NOVA SCHOOL OF SCIENCE & TECHNOLOGY

DEPARTAMENT OF LIFE SCIENCES

CteG, a *Chlamydia trachomatis* protein involved in host cell lytic exit

Inês Isabel Serrano Pereira

Master in Bacterial Pathogenesis and Infection

Doctorate in Biology

NOVA School of Science and Technology

September, 2022

NOVA SCHOOL OF SCIENCE & TECHNOLOGY

DEPARTAMENT OF LIFE SCIENCES

CteG, a *Chlamydia trachomatis* protein involved in host cell lytic exit

Inês Isabel Serrano Pereira

Master in Bacterial Pathogenesis and Infection

Doctorate in Biology

NOVA School of Science and Technology

September, 2022

DEPARTAMENT OF
LIFE SCIENCES

CteG, a *Chlamydia trachomatis* protein involved in host cell lytic exit

Inês Isabel Serrano Pereira

Master in Bacterial Pathogenesis and Infection

Adviser: Luís Jaime Gomes Ferreira da Silva Mota
Assistant Professor, NOVA School of Science and Technology, Portugal

Examination Committee:

Chair: Isabel Maria Godinho de Sá Nogueira
Full Professor, NOVA School of Science and Technology, Portugal

Rapporteurs: Anja Lührmann
Principal Investigator, Friedrich-Alexander-Universität Erlangen-Nürnberg, Universitätsklinikum Erlangen, Germany

Barbara Susanne Sixt
Principal Investigator, Umeå Universitet, Sweden

Adviser: Luís Jaime Gomes Ferreira da Silva Mota
Assistant Professor, NOVA School of Science and Technology, Portugal

Members: Isabel Maria Godinho de Sá Nogueira
Full Professor, NOVA School of Science and Technology, Portugal

Elsa Maria Ribeiro dos Santos Anes
Associate Professor with Habilitation, Faculty of Pharmacy, University of Lisbon, Portugal

João Paulo dos Santos Gomes
Assistant Researcher with Habilitation, National Institute of Health Doutor Ricardo Jorge (INSA), Portugal

Doctorate in Biology

NOVA School of Science and Technology
September, 2022

CteG, a *Chlamydia trachomatis* protein involved in host cell lytic exit

I dedicate this PhD thesis to my father,

who passed away before I completed this journey.

ACKNOWLEDGMENTS

With all due respect to non-Portuguese readers, I will write this section in Portuguese to better express myself.

Com todo o respeito pelos leitores não portugueses, irei escrever esta secção na minha língua materna por nela ser mais fácil expressar-me.

Em primeiro lugar, gostaria de agradecer ao meu tutor e orientador, o professor Jaime Mota, por me receber mais que uma vez sem hesitar na sua equipa, pela sua dedicação e presença incansáveis, pelo profissionalismo e seriedade e pela ajuda incondicional durante a minha difícil jornada. Toda a sua conduta foi determinante para o meu crescimento enquanto pessoa, cientista e profissional. Além de ter estado presente direta e indiretamente em quase toda a extensão do meu percurso académico, foi ele quem marcou o ponto de partida da minha jornada pela microbiologia e me incutiu o gosto pelo estudo dos mecanismos moleculares da patogénese bacteriana. Estar-lhe-ei sempre grata pela paciência que teve comigo nos meus momentos mais difíceis, e pela confiança que depositou em mim.

Deixo uma palavra de agradecimento à minha instituição de acolhimento, a UCIBIO – FCT NOVA, e ao departamento de Ciências da Vida (DCV) por me acolherem ao longo dos últimos anos. Várias foram as pessoas com quem contactei durante todo o meu percurso, que se estendeu além do doutoramento, incluindo também enriquecedoras colaborações com outros laboratórios. Agradeço à professora Paula Gonçalves e ao Dr. João Rodrigues por aceitarem fazer parte da minha comissão de tese, assim como à professora Isabel Sá Nogueira pela coordenação do Programa Doutoral em Biologia e pelo seu apoio e tutoria. Agradeço ainda à Carla Gonçalves e à professora Paula Gonçalves por toda a sua ajuda e discussões sobre filogenia. À Nicole Soares, deixo também um agradecimento pela simpatia,

disponibilidade e dedicação, por ter o material de laboratório sempre pronto (e em executar a entusiasmante tarefa de encher caixas de pontas), e pelas sempre animadas idas ao armazém. Aos colegas entusiastas de micróbios do DCV, onde se incluem, além das pessoas já mencionadas, a Ana Pontes, a Bárbara Gonçalves, o Gonçalo Cavaco, a Inês Grilo, a Lia Godinho, a Inês Gonçalves e a Raquel Portela, obrigada pelo sempre presente espírito de cooperação e companheirismo.

Quero ainda agradecer a todas as pessoas que passaram pelo laboratório de Biologia da Infeção e que contribuíram, de alguma forma, para o meu sucesso. À Carolina Condez, que aceitou os meus ensinamentos, ao Filipe Almeida, à Inês Leal, à Irina Franco, à Joana Bugalhão, à Maria Cunha, à Maria Luís e à Sara Pais, obrigada me proporcionarem conhecimentos e discussões valiosos, e pelo apoio em momentos difíceis. Um muito especial agradecimento à Maria Luís pelas discussões e partilhas científicas, pela compreensão e paciência em alturas de maior aperto e pelas galhofas durante os congressos.

Nada disto teria sido possível sem a cumplicidade e a partilha de risadas com amigos incondicionais. Por isso, dou um abraço bem apertado e cheio de carinho à Joana Valente e à Inês Barroso, amigas de longa data que foram pilares de apoio essenciais, quando mais precisei. Um enorme abraço também à Tânia Santos e à Maria João, cuja presença e sábios conselhos foram (e são!) preciosos. Aos amigos e quase familiares que me conhecem desde tenra idade, a quem eu fui muitas vezes pedir conselhos, e que me "arrancavam" dos "livros" quando eu precisava de uma pausa e não o percebia – o José e a Cláudia, a Luzia e o Achim e os "avós" Conceição e Óscar, – manifesto o meu maior apreço e gratidão. Aos primos Carlos e Rita, deixo o meu agradecimento pelo extensivo apoio.

Faço questão de deixar o meu tributo à professora Nair Enxerto, que me contagiou com uma "bactéria" extremamente resistente - a da Biologia –, e a quem devo o início do meu percurso académico na biologia celular e molecular.

Por último, um ternurento beijinho aos meus pais e aos meus avós maternos, que me transmitiram valores de integridade, consciência e persistência, e me ensinaram a não ter medo de perseguir os meus objetivos e de almejar conquistas, sempre com um pé na terra. Ao meu pai, a pessoa mais reta que já conheci e que partiu no decorrer deste doutoramento, e à minha mãe, cuja força e resiliência admiro, devo todo o carinho e a dedicação com que cresci e o facto de me chamarem à razão sempre que precisei. Era seu desejo ver-me cumprir esta etapa, e a presente tese reflete o seu investimento em mim.

Inês Serrano Pereira

FUNDING

This work was supported by Fundação para a Ciência e Tecnologia (FCT) through grants PTDC/IMI-MIC/1300/2014 and PTDC/BIA-MIC/28503/2017, and in the scope of the projects UIDP/04378/2020 and UIDB/04378/2020 of the Research Unit on Applied Molecular Biosciences – UCIBIO, and LA/P/0140/2020 of the Associate Laboratory Institute for Health and Bioeconomy - i4HB. ISP was supported by PhD fellowship SFRH/BD/129756/2017, also funded by FCT.

WORK CONTRIBUTIONS

This dissertation contains images and text written taken from a recently published article where Inês Serrano Pereira is first author (Pereira *et al.*, 2022). This is indicated as footnote at the beginning of each relevant Chapter (Materials and Methods, section 3.2 of the Results, and Discussion).

The experimental work performed throughout this PhD thesis was essentially performed by Inês Serrano Pereira. Few exceptions (also indicated as footnote at the beginning of each Chapter) are the following: Dr. Sara Vilela Pais performed experiments to analyze the lack of complementation in the growth defect of the *C. trachomatis cteG* mutant strain and made the first preliminary observations of the lytic exit deficiency phenotype of the *cteG* mutant strain (section 3.2 of the Results); Dr. Vítor Borges, Dr. Maria José Borrego, and Dr. João Paulo Gomes sequenced the *C. trachomatis cteG* mutant strain and corresponding parental strain (section 3.2 of the Results); the mass spectrometry analysis (LC-MS/MS) was performed at Clarify Analytical, Portugal (section 3.3 of the Results); Dr. Carla Gonçalves performed with Inês Serrano Pereira the analysis of the homologs of CteG (section 3.4 of the Results).

ABSTRACT

The Phylum *Chlamydiae* comprises bacteria that only multiply inside eukaryotic host cells, within a membrane-bound vacuole. Among *Chlamydiae*, the Family *Chlamydiaceae* includes *Chlamydia trachomatis*, a major human pathogen causing ocular and genital infections. The characteristic infectious cycle of *Chlamydiae* involves chlamydial-mediated host cell invasion and egress. Throughout the cycle, *Chlamydiae* subvert host cell processes through effector proteins delivered into host cells by a type III secretion system. Previously, it was shown that the *C. trachomatis* CteG effector localizes at the Golgi and plasma membrane of infected cells. Moreover, the first 100 residues of CteG fused to EGFP (EGFP-CteG$_{100}$) localize at the Golgi upon their ectopic expression in mammalian cells. In this work, we found that CteG mediates *C. trachomatis* host cell lytic exit. Cells infected by a CteG-deficient strain showed less chlamydiae in the culture supernatant and displayed lower levels of cytotoxicity comparing to cells infected by CteG-producing wild-type and complemented strains. We further showed that CteG and Pgp4, a global regulator of transcription encoded in the *C. trachomatis* virulence plasmid, act on the same pathway leading to chlamydial host cell lytic exit. We also found a predicted α-helix on the N-terminal region of CteG that is essential for the localization of ectopically expressed EGFP-CteG$_{100}$ at the Golgi and plays a role in adequate targeting of CteG to the Golgi and plasma membrane in infected cells. Finally, we identified host cell proteins that may interact with CteG and provided insights into the evolutionary history of *cteG* by bioinformatics analysis of its homologs in *Chlamydiaceae*. In summary, this work revealed a role of CteG in *C. trachomatis* host cell exit, a crucial step of the chlamydial infectious cycle. Together with other findings, this expanded the knowledge on *C. trachomatis*-host cell interactions and opened avenues for future research.

Keywords: Bacterial pathogenesis, pathogen egress, *Chlamydia trachomatis*, type III secretion, effector proteins.

RESUMO

O Filo *Chlamydiae* abrange bactérias que se multiplicam exclusivamente em células de hospedeiros eucariontes, no interior de um vacúolo. *Chlamydia trachomatis*, da Família *Chlamydiaceae*, causa infeções genitais e oculares em humanos. O ciclo infecioso das *Chlamydiae* envolve processos de invasão e saída da célula hospedeira promovidos pela bactéria. Durante este ciclo, as *Chlamydiae* manipulam as células hospedeiras através de proteínas efetoras, transportadas para essas células hospedeiras por um sistema de secreção do tipo III. Previamente, observou-se que CteG, uma proteína efetora de *C. trachomatis*, se localiza no Golgi e na membrana plasmática de células infetadas. Adicionalmente, os primeiros 100 aminoácidos de CteG, fundidos a EGFP (EGFP-CteG$_{100}$), localizam-se no Golgi após a sua expressão ectópica em células de mamífero. Neste trabalho, mostrou-se que CteG intervém na saída lítica de *C. trachomatis* da célula hospedeira. Verificou-se também que CteG e Pgp4, uma proteína codificada no plasmídeo de virulência de *C. trachomatis*, atuam na mesma via que resulta na saída lítica desta bactéria da célula hospedeira. Noutra parte do trabalho, descobriu-se que uma possível hélice-α na região N-terminal de CteG é essencial para a localização de EGFP-CteG$_{100}$ no Golgi, após a sua expressão ectópica em células de mamífero. Também se mostrou que esta hélice-α é importante para um eficiente direcionamento de CteG para o Golgi e para a membrana plasmática de células infetadas. Finalmente, foram identificadas proteínas da célula hospedeira que podem interagir com CteG e foi investigada a história evolutiva de *cteG* através de uma análise bioinformática dos seus homólogos em *Chlamydiaceae*. Em resumo, este trabalho revelou uma função de CteG na saída lítica de *C. trachomatis*, um passo crucial no ciclo infecioso desta bactéria. Juntamente com outras descobertas, foi assim expandido o conhecimento sobre as interações entre *C. trachomatis* e a célula hospedeira e abriram-se várias novas linhas futuras de investigação.

Palavras chave: Patogénese bacteriana, saída de organismo patogénico, *Chlamydia trachomatis*, secreção do tipo III, proteínas efetoras.

CONTENTS

1 INTRODUCTION 1

1.1 The Phylum *Chlamydiae* 1

1.1.1 The expanding group of environmental *Chlamydiae* 2

1.1.2 The Family *Chlamydiaceae* 4

1.2 *Chlamydia trachomatis* 7

1.2.1 The biovars and serovars of *C. trachomatis* 7

1.2.2 Diagnosis, treatments, and vaccines 8

1.2.3 *C. trachomatis* genetics 9

1.2.3.1 The *C. trachomatis* chromosome 9

1.2.3.2 Genetic manipulation of *C. trachomatis* 11

1.2.3.2.1 Transformation of *C. trachomatis* 11

1.2.3.2.2 Random mutagenesis 13

1.2.3.2.3 Site-directed mutagenesis 14

1.3 The *Chlamydiae* plasmids 15

1.3.1 The virulence plasmid of *C. trachomatis* 15

1.3.2 Plasmid-mediated gene regulation 17

1.4 *C. trachomatis* pathogenesis 19

1.4.1 The unique developmental cycle 19

1.4.2 Secretion systems in *Chlamydiae* 21

1.4.2.1 The T3S system of *C. trachomatis* 23

1.4.2.2 *C. trachomatis* T3S effectors 24

1.4.2.2.1 Inclusion membrane proteins (Incs) 25

1.4.2.2.2 *C. trachomatis* non-Inc effector proteins 30

1.4.2.3 Non-Inc T3S system-independent/candidate effectors 33

1.4.2.4 Chlamydial proteins delivered into the inclusion lumen 35
1.5 Objectives 36
2 MATERIALS AND METHODS 39
2.1 DNA manipulation, oligonucleotides, and plasmids 39
2.2 *Escherichia coli* strains and growth conditions 40
2.3 Mammalian cell lines 40
2.4 Transient transfection of mammalian cells and treatment with 2-Bromopalmitate 40
2.5 Manipulation of *C. trachomatis* 41
2.5.1 Generation of *C. trachomatis* strains 41
2.5.1.1 Transformation of *C. trachomatis* 41
2.5.1.2 Plaque purification of *C. trachomatis* 42
2.5.2 Infection of HeLa cells with *C. trachomatis* 43
2.5.3 Assessment of relative progeny generation and quantification of IFUs in infected cells 43
2.5.4 Cell cytotoxicity assays 44
2.5.5 Internalization of *C. trachomatis* 44
2.5.6 Real-time quantitative PCR (RT-qPCR) 45
2.6 Antibodies and dyes 46
2.7 Immunofluorescence microscopy 46
2.8 Immunoblotting 47
2.9 Co-immunoprecipitation assays 47
2.10 Mass spectrometry (nano-LC-MS/MS) 48
2.11 Whole-genome sequencing 49
2.12 Bioinformatics 50
2.12.1 Reciprocal best hit BLAST 50
2.12.2 Protein alignments and phylogenetic analysis 51
2.13 Statistical analysis 51
3 RESULTS 53
3.1 Determinants of the subcellular localization of CteG 54
3.1.1 The first 20 amino acid residues of CteG are necessary for the localization of mEGFP-CteG$_{1\text{-}100}$ at the Golgi 54

3.1.2 Specific residues at a putative α-helix in the N-terminal region of CteG are necessary for the localization of mEGFP-CteG$_{1-100}$ at the Golgi 57
3.1.3 Specific residues at a putative α-helix in the N-terminal region of CteG are required for an efficient targeting of mEGFP-CteG to the Golgi and plasma membrane 61
3.1.4 Inhibition of host S-palmitoyltransferases causes a minor impact in the localization of mEGFP-CteG$_{1-100}$ at the Golgi 64
3.1.5 Analysis of the impact of different regions of CteG in the subcellular localization of mEGFP-CteG proteins at the plasma membrane 66
3.1.6 The substitution of specific amino acid residues at the N-terminal region of CteG-2HA affects its type III secretion (T3S) system-mediated delivery and subcellular localization during *C. trachomatis* infection of HeLa cells 68
3.1.7 Conclusions 73
3.2 CteG mediates host cell lytic exit of *Chlamydia trachomatis* 74
3.2.1 A *C. trachomatis cteG::aadA* insertional mutant has a defect in progeny generation 74
3.2.2 The *cteG::aadA* mutant strain shows a CteG-dependent defect in egress from infected host cells 76
3.2.3 The *cteG::aadA* mutant strain shows a CteG-dependent defect in host cells lysis.. 79
3.2.4 Production and localization of CteG are not regulated by *C. trachomatis* virulence plasmid encoded Pgp4 82
3.2.5 A *cteG::aadA pgp4* double mutant strain displays a defect in inducing host cell lysis identical to *cteG::aadA* or *pgp4* single mutants 87
3.2.6 Overproduction of CteG-2HA suppresses the defect of CteG- and Pgp4-deficient *C. trachomatis* to mediate host cell lysis 91
3.2.7 Conclusions 92
3.3 Identification of CteG host cell interacting partners 93
3.4 Distribution of CteG homologs among *Chlamydiaceae* 96
3.4.1 Identification of putative homologs of CteG within *Chlamydiaceae* by reciprocal best hit BLAST 96
3.4.2 Phylogenetic analysis of the homologs of CteG among *Chlamydiaceae* 99
3.4.3 Synteny of *cteG* 103
3.4.4 Identification of homologs among *Chlamydiaceae* of chlamydial non-Inc T3S substrates 106
3.4.5 Conclusions 110

4 DISCUSSION AND CONCLUSIONS 111

4.1 Identifying the determinants of the subcellular localization of CteG during *C. trachomatis* infection 112

4.1.1 Pinpointing the amino acids that mediate the localization of CteG at the Golgi and plasma membrane of host cells 112

4.1.2 The possible role of S-palmitoylation in targeting CteG to eukaryotic membranes 114

4.2 *C. trachomatis* lytic exit from host cells is CteG-dependent 115

4.2.1 CteG is a novel protein involved in the lytic exit of *C. trachomatis* from host cells 115

4.2.2 Possible timing and mode of action of CteG during infection by *C. trachomatis*.... 116

4.2.3 Host cell lytic exit among different *Chlamydiae* 117

4.2.4 Model of action of CteG during *C. trachomatis* infection 118

4.3 Identifying CteG host cell interacting partners: other possible roles of CteG during infection 119

4.4 Conservation of *cteG* among *Chlamydiaceae* 121

4.5 Future directions and concluding remarks 124

REFERENCES 127

ANNEXES 157

LIST OF FIGURES

Figure 1.1. Phylogenetic tree of species from the Phylum *Chlamydiae* based on full length *16S* rRNA gene sequences. 4
Figure 1.2. Phylogenetic tree of some species of the Genus *Chlamydia*..... 5
Figure 1.3. Developmental cycle of *C. trachomatis*. 20
Figure 1.4. Secretion systems used by *C. trachomatis* to transport proteins to distinct chlamydial and host cell compartments. 23
Figure 1.5. Schematic representation of a proposed model for the localization and mode of action of CteG during *C. trachomatis* infection of epithelial cells. 36
Figure 3.1. Analysis of the localization of ectopically produced mEGFP-CteG$_{1\text{-}100}$ truncated proteins. 55
Figure 3.2. A putative α-helix at the N-terminal region of CteG is important for the localization of mEGFP-CteG$_{1\text{-}100}$ at the Golgi. 58
Figure 3.3. Analysis of the subcellular localization of mEGFP-CteG$_{1\text{-}100}$ proteins containing two or three amino acid replacements. 59
Figure 3.4. Analysis of the subcellular localization of mEGFP-CteG$_{1\text{-}100}$ proteins containing single amino acid replacements. 60
Figure 3.5. Specific amino acids at a putative α-helix in the N-terminal region of CteG are important for the subcellular localization of ectopically produced mEGFP-CteG. 63
Figure 3.6. Effect of the inhibition of mammalian palmitoyltransferases in the localization of mEGFP-CteG$_{1\text{-}100}$ and mEGFP-CteG$_{FL}$ at the Golgi. 65
Figure 3.7. Analysis of the importance of different regions of CteG in the localization of mEGFP-CteG proteins at the plasma membrane. 67
Figure 3.8. Analysis of the effect of specific amino acid replacements at the N-terminal region of CteG-2HA on its production during infection of HeLa cells by *C. trachomatis*. 69
Figure 3.9. Analysis of the effect of specific amino acid replacements at the N-terminal region of CteG-2HA on its localization during infection of HeLa cells by *C. trachomatis*. 70

Figure 3.10. Methodology for the quantitative analysis of the production and T3S system-mediated delivery of CteG-2HA proteins during *C. trachomatis* infection. 71
Figure 3.11. Quantitative analysis of the effect of specific amino acid replacements at the N-terminal region of CteG-2HA on its production and T3S system-mediated delivery during *C. trachomatis* infection 72
Figure 3.12. Methodology for the quantitative analysis of the subcellular localization of CteG-2HA proteins during *C. trachomatis* infection. 73
Figure 3.13. A *C. trachomatis cteG::aadA* insertional mutant is defective in progeny generation. 75
Figure 3.14. *C. trachomatis cteG::aadA* displays a CteG-dependent defect in egress from infected host cells 77
Figure 3.15. The *C. trachomatis cteG* mutant strain displays a CteG-dependent defect in egress from infected host cells. 78
Figure 3.16. *C. trachomatis cteG::aadA* displays a CteG-dependent defect in host cell lysis. 80
Figure 3.17. The *C. trachomatis cteG* mutant strain displays a CteG-dependent host cell lysis defect. 81
Figure 3.18. The *C. trachomatis* virulence plasmid contributes to lytic exit from host cells. 82
Figure 3.19. Verification of the accuracy of the *C. trachomatis* strains generated in this work. 83
Figure 3.20. Verification of the replacement of the *C. trachomatis* native plasmid by recombinant plasmids with or without *pgp4*. 84
Figure 3.21. Pgp4 does not modulate the production or the localization of CteG during *C. trachomatis* infection 86
Figure 3.22. Characterization of *C. trachomatis* strains used and generated in this study. 89
Figure 3.23. *C. trachomatis* mediates host cell lysis via a common pathway involving both CteG and Pgp4. 90
Figure 3.24. Simplified illustration of the co-immunoprecipitation assays using extracts of HeLa cells ectopically producing mEGFP-CteG$_{FL}$ or mEGFP alone for mass spectrometry analysis 94
Figure 3.25. Analysis of samples collected from different steps of the co-immunoprecipitation assays by SDS-PAGE and Coomassie staining, or immunoblotting. 95
Figure 3.26. Graphical summary of the reciprocal best hit BLAST and phylogenetic analyses to identify putative homologs of CteG. 97
Figure 3.27. Individual alignments between CteG and its putative homologs. 100
Figure 3.28. Phylogenetic tree of CteG and its putative homologs in species of the Family *Chlamydiaceae*. 101
Figure 3.29. Phylogenetic relationships between species from the Family *Chlamydiaceae* and from other *Chlamydiae*. 102

Figure 3.30. Organization of the genomic region of *cteG* and of each of its putative homologs. 104
Figure 4.1. Hypothetical model for the mode of action of CteG and Pgp4 in promoting host cell lysis. 119
Figure 4.2. Predicted tertiary structure of CteG and of two of its putative homologs. 123

LIST OF TABLES

Table 3.1. Possible homologs of CteG in other *Chlamydiae* before and after both bioinformatics and phylogenetic analyses. 98
Table 3.2. Analysis of the number of homologs of different non-Inc *C. trachomatis* proteins in other *Chlamydiae* by reciprocal best hit BLAST (found with protein BLAST / recovered with reciprocal BLAST)..... 108

ABBREVIATIONS

2BP	2-Bromopalmitate
ATP	Adenosine triphosphate
BLAST	Basic Local Alignment Search Tool
BLASTp	Protein BLAST
BSA	Bovine serum albumin
Cas	CRISPR associated protein
CC	*Chlamydia* clade
cGAS-STING	Cyclic GMP–AMP synthase–stimulator of interferon genes pathway
CHX	Cycloheximide
CI-M6PR	Cation-independent mannose 6-phosphate receptor
co-IP	Co-immunoprecipitation
CPAF	Chlamydial protease/proteasome-like activity factor
CRISPR	Clustered regularly interspaced short palindromic repeats
DMEM	Dulbecco's modified eagle medium
DNA	Deoxyribonucleic acid
EB	Elementary body
EMS	Ethylmethanesulfonate
ESCRT	Endosomal sorting complexes required for transport
FBS	Fetal bovine serum
FRAEM	Fluorescence-reported allelic exchange mutagenesis

GFP	Green fluorescent protein
GM130	Golgi matrix protein 130
GSK	Glycogen synthase kinase
GTPase	Guanosine triphosphate hydrolase
HA	Hemagglutinin
HBSS	Hank's balanced salt solution
HRP	Horseradish peroxidase
Hsp60	Heat-shock protein 60
IFN-γ	Interferon-gamma
Inc	Inclusion membrane protein
INSA	Portuguese National Institute of Health
LDH	Lactate dehydrogenase
LGV	Lymphogranuloma venereum
mEGFP	Monomeric enhanced green fluorescent protein
ML	Maximum likelihood
MLD	Membrane localization domain
MOI	Multiplicity of infection
MOMP	Major outer membrane protein
mRNA	Messenger RNA
MTOC	Microtubule-organizing center
NAAT	Nucleic acid amplification test
NCBI	National Center for Biotechnology Information
ORF	Open reading frame
p.i.	Post-infection
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PenG	Penicillin G

Pgp	Plasmid glycoprotein
Pmp	Polymorphic membrane protein
PMSF	Phenylmethylsulfonyl fluoride
PZ	Plasticity zone
RB	Reticulate body
RNA	Ribonucleic acid
ROI	Region of interest
rRNA	Ribosomal RNA
RT-qPCR	Real-time quantitative PCR
SDS	Sodium dodecyl sulphate
SDS-PAGE	SDS-polyacrylamide gel electrophoresis
SEM	Standard error of the mean
SNARE	Soluble N-ethylmaleimide-sensitive factor attachment protein receptors
SNX	Sorting nexin
sRNA	Small anti-sense RNA
T3S	Type III secretion (the same applied to type I, II, IV, V and VI secretion)
TGN	*trans*-Golgi network
tRNA	Transfer RNA
Trp	Tryptophan
WGS	Whole-genome sequencing
WHO	World Health Organization
α	Alpha
β	Beta
γ	Gamma

1

INTRODUCTION[1]

1.1 The Phylum *Chlamydiae*

Chlamydiae are a group of Gram-negative bacteria that are obligate intracellular and share a characteristic biphasic developmental cycle. This Phylum includes organisms responsible for relevant pathologies such as trachoma, for which diagnosis and treatment procedures have been documented since ancient times (Taylor, 2008). The first experimental report of a chlamydial infection was published in 1907, after the observation an infectious agent in the cytoplasm of cells scraped from the conjunctiva of orangutans with trachoma (Halberstädter and von Prowazeck, 1907). The intracytoplasmic bodies formed by this organism, now known as inclusions, appeared as "cloaks" that covered the nucleus of infected cells, originating the designation *Chlamydozoa* (from the Greek word meaning "cloak"). Such morphology also led to the misconception that these organisms were protozoan parasites (Halberstädter and von Prowazeck, 1907). Subsequently, for many years, the chlamydial etiological agents of psittacosis and lymphogranuloma venereum have been referred to and handled as viruses, which were difficult to propagate due to their intracellular lifestyle (Sanders, 1940; John and Gordon, 1946). In 1966, Moulder revised the main characteristics of chlamydial organisms and concluded that many are not compatible with viruses: a developmental cycle comprising two morphologically different forms, a cell wall resembling that of Gram-negative bacteria, non-compartmentalized DNA and RNA, ribosomes, and replication by binary fission (Moulder, 1966). *Chlamydiae* were therefore recognized as bacteria and were included in a single taxon.

[1] This Chapter was written by Inês Serrano Pereira, based on the cited bibliographic references.

1.1.1 The expanding group of environmental *Chlamydiae*

The Phylum *Chlamydiae* is so far exclusively represented by obligate intracellular pathogens of eukaryotic hosts, and only possesses one Class (*Chlamydiia*) and consensually comprises one Order (*Chlamydiales*). However, the taxonomy of the Phylum *Chlamydiae* is in constant modification as various comparative analyses are performed and new species are identified, as exemplified by a phylogenomic study proposing the division of the Class *Chlamydiia* into two Orders [*Chlamydiales* and *Parachlamydiales* ord. nov. (Gupta *et al.*, 2015)].

The Phylum *Chlamydiae* comprises the well-known Family *Chlamydiaceae*, a group of organisms that are all pathogenic to humans and to other animals, and species that are primarily symbionts of microbial eukaryotes, the so-called "environmental *Chlamydiae*" or "*Chlamydia*-like organisms" (CLOs; Figure 1.1). Comparing to the *Chlamydiaceae* species, these organisms typically possess larger genomes that encode genes deeply conserved among the Phylum *Chlamydiae* (housekeeping genes) and genes that are family- and genus-specific, which account for the variability within the *Chlamydiae* (Collingro *et al.*, 2011). Correlated to these larger genomes are extended metabolic abilities comparing to species pathogenic to animals, which reflects an adaptation to amoebal cell hosts (Taylor-Brown *et al.*, 2015).

Challenges on the isolation and cultivation of environmental *Chlamydiae* due to their strict dependency on a eukaryotic host cell, and the fact that their natural hosts and growth conditions are unknown (Taylor-Brown *et al.*, 2018a) have resulted in very limited knowledge about their existence and biology. However, the recent development of whole-genome sequencing (WGS), metagenomics and single cell genomics have allowed the determination of single cell amplified genomes (SAGs) and metagenome-assembled genomes (MAGs) from uncultured *Chlamydiae* that inhabit complex and diverse microbial environments. This has led to a significant expansion of the number of identified environmental *Chlamydiae* over the last few years, accompanied by an increase in comparative genomics and phylogenomic studies (Lagkouvardos *et al.*, 2014; Pillonel *et al.*, 2018; Taylor-Brown *et al.*, 2018b; Stairs *et al.*, 2020). Moreover, unprecedent features were identified in some of these new species, such as genes encoding complete flagellar apparatuses (Collingro *et al.*, 2017) or genes related to anaerobic metabolism (Dharamshi *et al.*, 2020; Stairs *et al.*, 2020; Köstlbacher *et al.*, 2021).

Several Families are included in the group of environmental *Chlamydiae* (Figure 1.1) and many include species whose primary hosts are free-living amoebae: the Family *Simkaneaceae* comprises *Simkania negevensis*, which infects a wide range of hosts from amoebae to mammalian cells (Kahane *et al.*, 2002; Koch *et al.*, 2020); *Parachlamydiaceae* is another Family that includes *Protochlamydia naegleriophila* and *Parachlamydia acanthamoebae*; the Family *Criblamydiaceae*, which comprises *Estrella lausannensis*; and the Family *Waddliaceae*, whose type species is *Waddlia chondrophila* (Figure 1.1). *S. negevensis*, *E. lausannensis* and the exemplified

Parachlamydiaceae species are associated to human respiratory diseases, and the former two species are considered potential emerging pathogens (Friedman *et al.*, 2003; de Barsy *et al.*, 2014), even though these facts are controversial (Tagini and Greub, 2018). Additional Families comprehend the *Rhabdochlamydiaceae*, the *Clavichlamydiaceae*, the *Piscichlamydiaceae* (Figure 1.1) and the *Parilichlamydiaceae*, whose members infect arthropods, fish, or potentially mammalians, including humans (Daniele and Gilbert, 2006; Taylor-Brown *et al.*, 2015, 2018b).

As above-mentioned, new chlamydial species have been unveiled over the last few years. For instance, Dharamshi and colleagues identified new chlamydial species in marine sediments and proposed new clades comprising already known species or the identified ones (Dharamshi *et al.*, 2020). The new clades comprise, besides *Chlamydiaceae* and the classic environmental *Chlamydiae*, the *Chlamydia* clades (CC) I to IV and the anaerobic clade of bacteria *Candidatus* Anoxychlamydiales. CC-IV is described as a sister clade of the Family *Chlamydiaceae*, and its species possess intermediate features between the *Chlamydiaceae* species and environmental *Chlamydiae*, as the size of their genomes and certain putative metabolism-associated or flagellar genes. CC-IV shares a large set of genes with the *Chlamydicaceae* that are absent in all other lineages, but most have unknown function (Dharamshi *et al.*, 2020). Therefore, the CC-IV clade supports the idea that the evolution of *Chlamydiaceae* species was marked by gene loss as an adaptation to an intracellular lifestyle (Collingro *et al.*, 2011; Nunes and Gomes, 2014).

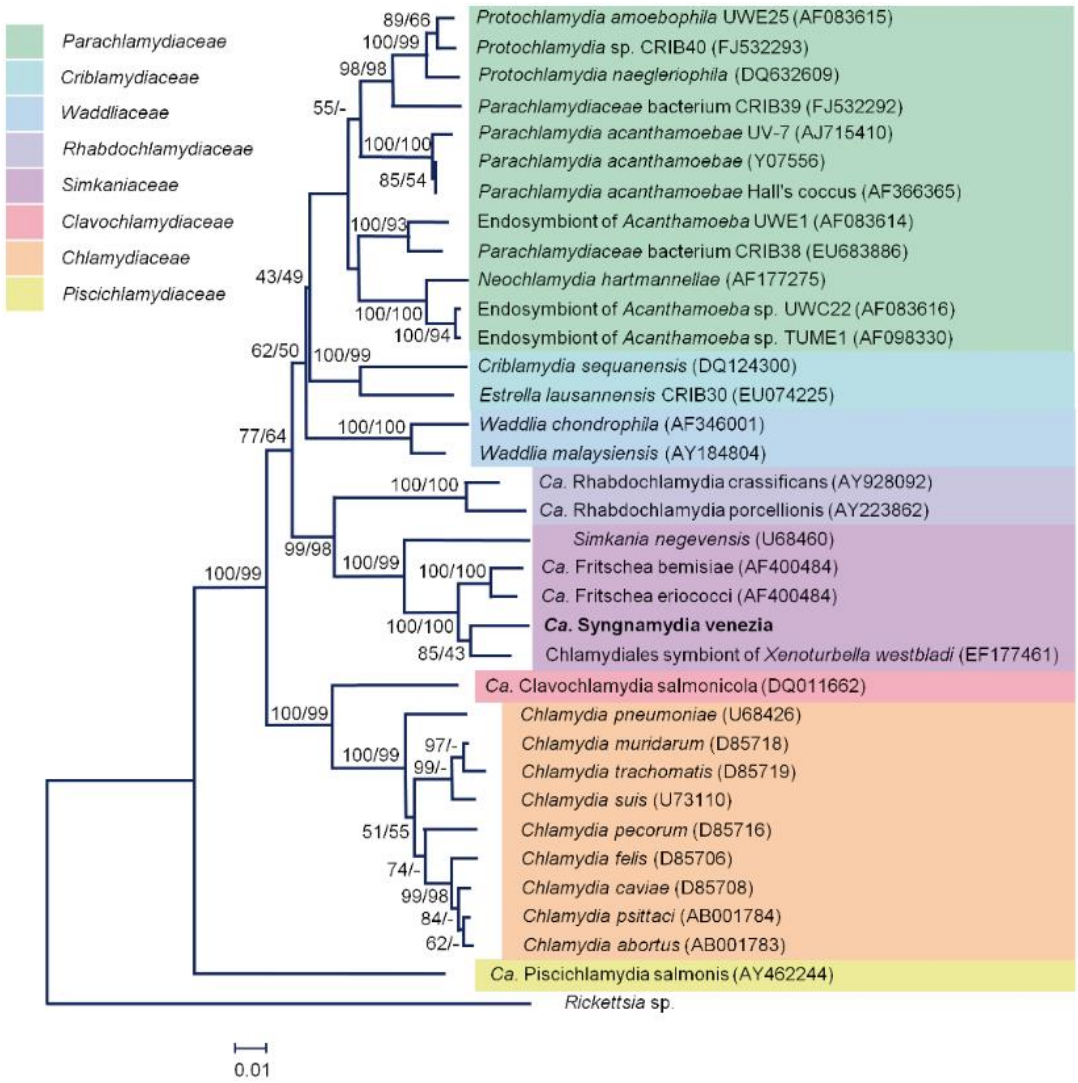


Figure 1.1. Phylogenetic tree of species from the Phylum *Chlamydiae* based on full length *16S* rRNA gene sequences. Phylogenies inferred with the neighbour-joining/maximum parsimony methods. Numbers represent the percentage of replicate trees in which the associated taxa cluster together in 1000 bootstrap replicates. Reprinted from (Fehr *et al.*, 2013).

1.1.2 The Family *Chlamydiaceae*

The Family *Chlamydiaceae* includes the Genus *Chlamydia* (Figure 1.2), which comprises to this date 19 documented species, of which 11 are well established (Bachmann *et al.*, 2014). *Chlamydia trachomatis* is an exclusively human pathogen and is the leading cause of bacterial infections of the conjunctiva and of the urogenital tract worldwide, conditions that may have serious outcomes (see section 1.2.1 below). *Chlamydia pneumoniae* is also a human pathogen acquired through inhalation of contaminated droplets and is associated to respiratory disease in humans and to community-acquired pneumonia (Roulis *et al.*, 2013; Aliberti *et al.*, 2021). Infections by *C. pneumoniae* are also thought to be relevant in the context of human pathologies

such as atherosclerosis (Belland *et al.*, 2004; Tumurkhuu *et al.*, 2018), or chronic conditions such as Alzheimer's disease (Balin *et al.*, 2018), arthritis or lung cancer, but this still requires further investigation. *C. pneumoniae* also causes vascular, respiratory, or systemic disease in other animals such as reptiles, koalas, horses, and frogs, which have the potential to transmit the pathogen to humans (Roulis *et al.*, 2013).

The other 9 species of the Genus *Chlamydia* have non-human animals as main hosts, although some may have zoonotic potential and be transmitted from animals to humans (Figure 1.2). Importantly, many species cause asymptomatic infections in their hosts that only occasionally evolve into more serious illnesses.

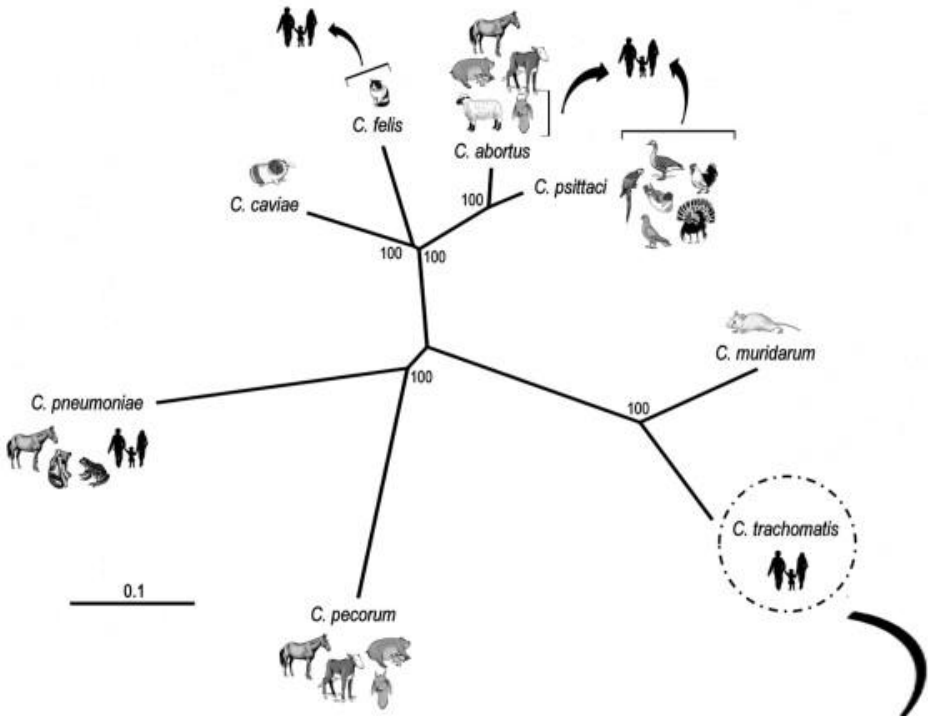


Figure 1.2. Phylogenetic tree of some species of the Genus *Chlamydia*. Neighbor-joining tree based on 600 orthologous genes, which illustrates *Chlamydia* species pathogenic only to humans (*C. trachomatis*), animals or to both humans and animals. The brackets represent demonstrated zoonotic transmission of the *Chlamydia* species (*C. abortus*, *C. psittaci* and *C. felis*) from animals to humans at the time of the analysis. Reprinted with permission from (Nunes and Gomes, 2014).

Chlamydia suis is one of the most phylogenetically closely related to *C. trachomatis* and yet its genome was only recently sequenced (Manuela *et al.*, 2014). *C. suis* is endemic to pigs and is the cause of conjunctivitis, pneumonia, and various reproductive and fertility disorders. *C. suis* has a high resistance rate to tetracycline conferred by a *tet*(C)-containing cassette, which raises concerns as to a possible transference of resistance from *C. suis* to *Chlamydia* species that infect humans upon zoonotic transmission from pigs to humans (Joseph *et al.*, 2016).

Chlamydia muridarum causes pneumonitis in mice (Williams *et al.*, 1981), and is also able to establish persistent infections in the gastrointestinal tract of these animals. *C. muridarum*

infection of the genital tract of mice is used as a model to study the mechanisms of *C. trachomatis* pathogenesis in the human genital tract (Barron *et al.*, 1981; Yeruva *et al.*, 2013).

Chlamydia psittaci infects typically avian hosts and is especially relevant in domestic birds, leading to psittacosis/ornithosis. It is also found in non-avian animals as sheep, horses, and cats, and can also be responsible for atypical pneumonia in humans after zoonotic transmission (Knittler and Sachse, 2015). Other *Chlamydia* species infecting avian hosts are *Chlamydia avium* and *Chlamydia gallinacea* which were recently identified from poultry that had respiratory and gastrointestinal symptoms (Sachse *et al.*, 2014), but information about the pathogenesis of these strains is still limited. A recent study has described that, comparing to *C. psittaci*, *C. gallinacea* possesses genes encoding key virulence factors but to a reduced number of homologs, which hypothetically translates into lower mortality rates (Heijne *et al.*, 2021).

Chlamydia abortus is a zoonotic pathogen that infects small ruminants, especially sheep and goats, and ultimately leads to fetal loss during late stages of gestation, which translates into important economic losses (Longbottom and Coulter, 2003). *C. abortus* is also a concern in humans for its ability to infect the placenta, leading to miscarriage (Pichon *et al.*, 2020), and its capacity to induce atypical pneumonia with severe outcomes (Liu *et al.*, 2022).

Chlamydia pecorum primarily infects the ocular, respiratory, or urogenital tissues of koalas and is a cause for the reduction of populations and infertility of this species (Fabijan *et al.*, 2019). However, *C. pecorum* also disseminates to other livestock animals, causing respiratory, gastrointestinal, or urogenital infections (Sait *et al.*, 2014; Jelocnik *et al.*, 2015).

Chlamydia felis mainly causes conjunctivitis in cats, although it may also target the respiratory and genital tracts. This pathogen may be transmitted to humans by close contact with ocular secretions or aerosols of infected cats (Hartley *et al.*, 2001; Wons *et al.*, 2017).

Chlamydia caviae is commonly associated to infections of the conjunctiva (Murray, 1964) and genital tract of guinea pigs, being used as a model to study human chlamydial infections (Neuendorf *et al.*, 2015). However, it has also been recently identified as a zoonotic pathogen responsible for community-acquired pneumoniae (Ramakers *et al.*, 2017).

During the last years, at least 8 other *Chlamydia* species have been discovered and are proposed to be integrated in the Family *Chlamydiaceae*. These comprehend species that infect snakes, *Ca.* Chlamydia corallus (Taylor-Brown *et al.*, 2017), *Chlamydia serpentis*, *Chlamydia poikilotherma* (Staub *et al.*, 2018) and *Chlamydia sanzinia* (Taylor-Brown *et al.*, 2016); birds, *Chlamydia ibidis* (Vorimore *et al.*, 2013) and *Chlamydia buteonis* (Laroucau *et al.*, 2019); tortoises, *Ca.* Chlamydia testudinis (Laroucau *et al.*, 2020) and Siamese crocodiles, *Chlamydia crocodilis* (Chaiwattanarungruengpaisan *et al.*, 2021). Additionally, new species pathogenic to flamingoes have been discovered, and even proposed to form a new Genus within the Family *Chlamydiaceae* (Vorimore *et al.*, 2021). However, since these are newly identified species, their pathogenesis and phylogenetic relation with other *Chlamydia* needs to be further assessed.

1.2 *Chlamydia trachomatis*

C. trachomatis shares with other *Chlamydiae* a unique and remarkably host cell-dependent developmental cycle involving two distinct forms - an elementary body (EB) and a reticulate body (RB) (see Figure 1.3 in section 1.4.1 below). During infection of host cells, chalmydiae reside and replicate inside a membrane-bound vacuole, while manipulating several host cell processes to acquire host nutrients and survive. This leads to the formation of a large *Chlamydia*-containing vacuole known as inclusion. At late stages of infection, *Chlamydia*e exit the intracellular niche by lysis of host cells or by extrusion of the entire inclusion (see Figure 1.3 below). Understanding this distinctive developmental cycle is a key aspect to decipher the virulence of *C. trachomatis* and allows one to appreciate the challenges associated to the investigation of this bacterium.

Completion of the developmental cycle by all *Chlamydiae* relies on a type III secretion (T3S) system (see Figure 1.3 and section 1.4.2.1 below), a protein transport mechanism present in many Gram-negative bacteria. This system allows the delivery of bacterially-produced effector proteins into eukaryotic host cells (see Figure 1.3 below) known as T3S effectors, which manipulate a variety of host cell processes [(Bugalhão and Mota, 2019); see section 1.4.2.2 below]. The T3S substrates may be delivered into the inclusion lumen, into the inclusion membrane (also known as Incs) or into the host cell cytoplasm. Chlamydial proteins detected in the host cell cytosol also include type III-independent secreted proteins such as CPAF (chlamydial protease/proteasome-like activity factor) (see section 1.4.2.3 below). While much about the functions of chlamydial effectors remains to be clarified, it is certain that they play key roles on the completion of different stages of the developmental cycle (Elwell *et al.*, 2016; Bugalhão and Mota, 2019).

1.2.1 The biovars and serovars of *C. trachomatis*

C. trachomatis strains are clinically relevant because of their impact in public health. Therefore, they are the most studied organisms within the Phylum *Chlamydiae*. *C. trachomatis* strains comprise the trachoma, the genital tract, and the lymphogranuloma venereum (LGV) biovars, depending on their tropism for different human tissues. Each biovar is further divided into serovars, which are established typically according to the antigenic diversity of chlamydial major outer membrane protein (MOMP) (Stephens *et al.*, 1987).

C. trachomatis trachoma biovars (serovars A to C) normally cause ocular infections in humans and are transmitted by eye or nose discharges. Recurrent infections may result in trachoma, which affects ~ 21 million people worldwide and is the leading cause of infectious blindness globally [~ 2.2 million individuals; (Taylor *et al.*, 2014)]. Trachoma is endemic in the

poorest areas of Africa, Asia, Australia, and Middle East, where hygiene conditions and access to healthcare are more limited (Taylor *et al.*, 2014). Global strategies have reduced the risk of blindness by trachoma in developing countries by 91% in the last two decades, but this disease remains a public health problem that needs efforts to eradicate (WHO Alliance for the Global Elimination of Trachoma by 2020: progress report, 2019).

Genital tract biovars of *C. trachomatis* (serovars D to K) are the major cause of human urogenital tract infections of bacterial origin worldwide, with an estimated global prevalence of 3.4% in women and 3.3% in men, and a total of 127.2 million new infections in 2016 (15-49-year-old individuals; Rowley *et al.*, 2019). Infections are often asymptomatic (~80%) and may persist for several years if left untreated, which account for their high prevalence. In women, genital infections may ascend to the cervix and to the upper genital tract causing pelvic inflammatory disease, tubal infertility, pelvic pain, and adverse pregnancy outcomes. In men, infections can lead to urethritis, epididymitis, and inflammation of the rectal mucosa (proctitis) (O'Connell and Ferone, 2016). In Portugal, data on infections by *C. trachomatis* is scarce due to the lack of a surveillance program. The Portuguese National Institute of Health (INSA) reported 2.7% incidence of *C. trachomatis* among 931 analyzed individuals in 2016 (INSA Serological Inquiry, 2015-2016). According to a more recent study carried between 2015 and 2018, the prevalence of *C. trachomatis* in Portugal is rising, with a total of 1267 infections identified (43% were symptomatic and 70% occurred in males) (Pinho-Bandeira *et al.*, 2020).

The LGV biovars (serovars L1 to L3) are responsible for genital, anorectal and more invasive infections that are usually symptomatic, although 25% of the anorectal infections are asymptomatic (de Vries *et al.*, 2019). In Europe and North America, outbreaks of LGV-associated anorectal infections in men that have sex with men are ascending (Mejuto *et al.*, 2013; O'Connell and Ferone, 2016; Peuchant *et al.*, 2020). Moreover, co-infections are remarkably frequent [for example, with other *C. trachomatis* serovars (Rodriguez-Dominguez *et al.*, 2015) or HIV (Marangoni *et al.*, 2021)]. The primary stage of infection is characterized by the appearance of ulcers in the inoculation site, and, on a secondary stage, *C. trachomatis* disseminates from the mucosa to regional lymph nodes and the infection evolves into inguinal lymphadenopathy, which is common in developing countries (de Vries *et al.*, 2019).

1.2.2 Diagnosis, treatments, and vaccines

Diagnostic of localized infections by *C. trachomatis* can be performed by cell culture, antigen tests, nucleic acid hybridization and amplification tests that directly detect bacteria, usually using swabs from different anatomical sites. Nucleic acids amplification test (NAAT) is the most specific and sensitive detection method and does not require viable bacteria as cell culture-based methods. NAATs based on real-time detection of amplification products

(usually from chlamydial plasmid genes or rRNA gene transcripts), along with automated nucleic acid extraction, may generate results in a few hours. Alternatively, antigen-based rapid detection tests (RDTs) allow diagnosis in a few minutes and, consequently, a rapid treatment initiation but are significantly less sensitive and accurate. On the other hand, indirect methods comprise detection of antibodies against *C. trachomatis* in the serum of patients, which is more reliable for persistent/invasive infections due to the presence of high antibody titers (Domeika *et al.*, 2009; Meyer, 2016).

C. trachomatis infections are usually treated with azithromycin and doxycycline in different dosages and for variable times, depending on the infection type and complication. Second-line antibiotics include erythromycin, levofloxacin and ofloxacin, and josamycin can be used as a third-line treatment (Lanjouw *et al.*, 2016). Although *C. trachomatis* resistance to antibiotics is not confirmed, it potentially remains a threat in human chlamydial infections as cases of therapeutic failure and reinfection are documented (Borel *et al.*, 2016), urging the creation of novel anti-chlamydial therapies. The development of an effective chlamydial vaccine has also been challenging since most studies are performed in animal models, whose immune responses to infection differ from humans. Yet, a novel vaccine based on the antigen CTH522, a recombinant version of *C. trachomatis* MOMP, was recently developed and proved to be immunogenic and safe on a first-in-human phase I clinical trial (Abraham *et al.*, 2019).

1.2.3 *C. trachomatis* genetics

1.2.3.1 The *C. trachomatis* chromosome

The first complete *C. trachomatis* genome to be published was from a serovar D strain. The small chromosome of this bacterium consists of over 1 million base pairs and contains about 894 protein-encoding genes, which corresponds to a coding density of 90% (for serovar L2, these values are 889 and 89%, respectively). 28% of these genes do not share similarities with other sequences deposited in GenBank (Stephens *et al.*, 1998; Thomson *et al.*, 2008). The small size of the genome of *C. trachomatis* suggests gene loss might have occurred during its evolution into an intracellular pathogen. Indeed, many metabolic pathways such as nucleotide and amino acid biosynthesis, glycolysis, and tricarboxylic acid (TCA) cycle are either absent or incomplete, which translates into dependency on the host nutrients. On the other hand, essential genes related to DNA replication, transcription, and translation, components of DNA repair/recombination, aerobic respiration, peptidoglycan and lipid synthesis and membrane transport systems (ABC transporters) are present (Stephens *et al.*, 1998; Bachmann *et al.*, 2014).

The genomes of *C. trachomatis* are very similar in size and highly conserved between strains (> 99% similarity), with a high degree of synteny, few indels and no known variably

present genomic islands (Stephens *et al.*, 2009; Harris *et al.*, 2012; Seth-Smith *et al.*, 2013). Therefore, differences between serovars and their tropism for different human tissues are commonly attributed to genetic variations in a few genes. Some of these genes include *ompA*, polymorphic membrane protein (Pmp) genes *pmpA* to *pmpI*, genes encoding several type III secretion effectors, including Incs (*incD-G* and *incA-C*), genes associated to some metabolic pathways and possibly genes with unknown function (Abdelsamed *et al.*, 2013; Nunes *et al.*, 2013). Additionally, a hypervariable genomic region named "plasticity zone" (PZ) is present in several *Chlamydia* species and is thought to be important for their tropism for different tissues or organisms. This region comprises several genes as those encoding a cytotoxin (CT166 in serovar D) and the tryptophan (Trp) operon (Thomson *et al.*, 2008; Bachmann *et al.*, 2014). CT166 was found to be a chlamydial cytotoxin (Belland *et al.*, 2001) that induces actin reorganization and loss of host cell shape (see section 1.4.2.3 below) (Thalmann *et al.*, 2010). The last genes involved in Trp biosynthesis (*trpRBA*) display important differences between genital and ocular strains of *C. trachomatis*. *trpA* and *trpB* encode both subunits of a tryptophan synthase (TrpA and TrpB, respectively). While genital strains encode both intact *trpA* and *trpB*, ocular strains have frameshift mutations in *trpA* and, thus, are not able to produce a functional tryptophan synthase. This is thought to be biologically relevant for *C. trachomatis* genital strains *in vivo*, whereby Trp synthesis is crucial. In response to chlamydial infection, the host produces interferon-γ (IFN-γ) which binds to infected cells and stimulates them to degrade Trp. Trp deprivation would inhibit replication of *C. trachomatis*, but genital strains can produce Trp from indole that is provided by the genital tract microbiome, allowing chlamydiae to evade the host immune response (Fehlner-Gardiner *et al.*, 2002; Caldwell *et al.*, 2003). This role of the Trp operon during infection was further validated upon the observation that an ocular strain of *C. trachomatis* (harboring truncated *trpA*) complemented with tryptophan biosynthesis genes from a *C. trachomatis* urogenital strain could grow deprived of Trp and in the presence of indole (O'Neill *et al.*, 2018).

Contrary to previous conventions, chromosomal recombination is a natural and common source of variation within species and between biovars of *Chlamydia* and is consistent with the presence of DNA repair/recombination genes in the genome of *C. trachomatis* (Stephens *et al.*, 1998). Recombination events are observed in a large portion of the genome, rather than in just a few limited "hotspots" as once hypothesized (Gomes *et al.*, 2007; Harris *et al.*, 2012). Regions that are subjected to high recombination events include *ompA*, a genomic segment within the plasticity zone and a region encoding Pmps (Harris *et al.*, 2012). This has implications on the phylogenetic study or typing of *C. trachomatis* strains relying on such regions, that may not reflect the true genetic relationship within species. For those purposes, whole-genome analysis seems to be a more accurate option. Despite recombination events are observed, the genome of *Chlamydia* organisms lacks mobile genetic elements, phages and

genes encoding DNA restriction and modification systems, suggesting rare acquisition of foreign DNA by horizontal gene transfer. However, this event should have occurred in ancestral chlamydial species as *C. trachomatis* encodes genes exclusively found in eukaryotic cells like SET-domain containing proteins (NUE/CT737) or DNA helicases (CT555 and CT708) (Stephens *et al.*, 1998; Pennini *et al.*, 2010; Bastidas *et al.*, 2013).

1.2.3.2 Genetic manipulation of *C. trachomatis*

The obligate intracellular lifestyle and the unique developmental cycle of *C. trachomatis* have posed many challenges to the development of tools to genetically manipulate this bacterium. Until some years ago, mutagenesis and recombinant protein expression were not available in *C. trachomatis*, limiting the study of the molecular mechanisms underlying infection by this pathogen. However, genetic analysis tools that are generally used in other bacteria have been adapted to *C. trachomatis* and, despite being time-consuming and somehow prone to failure, they have propelled the development of the *Chlamydia* research field.

1.2.3.2.1 Transformation of *C. trachomatis*

Chlamydial EBs are encased by a cell wall composed of tightly cross-linked proteins that provide rigidity to this structure and protect EBs from environmental stress during dissemination (Hatch, 1996). Such feature has challenged the delivery of exogenous DNA into chlamydial EBs, which was first accomplished by Tam and colleagues using electroporation (Tam *et al.*, 1994). They successfully transformed EBs with a chimeric plasmid containing portions of both a *C. trachomatis* serovar E endogenous plasmid and an *E. coli* plasmid, and a chloramphenicol acetyltransferase (*cat*) cassette fused to a chlamydial promoter. However, chloramphenicol-resistant bacteria were only transiently recovered and were lost after four passages, which was attributed to transient expression of the *cat* cassette only during early stages of the development of RBs. Some years later, electroporation was used to introduce in *C. psittaci* EBs a plasmid harboring a synthetic *16S* rRNA gene homologous to the chromosomal copy (Binet and Maurelli, 2009). This *16S* rRNA allele contained four nucleotide substitutions, two of which conferred resistance to kasugamycin and spectinomycin. Cell monolayers were infected with these EBs and transformants resistant to both antibiotics were successfully recovered by plaque assay, meaning that they acquired the mutant *16S* rRNA allele by allelic exchange. Despite fruitful, these electroporation-based methods were not widespread likely due to their low efficiencies and to the high amount of DNA they required.

Delivery of DNA by complexation with dendrimers was also used to transform *C. pneumoniae* and *C. trachomatis* RBs. Transformation of a *C. trachomatis* plasmidless strain was accomplished by adding a *C. trachomatis*-*E. coli* vector complexed with dendrimers at 16 h post-

infection for three hours, and in the absence of antibiotics selection (Kannan *et al.*, 2013). Plasmid replication and expression of its eight plasmid open reading frames (ORFs) and of *glgA* (a gene encoding a glycogen synthase, whose expression was known to be diminished in a plasmidless strain) were confirmed. Moreover, transformation of that plasmidless strain with a plasmid enabling the production of GFP revealed an efficiency of about 80%. However, this transformation method was not broadly adopted likely due to its technical requirements.

A breakthrough in the *Chlamydia* research field was the development of a calcium chloride ($CaCl_2$)-based treatment for transformation of *C. trachomatis* EBs, by Wang and his colleagues (Wang *et al.*, 2011). This protocol was the first allowing the stable introduction of a shuttle plasmid into chlamydiae and is currently widespread due to its simplicity and reproducibility. For transformation of *C. trachomatis*, a *C. trachomatis*-*E. coli* shuttle vector was generated, and a penicillin resistance marker (specifically a β-lactamase resistance gene) was chosen due to the well-known effects penicillin has on *C. trachomatis*. This antibiotic hampers the conversion of RBs into EBs and induces the formation of giant, aberrant RBs that are easily distinguishable by phase-contrast microscopy. Therefore, plasmid acquisition by *C. trachomatis* enables the production of β-lactamase and consequent rescue from the penicillin-induced aberrant state, a phenotype that can be observed by microscopy. Furthermore, transformants can be selected by increasing the concentration of penicillin on each passage because aberrant RBs are not able to infect other host cells. The method used to transform *C. trachomatis* with the desired plasmid was adapted from the one used for chemical transformation of *E. coli*. Essentially, *C. trachomatis* EBs are incubated with the plasmid DNA in a $CaCl_2$ buffer and then added to human epithelial (HeLa) cells, and selection of transformants is done by applying increasing concentrations of penicillin on each round of infection (Wang *et al.*, 2011). Throughout this process, the recombinant plasmid replaces the native plasmid due to plasmid incompatibility and is stably propagated for several passages in the absence of penicillin.

With the availability of a reliable transformation protocol, several *C. trachomatis*-*E. coli* shuttle vectors enabling protein expression in *C. trachomatis* have been designed. Besides the vector generated for the development of the $CaCl_2$-mediated transformation protocol (Wang *et al.*, 2011), shuttle vectors allowing the expression of fluorescent proteins in *C. trachomatis* under the control of endogenous or inducible promoters, and with new selectable markers have been engineered (Agaisse and Derré, 2013; Wickstrum *et al.*, 2013; Bauler and Hackstadt, 2014). These vectors have subsequently been adapted to other purposes as targeted mutagenesis (see section 1.2.3.2.3 below) and to monitor aspects of the chlamydial developmental cycle as RB-to-EB conversion (Cortina *et al.*, 2019). Recently, a broad-host-range plasmid isolated from *Bordetella pertussis* that does not have a chlamydial origin of replication was transformed into, and successfully maintained by *C. trachomatis*, either in a free form or

integrated in the chromosome by allelic exchange (Garvin *et al.*, 2021). This could also be a useful and versatile tool to study *C. trachomatis*.

1.2.3.2.2 Random mutagenesis

Genetic alterations can be introduced on the DNA by random mutagenesis using chemical compounds such as ethylmethanesulfonate (EMS) or N-ethyl-N-nitrosourea (ENU). This method was used by Kari and colleagues on a first reverse genetic approach to generate isogenic *C. trachomatis* null mutants (Kari *et al.*, 2011a). By using low amounts of EMS, a mutant library of chlamydiae with ~0.5 mutations per genome was generated. Subpopulations were then expanded, and PCR was used to obtain amplicons spanning their *trpRBA* loci, which were hybridized with the wild-type *trpRBA* locus. A mutant *trpRBA* locus forms DNA heteroduplexes with mismatches that can be detected by digestion with the mismatch specific endonuclease CEL1. This allowed the detection of a *trpB* null mutant that was unable to survive under IFN-γ-induced tryptophan starvation. This method does not allow screening of essential genes though, while being laborious and costly and time-consuming.

In other report, a forward genetic approach to produce *C. trachomatis* mutants was developed (Nguyen and Valdivia, 2012). First, a library of rifampin-resistant (Rif^R) mutants was generated by chemical mutagenesis using EMS. Clones of chlamydial mutants were then isolated by plaque assay, and WGS was performed in mutants whose plaques displayed the same morphology to identify common genetic lesions. Then, cells were co-infected with mutant and wild-type spectinomycin-resistant (Spc^R) strains to generate Rif^R-Spc^R recombinants strains. Finally, the recombinant progeny was analyzed to establish a link between specific point mutations and the occurrence of a specific plaque morphology. This method was later improved to generate, isolate, and sequence the genomes of an arrayed collection of mutants, which can be used in reverse genetic applications (Kokes *et al.*, 2015).

Besides chemical compounds, the use of transposons to generate random mutations was also recently applied to *Chlamydia*. LaBrie and colleagues were the first to successfully adapt to *C. trachomatis* the Himar C9 transposon system (LaBrie *et al.*, 2019), which is widely used to study other bacteria (Beare *et al.*, 2009; Clark *et al.*, 2011). In this study, *C. trachomatis* was transformed with a suicide vector encoding the C9 Himar1 transposase and a β-lactamase (*bla*) gene flanked by inverted repeat sequences that are recognized by the transposase. A total of 105 β-lactam-resistant mutants containing nonspecific, stable insertions across the entire genome were recovered. However, the efficiency of this process is very low, and the authors of this study are currently trying to improve it by developing a self-replicating vector while limiting an eventual transposition-mediated toxicity (O'Neill *et al.*, 2021).

1.2.3.2.3 Site-directed mutagenesis

Replacement of a specific gene had already been reported by Binet and Maurelli in 2009, who successfully replaced by allelic exchange the *16S* rRNA allele of *C. psittaci* by a mutant allele (Binet and Maurelli, 2009) (see section 1.2.3.2.1 above). However, the first protocol allowing targeted disruption of a chromosomal gene was developed in 2013 (Johnson and Fisher, 2013). Johnson and Fisher adapted the TargeTron™ system (Sigma-Aldrich) for selectable gene inactivation in *C. trachomatis* by modifying a group II intron to target a specific location of a gene of interest, and to carry a resistance gene that allows selection of mutants. LtrA is an intron encoded protein that mediates insertion of the intron itself within the target gene, while also having the ability to splice it. However, in this adapted system, the *ltrA* gene was included in a suicide plasmid along with the retargeted intron under the control of a chlamydial promoter. When *C. trachomatis* is transformed with this plasmid, LtrA and the retargeted intron are expressed and LtrA mediates insertion of the intron in the target gene, disrupting it. As chlamydiae do not replicate this vector, LtrA is not produced and intron splice is not possible, yielding a stable and inheritable gene inactivation. As a proof of principle, the gene encoding inclusion membrane protein IncA was chosen for inactivation, as naturally occurring *incA* mutants already existed and were known to be viable (Johnson and Fisher, 2013). Limitations of this approach include limited positions within a gene that can be chosen for intron retargeting and a potential polar effect caused by intron insertion on neighboring genes. Nevertheless, this method has successfully allowed the generation of several *C. trachomatis* mutant strains [some examples are reported in (Thompson *et al.*, 2015; Sixt *et al.*, 2017; Weber *et al.*, 2017; Wesolowski *et al.*, 2017; Almeida *et al.*, 2018; Pais *et al.*, 2019)].

Subsequently, Mueller and colleagues developed a system for fluorescence-reported allelic exchange mutagenesis (FRAEM) in *C. trachomatis* (Mueller *et al.*, 2016). The vector generated for mutagenesis comprised a cassette encoding a selectable marker (β-lactamase) and a fluorescent marker (GFP) flanked by the DNA sequences corresponding to approximately 3 Kb of the genomic region upstream and downstream the gene of interest. Moreover, plasmid gene *pgp6* (essential for plasmid propagation) was placed under the control of a Tet-inducible promoter, enabling controlled plasmid removal. Upon transformation of *C. trachomatis*, the targeted chromosomal locus is replaced by the plasmid cassette by rare allelic exchange events. Mutants that have acquired the cassette can be isolated by antibiotic selection and will become detectable by fluorescence microscopy. FRAEM proved to be effective, as stable mutations were successfully obtained for *trpA* and for *tmeA*/*ct694*/*ctl0063*, *tmeB*/*ct695*/*ctl0064* and *ct696*/*ctl0065* (Mueller *et al.*, 2016).

The techniques described above only allow gene depletion by interruption, meaning that mutation of essential genes will render non-viable organisms. Recently, Ouellette proposed a CRISPR interference ($CRISPR_i$)-based conditional knockout system for *C. trachomatis*

(Ouellette, 2018) to bypass that limitation. CRISPR$_i$ is a technology in which a catalytically inactive version of the enzyme Cas9 recognizes a genomic target sequence through a cognate guiding RNA but is not able to cleave it. Cas9 binding to the target gene produces a steric block to the transcriptional machinery and, therefore, the gene is not transcribed. In this case, transcription of a target gene sequence can be selectively and reversibly blocked by encoding *cas9* in a vector under the control of an inducible promoter. As a proof-of-concept, the author used this method to try to knockout *incA* from *C. trachomatis* but observed chlamydial inclusions that did not produce IncA after removal of the inducer. This indicated an erratic reversibility of the *incA* repression likely due to plasmid instability, existence of off-target effects and leaky expression of *cas9* even in the absence of induction (Ouellette, 2018). While some of these issues were addressed in a subsequent study by modifying the vector backbone or controlling the production/degradation of Cas9, other system based on Cas12 was also shown to be an efficient tool for gene conditional knockout (Ouellette *et al.*, 2021).

1.3 The *Chlamydiae* plasmids

Most members of the Family *Chlamydiaceae* harbor a highly conserved plasmid of approximately 7.5 kBp (Carlson *et al.*, 2005; Zhong, 2017). Despite this fact, naturally existing plasmid-deficient strains have been described for the L2 serovar (Peterson *et al.*, 1990), and *C. pneumoniae* strains are usually devoid of a virulence plasmid (Read *et al.*, 2000). Besides the *Chlamydiaceae*, a plasmid was identified in the Genra *Simkania* and *Waddlia* but not in the *Parachlamydia* or *Protochlamydia* (Collingro *et al.*, 2011). *W. chondrophila* str WSU 86-1044 carries a 15.6-kBp plasmid (pWc) containing 22 predicted proteins, 2 of them sharing similarities with *Chlamydiaceae* plasmid encoded proteins (Bertelli *et al.*, 2010). On the other hand, a 132-kBp plasmid (pSn) was described for *S. negevensis* str Z (Collingro *et al.*, 2011). pSn encodes 138 predicted proteins, some of them possibly involved in pathogenicity, host adaptation or metabolic processes. Remarkably, pSn encodes a type IV secretion system and other proteins involved in F-type conjugation, suggesting this mechanism is present in *Simkania*. These features were likely acquired by a chlamydial ancestor but were lost during evolution of the pathogenic *Chlamydiaceae*, as conjugation-related genes are not encoded in the genomes or plasmids of its members (Collingro *et al.*, 2011).

1.3.1 The virulence plasmid of *C. trachomatis*

The prevalence of a virulence plasmid in different chlamydial strains and in most clinical isolates suggests that it could confer adaptive advantages. However, when plasmid-deficient *C. trachomatis* urogenital or L2 strains were isolated from patients, their ability to grow *in vitro*

was comparable to that of a plasmid-bearing L2 strain (Peterson *et al.*, 1990; Farencena *et al.*, 1997). This suggested that the plasmid is not necessary for *C. trachomatis* growth at least *in vitro*. Only when generating and studying a plasmid-deficient *C. muridarum* strain it was noted that it displays a different phenotype comparing to a wild-type strain, namely less accumulation of glycogen within the inclusion and the formation of smaller plaques during plaque assay (Matsumoto *et al.*, 1998; O'Connell and Nicks, 2006), which result from lower chlamydial-induced lysis of host cells. Indeed, the plasmid is now known to mediate *C. trachomatis* host cell lysis (Yang *et al.*, 2015) and to regulate the expression of the chromosomal genes required for glycogen synthesis (Carlson *et al.*, 2008; Song *et al.*, 2013; Patton *et al.*, 2018), which is consistent with the phenotypes observed for plasmid-deficient *C. muridarum*. Importantly, plasmid-deficient *C. muridarum* failed to cause infection in the upper genital tract of mice (O'Connell *et al.*, 2007), and the same was observed for a urogenital *C. trachomatis* isolate (Sigar *et al.*, 2014) and for an ocular *C. trachomatis* strain in the case of ocular infections in macaques (Kari *et al.*, 2011b). Altogether, these studies point to an important role of the virulence plasmid in the pathogenicity of *C. trachomatis* and *C. muridarum* in animal models. However, plasmid-free organisms are still able to infect human or animal tissues, which indicates that plasmid-independent factors are also involved in chlamydial pathogenesis.

The *C. trachomatis* virulence plasmid (pL2) is low copy, with about 4 or 8 unities per chlamydial EB or RB, respectively (Pickett *et al.*, 2005; Ferreira *et al.*, 2013). pL2 encodes eight ORFs (Thomas *et al.*, 1997) designated pORF1 to 8, which are all translated into proteins during infection (Li *et al.*, 2008b), and two small anti-sense RNAs (sRNAs; ORF2/Pgp8 and ORF7/Pgp5) (Ricci *et al.*, 1993; Albrecht *et al.*, 2010; Ferreira *et al.*, 2013). The designations of pL2 ORFs may be ambiguous in the literature as for example pORF1 is also named Pgp7 (and not Pgp1). To simplify, plasmid ORFs will be hereafter designated by Pgp1 to 8.

Mutagenesis by deletion or introduction of a premature termination sequence on each plasmid ORF have shown that Pgp1, 2, 6 and the DNA sequence of *pgp8*, but not the Pgp8 protein are essential for plasmid maintenance, whereas the other ORFs are not (Gong *et al.*, 2013; Song *et al.*, 2013). The sRNA anti-sense to *pgp8* is one of the most abundant transcripts in *C. trachomatis* (Albrecht *et al.*, 2010) and it could have a role in plasmid viability (Gong *et al.*, 2013). Pgp1 is homologous to a DnaB-like helicase, and Pgp7 and Pgp8 are putative recombinases/integrases (Albrecht *et al.*, 2010; Zhong, 2017). The functions of Pgp2 and Pgp6 are unknown (Thomas *et al.*, 1997; Zhong, 2017).

Pgp5 may be a negative regulator of the expression of genes that are positively regulated by Pgp4 (see section 1.3.2 below) (Liu *et al.*, 2014). When *pgp5* was deleted or interrupted, the expression of plasmid-dependent chromosomal genes increased, indicating that the Pgp5 protein is directly responsible for their inhibition. The sRNA anti-sense to *pgp5* is not crucial for this inhibition of chromosomal genes (Song *et al.*, 2013; Liu *et al.*, 2014), nor

for plasmid maintenance (Gong *et al.*, 2013; Song *et al.*, 2013), and its function is unknown. *pgp5*-deficient C. *muridarum* fails to ascend to the upper genital tract of mice and to induce inflammation, meaning that Pgp5 is important for the pathogenicity of C. *muridarum* in mice (Huang *et al.*, 2015).

Pgp3 is produced as a 28 kDa protein that localizes at the bacterial outer membrane (Comanducci *et al.*, 1993) and at the cytosol of host cells during infection of several plasmid-harboring *Chlamydia* strains (Li *et al.*, 2008b). Pgp3 is thought to be delivered into the host cell cytosol by a plasmid-dependent mechanism, in globular structures containing Pgp4-regulated proteins as GlgA, Pgp3 itself and other chlamydial proteins (Lei *et al.*, 2021). Most of the roles Pgp3 is thought to play *in vivo* have been studied using *C. muridarum* infection model of the genital tract of mice. The virulence plasmid and specifically Pgp3 are key players on gastrointestinal (Shao *et al.*, 2018; Ma *et al.*, 2020) and urogenital infections by *C. muridarum* in mice [reviewed by Zhong, (Zhong, 2017)]. A *C. muridarum pgp3*-deficient strain displays reduced pathogenicity, failing to ascend to, survive and induce inflammation in the mouse upper genital tract. Importantly, Pgp3 is highly immunogenic and contributes to anti-chlamydial immune responses in the host, thus being a target for the development of vaccines against *Chlamydia* (Zhou *et al.*, 2022). *C. trachomatis* Pgp3 was described to target host antimicrobial peptides *in vitro* (Hou *et al.*, 2015) and in mouse and non-human primate infection models, which is thought to be important in persistent infections (Yang *et al.*, 2020).

Pgp4 is the smallest plasmid encoded protein, with 102 amino acid residues. Pgp4 is a virulence factor and a key regulator of plasmid and chromosomal genes, and its regulatory function is going to be discussed in the next section.

1.3.2 Plasmid-mediated gene regulation

The eminent role of the virulence plasmid in *Chlamydia* pathogenesis has been correlated with the regulation of several plasmid and chromosomal genes. Carlson and colleagues observed that a clinically isolated plasmid-deficient strain [L2(25667R); (Peterson *et al.*, 1990)] can infect tissue culture cells as a wild-type strain, as already mentioned (see section 1.3.1 above), but fails to colonize the genital tract of mice (Carlson *et al.*, 2008). Moreover, transcriptomic analysis of both strains revealed that the gene *glgA*, which encodes the glycogen synthase GlgA, was downregulated in L2(25667R). This correlated with an observed lack of glycogen accumulation for this strain (Matsumoto *et al.*, 1998). In total, 22 chromosomally encoded genes showed significantly altered expression levels in L2(25667R) comparing to plasmid-bearing, wild-type L2 (Carlson *et al.*, 2008). Genes that were downregulated included *glgA*, whereas genes that were upregulated comprised *ctl0233/ct858/cpaf* which encodes the well-known

chlamydial protein CPAF. These observations indicated that the plasmid of *C. trachomatis* is a transcriptional regulator of chromosomal genes (Carlson *et al.*, 2008).

Five years later, with the development of a transformation protocol for *C. trachomatis*, the same group generated *C. trachomatis* strains harboring plasmids deleted in each of the eight ORFs (Song *et al.*, 2013). They observed that a *C. trachomatis* plasmid-deficient strain (L2R) was phenocopied by a *pgp4* mutant strain (L2RpΔ*pgp4*) in terms of inclusion morphology and glycogen accumulation. Based on this, the transcriptional profiles of L2R and of L2RpΔ*pgp4* were analyzed by comparison with their respective wild-type strains harboring an intact plasmid. Curiously, eight of the chromosomal genes found to have differential expression were consistent for both L2R and L2RpΔ*pgp4* strains, showing that *pgp4* is the plasmid element involved in regulation of those genes. Additionally, the expression of *pgp3* was also downregulated in L2RpΔ*pgp4* (Song *et al.*, 2013), suggesting that *pgp4* also modulates the expression of plasmid genes. A posterior proteomics study showed that the production of Pgp4 is associated with increased expression levels of a conserved set of chromosomally and plasmid encoded proteins (Patton *et al.*, 2018), consistent with the previous transcriptomics studies (Carlson *et al.*, 2008; Song *et al.*, 2013). In agreement with earlier results (Carlson *et al.*, 2008), this study also showed that the plasmid is involved in the negative regulation of *C. trachomatis* CPAF, but this does not seem to be Pgp4-dependent (Patton *et al.*, 2018).

Recently, a novel role for Pgp4 as a negative regulator of chlamydial late gene expression has been described (Zhang *et al.*, 2020). In this study, Pgp4 was found to bind to the repressor of chlamydial late genes EUO and enhance its ability to bind and repress EUO-regulated promoters, avoiding their premature expression. Nevertheless, and consistent with previous reports, Pgp4 also acts as a positive gene regulator of chlamydial transcription in a EUO-independent manner (Zhang *et al.*, 2020). Altogether, these findings support that Pgp4 is a master regulator of the expression of plasmid and chromosomal genes in *C. trachomatis*.

Pgp4 is a virulence factor that has also been implied in plasmid-mediated *C. trachomatis* lytic exit from host cells (Yang *et al.*, 2015). Similar to mutation of *pgp4*, presumed inhibition of the type III secretion system of *C. trachomatis* by the compound C1 prevented chlamydial host cell lytic exit. Hence, the authors suggested that *C. trachomatis* Pgp4-dependent lytic exit possibly involves regulation of a chromosomal type III secretion-related gene (Yang *et al.*, 2015).

1.4 *C. trachomatis* pathogenesis

1.4.1 The unique developmental cycle

As already mentioned, the *C. trachomatis* developmental cycle is shared among *Chlamydiae* and comprises two interconvertible, morphologically distinct forms - EBs and RBs [Figure 1.3; reviewed in (AbdelRahman and Belland, 2005; Elwell *et al.*, 2016)]. Additionally, an RB-to-EB conversion intermediate form has also been recognized and named intermediate body (IB) (Phillips *et al.*, 1984; Lee *et al.*, 2018). The small EB [~ 0.35 μm in diameter (Lee *et al.*, 2018)] is infectious and non-replicative and is the extracellular form of *C. trachomatis*. EBs possess highly condensed nucleic acid content and a rigid outer membrane complex majorly composed of the abundant MOMP and of the two cysteine-rich proteins OmcA and OmcB (Caldwell *et al.*, 1981; Everett and Hatch, 1991). These proteins cross-link by disulfide bonds, forming a lattice that confers resistance to osmotic and mechanical stress during dissemination (Hackstadt *et al.*, 1985; Hatch, 1996; AbdelRahman and Belland, 2005; Omsland *et al.*, 2014). Proteomics studies have shown that EBs, once thought to be metabolically inactive, are primed for type III secretion capacity and glucose catabolism likely required in the initial events of infection (Saka *et al.*, 2011). On the other hand, the RB [~ 1.2 μm in diameter (Lee *et al.*, 2018)] develops from EB differentiation and is the non-infectious, replicative form of *C. trachomatis*. RBs have a less condensed genetic material and are more fragile than EBs due to a lower degree of protein cross-linking in their outer membrane complex (Hackstadt *et al.*, 1985; Omsland *et al.*, 2014). RBs have a high metabolic activity of protein synthesis, nutrient transport, and accumulation of ATP necessary for replication and transition into EBs (Saka *et al.*, 2011).

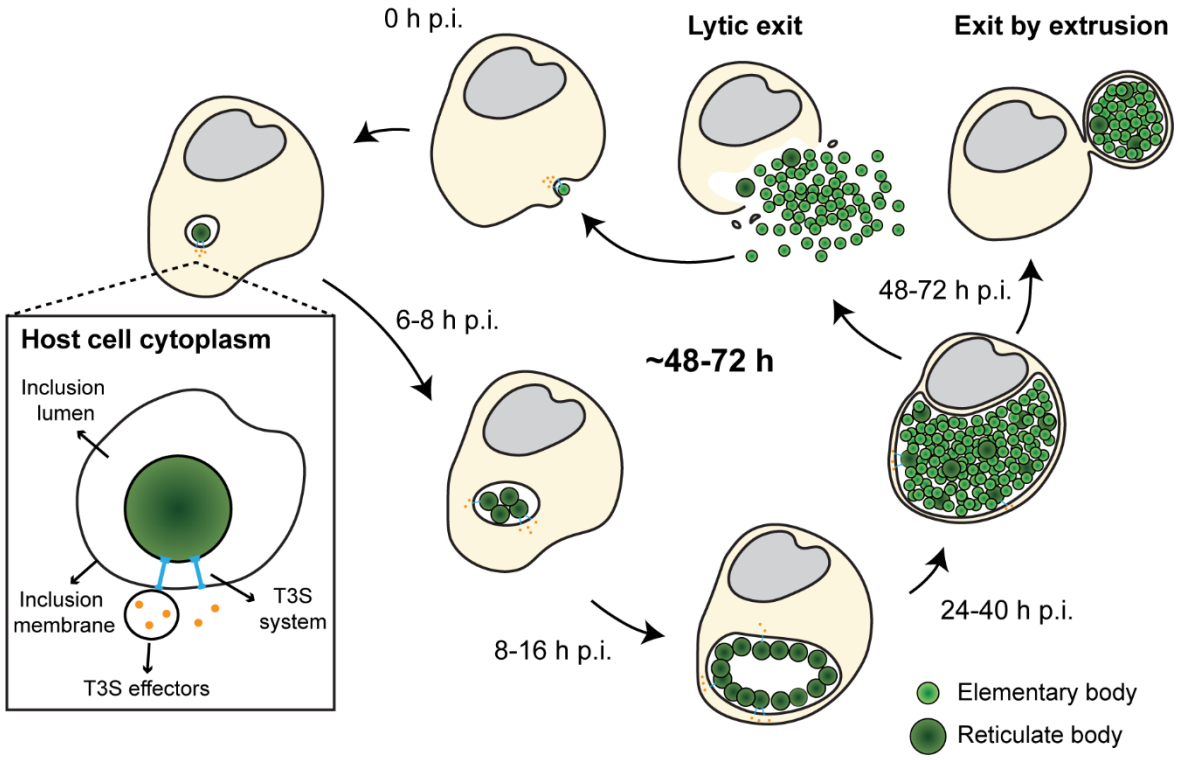


Figure 1.3. Developmental cycle of *C. trachomatis*. The developmental cycle shared among *Chlamydiae* species alternates between an infectious, non-replicative elementary body (EB, light green) and a replicative, non-infectious reticulate body (RB, dark green). After attachment to host cells, EBs are internalized in a membrane-bound vacuole, forming the nascent inclusion. At 6-8 h post-infection (p.i.), EBs differentiate into RBs which actively replicate (8-16 h p.i.) and form a mature inclusion. Between ~ 24 and 40 h p.i., RBs re-differentiate back into EBs which ultimately exit the intracellular niche by extrusion of the entire inclusion or lysis of the host cell (48-72 h p.i., depending on the chlamydial species). Throughout this cycle, *C. trachomatis* uses a T3S system (blue) to deliver virulence proteins known as T3S effectors (orange) across the bacterial and inclusion membranes into the host cell cytosol. These effectors interfere with numerous host cell pathways to promote bacterial growth.

C. trachomatis EBs attach to host cells through low- and high-affinity interactions between chlamydial adhesins (such as OmcB, MOMP, chlamydial LPS and Pmps) and diverse host cell receptors [as heparan sulfate proteoglycans (HSPGs), the mannose 6-phosphate receptor and β1 integrin] (Elwell *et al.*, 2016). Upon attachment, *C. trachomatis* induces host cell actin remodeling which leads to internalization of EBs in a membrane-bound vacuole (Figure 1.3), forming the nascent inclusion (Elwell *et al.*, 2016). About 2 h after invasion, the nascent inclusion is transported along microtubules to the centrosome/microtubule-organizing center (MTOC) at the perinuclear region (Richards *et al.*, 2013) and the content of the vacuolar membrane is modified to avoid degradation via the host endocytic pathway (late endosomes and lysosomes) (Hackstadt, 2014). Simultaneously, fusion of the vacuole with vesicles rich in nutrients as sphingomyelin and cholesterol is promoted (Fields and Hackstadt, 2002). The EBs differentiate into RBs and at about 6-8 h post-infection RBs start multiplying (Figure 1.3) likely by binary fission (Lee *et al.*, 2018), although polarized cell division has also been proposed (Abdelrahman *et al.*, 2016). Multiple rounds of chlamydial replication cause the vacuole to

increase in volume and form an inclusion (Figure 1.3). These events are sustained by the acquisition of host cell nutrients such as lipids and amino acids. In fact, the inclusion interacts with several host compartments or pathways for nutrient scavenge, namely with the Golgi apparatus and multivesicular bodies for lipid acquisition (Cocchiaro *et al.*, 2008; Robertson *et al.*, 2009; Capmany and Damiani, 2010), with lipid droplets and peroxisomes possibly for uptake of triacylglycerides and metabolic enzymes (Boncompain *et al.*, 2014; Bugalhão *et al.*, 2022), and with the endoplasmic reticulum to access host proteins involved in lipid transport and signaling (Derré *et al.*, 2011; Agaisse and Derré, 2015) [reviewed in (Elwell *et al.*, 2016)]. From about 24 h post-infection, RBs asynchronously re-differentiate into EBs (Figure 1.3) by a mechanism that possibly depends on the size of RBs (Lee *et al.*, 2018). Finally, the produced EBs exit the host cell and infect neighboring cells. *C. trachomatis* may exit host cells by two mutually exclusive pathways - lytic exit or extrusion [Figure 1.3; (Hybiske and Stephens, 2007)]. On the pathway of lytic exit, there is sequential lysis of the inclusion and host cell membranes followed by release of chlamydial EBs and host cell death. This pathway was shown to depend on intracellular calcium levels and proteases, specifically cysteine proteases (Hybiske and Stephens, 2007). The extrusion pathway involves release of the entire inclusion to the extracellular milieu without host cell lysis, in a process where actin polymerization, small GTPases of the Rho family, and myosin II have been shown to play a role (Hybiske and Stephens, 2007). Although the mechanisms underlying both exit routes are still poorly understood, it is known that they are mechanistically different and may involve different host and chlamydial proteins (Hybiske and Stephens, 2007; Yang *et al.*, 2015; Nguyen *et al.*, 2018; Shaw *et al.*, 2018; Zuck and Hybiske, 2019).

The duration of the *C. trachomatis* developmental cycle (Figure 1.3) depends on the strain. Genital strains take between 30 and 48 h to complete their developmental cycle, whereas ocular strains do not complete theirs until 48 to 68 h. The developmental cycle of *C. trachomatis* serovar L2 strains is around 48 h (Miyairi *et al.*, 2006). *C. trachomatis* RBs may enter in a reversible persistent state during the developmental cycle where they remain viable, but display an enlarged, aberrant morphology (known as aberrant bodies) and decreased metabolism. This persistent state may allow chlamydiae to survive under conditions unfavorable to replication such as antibiotic pression, nutritional deprivation or exposure to host cell immune molecules (as IFN-γ) (Hogan *et al.*, 2004).

1.4.2 Secretion systems in *Chlamydiae*

Bacteria have evolved several specialized macromolecular nanomachineries known as secretion systems to transport a wide range of substrates across their membranes, including proteins, DNA, and small molecules [reviewed in (Costa *et al.*, 2015)]. These substrates allow

bacteria to interact with the environment and participate in several physiological processes such as adhesion, adaptation, pathogenicity, nutrient scavenging, and competition with other species. In the case of Gram-negative bacteria, proteins must cross the inner and outer bacterial membranes and, in some cases, a peptidoglycan layer to reach the outer membrane, the extracellular space, or a target eukaryotic or bacterial cell. Protein secretion systems either span both the inner and outer membranes, as the type I secretion (T1S), T2S, T3S, T4S and T6S systems, or just the outer membrane, being this the case of the T5S system (Costa *et al.*, 2015). Moreover, translocation of substrates can occur in a one-step or in a two-step manner. In the former case, substrates are transported directly from the bacterial cytoplasm into the extracellular space or target cell, as observed for the T1S, T3S, T4S and T6S systems, and in the latter case substrates are first transported into the periplasm via an inner membrane-spanning transporter and then to the outer membrane or extracellular space by outer-membrane transporters (Costa *et al.*, 2015), as in the case of T2S and T5S systems. Inner membrane transporters include the Sec translocon and the twin-arginine translocation (Tat) system (Palmer and Berks, 2012; Park and Rapoport, 2012), which also exist in Gram-positive bacteria.

Within the context of bacterial pathogenesis, the most studied protein secretion machines are the T3S, T4S and T6S systems, which function as protein-injection machineries. The T3S and T4S systems are evolutionarily and structurally different but have the same aim of delivering bacterial effectors into eukaryotic host cells to modulate their processes. The unrelated T6S system mainly delivers proteins into other bacteria for inter-bacterial competition purposes, although a role in host-pathogen interactions is not discarded (Galán and Waksman, 2018). Overall, these functions are related but different, indicating that these protein secretion systems have evolved from ancestors with various roles such as motility (T3S system), DNA transfer (T4S system) or bacteriophage infection (T6S system) (Galán and Waksman, 2018).

The *C. trachomatis* genome encodes a T2S system, a T3S system, a T5S system and components of the Sec pathway (Figure 1.4), namely an ortholog of the SecY translocase (CT510) and predicted signal peptidases (CT020 and CT408) that cleave Sec-dependent secretion signals (present in Pmps, for example). Conversely, *C. trachomatis* lacks a T4S secretion system, which exists among environmental *Chlamydiae*, and there is no evidence for the presence of a Tat pathway or a T6S system (Stephens *et al.*, 1998; Fields, 2012).

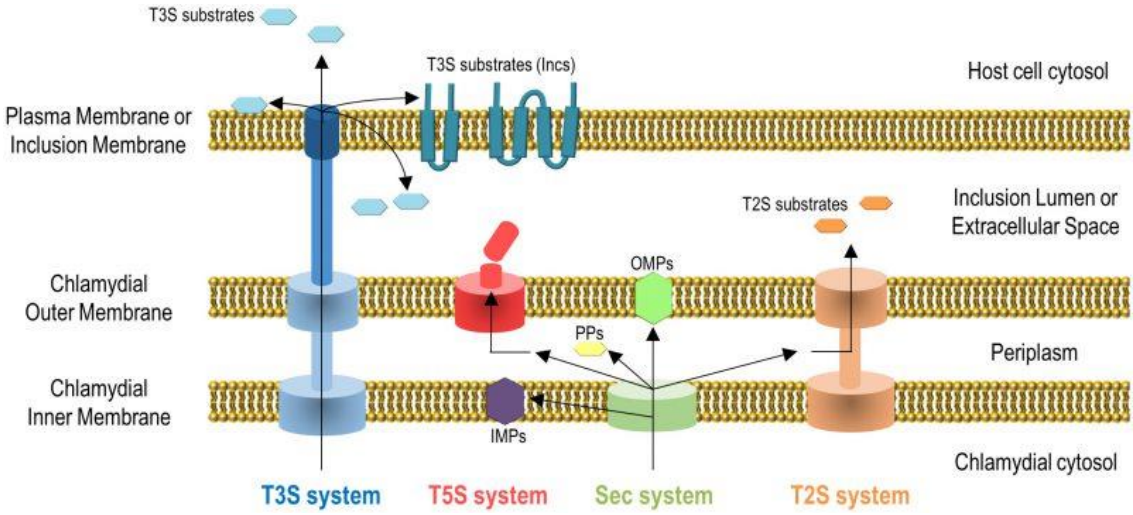


Figure 1.4. Secretion systems used by *C. trachomatis* to transport proteins to distinct chlamydial and host cell compartments. Secretion systems encoded by *C. trachomatis* comprise the T2S, T3S and T5S systems and the Sec system. The T3S system directly transports proteins (T3S substrates) from the chlamydial cytoplasm and across the bacterial membranes into the host cell cytoplasm, into the inclusion lumen or into the inclusion membrane (Incs). The mechanism of transport and/or membrane insertion of some of these proteins is still unclear. IMPs, inner membrane proteins; PPs, periplasmic proteins; OMPs, outer membrane proteins. Reprinted from (Bugalhão and Mota, 2019).

1.4.2.1 The T3S system of *C. trachomatis*

The T3S system is encoded by a wide range of Gram-negative bacteria that establish symbiotic or pathogenic interactions with a variety of hosts including vertebrates, plants, and insects (Hu *et al.*, 2017). Non-flagellar T3S systems are thought to have evolved from flagellar structures that experienced a series of feature loss or acquisition from other molecular systems to become proficient in translocating proteins (Cornelis, 2006; Abby and Rocha, 2012). The main function of these T3S systems is to directly deliver effector proteins from the bacteria into the cytoplasm of host eukaryotic cells (Figure 1.4) to modulate host cellular pathways. T3S genes are typically acquired by horizontal gene transfer (Gophna *et al.*, 2003) and, in many bacteria, they are encoded in virulence plasmids (as in *Shigella* spp. and *Yersinia* spp.) or within the same chromosomal locus, in pathogenicity islands (as in *Salmonella* spp.) (Hueck, 1998).

The first observation of the *Chlamydia* T3S system was likely done by Matsumoto and colleagues, who observed rosette-like structures on the surface of *C. psittaci* EBs and RBs (Matsumoto, 1973). Indeed, it was subsequently confirmed that not only the genomes of all *Chlamydiaceae* species, but also those of members of the environmental *Chlamydiae* encode T3S genes (Stephens *et al.*, 1998; Horn *et al.*, 2004; Fields, 2012; Dharamshi *et al.*, 2020). In *C. trachomatis*, T3S genes account for roughly 10% of the encoding capacity of its genome (Betts-Hampikian and Fields, 2010). Contrary to a plasmid location or a pathogenicity island configuration observed in other species, the T3S genes of *Chlamydia* are encoded in separated

regions of the chromosome, possibly in ten different operons (Stephens *et al.*, 1998; Hefty and Stephens, 2007). Usually the G/C content of T3S genes is lower comparing to the rest of the genome, but the G/C content of *C. trachomatis* T3S genes equals that of the genome (~40%) (Stephens *et al.*, 1998). Along with an absence of gene integration remnant sequences flanking the T3S genes, these observations support the fact that the acquisition of T3S genes was likely very ancient (Abby and Rocha, 2012).

Several bacterial species encode more than one type of T3S system, but the architecture of T3S system components is highly conserved (Hu *et al.*, 2017). The T3S system core structure comprises a multiprotein injectisome consisting of a basal body made of several stacked rings that span the inner and outer bacterial membranes, and that are linked by a central tube (also named neck). A needle-like filament extends from the basal body, protruding from the bacterial surface into the extracellular space, and forms a helical tubular structure that serves as a channel for protein delivery into the interior of host cells. In contact with the host cell, translocator proteins are secreted through the channel and form a pore in the host cell membrane named translocon. In some cases, the needle-like filament terminates in a needle tip complex which upon contact with host cells may generate an activating signal for the secretion-machine components in the bacterial cytoplasm or serve as an assembly platform for the pore itself (Cornelis, 2006; Costa *et al.*, 2015; Galán and Waksman, 2018).

Besides T3S components of the injectisome and T3S effectors, T3S chaperones also play an important role in adequate substrate secretion. The functions of T3S chaperones towards their substrates include: i) keeping them in a partially unfolded state, conformationally suitable for secretion; ii) preventing their degradation; iii) avoiding unproductive/premature interactions with their substrates; iv) directly target them for secretion by direct interaction with the T3S components (Mueller *et al.*, 2014).

1.4.2.2 *C. trachomatis* T3S effectors

T3S effectors act by direct association with or by enzymatic modification of their targets, or by mimicking host cell proteins (Mueller *et al.*, 2014). Although the T3S apparatus is very conserved among bacteria, T3S effectors vary greatly in terms of sequence identity. T3S effectors possess within the first ~20 N-terminal amino acids a secretion signal that, although poorly conserved, is characterized for being non-cleavable, intrinsically unstructured, and rich in serine, threonine, isoleucine, and proline residues (Galán and Waksman, 2018). Apart from the absence of genetic tools to study *Chlamydia*, the lack of similarity with proteins of known function and of an easily distinguishable secretion signal challenged early identification of chlamydial T3S effectors. Nevertheless, complex bioinformatics analyses relying on the identification of putative secretion signals have been used to identify new T3S effectors

(Arnold *et al.*, 2009; Löwer and Schneider, 2009; Samudrala *et al.*, 2009; Dehoux *et al.*, 2011). Additional methodologies to identify new T3S effectors comprised the analysis of the effect of the overexpression of individual proteins in *Saccharomyces cerevisiae* (Sisko *et al.*, 2006) or the use of genetically tractable heterologous host bacteria as *Shigella* (Subtil *et al.*, 2005; Dehoux *et al.*, 2011), *Salmonella* (Ho and Starnbach, 2005) and *Yersinia* (da Cunha *et al.*, 2014).

Validation or further investigation of the predicted chlamydial T3S effectors was carried by direct observation of their localization at the host cell cytoplasm or at the inclusion membrane of infected cells by immunofluorescence microscopy [examples of studies, (Lu *et al.*, 2013; Dumoux *et al.*, 2015; Weber *et al.*, 2015; Pais *et al.*, 2019)] or by using a reporter assay to monitor their delivery into eukaryotic host cells [for example, by fusing proteins to β-lactamase (Mueller and Fields, 2015), Cya or GSK (Bauler and Hackstadt, 2014), or by using the split-GFP method (Wang *et al.*, 2018)]. Additionally, possible functions and/or interacting partners of T3S effectors have been determined by protein-protein interaction assays (Scidmore and Hackstadt, 2001; Almeida *et al.*, 2018) or by ectopic expression in mammalian cells (Mirrashidi *et al.*, 2015; Pais *et al.*, 2019; Bugalhão *et al.*, 2022). With the development of new genetic tools to manipulate *C. trachomatis*, the knowledge about the function of chlamydial T3S effectors during infection has been considerably expanding.

1.4.2.2.1 Inclusion membrane proteins (Incs)

A major class of *C. trachomatis* T3S effectors are the inclusion membrane proteins (Incs). Incs are present in all *Chlamydiae* (Heinz *et al.*, 2010; Dehoux *et al.*, 2011; Kebbi-Beghdadi *et al.*, 2019), but only some are conserved among *Chlamydia* species (Dehoux *et al.*, 2011; Lutter *et al.*, 2012). As the name indicates, Incs insert into the inclusion membrane, in the interface between the inclusion and the host cell cytoplasm, being thus key mediators of chlamydiae-host cell interactions. Incs have distinctive single or multiple bilobed hydrophobic domains which likely mediate their insertion into the inclusion membrane (Rockey *et al.*, 2002). These domains consist of two closely spaced transmembrane regions that are separated by a short hairpin loop, with their N-terminal and/or C-terminal regions predicted to face the cytoplasm of the host cell. While Incs share minimal primary sequence identity with each other, the characteristic bilobed hydrophobic domain enabled the identification of putative Incs by bioinformatics analyses (Bannantine *et al.*, 2000; Dehoux *et al.*, 2011). To date, ~60 putative Incs were identified in *C. trachomatis* but only about 36 have their predicted inclusion membrane localization experimentally confirmed (Bugalhão and Mota, 2019).

Incs are the most studied T3S effectors of *C. trachomatis*. They have been categorized in terms of the chlamydial developmental cycle stage in which their corresponding messenger RNA (mRNA) levels are higher into early-cycle (between about 2 to 6 h post-infection), mid-cycle (between about 6 to 20 h post-infection) or late-cycle (after about 20 h post-infection) Incs

(Shaw *et al.*, 2000; Nicholson *et al.*, 2003). Incs participate in the modulation of several host organelles and/or pathways such as: i) the host cell cytoskeleton [CT223/IPAM, (Dumoux *et al.*, 2015); CT813/InaC, (Kokes *et al.*, 2015; Wesolowski *et al.*, 2017); CT850, (Mital *et al.*, 2015)], ii) the host cell vesicular and non-vesicular trafficking [CT229/CpoS, (Faris *et al.*, 2019); CT115/IncD (Derré *et al.*, 2011); CT116/IncE, (Mirrashidi *et al.*, 2015; Elwell *et al.*, 2017); CT119/IncA, (Delevoye *et al.*, 2008)], iii) host cell death [CT229/CpoS, (Sixt *et al.*, 2017); CT135 (Bishop and Derré, 2022)], iv) the stability [CT229/CpoS, CT383, and CT233/IncC, (Weber *et al.*, 2017)] and positioning [CT850, (Mital *et al.*, 2015)] of the inclusion, and v) chlamydial exit from host cells [CT228 (Lutter *et al.*, 2013; Shaw *et al.*, 2018); CT101/MrcA (Nguyen *et al.*, 2018); CT135 (Bishop and Derré, 2022)] [the main confirmed or putative functions of all known Incs are reviewed in (Bugalhão and Mota, 2019)]. Because Incs were not the focus of this work, a selection of the most studied Incs that are described to manipulate relevant host pathways and/or to participate in important chlamydial events during different stages of infection will be next described.

Incs modulating the host cell cytoskeleton and the Golgi

C. trachomatis manipulates and remodels the host cell cytoskeleton, including microtubules (Al-Zeer *et al.*, 2014; Dumoux *et al.*, 2015), actin filaments and intermediate filaments (Kumar and Valdivia, 2008; Tarbet *et al.*, 2018), and septins (Volceanov *et al.*, 2014). During *C. trachomatis* infection, microtubules and actin filaments accumulate around the inclusion, forming a "cage" that possibly contributes to inclusion stability. Moreover, the manipulation of microtubules is associated with maintenance and positioning of the inclusion during infection (Mital *et al.*, 2015). Besides the cytoskeleton, the host cell Golgi complex is extensively remodeled and distributes around the inclusion, which is thought to facilitate the acquisition of lipids by *C. trachomatis* during infection (Heuer *et al.*, 2009; Al-Zeer *et al.*, 2014).

Inc CT223, also named inclusion protein acting on microtubules (IPAM), was identified bioinformatically through the unique bilobed hydrophobic domain of Incs and it was observed at the chlamydial inclusion membrane by immunofluorescence microscopy (Bannantine *et al.*, 2000; Li *et al.*, 2008a; Weber *et al.*, 2015). The ectopic expression of IPAM leads to the appearance of multinucleated cells and to the occurrence of centrosomal supranumeracy, indicating a role of IPAM in blocking host cell cytokinesis (Alzhanov *et al.*, 2009). Upon the observation that microtubules assemble in a cage around the inclusion and assuming that a chlamydial protein should be involved in this process, IPAM was found to co-localize in patches with the centrosome and to destabilize the pericentriolar matrix, perturbing microtubule organization and assembly (which originated the designation IPAM) (Dumoux *et al.*, 2015). The centriolar protein CEP170 was identified as an interacting partner of IPAM and is necessary for its action on microtubule disruption. The interplay between IPAM and CEP170 on the reorganization of

host microtubules was further confirmed in *C. trachomatis*-infected cells (Dumoux *et al.*, 2015). Hence, a possible role of IPAM is to remodel microtubules in infected cells through CEP170.

Another Inc named CT813/InaC was first observed in the chlamydial inclusion membrane by Chen and colleagues using antibodies raised against the endogenous protein (Chen *et al.*, 2006), which was posteriorly validated (Li *et al.*, 2008a). Ectopically expressed CT813 displayed a cytoskeleton-like distribution, suggesting its interaction with this host cell structure (Chen *et al.*, 2006). Kokes and colleagues identified chemically induced mutants that failed to recruit F-actin to the inclusion and one of them had a nonsense mutation in *ct813* (Kokes *et al.*, 2015). After associating the absence of CT813 with a defect in recruitment and assembly of F-actin around the inclusion, the authors named it inclusion membrane protein for actin assembly (InaC). CT813 and intact F-actin filaments were necessary for Golgi redistribution around the inclusion, and CT813 was found to interact with eukaryotic ARF GTPases and with 14-3-3 proteins (Kokes *et al.*, 2015). In other study, using a *ct813* knockout *C. trachomatis* strain, CT813 was shown to directly mediate the recruitment of ARF1 and ARF4 GTPases to the inclusion membrane. These ARF GTPases were proposed to induce posttranslationally modified microtubules that distribute around the inclusion. This microtubule configuration led to a redistribution of the host Golgi complex around the inclusion in a process that was independent on actin filaments (Wesolowski *et al.*, 2017). Therefore, CT813 likely plays a role in inducing Golgi redistribution around the inclusion during *C. trachomatis* infection.

Incs modulating the host cell vesicular trafficking

During infection, *C. trachomatis* manipulates host cell vesicular trafficking to avoid degradation by the endolysosomal pathway and to hijack vesicles containing nutrients necessary for growth and survival (see section 1.4.1 above). At least *C. trachomatis* Incs CT229/CpoS, CT119/IncA and CT116/IncE have been shown to be directly involved in these processes.

IncA, IncE and CpoS were bioinformatically and experimentally confirmed to be Incs before the development of genetic manipulation techniques in *C. trachomatis* (Bannantine *et al.*, 1998, 2000; Scidmore-Carlson *et al.*, 1999; Li *et al.*, 2008a). IncA has a well-established role in promoting the fusion of multiple chlamydial inclusions within a single eukaryotic cell into a single inclusion, in a mechanism known as homotypic fusion (Suchland *et al.*, 2000; Johnson and Fisher, 2013). In eukaryotic cells, membrane fusion involves soluble N-ethylmaleimide-sensitive factor attachment protein receptors (SNAREs), and IncA possesses two SNARE-like motifs that are required for fusion between inclusions (Weber *et al.*, 2016). A *C. trachomatis incA* mutant strain has lower levels of SNARE proteins localizing at the inclusion membrane and it was shown that IncA interacts with and recruits host cell SNAREs to that compartment

(Delevoye *et al.*, 2008). Moreover, both SNARE-like domains of IncA are also involved in inhibition of SNAREs-mediated membrane fusion likely to avoid fusion with the endocytic pathway (Paumet *et al.*, 2009; Ronzone *et al.*, 2014). Thus, IncA has a dual function of promoting fusion between *C. trachomatis* inclusions and inhibiting fusion with endocytic vesicles.

In a large proteomics study using cells ectopically expressing several chlamydial proteins, IncE was found to interact with and recruit to the chlamydial inclusion sorting nexin (SNX) proteins, including SNXs 5 and 6 (Mirrashidi *et al.*, 2015). SNXs are proteins that belong to the eukaryotic retromer, a trafficking pathway that recycles cargo from endosomes to the plasma membrane or to the trans-Golgi network (TGN). IncE was shown to interfere with the normal retromer function, which was associated with recruitment of SNXs 5 and 6 to the inclusion membrane. Moreover, depletion of SNXs 5 and 6 enhanced the production of infectious progeny by *C. trachomatis*, suggesting that the retromer restricts chlamydial growth, and IncE may play a role in counteracting this effect (Mirrashidi *et al.*, 2015). Subsequently, the crystal structure of the C-terminus of IncE complexed with SNX5 was solved and an interaction between the cation-independent mannose-6-phosphate receptor (CI-M6PR, a source of newly synthesized lysosomal enzymes that is recycled via the retromer) and SNX5 was unveiled. Moreover, it was found that IncE competes with CI-M6PR for binding to SNX5 and that the SNX5:CI-M6PR interaction is inhibited during *C. trachomatis* infection (Elwell *et al.*, 2017; Paul *et al.*, 2017; Sun *et al.*, 2017). These studies suggest that IncE may interfere with retromer and lysosomal functions by binding to host cell SNXs.

CpoS was first found to interact with Rab4 by yeast two-hybrid and pull-down assays (Rzomp *et al.*, 2006), and later with several other Rab proteins by co-immunoprecipitation of cells ectopically expressing CpoS or cells infected by *C. trachomatis* (Mirrashidi *et al.*, 2015; Sixt *et al.*, 2017; Faris *et al.*, 2019). Rabs belong to a large family of GTPases which coordinate several steps of the host cell vesicular trafficking such as vesicle formation/motility and tethering of vesicles with their target membranes (Zerial and McBride, 2001). While searching for modulators of chlamydial cell death, Sixt and colleagues observed that mutation of *cpoS* caused enhanced *C. trachomatis*-induced cytotoxicity and sought to identify CpoS interacting partners. CpoS was confirmed to interact with several Rabs (for example, with Rabs 1, 2, 6, 8, 10, 14, 34 and 35). Inhibition of vesicular transport was further shown to block STING-mediated cell death, leading to the conclusion that CpoS suppresses this type of cell death possibly by interfering with the host cell trafficking (Sixt *et al.*, 2017). In other report, Faris and colleagues confirmed the recruitment of some of the Rabs found by Sixt *et al.* to the inclusion membrane (Faris *et al.*, 2019). A decrease in that recruitment was observed for a generated *C. trachomatis cpoS* mutant strain in both studies. CpoS is required for the accumulation of transferrin (via a transferrin receptor, which normally traffics from the plasma membrane to endosomes in a clathrin-dependent transport and is then recycled in a process dependent of

Rabs) and of CI-M6PR (which traffics from the TGN to endosomes in a clathrin- and Rab-dependent transport and is recycled by a Rab-dependent process) near the inclusion membrane (Bugalhão and Mota, 2019; Faris *et al.*, 2019). Overall, these observations suggest that CpoS targets multiple host Rabs, possibly allowing *C. trachomatis* to avoid the endolysosomal pathway while promoting interactions with the recycling pathway for nutrient acquisition.

Incs controlling chlamydial exit from host cells

Egress from host cells is still a poorly studied feature of the *C. trachomatis* developmental cycle, but it has been subject of research over the last few years. As already referred, *C. trachomatis* exits host cells either by lytic exit or extrusion (see section 1.4.1 above). Two Incs have been described to play a role in the extrusion pathway of *C. trachomatis* egress - CT228 and CT101/MrcA. Extrusion was found to depend on the actin motor protein non-muscle myosin II (Hybiske and Stephens, 2007). Myosin II comprises a motor and contractile subunit (myosin isoforms IIA and IIB) and a regulatory light chain subunit [myosin light chain 2 (MLC2)]. When MLC2 is phosphorylated, the motor activity of myosin II is enhanced. The phosphorylation state of MLC2 is determined by the countering activities of myosin phosphatase (which dephosphorylates MLC2, inhibiting myosin II) and myosin kinase (MLCK, which phosphorylates MLC2, activating myosin II). Conversely, the activity of these enzymes is also regulated - MLCK is activated by Ca^{2+}/calmodulin, whereas the myosin phosphatase is inactive when its subunit MYPT1 is phosphorylated (Vicente-Manzanares *et al.*, 2009). It was found by yeast-two hybrid that the Inc CT228 interacts with MYPT1, which is recruited to the periphery of and around the chlamydial inclusion (Lutter *et al.*, 2013). Phosphorylated, inactive MYPT1 localizes in microdomains at the inclusion membrane along with myosin IIA/IIB, MLC2 and MLCK, whose depletion causes reduced chlamydial extrusion and enhanced lytic exit (Lutter *et al.*, 2013). A *C. trachomatis* strain with *ct228* inactivated displays a lack of MYPT1 recruitment to the inclusion membrane and increased extrusion production, and causes a delayed systemic humoral response in mice, producing longer infections (Shaw *et al.*, 2018). Inc CT101, which also localizes at the inclusion membrane in microdomains, was shown by yeast-two hybrid to interact with the Ca^{2+} channel inositol 1,4,5-trisphosphate receptor type 3 (ITPR3) (Nguyen *et al.*, 2018). CT101 co-localizes with ITPR3 at the inclusion membrane and with STIM1, which also senses and controls Ca^{2+} levels. Depletion of *ct101* in *C. trachomatis* resulted in decreased ITPR3 recruitment and a reduction of chlamydial release by extrusion, a phenotype that was complemented. Reduced extrusion formation was also observed when ITPR3 or STIM1 were depleted or when Ca^{2+} was chelated, which indicates loss of myosin motor activity (Nguyen *et al.*, 2018). Altogether, these studies

indicate that *C. trachomatis* controls its egress by extrusion or host cell lysis by controlling the host cell myosin II.

In a very recent study, Bishop and Derré showed that Inc CTL0390/CT135 is involved in *C. trachomatis* host cell lytic exit (Bishop and Derré, 2022). The authors showed that a *C. trachomatis ctl0390* mutant strain is defective in inducing host cell lysis and that the overproduction of CTL0390 leads to premature host cell lysis and nuclear condensation. Moreover, by using inhibitors of several pathways related to host cell immunity, they showed that CTL0390 induces host cell lysis via the cyclic GMP–AMP synthase (cGAS)–stimulator of interferon genes (STING; cGAS-STING) DNA sensing pathway, a key mediator of inflammatory processes. This was the first Inc showed to participate in *C. trachomatis*-mediated host cell lytic exit.

1.4.2.2.2 *C. trachomatis* non-Inc effector proteins

Non-Inc effector proteins do not possess the characteristic bilobed hydrophobic domain of Incs and the absence of distinguishable features on their primary amino acid sequence makes their identification more challenging. Some of these proteins have been detected within chlamydial EBs, while others have been shown to be delivered to the outside of the inclusion and localize in the host cell cytoplasm, plasma membrane, or nucleus, and/or at the inclusion membrane (Bugalhão and Mota, 2019).

Non-Inc T3S effectors packed in chlamydial EBs

Some of the identified non-Inc T3S effectors are pre-packed inside EBs and are delivered into the host cell cytosol during or after their internalization. CT456/chlamydial translocated actin-recruiting phosphoprotein (TarP) is synthesized at late stages of infection and is carried within chlamydial EBs. Shortly after attachment of the EB, TarP is discharged into the host cell cytosol where it is tyrosine phosphorylated, promoting the recruitment of host actin to the entry site, which ultimately leads to internalization (Clifton *et al.*, 2004). Tarp has been shown to contain a conserved actin nucleating domain at its C-terminal region, which binds both globular (G-) and filamentous (F-) actin and has both nucleating and bundling activity towards actin (Jewett *et al.*, 2006, 2010; Jiwani *et al.*, 2013). At the N-terminal region, TarP contains a tyrosine-rich repeat domain that has been shown to be phosphorylated by a complex set of host tyrosine kinases [as Src and Abl kinases; (Jewett *et al.*, 2006, 2008, 2010; Mehlitz *et al.*, 2008)]. Both the actin binding and phosphorylation domains are crucial for invasion, as *C. trachomatis* strains lacking each of them display an invasion defect (Parrett *et al.*, 2016). Actin recruitment involves the Arp2/3 complex, which is activated by a cascade of events. Phosphorylated TarP interacts with the Rac guanine nucleotide exchange factors (GEFs) Sos1 and Vav2, activating the small GTPase Rac. Rac interacts with Abi-1 and WAVE2, promoting Arp2/3-dependent actin

recruitment (Carabeo *et al.*, 2007; Lane *et al.*, 2008). A role of TarP in avoiding host cell apoptotic death by interacting with the host protein SHC1 has also been described (Mehlitz *et al.*, 2010).

CT875/Translocated early phosphoprotein (TepP) is translocated into the host cell cytosol early during *C. trachomatis* entry and is tyrosine phosphorylated by host kinases, as TarP. TepP interacts with and recruits to the nascent inclusion Crk, a scaffolding protein involved in several processes such as regulation of cytoskeletal dynamics and cell signaling (Chen *et al.*, 2014; Carpenter *et al.*, 2017). TepP was also shown to recruit class I phosphoinositide 3-kinases (PI3K) to the nascent inclusion, and both proteins lead to decreased transcription of type I IFN-induced genes early during infection (Carpenter *et al.*, 2017). This suggests that TepP modulates host cell signaling and membrane trafficking events.

CT694/Translocated membrane-associated effector (Tme) A and CT695/TmeB are encoded by two genes that are co-transcribed. Shortly after internalization, TmeA is detected in purified chlamydial EBs and both TmeA and TmeB are observed in close association with invading EBs, near the nascent inclusion (Hower *et al.*, 2009; Mueller and Fields, 2015). At 24 h post-infection, TmeA and TmeB localize at the host cell cytosol and TmeA is also detected at the host cell plasma membrane of both infected and transfected cells (Hower *et al.*, 2009; Mueller and Fields, 2015; Wang *et al.*, 2018). TmeB accumulates in the host cell cytosol adjacent to the inclusion membrane, but it has unknown functions (Mueller and Fields, 2015). TmeA was shown to interact with the mammalian AHNAK nucleoprotein (AHNAK) and to interfere with the AHNAK-mediated F-actin bundling activity *in vitro* (Haase *et al.*, 2004; Hower *et al.*, 2009; McKuen *et al.*, 2017). However, the recruitment of AHNAK to the nascent inclusion observed during *C. trachomatis* infection was shown to be independent of TmeA. Moreover, a *C. trachomatis tmeA* mutant strain displayed a defect on invasion but this was unrelated to the TmeA-AHNAK interaction, which indicates that TmeA but not AHNAK is required for chlamydial invasion (McKuen *et al.*, 2017). Recently, two studies have shown that TmeA recruits and directly activates the nucleation-promoting factor N-WASP through its GTPase domain, which leads to the recruitment of the Arp2/3 complex to the site of chlamydial invasion. Importantly, TarP and TmeA are both required for efficient chlamydial invasion by acting synergistically on two separate pathways that culminate in Arp2/3-dependent actin polymerization (Faris *et al.*, 2020; Keb *et al.*, 2021).

CT622/TaiP is delivered into the host cell cytosol during *C. trachomatis* infection (Gong *et al.*, 2011). CT622 was found to bind CT635 through its N-terminal region, but this interaction was not further studied. Moreover, a *C. trachomatis ct622* mutant displays less infectivity and growth *in vitro* (Cossé *et al.*, 2018). Recently, an interaction between CT622 and the autophagy-related protein protein 16-1 (ATG16L1) was found, which led to the designation of translocated ATG16L1 interacting protein (TaiP). ATG16L1 is thought to interact with TMEM59 and supply TMEM59-containing vesicles with Rab6-positive compartments. This

interaction restricts chlamydial growth, and TaiP counteracts this effect by disrupting the ATG16L1-TMEM59 interaction (Hamaoui *et al.*, 2020).

Non-Inc T3S effectors targeting the host cell cytoplasm, plasma membrane, or nucleus, or the inclusion membrane

CT105/*Chlamydia trachomatis* effector associated with the Golgi (CteG) was first identified as a T3S substrate using *Y. enterocolitica* as a heterologous organism (da Cunha *et al.*, 2014). CteG was later shown to be delivered by *C. trachomatis* into the host cell cytoplasm, where it localizes mainly at the Golgi at ~20 h post-infection and mainly at the host plasma membrane from ~30 h post-infection. CteG has a Golgi-targeting motif within its first 100 amino acid residues and induces a vacuolar protein sorting defect when ectopically expressed in *S. cerevisiae* (Pais *et al.*, 2019). However, the target(s) and mechanism(s) of action of CteG during *C. trachomatis* infection are still unknown.

The CT737/nuclear effector (NUE) protein contains a SET domain that is found in eukaryotic histone methyltransferases, proteins that control gene expression and perform chromatin modifications. NUE was observed in the nucleus of both *C. trachomatis*-infected cells and transfected cells. Moreover, NUE was able to automethylate and to methylate mammalian histones H2B, H3 and H4 *in vitro*. Therefore, the effector NUE may act as a methyltransferase that methylates histones and remodels the chromatin of host cells (Pennini *et al.*, 2010).

CT619, CT620, CT621, CT711 and CT712 are a group of T3S effectors that contain a unique domain of unknown function (DUF582) characteristic of the *Chlamydiaceae*. This C-terminal domain is predicted to be mainly α-helical and to contain a segment of coiled-coil conformation and shares low identity (18 to 39%) among the five *C. trachomatis* DUF582-containing proteins. CT620 and CT621 have been shown to localize mainly in the host cell cytosol and inclusion lumen, and to some extent in the nucleus of *C. trachomatis*-infected cells. CT711 was only detected in the nucleus of transfected cells (Hobolt-Pedersen *et al.*, 2009; Muschiol *et al.*, 2011; Vromman *et al.*, 2016). The DUF582 domain of all proteins (except the one of CT621) was shown to interact with the protein Hrs, and the N-terminal part of CT619 also interacted with Tsg101. Both Hrs and Tsg101 are proteins that belong to the endosomal sorting complex required for transport (ESCRT) machinery, which has a pivotal role in host cell vesicular trafficking and membrane constriction. However, disruption of Hrs, Tsg101 or other proteins involved in ESCRT-driven processes did not affect the *C. trachomatis* developmental cycle (Vromman *et al.*, 2016). Conversely, a role for ESCRT machinery abscission proteins in the extrusion pathway of chlamydial egress has been proposed, as depletion of some of these proteins decreased the production of extrusions (Zuck and Hybiske, 2019).

CT089/CopN and CT529/Cap1 localize in the inclusion membrane (Fields and Hackstadt, 2000; Fling *et al.*, 2001) and they are considered Incs, although they do not possess

a characteristic bilobed hydrophobic domain (Li *et al.*, 2008a). CopN has been proposed to act as a T3S system gatekeeper in complex with the chaperone Scc3 (Archuleta and Spiller, 2014). Interestingly, this function is disparate from that observed for *C. pneumoniae* CopN, which has been shown to dampen the assembly of host cell microtubules (Archuleta *et al.*, 2011). Little is known about Cap1, but studies have shown that it is an antigen that triggers the activation of MHC class I restricted $CD8^+$ T cells during infection of mice with *C. trachomatis* (Fling *et al.*, 2001), and that it associates with lipid droplets in transfected cells (Saka *et al.*, 2015).

CT847 was found by yeast-two hybrid to interact with mammalian Grap2 cyclin D-interacting protein (GCIP), a protein likely involved in the regulation of cell differentiation and proliferation. The levels of GCIP were shown to be decreased in cells infected by *C. trachomatis*, a phenomenon that was dependent on bacterial protein synthesis and on a functional T3S system. Moreover, siRNA-mediated depletion of GCIP led to enhanced production of chlamydial infectious progeny (Chellas-Géry *et al.*, 2007). Although a direct role of CT847 in GCIP depletion has not been confirmed, it is thought that CT847 might mediate this process, which appears to favor chlamydial infection (Chellas-Géry *et al.*, 2007).

1.4.2.3 Non-Inc T3S system-independent/candidate effectors

Some *C. trachomatis* proteins that were found to be or might be relevant during infection are transported by mechanisms that are still unknown (Bugalhão and Mota, 2019).

CT166, as already mentioned in section 1.2.3.1, was found to be a chlamydial cytotoxin (Belland *et al.*, 2001). Ocular and urogenital serovars (except for serovar B) encode a total or partial *ct166* coding sequence that it is completely absent in serovars L1-L3 (Carlson *et al.*, 2004). CT166 is present in EBs and is detected in extracts of *C. trachomatis*-infected cells during the first hour of infection, after which it is likely degraded. Inhibition of new bacterial transcription or translation prior to infection did not impact the cytotoxic effect of EBs, which supports the fact that EBs carry a preformed cytotoxin (Belland *et al.*, 2001). Moreover, species that contain homologs of CT166, such as *C. muridarum* or *C. trachomatis* serovar D, were observed to induce cytopathic effects (as cell rounding) and dramatic alterations in the host cell actin cytoskeleton as opposed to *C. trachomatis* serovar L2, which does not possess a homolog of CT166 (Belland *et al.*, 2001; Thalmann *et al.*, 2010). These cytopathic effects were equally observed in HeLa cells ectopically expressing CT166 (Thalmann *et al.*, 2010), and were associated to a putative glucosyltransferase activity of this protein. CT166 is thought to glucosylate and inactivate the Rho-family protein Rac1, leading to actin cytoskeleton rearrangements (Thalmann *et al.*, 2010).

CT868/*Chlamydia* deubiquitinase (Cdu) 1/ChlaDUB1 and CT867/Cdu2/ChlaDUB2 were shown to possess deubiquitinating and deneddylating activities in mammalian

transfected cells (Misaghi *et al.*, 2006), while Cdu1 also displays acetyltransferase activity (Pruneda *et al.*, 2018). In *C. trachomatis*-infected cells, both Cdu1 and Cdu2 localize at the inclusion membrane at 24 h post-infection. At 48 h post-infection, this localization is maintained by Cdu1 whereas Cdu2 localizes at the host cell plasma membrane (Fischer *et al.*, 2017; Wang *et al.*, 2018). Cdu1 faces the host cell cytosol from where it interacts with the anti-apoptotic protein Mcl-1, maintaining this protein deubiquitinated and stabilized, which possibly contributes to *C. trachomatis* resistance to apoptosis (Sharma *et al.*, 2011; Fischer *et al.*, 2017). Cdu1 was also shown to prevent activation of the nuclear factor (NF)-kB pathway, which is involved in host inflammatory responses (Le Negrate *et al.*, 2008). Remarkably, fragmentation and redistribution of the Golgi, a hallmark of chlamydial infection, did not occur in cells infected with *C. trachomatis cdu1* and *cdu2* mutant strains. This alteration in the morphology of the Golgi was observed in cells ectopically expressing Cdu1 and Cdu2, further supporting a role of these proteins in this process (Pruneda *et al.*, 2018). However, a report suggests that a *cdu1* mutant is still able to induce Golgi fragmentation but fails to recruit Golgi-derived vesicles (Auer *et al.*, 2020).

CT156/lipid droplet-associated protein (Lda) 1, CT163/Lda2, CT473/Lda3 and CT257/Lda4 were identified as proteins with tropism for eukaryotic lipid droplets in yeasts. Lda1, Lda2 and Lda3 were further shown to co-localize with eukaryotic lipid droplets in transfected and *C. trachomatis*-infected mammalian cells. Lipid droplets have been observed in the inclusion lumen and may be used by chlamydiae as a source of lipids, and Lda proteins could be involved in the process of lipid droplet acquisition (Kumar *et al.*, 2006; Sisko *et al.*, 2006; Cocchiaro *et al.*, 2008).

CT311 and CT795 were shown to be delivered into the cytosol of *C. trachomatis*-infected cells (Lei *et al.*, 2011). CT311 is targeted to the host cell nucleus during late stages of infection by a nuclear localization signal composed of two clusters of basic amino acids (Lei *et al.*, 2011, 2013). So far, the function of these two proteins during *C. trachomatis* infection is still unknown.

Other *C. trachomatis* proteins have been shown to be delivered into the host cell cytoplasm by mechanisms independent of the T3S system. One of them is plasmid encoded Pgp3, an important virulence factor that is secreted into the host cell cytosol possibly by a plasmid-dependent mechanism, which has already been described (see section 1.3.1 above).

CT858/CPAF is an extensively studied chlamydial protein with serine protease activity. CPAF has been proposed to have very different targets and functions, which have posed challenges in understanding its function during *C. trachomatis* infection. Initial studies have described that CPAF cleaved a wide range of host and chlamydial substrates (Zhong *et al.*, 2001; Zhong, 2011), but it was later demonstrated that cleavage of these substrates was caused by enzymatic activity after cell lysis (Chen *et al.*, 2012). However, studies have confirmed that CPAF is produced in chlamydiae as a 70-kDa inactive zymogen that activates

itself by a series of autocatalytic events, yielding fully matured CPAF (Dong *et al.*, 2004; Huang *et al.*, 2008; Paschen *et al.*, 2008). CPAF was shown to be essential for chlamydial survival in the lower genital tract of mice (Yang *et al.*, 2016). CPAF was also implied in the cleavage of vimentin and lamin-associated protein-1 (LAP1) in the host cell cytosol after loss of inclusion membrane integrity, likely to induce host cell lysis (Snavely *et al.*, 2014). Moreover, a *C. trachomatis cpaf* null mutant is defective in host cell lytic exit, suggesting that CPAF is involved in this pathway of *C. trachomatis* egress (Yang *et al.*, 2015).

CT823/chlamydial high temperature requirement protein A (HtrA) is a chaperone and serine protease involved in multiple processes (Huston *et al.*, 2007; Marsh *et al.*, 2017). Although HtrA is a periplasmic protein, it has been detected at the cytosol of cells infected by *C. trachomatis* where it may have unidentified host-pathogen interaction roles (Wu *et al.*, 2011).

1.4.2.4 Chlamydial proteins delivered into the inclusion lumen

A few *C. trachomatis* proteins of undetermined function have been observed within the inclusion lumen, outside of the chlamydiae, but not in the host cell cytoplasm. This is the case of the T3S substrates CT142, CT143 and CT144, whose expression is regulated by the *Chlamydia* virulence plasmid (Patton *et al.*, 2018). These proteins are encoded in an operon and appear in globular structures outside of the chlamydiae and within the inclusion lumen during infection by *C. trachomatis* (da Cunha *et al.*, 2017). CT143 was also described to induce the secretion of pro-inflammatory cytokines in THP-1 macrophages (Jia *et al.*, 2019).

CT042/GlgX and CT798/GlgA are respectively the glycogen debranching enzyme and the glycogen synthase, enzymes involved in the metabolism of glycogen, and they are possible T3S substrates that mainly localize at the inclusion lumen. However, while GlgX also localizes at the inclusion membrane, GlgA was observed at the host cell cytosol during *C. trachomatis* infection, but the significance of its presence in this compartment remains unclear (Lu *et al.*, 2013; Gehre *et al.*, 2016).

CT049/Pls1 and CT050/Pls2 are potential inclusion membrane-associating proteins that accumulate in the inclusion lumen by a yet unknown mechanism. They lack the bilobed hydrophobic domain of Incs and share homology with a specific domain of PmpC. Their role during infection might be related with an efficient expansion of the inclusion (Jorgensen and Valdivia, 2008).

CT806/Ptr is a putative secreted protease that is produced during mid to late stages of *C. trachomatis* infection and localizes in the inclusion lumen. Ptr was found in a screening of *C. trachomatis* mutants displaying impaired recovery from IFN-γ-induced stress and has therefore been associated to this mechanism (Panzetta *et al.*, 2019).

1.5 Objectives

The overall objective of this PhD project was to expand the knowledge on the mechanisms used by *C. trachomatis* to subvert host cells during infection. For this, we focused on further characterizing the effector protein CteG/CT105.

CteG is a protein of 656 amino acid residues with no obvious motifs/signatures that could give insights about its function and whose sequence does not show significant similarity to other proteins except for putative homologs in other *Chlamydia* spp. (Pais *et al.*, 2019). In previous studies, we found that CteG mainly localizes at the host cell Golgi between 16 and 20 h post-infection, and accumulates at the host cell plasma membrane at the end of infection (Figure 1.5) (Pais *et al.*, 2019). We also found that the first 100 amino acids of CteG were crucial for its targeting to the Golgi of transfected cells. Moreover, we identified a CteG-induced defect in protein traffic to the yeast vacuole upon its expression in *Saccharomyces cerevisiae*, which suggests that CteG may subvert host cell vesicular trafficking [Figure 1.5; (Pais *et al.*, 2019)]. Finally, we observed that a *C. trachomatis cteG* insertional mutant strain (*cteG::aadA*) develops smaller inclusions comparing to the correspondent parental strain, a phenotype that could not be complemented by expression of *cteG* from a plasmid (Pais *et al.*, 2019).

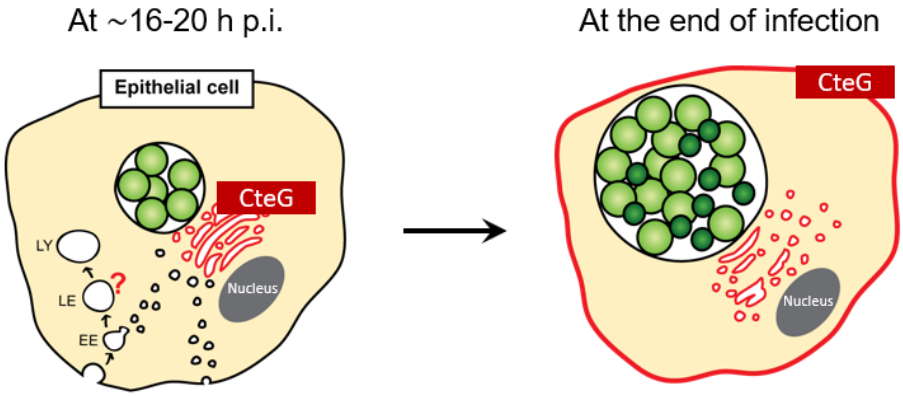


Figure 1.5. Schematic representation of a proposed model for the localization and mode of action of CteG during *C. trachomatis* infection of epithelial cells. After being T3S system-delivered into the host cell cytoplasm by the chlamydiae (green), CteG (represented in red) mainly associates with the Golgi complex between 16 and 20 h post-infection (p.i). While in this compartment, CteG could modify the host cell vesicular trafficking pathways (EE, early endosomes; LE, late endosomes; LY; lysosomes). As the developmental cycle progresses, CteG accumulates at the host cell plasma membrane where it might contribute to *C. trachomatis* exit from host cells at the end of infection (Pais *et al.*, 2019). This illustration was kindly provided by Dr. Sara Pais (Pais, 2018).

Despite these previous insights, several questions about CteG remained to be answered. Specifically, the identification of possible host cell interacting partners of CteG, the phenotypes associated to host cell infection by *cteG::aadA* strain, which would reflect CteG function(s), and the mechanism by which CteG is directed to both the Golgi and plasma membrane of host

cells at different times of infection remained to be investigated. Addressing these points could give important clues about the biological role of CteG during *C. trachomatis* infection. Therefore, in this work, our specific objectives were:

- To identify the determinants of the subcellular localization of CteG during *C. trachomatis* infection.
- To identify and characterize phenotypes associated to the *cteG::aadA* strain.
- To identify CteG host cell interacting partners and validate possible interactions.

In parallel, it was also planned as possible contingency to study the putative homologs of CteG within *Chlamydiaceae*. As a consequence of lockdown due to COVID-19, this part was also done. We previously noted that full-length homologs of CteG were only detected in *C. muridarum* and *C. suis* (Pais *et al.*, 2019). In *C. suis* and in other *Chlamydia* spp. different open reading frames might encode proteins with some identity (between ~30-22%) to only some parts of the amino acid sequence of CteG. Understanding the evolutionary history of CteG could also give insights on its function and how it is targeted to the plasma membrane and Golgi.

2

MATERIALS AND METHODS[2]

2.1 DNA manipulation, oligonucleotides, and plasmids

The plasmids used in this work, and a description of their main characteristics are in Annexes Table 1. The DNA oligonucleotides used in plasmid construction and in other molecular biology procedures are listed in Annexes Table 2. In general, plasmids were generated by cloning with restriction enzymes using standard molecular biology procedures. Briefly, DNA sequences were amplified with proof-reading Phusion high-fidelity DNA polymerase (Thermo Fisher Scientific). DNA sequences and backbone plasmids were then digested with FastDigest restriction enzymes (Thermo Fisher Scientific) and ligated with T4 DNA ligase (Thermo Fisher Scientific). NZYTaq II DNA polymerase (NZYTech) was used for screening of positive clones. DNA fragments and plasmids were purified with DNA Clean & Concentrator-5™ kit (Zymo Research), Zymoclean™ Gel DNA recovery kit (Zymo Research), NZYMiniprep kit (NZYTech) or NZYMidiprep kit (NZYTech) following the manufacturer's instructions. Some of the plasmids (specified in Annexes Table 1) were generated by total plasmid amplification with Phusion high-fidelity DNA polymerase, followed by digestion of the parental plasmids with DpnI (Thermo Fisher Scientific).

The backbone plasmids used in this work included p2TK2-SW2 (Agaisse and Derré, 2013), a cloning vector suitable for transformation of *C. trachomatis*, and its derivative pSVP247/pVector[Pgp4^{+}] (da Cunha *et al.*, 2017) which enables the expression of proteins with a C-terminal double-hemagglutinin (2HA) tag in *C. trachomatis*. The accuracy of the nucleotide sequence of all the inserts or plasmids was confirmed by Sanger sequencing. In case of plasmid pIP68/pVector[Pgp4^{-}] with *pgp4* deleted (Annexes Table 1), the accuracy of the nucleotide sequence of the entire plasmid was confirmed.

[2] This Chapter was written by Inês Serrano Pereira. Parts of text were transcribed from Pereira *et al.* [(Pereira *et al.*, 2022)], a recent publication that includes significant data from this PhD thesis.

2.2 *Escherichia coli* strains and growth conditions

Escherichia coli NEB® 10β (New England Biolabs) was used for plasmid construction and purification, and *E. coli* ER2925 (New England Biolabs) was used to replicate and purify plasmids for transformation of *C. trachomatis*. *E. coli* strains were grown in liquid or agar lysogeny broth (LB; NZYTech) with the appropriate selective antibodies and supplements. *E. coli* cells were transformed with the plasmids by electroporation.

2.3 Mammalian cell lines

HeLa 229 and Vero cells (from the European Collection of Authenticated Cell Culture; ECACC) were passaged in 4.5 g/L glucose, L-glutamine Dulbecco's Modified Eagle's Medium (DMEM; Corning) supplemented with heat-inactivated 10% (v/v) fetal bovine serum (FBS; Thermo Fisher Scientific) at 37ºC in a humidified atmosphere of 5% (v/v) CO_2 and detached from culture plates or flasks with TrypLE™ Express (Thermo Fisher Scientific). Cell cultures were regularly tested for *Mycoplasma* by conventional PCR as described (Uphoff and Drexler, 2011).

2.4 Transient transfection of mammalian cells and treatment with 2-Bromopalmitate

For immunofluorescence microscopy analysis and immunoblotting, 5×10^4 or 1×10^5 HeLa 229 cells per well were seeded in 24-well plates, respectively. In the former case, cells were seeded onto 13 mm glass coverslips (VWR). For co-immunoprecipitation (co-IP) assays, 5×10^5 HeLa 229 cells per well were seeded in 6-well plates. The next day, cells were transfected with plasmids using jetPEI™ reagent (Polyplus-Transfection) according to the manufacturer's instructions. Briefly, to each well of 24-well plates, 250 ng of plasmid DNA and 1.5 μL of jetPEI were added onto seeded cells. For co-IP assays, the amount of plasmid DNA and jetPEI was scaled up according to well surface area. Plates were then centrifuged at 180 x *g* for 5 min at room temperature and incubated at 37°C in a humidified atmosphere of 5% (v/v) CO_2 for 4 h. After this period, the medium was replaced by fresh DMEM and transfections proceeded for a total of 24 h, after which cells were fixed for immunofluorescence microscopy or processed for immunoblotting. For co-IP assays, cells were transfected for 16 h without replacing the transfection medium before being collected. For experiments with the inhibitor 2-Bromopalmitate (2BP), cells were transfected with plasmids for 16 h as described above. The

transfection medium was then removed, and cells were incubated for 5 h in DMEM supplemented with dimethyl sulfoxide (DMSO) or 40 μM 2BP (in DMSO) before fixation.

2.5 Manipulation of *C. trachomatis*

2.5.1 Generation of *C. trachomatis* strains

The *C. trachomatis* strains used and generated in this work are listed in Annexes Table 3. They were propagated in HeLa 229 cells using standard procedures (Scidmore, 2005). *Chlamydia* stocks were tested for *Mycoplasma* by conventional PCR (Uphoff and Drexler, 2011) and Sanger sequencing techniques. All newly generated *C. trachomatis* strains were checked for loss of the native plasmid and acquisition of the desired plasmid by conventional PCR.

2.5.1.1 Transformation of *C. trachomatis*

Transformation of *C. trachomatis* was performed essentially as described by Agaisse and Derré (Agaisse and Derré, 2013). Approximately 1×10^7 *C. trachomatis* EBs (from extracts of infected cells) and 6 μg of plasmid DNA were mixed in 200 μL $CaCl_2$ buffer (10 mM Tris pH 7.4, 50 mM $CaCl_2$) and incubated at room temperature for 30 min. In the meantime, 4×10^6 HeLa 229 cells per transformation were freshly trypsinized and washed in phosphate buffered saline (PBS) by centrifugation at 180 x *g* for 5 min. Each cell pellet was resuspended in 200 μL $CaCl_2$ buffer and added to the DNA/EBs mixture. The HeLa/DNA/EBs mixture was incubated at room temperature for 20 min, with gentle resuspension each 5 min, and then distributed by two wells of a 6-well plate containing pre-warmed DMEM. Cells were then incubated at 37°C in a humidified atmosphere of 5% (v/v) CO_2 and, at 16 h post-transformation, the medium was replaced by fresh DMEM supplemented with 0.3U/mL penicillin G (PenG). At ~44 h post-transformation, infected cells were lysed by incubation with 500 μL of sterile dH_2O per well for 15 min at room temperature and added to 500 μL 2x sucrose-phosphate-glutamate buffer (SPG 2x; 0.44 M sucrose, 34 mM Na_2HPO_4, 6 mM NaH_2PO_4, 10 mM L-glutamic acid). This lysate was centrifuged at 270 x *g* for 5 min at room temperature, and the supernatant was added to HeLa 229 cells seeded in T25 flasks (~ 2.5×10^6 cells/flask) and pre-incubated in Hanks' Balanced Salt Solution (HBSS, Corning). After incubation with cells for 1 h at room temperature under gentle agitation, the inoculum was removed and replaced by DMEM supplemented with 0.3 U/mL PenG and 1 μg/mL cycloheximide (CHX), and cells were incubated at 37°C in a humidified atmosphere of 5% (v/v) CO_2 for ~44 h. For selection of transformants, infected cells were processed as described above but adding to the infection

medium increasing concentrations of PenG (0.3, 1 or 10 U/mL) on each passage, and 250 µg/mL of spectinomycin when transforming the *cteG::aadA* insertional mutant strain. When "wild-type" chlamydial inclusions were predominantly observed in infected HeLa 229 cells by phase-contrast microscopy, indicating a successful transformation (at least after 3 passages under the presence of 10 U/mL PenG), *C. trachomatis* EBs were collected and frozen in aliquots at −80°C before proceeding to clone isolation.

2.5.1.2 Plaque purification of *C. trachomatis*

Clone isolation of *C. trachomatis* strains was performed by plaque purification as described elsewhere (Nguyen and Valdivia, 2013). 4×10^5 Vero cells were seeded in each well of 6-well plates and incubated at 37°C in a humidified atmosphere of 5% (v/v) CO_2. The next day, cells were incubated in HBSS while *C. trachomatis* EBs, previously stored at −80°C, were briefly thawed at 37°C and 10-fold diluted in HBSS. Vero cells were infected with 500 µL of each dilution for 30 min at 37°C in a humidified atmosphere of 5% (v/v) CO_2. The inoculum was then replaced by fresh DMEM supplemented with 1 U/mL PenG, 10 µg/mL gentamicin, 1 µg/mL CHX, and 250 µg/mL spectinomycin when isolating strains derived from the *C. trachomatis cteG::aadA* mutant strain. After a 24-hour incubation at 37°C in a humidified atmosphere of 5% (v/v) CO_2, the medium was aspired from infected cells to add an overlay medium composed of 0.6% (w/v) agarose in complete DMEM (without phenol red) supplemented with 10% (v/v) FBS, 1 U/mL PenG, 1 µg/mL CHX, and 250 µg/mL spectinomycin when relevant. After polymerization of this overlay medium at room temperature, complete liquid DMEM (without phenol red) was added on top of the overlay medium. Infected Vero cells were incubated at 37°C in a humidified atmosphere of 5% (v/v) CO_2 for 3-4 days to allow the formation of "plaques" (locations devoid of cells due to lysis caused by *C. trachomatis* infection) discernible by phase-contrast microscopy. Each plaque, which was generated by a single *C. trachomatis* infectious particle, was picked with a barrier pipette tip, and swirled in 100 µL of DMEM supplemented with 1 U/mL PenG and 1 µg/mL CHX. Then, separate wells of a 96-well plate containing Vero cells seeded the day before were infected with the 100 µL of DMEM containing each collected plaque. The 96-well plate was centrifuged at 2250 x *g* for 30 min at 15°C, and then incubated at 37°C in a humidified atmosphere of 5% (v/v) CO_2. After two days, 100 µL dH_2O were added to each well to lyse infected cells, and the lysate was resuspended in 100 µL SPG 2x. Each lysate corresponding to separate chlamydial plaques was used to infect Vero cells previously seeded in 24-well plates. Clones were expanded by repeating this procedure in T25 or T75 flasks seeded with HeLa 229 cells until a higher number of infectious particles could be recovered and stored at −80°C.

2.5.2 Infection of HeLa cells with *C. trachomatis*

Infections for quantification of inclusion forming units (IFUs), cell cytotoxicity assays, determination of inclusion size and assessment of protein levels by immunoblotting were carried out by seeding 1×10^5 HeLa 229 cells per well in 24-well plates. For immunofluorescence experiments, cells were seeded onto 13 mm glass coverslips (VWR). The day after seeding, cells were incubated in HBSS for ~ 15 min at 37°C in a humidified atmosphere of 5% (v/v) CO_2 while *Chlamydia* inocula were prepared in SPG 1x (0.22 M sucrose, 17 mM Na_2HPO_4, 3 mM NaH_2PO_4, 5 mM L-glutamic acid) at various multiplicities of infection (MOIs). Cells where infected with these inocula for 30 min at 37°C in a humidified atmosphere of 5% (v/v) CO_2, which was then replaced by plain DMEM, and incubated at 37°C in a humidified atmosphere of 5% (v/v) CO_2. For all experiments, this was considered the time zero of infection. To determine the effect of the addition of gentamicin in the number of recovered IFUs or in cytotoxicity levels, DMEM supplemented with 10 μg/mL of gentamicin was added to cells at 0 h of infection. At 24 h post-infection, cells were washed once with DMEM supplemented with 10% (v/v) FBS and left in fresh media without gentamicin for the remainder time of infection. For determination of inclusion size, cells were infected at a MOI of 0.06 for 24 h before fixation. To assess protein levels by immunoblotting, cells were infected with a MOI of 6 and incubated for 16, 24, 30 or 40 h in DMEM supplemented with 10% (v/v) FBS and 10 μg/mL gentamicin.

For detection by immunoblotting of bacteria in the supernatant of infected cells, 5×10^5 HeLa 229 cells per well were seeded in 6-well plates. Cells were infected as described above with a MOI of 0.06 and incubated at 37°C for 38 h in DMEM supplemented with 10% (v/v) FBS. At this time, the medium was replaced by DMEM without FBS, and the infection allowed to proceed up to 48 h.

2.5.3 Assessment of relative progeny generation and quantification of IFUs in infected cells

Assessment of progeny generation was performed essentially as previously described (Sixt *et al.*, 2017). Briefly, two identical 24-well plates seeded with HeLa 229 cells were infected with *C. trachomatis* strains at a MOI of 0.06. In one of the plates, infected cells were fixed at 24 h post-infection with methanol for 7 min at -20ºC (input). In the other plate, infection was allowed to proceed for 40 h, after which cells were washed very briefly with dH_2O and then osmotically lysed by incubation for 15 min at room temperature with 500 μL of dH_2O. The lysed cells were vigorously resuspended by pipetting up and down several times and the suspension was added to 500 μL of SPG 2x. The lysates obtained were homogenized by

vortexing, serial diluted in SPG 1x (0.22 M sucrose, 17 mM Na_2HPO_4, 3 mM NaH_2PO_4, 5 mM L-glutamic acid), and used to infect a fresh layer of HeLa 229 cells. These cells were fixed with methanol 24 h post-infection for 7 min at -20ºC (output), and immunolabelled. Inclusions were counted by fluorescence microscopy in ≥30 fields of duplicated samples, using a total amplification of 400x, and IFUs/mL were determined as previously described (Scidmore, 2005). For each strain, the relative progeny generation was obtained by dividing the number of IFUs in the output by those in the input.

For quantification of IFUs in the cell culture supernatant and cell lysate fractions of infected cells, the supernatants (1 mL) were collected and vortexed to homogenize extracellular bacteria (supernatant fraction). Attached cells were washed once with dH_2O and lysed by osmotic shock, as described above for assessment of progeny generation. Lysed cells were resuspended, added to SPG 2x and vortexed to homogenize recovered intracellular bacteria (lysate fraction). Both fractions were serial diluted in SPG 1x and the quantification of IFUs was done as for assessment of progeny generation.

2.5.4 Cell cytotoxicity assays

The supernatants of infected HeLa 229 cells were assayed for released lactate dehydrogenase (LDH) with the CytoScan™ LDH Cytotoxicity Assay kit (G-Biosciences), following the manufacturer's instructions and including the appropriate controls. To calculate the % of LDH released in each assay and time-point, the amount of LDH activity detected in uninfected cells after lysis with 1 % (v/v) Triton X-100, and the amount of LDH activity released from uninfected cells, were determined. The % of LDH released was then calculated as 100 x [(LDH activity released from infected cells − LDH activity released from uninfected cells) / (LDH activity detected in uninfected cells after lysis with Triton X-100 − LDH activity released from uninfected cells)]. Absorbance at 490 nm was measured in a SpectraMax 190 microplate reader (Molecular Devices) and data was acquired using the SoftMax Pro 7.1 software (Molecular Devices).

2.5.5 Internalization of *C. trachomatis*

To assess internalization of *Chlamydia*e by host cells, 4×10^4 HeLa 229 cells per well were seeded in 24-well plates and incubated at 37°C in a humidified atmosphere of 5% (v/v) CO_2. The day after seeding, HeLa 229 cells were incubated at 4°C for 15 min in pre-cooled DMEM. *Chlamydia* inocula were prepared in pre-cooled SPG 1x at a MOI of 20 and added to cells, which were then incubated at 4°C for 30 min to allow bacterial attachment. Inocula were then

replaced by pre-warmed DMEM, and cells were incubated for 45 min at 37°C in a humidified atmosphere of 5% (v/v) CO_2 to promote bacterial internalization. The medium was then removed, and cells were washed 3x in ice-cold PBS before being fixed with paraformaldehyde (PFA) 4% (w/v) at room temperature for 30 min. Immunolabelling of infected cells was then performed in two steps: first, non-internalized, attached bacteria were immunolabelled with primary anti-chlamydial MOMP and secondary AF488-conjugated antibodies in the absence of a permeabilizing agent; second, both non-internalized and internalized bacteria were immunolabelled with primary anti-chlamydial MOMP and secondary AF594-conjugated antibodies in the presence of Triton™ X-100 (Sigma-Aldrich) to allow cell permeabilization. For each strain, images of ≥ 40 individual infected cells were collected by fluorescence microscopy. Bacteria were manually counted and the percentage of internalization for each *C. trachomatis* strain was determined as 100 x [(Total number of bacteria − Number of non-internalized bacteria)/ Total number of bacteria], whereby Total number of bacteria = number of non-internalized bacteria + number of internalized bacteria.

2.5.6 Real-time quantitative PCR (RT-qPCR)

To determine the level of *cteG* expression in different *C. trachomatis* strains, 5×10^5 HeLa 229 cells per well were seeded in a 6-well plate. Cells were infected the next day with each *C. trachomatis* strain at a MOI of 2.5 as above-mentioned, in duplicated wells. At 2 and 8 h post-infection, cells were detached with TrypLE™ Express (Thermo Fisher Scientific), collected and centrifuged at 270 x *g* for 5 min, at 4ºC. After two washes in ice-cold PBS, cell pellets were frozen in liquid nitrogen and kept at -80ºC before RNA extraction. Total RNA was extracted using NZY Total RNA Isolation kit (NZYTech) following the manufacturer's instructions but performing a 90 min incubation step with DNase I at room temperature. The extracted RNA was then quantified, appropriately diluted, and stored at -80ºC. RT-qPCR reactions were performed with SensiFast™ SYBR® No-ROX One-Step kit (Bioline) in a Corbett Rotor-Gene™ 6000 (Qiagen) cycler. Primers for *cteG* and *16S* were used in previous studies (Pais et al, 2019) and are listed in Annexes Table 2. For each condition, a negative control where reverse transcriptase (RT) was not added to the reaction mixture was included. Ct values were obtained with Rotor-Gene 6000 Series Software 1.7 using the same threshold on each experiment. Ratios to the *16S* rRNA transcript were considered for normalization of bacterial load.

2.6 Antibodies and dyes

For immunoblotting, the following primary antibodies were used: rabbit anti-GFP (Abcam; 1:1,000); rat monoclonal anti-hemagglutinin (HA) (3F10, Roche; 1:1,000), mouse monoclonal anti-chlamydial heat-shock protein 60 (Hsp60) (A57-B9; Thermo Fisher Scientific; 1:1,000), mouse monoclonal anti-α-tubulin (clone B-5-1-2, Sigma-Aldrich; 1:1,000). Anti-mouse, anti-rabbit or anti-rat secondary antibodies were all horseradish peroxidase (HRP)-conjugated (GE Healthcare and Jackson ImmunoResearch; 1:10,000).

For immunofluorescence microscopy, the following primary antibodies were used: goat polyclonal anti-chlamydial MOMP (Abcam, 1:200), rat monoclonal anti-HA (3F10, Roche; 1:200), rabbit polyclonal anti-GM130 (Sigma-Aldrich; 1:200), mouse monoclonal anti-TGN46 (clone TGN46-8, Sigma-Aldrich; 1:200) and goat anti-*C. trachomatis*-EB, FITC-conjugated polyclonal antibody (Sigma-Aldrich, 1:150). The fluorophore-conjugated secondary antibodies were all purchased from Jackson ImmunoResearch and diluted 1:200: Rhodamine Red-X-conjugated anti-rat, AF568-conjugated anti-mouse, DyLight 405-conjugated anti-goat, Cyanine 3 (Cy3)-conjugated anti-rabbit, AF488-conjugated anti-goat, AF594-conjugated anti-goat. DAPI (4',6-Diamidino-2-phenylindole; 1:30,000) was used to label DNA, and actin staining was carried out by incubating HeLa 229 cells with Phalloidin-AF488 (Thermo Fisher Scientific; 1:100).

2.7 Immunofluorescence microscopy

Transfected HeLa 229 cells were fixed with PFA 4% (w/v) for 15 min at room temperature. Infected HeLa 229 cells were fixed either with freezing methanol (-20ºC) for 7 min or with PFA 4% (w/v) for 15 min at room temperature, as specified in figure legends. For immunolabelling, antibodies were diluted in PBS containing 10% (v/v) horse serum. Cells fixed with PFA or methanol were permeabilized with 0.1% (w/v) saponin or 0.1% (v/v) Triton™ X-100, respectively. All incubations were done for 1 h at room temperature. Cells were washed with PBS or PBS containing saponin or Triton between incubation with each antibody, and finally in PBS and H_2O. The coverslips were assembled on microscopy glass slides using Aqua-Poly/Mount (Polysciences) mounting medium and cells were examined by fluorescence microscopy in a Axio Imager.D2 (Zeiss) upright microscope. Images were collected by an Axiocam MRm (Zeiss) camera and processed with Zeiss ZEN (Zeiss) software, Fiji software (Schindelin *et al.*, 2012) and Adobe Illustrator.

2.8 Immunoblotting

For immunoblotting, detachment of transfected of infected HeLa 229 cells was done with TrypLE™ Express (Thermo Fisher Scientific) for 5 min at 37°C in a humidified atmosphere of 5% (v/v) CO_2. Cells were then collected to microtubes after being resuspended in DMEM, which was removed by centrifugation at 12,000 x *g* for 5 min at 4°C. Next, cells were washed 2x in ice-cold PBS and immediately re-suspended and boiled in SDS-PAGE Laemmli buffer (SDS loading buffer). Cell extracts were treated with benzonase (Novagen) for 20 min at room temperature to degrade DNA and reduce sample viscosity. For the analysis of supernatants of infected cells, SDS loading buffer was added immediately and directly to the supernatants. All samples were boiled for 5 min at 100°C and then separated by 12% (v/v) SDS-PAGE and transferred onto 0.2 mm nitrocellulose membranes (Bio-Rad) using Trans-Blot Turbo Transfer System (BioRad). Detection was done with SuperSignal West Pico Chemiluminescent Substrate (Thermo Fisher Scientific) or SuperSignal West Femto Maximum Sensitivity Substrate (Thermo Fisher Scientific), as specified in figure legends, and exposure to Amersham Hyperfilm ECL (GE Healthcare) as previously described (da Cunha *et al.*, 2017; Pais *et al.*, 2019).

2.9 Co-immunoprecipitation assays

Samples for mass spectrometry analysis were obtained by co-immunoprecipitation of extracts of HeLa 229 cells ectopically expressing mEGFP or mEGFP-CteG$_{FL}$. Transfected cells were washed 1x with PBS, scraped and collected to pre-cooled microtubes. Cells were spun at 1.000 *x g* for 5 min at 4°C, washed 2x in ice-cold PBS and lysed in ice-cold lysis buffer [20 mM Tris pH 7.5, 150 mM NaCl, 0.5 mM EDTA, 0.5% NP-40, 1x protease inhibitor cocktail of general use (AMRESCO, VWR) and 1 mM phenylmethylsulfonyl fluoride (PMSF)] for 30 min, with a brief homogenization each 10 min. In the meantime, GFP-Trap® beads (ChromoTek) were equilibrated by washing 2x in ice-cold dilution buffer (10 mM Tris pH 7.5, 150 mM NaCl, 0.5 mM EDTA). Between washes, beads were centrifuged at 3.000 *x g* for 5 min at 4°C. After lysis, cells were centrifuged at 13.000 *x g* for 10 min at 4°C and then incubated with the pre-equilibrated beads under the presence of 1x protease inhibitor cocktail and 1 mM PMSF. Interaction between mEGFP or mEGFP-CteG$_{FL}$ with the beads was allowed to occur by incubation in a rotating shaker for 4 h at 4°C. After this time, beads were settled at 3.000 *x g* for 5 min at 4°C and washed 6x in ice-cold dilution buffer. Beads were resuspended in SDS loading buffer 5x and boiled at 95°C for 10 min. Finally, beads were removed and samples were frozen at −80°C until analysis by mass spectrometry.

2.10 Mass spectrometry (nano-LC-MS/MS)

Mass spectrometry was performed using an Ultimate 3000 liquid chromatography system coupled to a Q-Exactive Hybrid Quadrupole-Orbitrap mass spectrometer (Thermo Scientific, Bremen, Germany). Samples were loaded onto a trapping cartridge (Acclaim PepMap C18 100Å, 5 mm x 300 μm i.d., 160454, Thermo Scientific) in a mobile phase of 2% ACN, 0.1% FA at 10 μL/min. After 3 min loading, the trap column was switched in-line to a 50 cm by 75µm inner diameter EASY-Spray column (ES803, PepMap RSLC, C18, 2 µm, Thermo Scientific, Bremen, Germany) at 250 nL/min. Separation was generated by mixing A: 0.1% FA, and B: 80% ACN, with the following gradient: 5 min (2.5% B to 10% B), 120 min (10% B to 30% B), 20 min (30% B to 50% B), 5 min (50% B to 99% B) and 10 min (hold 99% B). Subsequently, the column was equilibrated with 2.5% B for 17 min. Data acquisition was controlled by Xcalibur 4.0 and Tune 2.9 software (Thermo Scientific, Bremen, Germany).

The mass spectrometer was operated in data-dependent (dd) positive acquisition mode alternating between a full scan (m/z 380-1580) and subsequent HCD MS/MS of the 10 most intense peaks from full scan (normalized collision energy of 27%). ESI spray voltage was 1.9 kV. Global settings: use lock masses best (m/z 445.12003), lock mass injection Full MS, chrom. peak width (FWHM) 15s. Full scan settings: 70k resolution (m/z 200), AGC target 3e6, maximum injection time 120 ms. dd settings: minimum AGC target 8e3, intensity threshold 7.3e4, charge exclusion: unassigned, 1, 8, >8, peptide match preferred, exclude isotopes on, dynamic exclusion 45s. MS2 settings: microscans 1, resolution 35k (m/z 200), AGC target 2e5, maximum injection time 110 ms, isolation window 2.0 m/z, isolation offset 0.0 m/z, spectrum data type profile.

The raw data was processed using Proteome Discoverer 2.4.0.305 software (Thermo Scientific) and searched against a spectral database (NIST_Human_Orbitrap_HCD_20160923) the UniProt database for the Homo sapiens Proteome 2019_09. The MSPepSearch and Sequest HT search engines were used to identify tryptic peptides. The ion mass tolerance was 10 ppm for precursor ions and, respectively, 20 ppm or 0.02 Da for fragment ions. Maximum allowed missing cleavage sites was set 2. Cysteine carbamidomethylation was defined as constant modification. Methionine oxidation, protein N-terminus acetylation, methionine loss and methionine loss plus acetylation were defined as variable modifications. Peptide confidence was set to high. The processing node Percolator was enabled with the following settings: maximum delta Cn 0.05; decoy database search target FDR 1%, validation based on q-value. Protein label free quantitation was performed with the Minora feature detector node at the processing step. Precursor ions quantification was performing at the processing step including unique and razor peptides. The precursor abundance was based on intensity.

2.11 Whole-genome sequencing

C. trachomatis L2/434 and its derivative *cteG::aadA* (Pais *et al.*, 2019) (Annexes Table 3) were subjected to whole-genome sequencing (WGS). For this, an optimized DNA purification procedure was used to ensure depletion of human nucleic acids. First, suspensions of infected HeLa 229 cells were sonicated (3x10s, 50%, 5 K cycles/s; VibraCell, Bioblock Scientific) and the cell debris were discarded through low-speed centrifugation. Subsequently, the *Chlamydia*e in the supernatant were pelleted by high-speed centrifugation, followed by resuspension in a DNase/RNase cocktail [stock solution with 4.6 mg/ml DNase (Sigma; 400 Kunitz U/mg) and 13 mg/ml RNase (Applichem; 100.8 Kunitz U/mg), in Hanks' Balanced Salt Solution (HBSS), diluted 1:10 in HBSS], sonication (2x20s; S30 Elmasonic), and incubation at 37°C for 20 min. The DNase and RNase were then inactivated by incubation at 65°C for 15 min, followed by chilling 1 min on ice. The *Chlamydia*e in the suspensions were again pelleted, resuspended in PBS, and then added over a layer of 30% (v/v) urographin [diluted from 76% (v/v) urographin, sodium amidotrizoate (0.1 g/ml) and meglumine amidotrizoate (0.66 g/ml); Bayer, Portugal)]. A high-speed centrifugation step was then carried out and the *Chlamydia*e-enriched fraction was collected, resuspended in PBS, and subjected to a second round of DNase/RNase digestion and inactivation. The *Chlamydia*e in these suspensions were then pelleted, washed with PBS, and further processed for DNA isolation using Proteinase K (20 mg/mL) lysis and the QIAamp DNA Mini Kit (Qiagen, Valencia, CA, USA), according to manufacturer's instructions. Purified DNA was subsequently subjected to Nextera XT library preparation and subsequent paired-end sequencing (2x250 bp) in Illumina MiSeq (Illumina Inc., San Diego, CA, USA), according to the manufacturer's instructions.

Bioinformatics analysis involved: i) reads' quality analysis and cleaning/improvement using FastQC (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/) and Trimmomatic (http://www.usadellab.org/cms/?page=trimmomatic) (Bolger *et al.*, 2014); ii) reference-based mapping and SNP/indel analysis against the *C. trachomatis* L2/434/Bu reference genome sequences (NCBI accession numbers: AM884176.1 for chromosome and AM886278.1/X07547.1 for the plasmid) using Snippy v 3.2 (https://github.com/tseemann/snippy); iii) *de novo* genome assembly using SPAdes v 3.11.0 (Bankevich *et al.*, 2012); iv) whole genome alignment and inspection using Mauve v 2.3.1 (Darling *et al.*, 2010); and v) SNP/indel inspection using the Integrative Genomics Viewer (http://www.broadinstitute.org/igv) (Robinson *et al.*, 2011). This procedure for *C. trachomatis* enrichment allowed obtaining a percentage of "*on-target*" reads above 99.5% for both strains. WGS raw reads were submitted to the European Nucleotide Archive (ENA) under the BioProject accession number PRJEB51643.

2.12 Bioinformatics

2.12.1 Reciprocal best hit BLAST

Putative *cteG* homolog genes were investigated by reciprocal best hit BLAST. First, nucleotide databases were generated for each *Chlamydia* species using assembled genomes deposited in the National Center for Biotechnology Information (NCBI). The reference strains used for all the analyses performed in this study and their respective GenBank assembly accession numbers, indicated in parenthesis, are: *C. abortus* str. S26/3 (GCA_000026025.1); *C. avium* str. 10DC88 (GCA_000583875.1); *Ca.* C. corallus str. G3/2742-324 (GCA_002817655.1); *C. caviae* str. GPIC (GCA_000007605.1); *C. felis* str. Fe/C-56 (GCA_000009945.1); *C. gallinacea* str. 08-1274/3 (GCA_000471025.2); *C. ibidis* str. 10-1398 (GCA_000454725.1); *C. pneumoniae* str. CWL029 (GCA_000008745.1); *C. pecorum* str. E58 (GCA_000204135.1); *C. poikilotherma* str. S15-834K (GCA_900239975.1); *C. muridarum* str. Nigg (GCA_000006685.1); *C. psittaci* str. 6BC (GCA_000204255.1); *C. sanzinia* str. 2742-308 (GCA_001653975.1); *C. serpentis* str. H15-1957-10C (GCA_900239945.1); *C. suis* str. MD56 (GCA_000493885.1); *Estrella lausannensis* str. CRIB 30 (GCA_900000175.1); *S. negevensis* str. Z (GCA_000237205.1); *Chlamydiae bacterium*, isolate K940_chlam_9 (GCA_011064985.1); *Chlamydiae bacterium*, isolate KR126_chlam_2 (GCA_011064935.1); *Chlamydiae bacterium*, isolate K1000_chlam_4 (GCA_011065205.1); *Ca.* Protochlamydia naegleriophila str. KNic (GCA_001499655.1).

Putative *cteG* homolog genes were searched with tBLASTx using the nucleotide sequence of *cteG* from *C. trachomatis* L2/434/Bu (gene ref. CAP03800; GenBank genome accession ref. CP003963) as query, against the databases created for each bacterial species. tBLASTx is a tool that converts a nucleotide query sequence into protein sequences in all six reading frames and compares them to a nucleotide database that has also been translated in all six reading frames. Genomic regions of each species encoding proteins that had similarities with proteins encoded by all reading frames of *cteG* were considered potential hits when e-values were lower than 0.001. The nucleotide sequence corresponding to 2 kBp upstream the start, and 2 kBp downstream the end of these genomic regions was used for *ab initio* gene prediction with AUGUSTUS (Keller *et al.*, 2011), using *Staphylococcus aureus* as the bacterial reference organism. Best hit proteins potentially encoded by genes predicted at these regions were blasted against the NCBI non-redundant (nr) standard database of the corresponding *Chlamydia* species. Whenever the best hit proteins in NCBI corresponded to the identity of the protein encoded by the query gene, it was assumed that the gene was present and may therefore encode a CteG homolog.

2.12.2 Protein alignments and phylogenetic analysis

To visualize identity between protein sequences, the amino acid sequences of CteG and of its putative homologs were individually aligned using the protein BLAST (BLASTp) tool. The amino acid sequences of CteG and of its putative homologs sorted by reciprocal best hit BLAST were then used to perform multiple alignments with MAFFT v7.222 (Katoh and Standley, 2014) using an iterative refinement algorithm (L-INS-i). Poorly aligned regions were removed with trimAl v1.2 using the "*gappyout*" option (Capella-Gutiérrez *et al.*, 2009). A Maximum Likelihood (ML) tree was inferred with IQ-TREE v2.0 (Nguyen *et al.*, 2015) using an automatic detection of the best-fitting model of amino acid evolution and ultrafast bootstraps (-bb 1,000) (Hoang *et al.*, 2018). Proteins that did not group with *C. trachomatis* CteG were considered not to be its homologs.

To reconstruct the species tree, single copy orthologs (SCO) were retrieved using Orthofinder 2 (Emms and Kelly, 2019) from the predicted proteomes of the studied *Chlamydiaceae* species, CC-IV clade species and more distantly related species. The resulting concatenated alignment contained 201.091 amino acid positions that were subsequently used to infer a ML tree with IQ-TREE v2.0 (Nguyen *et al.*, 2015) using an automatic detection of the best-fitting model of amino acid evolution and ultrafast bootstraps (-bb 1,000) (Hoang *et al.*, 2018). Five independent tree searches were performed in total (--runs 5) and the tree with the highest likelihood score was considered the one representing the most likely phylogenetic relationships between species.

2.13 Statistical analysis

All statistical analysis was performed with GraphPad Prism, version 9 for MacOS (GraphPad Software, San Diego, California, USA, https://www.graphpad.com). Statistic tests are specified in the legend of each figure. When necessary, data for statistical analysis was transformed by applying the natural logarithm, which rendered the distribution of populations Gaussian. Statistical differences were considered significant when $p < 0.05$.

3

RESULTS[3]

The Results section will be subdivided in four parts. The first two parts (sections 3.1 and 3.2) will describe the results corresponding to the main topics of research, which comprise the determinants of the subcellular localization of CteG and the CteG-dependent host cell lytic exit of *C. trachomatis*. The two last parts (sections 3.3 and 3.4) comprise the identification of CteG host cell interacting partners, which is more preliminary, and the distribution of CteG homologs among *Chlamydiaceae*.

Regarding the investigation of CteG interacting partners (section 3.3), we obtained results from mass spectrometry but did not follow-up with further analyses. The results shown are preliminary and still need to be further validated.

The bioinformatic analysis of the distribution of CteG homologs among *Chlamydiaceae* (section 3.4) was mainly performed during the periods of lockdown due to COVID-19. This was planned as a possible back-up plan. The primary goal of this study was to set the foundations for an experimental analysis of the homologs of CteG among *Chlamydiaceae*, which is currently being carried out.

[3] This Chapter was written by Inês Serrano Pereira, based on the data obtained. The text and figures in section 3.2 are based on data from a recently published article (Pereira *et al.*, 2022) and are transcribed from the paper, with modifications.

3.1 Determinants of the subcellular localization of CteG[4]

In a previous study, it was shown that CteG is targeted to the Golgi complex and plasma membrane at different times of infection of host cells by *C. trachomatis* (Pais *et al.*, 2019). In mammalian cells ectopically producing CteG, the protein is directed to the same eukaryotic cell compartments and targeting to the Golgi appears to require its first 100 amino acid residues (Pais *et al.*, 2019). In this chapter, we aimed to identify specific residues of CteG involved in its subcellular targeting aiming at eventually testing their role in follow-up studies on the function of CteG. To determine the specific amino acid residues within the first 100 amino acid residues of CteG that mediate its targeting to the Golgi complex, deletion and alanine scanning analysis using ectopically expressed CteG fusion proteins was used. Then, the importance of the identified residues in the subcellular localization of CteG produced and delivered into host cells by *C. trachomatis* was assessed. A possible role of amino acid residues 350 to 656 (which we termed C-terminal region) of CteG on its localization at the plasma membrane of transfected mammalian cells was also analyzed.

3.1.1 The first 20 amino acid residues of CteG are necessary for the localization of mEGFP-CteG$_{1\text{-}100}$ at the Golgi

To define the region within the first 100 amino acid residues of CteG (CteG$_{1\text{-}100}$) responsible for the localization of the protein in the Golgi (Pais *et al.*, 2019), we transfected HeLa cells with plasmids encoding truncated versions of CteG$_{1\text{-}100}$ fused to the C-terminus of monomeric EGFP (mEGFP; 27 kDa), mEGFP-CteG$_{1\text{-}100}$ (38 kDa), mEGFP-CteG$_{20\text{-}100}$ (36 kDa), mEGFP-CteG$_{40\text{-}100}$ (34 kDa), mEGFP-CteG$_{60\text{-}100}$ (32 kDa), mEGFP-CteG$_{1\text{-}80}$ (35 kDa), mEGFP-CteG$_{1\text{-}60}$ (33 kDa), mEGFP-CteG$_{1\text{-}40}$ (32 kDa), mEGFP-CteG$_{1\text{-}30}$ (31 kDa), and mEGFP-CteG$_{1\text{-}20}$ (30 kDa) (Figure 3.1A).

[4] The experiments described in this section, which contains unpublished data, were entirely performed by Inês Serrano Pereira.

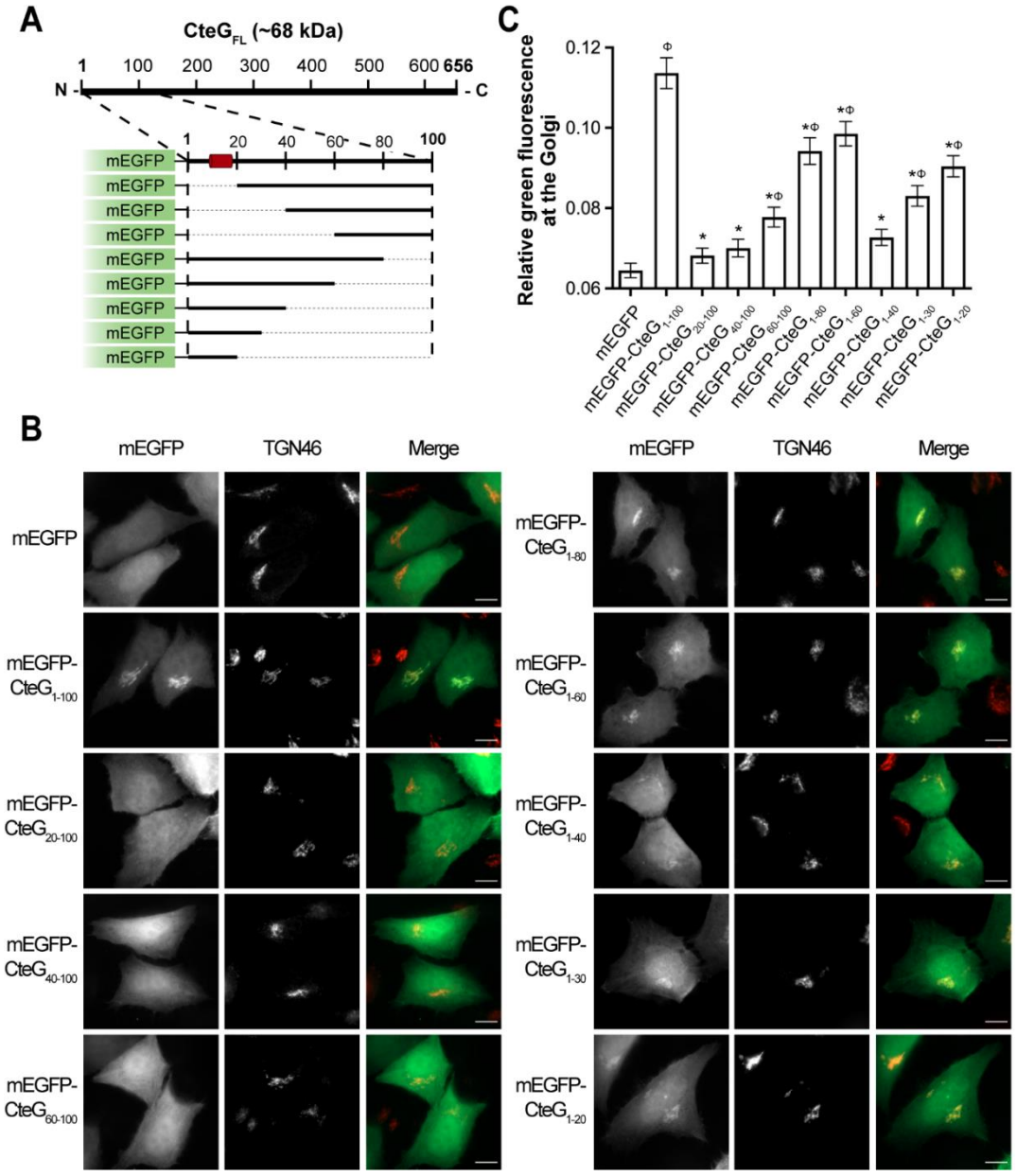


Figure 3.1. Analysis of the localization of ectopically produced mEGFP-CteG$_{1-100}$ truncated proteins. (A) Full-length (CteG$_{FL}$; 656 amino acid residues) possesses a Golgi-targeting motif within its first 100 residues (CteG$_{1-100}$), which is predicted to contain an α-helix (red; see Figure 3.2 below). HeLa cells were transfected with plasmids encoding CteG$_{1-100}$ truncations fused to the C-terminus of mEGFP for 24 h. **(B)** Transfected cells were fixed with PFA 4% (w/v) and immunolabelled with an antibody against TGN46 (which concentrates at the *trans*-Golgi network), and with the appropriate fluorophore-conjugated antibody. mEGFP-CteG$_{1-100}$ truncations (green) and TGN46-immunolabelled Golgi (red) were analyzed by fluorescence microscopy. Scale bar, 10 μm. **(C)** In each cell, the level of fluorescence of the mEGFP-CteG$_{1-100}$ truncated proteins co-localizing with the TGN46-immunolabelled Golgi was quantified relatively to their fluorescence levels in the whole cell (detailed in Annexes Figure 2). Data are mean ± standard error of the mean (SEM) of at least 30 individual cells from 3 independent experiments. Statistical analysis was performed with one-way ANOVA and Dunnett's post-test. $*$ and ϕ represent $p<0.05$ by comparison with mEGFP-CteG$_{1-100}$ or mEGFP alone, respectively.

Whole extracts of transfected cells were analyzed by immunoblotting, which confirmed the production of all fusion proteins and their migration on SDS-PAGE according to their predicted molecular mass (Annexes Figure 1A), except mEGFP-CteG$_{20-100}$ and mEGFP-CteG$_{1-80}$ that migrated slightly above their expected molecular mass (Annexes Figure 1A). Bands corresponding to species of lower molecular mass were observed, also in similar experiments throughout this work, and likely resulted from protein degradation (Annexes Figure 1A). A preliminary analysis by immunofluorescence microscopy of transfected cells revealed that some proteins lost their co-localization with the Golgi complex, whereas others still maintained that localization to some degree (Figure 3.1B). To quantify this, we analyzed in each cell the level of fluorescence of the mEGFP hybrid proteins co-localizing with immunolabelled TGN46 (which concentrates at the *trans*-Golgi network) relative to their fluorescence levels in the whole cell (Annexes Figure 2). This revealed that, when comparing with mEGFP-CteG$_{1-100}$, all ectopically produced CteG truncated versions displayed a lower degree of co-localization with the TGN46-immunollabeled Golgi (Figure 3.1C). Despite this, almost all fusion proteins lacking amino acid residues from the C-terminus of CteG$_{1-100}$ (mEGFP-CteG$_{1-80}$, mEGFP-CteG$_{1-60}$, mEGFP-CteG$_{1-30}$, and mEGFP-CteG$_{1-20}$; Figure 3.1A) still showed a marked co-localization with the Golgi (Figure 3.1B, C). The exception was mEGFP-CteG$_{1-40}$, which barely localized at the Golgi (Figure 3.1B, C). In contrast, the truncated proteins lacking amino acid residues from the N-terminal region of CteG$_{1-100}$ (mEGFP-CteG$_{20-100}$, mEGFP-CteG$_{40-100}$, and mEGFP-CteG$_{60-100}$) displayed essentially a cytosolic distribution (Figure 3.1B, C). Overall, this showed that the first 20 amino acid residues of CteG are crucial for the localization of mEGFP-CteG$_{1-100}$ at the Golgi in transfected HeLa cells. However, amino acid residues from position 21 to 100 also play a role for optimal targeting of mEGFP-CteG$_{1-100}$ to the Golgi.

3.1.2 Specific residues at a putative α-helix in the N-terminal region of CteG are necessary for the localization of mEGFP-CteG$_{1\text{-}100}$ at the Golgi

The secondary structure of CteG (deduced with PSIPRED, University College London, UK; http://bioinf.cs.ucl.ac.uk/psipred/; Annexes Figure 3) revealed a predicted α-helix between amino acid residues 9 to 17 (Figure 3.2A). A helical wheel diagram generated bioinformatically with these amino acid residues (using pepwheel, EMBOSS; https://www.bioinformatics.nl/cgi-bin/emboss/pepwheel) exhibited an amphipathic configuration with a non-polar face composed mostly of leucine and tryptophan residues, and a polar one rich in serine residues (Figure 3.2B). As amphipathic α-helixes can facilitate membrane insertion (Ulmschneider *et al.*, 2014), we questioned whether the amino acid residues of CteG that compose this hypothetical amphipathic α-helix could be important for the targeting of mEGFP-CteG$_{1\text{-}100}$ to the Golgi. For this, we started by transfecting HeLa cells with plasmids encoding mEGFP-CteG$_{1\text{-}100}$ mutant proteins in which groups of three or two amino acid residues between positions 2 and 19 of CteG were replaced by alanines (mEGFP-CteG$_{1\text{-}100\ (2\text{-}4\rightarrow AAA)}$, mEGFP-CteG$_{1\text{-}100\ (5\text{-}7\rightarrow AAA)}$, mEGFP-CteG$_{1\text{-}100\ (9\text{-}10\rightarrow AA)}$, mEGFP-CteG$_{1\text{-}100\ (11\text{-}13\rightarrow AAA)}$, mEGFP-CteG$_{1\text{-}100\ (14\text{-}16\rightarrow AAA)}$ and mEGFP-CteG$_{1\text{-}100\ (17\text{-}19\rightarrow AAA)}$; Figure 3.2A). Immunoblotting of whole cell extracts confirmed production of all proteins and their migration on SDS-PAGE according to the predicted molecular mass (Annexes Figure 1B). Immunofluorescence microscopy of transfected cells was used to visually assess the localization of mutant proteins (Figure 3.3) and to quantify their relative fluorescence at the TGN46-immunolabelled Golgi of transfected HeLa cells (Figure 3.2C). Comparing to mEGFP-CteG$_{1\text{-}100}$, the localization of mEGFP-CteG$_{1\text{-}100\ (2\text{-}4\rightarrow AAA)}$, mEGFP-CteG$_{1\text{-}100\ (5\text{-}7\rightarrow AAA)}$, mEGFP-CteG$_{1\text{-}100\ (8\text{-}10\rightarrow AAA)}$ and mEGFP-CteG$_{1\text{-}100\ (14\text{-}16\rightarrow AAA)}$ at the Golgi was moderately affected (~1.3 to 1.8-fold) (Figures 3.2C and 3.3). In contrast, the localization of mEGFP-CteG$_{1\text{-}100\ (11\text{-}13\rightarrow AAA)}$ and mEGFP-CteG$_{1\text{-}100\ (17\text{-}19\rightarrow AAA)}$ was severely affected (> 2-fold) by the amino acid replacements as these proteins often presented a cytosolic distribution resembling mEGFP alone (Figures 3.2C and 3.3). Overall, this indicated that the region between amino acid residues 11 and 19 of CteG, which overlaps with the putative amphipathic α-helix (Figure 3.2A), is essential for the localization of mEGFP-CteG$_{1\text{-}100}$ at the Golgi.

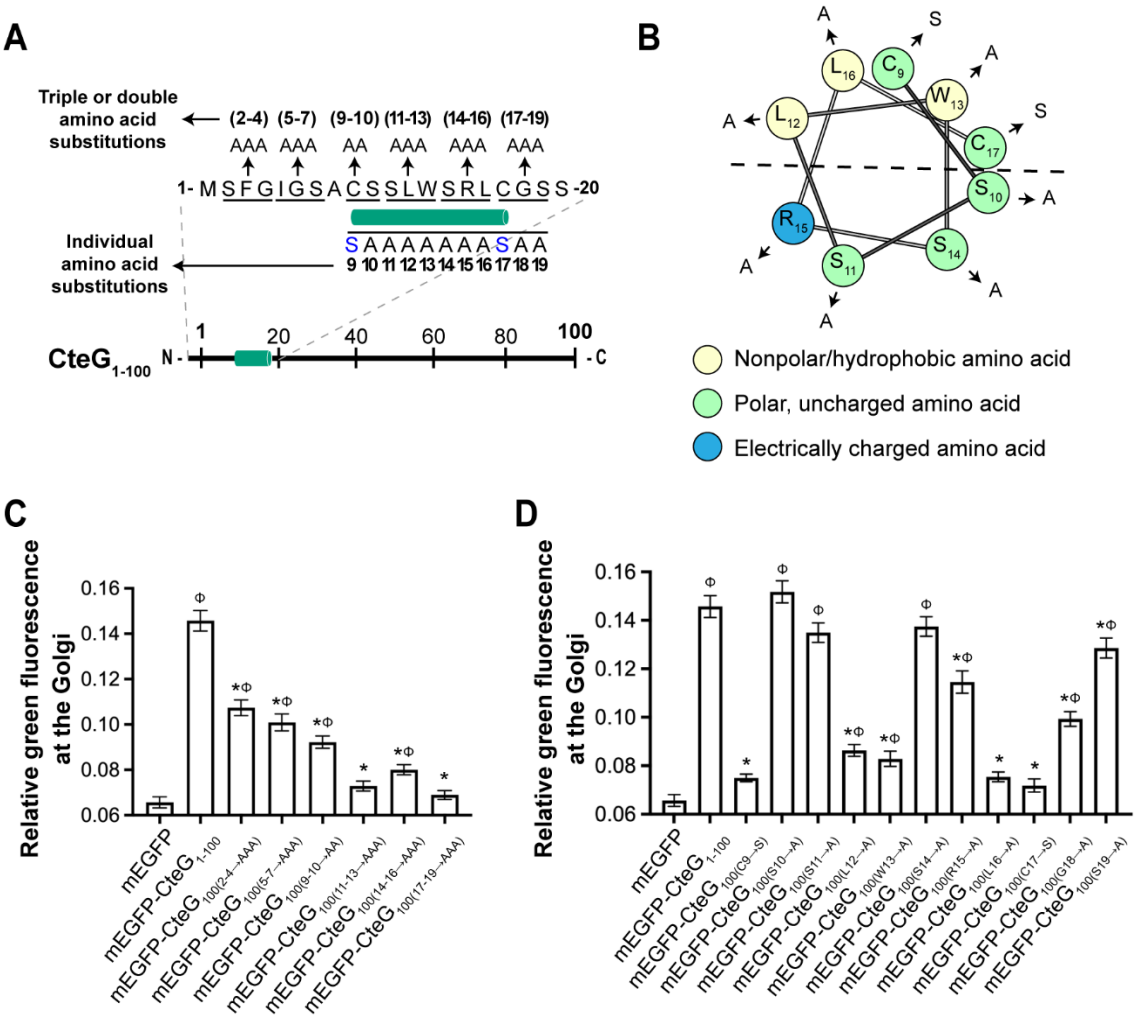


Figure 3.2. A putative α-helix at the N-terminal region of CteG is important for the localization of mEGFP-CteG$_{1\text{-}100}$ at the Golgi. **(A)** Schematic representation of CteG$_{1\text{-}100}$ featuring the first 20 amino acid residues of CteG and the mutagenesis performed within this region. Substitutions between positions 2 and 19 comprised either groups of three or two residues that were replaced by alanines, or individual residues that were replaced by serine (blue) or alanine (also see panel B). An α-helix, which is predicted to occur between amino acid residues 9 and 17 (Annexes Figure 3), is depicted (green). **(B)** A helical wheel diagram generated with amino acids 9 to 17 of CteG (using pepwheel, EMBOSS; https://www.bioinformatics.nl/cgi-bin/emboss/pepwheel) reveals an amphipathic α-helix whose possible hydrophobic and hydrophilic faces are divided by the dashed line. Replacements of each individual amino acid residue by serine or alanine are represented (see panel A). **(C, D)** Quantification of the relative fluorescence of each mEGFP-CteG$_{1\text{-}100}$ protein containing multiple **(C)** or single **(D)** amino acid replacements at the TGN46-immunolabelled Golgi of transfected HeLa cells (detailed in Annexes Figure 2). Data correspond to mean ± SEM of at least 30 individual cells from 3 independent experiments. Statistical analysis was performed with one-way ANOVA and Dunnett's post-test. ∗ and ϕ represent $p<0.05$ by comparison with mEGFP-CteG$_{1\text{-}100}$ or mEGFP alone, respectively.

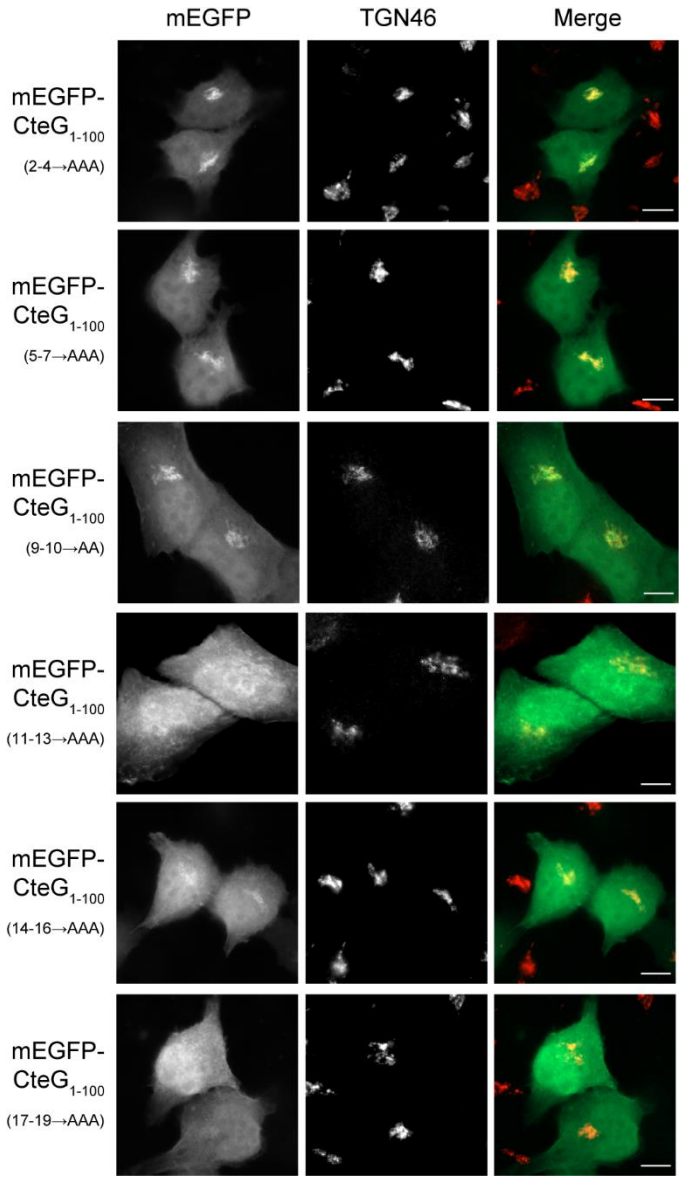


Figure 3.3. Analysis of the subcellular localization of mEGFP-CteG$_{1\text{-}100}$ proteins containing two or three amino acid replacements. HeLa cells ectopically producing mEGFP-CteG$_{1\text{-}100}$ proteins with two or three amino acid replacements were fixed at 24 h post-transfection with PFA 4% (w/v) and immunolabelled with an antibody against TGN46 (*trans*-Golgi network), and with the appropriate fluorophore-conjugated antibody. mEGFP-CteG$_{1\text{-}100}$ mutant proteins (green) and TGN46-immunolabelled Golgi (red) were analyzed by fluorescence microscopy. Scale bar, 10 μm.

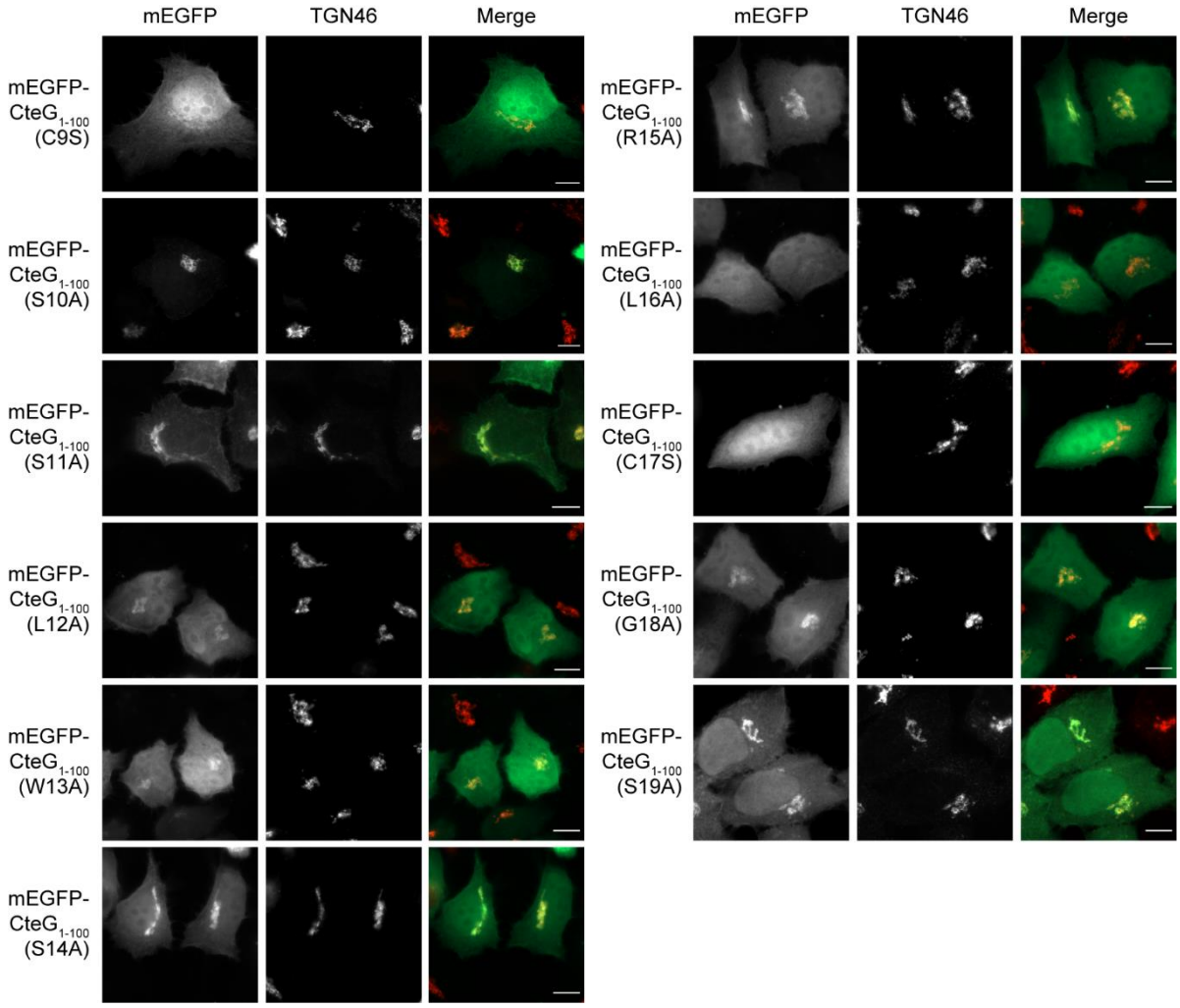


Figure 3.4. Analysis of the subcellular localization of mEGFP-CteG$_{1\text{-}100}$ proteins containing single amino acid replacements. HeLa cells ectopically producing mEGFP-CteG$_{1\text{-}100}$ proteins with single amino acid replacements were fixed with PFA 4% (w/v) at 24 h post-transfection and immunolabelled with an antibody against TGN46 (*trans*-Golgi network), and with the appropriate fluorophore-conjugated antibody. mEGFP-CteG$_{1\text{-}100}$ mutant proteins (green) and TGN46-immunolabelled Golgi (red) were examined by fluorescence microscopy. Scale bar, 10 μm.

To finely map the residues that target ectopically produced mEGFP-CteG$_{1\text{-}100}$ to the Golgi, we substituted each amino acid residue between the positions 9 and 19 by alanines or serines (mEGFP-CteG$_{1\text{-}100\ (C9\rightarrow S)}$, mEGFP-CteG$_{1\text{-}100\ (S10\rightarrow A)}$, mEGFP-CteG$_{1\text{-}100\ (S11\rightarrow A)}$, mEGFP-CteG$_{1\text{-}100\ (L12\rightarrow A)}$, mEGFP-CteG$_{1\text{-}100\ (W13\rightarrow A)}$, mEGFP-CteG$_{1\text{-}100\ (S14\rightarrow A)}$, mEGFP-CteG$_{1\text{-}100\ (R15\rightarrow A)}$, mEGFP-CteG$_{1\text{-}100\ (L16\rightarrow A)}$, mEGFP-CteG$_{1\text{-}100\ (C17\rightarrow S)}$, mEGFP-CteG$_{1\text{-}100\ (G18\rightarrow A)}$ and mEGFP-CteG$_{1\text{-}100\ (S19\rightarrow A)}$; Figure 3.2A). Specifically, residues C_9 and C_{17} were substituted by serine (Figures 3.2A, B) as cysteine and serine only differ in their thiol and hydroxyl groups. Residues C_9 and S_{10} were also considered for the analysis as they are still part of the α-helix predicted in that region (Figures 3.2A, B). Likewise, residues G_{18} and S_{19} were also considered as mEGFP-CteG$_{1\text{-}100\ (17\text{-}19\rightarrow AAA)}$ displayed significantly decreased co-localization with the TGN46-immunolabelled

Golgi (Figure 3.2C). Analysis by immunoblotting of whole extracts of transfected cells confirmed the production of all proteins with the predicted molecular mass (Annexes Figure 1C). Immunofluorescence microscopy analysis of cells ectopically producing the proteins (Figure 3.4) and quantification of their relative fluorescence at the TGN46-immunolabelled Golgi (Figure 3.2D) revealed that the localization of mEGFP-CteG$_{1\text{-}100}$ was mostly affected by the substitution of amino acid residues in the hydrophobic face of the putative amphipathic α-helix [mEGFP-CteG$_{1\text{-}100\ (C9\rightarrow S)}$, mEGFP-CteG$_{1\text{-}100\ (L12\rightarrow A)}$, mEGFP-CteG$_{1\text{-}100\ (W13\rightarrow A)}$, mEGFP-CteG$_{1\text{-}100\ (L16\rightarrow A)}$, mEGFP-CteG$_{1\text{-}100\ (C17\rightarrow S)}$] (Figures 3.2D and 3.4), and not affected or only slightly/moderately affected by the modifications in the hydrophilic face of the putative amphipathic α-helix [mEGFP-CteG$_{1\text{-}100\ (S10\rightarrow A)}$, mEGFP-CteG$_{1\text{-}100\ (S11\rightarrow A)}$, mEGFP-CteG$_{1\text{-}100\ (S14\rightarrow A)}$, mEGFP-CteG$_{1\text{-}100\ (R15\rightarrow A)}$] or in residues G$_{18}$ or S$_{19}$ (Figures 3.2D and 3.4). Therefore, these results suggest that the hydrophobic face of the putative α-helix present at the N-terminal region of CteG (Figures 3.2A, B) could be a structure that targets mEGFP-CteG$_{1\text{-}100}$ to the Golgi.

3.1.3 Specific residues at a putative α-helix in the N-terminal region of CteG are required for an efficient targeting of mEGFP-CteG to the Golgi and plasma membrane

Our previous work showed that while ectopically expressed full-length CteG (CteG$_{FL}$; 656 amino acid residues; 68 kDa; Figure 3.1A) fused to mEGFP (mEGFP-CteG$_{FL}$) predominantly localizes at the plasma membrane of HeLa cells, it also localizes at the Golgi in some cells (Pais *et al.*, 2019). However, when analyzing an ectopically produced mEGFP-CteG hybrid protein without its first 100 residues (mEGFP-CteG$_{101\text{-}656}$), its localization was mostly cytosolic (Pais *et al.*, 2019).

Following our previous work and the mutagenesis analysis of CteG$_{1\text{-}100}$ (Figures 3.1-3.4), we sought to test the effect of smaller deletions and specific amino acid substitutions in the first 100 amino acids of CteG on the localization of mEGFP-CteG hybrid proteins. First, we transfected HeLa cells with plasmids encoding mEGFP-CteG$_{FL}$ (96 kDa), or mEGFP-CteG proteins lacking the first 20, 40 or 60 N-terminal amino acid residues of CteG [mEGFP-CteG$_{20\text{-}656}$ (94 kDa); mEGFP-CteG$_{40\text{-}656}$ (92 kDa); and mEGFP-CteG$_{60\text{-}656}$ (90 kDa)]. Immunoblotting confirmed that all proteins were produced and migrated on SDS-PAGE according to their predicted molecular mass (Annexes Figure 4A). Then, the localization of each protein was assessed by immunofluorescence microscopy in terms of their localization at the TGN46-immunolabelled Golgi, plasma membrane or cytosol of the transfected cells. Consistent with previous results (Pais *et al.*, 2019), the predominant localization of mEGFP-CteG$_{FL}$ was at the plasma membrane (89 ± 1% of transfected cells), but the protein was also detected at the Golgi

(26 ± 2%) and more rarely it was considered essentially cytosolic (6 ± 1%) (Figures 3.5A, B). The categorization of the protein into plasma membrane or Golgi localization includes cases where the protein was seen in both compartments (21 ± 3%) (Figure 3.5A). Therefore, while mEGFP-CteG$_{FL}$ mostly localizes at the plasma membrane its distribution varies among transfected cells. By comparison to mEGFP-CteG$_{FL}$, the localization of ectopically produced mEGFP-CteG$_{20-656}$, mEGFP-CteG$_{40-656}$, or mEGFP-CteG$_{60-656}$ was more frequently cytosolic (~50%) and less often at the Golgi (~10%) or at the plasma membrane (~40%) (Figure 3.5B and Annexes Figure 5), without significant differences among these three truncated proteins (Figure 3.5B and Annexes Figure 5). These results indicate that the first 20 amino acid residues of CteG contain determinants required for an efficient targeting of mEGFP-CteG$_{FL}$ to the plasma membrane and Golgi of transfected cells.

To assess the impact of specific amino acid residues at the N-terminal region of CteG in the localization of mEGFP-CteG$_{FL}$, HeLa cells were transfected with plasmids encoding mEGFP-CteG$_{FL}$ mutant proteins [mEGFP-CteG$_{FL\ (C9S,\ C17S)}$, mEGFP-CteG$_{FL\ (W13A,\ C17S)}$, mEGFP-CteG$_{FL\ (L16A,\ C17S)}$ and mEGFP-CteG$_{FL\ (C9S,\ L12A,\ W13A,\ L16A,\ C17S)}$], based on our previous site-directed mutagenesis analysis of the determinants of the localization of mEGFP-CteG$_{1-100}$ at the Golgi (Figure 3.2). Immunoblotting of whole extracts of HeLa cells transfected with these plasmids confirmed expression of all fusion mutant proteins and their migration on SDS-PAGE according to the predicted molecular mass (Annexes Figure 4B). Then, the subcellular localization of each protein was evaluated by immunofluorescence microscopy (Figure 3.5C and Annexes Figure 6). In general, all four mutant proteins appeared more frequently cytosolic than mEGFP-CteG$_{FL}$ and less often at the Golgi or at the plasma membrane (Figure 3.5C and Annexes Figure 6). Overall, this indicates that the putative amphipathic α-helix is important for the localization of ectopically produced mEGFP-CteG$_{FL}$ at the plasma membrane and Golgi.

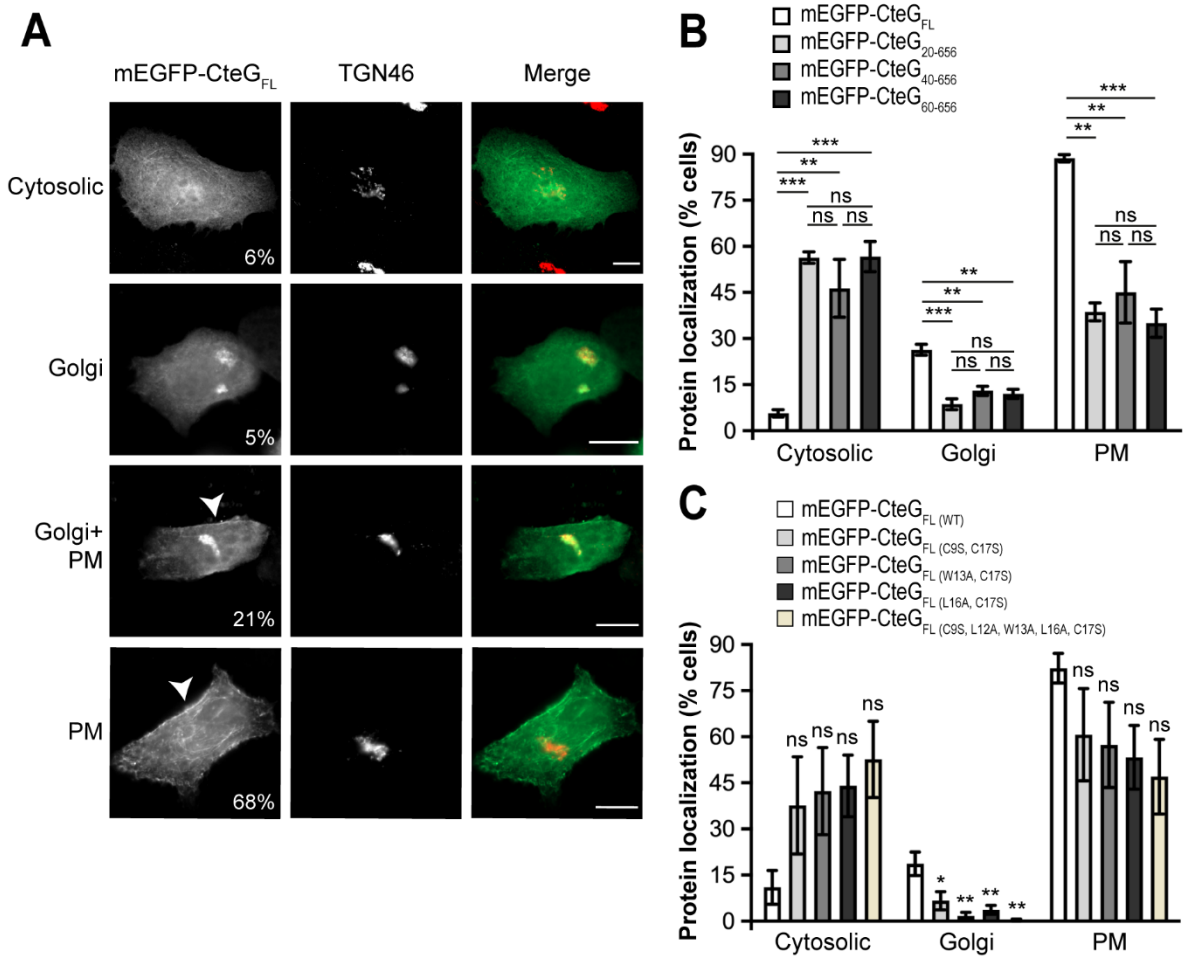


Figure 3.5. Specific amino acids at a putative α-helix in the N-terminal region of CteG are important for the subcellular localization of ectopically produced mEGFP-CteG. HeLa cells ectopically producing mEGFP-CteG proteins with deletions or specific amino acid replacements in CteG were fixed with PFA 4% (w/v) at 24 h post-transfection, immunolabelled with an antibody against TGN46 (*trans*-Golgi network) and with the appropriate fluorophore-conjugated antibody and analyzed by fluorescence microscopy. **(A)** HeLa cells ectopically producing and displaying mEGFP-CteG$_{FL}$ at the cell cytoplasm, TGN46-immunolabelled Golgi, plasma membrane (PM) or both at TGN46-immunolabelled Golgi and PM. Percentages refer to the cell population where mEGFP-CteG$_{FL}$ exhibits that specific localization (and correspond to data from panel B for mEGFP-CteG$_{FL}$). Arrows point to mEGFP-CteG$_{FL}$ localizing at the PM. Scale bar, 10 μm. **(B, C)** All mEGFP hybrid proteins were categorized into the localizations shown in panel A: **(B)** mEGFP-CteG proteins containing deletions at the N-terminal region of CteG, and **(C)** mEGFP-CteG proteins containing specific amino acid replacements in CteG. In (B) and (C), data are mean ± SEM of 3 independent experiments (N=100 in each experiment). Statistical analysis was performed with one-way ANOVA and Tukey's post-test in (B), and with one-way ANOVA and Dunnett's post-test relatively to "wild-type" mEGFP-CteG$_{FL}$ in (C). ns, non-significant; ∗p<0.05; ∗∗p<0.01; ∗∗∗p<0.001.

3.1.4 Inhibition of host S-palmitoyltransferases causes a minor impact in the localization of mEGFP-CteG$_{1\text{-}100}$ at the Golgi

S-palmitoylation is a protein posttranslational modification characterized by the reversible addition of a lipid chain to cysteine residues by enzymes termed palmitoyltransferases. These modifications can increase the affinity of proteins to membranes (Resh, 2006). A palmitoylation site prediction tool [CSS-Palm v.4.0, (Ren *et al.*, 2008)] identified several cysteine residues on CteG which can be palmitoylated, including C_9 and C_{17} within its first 100 amino acid residues (Annexes Figure 7), which when replaced by serines resulted in a major impact in the localization of mEGFP-CteG$_{1\text{-}100}$ (Figure 3.2) and affected the localization of mEGFP-CteG$_{FL}$ (Figure 3.5). Therefore, we assessed the effect of S-palmitoylation inhibition on the localization of ectopically expressed CteG at the Golgi. For this, HeLa cells were transfected with plasmids encoding mEGFP-CteG$_{1\text{-}100}$, mEGFP-CteG$_{FL}$ and mEGFP alone, in the presence or absence of the mammalian palmitoyltransferase inhibitor 2-Bromopalmitate (2BP). Cells ectopically expressing the *L. pneumophila* protein GobX fused to the C-terminus of mEGFP (mEGFP-GobX) were also analyzed as control, as this protein is targeted to the Golgi membrane by S-palmitoylation of a cysteine residue localized in an amphipathic alpha helix and this is inhibited by 2BP (Lin *et al.*, 2015). Immunofluorescence microscopy of transfected cells followed by qualitative and quantitative analysis of localization at the Golgi confirmed the redistribution of mEGFP-GobX from the Golgi to the cell cytosol under the presence of 2BP (Figure 3.6) (Lin *et al.*, 2015), which confirmed the functionality of 2BP. However, the analyses did not reveal any discernible difference on the localization of mEGFP-CteG$_{FL}$ at the Golgi complex under the presence or absence of 2BP (Figure 3.6). On the other hand, although an obvious difference in the localization of mEGFP-CteG$_{1\text{-}100}$ was not noticeable (Figure 3.6A), the quantification of the relative fluorescence of mEGFP-CteG$_{1\text{-}100}$ at the Golgi revealed a slight but significant difference between cells treated or not with 2BP (Figure 3.6B).

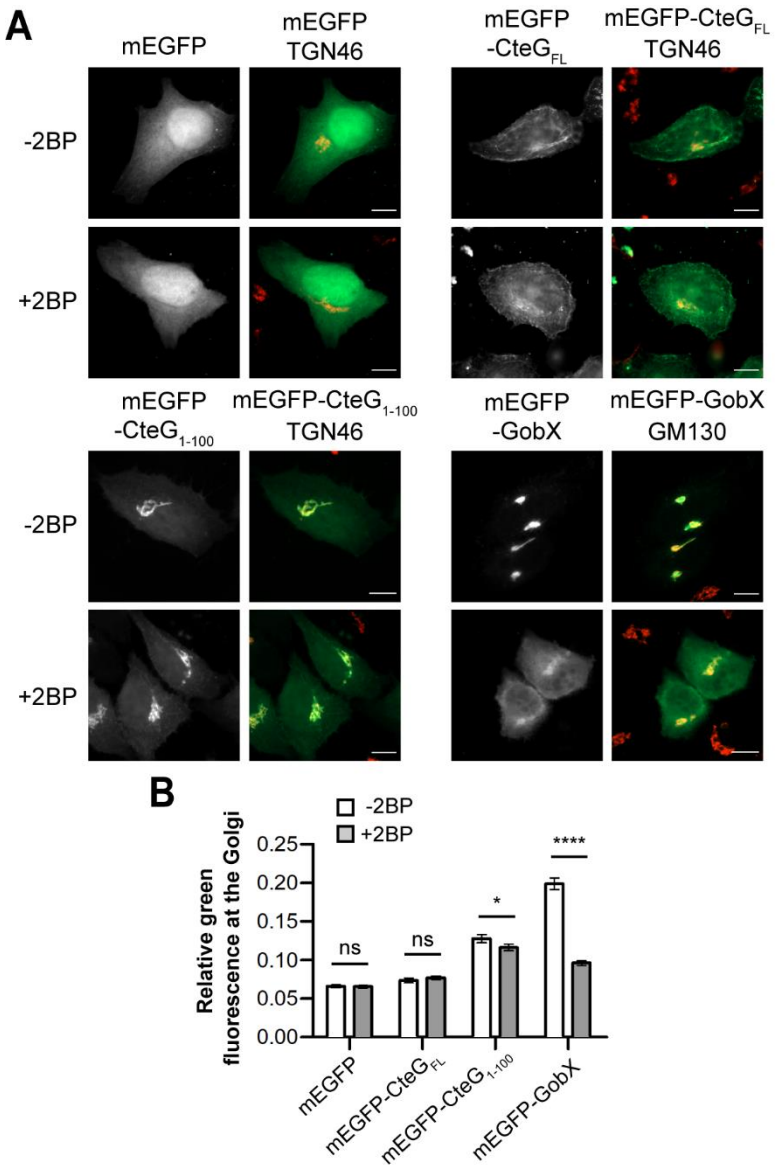


Figure 3.6. Effect of the inhibition of mammalian palmitoyltransferases in the localization of mEGFP-CteG$_{1-100}$ and mEGFP-CteG$_{FL}$ at the Golgi. (A) HeLa cells were transfected with plasmids encoding mEGFP-CteG$_{FL}$, mEGFP-CteG$_{1-100}$, mEGFP-CteG alone or mEGFP-GobX (control) in the presence or absence of 2-bromopalmitate (2BP), an inhibitor of mammalian palmitoyltransferases. Cells were fixed with PFA 4% (w/v), immunolabelled with an antibody against TGN46 (*trans*-Golgi network) or the *cis*-Golgi matrix protein GM130 (red) in the case of mEGFP-GobX, and with appropriate fluorophore-conjugated antibodies, and analyzed by fluorescence microscopy. Scale bar, 10 μm. **(B)** Quantification of the relative fluorescence of each protein at the TGN46- (mEGFP alone, mEGFP-CteG$_{FL}$ or mEGFP-CteG$_{1-100}$) or at the GM130-immunolabelled Golgi (mEGFP-GobX) of transfected HeLa cells (detailed in Annexes Figure 2). Data are mean ± SEM of at least 20 individual cells from 3 independent experiments. For each protein, statistical analysis was performed with a two-tailed unpaired Student's t-test between presence or absence of 2BP. ns, non-significant; ∗p<0.05; ∗∗∗∗p<0.0001.

Overall, these results suggest that while S-palmitoylation of cysteine residues may play a role in targeting ectopically expressed mEGFP-CteG$_{1-100}$ to the Golgi, other mechanisms should be involved. Moreover, the inhibition of this posttranslational modification does not affect the localization of ectopically produced mEGFP-CteG$_{FL}$ at the Golgi.

3.1.5 Analysis of the impact of different regions of CteG in the subcellular localization of mEGFP-CteG proteins at the plasma membrane

To analyze other possible determinants of the localization of mEGFP-CteG$_{FL}$ at the plasma membrane, we generated several transfection plasmids encoding hybrid proteins with truncations in CteG (Figure 3.7A), designed based on its predicted secondary structure (Annexes Figure 3). These plasmids encoding mEGFP-CteG$_{1-350}$ (62 kDa), mEGFP-CteG$_{1-443}$ (73 kDa), mEGFP-CteG$_{1-538}$ (83 kDa), mEGFP-CteG$_{1-598}$ (90 kDa), mEGFP-CteG$_{1-628}$ (93 kDa), mEGFP-CteG$_{350-656}$ (62 kDa), mEGFP-CteG$_{443-656}$ (52 kDa), mEGFP-CteG$_{538-656}$ (41 kDa), or mEGFP-CteG$_{598-656}$ (35 kDa) were then used to transfect HeLa cells. Immunoblotting of whole cell extracts confirmed production of most fusion proteins with the predicted molecular mass (Annexes Figure 8). mEGFP-CteG$_{1-628}$ was poorly produced but still migrated on SDS-PAGE according to the predicted molecular mass (Annexes Figure 8, left-hand panel, lane 7). The production of mEGFP-CteG$_{1-350}$ could not be detected by immunoblotting (Annexes Figure 8, left-hand panel, lane 4), although a fluorescence signal was detected when analyzing mEGFP-CteG$_{1-350}$ by immunofluorescence microscopy (Figure 3.7B). In some cases, an appreciable degree of degradation was observable (Annexes Figure 8; mEGFP-CteG$_{1-443}$, left-hand panel, lane 5; mEGFP-CteG$_{350-656}$ and mEGFP-CteG$_{598-656}$, right-hand panel, lanes 3 and 6). Immunofluorescence microscopy of transfected cells was employed to collect images of cells representing the most predominant localization of each protein (Figure 3.7B) and to quantify the number of cells where each hybrid protein was observed at the cell plasma membrane (example in Figure 3.5A, and Figure 3.7C). Relative to mEGFP-CteG$_{FL}$, all truncated mEGFP-CteG proteins were defective for the frequency at which they appear at the plasma membrane (Figure 3.7C). However, many of the truncated mEGFP-CteG proteins analyzed appeared in the form of puncta scattered in the cytoplasm of transfected cells, which may indicate misfolding or degradation after protein production (Figure 3.7B). In some cases, this was supported by the appearance of fast migrating species in immunoblotting (Annexes Figure 8). Therefore, it was not possible from these experiments to pinpoint a specific region of CteG that is critical for mEGFP-CteG$_{FL}$ localization at the plasma membrane.

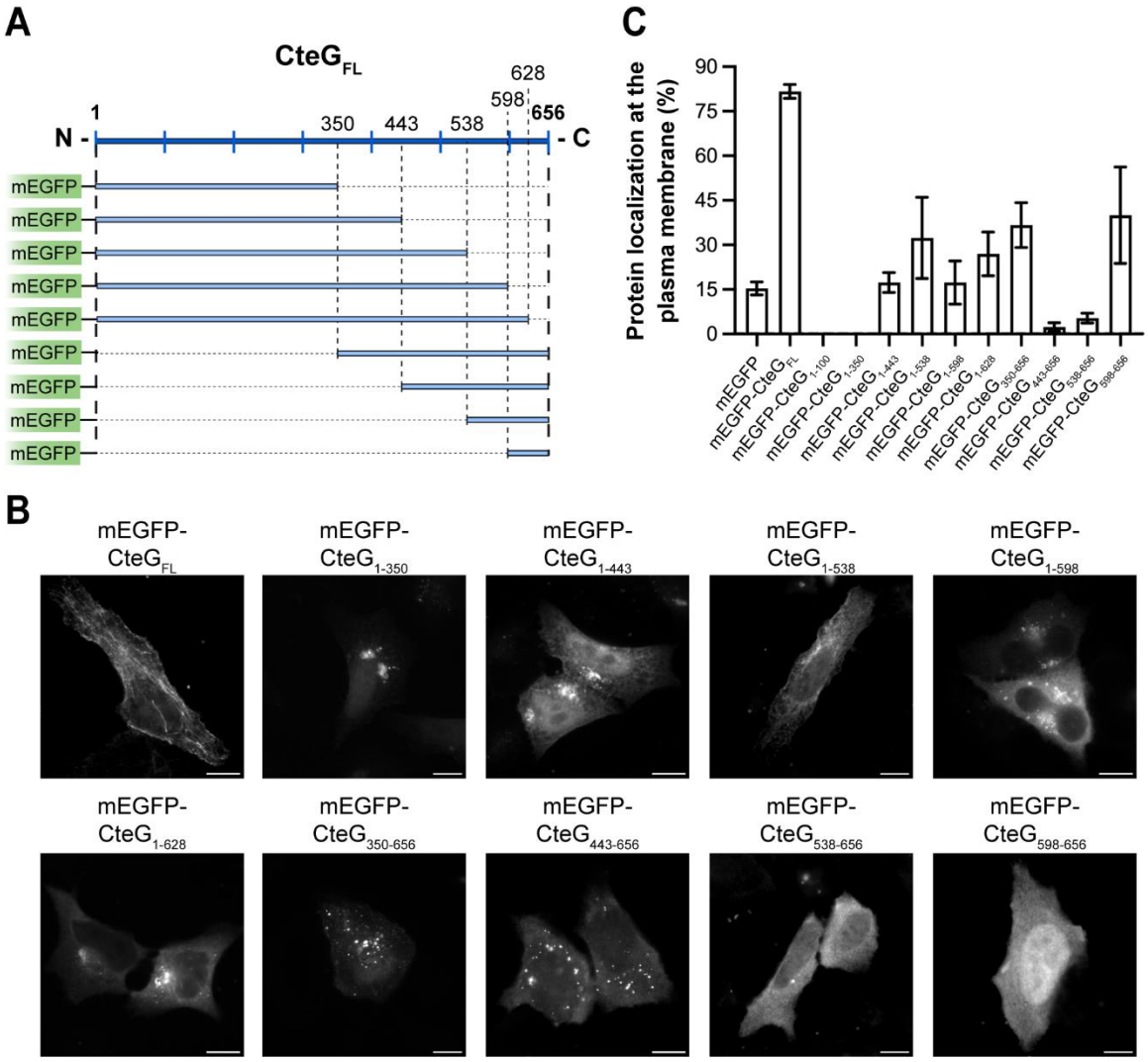


Figure 3.7. Analysis of the importance of different regions of CteG in the localization of mEGFP-CteG proteins at the plasma membrane. (A) Mammalian expression vectors encoding hybrid proteins with truncations in CteG as depicted, fused to the C-terminus of mEGFP, were constructed. **(B, C)** Hela cells were transfected with these plasmids, fixed with PFA 4% (w/v) at 24 h post-transfection and analyzed by fluorescence microscopy. **(B)** Images of HeLa cells representing the most predominant localization of each mEGFP hybrid protein. Scale bar, 10 μm. **(C)** Quantification of the number of cells where each of the mEGFP hybrid proteins localized at the cell plasma membrane. Data are mean ± SEM of 3 independent experiments (N=100 for each experiment).

3.1.6 The substitution of specific amino acid residues at the N-terminal region of CteG-2HA affects its type III secretion (T3S) system-mediated delivery and subcellular localization during *C. trachomatis* infection of HeLa cells

To test if the residues in the putative α-helix of CteG that are important for accurate localization of mEGFP-CteG$_{1\text{-}100}$ or mEGFP-CteG$_{FL}$ are also relevant for the localization of CteG in infected cells, we transformed a previously generated CteG-deficient *C. trachomatis cteG::aadA* strain (Pais *et al.*, 2019) with plasmids encoding double mutant CteG$_{C9S,\ C17S}$ with a double hemagglutinin tag (CteG$_{2aa}$-2HA) or quintuple mutant CteG$_{C9S,\ L12A,\ W13A,\ L16A,\ C17S}$-2HA (CteG$_{5aa}$-2HA) proteins. A previously generated *cteG::aadA*-derived strain producing wild-type CteG-2HA from a plasmid [CteG$_{WT}$-2HA] was used for comparison. HeLa cells were infected with these strains for 24 or 40 h and the infected cells were then analyzed by immunoblotting (Figure 3.8) and immunofluorescence microscopy (Figure 3.9A) for production and subcellular localization of the CteG-2HA proteins. Immunoblotting of extracts of HeLa cells infected with these strains revealed that CteG$_{WT}$-2HA and CteG$_{2aa}$-2HA migrated on SDS-PAGE according to the predicted molecular mass (68 kDa; Figure 3.8), and both presented a characteristic pattern of fast migrating species that has previously been observed for CteG$_{WT}$-2HA (Pais *et al.*, 2019). A CteG$_{5aa}$-2HA species with the predicted molecular mass was barely detected at 40 h post-infection, while a predominant species that migrated below the predicted molecular mass was clearly detected at 24 and 40 h post-infection (Figure 3.8).

Immunofluorescence analysis of HeLa cells infected with the same *C. trachomatis* strains confirmed production of the proteins at both 24 and 40 h post-infection (Figure 3.9A). As CteG$_{WT}$-2HA, CteG$_{2aa}$-2HA concentrated mainly at the Golgi complex and plasma membrane of host cells at 24 and 40 h post-infection, respectively (Figure 3.9A) (Pais *et al.*, 2019). However, at a first glance, CteG$_{2aa}$-2HA appeared to have a weaker host cell cytoplasmic signal which was consistent with the immunoblotting data (Figure 3.8). In contrast, CteG$_{5aa}$-2HA appeared to be retained within chlamydiae at both time-points analyzed (Figure 3.9A). To clarify these observations, we performed a quantification of the mean fluorescence signal of each protein in the cytoplasm of cells infected for 24 or 40 h, excluding the chlamydial inclusion (Figure 3.10A). In each case, non-infected neighboring cells were used for quantification of residual non-specific signal (Figure 3.10A). In general, this confirmed that CteG$_{2aa}$-2HA is slightly less abundant than CteG$_{WT}$-2HA in the cytoplasm of infected host cells, and that CteG$_{5aa}$-2HA is not detected in the cytoplasm of infected host cells (Figure 3.11A). Therefore, CteG$_{5aa}$-2HA was not considered for further analysis.

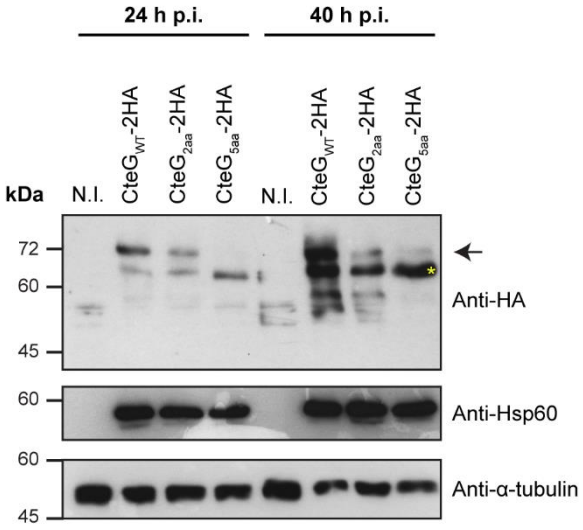


Figure 3.8. Analysis of the effect of specific amino acid replacements at the N-terminal region of CteG-2HA on its production during infection of HeLa cells by *C. trachomatis*. HeLa cells were left non-infected (N.I.), or were infected for 24 h or 40 h with *cteG::aadA*-derived *C. trachomatis* strains producing from plasmids wild-type CteG-2HA (CteG$_{WT}$-2HA), double mutant CteG$_{C9S,C17S}$ fused to a double hemagglutinin tag (CteG$_{2aa}$-2HA) or quintuple mutant CteG$_{C9S, L12A, W13A, L16A, C17S}$-2HA (CteG$_{5aa}$-2HA), at a multiplicity of infection of 0.3. Whole cell extracts were analyzed by immunoblotting using antibodies against HA, chlamydial heat shock protein 60 (Hsp60) and human α-tubulin, and appropriate HRP-conjugated secondary antibodies. CteG-2HA proteins were detected using the SuperSignal West Femto detection kit (Thermo Fisher Scientific), and Hsp60 and α-tubulin were detected using SuperSignal West Pico detection kit (Thermo Fisher Scientific). The arrow indicates the expected molecular mass for all CteG-2HA proteins (68 kDa). $*$ indicates a species of lower molecular mass predominantly detected for CteG$_{5aa}$-2HA.

A weaker signal of CteG$_{2aa}$-2HA in the cytoplasm of infected cells could be linked to a lower protein production or to a decreased delivery by the type III secretion (T3S) machinery. To discriminate between these two possibilities, we determined the total fluorescence signal of CteG$_{WT}$-2HA or CteG$_{2aa}$-2HA in the cytoplasm of each infected cell but excluding the signal co-localizing with the chlamydial inclusion (Figure 3.10B). This value was divided by the total protein fluorescence determined for the same cell (Figure 3.10B). Interestingly, the proportion of CteG$_{2aa}$-2HA in the cytoplasm of infected cells even slightly surpassed that of CteG$_{WT}$-2HA (Figure 3.11B). Therefore, the substitution of C_9 and C_{17} affected the global production of CteG-2HA rather than its T3S system-mediated transport into the cell cytoplasm.

Finally, we sought to compare quantitatively the localization of CteG$_{WT}$-2HA and CteG$_{2aa}$-2HA at the Golgi complex of cells infected for 24 h and at the plasma membrane of cells infected for 40 h. For this, we determined the total fluorescence intensity of the two proteins in these compartments relative to their total host cytoplasmic fluorescence (excluding the inclusion) (Figures 3.12A, B). We observed that the co-localization of CteG$_{2aa}$-2HA with the Golgi was significantly diminished by comparison with CteG$_{WT}$-2HA (Figure 3.9B). Similarly, the localization at the host cell plasma membrane was also significantly affected by the amino acid replacements analyzed (Figure 3.9C).

Overall, these results suggest that the replacement of C_9 and C_{17} of CteG-2HA affect its localization at the host cell Golgi or at the plasma membrane during infection but does not interfere with its T3S system-mediated delivery into host cells. However, the substitution of C_9, L_{12}, W_{13}, L_{16} and C_{17} prevented the T3S system-mediated delivery of CteG-2HA into the host cell.

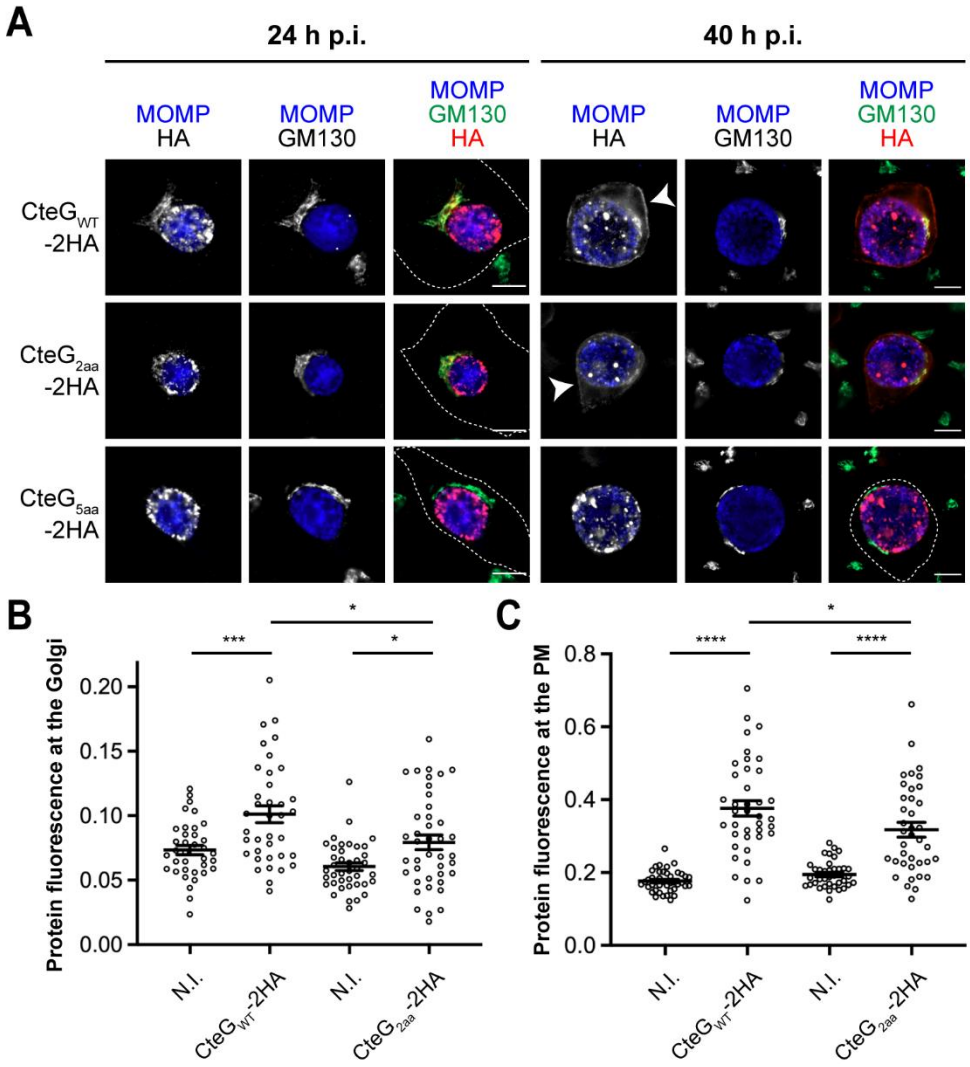


Figure 3.9. Analysis of the effect of specific amino acid replacements at the N-terminal region of CteG-2HA on its localization during infection of HeLa cells by *C. trachomatis*. HeLa cells were infected for 24 h or 40 h with *cteG::aadA*-derived *C. trachomatis* strains producing CteG$_{WT}$-2HA, double mutant CteG$_{2aa}$-2HA or quintuple mutant CteG$_{5aa}$-2HA, at a multiplicity of infection of 0.3. **(A)** Infected cells were fixed with methanol and immunolabelled with antibodies against HA (red), the *cis*-Golgi matrix protein GM130 (green) and *C. trachomatis* MOMP (blue), and appropriate fluorophore-conjugated antibodies. Arrows indicate CteG$_{WT}$-2HA and CteG$_{2aa}$-2HA accumulated at the host cell plasma membrane. Scale bar, 10 μm. **(B, C)** Quantitative analysis of the mean fluorescence of CteG$_{WT}$-2HA and CteG$_{2aa}$-2HA at the GM130-immunolabelled Golgi of cells infected for 24 h **(B)**, or at the plasma membrane (PM) of cells infected for 40 h **(C)** by comparison with residual non-specific signal in non-infected (N.I.) neighboring cells (see Figure 3.12 below). Data are the mean and SEM of at least 36 individual cells from 2 independent experiments. In (B) and (C), statistical analysis was performed with one-way ANOVA + Tukey's post-test. $*p<0.05$; $***p<0.001$; $****p<0.0001$.

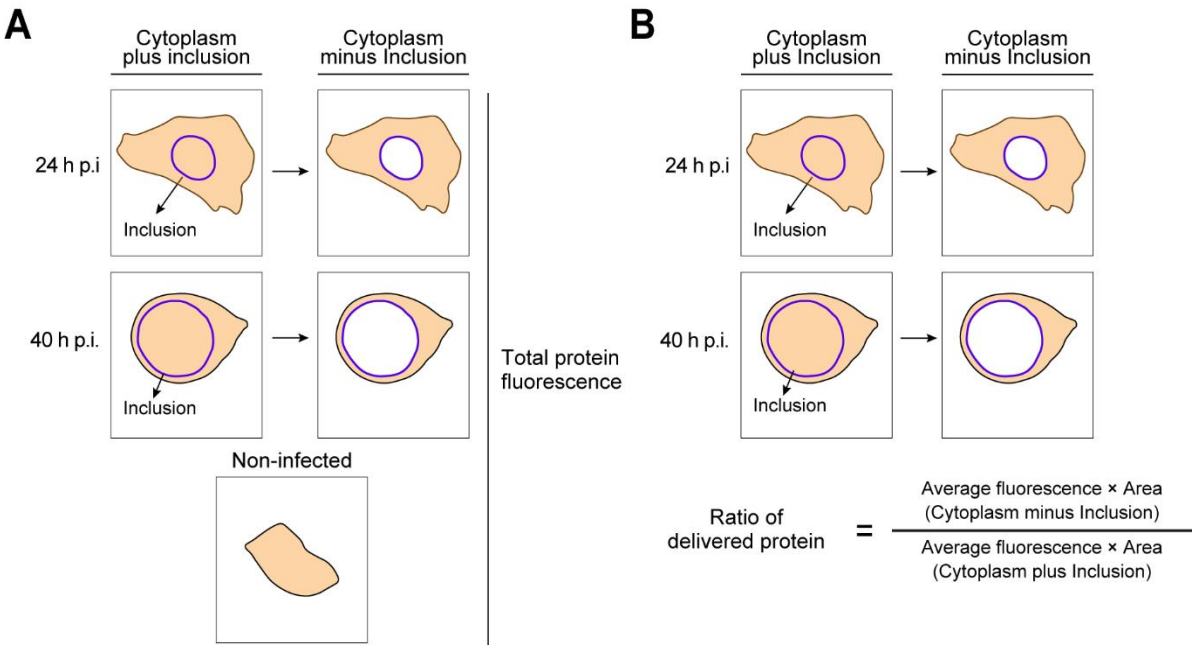


Figure 3.10. Methodology for the quantitative analysis of the production and T3S system-mediated delivery of CteG-2HA proteins during *C. trachomatis* infection. Images of HeLa cells infected by *C. trachomatis* were collected by immunofluorescence microscopy, and the distribution of each CteG-2HA protein (represented here in orange) in discrete cells was quantitatively analyzed using average fluorescence and area values determined with Fiji. **(A)** The total protein fluorescence in the cytoplasm of cells infected for 24 h or 40 h shown in Figure 3.11A corresponds to the average protein fluorescence at the cell cytoplasm but excluding the inclusion ("Cytoplasm minus Inclusion"). **(B)** The ratio of protein that was delivered into the cytoplasm of cells infected for 24 h or 40 h shown in Figure 3.11B was calculated by dividing the product between the average protein fluorescence at the cell cytoplasm but excluding the inclusion (values determined in panel A) and the area of this compartment ("Cytoplasm minus Inclusion"), by the product between the average protein fluorescence in the cytoplasm including the inclusion and the total cell area ("Cytoplasm plus Inclusion"). In (A), the same parameter was determined for non-infected neighboring cells ("Non-infected") of the infected ones to account for residual non-specific signal.

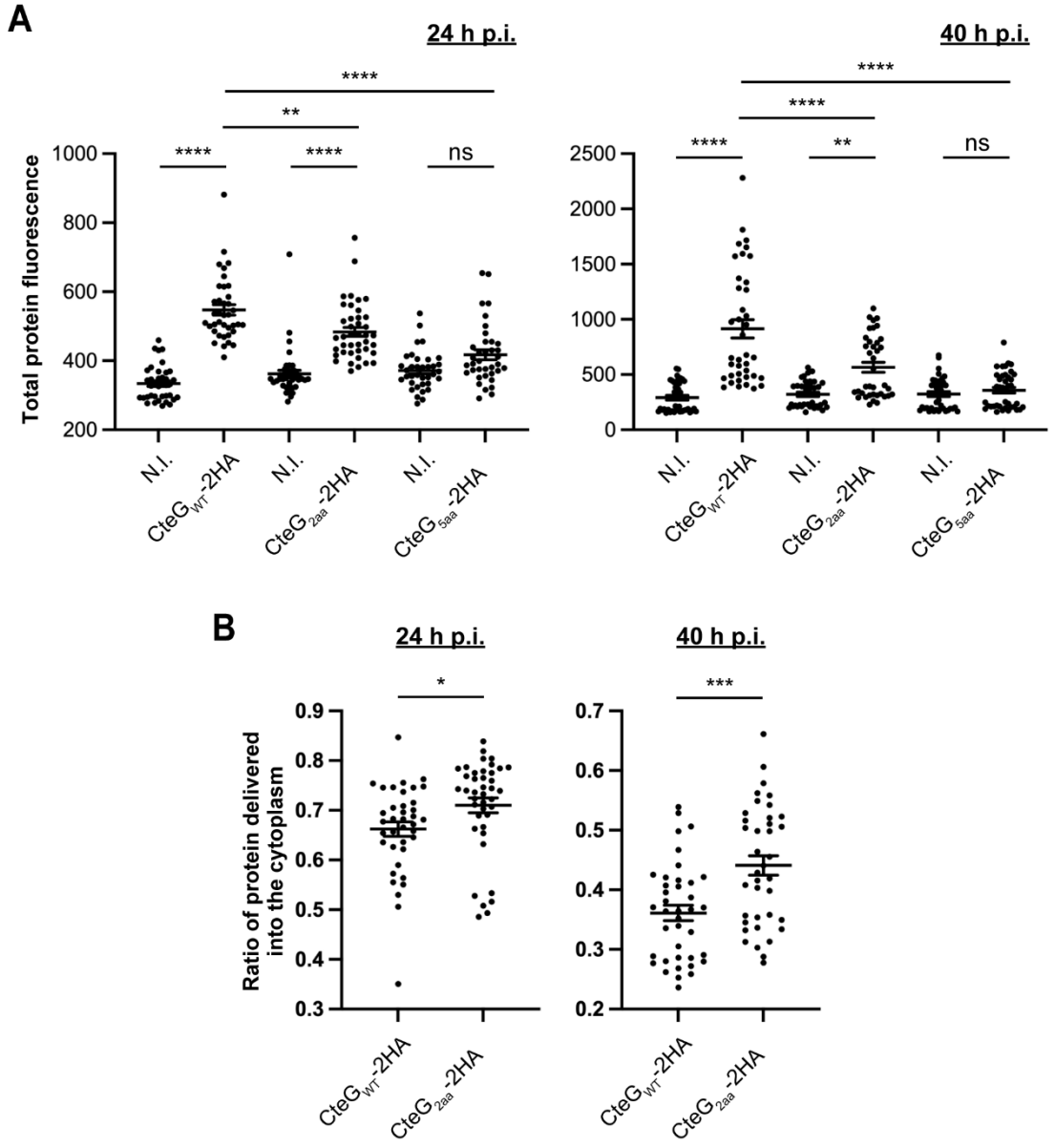


Figure 3.11. Quantitative analysis of the effect of specific amino acid replacements at the N-terminal region of CteG-2HA on its production and T3S system-mediated delivery during *C. trachomatis* infection. Cells were infected for 24 h or 40 h with *C. trachomatis* strains producing CteG$_{WT}$-2HA, CteG$_{2aa}$-2HA and CteG$_{5aa}$-2HA **(A)** or CteG$_{WT}$-2HA and CteG$_{2aa}$-2HA **(B)**, at a multiplicity of infection of 0.3. **(A)** Total protein fluorescence in the cytoplasm of cells infected for 24 h (left panel) or 40 h (right panel) by comparison with residual non-specific signal in non-infected (N.I.) neighboring cells. **(B)** Amount of protein delivered into the cytoplasm of cells infected for 24 h (left panel) or 40 h (right panel) relatively to the total protein fluorescence determined for the same cell. A more complete description of the quantification methodology can be found in Figure 3.10 above. Data are mean ± SEM of at least 36 individual cells from 2 independent experiments. Statistical analysis was performed with (A) one-way ANOVA + Tukey's post-test and (B) two-tailed unpaired Student's t-test between strains in each time point. ns, non-significant, ∗p<0.05; ∗∗p<0.01; ∗∗∗p<0.001; ∗∗∗∗p<0.0001.

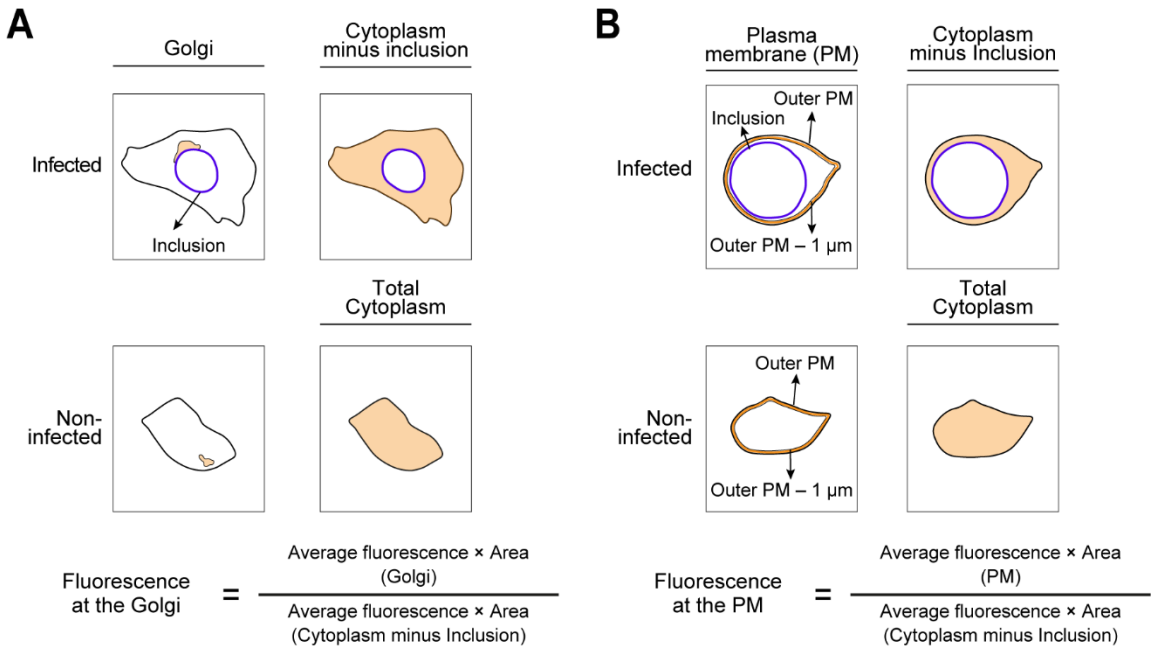


Figure 3.12. Methodology for the quantitative analysis of the subcellular localization of CteG-2HA proteins during *C. trachomatis* infection. Images of HeLa cells infected by *C. trachomatis* were collected as referred in Figure 3.10. **(A)** The protein fluorescence at the Golgi complex of cells infected for 24 h shown in Figure 3.9B was determined by dividing the product between the average protein fluorescence at the Golgi and its area (upper row, "Golgi"), by the product between the average protein fluorescence at the cell cytoplasm but excluding the inclusion and the area of this compartment (upper row, "Cytoplasm minus Inclusion"). **(B)** To calculate the protein fluorescence at the plasma membrane (PM) of cells infected for 40 h shown in Figure 3.9C, a line along the outer PM and another line parallel to that one but 1 μm shorter (outer PM - 1 μm; upper row, "Plasma membrane (PM)") were traced, and both the average protein fluorescence and the area between these two lines were calculated. The product between these two values was divided by the product between the average protein fluorescence at the cell cytoplasm but excluding the inclusion and the area of this compartment (upper row, "Cytoplasm minus Inclusion"). In (A) and (B), the same parameter was determined for non-infected neighboring cells ("Non-infected") of the infected ones to account for residual non-specific signal.

3.1.7 Conclusions

In summary, the experiments described in this section provided further insights about the determinants playing a role in the localization of CteG. We identified specific amino acid residues in a putative α-helix on the N-terminal region of CteG (C_9 and C_{17}) that are essential for targeting EGFP-$CteG_{1\text{-}100}$ to the Golgi in transfected cells. These same residues are also important for adequate localization of CteG at the Golgi and plasma membrane, after *C. trachomatis* infection of HeLa cells. However, they do not play an essential role, as could be thought from the studies done in transfected cells. This indicates that other determinants are involved in directing CteG to the Golgi in infected cells. Unfortunately, our analysis did not allow the identification of specific sequences on CteG that mediate the localization of this protein at the host cell plasma membrane. On the other hand, while analyzing the role of the N-terminal region of CteG in Golgi targeting, we found residues that may be important for an efficient delivery of CteG by the T3S machinery during *C. trachomatis* infection.

3.2 CteG mediates host cell lytic exit of *Chlamydia trachomatis*[5]

3.2.1 A *C. trachomatis cteG::aadA* insertional mutant has a defect in progeny generation

In a previous report, we observed that a *C. trachomatis cteG::aadA* mutant strain, generated with a modified group II intron containing a spectinomycin-resistance gene (*aadA*) (Figure 3.13A), produces smaller inclusions comparing to its parental L2/434 strain (Pais *et al.*, 2019). This could not be complemented by CteG with a C-terminal double hemagglutinin tag (CteG-2HA) encoded in a plasmid, with the gene expressed from the endogenous *cteG* promoter (Pais *et al.*, 2019). However, in this previous work, we did not observe significant differences in the generation of infectious progeny of the mutant and complemented strains relative to the L2/434 strain (Pais *et al.*, 2019). To clarify these issues, we started by reassessing the generation of infectious progeny, but also quantifying in each assay the number of internalized IFUs for each strain. This revealed a slight (~1.5-fold) but significant difference between the L2/434 parental strain and the mutant and complemented strains (Figure 3.13B). These differences were not caused by a lower ability of the mutant and complemented strains to invade cells, as they presented percentages of internalization comparable to the L2/434 parental strain (Figure 3.13C). To address the reason for lack of complementation (Figure 3.13B), a possible interference of the 2HA tag on the activity of CteG was excluded, as a *C. trachomatis cteG::aadA* strain encoding native CteG in a plasmid also produced smaller inclusions comparing to the parental strain (Figure 3.13A and Annexes Figure 9A). Furthermore, *C. trachomatis cteG::aadA* strains harboring plasmids encoding *cteG* and one (*ctl0359*/*fabI*) or both (*ctl0359*/*fabI* and *ctl0361*) of its flanking genes (Figure 3.13A) did not complement the defects in progeny generation (Figure 3.13D) or in inclusion area (Annexes Figure 9B). This excluded a polar effect in *ctl0359*/*fabI* or *ctl0361* (Figure 3.13A) arising from the insertion of the intron within *cteG*. Finally, WGS of the parental and mutant strains was performed, revealing four nucleotide changes in the *cteG::aadA* strain that led to missense mutations, for example in an Inc (CT618) and in a T3S gene (LcrH/CT862), as well as two nucleotide changes in non-coding regions (Annexes Table 4).

[5] The experiments described in this section were performed by Inês Serrano Pereira with the exception of preliminary analyses of the *cteG* mutant phenotypes performed by Sara Pais, and of the sequencing of the *cteG::aadA* mutant strain and corresponding parental strain, which was performed by Vítor Borges, Maria José Borrego and João Paulo Gomes at INSA Ricardo Jorge.

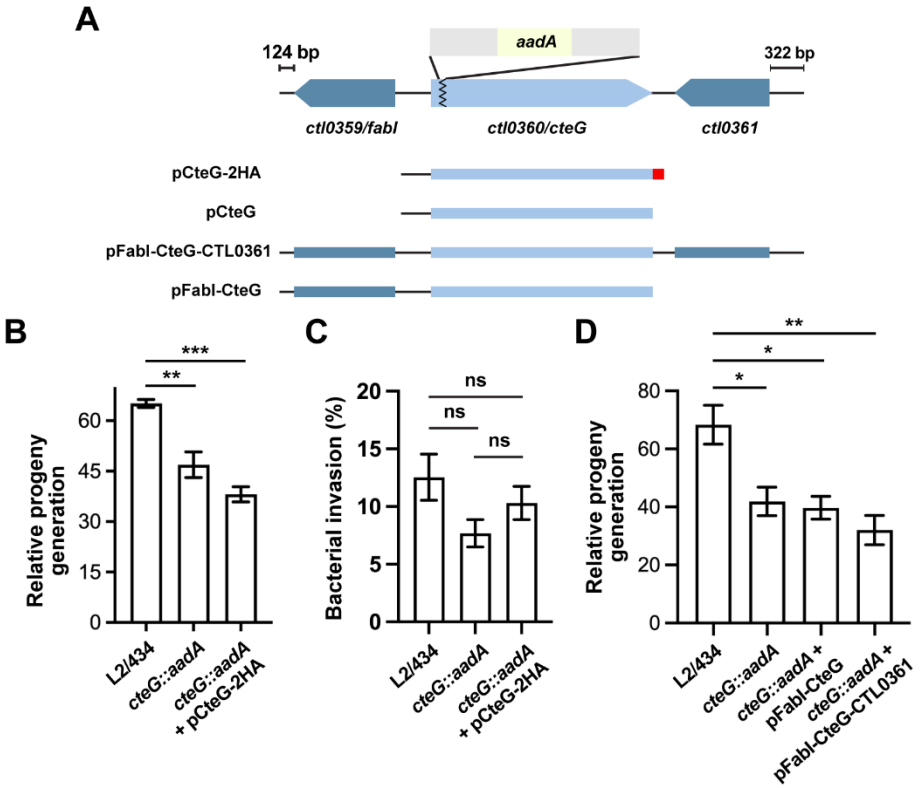


Figure 3.13. A *C. trachomatis cteG::aadA* insertional mutant is defective in progeny generation. (A) Illustration of the genomic region of *ctl0360/cteG* (light blue), which was disrupted by insertion of a modified group II intron (grey) carrying a spectinomycin-resistance gene, *aadA* (yellow) to generate a *C. trachomatis cteG* mutant strain (*cteG::aadA*) (Pais *et al.*, 2019). The *cteG::aadA* mutant strain was transformed with plasmids encoding CteG fused to a 2HA tag (red; pCteG-2HA; also named pCteG-2HA[Pgp4^{+}] in Annexes Table 3), native CteG (pCteG) (Annexes Figure 9), or CteG and two (*ctl0359/fabI* and *ctl0361*; pFabI-CteG-CTL0361) or one (*ctl0359/fabI*; pFabI-CteG) of its flanking genes (dark blue). **(B)** Two identical plates seeded with HeLa cells were infected with *C. trachomatis* L2/434, *cteG::aadA* mutant and complemented (*cteG::aadA* harboring pCteG-2HA, also named pCteG-2HA[Pgp4^{+}] in Annexes Table 3) strains at a MOI of 0.06. In one plate (input) cells were fixed at 24 h p.i. and in the second plate (output) cells were lysed at 40 h p.i. to infect a new plate seeded with HeLa cells for 24 h, after which cells were fixed. Chlamydiae were immunolabelled to quantify the number of released infectious particles by fluorescence microscopy. For each strain, the relative progeny generation was obtained by dividing the number of IFUs in the output by those in the input (see more details in Materials and Methods). **(C)** Images of individual cells infected with each *C. trachomatis* strain were collected to determine the percentage of bacterial internalization. Data correspond to mean ± SEM of at least 63 cells for each strain, from two independent experiments. **(D)** *cteG::aadA* mutant strains harboring pFabI-CteG or pFabI-CteG-CTL0361 (see Panel A) were assessed in terms of infectious progeny generation as in (B) by comparison with the L2/434 and *cteG::aadA* strains. Data in (B, D) correspond to the mean ± SEM (n=3). In (B, C, D) statistics was performed with ordinary one-way ANOVA and Dunnett's post-test analysis relative to the L2/434 strain (B, D) or Tukey's post-test analysis (C). *p<0.05; **p<0.01; ***p<0.001.

Therefore, the observed differences in inclusion size and progeny generation between the wild-type strain and the *cteG::aadA* mutant strain are not due to the disruption of *cteG*, or to an indirect effect of the disruption on neighboring genes, but possibly because of at least one of the nucleotide changes detected in the *cteG::aadA* mutant strain. However, within the scope of this work we did not study how the identified mutations may result in the observed defects.

3.2.2 The *cteG::aadA* mutant strain shows a CteG-dependent defect in egress from infected host cells

The localization of CteG at the host cell plasma membrane at later times of host cell infection led us to hypothesize that this effector could be involved in *C. trachomatis* egress. To analyze this, HeLa cells were infected with the *C. trachomatis* parental (L2/434), mutant (*cteG::aadA*), and complemented (*cteG::aadA* harboring a plasmid encoding CteG-2HA) strains for 48, 72 or 96 h at different multiplicities of infection (MOIs). The experiments were performed in the absence of gentamicin, to avoid possible killing of externalized chlamydiae. At each time point, tissue culture cell supernatants were collected (supernatant fraction). Adherent cells were subsequently lysed by osmotic shock, enabling recovery of chlamydiae that remained intracellular (lysate fraction). As shown in Figure 3.14A, with a MOI of 0.06, and at 48 and 72 h post-infection, significantly less IFUs were present in the supernatant fraction of cells infected with the *cteG* mutant strain comparing to the L2/434 strain. This phenotype was restored in the complemented strain (Figure 3.14A). Therefore, the presence of less *C. trachomatis* infectious particles in the culture supernatant of cells infected by the *cteG::aadA* mutant strain is CteG-dependent and is not related to the defect in progeny generation of the mutant as this latter defect is also displayed by the complemented strain (Figures 3.13B and 3.14B). The lower abundance of chlamydiae in the supernatant of cells infected by the *cteG::aadA* mutant strain is also not due to an invasion defect by this strain, which was as able to infect cells as the L2/434 and complemented strains (Figure 3.13C). Furthermore, the calculated ratios between the number of IFUs in the supernatant (Figure 3.14A) against the total IFUs (supernatant and lysates; Figures 3.14A, B) at 48 h post-infection were 2.6 ± 0.6% (mean±SEM) for the L2/434 strain and 6.0 ± 1.6% for the complemented strain, while only 0.6 ± 0.1% for the *cteG::aadA* mutant strain. This CteG-dependent phenotype was also observed at higher MOIs (0.3, 1.5, or 3; Figure 3.15A), and regardless of the presence or absence of gentamicin in the culture medium between 0 and 24 h post-infection (Annexes Figure 10A). At 96 h post-infection the phenotype was less obvious (Figures 3.14A and 3.15A), in part possibly because of re-infection events that interfere with the quantification of IFUs in the culture supernatants. In addition, at 72 and 96 h post-infection, at a higher MOI of 3, there were significantly more IFUs in the lysates of cells infected by the *cteG::aadA* mutant strain than in cells infected by the parental or complemented strains (Figure 3.15B), as a consequence of the progressive destruction of the cell monolayer in cells infected by chlamydiae producing CteG (see Figure 3.16 below). This gradual reduction in viable host cells might also interfere with measurements of IFUs in the supernatant of cells infected at higher MOIs and for longer times. Finally, analyzing the proteins in the supernatant and lysate fractions by immunoblotting with an anti-*C. trachomatis* Hsp60 antibody confirmed that the culture supernatant of cells infected by the *C. trachomatis*

cteG::aadA mutant contains less chlamydiae relative to cells infected by the L2/434 strain, and that this is CteG-dependent (Figure 3.14C). Overall, this indicated that CteG is involved in *C. trachomatis* egress from host cells, likely by contributing to host cell lysis.

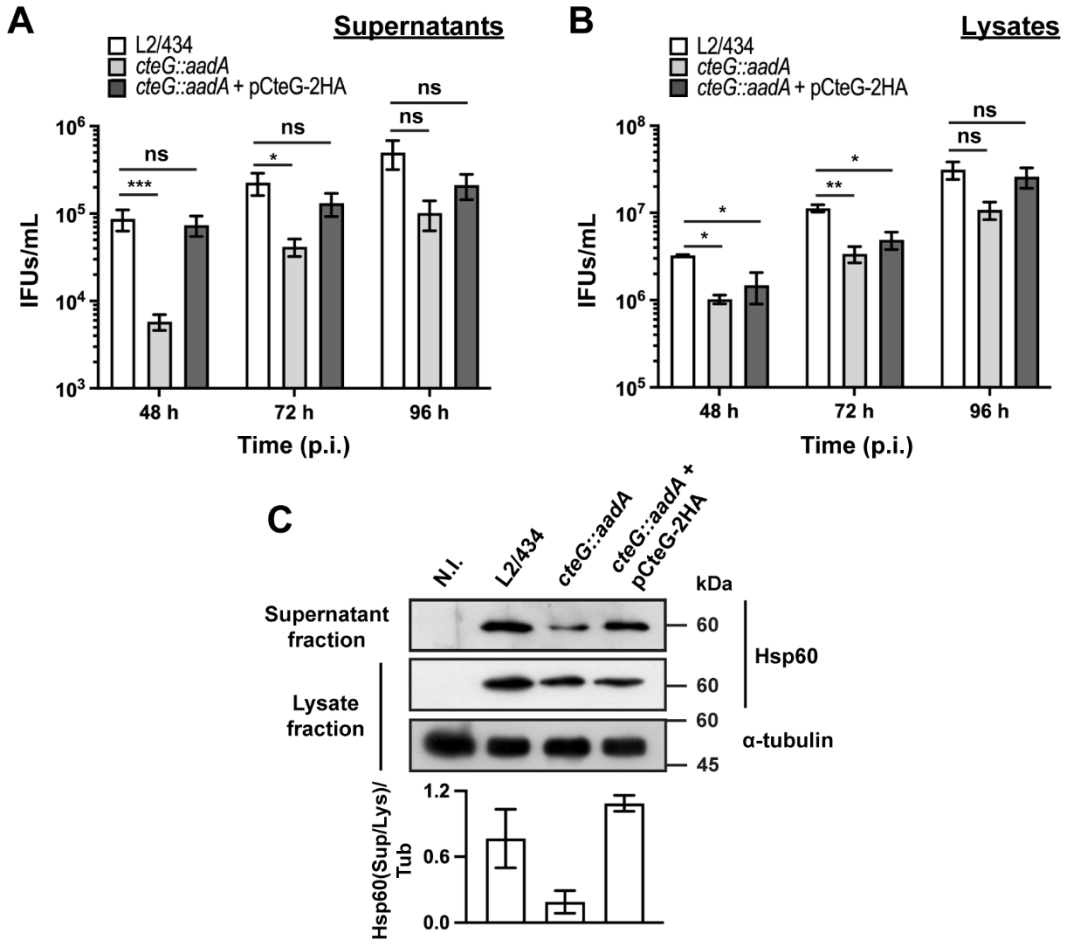


Figure 3.14. *C. trachomatis cteG::aadA* displays a CteG-dependent defect in egress from infected host cells. HeLa cells were infected with *C. trachomatis* parental (L2/434), mutant (*cteG::aadA*), and complemented (*cteG::aadA* harboring pCteG-2HA; also named pCteG-2HA[Pgp4^{+}] in Annexes Table 3) strains at an MOI of 0.06 for 48, 72 or 96 h. At each time post-infection, cell supernatants were collected (supernatant fraction) and adherent cells were lysed to recover intracellular *Chlamydia*e (lysate fraction). Fresh layers of HeLa cells were infected with serial dilutions of both supernatant **(A)** and lysate **(B)** fractions to quantify the number of recoverable inclusion-forming units (IFUs/mL). Data correspond to mean ± SEM (n≥3). For each time point, statistics was performed with ordinary one-way ANOVA and Dunnett's post-test analysis relative to the L2/434 parental strain (ns, non-significant; *p<0.5; **p<0.01, ***p<0.001), and the natural logarithm was applied to data to ensure normality of the populations. **(C)** HeLa cells were left non-infected (N.I.) or were infected for 48 h with *C. trachomatis* L2/434, *cteG::aadA* or *cteG::aadA* harboring pCteG-2HA (also named pCteG-2HA[Pgp4^{+}] in Annexes Table 3) at a MOI of 0.06. The proteins in the supernatant fraction (containing extracellular bacteria) were analyzed by immunoblotting with an antibody against *C. trachomatis* Hsp60 and the lysate fraction (intracellular bacteria) was analyzed by immunoblotting with antibodies against *C. trachomatis* Hsp60 and human α-tubulin (cell loading control), and using SuperSignal West Pico detection kit (Thermo Fisher Scientific) to detect proteins in the lysate fraction or SuperSignal West Femto detection kit (Thermo Fisher Scientific) to detect proteins in the supernatant fraction. Bands were quantified using the software Fiji, and the Hsp60 signal in the supernatant fraction (Sup) was normalized to that in the lysate fraction (Lys) and to α-tubulin signal (Tub). Bars correspond to mean ± SEM (n=3).

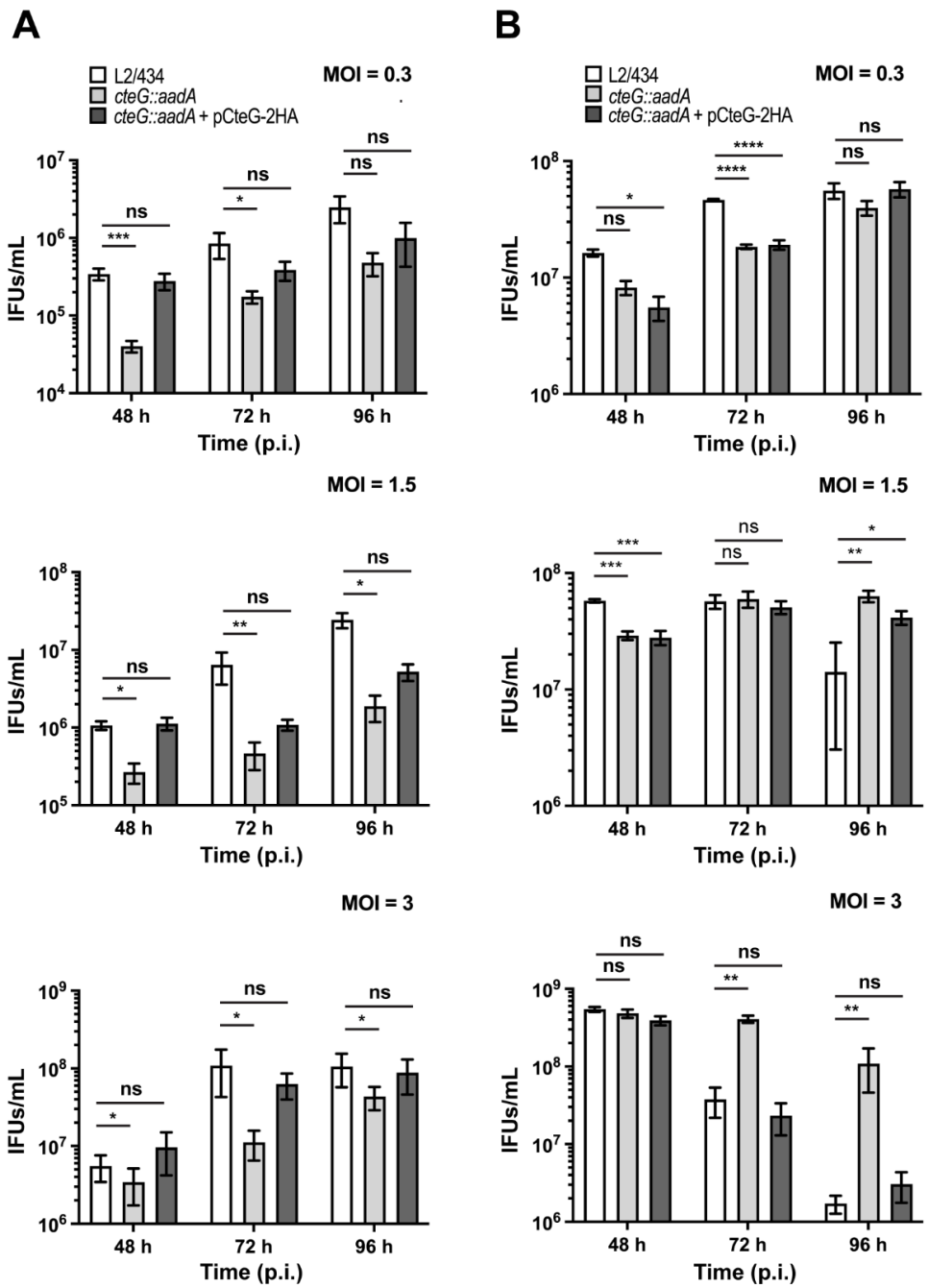


Figure 3.15. The *C. trachomatis cteG* mutant strain displays a CteG-dependent defect in egress from infected host cells. HeLa cells were infected with *C. trachomatis* parental (L2/434), mutant (*cteG::aadA*) and complemented (*cteG::aadA* harboring a plasmid encoding CteG-2HA; pCteG-2HA, also named pCteG-2HA/[Pgp4^{+}]) for 48, 72 or 96 h at the multiplicity of infection (MOI) of 0.3, 1.5 or 3. At each time post-infection, supernatant and lysate fractions were collected as in Figure 3.14 (see Materials and Methods). Fresh layers of HeLa cells were infected with serial dilutions of the supernatant (A) or lysate (B) fractions, and the number of recoverable inclusion-forming units (IFUs) was determined by immunofluorescence microscopy. Data correspond to mean ± SEM (n≥3). For each time point, statistical significance was determined by using ordinary one-way ANOVA and Dunnett's post-test analysis relative to the L2/434 parental strain. For statistical analysis, natural logarithm was applied to data to ensure normality of the populations. (ns, non-significant; *p<0.5; **p<0.01, ***p<0.001; ****p<0.0001).

3.2.3 The *cteG::aadA* mutant strain shows a CteG-dependent defect in host cells lysis

As mentioned above, preliminary phase-contrast microscopy observations indicated that the significantly higher amounts of recoverable IFUs in lysates of HeLa cells infected with a MOI of 3 for 72 or 96 h by the *cteG::aadA* mutant relative to the parental L2/434 and complemented strains (Figure 3.15B) were a direct consequence of a much more pronounced destruction of the HeLa cell monolayer by the parental and complemented strains. To visualize this directly, we analyzed cells infected (MOI of 0.3) for 48, 72 or 96 h with the parental, mutant, and complemented strains by immunofluorescence microscopy. Infected cells were fixed and stained with an anti-chlamydial Hsp60 antibody (to visualize chlamydial inclusions), fluorophore-conjugated phalloidin (to visualize the host actin cytoskeleton) and with DAPI (to visualize the host cell nuclei). The representative images on Figure 3.16 illustrate that the monolayer of HeLa cells infected with the mutant strain remained relatively intact, even at 96 h post-infection, whereas the monolayer of cells infected with either the parental or the complemented strain was visibly destroyed from 72 h post-infection.

To perform a direct measurement of host cell lysis, we monitored the release of LDH into the supernatant of HeLa cells infected at different MOIs, for 48, 72, and 96 h, by the parental, mutant, and complemented strains. As illustrated in Figure 3.17, at a MOI of 0.3, the *cteG::aadA* mutant strain showed a lower ability to cause host cell lysis by comparison with the parental and complemented strains. Similar observations were made at lower (0.06) and higher MOIs (1.5 or 3) (Figure 3.17), and regardless of the presence of gentamicin in the cell culture media from 0 to 24 h post-infection (Annexes Figure 10B). While part of the chlamydiae that we detected in the culture supernatant of infected cells (Figure 3.14) could be released by extrusion, altogether, these data indicated that, from ~48 h post-infection of HeLa cells, CteG promotes chlamydial egress by contributing to host cell lysis by *C. trachomatis*.

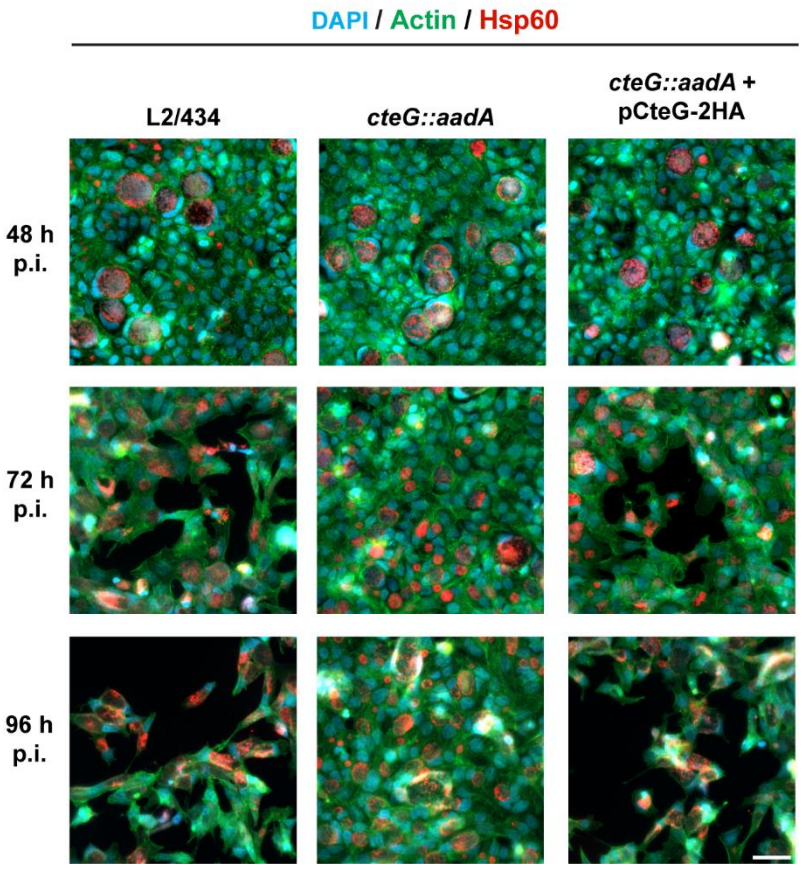


Figure 3.16. *C. trachomatis cteG::aadA* displays a CteG-dependent defect in host cell lysis. HeLa 229 cells were infected with *C. trachomatis* parental (L2/434), mutant (*cteG::aadA*), and complemented (*cteG::aadA* harboring a plasmid encoding CteG-2HA; also named pCteG-2HA[Pgp4^{+}], Annexes Table 3) strains at a MOI of 0.3. **(A)** At 48, 72 or 96 h post-infection (p.i.), cells were fixed with methanol, immunolabelled with antibodies against *C. trachomatis* Hsp60 (red) and appropriate fluorophore-conjugated secondary antibodies, and stained with DAPI (host cell nuclei and chlamydial inclusions; blue) and with fluorophore-conjugated phalloidin (host actin cytoskeleton; green). Scale bar, 40 μm.

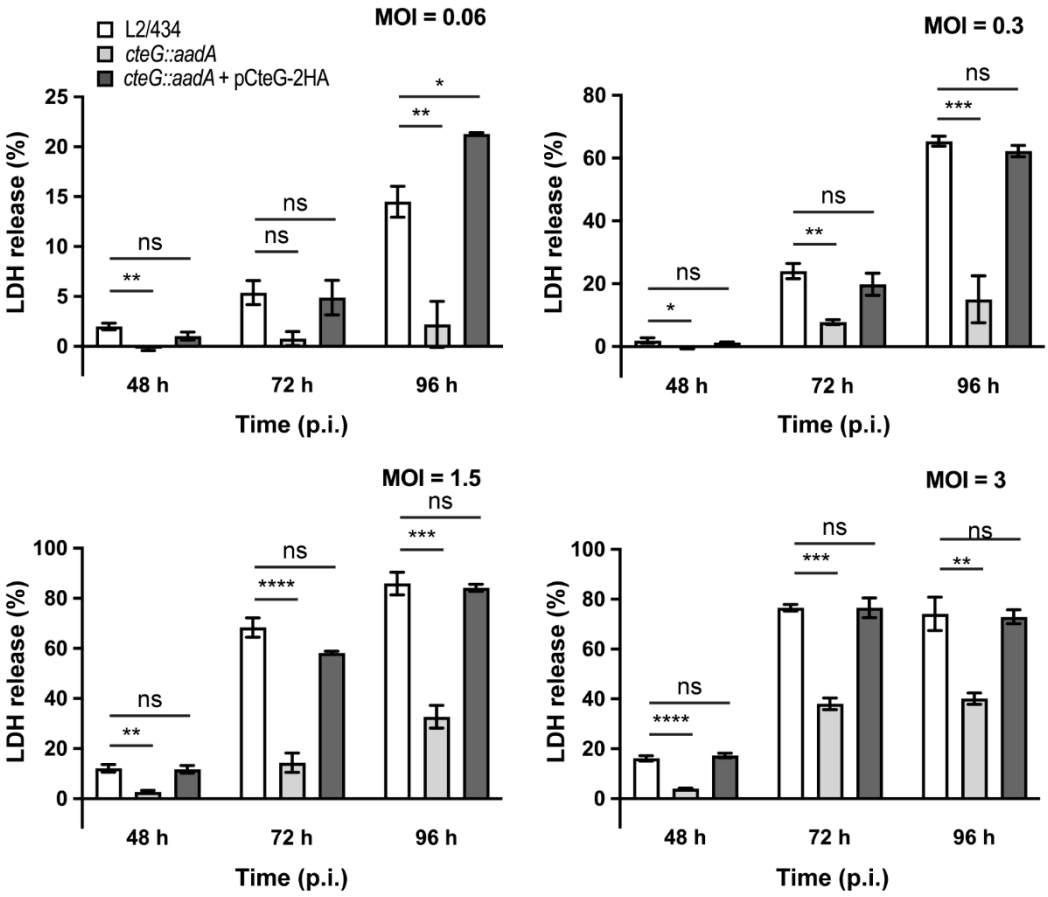


Figure 3.17. The *C. trachomatis cteG* mutant strain displays a CteG-dependent host cell lysis defect. HeLa cells were infected with *C. trachomatis* strains L2/434, *cteG::aadA* and *cteG::aadA* harboring a plasmid encoding CteG-2HA (pCteG-2HA, also named pCteG-2HA/[Pgp4^{+}]), using a multiplicity of infection (MOI) of 0.06, 0.3, 1.5 or 3, for 48, 72 or 96 h. The release of host LDH into the supernatant of infected cells was quantified using a CytoScan™ LDH Cytotoxicity Assay kit (G-Biosciences). Data are representative of five independent experiments and correspond to the mean ± SEM of three biological replicates. For each time post-infection, statistical significance was determined by using ordinary one-way ANOVA and Dunnett's post-test analysis relative to the L2/434 parental strain (ns, non-significant; *p<0.05; **p<0.01; ***p<0.001; ****p<0.0001).

3.2.4 Production and localization of CteG are not regulated by *C. trachomatis* virulence plasmid encoded Pgp4

We confirmed previous observations that the virulence plasmid contributes to *C. trachomatis* lytic exit (Yang *et al.*, 2015) by infecting HeLa cells with the L2/434 strain side by side with a plasmidless *C. trachomatis* strain (25667R) and monitoring LDH release at 48, 72, and 96 h post-infection (Figure 3.18). The plasmid-dependent role on host cell lysis has been shown to be due to plasmid encoded Pgp4 (Yang *et al.*, 2015), which mediates transcriptional regulation of several plasmid and chromosomal genes (Carlson *et al.*, 2008; Song *et al.*, 2013). However, expression of *cteG* is not regulated by Pgp4 (Song *et al.*, 2013; Patton *et al.*, 2018).

To study how CteG and Pgp4 contribute to chlamydial lytic exit, we generated several *C. trachomatis* strains harboring recombinant plasmids carrying ($Pgp4^+$) or lacking ($Pgp4^-$) the *pgp4* gene (Annexes Table 3). It has been previously shown that during chlamydial transformation the *C. trachomatis* native plasmid is eventually lost by exchange with the novel plasmid (Wang *et al.*, 2011; Mueller *et al.*, 2016). To ensure that the newly generated $Pgp4^+$ or $Pgp4^-$ *C. trachomatis* strains lost the native plasmid, we verified both the presence of the desired recombinant plasmid (Figure 3.19) and the loss of the native plasmid (Figure 3.20).

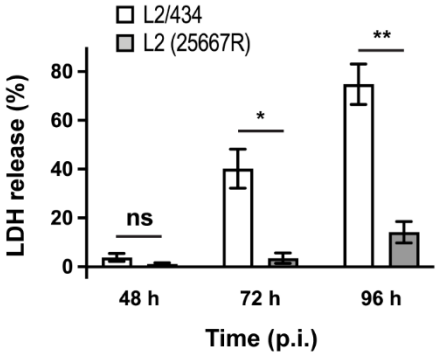


Figure 3.18. The *C. trachomatis* virulence plasmid contributes to lytic exit from host cells. HeLa cells were infected with *C. trachomatis* L2/434 and with plasmid-deficient 25667R strain for 48, 72 or 96 h at a MOI of 0.3. At each time post-infection (p.i.), the release of host lactate dehydrogenase (LDH) into the supernatant of infected HeLa cells was measured using a CytoScan™ LDH Cytotoxicity Assay kit (G-Biosciences). Data correspond to the mean ± SEM of three independent experiments. Statistical significance was determined for each time point by using a two-tailed unpaired Student's t-test (ns, non-significant; *$p<0.05$; **$p<0.01$).

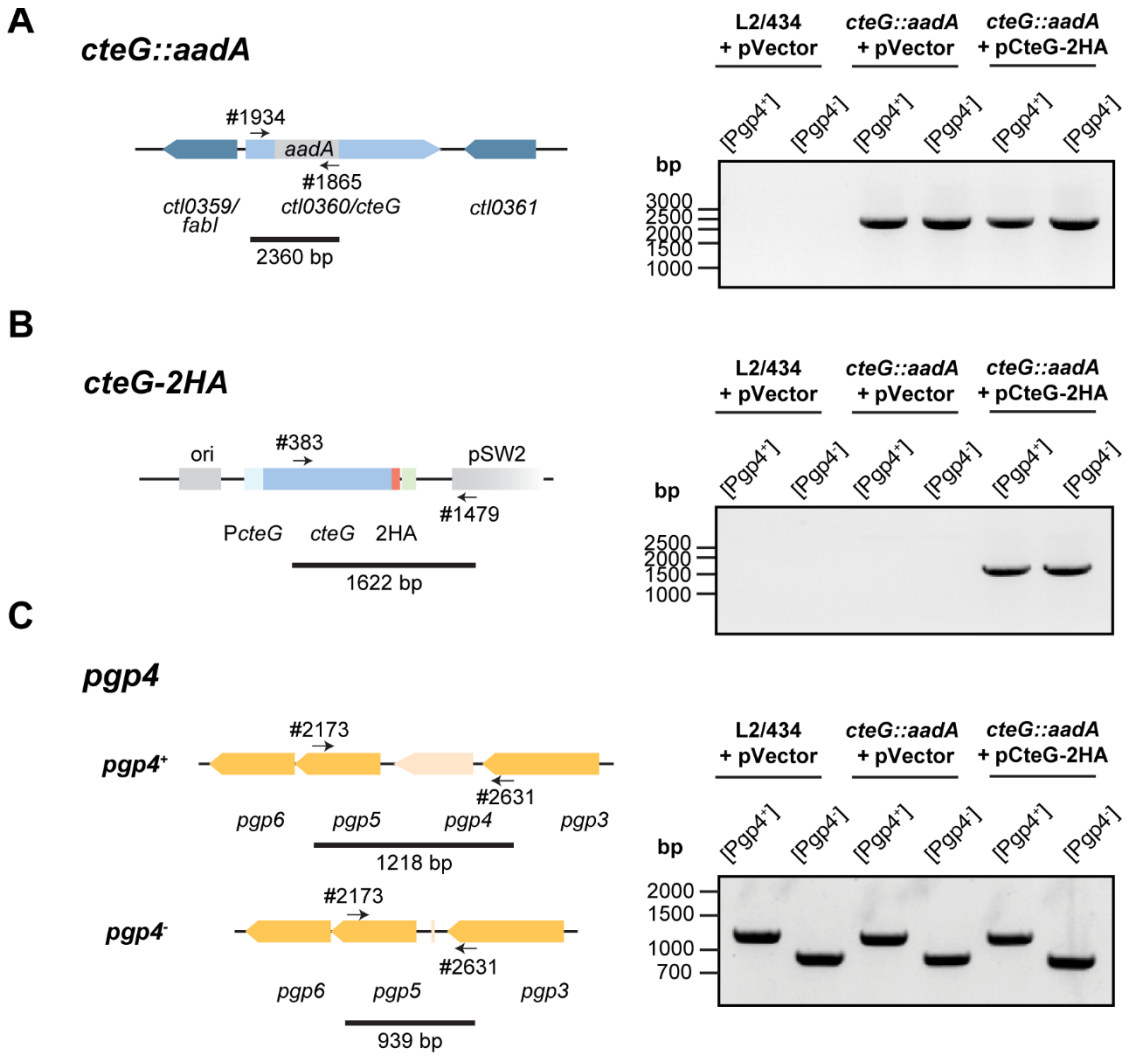


Figure 3.19. Verification of the accuracy of the *C. trachomatis* strains generated in this work. The indicated *C. trachomatis* strains (see plasmid and strain nomenclature in Annexes Table 3) were verified by PCR **(A)** for the presence or absence of a group II intron interrupting *cteG* (*cteG::aadA*), **(B)** for the presence of the plasmid carrying *cteG-2HA*, or **(C)** for the presence or absence of intact plasmid encoded *pgp4* (*pgp4*). Plasmid pSVP264 is the complementing plasmid used throughout this work (also named pCteG-2HA or pCteG-2HA[Pgp4$^+$]; Annexes Tables 1, 3). Illustrations of all features are depicted. The arrows and numbers indicate the approximate hybridization site of DNA oligonucleotides (Annexes Table 2) used in PCR reactions, yielding amplification products of the indicated lengths in base pairs (bp).

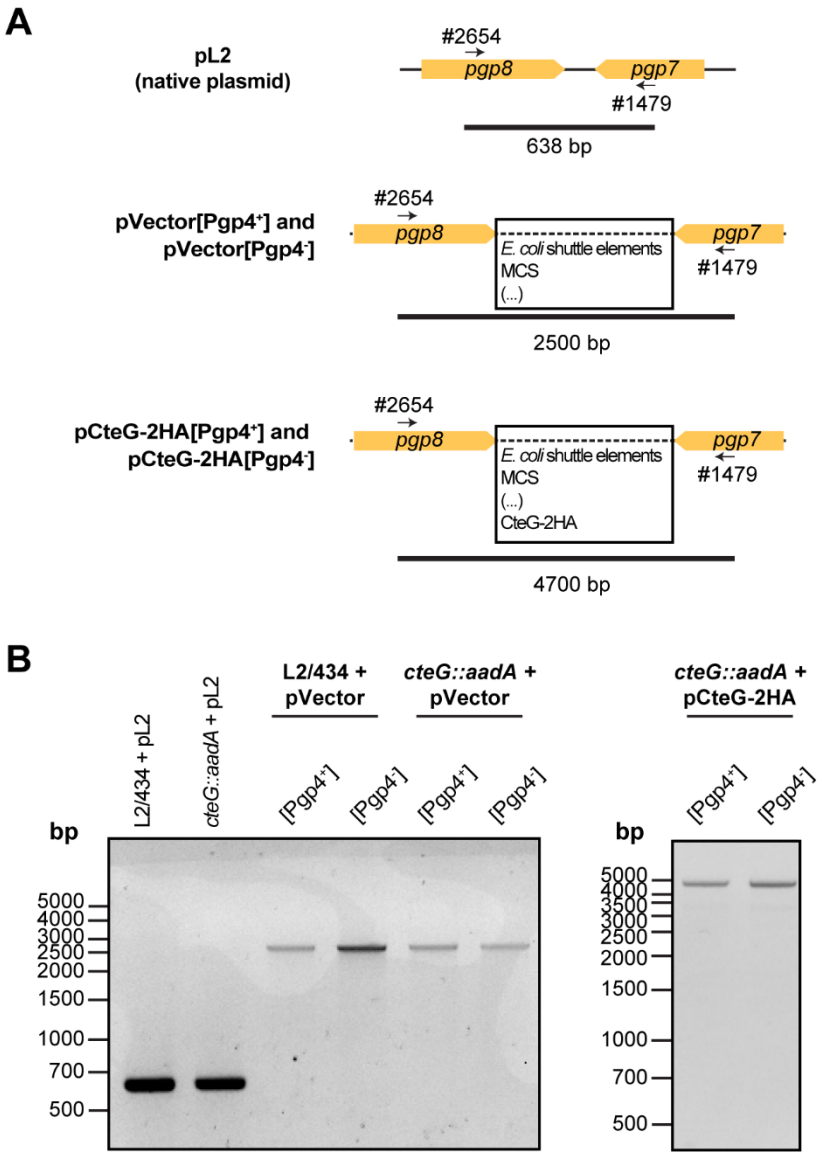


Figure 3.20. Verification of the replacement of the *C. trachomatis* native plasmid by recombinant plasmids with or without *pgp4*. (A) DNA organization of the indicated plasmids. The arrows and numbers indicate the approximate hybridization site of DNA oligonucleotides (Annexes Table 2) used in PCR reactions, yielding amplification products of the indicated lengths in base pairs (bp). **(B)** Presence of the *C. trachomatis* native plasmid (pL2) in L2/434 and *cteG::aadA* strains and its loss and replacement in strains carrying instead plasmids pVector[Pgp4$^+$], pVector[Pgp4$^-$], pCteG-2HA[Pgp4$^+$] or pCteG-2HA[Pgp4$^-$] (Annexes Table 1) was confirmed by PCR.

First, to analyze if Pgp4 influences the production or the subcellular localization of CteG, HeLa cells were infected for 16, 24, 30 and 40 h with *C. trachomatis cteG::aadA* strains encoding CteG-2HA on a plasmid, but in one case with a plasmid carrying Pgp4 [pCteG-2HA[Pgp4^{+}]; (Pais *et al.*, 2019)] and in the other not (pCteG-2HA[Pgp4^{-}]; Annexes Table 3), followed by immunoblotting of whole cell extracts. In these two plasmids the expression of the hybrid *cteG-2HA* gene is under the control of the *cteG* promoter, mimicking endogenous regulation. Quantitative analysis of immunoblots revealed that Pgp4 does not regulate the timing or amount of CteG-2HA production (Figure 3.21A). Previously, we reported the appearance of faster migrating species of CteG-2HA detected by immunoblotting with the anti-HA antibody of extracts of HeLa cells infected with *C. trachomatis* producing CteG-2HA for more than 20-24 h (Pais *et al.*, 2019). These faster migrating species are indicative of CteG degradation or processing occurring within the chlamydiae (Pais *et al.*, 2019). It is currently unknown if they have functional relevance or are a consequence of plasmid-mediated overexpression of CteG-2HA, but Pgp4 does not influence their appearance (Figure 3.21A). We then performed immunofluorescence microscopy of cells infected by these two strains, followed by quantitative analysis of the localization of CteG-2HA and showed that CteG-2HA localizes at the Golgi (at 24 h post-infection) or at the host cell plasma membrane (at 40 h post-infection) during infection of host cells by *C. trachomatis* regardless of the presence or absence of Pgp4 (Figure 3.21B and Annexes Figure 11A). We found only minor and no significant differences in Golgi distribution around the inclusion when analyzing cells infected by the *cteG::aadA*(pCteG-2HA[Pgp4^{+}]) or *cteG::aadA*(pCteG-2HA[Pgp4^{-}]) strains by immunofluorescence microscopy (Annexes Figures 11B, C). Therefore, and in summary, Pgp4 does not control the production or localization of CteG.

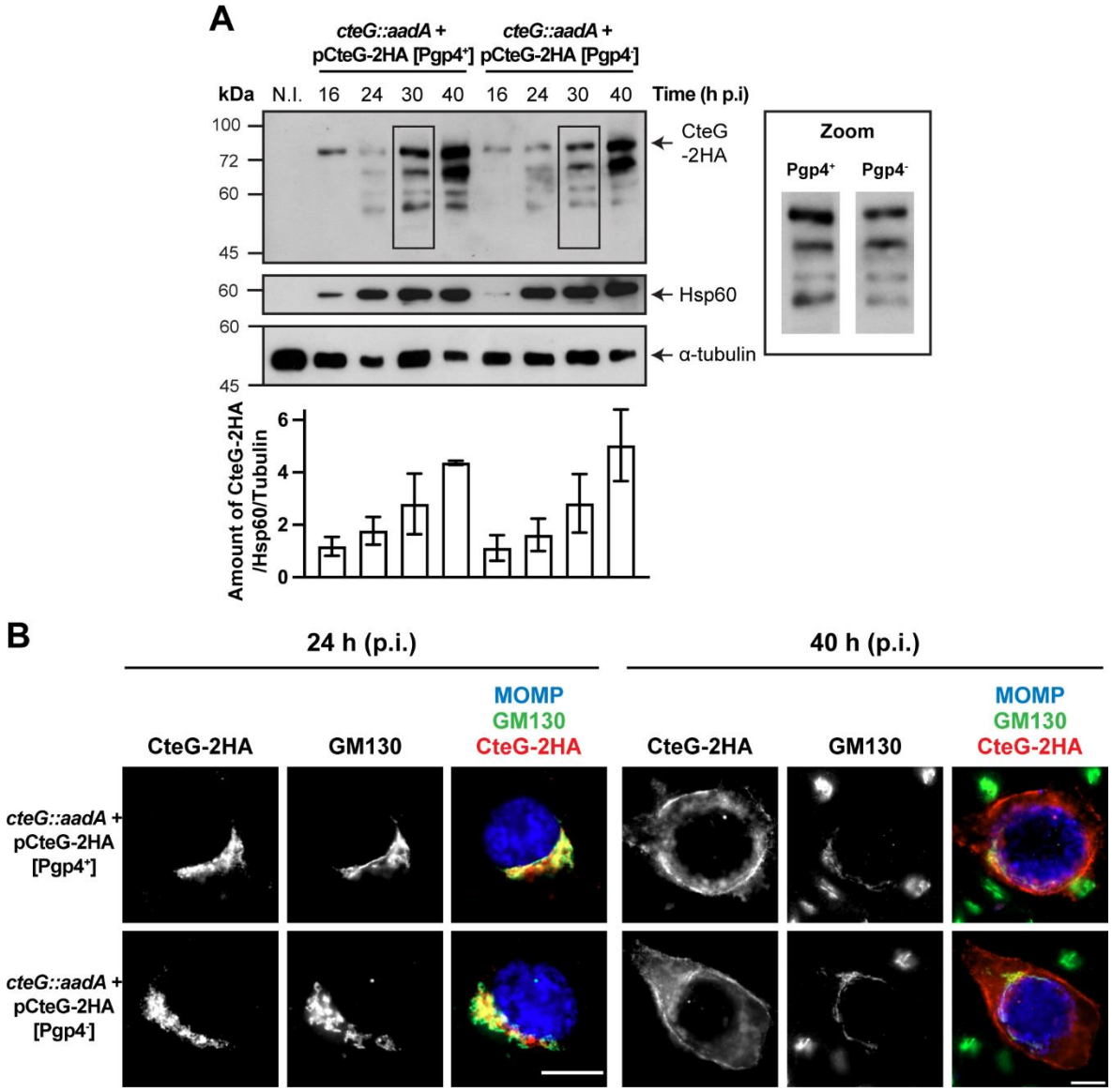


Figure 3.21. Pgp4 does not modulate the production or the localization of CteG during *C. trachomatis* infection. HeLa cells were either left non-infected (N.I.) or infected with *C. trachomatis cteG::aadA* strains carrying pCteG-2HA[Pgp4$^+$] (also named pCteG-2HA in Annexes Table 3; Pgp4$^+$/CteG-2HA$^+$) or pCteG-2HA[Pgp4$^-$] (Pgp4$^-$/CteG-2HA$^+$). **(A)** At 16, 24, 30 or 40 h p.i., whole cell extracts were prepared and then analyzed by immunoblotting with antibodies against HA (CteG-2HA), *C. trachomatis* Hsp60 (bacterial loading control) and human α-tubulin (HeLa cell loading control), and using SuperSignal West Pico detection kit (Thermo Fisher Scientific) to detect Hsp60 or α-tubulin, or SuperSignal West Femto detection kit (Thermo Fisher Scientific) to detect CteG-2HA. The band corresponding to full-length CteG-2HA is indicated with an arrow. Zooms of the band pattern of CteG-2HA species at 30 h p.i. in both Pgp4$^+$ and Pgp4$^-$ backgrounds are displayed. The intensity of all bands on each lane was quantified using the software Fiji and summed to obtain the intensity of all CteG-2HA species at a given time point. Each value was normalized to the bacterial and HeLa cell loading controls. Bars correspond to mean ± SEM (n=3). **(B)** Cells were fixed with PFA 4% (w/v) at 24 or 40 h p.i., immunolabelled with antibodies against *C. trachomatis* MOMP (blue), *cis*-Golgi network (GM130; green) and HA (CteG-2HA; red), and appropriate fluorophore-conjugated secondary antibodies, and analyzed by fluorescence microscopy. Scale bar, 10 μm.

To try to assess if the presence or absence of Pgp4 influenced the mRNA levels of *cteG*, we performed RT-qPCR of *cteG* transcripts in L2/434-derived strains carrying pVector[Pgp4$^+$] (CteG$^+$/Pgp4$^+$) or pVector[Pgp4$^-$] (CteG$^+$/Pgp4$^-$) (Annexes Table 3) by comparison with the L2/434 parental strain. Because we previously found that *cteG* mRNA levels peak at 2 h post-infection and remain low from 8 h post-infection (Pais *et al.*, 2019), we analyzed these two time points. *cteG* mRNA levels for the L2/434 parental strain were ~34.2 ± 7.1 (mean±SEM) at 2 h post-infection and 2.4 ± 0.9 at 8 h post-infection (Annexes Figure 12), which was consistent with our previous results (Pais *et al.*, 2019). However, we did not find a significant difference of *cteG* mRNA levels between the L2/434 parental strain and L2/434 carrying pVector[Pgp4$^+$] or L2/434 carrying pVector[Pgp4$^-$] strains at 2 or 8 h post-infection (Annexes Figure 12), suggesting that, at least at these time-points, Pgp4 does not regulate the expression of *cteG* at the transcriptional level.

3.2.5 A *cteG::aadA pgp4* double mutant strain displays a defect in inducing host cell lysis identical to *cteG::aadA* or *pgp4* single mutants

If Pgp4 does not control the production or localization of CteG, then CteG and Pgp4 may function independently to promote host cells lysis. If this was the case, then a *C. trachomatis* strain lacking both CteG and Pgp4 would be more defective in host cells lysis than strains lacking only CteG or Pgp4. Alternatively, CteG and Pgp4 may act on the same pathway to promote host cell lysis. In this scenario a Pgp4-regulated gene could influence CteG activity, and the double mutant would be indistinguishable from the single mutants in its ability to promote host cell lysis. To analyze this, we used *C. trachomatis* L2/434 or *cteG::aadA*-derived strains carrying Pgp4 but not CteG (in pVector[Pgp4$^+$], a recombinant derivative of the endogenous virulence plasmid; Annexes Table 3 and Figures 3.19 and 3.20), or neither Pgp4 nor CteG (in pVector[Pgp4$^-$], a derivative of pVector[Pgp4$^+$] with *pgp4* deleted; Annexes Table 3 and Figures 3.19 and 3.20). When analyzing these strains for their ability to generate infectious particles, the ones lacking CteG and/or Pgp4 revealed a defect relative to the strain carrying chromosomally encoded *cteG* and plasmid encoded *pgp4* (Figure 3.22A). This defect was more pronounced for the strain lacking both CteG and Pgp4 (Figure 3.22A). Initial experiments also indicated that, for unknown reasons, the levels of host cell lysis mediated by *C. trachomatis* strains (L2/434 and *cteG::aadA*) carrying pVector[Pgp4$^+$] were higher than in similar strains carrying the endogenous virulence plasmid (Figure 3.22B). As a consequence, the difference between the IFUs released by cells infected by the L2/434 and *cteG::aadA* strains carrying pVector[Pgp4$^+$] was less pronounced than in cells infected by similar strains carrying the native plasmid (Figure 3.22C). Although this difference in released IFUs is still detectable

and can be complemented (Figures 3.22C, D), in subsequent experiments analyzing strains carrying pVector[Pgp4$^+$] or pVector[Pgp4$^+$]-derived plasmids we focused on the more robust monitoring of LDH release into the supernatant of infected cells as a measure of the ability of CteG to mediate chlamydial egress by host cell lysis. HeLa cells were then infected at a MOI of 0.3, for 72 h, by *C. trachomatis* L2/434 carrying pVector[Pgp4$^+$] (CteG$^+$/Pgp4$^+$), L2/434 carrying pVector[Pgp4$^-$] (CteG$^+$/Pgp4$^-$), *cteG::aadA* carrying pVector[Pgp4$^+$] (CteG$^-$/Pgp4$^+$), or *cteG::aadA* carrying pVector[Pgp4$^-$] (CteG$^-$/Pgp4$^-$). This further confirmed that both CteG and Pgp4 contribute to *C. trachomatis*-mediated host cell lysis (Figure 3.23A). However, the CteG$^-$/Pgp4$^-$ strain showed a defect in host cell lysis similar to the CteG$^+$/Pgp4$^-$ or CteG$^-$/Pgp4$^+$ strains (Figure 3.23A). The same trend was observed at earlier (57 h) or later (81 h) times post-infection, although the CteG$^-$/Pgp4$^-$ double mutant strain displayed a slight defect in inducing host cell lysis relatively to the CteG$^-$/Pgp4$^+$ strain at 81 h post-infection (Annexes Figure 13). This could be related with the defect of the CteG$^-$/Pgp4$^-$ strain in producing infectious progeny when comparing to the CteG$^-$/Pgp4$^+$ strain (Figure 3.22A), which results in a cumulative defect of the CteG$^-$/Pgp4$^-$ strain in inducing host cell lysis at later times post-infection. Consistent with these results, the number of infectious particles at the supernatant of cells infected with the CteG$^-$/Pgp4$^-$ strain was equivalent to that of cells infected with the CteG$^+$/Pgp4$^-$ or CteG$^-$/Pgp4$^+$ single mutant strains (Annexes Figure 14). These results indicate that CteG and Pgp4 act on the same pathway to promote host cell lysis mediated by *C. trachomatis*.

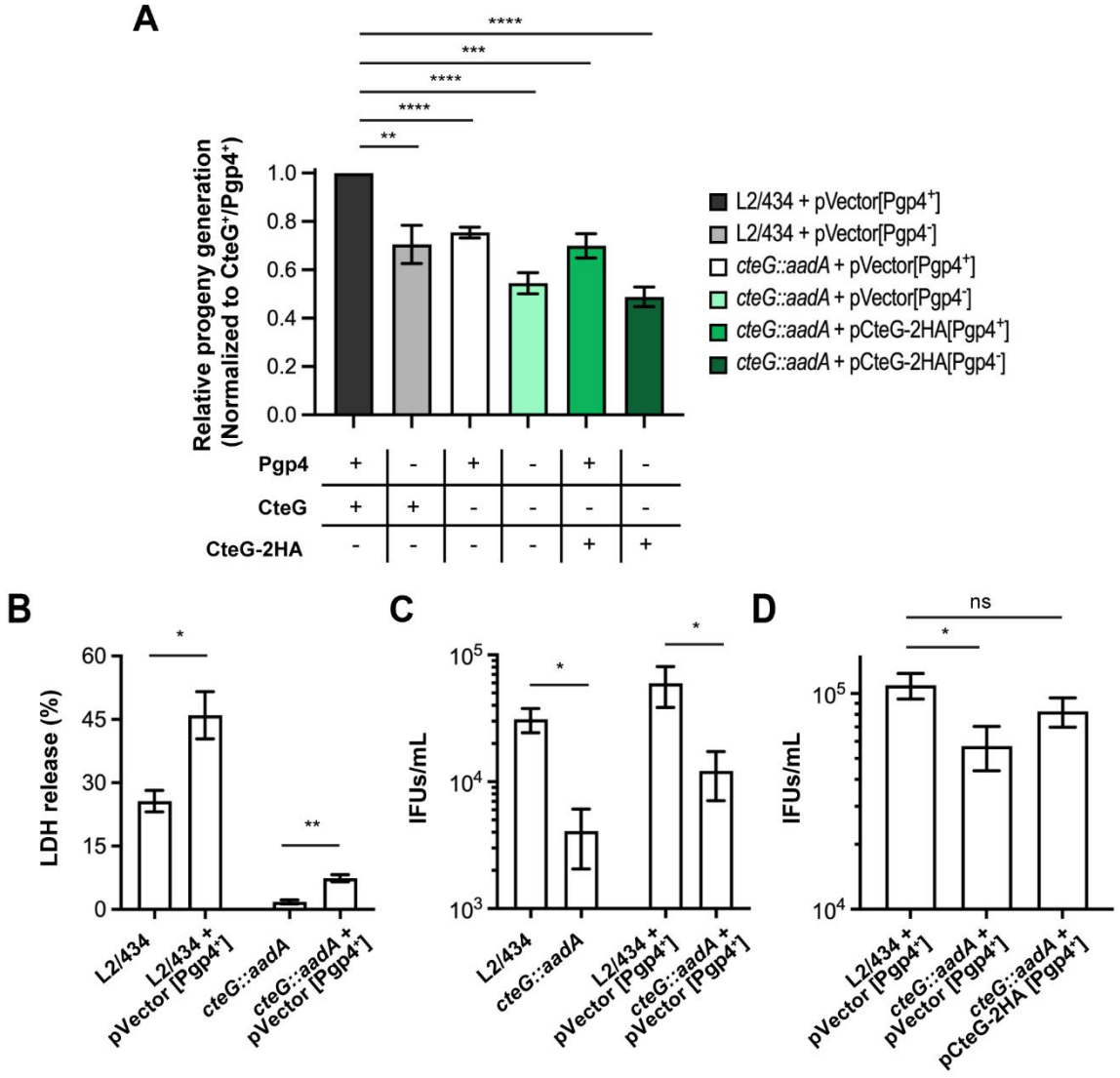


Figure 3.22. Characterization of *C. trachomatis* strains used and generated in this study. (A) The ability of *C. trachomatis* strains used and constructed in this study (see Annexes Table 3) to generate infectious progeny was tested, as described in Materials and Methods. In each assay, the values were normalized to that of the CteG$^+$/Pgp4$^+$ strain. Data correspond to the mean ± SEM (n=4). **(B, C, and D)** HeLa cells were infected with the indicated *C. trachomatis* strains with a multiplicity of infection (MOI) of 0.3 for 72 h (B), or a MOI of 0.06 for 48 h (C and D). **(B)** The release of host lactate dehydrogenase (LDH) into the supernatant of infected cells was quantified using a CytoScan™ LDH Cytotoxicity Assay kit (G-Biosciences). Data correspond to mean ± SEM of the mean (n=3). **(C and D)** The number of inclusion forming units (IFUs) in the supernatant fraction was quantified as described in Materials and Methods. Data correspond to mean ± SEM of three (C) or seven (D) independent assays. Statistical significance was determined using two-tailed unpaired Student's t-test (A, B, and C) or by using ordinary one-way ANOVA and Dunnett's post-test analysis relative to the CteG$^+$/Pgp4$^+$ strain. In (C) and (D), data was transformed by applying the natural logarithm to ensure normality of the populations. (ns, non-significant; *p<0.05; **p<0.01; ***p<0.001; ****p<0.0001).

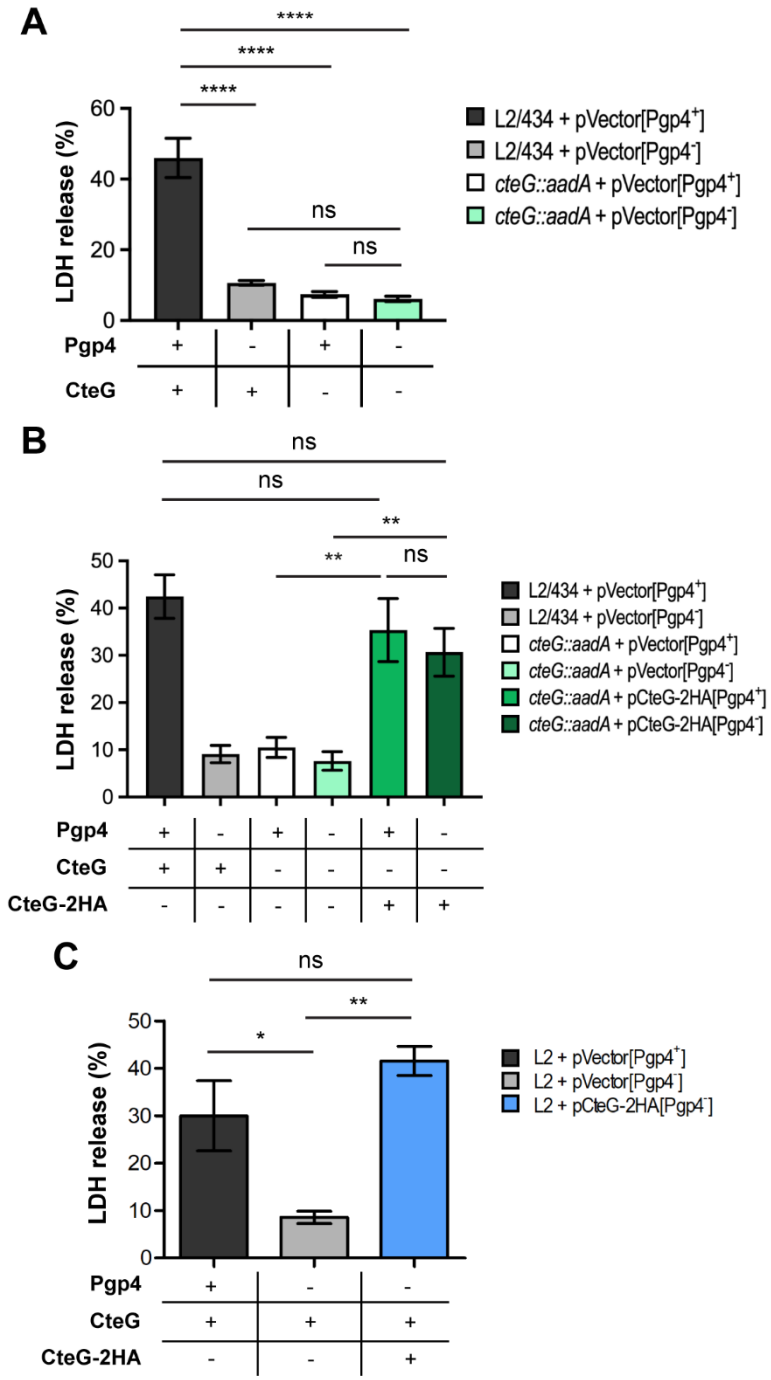


Figure 3.23. ***C. trachomatis*** **mediates host cell lysis via a common pathway involving both CteG and Pgp4.** **(A)** HeLa cells were infected with *C. trachomatis* L2/434 carrying pVector[Pgp4$^+$] (CteG$^+$/Pgp4$^+$) or pVector[Pgp4$^-$] (CteG$^+$/Pgp4$^-$), and *cteG::aadA* carrying pVector[Pgp4$^+$] (CteG$^-$/Pgp4$^+$) or pVector[Pgp4$^-$] (CteG$^-$/Pgp4$^-$) for 72 h at an MOI of 0.3, and the LDH released by lysed host cells was measured using a CytoScan™ LDH Cytotoxicity Assay kit (G-Biosciences). **(B)** As in panel A, but cells were also infected with *cteG::aadA* carrying pCteG-2HA[Pgp4$^+$] (also named pCteG-2HA in Annexes Table 3; CteG$^-$/CteG-2HA$^+$/Pgp4$^+$), or pCteG-2HA[Pgp4$^-$] (CteG$^-$/CteG-2HA$^+$/Pgp4$^-$). **(C)** The LDH released by lysed host cells was measured as in panels A, B for HeLa cells infected with *C. trachomatis* L2/434 carrying pVector[Pgp4$^+$] (CteG$^+$/Pgp4$^+$), pVector[Pgp4$^-$] (CteG$^+$/Pgp4$^-$) or pCteG-2HA[Pgp4$^-$] (CteG$^+$/CteG-2HA$^+$/Pgp4$^-$). Statistical significance was assessed by using ordinary one-way ANOVA and Tukey's post-test analysis. Data correspond to mean ± SEM (n = 4 in panel A, n = 7 in panel B and n ≥ 3 in panel C; ns, non-significant; *p<0.05; **p<0.01; ****p<0.0001).

3.2.6 Overproduction of CteG-2HA suppresses the defect of CteG- and Pgp4-deficient *C. trachomatis* to mediate host cell lysis

CteG can be detected in the cytoplasm of host cells infected by *C. trachomatis* from 16 h post-infection (Pais *et al.*, 2019), but the CteG-dependent increase in host cell lysis occurs from 48 to 72 h post-infection (Figures 3.16 and 3.17). This suggests a mechanism that initially keeps CteG inhibited in its ability to promote host cell lysis. One hypothesis to explain how CteG and Pgp4 act on the same pathway to promote host cell lysis would be that the product of a Pgp4-regulated gene could be involved in a process of activation or inhibition relief of CteG. We reasoned that if the product of a Pgp4-regulated gene mediates activation or inhibition relief of CteG, then this might be surpassed by overproduction of CteG-2HA. Our previous data indicate that in a *C. trachomatis* strain with plasmid encoded CteG-2HA its mRNA levels are ~10-fold higher than of chromosomal *cteG* (Pais *et al.*, 2019). Therefore, we infected HeLa cells with the same strains as before [*C. trachomatis* L2/434 carrying pVector[Pgp4$^+$] (CteG$^+$/Pgp4$^+$), L2/434 carrying pVector[Pgp4$^-$] (CteG$^+$/Pgp4$^-$), *cteG::aadA* carrying pVector[Pgp4$^+$] (CteG$^-$/Pgp4$^+$), or *cteG::aadA* carrying pVector[Pgp4$^-$] (CteG$^-$/Pgp4$^-$); Figure 3.23A] but also with Pgp4$^+$ and Pgp4$^-$ strains carrying a plasmid encoding CteG-2HA [*cteG::aadA* carrying plasmid CteG-2HA[Pgp4$^+$] (CteG$^-$/CteG-2HA$^+$/Pgp4$^+$), corresponding to the complemented strain used in other experiments, or *cteG::aadA* carrying plasmid CteG-2HA[Pgp4$^-$] (CteG$^-$/CteG-2HA$^+$/Pgp4$^-$)] (Annexes Table 3), and monitored the release of LDH at 72 h post-infection. This further confirmed that the defect in host cell lysis of the CteG$^-$/Pgp4$^-$ strain is similar to the CteG$^+$/Pgp4$^-$ or CteG$^-$/Pgp4$^+$ strains (Figure 3.23B) and revealed that the Pgp4$^-$ strain carrying plasmid encoded CteG-2HA (CteG$^-$/CteG-2HA$^+$/Pgp4$^-$) displays an ability to promote host cell lysis identical to the CteG$^+$/Pgp4$^+$ and CteG$^-$/CteG-2HA$^+$/Pgp4$^+$ strains (Figure 3.23B). This suggested that the overproduction of CteG-2HA could restore the ability of a Pgp4-deficient strain to induce host cell lysis. To corroborate this hypothesis, we generated a *C. trachomatis* L2/434-derived strain, which also harbors an intact copy of *cteG* in the chromosome, carrying pCteG-2HA[Pgp4$^-$] (Annexes Table 3; CteG$^+$/CteG-2HA$^+$/Pgp4$^-$) and measured the release of LDH in cells infected by this strain or by the *C. trachomatis* L2/434 carrying pVector[Pgp4$^+$] (CteG$^+$/Pgp4$^+$) or L2/434 carrying pVector[Pgp4$^-$] (CteG$^+$/Pgp4$^-$) strains at 72 h post-infection. We observed that the CteG$^+$/CteG-2HA$^+$/Pgp4$^-$ strain has an ability to promote host cell lysis comparable to the CteG$^+$/Pgp4$^+$ strain (Figure 3.23C). Therefore, the overproduction of CteG-2HA can compensate for the lack of Pgp4 regarding the ability of *C. trachomatis* to induce host cell lysis.

3.2.7 Conclusions

In summary, throughout the results described in this section, we found less infectious chlamydiae present in the culture supernatant of a *C. trachomatis cteG* mutant strain comparing to the wild-type strain, and that the *cteG* mutant strain has a defect in inducing host cell cytotoxicity comparing to the wild-type strain. Both phenotypes were complemented in a *C. trachomatis cteG* mutant strain overproducing CteG-2HA from a plasmid, confirming that CteG mediates *C. trachomatis* lytic exit from host cells. Moreover, although Pgp4 does not appear to regulate the production or subcellular localization of CteG, both proteins act on the same pathway to promote host cell lysis. Finally, the overproduction of CteG-2HA from a plasmid circumvented the defect in host cell cytotoxicity caused by the absence of Pgp4. Overall, this revealed the role of CteG in an essential step of the *Chlamydia* infectious cycle and provided mechanistic insights of its possible functional relation to Pgp4.

3.3 Identification of CteG host cell interacting partners[6]

To identify host cell interacting proteins of CteG, we performed co-IP assays followed by liquid chromatography-tandem mass spectrometry (LC-MS/MS) analysis (Figure 3.24). First, extracts of HeLa cells ectopically producing mEGFP-CteG$_{FL}$ or mEGFP alone as control were affinity-purified with anti-GFP (GFP-Trap) beads (Figure 3.24). Samples from several steps of the co-IP procedure were collected, including the pellet of lysed cells ("pellet" fraction), the supernatant of lysed cells before ("input" fraction) or after ("non-bound" fraction) incubation with anti-GFP beads, and washed anti-GFP beads containing immunoprecipitated proteins ("bound" fraction) (Figure 3.24). These samples were analyzed by SDS-PAGE followed by Coomassie blue staining or immunoblotting (Figure 3.25). By Coomassie staining, the presence of clear bands that could correspond to mEGFP (27 kDa) or mEGFP-CteG$_{FL}$ (96 kDa) was not detected in the cell pellet (Figure 3.25A, lanes 1 and 2), input (Figure 3.25A, lanes 3 and 4) or non-bound (Figure 3.25A, lanes 5 and 6) fractions. By Coomassie staining, a band likely corresponding to mEGFP was barely detected in the bound fraction (Figure 3.25A, lane 7), but the same did not happen for mEGFP-CteG$_{FL}$ (Figure 3.25A, lane 8). Furthermore, no additional bands that could correspond to possible interacting partners of both proteins were observed. Nevertheless, detection of both mEGFP and mEGFP-CteG$_{FL}$ in the input and bound fractions by immunoblotting using an anti-GFP antibody confirmed their production and immunoprecipitation (Figure 3.25B).

The approximate total amount of protein present in the co-IP eluates (two for mEGFP, and three for mEGFP-CteG$_{FL}$) was determined in Coomassie-stained SDS-PAGE by using a standard curve of bovine serum albumin (BSA) (exemplified in Figure 3.25C). The eluates were then subjected to LC-MS/MS (Figure 3.24), and the raw data was searched against a human proteome database as described in Materials and Methods. The hits obtained for the two and the three samples of mEGFP (>17,090) and mEGFP-CteG$_{FL}$ (>14,615), respectively, were sorted in terms of their occurrence in all the mEGFP-CteG$_{FL}$ eluates and in none of the mEGFP eluates (Annexes Table 5). Generally, sorted hit proteins were associated with low numbers of identified peptide sequences (#Peptides and #Unique Peptides) and of total matching peptide sequences across the three mEGFP-CteG$_{FL}$ samples (Annexes Table 5). Additionally, even among hit proteins that presented higher values for those parameters, many either have unknown functions, or participate in cellular processes or have predicted subcellular localizations that are not compatible with the information we currently have about CteG, namely ribosomal (as 40S or 28S ribosomal proteins) or nuclear (as histone H1 or lamin-B2)

[6] The experiments described in this section were performed by Inês Serrano Pereira. The LC-MS/MS analysis was performed at Clarify Analytical (Portugal).

proteins (Annexes Table 5). Yet, it is possible that some of the identified proteins interact with CteG either to promote host cell lysis or in the context of other function CteG might exert.

Overall, we were unable to draw clear conclusions from this data about possible interacting proteins of CteG that would be further validated and studied. However, in future experiments, some of the proteins obtained could still be selected to investigate a possible interaction with CteG or, conversely, other methodologies could be employed to search for CteG interacting proteins.

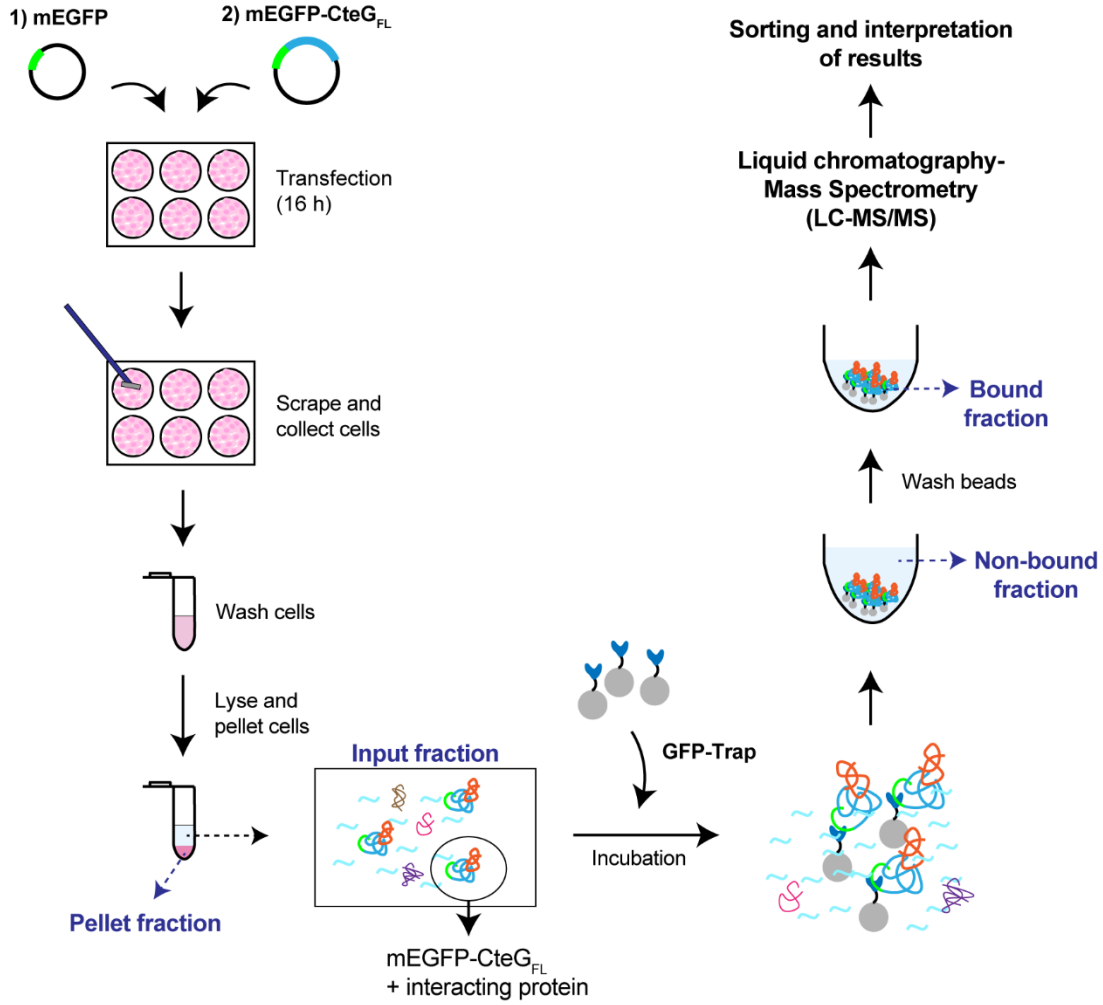


Figure 3.24. Simplified illustration of the co-immunoprecipitation assays using extracts of HeLa cells ectopically producing mEGFP-CteG$_{FL}$ or mEGFP alone for mass spectrometry analysis. HeLa cells were transfected with plasmids encoding mEGFP-CteG$_{FL}$ or mEGFP alone. After 16 h, cells were scraped, collected to tubes, washed, and lysed. The cell lysate was then centrifuged to separate the cellular debris and insoluble proteins (pellet fraction) from the supernatant fraction containing soluble proteins, including mEGFP-CteG$_{FL}$ or mEGFP alone and their possible interacting partners (input fraction). Pre-washed GFP-Trap beads (ChromoTek) were incubated for 4 h with the input fractions to allow interaction with mEGFP proteins. Beads were then settled by centrifugation, and the supernatant containing proteins that did not interact with the beads was removed (non-bound fraction). Beads were washed 6 times and finally resuspended in Laemmli Sample Buffer 5x (bound fraction). Eluates were analyzed by LC-MS/MS and hit proteins were sorted in terms of their occurrence in all the mEGFP-CteG$_{FL}$ eluates and in none of the mEGFP eluates (Annexes Table 5).

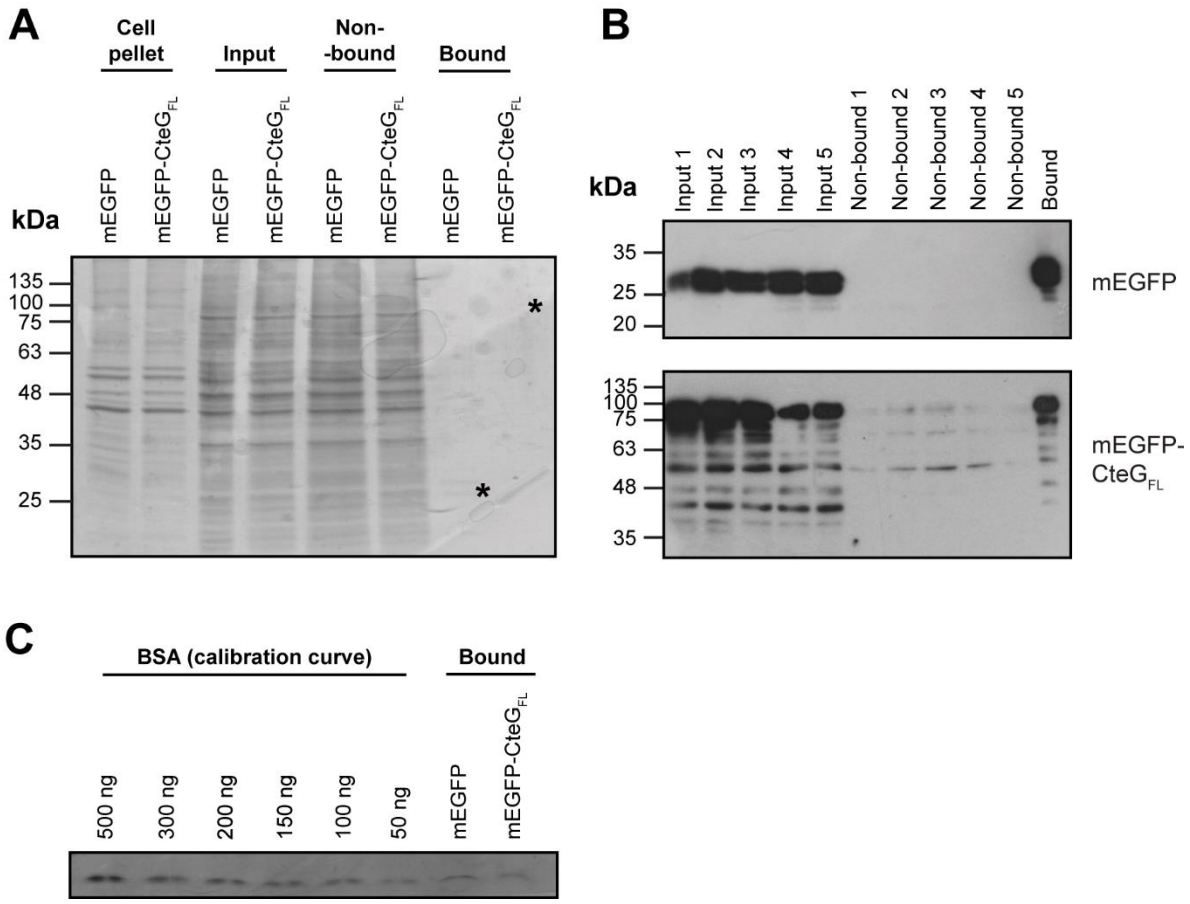


Figure 3.25. Analysis of samples collected from different steps of the co-immunoprecipitation assays by SDS-PAGE and Coomassie staining, or immunoblotting. (A) Samples from different steps of one representative co-immunoprecipitation (co-IP) assay using extracts of HeLa cells producing mEGFP alone or mEGFP-CteG$_{FL}$ were collected. The cell pellet, input, non-bound, and bound fractions were analyzed by SDS-PAGE followed by Coomassie staining. ∗ indicates the expected molecular mass for mEGFP alone or mEGFP-CteG$_{FL}$ which should be visible in the bound fraction. **(B)** co-IP assays were scaled up to obtain enough protein eluates for the LC-MS/MS analysis. Each two wells from a 6-well plate of transfected cells were processed separately following the procedure described in Figure 3.24. Samples from different input and non-bound fractions (each corresponding to two wells of transfected cells), and from the final bound fraction (where all the eluates were mixed to form a "final" eluate) were analyzed by immunoblotting using an antibody against GFP and the appropriate HRP-conjugated secondary antibody. Proteins were detected using SuperSignal West Pico detection kit (Thermo Fisher Scientific). **(C)** Quantification of the approximate total amount of protein present in mEGFP alone or mEGFP-CteG$_{FL}$ "final" eluates of one of the performed co-IP assays by Coomassie-stained SDS-PAGE and BSA standard curve.

3.4 Distribution of CteG homologs among *Chlamydiaceae*[7]

An in-depth analysis of CteG homologs among *Chlamydiaceae*, particularly their subcellular localization and conservation of specific amino acid sequences, could provide further clues about the determinants of the localization and function of CteG. In our previous work, we identified putative CteG homologs among *Chlamydiaceae* by performing a Position-Specific Iterated (PSI)-protein BLAST (BLASTp) (Pais *et al.*, 2019). We also noticed that CteG appeared to be unique in terms of number of its putative homologs among *Chlamydiaceae* when comparing with other chlamydial T3S effectors.

3.4.1 Identification of putative homologs of CteG within *Chlamydiaceae* by reciprocal best hit BLAST

To identify putative homologs of CteG among *Chlamydiaceae,* we performed reciprocal best hit BLAST analysis (detailed in Materials and Methods). First, we used genome assemblies annotated in National Center for Biotechnology Information (Genbank, NCBI) databases for different species of the Family *Chlamydiaceae* to generate genome nucleotide databases (Figure 3.26). We also included possible species of the Family *Chlamydiaceae* that were more recently identified and whose genomes were deposited in NCBI, namely *Ca.* C. corallus (*C. corallus*), *C. ibidis*, *C. poikilotherma*, *C. sanzinia* and *C. serpentis* (Table 3.1). Furthermore, we considered species from other Families of the Phylum *Chlamydiae*, namely *Estrella lausannensis* from the Family *Criblamydiaceae* and *S. negevensis* from the Family *Simkaniaceae*, and three members of a non-pathogenic sister clade of *Chlamydiaceae* [*Chlamydia* clade (CC)-IV] denominated *Chlamydiae* bacterium K940_chlam_9, *Chlamydiae* bacterium K1000_chlam_4 and *Chlamydiae* bacterium KR126_chlam_2, whose genomes are also deposited in NCBI. We then performed a tBLASTx using the nucleotide sequence of *cteG* against each genome nucleotide database. Hit genomic regions were then analyzed using an *ab initio* gene prediction tool [AUGUSTUS, (Keller *et al.*, 2011)] to identify putative genes and their respective encoded proteins (Figure 3.26). Putative homologs that were identified for CteG [(Pais *et al.*, 2019) and this work; Table 3.1] by BLASTp but were not recovered after reciprocal BLAST (Figure 3.26) were not further considered. This was the case of M832_01180 and M832_01160 from *C. avium*, CCA_00297 from *C. caviae*, M787_003335 and M787_003340 from *C. gallinacea* and CPSIT_0422 from *C. psittaci* (Table 3.1).

[7] The analysis described in this section was performed by Inês Serrano Pereira, in collaboration with Dr. Carla Gonçalves and Prof. Paula Gonçalves (from the Yeast Genomics Laboratory at the host institution).

We identified between one and four CteG putative homologs within each species of the Family *Chlamydiaceae* (Table 3.1). BLASTp of CteG against a *C. ibidis* protein database identified only one putative homolog, H359_0725, that was recovered after reciprocal BLAST (Table 3.1). Another protein named H359_0450 was also identified by reciprocal BLAST with a high associated e-value, but not by BLASTp (Figure 3.26 and Table 3.1). Indeed, a direct alignment between the amino acid sequences of CteG and H359_0450 using the BLASTp tool detected similarities between both sequences (21% of identity; see Figure 3.27 below). Therefore, H359_0450 was considered a putative CteG homolog and was included in posterior analyses. In total, 27 best hit putative homologs of CteG were identified (Table 3.1).

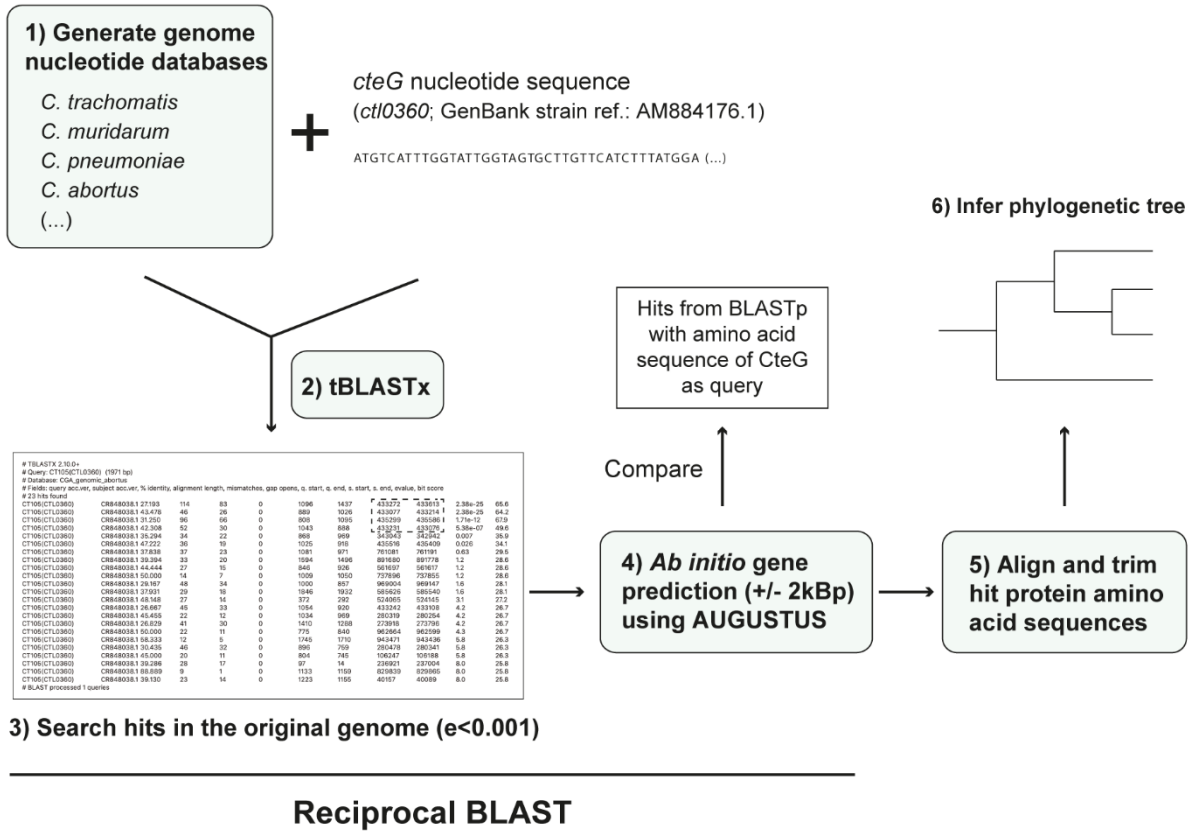


Figure 3.26. Graphical summary of the reciprocal best hit BLAST and phylogenetic analyses to identify putative homologs of CteG. Reciprocal best hit BLAST analysis involved the following steps: 1) genome databases were generated for each *Chlamydiae* species from Genbank (NCBI) genome assemblies; 2) the nucleotide sequence of *cteG*/*ctl0360* was used to perform a tBLASTx against each genome database; 3) best hit genomic regions with an e-value lower than 0.001 were localized in the original genome; 4) the nucleotide sequence 2 kBp upstream and 2 kBp downstream of the hit genomic regions were submitted in AUGUSTUS for *ab initio* gene prediction. Proteins potentially encoded by *ab initio* predicted genes were compared with the putative homologs of CteG found by BLASTp [(Pais *et al.*, 2019) and this work, Table 3.1]. Proteins that were identified by BLASTp but were not recovered with best hit reciprocal BLAST were not considered homologs of CteG. Then, 5) multiple alignments were performed with the amino acid sequences of the best hit putative homologs of CteG, followed by trimming of poorly aligned sequences using TrimAl. Finally, 6) a phylogenetic tree of CteG was inferred with the aligned and trimmed sequences using IQ-TREE (see Figure 3.28 below). A more detailed description of this methodology can be found in Materials and Methods.

Table 3.1. Possible homologs of CteG in other *Chlamydiae* before and after both bioinformatics and phylogenetic analyses.

	Protein (hits found in a PSI-BLAST)	Protein ID (NCBI)	Recovered after reciprocal BLAST?	Grouped with CteG in a phylogenetic tree?
C. abortus [a]	CAB376	CAH63829.1	Yes	Yes
C. avium [a]	M832_01180	AHK62988.1	No	-
	M832_01160	AHK62986.1	No	-
C. caviae [a]	CCA_00389	AAP05136.1	Yes	Yes
	CCA_00390	AAP05137.1	Yes	Yes
	CCA_00297	AAP05046.1	No	-
	CCA_00298	AAP05047.1	Yes	No
C. corallus [b]	WP_151899123	WP_151899123	Yes	Yes
C. felis [a]	CF0619	BAE81391.1	Yes	Yes
	CF0618	BAE81390.1	Yes	Yes
	CF0705	BAE81477.1	Yes	No
	CF0706	BAE81478.1	No*	No
C. ibidis [b]	H359_0725	EQM62726.1	Yes	Yes
	H359_0450	EQM62777.1	Yes**	Yes
C. gallinacea [a]	M787_003335	ANG66342.1	No	-
	M787_003340	ANG66343.1	No	-
C. muridarum [a]	TC_0381	AAF39239.1	Yes	Yes
C. pecorum [a]	G5S_0729	AEB41680.1	Yes	Yes
	G5S_0731	AEB41682.1	Yes	Yes
	G5S_0733	AEB41684.1	Yes	Yes
C. pneumoniae [a]	CPn_0405	AAD18549.1	Yes	Yes
	CPn_0404	AAD18548.1	Yes	Yes
C. poikilotherma [b]	C834K_0411	SYX08872.1	Yes	Yes
	C834K_0412	SYX08873.1	Yes	Yes
	C834K_0321	SYX08786.1	Yes	No
	C834K_0322	SYX08787.1	Yes	No
C. psittaci [a]	CPSIT_0421	ADZ18264.1	Yes	Yes
	CPSIT_0422	ADZ18833.1	No	-
C. sanzinia [b]	Cs308_0967	ANH79137.1	Yes	Yes
C. serpentis [b]	C10C_0369	SPN73539.1	Yes	Yes
	C10C_1043	SPN74174.1	Yes	No
C. suis [a]	Q499_0113	ESN89684.1	Yes	Yes

	Q499_0114	ESN89662.1	Yes	Yes
Chlamydiae bacterium K940_chlam_9 §	Not found	-	-	-
Chlamydiae bacterium K1000_chlam_4 §	Not found	-	-	-
Chlamydiae bacterium KR126_chlam_2 §	Not found	-	-	-
S. negevensis	Not found	-	-	-
Estrella lausannensis	Not found	-	-	-

[a] CteG (CTL0360; Protein ID: CAP03800) homologs for these species were previously identified by PSI-BLAST (Pais *et al.*, 2019).

[b] A PSI-BLAST as described in (Pais *et al.*, 2019) was performed to search possible CteG homologs in representative strains of additional *Chlamydia* spp. in the Family *Chlamydiaceae* – *C. corallus* strain G3/2742-324, *C. ibidis* strain 10-1398/6, *C. sanzinia* strain 2742-308, *C. serpentis* strain H15-1957-10C and *C. poikilotherma* strain S15-834K.

**C. felis* CF0706 tBLASTx e-value was slightly above 0.001, but the protein was still included in the phylogenetic analysis.

** *C. ibidis* H359_0450 was not identified by protein BLAST of CteG against NCBI protein databases but was found with best hit reciprocal BLAST.

§ Species that belong to CC-IV, a sister clade of the Family *Chlamydiaceae* (Dharamshi *et al.*, 2020).

No putative CteG homologs were identified for species from the Families *Criblamydiaceae* or *Simkaniaceae*, or from the CC-IV clade (Table 3.1). Overall, this indicates that *cteG* should have been acquired by an ancestral species that originated pathogenic *Chlamydiaceae*.

3.4.2 Phylogenetic analysis of the homologs of CteG among *Chlamydiaceae*

We next performed individual and multiple alignments using the amino acid sequences of CteG and of its best hit putative homologs among *Chlamydiaceae* (Figure 3.26). Because the two short proteins CPn_0404 and CPn_0405 of *C. pneumoniae* are homologs to CteG but not to each other (Figure 3.27), we selected only CPn_0404 to perform multiple alignments. These analyses confirmed that among the CteG putative homologs among *Chlamydiaceae* only TC0381 (from *C. muridarum*) and Q499_0113 (from *C. suis*) displayed a high degree of identity (53% and 47%, respectively) to the entire polypeptide sequence of CteG (Figure 3.27). All the other CteG putative homologs share between 20% and 30% identity with CteG but have significant parts of their amino acid sequences that are unrelated to CteG (Figure 3.27). Moreover, we did not identify any region of CteG that was common to all putative homologs (Figure 3.27). Along with analyses of the subcellular localization of these putative homologs, this information could

be considered in future studies to pinpoint the determinants of the localization of CteG during *C. trachomatis* infection.

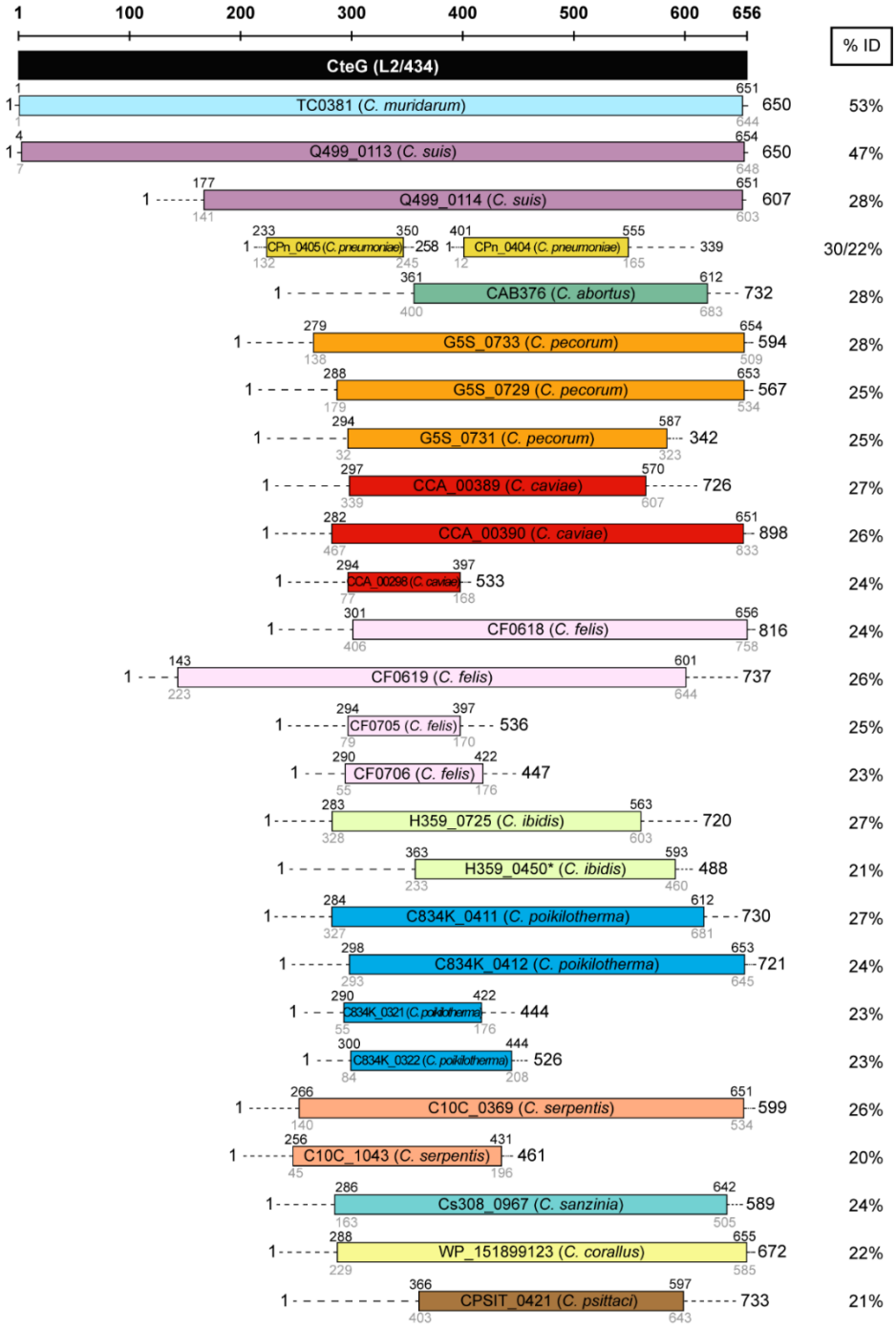


Figure 3.27. Individual alignments between CteG and its putative homologs. The amino acid sequences of CteG from *C. trachomatis* strain L2/434 and of its putative homologs were individually aligned using BLASTp and the region of each CteG putative homolog that possesses identity with CteG is shown. The following features for each CteG putative homolog are depicted: its length in amino acids (bigger black numbers); the region of its sequence (smaller grey numbers) that shares identity with a region of CteG (smaller black numbers); the percentage of identity between both aligned sequences (% ID). * H359_0450 was identified with reciprocal best hit BLAST analysis and was considered a putative CteG homolog.

A phylogenetic tree was then inferred with the multiple alignments of the trimmed amino acid sequences that showed significant identity to CteG (Figures 3.26 and 3.28).

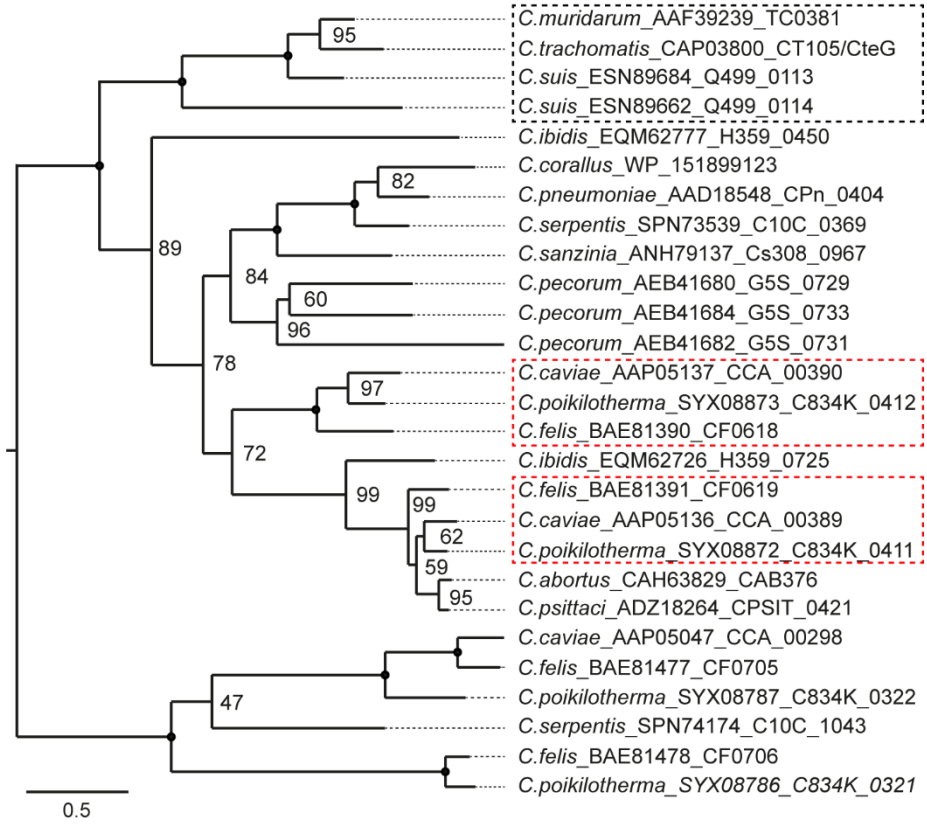


Figure 3.28. Phylogenetic tree of CteG and its putative homologs in species of the Family *Chlamydiaceae*. Phylogenetic relationships between *C. trachomatis* CteG and its putative homologs in other species from the Family *Chlamydiaceae* are depicted, based on a maximum likelihood phylogeny constructed as described in Materials and Methods. Each protein is indicated by "Species name_NCBI protein ID_NCBI protein designation" (see Table 3.1 above). The black box refers to the group formed by *C. trachomatis* CteG and its putative homologs in *C. suis* and *C. muridarum*, and the red boxes indicate putative homologs of CteG in *C. caviae*, *C. poikilotherma* and *C. felis* that could be paralogs. Branches with 100% support are indicated by black dots.

Reasoning that the tree root corresponds to an ancient common ancestor protein, a first duplication event originated a precursor of CteG and of 20 of its putative homologs among *Chlamydiaceae* (upper clade), and another that was the ancestor of the other 6 putative homologs (lower clade). The latter include one protein from *C. caviae* (CCA_00298), one protein from *C. serpentis* (C10C_1043), two proteins from *C. felis* (CF0705 and CF0706) and two proteins from *C. poikilotherma* (C834K_0321 and C834K_0322). While CF0705 and CF0706, or C834K_0321 and C834K_0322 could be paralogs, the clade where all 6 proteins were grouped is different from that where CteG was included (Figure 3.28). These genes possibly originated from an early duplication event that occurred in the ancestor of pathogenic *Chlamydiaceae* (see Figure 1.1 above and Figure 3.29 below) prior to speciation. This suggests that the *cteG* paralog originating from this early duplication event was eventually lost in all other *Chlamydia* species (except *C. caviae*, *C. serpentis*, *C. felis*, *and C. poikilotherma*).

To better compare the evolution of CteG and its putative homologs among *Chlamydiaceae* with the phylogenetic relationship between the corresponding species, we inferred a phylogenetic tree based on 215 concatenated proteins from all the studied members of the Family *Chlamydiaceae*, species of the CC-IV clade and species that belong to more distant lineages, including *S. negevensis*, *E. lausannensis* and *Ca.* Protochlamydia naegleriophila (*P. naegleriophila*) from the Family *Parachlamydiaceae* (Figure 3.29). Consistent with other studies, species from the CC-IV clade are more closely related to the *Chlamydiaceae* than *S. negevensis*, *E. lausannensis* or *P. naegleriophila*, which were grouped in a distinct clade (Phillips *et al.*, 2019; Dharamshi *et al.*, 2020) (Figure 3.29). Moreover, the phylogenetic relation of the *Chlamydiaceae* members is either consistent (Nunes and Gomes, 2014) or slightly differs from that described in previous studies (Phillips *et al.*, 2019; Dharamshi *et al.*, 2020) (Figure 3.29). Such variations could be attributed to the inclusion of more species, especially in more recent studies, or to the use of concatenated proteins or gene nucleotide sequences in phylogeny construction.

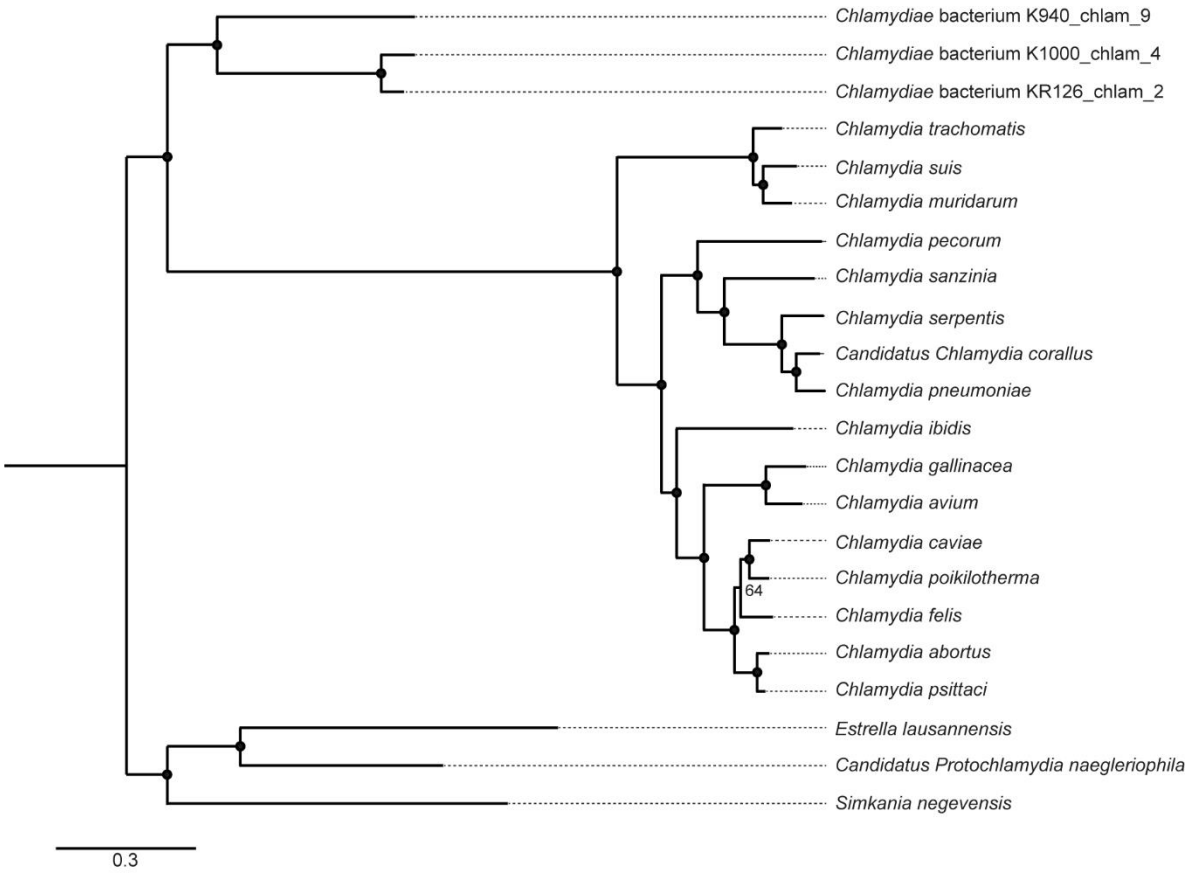


Figure 3.29. Phylogenetic relationships between species from the Family *Chlamydiaceae* and from other *Chlamydiae*. A species tree was generated by maximum likelihood using a concatenated alignment of 215 proteins as described in Materials and Methods. The analysis comprised all the studied species from the pathogenic Family *Chlamydiaceae*, species from the closely related CC-IV clade (upper clade) and species from more phylogenetically distant Families (*E. lausannensis*, *P. naegleriophila* and *S. negevensis*). Branches with 100% support are indicated by black dots.

The putative homologs of CteG in *C. muridarum* (TC0381) and *C. suis* (Q499_0113 and Q499_0114) group with CteG itself according to the species phylogenetic relation (Figure 3.28, black dashed box; Figure 3.29). The fact that Q499_0114 appears on the same clade as Q499_0113 suggests that their corresponding genes are paralogs arising by gene duplication in an ancestor of all three species (Figure 3.28). Again, this suggests that the *cteG* paralog originating from this duplication event was lost in *C. trachomatis* and *C. muridarum*.

Likewise, *C. pecorum* G5S_0729, G5S_0731 and G5S_0733 belong in the same clade and could also be classified as paralogs (Figure 3.28). Moreover, *C. caviae*, *C. felis* and *C. poikilotherma* encode two putative CteG homologs each (CCA_00390/CCA_00389, CF0618/CF0619 and C834K_0412/C834K_0411, respectively) that could be paralogs to each other, as they group in two separate clades that are closely related (Figure 3.28, red dashed boxes). Again, this indicates the occurrence of multiple gene duplication and gene loss events. On the other hand, *C. ibidis* encodes two CteG putative homologs that appear to be unrelated to each other: while H359_0725 groups with proteins from other species, H359_0450 forms itself a separate clade (Figure 3.28) which could mean that this protein is experiencing divergence. Apart from this, CteG and its identified putative homologs were grouped essentially according to the species phylogeny (Figure 3.29), suggesting that their evolution progressed with the evolution of the *Chlamydiaceae* species. Overall, this suggests that the evolutionary history of *cteG* is complex and marked by several events of gene deletion and gene loss.

3.4.3 Synteny of *cteG*

We next studied the genomic region of *cteG* by comparison with the genomic region of its putative homologs in *Chlamydiaceae* in terms of gene organization and conservation. We started by identifying ~11-12 proteins potentially encoded upstream of *cteG* and ~10-15 proteins potentially encoded downstream of *cteG* or of its putative homologs using AUGUSTUS (Keller *et al.*, 2011). For each protein encoded by a gene identified in the vicinity of *cteG*, we used BLASTp to locate their putative homolog in each *Chlamydia* species. Although no putative CteG homologs were found in *C. avium* and *C. gallinacea*, we also analyzed the genomic regions of these species encoding possible homologs of *cteG* neighboring genes. Figure 3.30 shows an illustration of all the analyzed genomic regions, in which a different color was assigned to each gene and their respective homolog in each *Chlamydia* species.

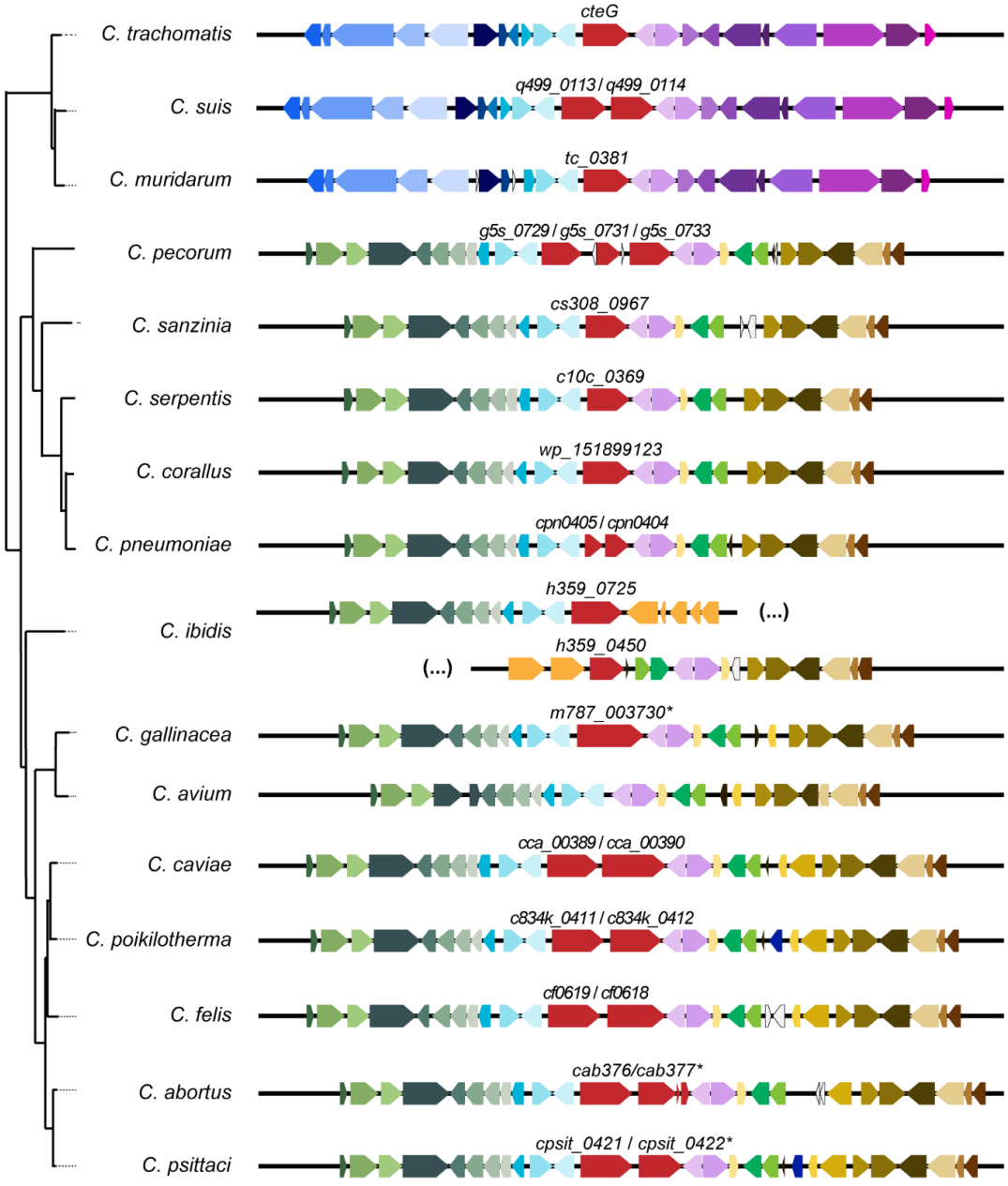


Figure 3.30. Organization of the genomic region of *cteG* and of each of its putative homologs. Putative genes encoded in the genomic region of *cteG* or of the genes encoding its putative homologs were identified by *ab initio* prediction using AUGUSTUS. For each protein potentially encoded by the predicted genes, the identity of its putative homologs was confirmed by BLASTp and a different color was given to each group of putative homologs. Genomic regions are depicted according to the species tree in Figure 3.29. For both *cteG* putative homologs in *C. ibidis*, the two corresponding genomic regions are represented separately. Genes for which no putative homologs were found are colored in white. **cab377* was not identified as a putative homolog of *cteG* and is annotated in NCBI databases as a fragmented pseudogene (the 3 fragments are shown); *cpsit_0422* and *m787_003730* were discarded with the reciprocal best hit BLAST and were not considered in further analyses.

Based on the similarity of their genomic regions and apart minor variations on gene presence and order, *Chlamydia* species may be divided into two distinct groups: one comprising *C. trachomatis*, *C. suis* and *C. muridarum* and other comprising all the other species (Figure 3.30). The first three proteins encoded upstream of *cteG* are conserved between all species in terms of identity and location (Figure 3.30). The first two correspond to a putative enoyl-[acyl-carrier-protein] reductase (CT104/CTL0359/FabI), shown to be involved in fatty acid biosynthesis in other bacteria (Jiangwei *et al.*, 2016; Radka *et al.*, 2020), and to a putative hydrolase/phosphatase (CT103/CTL0358) from the haloacid dehydrogenase (HAD) superfamily, whose proteins participate in diverse metabolic reactions (Kuznetsova *et al.*, 2006). The third hypothetical protein (CT102/CTL0357) has unknown function but is annotated as a putative membrane protein. Regarding putative proteins encoded downstream of *cteG*, the first two are conserved between species but slightly differ in their genomic position (Figure 3.30). These proteins were identified as a putative tRNA pseudouridine synthase (CT106/CTL0361), a ubiquitous type of enzymes involved in tRNA posttranscriptional modification (Hamma and Ferré-D'Amaré, 2006), and a putative A/G-specific adenine glycosylase (CT107/CTL0362), an enzyme involved in DNA mismatch repairing (Au *et al.*, 1989). Although most of these functions were deduced by homology with other known proteins, their preservation across all species suggests they may be important for chlamydial survival and/or replication.

Consistent with the phylogenetic proximity between *C. trachomatis*, *C. suis* and *C. muridarum* (Figure 3.29), the genomic region of these species is well conserved (Figure 3.30). While the function of most proteins encoded by genes in this genomic region is unknown or was inferred by homology with other proteins and thus requires confirmation, many are described to be T3S system substrates [as Incs CT101/MrcA, IncD/E/F/G and IncA (Bugalhão and Mota, 2019), most of which are not depicted], chaperones [CT110/GroEL/Hsp60 and CT111/GroES/Hsp10 (Wilson *et al.*, 2005)], or proteins involved in gene expression and its regulation [CT096/InfB and CT097/NusA, (Huang *et al.*, 2022)] (Figure 3.30). Being CteG a T3S substrate involved in host cell lysis (see section 3.2 above), the genomic proximity of *cteG* with genes encoding other T3S effectors, including CT101/MrcA, might be related with a hypothetical acquisition of several virulence-associated genes by an ancestral species.

Regarding *C. pecorum*, *C. sanzinia*, *C. serpentis*, *C. corallus*, *C. pneumoniae*, *C. caviae*, *C. poikilotherma*, *C. felis*, *C. psittaci* and *C. abortus*, they also exhibit conserved genomic vicinities (Figure 3.30) that were likely acquired from a common ancestor (Figure 3.29). *C. ibidis* is also included in this group but exhibits a peculiar difference: genes that share homology to others from the former ten species are located upstream from *h359_0725* and downstream from *h359_0450* (Table 3.1 and Figure 3.30). *C. ibidis* is closely related to *C. avium* and *C. gallinacea* (Figure 3.29), for which we also observed analogous genomic regions but did not find putative

homologs of CteG (Table 3.1), suggesting that the genomic organization and structure of these three species is deviating from those of other close *Chlamydiaceae* species. Interestingly, in the case of *C. gallinacea*, the genomic position in which *cteG* putative homolog genes are usually located encodes a putative protein (M787_003730) that may be evolutionarily related to CteG but currently has no defined function (Figure 3.30). Conversely, the same genomic position in *C. avium* is empty suggesting that if once present, the gene encoding a CteG homolog was lost (Figure 3.30). The gene encoding *C. abortus* CAB377 is annotated as a fragmented pseudogene, but its proximity to the gene encoding CAB376 suggests that *cab377* could have been a *cteG* homolog and that both genes could have been paralogs. The same argument could be applied to *C. psittaci* CPSIT_0422 which was discarded by the reciprocal BLAST analysis (Figure 3.30). Overall, putative proteins encoded in the vicinity of *cteG* homologs in all these species possess homologs in *C. trachomatis*. Some of these *C. trachomatis* homologs have been studied, as the chaperone CT260/CTL0409 (Spaeth *et al*., 2009), CT262/CTL0514 and CT263/CTL0515 which participate in the quinone biosynthesis pathway (Barta *et al*., 2014) or CT257/CTL0509/Lda4 which has been proposed to associate with yeast lipid droplets (Kumar *et al*., 2006). Otherwise, *C. trachomatis* homologs of most proteins have inferred functions based on homologs in other bacterial species such as CT261/CTL0513 which is a putative epsilon subunit of DNA polymerase, the potential ATP-binding ABC-transporter CT264/CTL0516, or putative enzymes involved in several metabolic pathways as CT265/CTL0517 (lipid metabolism) and CT258/ CTL0510 (cysteine desulfurase).

In summary, this analysis provided a wider perspective about the genomic region where *cteG* and its putative homologs localize. In future studies, it would be interesting to have a deeper understanding about the evolution of these proteins and to correlate it with the evolution of their genomic vicinity.

3.4.4 Identification of homologs among *Chlamydiaceae* of chlamydial non-Inc T3S substrates

Finally, we analyzed the conservation of other chlamydial T3S substrates among *Chlamydiaceae*. We sought to analyze if other chlamydial T3S substrates possessed duplicated homologs/paralogs among *Chlamydia* species as CteG. Using a procedure identical to that described above for *cteG* (Figure 3.26) but using the nucleotide sequence of *C. trachomatis* genes that encode well-known non-Inc or putative T3S effectors (Bugalhão and Mota, 2019) against each chlamydial genome nucleotide database (Table 3.2). This revealed that except for CT620, CT621, and CT711, all other chlamydial proteins analyzed generally just have one homolog in each species (Table 3.2). CT620, CT621, and CT711, together with CT619 and CT712, belong to a family of chlamydial T3S effectors characterized by a DUF582 domain of unknown function

(Muschiol *et al.*, 2011). As these three proteins are paralogs in *C. trachomatis*, the identification of several homologs in other *Chlamydia* species indicates the presence of this family of effectors in these species. This is different from the case of *cteG* in which paralog genes were found in seven chlamydial species (Table 3.2), while in *C. trachomatis* the gene is unique.

In addition, we did not identify putative homologs for most of these proteins in *S. negevensis*, *E. lausannensis* or species from the CC-IV clade, except for CT089/CopN [a possible regulator of the T3S system (Fields and Hackstadt, 2000)], CT473/Lda3 [targets and modulates host lipid droplets (Kumar *et al.*, 2006); putative homolog only detected in *E. lausannensis*], CT042/GlgX [glycogen hydrolase (Gehre *et al.*, 2016)], CT737/NUE [histone methyltransferase; interferes with host cell transcription (Pennini *et al.*, 2010)] and CT798/GlgA [glycogen synthase (Lu *et al.*, 2013)] (Table 3.2). Putative homologs of the four latter proteins, as well as of CT620, CT621 and CT711 are also present in members of the CC-IV clade (Table 3.2). The presence of these homologs in the CC-IV clade would be consistent with a gradual acquisition of virulence-associated genes by ancestral species common to both the CC-IV clade and the pathogenic *Chlamydiaceae* members (Dharamshi *et al.*, 2020), which is a hallmark of co-evolution with host eukaryotic cells. Taken together, these results support our previous observations that CteG possesses the most variable number of putative homologs (orthologs and paralogs) among *Chlamydia* effector genes.

Table 3.2. Analysis of the number of homologs of different non-Inc *C. trachomatis* proteins in other *Chlamydiae* by reciprocal best hit BLAST (found with protein BLAST / recovered with reciprocal BLAST).

		Chlamydiaceae															
Protein (serovar D)	**Protein (serovar L2)**	*C. trachomatis*	*C. abortus*	*C. avium*	*C. caviae*	*C. corallus*	*C. felis*	*C. ibidis*	*C. gallinacea*	*C. muridarum*	*C. pecorum*	*C. pneumoniae*	*C. poikilothermia*	*C. psittaci*	*C. sanzinia*	*C. serpentis*	*C. suis*
CT105	CTL0360	1 / 1	1 / 1	2 / 0	4 / 2	1 / 1	4 / 2	1 / 2	2 / 0	1 / 1	3 / 3	2 / 2	4 / 2	2 / 1	1 / 1	2 / 1	2 / 2
CT042	CTL0298	1 / 1	1 / 1	1 / 1	1 / 1	1 / 1	1 / 1	1 / 1	1 / 1	1 / 1	1 / 1	1 / 1	1 / 1	1 / 1	1 / 1	1 / 1	1 / 0
CT089	CTL0344	1 / 1	1 / 1	1 / 1	1 / 1	1 / 1	1 / 1	1 / 1	1 / 1	1 / 1	1 / 1	1 / 1	1 / 1	1 / 1	1 / 1	1 / 1	1 / 1
CT142	CTL0397	1 / 1	1 / 1	1 / 1	1 / 1	1 / 1	1 / 1	1 / 1	1 / 1	1 / 1	1 / 1	1 / 1	1 / 1	1 / 1	1 / 1	1 / 1	1 / 1
CT143	CTL0398	1 / 1	1 / 1	1 / 1	1 / 1	1 / 1	1 / 1	1 / 1	1 / 1	1 / 1	1 / 1	2 / 2	1 / 1	1 / 1	1 / 1	1 / 1	1 / 1
CT144	CTL0399	1 / 1	1 / 1	1 / 1	1 / 1	1 / 1	1 / 1	1 / 1	1 / 1	1 / 1	1 / 1	1 / 1	1 / 1	1 / 1	1 / 1	1 / 1	1 / 1
CT156*	-	1 / 1*	0 / 0	0 / 0	0 / 0	0 / 0	0 / 0	0 / 0	0 / 0	0 / 0	0 / 0	0 / 0	0 / 0	0 / 0	0 / 0	0 / 0	0 / 0
CT163	CTL0419	1 / 1	0 / 0	0 / 0	0 / 0	0 / 0	0 / 0	0 / 0	0 / 0	2 / 0	0 / 0	0 / 0	0 / 0	0 / 0	0 / 0	0 / 0	3 / 3
CT456	CTL0716	1 / 1	1 / 1	1 / 1	1 / 1	1 / 1	1 / 1	1 / 1	1 / 1	1 / 1	1 / 1	1 / 1	1 / 1	1 / 1	2 / 1	1 / 1	2 / 1
CT473	CTL0734	1 / 1	1 / 1	1 / 1	1 / 1	1 / 1	1 / 1	1 / 1	0 / 1*	1 / 1	1 / 1	1 / 1	1 / 1	1 / 1	1 / 1	1 / 1	1 / 1
CT529	CTL0791	1 / 1	1 / 1	1 / 1	1 / 1	1 / 1	1 / 1	1 / 1	1 / 1	1 / 1	1 / 1	1 / 1	1 / 1	1 / 1	1 / 1	1 / 1	1 / 1
CT620	CTL0884	3 / 3	4 / 4	4 / 4	4 / 4	3 / 3	4 / 4	4 / 4	4 / 4	4 / 4	4 / 4	3 / 3	3 / 3	5 / 4	4 / 4	4 / 4	4 / 4
CT621	CTL0885	3 / 2	3 / 3	4 / 3	4 / 3	4 / 4	4 / 4	4 / 4	4 / 3	4 / 3	4 / 4	4 / 4	4 / 3	4 / 4	4 / 4	4 / 4	4 / 3
CT622	CTL0886	1 / 1	1 / 1	1 / 1	1 / 1	1 / 1	1 / 1	1 / 1	1 / 1	1 / 1	1 / 1	1 / 1	1 / 1	1 / 1	1 / 1	1 / 1	1 / 1
CT694	CTL0063	1 / 1	0 / 0	0 / 0	0 / 0	0 / 0	1 / 0	0 / 0	1 / 0	1 / 1	1 / 0	0 / 0	0 / 0	0 / 0	0 / 0	0 / 0	1 / 1
CT695	CTL0064	1 / 1	1 / 1	1 / 1	1 / 1	0 / 0	1 / 1	1 / 1	1 / 1	1 / 1	0 / 0	0 / 0	1 / 1	1 / 1	1 / 1	0 / 0	1 / 1
CT711	CTL0080	3 / 3	2 / 2	3 / 3	2 / 3	3 / 3	2 / 3	3 / 3	3 / 3	3 / 3	2 / 3	3 / 3	2 / 3	2 / 3	2 / 3	3 / 3	4 / 4
CT737	CTL0106	1 / 1	1 / 1	1 / 1	1 / 1	1 / 1	1 / 1	1 / 1	1 / 1	1 / 1	1 / 1	1 / 1	2 / 1	1 / 1	2 / 1	1 / 1	1 / 1
CT798	CTL0167	1 / 1	1 / 1	1 / 1	1 / 1	1 / 1	1 / 1	1 / 1	1 / 1	1 / 1	1 / 1	1 / 1	1 / 1	1 / 1	1 / 1	1 / 1	1 / 1
CT867	CTL0246	1 / 1	1 / 1	1 / 1	1 / 1	0 / 0	1 / 1	1 / 1	1 / 1	1 / 1	0 / 0	0 / 0	1 / 1	1 / 1	0 / 0	0 / 0	3 / 3
CT868	CTL0247	1 / 1	1 / 1	1 / 1	1 / 1	1 / 0	1 / 1	1 / 1	1 / 1	2 / 3	0 / 0	0 / 0	1 / 1	1 / 1	0 / 0	0 / 0	3 / 3
CT875	CTL0255	1 / 1	0 / 0	1 / 0	0 / 0	0 / 0	1 / 0	0 / 0	0 / 0	1 / 1	0 / 0	0 / 0	0 / 0	0 / 0	0 / 0	1 / 0	1 / 1

**ct156* was only found in *C. trachomatis*; although it is annotated in NCBI databases only for serovars A and D, we also found a putative ORF in serovar L2.

Table 3.2 (continuation) Analysis of the number of homologs of different non-Inc *C. trachomatis* proteins in other *Chlamydiae* by reciprocal best hit BLAST (found with protein BLAST / recovered with reciprocal BLAST).

		Chlamydia clade (CC)-IV			*Simkaniaceae*	*Criblamydiaceae*
Protein (serovar D)	**Protein (serovar L2)**	***C. bacterium K940_chlam9***	***C. bacterium K1000_chlam4***	***C. bacterium KR126_chlam2***	***S. negevensis***	***E. lausannensis***
CT105	CTL0360	0 / 0	0 / 0	0 / 0	0 / 0	0 / 0
CT042	CTL0298	1 / 1	1 / 0	1 / 1	1 / 1	1 / 1
CT089	CTL0344	0 / 0	0 / 0	0 / 0	1 / 1	1 / 1
CT142	CTL0397	0 / 0	0 / 0	0 / 0	0 / 0	0 / 0
CT143	CTL0398	0 / 0	0 / 0	0 / 0	0 / 0	0 / 0
CT144	CTL0399	0 / 0	0 / 0	1 / 0	0 / 0	0 / 0
CT156*	-	0 / 0	0 / 0	0 / 0	0 / 0	0 / 0
CT163	CTL0419	1 / 0	0 / 0	1 / 0	0 / 0	0 / 0
CT456	CTL0716	0 / 0	0 / 0	0 / 0	0 / 0	0 / 0
CT473	CTL0734	2 / 1	0 / 0	1 / 0	1 / 0	1 / 1
CT529	CTL0791	1 / 0	0 / 0	1 / 0	0 / 0	0 / 0
CT620	CTL0884	1 / 1	0 / 0	1 / 0	0 / 0	0 / 0
CT621	CTL0885	1 / 1	0 / 0	0 / 0	0 / 0	0 / 0
CT622	CTL0886	0 / 0	0 / 0	0 / 0	0 / 0	0 / 0
CT694	CTL0063	0 / 0	1 / 0	0 / 0	0 / 0	0 / 0
CT695	CTL0064	1 / 0	0 / 0	0 / 0	0 / 0	0 / 0
CT711	CTL0080	1 / 1	0 / 0	0 / 0	0 / 0	0 / 0
CT737	CTL0106	2 / 1	1 / 1	1 / 1	2 / 2	3 / 1
CT798	CTL0167	1 / 0	1 / 0	1 / 1	1 / 1	2 / 1
CT867	CTL0246	0 / 0	0 / 0	0 / 0	1 / 0	0 / 0
CT868	CTL0247	0 / 0	0 / 0	0 / 0	0 / 0	0 / 0
CT875	CTL0255	1 / 0	0 / 0	1 / 0	0 / 0	0 / 0

**ct156* was only found in *C. trachomatis*; although it is annotated in NCBI databases only for serovars A and D, we also found a putative ORF in serovar L2.

3.4.5 Conclusions

To summarize, we identified several putative homologs of CteG which are only present among species of the Family *Chlamydiaceae*. When comparing the evolution of CteG and its homologs with the evolution of their corresponding chlamydial species, we observed that they were generally consistent. However, the existence of paralogs in some species but not in others suggests gene duplications events resulting in genes that were maintained in some species but not in others. As such, while many of the CteG homologs among *Chlamydiaceae* are CteG orthologs, several others are evolutionary related to CteG but did not evolve by speciation. By analyzing the genomic vicinities of *cteG* and of its putative homologs, we found that they are grouped in two distinct sets of conserved genomic regions. Overall, the preservation of these genes among these *Chlamydiaceae* species indicates that they may have important functions for their virulence. Finally, we found that CteG is, among other analyzed chlamydial non-Inc T3S substrates, the one that presents the most variable number of putative homologs.

4

DISCUSSION AND CONCLUSIONS[8]

The general goal of this project was to understand how *C. trachomatis* subverts host cells during infection, by further characterizing the previously found T3S effector CteG (Pais *et al.*, 2019). Using transfection vectors encoding different combinations including the first 100 amino acid residues of CteG fused to mEGFP, we found that the first 20 residues of CteG are crucial for the localization of that hybrid protein (mEGFP-CteG$_{1\text{-}100}$) at the Golgi in transfected mammalian cells. We also found that specific N-terminal amino acid residues localizing at a putative α-helix are essential for the localization of ectopically expressed mEGFP-CteG$_{1\text{-}100}$ at the Golgi. Additionally, bioinformatics suggested that S-palmitoylation of CteG could play a role in targeting mEGFP-CteG$_{FL}$ to the Golgi, but this was not confirmed by using a broad inhibitor of S-palmitoylation [2-Bromopalmitate (2BP)]. The amino acids residues that are essential for subcellular localization of mEGFP-CteG$_{1\text{-}100}$ at the Golgi appeared to be important, but not essential for the adequate localization of ectopically expressed full-length CteG (mEGFP-CteG$_{FL}$) at the mammalian cell Golgi and plasma membrane. Likewise, these residues are important, but not essential for the localization of CteG at those two compartments during *C. trachomatis* infection. Overall, these results deepened our knowledge about the determinants of the subcellular localization of CteG and suggested that contrary to what was suggested by transfection experiments, the first 100 amino acid residues of CteG do not correspond to a major Golgi targeting region of the full-length protein.

When analyzing a previously generated *C. trachomatis cteG* mutant strain, we found that CteG is involved in chlamydial lytic exit from host cells (Pereira *et al.*, 2022). Moreover,

[8] This section was written by Inês Serrano Pereira, based on the data obtained and on the cited bibliographic references. Parts of this Discussion were transcribed from Pereira *et al.* [(Pereira *et al.*, 2022)], a recent publication that includes data from this PhD thesis.

we observed that plasmid encoded Pgp4, which has also been described to mediate the lytic exit of *C. trachomatis* (Yang *et al.*, 2015), does not regulate the production or the subcellular localization of CteG. However, CteG and Pgp4 participate in the same cascade of events leading to lysis of host cells and release of the chlamydiae. Additionally, the defect of a *pgp4* mutant *C. trachomatis* strain in host cell cytotoxicity could be rescued by the overproduction of CteG from a plasmid. Overall, we discovered a novel role for CteG in mediating *C. trachomatis* host cell lytic exit (Pereira *et al.*, 2022), an essential step of the *C. trachomatis* infectious cycle.

Another objective of this project was to identify possible CteG host cell interacting partners. However, the results obtained by mass spectrometry of extracts of HeLa cells ectopically expressing full-length CteG fused to mEGFP (mEGFP-CteG$_{FL}$) still need further investigation and possible validation. Moreover, we studied the distribution of CteG homologs among *Chlamydiaceae*. The results obtained provided further insights about the evolutionary history of CteG and set the basis for an ongoing experimental analysis of the homologs of CteG.

4.1 Identifying the determinants of the subcellular localization of CteG during *C. trachomatis* infection

4.1.1 Pinpointing the amino acids that mediate the localization of CteG at the Golgi and plasma membrane of host cells

In previous studies, we have shown that ectopically expressed mEGFP-CteG$_{FL}$ localizes mostly at the cell plasma membrane, but also at the Golgi in some cells (Pais *et al.*, 2019). Moreover, we also showed that the first 100 amino acid residues of CteG contain a Golgi-targeting region (Pais *et al.*, 2019), as observed in transfected cells. Here, we found that the first 20 N-terminal amino acid residues of CteG contain specific residues at a putative α-helix which are essential for the localization of mEGFP-CteG$_{1\text{-}100}$ at the Golgi of transfected cells. This contrasts with other proteins that are targeted to or interact with the Golgi through their C-terminal or more central regions such as *L. pneumophila* GobX (Lin *et al.*, 2015) and SdhA (Ge *et al.*, 2012), or *S. enterica* serovar Typhimurium SseF and SseG (Salcedo, 2003; Deiwick *et al.*, 2006). Curiously, our previous studies indicate that the secretion signal of CteG recognized by the T3S system of *Y. enterocolitica* is precisely contained within its first 20 residues (da Cunha *et al.*, 2014; Pais, 2018). One could reason that the sequence of CteG that possibly includes the T3S system signal also mediates targeting to the Golgi. For example, the *Brucella abortus* type 4 secretion (T4S) system effector BspB contains an N-terminal transmembrane motif essential for its targeting to the Golgi upon ectopic expression and possibly for its stability and T4S system delivery

during infection (Miller *et al.*, 2017). Nevertheless, we also found that the mutations affecting the localization of mEGFP-CteG$_{1\text{-}100}$ at the Golgi produced a different effect when introduced in mEGFP-CteG$_{FL}$. Ectopically expressed mEGFP-CteG hybrid proteins lacking the first 20 amino acid residues of CteG or with specific amino acid replacements in this region had their Golgi and plasma membrane localization slightly affected, but their targeting to the Golgi was not completely blocked. This indicates that the region and/or amino acid residues targeting mEGFP-CteG$_{1\text{-}100}$ to the Golgi may also play a role in directing mEGFP-CteG$_{FL}$ to this organelle but other regions of CteG are likely involved and have a more prominent role in this process.

A prediction of the tertiary structure of CteG has recently become available (AlphaFold; DeepMind, EMBL-EBI; see Figure 4.2A below). The predicted model still has very low confidence scores but corroborates the presence of a putative α-helix in the first 20 amino acid residues of CteG and suggests that its N-terminal region (approximately residues 21 to 138) is unstructured (see Figure 4.2A below). Conversely, other parts of the amino acid sequence of CteG have predicted helical structures that could be of amphipathic nature and aid in the association of CteG with host cell membranes. It is therefore possible that a peptide comprising only the first 100 residues of CteG is targeted to the Golgi via the predicted α-helix because no other putative interacting regions are available, and that other regions are involved in targeting native CteG to the Golgi.

We reasoned that the determinants of the localization of CteG at the plasma membrane should localize along the amino acid sequence of CteG that we denominated the C-terminal region (amino acids 350 to 656; see section 3.1 above). We analyzed ectopically expressed mEGFP-CteG proteins with truncations at this region of CteG and verified that they were less frequently detected at the plasma membrane by comparison with mEGFP-CteG$_{FL}$. However, many truncations were not stable and/or were poorly expressed even when only the last 28 C-terminal amino acid residues of CteG were missing. This suggests that the determinants of the localization of CteG at the plasma membrane may be dispersed along the C-terminal region or even along the entire amino acid sequence of CteG. The *C. trachomatis* T3S effector TmeA has also been described to be targeted to the plasma membrane upon ectopic expression (Hower *et al.*, 2009) by a membrane localization domain (MLD) at its N-terminal region [specifically within residues 40 to 80; (Bullock *et al.*, 2012)]. This MLD is also present in the N-terminal regions of the T3S effectors YopE from *Yersinia* and ExoS from *P. aeruginosa* (Krall *et al.*, 2004), which are Rho GTPase-activating proteins (RhoGAPs) for which a plasma membrane localization is essential (Black and Bliska, 2000; Zhang *et al.*, 2007; Isaksson *et al.*, 2009). Interestingly, there is no significant amino acid similarity between the overall amino acid sequence of TmeA and of YopE or ExoS and the same applies for their MLDs (Bullock *et al.*, 2012). Therefore, CteG, which does not possess similarity to other known proteins, may also

encode a MLD that could potentially be important for its localization and for its yet poorly studied role during infection (see section 4.2 below).

We have previously observed that a CteG protein lacking its first 20 residues produced by *C. trachomatis* remained within the chlamydiae and was not detected at the cytoplasm of infected HeLa cells, possibly because of the absence of a T3S signal [data not published; (Pais, 2018)]. Therefore, we could only analyze CteG proteins harboring specific amino acid replacements within the first 20 residues of CteG. Curiously, the substitution of five specific amino acid residues was sufficient to affect the T3S system delivery of CteG, suggesting that the T3S signal was possibly affected. Conversely, the substitution of only two of those residues (C_9 and C_{17}) did not dampen the T3S delivery of CteG but caused a slight decreased degree of the localization of CteG at the Golgi and plasma membrane of infected cells at 24 and 40 h post-infection, respectively. In the future, to pinpoint the determinants of the subcellular localization of CteG during infection, other amino acid replacements that affect the localization of CteG but not its T3S system-mediated delivery should be tested.

Overall, the determinants of the localization of CteG at the Golgi or at the plasma membrane could be fine-tuned by applying other deletions and/or amino acid substitutions on CteG, and by analyzing these mutants in the context of infection by *C. trachomatis*.

4.1.2 The possible role of S-palmitoylation in targeting CteG to eukaryotic membranes

The affinity of certain proteins to eukaryotic membranes may be increased by the posttranslational addition of lipids to specific amino acid residues. S-palmitoylation is one of the mechanisms of lipid attachment to proteins (Resh, 2006; Sobocińska *et al.*, 2018), and has been shown to occur in different bacterial effector proteins. The already mentioned *L. pneumophila* GobX is a protein with E3 ubiquitin ligase activity and has a predicted α-helix at its C-terminal region (Lin *et al.*, 2015). The hydrophobic face of this α-helix mediates insertion of GobX into the Golgi membrane, a process that is simultaneously enhanced by S-palmitoylation of a specific cysteine residue (Lin *et al.*, 2015). LpdA is a *L. pneumophila* protein with phospholipase D activity that is S-palmitoylated in five cysteine residues at its C-terminal region (Schroeder *et al.*, 2015). In transfected cells, this promotes LpdA targeting to the plasma membrane where it hydrolyzes eukaryotic lipids to generate phosphatidic acid (Schroeder *et al.*, 2015). Moreover, *S. enterica* Typhimurium T3S system effectors SspH2, which belongs to the family of E3 ubiquitin ligases, and SseI, which is involved in directed migration of cells infected by *Salmonella*, are also S-palmitoylated at their N-terminal regions (Hicks *et al.*, 2011). This promotes their targeting to the plasma membrane of cells infected by *Salmonella* and is crucial for the effector activity of SseI (Hicks *et al.*, 2011). We hypothesized that CteG could

also be S-palmitoylated because of its dual localization at the host cell Golgi and plasma membranes during infection and due to the presence of putative S-palmitoylation sites detected with CSS-Palm [Annexes Figure 7; (Ren *et al.*, 2008)]. Nevertheless, there were no perceptible differences in the localization of ectopically expressed mEGFP-CteG$_{FL}$ when using 2BP, an inhibitor of eukaryotic palmitoyltransferases. There was however a slight effect on the localization of mEGFP-CteG$_{1-100}$ at the Golgi in the presence of 2BP, suggesting that although S-palmitoylation might occur it is not the main mechanism targeting CteG to the Golgi and plasma membrane. In the future, specific techniques could be applied to determine whether CteG is S-palmitoylated and targeted by this mechanism to eukaryotic membranes. Labelling of CteG with a palmitic acid analogue [as 17-octadecynoic acid (ODYA), or alkynyl-16] followed by click-chemistry (Hicks *et al.*, 2011; Lin *et al.*, 2015; Schroeder *et al.*, 2015), or the use of acyl-resin-assisted capture (acyl-RAC) assays (Forrester *et al.*, 2011) are possible methods to assess these aspects.

4.2 *C. trachomatis* lytic exit from host cells is CteG-dependent

4.2.1 CteG is a novel protein involved in the lytic exit of *C. trachomatis* from host cells

Most studies on chlamydial effectors have focused on those involved in the initial steps of host cell infection and on Incs (Bugalhão and Mota, 2019), including those that mediate exit by extrusion (Lutter *et al.*, 2013; Nguyen *et al.*, 2018; Shaw *et al.*, 2018). In this study, we found that the T3S effector CteG (Pais *et al.*, 2019) is involved in the lytic exit of *C. trachomatis* from host cells. This is one of the first chlamydial T3S effectors shown to be involved in this process, thus filling the gap of the previously proposed link between the *C. trachomatis* virulence plasmid and its T3S system (Yang *et al.*, 2015) in mediating this essential step of the chlamydial infectious cycle. Recently, simultaneously to our publication of the role of CteG on the lytic exit of *C. trachomatis* (Pereira *et al.*, 2022), Inc CT135/CTL0390 has also been shown to mediate this process via induction of the innate immune cGAS-STING pathway (Bishop and Derré, 2022). Our work, together with previous related studies (Hybiske and Stephens, 2007; Yang *et al.*, 2015), indicates that similar to chlamydial egress by extrusion of the inclusion (Lutter *et al.*, 2013; Nguyen *et al.*, 2018; Shaw *et al.*, 2018; Zuck and Hybiske, 2019), *C. trachomatis* lytic exit involves different chlamydial players [CteG, Pgp4, at least one of the several Pgp4-regulated genes (Song *et al.*, 2013; Patton *et al.*, 2018), CPAF and Inc CT135/CTL0390 (Bishop and Derré, 2022)], likely host cell factors and different layers of regulation.

A limitation of our study is that the *cteG::aadA* mutant strain is not isogenic to the parental L2/434 strain and displays a slight growth defect that is CteG-independent. Comparing to the L2/434 strain, and besides the inactivation of *cteG*, we identified six nucleotide differences in the CteG-deficient strain leading to four missense mutations and two alterations in non-coding regions. As detailed in Annexes Table 4, the missense mutations are in a putative lipoprotein (CT734), in a putative integral membrane protein (CT853), and in two known virulence proteins: a T3S chaperone (CT862/LcrH/Scc3) (Fields *et al.*, 2005), and an Inc (CT618) (Bugalhão and Mota, 2019). However, as the defect of the CteG-deficient strain in promoting host cell lysis can be complemented, this confirms the role of CteG in this process. It is presently unclear which of the referred mutations leads to the slight growth defect of the *cteG* mutant strain. To assess this, *C. trachomatis cteG::aadA*-derived strains overproducing from plasmids the possibly affected proteins either individually or in combinations should be constructed and their progeny generation ability assessed.

4.2.2 Possible timing and mode of action of CteG during infection by *C. trachomatis*

C. trachomatis late-stage host cell lysis is a bacterially induced process that initiates with rupture of the inclusion membrane in a chlamydial-dependent manner and culminates with destruction of the host plasma membrane (Hybiske and Stephens, 2007; Yang *et al.*, 2015; Kerr *et al.*, 2017), a mechanism used by other pathogens (Andreadaki *et al.*, 2018; Flieger *et al.*, 2018). Therefore, CteG could promote lysis of the inclusion or host cell plasma membranes either directly or indirectly by activating other chlamydial proteins and/or host cell factors. However, at the present, we are not able to distinguish between these different possibilities. As CteG concentrates at the plasma membrane at late stages of infection, a possible action on the integrity of the plasma membrane to mediate host cell lysis would appear more likely. On the other hand, it has been shown that laser-mediated rupture of the chlamydial inclusion leads to a rapid necrotic cell death-dependent pathway that appears to be mostly mediated by the host (Kerr *et al.*, 2017). This was proposed due to the inefficacy of bacterial protein synthesis inhibitors added at 24 h post-infection to prevent host cell lysis after laser-mediated inclusion rupture (Kerr *et al.*, 2017). However, bacterial proteins such as CteG, already present in the host cell cytoplasm at 24 h post-infection, could still mediate plasma membrane rupture if activated or relieved from inhibition upon inclusion rupture.

To clarify the afore-mentioned aspects, it will be crucial to determine host cell proteins CteG might interact with and to elucidate if it has an enzymatic activity. For example, cysteine proteases have been proposed to mediate chlamydial lytic exit by promoting rupture of the inclusion membrane (Hybiske and Stephens, 2007). While CteG does not seem to possess in

its amino acid sequence any consensus motif characteristic of such proteases, it could have a yet undescribed cysteine protease domain. Furthermore, it is also conceivable that CteG could activate a host cell protease such as calpains, which are cysteine proteases that have been suggested to be involved in inclusion membrane rupture (Kerr *et al.*, 2017). In line with a role of CteG in interfering with vesicular trafficking in *S. cerevisiae* (Pais *et al.*, 2019), it is also possible that CteG modifies Rho GTPases as described for other bacterial effectors (Fu and Galán, 1999; Pederson *et al.*, 1999; Black and Bliska, 2000) to alter the integrity of the cytoskeleton or the composition of the plasma membrane of host cells to promote lysis.

Some intracellular pathogens exit host cells by inducing active membrane destruction in a process unrelated to programmed host cell death (Kafsack *et al.*, 2009; Glushakova *et al.*, 2018). Other pathogens exploit programmed cell death mechanisms such as pyroptosis or necroptosis to exit the intracellular niche [(Lindgren *et al.*, 1996; Uwamahoro *et al.*, 2014; Dallenga *et al.*, 2017); reviewed in (Flieger *et al.*, 2018)]. A possible role of these forms of programmed cell death in the exit of *C. trachomatis* is still unexplored, although it is known that this bacterium interacts with such pathways during its developmental cycle. For example, an extensive number of reports suggests that *C. trachomatis* inhibits apoptosis to avoid clearance and maintain host cell viability [a feature not observed in other *Chlamydiae* such as *W. chondrophila* (Dille *et al.*, 2015)], although the resistance of *C. trachomatis* to pro-apoptotic insults is still a matter of discussion [reviewed in (Sixt, 2021)]. Recently, as mentioned above, Inc CT135/CTL0390 has been shown to interact with the cGAS-STING pathway to promote lysis of host cells (Bishop and Derré, 2022). Although the downstream events leading to host cell lysis are still unknown, it is possible that STING activation leads to a type I interferon response or to the activation of the NF-kB pathway, autophagy, or the afore-mentioned programmed cell death pathways (Sixt, 2021; Bishop and Derré, 2022). Similarly, CteG could directly or indirectly manipulate such pathways to promote *C. trachomatis* host cell lytic exit. In general, all these possibilities should be considered in future experiments to investigate the mechanism of action of CteG in promoting host cell lysis.

4.2.3 Host cell lytic exit among different *Chlamydiae*

Reports indicate that *C. trachomatis* LGV strains (such as the strain we used in our study) and *C. muridarum* are prone to be more lytic than ocular and urogenital *C. trachomatis* strains that would exit predominantly by extrusion (Todd and Caldwell, 1985; Yang *et al.*, 2015). In contrast, the rate of host lysis was observed to be identical in cells infected by LGV or urogenital strains in another study (Hybiske and Stephens, 2007). While chlamydial lytic exit should enable a rapid re-infection of host cells, exit by extrusion should facilitate dissemination and the release of chlamydiae that remain infectious for longer (Zuck *et al.*,

2017). To understand the advantages of each egress pathway, it would be interesting to correlate possible differences in these processes between *C. trachomatis* serovars or amongst *Chlamydia* species to the genetic and transcriptomic variability we previously described for *cteG* (Pais *et al.*, 2019). Furthermore, it remains unknown if the rate of extrusion is altered in CteG-deficient *C. trachomatis* or if the rate of lytic exit is affected in chlamydiae deficient for Incs (CT228 and MrcA) that mediate extrusion (Lutter *et al.*, 2013; Nguyen *et al.*, 2018; Shaw *et al.*, 2018). Future clarification of these aspects could also help understanding the advantages of both exit pathways and how are they controlled by *C. trachomatis*.

Host cell lytic exit is conserved in *Chlamydiae* as this was observed for the *Chlamydia*-like microorganisms and emerging pathogens *S. negevensis* (Koch *et al.*, 2020) and *W. chondrophila* (Goy *et al.*, 2008; Dille *et al.*, 2015). However, a simple BLAST analysis indicates there is no CteG or even CT135 (Bishop and Derré, 2022) homologs in the genomes of *S. negevensis* or *W. chondrophila*. The infectious cycle of *S. negevensis* takes much longer than the one of *C. trachomatis*, with infectious forms appearing only about three days after infection (Kahane *et al.*, 2002). In turn, *W. chondrophila* displays a marked cytotoxic effect when infecting macrophage and different types of epithelial cell lines, with an increase of 2.5-3 logs in the number of bacteria per cell in the first 24 h of infection (Goy *et al.*, 2008; de Barsy and Greub, 2013). This cytotoxic effect of *W. chondrophila* is more pronounced than that of *C. trachomatis* during infection of HeLa cells (Dille *et al.*, 2015). Altogether, these observations suggest that *C. trachomatis*, *S. negevensis* and *W. chondrophila* evolved different host cell lytic exit mechanisms.

4.2.4 Model of action of CteG during *C. trachomatis* infection

While the mechanistic details need to be experimentally tested and other possibilities for how CteG mediates host cell lysis might be envisioned, we propose a hypothetical model in which generation of a CteG protein capable of mediating host cell lysis requires Pgp4-dependent activation, a step that can be suppressed by overproduction of CteG (Figure 4.1). As *cteG* expression is not regulated by Pgp4 [this work, (Pereira *et al.*, 2022); (Carlson *et al.*, 2008; Song *et al.*, 2013; Patton *et al.*, 2018)], which we confirmed also at the level of CteG production and localization in infected host cells, the product of at least one Pgp4-regulated gene should be involved in this activation of CteG. At the present, we cannot discriminate whether the hypothetical Pgp4-dependent activation of CteG occurs within chlamydiae or after delivery of CteG into the host cell cytoplasm. In any event, as CteG is type III-delivered into the cytoplasm of host cells from about 16-20 h post-infection and as host cell lysis is only detected from 48 h post-infection, there should be an inhibitor in infected host cells that prevents CteG-mediated host cell lysis until later in the chlamydial infectious cycle (Figure 4.1). This inhibitor could be another chlamydial effector or a host cell factor. EUO is a repressor of chlamydial late genes

whose activity has been shown to be enhanced by Pgp4 (Zhang *et al.*, 2020). It is tempting to speculate that Pgp4 regulates the expression of *cteG* indirectly through EUO. However, while it has been shown that the mRNA levels of *cteG* peak at 2 h post-infection [(Pais, 2018; Pais *et al.*, 2019); this work], our model suggests that Pgp4 possibly activates CteG in a mechanism upstream the potential inhibition of CteG (Figure 4.1). Moreover, if Pgp4 promoted the repression of *cteG* via EUO, a *C. trachomatis pgp4* mutant strain should have increased cytotoxicity potential, and our and previous (Yang *et al.*, 2015) data suggest the opposite. Therefore, if our model is proven to be valid, EUO should not be the factor that inhibits premature CteG-dependent host cell lysis. CPAF, another *C. trachomatis* protein that has been shown to be involved in chlamydial lytic exit (Yang *et al.*, 2015), could cleave the hypothetical inhibitor of CteG, thus liberating CteG to exert its lytic action. However, there is conflicting evidence on whether CPAF is secreted into the host cell cytoplasm during the chlamydial infectious cycle or if it is only released after inclusion rupture later in the cycle (Prusty *et al.*, 2018; Bugalhão and Mota, 2019). Furthermore, inhibition of CPAF activity after laser-mediated rupture of the chlamydial inclusion membrane does not prevent subsequent host cell lysis (Kerr *et al.*, 2017). Therefore, how and if CPAF might contribute to *C. trachomatis* lytic exit mediated by CteG remains to be directly analyzed.

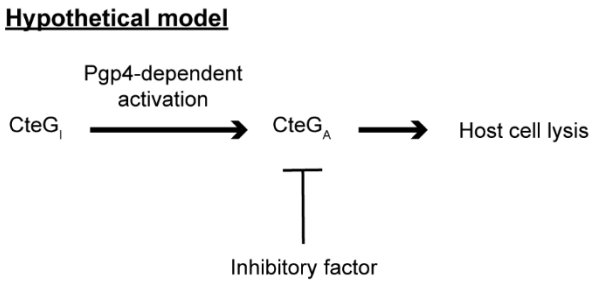


Figure 4.1. Hypothetical model for the mode of action of CteG and Pgp4 in promoting host cell lysis. After CteG is produced by chamydiae in an inactive form (CteG$_I$), CteG is activated (CteG$_A$) in a Pgp4-dependent manner, which could occur within chlamydiae or after delivery of CteG into the host cell cytoplasm. Premature CteG$_A$-mediated host cell lysis is prevented by the action of an unknown inhibitory factor, which could be another chlamydial effector or a host cell factor. At late stages of infection, the effect of this inhibitor in CteG$_A$ is alleviated by an unknown mechanism and host cell lysis is triggered.

4.3 Identifying CteG host cell interacting partners: other possible roles of CteG during infection

We attempted to identify possible host cell interacting partners of CteG by mass spectrometry using lysates of cells ectopically producing mEGFP-CteG$_{FL}$. However, we considered the results inconclusive as no obvious specific interaction was found, and did not proceed with

the validation of any of the identified proteins as interacting partners of CteG. However, by using an equivalent procedure, a recent unpublished study has shown that a CteG protein fused to a FLAG tag interacts with the centrosomal protein centrin-2 [bioRxiv; (Steiert *et al.*, 2022)], which was also recovered in our analysis (Annexes Table 5). Centrin-2 is a component of the host cell centrosomes and a regulator of centriole duplication and was shown to interact with the C-terminal region of CteG, specifically its 17 last residues [bioRxiv; (Steiert *et al.*, 2022)]. Moreover, it was concluded that CteG promotes centrosome amplification but is not involved in host cell multinucleation or in centrosome positioning [bioRxiv; (Steiert *et al.*, 2022)]. Therefore, besides the role we propose for CteG in chlamydial host cell lytic exit (see section 3.2 above), this report describes a novel function of CteG in targeting and modulating host cell centrins, being the first bacterial effector to have such function. The ability of a single effector to perform different functions is not surprising, as previous studies have suggested that different chlamydial proteins have the potential to interact with a multitude of host cell proteins (Mirrashidi *et al.*, 2015). Moreover, specific chlamydial proteins were shown to be involved in different processes and/or in recruiting different organelles to the inclusion, such as the Incs CT192/Dynactin Recruiting Effector 1 (Dre1) [bioRxiv; (Sherry *et al.*, 2022)], CT223/IPAM (Alzhanov *et al.*, 2009; Dumoux *et al.*, 2015) or CT229/CpoS (Sixt *et al.*, 2017; Weber *et al.*, 2017; Faris *et al.*, 2019). This could also be the case of CteG, which was a recently identified T3S effector (Pais *et al.*, 2019) and for which many features remain to be elucidated. Importantly, this was a valuable reminder that our mass spectrometry results may include potential host cell interacting partners of CteG. Therefore, their interactions should be analyzed more carefully by performing co-IP or protein-protein interaction assays. Furthermore, in the host laboratory CteG interacting partners were also searched by using a yeast two-hybrid screen. Several hits were found that remain to be validated. In summary, although we did not identify a host cell protein interacting with CteG, several putative interactors were found and remain to be further analyzed.

Interestingly, a possible interaction of CteG with the host cell lipid phosphatidylinositol 3-phosphate [PI(3)P] was also identified in the host laboratory while supervising the mini-project of an undergraduate student (Condez *et al.*, unpublished data). Phosphatidylinositol phosphates (PIPs) are key elements in vesicular trafficking processes and PI(3)P is the most predominant species of those molecules found on early endosomes (Marat and Haucke, 2016). Pathogens as *L. pneumophila* and *Francisella* are known to use effectors to modulate the host cell membrane trafficking by directly binding PI(3)P (Nachmias *et al.*, 2019; Pike *et al.*, 2019) or by mimicking phosphatidylinositol 3-kinases or phosphatases (Ledvina *et al.*, 2018; Hsieh *et al.*, 2021), to block interference of the host cell with the pathogen-containing vacuole in the case of *Legionella* or to promote escape into the host cell cytoplasm in the case of *Francisella*. Although an interaction between CteG and PI(3)P needs to be further validated, it is

reasonable that CteG could interact with such molecule or with other host cell lipids to alter host membrane dynamics, possibly leading to host cell lysis and *C. trachomatis* egress, or to modulate host cell vesicular trafficking.

4.4 Conservation of *cteG* among *Chlamydiaceae*

The conservation of a given gene in a group of related organisms is suggestive of its importance along the evolution of these species. As already mentioned, T3S system genes are highly conserved among bacteria which is elusive of the importance of this system during bacterial evolution. Despite associated to a higher variability, T3S system-delivered substrates are crucial components of the T3S machinery, many of them being recognized effectors that modulate specific host cell processes. Therefore, understanding their conservation among related species is key to understand how pathogens evolve their virulence mechanisms according to the niche or organism where they survive and multiply. Based on this and on the important role we identified for CteG in *C. trachomatis* host cell lytic exit, we decided to evaluate the conservation of *cteG* among different species of the Phylum *Chlamydiae*. In many of the studied species, more than one CteG homolog was identified, and the number of homologs was heterogeneous between species from the *Chlamydiaceae*. This was different from some of the other analyzed chlamydial non-Inc T3S substrates, for which more than one homolog was identified but this number was maintained between species. Moreover, species as *C. avium* and *C. gallinacea* seem not to harbor a gene homolog to c*teG*, indicating it might have been lost or, if still present, it diverged significantly from *cteG*. These observations suggest that the evolutionary history of *cteG* was likely marked by the occurrence of duplications (originating paralogs) and/or gene loss events, which is also supported by our phylogenetic analysis of CteG and its homologs (see section 3.4.2 above). This also indicates that only some of the identified proteins may be orthologs of CteG.

We did not find *cteG* homologs in the analyzed *Simkaniaceae* and C*riblamydiaceae* species or in a group of species closely related to the *Chlamydiaceae* (Dharamshi *et al.*, 2020). This is consistent with the acquisition of *cteG* by an ancestral species that originated chlamydiae pathogenic to animals. Conversely, genes that share homology with *cteG* and that could be its orthologs are present among most species of the Family *Chlamydiaceae*, which suggests that at least during part of the evolution of these species CteG was and still seems to be an important virulence factor. Intriguingly, *C. gallinacea*, which does not seem to possess a *cteG* homolog, has been shown to encode a lower number of homologs of virulence-associated genes and is associated to lower mortality rates (Heijne *et al.*, 2021). Within this scope, it would be interesting to investigate whether the role of CteG in chlamydial lytic exit is also conserved in

its identified putative homologs and, if so, correlate the maintenance of *cteG* among the respective species with their higher or lower virulence potential comparatively with the species where CteG might have lost function or play a different role. However, the information we currently have about the biological role of CteG is still limited and needs to be further investigated.

When performing a simple protein BLAST between CteG and each of its identified homologs, we verified that parts of the amino acid sequences of these homologs share amino acid similarities exclusively with the C-terminal region of CteG. This observation suggests that this region may be important for the functions these proteins may exert in their biological context, which, at the present, are still unknown or poorly studied in the case of CteG. On the other hand, the N-terminal region of CteG is not conserved among its homologs. If this region of CteG comprises a T3S signal as already mentioned (da Cunha *et al.*, 2014), and if these homologs function as effectors, they either possess entirely different T3S signals or they are delivered by other unknown mechanisms. Alternatively, they could have lost the ability to be T3S-delivered and no longer function as effectors.

Whether some of the identified proteins are orthologs of CteG, meaning that they exert a similar function in their native organisms, needs to be investigated and experimentally confirmed. Curiously, AlphaFold (DeepMind, EMBL-EBI) tertiary structure predictions show that as CteG (Figure 4.2A), the N-terminal region of many of these homologs is highly unstructured. This is the example of *C. muridarum* TC_0381 (Figure 4.2B) or *C. felis* CF0619 (Figure 4.2C). While these predictions support the idea that some of the identified homologs may have a conformation similar to CteG and could thus have similar roles, they should be considered cautiously due to their low confidence levels and need of improvement.

While analyzing the conservation of the genomic region of *cteG* and its homologs, it was perceptible that the *Chlamydiaceae* species have experienced genomic rearrangements throughout evolution as it should be expected. Examples of this are: 1) the existence of two distinct groups of species based on the conservation of their genomic regions, which was consistent with their phylogenetic relationships; 2) the fact that the putative homologs of *cteG* in *C. ibidis* lay at different locations of its genome and are one-sided flanked by genomic regions that resemble those of the *cteG* homologs in other species. Curiously, *C. avium* no longer encodes a *cteG* homolog in the analyzed genomic region which supports loss of this gene. Nevertheless, it is prudent to consider that homologs of some or all the identified genes, including *cteG*, could be present in all species but in other genomic locations. In the future, a more detailed analysis about how the genomic regions of *cteG* and of its putative homologs rearranged, and how the genes there encoded evolved should provide a better understanding of how the corresponding species are related.

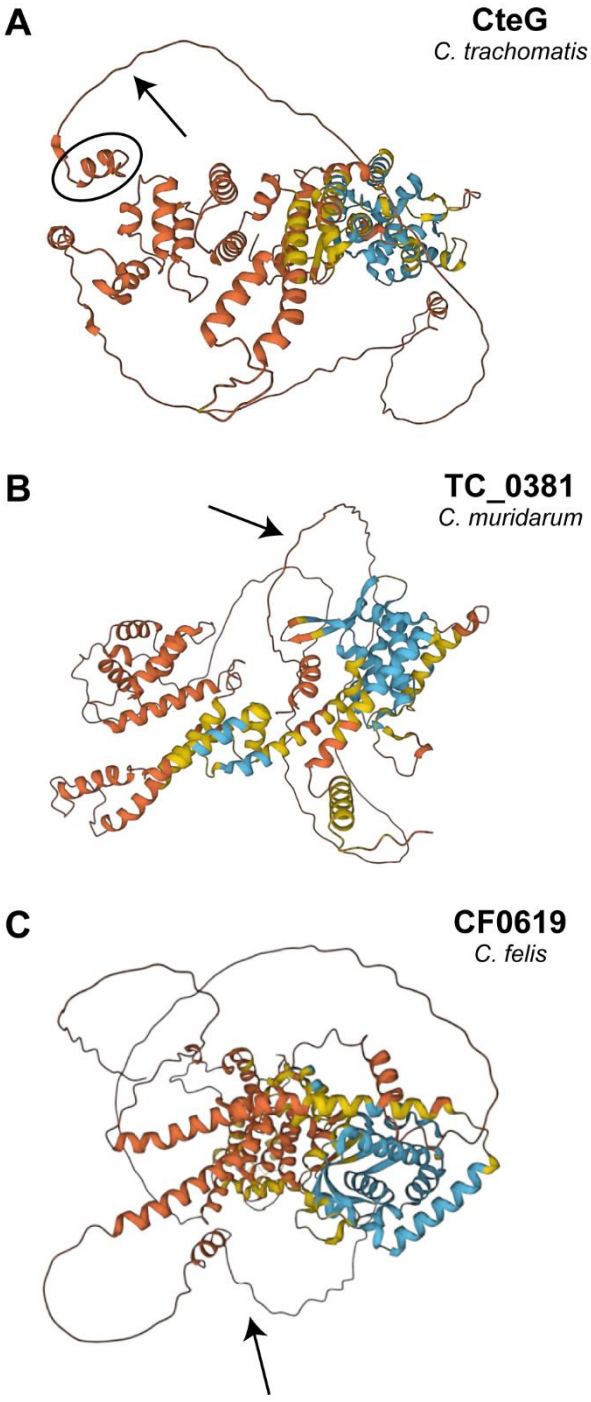


Figure 4.2. Predicted tertiary structure of CteG and of two of its putative homologs. Models of the tertiary structure of **(A)** *C. trachomatis* CteG (https://alphafold.ebi.ac.uk/entry/A0A0H3MKJ7), **(B)** *C. muridarum* TC_0381 (https://alphafold.ebi.ac.uk/entry/Q9PKT1) and **(C)** *C. felis* CF0619 (https://alphafold.ebi.ac.uk/entry/Q253Z7) predicted with AlphaFold (DeepMind, EMBL-EBI). Structures were retrieved in a perspective enabling the visualization of similar three-dimensional features and of the disorganized N-terminal region of each protein (represented by arrows). In (A), the black circle indicates the presence of a putative α-helix in the first 20 amino acid residues of CteG. Colors represent the confidence of the model: orange, very low; yellow, low; blue, confident.

Overall, many of these considerations are hypothetical. Most of them still need to be thoroughly investigated and lack experimental support to be considered as valid or not. Nevertheless, the analysis we performed has broadened our knowledge about the conservation of *cteG* among the Phylum *Chlamydiae* and will support our research for the characterization of CteG.

4.5 Future directions and concluding remarks

Understanding how *C. trachomatis* effectors mediate the different stages of the chlamydial infectious cycle is key to understand how this pathogen survives and replicates within host cells. In this work, we describe a new function for the *C. trachomatis* T3S effector CteG in mediating host cell lysis and concomitant chlamydial exit, which is a critical aspect of the *C. trachomatis* infectious cycle. We propose that CteG achieves this through a cascade of events in which at least a Pgp4-regulated gene participates. This contributes to a better understanding of chlamydial molecular pathogenesis. Despite this, many questions remain to be answered about the general function(s) of CteG and how it promotes host cell lysis. Although we have characterized in more detail the determinants of the subcellular localization of CteG, it remains unknown how CteG is targeted to the Golgi and plasma membrane in infected host cells, how this is controlled during infection, and if this dual localization corresponds to distinct functions. More specifically, it is unknown if a particular localization of CteG is required for its ability to induce host cell lysis. So far, only one host cell protein interacting with CteG (centrin-2) has been identified, but how and if this interaction relates with the role of CteG in chlamydial host cell lytic exit as suggested by our data is unclear. Furthermore, as already mentioned, the biochemical activity enabling CteG to mediate host cell lysis is unknown and it is also unclear if CteG mediates host cell lysis directly or indirectly. Answering most of these questions is required to provide detailed mechanistic insights on how the action of CteG leads to host cell lysis. Other important questions that remain to be answered include the fine-tuning of CteG activation or availability with the timing of host cell lysis specifically at late stages of infection, the role of Pgp4, Pgp4-regulated proteins and CPAF in these events and whether the recently identified Inc CT135/CTL0390 is also implied.

Additionally, we have gained further insights into the uniqueness of *cteG* in terms of the number of homologs in other species of the *Chlamydiaceae* and into the phylogenetic evolution of CteG and its homologs. Our bioinformatics analysis will be the foundation for an experimental investigation of the identified homologs of CteG in terms of their localization and the respective determinants, which will complement the analysis of those features in the case of CteG itself. In the future, determining the biological function of the identified proteins will be key to define whether some of these homologs conserved the function(s) of CteG or

evolved different functionalities. Finally, we earned a better understanding of the conservation of the genomic region of *cteG* and its homologs. Overall, the evolution of these genomic regions that translated into differences observed among groups of *Chlamydiaceae* species need to be further investigated and related with the evolution of the proteins encoded in those regions.

Overall, this PhD expanded the knowledge about several features of the T3S effector CteG of *C. trachomatis*. Specifically, besides deciphering a function of CteG on an essential step of the *C. trachomatis* infectious cycle, we have a deeper knowledge about the determinants of the subcellular localization of CteG, its role during *C. trachomatis* infection and the conservation and phylogenetic evolution of CteG in other *Chlamydia* species. However, many important questions remain to be answered and future experiments should aim at pinpointing the subcellular targeting cues of CteG and determining the mechanism of action and additional interacting partners of CteG during *C. trachomatis* infection. Altogether, this will enable a better understanding of how this bacterium establishes infections.

REFERENCES

Abby, S. S., and Rocha, E. P. C. (2012). The Non-Flagellar Type III Secretion System Evolved from the Bacterial Flagellum and Diversified into Host-Cell Adapted Systems. *PLoS Genet* 8, e1002983. doi: 10.1371/journal.pgen.1002983.

AbdelRahman, Y. M., and Belland, R. J. (2005). The chlamydial developmental cycle. *FEMS Microbiol Rev* 29, 949–959. doi: 10.1016/j.femsre.2005.03.002.

Abdelrahman, Y., Ouellette, S. P., Belland, R. J., and Cox, J. v. (2016). Polarized Cell Division of *Chlamydia trachomatis*. *PLoS Pathog* 12, e1005822. doi: 10.1371/journal.ppat.1005822.

Abdelsamed, H., Peters, J., and Byrne, G. I. (2013). Genetic variation in *Chlamydia trachomatis* and their hosts: impact on disease severity and tissue tropism. *Future Microbiol* 8, 1129–1146. doi: 10.2217/fmb.13.80.

Abraham, S., Juel, H. B., Bang, P., Cheeseman, H. M., Dohn, R. B., Cole, T., *et al.* (2019). Safety and immunogenicity of the chlamydia vaccine candidate CTH522 adjuvanted with CAF01 liposomes or aluminium hydroxide: a first-in-human, randomised, double-blind, placebo-controlled, phase 1 trial. *Lancet Infect Dis* 19, 1091–1100. doi: 10.1016/S1473-3099(19)30279-8.

Agaisse, H., and Derré, I. (2013). A *C. trachomatis* Cloning Vector and the Generation of *C. trachomatis* Strains Expressing Fluorescent Proteins under the Control of a *C. trachomatis* Promoter. *PLoS One* 8, e57090. doi: 10.1371/journal.pone.0057090.

Agaisse, H., and Derré, I. (2015). STIM1 Is a Novel Component of ER-*Chlamydia trachomatis* Inclusion Membrane Contact Sites. *PLoS One* 10, e0125671. doi: 10.1371/journal.pone.0125671.

Albrecht, M., Sharma, C. M., Reinhardt, R., Vogel, J., and Rudel, T. (2010). Deep sequencing-based discovery of the *Chlamydia trachomatis* transcriptome. *Nucleic Acids Res* 38, 868–877. doi: 10.1093/nar/gkp1032.

Aliberti, S., dela Cruz, C. S., Amati, F., Sotgiu, G., and Restrepo, M. I. (2021). Community-acquired pneumonia. *The Lancet* 398, 906–919. doi: 10.1016/S0140-6736(21)00630-9.

Almeida, F., Luís, M. P., Pereira, I. S., Pais, S. v., and Mota, L. J. (2018). The Human Centrosomal Protein CCDC146 Binds *Chlamydia trachomatis* Inclusion Membrane Protein CT288 and Is

Recruited to the Periphery of the *Chlamydia*-Containing Vacuole. *Front Cell Infect Microbiol* 8, 254. doi: 10.3389/fcimb.2018.00254.

Al-Zeer, M. A., Al-Younes, H. M., Kerr, M., Abu-Lubad, M., Gonzalez, E., Brinkmann, V., *et al.* (2014). *Chlamydia trachomatis* remodels stable microtubules to coordinate Golgi stack recruitment to the chlamydial inclusion surface. *Mol Microbiol* 94, 1285–1297. doi: 10.1111/mmi.12829.

Alzhanov, D. T., Weeks, S. K., Burnett, J. R., and Rockey, D. D. (2009). Cytokinesis is blocked in mammalian cells transfected with *Chlamydia trachomatis* gene CT223. *BMC Microbiol* 9, 2. doi: 10.1186/1471-2180-9-2.

Andreadaki, M., Hanssen, E., Deligianni, E., Claudet, C., Wengelnik, K., Mollard, V., *et al.* (2018). Sequential Membrane Rupture and Vesiculation during *Plasmodium berghei* Gametocyte Egress from the Red Blood Cell. *Sci Rep* 8, 3543. doi: 10.1038/s41598-018-21801-3.

Archuleta, T. L., Du, Y., English, C. A., Lory, S., Lesser, C., Ohi, M. D., *et al.* (2011). The *Chlamydia* Effector Chlamydial Outer Protein N (CopN) Sequesters Tubulin and Prevents Microtubule Assembly. *Journal of Biological Chemistry* 286, 33992–33998. doi: 10.1074/jbc.M111.258426.

Archuleta, T. L., and Spiller, B. W. (2014). A Gatekeeper Chaperone Complex Directs Translocator Secretion during Type Three Secretion. *PLoS Pathog* 10, e1004498. doi: 10.1371/journal.ppat.1004498.

Arnold, R., Brandmaier, S., Kleine, F., Tischler, P., Heinz, E., Behrens, S., *et al.* (2009). Sequence-Based Prediction of Type III Secreted Proteins. *PLoS Pathog* 5, e1000376. doi: 10.1371/journal.ppat.1000376.

Au, K. G., Clark, S., Miller, J. H., and Modrich, P. (1989). *Escherichia coli mutY* gene encodes an adenine glycosylase active on G-A mispairs. *Proceedings of the National Academy of Sciences* 86, 8877–8881. doi: 10.1073/pnas.86.22.8877.

Auer, D., Hügelschäffer, S. D., Fischer, A. B., and Rudel, T. (2020). The chlamydial deubiquitinase Cdu1 supports recruitment of Golgi vesicles to the inclusion. *Cell Microbiol* 22. doi: 10.1111/cmi.13136.

Bachmann, N. L., Polkinghorne, A., and Timms, P. (2014). *Chlamydia* genomics: Providing novel insights into chlamydial biology. *Trends Microbiol* 22, 464–472. doi: 10.1016/j.tim.2014.04.013.

Balin, B. J., Hammond, C. J., Little, C. S., Hingley, S. T., Al-Atrache, Z., Appelt, D. M., *et al.* (2018). *Chlamydia pneumoniae*: An Etiologic Agent for Late-Onset Dementia. *Front Aging Neurosci* 10. doi: 10.3389/fnagi.2018.00302.

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A. A., Dvorkin, M., Kulikov, A. S., *et al.* (2012). SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol* 19, 455–477. doi: 10.1089/cmb.2012.0021.

Bannantine, Griffiths, Viratyosin, Brown, and Rockey (2000). A secondary structure motif predictive of protein localization to the chlamydial inclusion membrane. *Cell Microbiol* 2, 35–47. doi: 10.1046/j.1462-5822.2000.00029.x.

Bannantine, J. P., Stamm, W. E., Suchland, R. J., and Rockey, D. D. (1998). *Chlamydia trachomatis* IncA Is Localized to the Inclusion Membrane and Is Recognized by Antisera from Infected Humans and Primates. *Infect Immun* 66, 6017–6021. doi: 10.1128/IAI.66.12.6017-6021.1998.

Barron, A. L., White, H. J., Rank, R. G., Soloff, B. L., and Moses, E. B. (1981). A New Animal Model for the Study of *Chlamydia trachomatis* Genital Infections: Infection of Mice with the Agent of Mouse Pneumonitis. *J Infect Dis* 143, 63–66. doi: 10.1093/infdis/143.1.63.

Barta, M. L., Thomas, K., Yuan, H., Lovell, S., Battaile, K. P., Schramm, V. L., *et al.* (2014). Structural and Biochemical Characterization of *Chlamydia trachomatis* Hypothetical Protein CT263 Supports That Menaquinone Synthesis Occurs through the Futalosine Pathway. *Journal of Biological Chemistry* 289, 32214–32229. doi: 10.1074/jbc.M114.594325.

Bastidas, R. J., Elwell, C. A., Engel, J. N., and Valdivia, R. H. (2013). Chlamydial intracellular survival strategies. *Cold Spring Harb Perspect Med* 3. doi: 10.1101/cshperspect.a010256.

Bauler, L. D., and Hackstadt, T. (2014). Expression and Targeting of secreted proteins from *Chlamydia trachomatis*. *J Bacteriol* 196, 1325–1334. doi: 10.1128/JB.01290-13.

Beare, P. A., Howe, D., Cockrell, D. C., Omsland, A., Hansen, B., and Heinzen, R. A. (2009). Characterization of a *Coxiella burnetii ftsZ* Mutant Generated by Himar1 Transposon Mutagenesis. *J Bacteriol* 191, 1369–1381. doi: 10.1128/JB.01580-08.

Belland, R. J., Ouellette, S. P., Gieffers, J., and Byrne, G. I. (2004). *Chlamydia pneumoniae* and atherosclerosis. *Cell Microbiol* 6, 117–127. doi: 10.1046/j.1462-5822.2003.00352.x.

Belland, R. J., Scidmore, M. A., Crane, D. D., Hogan, D. M., Whitmire, W., McClarty, G., *et al.* (2001). *Chlamydia trachomatis* cytotoxicity associated with complete and partial cytotoxin genes. *Proceedings of the National Academy of Sciences* 98, 13984–13989. doi: 10.1073/pnas.241377698.

Bertelli, C., Collyn, F., Croxatto, A., Rückert, C., Polkinghorne, A., Kebbi-Beghdadi, C., *et al.* (2010). The *Waddlia* Genome: A Window into Chlamydial Biology. *PLoS One* 5, e10890. doi: 10.1371/journal.pone.0010890.

Betts-Hampikian, H. J., and Fields, K. A. (2010). The Chlamydial Type III Secretion Mechanism: Revealing Cracks in a Tough Nut. *Front Microbiol* 1. doi: 10.3389/fmicb.2010.00114.

Binet, R., and Maurelli, A. T. (2009). Transformation and isolation of allelic exchange mutants of *Chlamydia psittaci* using recombinant DNA introduced by electroporation. *Proceedings of the National Academy of Sciences* 106, 292–297. doi: 10.1073/pnas.0806768106.

Bishop, R. C., and Derré, I. (2022). The *Chlamydia trachomatis* Inclusion Membrane Protein CTL0390 Mediates Host Cell Exit via Lysis through STING Activation. *Infect Immun* 90. doi: 10.1128/iai.00190-22.

Black, D. S., and Bliska, J. B. (2000). The RhoGAP activity of the *Yersinia pseudotuberculosis* cytotoxin YopE is required for antiphagocytic function and virulence. *Mol Microbiol* 37, 515–527. doi: 10.1046/j.1365-2958.2000.02021.x.

Bolger, A. M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30, 2114–2120. doi: 10.1093/bioinformatics/btu170.

Boncompain, G., Müller, C., Meas-Yedid, V., Schmitt-Kopplin, P., Lazarow, P. B., and Subtil, A. (2014). The Intracellular Bacteria *Chlamydia* Hijack Peroxisomes and Utilize Their Enzymatic Capacity to Produce Bacteria-Specific Phospholipids. *PLoS One* 9, e86196. doi: 10.1371/journal.pone.0086196.

Borel, N., Leonard, C., Slade, J., and Schoborg, R. v. (2016). Chlamydial Antibiotic Resistance and Treatment Failure in Veterinary and Human Medicine. *Curr Clin Microbiol Rep* 3, 10–18. doi: 10.1007/s40588-016-0028-4.

Bugalhão, J. N., Luís, M. P., Pereira, I. S., da Cunha, M., Pais, S. v, and Mota, L. J. (2022). The *Chlamydia trachomatis* inclusion membrane protein CT006 associates with lipid droplets in eukaryotic cells. *PLoS One* 17, 1–27. doi: 10.1371/journal.pone.0264292.

Bugalhão, J. N., and Mota, L. J. (2019). The multiple functions of the numerous *Chlamydia trachomatis* secreted proteins: The tip of the iceberg. *Microbial Cell* 6, 414–449. doi: 10.15698/mic2019.09.691.

Bullock, H. D., Hower, S., and Fields, K. A. (2012). Domain Analyses Reveal That *Chlamydia trachomatis* CT694 Protein Belongs to the Membrane-localized Family of Type III Effector Proteins. *Journal of Biological Chemistry* 287, 28078–28086. doi: 10.1074/jbc.M112.386904.

Caldwell, H. D., Kromhout, J., and Schachter, J. (1981). Purification and partial characterization of the major outer membrane protein of *Chlamydia trachomatis*. *Infect Immun* 31, 1161–1176. doi: 10.1128/iai.31.3.1161-1176.1981.

Caldwell, H. D., Wood, H., Crane, D., Bailey, R., Jones, R. B., Mabey, D., *et al.* (2003). Polymorphisms in *Chlamydia trachomatis* tryptophan synthase genes differentiate between genital and ocular isolates. *Journal of Clinical Investigation* 111, 1757–1769. doi: 10.1172/JCI17993.

Capella-Gutiérrez, S., Silla-Martínez, J. M., and Gabaldón, T. (2009). trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* 25, 1972–1973. doi: 10.1093/bioinformatics/btp348.

Capmany, A., and Damiani, M. T. (2010). *Chlamydia trachomatis* Intercepts Golgi-Derived Sphingolipids through a Rab14-Mediated Transport Required for Bacterial Development and Replication. *PLoS One* 5, e14084. doi: 10.1371/journal.pone.0014084.

Carabeo, R. A., Dooley, C. A., Grieshaber, S. S., and Hackstadt, T. (2007). Rac interacts with Abi-1 and WAVE2 to promote an Arp2/3-dependent actin recruitment during chlamydial invasion. *Cell Microbiol* 9, 2278–2288. doi: 10.1111/j.1462-5822.2007.00958.x.

Carlson, J. H., Hughes, S., Hogan, D., Cieplak, G., Sturdevant, D. E., McClarty, G., *et al.* (2004). Polymorphisms in the *Chlamydia trachomatis* Cytotoxin Locus Associated with Ocular and Genital Isolates. *Infect Immun* 72, 7063–7072. doi: 10.1128/IAI.72.12.7063-7072.2004.

Carlson, J. H., Porcella, S. F., McClarty, G., and Caldwell, H. D. (2005). Comparative Genomic Analysis of *Chlamydia trachomatis* Oculotropic and Genitotropic Strains. *Infect Immun* 73, 6407–6418. doi: 10.1128/IAI.73.10.6407-6418.2005.

Carlson, J. H., Whitmire, W. M., Crane, D. D., Wicke, L., Virtaneva, K., Sturdevant, D. E., *et al.* (2008). The *Chlamydia trachomatis* Plasmid Is a Transcriptional Regulator of Chromosomal Genes and a Virulence Factor. *Infect Immun* 76, 2273–2283. doi: 10.1128/IAI.00102-08.

Carpenter, V., Chen, Y.-S., Dolat, L., and Valdivia, R. H. (2017). The Effector TepP Mediates Recruitment and Activation of Phosphoinositide 3-Kinase on Early *Chlamydia trachomatis* Vacuoles. *mSphere* 2. doi: 10.1128/mSphere.00207-17.

Chaiwattanarungruengpaisan, S., Thongdee, M., Anuntakarun, S., Payungporn, S., Arya, N., Punchukrang, A., *et al.* (2021). A new species of *Chlamydia* isolated from Siamese crocodiles (*Crocodylus siamensis*). *PLoS One* 16, e0252081. doi: 10.1371/journal.pone.0252081.

Chellas-Géry, B., Linton, C. N., and Fields, K. A. (2007). Human GCIP interacts with CT847, a novel *Chlamydia trachomatis* type III secretion substrate, and is degraded in a tissue-culture infection model. *Cell Microbiol* 9, 2417–2430. doi: 10.1111/j.1462-5822.2007.00970.x.

Chen, A. L., Johnson, K. A., Lee, J. K., Sütterlin, C., and Tan, M. (2012). CPAF: A Chlamydial Protease in Search of an Authentic Substrate. *PLoS Pathog* 8, e1002842. doi: 10.1371/journal.ppat.1002842.

Chen, C., Chen, D., Sharma, J., Cheng, W., Zhong, Y., Liu, K., *et al.* (2006). The Hypothetical Protein CT813 Is Localized in the *Chlamydia trachomatis* Inclusion Membrane and Is Immunogenic in Women Urogenitally Infected with *C. trachomatis*. *Infect Immun* 74, 4826–4840. doi: 10.1128/IAI.00081-06.

Chen, Y.-S., Bastidas, R. J., Saka, H. A., Carpenter, V. K., Richards, K. L., Plano, G. v., *et al.* (2014). The *Chlamydia trachomatis* Type III Secretion Chaperone Slc1 Engages Multiple Early Effectors, Including TepP, a Tyrosine-phosphorylated Protein Required for the Recruitment of CrkI-II to Nascent Inclusions and Innate Immune Signaling. *PLoS Pathog* 10, e1003954. doi: 10.1371/journal.ppat.1003954.

Clark, T. R., Lackey, A. M., Kleba, B., Driskell, L. O., Lutter, E. I., Martens, C., *et al.* (2011). Transformation Frequency of a mariner-Based Transposon in *Rickettsia rickettsii*. *J Bacteriol* 193, 4993–4995. doi: 10.1128/JB.05279-11.

Clifton, D. R., Fields, K. A., Grieshaber, S. S., Dooley, C. A., Fischer, E. R., Mead, D. J., *et al.* (2004). A chlamydial type III translocated protein is tyrosine-phosphorylated at the site of entry and associated with recruitment of actin. *Proceedings of the National Academy of Sciences* 101, 10166–10171. doi: 10.1073/pnas.0402829101.

Cocchiaro, J. L., Kumar, Y., Fischer, E. R., Hackstadt, T., and Valdivia, R. H. (2008). Cytoplasmic lipid droplets are translocated into the lumen of the *Chlamydia trachomatis* parasitophorous vacuole. *Proceedings of the National Academy of Sciences* 105, 9379–9384. doi: 10.1073/pnas.0712241105.

Collingro, A., Köstlbacher, S., Mussmann, M., Stepanauskas, R., Hallam, S. J., and Horn, M. (2017). Unexpected genomic features in widespread intracellular bacteria: evidence for motility of marine chlamydiae. *ISME J* 11, 2334–2344. doi: 10.1038/ismej.2017.95.

Collingro, A., Tischler, P., Weinmaier, T., Penz, T., Heinz, E., Brunham, R. C., *et al.* (2011). Unity in Variety—The Pan-Genome of the *Chlamydiae*. *Mol Biol Evol* 28, 3253–3270. doi: 10.1093/molbev/msr161.

Comanducci, M., Cevenini, R., Moroni, A., Giuliani, M. M., Ricci, S., Scarlato, V., *et al.* (1993). Expression of a plasmid gene of *Chlamydia trachomatis* encoding a novel 28 kDa antigen. *J Gen Microbiol* 139, 1083–1092. doi: 10.1099/00221287-139-5-1083.

Cornelis, G. R. (2006). The type III secretion injectisome. *Nat Rev Microbiol* 4, 811–825. doi: 10.1038/nrmicro1526.

Cortina, M. E., Ende, R. J., Bishop, R. C., Bayne, C., and Derré, I. (2019). *Chlamydia trachomatis* and *Chlamydia muridarum* spectinomycin resistant vectors and a transcriptional fluorescent reporter to monitor conversion from replicative to infectious bacteria. *PLoS One* 14, e0217753. doi: 10.1371/journal.pone.0217753.

Cossé, M. M., Barta, M. L., Fisher, D. J., Oesterlin, L. K., Niragire, B., Perrinet, S., *et al.* (2018). The Loss of Expression of a Single Type 3 Effector (CT622) Strongly Reduces *Chlamydia trachomatis* Infectivity and Growth. *Front Cell Infect Microbiol* 8. doi: 10.3389/fcimb.2018.00145.

Costa, T. R. D., Felisberto-Rodrigues, C., Meir, A., Prevost, M. S., Redzej, A., Trokter, M., *et al.* (2015). Secretion systems in Gram-negative bacteria: structural and mechanistic insights. *Nat Rev Microbiol* 13, 343–359. doi: 10.1038/nrmicro3456.

da Cunha, M., Milho, C., Almeida, F., Pais, S. v, Borges, V., Maurício, R., *et al.* (2014). Identification of type III secretion substrates of *Chlamydia trachomatis* using *Yersinia enterocolitica* as a heterologous system. *BMC Microbiol* 14, 40. doi: 10.1186/1471-2180-14-40.

da Cunha, M., Pais, S. v., Bugalhão, J. N., and Mota, L. J. (2017). The *Chlamydia trachomatis* type III secretion substrates CT142, CT143, and CT144 are secreted into the lumen of the inclusion. *PLoS One* 12, e0178856. doi: 10.1371/journal.pone.0178856.

Dallenga, T., Repnik, U., Corleis, B., Eich, J., Reimer, R., Griffiths, G. W., *et al.* (2017). *M. tuberculosis*-Induced Necrosis of Infected Neutrophils Promotes Bacterial Growth Following Phagocytosis by Macrophages. *Cell Host Microbe* 22, 519-530.e3. doi: 10.1016/j.chom.2017.09.003.

Daniele, C., and Gilbert, G. (2006). Pathogenic Potential of Novel *Chlamydiae* and Diagnostic Approaches to Infections Due to These Obligate Intracellular Bacteria. *Clin Microbiol Rev* 19, 283–297. doi: 10.1128/CMR.19.2.283-297.2006.

Darling, A. E., Mau, B., and Perna, N. T. (2010). progressiveMauve: multiple genome alignment with gene gain, loss and rearrangement. *PLoS One* 5, e11147. doi: 10.1371/journal.pone.0011147.

de Barsy, M., Bottinelli, L., and Greub, G. (2014). Antibiotic susceptibility of *Estrella lausannensis*, a potential emerging pathogen. *Microbes Infect* 16, 746–754. doi: 10.1016/j.micinf.2014.08.003.

de Barsy, M., and Greub, G. (2013). *Waddlia chondrophila*: from biology to pathogenicity. *Microbes Infect* 15, 1033–1041. doi: 10.1016/j.micinf.2013.09.010.

de Vries, H. J. C., de Barbeyrac, B., de Vrieze, N. H. N., Viset, J. D., White, J. A., Vall-Mayans, M., *et al.* (2019). 2019 European guideline on the management of lymphogranuloma venereum. *Journal of the European Academy of Dermatology and Venereology* 33, 1821–1828. doi: 10.1111/jdv.15729.

Dehoux, P., Flores, R., Dauga, C., Zhong, G., and Subtil, A. (2011). Multi-genome identification and characterization of chlamydiae-specific type III secretion substrates: The Inc proteins. *BMC Genomics* 12. doi: 10.1186/1471-2164-12-109.

Deiwick, J., Salcedo, S. P., Boucrot, E., Gilliland, S. M., Henry, T., Petermann, N., *et al.* (2006). The Translocated *Salmonella* Effector Proteins SseF and SseG Interact and Are Required To Establish an Intracellular Replication Niche. *Infect Immun* 74, 6965–6972. doi: 10.1128/IAI.00648-06.

Delevoye, C., Nilges, M., Dehoux, P., Paumet, F., Perrinet, S., Dautry-Varsat, A., *et al.* (2008). SNARE Protein Mimicry by an Intracellular Bacterium. *PLoS Pathog* 4, e1000022. doi: 10.1371/journal.ppat.1000022.

Derré, I., Swiss, R., and Agaisse, H. (2011). The Lipid Transfer Protein CERT Interacts with the *Chlamydia* Inclusion Protein IncD and Participates to ER-*Chlamydia* Inclusion Membrane Contact Sites. *PLoS Pathog* 7, e1002092. doi: 10.1371/journal.ppat.1002092.

Dharamshi, J. E., Tamarit, D., Eme, L., Stairs, C. W., Martijn, J., Homa, F., *et al.* (2020). Marine Sediments Illuminate *Chlamydiae* Diversity and Evolution. *Current Biology* 30, 1032-1048.e7. doi: 10.1016/j.cub.2020.02.016.

Dille, S., Kleinschnitz, E.-M., Kontchou, C. W., Nölke, T., and Häcker, G. (2015). In Contrast to *Chlamydia trachomatis*, *Waddlia chondrophila* Grows in Human Cells without Inhibiting Apoptosis, Fragmenting the Golgi Apparatus, or Diverting Post-Golgi Sphingomyelin Transport. *Infect Immun* 83, 3268–3280. doi: 10.1128/IAI.00322-15.

Domeika, M., Savicheva, A., Sokolovskiy, E., Frigo, N., Brilene, T., Hallén, A., *et al.* (2009). Guidelines for the laboratory diagnosis of *Chlamydia trachomatis* infections in East European

countries. *Journal of the European Academy of Dermatology and Venereology* 23, 1353–1363. doi: 10.1111/j.1468-3083.2009.03328.x.

Dong, F., Pirbhai, M., Zhong, Y., and Zhong, G. (2004). Cleavage-dependent activation of a chlamydia-secreted protease. *Mol Microbiol* 52, 1487–1494. doi: 10.1111/j.1365-2958.2004.04072.x.

Dumoux, M., Menny, A., Delacour, D., and Hayward, R. D. (2015). A *Chlamydia* effector recruits CEP170 to reprogram host microtubule organization. *J Cell Sci*. doi: 10.1242/jcs.169318.

Elwell, C. A., Czudnochowski, N., von Dollen, J., Johnson, J. R., Nakagawa, R., Mirrashidi, K., *et al.* (2017). *Chlamydia* interfere with an interaction between the mannose-6-phosphate receptor and sorting nexins to counteract host restriction. *Elife* 6. doi: 10.7554/eLife.22709.

Elwell, C., Mirrashidi, K., and Engel, J. (2016). *Chlamydia* cell biology and pathogenesis. *Nat Rev Microbiol* 14, 385–400. doi: 10.1038/nrmicro.2016.30.

Emms, D. M., and Kelly, S. (2019). OrthoFinder: phylogenetic orthology inference for comparative genomics. *Genome Biol* 20, 238. doi: 10.1186/s13059-019-1832-y.

Everett, K. D., and Hatch, T. P. (1991). Sequence analysis and lipid modification of the cysteine-rich envelope proteins of *Chlamydia psittaci* 6BC. *J Bacteriol* 173, 3821–3830. doi: 10.1128/jb.173.12.3821-3830.1991.

Fabijan, J., Caraguel, C., Jelocnik, M., Polkinghorne, A., Boardman, W. S. J., Nishimoto, E., *et al.* (2019). *Chlamydia pecorum* prevalence in South Australian koala (*Phascolarctos cinereus*) populations: Identification and modelling of a population free from infection. *Sci Rep* 9, 6261. doi: 10.1038/s41598-019-42702-z.

Farencena, A., Comanducci, M., Donati, M., Ratti, G., and Cevenini, R. (1997). Characterization of a new isolate of *Chlamydia trachomatis* which lacks the common plasmid and has properties of biovar trachoma. *Infect Immun* 65, 2965–2969. doi: 10.1128/iai.65.7.2965-2969.1997.

Faris, R., McCullough, A., Andersen, S. E., Moninger, T. O., and Weber, M. M. (2020). The *Chlamydia trachomatis* secreted effector TmeA hijacks the N-WASP-ARP2/3 actin remodeling axis to facilitate cellular invasion. *PLoS Pathog* 16, e1008878. doi: 10.1371/journal.ppat.1008878.

Faris, R., Merling, M., Andersen, S. E., Dooley, C. A., Hackstadt, T., and Weber, M. M. (2019). *Chlamydia trachomatis* CT229 Subverts Rab GTPase-Dependent CCV Trafficking Pathways to Promote Chlamydial Infection. *Cell Rep* 26, 3380-3390.e5. doi: 10.1016/j.celrep.2019.02.079.

Fehlner-Gardiner, C., Roshick, C., Carlson, J. H., Hughes, S., Belland, R. J., Caldwell, H. D., *et al.* (2002). Molecular Basis Defining Human *Chlamydia trachomatis* Tissue Tropism. *Journal of Biological Chemistry* 277, 26893–26903. doi: 10.1074/jbc.M203937200.

Fehr, A., Walther, E., Schmidt-Posthaus, H., Nufer, L., Wilson, A., Svercel, M., *et al.* (2013). *Candidatus* Syngnamydia Venezia, a Novel Member of the Phylum *Chlamydiae* from the

Broad Nosed Pipefish, *Syngnathus typhle*. *PLoS One* 8, e70853. doi: 10.1371/journal.pone.0070853.

Ferreira, R., Borges, V., Nunes, A., Borrego, M. J., and Gomes, J. P. (2013). Assessment of the load and transcriptional dynamics of *Chlamydia trachomatis* plasmid according to strains' tissue tropism. *Microbiol Res* 168, 333–339. doi: 10.1016/j.micres.2013.02.001.

Fields, K. A. (2012). "Protein Secretion and *Chlamydia* Pathogenesis," in *Intracellular Pathogens I: Chlamydiales*, eds. M. Tan and P. Bavoil (Washington, DC, USA: ASM Press), 192–216. doi: 10.1128/9781555817329.

Fields, K. A., Fischer, E. R., Mead, D. J., and Hackstadt, T. (2005). Analysis of putative *Chlamydia trachomatis* chaperones Scc2 and Scc3 and their use in the identification of type III secretion substrates. *J Bacteriol* 187, 6466–6478. doi: 10.1128/JB.187.18.6466-6478.2005.

Fields, K. A., and Hackstadt, T. (2000). Evidence for the secretion of *Chlamydia trachomatis* CopN by a type III secretion mechanism. *Mol Microbiol* 38, 1048–1060. doi: 10.1046/j.1365-2958.2000.02212.x.

Fields, K. A., and Hackstadt, T. (2002). The Chlamydial Inclusion: Escape from the Endocytic Pathway. *Annu Rev Cell Dev Biol* 18, 221–245. doi: 10.1146/annurev.cellbio.18.012502.105845.

Fischer, A., Harrison, K. S., Ramirez, Y., Auer, D., Chowdhury, S. R., Prusty, B. K., *et al.* (2017). *Chlamydia trachomatis*-containing vacuole serves as deubiquitination platform to stabilize Mcl-1 and to interfere with host defense. *Elife* 6. doi: 10.7554/eLife.21465.

Flieger, A., Frischknecht, F., Häcker, G., Hornef, M. W., and Pradel, G. (2018). Pathways of host cell exit by intracellular pathogens. *Microbial Cell* 5, 525–544. doi: 10.15698/mic2018.12.659.

Fling, S. P., Sutherland, R. A., Steele, L. N., Hess, B., D'Orazio, S. E. F., Maisonneuve, J.-F., *et al.* (2001). CD8+ T cells recognize an inclusion membrane-associated protein from the vacuolar pathogen *Chlamydia trachomatis*. *Proceedings of the National Academy of Sciences* 98, 1160–1165. doi: 10.1073/pnas.98.3.1160.

Forrester, M. T., Hess, D. T., Thompson, J. W., Hultman, R., Moseley, M. A., Stamler, J. S., *et al.* (2011). Site-specific analysis of protein S-acylation by resin-assisted capture. *J Lipid Res* 52, 393–398. doi: 10.1194/jlr.D011106.

Friedman, M. G., Dvoskin, B., and Kahane, S. (2003). Infections with the chlamydia-like microorganism *Simkania negevensis*, a possible emerging pathogen. *Microbes Infect* 5, 1013–1021. doi: 10.1016/S1286-4579(03)00188-6.

Fu, Y., and Galán, J. E. (1999). A *Salmonella* protein antagonizes Rac-1 and Cdc42 to mediate host-cell recovery after bacterial invasion. *Nature* 401, 293–297. doi: 10.1038/45829.

Galán, J. E., and Waksman, G. (2018). Protein-Injection Machines in Bacteria. *Cell* 172, 1306–1318. doi: 10.1016/j.cell.2018.01.034.

Garvin, L., vande Voorde, R., Dickinson, M., Carrell, S., Hybiske, K., and Rockey, D. (2021). A broad-spectrum cloning vector that exists as both an integrated element and a free plasmid in *Chlamydia trachomatis*. *PLoS One* 16, e0261088. doi: 10.1371/journal.pone.0261088.

Ge, J., Gong, Y.-N., Xu, Y., and Shao, F. (2012). Preventing bacterial DNA release and absent in melanoma 2 inflammasome activation by a *Legionella* effector functioning in membrane trafficking. *Proceedings of the National Academy of Sciences* 109, 6193–6198. doi: 10.1073/pnas.1117490109.

Gehre, L., Gorgette, O., Perrinet, S., Prevost, M.-C., Ducatez, M., Giebel, A. M., *et al.* (2016). Sequestration of host metabolism by an intracellular pathogen. *Elife* 5, e12552. doi: 10.7554/eLife.12552.

Glushakova, S., Beck, J. R., Garten, M., Busse, B. L., Nasamu, A. S., Tenkova-Heuser, T., *et al.* (2018). Rounding precedes rupture and breakdown of vacuolar membranes minutes before malaria parasite egress from erythrocytes. *Cell Microbiol* 20, e12868. doi: 10.1111/cmi.12868.

Gomes, J. P., Bruno, W. J., Nunes, A., Santos, N., Florindo, C., Borrego, M. J., *et al.* (2007). Evolution of *Chlamydia trachomatis* diversity occurs by widespread interstrain recombination involving hotspots. *Genome Res* 17, 50–60. doi: 10.1101/gr.5674706.

Gong, S., Lei, L., Chang, X., Belland, R., and Zhong, G. (2011). *Chlamydia trachomatis* secretion of hypothetical protein CT622 into host cell cytoplasm via a secretion pathway that can be inhibited by the type III secretion system inhibitor compound 1. *Microbiology (N Y)* 157, 1134–1144. doi: 10.1099/mic.0.047746-0.

Gong, S., Yang, Z., Lei, L., Shen, L., and Zhong, G. (2013). Characterization of *Chlamydia trachomatis* plasmid-encoded open reading frames. *J Bacteriol* 195, 3819–3826. doi: 10.1128/JB.00511-13.

Gophna, U., Ron, E. Z., and Graur, D. (2003). Bacterial type III secretion systems are ancient and evolved by multiple horizontal-transfer events. *Gene* 312, 151–163. doi: 10.1016/S0378-1119(03)00612-7.

Goy, G., Croxatto, A., and Greub, G. (2008). *Waddlia chondrophila* enters and multiplies within human macrophages. *Microbes Infect* 10, 556–562. doi: 10.1016/j.micinf.2008.02.003.

Gupta, R. S., Naushad, S., Chokshi, C., Griffiths, E., and Adeolu, M. (2015). A phylogenomic and molecular markers based analysis of the phylum *Chlamydiae*: proposal to divide the class *Chlamydiia* into two orders, *Chlamydiales* and *Parachlamydiales* ord. nov., and emended description of the class *Chlamydiia*. *Antonie Van Leeuwenhoek* 108, 765–781. doi: 10.1007/s10482-015-0532-1.

Haase, H., Pagel, I., Khalina, Y., Zacharzowsky, U., Person, V., Lutsch, G., *et al.* (2004). The carboxyl-terminal ahnak domain induces actin bundling and stabilizes muscle contraction. *The FASEB Journal* 18, 839–841. doi: 10.1096/fj.03-0446fje.

Hackstadt, T. (2014). "Initial Interactions of *Chlamydiae* with the Host Cell," in *Intracellular Pathogens I: Chlamydiales*, eds. M. Tan and P. Bavoli (Washington, DC, USA: ASM Press), 126–148. doi: 10.1128/9781555817329.

Hackstadt, T., Todd, W. J., and Caldwell, H. D. (1985). Disulfide-mediated interactions of the chlamydial major outer membrane protein: role in the differentiation of chlamydiae? *J Bacteriol* 161, 25–31. doi: 10.1128/jb.161.1.25-31.1985.

Halberstädter, L., and von Prowazeck, S. (1907). Über Zelleinschlüsse parasitärer Natur beim Trachom. 26, 44–47.

Hamaoui, D., Cossé, M. M., Mohan, J., Lystad, A. H., Wollert, T., and Subtil, A. (2020). The *Chlamydia* effector CT622/TaiP targets a nonautophagy related function of ATG16L1. *Proceedings of the National Academy of Sciences* 117, 26784–26794. doi: 10.1073/pnas.2005389117.

Hamma, T., and Ferré-D'Amaré, A. R. (2006). Pseudouridine Synthases. *Chem Biol* 13, 1125–1135. doi: 10.1016/j.chembiol.2006.09.009.

Harris, S. R., Clarke, I. N., Seth-Smith, H. M. B., Solomon, A. W., Cutcliffe, L. T., Marsh, P., *et al.* (2012). Whole-genome analysis of diverse *Chlamydia trachomatis* strains identifies phylogenetic relationships masked by current clinical typing. *Nat Genet* 44, 413–419. doi: 10.1038/ng.2214.

Hartley, J. C., Stevenson, S., Robinson, A. J., Littlewood, J. D., Carder, C., Cartledge, J., *et al.* (2001). Conjunctivitis Due to *Chlamydophila felis* (*Chlamydia psittaci* Feline Pneumonitis Agent) Acquired From a Cat: Case Report with Molecular Characterization of Isolates from the Patient and Cat. *Journal of Infection* 43, 7–11. doi: 10.1053/jinf.2001.0845.

Hatch, T. P. (1996). Disulfide cross-linked envelope proteins: the functional equivalent of peptidoglycan in chlamydiae? *J Bacteriol* 178, 1–5. doi: 10.1128/jb.178.1.1-5.1996.

Hefty, P. S., and Stephens, R. S. (2007). Chlamydial Type III Secretion System Is Encoded on Ten Operons Preceded by Sigma 70-Like Promoter Elements. *J Bacteriol* 189, 198–206. doi: 10.1128/JB.01034-06.

Heijne, M., Jelocnik, M., Umanets, A., Brouwer, M. S. M., Dinkla, A., Harders, F., *et al.* (2021). Genetic and phenotypic analysis of the pathogenic potential of two novel *Chlamydia gallinacea* strains compared to *Chlamydia psittaci*. *Sci Rep* 11, 16516. doi: 10.1038/s41598-021-95966-9.

Heinz, E., Rockey, D. D., Montanaro, J., Aistleitner, K., Wagner, M., and Horn, M. (2010). Inclusion Membrane Proteins of *Protochlamydia amoebophila* UWE25 Reveal a Conserved Mechanism for Host Cell Interaction among the *Chlamydiae*. *J Bacteriol* 192, 5093–5102. doi: 10.1128/JB.00605-10.

Heuer, D., Lipinski, A. R., Machuy, N., Karlas, A., Wehrens, A., Siedler, F., *et al.* (2009). *Chlamydia* causes fragmentation of the Golgi compartment to ensure reproduction. *Nature* 457, 731–735. doi: 10.1038/nature07578.

Hicks, S. W., Charron, G., Hang, H. C., and Galán, J. E. (2011). Subcellular Targeting of *Salmonella* Virulence Proteins by Host-Mediated S-Palmitoylation. *Cell Host Microbe* 10, 9–20. doi: 10.1016/j.chom.2011.06.003.

Ho, T. D., and Starnbach, M. N. (2005). The *Salmonella enterica* Serovar Typhimurium-Encoded Type III Secretion Systems Can Translocate *Chlamydia trachomatis* Proteins into the Cytosol of Host Cells. *Infect Immun* 73, 905–911. doi: 10.1128/IAI.73.2.905-911.2005.

Hoang, D. T., Chernomor, O., von Haeseler, A., Minh, B. Q., and Vinh, L. S. (2018). UFBoot2: Improving the Ultrafast Bootstrap Approximation. *Mol Biol Evol* 35, 518–522. doi: 10.1093/molbev/msx281.

Hobolt-Pedersen, A.-S., Christiansen, G., Timmerman, E., Gevaert, K., and Birkelund, S. (2009). Identification of *Chlamydia trachomatis* CT621, a protein delivered through the type III secretion system to the host cell cytoplasm and nucleus. *FEMS Immunol Med Microbiol* 57, 46–58. doi: 10.1111/j.1574-695X.2009.00581.x.

Hogan, R. J., Mathews, S. A., Mukhopadhyay, S., Summersgill, J. T., and Timms, P. (2004). Chlamydial Persistence: beyond the Biphasic Paradigm. *Infect Immun* 72, 1843–1855. doi: 10.1128/IAI.72.4.1843-1855.2004.

Horn, M., Collingro, A., Schmitz-Esser, S., Beier, C. L., Purkhold, U., Fartmann, B., *et al.* (2004). Illuminating the Evolutionary History of *Chlamydiae*. *Science (1979)* 304, 728–730. doi: 10.1126/science.1096330.

Hou, S., Dong, X., Yang, Z., Li, Z., Liu, Q., and Zhong, G. (2015). Chlamydial Plasmid-Encoded Virulence Factor Pgp3 Neutralizes the Antichlamydial Activity of Human Cathelicidin LL-37. *Infect Immun* 83, 4701–4709. doi: 10.1128/IAI.00746-15.

Hower, S., Wolf, K., and Fields, K. A. (2009). Evidence that CT694 is a novel *Chlamydia trachomatis* T3S substrate capable of functioning during invasion or early cycle development. *Mol Microbiol* 72, 1423–1437. doi: 10.1111/j.1365-2958.2009.06732.x.

Hsieh, T.-S., Lopez, V. A., Black, M. H., Osinski, A., Pawłowski, K., Tomchick, D. R., *et al.* (2021). Dynamic remodeling of host membranes by self-organizing bacterial effectors. *Science (1979)* 372, 935–941. doi: 10.1126/science.aay8118.

Hu, Y., Huang, H., Cheng, X., Shu, X., White, A. P., Stavrinides, J., *et al.* (2017). A global survey of bacterial type III secretion systems and their effectors. *Environ Microbiol* 19, 3879–3895. doi: 10.1111/1462-2920.13755.

Huang, Y., Wurihan, W., Lu, B., Zou, Y., Wang, Y., Weldon, K., *et al.* (2022). Robust Heat Shock Response in *Chlamydia* Lacking a Typical Heat Shock Sigma Factor. *Front Microbiol* 12. doi: 10.3389/fmicb.2021.812448.

Huang, Y., Zhang, Q., Yang, Z., Conrad, T., Liu, Y., and Zhong, G. (2015). Plasmid-Encoded Pgp5 Is a Significant Contributor to *Chlamydia muridarum* Induction of Hydrosalpinx. *PLoS One* 10, e0124840. doi: 10.1371/journal.pone.0124840.

Huang, Z., Feng, Y., Chen, D., Wu, X., Huang, S., Wang, X., *et al.* (2008). Structural Basis for Activation and Inhibition of the Secreted *Chlamydia* Protease CPAF. *Cell Host Microbe* 4, 529–542. doi: 10.1016/j.chom.2008.10.005.

Hueck, C. J. (1998). Type III Protein Secretion Systems in Bacterial Pathogens of Animals and Plants. *Microbiology and Molecular Biology Reviews* 62, 379–433. doi: 10.1128/MMBR.62.2.379-433.1998.

Huston, W. M., Swedberg, J. E., Harris, J. M., Walsh, T. P., Mathews, S. A., and Timms, P. (2007). The temperature activated HtrA protease from pathogen *Chlamydia trachomatis* acts as both a chaperone and protease at 37°C. *FEBS Lett* 581, 3382–3386. doi: 10.1016/j.febslet.2007.06.039.

Hybiske, K., and Stephens, R. S. (2007). Mechanisms of host cell exit by the intracellular bacterium *Chlamydia*. *Proceedings of the National Academy of Sciences* 104, 11430–11435. doi: 10.1073/pnas.0703218104.

Isaksson, E. L., Aili, M., Fahlgren, A., Carlsson, S. E., Rosqvist, R., and Wolf-Watz, H. (2009). The Membrane Localization Domain Is Required for Intracellular Localization and Autoregulation of YopE in *Yersinia pseudotuberculosis*. *Infect Immun* 77, 4740–4749. doi: 10.1128/IAI.00333-09.

Jelocnik, M., Bachmann, N. L., Kaltenboeck, B., Waugh, C., Woolford, L., Speight, K. N., *et al.* (2015). Genetic diversity in the plasticity zone and the presence of the chlamydial plasmid differentiates *Chlamydia pecorum* strains from pigs, sheep, cattle, and koalas. *BMC Genomics* 16, 893. doi: 10.1186/s12864-015-2053-8.

Jewett, T. J., Dooley, C. A., Mead, D. J., and Hackstadt, T. (2008). *Chlamydia trachomatis* tarp is phosphorylated by src family tyrosine kinases. *Biochem Biophys Res Commun* 371, 339–344. doi: 10.1016/j.bbrc.2008.04.089.

Jewett, T. J., Fischer, E. R., Mead, D. J., and Hackstadt, T. (2006). Chlamydial TARP is a bacterial nucleator of actin. *Proceedings of the National Academy of Sciences* 103, 15599–15604. doi: 10.1073/pnas.0603044103.

Jewett, T. J., Miller, N. J., Dooley, C. A., and Hackstadt, T. (2010). The Conserved Tarp Actin Binding Domain Is Important for Chlamydial Invasion. *PLoS Pathog* 6, e1000997. doi: 10.1371/journal.ppat.1000997.

Jia, L., Sun, F., Wang, J., Gong, D., and Yang, L. (2019). *Chlamydia trachomatis* ct143 stimulates secretion of proinflammatory cytokines via activating the p38/MAPK signal pathway in THP-1 cells. *Mol Immunol* 105, 233–239. doi: 10.1016/j.molimm.2018.12.007.

Jiangwei, Y., Ericson, M., Frank, M., and Rock, C. (2016). Enoyl-Acyl Carrier Protein Reductase I (FabI) Is Essential for the Intracellular Growth of *Listeria monocytogenes*. *Infect Immun* 84, 3597–3607. doi: 10.1128/IAI.00647-16.

Jiwani, S., Alvarado, S., Ohr, R. J., Romero, A., Nguyen, B., and Jewett, T. J. (2013). *Chlamydia trachomatis* Tarp Harbors Distinct G and F Actin Binding Domains That Bundle Actin Filaments. *J Bacteriol* 195, 708–716. doi: 10.1128/JB.01768-12.

John, E., and Gordon, F. B. (1946). Studies on antigenic relationships within the psittacosis-lymphogranuloma group of viruses. *J Bacteriol* 51, 617.

Johnson, C. M., and Fisher, D. J. (2013). Site-Specific, Insertional Inactivation of *incA* in *Chlamydia trachomatis* Using a Group II Intron. *PLoS One* 8, e83989. doi: 10.1371/journal.pone.0083989.

Jorgensen, I., and Valdivia, R. H. (2008). Pmp-Like Proteins Pls1 and Pls2 Are Secreted into the Lumen of the *Chlamydia trachomatis* Inclusion. *Infect Immun* 76, 3940–3950. doi: 10.1128/IAI.00632-08.

Joseph, S. J., Marti, H., Didelot, X., Read, T. D., and Dean, D. (2016). Tetracycline Selective Pressure and Homologous Recombination Shape the Evolution of *Chlamydia suis*: A Recently Identified Zoonotic Pathogen. *Genome Biol Evol* 8, 2613–2623. doi: 10.1093/gbe/evw182.

Kafsack, B. F. C., Pena, J. D. O., Coppens, I., Ravindran, S., Boothroyd, J. C., and Carruthers, V. B. (2009). Rapid membrane disruption by a perforin-like protein facilitates parasite exit from host cells. *Science* 323, 530–533. doi: 10.1126/science.1165740.

Kahane, S., Kimmel, N., and Friedman, M. G. (2002). The growth cycle of *Simkania negevensis*. *Microbiology (Reading)* 148, 735–742. doi: 10.1099/00221287-148-3-735.

Kannan, R. M., Gérard, H. C., Mishra, M. K., Mao, G., Wang, S., Hali, M., *et al.* (2013). Dendrimer-enabled transformation of *Chlamydia trachomatis*. *Microb Pathog* 65, 29–35. doi: 10.1016/j.micpath.2013.08.003.

Kari, L., Goheen, M. M., Randall, L. B., Taylor, L. D., Carlson, J. H., Whitmire, W. M., *et al.* (2011a). Generation of targeted *Chlamydia trachomatis* null mutants. *Proceedings of the National Academy of Sciences* 108, 7189–7193. doi: 10.1073/pnas.1102229108.

Kari, L., Whitmire, W. M., Olivares-Zavaleta, N., Goheen, M. M., Taylor, L. D., Carlson, J. H., *et al.* (2011b). A live-attenuated chlamydial vaccine protects against trachoma in nonhuman primates. *Journal of Experimental Medicine* 208, 2217–2223. doi: 10.1084/jem.20111266.

Katoh, K., and Standley, D. M. (2014). MAFFT: iterative refinement and additional methods. *Methods Mol Biol* 1079, 131–146. doi: 10.1007/978-1-62703-646-7_8.

Keb, G., Ferrell, J., Scanlon, K. R., Jewett, T. J., and Fields, K. A. (2021). *Chlamydia trachomatis* TmeA Directly Activates N-WASP To Promote Actin Polymerization and Functions Synergistically with TarP during Invasion. *mBio* 12. doi: 10.1128/mBio.02861-20.

Kebbi-Beghdadi, C., Pilloux, L., Croxatto, A., Tosetti, N., Pillonel, T., and Greub, G. (2019). A predation assay using amoebae to screen for virulence factors unearthed the first *W. chondrophila* inclusion membrane protein. *Sci Rep* 9, 19485. doi: 10.1038/s41598-019-55511-1.

Keller, O., Kollmar, M., Stanke, M., and Waack, S. (2011). A novel hybrid gene prediction method employing protein multiple sequence alignments. *Bioinformatics* 27, 757–763. doi: 10.1093/bioinformatics/btr010.

Kerr, M. C., Gomez, G. A., Ferguson, C., Tanzer, M. C., Murphy, J. M., Yap, A. S., *et al.* (2017). Laser-mediated rupture of chlamydial inclusions triggers pathogen egress and host cell necrosis. *Nat Commun* 8, 14729. doi: 10.1038/ncomms14729.

Knittler, M. R., and Sachse, K. (2015). *Chlamydia psittaci*: update on an underestimated zoonotic agent. *Pathog Dis* 73, 1–15. doi: 10.1093/femspd/ftu007.

Koch, R. D., Hörner, E. M., Münch, N., Maier, E., and Kozjak-Pavlovic, V. (2020). Modulation of Host Cell Death and Lysis Are Required for the Release of *Simkania negevensis*. *Front Cell Infect Microbiol* 10, 594932. doi: 10.3389/fcimb.2020.594932.

Kokes, M., Dunn, J. D., Granek, J. A., Nguyen, B. D., Barker, J. R., Valdivia, R. H., *et al.* (2015). Integrating Chemical Mutagenesis and Whole-Genome Sequencing as a Platform for Forward and Reverse Genetic Analysis of *Chlamydia*. *Cell Host Microbe* 17, 716–725. doi: 10.1016/j.chom.2015.03.014.

Köstlbacher, S., Collingro, A., Halter, T., Schulz, F., Jungbluth, S. P., and Horn, M. (2021). Pangenomics reveals alternative environmental lifestyles among chlamydiae. *Nat Commun* 12, 4021. doi: 10.1038/s41467-021-24294-3.

Krall, R., Zhang, Y., and Barbieri, J. T. (2004). Intracellular Membrane Localization of *Pseudomonas* ExoS and *Yersinia* YopE in Mammalian Cells. *Journal of Biological Chemistry* 279, 2747–2753. doi: 10.1074/jbc.M301963200.

Kumar, Y., Cocchiaro, J., and Valdivia, R. H. (2006). The Obligate Intracellular Pathogen *Chlamydia trachomatis* Targets Host Lipid Droplets. *Current Biology* 16, 1646–1651. doi: 10.1016/j.cub.2006.06.060.

Kumar, Y., and Valdivia, R. H. (2008). Actin and Intermediate Filaments Stabilize the *Chlamydia trachomatis* Vacuole by Forming Dynamic Structural Scaffolds. *Cell Host Microbe* 4, 159–169. doi: 10.1016/j.chom.2008.05.018.

Kuznetsova, E., Proudfoot, M., Gonzalez, C. F., Brown, G., Omelchenko, M. v, Borozan, I., *et al.* (2006). Genome-wide Analysis of Substrate Specificities of the *Escherichia coli* Haloacid Dehalogenase-like Phosphatase Family. *Journal of Biological Chemistry* 281, 36149–36161. doi: 10.1074/jbc.M605449200.

LaBrie, S. D., Dimond, Z. E., Harrison, K. S., Baid, S., Wickstrum, J., Suchland, R. J., *et al.* (2019). Transposon Mutagenesis in *Chlamydia trachomatis* Identifies CT339 as a ComEC Homolog Important for DNA Uptake and Lateral Gene Transfer. *mBio* 10. doi: 10.1128/mBio.01343-19.

Lagkouvardos, I., Weinmaier, T., Lauro, F. M., Cavicchioli, R., Rattei, T., and Horn, M. (2014). Integrating metagenomic and amplicon databases to resolve the phylogenetic and ecological diversity of the *Chlamydiae*. *ISME J* 8, 115–125. doi: 10.1038/ismej.2013.142.

Lane, B. J., Mutchler, C., al Khodor, S., Grieshaber, S. S., and Carabeo, R. A. (2008). Chlamydial Entry Involves TARP Binding of Guanine Nucleotide Exchange Factors. *PLoS Pathog* 4, e1000014. doi: 10.1371/journal.ppat.1000014.

Lanjouw, E., Ouburg, S., de Vries, H., Stary, A., Radcliffe, K., and Unemo, M. (2016). 2015 European guideline on the management of *Chlamydia trachomatis* infections. *Int J STD AIDS* 27, 333–348. doi: 10.1177/0956462415618837.

Laroucau, K., Ortega, N., Vorimore, F., Aaziz, R., Mitura, A., Szymanska-Czerwinska, M., *et al.* (2020). Detection of a novel *Chlamydia* species in captive spur-thighed tortoises (*Testudo graeca*) in southeastern Spain and proposal of *Candidatus* Chlamydia testudinis. *Syst Appl Microbiol* 43, 126071. doi: 10.1016/j.syapm.2020.126071.

Laroucau, K., Vorimore, F., Aaziz, R., Solmonson, L., Hsia, R. C., Bavoil, P. M., *et al.* (2019). *Chlamydia buteonis*, a new *Chlamydia* species isolated from a red-shouldered hawk. *Syst Appl Microbiol* 42, 125997. doi: 10.1016/j.syapm.2019.06.002.

le Negrate, G., Krieg, A., Faustin, B., Loeffler, M., Godzik, A., Krajewski, S., *et al.* (2008). ChlaDub1 of *Chlamydia trachomatis* suppresses NF-κB activation and inhibits IκBα ubiquitination and degradation. *Cell Microbiol* 10, 1879–1892. doi: 10.1111/j.1462-5822.2008.01178.x.

Ledvina, H. E., Kelly, K. A., Eshraghi, A., Plemel, R. L., Peterson, S. B., Lee, B., *et al.* (2018). A Phosphatidylinositol 3-Kinase Effector Alters Phagosomal Maturation to Promote Intracellular Growth of *Francisella*. *Cell Host Microbe* 24, 285-295.e8. doi: 10.1016/j.chom.2018.07.003.

Lee, J. K., Enciso, G. A., Boassa, D., Chander, C. N., Lou, T. H., Pairawan, S. S., *et al.* (2018). Replication-dependent size reduction precedes differentiation in *Chlamydia trachomatis*. *Nat Commun* 9, 1–9. doi: 10.1038/s41467-017-02432-0.

Lei, L., Chunfu, Y., John, P. M., Margery, S., David, D., Li, M., *et al.* (2021). A Chlamydial Plasmid-Dependent Secretion System for the Delivery of Virulence Factors to the Host Cytosol. *mBio* 12, e01179-21. doi: 10.1128/mBio.01179-21.

Lei, L., Dong, X., Li, Z., and Zhong, G. (2013). Identification of a Novel Nuclear Localization Signal Sequence in *Chlamydia trachomatis*-Secreted Hypothetical Protein CT311. *PLoS One* 8, e64529. doi: 10.1371/journal.pone.0064529.

Lei, L., Qi, M., Budrys, N., Schenken, R., and Zhong, G. (2011). Localization of *Chlamydia trachomatis* hypothetical protein CT311 in host cell cytoplasm. *Microb Pathog* 51, 101–109. doi: 10.1016/j.micpath.2011.05.002.

Li, Z., Chen, C., Chen, D., Wu, Y., Zhong, Y., and Zhong, G. (2008a). Characterization of Fifty Putative Inclusion Membrane Proteins Encoded in the *Chlamydia trachomatis* Genome. *Infect Immun* 76, 2746–2757. doi: 10.1128/IAI.00010-08.

Li, Z., Chen, D., Zhong, Y., Wang, S., and Zhong, G. (2008b). The Chlamydial Plasmid-Encoded Protein Pgp3 Is Secreted into the Cytosol of *Chlamydia*-Infected Cells. *Infect Immun* 76, 3415–3428. doi: 10.1128/IAI.01377-07.

Lin, Y.-H., Doms, A. G., Cheng, E., Kim, B., Evans, T. R., and Machner, M. P. (2015). Host Cell-catalyzed S-Palmitoylation Mediates Golgi Targeting of the *Legionella* Ubiquitin Ligase GobX. *Journal of Biological Chemistry* 290, 25766–25781. doi: 10.1074/jbc.M115.637397.

Lindgren, S. W., Stojiljkovic, I., and Heffron, F. (1996). Macrophage killing is an essential virulence mechanism of *Salmonella* Typhimurium. *Proceedings of the National Academy of Sciences* 93, 4197–4201. doi: 10.1073/pnas.93.9.4197.

Liu, M., Wen, Y., Ding, H., and Zeng, H. (2022). Septic shock with *Chlamydia abortus* infection. *Lancet Infect Dis* 22, 912. doi: 10.1016/S1473-3099(21)00756-8.

Liu, Y., Chen, C., Gong, S., Hou, S., Qi, M., Liu, Q., *et al.* (2014). Transformation of *Chlamydia muridarum* Reveals a Role for Pgp5 in Suppression of Plasmid-Dependent Gene Expression. *J Bacteriol* 196, 989–998. doi: 10.1128/JB.01161-13.

Longbottom, D., and Coulter, L. J. (2003). Animal Chlamydioses and Zoonotic Implications. *J Comp Pathol* 128, 217–244. doi: 10.1053/jcpa.2002.0629.

Löwer, M., and Schneider, G. (2009). Prediction of Type III Secretion Signals in Genomes of Gram-Negative Bacteria. *PLoS One* 4, e5917. doi: 10.1371/journal.pone.0005917.

Lu, C., Lei, L., Peng, B., Tang, L., Ding, H., Gong, S., *et al.* (2013). *Chlamydia trachomatis* GlgA Is Secreted into Host Cell Cytoplasm. *PLoS One* 8, e68764. doi: 10.1371/journal.pone.0068764.

Lutter, E. I., Barger, A. C., Nair, V., and Hackstadt, T. (2013). *Chlamydia trachomatis* Inclusion Membrane Protein CT228 Recruits Elements of the Myosin Phosphatase Pathway to Regulate Release Mechanisms. *Cell Rep* 3, 1921–1931. doi: 10.1016/j.celrep.2013.04.027.

Lutter, E. I., Martens, C., and Hackstadt, T. (2012). Evolution and Conservation of Predicted Inclusion Membrane Proteins in *Chlamydiae*. *Comp Funct Genomics* 2012, 1–13. doi: 10.1155/2012/362104.

Ma, J., He, C., Huo, Z., Xu, Y., Arulanandam, B., Liu, Q., *et al.* (2020). The Cryptic Plasmid Improves *Chlamydia* Fitness in Different Regions of the Gastrointestinal Tract. *Infect Immun* 88, 1–10. doi: 10.1128/IAI.00860-19.

Manuela, D., Heather, H.-C., Michael, H., Maria, D. P., Antonietta, D. F., and A, M. G. S. (2014). Genome Sequence of *Chlamydia suis* MD56, Isolated from the Conjunctiva of a Weaned Piglet. *Genome Announc* 2, e00425-14. doi: 10.1128/genomeA.00425-14.

Marangoni, A., Foschi, C., Tartari, F., Gaspari, V., and Re, M. C. (2021). Lymphogranuloma venereum genovariants in men having sex with men in Italy. *Sex Transm Infect* 97, 441. doi: 10.1136/sextrans-2020-054700.

Marat, A. L., and Haucke, V. (2016). Phosphatidylinositol 3-phosphates—at the interface between cell signalling and membrane traffic. *EMBO J* 35, 561–579. doi: 10.15252/embj.201593564.

Marsh, J. W., Ong, V. A., Lott, W. B., Timms, P., Tyndall, J. da, and Huston, W. M. (2017). CtHtrA: the lynchpin of the chlamydial surface and a promising therapeutic target. *Future Microbiol* 12, 817–829. doi: 10.2217/fmb-2017-0017.

Matsumoto, A. (1973). Fine Structures of Cell Envelopes of *Chlamydia* Organisms as Revealed by Freeze-Etching and Negative Staining Techniques. *J Bacteriol* 116, 1355–1363. doi: 10.1128/jb.116.3.1355-1363.1973.

Matsumoto, A., Izutsu, H., Miyashita, N., and Ohuchi, M. (1998). Plaque Formation by and Plaque Cloning of *Chlamydia trachomatis* Biovar Trachoma. *J Clin Microbiol* 36, 3013–3019. doi: 10.1128/JCM.36.10.3013-3019.1998.

McKuen, M. J., Mueller, K. E., Bae, Y. S., and Fields, K. A. (2017). Fluorescence-Reported Allelic Exchange Mutagenesis Reveals a Role for *Chlamydia trachomatis* TmeA in Invasion That Is Independent of Host AHNAK. *Infect Immun* 85. doi: 10.1128/IAI.00640-17.

Mehlitz, A., Banhart, S., Hess, S., Selbach, M., and Meyer, T. F. (2008). Complex kinase requirements for *Chlamydia trachomatis* Tarp phosphorylation. *FEMS Microbiol Lett* 289, 233–240. doi: 10.1111/j.1574-6968.2008.01390.x.

Mehlitz, A., Banhart, S., Mäurer, A. P., Kaushansky, A., Gordus, A. G., Zielecki, J., *et al.* (2010). Tarp regulates early *Chlamydia*-induced host cell survival through interactions with the human adaptor protein SHC1. *Journal of Cell Biology* 190, 143–157. doi: 10.1083/jcb.200909095.

Mejuto, P., Boga, J. A., Junquera, M., Torreblanca, A., and Leiva, P. S. (2013). Genotyping *Chlamydia trachomatis* strains among men who have sex with men from a Northern Spain region: a cohort study. *BMJ Open* 3, e002330. doi: 10.1136/bmjopen-2012-002330.

Meyer, T. (2016). Diagnostic Procedures to Detect *Chlamydia trachomatis* Infections. *Microorganisms* 4, 25. doi: 10.3390/microorganisms4030025.

Miller, C. N., Smith, E. P., Cundiff, J. A., Knodler, L. A., Bailey Blackburn, J., Lupashin, V., *et al.* (2017). A *Brucella* Type IV Effector Targets the COG Tethering Complex to Remodel Host Secretory Traffic and Promote Intracellular Replication. *Cell Host Microbe* 22, 317-329.e7. doi: 10.1016/j.chom.2017.07.017.

Mirrashidi, K. M., Elwell, C. A., Verschueren, E., Johnson, J. R., Frando, A., von Dollen, J., *et al.* (2015). Global Mapping of the Inc-Human Interactome Reveals that Retromer Restricts *Chlamydia* Infection. *Cell Host Microbe* 18, 109–121. doi: 10.1016/j.chom.2015.06.004.

Misaghi, S., Balsara, Z. R., Catic, A., Spooner, E., Ploegh, H. L., and Starnbach, M. N. (2006). *Chlamydia trachomatis*-derived deubiquitinating enzymes in mammalian cells during infection. *Mol Microbiol* 61, 142–150. doi: 10.1111/j.1365-2958.2006.05199.x.

Mital, J., Lutter, E. I., Barger, A. C., Dooley, C. A., and Hackstadt, T. (2015). *Chlamydia trachomatis* inclusion membrane protein CT850 interacts with the dynein light chain DYNLT1 (Tctex1). *Biochem Biophys Res Commun* 462, 165–170. doi: 10.1016/j.bbrc.2015.04.116.

Miyairi, I., Mahdi, O. S., Ouellette, S. P., Belland, R. J., and Byrne, G. I. (2006). Different Growth Rates of *Chlamydia trachomatis* Biovars Reflect Pathotype. *J Infect Dis* 194, 350–357. doi: 10.1086/505432.

Moulder, J. W. (1966). The Relation of the Psittacosis Group (*Chlamydiae*) to Bacteria and Viruses. *Annu Rev Microbiol* 20, 107–130. doi: 10.1146/annurev.mi.20.100166.000543.

Mueller, K. E., and Fields, K. A. (2015). Application of β-Lactamase Reporter Fusions as an Indicator of Effector Protein Secretion during Infections with the Obligate Intracellular Pathogen *Chlamydia trachomatis*. *PLoS One* 10, e0135295. doi: 10.1371/journal.pone.0135295.

Mueller, K. E., Plano, G. v., and Fields, K. A. (2014). New frontiers in type III secretion biology: The *Chlamydia* perspective. *Infect Immun* 82, 2–9. doi: 10.1128/IAI.00917-13.

Mueller, K. E., Wolf, K., and Fields, K. A. (2016). Gene deletion by fluorescence-reported allelic exchange mutagenesis in *Chlamydia trachomatis*. *mBio* 7, e01817-15. doi: 10.1128/mBio.01817-15.

Murray, E. S. (1964). Guinea Pig Inclusion Conjunctivitis Virus: I. Isolation and Identification as a Member of the Psittacosis-Lymphogranuloma-trachoma Group. *J Infect Dis* 114, 1–12. doi: 10.1093/infdis/114.1.1.

Muschiol, S., Boncompain, G., Vromman, F., Dehoux, P., Normark, S., Henriques-Normark, B., *et al.* (2011). Identification of a Family of Effectors Secreted by the Type III Secretion System That Are Conserved in Pathogenic *Chlamydiae*. *Infect Immun* 79, 571–580. doi: 10.1128/IAI.00825-10.

Nachmias, N., Zusman, T., and Segal, G. (2019). Study of *Legionella* Effector Domains Revealed Novel and Prevalent Phosphatidylinositol 3-Phosphate Binding Domains. *Infect Immun* 87. doi: 10.1128/IAI.00153-19.

Neuendorf, E., Gajer, P., Bowlin, A. K., Marques, P. X., Ma, B., Yang, H., *et al.* (2015). *Chlamydia caviae* infection alters abundance but not composition of the guinea pig vaginal microbiota. *Pathog Dis* 73, ftv019. doi: 10.1093/femspd/ftv019.

Nguyen, B. D., and Valdivia, R. H. (2012). Virulence determinants in the obligate intracellular pathogen *Chlamydia trachomatis* revealed by forward genetic approaches. *Proceedings of the National Academy of Sciences* 109, 1263–1268. doi: 10.1073/pnas.1117884109.

Nguyen, B. D., and Valdivia, R. H. (2013). Forward genetic approaches in *Chlamydia trachomatis*. *J Vis Exp*, e50636. doi: 10.3791/50636.

Nguyen, L.-T., Schmidt, H. A., von Haeseler, A., and Minh, B. Q. (2015). IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol Biol Evol* 32, 268–274. doi: 10.1093/molbev/msu300.

Nguyen, P. H., Lutter, E. I., and Hackstadt, T. (2018). *Chlamydia trachomatis* inclusion membrane protein MrcA interacts with the inositol 1,4,5-trisphosphate receptor type 3 (ITPR3) to regulate extrusion formation. *PLoS Pathog* 14, e1006911. doi: 10.1371/journal.ppat.1006911.

Nicholson, T. L., Olinger, L., Chong, K., Schoolnik, G., and Stephens, R. S. (2003). Global Stage-Specific Gene Regulation during the Developmental Cycle of *Chlamydia trachomatis*. *J Bacteriol* 185, 3179–3189. doi: 10.1128/JB.185.10.3179-3189.2003.

Nunes, A., Borrego, M. J., and Gomes, J. P. (2013). Genomic features beyond *Chlamydia trachomatis* phenotypes: What do we think we know? *Infection, Genetics and Evolution* 16, 392–400. doi: 10.1016/j.meegid.2013.03.018.

Nunes, A., and Gomes, J. P. (2014). Evolution, phylogeny, and molecular epidemiology of *Chlamydia*. *Infection, Genetics and Evolution* 23, 49–64. doi: 10.1016/j.meegid.2014.01.029.

O'Connell, C. M., and Ferone, M. E. (2016). *Chlamydia trachomatis* Genital Infections. *Microb Cell* 3, 390–403. doi: 10.15698/mic2016.09.525.

O'Connell, C. M., Ingalls, R. R., Andrews, C. W., Scurlock, A. M., and Darville, T. (2007). Plasmid-Deficient *Chlamydia muridarum* Fail to Induce Immune Pathology and Protect against Oviduct Disease. *The Journal of Immunology* 179, 4027–4034. doi: 10.4049/jimmunol.179.6.4027.

O'Connell, C. M., and Nicks, K. M. (2006). A plasmid-cured *Chlamydia muridarum* strain displays altered plaque morphology and reduced infectivity in cell culture. *Microbiology (N Y)* 152, 1601–1607. doi: 10.1099/mic.0.28658-0.

Omsland, A., Sixt, B. S., Horn, M., and Hackstadt, T. (2014). Chlamydial metabolism revisited: interspecies metabolic variability and developmental stage-specific physiologic activities. *FEMS Microbiol Rev* 38, 779–801. doi: 10.1111/1574-6976.12059.

O'Neill, C. E., Skilton, R. J., Forster, J., Cleary, D. W., Pearson, S. A., Lampe, D. J., *et al.* (2021). An inducible transposon mutagenesis approach for the intracellular human pathogen *Chlamydia trachomatis*. *Wellcome Open Res* 6, 312. doi: 10.12688/wellcomeopenres.16068.1.

O'Neill, C. E., Skilton, R. J., Pearson, S. A., Filardo, S., Andersson, P., and Clarke, I. N. (2018). Genetic Transformation of a *C. trachomatis* Ocular Isolate With the Functional Tryptophan Synthase Operon Confers an Indole-Rescuable Phenotype. *Front Cell Infect Microbiol* 8. doi: 10.3389/fcimb.2018.00434.

Ouellette, S. P. (2018). Feasibility of a conditional knockout system for *Chlamydia* based on CRISPR interference. *Front Cell Infect Microbiol* 8. doi: 10.3389/fcimb.2018.00059.

Ouellette, S. P., Blay, E. A., Hatch, N. D., and Fisher-Marvin, L. A. (2021). CRISPR Interference To Inducibly Repress Gene Expression in *Chlamydia trachomatis*. *Infect Immun* 89. doi: 10.1128/IAI.00108-21.

Pais, S. V. (2018). Identification and characterization of CteG, a novel *Chlamydia trachomatis* type III secretion effector protein.

Pais, S. v, Key, C. E., Borges, V., Pereira, I. S., Gomes, J. P., Fisher, D. J., *et al.* (2019). CteG is a *Chlamydia trachomatis* effector protein that associates with the Golgi complex of infected host cells. *Sci Rep* 9, 6133. doi: 10.1038/s41598-019-42647-3.

Palmer, T., and Berks, B. C. (2012). The twin-arginine translocation (Tat) protein export pathway. *Nat Rev Microbiol* 10, 483–496. doi: 10.1038/nrmicro2814.

Panzetta, M. E., Luján, A. L., Bastidas, R. J., Damiani, M. T., Valdivia, R. H., and Saka, H. A. (2019). Ptr/CTL0175 Is Required for the Efficient Recovery of *Chlamydia trachomatis* From Stress Induced by Gamma-Interferon. *Front Microbiol* 10. doi: 10.3389/fmicb.2019.00756.

Park, E., and Rapoport, T. A. (2012). Mechanisms of Sec61/SecY-Mediated Protein Translocation Across Membranes. *Annu Rev Biophys* 41, 21–40. doi: 10.1146/annurev-biophys-050511-102312.

Parrett, C. J., Lenoci, R. v., Nguyen, B., Russell, L., and Jewett, T. J. (2016). Targeted Disruption of *Chlamydia trachomatis* Invasion by in Trans Expression of Dominant Negative Tarp Effectors. *Front Cell Infect Microbiol* 6. doi: 10.3389/fcimb.2016.00084.

Paschen, S. A., Christian, J. G., Vier, J., Schmidt, F., Walch, A., Ojcius, D. M., *et al.* (2008). Cytopathicity of *Chlamydia* is largely reproduced by expression of a single chlamydial protease. *Journal of Cell Biology* 182, 117–127. doi: 10.1083/jcb.200804023.

Patton, M. J., Chen, C.-Y., Yang, C., McCorrister, S., Grant, C., Westmacott, G., *et al.* (2018). Plasmid Negative Regulation of CPAF Expression Is Pgp4 Independent and Restricted to Invasive *Chlamydia trachomatis* Biovars. *mBio* 9, e02164-17. doi: 10.1128/mBio.02164-17.

Paul, B., Kim, H. S., Kerr, M. C., Huston, W. M., Teasdale, R. D., and Collins, B. M. (2017). Structural basis for the hijacking of endosomal sorting nexin proteins by *Chlamydia trachomatis*. *Elife* 6. doi: 10.7554/eLife.22311.

Paumet, F., Wesolowski, J., Garcia-Diaz, A., Delevoye, C., Aulner, N., Shuman, H. A., *et al.* (2009). Intracellular Bacteria Encode Inhibitory SNARE-Like Proteins. *PLoS One* 4, e7375. doi: 10.1371/journal.pone.0007375.

Pederson, K. J., Vallis, A. J., Aktories, K., Frank, D. W., and Barbieri, J. T. (1999). The amino-terminal domain of *Pseudomonas aeruginosa* ExoS disrupts actin filaments via small-molecular-weight GTP-binding proteins. *Mol Microbiol* 32, 393–401. doi: 10.1046/j.1365-2958.1999.01359.x.

Pennini, M. E., Perrinet, S., Dautry-Varsat, A., and Subtil, A. (2010). Histone Methylation by NUE, a Novel Nuclear Effector of the Intracellular Pathogen *Chlamydia trachomatis*. *PLoS Pathog* 6, e1000995. doi: 10.1371/journal.ppat.1000995.

Pereira, I. S., Pais, S. V., Borges, V., Borrego, M. J., Gomes, J. P., and Mota, L. J. (2022). The Type III Secretion Effector CteG Mediates Host Cell Lytic Exit of *Chlamydia trachomatis*. *Front Cell Infect Microbiol* 12. doi: 10.3389/fcimb.2022.902210.

Peterson, E. M., Markoff, B. A., Schachter, J., and de la Maza, L. M. (1990). The 7.5-kb plasmid present in *Chlamydia trachomatis* is not essential for the growth of this microorganism. *Plasmid* 23, 144–148. doi: 10.1016/0147-619X(90)90033-9.

Peuchant, O., Touati, A., Laurier-Nadalié, C., Hénin, N., Cazanave, C., Bébéar, C., *et al.* (2020). Prevalence of lymphogranuloma venereum among anorectal *Chlamydia trachomatis*-positive

MSM using pre-exposure prophylaxis for HIV. *Sex Transm Infect* 96, 615. doi: 10.1136/sextrans-2019-054346.

Phillips, D. M., Swenson, C. E., and Schachter, J. (1984). Ultrastructure of *Chlamydia trachomatis* infection of the mouse oviduct. *J Ultrastruct Res* 88, 244–256. doi: 10.1016/S0022-5320(84)90122-9.

Phillips, S., Quigley, B. L., and Timms, P. (2019). Seventy Years of *Chlamydia* Vaccine Research – Limitations of the Past and Directions for the Future. *Front Microbiol* 10. doi: 10.3389/fmicb.2019.00070.

Pichon, N., Guindre, L., Laroucau, K., Cantaloube, M., Nallatamby, A., and Parreau, S. (2020). *Chlamydia abortus* in Pregnant Woman with Acute Respiratory Distress Syndrome. *Emerging Infectious Disease journal* 26, 628. doi: 10.3201/eid2603.191417.

Pickett, M. A., Everson, J. S., Pead, P. J., and Clarke, I. N. (2005). The plasmids of *Chlamydia trachomatis* and *Chlamydophila pneumoniae* (N16): accurate determination of copy number and the paradoxical effect of plasmid-curing agents. *Microbiology (N Y)* 151, 893–903. doi: 10.1099/mic.0.27625-0.

Pike, C. M., Boyer-Andersen, R., Kinch, L. N., Caplan, J. L., and Neunuebel, M. R. (2019). The *Legionella* effector RavD binds phosphatidylinositol-3-phosphate and helps suppress endolysosomal maturation of the *Legionella*-containing vacuole. *Journal of Biological Chemistry* 294, 6405–6415. doi: 10.1074/jbc.RA118.007086.

Pillonel, T., Bertelli, C., and Greub, G. (2018). Environmental Metagenomic Assemblies Reveal Seven New Highly Divergent Chlamydial Lineages and Hallmarks of a Conserved Intracellular Lifestyle. *Front Microbiol* 9. doi: 10.3389/fmicb.2018.00079.

Pinho-Bandeira, T., Cabral Veríssimo, V., and Sá Machado, R. (2020). The epidemiology of chlamydia, gonorrhoea and syphilis in Portugal: where do we go from now? *Eur J Public Health* 30. doi: 10.1093/eurpub/ckaa165.538.

Pruneda, J. N., Bastidas, R. J., Bertsoulaki, E., Swatek, K. N., Santhanam, B., Clague, M. J., *et al.* (2018). A *Chlamydia* effector combining deubiquitination and acetylation activities induces Golgi fragmentation. *Nat Microbiol* 3, 1377–1384. doi: 10.1038/s41564-018-0271-y.

Prusty, B. K., Chowdhury, S. R., Gulve, N., and Rudel, T. (2018). Peptidase Inhibitor 15 (PI15) Regulates Chlamydial CPAF Activity. *Front Cell Infect Microbiol* 8, 183. doi: 10.3389/fcimb.2018.00183.

Radka, C. D., Frank, M. W., Yao, J., Seetharaman, J., Miller, D. J., and Rock, C. O. (2020). The genome of a Bacteroidetes inhabitant of the human gut encodes a structurally distinct enoyl-acyl carrier protein reductase (FabI). *Journal of Biological Chemistry* 295, 7635–7652. doi: 10.1074/jbc.RA120.013336.

Ramakers, B. P., Heijne, M., Lie, N., Le, T.-N., van Vliet, M., Claessen, V. P. J., *et al.* (2017). Zoonotic *Chlamydia caviae* Presenting as Community-Acquired Pneumonia. *New England Journal of Medicine* 377, 992–994. doi: 10.1056/NEJMc1702983.

Read, T. D., Brunham, R. C., Shen, C., Gill, S. R., Heidelberg, J. F., White, O., *et al.* (2000). Genome sequences of *Chlamydia trachomatis* MoPn and *Chlamydia pneumoniae* AR39. *Nucleic Acids Res* 28, 1397–1406. doi: 10.1093/nar/28.6.1397.

Ren, J., Wen, L., Gao, X., Jin, C., Xue, Y., and Yao, X. (2008). CSS-Palm 2.0: an updated software for palmitoylation sites prediction. *Protein Engineering, Design and Selection* 21, 639–644. doi: 10.1093/protein/gzn039.

Resh, M. D. (2006). Trafficking and signaling by fatty-acylated and prenylated proteins. *Nat Chem Biol* 2, 584–590. doi: 10.1038/nchembio834.

Ricci, S., Cevenini, R., Cosco, E., Comanducci, M., Ratti, G., and Scarlato, V. (1993). Transcriptional analysis of the *Chlamydia trachomatis* plasmid pCT identifies temporally regulated transcripts, anti-sense RNA and σ^{70}-selected promoters. *Mol Gen Genet* 237, 318–326. doi: 10.1007/BF00279434.

Richards, T. S., Knowlton, A. E., and Grieshaber, S. S. (2013). *Chlamydia trachomatis* homotypic inclusion fusion is promoted by host microtubule trafficking. *BMC Microbiol* 13, 1. doi: 10.1186/1471-2180-13-185.

Robertson, D. K., Gu, L., Rowe, R. K., and Beatty, W. L. (2009). Inclusion Biogenesis and Reactivation of Persistent *Chlamydia trachomatis* Requires Host Cell Sphingolipid Biosynthesis. *PLoS Pathog* 5, e1000664. doi: 10.1371/journal.ppat.1000664.

Robinson, J. T., Thorvaldsdóttir, H., Winckler, W., Guttman, M., Lander, E. S., Getz, G., *et al.* (2011). Integrative genomics viewer. *Nat Biotechnol* 29, 24–26. doi: 10.1038/nbt.1754.

Rockey, D. D., Scidmore, M. A., Bannantine, J. P., and Brown, W. J. (2002). Proteins in the chlamydial inclusion membrane. *Microbes Infect* 4, 333–340. doi: 10.1016/S1286-4579(02)01546-0.

Rodriguez-Dominguez, M., Gonzalez-Alba, J. M., Puerta, T., Menendez, B., Sanchez-Diaz, A. M., Canton, R., *et al.* (2015). High Prevalence of Co-Infections by Invasive and Non-Invasive *Chlamydia trachomatis* Genotypes during the Lymphogranuloma Venereum Outbreak in Spain. *PLoS One* 10, e0126145. doi: 10.1371/journal.pone.0126145.

Ronzone, E., Wesolowski, J., Bauler, L. D., Bhardwaj, A., Hackstadt, T., and Paumet, F. (2014). An α-Helical Core Encodes the Dual Functions of the Chlamydial Protein IncA. *Journal of Biological Chemistry* 289, 33469–33480. doi: 10.1074/jbc.M114.592063.

Roulis, E., Polkinghorne, A., and Timms, P. (2013). *Chlamydia pneumoniae*: modern insights into an ancient pathogen. *Trends Microbiol* 21, 120–128. doi: 10.1016/j.tim.2012.10.009.

Rowley, J., vander Hoorn, S., Korenromp, E., Low, N., Unemo, M., Abu-Raddad, L. J., *et al.* (2019). Chlamydia, gonorrhoea, trichomoniasis and syphilis: global prevalence and incidence estimates, 2016. *Bull World Health Organ* 97, 548-562P. doi: 10.2471/BLT.18.228486.

Rzomp, K. A., Moorhead, A. R., and Scidmore, M. A. (2006). The GTPase Rab4 Interacts with *Chlamydia trachomatis* Inclusion Membrane Protein CT229. *Infect Immun* 74, 5362–5373. doi: 10.1128/IAI.00539-06.

Sachse, K., Laroucau, K., Riege, K., Wehner, S., Dilcher, M., Creasy, H. H., *et al.* (2014). Evidence for the existence of two new members of the family *Chlamydiaceae* and proposal of *Chlamydia avium* sp. nov. and *Chlamydia gallinacea* sp. nov. *Syst Appl Microbiol* 37, 79–88. doi: 10.1016/j.syapm.2013.12.004.

Sait, M., Livingstone, M., Clark, E. M., Wheelhouse, N., Spalding, L., Markey, B., *et al.* (2014). Genome sequencing and comparative analysis of three *Chlamydia pecorum* strains associated with different pathogenic outcomes. *BMC Genomics* 15, 23. doi: 10.1186/1471-2164-15-23.

Saka, H. A., Thompson, J. W., Chen, Y.-S., Dubois, L. G., Haas, J. T., Moseley, A., *et al.* (2015). *Chlamydia trachomatis* Infection Leads to Defined Alterations to the Lipid Droplet Proteome in Epithelial Cells. *PLoS One* 10, e0124630. doi: 10.1371/journal.pone.0124630.

Saka, H. A., Thompson, J. W., Chen, Y.-S., Kumar, Y., Dubois, L. G., Moseley, M. A., *et al.* (2011). Quantitative proteomics reveals metabolic and pathogenic properties of *Chlamydia trachomatis* developmental forms. *Mol Microbiol* 82, 1185–1203. doi: 10.1111/j.1365-2958.2011.07877.x.

Salcedo, S. P. (2003). SseG, a virulence protein that targets *Salmonella* to the Golgi network. *EMBO J* 22, 5003–5014. doi: 10.1093/emboj/cdg517.

Samudrala, R., Heffron, F., and McDermott, J. E. (2009). Accurate Prediction of Secreted Substrates and Identification of a Conserved Putative Secretion Signal for Type III Secretion Systems. *PLoS Pathog* 5, e1000375. doi: 10.1371/journal.ppat.1000375.

Sanders, M. (1940). Studies on the Cultivation of the Virus of Lymphogranuloma Venereum. *Journal of Experimental Medicine* 71, 113–128. doi: 10.1084/jem.71.1.113.

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., *et al.* (2012). Fiji: an open-source platform for biological-image analysis. *Nat Methods* 9, 676–682. doi: 10.1038/nmeth.2019.

Schroeder, G. N., Aurass, P., Oates, C. v., Tate, E. W., Hartland, E. L., Flieger, A., *et al.* (2015). *Legionella pneumophila* Effector LpdA Is a Palmitoylated Phospholipase D Virulence Factor. *Infect Immun* 83, 3989–4002. doi: 10.1128/IAI.00785-15.

Scidmore, M. A. (2005). Cultivation and Laboratory Maintenance of *Chlamydia trachomatis*. *Curr Protoc Microbiol* Chapter 11, Unit 11A.1. doi: 10.1002/9780471729259.mc11a01s00.

Scidmore, M. A., and Hackstadt, T. (2001). Mammalian 14-3-3β associates with the *Chlamydia trachomatis* inclusion membrane via its interaction with IncG. *Mol Microbiol* 39, 1638–1650. doi: 10.1046/j.1365-2958.2001.02355.x.

Scidmore-Carlson, M. A., Shaw, E. I., Dooley, C. A., Fischer, E. R., and Hackstadt, T. (1999). Identification and characterization of a *Chlamydia trachomatis* early operon encoding four novel inclusion membrane proteins. *Mol Microbiol* 33, 753–765. doi: 10.1046/j.1365-2958.1999.01523.x.

Shao, L., Zhang, T., Melero, J., Huang, Y., Liu, Y., Liu, Q., *et al.* (2018). The Genital Tract Virulence Factor pGP3 Is Essential for *Chlamydia muridarum* Colonization in the Gastrointestinal Tract. *Infect Immun* 86. doi: 10.1128/IAI.00429-17.

Sharma, M., Machuy, N., Böhme, L., Karunakaran, K., Mäurer, A. P., Meyer, T. F., *et al.* (2011). HIF-1α is involved in mediating apoptosis resistance to *Chlamydia trachomatis*-infected cells. *Cell Microbiol* 13, 1573–1585. doi: 10.1111/j.1462-5822.2011.01642.x.

Shaw, E. I., Dooley, C. A., Fischer, E. R., Scidmore, M. A., Fields, K. A., and Hackstadt, T. (2000). Three temporal classes of gene expression during the *Chlamydia trachomatis* developmental cycle. *Mol Microbiol* 37, 913–925. doi: 10.1046/j.1365-2958.2000.02057.x.

Shaw, J. H., Key, C. E., Snider, T. A., Sah, P., Shaw, E. I., Fisher, D. J., *et al.* (2018). Genetic Inactivation of *Chlamydia trachomatis* Inclusion Membrane Protein CT228 Alters MYPT1 Recruitment, Extrusion Production, and Longevity of Infection. *Front Cell Infect Microbiol* 8, 415. doi: 10.3389/fcimb.2018.00415.

Sherry, J., Dolat, L., McMahon, E., Swaney, D. L., Bastidas, R. J., Johnson, J. R., *et al.* (2022). *Chlamydia trachomatis* effector Dre1 interacts with dynactin to reposition host organelles during infection. *bioRxiv*. doi: 10.1101/2022.04.15.488217.

Sigar, I. M., Schripsema, J. H., Wang, Y., Clarke, I. N., Cutcliffe, L. T., Seth-Smith, H. M. B., *et al.* (2014). Plasmid deficiency in urogenital isolates of *Chlamydia trachomatis* reduces infectivity and virulence in a mouse model. *Pathog Dis* 70, 61–69. doi: 10.1111/2049-632X.12086.

Sisko, J. L., Spaeth, K., Kumar, Y., and Valdivia, R. H. (2006). Multifunctional analysis of *Chlamydia*-specific genes in a yeast expression system. *Mol Microbiol* 60, 51–66. doi: 10.1111/j.1365-2958.2006.05074.x.

Sixt, B. S. (2021). Host cell death during infection with *Chlamydia*: a double-edged sword. *FEMS Microbiol Rev* 45, fuaa043. doi: 10.1093/femsre/fuaa043.

Sixt, B. S., Bastidas, R. J., Finethy, R., Baxter, R. M., Carpenter, V. K., Kroemer, G., *et al.* (2017). The *Chlamydia trachomatis* Inclusion Membrane Protein CpoS Counteracts STING-Mediated Cellular Surveillance and Suicide Programs. *Cell Host Microbe* 21, 113–121. doi: 10.1016/j.chom.2016.12.002.

Snavely, E. A., Kokes, M., Dunn, J. D., Saka, H. A., Nguyen, B. D., Bastidas, R. J., *et al.* (2014). Reassessing the role of the secreted protease CPAF in *Chlamydia trachomatis* infection through genetic approaches. *Pathog Dis* 71, 336–351. doi: 10.1111/2049-632X.12179.

Sobocińska, J., Roszczenko-Jasińska, P., Ciesielska, A., and Kwiatkowska, K. (2018). Protein Palmitoylation and Its Role in Bacterial and Viral Infections. *Front Immunol* 8. doi: 10.3389/fimmu.2017.02003.

Song, L., Carlson, J. H., Whitmire, W. M., Kari, L., Virtaneva, K., Sturdevant, D. E., *et al.* (2013). *Chlamydia trachomatis* Plasmid-Encoded Pgp4 Is a Transcriptional Regulator of Virulence-Associated Genes. *Infect Immun* 81, 636–644. doi: 10.1128/IAI.01305-12.

Spaeth, K. E., Chen, Y.-S., and Valdivia, R. H. (2009). The *Chlamydia* Type III Secretion System C-ring Engages a Chaperone-Effector Protein Complex. *PLoS Pathog* 5, e1000579-. doi: 10.1371/journal.ppat.1000579.

Stairs, C. W., Dharamshi, J. E., Tamarit, D., Eme, L., Jørgensen, S. L., Spang, A., *et al.* (2020). Chlamydial contribution to anaerobic metabolism during eukaryotic evolution. *Sci Adv* 6, eabb7258. doi: 10.1126/sciadv.abb7258.

Staub, E., Marti, H., Biondi, R., Levi, A., Donati, M., Leonard, C. A., *et al.* (2018). Novel *Chlamydia* species isolated from snakes are temperature-sensitive and exhibit decreased susceptibility to azithromycin. *Sci Rep* 8, 5660. doi: 10.1038/s41598-018-23897-z.

Steiert, B., Icardi, C. M., Faris, R., Klingelhutz, A. J., Yau, P. M., and Weber, M. M. (2022). The *Chlamydia trachomatis* type III secreted effector protein CteG induces centrosome amplification through interactions with centrin-2. *bioRxiv*. doi: 10.1101/2022.06.23.496711.

Stephens, R. S., Kalman, S., Lammel, C., Fan, J., Marathe, R., Aravind, L., *et al.* (1998). Genome sequence of an obligate intracellular pathogen of humans: *Chlamydia trachomatis*. *Science* 282, 754–759. doi: 10.1126/science.282.5389.754.

Stephens, R. S., Myers, G., Eppinger, M., and Bavoil, P. M. (2009). Divergence without difference: phylogenetics and taxonomy of *Chlamydia* resolved. *FEMS Immunol Med Microbiol* 55, 115–119. doi: 10.1111/j.1574-695X.2008.00516.x.

Stephens, R. S., Sanchez-Pescador, R., Wagar, E. A., Inouye, C., and Urdea, M. S. (1987). Diversity of *Chlamydia trachomatis* major outer membrane protein genes. *J Bacteriol* 169, 3879–3885. doi: 10.1128/jb.169.9.3879-3885.1987.

Subtil, A., Delevoye, C., Balañá, M. E., Tastevin, L., Perrinet, S., and Dautry-Varsat, A. (2005). A directed screen for chlamydial proteins secreted by a type III mechanism identifies a translocated protein and numerous other new candidates. *Mol Microbiol* 56, 1636–1647. doi: 10.1111/j.1365-2958.2005.04647.x.

Suchland, R. J., Rockey, D. D., Bannantine, J. P., and Stamm, W. E. (2000). Isolates of *Chlamydia trachomatis* That Occupy Nonfusogenic Inclusions Lack IncA, a Protein Localized to the Inclusion Membrane. *Infect Immun* 68, 360–367. doi: 10.1128/IAI.68.1.360-367.2000.

Sun, Q., Yong, X., Sun, X., Yang, F., Dai, Z., Gong, Y., *et al.* (2017). Structural and functional insights into sorting nexin 5/6 interaction with bacterial effector IncE. *Signal Transduct Target Ther* 2, 17030. doi: 10.1038/sigtrans.2017.30.

Tagini, F., and Greub, G. (2018). Infections à chlamydias : épidémiologie, pathogenèse, diagnostic et traitements. *Rev Med Suisse* 14, 1620–1625. doi: 10.53738/REVMED.2018.14.618.1620.

Tam, J. E., Davis, C. H., and Wyrick, P. B. (1994). Expression of recombinant DNA introduced into *Chlamydia trachomatis* by electroporation. *Can J Microbiol* 40, 583–591. doi: 10.1139/m94-093.

Tarbet, H. J., Dolat, L., Smith, T. J., Condon, B. M., O'Brien, E. T., Valdivia, R. H., *et al.* (2018). Site-specific glycosylation regulates the form and function of the intermediate filament cytoskeleton. *Elife* 7. doi: 10.7554/eLife.31807.

Taylor, H. R. (2008). "Trachoma is an ancient disease and a weapon of mass destruction," in *Trachoma: a blinding scourge from the Bronze Age to the twenty-first century*, ed. H. R. Taylor (East Melbourne, Australia: Centre for Eye Research Australia), 1–15.

Taylor, H. R., Burton, M. J., Haddad, D., West, S., and Wright, H. (2014). Trachoma. *The Lancet* 384, 2142–2152. doi: 10.1016/S0140-6736(13)62182-0.

Taylor-Brown, A., Bachmann, N. L., Borel, N., and Polkinghorne, A. (2016). Culture-independent genomic characterisation of *Candidatus* Chlamydia sanzinia, a novel uncultivated bacterium infecting snakes. *BMC Genomics* 17, 710. doi: 10.1186/s12864-016-3055-x.

Taylor-Brown, A., Madden, D., and Polkinghorne, A. (2018a). Culture-independent approaches to chlamydial genomics. *Microb Genom* 4. doi: 10.1099/mgen.0.000145.

Taylor-Brown, A., Pillonel, T., Greub, G., Vaughan, L., Nowak, B., and Polkinghorne, A. (2018b). Metagenomic Analysis of Fish-Associated *Ca.* Parilichlamydiaceae Reveals Striking Metabolic Similarities to the Terrestrial *Chlamydiaceae*. *Genome Biol Evol* 10, 2587–2595. doi: 10.1093/gbe/evy195.

Taylor-Brown, A., Spang, L., Borel, N., and Polkinghorne, A. (2017). Culture-independent metagenomics supports discovery of uncultivable bacteria within the genus *Chlamydia*. *Sci Rep* 7, 10661. doi: 10.1038/s41598-017-10757-5.

Taylor-Brown, A., Vaughan, L., Greub, G., Timms, P., and Polkinghorne, A. (2015). Twenty years of research into *Chlamydia*-like organisms: A revolution in our understanding of the biology and pathogenicity of members of the phylum *Chlamydiae*. *Pathog Dis* 73. doi: 10.1093/femspd/ftu009.

Thalmann, J., Janik, K., May, M., Sommer, K., Ebeling, J., Hofmann, F., *et al.* (2010). Actin Re-Organization Induced by *Chlamydia trachomatis* Serovar D - Evidence for a Critical Role of the Effector Protein CT166 Targeting Rac. *PLoS One* 5, e9887. doi: 10.1371/journal.pone.0009887.

Thomas, N. S., Lusher, M., Storey, C. C., and Clarke, I. N. (1997). Plasmid diversity in *Chlamydia*. *Microbiology (N Y)* 143, 1847–1854. doi: 10.1099/00221287-143-6-1847.

Thompson, C. C., Griffiths, C., Nicod, S. S., Lowden, N. M., Wigneshweraraj, S., Fisher, D. J., *et al.* (2015). The Rsb Phosphoregulatory Network Controls Availability of the Primary Sigma Factor in *Chlamydia trachomatis* and Influences the Kinetics of Growth and Development. *PLoS Pathog* 11, e1005125. doi: 10.1371/journal.ppat.1005125.

Thomson, N. R., Holden, M. T. G., Carder, C., Lennard, N., Lockey, S. J., Marsh, P., *et al.* (2008). *Chlamydia trachomatis*: Genome sequence analysis of lymphogranuloma venereum isolates. *Genome Res* 18, 161–171. doi: 10.1101/gr.7020108.

Todd, W. J., and Caldwell, H. D. (1985). The interaction of *Chlamydia trachomatis* with host cells: ultrastructural studies of the mechanism of release of a biovar II strain from HeLa 229 cells. *J Infect Dis* 151, 1037–1044. doi: 10.1093/infdis/151.6.1037.

Tumurkhuu, G., Dagvadorj, J., Porritt, R. A., Crother, T. R., Shimada, K., Tarling, E. J., *et al.* (2018). *Chlamydia pneumoniae* Hijacks a Host Autoregulatory IL-1β Loop to Drive Foam Cell Formation and Accelerate Atherosclerosis. *Cell Metab* 28, 432-448.e4. doi: 10.1016/j.cmet.2018.05.027.

Ulmschneider, M. B., Ulmschneider, J. P., Schiller, N., Wallace, B. A., von Heijne, G., and White, S. H. (2014). Spontaneous transmembrane helix insertion thermodynamically mimics translocon-guided insertion. *Nat Commun* 5, 4863. doi: 10.1038/ncomms5863.

Uphoff, C. C., and Drexler, H. G. (2011). Detecting *Mycoplasma* Contamination in Cell Cultures by Polymerase Chain Reaction. *Methods in Molecular Biology* 731, 93–103. doi: 10.1007/978-1-61779-80-5_8.

Uwamahoro, N., Verma-Gaur, J., Shen, H.-H., Qu, Y., Lewis, R., Lu, J., *et al.* (2014). The Pathogen *Candida albicans* Hijacks Pyroptosis for Escape from Macrophages. *mBio* 5, e00003-14. doi: 10.1128/mBio.00003-14.

Vicente-Manzanares, M., Ma, X., Adelstein, R. S., and Horwitz, A. R. (2009). Non-muscle myosin II takes centre stage in cell adhesion and migration. *Nat Rev Mol Cell Biol* 10, 778–790. doi: 10.1038/nrm2786.

Volceanov, L., Herbst, K., Biniossek, M., Schilling, O., Haller, D., Nölke, T., *et al.* (2014). Septins Arrange F-Actin-Containing Fibers on the *Chlamydia trachomatis* Inclusion and Are Required for Normal Release of the Inclusion by Extrusion. *mBio* 5. doi: 10.1128/mBio.01802-14.

Vorimore, F., Hölzer, M., Liebler-Tenorio, E. M., Barf, L.-M., Delannoy, S., Vittecoq, M., *et al.* (2021). Evidence for the existence of a new genus *Chlamydiifrater* gen. nov. inside the family *Chlamydiaceae* with two new species isolated from flamingo (*Phoenicopterus roseus*): *Chlamydiifrater phoenicopteri* sp. nov. and *Chlamydiifrater volucris* sp. nov. *Syst Appl Microbiol* 44, 126200. doi: 10.1016/j.syapm.2021.126200.

Vorimore, F., Hsia, R., Huot-Creasy, H., Bastian, S., Deruyter, L., Passet, A., *et al.* (2013). Isolation of a New *Chlamydia* species from the Feral Sacred Ibis (*Threskiornis aethiopicus*): *Chlamydia ibidis*. *PLoS One* 8, e74823. doi: 10.1371/journal.pone.0074823.

Vromman, F., Perrinet, S., Gehre, L., and Subtil, A. (2016). The DUF582 Proteins of *Chlamydia trachomatis* Bind to Components of the ESCRT Machinery, Which Is Dispensable for Bacterial Growth In vitro. *Front Cell Infect Microbiol* 6, 123. doi: 10.3389/fcimb.2016.00123.

Wang, X., Hybiske, K., and Stephens, R. S. (2018). Direct visualization of the expression and localization of chlamydial effector proteins within infected host cells. *Pathog Dis* 76. doi: 10.1093/femspd/fty011.

Wang, Y., Kahane, S., Cutcliffe, L. T., Skilton, R. J., Lambden, P. R., and Clarke, I. N. (2011). Development of a Transformation System for *Chlamydia trachomatis*: Restoration of Glycogen Biosynthesis by Acquisition of a Plasmid Shuttle Vector. *PLoS Pathog* 7, e1002258. doi: 10.1371/journal.ppat.1002258.

Weber, M. M., Bauler, L. D., Lam, J., and Hackstadt, T. (2015). Expression and Localization of Predicted Inclusion Membrane Proteins in *Chlamydia trachomatis*. *Infect Immun* 83, 4710–4718. doi: 10.1128/IAI.01075-15.

Weber, M. M., Lam, J. L., Dooley, C. A., Noriea, N. F., Hansen, B. T., Hoyt, F. H., *et al.* (2017). Absence of Specific *Chlamydia trachomatis* Inclusion Membrane Proteins Triggers Premature Inclusion Membrane Lysis and Host Cell Death. *Cell Rep* 19, 1406–1417. doi: 10.1016/j.celrep.2017.04.058.

Weber, M. M., Noriea, N. F., Bauler, L. D., Lam, J. L., Sager, J., Wesolowski, J., *et al.* (2016). A Functional Core of IncA Is Required for *Chlamydia trachomatis* Inclusion Fusion. *J Bacteriol* 198, 1347–1355. doi: 10.1128/JB.00933-15.

Wesolowski, J., Weber, M. M., Nawrotek, A., Dooley, C. A., Calderon, M., st. Croix, C. M., *et al.* (2017). *Chlamydia* Hijacks ARF GTPases To Coordinate Microtubule Posttranslational Modifications and Golgi Complex Positioning. *mBio* 8. doi: 10.1128/mBio.02280-16.

Wickstrum, J., Sammons, L. R., Restivo, K. N., and Hefty, P. S. (2013). Conditional Gene Expression in *Chlamydia trachomatis* Using the Tet System. *PLoS One* 8, e76743. doi: 10.1371/journal.pone.0076743.

Williams, D. M., Schachter, J., Drutz, D. J., and Sumaya, C. v (1981). Pneumonia Due to *Chlamydia trachomatis* in the Immunocompromised (Nude) Mouse. *J Infect Dis* 143, 238–241. Available at: http://www.jstor.org/stable/30113553.

Wilson, A. C., Wu, C. C., Yates, J. R., and Tan, M. (2005). Chlamydial GroEL Autoregulates Its Own Expression through Direct Interactions with the HrcA Repressor Protein. *J Bacteriol* 187, 7535–7542. doi: 10.1128/JB.187.21.7535-7542.2005.

Wons, J., Meiller, R., Bergua, A., Bogdan, C., and Geißdörfer, W. (2017). Follicular Conjunctivitis due to *Chlamydia felis* - Case Report, Review of the Literature and Improved Molecular Diagnostics. *Front Med (Lausanne)* 4. doi: 10.3389/fmed.2017.00105.

Wu, X., Lei, L., Gong, S., Chen, D., Flores, R., and Zhong, G. (2011). The chlamydial periplasmic stress response serine protease cHtrA is secreted into host cell cytosol. *BMC Microbiol* 11, 87. doi: 10.1186/1471-2180-11-87.

Yang, C., Kari, L., Lei, L., Carlson, J. H., Ma, L., Couch, C. E., *et al.* (2020). *Chlamydia trachomatis* Plasmid Gene Protein 3 Is Essential for the Establishment of Persistent Infection and Associated Immunopathology. *mBio* 11. doi: 10.1128/mBio.01902-20.

Yang, C., Starr, T., Song, L., Carlson, J. H., Sturdevant, G. L., Beare, P. A., *et al.* (2015). Chlamydial lytic exit from host cells is plasmid regulated. *mBio* 6, e01648-15. doi: 10.1128/mBio.01648-15.

Yang, Z., Tang, L., Shao, L., Zhang, Y., Zhang, T., Schenken, R., *et al.* (2016). The *Chlamydia*-Secreted Protease CPAF Promotes Chlamydial Survival in the Mouse Lower Genital Tract. *Infect Immun* 84, 2697–2702. doi: 10.1128/IAI.00280-16.

Yeruva, L., Melnyk, S., Spencer, N., Bowlin, A., and Rank, R. G. (2013). Differential Susceptibilities to Azithromycin Treatment of Chlamydial Infection in the Gastrointestinal Tract and Cervix. *Antimicrob Agents Chemother* 57, 6290–6294. doi: 10.1128/AAC.01405-13.

Zerial, M., and McBride, H. (2001). Rab proteins as membrane organizers. *Nat Rev Mol Cell Biol* 2, 107–117. doi: 10.1038/35052055.

Zhang, Q., Rosario, C. J., Sheehan, L. M., Rizvi, S. M., Brothwell, J. A., He, C., *et al.* (2020). The Repressor Function of the *Chlamydia* Late Regulator EUO Is Enhanced by the Plasmid-Encoded Protein Pgp4. *J Bacteriol* 202, e00793-19. doi: 10.1128/JB.00793-19.

Zhang, Y., Deng, Q., Porath, J. A., Williams, C. L., Pederson-Gulrud, K. J., and Barbieri, J. T. (2007). Plasma membrane localization affects the RhoGAP specificity of *Pseudomonas* ExoS. *Cell Microbiol* 9, 2192–2201. doi: 10.1111/j.1462-5822.2007.00949.x.

Zhong, G. (2011). *Chlamydia trachomatis* Secretion of Proteases for Manipulating Host Signaling Pathways. *Front Microbiol* 2. doi: 10.3389/fmicb.2011.00014.

Zhong, G. (2017). Chlamydial Plasmid-Dependent Pathogenicity. *Trends Microbiol* 25, 141–152. doi: 10.1016/j.tim.2016.09.006.

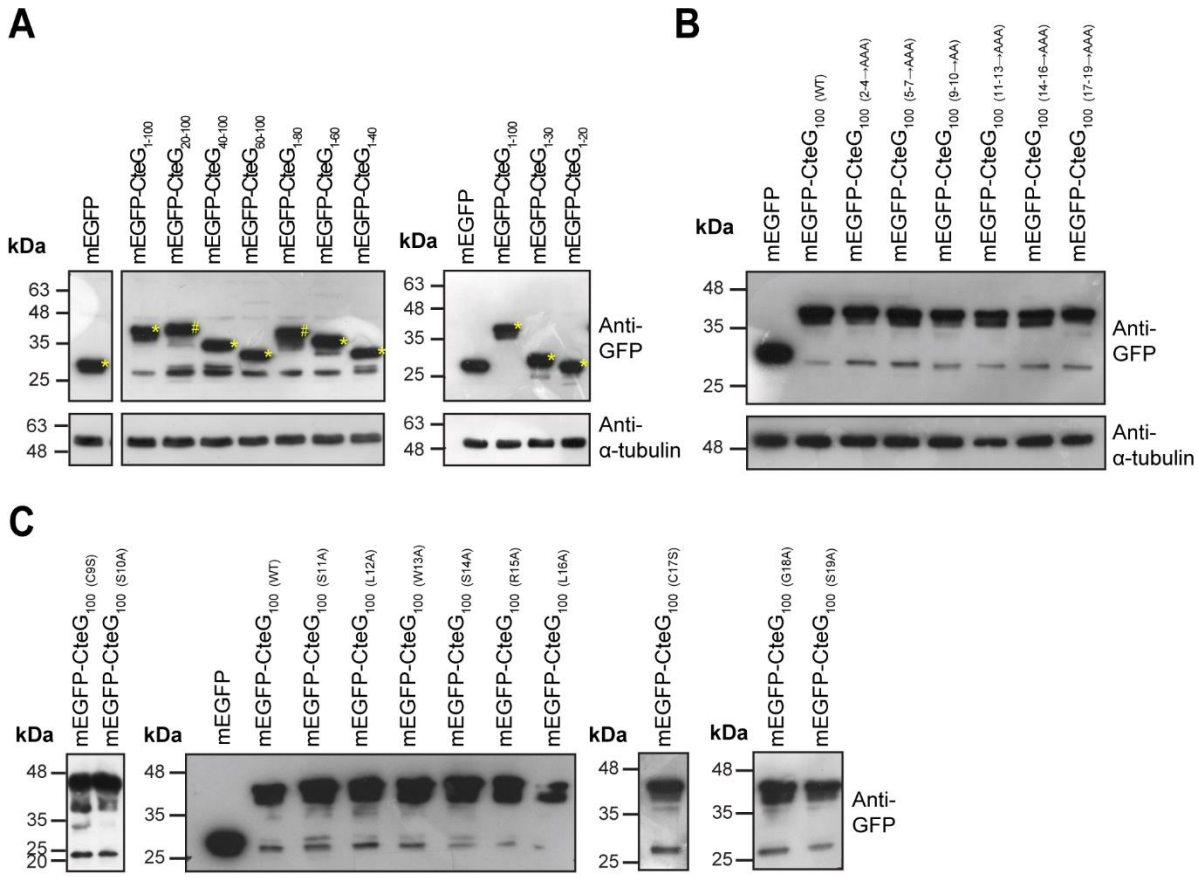
Zhong, G., Fan, P., Ji, H., Dong, F., and Huang, Y. (2001). Identification of a Chlamydial Protease–Like Activity Factor Responsible for the Degradation of Host Transcription Factors. *Journal of Experimental Medicine* 193, 935–942. doi: 10.1084/jem.193.8.935.

Zhou, Z., Tian, Q., Wang, L., and Zhong, G. (2022). *Chlamydia* Deficient in Plasmid-Encoded Glycoprotein 3 (pGP3) as an Attenuated Live Oral Vaccine. *Infect Immun* 90. doi: 10.1128/iai.00472-21.

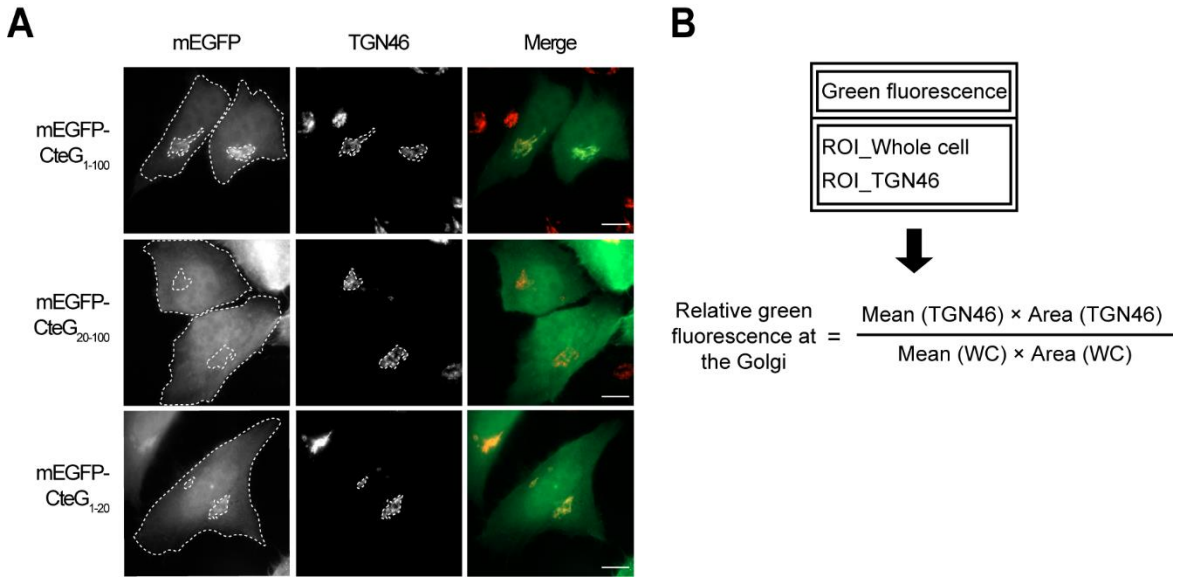
Zuck, M., Ellis, T., Venida, A., and Hybiske, K. (2017). Extrusions are phagocytosed and promote *Chlamydia* survival within macrophages. *Cell Microbiol* 19, 1–12. doi: 10.1111/cmi.12683.

Zuck, M., and Hybiske, K. (2019). The *Chlamydia trachomatis* Extrusion Exit Mechanism Is Regulated by Host Abscission Proteins. *Microorganisms* 7, 149. doi: 10.3390/microorganisms7050149.

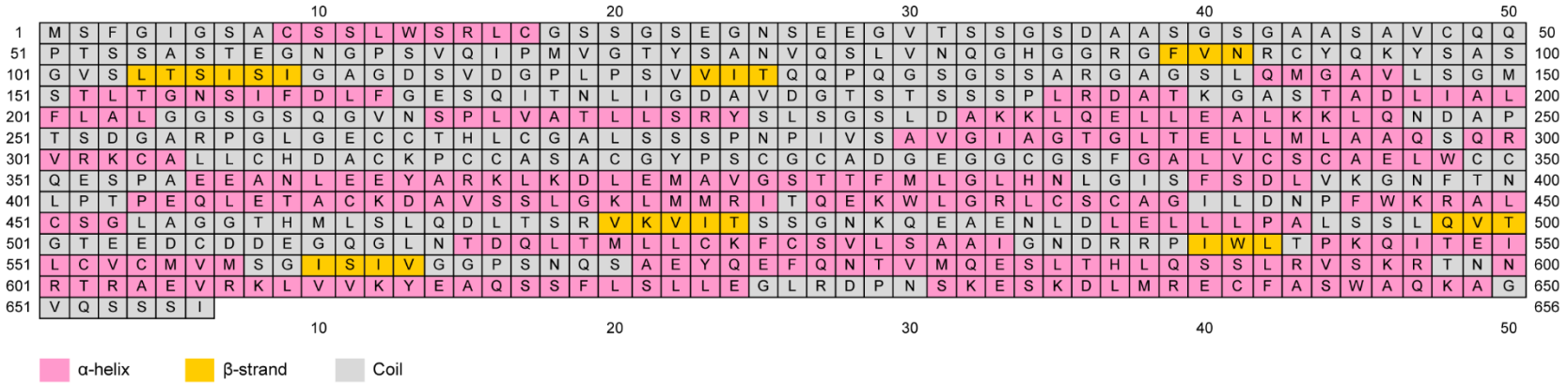
ANNEXES



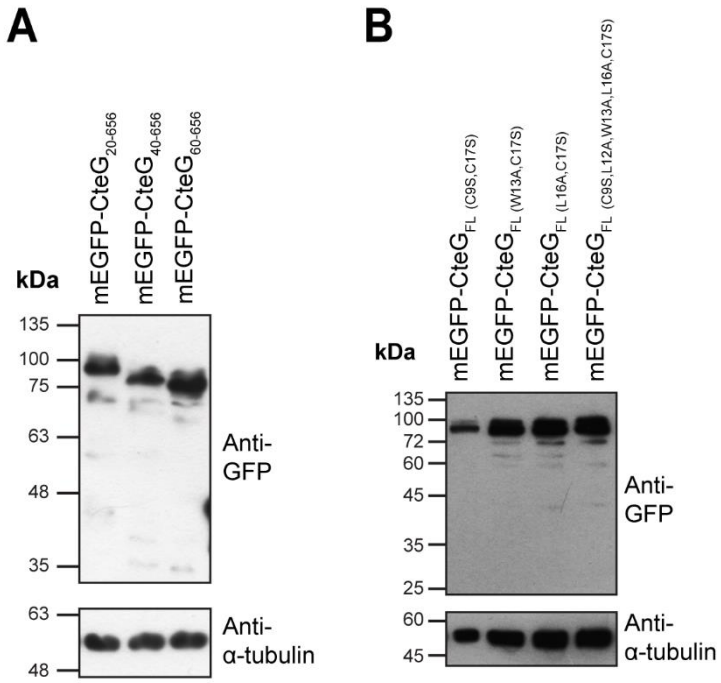
Annexes Figure 1. Analysis of the production of mEGFP-CteG$_{1\text{-}100}$ truncations, and mEGFP-CteG$_{1\text{-}100}$ proteins with amino acid replacements by immunoblotting. Whole extracts of HeLa cells producing mEGFP-CteG hybrid proteins were analyzed by immunoblotting using antibodies against GFP, and human α-tubulin (cell loading control) in (A) and (B), and appropriate HRP-conjugated secondary antibodies. **(A)** mEGFP-CteG$_{1\text{-}100}$ truncations. $*$ indicates proteins that were produced according to the predicted molecular mass; # indicates proteins that were produced above the predicted molecular mass. **(B)** mEGFP-CteG$_{1\text{-}100}$ proteins (38 kDa) with groups of three or two amino acids replaced by alanines. **(C)** mEGFP-CteG$_{1\text{-}100}$ proteins (38 kDa) with single amino acids replaced by alanine or serine (C_9 or C_{17}). Proteins were detected using the SuperSignal West Pico detection kit (Thermo Fisher Scientific).



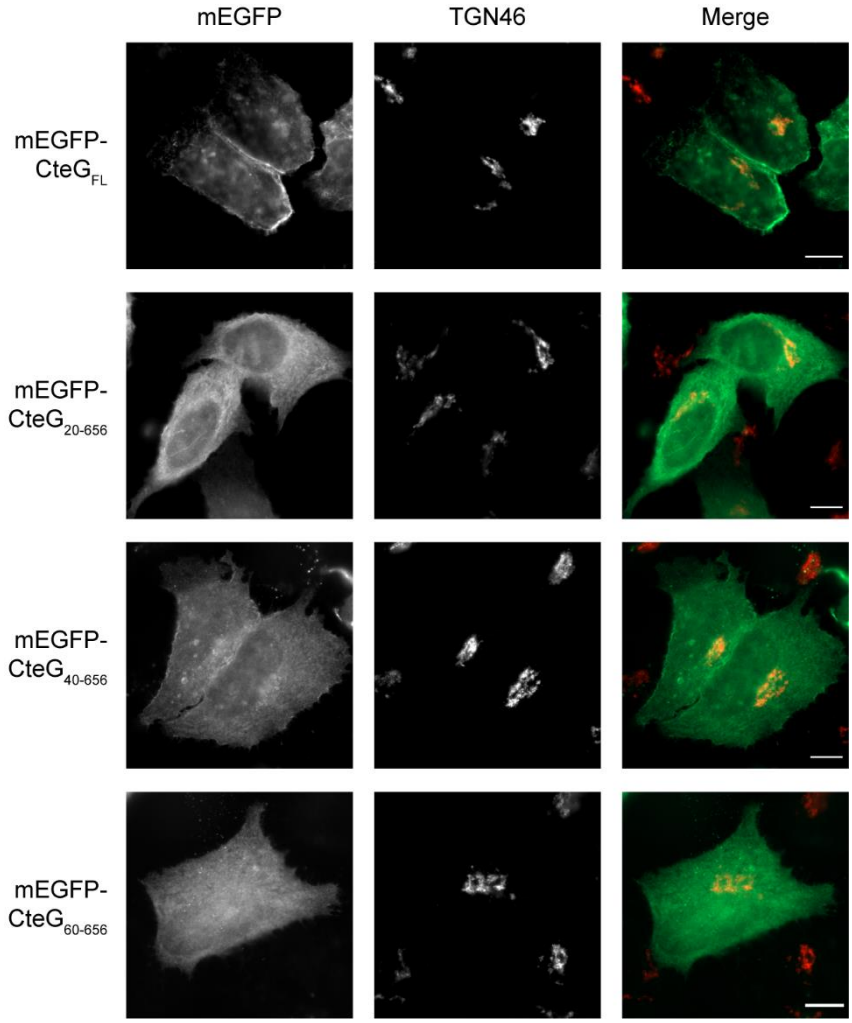
Annexes Figure 2. Method for quantification of the fluorescence of mEGFP-CteG proteins co-localizing with the Golgi of HeLa cells. HeLa cells ectopically producing each mEGFP-CteG protein were immunolabelled with an antibody against TGN46 (which concentrates at the *trans*-Golgi network) and with the appropriate fluorophore-conjugated antibody. **(A)** Images of individual cells were collected by immunofluorescence microscopy, as exemplified for mEGFP-CteG$_{1\text{-}100}$, mEGFP-CteG$_{20\text{-}100}$ and mEGFP-CteG$_{1\text{-}20}$ (green). **(B)** Each individual cell and the corresponding TGN46-immunolabelled Golgi (red) were delineated (region of interest, ROI) as shown by the white dashed lines in panel A. The mean green fluorescence of each protein in the entire cell (ROI_Whole cell, WC) and at the TGN46-immunolabelled Golgi (ROI_TGN46), and the area of each ROI were determined with Fiji (Schindelin *et al*., 2012). The product of the mean green fluorescence of the protein at the TGN46-immunolabelled Golgi by the area of this compartment was divided by the product of the mean green fluorescence of the protein in the whole cell by the area of the whole cell to obtain the relative green fluorescence at the Golgi.



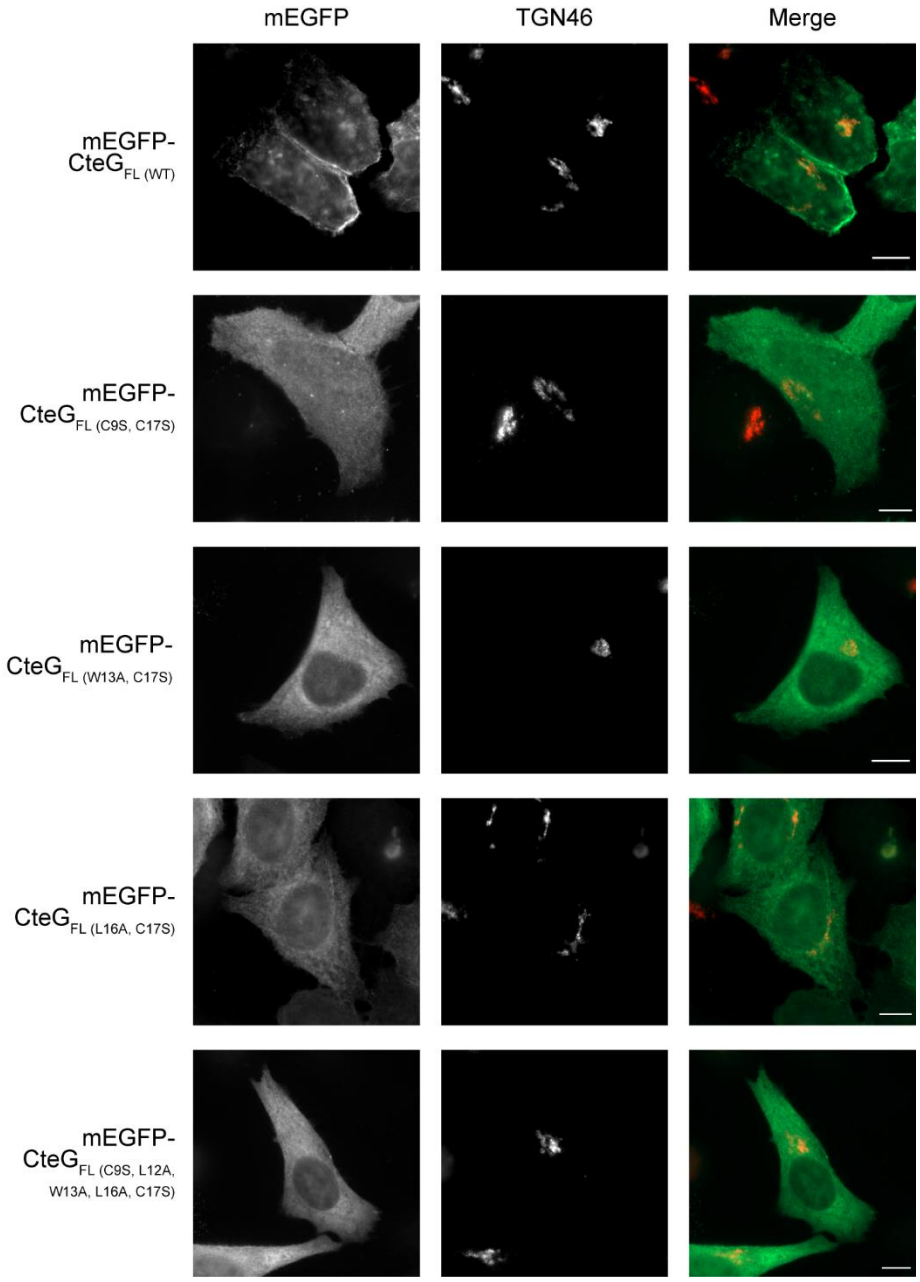
Annexes Figure 3. Prediction of the secondary structure of CteG. The secondary structure of CteG (656 amino acid residues) was predicted with PSIPRED v4.0 (http://bioinf.cs.ucl.ac.uk/psipred/; University College London, UK) using the default settings. A simplified scheme of the α-helixes (pink), β-strands (yellow) and coils (grey) predicted to occur in CteG are illustrated.



Annexes Figure 4. Analysis of the production in HeLa cells of mEGFP-CteG truncations and mEGFP-CteG$_{FL}$ proteins with amino acid replacements in CteG by immunoblotting. Whole extracts of HeLa cells producing mEGFP-CteG hybrid proteins were analyzed by immunoblotting using antibodies against GFP and human α-tubulin (cell loading control), and appropriate HRP-conjugated secondary antibodies. **(A)** mEGFP-CteG truncations. **(B)** mEGFP-CteG$_{FL}$ proteins (96 kDa) with groups of two or five amino acids replaced by alanines or serines. Proteins were detected using the SuperSignal West Pico detection kit (Thermo Fisher Scientific).



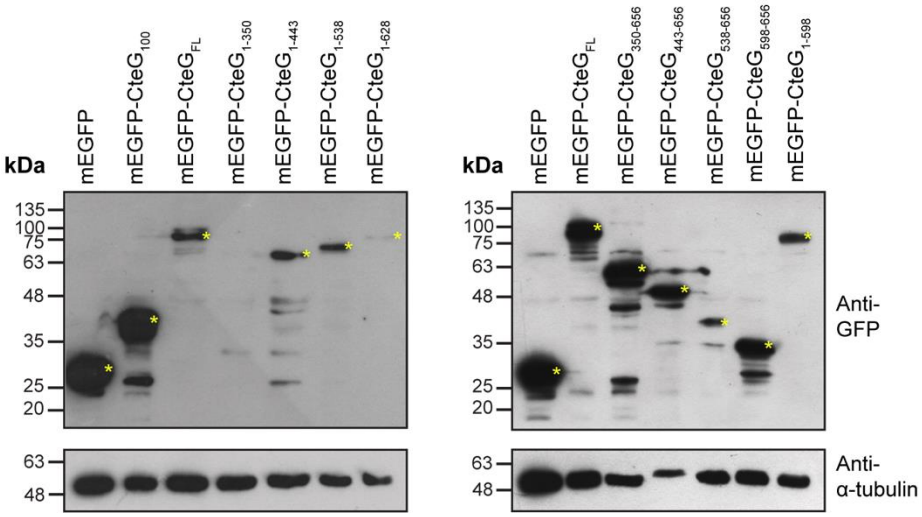
Annexes Figure 5. Analysis of the subcellular localization of mEGFP-CteG proteins lacking N-terminal amino acid portions of CteG by immunofluorescence microscopy. Mammalian expression vectors encoding mEGFP-CteG hybrid proteins lacking the first 20, 40 or 60 N-terminal residues of CteG were constructed. HeLa cells were transfected with these plasmids for 24 h before being fixed with PFA 4% (w/v) and immunolabelled with an antibody against TGN46 (*trans*-Golgi network, red), and the appropriate fluorophore-conjugated antibody. Images of cells producing each protein (green) were obtained by fluorescence microscopy and illustrate the predominant cytosolic localization of all proteins, and a less frequent co-localization with the cell plasma membrane (perceptible for mEGFP-CteG$_{40-656}$). Scale bar, 10 μm.



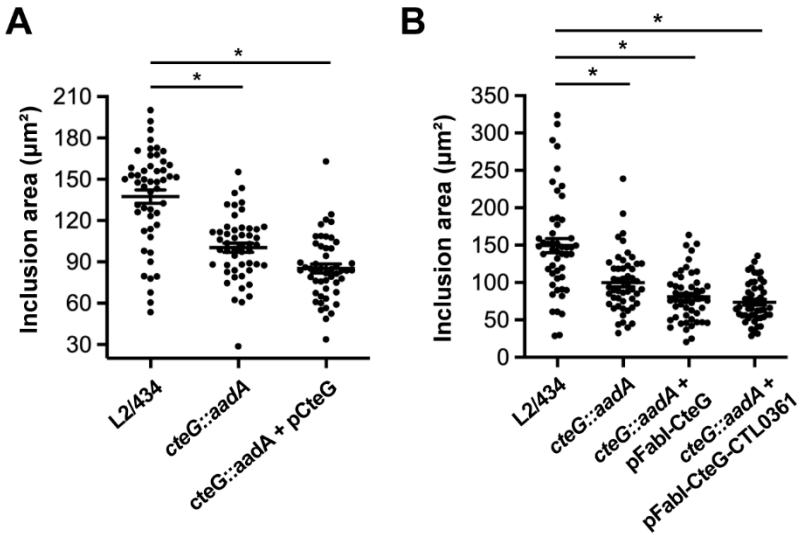
Annexes Figure 6. Analysis of the subcellular localization of mEGFP-CteG proteins containing specific amino acid replacements at the N-terminal region of CteG by immunofluorescence microscopy. HeLa cells ectopically producing mEGFP-CteG hybrid proteins carrying specific amino acid substitutions at the N-terminal region of CteG were fixed with PFA 4% (w/v) at 24 h post-transfection and immunolabelled with an antibody against TGN46 (*trans*-Golgi network, red) and with the appropriate fluorophore-conjugated antibody. Images of cells producing each protein (green) were obtained by fluorescence microscopy and illustrate the predominant cytosolic localization of all proteins. Scale bar, 10 μm.

ID	Position	Peptide	Score	Cut-off
CteG	9	SFGIGSACSSLWSRL	12.62	10.722
	17	SSLWSRLCGSSGSEG	13.747	10.722
	262	ARPGLGECCTHLCGA	5.988	2.412
	263	RPGLGECCTHLCGAL	4.234	3.717
	308	VRKCALLCHDACKPC	4.465	2.412
	315	CHDACKPCCASACGY	5.929	2.412
	316	HDACKPCCASACGYP	8.052	3.717
	320	KPCCASACGYPSCGC	3.316	2.412
	325	SACGYPSCGCADGEG	12.892	10.722
	349	CSCAELWCCQESPAE	11.541	2.412
	350	SCAELWCCQESPAEE	4.498	3.717
	523	DQLTMLLCKFCSVLS	4.844	3.717
	554	ITEILCVCMVMSGIS	4.63	3.717
	641	SKDLMRECFASWAQK	4.077	3.717

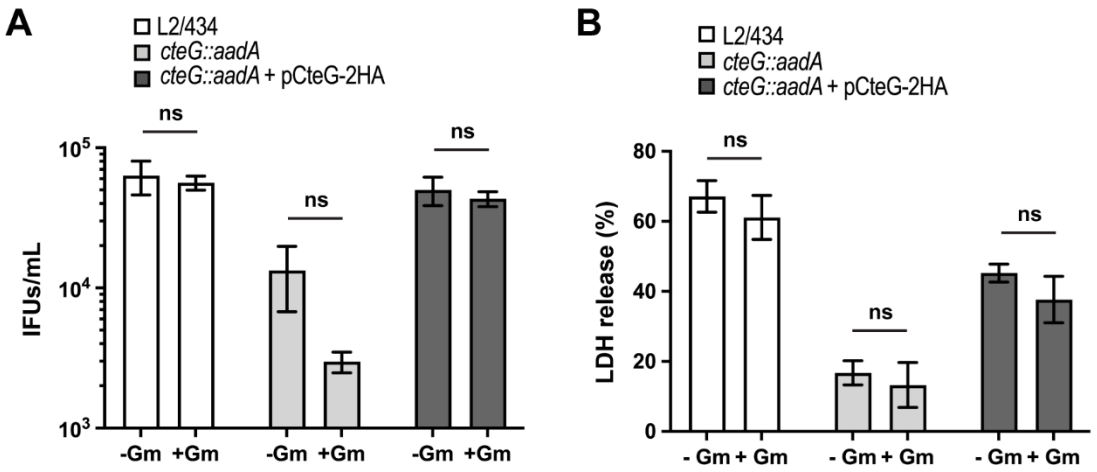
Annexes Figure 7. Prediction of potentially S-palmitoylated cysteine residues on CteG. The amino acid sequence of CteG was used as input in the online interface of the palmitoylation site prediction tool CSS-Palm [v4.0; http://csspalm.biocuckoo.org/online.php; (Ren *et al.*, 2008)]. The threshold was set to "medium". Cysteine residues that scored above the cut-off value and their respective positions (highlighted in red) are presented.



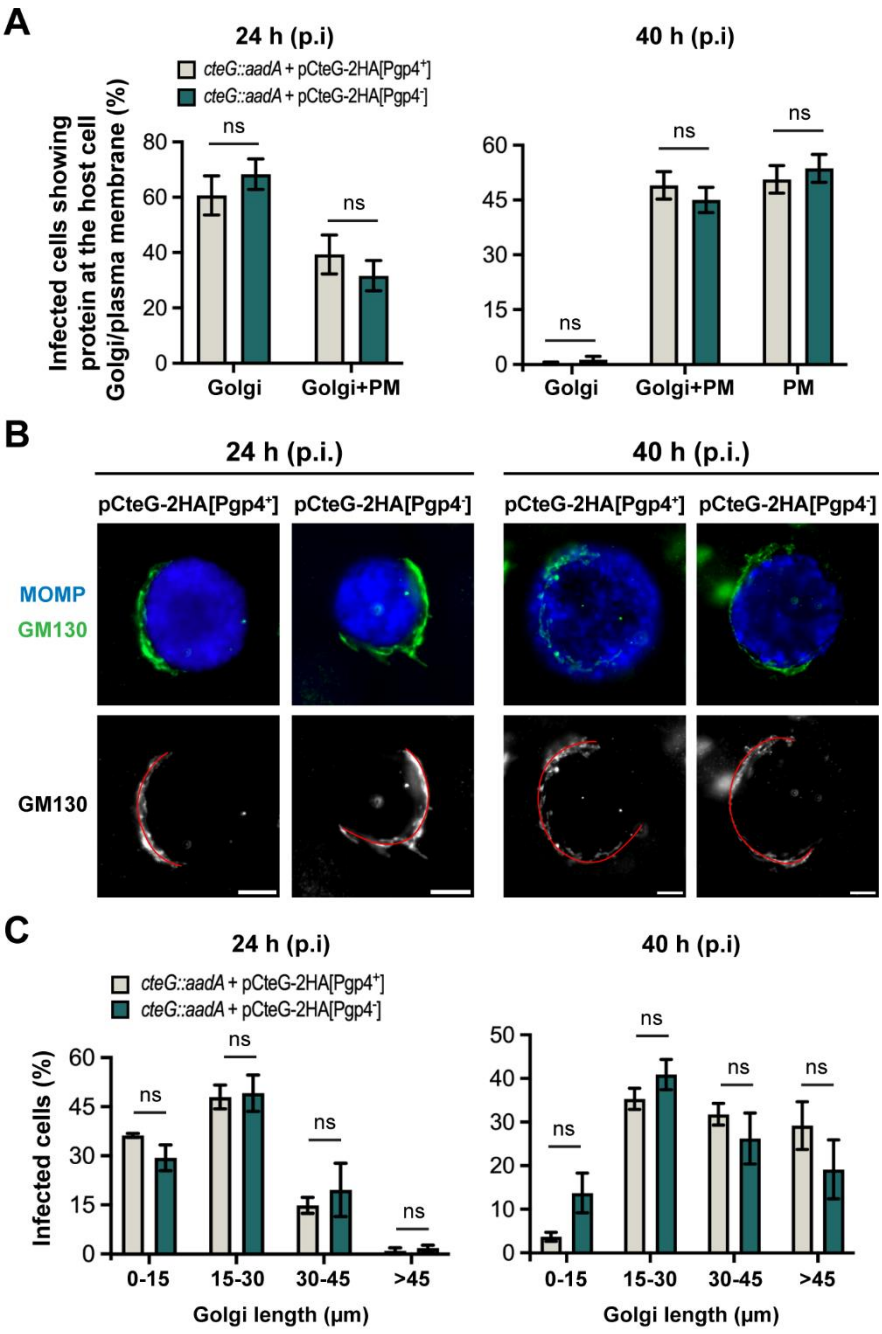
Annexes Figure 8. Analysis of the production of mEGFP-CteG hybrid proteins with truncations in CteG by immunoblotting. Whole extracts of HeLa cells producing mEGFP-CteG hybrid proteins with truncations in CteG were analyzed by immunoblotting using antibodies against GFP and human α-tubulin (cell loading control), and appropriate HRP-conjugated secondary antibodies. Proteins were detected using the SuperSignal West Pico detection kit (Thermo Fisher Scientific). $*$ indicates proteins that were produced according to the predicted molecular mass. A band corresponding to mEGFP-CteG$_{1\text{-}350}$ was not detected.



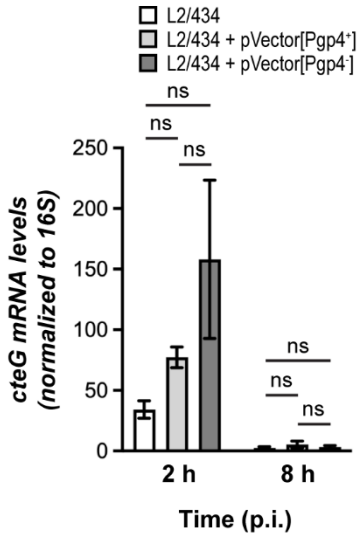
Annexes Figure 9. Plasmids encoding native *cteG* or *cteG* and its flanking genes do not complement the growth defect of the *C. trachomatis cteG::aadA* strain. HeLa cells were infected for 24 h with the indicated *C. trachomatis* strains. Cells were then fixed with methanol and immunolabelled with an antibody against *C. trachomatis* MOMP and with the appropriate fluorophore-conjugated secondary antibody. Independent photos of infected cells were obtained by immunofluorescence microscopy and the area of 50 randomly chosen inclusions was measured using the software Fiji. **(A)** pCteG carries *cteG* alone. **(B)** pFabI-CteG carries *cteG* and *ctl0359/fabI*; pFabI-CteG-CTL0361 carries *cteG*, *ctl0359/fabI* and *ctl0361* (Figure 3.13A). Statistical significance was determined by using ordinary one-way ANOVA and Dunnett post-test analysis (*p<0.0001).



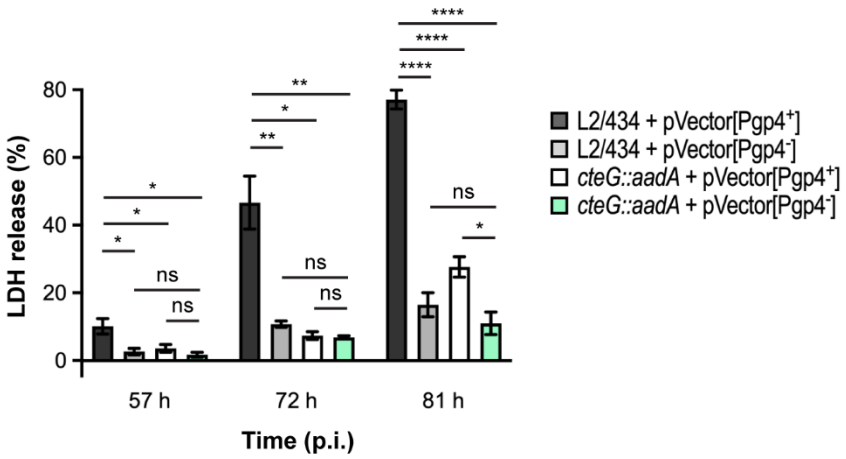
Annexes Figure 10. The presence of gentamicin in the infection medium during the first 24 h of infection does not significantly alter the number of released inclusion forming units or the ability of the strains to cause host cell lysis. HeLa cells were infected with *C. trachomatis* strains L2/434, *cteG::aadA* and *cteG::aadA* harboring a plasmid encoding CteG-2HA (pCteG-2HA, also named pCteG-2HA[Pgp4^{+}]) with a multiplicity of infection (MOI) of 0.06 for 48 h (A), or a MOI of 0.3 for 72 h (B) and incubated under the presence of 10 μg/mL gentamicin during the first 24 h of infection. At this time point, the medium was replaced by one without gentamicin. **(A)** The supernatant fraction of infected cells was collected and IFUs were quantified by immunofluorescence microscopy. **(B)** The release of host lactate dehydrogenase (LDH) into the supernatant of infected cells was measured using a CytoScan™ LDH Cytotoxicity Assay kit (G-Biosciences). Data in (A) and (B) correspond to the mean ± SEM (n=3). Statistical significance was determined for each strain between presence and absence of gentamicin conditions using two-tailed unpaired Student's t-test (ns, non-significant). For statistical analysis in (A), natural logarithm was applied to data to ensure normality of the populations.



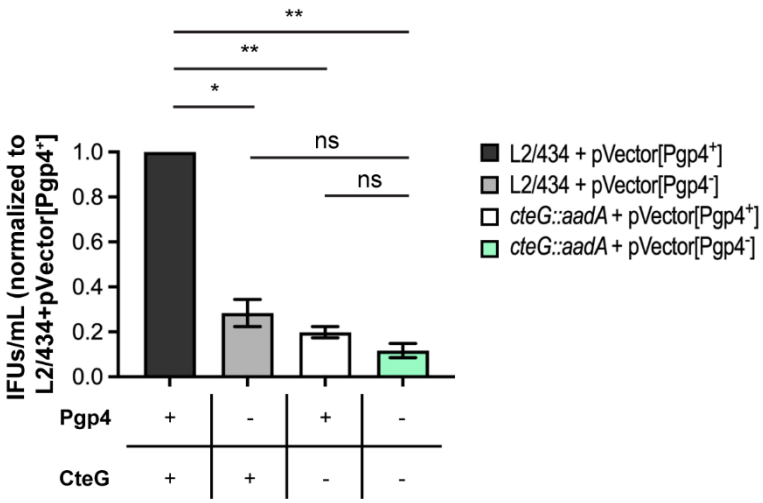
Annexes Figure 11. Localization of CteG and Golgi distribution around the inclusion are indistinguishable in cells infected by *C. trachomatis* Pgp4⁺ or Pgp4⁻ strains. HeLa cells were infected by *C. trachomatis* strains harboring pCteG-2HA[Pgp4⁺] or pCteG-2HA[Pgp4⁻] at a MOI of 0.3 for 24 or 40 h, fixed with methanol and immunolabelled with antibodies against *cis*-Golgi marker GM130 (green) and HA (red), and against MOMP (blue) when applicable, and with the appropriate fluorophore-conjugated secondary antibodies. **(A)** Fluorescence microscopy was used to enumerate cells showing CteG-2HA only at the Golgi, only at the plasma membrane (PM) or both at the Golgi and at the plasma membrane (Golgi+PM). Data correspond to the mean ± SEM of three independent experiments (N=100 for each assay). **(B)** Images of random infected cells were collected by fluorescence microscopy and the extension of Golgi distribution around the inclusion was measured as previously described (Pais *et al.*, 2019), and as exemplified by the red line in the lower panel. Scale bars, 5 µm. **(C)** Percentage of infected cells with Golgi around the inclusion of the indicated length. Data are mean ± SEM of three independent experiments (N≥35 per condition). In (A) and (B), statistical significance was determined for each time point by using a two-tailed unpaired Student's t-test (ns, non-significant).



Annexes Figure 12. The mRNA levels of *cteG* are similar in cells infected by *C. trachomatis* Pgp4$^+$ or Pgp4$^-$ strains. HeLa cells were infected with *C. trachomatis* L2/434 parental strain or with its derivative strains carrying pVector[Pgp4$^+$] (CteG$^+$/Pgp4$^+$) or pVector[Pgp4$^-$] (CteG$^+$/Pgp4$^-$) (Annexes Table 3) at a MOI of 2.5 for 2 or 8 h. Infected cells were collected and total RNA was extracted to perform RT-qPCR of the *cteG* transcript. *cteG* mRNA levels were normalized against those of the *16S* mRNA in the corresponding sample. Data corresponds to mean ± SEM (n=3). Statistical significance was determined for each time point with ordinary one-way ANOVA and Tukey's post-test analysis (ns, non-significant).



Annexes Figure 13. A *C. trachomatis* CteG$^-$/Pgp4$^-$ double mutant strain displays a defect in inducing host cell lysis similar to CteG$^+$/Pgp4$^-$ and CteG$^-$/Pgp4$^+$ single mutant strains. HeLa cells were infected with *C. trachomatis* L2/434 carrying pVector[Pgp4$^+$] (CteG$^+$/Pgp4$^+$), L2/434 carrying pVector[Pgp4$^-$] (CteG$^+$/Pgp4$^-$), *cteG::aadA* carrying pVector[Pgp4$^+$] (CteG$^-$/Pgp4$^+$), or *cteG::aadA* carrying pVector[Pgp4$^-$] (CteG$^-$/Pgp4$^-$) for 57, 72 or 81 h at an MOI of 0.3. At each time point, the release of host LDH into the supernatant of infected cells was quantified using a CytoScan™ LDH Cytotoxicity Assay kit (G-Biosciences). Data correspond to mean ± SEM (n=3). For each time post-infection (p.i.), statistical significance was determined by using ordinary one-way ANOVA and Tukey's post-test analysis (ns, non-significant; *p<0.05; **p<0.01; ***p<0.001; ****p<0.0001).



Annexes Figure 14. A *C. trachomatis* CteG$^-$/Pgp4$^-$ double mutant strain displays a defect in releasing infectious particles to the cell culture supernatant identical to CteG$^+$/Pgp4$^-$ and CteG$^-$/Pgp4$^+$ single mutant strains. HeLa cells were infected with *C. trachomatis* L2/434 carrying pVector[Pgp4$^+$] (CteG$^+$/Pgp4$^+$), L2/434 carrying pVector[Pgp4$^-$] (CteG$^+$/Pgp4$^-$), *cteG::aadA* carrying pVector[Pgp4$^+$] (CteG$^-$/Pgp4$^+$), or *cteG::aadA* carrying pVector[Pgp4$^-$] (CteG$^-$/Pgp4$^-$) for 48 h at an MOI of 0.06. The supernatant fractions of infected cells were collected as described in Materials and Methods, and the number of recoverable inclusion forming units (IFUs) was determined by immunofluorescence microscopy. In each experiment, the number of IFUs determined for each strain was divided by the number of IFUs obtained for the CteG$^+$/Pgp4$^+$ strain. Data correspond to mean ± SEM (n=3). Statistical significance was determined by using a two-tailed unpaired Student's t-test between each pair of strains (ns, non-significant; *p<0.001; **p<0.0001).

Annexes Table 1. Plasmids used in this work.

Plasmid	Description	Source/Reference	Chapter
p2TK2--SW2	*E. coli* - *C. trachomatis* shuttle vector which enables the expression of proteins in *C. trachomatis* (Amp^R).	(Agaisse and Derré, 2013).	**3.1, 3.2**
pSVP247/pVector[$Pgp4^+$]	Derivative of p2TK2--SW2 for expression of proteins with a C-terminal double hemagglutinin (2HA) tag. Contains the terminator of the *incDEFG* operon (T_{incD}) of *C. trachomatis* L2/434 Bu (Amp^R).	(da Cunha *et al.*, 2017).	**3.1, 3.2**
pSVP264/pCteG-2HA/pCteG-2HA[$Pgp4^+$]	Derivative of pSVP247/pVector[$Pgp4^+$] for the over-production of CteG-2HA under the control of the predicted *cteG* promoter (P_{cteG};Amp^R).	(Pais *et al.*, 2019).	**3.1, 3.2**
pIP40/pCteG	Enables the expression of native $CteG_{FL}$ under the control of P_{cteG}. Contains the T_{incD} sequence. A DNA fragment comprising P_{cteG}-*cteG* was amplified from pSVP264/pCteG-2HA using primers 1680 and 2321. Another DNA fragment containing the T_{incD} sequence was amplified from pSVP247/pVector[$Pgp4^+$] with primers 2320 and 1483. Both fragments were then fused by overlapping PCR using primers 1680 and 1483. The resulting DNA product was digested with KpnI and SalI and inserted into those sites of p2TK2--SW2 (Amp^R).	This work.	**3.2**
pIP53/pFabI-CteG-CTL0361	Enables the expression of CTL0359 (FabI), CTL0360 (CteG) and CTL0361. A DNA fragment containing the DNA sequences of *ctl0360* (*cteG*), and of its neighboring genes *ctl0359* and *ctl0361* plus some nucleotides upstream and downstream these genes was amplified from *C. trachomatis* L2/434 chromosomal DNA using primers 2388 and 2389. The resulting DNA product was digested with KpnI and SalI and inserted into those sites of p2TK2--SW2 (Amp^R).	This work.	**3.2**
pIP54/pFabI-CteG	Enables the expression of CTL0359 (FabI) and CTL0360 (CteG). A DNA fragment containing *ctl0359* and *ctl0360* was	This work.	**3.2**

	amplified from *C. trachomatis* L2/434 chromosomal DNA using primers 2388 and 2321. A DNA fragment containing the sequence of T_{incD} was amplified from pSVP247/pVector[Pgp4^{+}] with primers 2320 and 1483. Both DNA fragments were then fused by overlapping PCR using primers 2388 and 1483. The resulting DNA product was digested with KpnI and SalI and inserted into those sites of p2TK2--SW2 (Amp^{R}).		
pIP68/pVector[Pgp4^{-}]	Derivative of pSVP247/pVector[Pgp4^{+}] where *pgp4/porf6* is deleted. pSVP247/pVector[Pgp4^{+}] was amplified using primers 2554 and 2555. The resulting DNA product was digested with AscI and ligated, enabling plasmid circularization.	This work.	**3.2**
pIP69/pCteG-2HA[Pgp4^{-}]	Enables the expression of CteG-2HA under the control of P_{cteG} in a *pgp4*$^{-}$ background. A DNA fragment containing P_{cteG}-*cteG* was amplified from pSVP264/pCteG-2HA using primers 1680 and 1552. The resulting DNA product was digested with KpnI and NotI and inserted into those sites of pIP68/pVector[Pgp4^{-}] (Amp^{R}).	This work.	**3.2**
pIP72	Enables the expression of $CteG_{FL}$ carrying substitutions of amino acid residues 9 and 17 by serine, under the control of P_{cteG}. Contains the T_{incD} sequence. A DNA fragment encoding P_{cteG}-CteG(C17) was amplified from pSVP264 using primers 1680 and 2826. Another DNA fragment encoding CteG(C17)-CteG(I656) was amplified from pSVP264 using primers 2825 and 1552. Both fragments were then fused by overlapping PCR using primers 1680 and 1552. The resulting DNA product was digested with KpnI and NotI and inserted into those sites of pSVP247/pVector[Pgp4^{+}] (Amp^{R}).	This work.	**3.1**
pIP73	Enables the expression of $CteG_{FL}$ carrying substitutions of amino acid residues 9 and 17 by serine, and of amino acid residues	This work.	**3.1**

	12, 13 and 16 by alanine under the control of P_{cteG}. Contains the T_{incD} sequence. A DNA fragment encoding P_{cteG}-CteG(C17) was amplified from pSVP264 using primers 1680 and 2828. Another DNA fragment encoding CteG(C17)-CteG(I656) was amplified from pSVP264 using primers 2827 and 1552. Both fragments were then fused by overlapping PCR using primers 1680 and 1552. The resulting DNA product was digested with KpnI and NotI and inserted into those sites of pSVP247/pVector[Pgp4^{+}] (Amp^R).		
pmEGFP-C1	Mammalian transfection vector for expression of proteins fused to the N-terminus of monomeric EGFP (mEGFP) under the control of the CMV promoter (Kan^R).	(Pais *et al.*, 2019).	**3.1, 3.3**
pSVP310	Transfection vector derived from pmEGFP-C1 which encodes mEGFP-$CteG_{1-100}$ (Km^R).	(Pais *et al.*, 2019).	**3.1**
pALT4	Transfection vector derived from pmEGFP-C1 which encodes mEGFP-$CteG_{FL}$ (Km^R).	(Pais *et al.*, 2019).	**3.1, 3.3**
pIP51	Transfection vector encoding mEGFP-$GobX_{FL}$. A DNA fragment containing *gobX* was amplified from *Legionella pneumophila* str. Paris chromosomal DNA using primers 2333 and 2334. The resulting DNA product was digested with SalI and KpnI and inserted into those sites of pmEGFP-C1 (Km^R).	This work.	**3.1**
pIP15	Transfection vector encoding mEGFP-$CteG_{20-100}$. A DNA fragment encoding amino acid residues 20 to 100 of CteG was amplified from pSVP310 using primers 2180 and 2128. The resulting DNA product was digested with SalI and KpnI and inserted into those sites of pmEGFP-C1 (Km^R).	This work.	**3.1**
pIP16	Transfection vector encoding mEGFP-$CteG_{40-100}$. A DNA fragment encoding amino acid residues 40 to 100 of CteG was amplified from pSVP310 using primers	This work.	**3.1**

	2181 and 2128. The resulting DNA product was digested with SalI and KpnI and inserted into those sites of pmEGFP-C1 (Km^R).		
pIP17	Transfection vector encoding mEGFP-$CteG_{60-100}$. A DNA fragment encoding amino acid residues 60 to 100 of CteG was amplified from pSVP310 using primers 2182 and 2128. The resulting DNA product was digested with SalI and KpnI and inserted into those sites of pmEGFP-C1 (Km^R).	This work.	**3.1**
pIP18	Transfection vector encoding mEGFP-$CteG_{1-80}$. A DNA fragment encoding amino acid residues 1 to 80 of CteG was amplified from pSVP310 using primers 652 and 2183. The resulting DNA product was digested with SalI and KpnI and inserted into those sites of pmEGFP-C1 (Km^R).	This work.	**3.1**
pIP19	Transfection vector encoding mEGFP-$CteG_{1-60}$. A DNA fragment encoding amino acid residues 1 to 60 of CteG was amplified from pSVP310 using primers 652 and 2184. The resulting DNA product was digested with SalI and KpnI and inserted into those sites of pmEGFP-C1 (Km^R).	This work.	**3.1**
pIP20	Transfection vector encoding mEGFP-$CteG_{1-40}$. A DNA fragment encoding amino acid residues 1 to 40 of CteG was amplified from pSVP310 using primers 652 and 2185. The resulting DNA product was digested with SalI and KpnI and inserted into those sites of pmEGFP-C1 (Km^R).	This work.	**3.1**
pIP21	Transfection vector encoding mEGFP-$CteG_{1-30}$. A DNA fragment encoding amino acid residues 1 to 30 of CteG was amplified from pSVP310 using primers 652 and 2204. The resulting DNA product was digested with SalI and KpnI and inserted into those sites of pmEGFP-C1 (Km^R).	This work.	**3.1**

pIP22	Transfection vector encoding mEGFP-CteG$_{1-20}$. A DNA fragment encoding amino acid residues 1 to 20 of CteG was amplified from pSVP310 using primers 652 and 2205. The resulting DNA product was digested with SalI and KpnI and inserted into those sites of pmEGFP-C1 (Km^R).	This work.	**3.1**
pIP47	Transfection vector encoding mEGFP-CteG$_{1-350}$. A DNA fragment encoding amino acid residues 1 to 350 of CteG was amplified from pSVP264 using primers 652 and 2335. The resulting DNA product was digested with SalI and KpnI and inserted into those sites of pmEGFP-C1 (Km^R).	This work.	**3.1**
pIP48	Transfection vector encoding mEGFP-CteG$_{1-443}$. A DNA fragment encoding amino acid residues 1 to 443 of CteG was amplified from pSVP264 using primers 652 and 2336. The resulting DNA product was digested with SalI and KpnI and inserted into those sites of pmEGFP-C1 (Km^R).	This work.	**3.1**
pIP49	Transfection vector encoding mEGFP-CteG$_{1-538}$. A DNA fragment encoding amino acid residues 1 to 538 of CteG was amplified from pSVP264 using primers 652 and 2337. The resulting DNA product was digested with SalI and KpnI and inserted into those sites of pmEGFP-C1 (Km^R).	This work.	**3.1**
pIP60	Transfection vector encoding mEGFP-CteG$_{1-598}$. A DNA fragment encoding amino acid residues 1 to 598 of CteG was amplified from pSVP264 using primers 652 and 2422. The resulting DNA product was digested with SalI and KpnI and inserted into those sites of pmEGFP-C1 (Km^R).	This work.	**3.1**
pIP50	Transfection vector encoding mEGFP-CteG$_{1-628}$. A DNA fragment encoding amino acid residues 1 to 628 of CteG was amplified from pSVP264 using primers 652 and 2338. The resulting DNA product	This work.	**3.1**

	was digested with SalI and KpnI and inserted into those sites of pmEGFP-C1 (Km^R).		
pIP56	Transfection vector encoding mEGFP-$CteG_{350-656}$. A DNA fragment encoding amino acid residues 350 to 656 of CteG was amplified from pSVP264 using primers 2418 and 653. The resulting DNA product was digested with SalI and KpnI and inserted into those sites of pmEGFP-C1 (Km^R).	This work.	**3.1**
pIP57	Transfection vector encoding mEGFP-$CteG_{443-656}$. A DNA fragment encoding amino acid residues 443 to 656 of CteG was amplified from pSVP264 using primers 2419 and 653. The resulting DNA product was digested with SalI and KpnI and inserted into those sites of pmEGFP-C1 (Km^R).	This work.	**3.1**
pIP58	Transfection vector encoding mEGFP-$CteG_{538-656}$. A DNA fragment encoding amino acid residues 538 to 656 of CteG was amplified from pSVP264 using primers 2420 and 653. The resulting DNA product was digested with SalI and KpnI and inserted into those sites of pmEGFP-C1 (Km^R).	This work.	**3.1**
pIP59	Transfection vector encoding mEGFP-$CteG_{598-656}$. A DNA fragment encoding amino acid residues 598 to 656 of CteG was amplified from pSVP264 using primers 2421 and 653. The resulting DNA product was digested with SalI and KpnI and inserted into those sites of pmEGFP-C1 (Km^R).	This work.	**3.1**
pIP23	Transfection vector encoding mEGFP-$CteG_{1-100}$ with amino acid residues 2 to 4 replaced by alanines. pSVP310 was amplified using mutagenic primers 2206 and 2207. Methylated, non-mutated parental DNA templates were digested with DpnI. Plasmids carrying the desired mutations (Km^R) were recovered from *E. coli* cells transformed with DpnI-digested plasmid mixture.	This work.	**3.1**

pIP24	Transfection vector encoding mEGFP-$CteG_{1-100}$ with amino acid residues 5 to 7 replaced by alanines. pSVP310 was amplified using mutagenic primers 2208 and 2209. Methylated, non-mutated parental DNA templates were digested with DpnI. Plasmids carrying the desired mutations (Km^R) were recovered from *E. coli* cells transformed with DpnI-digested plasmid mixture.	This work.	**3.1**
pIP25	Transfection vector encoding mEGFP-$CteG_{1-100}$ with amino acid residues 8 to 10 replaced by alanines. pSVP310 was amplified using mutagenic primers 2210 and 2211. Methylated, non-mutated parental DNA templates were digested with DpnI. Plasmids carrying the desired mutations (Km^R) were recovered from *E. coli* cells transformed with DpnI-digested plasmid mixture.	This work.	**3.1**
pIP26	Transfection vector encoding mEGFP-$CteG_{1-100}$ with amino acid residues 11 to 13 replaced by alanines. A DNA fragment encoding amino acid residues 1 to 13 of CteG was amplified from pSVP310 using primers 652 and 2213. Another DNA fragment encoding amino acid residues 11 to 100 of CteG was amplified from pSVP310 using primers 2212 and 2128. Both fragments were then fused by overlapping PCR using primers 652 and 2128. The resulting DNA product was digested with KpnI and SalI and inserted into those sites of pmEGFP-C1 (Km^R).	This work.	**3.1**
pIP27	Transfection vector encoding mEGFP-$CteG_{1-100}$ with amino acid residues 14 to 16 replaced by alanines. A DNA fragment encoding amino acid residues 1 to 16 of CteG was amplified from pSVP310 using primers 652 and 2215. Another DNA fragment encoding amino acid residues 14 to 100 of CteG was amplified from pSVP310 using primers 2214 and 2128. Both fragments were then fused by overlapping PCR using primers 652 and	This work.	**3.1**

	2128. The resulting DNA product was digested with KpnI and SalI and inserted into those sites of pmEGFP-C1 (Km^R).		
pIP28	Transfection vector encoding mEGFP-$CteG_{1-100}$ with amino acid residues 17 to 19 replaced by alanines. A DNA fragment encoding amino acid residues 1 to 19 of CteG was amplified from pSVP310 using primers 652 and 2217. Another DNA fragment encoding amino acid residues 17 to 100 of CteG was amplified from pSVP310 using primers 2216 and 2128. Both fragments were then fused by overlapping PCR using primers 652 and 2128. The resulting DNA product was digested with KpnI and SalI and inserted into those sites of pmEGFP-C1 (Km^R).	This work.	**3.1**
pIP63	Transfection vector encoding mEGFP-$CteG_{1-100}$ with amino acid residue 9 replaced by a serine. A DNA fragment encoding amino acid residues 1 to 9 of CteG was amplified from pSVP310 using primers 652 and 2432. Another DNA fragment encoding amino acid residues 9 to 100 of CteG was amplified from pSVP310 using primers 2431 and 2128. Both fragments were then fused by overlapping PCR using primers 652 and 2128. The resulting DNA product was digested with KpnI and SalI and inserted into those sites of pmEGFP-C1 (Km^R).	This work.	**3.1**
pIP64	Transfection vector encoding mEGFP-$CteG_{1-100}$ with amino acid residue 10 replaced by an alanine. A DNA fragment encoding amino acid residues 1 to 10 of CteG was amplified from pSVP310 using primers 652 and 2434. Another DNA fragment encoding amino acid residues 10 to 100 of CteG was amplified from pSVP310 using primers 2433 and 2128. Both fragments were then fused by overlapping PCR using primers 652 and 2128. The resulting DNA product was digested with KpnI and SalI and inserted into those sites of pmEGFP-C1 (Km^R).	This work.	**3.1**

pIP41	Transfection vector encoding mEGFP-$CteG_{1\text{-}100}$ with amino acid residue 11 replaced by an alanine. A DNA fragment encoding amino acid residues 1 to 11 of CteG was amplified from pSVP310 using primers 652 and 2342. Another DNA fragment encoding amino acid residues 11 to 100 of CteG was amplified from pSVP310 using primers 2341 and 2128. Both fragments were then fused by overlapping PCR using primers 652 and 2128. The resulting DNA product was digested with KpnI and SalI and inserted into those sites of pmEGFP-C1 (Km^R).	This work.	**3.1**
pIP42	Transfection vector encoding mEGFP-$CteG_{1\text{-}100}$ with amino acid residue 12 replaced by an alanine. A DNA fragment encoding amino acid residues 1 to 12 of CteG was amplified from pSVP310 using primers 652 and 2344. Another DNA fragment encoding amino acid residues 12 to 100 of CteG was amplified from pSVP310 using primers 2343 and 2128. Both fragments were then fused by overlapping PCR using primers 652 and 2128. The resulting DNA product was digested with KpnI and SalI and inserted into those sites of pmEGFP-C1 (Km^R).	This work.	**3.1**
pIP43	Transfection vector encoding mEGFP-$CteG_{1\text{-}100}$ with amino acid residue 13 replaced by an alanine. A DNA fragment encoding amino acid residues 1 to 13 of CteG was amplified from pSVP310 using primers 652 and 2346. Another DNA fragment encoding amino acid residues 13 to 100 of CteG was amplified from pSVP310 using primers 2345 and 2128. Both fragments were then fused by overlapping PCR using primers 652 and 2128. The resulting DNA product was digested with KpnI and SalI and inserted into those sites of pmEGFP-C1 (Km^R).	This work.	**3.1**
pIP44	Transfection vector encoding mEGFP-$CteG_{1\text{-}100}$ with amino acid residue 14 replaced by an alanine. A DNA fragment	This work.	**3.1**

	encoding amino acid residues 1 to 14 of CteG was amplified from pSVP310 using primers 652 and 2348. Another DNA fragment encoding amino acid residues 14 to 100 of CteG was amplified from pSVP310 using primers 2347 and 2128. Both fragments were then fused by overlapping PCR using primers 652 and 2128. The resulting DNA product was digested with KpnI and SalI and inserted into those sites of pmEGFP-C1 (Km^R).		
pIP45	Transfection vector encoding mEGFP-$CteG_{1-100}$ with amino acid residue 15 replaced by an alanine. A DNA fragment encoding amino acid residues 1 to 15 of CteG was amplified from pSVP310 using primers 652 and 2350. Another DNA fragment encoding amino acid residues 15 to 100 of CteG was amplified from pSVP310 using primers 2349 and 2128. Both fragments were then fused by overlapping PCR using primers 652 and 2128. The resulting DNA product was digested with KpnI and SalI and inserted into those sites of pmEGFP-C1 (Km^R).	This work.	**3.1**
pIP46	Transfection vector encoding mEGFP-$CteG_{1-100}$ with amino acid residue 16 replaced by an alanine. A DNA fragment encoding amino acid residues 1 to 16 of CteG was amplified from pSVP310 using primers 652 and 2352. Another DNA fragment encoding amino acid residues 16 to 100 of CteG was amplified from pSVP310 using primers 2351 and 2128. Both fragments were then fused by overlapping PCR using primers 652 and 2128. The resulting DNA product was digested with KpnI and SalI and inserted into those sites of pmEGFP-C1 (Km^R).	This work.	**3.1**
pIP37	Transfection vector encoding mEGFP-$CteG_{1-100}$ with amino acid residue 17 replaced by a serine. A DNA fragment encoding amino acid residues 1 to 17 of CteG was amplified from pSVP310 using primers 652 and 2327. Another DNA	This work.	**3.1**

	fragment encoding amino acid residues 17 to 100 of CteG was amplified from pSVP310 using primers 2326 and 2128. Both fragments were then fused by overlapping PCR using primers 652 and 2128. The resulting DNA product was digested with KpnI and SalI and inserted into those sites of pmEGFP-C1 (Km^R).		
pIP32	Transfection vector encoding mEGFP-$CteG_{1-100}$ with amino acid residue 18 replaced by an alanine. A DNA fragment encoding amino acid residues 1 to 18 of CteG was amplified from pSVP310 using primers 652 and 2284. Another DNA fragment encoding amino acid residues 18 to 100 of CteG was amplified from pSVP310 using primers 2283 and 2128. Both fragments were then fused by overlapping PCR using primers 652 and 2128. The resulting DNA product was digested with KpnI and SalI and inserted into those sites of pmEGFP-C1 (Km^R).	This work.	**3.1**
pIP33	Transfection vector encoding mEGFP-$CteG_{1-100}$ with amino acid residue 19 replaced by an alanine. A DNA fragment encoding amino acid residues 1 to 19 of CteG was amplified from pSVP310 using primers 652 and 2286. Another DNA fragment encoding amino acid residues 19 to 100 of CteG was amplified from pSVP310 using primers 2285 and 2128. Both fragments were then fused by overlapping PCR using primers 652 and 2128. The resulting DNA product was digested with KpnI and SalI and inserted into those sites of pmEGFP-C1 (Km^R).	This work.	**3.1**
pIP55	Transfection vector encoding mEGFP-$CteG_{20-656}$. A DNA fragment encoding amino acid residues 20 to 656 of CteG was amplified from pSVP264 using primers 2180 and 653. The resulting DNA product was digested with SalI and KpnI and inserted into those sites of pmEGFP-C1 (Km^R).	This work.	**3.1**

pIP61	Transfection vector encoding mEGFP-$CteG_{40-656}$. A DNA fragment encoding amino acid residues 40 to 656 of CteG was amplified from pSVP264 using primers 2181 and 653. The resulting DNA product was digested with SalI and KpnI and inserted into those sites of pmEGFP-C1 (Km^R).	This work.	**3.1**
pIP62	Transfection vector encoding mEGFP-$CteG_{60-656}$. A DNA fragment encoding amino acid residues 60 to 656 of CteG was amplified from pSVP264 using primers 2182 and 653. The resulting DNA product was digested with SalI and KpnI and inserted into those sites of pmEGFP-C1 (Km^R).	This work.	**3.1**
pIP65	Transfection vector encoding mEGFP-$CteG_{FL}$ with amino acid residues 9 and 17 replaced by serines. A DNA fragment encoding amino acid residues 1 to 9 of CteG was amplified from pIP38 using primers 652 and 2432. Another DNA fragment encoding amino acid residues 9 to 656 of CteG was amplified from pIP38 using primers 2431 and 653. Both fragments were then fused by overlapping PCR using primers 652 and 653. The resulting DNA product was digested with KpnI and SalI and inserted into those sites of pmEGFP-C1 (Km^R).	This work.	**3.1**
pIP66	Transfection vector encoding mEGFP-$CteG_{FL}$ with amino acid residues 13 and 17 replaced by alanine and serine, respectively. A DNA fragment encoding amino acid residues 1 to 13 of CteG was amplified from pIP38 using primers 652 and 2436. Another DNA fragment encoding amino acid residues 13 to 656 of CteG was amplified from pIP38 using primers 2435 and 653. Both fragments were then fused by overlapping PCR using primers 652 and 653. The resulting DNA product was digested with KpnI and SalI and inserted into those sites of pmEGFP-C1 (Km^R).	This work.	**3.1**

pIP67	Transfection vector encoding mEGFP-$CteG_{FL}$ with amino acid residues 16 and 17 replaced by alanine and serine, respectively. A DNA fragment encoding amino acid residues 1 to 16 of CteG was amplified from pIP38 using primers 652 and 2494. Another DNA fragment encoding amino acid residues 16 to 656 of CteG was amplified from pIP38 using primers 2493 and 653. Both fragments were then fused by overlapping PCR using primers 652 and 653. The resulting DNA product was digested with KpnI and SalI and inserted into those sites of pmEGFP-C1 (Km^R).	This work.	**3.1**
pIP74	Transfection vector encoding mEGFP-$CteG_{FL}$ with amino acid residues 9 and 17 replaced by serine, and 12, 13 and 16 replaced by alanine. A DNA fragment encoding CteG with said amino acid substitutions was amplified from pIP73 using primers 652 and 653. The resulting DNA product was digested with KpnI and SalI and inserted into those sites of pmEGFP-C1 (Km^R).	This work.	**3.1**

[a]Amp^R: ampicillin resistance; Kan^R: kanamycin resistance.

Annexes Table 2. Oligonucleotides used in this work.

Number	Description	Sequence (5′ → 3′)	Restriction site
383	Forward primer to verify the presence of the transformation plasmid encoding CteG-2HA in *C. trachomatis* strains.	AAGCTCCAAGAGTTATTGG	-
652	Forward primer to construct pIP18, pIP19, pIP20, pIP21, pIP22, pIP26, pIP27, pIP28, pIP32, pIP33, pIP37, pIP38, pIP41, pIP42, pIP43, pIP44, pIP45, pIP46, pIP47, pIP48, pIP49, pIP50, pIP60, pIP63, pIP64 and pIP74.	GATCGATCGTCGACTCATTTGGTATTGGTAGTGC	SalI
653	Reverse primer to construct pIP38, pIP55, pIP56, pIP57, pIP58, pIP59, pIP61, pIP62 and pIP74.	GATCGGTACCCTAGATAGAGGAGCTTTGCACACC	KpnI
1483	Reverse primer to construct pIP40 and pIP54.	GATCGTCGACGTCTTAGGAGCTTTTTGCAATGC	SalI
1552	Forward primer to construct pIP69, pIP72 and pIP73.	GATCGCGGCCGCGGATAGAGGAGCTTTGCACACC	NotI
1680	Forward primer to construct pIP40, pIP69, pIP72 and pIP73.	GATCGGTACCTTCTTTATTATTGAGAAACG	KpnI
1865	Reverse primer to verify the presence of a group II intron within *cteG* in *C. trachomatis* strains.	TCTCGGAGTATACGGCTCTG	-
1932	Forward primer used in *16S* RT-qPCR.	GCGAAGGCGCTTTTCTAATTTAT	-
1933	Reverse primer used in *16S* RT-qPCR.	CCAGGGTATCTAATCCTGTTTGCT	-
1934	Forward primer used in *cteG* RT-qPCR, and to verify the presence of a group II intron within *cteG* in *C. trachomatis* strains.	ATGGAGCCGTTTGTGTGGTT	-
1935	Reverse primer for RT-qPCR of *cteG*.	CCTTCTTCGCTGTTACCCTCACT	-
2128	Forward primer to construct pIP15, pIP16, pIP17, pIP26, pIP27, pIP28, pIP32, pIP33, pIP37, pIP41, pIP42, pIP43, pIP44, pIP45, pIP46, pIP63 and pIP64.	GATCGGTACCCTAACTAGCAGAATATTTTTGGTAGC	KpnI

2173	Forward primer to verify the deletion of *pgp4* in *C. trachomatis* strains.	TGCGGCCCTAGAATTTGG	-
2180	Forward primer to construct pIP15 and pIP55.	GATCGATCGTCGACTCAGGCAGTGAGGGTAACAG	SalI
2181	Forward primer to construct pIP16 and pIP61.	GATCGATCGTCGACGGTTCTGGTGCTGCTTCTGC	SalI
2182	Forward primer to construct pIP17 and pIP62.	GATCGATCGTCGACAATGGTCCTAGTGTACAGATACC	SalI
2183	Forward primer to construct pIP18.	GATCGGTACCCTACACAAGGCTTTGCACATTAGC	KpnI
2184	Forward primer to construct pIP19.	GATCGGTACCCTAATTCCCTTCTGTAGAAGCGC	KpnI
2185	Forward primer to construct pIP20.	GATCGGTACCCTAACCTGAGGCGGCGTCTGAACCC	KpnI
2204	Forward primer to construct pIP21.	GATCGGTACCCTACACTCCTTCTTCGCTGTTACCC	KpnI
2205	Forward primer to construct pIP22.	GATCGGTACCCTATGATGAACCACACAAACGGCTCC	KpnI
2206	Forward overlap primer to construct pIP23.	CGAATTCTGCAGTCGACGCAGCTGCTATTGGTAGTGCTTGTTC	-
2207	Reverse overlap primer to construct pIP23.	GAACAAGCACTACCAATAGCAGCTGCGTCGACTGCAGAATTCG	-
2208	Forward overlap primer to construct pIP24.	GCAGTCGACTCATTTGGTGCTGCTGCTGCTTGTTCATCTTTATGG	-
2209	Reverse overlap primer to construct pIP24.	CCATAAAGATGAACAAGCAGCAGCAGCACCAAATGAGTCGACTGC	-
2210	Forward overlap primer to construct pIP25.	CGACTCATTTGGTATTGGTAGTGCTGCTGCATCTTTATGGAGCCG	-
2211	Reverse overlap primer to construct pIP25.	CGGCTCCATAAAGATGCAGCAGCACTACCAATACCAAATGAGTCG	-
2212	Forward overlap primer to construct pIP26.	GGTATTGGTAGTGCTTGTTCAGCTGCAGCGAGCCGTTTGTGTGGTTC	-
2213	Reverse overlap primer to construct pIP26.	GAACCACACAAACGGCTCGCTGCAGCTGAACAAGCACTACCAATACC	-
2214	Forward overlap primer to construct pIP27.	GCTTGTTCATCTTTATGGGCCGCTGCGTGTGGTTCATCAGGC	-
2215	Reverse overlap primer to construct pIP27.	GCCTGATGAACCACACGCAGCGGCCCATAAAGATGAACAAGC	-
2216	Forward overlap primer to construct pIP28.	CTTTATGGAGCCGTTTGGCTGCTGCATCAGGCAGTGAGGG	-
2217	Reverse overlap primer to construct pIP28.	CCCTCACTGCCTGATGCAGCAGCCAAACGGCTCCATAAAG	-

2283	Forward overlap primer to construct pIP32.	CTTTATGGAGCCGTTTGTGTGCTTCATCAGGCAGTGAGGG	-
2284	Reverse overlap primer to construct pIP32.	CCCTCACTGCCTGATGAAGCACACAAACGGCTCCATAAAG	-
2285	Forward overlap primer to construct pIP33.	GGAGCCGTTTGTGTGGTGCATCAGGCAGTGAGGGTAACAGC	-
2286	Reverse overlap primer to construct pIP33.	GCTGTTACCCTCACTGCCTGATGCACCACACAAACGGCTCC	-
2320	Forward overlap primer to construct pIP40 and pIP54.	GGTGTGCAAAGCTCCTCTATCTAGGGATGACATGTGATTCGCG	-
2321	Reverse overlap primer to construct pIP40 and pIP54.	CGCGAATCACATGTCATCCCTAGATAGAGGAGCTTTGCACACC	-
2326	Forward overlap primer to construct pIP37 and pIP38.	CTTTATGGAGCCGTTTGAGTGGTTCATCAGGCAGTGAGGG	-
2327	Reverse overlap primer to construct pIP37 and pIP38.	CCCTCACTGCCTGATGAACCACTCAAACGGCTCCATAAAG	-
2333	Forward primer to construct pIP51.	GATCGATCGTCGACACGAAAATTGTTTATCTACAC	SalI
2334	Forward primer to construct pIP51.	GATCGGTACCTTAATGATGGGGCTGTATATCATACG	KpnI
2335	Forward primer to construct pIP47.	GATCGGTACCCTAACAGCACCATAGTTCGGCACAAC	KpnI
2336	Forward primer to construct pIP48.	GATCGGTACCCTAGTTATCTAAAATTCCTGCACAAGAG	KpnI
2337	Forward primer to construct pIP49.	GATCGGTACCCTATCTGCGATCATTTCCTATAGCTGCGG	KpnI
2338	Forward primer to construct pIP50.	GATCGGTACCCTAATCTCTTAGACCCTCTAATAGGG	KpnI
2341	Forward overlap primer to construct pIP41.	GGTATTGGTAGTGCTTGTTCAGCTTTATGGAGCCGTTTGTGTGG	-
2342	Reverse overlap primer to construct pIP41.	CACACAAACGGCTCCATAAAGCTGAACAAGCACTACCAATACC	-
2343	Forward overlap primer to construct pIP42.	GTAGTGCTTGTTCATCTGCATGGAGCCGTTTGTGTG	-
2344	Reverse overlap primer to construct pIP42.	CACACAAACGGCTCCATGCAGATGAACAAGCACTAC	-
2345	Forward overlap primer to construct pIP43.	GGTAGTGCTTGTTCATCTTTAGCGAGCCGTTTGTGTGGTTCATCAGG	-
2346	Reverse overlap primer to construct pIP43.	CCTGATGAACCACACAAACGGCTCGCTAAAGATGAACAAGCACTACC	-
2347	Forward overlap primer to construct pIP44.	GCTTGTTCATCTTTATGGGCCCGTTTGTGTGGTTCATCAGG	-
2348	Reverse overlap primer to construct pIP44.	CCTGATGAACCACACAAACGGGCCCATAAAGATGAACAAGC	-

2349	Forward overlap primer to construct pIP45.	GCTTGTTCATCTTTATGGAGCGCTTTG TGTGGTTCATCAGGC	-
2350	Reverse overlap primer to construct pIP45.	GCCTGATGAACCACACAAAGCGCTC CATAAAGATGAACAAGC	-
2351	Forward overlap primer to construct pIP46.	GTTCATCTTTATGGAGCCGTGCGTGT GGTTCATCAGGCAGTGAGG	-
2352	Reverse overlap primer to construct pIP46.	CCTCACTGCCTGATGAACCACACGCA CGGCTCCATAAAGATGAAC	-
2388	Forward primer to construct pIP53 and pIP54.	GATCGGTACCAAAATTGTTATACAGA CGGC	KpnI
2389	Reverse primer to construct pIP53.	GATCGATGTCGACCGAGAATAATAA CCCAGCCC	SalI
2418	Forward primer to construct pIP56.	GATCGATCGTCGACTGTCAAGAGTCT CCTGCAGAAG	SalI
2419	Forward primer to construct pIP57.	GATCGATCGTCGACAACCCCTTCTGG AAAAGAGC	SalI
2420	Forward primer to construct pIP58.	GATCGATCGTCGACAGACCTATTTGG TTAACACCTAAAC	SalI
2421	Forward primer to construct pIP59.	GATCGATCGTCGACACAAACAACAG AACTCGAGC	SalI
2422	Forward primer to construct pIP60.	GATCGGTACCCTATGTTCTCTTAGAA ACTCTCAAGG	KpnI
2431	Forward overlap primer to construct pIP63 and pIP65.	ATTTGGTATTGGTAGTGCTAGTTCATC TTTATGGAGCCG	-
2432	Reverse overlap primer to construct pIP63 and pIP65.	CGGCTCCATAAAGATGAACTAGCAC TACCAATACCAAATG	-
2433	Forward overlap primer to construct pIP64.	GGTATTGGTAGTGCTTGTGCATCTTTA TGGAGCCGTTTG	-
2434	Reverse overlap primer to construct pIP64.	CAAACGGCTCCATAAAGATGCACAA GCACTACCAATACC	-
2435	Forward overlap primer to construct pIP66.	GTGCTTGTTCATCTTTAGCGAGCCGTT TGAGTGGTTCATC	-
2436	Reverse overlap primer to construct pIP66.	GATGAACCACTCAAACGGCTCGCTA AAGATGAACAAGCAC	-
2493	Forward overlap primer to construct pIP67.	GTTCATCTTTATGGAGCCGTGCGAGT GGTTCATCAGGCAGTG	-
2494	Reverse overlap primer to construct pIP67.	CACTGCCTGATGAACCACTCGCACGG CTCCATAAAGATGAAC	-
2554	Forward primer to construct pIP68.	GATCGGCGCGCCAATTTTGCATAACA AACCCCGTAATTC	AscI
2555	Reverse primer to construct pIP68.	GATCGGCGCGCCAAGGCTGAATAGA CAACTTACTC	AscI

2631	Reverse primer to verify the deletion of *pgp4* in *C. trachomatis* strains.	GTGGTATGGGTTAATGCCC	
2825	Forward overlap primer to construct pIP72.	GGTATTGGTAGTGCTTCTTCATCTTTATGGAGCCGTTTGTCTGGTTCATCAGGCAGTGAGG	-
2826	Reverse overlap primer to construct pIP72.	CCTCACTGCCTGATGAACCAGACAAACGGCTCCATAAAGATGAAGAAGCACTACCAATACC	-
2827	Forward overlap primer to construct pIP73.	GGTATTGGTAGTGCTTCTTCATCTGCAGCGAGCCGTGCGTCTGGTTCATCAGGCAGTGAGG	-
2828	Reverse overlap primer to construct pIP73.	CCTCACTGCCTGATGAACCAGACGCACGGCTCGCTGCAGATGAAGAAGCACTACCAATACC	-

Annexes Table 3. *Chlamydia trachomatis* strains used and constructed in this study.

Strains	Description	Relevant genotype	Source/Refs.	Chapter
L2/434/Bu ACE051	Wild-type strain.	*cteG*$^{+}$ *pgp4*$^{+}$	From Derek J. Fisher (originating from Tony Maurelli's lab; University of Florida).	**3.2**
L2/434 (pSVP247/pVector[Pgp4^{+}])	Derivative of L2/434/Bu ACE051 carrying plasmid pSVP247/pVector[Pgp4^{+}].	*cteG*$^{+}$ *pgp4*$^{+}$	This work.	**3.2**
L2/434 (pIP68/pVector[Pgp4^{-}])	Derivative of L2/434/Bu ACE051 carrying plasmid pIP68/pVector[Pgp4^{-}].	*cteG*$^{+}$ *pgp4*$^{-}$	This work.	**3.2**
L2/434 (pIP69/pCteG-2HA[Pgp4^{-}])	Derivative of L2/434/Bu ACE051 carrying plasmid pIP69/pCteG-2HA[Pgp4^{-}].	*cteG*$^{+}$/*cteG-2HA*$^{+}$ *pgp4*$^{-}$	This work.	**3.2**
cteG::aadA	Derivative of L2/434/Bu ACE051 with *cteG* inactivated.	*cteG*$^{-}$ *pgp4*$^{+}$	(Pais *et al.*, 2019).	**3.2**
cteG::aadA (pIP40/pCteG)	Derivative of *cteG::aadA* carrying plasmid pIP40/pCteG.	*cteG*$^{-}$/*cteG*$^{+}$ *pgp4*$^{+}$	This work.	**3.2**
cteG::aadA (pIP54/pFabI-CteG)	Derivative of *cteG::aadA* carrying plasmid pIP54/pFabI-CteG (harboring *cteG* and *ctl0359/fabI*).	*cteG*$^{-}$/*cteG*$^{+}$ *pgp4*$^{+}$	This work.	**3.2**
cteG::aadA (pIP53/pFabI-CteG-CTL0361)	Derivative of *cteG::aadA* carrying plasmid pIP53/pFabI-CteG-CTL0361 (harboring *cteG, ctl0359/fabI* and *ctl0361*).	*cteG*$^{-}$/*cteG*$^{+}$ *pgp4*$^{+}$	This work.	**3.2**
cteG::aadA (pSVP247/pVector[Pgp4^{+}])	Derivative of *cteG::aadA* carrying plasmid pSVP247/pVector [Pgp4^{+}].	*cteG*$^{-}$ *pgp4*$^{+}$	This work.	**3.2**
cteG::aadA (pIP68/pVector[Pgp4^{-}])	Derivative of *cteG::aadA* carrying plasmid pIP68/pVector[Pgp4^{-}].	*cteG*$^{-}$ *pgp4*$^{-}$	This work.	**3.2**
cteG::aadA (pSVP264/pCteG$_{W1}$-2HA/pCteG-2HA/ pCteG-2HA[Pgp4^{+}])	Derivative of *cteG::aadA* carrying plasmid pSVP264/ pCteG$_{W1}$-2HA/pCteG-2HA/pCteG-2HA [Pgp4^{+}].	*cteG*$^{-}$/*cteG-2HA*$^{+}$ *pgp4*$^{+}$	(Pais *et al.*, 2019).	**3.1, 3.2**
cteG::aadA (pIP69/pCteG-2HA[Pgp4^{-}])	Derivative of *cteG::aadA* carrying plasmid pIP69/pCteG-2HA [Pgp4^{-}].	*cteG*$^{-}$/*cteG-2HA*$^{+}$ *pgp4*$^{-}$	This work.	**3.2**
L2/25667R	Plasmidless L2 strain.	*cteG*$^{+}$ *pgp4*$^{-}$	(From Agathe Subtil; Peterson *et al.*, 1990).	**3.2**
cteG::aadA (pIP72/pCteG$_{C9S,\ C17S}$-2HA/ pCteG$_{2aa}$-2HA)	Derivative of *cteG::aadA* carrying plasmid pIP72/pCteG$_{C9S,\ C17S}$-2HA/pCteG$_{2aa}$-2HA.	*cteG*$^{-}$/*cteG*$_{C9S,\ C17S}$-*2HA*$^{+}$	This work.	**3.1**
cteG::aadA (pIP73/pCteG$_{C9S,\ L12A,\ W13A,\ L16A,\ C17S}$-2HA/pCteG$_{5aa}$-2HA)	Derivative of *cteG::aadA* carrying plasmid pIP73/pCteG$_{C9S,\ L12A,\ W13A,\ L16A,\ C17S}$-2HA/pCteG$_{5aa}$-2HA.	*cteG*$^{-}$/*cteG*$_{C9S,\ L12A,\ W13A, L16A, C17S}$-*2HA*$^{+}$	This work.	**3.1**

Annexes Table 4. Nucleotide changes between *C. trachomatis* strains L2 434/Bu (*cteG*$^+$) and *cteG::aadA* (*cteG*$^-$) used in this study.

Reference genome position[1]	Base pair change	Amino acid change	Variant type[2]	Nt in reference strain[3]	L2/434/Bu locus tag[4]	Serovar D locus tag[5]	Gene	Gene product	L2/434/Bu (*cteG*$^+$) (frequency)	*cteG::aadA* (*cteG*$^-$) (frequency)
127,340	*C659T*	A220V	Missense	C	CTL0103	CT734		Putative lipoprotein	C (100%)	T (100%)
282,299	*C11T*	S4L	Missense	G	CTL0226	CT853		Putative integral membrane protein	G (100%)	A (100%)
295,554	*G592T*	A198S	Missense	C	CTL0237	CT862	*lcrH*	Type III secretion chaperone	C (90%) A (10%)	A (100%)
929,598	C→A		Noncoding	C	Upstream CTL0805 /CTL0806		*hisS* /*uhpC*	*hisS*: histidyl-tRNA synthetase; *uhpC*: putative sugar phosphate permease	C (84%) A (16%)	A (100%)
929,717	A→C		Noncoding	C	Upstream CTL0805 /CTL0806	Upstream CT543 /CT544	*hisS* /*uhpC*	*hisS*: histidyl-tRNA synthetase; *uhpC*: putative sugar phosphate permease	A(66%); C(34%)	C (100%)
1,017,120	*C391A*	R131S	Missense	C	CTL0882	CT618		Putative membrane protein	C (58%) A (25%) T (17%)	A (100%)

[1] Nucleotide position in *C. trachomatis* L2/434/Bu reference strain genome (GenBank accession number AM884176.1).

[2] Single nucleotide variant classification: Noncoding – outside coding sequence; Missense – substitution.

[3] Nucleotide described for *C. trachomatis* L2/434/Bu reference strain.

[4] Locus designation in *C. trachomatis* L2/434/Bu reference strain.

[5] Locus designation in *C. trachomatis* D/UW-3/CX (Serotype D) reference strain (GenBank accession number AAC68146.1).

Annexes Table 5. Mass spectrometry results - proteins recovered in all mEGFP-CteG$_{FL}$ (CteG_1, CteG_2 and CteG_3) and none of the mEGFP (EGFP_1 and EGFP_2) eluates, and their respective parameters for each sample.

	Protein FDR Confidence	Accession	Description	Exp. q-value	Sum PEP Score	% Coverage	# Peptides	# PSMs	# Unique Peptides	Biological Process	Cellular Component	Molecular Function
CteG_1	High	P10412	Histone H1.4	0	21.201	22	4	9	1	Cell organization and biogenesis; metabolic process; regulation of biological process	Chromosome; nucleus	DNA binding; metal ion binding; nucleotide binding; protein binding; RNA binding
CteG_2	High			0	25.839	22	4	9	1			
CteG_3	High			0	28.441	22	4	13	1			
CteG_1	High	P62753	40S ribosomal protein S6	0	20.67	14	4	5	4	Cell death; cell differentiation; cell organization and biogenesis; cell proliferation; metabolic process; regulation of biological process; response to stimulus; transport	Cytoplasm; cytosol; membrane; nucleus; ribosome	Protein binding; RNA binding; structural molecule activity
CteG_2	High			0	33.272	22	6	9	6			
CteG_3	High			0	44.743	18	6	10	6			
CteG_1	High	C9J3L8	Translocon-associated protein subunit alpha	0	29.025	31	4	5	4		Membrane	
CteG_2	High			0	36.142	35	5	7	5			
CteG_3	High			0	22.168	22	4	6	4			
CteG_1	High	G5EA06	28S ribosomal protein S27, mitochondrial	0	18.459	21	6	7	6	Cell organization and biogenesis; cell proliferation; metabolic process; regulation of biological process	Cytoplasm; membrane; mitochondrion; ribosome	Protein binding; RNA binding
CteG_2	High			0	8.911	9	3	3	3			
CteG_3	High			0	14.946	14	3	3	3			

CteG_1	High	E9PCT5	Caveolin	0	9.718	19	2	4	2	Cell organization and biogenesis	Golgi; membrane	
CteG_2	Medium			0.016	1.287	13	1	1	1			
CteG_3	High			0	7.273	19	2	3	2			
CteG_1	High	H7BXY3	ATP-dependent RNA helicase DHX30	0	3.968	2	2	2	2	Cell organization and biogenesis; metabolic process	Cytosol; mitochondrion	Catalytic activity; nucleotide binding; protein binding; RNA binding
CteG_2	High			0	12.231	4	4	5	4			
CteG_3	High			0	34.138	8	8	10	8			
CteG_1	High	Q02750	Dual specificity mitogen-activated protein kinase kinase 1	0	13.223	12	3	4	3	Cell differentiation; cell proliferation; cellular component movement; metabolic process; regulation of biological process; response to stimulus	Cytoplasm; cytoskeleton; cytosol; endoplasmic reticulum; endosome; Golgi; membrane; mitochondrion; nucleus	Catalytic activity; enzyme regulator activity; nucleotide binding; protein binding; signal transducer activity
CteG_2	High			0	9.351	16	4	4	4			
CteG_3	Medium			0.034	0.885	7	1	1	1			
CteG_1	High	B1AHE3	Ataxin-10	0	3.151	12	1	1	1	Cell organization and biogenesis	Cytoplasm; cytosol; membrane	Protein binding
CteG_2	High			0	14.481	27	3	3	2			
CteG_3	High			0	5.521	12	1	1	1			
CteG_1	High	Q8WVZ9	Kelch repeat and BTB domain-containing protein 7	0.005	1.776	1	1	1	1	Metabolic process; regulation of biological process; response to stimulus	Cytosol	Catalytic activity; protein binding
CteG_2	High			0.006	1.814	1	1	2	1			
CteG_3	High			0.009	1.421	1	1	1	1			
CteG_1	High	Q03252	Lamin-B2	0	7.623	4	3	3	2			

CteG_2	High			0	9.118	5	4	5	2		Membrane; nucleus	Catalytic activity; motor activity; protein binding; structural molecule activity
CteG_3	High			0	10.924	6	5	5	3			
CteG_1	High	Q08945	FACT complex subunit SSRP1	0.004	1.904	2	1	1	1	Metabolic process; regulation of biological process; response to stimulus	Chromosome; nucleus	DNA binding; protein binding; RNA binding
CteG_2	High			0.003	2.255	2	1	1	1			
CteG_3	High			0	22.42	11	6	6	6			
CteG_1	High	E9PKP7	Nucleolar transcription factor 1	0	6.534	2	1	2	1		Nucleus	DNA binding
CteG_2	High			0	6.619	3	2	2	2			
CteG_3	High			0	13.659	7	4	4	4			
CteG_1	High	M0R299	rRNA 2'-O-methyltransferase fibrillarin (Fragment)	0	6.208	12	2	2	2	Metabolic process	Nucleus	Catalytic activity; RNA binding
CteG_2	High			0	6.169	12	2	2	2			
CteG_3	High			0	22.993	39	6	7	6			
CteG_1	High	Q9UKV3	Apoptotic chromatin condensation inducer in the nucleus	0.002	2.418	1	1	1	1	Cell death; cell differentiation; cell organization and biogenesis; metabolic process; regulation of biological process	Cytosol; membrane; nucleus	Catalytic activity; protein binding; RNA binding
CteG_2	Medium			0.024	1.066	1	1	1	1			
CteG_3	Medium			0.014	1.211	1	1	1	1			
CteG_1	High	A0A3B3IRR6	UPF0488 protein C8orf33	0	7.082	7	1	1	1			
CteG_2	High			0	6.346	7	1	1	1			
CteG_3	High			0	6.977	7	1	1	1			

CteG_1	High	P38919	Eukaryotic initiation factor 4A-III	0	11.639	5	2	4	1	Metabolic process; regulation of biological process; response to stimulus; transport	Cytoplasm; cytosol; membrane; nucleus; spliceosomal complex	Catalytic activity; nucleotide binding; protein binding; RNA binding
CteG_2	High			0	13.844	8	3	4	2			
CteG_3	High			0	49.737	26	8	10	7			
CteG_1	High	Q99459	Cell division cycle 5-like protein	0	3.549	3	2	2	2	Cell differentiation; cell organization and biogenesis; metabolic process; regulation of biological process; response to stimulus	Cytoplasm; membrane; nucleus; spliceosomal complex	DNA binding; protein binding; RNA binding
CteG_2	High			0.003	2.161	1	1	1	1			
CteG_3	High			0	5.676	6	3	3	3			
CteG_1	High	C9J384	Protein CMSS1 (Fragment)	0.005	1.76	3	1	1	1			Nucleotide binding
CteG_2	High			0	3.808	9	2	2	2			
CteG_3	High			0.005	1.863	3	1	1	1			
CteG_1	High	Q99848	Probable rRNA-processing protein EBP2	0.004	2.207	3	1	1	1	Metabolic process	Nucleus	Protein binding; RNA binding
CteG_2	High			0	4.354	4	1	2	1			
CteG_3	High			0	30.859	17	5	8	5			
CteG_1	High	H3BV03	Protein fantom	0.009	1.624	1	1	1	1	Metabolic process; regulation of biological process	Cytoplasm; cytosol; membrane	Catalytic activity; motor activity; protein binding
CteG_2	High			0	4.336	3	2	2	2			
CteG_3	High			0.002	2.078	1	1	1	1			
CteG_1	High	Q6DCA0	AMMECR1-like protein	0.004	2.148	3	1	1	1			
CteG_2	High			0	5.661	9	2	3	1			
CteG_3	Medium			0.034	0.873	3	1	1	1			
CteG_1	High			0.004	1.888	1	1	1	1	Transport		

CteG_2	Medium	A0A0A0MSB8	Exocyst complex component 7	0.011	1.516	1	1	2	1			
CteG_3	Medium			0.027	0.963	1	1	1	1			
CteG_1	High	Q9Y3D7	Mitochondrial import inner membrane translocase subunit TIM16	0.004	1.908	18	1	1	1	Regulation of biological process; transport	Membrane; mitochondrion; organelle lumen	Protein binding
CteG_2	High			0	4.272	18	1	1	1			
CteG_3	High			0	2.534	18	1	2	1			
CteG_1	High	Q9BRQ8	Apoptosis-inducing factor 2	0	4.455	3	1	1	1	Cell death; cell organization and biogenesis; metabolic process; regulation of biological process	Cytoplasm; cytosol; membrane; mitochondrion	Catalytic activity; DNA binding; nucleotide binding
CteG_2	High			0	16.298	21	5	5	5			
CteG_3	High			0	8.195	8	2	2	2			
CteG_1	High	Q9Y3P9	Rab GTPase-activating protein 1	0	5.342	1	1	1	1	Regulation of biological process; transport	Cytoplasm; cytoskeleton; cytosol	Enzyme regulator activity; protein binding
CteG_2	High			0	15.697	5	4	5	4			
CteG_3	Medium			0.015	1.128	1	1	1	1			
CteG_1	High	P41208	Centrin-2	0	12.661	19	1	1	1	Cell division; cell organization and biogenesis; metabolic process; regulation of biological process; response to stimulus	Cytoplasm; cytoskeleton; cytosol; nucleus	Metal ion binding; protein binding
CteG_2	High			0	21.835	37	3	4	3			
CteG_3	High			0	12.72	29	2	2	2			
CteG_1	Medium	Q9NQT5	Exosome complex component RRP40	0.014	1.314	6	1	1	1	Metabolic process; regulation of biological process	Cytoplasm; cytosol; nucleus	Catalytic activity; protein binding; RNA binding
CteG_2	High			0	3.739	8	1	1	1			
CteG_3	High			0	4.826	13	1	1	1			

CteG_1	Medium	Q9BQ67	Glutamate-rich WD repeat-containing protein 1	0.019	1.215	4	1	1	1	Cell organization and biogenesis; metabolic process	Chromosome; cytosol; nucleus	DNA binding; protein binding; RNA binding
CteG_2	High			0	4.07	8	2	2	2			
CteG_3	High			0	21.428	11	3	3	3			
CteG_1	High	H7C0G1	Transmembrane protein 245 (Fragment)	0.004	1.93	1	1	1	1		Membrane	
CteG_2	High			0	6.408	3	1	2	1			
CteG_3	High			0.006	1.637	6	1	1	1			

2022

Inês Isabel Serrano Pereira

CteG, a *Chlamydia trachomatis* protein involved in host cell lytic exit