *et al.* 2021). This makes the pO83-T4SS unusual as it lacks some important genes that encode components otherwise determined as crucial for T4SS assembly and operation, such as VirB7-like TraV, an essential protein for assembly of the outer membrane core complex (OMCC) or the VirB4-like ATPase TraC, the largest component of the inner membrane complex (IMC) (Figures 7&8) (Hu *et al.* 2019a, Amin *et al.* 2021). Nevertheless, Elhenawy *et al.* (2021) showed that a knockout mutation of *traA* – the gene encoding the F-pilin subunit (Figure 10) - is detrimental to the persistence of NRG857c AIEC strain in the mouse gut and abrogates biofilm formation of this strain.

The unusual organisation of the pO83 *tra* operon prompted my interest in the pO83-T4SS, which despite being devoid of core structural components comprising the OMCC and the IMC, it somehow still functions in the strain virulence. Therefore, the aim of this study is to understand how the pO83-T4SS overcomes the lack of these subunits.

My first hypothesis is that the pO83-T4SS forms a highly unusual apparatus, capable of elaborating F-pili, which in turn play a role in the formation of biofilms and thus explaining the importance of the *traA* gene in the persistence of NRG857c in the mouse gut. This is however highly unlikely since it is well established in the literature that both VirB7-like TraV and the VirB4-like ATPase TraC are core genes needed for the establishment of the T4SS and hence pilus formation (Bragagnolo *et al.* 2020, Costa *et al.* 2021, Sheedlo *et al.* 2022, Costa *et al.* 2023). Therefore, my second hypothesis is that there is another yet unidentified T4SS in the genome of NRG857 which can provide its OMCC and IMC *in trans* and aid in the establishment of biofilms, and hence enhance the persistence of the strain. Such chimeric T4SS have been artificially generated in the past (Gordon *et al.* 2017, Kishida *et al.* 2024) and shown to be functional, suggesting that they can occur in nature.

Here, I show that the previously described pO83-T4SS does not form a typical OMCC capable of accommodating a growing pilus, hence rendering this system non-functional on its own.

However, I report the presence of a previously undetected T4SS in the NRG857c strain, which is encoded chromosomally within an integrative and conjugative element (ICE). I furthermore classify this ICE-T4SS to the group of G-like T4SS and I establish that it forms an OMCC which structurally resembles the archetypal OMCC structures, suggesting the formation of a functional ICE-T4SS. Lastly, I show that this T4SS is essential for the colonisation of human epithelial cells by the NRG857c AIEC. Altogether, my findings suggest the formation of a chimeric T4SS, made up of the interacting incomplete pO83-T4SS and ICE-T4SS, which is essential for pathogenesis of this strain. These findings underscore the importance of T4SS in L-CD pathogenesis as a virulence factor, and also document the possibility of a first naturally occurring chimeric T4SS.

5.2 Experimental Procedures

5.2.1 Materials

Table 8. Bacterial strains

Strain	Information	Source
JE2571	*E. coli* strain. Harbours the plasmid pED208 encoding the F-T4SS	(Costa *et al.* 2016)
Top10	Common cloning and expression strain of *E. coli*	Costa Lab
BL21 (DE3)	Common expression strain of *E. coli*	Costa Lab
NRG857c	AIEC *E. coli* harbouring pO83 plasmid that encodes a T4SS	(Elhenawy *et al.* 2021)
NRG857c*Δtfc16*	NRG857c AIEC *E. coli* with *Δtfc16* mutation.	Elhenawy Lab

Table 9. Plasmids and genetic constructs used in this study

Plasmid	Information	Source
pED208	Derepressed IncF plasmid	(Costa *et al.* 2016)
pBAD24	Plasmid used for protein overexpression. Contains an arabinose-inducible promoter. Amp^R	Costa Lab
pASK	Plasmid used for protein overexpression. Contains a tetracycline-inducible promoter. Amp^R	Ilangovan Lab

pASK_pO83-OMCC	Plasmid used for overexpression of the OMCC (*traK* and *traBH* with a 10xhis tag on *traBH)* encoded on the pO83 plasmid of NRG857c under tetracycline-inducible promoter. Amp^R	This study
pBAD_ICE-OMCC	Plasmid used for overexpression of the OMCC (*tfc13, tfc14* and *tfc15,* with a Strep-tag II on *tfc14) encoded* on the ICE of NRG857c under arabinose-inducible promoter. Amp^R	This study

Table 10. Primers used in this study

Name	DNA sequence 5' to 3'	Purpose
pBAD F	TGAGTACCCGGGGATCCTCT	pBAD backbone amplification
pBAD R	GGTGAATTCCTCCTGCTAGCC	
pASK F	GCTTGACCTGTGAAGTGAAAAATGGC	pASK backbone amplification
pASK R	CATTTTTTGCCCTCGTTATCTAGATTTTTGTCG	
pO83_TraK-Bhis F	CTTCACAGGTCAAGCTCAGTGGTGGTGGTGATGGTGATGATGAT GATGCCGCGTGCTCCCTCCGAAG	Amplification of the DNA stretch between *traK* and *traBH* from pO83-T4SS locus, introduction of a 10xhis tag before the STOP codon of *traBH*.
pO83_TraK-Bhis R	CGAGGGCAAAAAATGATGAGAAAAAATAATACGGCAATAATATT CGGCAG	
NRG1_TraK-Bst F	CAGGAGGAATTCACCATGAGTTTACGACGATTTTTACTGTCGC	Amplification of the DNA stretch between *tfc13* and *tfc14* from the ICE-T4SS locus, introduction of a Strep Tag II before the STOP codon of *tfc14*.
NRG1_TraK-Bst R	CCTTTCATTTACTTTTCGAACTGCGGGTGGCTCCAATCCAGCTCG TGTTGCTGGCTGACATC	
NRG2_TraV F	TGGAGCCACCCGCAGTTCGAAAAGTAAATGAAAGGATACGGTTT ATTCATGCTGG	Amplification of the DNA of *tfc15* from the ICE T4SS locus.
NRG2_TraV R	ATCCCCGGGTACTCATCAGAGTGCCTCCAGCC	

5.2.2 Genomic search for T4SS-associated loci in NRG857c

Both, the chromosomal (NCBI ID: NC_017634.1) and the pO83 plasmid (NCBI ID: NC_017659.1) nucleotide sequences associated with NRG857c strain were subjected to identification of T4SS locus using SecReT4 database (Bi *et al.* 2013). Genes with high identity to known T4SS operons have been provided, and based on that, new labelling of the genes has been proposed.

5.2.3 Classification of the genomic T4SS locus

To classify both T4SS loci of NRG857c, the amino acid sequence of T4CP from the pO83-encoded T4SS (NCBI ID: WP_000009324.1) and T4CP from the ICE-encoded T4SS (NCBI ID :WP_001283927.1) were aligned to amino acid sequences of coupling proteins of T4SS representative of different incompatibility groups, that is: St_pED (NCBI ID: WP_032490373.1); Kp_pCT (NCBI ID: WP_059274178.1); Ec_pOX38 (NCBI ID: WP_000009350.1); Ng_pMS11 (NCBI ID: AAW83057.1); St_pR27 (NCBI ID: WP_000167420.1); Ec_pYD626E (NCBI ID: ADL14203.1); Ec_pKM101 (NCBI ID: AAB58711.1); Ec_R388 (NCBI ID: CAA44852.1); At_Ti (NCBI ID: P09817.1); Ec_pRP4 (NCBI ID: WP_011205817.1); Lp_Dot/Icm (NCBI ID: AAC38193.1); St_pR64 (NCBI ID: WP_001289276.1) using Clustal Omega multiple sequence alignment (Sievers *et al.* 2011). The root tree was generated and visualised in NCBI Tree Viewer (https://www.ncbi.nlm.nih.gov/tools/treeviewer/).

5.2.4 Protein structure prediction of the OMCC components

Structure predictions of TraBH and TraK and of Tfc13, Tfc14 and Tfc15 were performed by AlphaFold-2 via the ColabFold platform (Mirdita *et al.* 2022).

5.2.5 Generation of pASK_pO83-OMCC and pBAD_ICE-OMCC

To generate pASK_pO83-OMCC, the DNA of *traK* and *traBH* from the pO83 plasmid along with the vector pASK, were amplified through polymerase chain reaction (PCR) with primers detailed in Table 10. with a C-terminal 10x poly-histidine tag (His-tag) encoding sequence added before the stop codon of *traBH*. Post-PCR, Dpn1 digestion was applied to eliminate template DNA from the reaction products. Gel electrophoresis of the PCR products was executed on a 0.9% agarose gel supplemented with SYBR safe DNA stain (Thermo-Fisher) to validate accurate amplification. The samples were electrophoresed in TAE (Tris-acetate-EDTA) buffer for 35 minutes at 90 V and visualised under UV light. Subsequently, PCR clean-

up was performed using a QIAquick PCR purification kit (QIAGEN) to eliminate impurities. The pASK and *traK-BH* fragments were ligated through in-fusion cloning In-Fusion HD Cloning Kit User Manual (TakaraBio). For transformation, competent TOP10 cells transformed with the pASK-TraK-TraBHHis construct via a one-minute heat shock at 42°C, followed by a 1-hour incubation in 1 L of LB broth at 37°C. Subsequently, the cells were streaked onto agar plates containing the ampicillin derivative carbenicillin (100 μg/mL) and left to incubate overnight at 37°C. A single successfully transformed colony was chosen and cultured in LB with carbenicillin (100 μg/mL) overnight at 37°C and the construct DNA was purified with MiniPrep Kit (QIAGEN), and sequence to confirm correct cloning.

To generate pBAD_ICE-OMCC, the DNA of *tfc13, tfc14* and *tfc15* along with the pBAD backbone were amplified using primers as in Table 10, with a C-terminal Strep-tag II (amino acid sequence: WSHPQFEK). The following procedures were the same as outlined above for pBAD_pO83-OMCC.

5.2.6 Purification of the pO83- and ICE-encoded OMCC

Purification of both constructs followed a similar protocol with minor changes where indicated. A single TOP10 colony harbouring either the pBAD_pO83-OMCC or pBAD_ICE-OMCC was inoculated into LB supplemented with ampicillin (100 μg/mL) at 37°C overnight. On the next day, 40 ml of the overnight culture was transferred to 8L of LB with ampicillin (100 μg/ml) and grown at 37°C until the cultures reached an OD_{600} of 0.6-0.8. The expression of pO83-OMCC was induced by addition of anhydrotetracycline (200 μg/l), whereas expression of ICE-OMCC was induced by addition of 0.2% L-arabinose. The expression was maintained overnight at 18°C. The induced cells were then harvested through centrifugation at 7,000 rpm for 7 minutes. The cell pellet was resuspended in 50 mM HEPES (pH 7.6), 0.3 M NaCl, 1 mM EDTA, and a protease inhibitor tablet (Roche). Subsequently, cell lysis was achieved through double cell disruption at 27 kpsi. To eliminate cell debris, the lysed mixture underwent

centrifugation at 35,000 g for 30 minutes at 4°C. The resulting membrane fraction was isolated by subjecting the supernatant to ultracentrifugation at 120,000 g for 1 hour at 4°C. The membrane pellet was then solubilised for 1h at 4°C in 50 mM HEPES (pH 7.6), 0.3 M NaCl 0.8% w/v lauryl dimethylamine oxide (LDAO) and 1% w/v n-dodecyl-B-D-maltoside (DDM). Following this, the sample was once again subjected to ultracentrifugation at 120,000 g for 30 minutes, and the resulting pellet was discarded to eliminate insolubilised membranes.

For purification of pO83-OMCC, the supernatant containing solubilised was applied to TALON cobalt column (Cytiva) and subjected to washing with 50 mM HEPES buffer (pH 7.6) containing 0.3 M NaCl, 10 mM imidazole, and 0.23% (w/v) LDAO, the proteins bound to the column were then eluted using a linear imidazole gradient ranging from 10 mM to 500 mM over 20 ml at a flow rate of 1 ml/min. Elution fractions were collected at 0.5 ml intervals. The elution fractions of pO83-OMCC obtained from affinity chromatography were pooled and concentrated to 500 µl using a 10 kDa MWCO protein concentrator (Merck). Subsequently, the membrane protein sample underwent separation through a Superose 6 Increase 10/300 GL column (Cytiva) in a 50 mM HEPES buffer (pH 7.6) with 0.3 M NaCl, 10 mM imidazole, and 0.23% (w/v) LDAO. Elution fractions were collected at 0.5 ml intervals and visualised using SDS-PAGE.

For purification of ICE-OMCC, the supernatant containing solubilised was applied to Strep Trap XT column (Cytiva) or and subjected to washing with 50 mM HEPES buffer (pH 7.6) containing 0.3 M NaCl and 0.23% (w/v) LDAO, the proteins bound to the column were then eluted using the same buffer supplemented with 10mM desthiobiotin at a flow rate of 1 ml/min. Elution fractions were collected at 0.5 ml intervals and visualised using SDS-PAGE.

SDS-PAGE was conducted using Novex WedgeWell SDS-polyacrylamide gels (10-20%) from Thermofisher. SDS-PAGE loading buffer (Thermo-Fisher) was added, and the samples were heat-treated at 95°C for 3 minutes. Gel electrophoresis was conducted at 200 V for 50 minutes

in SDS-PAGE running buffer (Tris, glycine, SDS), and visualisation was accomplished through overnight staining with Coomassie blue stain (Thermo-Fisher).

5.2.7 Western Blotting

The elution fractions obtained from IMAC were subjected to SDS-PAGE, following established protocols. The presence of TraBH-6xHis was determined through Western blot analysis. A methanol-soaked polyvinylidene fluoride (PVDF) transfer membrane (Merck) was utilised for Western transfer at 200V for 50 minutes in Tris-buffer saline buffer with 10% Tween 20 (TSB-T). To prevent non-specific antibody binding, the membrane was incubated in TSB-T with 5% (w/v) skimmed milk powder for 1 hour. After washing with TSB-T, the membrane was incubated for 1 hour with anti-His antibodies, followed by horseradish peroxidase (HRP)-linked anti-mouse antibodies. After thorough washing with TSB-T, the membrane underwent a 5-minute incubation with SuperSignal chemiluminescent substrate (Thermo-fisher) and was visualised by chemiluminescence using a ChemiDoc XRS+ System (Bio-rad) with an exposure time of 90 seconds.

5.2.8 NS-EM imaging

10 µl of elution fractions of purified OMCC were individually loaded onto glow-discarded 300 mesh carbon-coated copper grids (Agar Scientific). Excess sample was blotted off with filter paper, and the grids were then left at room temperature for 1 minute. Each loaded grid underwent three washes with 10 µl of water, was blotted dry, stained with 10 µl of 2% (w/v) uranyl acetate, left for 1 minute, and finally imaged using a Technai12 Spirit transmission electron microscope (FEI) at 120 kV with a 2 K Eagle camera (FEI).

5.2.9 Imaging of NRG857c infection on Caco-2 cells

Immunostaining of infected cells and tissues involved cultivating Caco-2 cells in Dulbecco's Modified Eagle's Medium (DMEM, Gibco) supplemented with 20% foetal bovine serum (Gibco), 1× minimum non-essential amino acids (Gibco), 10 units/ml penicillin, 10 µg/ml

streptomycin (Gibco), and 1 mM HEPES (Gibco). The cells, grown to approximately 90% confluency on poly-L-lysine-coated coverslips, were infected with either wild-type or the ICE-T4SS:*Δtfc16* mutant of NRG857c strains at a multiplicity of infection (MOI) of 25:1 for 1.5 hours. After removing non-adherent bacteria and changing the media to Iscove's Modified Dulbecco's Medium (Gibco), a subsequent 6-hour incubation took place, followed by removal of the media, washing, and fixation with 4% paraformaldehyde (Sigma-Aldrich) for 20 minutes at room temperature. Coverslips were blocked overnight at 4°C with 3% goat serum (Invitrogen) and 1% bovine serum albumin (Sigma-Aldrich). AIEC staining was accomplished by incubating the samples with a primary antibody anti-O83 (1:300 dilution, Statens Serum Institute, CAT: 85077). Mucin staining was achieved by incubating slides with wheat-germ agglutinin (1:200 dilution, 1% goat serum, 0.1% BSA, Invitrogen, CAT: W11263) for 1 hour at 37°C. Subsequent steps included permeabilisation, incubation with rabbit anti-O83 (1:300 dilution), and mouse anti-β actin (1:500 dilution in 1% goat serum, 1% Tween 20, 0.1% BSA, Abcam, CAT: ab8226) for 40 minutes at room temperature. After washing, slides were incubated with anti-rabbit Alexa-Fluor 488 and anti-mouse Alexa-Fluor 568 (both at 1:200 dilution, Invitrogen) as secondary antibodies for 1 hour at room temperature. Microscopy analysis was conducted on a Zeiss Axio Imager with 20× and 40× objectives, utilising a Hamamatsu Digital camera ORCA-R2.

5.3 Results

5.3.1 The pO83-T4SS does not form a typical OMCC

The pO83 plasmid identified in the NRG857c strain belongs to the IncF family and shows over 80% nucleotide sequence identity between the shared genes to the *tra* operon of the F-plasmid (pOX38). Strikingly, the pO83 *tra* operon exhibits an extensive deletion of a region between *traB* and *traH* genes, resulting in a loss of approximately 20 genes when compared to the canonical F-plasmid operon. To visualise this, an Easyfig representation of the sequence

alignment together with corresponding gene tags is shown (Figure 46). The "lost" genes include important structural genes of T4SS, such as the *virB7*-like *traV,* essential for the establishment of an OMCC or *traC*, essential for the establishment of IMC. Moreover, the gene encoding the *virB10-like traB*, another OMCC component, is a truncated fusion of the canonical *traB* and *traH* genes of the F-plasmid (Figure 46), hence it will be referred to as *traBH.*

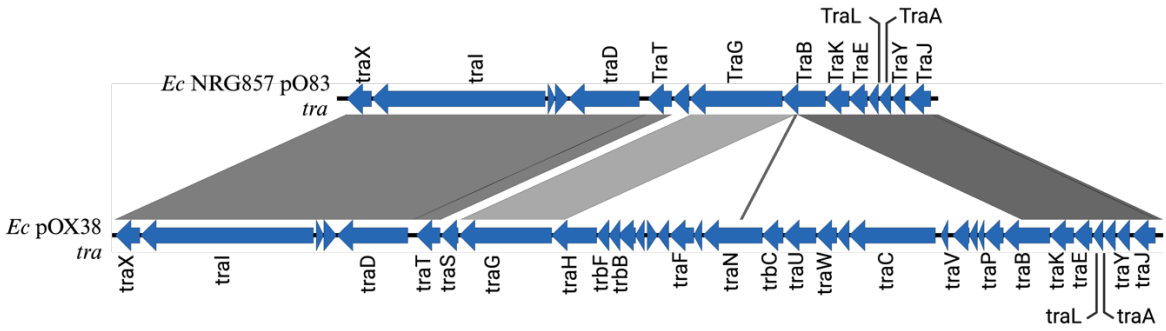


Figure 46. Homology between pO83-T4SS and the F-T4SS

The *tra* operon of pO83-T4SS exhibits high homology to that of F-plasmid (pOX38), sharing over 80% sequence identity between the shared components. A large segment between genes *traB* and *traH* genes is absent in the pO83 *tra* operon when compared to the F-plasmid encoded operon, leading to the lack of important genes, such as *traC* or *traV*, which upon expression are fundamental components for the assembly of the OMCC, and the IMC, respectively. Additionally, genes encoding *traB* and *traH* were fused, generating a *traBH* gene.

As the OMCC is the largest subcomplex and serves as a platform during the T4SS assembly upon which the remainder of the T4SS will be constructed (Khara *et al.* 2021), the presence of an OMCC is a good indicator of whether a T4SS is formed or not, hence the ability of pO83-T4SS to form an OMCC was assessed by a combination of size-exclusion chromatography (SEC) and negative-stain electron microscopy (NS-EM).

Briefly, *traK* and *traBH* genes were cloned into a pASK vector with a C-terminal his-tag on *traBH*. Proteins were then expressed in BL21 *E. coli*, membranes were extracted and solubilised in buffers containing 1% LDAO and 0.8% DDM – a detergent mixture previously

employed for extracting F-plasmid encoded OMCC from the membranes (Amin et al., 2021). The TraK-TraBH complex was purified via affinity chromatography using a TALON column with affinity to the his-tag on TraBH, yielding a pure fraction of TraK-TraBH as assessed by SDS-PAGE (Figure 47). The resulting fractions were applied on Superose 6 Increase 10/300 GL SEC column (Cytiva) to determine the size of the formed complex. Based on the elution volume of the SEC fractions, the TraK-TraBH complex eluted as two peaks (P1 and P2), both at a volume corresponding to a size between 440 and 66 kDa (Figure 47), much smaller than the canonical multi-megadalton OMCC.

To assess the morphology of the purified complex, NS-EM imaging was employed of fractions corresponding to both peaks. Images revealed the presence of a small, circular complex with a diameter of less than 10 nm (Figure 48), much smaller compared to a typical OMCC (> 20 nm in diameter (Sheedlo *et al.* 2022, Costa *et al.* 2023). The size of the formed complex is insufficient to accommodate the growing F-pilus because an OMCC lumen of at least 9 nm is needed, as this is the size of the F-pilus diameter (Figure 11) (Costa *et al.* 2016). Therefore, this suggests that a complete OMCC is not formed by the pO83-T4SS.

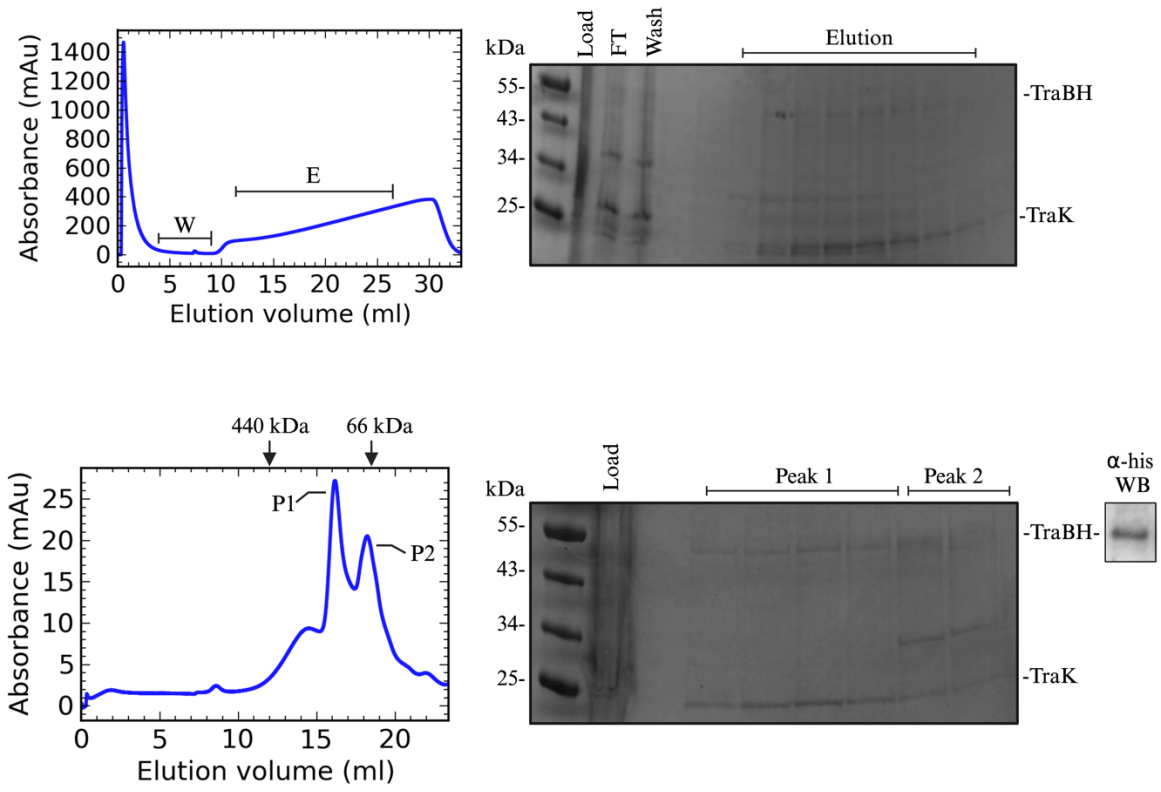


Figure 47. pO83-OMCC does not assemble large rings

Top: The resulting chromatogram of TraK-TraBH purification alongside the corresponding SDS-PAGE gel showing a pure fraction of TraK-TraBH complex. Bottom: The resulting chromatogram and corresponding SDS-PAGE gel of size-exclusion chromatography. The estimated size of the resulting particles based on the elution volume is between 440 and 66 kDa. Western blot confirms the presence of TraBH. Calculated theoretical molecular weights: TraBH - 47.5 kDa, TraK - 25.5 kDa. FT – flow-through; E – elution; W – wash, P1 – peak 1; P2 – peak 2, WB – Western Blot.

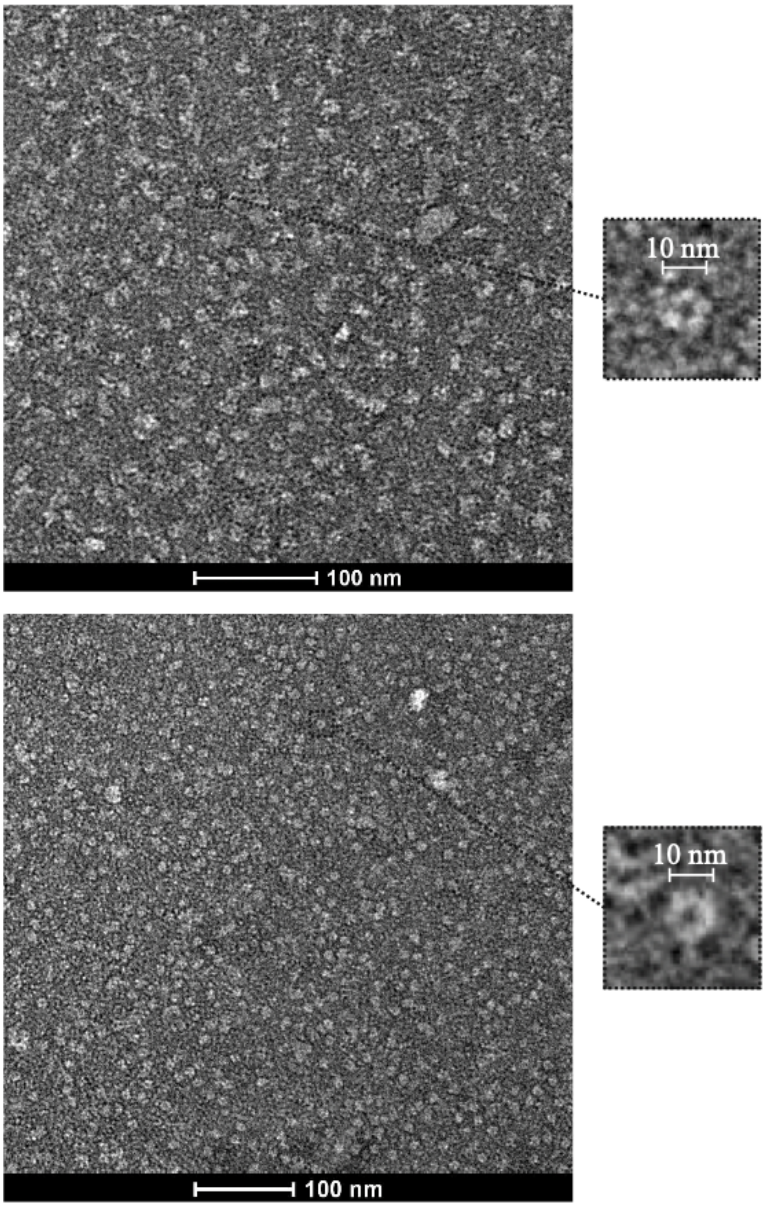


Figure 48. NS-EM imaging of the pO83-OMCC SEC fractions via NS-EM

Top: fraction corresponding to P1 SEC peak. Bottom: fraction corresponding to P2 SEC peak. Both micrographs show small rings, with a diameter of approximately 10 nm, which is much smaller than typical OMCC (25 nm) and insufficient to accommodate a growing F-pilus which requires at least 9 nm of the OMCC lumen diameter, based on its dimensions.

5.3.2 Discovery of a novel T4SS in NRG857c AIEC

Due to the absence of a large, ring-shaped OMCC assembled by the pO83-T4SS, I hypothesised the presence of an alternative T4SS within NRG857c, which could potentially provide its own OMCC *in trans*. VRprofile2 analysis failed to detect a T4SS locus in the NRG857 chromosome

(Supplementary Figure 2) (Wang *et al.* 2022). However, the SecReT4 (Bi *et al.* 2013) database was employed to interrogate the NRG857c genome for another T4SS.

Initially, a homologue of the *virB4* gene was identified on the chromosome (nucleotide coordinates 4447552-4449636), suggesting a potential presence of a T4SS operon. The subsequent analysis of the genes in the immediate vicinity of this gene revealed a full T4SS operon exhibiting amino acid homology to the components of the ICE-encoded *tfc* operon from a hospital-acquired multidrug-resistant strain of *K. pneumoniae* HS11286, isolated in Shanghai (Liu *et al.* 2012), which functions in the propagation of genomic islands through bacterial conjugation (Juhas *et al.* 2007)

The resulting analysis of the novel putative T4SS operon in the NRG857c chromosome has been summarised in Table 11, denoting the coordinates of the genes identified as potentially encoding a T4SS; providing NCBI reference sequences to the protein of highest homology, together with % identity; proposing a new locus tag (based on homology to the *tfc* operon of the ICE-encoded T4SS in *K. pneumoniae* HS11286) and indicating a putative function, based on the function of the protein of highest homology.

Accompanying Table 11 is an Easyfig representation of the nucleotide sequence homology between both *tfc* operons: those of the newly discovered T4SS operon of NRG857c and that of *K. pneumoniae* HS11286, representing an overall 20% sequence identity (Figure 49A). The new proposed loci tags of the NRG857c operon were included in the figure. Analysis of the novel *tfc* operon of NRG857c revealed the presence of genes encoding common T4SS components (Table 11, Figure 49A). Notably, it encodes homologues of both proteins absent in the pO83-T4SS, that is the VirB7/TraV-homologue Tfc15 and VirB4/TraC-homologue, Tfc16. Furthermore, a broad range of proteins exclusive to *tfc* systems were identified (Figure 49), with their functions largely unknown. Nevertheless, all three components of a putative OMCC were identified: VirB9/TraK-like Tfc13, VirB10/TraB-like Tfc14, and VirB7/TraV-like

Tfc15 (Table 11, Figure 49A). Interestingly, proteins homologous to type IV pilus assembly subunits, such as PilL-like Tfc2 and PilT-like Tfc22, were also found, a feature characteristic to *tfc* systems.

Table 11. Identification of a chromosomal T4SS in AIEC strain NRG857c

Original locus tag	Locus coordinates	Protein of highest homology	Identity to protein of highest homology [%]	New proposed locus tag	Putative function (based on homology)
Unlabelled	4444376 4444972 [+], 597	WP 001048687.1	68.24	*tfc2*	PilL like protein, type IV pilus assembly
Unlabelled	4444969 4445688 [+], 720	WP 001005792.1	35.29	*tfc2*	PilL like protein, type IV pilus assembly
Unlabelled	4445699 4446403 [+], 705	WP 000734461.1	60.71	*tfc3*	Unknown
Unlabelled	4446385 4446951 [+], 567	WP 000741281.1	71.51	*tfc4*	TrbI-like transglycosylase domain-containing protein
Unlabelled	4446948 4447448 [+], 501	WP 000980269.1	57.58	*tfc5*	Unknown
Unlabelled	4447552..4449636 [+], 2085	WP 001283927.1	83.87	*tfc6*	VirD4-like protein, type IV coupling protein
Unlabelled	4449929.. 4450693 [+], 765	WP 000081046.1	58.27	*tfc8*	Unknown
Unlabelled	4452171..4452482 [+], 312	WP 001242997.1	59.8	*tfc9*	Unknown
Unlabelled	4452744..4453115 [+], 372	WP 000458747.1	64.89	*tfc10*	Unknown
Unlabelled	4453125..4453505 [+], 381	WP 000190871.1	65.05	*tfc11*	Unknown
Unlabelled	4453498..4454154 [+], 657	WP 000032826.1	69.59	*tfc12*	Unknown
Unlabelled	4454154..4455029 [+], 876	WP 000058700.1	63.97	*tfc13*	VirB9/TraK-like component of OMCC
Unlabelled	4455019..4456485 [+], 1467	WP 001166082.1	63.89	*tfc14*	VirB10/TraB-like component of OMCC
Unlabelled	4456495..4456896 [+], 402	WP 000679470.1	61.72	*tfc15*	VirB7/TraV-like component of OMCC
Unlabelled	4456893..4459664 [+], 2772	WP 162740568.1	77.79	*tfc16*	VirB4/TraC-like ATPase, component of the IMC
Unlabelled	4459661..4460059 [+], 399	WP 000984473.1	38	*tfc17*	Unknown
Unlabelled	4462013..4462411 [+], 399	WP 000672509.1	50.75	*tfc24*	Unknown
Unlabelled	4462408..4463361 [+], 954	WP 000977625.1	57.75	*tfc23*	Unknown
Unlabelled	4463376..4464791 [+], 1416	WP 000153601.1	66.59	*tfc22*	PilT-like protein, type IV pilus assembly
Unlabelled	4465125..4466654 [+], 1530	WP 000123102.1	54.88	*tfc19*	TraG-like protein, component of the IMC

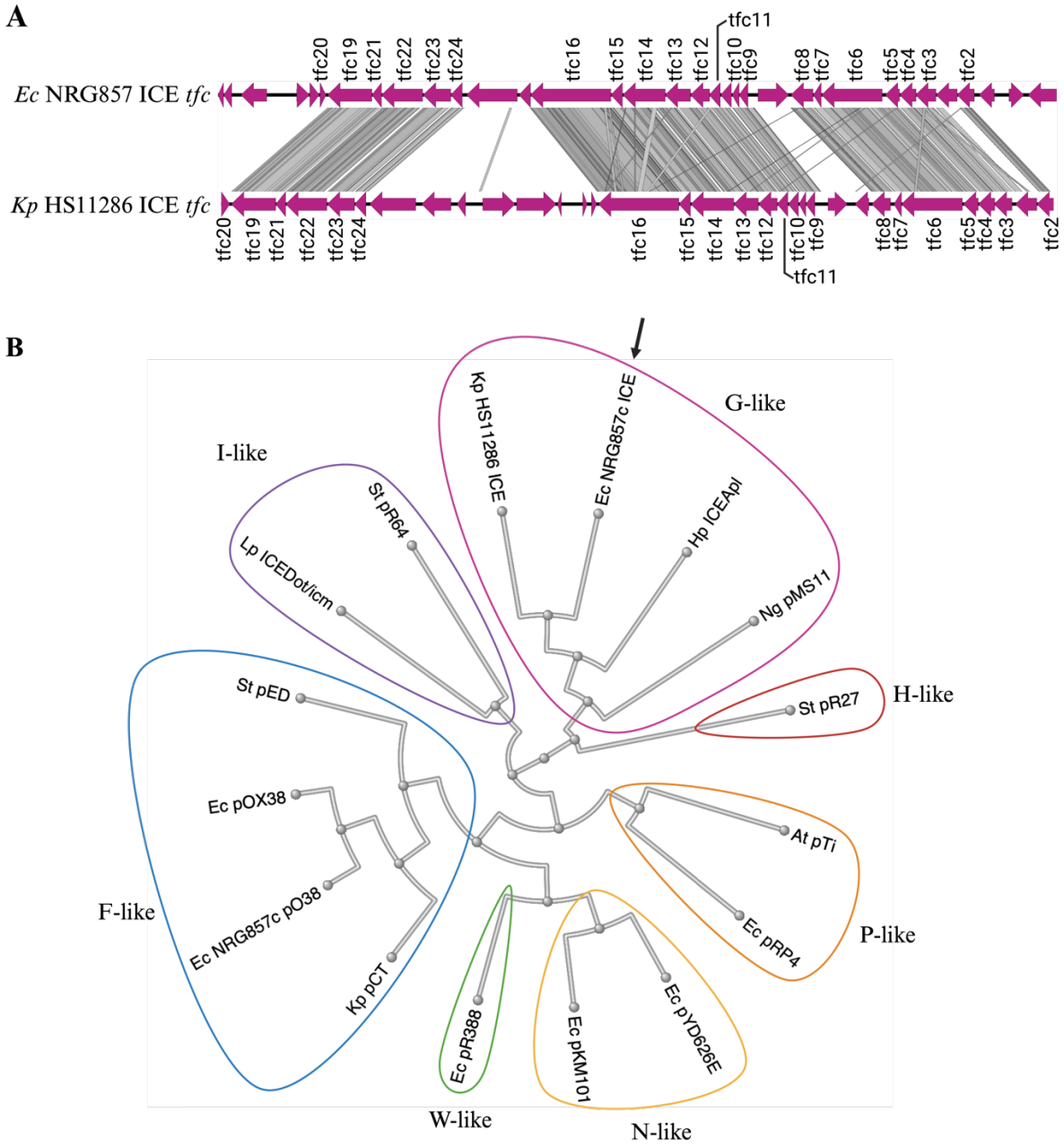


Figure 49. Organisation and classification of ICE-T4SS

A| Easyfig alignment showing the nucleotide sequence similarity between the *tfc* operons of the putative ICE-T4SS of NRG857c strain and the *K. pneumoniae* HS11286 ICE, with an overall 20% sequence identity. Despite some differences between the operons, all core genes for the assembly of T4SS are present in the NRG857c ICE *tra* operon, together with a broad range of *tfc*-exclusive genes. **B|** To classify the putative ICE-T4SS into broader incompatibility groups, the sequences of coupling proteins from both T4SS were aligned to sequences of coupling proteins from T4SS representative of different incompatibility groups – including H-, P-, N-, W-, F-, I-, and G-like T4SS. The pO83-T4SS of NRG857c is homologous to conjugative F-like systems, whereas the chromosomal ICE-T4SS of

NRG857c is homologous to G-like systems (black arrow). Ec – *E. coli*, Kp – *K. pneumoniae*, St – *S. typhimurium*, Lp – *L. pneumophila*, Hp – *H. pylori*, Ng – *N. gonorrhoeae*, At – *A. tumefaciens*.

T4SS can be commonly classified based on the incompatibility of the plasmids/ICE encoding them, forming H-, P-, N-, W-, F-, I-, and G-like T4SS (Lawley *et al.* 2003), with the expanded T4SS primarily falling into the F, I and G groups. To classify the putative NRG857c ICE-T4SS, amino acid alignment analysis was performed. Relationships within and between T4SS clusters are typically established based on the sequences of conserved proteins, such as the ATPases, including the VirB4-like and VirD4-like ATPases. Here, Clustal Omega amino acid alignments of VirD4-like representatives from each incompatibility group were plotted on a phylogenetic tree (Figure 49B) (Sievers *et al.* 2011). The ICE-T4SS of NRG857c clustered together with G-like systems, which comprise a group of largely understudied T4SS, characteristically encoded within genomic islands. Despite these systems being present in a range of bacterial species, including *L. pneumophila*, *S. enterica*, or *Pseudomonas fluorescens* (Bi et al. 2013), their function remains unknown, however some have been shown to successfully excise from the chromosomes and mobilise, hence function as conjugative systems (Flannery *et al.* 2011).

5.3.3 Comparison of OMCC subunits between pO83-T4SS and ICE-T4SS

Thanks to AlphaFold-2, structures of proteins can now be predicted with high confidence based only on amino acid sequence of any protein, making it a powerful tool in analysing putative proteins and assessing their similarity to known homologues. Here, I use AlphaFold-2 predictions to firstly, confirm that Tfc13, Tfc14 and Tfc15 of the putative ICE-T4SS indeed exhibit folds characteristic to OMCC as is expected by the gene homology (Table 11) and secondly, to examine similarities and differences between three OMCC, that is the chromosomally encoded ICE-OMCC, plasmid-encoded pO83-OMCC – both of which are found in NRG857c; together with the canonical F-OMCC encoded on the F-plasmid. Predictions of all subunits of those OMCC were generated and analysed, namely the ICE-

OMCC subunits: Tfc14 (homologue of VirB10), Tfc13 (VirB9), and Tfc15 (VirB7); pO83-OMCC subunits: TraBH (VirB10) and TraK (VirB9); and the canonical F-OMCC subunits: TraB (VirB10), TraK (VirB9), and TraV (VirB7).

Similar topology can be observed among all the shared subunits, owing to the high conservation of OMCC subunits across all, even distantly related T4SS (Figure 8), however some differences are evident. Firstly, between the VirB10-like subunits, TraBH of the pO83-OMCC stands out as significantly shorter than TraB or Tfc14 (Figure 50), leading to a reduced β-barrel size in TraBH. In TraB and Tfc14, the α_2-helix folds back to form the large β-barrel domain (Figure 50), however, as the C-terminus of TraBH is formed by a fusion with TraH, its α_2-helix fails to turn back into the β-barrel. Interestingly, Tfc14 features an additional β-hairpin structure between its α_1 and α_2 helices, absent in TraB or TraBH (Figure 50). Secondly, between the VirB9-like subunits, a similar topology is conserved, with the only significant difference being noticeable in the Tfc14 subunit, which displays an elongated loop between the two β-barrels (Figure 50). Lastly, the VirB7-like subunits comprise an N-terminal α-helix domain and a C-terminal β-hairpin domain, which is conserved between TraV and Tfc15. Striking however is the lack of a VirB7-like subunit encoded in the pO83 *tra* operon.

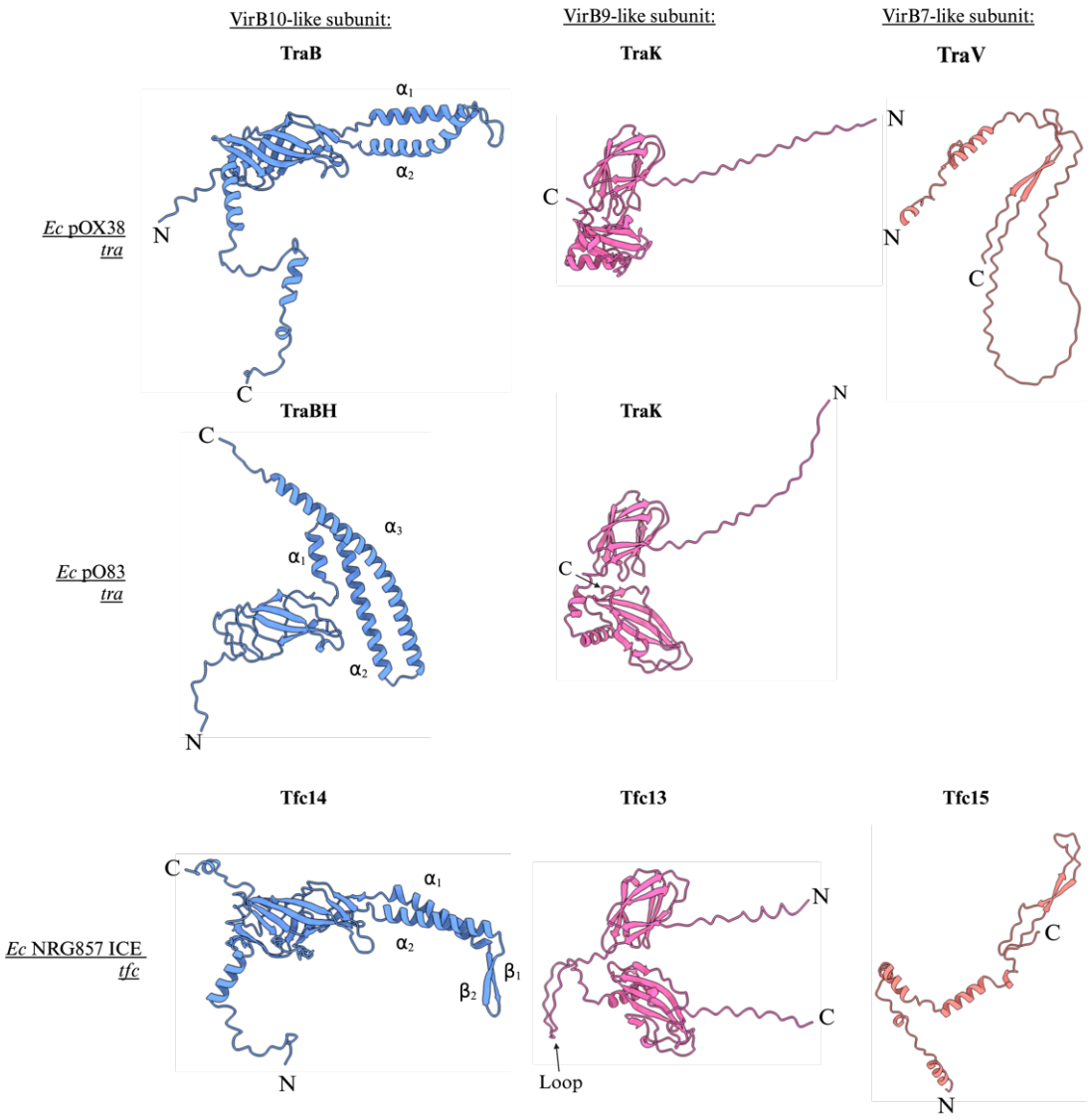


Figure 50. Comparison of OMCC between ICE-T4SS, pO83-T4SS and F-T4SS

The components of an OMCC consist of three key elements: VirB10-like, VirB9-like, and VirB7-like proteins. Presented here are the AlphaFold-2 predictions of the molecular structures of these OMCC components from three systems: the canonical F-plasmid (pOX38) encoded T4SS, the T4SS encoded on the pO83 plasmid, and the newly discovered putative ICE-T4SS. Structures present overall similar topology, owing to the high conservation of these subunits among T4SS. However, some differences are evident. Within the pO83-T4SS striking is the absence of TraV and the smaller β-barrel of its TraBH, lacking a turn after the α_2 helix back into the β-barrel. In the putative ICE-T4SS, Tfc14 forms a β-hairpin between the α_1 and α_2 helices, a feature not observed in the canonical TraB subunit, whereas its Tfc13 presents an extensive loop between its two β-barrels, absent in the TraK subunits of both pO83 and pOX38.

In summary, the predictions presented here suggest that indeed the three subunits comprising the putative ICE-OMCC of NRG857c exhibit folds that resemble those of the F-T4SS, indicating the formation of a typical ICE-OMCC. Secondly, the predictions also make it clear that the pO83-OMCC stands out as the most altered OMCC in terms of its subunits, with a much smaller TraBH that assumes a non-canonical topology attributed to being a fusion protein and lacking a TraV homologue. These properties are likely the reason behind the previously described inability of this system to assemble a large OMCC (Figures 47&48).

5.3.4 The ICE-T4SS forms a typical OMCC

Since the AlphaFold-2 prediction results show that Tfc13, Tfc14 and Tfc15 assume a fold characteristic of OMCC, whether they assemble a large OMCC *in vitro* will be tested here. Therefore, the genes *tfc13*, *tfc14* and *tfc15* were cloned on a pASK vector with a C-terminal strep-tag II sequence on *tfc14*. Following expression in *E. coli* BL21, resuspended cell pellets were lysed, clarified by low-speed centrifugation, and the membrane fraction was collected by ultracentrifugation. Membranes were solubilised with the addition of 0.8% LDAO and 1% DDM and applied on the Strep Trap XT column (Cytiva). The column was washed with the buffer containing 0.23% LDAO, and proteins were eluted with the same buffer containing 10 mM desthiobiotin. Such purified fraction was then observed on NS-EM.

Following the affinity purification, the SDS-PAGE analysis revealed the presence of all three purified OMCC subunits, that is Tfc13, Tfc14 and Tfc15 in the elution fraction. This signifies that all three interact and form a complex that was maintained throughout purification (Figure 51A). The subsequent NS-EM analysis unveils large ring-like structures with approximately 24 nm in diameter (Figure 51B). The size and the shape of the complex are canonical to T4SS (Sheedlo *et al.* 2022, Costa *et al.* 2023), which ultimately confirms the presence of an OMCC and is a confident indicator of the presence of a functional ICE-T4SS in the chromosome of NRG857c.

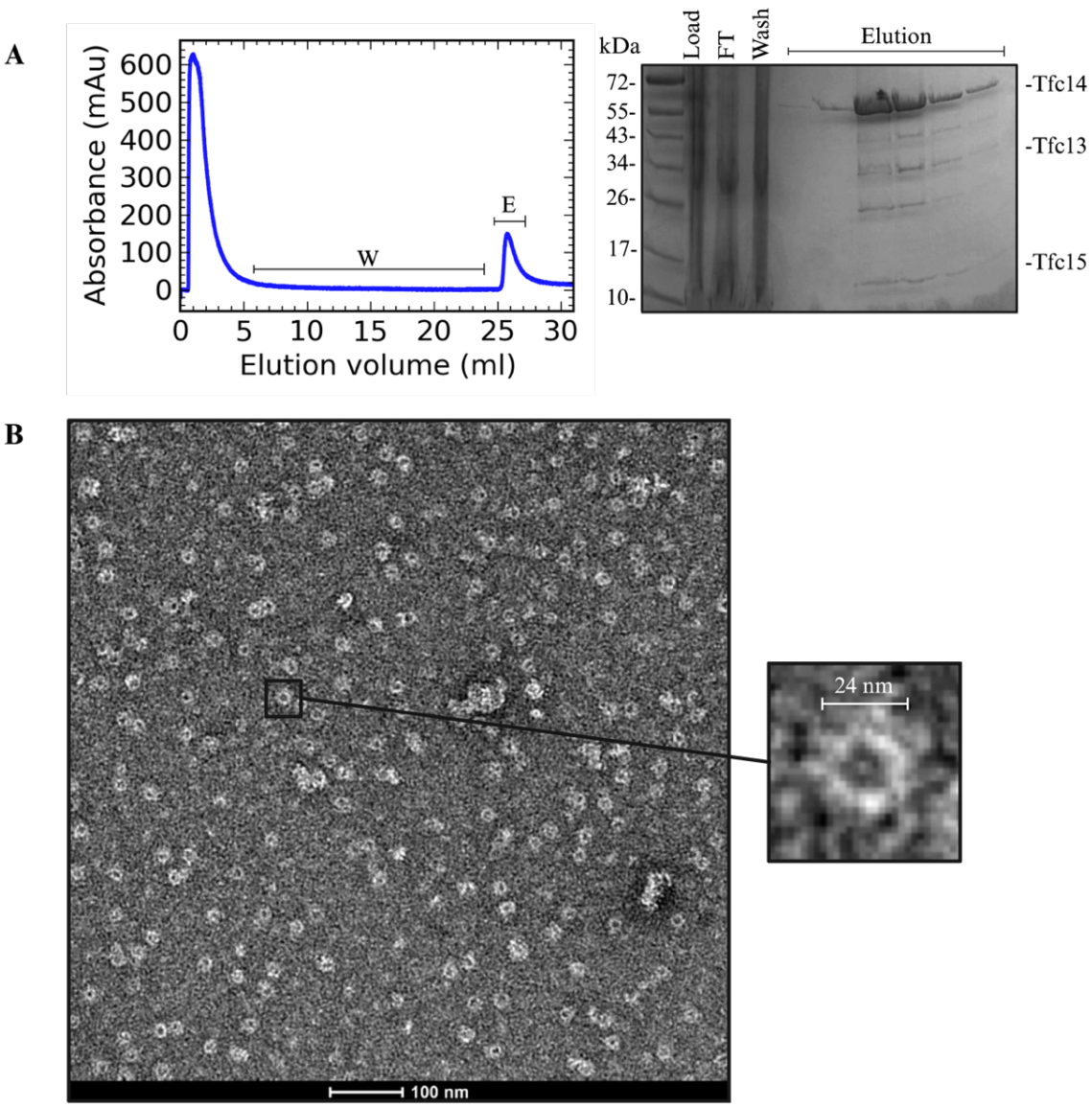


Figure 51. The ICE-OMCC forms typical large ring structures

A| Resulting strep-affinity chromatogram of Tfc13-Tfc14-Tfc15 purification with a corresponding SDS-PAGE gel showing a purified fraction of the complex. Calculated theoretical molecular weight of Tfc13 - 31,5 kDa, of Tfc14 - 51 kDa, of Tfc15 - 15kDa. **B|** NS-EM imaging of the elution fraction via NS-EM shows the assembly of large rings of 24 nm in diameter – a shape characteristic for OMCC across different T4SS. FT – flow-through; E- elution; W-wash

5.3.5 The novel ICE-T4SS is essential for biofilm formation

Experiments in this section were conducted by Wael Elhenawy (University of Alberta), following my discovery of the novel ICE-T4SS locus and request to examine its functional importance to the NRG857c strain.

To understand the significance of the ICE-T4SS to its host strain NRG857c, an ICE-T4SS:Δ*tfc16* mutant was generated. Δ*tfc16* is the deletion of the essential VirB4-like ATPase, central to the establishment of the IMC of T4SS, the removal of which renders a T4SS non-functional (Kerr and Christie 2010, Macé *et al.* 2022). The ability of this deletion mutant to form biofilms on Caco-2 cells – human epithelial cells isolated from colon tissue, was tested, and compared to wild-type NRG857c cells by fluorescence microscopy.

Briefly, Caco-2 cells were grown to 90% confluency on poly-L-lysine-coated coverslips, and infected with either wild-type NRG857c or the ICE-T4SS:Δ*tfc16* mutant at MOI of 25:1. After 1.5h incubation, non-adherent cells were washed off, and cells were stained. For NRG857c staining, specific anti-O83 antibodies were used, whereas for Caco-2 staining, anti-β-actin and mucin staining with wheat-germ agglutinin were used. After incubation and washing off the excess antibodies, coverslips were incubated with fluorescently labelled secondary antibodies and imaged using light microscopy.

The wild-type NRG857c strain was proficient in forming biofilms on Caco-2 cells, as exhibited by extensive clusters of green puncta on Caco-2 cells (Figure 52). However, the ICE-T4SS:*Δtfc16* mutation was detrimental to the strain biofilm formation ability on epithelial cells, as no large clusters of NRG857c could be observed, suggesting that the ICE-T4SS functions in the establishment of biofilms by NRG857c on epithelial cells. This ability is a key virulence factor of AIEC strains (Martinez-Medina *et al.* 2009, Prudent *et al.* 2021, Mansour *et al.* 2023).

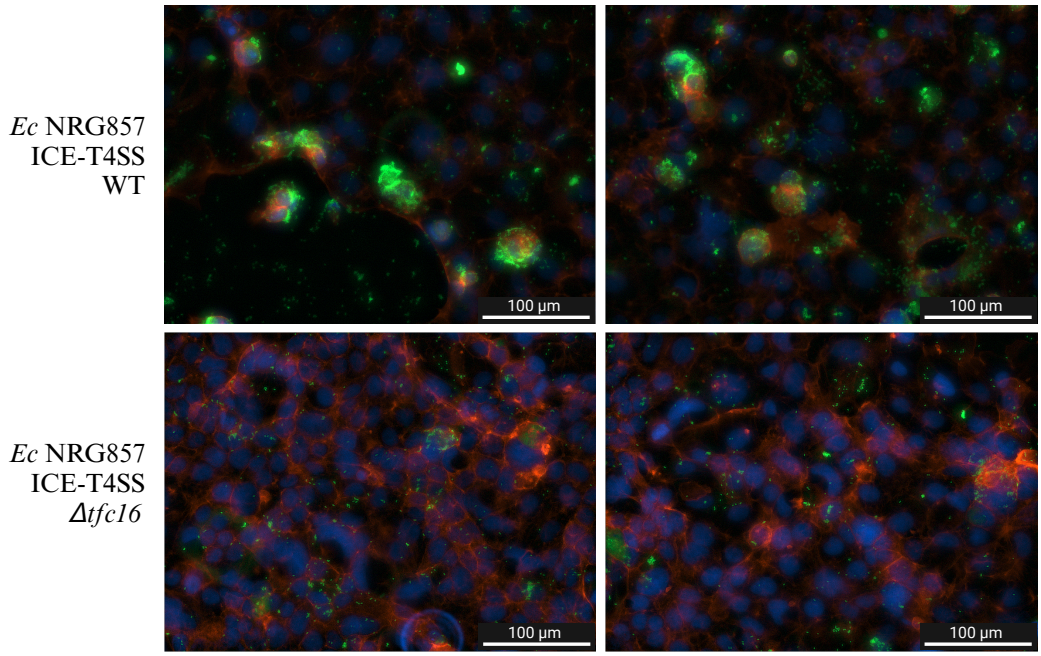


Figure 52. The ICE-T4SS is involved in biofilm formation on Caco-2 cells

The wild type (WT) NRG857c bacteria co-cultured with epithelial Caco-2 cells (mucin in blue, cytoskeleton in red) exhibit the formation of proficient biofilms to colonise the cells (green patches). Deletion of the ICE-T4SS *tfc16* however is detrimental to this process, resulting in lower rate of biofilm formation displayed by more sparsely located bacteria (green puncta).

5.4 Discussion

The hallmark of L-CD is the decline in bacterial diversity within the gut microbiome, accompanied by an increase in the share of *E. coli* strains, particularly AIEC. It is now evident that the ability to form biofilms is crucial for AIEC persistence in the human gut (Martinez-Medina *et al.* 2009, Elhenawy *et al.* 2021, Prudent *et al.* 2021, Mansour *et al.* 2023). Certain genetic determinants have been linked to biofilm formation by AIEC, for instance, the LF82 AIEC strain in L-CD patients relies on the σ^E stress response for biofilm formation on epithelial cells (Chassaing and Darfeuille-Michaud 2013), and more recently, the involvement of T4SS has been also documented (Elhenawy *et al.* 2021).

Conventionally, T4SS are associated with bacterial conjugation, however, they comprise a ubiquitous and diverse family of secretion systems in bacteria, and some have been shown to facilitate biofilm formation through the involvement of conjugative pili in the establishment of interbacterial interactions (Ghigo 2001). Therefore, the acquisition of a T4SS is a beneficial virulence strategy of AIEC as confirmed by the clear enrichment of genes encoding T4SS in microbial metagenomes of L-CD patients (Elhenawy *et al.* 2021).

I focused on the pO83-T4SS, which was recently found in NRG857c AIEC strain to be involved in biofilm formation on epithelial cells (Elhenawy *et al.* 2021). It has a highly unusual organisation of the *tra* operon encoding it, which contains severe deletions, resulting in the loss of many genes encoding otherwise important structural subunits for T4SS assembly and pilus biogenesis, such as the OMCC subunit TraV and IMC subunit TraC (Figure 46). Despite that, the system surprisingly maintains its ability to form F-pilus-mediated biofilms as evidenced by significantly hindered infection persistence and biofilm formation of this strain upon knockout mutation of the gene *traA*, which encodes the F-pilin subunits (Elhenawy *et al.* 2021). Based on this finding, I formed two hypotheses that could explain how the pO83-T4SS can support the formation of biofilms despite lacking key genes. Firstly, even without TraV, the pO83-T4SS can still form an OMCC capable of elaborating F-pili without the need for the otherwise essential TraC, or secondly, there is another T4SS locus in the genome of NRG857c that would provide its OMCC and a TraC homologue *in trans*.

Here, the evidence suggests that the pO83-T4SS does not form the canonical large, multi-megadalton OMCC, attributed to the absence of TraV-encoding gene in its operon and the inability of TraBH to form large oligomers (Figures 47&48). The observed TraK-TraBH complex presents a molecular mass lower than 440 kDa based on SEC results (Figure 47) and, upon imaging, revealed a ring-shaped topology of 10 nm in diameter (Figure 48) - too narrow to accommodate a growing F-pilus, which requires approximately 9 nm of OMCC lumen

diameter, based on its dimensions (Figure 11). This evidence, together with the lack of the largest IMC subunit, the ATPase TraC, is sufficient to reject the first hypothesis.

It is possible that the pO83-T4SS relies on another T4SS within the NRG857c genome, forming a chimeric T4SS that would explain the observed importance of TraA in biofilm formation of the strain, despite lacking a cognate OMCC (Elhenawy *et al.* 2021). Such chimeric T4SS were artificially generated in the past twice, in 2017 and 2024, and in both cases proved to be functional (Gordon *et al.* 2017, Kishida *et al.* 2024), suggesting that such chimeric T4SS can exist in nature, however, they have never been documented.

Here, I identified a novel T4SS in NRG857c encoded on the *tfc* operon (Table 11, Figure 49A) and classified it as an ICE-T4SS belonging to the G-like family of T4SS, closely related to the ICE-encoded T4SS of the multidrug-resistant strain *K. pneumoniae* HS11286 (Figure 49). G-like T4SS are a relatively young and largely understudied group of T4SS, however they have been shown to be involved in bacterial conjugation (Juhas *et al.* 2007). The analysis of AlphaFold-2 predictions of proteins encoded on the *tfc* operon allowed for the identification of all structural components of the ICE-OMCC of NRG857c (Figure 50), which were then purified and visualised by NS-EM, displaying the formation of rings typical of T4SS, with a diameter of 24 nm (Figure 51), and bearing resemblance to the OMCC encoded by the F-plasmid (Amin *et al.* 2021). As the OMCC represents the first independent subunit of T4SS that assembles during the lifecycle of a T4SS (Khara *et al.* 2021), its successful formation implies the presence of a functional ICE-T4SS assembled in the membranes of NRG857c. To understand its function, a deletion mutant of the main constituent of the IMC, the VirB4-like ATPase Tfc16, was generated and proved to severely impair the formation of biofilms by NRG857c on Caco-2 cells (Figure 52). In absence of Tfc16, the biofilm formation ability of the strain is significantly hindered (Figure 52) – displaying the same phenotype that was observed for the pO83Δ*traA* deletion mutant (Elhenawy *et al.* 2021). Considering that the

pO83-T4SS does not encode its own VirB4-like homologue and it must rely on Tfc16 for the biogenesis of its F-pilus, the similarity between pO83Δ*traA* and ICEΔ*tfc16* phenotypes suggests that both TraA from the pO83-T4SS and Tfc16 from the ICE-T4SS have a synergistic effect on the formation of biofilm by the strain NRG857c, suggesting the formation of a chimeric T4SS. However, further study is needed to directly show protein-protein interactions between pO83-T4SS and ICE-T4SS and thereby confirm the formation of a chimeric system beyond reasonable doubt. Nevertheless, this phenotype clearly manifests that the novel ICE-T4SS identified here is essential for biofilm formation of the NRG857c strain on human epithelial cells and, hence, highlights the role of T4SS as important virulence factors of AIEC in L-CD pathogenesis.

Chapter 6 Concluding remarks and outlook

6.1 Overview of the main findings

T4SS are incredibly complex and diverse nanomachines that drive the dissemination of ARG by their involvement in bacterial conjugation, therefore causing the worldwide AMR crisis. The scope of this thesis was to understand the structure-function relationships within expanded T4SS, with a focus on the F-plasmid encoded type IV coupling protein, the biomechanics of the conjugative F-pilus, its role during the distant route of conjugation, and finally - the role of T4SS in L-CD pathogenesis.

In Chapter 2, I provide molecular detail into the precise sequence of events required for a successful transfer of F-plasmid DNA by TraD during bacterial conjugation. I present the high-confidence prediction model of TraD, identifying domains responsible for the binding of DNA and hydrolysing ATP, two functionalities that I also confirmed biochemically. Based on that, I propose a model in which ATP hydrolysis is needed for DNA translocation by TraD in a mechanism reminiscent of FtsK, that is via the interaction of positively charged residues of the lumen with the travelling DNA. I furthermore document the ATP-induced oligomerisation of TraD, which perhaps functions as a safety mechanism that disassembles the TraD oligomer once DNA transfer is finished to prevent maintaining an unnecessary TraD pore in the inner membrane. My findings not only provide a more detailed insight into the mechanism of bacterial conjugation, but also establish the foundation for utilising TraD as a drug target against it.

In Chapter 3, I characterise the dynamics of bacterial conjugation, and the precise biomechanical adaptations of the iconic F-pilus that allow for successful conjugation and biofilm formation in highly turbulent environments, such as those present in human guts. I document the flexibility and spring-like properties of the F-pilus, which distinguish the F-pili

from other conjugative pili, that are rigid and incapable of conjugating in dynamic liquid environments. I present the incredible robustness and thermochemical resistance of the F-pilus, which can play a role in the protection of DNA during distant transfer in a variety of environments, including those resided by extremophiles. Moreover, I unveil for the first time the structural relevance of phospholipid molecules in structures of conjugative pili as essential for maintaining the filament robustness. I finish by proposing a model, in which the mixing of the donor-recipient populations by turbulent environments increases the probability of an encounter between them, however, to take advantage of this property, the conjugative pilus must be able to withstand such turbulent conditions without damage. The findings presented here help explain the observed ubiquity of IncF plasmids in human and animal microbiomes and their evolutionary success over other types of conjugative plasmids in those environments.

In Chapter 4, I develop an F-pilus variant that conforms with fluorescent maleimide labelling, which was subsequently used to visualise the DNA passage through the lumen of the F-pilus, thereby providing the first direct evidence of the distant transfer route of bacterial conjugation and the role of the F-pilus as a conduit for DNA passage.

Lastly, in Chapter 5, I identify a novel T4SS in the L-CD-associated AIEC strain NRG857c. This T4SS belongs to the understudied G-like family involved in conjugative propagation of genomic islands. I demonstrate that this novel T4SS functions in biofilm formation on human epithelial cells by the NRG857c strain. These findings highlight the importance of T4SS as a virulence factor in L-CD and might serve as a genetic marker in future diagnostics. Additionally, my research suggests that the coexistence of two T4SS in the NRG857c strain (one being severely mutilated albeit functional), might be the first example of a naturally occurring chimeric T4SS, where one system must take advantage of subunits of the coresiding system as it lacks its own homologues.

Overall, my thesis sheds light on a range of molecular mechanisms within T4SS and their role in bacterial conjugation and pathogenesis. My findings contribute novel and valuable insights into the spread of antibiotic resistance and help pave the way for the development of novel therapeutic strategies to combat AMR.

6.2 Future directions

The T4SS prove to be extremely versatile nanomachines which evolved sophisticated mechanisms to deliver substrate DNA to their neighbouring bacteria or eukaryotic cells in the process of bacterial conjugation. Although there is considerable knowledge about its operation, many questions remain unanswered.

The structural analysis of TraD presented here has opened the doors to documenting its fascinating mechanism of operation, with the most pressing being direct observation of DNA transfer through the lumen of TraD, perhaps by means described previously for the FtsK ATPase (Jean *et al.* 2020). The overarching goal of TraD investigation will be to trap TraD in its "pre-DNA translocation, post-capture" stage, which would be necessary for direct visualisation of the entire T4SS-TraD-relaxosome complex, thereby finally providing a detailed understanding of the molecular mechanism of bacterial conjugation. This would also finally help in understanding how T4CP docks to the T4SS, of which we still lack confident models.

It is also essential to understand how the role of VirB11 is provided by the F-plasmid encoded T4SS, which does not contain an obvious VirB11 homologue. Functionality of VirB11 must be therefore somehow conveyed either by TraD or VirB4-like ATPase TraC. Alternatively, the VirB11 function can be provided by the wide array of periplasmic proteins characteristic to the F-T4SS, which are known to be essential for F-plasmid conjugation, however, their precise functions remain unknown. Curious is also the lack of VirB5 and VirB6 homologues in the F-

T4SS, which in the R388-T4SS were recently found to provide the interface for the nucleation of the conjugative pilus. The AlphaFold-2 prediction structure of TraG shows its two-domain organisation with similarity to VirB5 and VirB6, suggesting it can provide homologous functions. It is plausible that proteolytic cleavage occurs between both domains, allowing for pilus outgrowth in a mechanism similar to that of VirB5-VirB6.

The F-pilus was shown to depend on the presence of phospholipids for its structural robustness, therefore it would be intriguing to see how changes in the type of phospholipids present in the F-pilus would affect the stability and structure of the pilus. This could answer why different conjugative pili, even those harboured by the same bacterial species, incorporate different phospholipids. It would be interesting to see how the presence of different phospholipids alters the overall structure of the filament. Such changes could explain why conjugative pili, despite having nearly identical pilins, differ so much in the global structure of the filament.

Moreover, the structural investigation of the F-pilus is far from finished. We lack the crucial structural information about what resides on the tip of the F-pilus and thus what is the precise mechanism of the F-pilus entry to the recipient. It must be noted, that while the entire T4SS machinery must be assembled to allow the F-pilus to transverse the donor cell membranes and peptidoglycan layer, penetrating the recipient cell would impose similar barriers, however no structure is currently known to aid the penetration of the recipient cell envelope. Other machineries such as the needle-like T6SS or the injection complex of bacteriophages have developed sophisticated strategies that allow for injecting various effectors/DNA into their target cell, none of those were observed in the T4SS. It might be so that the conjugative pili only contacts the surface of outer membrane, but then the question of how DNA makes its way through the periplasm, peptidoglycan and the inner membrane of the recipient cell appears.

Direct visualisation of F-pilus-mediated ssDNA transfer shows that when further retraction is impossible, transfer of DNA can be triggered. It is therefore a perfect foundation for further

investigation into how the F-pilus and its T4SS mechanically sense when to trigger it. Moreover, it is interesting how retraction is orchestrated in general. It seems that before F-pilus contacts the recipient cell, F-pilus extension and retraction occur randomly, however when the F-pilus contacts the recipient cell, only retraction can occur. It is unknown what mechanism senses the presence of the recipient cell and prevents further pilus extension.

The identification of T4SS as a virulence factor L-CD-associated AIEC leaves the question of how prevalent are such G-like T4SS among this pathotype of *E. coli.* Additionally, determining how the T4SS conveys the observed adhesion specifically to epithelial cells will also help us understand novel mechanisms of utilising T4SS in pathogenicity. Moreover, the hypothesis of chimeric T4SS circulating in nature needs to be directly evidenced by an observation of intersubunit interactions between the two co-residing T4SS. It would also be curious to understand the benefit of possessing multiple T4SS to the host, perhaps possessing multiple pilin copies generates multiple serological types of F-pili, hence masking them from host immune systems.

Lastly, an interesting aspect of the F-T4SS is the presence of F-specific phages that recognise the F-pili as their receptor. It is curious what is the mechanism of DNA injection into the cytoplasm by the phage, and whether it occurs through the T4SS. Specific mutations of TraC have been shown to hinder F-specific phage infection, thus proving an interaction between the phage and the IMC of the F-T4SS, the significance of which is unknown.

By addressing these key questions, we can gain a deeper understanding of bacterial conjugation and its role in bacterial evolution and pathogenesis.

References

Achtman, M. (1975). "Mating aggregates in Escherichia coli conjugation." Journal of bacteriology **123**(2): 505-515.

Achtman, M., Kennedy, N. and Skurray, R. (1977). "Cell--cell interactions in conjugating Escherichia coli: role of traT protein in surface exclusion." Proceedings of the National Academy of Sciences **74**(11): 5104-5108.

Afonine, P. V., Poon, B. K., Read, R. J., Sobolev, O. V., Terwilliger, T. C., Urzhumtsev, A. and Adams, P. D. (2018). "Real-space refinement in PHENIX for cryo-EM and crystallography." Acta Crystallogr D Struct Biol **74**(Pt 6): 531-544.

Alvarez-Martinez, C. E. and Christie, P. J. (2009). "Biological diversity of prokaryotic type IV secretion systems." Microbiol Mol Biol Rev **73**(4): 775-808.

Álvarez-Rodríguez, I., Arana, L., Ugarte-Uribe, B., Gómez-Rubio, E., Martín-Santamaría, S., Garbisu, C. and Alkorta, I. (2020). "Type IV Coupling Proteins as Potential Targets to Control the Dissemination of Antibiotic Resistance." Frontiers in Molecular Biosciences **7**.

Aly, K. A. and Baron, C. (2007). "The VirB5 protein localizes to the T-pilus tips in Agrobacterium tumefaciens." Microbiology (Reading) **153**(Pt 11): 3766-3775.

Amin, H., Ilangovan, A. and Costa, T. R. D. (2021). "Architecture of the outer-membrane core complex from a conjugative type IV secretion system." Nature Communications **12**(1): 6834.

Amro, J., Black, C., Jemouai, Z., Rooney, N., Daneault, C., Zeytuni, N., Ruiz, M., Bui, K. H. and Baron, C. (2023). "Cryo-EM structure of the Agrobacterium tumefaciens T-pilus reveals the importance of positive charges in the lumen." Structure **31**(4): 375-384.

Andersson, M., Björnham, O., Svantesson, M., Badahdah, A., Uhlin, B. E. and Bullitt, E. (2012). "A structural basis for sustained bacterial adhesion: biomechanical properties of CFA/I pili." J Mol Biol **415**(5): 918-928.

Arutyunov, D. and Frost, L. S. (2013). "F conjugation: Back to the beginning." Plasmid **70**(1): 18-32.

Asrat, S., de Jesus, D. A., Hempstead, A. D., Ramabhadran, V. and Isberg, R. R. (2014). "Bacterial pathogen manipulation of host membrane trafficking." Annu Rev Cell Dev Biol **30**: 79-109.

Atmakuri, K., Cascales, E. and Christie, P. J. (2004). "Energetic components VirD4, VirB11 and VirB4 mediate early DNA transfer reactions required for bacterial type IV secretion." Mol Microbiol **54**(5): 1199-1211.

Babić, A., Lindner, A. B., Vulić, M., Stewart, E. J. and Radman, M. (2008). "Direct Visualization of Horizontal Gene Transfer." Science **319**(5869): 1533-1536.

Backert, S. and Meyer, T. F. (2006). "Type IV secretion systems and their effectors in bacterial pathogenesis." Curr Opin Microbiol **9**(2): 207-217.

Baek, M., DiMaio, F., Anishchenko, I., Dauparas, J., Ovchinnikov, S., Lee, G. R., Wang, J., Cong, Q., Kinch, L. N., Schaeffer, R. D., Millán, C., Park, H., Adams, C., Glassman, C. R., DeGiovanni, A., Pereira, J. H., Rodrigues, A. V., van Dijk, A. A., Ebrecht, A. C., Opperman, D. J., Sagmeister, T., Buhlheller, C., Pavkov-Keller, T., Rathinaswamy, M. K., Dalwadi, U., Yip, C. K., Burke, J. E., Garcia, K. C., Grishin, N. V., Adams, P. D., Read, R.

J. and Baker, D. (2021). "Accurate prediction of protein structures and interactions using a three-track neural network." Science **373**(6557): 871-876.

Bagdasarian, M., Bailone, A., Angulo, J. F., Scholz, P., Bagdasarian, M. and Devoret, R. (1992). "PsiB, an anti‐SOS protein, is transiently expressed by the F sex factor during its transmission to an Escherichia coli K‐12 recipient." Molecular microbiology **6**(7): 885-893.

Baker, L. A. and Rubinstein, J. L. (2010). "Radiation damage in electron cryomicroscopy." Methods Enzymol **481**: 371-388.

Banta, L. M., Kerr, J. E., Cascales, E., Giuliano, M. E., Bailey, M. E., McKay, C., Chandran, V., Waksman, G. and Christie, P. J. (2011). "An Agrobacterium VirB10 Mutation Conferring a Type IV Secretion System Gating Defect." Journal of Bacteriology **193**(10): 2566-2574.

Barnich, N., Carvalho, F. A., Glasser, A. L., Darcha, C., Jantscheff, P., Allez, M., Peeters, H., Bommelaer, G., Desreumaux, P., Colombel, J. F. and Darfeuille-Michaud, A. (2007). "CEACAM6 acts as a receptor for adherent-invasive E. coli, supporting ileal mucosa colonization in Crohn disease." J Clin Invest **117**(6): 1566-1574.

Baumberg, S. and Dennison, S. (1975). "Variation in expression of sex factor genes in the Proteus-Providencia group relative to Escherichia coli." Journal of Bacteriology **123**(1): 278-286.

Bayer-Santos, E., Cenens, W., Matsuyama, B. Y., Oka, G. U., Di Sessa, G., Mininel, I. D. V., Alves, T. L. and Farah, C. S. (2019). "The opportunistic pathogen Stenotrophomonas maltophilia utilizes a type IV secretion system for interbacterial killing." PLOS Pathogens **15**(9): e1007651.

Becker, E. C. and Meyer, R. (2012). "Origin and fate of the 3′ ends of single-stranded DNA generated by conjugal transfer of plasmid R1162." Journal of bacteriology **194**(19): 5368-5376.

Beltrán, L., Torsilieri, H., Patkowski, J. B., Yang, J. E., Casanova, J., Costa, T. R. D., Wright, E. R. and Egelman, E. H. (2023). "The mating pilus of E. coli pED208 acts as a conduit for ssDNA during horizontal gene transfer." mBio **15**(1): e02857-02823.

Beltran, L. C., Cvirkaite-Krupovic, V., Miller, J., Wang, F., Kreutzberger, M. A. B., Patkowski, J. B., Costa, T. R. D., Schouten, S., Levental, I., Conticello, V. P., Egelman, E. H. and Krupovic, M. (2023). "Archaeal DNA-import apparatus is homologous to bacterial conjugation machinery." Nat Commun **14**(1): 666.

Bengtsson, S., Naseer, U., Sundsfjord, A., Kahlmeter, G. and Sundqvist, M. (2011). "Sequence types and plasmid carriage of uropathogenic Escherichia coli devoid of phenotypically detectable resistance." Journal of Antimicrobial Chemotherapy **67**(1): 69-73.

Bi, D., Liu, L., Tai, C., Deng, Z., Rajakumar, K. and Ou, H. Y. (2013). "SecReT4: a web-based bacterial type IV secretion system resource." Nucleic Acids Res **41**(Database issue): D660-665.

Biesen, T. v. and Frost, L. S. (1994). "The FinO protein of IncF plasmids binds FinP antisense RNA and its target, traJ mRNA, and promotes duplex formation." Molecular microbiology **14**(3): 427-436.

Binder, V. and Orholm, M. (1996). "Familial occurrence and inheritance studies in inflammatory bowel disease." Neth J Med **48**(2): 53-56.

Blair, J. M., Webber, M. A., Baylay, A. J., Ogbolu, D. O. and Piddock, L. J. (2015). "Molecular mechanisms of antibiotic resistance." Nature reviews microbiology **13**(1): 42-51.

Bock, L. V. and Grubmüller, H. (2022). "Effects of cryo-EM cooling on structural ensembles." Nature Communications **13**(1): 1709.

Böhm, M.-E., Huptas, C., Krey, V. M. and Scherer, S. (2015). "Massive horizontal gene transfer, strictly vertical inheritance and ancient duplications differentially shape the evolution of Bacillus cereus enterotoxin operons hbl, cytK and nhe." BMC evolutionary biology **15**(1): 1-17.

Boudeau, J., Barnich, N. and Darfeuille-Michaud, A. (2001). "Type 1 pili-mediated adherence of Escherichia coli strain LF82 isolated from Crohn's disease is involved in bacterial invasion of intestinal epithelial cells." Mol Microbiol **39**(5): 1272-1284.

Boudeau, J., Glasser, A. L., Masseret, E., Joly, B. and Darfeuille-Michaud, A. (1999). "Invasive ability of an Escherichia coli strain isolated from the ileal mucosa of a patient with Crohn's disease." Infect Immun **67**(9): 4499-4509.

Bradley, D., Coetzee, J., Bothma, T. and Hedges, R. (1981). "Phage X: a plasmid-dependent, broad host range, filamentous bacterial virus." Microbiology **126**(2): 389-396.

Bradley, D. E. (1979). "Morphology of pili determined by the N incompatibility group plasmid N3 and interaction with bacteriophages PR4 and IKe." Plasmid **2**(4): 632-636.

Bradley, D. E. (1980). "Morphological and serological relationships of conjugative pili." Plasmid **4**(2): 155-169.

Bradley, D. E. (1984). "Characteristics and function of thick and thin conjugative pili determined by transfer-derepressed plasmids of incompatibility groups I1, I2, I5, B, K and Z." Microbiology **130**(6): 1489-1502.

Bradley, D. E., Taylor, D. E. and Cohen, D. R. (1980). "Specification of surface mating systems among conjugative drug resistance plasmids in Escherichia coli K-12." Journal of Bacteriology **143**(3): 1466-1470.

Bragagnolo, N., Rodriguez, C., Samari-Kermani, N., Fours, A., Korouzhdehi, M., Lysenko, R. and Audette, G. F. (2020). "Protein Dynamics in F-like Bacterial Conjugation." Biomedicines **8**(9).

Brito, I. L. (2021). "Examining horizontal gene transfer in microbial communities." Nature Reviews Microbiology **19**(7): 442-453.

Buisson, A., Sokol, H., Hammoudi, N., Treton, X., Nachury, M., Fumery, M., Hébuterne, X., Rodrigues, M., Hugot, J.-P. and Boschetti, G. (2023). "Role of adherent and invasive Escherichia coli in Crohn's disease: lessons from the postoperative recurrence model." Gut **72**(1): 39-48.

Bunduc, C. M., Fahrenkamp, D., Wald, J., Ummels, R., Bitter, W., Houben, E. N. and Marlovits, T. C. (2021). "Structure and dynamics of a mycobacterial type VII secretion system." Nature **593**(7859): 445-448.

Burns, D. L. (2021). "Secretion of pertussis toxin from Bordetella pertussis." Toxins **13**(8): 574.

Büttner, D. (2012). "Protein export according to schedule: architecture, assembly, and regulation of type III secretion systems from plant- and animal-pathogenic bacteria." Microbiol Mol Biol Rev **76**(2): 262-310.

Cabezon, E. and de la Cruz, F. (2006). "TrwB: An F1-ATPase-like molecular motor involved in DNA transport during bacterial conjugation." Research in Microbiology **157**(4): 299-305.

Cabezón, E., Ripoll-Rozada, J., Peña, A., De La Cruz, F. and Arechaga, I. (2015). "Towards an integrated model of bacterial conjugation." FEMS microbiology reviews **39**(1): 81-95.

Cabezon, E., Sastre, J. I. and De La Cruz, F. (1997). "Genetic evidence of a coupling role for the TraG protein family in bacterial conjugation." Molecular and General Genetics MGG **254**: 400-406.

Callaghan, M. M., Heilers, J. H., van der Does, C. and Dillard, J. P. (2017). "Secretion of Chromosomal DNA by the Neisseria gonorrhoeae Type IV Secretion System." Curr Top Microbiol Immunol **413**: 323-345.

Callaway, E. (2022). "The entire protein universe': AI predicts shape of nearly every known protein." Nature **608**(7921): 15-16.

Cao, Y., Hallet, B., Sherratt, D. J. and Hayes, F. (1997). "Structure-function correlations in the XerD site-specific recombinase revealed by pentapeptide scanning mutagenesis." J Mol Biol **274**(1): 39-53.

Caparon, M. G. and Scott, J. R. (1989). "Excision and insertion of the conjugative transposon Tn916 involves a novel recombination mechanism." Cell **59**(6): 1027-1034.

Carraro, N., Poulin, D. and Burrus, V. (2015). "Replication and Active Partition of Integrative and Conjugative Elements (ICEs) of the SXT/R391 Family: The Line between ICEs and Conjugative Plasmids Is Getting Thinner." PLOS Genetics **11**(6): e1005298.

Cascales, E., Atmakuri, K., Sarkar, M. K. and Christie, P. J. (2013). "DNA Substrate-Induced Activation of the Agrobacterium VirB/VirD4 Type IV Secretion System." Journal of Bacteriology **195**(11): 2691-2704.

Cascales, E. and Christie, P. J. (2004a). "Agrobacterium VirB10, an ATP energy sensor required for type IV secretion." Proceedings of the National Academy of Sciences **101**(49): 17228-17233.

Cascales, E. and Christie, P. J. (2004b). "Definition of a bacterial type IV secretion pathway for a DNA substrate." Science **304**(5674): 1170-1173.

Chabbert, Y., Scavizzi, M., Witchitz, J., Gerbaud, G. and Bouanchaud, D. (1972). "Incompatibility groups and the classification of fi− resistance factors." Journal of Bacteriology **112**(2): 666-675.

Chandler, M., De La Cruz, F., Dyda, F., Hickman, A. B., Moncalian, G. and Ton-Hoang, B. (2013). "Breaking and joining single-stranded DNA: the HUH endonuclease superfamily." Nature Reviews Microbiology **11**(8): 525-538.

Chandran, V., Fronzes, R., Duquerroy, S., Cronin, N., Navaza, J. and Waksman, G. (2009). "Structure of the outer membrane complex of a type IV secretion system." Nature **462**(7276): 1011-1015.

Chassaing, B. and Darfeuille-Michaud, A. (2013). "The σE pathway is involved in biofilm formation by Crohn's disease-associated adherent-invasive Escherichia coli." J Bacteriol **195**(1): 76-84.

Chassaing, B., Rolhion, N., de Vallée, A., Salim, S. Y., Prorok-Hamon, M., Neut, C., Campbell, B. J., Söderholm, J. D., Hugot, J. P., Colombel, J. F. and Darfeuille-Michaud, A. (2011). "Crohn disease--associated adherent-invasive E. coli bacteria target mouse and human Peyer's patches via long polar fimbriae." J Clin Invest **121**(3): 966-975.

Chen, Y., Zhang, X., Manias, D., Yeo, H. J., Dunny, G. M. and Christie, P. J. (2008). "Enterococcus faecalis PcfC, a spatially localized substrate receptor for type IV secretion of the pCF10 transfer intermediate." J Bacteriol **190**(10): 3632-3645.

Chetrit, D., Hu, B., Christie, P. J., Roy, C. R. and Liu, J. (2018). "A unique cytoplasmic ATPase complex defines the Legionella pneumophila type IV secretion channel." Nature Microbiology **3**(6): 678-686.

Christie, P. J. (2016). "The Mosaic Type IV Secretion Systems." EcoSal Plus **7**(1).

Christie, P. J. (2023). "Illuminating type IV secretion-mediated DNA trafficking through long filaments." Proceedings of the National Academy of Sciences **120**(50): e2318508120.

Christie, P. J., Atmakuri, K., Krishnamoorthy, V., Jakubowski, S. and Cascales, E. (2005). "Biogenesis, architecture, and function of bacterial type IV secretion systems." Annu Rev Microbiol **59**: 451-485.

Christie, P. J. and Cascales, E. (2005). "Structural and dynamic properties of bacterial type IV secretion systems (review)." Mol Membr Biol **22**(1-2): 51-61.

Christie, P. J. and Vogel, J. P. (2000). "Bacterial type IV secretion: conjugation systems adapted to deliver effector molecules to host cells." Trends in microbiology **8**(8): 354-360.

Christie, P. J., Whitaker, N. and González-Rivera, C. (2014). "Mechanism and structure of the bacterial type IV secretion systems." Biochim Biophys Acta **1843**(8): 1578-1591.

Clarke, M., Maddera, L., Harris, R. L. and Silverman, P. M. (2008). "F-pili dynamics by live-cell imaging." Proc Natl Acad Sci U S A **105**(46): 17978-17981.

Coetzee, J., Bradley, D., Hedges, R., Fleming, J. and Lecatsas, G. (1983). "Bacteriophage M: an incompatibility group M plasmid-specific phage." Microbiology **129**(7): 2271-2276.

Coluzzi, C., Garcillán-Barcia, M. P., de la Cruz, F. and Rocha, E. P. C. (2022). "Evolution of Plasmid Mobility: Origin and Fate of Conjugative and Nonconjugative Plasmids." Mol Biol Evol **39**(6).

Coppens, F., Castaldo, G., Debraekeleer, A., Subedi, S., Moonens, K., Lo, A. and Remaut, H. (2018). "Hop-family Helicobacter outer membrane adhesins form a novel class of Type 5-like secretion proteins with an interrupted β-barrel domain." Molecular Microbiology **110**(1): 33-46.

Coppens, F., Iyyathurai, J., Ruer, S., Fioravanti, A., Taganna, J., Vereecke, L., De Greve, H. and Remaut, H. (2015). "Structural and adhesive properties of the long polar fimbriae protein LpfD from adherent-invasive Escherichia coli." Acta Crystallogr D Biol Crystallogr **71**(Pt 8): 1615-1626.

Costa, T., Ignatiou, A. and Orlova, E. V. (2017). Structural Analysis of Protein Complexes by Cryo Electron Microscopy. **1615**.

Costa, T. R., Felisberto-Rodrigues, C., Meir, A., Prevost, M. S., Redzej, A., Trokter, M. and Waksman, G. (2015). "Secretion systems in Gram-negative bacteria: structural and mechanistic insights." Nat Rev Microbiol **13**(6): 343-359.

Costa, T. R. D., Harb, L., Khara, P., Zeng, L., Hu, B. and Christie, P. J. (2021). "Type IV secretion systems: Advances in structure, function, and activation." Mol Microbiol **115**(3): 436-452.

Costa, T. R. D., Ilangovan, A., Ukleja, M., Redzej, A., Santini, J. M., Smith, T. K., Egelman, E. H. and Waksman, G. (2016). "Structure of the Bacterial Sex F Pilus Reveals an Assembly of a Stoichiometric Protein-Phospholipid Complex." Cell **166**(6): 1436-1444.e1410.

Costa, T. R. D., Patkowski, J. B., Macé, K., Christie, P. J. and Waksman, G. (2023). "Structural and functional diversity of type IV secretion systems." Nat Rev Microbiol.

Couturier, A., Virolle, C., Goldlust, K., Berne-Dedieu, A., Reuter, A., Nolivos, S., Yamaichi, Y., Bigot, S. and Lesterlin, C. (2023). "Real-time visualisation of the intracellular dynamics of conjugative plasmid transfer." Nature Communications **14**(1): 294.

Cover, T. L., Lacy, D. B. and Ohi, M. D. (2020). "The Helicobacter pylori Cag Type IV Secretion System." Trends Microbiol **28**(8): 682-695.

Crohn, B. B., Ginzburg, L. and Oppenheimer, G. D. (1932). "Regional ileitis: a pathologic and clinical entity." Journal of the American Medical Association **99**(16): 1323-1329.

Croll, T. I. (2018). "ISOLDE: a physically realistic environment for model building into low-resolution electron-density maps." Acta Crystallogr D Struct Biol **74**(Pt 6): 519-530.

Daehnel, K., Harris, R., Maddera, L. and Silverman, P. (2005). "Fluorescence assays for F-pili and their application." Microbiology **151**(11): 3541-3548.

Dahmen, S., Madec, J. Y. and Haenni, M. (2013). "F2:A-:B- plasmid carrying the extended-spectrum β-lactamase bla(CTX-M-55/57) gene in Proteus mirabilis isolated from a primate." Int J Antimicrob Agents **41**(6): 594-595.

Darbari, V. C., Ciccone, J., Patel, J. S., Islam, B., Agarwal, P. K. and Haider, S. (2020). "Electrostatic Switching Controls Channel Dynamics of the Sensor Protein VirB10 in A. tumefaciens Type IV Secretion System." ACS Omega **5**(7): 3271-3281.

Darby, E. M., Trampari, E., Siasat, P., Gaya, M. S., Alav, I., Webber, M. A. and Blair, J. M. A. (2023). "Molecular mechanisms of antibiotic resistance revisited." Nature Reviews Microbiology **21**(5): 280-295.

Darfeuille-Michaud, A., Neut, C., Barnich, N., Lederman, E., Di Martino, P., Desreumaux, P., Gambiez, L., Joly, B., Cortot, A. and Colombel, J. F. (1998). "Presence of adherent Escherichia coli strains in ileal mucosa of patients with Crohn's disease." Gastroenterology **115**(6): 1405-1413.

Datta, N. and Hedges, R. (1971). "Compatibility groups among fi− R factors." Nature **234**(5326): 222-223.

Datta, N. and Hughes, V. M. (1983). "Plasmids of the same Inc groups in Enterobacteria before and after the medical use of antibiotics." Nature **306**(5943): 616-617.

Davies, M. R., Shera, J., Van Domselaar, G. H., Sriprakash, K. S. and McMillan, D. J. (2009). "A novel integrative conjugative element mediates genetic transfer from group G Streptococcus to other beta-hemolytic Streptococci." J Bacteriol **191**(7): 2257-2265.

Davison, J. (1999). "Genetic Exchange between Bacteria in the Environment." Plasmid **42**(2): 73-91.

De La Cruz, F., Frost, L. S., Meyer, R. J. and Zechner, E. L. (2010). "Conjugative DNA metabolism in Gram-negative bacteria." FEMS Microbiology Reviews **34**(1): 18-40.

Decano, A. G., Pettigrew, K., Sabiiti, W., Sloan, D. J., Neema, S., Bazira, J., Kiiru, J., Onyango, H., Asiimwe, B. and Holden, M. T. G. (2021). "Pan-Resistome Characterization of Uropathogenic Escherichia coli and Klebsiella pneumoniae Strains Circulating in Uganda and Kenya, Isolated from 2017-2018." Antibiotics (Basel) **10**(12).

del Campo, I., Ruiz, R., Cuevas, A., Revilla, C., Vielva, L. and de la Cruz, F. (2012). "Determination of conjugation rates on solid surfaces." Plasmid **67**(2): 174-182.

delToro, D., Ortiz, D., Ordyan, M., Sippy, J., Oh, C. S., Keller, N., Feiss, M., Catalano, C. E. and Smith, D. E. (2016). "Walker-A Motif Acts to Coordinate ATP Hydrolysis with Motor Output in Viral DNA Packaging." J Mol Biol **428**(13): 2709-2729.

Dennison, S. and Baumberg, S. (1975). "Conjugational behaviour of N plasmids in Escherichia coli K12." Molecular and General Genetics MGG **138**(4): 323-331.

Dickson, C. J., Walker, R. C. and Gould, I. R. (2022). "Lipid21: complex lipid membrane simulations with AMBER." Journal of chemical theory and computation **18**(3): 1726-1736.

Dolejska, M. and Papagiannitsis, C. C. (2018). "Plasmid-mediated resistance is going wild." Plasmid **99**: 99-111.

Dubochet, J., Adrian, M., Chang, J.-J., Homo, J.-C., Lepault, J., McDowall, A. W. and Schultz, P. (1988). "Cryo-electron microscopy of vitrified specimens." Quarterly reviews of biophysics **21**(2): 129-228.

Durie, C. L., Sheedlo, M. J., Chung, J. M., Byrne, B. G., Su, M., Knight, T., Swanson, M., Lacy, D. B. and Ohi, M. D. (2020). "Structural analysis of the Legionella pneumophila Dot/Icm type IV secretion system core complex." Elife **9**.

Dutka, P., Liu, Y., Maggi, S., Ghosal, D., Wang, J., Carter, S. D., Zhao, W., Vijayrajratnam, S., Vogel, J. P. and Jensen, G. J. (2023). "Structure and Function of the Dot/Icm T4SS." bioRxiv: 2023.2003.2022.533729.

Eckburg, P. B. and Relman, D. A. (2007). "The role of microbes in Crohn's disease." Clin Infect Dis **44**(2): 256-262.

Eder, P., Dobrowolska, A. and Rydzewska, G. (2019). "Historical Upheavals and Eponyms in Crohn's Disease—Is There Any Relation?" Journal of Crohn's and Colitis **14**(5): 719-719.

Elhenawy, W., Hordienko, S., Gould, S., Oberc, A. M., Tsai, C. N., Hubbard, T. P., Waldor, M. K. and Coombes, B. K. (2021). "High-throughput fitness screening and transcriptomics identify a role for a type IV secretion system in the pathogenesis of Crohn's disease-associated Escherichia coli." Nature Communications **12**(1): 2032.

Elhenawy, W., Tsai, C. N. and Coombes, B. K. (2019). "Host-specific adaptive diversification of Crohn's disease-associated adherent-invasive Escherichia coli." Cell Host & Microbe **25**(2): 301-312. e305.

Emsley, P., Lohkamp, B., Scott, W. G. and Cowtan, K. (2010). "Features and development of Coot." Acta Crystallogr D Biol Crystallogr **66**(Pt 4): 486-501.

Erzberger, J. P. and Berger, J. M. (2006). "Evolutionary relationships and structural mechanisms of AAA+ proteins." Annual Review of Biophysics and Biomolecular Structure **35**(1): 93-114.

Fernandez-Lopez, R., de Toro, M., Moncalian, G., Garcillan-Barcia, M. P. and de la Cruz, F. (2016). "Comparative Genomics of the Conjugation Region of F-like Plasmids: Five Shades of F." Front Mol Biosci **3**: 71.

Filloux, A. (2022). "Bacterial protein secretion systems: Game of types." Microbiology **168**(5).

Fiocchi, C. (2005). "Inflammatory bowel disease pathogenesis: therapeutic implications." Chinese Journal of Digestive Diseases **6**(1): 6-9.

Flannery, E. L., Antczak, S. M. and Mobley, H. L. (2011). "Self-transmissibility of the integrative and conjugative element ICEPm1 between clinical isolates requires a functional integrase, relaxase, and type IV secretion system." J Bacteriol **193**(16): 4104-4112.

Flores-Ríos, R., Moya-Beltrán, A., Pareja-Barrueto, C., Arenas-Salinas, M., Valenzuela, S., Orellana, O. and Quatrini, R. (2019). "The type IV secretion system of ICEAfe1: formation of a conjugative pilus in Acidithiobacillus ferrooxidans." Frontiers in microbiology **10**: 434841.

Franken, L. E., Grünewald, K., Boekema, E. J. and Stuart, M. C. A. (2020). "A Technical Introduction to Transmission Electron Microscopy for Soft-Matter: Imaging, Possibilities, Choices, and Technical Developments." Small **16**(14): 1906198.

Frick-Cheng, A. E., Pyburn, T. M., Voss, B. J., McDonald, W. H., Ohi, M. D. and Cover, T. L. (2016). "Molecular and Structural Analysis of the Helicobacter pylori cag Type IV Secretion System Core Complex." mBio **7**(1): e02001-02015.

Fronzes, R., Schäfer, E., Wang, L., Saibil, H. R., Orlova, E. V. and Waksman, G. (2009). "Structure of a Type IV Secretion System Core Complex." Science **323**(5911): 266-268.

Frost, L. S., Leplae, R., Summers, A. O. and Toussaint, A. (2005). "Mobile genetic elements: the agents of open source evolution." Nature Reviews Microbiology **3**(9): 722-732.

Frost, L. S. and Paranchych, W. (1988). "DNA sequence analysis of point mutations in traA, the F pilin gene, reveal two domains involved in F-specific bacteriophage attachment." Molecular and General Genetics MGG **213**: 134-139.

Garcillán-Barcia, M. P., Jurado, P., González-Pérez, B., Moncalián, G., Fernández, L. A. and de la Cruz, F. (2007). "Conjugative transfer can be inhibited by blocking relaxase activity within recipient cells with intrabodies." Mol Microbiol **63**(2): 404-416.

Garcillán-Barcia, M. P., Redondo-Salvo, S. and de la Cruz, F. (2023). "Plasmid classifications." Plasmid **126**: 102684.

Ghigo, J. M. (2001). "Natural conjugative plasmids induce bacterial biofilm development." Nature **412**(6845): 442-445.

Gibiński, K. (1989). "Antoni Leśniowski and his contribution to regional enteritis (Crohn's disease)." Journal of the Royal Society of Medicine **82**(4): 243.

Gibold, L., Garenaux, E., Dalmasso, G., Gallucci, C., Cia, D., Mottet-Auselo, B., Faïs, T., Darfeuille-Michaud, A., Nguyen, H. T., Barnich, N., Bonnet, R. and Delmas, J. (2016). "The Vat-AIEC protease promotes crossing of the intestinal mucus layer by Crohn's disease-associated Escherichia coli." Cell Microbiol **18**(5): 617-631.

Glasser, A. L., Boudeau, J., Barnich, N., Perruchot, M. H., Colombel, J. F. and Darfeuille-Michaud, A. (2001). "Adherent invasive Escherichia coli strains from patients with Crohn's disease survive and replicate within macrophages without inducing host cell death." Infect Immun **69**(9): 5529-5537.

Goldlust, K., Ducret, A., Halte, M., Dedieu-Berne, A., Erhardt, M. and Lesterlin, C. (2023). "The F pilus serves as a conduit for the DNA during conjugation between physically distant bacteria." Proceedings of the National Academy of Sciences **120**(47): e2310842120.

Gomez-Valero, L., Rusniok, C., Carson, D., Mondino, S., Perez-Cobas, A. E., Rolando, M., Pasricha, S., Reuter, S., Demirtas, J., Crumbach, J., Descorps-Declere, S., Hartland, E. L., Jarraud, S., Dougan, G., Schroeder, G. N., Frankel, G. and Buchrieser, C. (2019). "More than 18,000 effectors in the Legionella genus genome provide multiple, independent combinations for replication in human cells." Proc Natl Acad Sci U S A **116**(6): 2265-2273.

Gomis-Rüth, F. X. and Coll, M. (2001). "Structure of TrwB, a gatekeeper in bacterial conjugation." The International Journal of Biochemistry & Cell Biology **33**(9): 839-843.

Gomis-Rüth, F. X., Moncalián, G., Pérez-Luque, R., González, A., Cabezón, E., de la Cruz, F. and Coll, M. (2001). "The bacterial conjugation protein TrwB resembles ring helicases and F1-ATPase." Nature **409**(6820): 637-641.

Gonzalez-Rivera, C., Bhatty, M. and Christie, P. J. (2016). "Mechanism and function of type IV secretion during infection of the human host." Microbiology spectrum **4**(3): 10.1128/microbiolspec. vmbf-0024-2015.

González-Torres, P., Rodríguez-Mateos, F., Antón, J. and Gabaldón, T. (2019). "Impact of Homologous Recombination on the Evolution of Prokaryotic Core Genomes." mBio **10**(1): 10.1128/mbio.02494-02418.

Gordon, J. E., Costa, T. R. D., Patel, R. S., Gonzalez-Rivera, C., Sarkar, M. K., Orlova, E. V., Waksman, G. and Christie, P. J. (2017). "Use of chimeric type IV secretion systems to define contributions of outer membrane subassemblies for contact-dependent translocation." Mol Microbiol **105**(2): 273-293.

Goulet, A. and Cambillau, C. (2022). "Present Impact of AlphaFold2 Revolution on Structural Biology, and an Illustration With the Structure Prediction of the Bacteriophage J-1 Host Adhesion Device." Front Mol Biosci **9**: 907452.

Grohmann, E., Christie, P. J., Waksman, G. and Backert, S. (2018). "Type IV secretion in Gram-negative and Gram-positive bacteria." Mol Microbiol **107**(4): 455-471.

Gruber, C. J., Lang, S., Rajendra, V. K. H., Nuk, M., Raffl, S., Schildbach, J. F. and Zechner, E. L. (2016). "Conjugative DNA Transfer Is Enhanced by Plasmid R1 Partitioning Proteins." Frontiers in Molecular Biosciences **3**.

Guglielmini, J., Quintais, L., Garcillán-Barcia, M. P., de la Cruz, F. and Rocha, E. P. C. (2011). "The Repertoire of ICE in Prokaryotes Underscores the Unity, Diversity, and Ubiquity of Conjugation." PLOS Genetics **7**(8): e1002222.

Gunton, J. E., Gilmour, M. W., Baptista, K. P., Lawley, T. D. and Taylor, D. E. (2007). "Interaction between the co-inherited TraG coupling protein and the TraJ membrane-associated protein of the H-plasmid conjugative DNA transfer system resembles chromosomal DNA translocases." Microbiology **153**(2): 428-441.

Guzmán-Herrador, D. L., Steiner, S., Alperi, A., González-Prieto, C., Roy, C. R. and Llosa, M. (2017). "DNA Delivery and Genomic Integration into Mammalian Target Cells through Type IV A and B Secretion Systems of Human Pathogens." Frontiers in Microbiology **8**.

Haft, R. J., Gachelet, E. G., Nguyen, T., Toussaint, L., Chivian, D. and Traxler, B. (2007). "In vivo oligomerization of the F conjugative coupling protein TraD." J Bacteriol **189**(18): 6626-6634.

Hare, S., Bayliss, R., Baron, C. and Waksman, G. (2006). "A Large Domain Swap in the VirB11 ATPase of Brucella suis Leaves the Hexameric Assembly Intact." Journal of Molecular Biology **360**(1): 56-66.

Harrington, L. C. and Rogerson, A. C. (1990). "The F pilus of Escherichia coli appears to support stable DNA transfer in the absence of wall-to-wall contact between cells." Journal of Bacteriology **172**(12): 7263-7264.

Hastak, P., Cummins, M. L., Gottlieb, T., Cheong, E., Merlino, J., Myers, G. S. A., Djordjevic, S. P. and Roy Chowdhury, P. (2020). "Genomic profiling of Escherichia coli isolates from bacteraemia patients: a 3-year cohort study of isolates collected at a Sydney teaching hospital." Microbial genomics **6**(5): e000371.

Hehemann, J.-H., Correc, G., Barbeyron, T., Helbert, W., Czjzek, M. and Michel, G. (2010). "Transfer of carbohydrate-active enzymes from marine bacteria to Japanese gut microbiota." Nature **464**(7290): 908-912.

Helinski, D. R. (2022). "A Brief History of Plasmids." EcoSal Plus **10**(1): eESP-0028-2021.

Hemme, C. L., Green, S. J., Rishishwar, L., Prakash, O., Pettenato, A., Chakraborty, R., Deutschbauer, A. M., Van Nostrand, J. D., Wu, L. and He, Z. (2016). "Lateral gene transfer in a heavy metal-contaminated-groundwater microbial community." MBio **7**(2): 10.1128/mbio. 02234-02215.

Ho, B. T., Dong, T. G. and Mekalanos, J. J. (2014). "A view to a kill: the bacterial type VI secretion system." Cell Host Microbe **15**(1): 9-21.

Ho, P. L., Lo, W. U., Yeung, M. K., Li, Z., Chan, J., Chow, K. H., Yam, W. C., Tong, A. H. Y., Bao, J. Y. J., Lin, C. H., Lok, S. and Chiu, S. S. (2012). "Dissemination of pHK01-like incompatibility group IncFII plasmids encoding CTX-M-14 in Escherichia coli from human and animal sources." Veterinary Microbiology **158**(1): 172-179.

Hofreuter, D., Odenbreit, S. and Haas, R. (2001). "Natural transformation competence in Helicobacter pylori is mediated by the basic components of a type IV secretion system." Molecular Microbiology **41**(2): 379-391.

Horiuchi, T. and Komano, T. (1998). "Mutational analysis of plasmid R64 thin pilus prepilin: the entire prepilin sequence is required for processing by type IV prepilin peptidase." Journal of bacteriology **180**(17): 4613-4620.

Hospenthal, M. K., Costa, T. R. D. and Waksman, G. (2017). "A comprehensive guide to pilus biogenesis in Gram-negative bacteria." Nature Reviews Microbiology **15**(6): 365-379.

Hu, B., Khara, P. and Christie, P. J. (2019a). "Structural bases for F plasmid conjugation and F pilus biogenesis in Escherichia coli." Proceedings of the National Academy of Sciences **116**(28): 14222-14227.

Hu, B., Khara, P., Song, L., Lin, A. S., Frick-Cheng, A. E., Harvey, M. L., Cover, T. L. and Christie, P. J. (2019b). "In Situ Molecular Architecture of the Helicobacter pylori Cag Type IV Secretion System." mBio **10**(3).

Humphrey, W., Dalke, A. and Schulten, K. (1996). "VMD: visual molecular dynamics." Journal of molecular graphics **14**(1): 33-38.

Ignatiou, A., Macé, K., Redzej, A., Costa, T. R. D., Waksman, G. and Orlova, E. V. (2024). "Structural Analysis of Protein Complexes by Cryo-Electron Microscopy." Methods Mol Biol **2715**: 431-470.

Ikuma, K. and Gunsch, C. K. (2012). "Genetic bioaugmentation as an effective method for in situ bioremediation: functionality of catabolic plasmids following conjugal transfers." Bioengineered **3**(4): 236-241.

Ilangovan, A., Kay, C. W. M., Roier, S., El Mkami, H., Salvadori, E., Zechner, E. L., Zanetti, G. and Waksman, G. (2017). "Cryo-EM Structure of a Relaxase Reveals the Molecular Basis of DNA Unwinding during Bacterial Conjugation." Cell **169**(4): 708-721.e712.

Jalajakumari, M., Guidolin, A., Buhk, H., Manning, P., Ham, L., Hodgson, A., Cheah, K. and Skurray, R. (1987). "Surface exclusion genes traS and traT of the F sex factor of Escherichia coli K-12: determination of the nucleotide sequence and promoter and terminator activities." Journal of molecular biology **198**(1): 1-11.

Jass, J., Schedin, S., Fällman, E., Ohlsson, J., Nilsson, U. J., Uhlin, B. E. and Axner, O. (2004). "Physical properties of Escherichia coli P pili measured by optical tweezers." Biophys J **87**(6): 4271-4283.

Jean, N. L., Rutherford, T. J. and Löwe, J. (2020). "FtsK in motion reveals its mechanism for double-stranded DNA translocation." Proceedings of the National Academy of Sciences **117**(25): 14202-14208.

Johnson, T. J. and Nolan, L. K. (2009). "Pathogenomics of the Virulence Plasmids of Escherichia coli." Microbiology and Molecular Biology Reviews **73**(4): 750-774.

Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W. and Klein, M. L. (1983). "Comparison of simple potential functions for simulating liquid water." The Journal of chemical physics **79**(2): 926-935.

Joung, I. S. and Cheatham III, T. E. (2008). "Determination of alkali and halide monovalent ion parameters for use in explicitly solvated biomolecular simulations." The journal of physical chemistry B **112**(30): 9020-9041.

Juhas, M., Crook, D. W., Dimopoulou, I. D., Lunter, G., Harding, R. M., Ferguson, D. J. and Hood, D. W. (2007). "Novel type IV secretion system involved in propagation of genomic islands." J Bacteriol **189**(3): 761-771.

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, S. A. A., Ballard, A. J., Cowie, A., Romera-Paredes, B., Nikolov, S., Jain, R., Adler, J. and Hassabis, D. (2021). "Highly accurate protein structure prediction with AlphaFold." Nature **596**(7873): 583-589.

Kamali Dolatabadi, R., Feizi, A., Halaji, M., Fazeli, H. and Adibi, P. (2021). "The Prevalence of Adherent-Invasive Escherichia coli and Its Association With Inflammatory Bowel Diseases: A Systematic Review and Meta-Analysis." Front Med (Lausanne) **8**: 730243.

Kanonenberg, K., Schwarz, C. K. and Schmitt, L. (2013). "Type I secretion systems - a story of appendices." Res Microbiol **164**(6): 596-604.

Ke, X., Bittencourt, C. and Van Tendeloo, G. (2015). "Possibilities and limitations of advanced transmission electron microscopy for carbon-based nanomaterials." Beilstein J Nanotechnol **6**: 1541-1557.

Kerr, J. E. and Christie, P. J. (2010). "Evidence for VirB4-Mediated Dislocation of Membrane-Integrated VirB2 Pilin during Biogenesis of the Agrobacterium VirB/VirD4 Type IV Secretion System." Journal of Bacteriology **192**(19): 4923-4934.

Khara, P., Song, L., Christie, P. J. and Hu, B. (2021). "In Situ Visualization of the pKM101-Encoded Type IV Secretion System Reveals a Highly Symmetric ATPase Energy Center." mBio **12**(5): e0246521.

Kim, S., Dallmann, H. G., McHenry, C. S. and Marians, K. J. (1996). "Coupling of a replicative polymerase and helicase: a tau-DnaB interaction mediates rapid replication fork movement." Cell **84**(4): 643-650.

Kishida, K., Li, Y. G., Ogawa-Kishida, N., Khara, P., Al Mamun, A. A. M., Bosserman, R. E. and Christie, P. J. (2024). "Chimeric systems composed of swapped Tra subunits between distantly-related F plasmids reveal striking plasticity among type IV secretion machines." PLoS Genet **20**(3): e1011088.

Kitao, T., Kubori, T. and Nagai, H. (2022). "Recent advances in structural studies of the Legionella pneumophila Dot/Icm type IV secretion system." Microbiol Immunol **66**(2): 67-74.

Klimke, W. A. and Frost, L. S. (1998). "Genetic analysis of the role of the transfer gene, traN, of the F and R100-1 plasmids in mating pair stabilisation during conjugation." Journal of bacteriology **180**(16): 4036-4043.

Klimke, W. A., Rypien, C. D., Klinger, B., Kennedy, R. A., Rodriguez-Maillard, J. M. and Frost, L. S. (2005). "The mating pair stabilisation protein, TraN, of the F plasmid is an outer-membrane protein with two regions that are important for its function in conjugation." Microbiology **151**(11): 3527-3540.

Koch, B., Callaghan, M. M., Tellechea-Luzardo, J., Seeger, A. Y., Dillard, J. P. and Krasnogor, N. (2020). "Protein interactions within and between two F-type type IV secretion systems." Mol Microbiol **114**(5): 823-838.

Koraimann, G. (2018). "Spread and Persistence of Virulence and Antibiotic Resistance Genes: A Ride on the F Plasmid Conjugation Module." EcoSal Plus **8**(1): 10.1128/ecosalplus.ESP-0003-2018.

Kostakioti, M., Hadjifrangiskou, M. and Hultgren, S. J. (2013). "Bacterial biofilms: development, dispersal, and therapeutic strategies in the dawn of the postantibiotic era." Cold Spring Harb Perspect Med **3**(4): a010306.

Kreida, S., Narita, A., Johnson, M. D., Tocheva, E. I., Das, A., Ghosal, D. and Jensen, G. J. (2023). "Cryo-EM structure of the Agrobacterium tumefaciens T4SS-associated T-pilus reveals stoichiometric protein-phospholipid assembly." Structure.

Kumar, R. B. and Das, A. (2002). "Polar location and functional domains of the Agrobacterium tumefaciens DNA transfer protein VirD4." Molecular Microbiology **43**(6): 1523-1532.

Kurbatfinski, N., Kramer, C. N., Goodman, S. D. and Bakaletz, L. O. (2023). "ESKAPEE pathogens newly released from biofilm residence by a targeted monoclonal are sensitized to killing by traditional antibiotics." Front Microbiol **14**: 1202215.

Lam, T., Ellison, C. K., Eddington, D. T., Brun, Y. V., Dalia, A. B. and Morrison, D. A. (2021). "Competence pili in Streptococcus pneumoniae are highly dynamic structures that retract to promote DNA uptake." Molecular Microbiology **116**(2): 381-396.

Larrea, D., de Paz, H. D., Matilla, I., Guzmán-Herrador, D. L., Lasso, G., de la Cruz, F., Cabezón, E. and Llosa, M. (2017). "Substrate translocation involves specific lysine residues of the central channel of the conjugative coupling protein TrwB." Molecular Genetics and Genomics **292**(5): 1037-1049.

Lawley, T. D., Klimke, W. A., Gubbins, M. J. and Frost, L. S. (2003). "F factor conjugation is a true type IV secretion system." FEMS Microbiol Lett **224**(1): 1-15.

Lawrence, J. G. (1999). "Gene transfer, speciation, and the evolution of bacterial genomes." Current opinion in microbiology **2**(5): 519-523.

Lederberg, J. (1952). "Cell genetics and hereditary symbiosis." Physiological reviews **32**(4): 403-430.

Lederberg, J. and Tatum, E. (1946). "Gene Recombination in Escherichia Coli." Nature **158**(4016): 558-558.

Lee, M. H., Kosuk, N., Bailey, J., Traxler, B. and Manoil, C. (1999). "Analysis of F factor TraD membrane topology by use of gene fusions and trypsin-sensitive insertions." J Bacteriol **181**(19): 6108-6113.

Leo, J. C., Grin, I. and Linke, D. (2012). "Type V secretion: mechanism(s) of autotransport through the bacterial outer membrane." Philos Trans R Soc Lond B Biol Sci **367**(1592): 1088-1101.

Leśniowski, A. (1903). "Przyczynek do chirurgii kiszek." Warszawa: Medycyna: 460-464.

Li, D. X., Zhang, S. M., Hu, G. Z., Wang, Y., Liu, H. B., Wu, C. M., Shang, Y. H., Chen, Y. X. and Du, X. D. (2012). "Tn3-associated rmtB together with qnrS1, aac(6')-Ib-cr and bla CTX-M-15 are co-located on an F49:A-:B- plasmid in an Escherichia coli ST10 strain in China." J Antimicrob Chemother **67**(1): 236-238.

Li, Y. G. and Christie, P. J. (2018). "The Agrobacterium VirB/VirD4 T4SS: Mechanism and Architecture Defined Through In Vivo Mutagenesis and Chimeric Systems." Curr Top Microbiol Immunol **418**: 233-260.

Liao, X.-P., Liu, B.-T., Yang, Q.-E., Sun, J., Li, L., Fang, L.-X. and Liu, Y.-H. (2013). "Comparison of Plasmids Coharboring 16S rRNA Methylase and Extended-Spectrum β-Lactamase Genes among Escherichia coli Isolates from Pets and Poultry." Journal of Food Protection **76**(12): 2018-2023.

Licht, T. R., Christensen, B. B., Krogfelt, K. A. and Molin, S. (1999). "Plasmid transfer in the animal intestine and other dynamic bacterial populations: the role of community structure and environment." Microbiology (Reading) **145 (Pt 9)**: 2615-2622.

Lin, Z., Akin, H., Rao, R., Hie, B., Zhu, Z., Lu, W., Smetanin, N., Verkuil, R., Kabeli, O. and Shmueli, Y. (2023). "Evolutionary-scale prediction of atomic-level protein structure with a language model." Science **379**(6637): 1123-1130.

Liu, P., Li, P., Jiang, X., Bi, D., Xie, Y., Tai, C., Deng, Z., Rajakumar, K. and Ou, H. Y. (2012). "Complete genome sequence of Klebsiella pneumoniae subsp. pneumoniae HS11286, a multidrug-resistant strain isolated from human sputum." J Bacteriol **194**(7): 1841-1842.

Liu, X., Khara, P., Baker, M. L., Christie, P. J. and Hu, B. (2022). "Structure of a type IV secretion system core complex encoded by multi-drug resistance F plasmids." Nature Communications **13**(1): 379.

Llosa, M. and Alkorta, I. (2017). Coupling Proteins in Type IV Secretion. Type IV Secretion in Gram-Negative and Gram-Positive Bacteria. S. Backert and E. Grohmann. Cham, Springer International Publishing**:** 143-168.

Llosa, M., Gomis-Rüth, F. X., Coll, M. and de la Cruz Fd, F. (2002). "Bacterial conjugation: a two-step mechanism for DNA transport." Mol Microbiol **45**(1): 1-8.

Lloyd-Price, J., Arze, C., Ananthakrishnan, A. N., Schirmer, M., Avila-Pacheco, J., Poon, T. W., Andrews, E., Ajami, N. J., Bonham, K. S., Brislawn, C. J., Casero, D., Courtney, H., Gonzalez, A., Graeber, T. G., Hall, A. B., Lake, K., Landers, C. J., Mallick, H., Plichta, D. R., Prasad, M., Rahnavard, G., Sauk, J., Shungin, D., Vázquez-Baeza, Y., White, R. A., 3rd, Braun, J., Denson, L. A., Jansson, J. K., Knight, R., Kugathasan, S., McGovern, D. P. B., Petrosino, J. F., Stappenbeck, T. S., Winter, H. S., Clish, C. B., Franzosa, E. A., Vlamakis, H., Xavier, R. J. and Huttenhower, C. (2019). "Multi-omics of the gut microbial ecosystem in inflammatory bowel diseases." Nature **569**(7758): 655-662.

Lockwood, D. C., Amin, H., Costa, T. R. D. and Schroeder, G. N. (2022). "The Legionella pneumophila Dot/Icm type IV secretion system and its effectors." Microbiology **168**(5).

Low, H. H., Gubellini, F., Rivera-Calzada, A., Braun, N., Connery, S., Dujeancourt, A., Lu, F., Redzej, A., Fronzes, R., Orlova, E. V. and Waksman, G. (2014). "Structure of a type IV secretion system." Nature **508**(7497): 550-553.

Low, W. W., Wong, J. L. C., Beltran, L. C., Seddon, C., David, S., Kwong, H. S., Bizeau, T., Wang, F., Peña, A., Costa, T. R. D., Pham, B., Chen, M., Egelman, E. H., Beis, K. and Frankel, G. (2022). "Mating pair stabilisation mediates bacterial conjugation species specificity." Nat Microbiol **7**(7): 1016-1027.

Lu, J., Wong, J. J., Edwards, R. A., Manchak, J., Frost, L. S. and Glover, J. N. (2008a). "Structural basis of specific TraD-TraM recognition during F plasmid-mediated bacterial conjugation." Mol Microbiol **70**(1): 89-99.

Lu, J., Wong, J. J. W., Edwards, R. A., Manchak, J., Frost, L. S. and Glover, J. N. M. (2008b). "Structural basis of specific TraD–TraM recognition during F plasmid-mediated bacterial conjugation." Molecular Microbiology **70**(1): 89-99.

Lu, S., Giuliani, M., Harvey, H., Burrows, L. L., Wickham, R. A. and Dutcher, J. R. (2015). "Nanoscale Pulling of Type IV Pili Reveals Their Flexibility and Adhesion to Surfaces over Extended Lengths of the Pili." Biophys J **108**(12): 2865-2875.

Macé, K., Vadakkepat, A. K., Redzej, A., Lukoyanova, N., Oomen, C., Braun, N., Ukleja, M., Lu, F., Costa, T. R. D., Orlova, E. V., Baker, D., Cong, Q. and Waksman, G. (2022). "Cryo-EM structure of a type IV secretion system." Nature **607**(7917): 191-196.

Maier, J. A., Martinez, C., Kasavajhala, K., Wickstrom, L., Hauser, K. E. and Simmerling, C. (2015). "ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB." J Chem Theory Comput **11**(8): 3696-3713.

Majdalani, N. and Ippen-Ihler, K. (1996). "Membrane insertion of the F-pilin subunit is Sec independent but requires leader peptidase B and the proton motive force." Journal of bacteriology **178**(13): 3742-3747.

Majdalani, N., Moore, D., Maneewannakul, S. and Ippen-Ihler, K. (1996). "Role of the propilin leader peptide in the maturation of F pilin." Journal of bacteriology **178**(13): 3748-3754.

Mansour, S., Asrar, T. and Elhenawy, W. (2023). "The multifaceted virulence of adherent-invasive Escherichia coli." Gut Microbes **15**(1): 2172669.

Manwaring, N. P., Skurray, R. A. and Firth, N. (1999). "Nucleotide sequence of the F plasmid leading region." Plasmid **41**(3): 219-225.

Marrero, J. and Waldor, M. K. (2007). "Determinants of Entry Exclusion within Eex and TraG Are Cytoplasmic." Journal of Bacteriology **189**(17): 6469-6473.

Martinez-Medina, M., Naves, P., Blanco, J., Aldeguer, X., Blanco, J. E., Blanco, M., Ponte, C., Soriano, F., Darfeuille-Michaud, A. and Garcia-Gil, L. J. (2009). "Biofilm formation as a novel phenotypic feature of adherent-invasive Escherichia coli (AIEC)." BMC Microbiol **9**: 202.

Masai, H. and Arai, K.-i. (1997). "Frpo: a novel single-stranded DNA promoter for transcription and for primer RNA synthesis of DNA replication." Cell **89**(6): 897-907.

May, T. and Okabe, S. (2008). "Escherichia coli Harboring a Natural IncF Conjugative F Plasmid Develops Complex Mature Biofilms by Stimulating Synthesis of Colanic Acid and Curli." Journal of Bacteriology **190**(22): 7479-7490.

May, T., Tsuruta, K. and Okabe, S. (2011). "Exposure of conjugative plasmid carrying Escherichia coli biofilms to male-specific bacteriophages." The ISME Journal **5**(4): 771-775.

Meir, A., Macé, K., Lukoyanova, N., Chetrit, D., Hospenthal, M. K., Redzej, A., Roy, C. and Waksman, G. (2020). "Mechanism of effector capture and delivery by the type IV secretion system from Legionella pneumophila." Nature communications **11**(1): 2864.

Merz, A. J., So, M. and Sheetz, M. P. (2000). "Pilus retraction powers bacterial twitching motility." Nature **407**(6800): 98-102.

Miller, J. M. and Enemark, E. J. (2016). "Fundamental Characteristics of AAA+ Protein Family Structure and Function." Archaea **2016**: 9294307.

Mirdita, M., Schütze, K., Moriwaki, Y., Heo, L., Ovchinnikov, S. and Steinegger, M. (2022). "ColabFold: making protein folding accessible to all." Nature Methods **19**(6): 679-682.

Mok, M. and Marians, K. (1987). "The Escherichia coli preprimosome and DNA B helicase can form replication forks that move at the same rate." Journal of Biological Chemistry **262**(34): 16644-16654.

Moncalián, G., Cabezón, E., Alkorta, I., Valle, M., Moro, F., Valpuesta, J. M. a., Goñi, F. M. and de la Cruz, F. (1999). "Characterization of ATP and DNA Binding Activities of TrwB, the Coupling Protein Essential in Plasmid R388 Conjugation *." Journal of Biological Chemistry **274**(51): 36117-36124.

Moore, D., Hamilton, C. M., Maneewannakul, K., Mintz, Y., Frost, L. S. and Ippen-Ihler, K. (1993). "The Escherichia coli K-12 F plasmid gene traX is required for acetylation of F pilin." J Bacteriol **175**(5): 1375-1383.

Moran, R. A., Anantham, S., Pinyon, J. L. and Hall, R. M. (2015). "Plasmids in antibiotic susceptible and antibiotic resistant commensal Escherichia coli from healthy Australian adults." Plasmid **80**: 24-31.

Mortezaei, N., Singh, B., Zakrisson, J., Bullitt, E. and Andersson, M. (2015). "Biomechanical and structural features of CS2 fimbriae of enterotoxigenic Escherichia coli." Biophys J **109**(1): 49-56.

Mshana, S. E., Hain, T., Domann, E., Lyamuya, E. F., Chakraborty, T. and Imirzalioglu, C. (2013). "Predominance of Klebsiella pneumoniae ST14 carrying CTX-M-15 causing neonatal sepsis in Tanzania." BMC Infect Dis **13**: 466.

Mukherjee, S. K. and Mukherjee, M. (2019). "Characterization and Bio-Typing of Multidrug Resistance Plasmids From Uropathogenic Escherichia coli Isolated From Clinical Setting." Frontiers in Microbiology **10**.

Mulani, M. S., Kamble, E. E., Kumkar, S. N., Tawre, M. S. and Pardesi, K. R. (2019). "Emerging strategies to combat ESKAPE pathogens in the era of antimicrobial resistance: a review." Frontiers in microbiology **10**: 539.

Murray, C. J. L., Ikuta, K. S., Sharara, F., Swetschinski, L., Robles Aguilar, G., Gray, A., Han, C., Bisignano, C., Rao, P., Wool, E., Johnson, S. C., Browne, A. J., Chipeta, M. G., Fell, F., Hackett, S., Haines-Woodhouse, G., Kashef Hamadani, B. H., Kumaran, E. A. P., McManigal, B., Agarwal, R., Akech, S., Albertson, S., Amuasi, J., Andrews, J., Aravkin, A., Ashley, E., Bailey, F., Baker, S., Basnyat, B., Bekker, A., Bender, R., Bethou, A., Bielicki, J., Boonkasidecha, S., Bukosia, J., Carvalheiro, C., Castañeda-Orjuela, C., Chansamouth, V., Chaurasia, S., Chiurchiù, S., Chowdhury, F., Cook, A. J., Cooper, B., Cressey, T. R., Criollo-Mora, E., Cunningham, M., Darboe, S., Day, N. P. J., De Luca, M., Dokova, K., Dramowski, A., Dunachie, S. J., Eckmanns, T., Eibach, D., Emami, A., Feasey, N., Fisher-Pearson, N., Forrest, K., Garrett, D., Gastmeier, P., Giref, A. Z., Greer, R. C., Gupta, V., Haller, S., Haselbeck, A., Hay, S. I., Holm, M., Hopkins, S., Iregbu, K. C., Jacobs, J., Jarovsky, D., Javanmardi, F., Khorana, M., Kissoon, N., Kobeissi, E., Kostyanev, T., Krapp, F., Krumkamp, R., Kumar, A., Kyu, H. H., Lim, C., Limmathurotsakul, D., Loftus, M. J., Lunn, M., Ma, J., Mturi, N., Munera-Huertas, T., Musicha, P., Mussi-Pinhata, M. M., Nakamura, T., Nanavati, R., Nangia, S., Newton, P., Ngoun, C., Novotney, A., Nwakanma, D., Obiero, C. W., Olivas-Martinez, A., Olliaro, P., Ooko, E., Ortiz-Brizuela, E., Peleg, A. Y., Perrone, C., Plakkal, N., Ponce-de-Leon, A., Raad, M., Ramdin, T., Riddell, A., Roberts, T., Robotham, J. V., Roca, A., Rudd, K. E., Russell, N., Schnall, J., Scott, J. A. G., Shivamallappa, M., Sifuentes-Osornio, J., Steenkeste, N., Stewardson, A. J., Stoeva, T., Tasak, N., Thaiprakong, A., Thwaites, G., Turner, C., Turner, P., van Doorn, H. R., Velaphi, S., Vongpradith, A., Vu, H., Walsh, T., Waner, S., Wangrangsimakul, T., Wozniak, T., Zheng, P., Sartorius, B., Lopez, A. D., Stergachis, A., Moore, C., Dolecek, C. and Naghavi, M. (2022). "Global burden of bacterial antimicrobial resistance in 2019: a systematic analysis." The Lancet **399**(10325): 629-655.

Nagai, H. and Kubori, T. (2011). "Type IVB Secretion Systems of Legionella and Other Gram-Negative Bacteria." Frontiers in Microbiology **2**.

Narciso, A., Nunes, F., Amores, T., Lito, L., Melo-Cristino, J. and Duarte, A. (2012). "Persistence of uropathogenic Escherichia coli strains in the host for long periods of time: relationship between phylogenetic groups and virulence factors." European Journal of Clinical Microbiology & Infectious Diseases **31**(6): 1211-1217.

NHS. (2023). "NHS steps up battle against life-threatening infections following successful world-first pilot. https://www.england.nhs.uk/2023/07/nhs-steps-up-battle-against-life-threatening-infections-following-successful-world-first-pilot/."

Nicolas, E., Lambin, M., Dandoy, D., Galloy, C., Nguyen, N., Oger, C. A. and Hallet, B. (2015). "The Tn3 - family of replicative transposons." Mobile DNA III: 693-726.

Nivaskumar, M. and Francetic, O. (2014). "Type II secretion system: a magic beanstalk or a protein escalator." Biochim Biophys Acta **1843**(8): 1568-1577.

Novick, R. P. (1987). "Plasmid incompatibility." Microbiological reviews **51**(4): 381-395.

Odakura, Y., Hashimoto, H. and Mitsuhashi, S. (1977). "Temperature sensitive R plasmids isolated from Proteus strains." Microbiology and Immunology **21**(11): 621-629.

Ogura, T., Whiteheart, S. W. and Wilkinson, A. J. (2004). "Conserved arginine residues implicated in ATP hydrolysis, nucleotide-sensing, and inter-subunit interactions in AAA and AAA+ ATPases." Journal of structural biology **146**(1-2): 106-112.

Olaitan, A. O. and Rolain, J. M. (2016). "Ancient Resistome." Microbiol Spectr **4**(4).

Olsen, R. H. and Shipley, P. L. (1975). "RP1 properties and fertility inhibition among P, N, W, and X incompatibility group plasmids." Journal of Bacteriology **123**(1): 28-35.

Ott, L. C. and Mellata, M. (2022). "Models for Gut-Mediated Horizontal Gene Transfer by Bacterial Plasmid Conjugation." Front Microbiol **13**: 891548.

Ourmazd, A., Moffat, K. and Lattman, E. E. (2022). "Structural biology is solved—now what?" Nature methods **19**(1): 24-26.

Palmela, C., Chevarin, C., Xu, Z., Torres, J., Sevrin, G., Hirten, R., Barnich, N., Ng, S. C. and Colombel, J. F. (2018). "Adherent-invasive Escherichia coli in inflammatory bowel disease." Gut **67**(3): 574-587.

Panicker, M. M. and Minkley, E. G., Jr. (1992). "Purification and properties of the F sex factor TraD protein, an inner membrane conjugal transfer protein." J Biol Chem **267**(18): 12761-12766.

Partridge, S. R., Kwong, S. M., Firth, N. and Jensen, S. O. (2018). "Mobile Genetic Elements Associated with Antimicrobial Resistance." Clinical Microbiology Reviews **31**(4): 10.1128/cmr.00088-00017.

Patkowski, J. B., Dahlberg, T., Amin, H., Gahlot, D. K., Vijayrajratnam, S., Vogel, J. P., Francis, M. S., Baker, J. L., Andersson, M. and Costa, T. R. D. (2023). "The F-pilus biomechanical adaptability accelerates conjugative dissemination of antimicrobial resistance and biofilm formation." Nature Communications **14**(1): 1879.

Penczek, P. A. (2010). "Resolution measures in molecular electron microscopy." Methods Enzymol **482**: 73-100.

Petrova, V., Chitteni-Pattu, S., Drees, J. C., Inman, R. B. and Cox, M. M. (2009). "An SOS inhibitor that binds to free RecA protein: the PsiB protein." Molecular cell **36**(1): 121-130.

Pitzschke, A. and Hirt, H. (2010). "New insights into an old story: Agrobacterium-induced tumour formation in plants by plant transformation." Embo j **29**(6): 1021-1032.

Prudent, V., Demarre, G., Vazeille, E., Wery, M., Quenech'Du, N., Ravet, A., Dauverd - Girault, J., van Dijk, E., Bringer, M.-A., Descrimes, M., Barnich, N., Rimsky, S., Morillon, A. and Espéli, O. (2021). "The Crohn's disease-related bacterial strain LF82 assembles biofilm-like communities to protect itself from phagolysosomal attack." Communications Biology **4**(1): 627.

Punjani, A., Rubinstein, J. L., Fleet, D. J. and Brubaker, M. A. (2017). "cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination." Nature Methods **14**(3): 290-296.

Purtschert-Montenegro, G., Cárcamo-Oyarce, G., Pinto-Carbó, M., Agnoli, K., Bailly, A. and Eberl, L. (2022). "Pseudomonas putida mediates bacterial killing, biofilm invasion and biocontrol with a type IVB secretion system." Nat Microbiol **7**(10): 1547-1557.

Quan, J., Li, X., Chen, Y., Jiang, Y., Zhou, Z., Zhang, H., Sun, L., Ruan, Z., Feng, Y. and Akova, M. (2017). "Prevalence of mcr-1 in Escherichia coli and Klebsiella pneumoniae recovered from bloodstream infections in China: a multicentre longitudinal study." The Lancet Infectious Diseases **17**(4): 400-410.

Ramakrishnan, C., Dani, V. S. and Ramasarma, T. (2002). "A conformational analysis of Walker motif A [GXXXXGKT (S)] in nucleotide-binding and other proteins." Protein Eng **15**(10): 783-798.

Ramsey, M. E., Woodhams, K. L. and Dillard, J. P. (2011). "The Gonococcal Genetic Island and Type IV Secretion in the Pathogenic Neisseria." Frontiers in Microbiology **2**.

Rashkova, S., Spudich, G. M. and Christie, P. J. (1997). "Characterization of membrane and protein interaction determinants of the Agrobacterium tumefaciens VirB11 ATPase." Journal of Bacteriology **179**(3): 583-591.

Ravatn, R., Studer, S., Zehnder, A. J. and Roelof van der Meer, J. (1998). "Int-B13, an unusual site-specific recombinase of the bacteriophage P4 integrase family, is responsible for chromosomal insertion of the 105-kilobase clc element of Pseudomonas sp. strain B13." Journal of bacteriology **180**(21): 5505-5514.

Redfield, R. J., Cameron, A. D., Qian, Q., Hinds, J., Ali, T. R., Kroll, J. S. and Langford, P. R. (2005). "A novel CRP-dependent regulon controls expression of competence genes in Haemophilus influenzae." Journal of molecular biology **347**(4): 735-747.

Redzej, A., Ukleja, M., Connery, S., Trokter, M., Felisberto-Rodrigues, C., Cryar, A., Thalassinos, K., Hayward, R. D., Orlova, E. V. and Waksman, G. (2017). "Structure of a VirD4 coupling protein bound to a VirB type IV secretion machinery." The EMBO Journal **36**(20): 3080-3095.

Reisner, A., Haagensen, J. A. J., Schembri, M. A., Zechner, E. L. and Molin, S. (2003). "Development and maturation of Escherichia coli K-12 biofilms." Molecular Microbiology **48**(4): 933-946.

Reisner, A., Krogfelt, K. A., Klein, B. M., Zechner, E. L. and Molin, S. (2006). "In Vitro Biofilm Formation of Commensal and Pathogenic Escherichia coli Strains: Impact of Environmental and Genetic Factors." Journal of Bacteriology **188**(10): 3572-3581.

Reyes, V. E. (2023). "Helicobacter pylori and Its Role in Gastric Cancer." Microorganisms **11**(5).

Reygaert, W. C. (2018). "An overview of the antimicrobial resistance mechanisms of bacteria." AIMS Microbiol **4**(3): 482-501.

Ripoll-Rozada, J., Zunzunegui, S., Cruz, F. d. l., Arechaga, I. and Cabezón, E. (2013). "Functional Interactions of VirB11 Traffic ATPases with VirB4 and VirD4 Molecular Motors in Type IV Secretion Systems." Journal of Bacteriology **195**(18): 4195-4201.

Rivera-Calzada, A., Fronzes, R., Savva, C. G., Chandran, V., Lian, P. W., Laeremans, T., Pardon, E., Steyaert, J., Remaut, H., Waksman, G. and Orlova, E. V. (2013). "Structure of a bacterial type IV secretion core complex at subnanometre resolution." The EMBO Journal **32**(8): 1195-1204.

Rozwandowicz, M., Brouwer, M., Fischer, J., Wagenaar, J., Gonzalez-Zorn, B., Guerra, B., Mevius, D. J. and Hordijk, J. (2018). "Plasmids carrying antimicrobial resistance genes in Enterobacteriaceae." Journal of Antimicrobial Chemotherapy **73**(5): 1121-1137.

Ruiz, E., Sáenz, Y., Zarazaga, M., Rocha-Gracia, R., Martínez-Martínez, L., Arlet, G. and Torres, C. (2012). "qnr, aac(6')-Ib-cr and qepA genes in Escherichia coli and Klebsiella spp.: genetic environments and plasmid and chromosomal location." J Antimicrob Chemother **67**(4): 886-897.

Ryckaert, J.-P., Ciccotti, G. and Berendsen, H. J. (1977). "Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes." Journal of computational physics **23**(3): 327-341.

Sagulenko, E., Sagulenko, V., Chen, J. and Christie, P. J. (2001). "Role of Agrobacterium VirB11 ATPase in T-Pilus Assembly and Substrate Selection." Journal of Bacteriology **183**(20): 5813-5825.

Salomon-Ferrer, R., Gotz, A. W., Poole, D., Le Grand, S. and Walker, R. C. (2013). "Routine microsecond molecular dynamics simulations with AMBER on GPUs. 2. Explicit solvent particle mesh Ewald." Journal of chemical theory and computation **9**(9): 3878-3888.

Sasisekharan, R. (1963). "Stereochemistry of polypeptide chain configurations J Mol Biol 79599. Ramachandran GN, Ramakrishnan C, and Sasisekharan (1963). Stereochemistry of polypeptide chain configurations." J Mol Biol **7**: 95-99.

Savvides, S. N., Yeo, H. J., Beck, M. R., Blaesing, F., Lurz, R., Lanka, E., Buhrdorf, R., Fischer, W., Haas, R. and Waksman, G. (2003). "VirB11 ATPases are dynamic hexameric assemblies: new insights into bacterial type IV secretion." The EMBO journal.

Scheres, S. H. (2012). "RELION: implementation of a Bayesian approach to cryo-EM structure determination." J Struct Biol **180**(3): 519-530.

Segura, R. L., Águila-Arcos, S., Ugarte-Uribe, B., Vecino, A. J., de la Cruz, F., Goñi, F. M. and Alkorta, I. (2014). "Subcellular location of the coupling protein TrwB and the role of its transmembrane domain." Biochimica et Biophysica Acta (BBA) - Biomembranes **1838**(1, Part B): 223-230.

Serfiotis-Mitsa, D., Herbert, A. P., Roberts, G. A., Soares, D. C., White, J. H., Blakely, G. W., Uhrín, D. and Dryden, D. T. (2010). "The structure of the KlcA and ArdB proteins reveals a novel fold and antirestriction activity against Type I DNA restriction systems in vivo but not in vitro." Nucleic acids research **38**(5): 1723-1737.

Sgro, G. G. and Costa, T. R. D. (2018). "Cryo-EM Grid Preparation of Membrane Protein Samples for Single Particle Analysis." Front Mol Biosci **5**: 74.

Sgro, G. G., Costa, T. R. D., Cenens, W., Souza, D. P., Cassago, A., Coutinho de Oliveira, L., Salinas, R. K., Portugal, R. V., Farah, C. S. and Waksman, G. (2018). "Cryo-EM structure of the bacteria-killing type IV secretion system core complex from Xanthomonas citri." Nat Microbiol **3**(12): 1429-1440.

Sgro, G. G., Oka, G. U., Souza, D. P., Cenens, W., Bayer-Santos, E., Matsuyama, B. Y., Bueno, N. F., Dos Santos, T. R., Alvarez-Martinez, C. E. and Salinas, R. K. (2019). "Bacteria-killing type IV secretion systems." Frontiers in microbiology **10**: 460432.

Shankar, C., Vasudevan, K., Jacob, J. J., Baker, S., Isaac, B. J., Neeravi, A. R., Sethuvel, D. P. M., George, B. and Veeraraghavan, B. (2022). "Hybrid Plasmids Encoding Antimicrobial Resistance and Virulence Traits Among Hypervirulent Klebsiella pneumoniae ST2096 in India." Front Cell Infect Microbiol **12**: 875116.

Sheedlo, M. J., Durie, C. L., Chung, J. M., Chang, L., Roberts, J., Swanson, M., Lacy, D. B. and Ohi, M. D. (2021). "Cryo-EM reveals new species-specific proteins and symmetry elements in the Legionella pneumophila Dot/Icm T4SS." eLife **10**: e70427.

Sheedlo, M. J., Ohi, M. D., Lacy, D. B. and Cover, T. L. (2022). "Molecular architecture of bacterial type IV secretion systems." PLoS Pathog **18**(8): e1010720.

Shen, Z., Tang, C. M. and Liu, G.-Y. (2022). "Towards a better understanding of antimicrobial resistance dissemination: what can be learnt from studying model conjugative plasmids?" Military Medical Research **9**(1): 3.

Sheth, L. K., Piotrowski, A. L. and Voss, N. R. (2015). "Visualization and quality assessment of the contrast transfer function estimation." Journal of Structural Biology **192**(2): 222-234.

Shiomi, D., Mori, H. and Niki, H. (2009). "Genetic mechanism regulating bacterial cell shape and metabolism." Commun Integr Biol **2**(3): 219-220.

Shirani, M., Pakzad, R., Haddadi, M. H., Akrami, S., Asadi, A., Kazemian, H., Moradi, M., Kaviar, V. H., Zomorodi, A. R., Khoshnood, S., Shafieian, M., Tavasolian, R., Heidary, M. and Saki, M. (2023). "The global prevalence of gastric cancer in Helicobacter pylori-infected individuals: a systematic review and meta-analysis." BMC Infectious Diseases **23**(1): 543.

Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Söding, J., Thompson, J. D. and Higgins, D. G. (2011) "Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega." Molecular systems biology **7**, 539 DOI: 10.1038/msb.2011.75.

Silver, R. P. and Cohen, S. N. (1972). "Nonchromosomal antibiotic resistance in bacteria V. Isolation and characterization of R factor mutants exhibiting temperature-sensitive repression of fertility." Journal of Bacteriology **110**(3): 1082-1088.

Smillie, C., Garcillán-Barcia, M. P., Francia, M. V., Rocha, E. P. C. and Cruz, F. d. l. (2010). "Mobility of Plasmids." Microbiology and Molecular Biology Reviews **74**(3): 434-452.

Souza, D. P., Oka, G. U., Alvarez-Martinez, C. E., Bisson-Filho, A. W., Dunger, G., Hobeika, L., Cavalcante, N. S., Alegria, M. C., Barbosa, L. R., Salinas, R. K., Guzzo, C. R. and Farah, C. S. (2015). "Bacterial killing via a type IV secretion system." Nat Commun **6**: 6453.

Stachurska, X., Roszak, M., Jabłońska, J., Mizielińska, M. and Nawrotek, P. (2021). "Double-Layer Agar (DLA) Modifications for the First Step of the Phage-Antibiotic Synergy (PAS) Identification." Antibiotics (Basel) **10**(11).

Stephens, C., Arismendi, T., Wright, M., Hartman, A., Gonzalez, A., Gill, M., Pandori, M. and Hess, D. (2020). "F Plasmids Are the Major Carriers of Antibiotic Resistance Genes in Human-Associated Commensal Escherichia coli." mSphere **5**(4).

Stephens, K. M., Roush, C. and Nester, E. (1995). "Agrobacterium tumefaciens VirB11 protein requires a consensus nucleotide-binding site for function in virulence." J Bacteriol **177**(1): 27-36.

Subramaniam, S. and Kleywegt, G. J. (2022). "A paradigm shift in structural biology." Nature Methods **19**(1): 20-23.

Tacconelli, E., Carrara, E., Savoldi, A., Harbarth, S., Mendelson, M., Monnet, D. L., Pulcini, C., Kahlmeter, G., Kluytmans, J. and Carmeli, Y. (2018). "Discovery, research, and development of new antibiotics: the WHO priority list of antibiotic-resistant bacteria and tuberculosis." The Lancet infectious diseases **18**(3): 318-327.

Tamang, M. D., Seol, S. Y., Oh, J.-Y., Kang, H. Y., Lee, J. C., Lee, Y. C., Cho, D. T. and Kim, J. (2008). "Plasmid-Mediated Quinolone Resistance Determinants qnrA, qnrB, and qnrS among Clinical Isolates of Enterobacteriaceae in a Korean Hospital." Antimicrobial Agents and Chemotherapy **52**(11): 4159-4162.

Terwilliger, T. C., Liebschner, D., Croll, T. I., Williams, C. J., McCoy, A. J., Poon, B. K., Afonine, P. V., Oeffner, R. D., Richardson, J. S., Read, R. J. and Adams, P. D. (2024). "AlphaFold predictions are valuable hypotheses and accelerate but do not replace experimental structure determination." Nature Methods **21**(1): 110-116.

Thomas, C. M. and Nielsen, K. M. (2005). "Mechanisms of, and barriers to, horizontal gene transfer between bacteria." Nature reviews microbiology **3**(9): 711-721.

Thornton, J. M., Laskowski, R. A. and Borkakoti, N. (2021). "AlphaFold heralds a data-driven revolution in biology and medicine." Nature Medicine **27**(10): 1666-1669.

Uchański, T., Masiulis, S., Fischer, B., Kalichuk, V., López-Sánchez, U., Zarkadas, E., Weckener, M., Sente, A., Ward, P., Wohlkönig, A., Zögg, T., Remaut, H., Naismith, J. H., Nury, H., Vranken, W., Aricescu, A. R., Pardon, E. and Steyaert, J. (2021). "Megabodies expand the nanobody toolkit for protein structure determination by single-particle cryo-EM." Nature Methods **18**(1): 60-68.

UN (2023). "Environment Programme. Bracing for Superbugs: Strengthening environmental action in the One Health response to antimicrobial resistance."

Usein, C. R., Papagheorghe, R., Oprea, M., Condei, M. and Strãuţ, M. (2016). "Molecular characterization of bacteremic Escherichia coli isolates in Romania." Folia Microbiol (Praha) **61**(3): 221-226.

Van Hootegem, P. and Travis, S. (2020). "Is Crohn's disease a rightly used eponym?" Journal of Crohn's and Colitis **14**(6): 867-871.

Veith, P. D., Glew, M. D., Gorasia, D. G. and Reynolds, E. C. (2017). "Type IX secretion: the generation of bacterial cell surface coatings involved in virulence, gliding motility and the degradation of complex biopolymers." Molecular microbiology **106**(1): 35-53.

Vestby, L. K., Grønseth, T., Simm, R. and Nesse, L. L. (2020). "Bacterial Biofilm and its Role in the Pathogenesis of Disease." Antibiotics (Basel) **9**(2).

Viasus, D., Gaia, V., Manzur-Barbur, C. and Carratalà, J. (2022). "Legionnaires' Disease: Update on Diagnosis and Treatment." Infect Dis Ther **11**(3): 973-986.

Virolle, C., Goldlust, K., Djermoun, S., Bigot, S. and Lesterlin, C. (2020). "Plasmid transfer by conjugation in Gram-negative bacteria: from the cellular to the community level." Genes **11**(11): 1239.

Waksman, G. (2019). "From conjugation to T4S systems in Gram-negative bacteria: a mechanistic biology perspective." EMBO Rep **20**(2).

Wang, M., Goh, Y. X., Tai, C., Wang, H., Deng, Z. and Ou, H. Y. (2022). "VRprofile2: detection of antibiotic resistance-associated mobilome in bacterial pathogens." Nucleic Acids Res **50**(W1): W768-w773.

Wang, R., Van Dorp, L., Shaw, L. P., Bradley, P., Wang, Q., Wang, X., Jin, L., Zhang, Q., Liu, Y. and Rieux, A. (2018). "The global distribution and spread of the mobilized colistin resistance gene mcr-1." Nature communications **9**(1): 1179.

Wang, Y. A., Yu, X., Silverman, P. M., Harris, R. L. and Egelman, E. H. (2009). "The Structure of F-Pili." Journal of Molecular Biology **385**(1): 22-29.

Wedel, E., Bernabe-Balas, C., Ares-Arroyo, M., Montero, N., Santos-Lopez, A., Mazel, D. and Gonzalez-Zorn, B. (2023). "Insertion Sequences Determine Plasmid Adaptation to New Bacterial Hosts." mBio **14**(3): e0315822.

Wentinck, K., Gogou, C. and Meijer, D. H. (2022). "Putting on molecular weight: Enabling cryo-EM structure determination of sub-100-kDa proteins." Curr Res Struct Biol **4**: 332-337.

Whitaker, N., Chen, Y., Jakubowski, S. J., Sarkar, M. K., Li, F., Christie, P. J. and Parkinson, J. S. (2015). "The All-Alpha Domains of Coupling Proteins from the Agrobacterium tumefaciens VirB/VirD4 and Enterococcus faecalis pCF10-Encoded Type IV Secretion Systems Confer Specificity to Binding of Cognate DNA Substrates." Journal of Bacteriology **197**(14): 2335-2349.

Whittle, G., Shoemaker, N. and Salyers, A. (2002). "The role of Bacteroides conjugative transposons in the dissemination of antibiotic resistance genes." Cellular and Molecular Life Sciences CMLS **59**: 2044-2054.

WHO (2017). "Prioritization of pathogens to guide discovery, research and development of new antibiotics for drug-resistant bacterial infections, including tuberculosis.".

WHO. (2021). "Call to Action on Antimicrobial Resistance." 2024, from https://www.who.int/news/item/30-07-2021-call-to-action-on-antimicrobial-resistance-2021.

Williams, C. J., Headd, J. J., Moriarty, N. W., Prisant, M. G., Videau, L. L., Deis, L. N., Verma, V., Keedy, D. A., Hintze, B. J., Chen, V. B., Jain, S., Lewis, S. M., Arendall, W. B., 3rd, Snoeyink, J., Adams, P. D., Lovell, S. C., Richardson, J. S. and Richardson, D. C. (2018). "MolProbity: More and better reference data for improved all-atom structure validation." Protein Sci **27**(1): 293-315.

Wong, J. J., Lu, J. and Glover, J. N. (2012a). "Relaxosome function and conjugation regulation in F-like plasmids - a structural biology perspective." Mol Microbiol **85**(4): 602-617.

Wong, J. J. W., Lu, J. and Glover, J. N. M. (2012b). "Relaxosome function and conjugation regulation in F-like plasmids – a structural biology perspective." Molecular Microbiology **85**(4): 602-617.

Woodford, N., Carattoli, A., Karisik, E., Underwood, A., Ellington, M. J. and Livermore, D. M. (2009). "Complete nucleotide sequences of plasmids pEK204, pEK499, and pEK516, encoding CTX-M enzymes in three major Escherichia coli lineages from the United Kingdom, all belonging to the international O25: H4-ST131 clone." Antimicrobial Agents and Chemotherapy **53**(10): 4472-4482.

Wozniak, R. A. and Waldor, M. K. (2010). "Integrative and conjugative elements: mosaic mobile genetic elements enabling dynamic lateral gene flow." Nature Reviews Microbiology **8**(8): 552-563.

Wu, W., Feng, Y., Tang, G., Qiao, F., McNally, A. and Zong, Z. (2019). "NDM metallo-β-lactamases and their bacterial producers in health care settings." Clinical microbiology reviews **32**(2): 10.1128/cmr. 00115-00118.

Wu, X., Zhao, Y., Zhang, H., Yang, W., Yang, J., Sun, L., Jiang, M., Wang, Q., Wang, Q., Ye, X., Zhang, X. and Wu, Y. (2023). "Mechanism of regulation of the Helicobacter pylori Cagβ ATPase by CagZ." Nature Communications **14**(1): 479.

Yang, Q. E., Sun, J., Li, L., Deng, H., Liu, B. T., Fang, L. X., Liao, X. P. and Liu, Y. H. (2015). "IncF plasmid diversity in multi-drug resistant Escherichia coli strains from animals in China." Front Microbiol **6**: 964.

Yang, X., Dong, N., Chan, E. W.-C., Zhang, R. and Chen, S. (2021). "Carbapenem resistance-encoding and virulence-encoding conjugative plasmids in Klebsiella pneumoniae." Trends in microbiology **29**(1): 65-83.

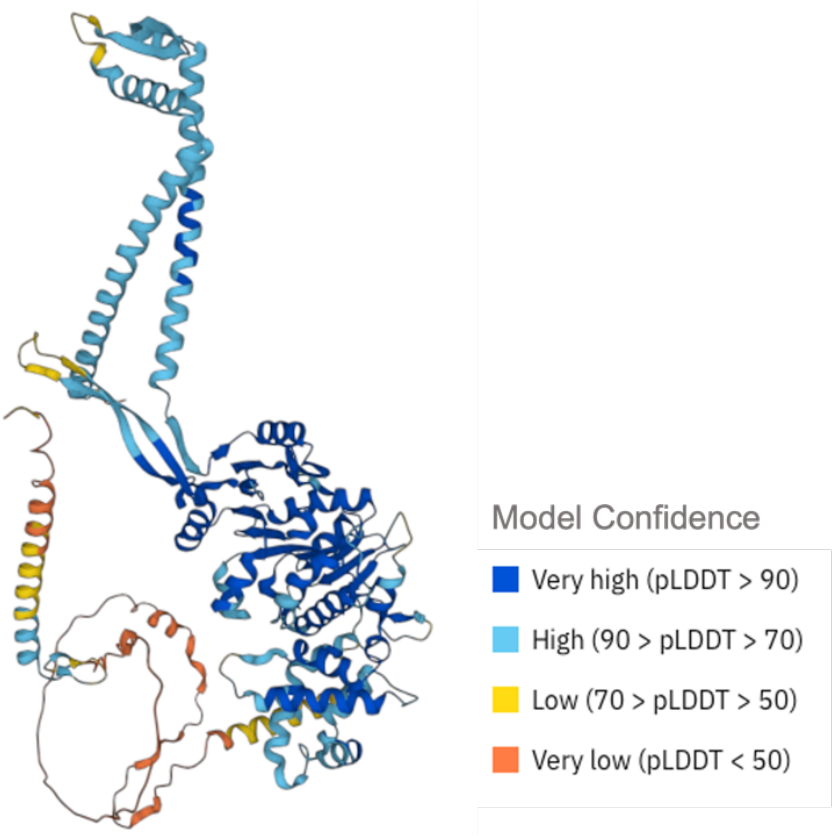
Yeates, T. O., Agdanowski, M. P. and Liu, Y. (2020). "Development of imaging scaffolds for cryo-electron microscopy." Curr Opin Struct Biol **60**: 142-149.

Yeo, H.-J., Savvides, S. N., Herr, A. B., Lanka, E. and Waksman, G. (2000). "Crystal Structure of the Hexameric Traffic ATPase of the Helicobacter pylori Type IV Secretion System." Molecular Cell **6**(6): 1461-1472.

Yoichi, H., Hiroshi, S. and Tohru, K. (1988). "Two single-strand DNA initiation signals located in the oriV region of plasmid RSF1010." Gene **68**(2): 221-228.

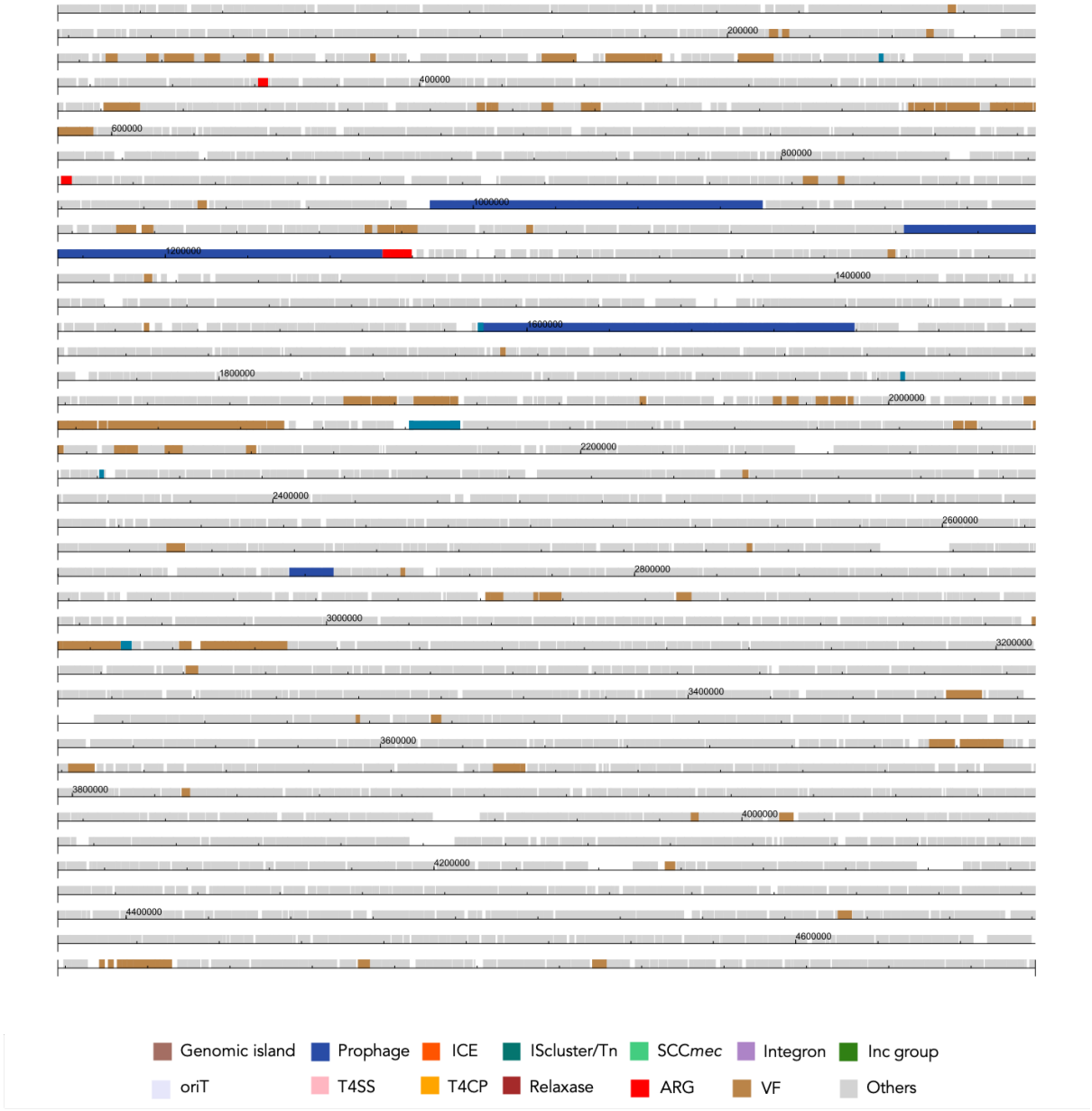
Zheng, W., Pena, A., Low, W. W., Wong, J. L. C., Frankel, G. and Egelman, E. H. (2020). "Cryoelectron-Microscopic Structure of the pKpQIL Conjugative Pili from Carbapenem-Resistant Klebsiella pneumoniae." Structure **28**(12): 1321-1328.e1322.

Supplementary Figures



Supplementary Figure 1. Confidence score of TraD prediction by AlphaFold-2.

Most of the TraD structure is predicted with very high (soluble domains) or high (transmembrane domain) confidence, with the exception of the unstructured, flexible C-terminal domain.



Supplementary Figure 2. MGE and virulence map of the NRG857c chromosome generated by VRprofile2 fails to detect any chromosomal T4SS.

Supplementary Tables

Supplementary Table 1. List of interacting residues between two adjacent TraD subunits

Hydrogen bonds

Subunit A	Distance [Å]	Subunit B
GLN 459	2.70	SER 471
SER 486	3.07	ASP 474
HIS 487	3.43	GLU 496
HIS 487	3.34	GLU 496
TYR 509	2.68	THR 522
TYR 511	2.75	VAL 520
ARG 517	2.24	SER 510
SER 533	2.56	GLU 499
TYR 534	3.63	GLU 499
TYR 534	3.61	GLY 495
ARG 597	2.47	GLY 280
ARG 597	2.84	THR 284
ASP 536	3.15	GLY 149
ALA 591	2.89	ASN 327
ASP 462	3.62	LYS 467
PRO 485	3.14	ARG 479
TYR 534	3.19	GLU 499
SER 535	2.72	GLU 499
TYR 511	2.54	VAL 520
TYR 509	2.53	THR 522
TYR 509	3.63	THR 522
THR 530	2.49	ARG 528
THR 530	2.93	ARG 528
GLU 609	3.16	ASP 606

Salt bridges

Subunit A	Distance [Å]	Subunit B
ARG 221	2.50	GLU 442
ARG 221	3.07	GLU 442
ARG 221	2.38	GLU 442
ARG 221	3.40	GLU 442
ASP 462	3.62	LYS 467