Functional and structural analyses of the terminal organelle of *Mycoplasma genitalium*

Luis González González

2015

Doctoral dissertation submitted to fulfil the requirements to obtain the Doctor of Phylosophy Degree in Bioquímica, Biologia Molecular i Biomedicina at Universitat Autònoma de Barcelona

This work has been performed at the Institut de Biotecnologia i Biomedicina in the Departament de Bioquimica i Biologia molecular under the supervision of Dr. Jaume Piñol Ribas and Dr. Enrique Querol Murillo

Universitat Autònoma de Barcelona

Functional and structural analyses of the terminal organelle of *Mycoplasma genitalium*

Luis González González

2015

Tesi doctoral

Doctorat de Bioquímica, Biologia Molecular i Biomedicina
Universitat Autònoma de Barcelona

Institut de Biotecnologia i Biomedicina
Departament de Bioquimica i Biologia molecular
Directors: Dr. Jaume Piñol Ribas and Dr. Enrique Querol Murillo

Dr. Jaume Piñol Ribas

Dr. Enrique Querol Murillo

Si sufres injusticias, consuélate, porque la verdadera desgracia es cometerlas

Pitágoras

A mis padres, Teresa y Jose Luis
A mi abuela Pilar

A mi hermano David
y a todos los que han hecho posible esta hazaña

SUMMARY

Mycoplamsa genitalium is and human pathogen and the causative agent of non-gonococcal non-chlamydial urethritis in men and pelvic inflammatory disease and cervicitis. Mycoplasmas, besides being interesting as minimal cells (given the small size of its genome), they also have unique features only present in its *genus*. In particular, it has been detected the presence of mechanisms of adhesion and motility that, in addition to being involved in the infection mechanism, are only found in this *genus*. Specially, the mechanism of motility of *M. genitalium* motility involves a polar structure containing characteristic cytoskeleton. It is known that the cytoskeleton is composed of several proteins involved in the of adhesion and motility processes. In the first three chapters of this thesis dissertation the role of three of MG219, MG318 (also known as P32) and MG386 have been stablished. The study was conducted by obtaining null mutants strains of these proteins. The MG219 protein has been found to be necessary for the proper functionality of the motility machinery. By fusion to fluorescent proteins it has been determined the subcellular localization of this MG219, which located at the nearest part of the terminal organelle relative to the cell body. Cells in the absence of MG219 move slower than half speedy (and half frequently) the cells of the wild type strain. This speed reduction is concomitant with the appearance of a greater number of dividing cells and cell with multiple terminal organelles. In a similar manner, cells of the strain lacking P32 move at lower speeds (and with half also half frequently) than cells of the wild type strain. Furthermore, it has been determined that the N-terminal P32 plays an important role in protein stability P110 and P140, the major adhesins of *M. genitalium*. It has also been established that the P32 protein is critical to the morphology of the most distal part of the terminal organelle relative to the cell body. The last mutant generated in this work, the null mutant for MG386, presents significant alterations in both cell morphology and motility. The cells of this strain show motility half frequently than the wild type strain but move to a velocities as greater as1.7 times than the reference strain. Surprisingly, this strain has a high number of terminal organelles detached from the cell body, suggesting an important role of protein MG386 anchoring the organelle to the cell body.

It has been observed that the membrane around the cytoskeleton is completely covered by the adhesion complex or "nap". By electron microscopy studies it has determined the structure by cryo-electron tomography at 3.5 nm and at 1.9 nm single particle by negative staining TEM of the purified P110 and P140 complex. Additionally, cryo-electron tomography also allowed to determine at low resolution the structure of the terminal button of the cytoskeleton, revealing that the plates forming most of the cytoskeleton are actually rings about 20 nm in diameter. In addition, 14 mutants lacking different proteins (or domains thereof) related to motility and / or adhesion have been analysed by cryo-electron tomography. All these data taken together provides an overview of the prior knowledge—in addition to the data generated in this work—of the role of the proteins involved motility and cytoskeleton formation in *M. geniatlium*.

Index

1. Abbreviations 1

2. General introduction 3
2.1 *Mycoplasma* genus 3

2.2 *Mycoplasma genitalium* 4

2.2.1 The Human pathogen 4
2.2.2 Genetic manipulation 6
2.2.3 Morphology and characteristic structures 7
2.2.4 Ultrastructure and components of the attachment organelle 10
2.2.5 Adherence and motility mechanisms 23
2.2.6 Cell division 25

3. Objectives 27

4. Chapter I: Functional analysis of a unique protein of *Mycoplasma genitalium*: MG219 29
4.1 Introduction 29
4.1.1 Previous knowledge of MG219 29

4.2 Results 30
4.2.1 Recombinant expression of MG219 and polyclonal antibody generation 30
4.2.2 MG_219 transcription start site 31
4.2.3 Obtaining a null mutant for MG219: Δmg219 strain 32
4.2.4 Confirmation of MG219 null mutant 33
4.2.5 Obtaining Δmg219 based mutant strains 35
4.2.6 Confirmation ofΔmg219 based mutant strains 39
4.2.7 Subcellular localization of MG219 and MG318 proteins 42
4.2.8 Characterization of MG219 null mutant 43
4.2.8.1 Gliding motility 44
4.2.8.2 Cellular morphology 47

4.3 Discussion 50

5. Chapter II: Functional analysis of the protein encoded in MG_318 locus: P32 54
5.1 Introduction 54
5.1.1 Previous knowledge of P32 protein 54

5.2 Results 56
5.2.1 Recombinant expression of P32 and polyclonal antibody generation 56
5.2.2 Plasmids for obtaining a P32 null mutant, a P32 C-ter lacking mutant and for reintroducing the WT allele 58
5.2.3 Confirmation of MG318 mutants 59
5.2.4 Characterization of MG318 mutants 67
5.2.4.1 Gliding motility 67
5.2.4.2 Cellular morphology 70

5.3 Discussion 73

6. Chapter III: Functional analysis of the high molecular weight protein MG386 76
6.1 Introduction 76
6.1.1 Previous knowledge of MG386 protein 76

6.2 Results 77
6.2.1 Obtaining MG386 null mutant
6.2.2 Confirmation of MG386 related mutants
6.2.3 Characterization of MG386 mutants 80
6.2.3.1 Gliding motility 80
6.2.3.2 Cellular morphology 83

6.3 Discussion 84

7. Chapter IV: Structural characterization of macromolecular complexes of *Mycoplasma genitalium* by Cryo-electron tomography 87

7.1 Introduction 87
7.1.1 Cryo-electron microscopy and cryo-electron tomography 87
7.1.2 Cryo electron tomography 90

7.2 Results 93
7.2.1 Cryo electron tomography of *M. genitalium* WT strain 93
7.2.2 Subtomogram averaging of *M. genitalium* WT strain macromolecular complexes 94
7.2.3 Isolation and structural characterization of the adhesion complex of *M. genitalium* 100
7.2.4 Cryo electron tomography of MG191 and MG192 null mutants 104
7.2.5 Cryo electron tomography of MG312 null mutant 105
7.2.6 Cryo electron tomography of MG217 null mutant 106
7.2.7 Cryo electron tomography of MG317 null mutant 107
7.2.8 Cryo electron tomography of MG491 mutants 108
7.2.9 Cryo electron tomography of MG200 null mutant 109
7.2.10 Cryo electron tomography of MG386 mutants 110
7.2.11 Cryo electron tomography of MG318 mutants 112
7.2.12 Cryo electron tomography of MG219 mutant 113
7.2.13 Cryo electron tomography of MG-P3 mutant 113

7.3 Discussion 114

8. General discussion 122

9. Conclusions 130

10. Materials and Methods 131
10.1 Bacterial growth and strains 131
10.2 Molecular cloning and RNA-DNA manipulations 137
10.3 Recombinant expression of MG219 and MG318 142
10.4 Transformation of *M. genitalium* 143

10.5 Southern blotting 144
10.6 SDS-PAGE and Western blotting 146
10.7 Microcinematographic studies 149
10.8 Epifluorescence microscopy 149
10.9 Purification of P110 and P140 from *M. genitalium* cells 150
10.10 Negative staining transmission electron microscopy 150
10.10.1 Single particle selection, clasiication and averaging 150
10.11 Scanning electron microscopy 151
10.12 Cryo-electron tomography 151
10.12.1 Sample preparation 151
10.12.2 Data acquisition 152
10.12.3 Data processing 152
10.12.3.1 3D volume reconstruction and analyses 152
10.12.3.2 Particle picking and subtomogram averaging 152

11. Acknowledgements 153

12. Bibliography 155

13. Appendices 167

1. Abbreviations

±	plus/minus: indicates standard error (SE)
Ø	diameter
aac-aph	acetyltransferase(6′)-Ie-phosphotransferase(2′′)-Ia [AAC(6′)-Ie-APH(2′′)-Ia
ATCC	American Type Culture Collection
ATP	adenosine triphosphate
BSA	bovine serum albumin
β-Gal	β–galactosidase
bla	ampicillin resistance gene.
CET	cryo-electron tomography
CmR	chloramphenicol resistance gene
CCD	scintillator-coupled devices
CMC	critical micelle concentration
CMOS	complementary metal-oxide semiconductors
CET	cryo-electron tomography
CFP	cyan fluorescent protein
cfu	colony-forming unit
crl	cytadherence regulatory locus
C-ter	carboxy-terminus
dNTP	deoxy-nucleotide triphosphate
EAGR	enriched aromatic glycine residues
EDTA	ethylenediamintetraacetic
ESRF	European Snychrotron Radiaton Facility
EYFP	enhanced yellow fluorescent protein
EGFP	enhanced green fluorescent protein
FBS	fetal bovine serum
GmR	gentamicin resistance gene
HMDS	Hexamethyldisiloxane
HA	haemadsoption
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HMW	high molecular weight
IBMB	Institut de Biologia Molecular de Barcelona
IBMC	Instituto de Biologia Molecular e Celular
IMAC	immobilized metal ion affinity chromatography
IPTG	isopropylβ-D-1-thyogalactopyianoside
kb	kilobase
kDa	kilodalton
LB	lysogenic broth
LHR	left homology region
mA	miliampere
Mw	molecular mass or molecular weight
NHBE	normal human bronchial epithelial (cell line)
NGS	next generation sequencing

N-ter	amino terminal
o/n	overnight
OG	n-octyl-β-D-glucopyranoside
OD	optical deasily
ORF	open reading frame
OriV	vegetative origin of replication
PBS	phosphate buffer saline
PCR	polymense chain reaction
pDNA	plasmidic DNA
PTM	post-translational modification
RBS	ribosome binding site
RHR	right homology region
RPM	revolutions per minute
SEC	size exclusion cromatography
SDS	sodium dodecyl sulphate
SDS-PAGE	SDS-polyacrilanúde gel electrophoresis
SEM	scanning electron microscopy
SP4	spiroplasma medium 4
TEMED	N,N,N,N'-tetramethhylethane-1,2-diamine
tetM438	tetracycline resistance gene with MG_438 promoter
Tris	2-Amino-2-hydroxymethyl-propane-1,3-diol
Trp	tryptophan
TSS	transcription start site
TO	terminal organelle
V	volt
v/v	volume fraction
WT	wild type
w/v	mass concentration
YFP	yellow fluorescent protein
X-Gal	5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside

2. General introduction

2.1 *Mycoplasma* genus

Mycoplasma genus belongs to the *Mollicutes* class. Etymologically, the word mollicutes is formed by the words *mollis*—meaning soft and *cutis*—meaning skin. The origin of this name comes from one of its most characteristic features; all *Mollicutes* lack the peptidoglycan layer resulting in the absence of cell wall thus exhibiting pleomorphic morphology. Other characteristic features from mollicutes can be observed in Table I.1. *Mycoplasma, Ureaplasma, Entomoplasma, Mesoplasma, Spiroplasma, Acholeplasma, Anaeroplasma* and *Asteroleplasma* belong to this class (Table I.2). All of them present a low G+C content (Fraser et al., 1995, Himmelreich et al., 1996). Phylogenetic analyses reveal that *Mollicutes* are closely related to Gram positive bacteria from *Lactobacillus* genus (Weisburg et al., 1989). One of the important traits to consider when performing molecular biology studies of *mollicutes* is that they use a distinct genetic code(Inamine et al., 1990). The *Spiroplasma/Mycoplasma* code uses the TGA codon as Trp codon. Therefore, recombinant expression of proteins encoded within the genome of *mollicutes* is routinely impossible when expressing ORFs if the common hosts for recombinant protein expression are used. Since TGA codon is the most common one encoding for tryptophan (6,2 ‰) directed mutagenesis to TGG is performed to allow protein expression in other organisms. Even specific mutagenesis tools for mycoplasma genes harbouring multiple TGA codons have been developed (Hames et al., 2005).

Table I.1. Disctintive traits of *mollicutes* (Razin et al., 1998).

Property	Mollicutes	Other eubacteria
Cell wall	Absent	Present
Plasma membrane	Cholesterol present in most species	Cholesterol absent
Genome size	580–2,220 kb	1,050–>10,000 kb
G+C content of genome	23–40 mol%	25–75 mol%
No. of rRNA operons	1 or 2*	1–10
5S rRNA length	104–113 nt	>114 nt
No. of tRNA genes	30 (*M. capricolum*), 33 (*M. pneumoniae*)	84 (*B. subtilis*), 86 (*E. coli*)
UGA codon usage	Tryptophan codon in *Mycoplasma, Ureaplasma, Spiroplasma, Mesoplasma*	Stop codon
RNA polymerase	Rifampin resistant	Rifampin sensitive

* Three rRNA operons in *Mesoplasma lactucae*

Mycoplasmas present the smallest genomes discovered to date ranging from 560 kb to 2.2 kb of organisms capable of self-replicate in axenic medium. Although it was thought that this small genomes represent a form of genetic simplicity (Fraser et al., 1995, Calderon-Copete et al., 2009, Vasconcelos et al., 2005, Minion et al., 2004, Jaffe et al., 2004b, Papazisi et al., 2003, Sasaki et al., 2002, Chambaud et al., 2001, Himmelreich et al., 1996) in the last years different studies revealed an unexpected complexity lying within the genome of these organisms (Peterson and Fraser, 2001, Guell et al., 2009, Yus et al., 2009, Kuhner et al., 2009, Lluch-Senar

et al., 2015a, Wodke et al., 2013, Lluch-Senar et al., 2013, Maier et al., 2013, van Noort et al., 2012, Yus et al., 2012). Understanding the process of life as a whole has always been one of the mayor goals of biology. A possible first step to attain this objective would be acquiring the knowledge of the minimal functions necessary for life. Although some cell models for mycoplasma have been recently developed, we are still far to fully comprehend the minimal components of life (Karr et al., 2012, Macklin et al., 2014).

2.2 *Mycoplasma genitalium*

M. genitalium was the second organism to be sequenced completely, after *Haemophilus influenzae* in 1995. Since the genome sequence of *Mycoplasma parvum* (564 kb (do Nascimento et al., 2013)), *M. genitalium* (580 kb (Fraser et al., 1995)) is no longer considered to be the smallest self-replicating organism. Despite this fact, *M. genitalium* is still an appealing model for minimal cell studies with its 527 predicted ORFs from which 482 encode for proteins and 43 for ribosomal RNA or transfer RNA. The boom of NGS methodologies has allowed the sequencing of many organisms of this class, even the sequences of different strains of *M. genitalium* are currently available in the public repositories (McGowin et al., 2012).

2.2.1 The human pathogen

M. genitalium is the causative agent of non-gonococcal urethritis and pelvic inflammatory disease (Jensen, 2004, McGowin and Anderson-Smits, 2011).It has been reported a prevalence of 7.3% in high-risk populations or 2.0 in low-risk population in women, and a 13-46% on male patients with acute non-gonococcal urethritis. It is remarkable that its asymptomatic presence is common (Cazanave et al., 2012). Those rates are between those of *Chlamydia trachomatis* and *Neisseria gonorrhoeae*. There are also correlations of pelvic inflammatory disease and infertility—indicating that upper tract inflammations might be attributed to this pathogen (Edwards et al., 2006, Hitti et al., 2010). Persistent urethritis after antibiotic treatment has also been linked to *M.* genitalium (Maeda et al., 2009).

Although the following results are preliminary and should be treated carefully, *M. genitalium* have been linked malignant cell transformation, enhanced HIV transmission, pneumonia, arthritis and the chronic fatigue syndrome (Namiki et al., 2009, Napierala Mavedzenge and Weiss, 2009, Jensen, 2004, Taylor-Robinson and Jensen, 2011, Anagrius et al., 2005).

Table I.2. Representative properties of the *Mollicutes*. (Dybvig and Voelker, 1996)

Organism	Genome size (kb)	Mol% G + C	Habitat	Phylogenetic group
Acholeplasma				
laidlawii	1600	32	Animals	Acholeplasma
oculi	1600	ND[b]	Animals	Acholeplasma
Anaeroplasma spp.	1600	29–33	Animals	Acholeplasma
Asteroleplasma				
anaerobium	1700	40	Pig	Asteroleplasma
Entomoplasma				
ellychniae	900	28	Insect	Mycoides
Mesoplasma				
florum	1600	26	Plant, insect	Mycoides
lactucae	900	30	Plant?	Mycoides
Mycoplasma				
arthritidis	840	31	Rodent	Hominis
bovis	1100	28–33	Cow	Fermentans
capricolum	1100	25	Goat	Mycoides
fermentans	1160	27	Human	Fermentans
flocculare	1200	ND	Pig	Hyorhinis
gallisepticum	1000	31	Bird	Pneumoniae
genitalium	580	32	Human	Pneumoniae
hominis	700	30	Human	Hominis
hyopneumoniae	1100	28	Pig	Hyorhinis
hyorhinis	820	27	Pig	Hyorhinis
iowae	1300	25	Bird	Pneumoniae
mobile	780	24	Fish	Hyorhinis
mycoides	1300	25	Cow, goat	Mycoides
pneumoniae	800	40	Human	Pneumoniae
pulmonis	950	28	Rodent	Fermentans
salivarium	900	29	Human	Hominis
Spiroplasma				
citri	1400–1800	26	Plants, insects	Spiroplasma
Ureaplasma				
urealyticum	800–900	26	Human	Pneumoniae

Remarkably, the lack of more than one randomized controlled clinical trial (Mena et al., 2009) causes difficulties to find a definitive the treatment of *M. genitalium* infections. Azithromycin cures between 79 and 87% of male or female patients. An better treatment have been described: 5-day azithromycin treatment after a 9-day doxycycline treatment raises the eradication rate up to 96% (Bjornelius et al., 2008).

M. genitalium was first isolated in 1981 in two male patients in the USA (strains G37 and M30) and were systematically characterized in 1983 (Tully et al., 1981, Tully et al., 1983). It has been isolated in urethra, cervix and pharynx. Its high prevalence in sexual workers lays in agreement with a sexual mechanism of transmission and suggests that *M. genitalium* is an emerging pathogen, especially in developing countries (Gomih-Alakija et al., 2014, Vandepitte et al., 2013, Johnston et al., 2012). Although the mechanism of transmission is not completely understood, the capacity of attaching to spermatozoids has been observed (Svenstrup et al., 2003).

M. genitalium isolation and growth is fastidious and it presents very slow growth rates, with doubling time as low as 12 h (Lluch-Senar et al., 2010b). Serology against this organism is not completely understood and although some serological tests have been developed, the detection threshold is not good enough (Ross and Jensen, 2006). So clinics and hospitals have no standard serological tests, and hence, PCR is the most practical and affordable test for its detection (Chalker et al., 2009). Molecular typing—using *mg191* gene as target—strongly supported the sexual transmission of this pathogen.

Many species of mycoplasmas have been found intracellularly (Momynaliev et al., 2000, Sasaki et al., 2002, Meseguer et al., 2003, Tarshis et al., 2004, Yavlovich et al., 2004, Dessi et al., 2005, Vancini and Benchimol, 2008). Specifically, *M. genitalium* have been found in Vero cells (Jensen et al., 1994, Tully et al., 1983) and in human epithelial vaginal and cervical cell lines (McGowin et al., 2009).

2.2.2 Genetic manipulation

The first successful genetic manipulation of *M. genitalium* was performed in 1996 by transposition, demonstrating that random insertion of foreign DNA was possible in *M. genitalium* by means of electroporation (Reddy et al., 1996). Transposons are based mainly in two genetic components: the inverted repeats and the transposase. The transposase is an enzyme capable of recognize de inverted repeat sequences and catalyse the mobility of the DNA within the inverted repeats generally to a random region in the genome (Polard and Chandler, 1995). This transposons were soon modified to be minitransposons (Pour-El et al., 2002). In minitransposon vectors the transposase responsible for the mobility of the foreign DNA to the genome is not included within the inverted repeats, generating in this way stable transposon mutants. This quickly allowed execution of "minimal" genome studies where analysis of the transposon insertions allowed an estimation non-essential genes (Glass et al.,

2006, Hutchison et al., 1999) which later on have been sophisticated through next generation sequencing (Lluch-Senar et al., 2015b).

This tool also allowed the identification of genes implicated in cytadherence and motility since most of insertions disrupt ORFs due to the minimized intergenic regions in this organism (Pich et al., 2006, Pich et al., 2008, Hasselbring et al., 2006a). Transposons and minitransposons have been proven useful for complementation studies which require the reintroduction of the WT allele in mutant strains (Pich et al., 2006). It has also been proven useful for reintroduction of mutated versions of several proteins (Calisto et al., 2012, Martinelli et al., 2015).

Later on, Baseman and colleagues (Dhandayuthapani et al., 1999) showed that targeted genetic modifications were possible through homologous recombination. These targeted genetic modifications allowed exchanging when double crossover events occurred. Replacing genes of interest for antibiotic resistance cassettes have been proved to be a useful methodology to study gene function in *M. genitalium* (Dhandayuthapani et al., 1999, Burgos et al., 2006, Burgos et al., 2008, Lluch-Senar et al., 2010b).

The lack of other genetic tools that are available for other microorganisms since many years ago (counter selection markers, inducible promoters, site specific recombinases, last generation genome modification tools) are not available in *M. genitalium*. Recently, a tetracycline-inducible gene expression system for *mollicutes* have been developed (Breton et al., 2010) but it stills need to be tested in *M. genitalium*.

2.2.3 Morphology and characteristic structures

M. genitalium cells present a characteristic flask shape (Fig. 1) that is conferred by its lack of cell wall and a tip structure called terminal organelle or attachment organelle. This structure is directly involved in adhesion to both solid surfaces and host cells and it is the responsible of its unique motility mechanism: the gilding motility (McBride, 2001). Noteworthy is the presence of a cytoskeleton within the terminal organelle that is revealed as an electrondense core when is observed by TEM (Fig. 1C). This cytoskeleton is a macromolecular complex of proteinaceous nature of an estimated Mw of 300 MDa. This structure even larger than the nuclear complex of eukaryotes, one of the largest macromolecular complexes observed in nature (Seybert et al., 2006).

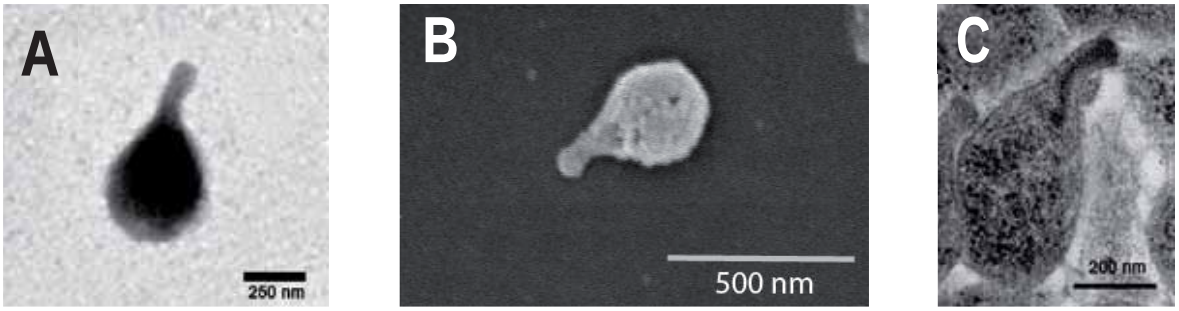


Figure I.1. Morphology of *M. genitalium*. **A.** TEM of positive stained *M. genitalium* cell.**B.** *M. gentalium cell* observed by SEM. **C.** Thin sections of *M. genitalium* cells observed by TEM.

Although all the cytoskeletons of the phylogenetically related motile mycoplasmas are similar there can be observed differences in its size and shape (Hatchel and Balish, 2008). This phenomenon is not factual for all the cytoskeletons in the different motile organisms on *mollicutes* class. In figure 2 striking differences can be observed between the different cytoskeletons described for *Mycoplasma mobile, Mycoplasma pneumoniae* and *Spiroplasma melliferum.*

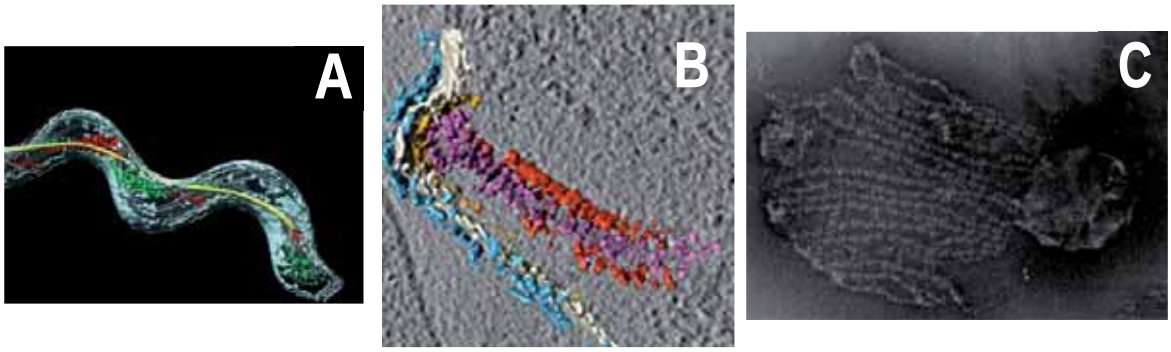


Figure I.2. Different cytoskeletons of motile *mollicutes* **A.** *S. melliferumum*3D reconstruction of its cytoskeleton obtained by CET(Kurner et al., 2005). **B.***M. pneumoniae* cytoskeleton 3D model obtained from CET data. Blue colour: adhesion complexes, white colour: cellular membrane, other colours: cytoskeleton. Notheworthy is the presence of an electronlucent area around the cytoskeleton(Seybert et al., 2006).. **C.** Cytoskeleton of *M. mobile*observed by negative staining TEM after detergent treatment (Triton X-100) (Nakane and Miyata, 2007).

For example, the cytoskeleton of *S. melliferum* is formed by two types of filaments arranged in ribbons. These filaments (red, green and yellow in Fig. I.2A) are parallel but different in length and they explain the helical morphology of the cell. *Mollicutes* with helical morphology are thought to propel themselves by means of coordinated length changes of these structures. Interestingly, the movement generated can be both: dextrorotary or levorotatory.

No data about the ultrastructure of *M. genitalium* is available, but it is thought to be very similar to the *M. pneumoniae* cytoskeleton due to its phylogenetic proximity and high identity degree between the proteins known to form this structure (Table I.1). Data about the ultrastructure and the components of *M. genitalium* also of its closest relative, *M. pneumoniae,* will be presented in next section. The energy source that makes capable this ultrastructure to promote adhesion and gliding have been not determined (Jurkovic et al., 2013).

On the other hand, the fish pathogen *M. mobile* presents a unique jellyfish-like cytoskeleton, which is only clearly visible after treatments with detergents to remove the cell membrane (Fig. I.2C). The ramifications of this structure are suggested to provide support to the *Ieg* proteins, both useful for adhesion to surfaces and generation of movement (Miyata, 2010). Recent results suggest that the energy source of this movement is GTP which is catalysed by one of the proteins forming *M. mobile* cytoskeleton (Jaffe et al., 2004a, Ohtani and Miyata, 2007, Uenoyama and Miyata, 2005).

2.2.4 Ultrastructure and components of the attachment organelle

Little is known about the ultrastructure of the cytoskeleton of *M. genitalium* which was firstly observed in its isolation in 1983. The cytoskeleton charactiric of the pneumoniae cluster was discovered in *M. gallisepticum* (Domermuth et al., 1964) and later characterized in *M.* pneumoniae (Biberfeld and Biberfeld, 1970). Although more than 45 years have passed and we are still far to completely understand how this structure promotes adhesion and gliding motility. As one can see in the micrographs obtained by TEM (Fig. I.1C) *M. genitalium* seems to share, at least, three of major components of the ultrastructure with *M. pneumoniae*. (Fig. I.3A). These components are the nap, the terminal button, the segmented paired plates and the wheel complex. Immunolofluorescent studies (Seto et al., 2001, Kenri et al., 2004, Hasselbring and Krause, 2007b) and translational fusions with fluorescent proteins (YFP, CYF) (Kenri et al., 2004, Balish et al., 2003) revealed that proteins involved in adhesion and motility are localized within this structure (Figure 3B).

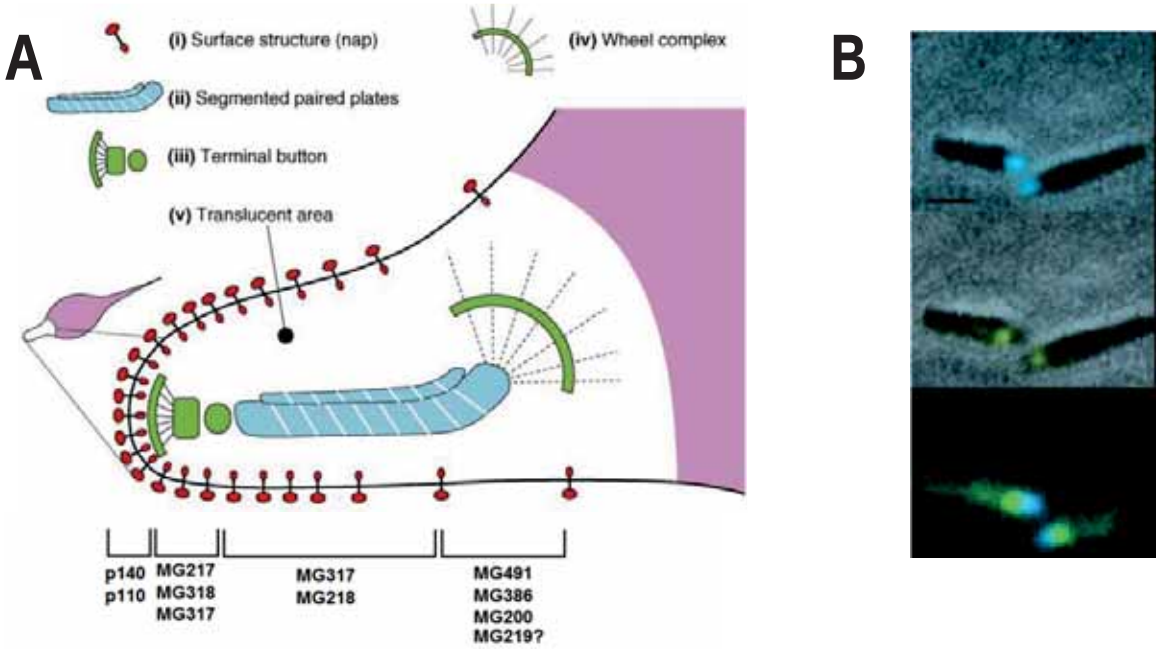


Figure I.3. The terminal organelle of *M. pneumoniae*. **A.** Cartoon representation of *M. pneumoniae* terminal organelle. The cellular body is represented in purple. For clarity reasons, the name of the orthologous proteins in *M. genitalium* is depicted in this image. The image has been modified from (Miyata, 2008). MG219 homology is under the threshold for being considered as the same protein (see Chapter 1). **B.** Example of an experiment on localising P30—blue—and P24—yellow/green—in *M. pneumoniae* (Hasselbring and Krause, 2007b). Scale bar is 2 μm.

Although it is likely that more proteins are present in this complex the nature of the ones depicted in figure I.3A make them to be the most likely to have a specific role in cytoskeleton formation and function. Sequence homology analyses showed that those proteins are only present in some of the mycoplasmas of the pneumoniae cluster (also called the pneumonieae phylogenetic group: *Mycoplasma gallisepticum*, *M. pneumoniae* and *M. genitalium*, *M. imitans*, *M. testudinis*, *M. amphoiforme*, *M. pirum* and *M. alvi*). More

specifically, only *M. genitalium*, *M. pneumoniae* and *M. galisepticum* present homologues of the proteins listed in Table I.3. The second reason is that mutants deficient in this proteins present severe deficiencies both in adhesion and gliding. In Table I.3 a comparison list between these proteins and the relative orthologous in *M. pneumoniae* is shown. The existence of other proteins cannot be discarded.

Table I.3. Equivalence list of ORFs and proteins of the terminal organelle of *M. genitalium* and *M. pneumoniae*.

Loci	Protein	Protein	*Loci*	Identity (%)
mg191	P140	P1	MPN141	45
mg192	P110	P90 and P40*	MPN142	50
mg200	MG200	TopJ	MPN119	35
mg217	MG217	P65	MPN309	42
mg218	HMW2	MG218	MPN310	57
C-ter mg218	MG218-s	HMW2-S	C-ter MPN310	
mg219	MG219	P24?	MPN312?	16
mg312	HMW1	MG312	MPN447	33
mg317	HMW3	MG317	MPN452	34
mg318	P32	P30	MPN453	43
mg386	MG386	P200	MPN567	30
mg491	MG491	P41	MPN311	53

* P90 and P40 are formed by a post-translational cleavage of the polypeptide encoded by MPN142 (Layhschmitt and Herrmann, 1992).

Different approaches in order to obtain structural data about the cytoskeleton have been taken in *M. pneumoniae*. In 1980 Meng and Pfister (Meng and Pfister, 1980) used different compounds (Triton X-100, KCl in a phosphate buffered solution with magnesium) to remove the membrane of attached mycoplasmas aiming to reveal the structure of the cytoskeleton by negative staining (Figure I.4). In 2006 two different groups analysed by cryo-ET the ultrastructure of *M. pneumoniae* macromolecular complexes (Seybert et al., 2006, Henderson and Jensen, 2006). In both cases a 3D model of the cytoskeleton, only including the nap, the terminal button and the segmented paired plates was generated (Figure I.2B and Figure I.4B).

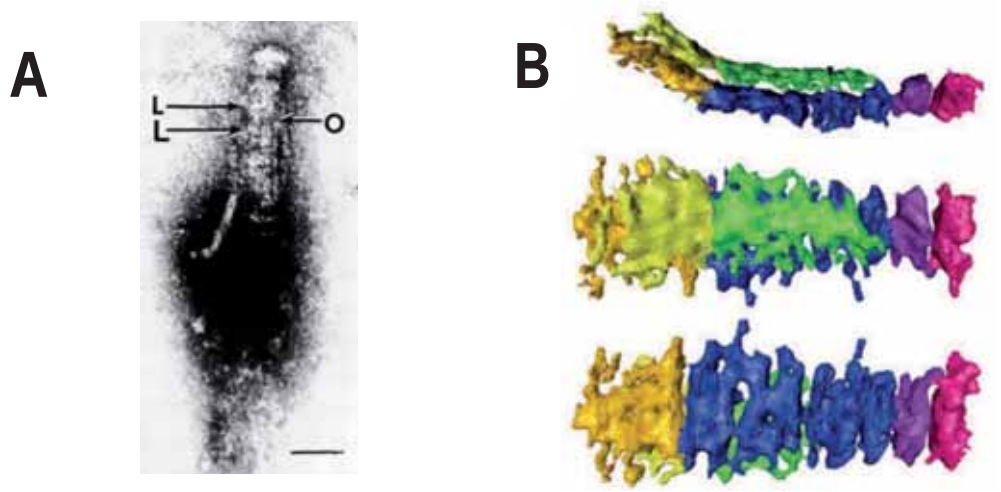


Figure I.4. Structure of the cytoskeleton of *M. pneumoniae*. **A**.Negative stained electrondense cores after membrane removal. L indicates electronlucent areas within the segmented paired plates. O indicates electronopaque areas. These areas probably correlate with the segmentation observable in B. Scale bar 100 nm. **B.** Different views of the Three-dimensional reconstruction of the electrondense core observed by CET. Note that the wheel complex is not included in this reconstruction.

The adhesion complexes are also called nap (Figure I.5) (Tully et al., 1983, Baseman et al., 1982b). The choice to use the term "nap" to name these structures comes from the word *noppe* in German, meaning a grouping of knobs of textile material, which is similar to the surface of the adhesion complexes observed by negative staining. The nap structure is thought to be composed by P1, P40 and P90 in *M. pneumoniae* and is critical for both adhesion and motility (Baseman et al., 1982b, Waldo and Krause, 2006, Seto et al., 2005, Layhschmitt and Herrmann, 1994, Nakane et al., 2011). Various studies link P1 to the major surface protein (Hu et al., 1982, Collier et al., 1983) in the membrane surrounding the terminal organelle (Figure I.5C). P140 and P110, the orthologues of these proteins in *M. genitalium* share the main structural features and localization with the ones of *M. pneumoniae* (Figure I.5B and I.5D). Moreover, this adhesion complex is the immunodominant agent in protein in natural infections occurring in humans (Leith et al., 1983, Svenstrup et al., 2006) Related to this, the coding genes for these proteins contain repeated regions that are also present in different positions along the genome. These regions are called RepMPs in *M. pneumoniae*(Wenzel and Herrmann, 1988) and MgPa islands in *M. genitalium* (Burgos et al., 2006). PP110 and P140 are encoded by MG_192 and MG_191, respectively, and they are also known as mgpC and mgpB. It has been demonstrated that these proteins are phase-variable proteins capable of shutting off their expression when performing single or double crossover homologous recombination events with these repeats. It has also been demonstrated that these genomic rearrangements are used to generate antigenic variation, even in the course of experimental infections (Musatovova et al., 2012, Musatovova et al., 2008, Wood et al., 2013, Totten et al., 2006, Iverson-Cabral et al.,

2006, Iverson-Cabral et al., 2007, Burgos et al., 2006). It is important to note that distinctively of most antigenic variation systems this mechanism is reversible, in other bacteria is usually unidirectional.

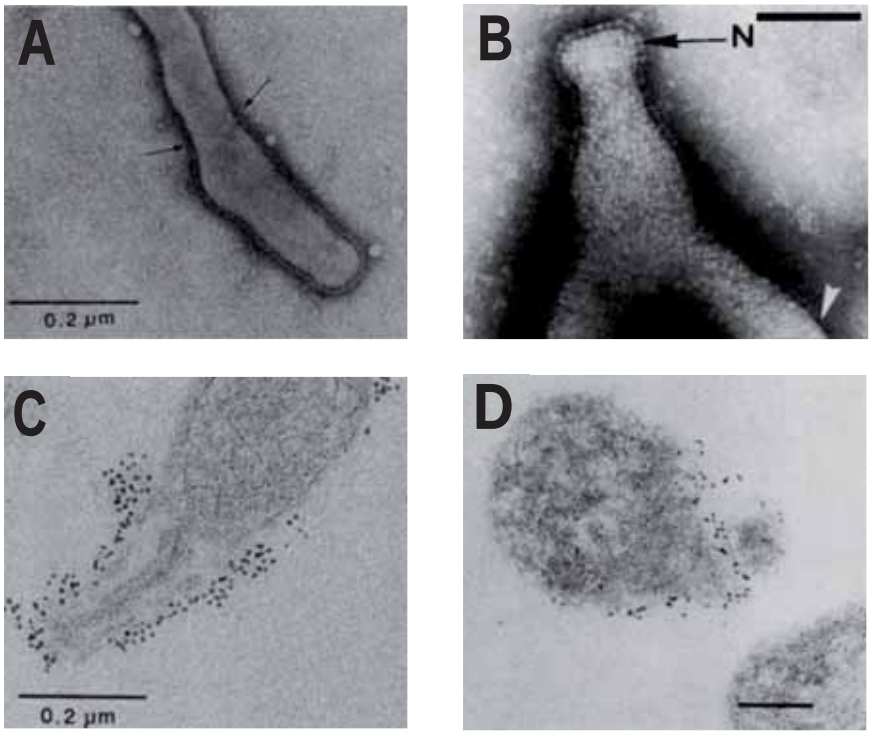


Figure I.5. Micrographs of thinsections of M. genitalium and M. pneumoniae. **A.** Negative staining with ammonium molybdate of *M. pneumoniae* whole cells(Hu et al., 1987). **B.** Negative staining with ammonium molybdate of *M. genitalium* whole cells(Tully et al., 1983, Fedotina, 1978, Henderson and Jensen, 2006).The nap is pointed with a black arrow with the capital letter N**.** Scale bar is 100 nm.**C.** Immunogold labelling localizing P1 with a monoclonal antibody (Hu et al., 1982). **D.** Immunogold labelling localizing of P140 with a monoclonal antibody (Hu et al., 1987). Scale bar is 100 nm.

Different spontaneous mutants lacking P1 protein have been isolated for *M. pneumoniae*. They were first isolated as HA^- mutants and further analysis demonstrated that they lack P1 and also the nap structure (Baseman et al., 1982b). Some of these spontaneous mutants also lack the nap but wild-type levels of P1 adhesin can be detected by western blot, questioning whether P1, in complex with P90 and P40, are the only components of the nap (Figure I.6). Phosphorylation of P1 has been detected experimentally, but it the impact of these phosphorylation in gliding motility or adhesion is not clear (van Noort et al., 2012). P1 proteins is a 170 kDa protein with transmembrane segments, exposed to the outer face of the membrane and responsible for binding to solid surfaces in a mechanism theoretically depending on the presence of sialylated compounds (Chandler et al., 1982, Gesner and Thomas, 1966, Baseman et al., 1982a, Roberts et al., 1989, Kasai et al., 2013). The direct interaction between the nap and those sialylated compounds has been not yet stabilised. Recently, P1 adhesin in complex with P90 have been isolated and characterized from *M.*

pneumoniae cells (Nakane et al., 2011). It has been demonstrated that it forms a tetramer by native polyacrylamide gel electrophoresis composed by two molecules of P1 and two molecules of P90 but not P40 forming a 480 kDa complex. Rotatory shadowing electron microscopy showed that the complex is a sphere of 20 of diameter.

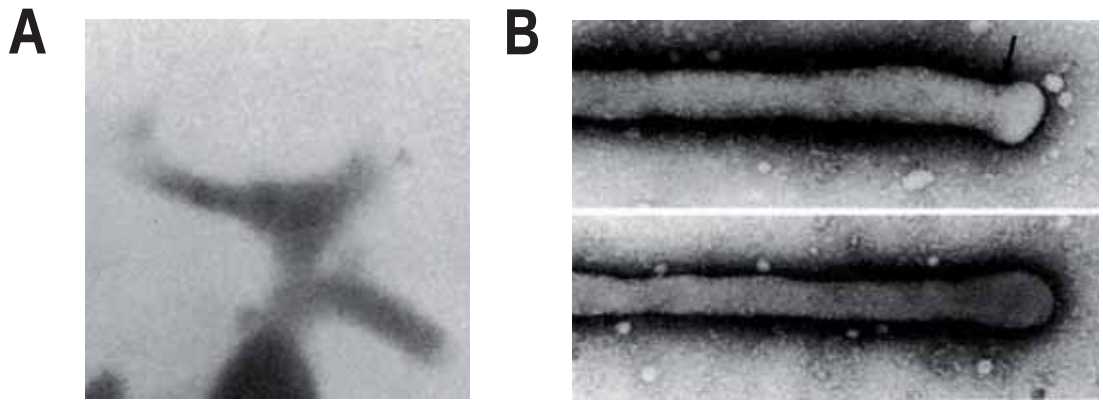


Figure I.6. TEM of HA$^-$ mutants of *M. pneumoniae (Baseman et al., 1982b)*. **A.** Immunoferritin electron microscopy of class IV mutant lacking P1 protein. **B.** Negative staining with ammonium molybdate of *M. pneumoniae* whole cells class III (upper panel) and class I (lower panel) mutants. These mutants have WT levels of P1 but lack the the nap structure.

M. genitalium mutants lacking P110 and P140 have been engineered (Burgos et al., 2006) and showed that this proteins in reciprocally stabilized, fact that also have been reported in *M. pneumoniae(Waldo and Krause, 2006)*. Although cytoskeletons can be observed in some of the cells in these mutants; most of them are rounded and bigger than the WT (Figure I.7A and B). Mutants lacking the adhesion complex are virtually non-adherent and in consequence they fail to perform gliding motility. No structural data of the adhesion complex is available in the literature.

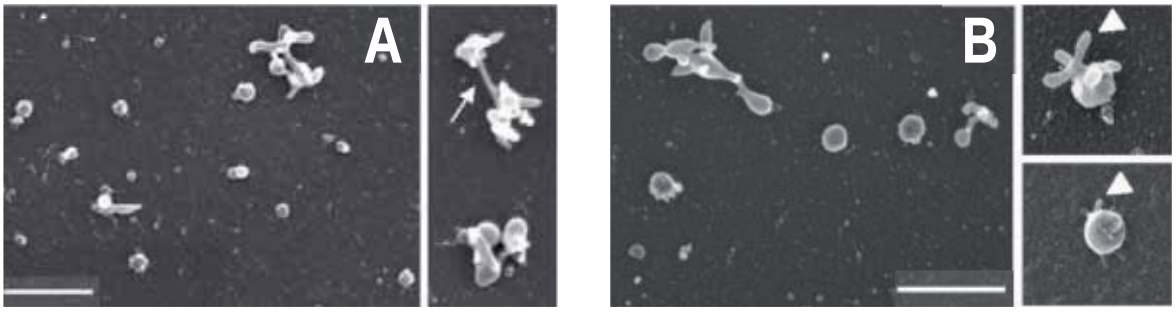


Figure I.7A. SEM of mutant strains of *M. genitalium* presenting abnormalities in cell morphology. **A** mg191 strain. **B.** mg192 strain.Scalebar is 1 µm

Some authors also claim that the nap would be colocalizing in the outer side of the membrane with P30 (Chaudhry et al., 2007, Krause and Balish, 2004, Rottem, 2003), but no actual evidence of that have been found.

The terminal button, the most distal part of the cytoskeleton relative to the cell body, is thought to be composed by P65, HMW3 and P30 (Kenri et al., 2004, Seto et al., 2001). P65 homologue is one of the only *M.* genitalium terminal organelle proteins truly localizedin the literatureby cryo-TEM immunogold localization(Burgos et al., 2008). This study showed that MG217 is directly involved in gliding behaviour but not in gliding frequency or gliding speed. A P65 mutant in *M.* pneumoniae is not available but limited information about the interplay of this protein with other components of the terminal organelle has been reported (Hasselbring et al., 2012, Jordan et al., 2001, Proft et al., 1995). The cells of mutant lacking MG217 protein present an increase in a 40% of the cells drawing an erratic path when observed by microcinemaptographic studies. Interestingly, P65 and MG217 have been both detected as phosphorylated proteins (Su et al., 2007). A role of protein phosphorylation in gliding behaviour can be hypothesized in that way. In addition, a single acetylation has been detected in Lys261 in *M. pneumoniae.*

The *M. genitalium mg317*⁻ transposon mutant lacking the HMW3 orthologue (Pich et al., 2008) showed that this protein in an essential element for the proper morphology of the cells (Fig. I.8A). The individual cells of this strain glide at fivefold reduced speeds and half frequently when compared with the WT. MG317 protein is thought to be located in the terminal button since immunogold experiments performed against *M. pneumoniae* HMW3 protein localized in this position (Stevens and Krause, 1992). HMW3 is suggested to be associated with the cytoplasmic surface of the cell membrane around the cytoskeleton (Krause, 1996). Changes in morphology, abnormal clustering of P1 in the membrane of the TO and atypical cytoskeleton (mainly deficiencies in the terminal button) are characteristic of the transposon mutant isolated in *M. pneumoniae* for HMW3 (Willby and Krause, 2002). HMW3 have been also detected as a phosphorylated protein (Schmidl et al., 2010) and a multiply acetylated protein (van Noort et al., 2012).

Little is known about P32 of *M. genitalium*, it has been annotated as an adhesin in the sequence databases by homology to adhesin P30 of *M. pneumoniae.* Its homology has been also confirmed by orthologous gene replacement (Relich and Balish, 2011). Several studies P30 have been linked to both adhesion and gliding (Hasselbring et al., 2005), for detailed information about this protein see chapter II.

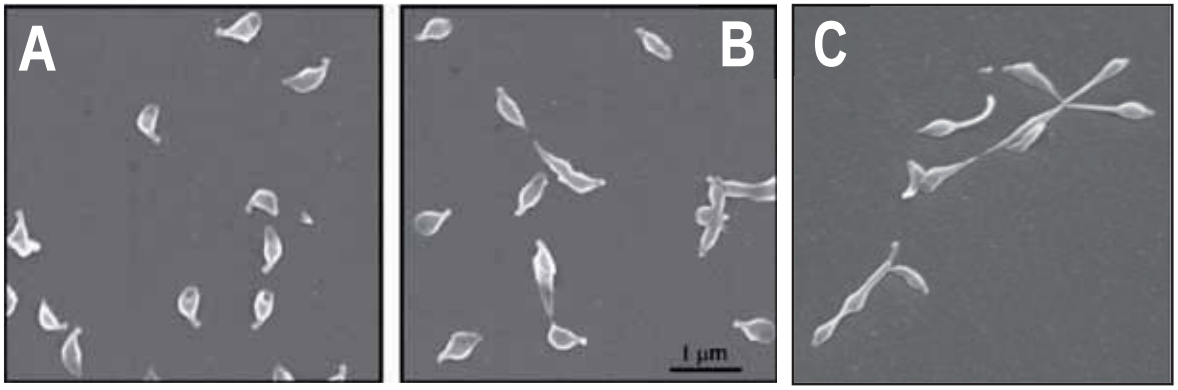


Figure I.8. SEM of different strains of *M. genitalium*. **A**. Wild type G37 strain. **B.** ΔMG_217 mg192 strain.**C.** mg317^{-} strain

The segmented paired plates, sometimes called simply rod, are the central and mayor part of the cytoskeleton and is surrounded by a remarkable electronlucent area in *M.pneumoniae* (Hu et al., 1982) which is not easy visible on *M. genitalium*(Figure I.5B and I.5D, Figure I.1C).

It is known that HMW1 and HMW2 are localised in the position corresponding to the segmented paired plates(Kenri et al., 2004, Seto et al., 2001, Balish et al., 2003) and are required for terminal organelle formation and adhesion trafficking (Seto et al., 2001, Hahn et al., 1998, Popham et al., 1997). Interestingly, HMW1 and HMW2 have been found as phosphorylated proteins (Su et al., 2007, van Noort et al., 2012, Lluch-Senar et al., 2015a, Dirksen et al., 1994, Krebes et al., 1995) and multiply acetylated (van Noort et al., 2012). It is also known that HMW2 encodes a smaller ORF sharing sequence with the C-ter of its coding sequence called HMW2-s or P28(Boonmee et al., 2009, Hasselbring and Krause, 2007b). This small ORF was also found in *M. genitalium* MG218 homolog (Broto A. manuscript in preparation). A null mutant for MG218-s but not for MG218 revealed a specific role of this protein in gliding motility. Cells lacking MG218-s showed a slight increase in gliding motility speed, in erratic motility tracks and in the diameter of the circular tracks (Broto A. manuscript in preparation).

A null mutant lacking MG218 in *M. genitalium* was obtained by gene replacement (Pich et al., 2008). This mutant showed a severe adhesion deficiency plus strong downstream destabilization events in MG386, MG312, P110 and P140 and mild destabilization events in MG491 and MG317. In consequence, the mutant lacking MG218 showed bigger rounded-shaped cells (Figure I.9A). Transposon mutants lying within the coding region of MG218 showed that truncated forms of the protein are stable in the cell, resulting in mutants with severe problems in cell adhesion and gliding motility.

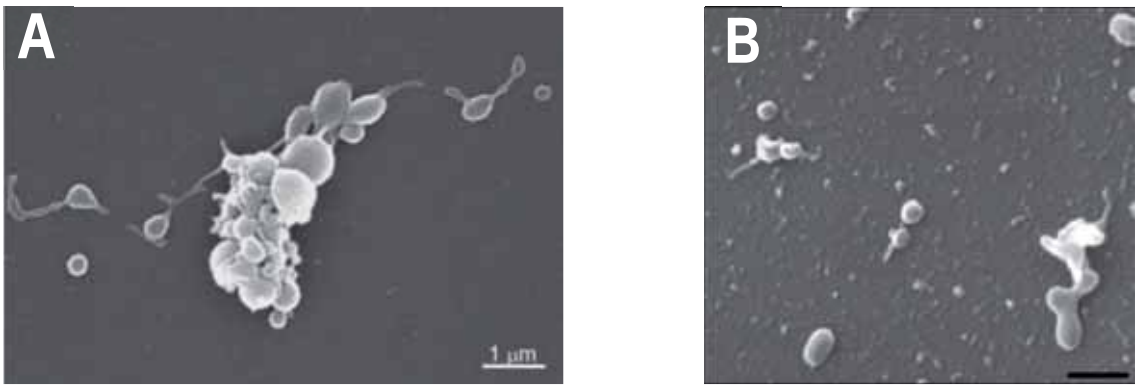


Figure I.9. SEM of mutants lacking the putative major components of the cytoskeleton. **A.** SEM of different strains of *M. genitalium*. **A**. Δmg218 strain. **B.** ΔMG_312 strain. Scale bar is 500 nm

A null mutant lacking MG312 protein in *M. genitalium* was also obtained by gene replacement (Burgos et al., 2007). This mutant showed a severe adhesion deficiency in addition to strong downstream destabilization events in P110 and P140, MG217 and MG317 and a slight protein decrease of MG386 and MG218. In consequence, the cells of the mutant lacking MG312 showed pleomorphic morphology (Figure I.9B). TEM analyses of cells lacking MG312 showed big rounded cells with no visible EDC. Domain functionality analysis of MG312 was carried out introducing engineered alleles without those domains. This analyses revealed that the C-ter of MG312 is the key component for the downstream destabilization events, but adhesion capabilities and motility speed remained near wild-type levels in this mutant but only near half of the cells attached to the surface performed gliding motility. On the other hand, N-terminal domain is critical for achieve wild type levels since cells in this mutant glide at half speed and half frequently, but no affectation in protein levels were detected. The EAGR domain, the central domain and the Walker A box were critical to glide at wild type levels, but the effect is milder than on the other domains or regions analysed. For more information about the different common domains shared between the terminal organelle proteins see Fig. I.10A.

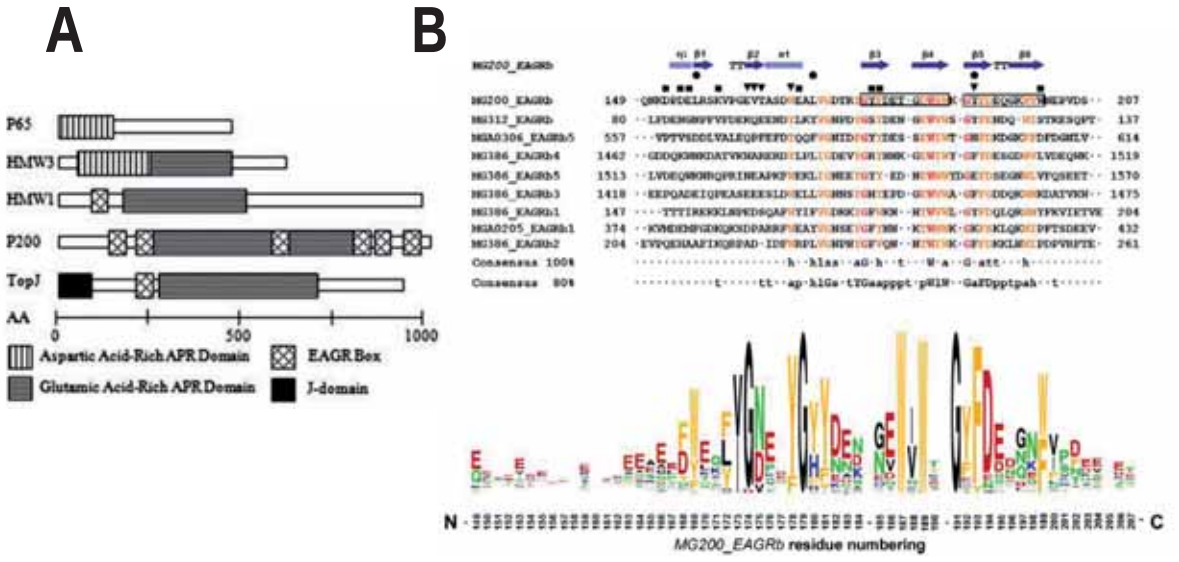


Figure I.10A. Domains presenting different proteins forming the attachment organelle. **A.** Common domains present in different terminal organelle proteins the distribution of domains is conserved in *M. pneumoniae* and *M. genitalium* except in MG386 which does not present the third EAGR box. The EAGR boxes are small domains of enriched aromatic and glycine residues APR equals to acidic and proline rich domains (Balish et al., 2001, Cloward and Krause, 2009b). **B.** Consensus sequence between all the EAGR boxes of *M. genitalium*. Modified from (Wolgemuth et al., 2003, Calisto et al., 2012)

The wheel complex (Hegermann et al., 2002), also called bowl (Henderson and Jensen, 2006) is composed in *M.* pneumoniae at least by P200, P41, P24 and topJ (Kenri et al., 2004, Cloward and Krause, 2009a).

M. pneumoniae P200⁻ transposon mutant showed remarkable difficulties to adhere to NHBE cells (normal human bronchial epithelial) in air-liquid interface culture, but not to lung adenocarcinoma cells A549 or erythrocytes. The cells of this mutant also showed near half gliding speed and no-motility periods twice longer during the micro-cinematographyc acquisition times when they compared to the wild type strain. TEM of cell of P200 transposon mutant strain showed no major differences in terminal organelle ultrastructure. Similarly, several MG386⁻ transposon mutants were previously described in *M.genitalium*[1]. The transposon mutant isolated with the largest part of the ORF truncated showed a majority of the cells motionless (around 80%) and from those moving 77% presented circular movements and 23% erratic movements both gliding at 29% of the gliding speed of the wild type strain. Other transposon insertions present stable larger fragments of P200 with milder affectations in gliding motility as expected. The mutants of this strain show flask-shape morphology as the wild-type (Pich et al., 2009). The domain containing the 3, 4 and 5 EAGR boxes from MG386 have been found to strongly interact with MG219 protein but now with

[1] More information about MG386 protein can be found in chapter III.

the N-terminal, suggesting a specific interaction between the C-ter of MG386 and the C-ter of MG219(Martinelli, 2014).

P41 was located in 2004 as well as the entire cytadherence regulatory locus—P65, HMW2, P41 and P24 in a location compatible with the wheel complex(Kenri et al., 2004) and later colocalized with P30 (Hasselbring and Krause, 2007b). A transposon mutant of P41 in *M. pneumoniae* revealed a striking phenotype; occasionally, the terminal organelle detaches from the cell body and glides independently in cells lacking MG491 (Hasselbring and Krause, 2007a). Moreover, nascent terminal organelles from cells lacking P41 acquire gliding capabilities quicker than the WT strain. All these data suggest that 41 is a key protein attaching the terminal organelle to the cell body. Recently, a null mutant lacking MG491 have been developed (Garcia-Morales, L, manuscript in preparation). This mutant presents a big contrast in the phenotype when it compared with the mutant of P41. In addition to hemadsorption negative phenotype, a massive downstream destabilization of all the known proteins of the terminal organelle was also observed. After the incubation of high amounts of biomass over glass coverslips large pleomorphic aggregates can be observed on the surface. Strikingly, protruding thin filaments extending from these aggregates can be observed suggesting that gliding activity is still active in those cells (Figure I.11A). Micro-cinematographic studies confirmed that notion; those filaments glided at speeds as low as 0.06 nm s^{-1}—0.05% of the wild-type velocity. As low as this velocity is, these observations required several hours to be detected. In addition, MG491 *M. genitalium* is phosphorylated (Su et al., 2007) suggesting that MG491 can be involved in control of gliding or terminal organelle duplication depending on its phosphorylation status.

More in-detail studies such as partial deletions or amino acid substitutions were performed to gain insight into MG491 function. Mutants lacking the N-terminal showed a very similar phenotype for the null mutant MG491, indicating that the N-terminal domain of MG491 is the main responsible for the downstream destabilization effects. After solving the crystal structure of the central domain of MG491 (aa from Asn67 to Ala203) at 3 Å different amino acid specific mutations were designed aiming to destabilize the quaternary structure of the protein (Garcia-Morales, L, manuscript in preparation) (Figure I.12A). Interestingly, mutants presenting the mutation C87S glided at frequencies fourfold lower than the WT and at 80% of the wild type gliding speed. Moreover, mutants replacing two Phe for Ala in the central domain resulted in the apparition of minute cells. These results indicate that MG491 oligomerization is important to maintain the proper functioning of MG491 protein, and suggest that each of the functions assigned to MG491 lye within different regions of the protein. In addition, interaction studies revealed that a small peptide of 25 aa near the C-ter

of MG491 (called peptide 1) and the EAGR domain of MG200 weakly interact (Martinelli et al., 2015). A mutant bearing a mutant allele lacking the peptide 1 but no the WT allele revealed that this peptide is involved in gliding behaviour and attachment of the terminal organelle to the cell body. The diameter of the tracks depicted by gliding motility was approximately 20% smaller from those of the WT strain. Besides the effects in gliding behaviour, a dramatic increase of cells showing a bipartite morphology and minute cells (20% in both cases) is also observed. It is likely that this minute cells were originated by asymmetric cell divisions (Figure I.11B).The ratio of motile cells or mean velocity was not affected by this mutation.

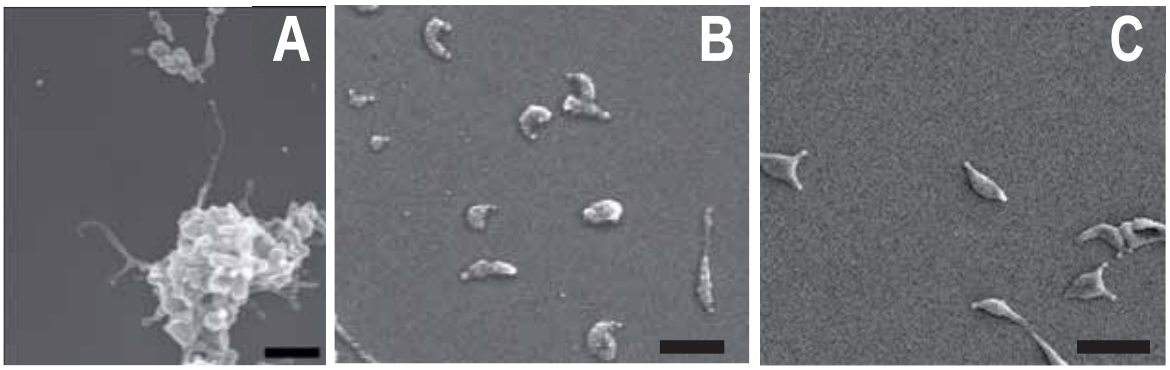


Figure I.11 SEM of mutants presenting mutations in the proteins of the wheel complex. **A.** Scanning electron microscopy of several mutants with alterations in proteins of the wheel complex of *M. genitalium.* **A.** ΔMG491 strain (Garcia-Morales, L, manuscript in preparation). Scale bar is 1 μm. **B.** Δp1c1. Scale bar is 1 μm (Martinelli et al., 2015) Minute cells, asymmetric cell divisions and bipartite cells can be observed. Scale bar is 1 μm **C.** MG200^{-} c12.

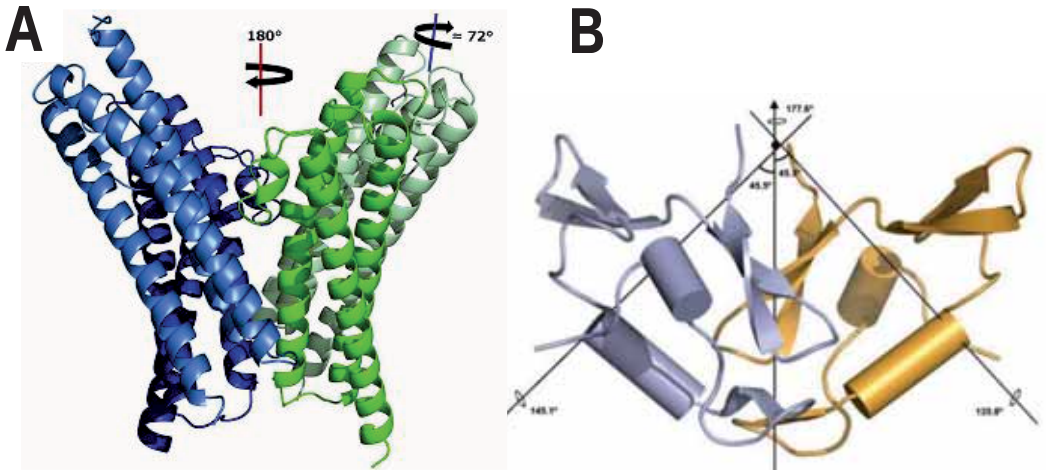


Figure I.12. Crystal structures of domains of terminal organelle proteins determined to date. **A.** MG491 central domain (Garcia-Morales, L, manuscript in preparation). Three alpha helixes can be observed interacting in a tetrameric form. **B.** Dimeric organization of the crystal structure of EAGR box of MG200. Two alpha helices and 6 beta sheets can be observed.

TopJ protein (meaning terminal organelle protein with a J-domain), have been related with co-chaperone activity necessary for terminal organelle maturation and function.(Cloward and Krause, 2009b, Cloward and Krause, 2011, Cloward and Krause, 2010). Although co-chaperone function was not clearly demonstrated it has been suggested by the absence of P24 (Cloward and Krause, 2009a) in a transposon mutant lacking TopJ. The co-chaperone function is reinforced by the view that J-domain itself acts as a co-chaperone in several proteins(Walsh et al., 2004, Hennessy et al., 2005, Mayer and Bukau, 2005, Bukau and Horwich, 1998, Mayer et al., 2000, Schmid et al., 1994). A transposon mutant lacking TopJ revealed that this protein is critical for adhesion to both erythrocytes and adenocarcinoma cells. In addition, gliding motility was not detected in this mutant. Immunofluorescence studies and translational fusions with fluorescent proteins revealed more impaired foci or paired non-polar foci of cytadherence related proteins (P41, P30, P41 and HMW1) than in the WT strain (Cloward and Krause, 2009a) suggesting that terminal organelle formation and maturation could be affected in this mutant. This data should be treated carefully since further studies are needed to clarify that (Cloward and Krause, 2010).Further characterization of topJ functionality was carried out by domain functionality analyses (Cloward and Krause, 2010). These experiments revealed that the HPD region within the J-domain is critical for cell adhesion but neither the C-ter nor the APR domains are necessary to the cells of *M. pneumoniae* to cytadhere. These analyses revealed that gliding motility is mildly affected in mutants lacking C-terminal domain or the APR region (60%,30% of the WT respectively) or completely absentin the HPD region within the J-domain. Remarkably, topJ has been found to be phosphorylated and acetylated in several positions (van Noort et al., 2012, Lluch-Senar et al., 2015a).

Despite the extensive studies carried out about topJ protein, the first evidence of the implication of the protein MG200 in *M. genitalium* was in 2006 (Pich et al., 2006). Two transposon insertions were obtained in MG200, which keep the entire J-domain intact. The mutants lacking the APR domain and the C-ter domain in *M. genitalium* presented only a 5% of motile cells and glided at a 10% of the WT gliding velocity. Cellular morphology is not directly affected in this mutant but the strong effect in motility is reflected in cell division (see 2.2.6. Cell division). An increase approximately in the 20% of the individual cells presenting more than one terminal organelle can be observed in this mutant when it compared to the WT (Figure I.11C) and it is also noteworthy the high prevalence of large aggregates on the sample. In order to gain further insight into the function of MG200 domain mutant analysis and chimeric studies have been performed. A mutant lacking the EAGR box domain of MG200 revealed that this domain is critical of cell gliding motility, the effect of lacking the EAGR box is not as severe as lacking the C-ter of the protein. Cells glided at half

the velocity of the WT and half frequently. Also the restring periods (cells that present no movement at least during 5 seconds during the 2 minutes of acquisition time) increased in a 25%(Calisto et al., 2012).The functional study of the MG200 EAGR box was nicely accompanied with the determination of the crystal structure of MG200 (Figure I.12B). Later on, an antibody generated against the J-domain of MG200 detected that this domain is independent of the C-terminal fully stable in the mutant obtained by transposition. Moreover, J-domain was shown to be required for P140 and P110 folding, stability or trafficking since only mutants lacking these proteins resulted in viable cells (Broto, A, manuscript in preparation). The nature or the lack of adhesins in this mutant proved to be due genome rearrangements between the MgPa operon and the MgPa islands. More in detail, the glycine and phenylalanine region (G/F) just at the C-ter region of the J-domain of MG200 was shown to be critical for gliding motility. Further experiments revealed that mutants bearing chimeric versions of MG200 shuffling the J-domain andthe G/F region of different J-domain containing proteins showed that any J domain seems capable of perform the function required for adhesion and gliding encoded in MG_200. In contrast the G/F region of MG200 is the only capable to maintain P140 and P110 stability and gliding motility at WT levels. Furthermore, if the J domain and the G/F domain of MG200 are not present in the same polypeptide the cells are still able to adhere but not to glide at WT levels suggesting that the J-domain and the G/F region need to act in *cis* to perform its intended function. The C-ter of MG200 have been found to strongly interact with the N-terminal of MG200 (Martinelli, 2014).

P24 have been located in the terminal organelle by different studies (Hasselbring and Krause, 2007b, Kenri et al., 2004). Analyses of mutant lacking P24 showed a decrease of a 60% in gliding frequency and the appearance of a 30% more of terminal organelles than in the WT (Hasselbring and Krause, 2007b). The fact that MPN312 have been found phosphorylated (van Noort et al., 2012, Lluch-Senar et al., 2015a), despite being a low molecular weight protein, suggest that its regulatory function in terminal organelle duplication can be controlled by protein kinases/phosphatases. MG219 protein, the putative homolog of MPN312 in *M. genitalium* has been not characterized[2]. N-terminal of MG219 have been shown to be strongly interacting *in vitro* with the C-ter of MG200 (Martinelli, 2014).

[2] Detailed information about MG219 can be found in chapter I.

2.2.5 Adherence and motility mechanisms

Adherence and motility have been widely linked to pathogenesis processes (REF) and it is not different for mycoplasmas (Baseman and Tully, 1997, Razin and Jacobs, 1992). The majority of the mycoplasmas found to adhere to host cells are parasites and/or pathogens and it is thought that adhesion is an initial an essential step for colonization and disease (Chandler et al., 1984). A direct recognition of the bacterial adhesins to carbohydrate compounds on the host membrane (Razin and Jacobs, 1992, Roberts et al., 1989) is expected to be the prior step to cell internalization when applicable. The discovery of the capacity to mycoplasmas to glass and plastic surfaces (Somerson et al., 1967, Taylor-Robinson and Manchee, 1967) allowed the *in vitro* analysis of gliding capabilities of different strains under great panoply of conditions.

Specifically, systematic studies of different sialylated compounds have been performed in *M. mobile* and *M. pneumoniae (Kasai et al., 2013).*The most effective sialylated compounds were Neu5Ac-α-2,6-galactoseand Neu5Ac-α-2,3-lactosamine respectively. These results showed that the receptors of both species catch and release the sialylated compounds in a cooperative fashion and they can generate a drag force after the stroke since gliding velocity decreased when adding sialylated compounds to the examination buffer. It is expected that P1 complex is responsible to the recognition of this sialylated compound. In the same way, P140 and P110 are expected to be the responsible to recognizing syalilated compounds in *M. genitalium* but since the epithelium of the respiratory tract and the urogenital tract present high differences in oligosaccharide composition (Varki, 2009) it is expected that other sialylated compounds have higher affinity to these adhesins. Similarly, monoclonal antibodies against P1 protein have been found to stop gliding motility prior to detachment of *M. pneumoniae* reinforcing the view that P1 acts as a "leg" protein (Seto et al., 2005).

The mechanism how the motility occurs in mycoplasmas is far to be completely understood. No flagella or pili have been observed in this genus and a lot of them are not motile. From more that 200 of mycoplasma detected species 14 of them have been found to perform gliding motility: *M. pneumoniae, M. insons, M. genitalium, M. gallisepticum, M. testudinis, M. amphoriforme, M. pirum, M. imitans, M. pentrans, M. testudineum, M. agassizii, M. pulmonis* y *M. mobile.* Interestingly, these species present a cytoskeletonsuggesting a direct relationship between cytoskeleton and motility (Hatchel and Balish, 2008). This study also showed that there is no direct correlation between cytoskeleton morphology, phylogenetic distance and gliding velocity.

There are two models proposed for mycoplasma gliding; the inchworm model and the centipede model. The inchworm model is based in the theory that changes in the length rod by curving causes repeated dissociation, displacement and associations of adhesins to the solid surface. These repeated contractions will lead to a net movement of the cell. This model is supported by cryo-TEM data where different curvatures within the segmented paired plates were observed (Figure I.13).

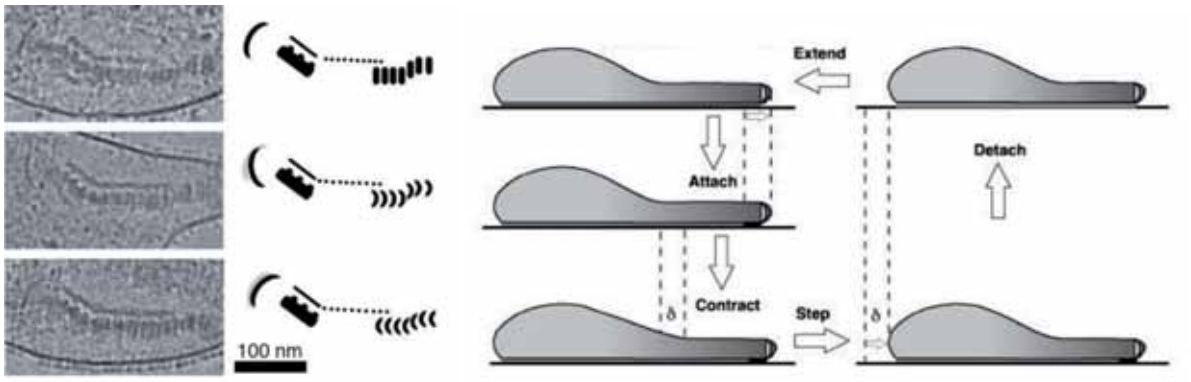
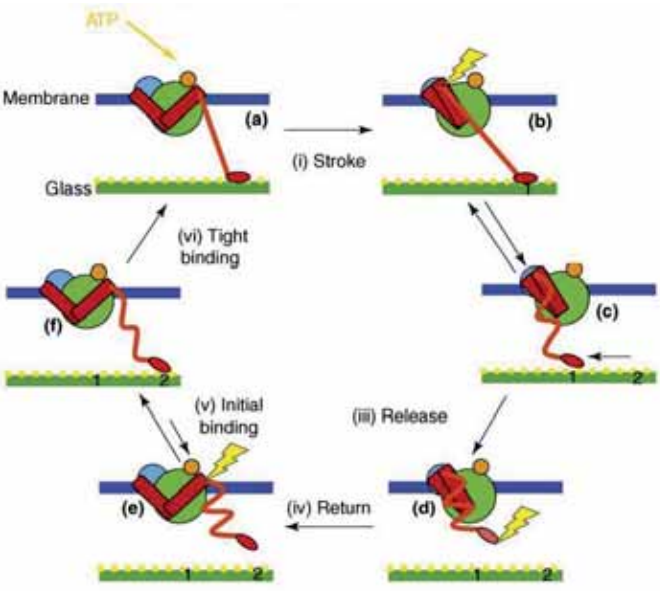


Figure I.13. Inchorm modelshypothesized for gliding motility(Henderson and Jensen, 2006) (Wolgemuth et al., 2003).

The centipede or power stroke model relies on the view that adhesins repeatedly catch and release from the surface repeating the following steps: stroke, movement, release, return, initial binding and tight binding (Figure I.14.). In this model cells will propel with small unitary steps. This model is supported by the fact that P1 antibodies against the adhesion complex of *M. pneumoniae* brake the cells prior to detachment of gliding surface.



None of these approaches have been properly studied and further experimental is needed for demonstration.

Figure I.14. Centipede model for gliding motility. Modified from (Miyata, 2008).

In addition, the fact that both proteins present in the segmented paired plates have this PTM suggests an important role of phosphorylation and acetylation for motility and/or terminal organelle formation. One of these views is reinforced by data suggesting that mutations in the kinase prkC and the phosphatase prpC alters gliding speed and frequency in *M. pneumoniae* but do not show strong impact in hemadsorption (Page and Krause, 2013). The conclusions derived from this work suggest that excess of phosphorylation in specific proteins causes *M. pneumoniae* cells to glide almost at twice the speed and trice frequently. The phosphorylation specific targets causing these alterations in gliding are not known.

2.2.6 Cell division

It is known that cell division is done by binary fission in mycoplasmas. In gliding mycoplasmas it has been found that cell motility is involved in cell division (Freundt, 1969). It is also known that the division of the terminal organelle occurs always prior to the initial stages of cell division, and it is thought to be performed in a semiconservative way (Nakane and Miyata, 2009, Hegermann et al., 2002) (Figure I.15A). In this model, the terminal button would be the first to be duplicated, followed by the segmented paired plates and finished by the duplication of the wheel complex. Experimental data suggest that the new terminal organelles are not motile, causing cells to be anchored to a fixed position in the gliding surface although the gliding apparatus of the pre-existent TO is still completely functional. This phenomenon causes the two terminal organelles to end in the opposite position of the cell. Later, an undetermined signal activates gliding motility on the new TO and both cells start gliding in opposite directions (Figure I.15B). The cell division concludes when there is not enough membrane to maintain both cells united which seem to be aided by ftsZ (Lluch-Senar et al., 2010b). (Balish, 2014, Bredt, 1968, Rosengarten and Kirchhoff, 1989). This model is supported by the evidence that cells prior to cytokinesis present higher amounts of DNA than individual cells, increasing in a time-dependent manner from the start of the TO duplication (Seto et al., 2001). It is also suggested that the TO could be responsible for chromosome segregation, but the missing link between the terminal organelle and the bacterial chromosome is yet to be found. In agreement with this model, mutants deficient in gliding motility present cells with multiple terminal organelles in higher proportions than in the WT(Hasselbring et al., 2006b, Pich et al., 2009) (Figure I.11C). It is important to note that in *M. pneumoniae*, the frequency of multiple TO cells is higher than in *M. genitalium*. In addition it has been found that low levels of P140 and P110 in *M. genitalium* causes important defects in cell division. Cells with diminished amounts of P110 and P140 presented 50% more cells with multiple terminal organelles. Although half frequently gliding, the mean gliding speed was not affected in the cells of this mutant suggesting a

specific role of the adhesion complex in terminal organelle duplication. Another possible explanation is that cells with one TO (those capable to glide) present wild-type levels of P110 and P140 which would explain the WT gliding speeds. In that scenario the cells with multiple TO will have very low amounts of adhesins, gliding poorly and unable to develop normal cell cycles. More experiments should be performed to clarify that.

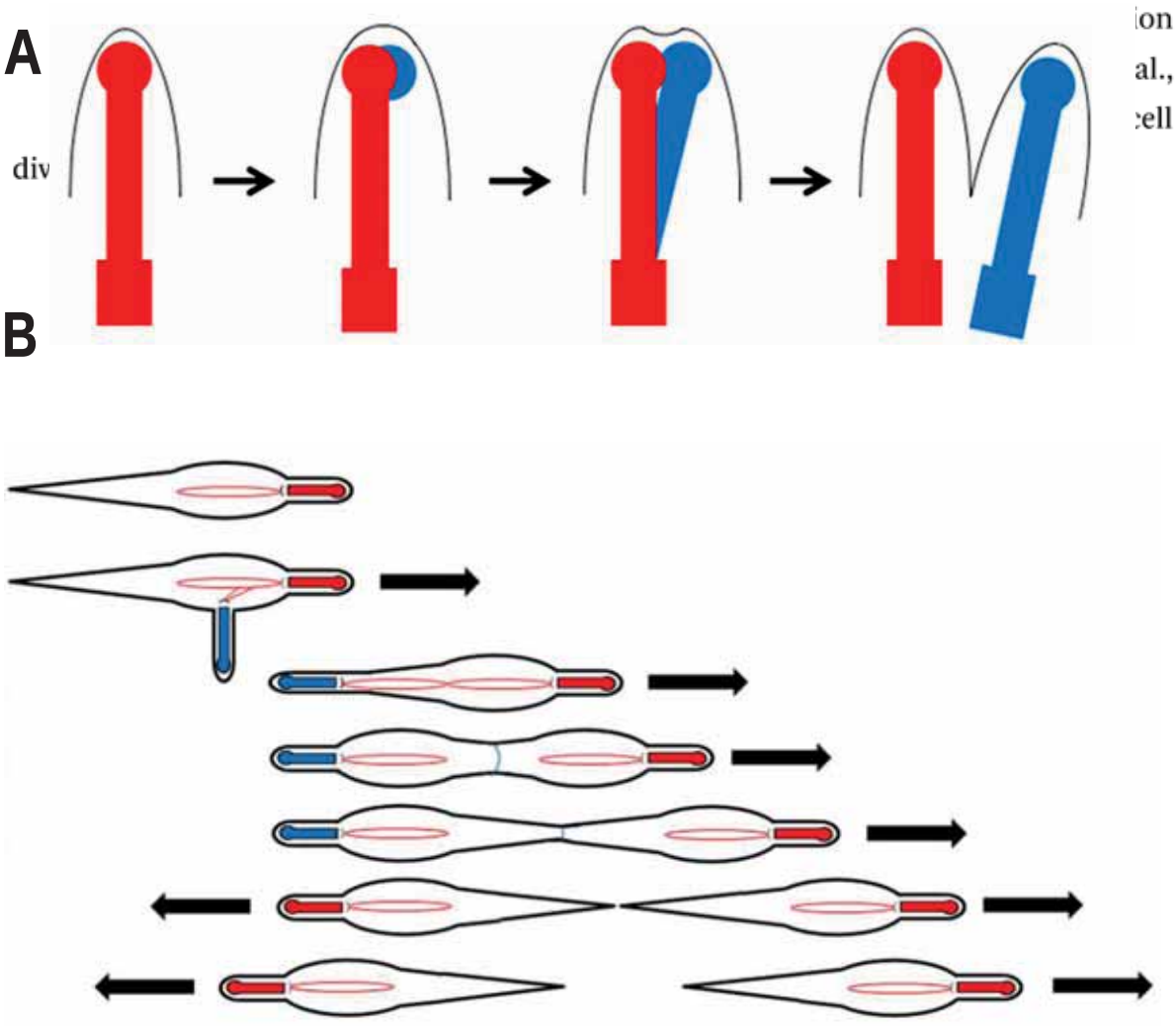


Figure I.15. Cell division in mycoplasmas with terminal organelle and gliding motility. Modified from (Balish, 2014). **A.** Model of semiconservative division of the TO in *M. genitalium* and its relatives. In red the initial core: motile TO. In blue the new non-gliding TO in development **B.** The cell division process in mycoplasmas with TO.

3. Objectives

Chapter I

Functional analysis of a unique protein of Mycoplasma genitalium: MG219

- Recombinant expression of MG219 in *Escherichia coli* aiming to obtain polyclonal antibodies against MG219.

- Obtaining a null mutant for MG219 coding gene in *M. genitalium* in order to study its possible implication in gliding motility.

- Characterization of the null mutant for MG219 coding gene by means of micro-cinematography and microscopy.

- Complementation studies to demonstrate that both the WT allele and Translational Fusions to other proteins or tags are capable to restore the wild type phenotype.

- Obtaining a mutant that has replaced MG219 for MPN312 (the putative homolog *M. pneumoniae*) to gain insight in the relationship between these two proteins

- Subcellular localization of MG219 protein as well as MG318.

Chapter II

Functional analysis of the coding gene for P32 protein: MG318

- Recombinant expression of MG318 in *Escherichia coli* to obtain polyclonal antibodies against MG318.

- Obtaining a null mutant for the entire P32 coding region and also a mutant lacking the C-ter of P32 of *M. genitalium*.

- Study of the implication of the P32 in gliding motility by electron microscopy and micro-cinematography.

- Study of the functionality of P32 N-ter and comparison with the well characterized mutant defective in P30 of *M. pneumoniae*.

Chapter III

Functional analysis of high molecular weight protein MG386.

- Obtaining a null mutant for MG386 coding gene of *M. genitalium*.

- Study of the effects of lacking MG386 in gliding motility and cell morphology.

Chapter IV

Structural characterization of macromolecular complexes of *Mycoplasma genitalium* by electron microscopy

- Use cryo-electron tomography to obtain ultrastructural information of *M. genitalium* whole cells.

- Analyse the structure of the cytoskeleton and other macromolecular complexes from the wild-type strain.

- Use cryo-electron tomography analyse the mutants previously obtained for proteins related with gliding motility and terminal organelle formation.

4. Chapter I: Functional analysis of a unique protein of *Mycoplasma genitalium*: MG219

4.1 Introduction:

4.1.1 Previous knowledge of MG219

Little is known about the MG219 protein of *M. genitalium*. In fact, its sequence is so distinctive to any other in the sequence databases that it has been used to develop a specific PCR to detect *M. genitalium* in clinical samples (Chalker et al., 2009). Interestingly, when analysing the cytadherence regulatory locus of *M. genitalum* and *M. pneumoniae* it is clear that MG_219 and MPN312 loci are related by synteny and size (Fig. 1.1A).

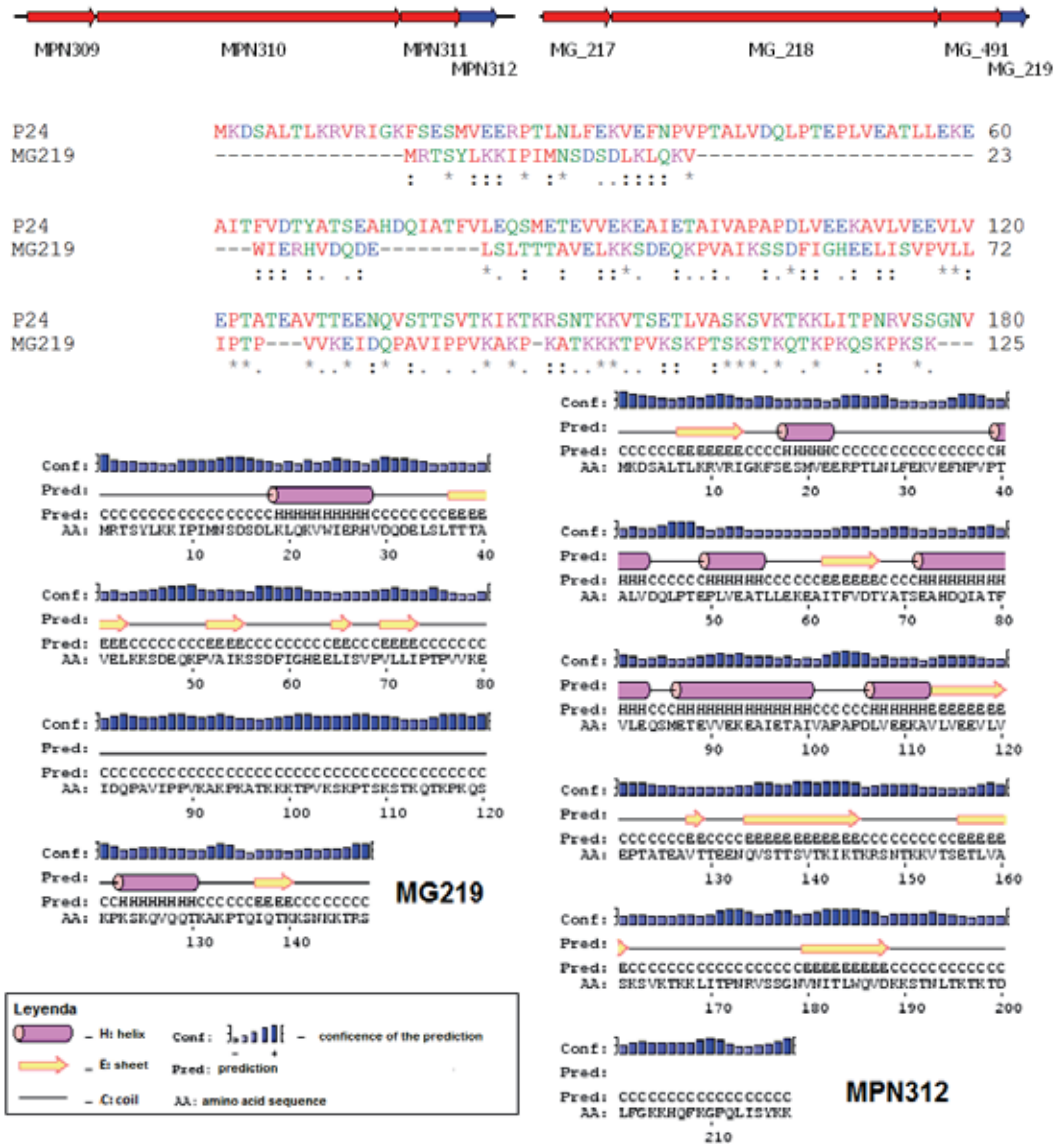


Figure 1.1. Comparison between MG219 and MPN312. **A.** Operon comparison n of the *crl* of *M. genitalium* and *M. pneumoniae.* **B.** Sequence comparison at amino acid level. **C.** Secondary structure prediction of MG219 and MPN312.

Besides this apparent similarity, amino acid sequence comparison showed only 16% of identity, which is significantly lower than the threshold to consider homology between two proteins (Fig. 1.1B). According to these data, secondary structure predictions show a big discrepancy (Fig. 1.1C). Hence, MG219 and MPN312 cannot be considered orthologues only by sequence comparison, but experimental derived data could help to solve this enigma.

Interestingly, NMR studies about the recombinantly expressed MG219 protein in this chapter showed that MG219 is an intrinsically disordered protein. The only secondary structure element is a short alpha helix from residue 175 to 180 (Calisto et al, manuscript in preparation).

4.2 Results

4.2.1 Recombinant expression of MG219 and polyclonal antibody generation

The entire coding region of MG219 protein was cloned in a pET-21a (from aa 1 until 148) with a histidine tag in the carboxyl-terminus of the protein. Both protein expression and purification yields were analysed at 16 ºC and 37 ºC (Fig. 1.2). All tested temperatures showed that recombinant MG219 were mostly soluble but higher yields were detected 37°C.

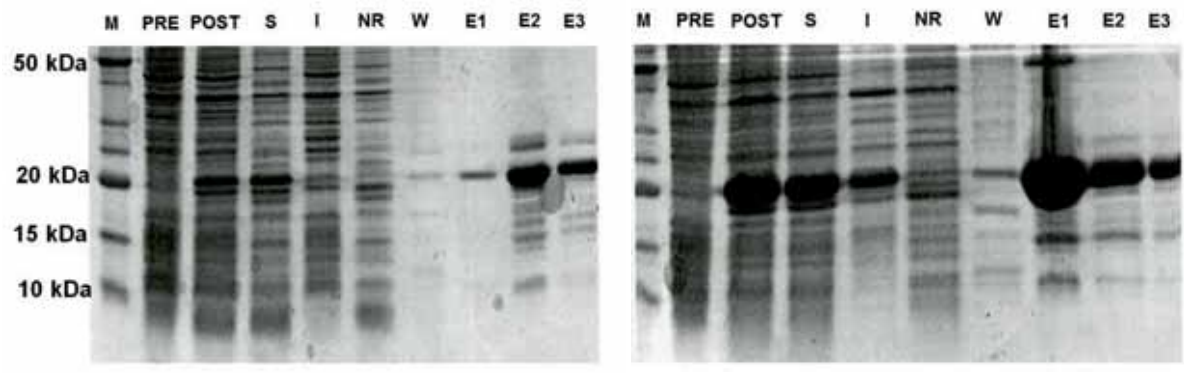


Figure 1.2. SDS-PAGE of recombinant expression and purification of M219 in *E.coli* BL21(DE3) cell. Left panel: Expression at 16ºC. Right panel**:** Expression at 37ºC. M: molecular weight marker. Benchmark ladder (Life technologies), PRE: sample of the culture before induction. POST: sample of the culture after induction S: soluble fraction after cell lysis. I. Insoluble fraction after cell lysis. NR: non-retained sample in the HisTrap HP (GE) column. W: washed low-affinity proteins with 50 mM imidazole. E1, E2, E3: different elution fractions at 500 mM of imidazole. Fraction size is 0.5 ml.

After dialysis in PBS the recombinant MG219 protein was intraperitoneally injected in BALB/c mice. Once the immunisation protocol was finished the serum was tested to be positive against MG219 protein (see section 4.2.4).

4.2.2 MG_219 transcription start site

Since the transcription start site of MG_219 has been not stablished experimentally, a primer extension experiment was performed. This experiment depends on the retrotranscription activity of a retrotranscriptase using a 5' 6-Carboxylfuorescein DNA oligomer priming near the 5' of the MG_219 coding RNA. The analysis of the amplicons resulting from this retrotranscription revealed three putative TTS (Fig. 1.3). Those TSS match to bases 265616, 265617 and 26518 of GenBank L43967.2. Heterogeneous transcriptional start sites have been previously found in mycoplasmas (Musatovova et al., 2003, Musatovova et al., 2006, Pich et al., 2006), reinforcing the view that these putative TSS are genuine. The previously mentioned putative TSSs are within the coding region of MG219 (between amino acid K7 and K8). At first glance this result could seem impossible or unreliable, but it is feasible that MG_219 suffers from translational start site improper assignation in the reference genome. This possibility is further explored in following sections.

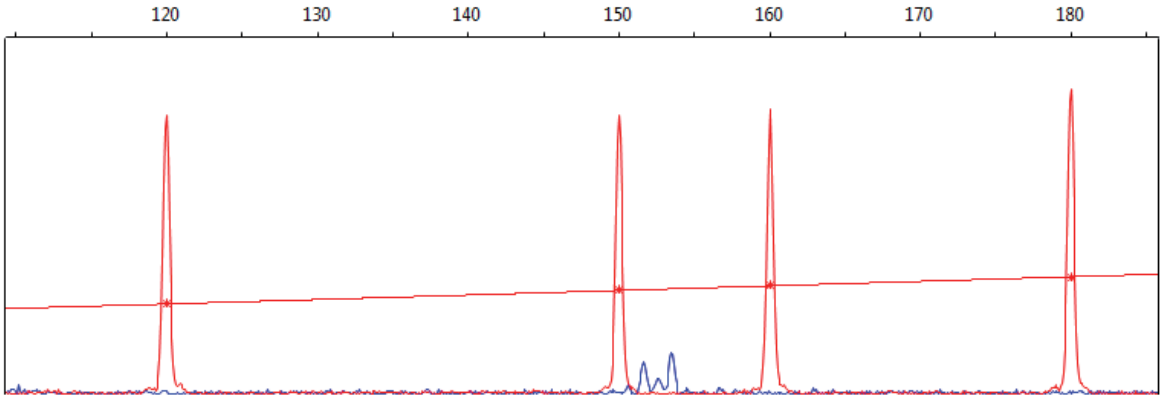


Figure 1.3. Primer extension analysis of MG_219 in WT strain. Red peaks correspond to Genescan 400HD ROX standard while blue peaks correspond to the primer extension products. 151, 152 and 153 bp blue peaks lead to the identification of the putative transcription initiation sites of MG_219.

4.2.3 Obtaining a null mutant for MG219: Δmg219 strain

To gain insight in the function of MG219, a mutant lacking this protein is conceived as a platform for phenotypic studies. This null mutant for MG219 was obtained by homologous recombination using the plasmid pΔmg219 (Fig. 1.4) and the generated strain was called Δmg219. The construct was designed to have flanking the TetM438 gene by 1021 bp immediately downstream from the stop codon of the MG219 coding sequence and 1000 pb upstream of MG219 coding sequence but keeping 41 bp of the 5' of this gene. This 41 bp from the 5' of MG_219 were aimed to be kept in Δmg219 strain because the 3' of MG_491 and 5' of MG_219 overlap 41 bp (see Fig. 1.1). This means that the first hypothetical 13 amino acids would be present in Δmg219 strain, but it is expected that this small fraction of the protein would have no impact in the phenotype of this strain. pΔmg219 was electroporated in *M. genitalium* G37 strain, and single cell colony forming units were isolated for further characterization.

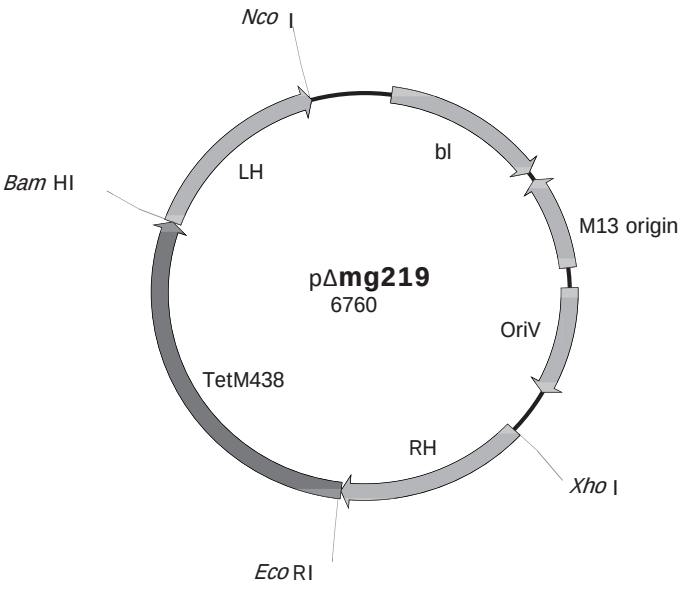


Figure 1.4. Schematic representation of pΔmg219 plasmid. Restriction sites used for the construction of the vector are highlighted in the map.

4.2.4 Confirmation of MG219 null mutant

The intended gene replacement in strain Δmg219 was confirmed by southern blot and the absence of the MG219 protein product was confirmed by western blot (Fig. 1.5). An expected 5006 bp band in the mutant strain confirmed the intended gene replacement. Western blot also confirmed the absence of MG219 protein.

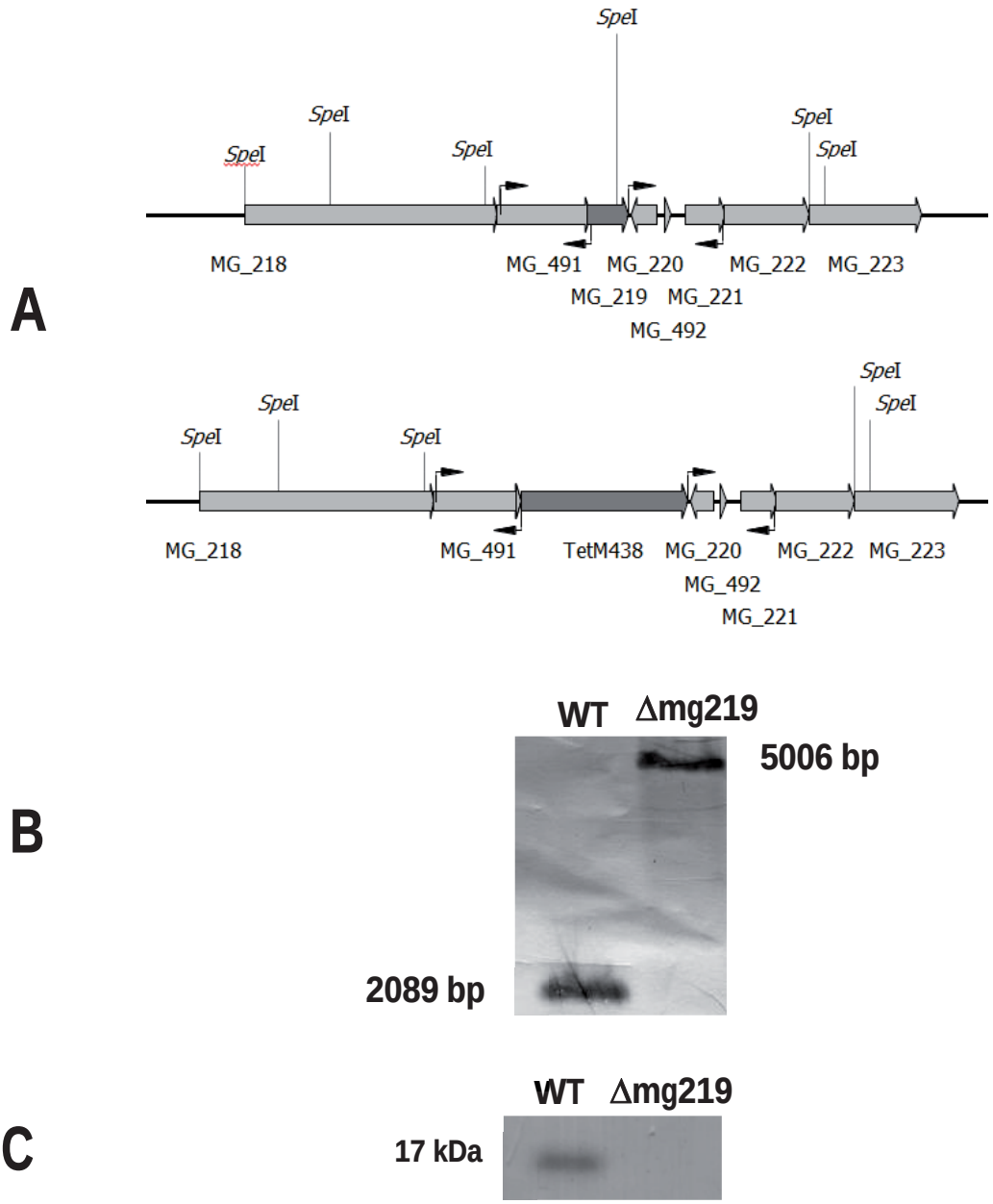


Figure 1.5. Confirmation analyses of strain Δmg219**. A.** Schematics of the genome before (upper panel) and after (lower panel) genetic manipulation with pΔmg219 . **B.** Southern blot using RHR has probe with a genome digested with *Spe*I. **C.** Western blot using MG219 antiserum.

Protein profile of MG219 mutant and western blot using antisera against all the known proteins of the terminal organelle did not show any relevant affectation except from the expected MG219 missing band in Δmg219 strain (Fig. 1.6).

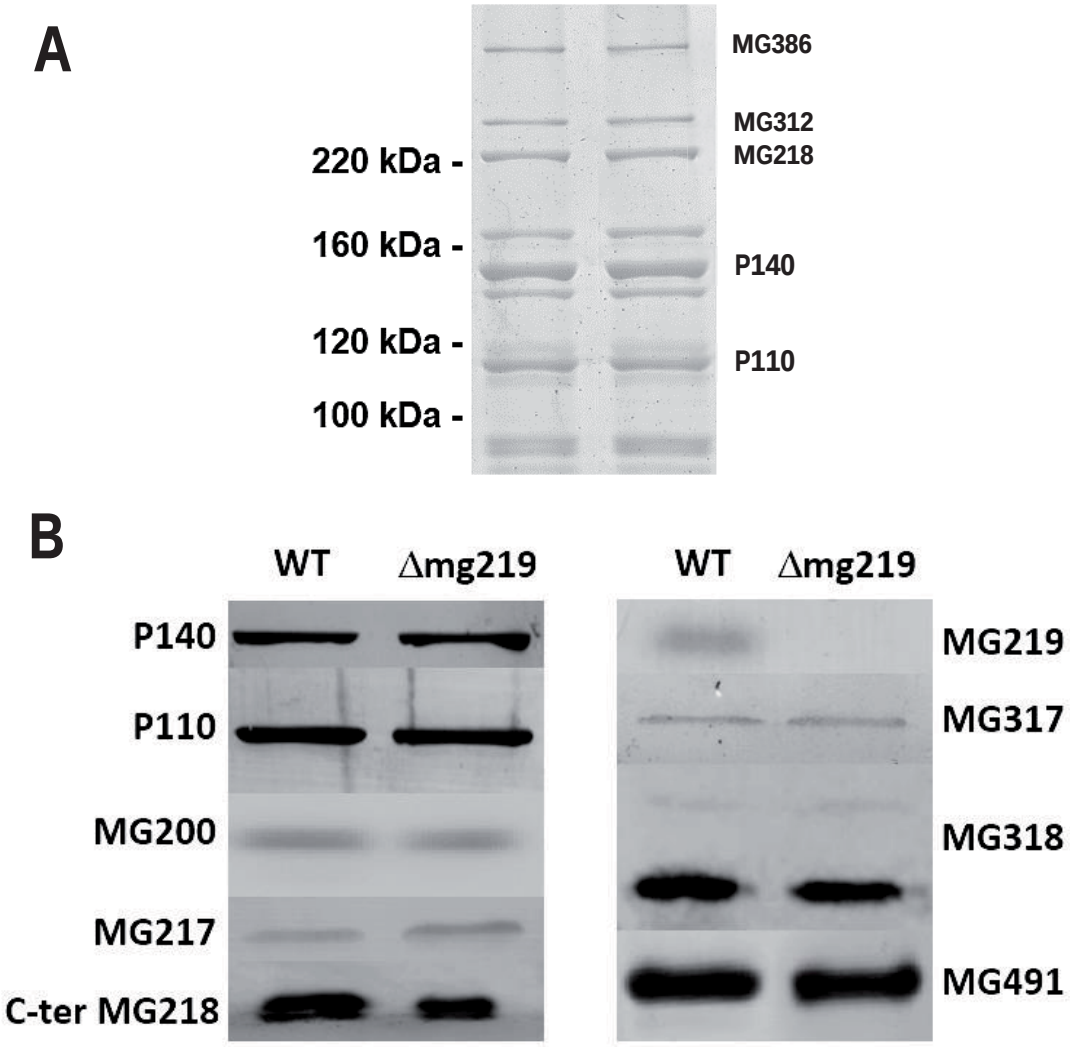


Figure 1.6. Protein levels of terminal organelle related proteins of **A.** Analysis of protein levels of high molecular weight proteins by SDS-PAGE. **B.** Western blot against from WT and Δmg219 strains.

4.2.5 Obtaining Δmg219 based mutant strains

One of the main issues when performing phenotyping studies on strains engineered to present mutations in the genome is to be certain that the effects observed are only consequence of the desired mutation. For this purpose a reintroduction of the WT allele is normally introduced in the genome in a replicative plasmid or in a random position in the genome. This approach is called complementation assay. It is normally required that this complementation is made in another position of the genome discarding any polar effects of the initial mutation. When this assay is performed by transposition, insertion sites of the transposon should be examined to discard disruption of other proteins relevant in latter phenotyping. For these complementation studies a minitransposon containing MG219 under the control of *ackA* promoter and the gentamicin resistance gene (aac-aph) between the inverted repeats was constructed and called pMTnGmRmg219 (Figure 1.7A). The *ackA* promoter has been proven useful for the expression of proteins in *M. pneumoniae*. (Schmidl et al., 2007). Since the translation start site of MG219 has been never stablished experimentally and the primer extension experiments suggest that the TSS of MG_219 lies within MG219 coding region, another minitransposon based plasmid harbouring a shortened version was also created. This plasmid is called pMTnGmRmg2192nd (Figure 1.7B) and carries the coding sequence for MG219 from methionine at position 12, the nearest possible ATG codon downstream of the putative TSSs identified.

MG386 and MG200 have been identified as interaction partners of MG219 *in vitro* but we have no information about the interaction partners *in vivo*. Since purification of tagged proteins from mycoplasmas has been previously proved effective (Schmidl et al., 2007), plasmidpMTNGmmg219Strep-tag (Figure 1.7C) was created to perform a Strep-tag affinity purification strategy. This strategy aimed to identify co-purification partners of MG219. This vector harbours between the inverted repeat regions the acc-aph gene and a translational fusion between MG219 and Strep-tag sequence (Schmidt and Skerra, 2007) at MG219 C-terminus.

Given the doubtful homology between P24 and MG219, the plasmid pMTNGmMPN312 harbouring a MPN312 copy under the control of *ackA* promoter was also created (Figure 1.7D). Since orthologous gene replacement has been proven useful for homology confirmation between P30 *M. pneumoniae* and P32 of *M. genitalium* (Relich and Balish, 2011), one can expect that the introduction of a MPN312 copy in Δmg219 would clarify wheter MG219 and P24 are orthologues or not.

Aiming to obtain data about the subcellular localization of MG219 in relation to the cell body and the terminal organelle, plasmids harbouring translational fusions with fluorescent proteins were created. Plasmids pMTnCmEYFP:mg219 and pMTnCmEYFP:mg2192nd present a translational fusion of EYFP and MG219 protein (longer or shorter version, respectively) at the N-terminal of MG219 (Figure191.7D and E). EYFP is a derived version of GFP from *Aequora Victoria* with 4 amino acids substitutions which a fluorescence roughly equivalent to the GFP generated by Clontech. These two N-terminal protein fusions carry a linker of 10 Asn residues between the two proteins. They also contain within the inverted repeats the *cat* gene under the control of MG_438 promoter.

Simultaneously, pmg219:mCherry vector was created. This plasmid is intended to perform homologous recombination in WT or in Δmg219 strain and replace the MG219 WT allele or tetM438 for a C-terminal translational fusion of MG219 with mCherry under the control of its own promoter (Figure191.7G). mCherry is a monomeric 27 kDa fluorescent red protein from *Discosoma sp.* which shows a good photostability compared to other monomeric fluorophores and usually cells present a low level autofluorescence under its excitation conditions (Shaner et al., 2004).

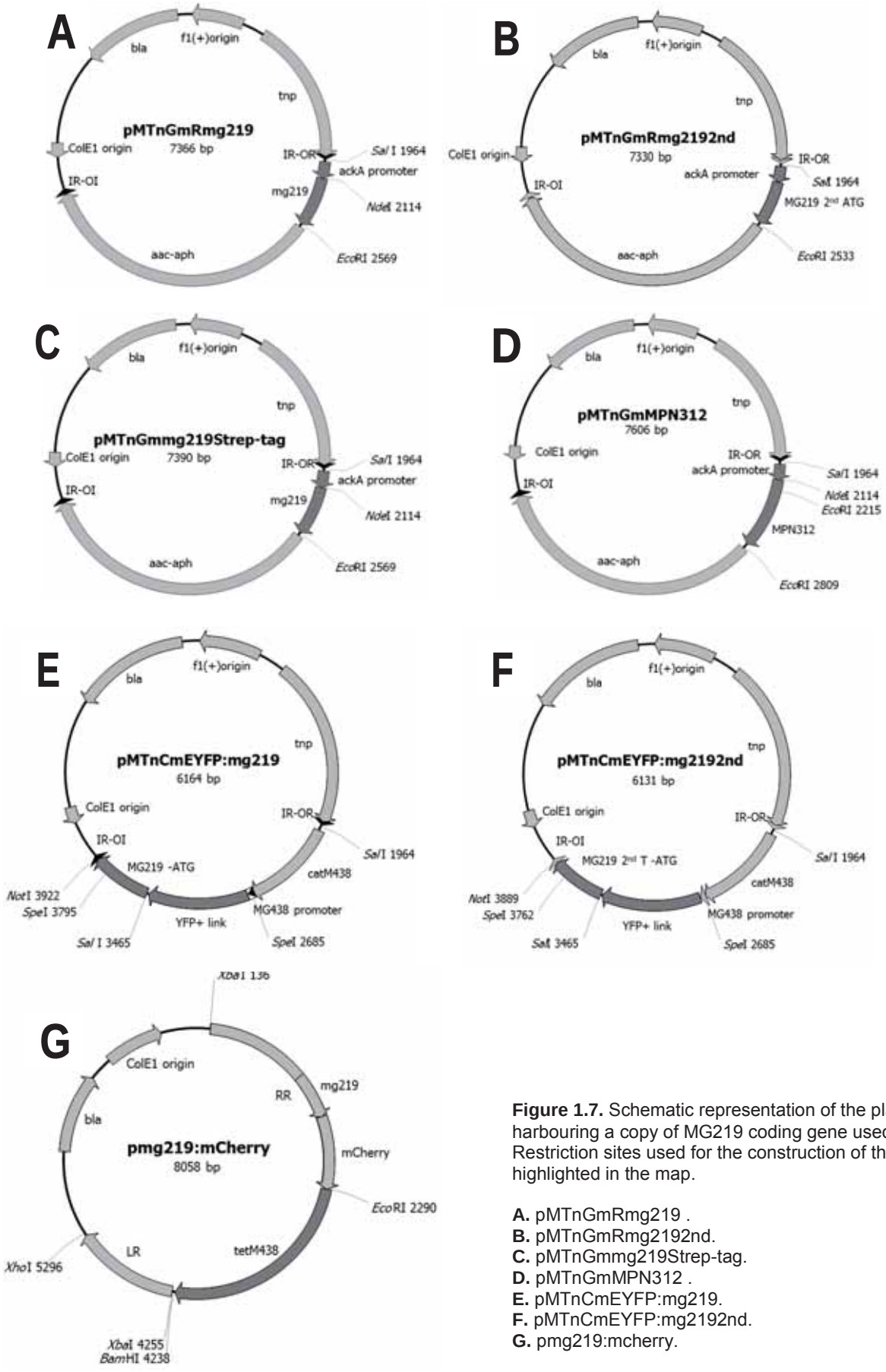


Figure 1.7. Schematic representation of the plasmids harbouring a copy of MG219 coding gene used in chapter I. Restriction sites used for the construction of the vector are highlighted in the map.

A. pMTnGmRmg219 .
B. pMTnGmRmg2192nd.
C. pMTnGmmg219Strep-tag.
D. pMTnGmMPN312 .
E. pMTnCmEYFP:mg219.
F. pMTnCmEYFP:mg2192nd.
G. pmg219:mcherry.

Terminal organelle is only around 70 nm in thickness, which is clearly under the resolution of conventional optical microscopy. In order to differentiate the cell poles presenting terminal organelles a translation fusion of P32 protein with mCherry or EGFP were created (Figure 1.8A and 1.8B). The orthologous of P32 is a profusely studied protein and has been stablished as marker of the terminal organelle in fluorescence studies in *M. pneumoniae*[3] (Relich and Balish, 2011, Hasselbring et al., 2005, Hasselbring and Krause, 2007b, Hasselbring and Krause, 2007a).. The EGFP protein is the green fluorescent 27 kDa protein from *A. Victoria* with the point mutation F64L which presents an acceptable stability for most applications at 37 ºC (Cormack et al., 1996).

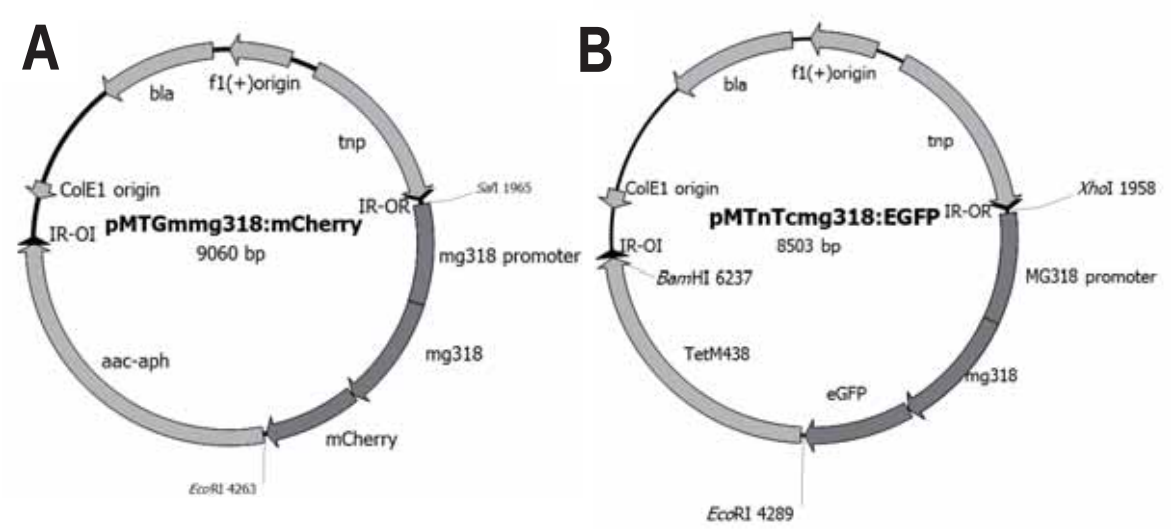


Figure 1.8. Plasmids containing translational fusions of P32 with fluorescent proteins. **A.** pMTGmmg318:mcherry. **B.** pMTnTcmg318:EGFP. Restriction sites used for the construction of the vector are highlighted in the map.

The plasmids were transformed in different combinations and in different strains to obtain the strains in table 1.1. A minimum of five single colonies were recovered from each transformation experiment and further analysed for selection of the most suitable for phenotyping studies.

[3] For more detailed information about P32 protein see chapter 2

Table 1.1. Different mutant strains generated in chapter I.

Parental strain	Plasmid	Strain generated	Antibiotic	Insertion base	Gene
G37 WT	pΔmg219	Δmg219	Tet	NA	-
Δmg219	pMTnGmRmg219	Δmg219 + MG219	TetGm	313362	MgPa island
Δmg219	pMTnGmRmg2192nd	Δmg219 + MG2192ndATG	TetGm	84142	Intergenic
Δmg219	pMTnCmEYFP:mg219	Δmg219 + MG219GYFP	TetCm	124280	Intergenic
Δmg219	pMTnCmEYFP:mg2192nd	Δmg219 + MG2192ndATGYFP	TetCm	349212	Intergenic
Δmg219	pMTNGmMPN312	Δmg219 + MPN312	TetGm	36842	Intergenic
Δmg219	pmg219:mCherry	mg219ch	Gm	NA	-
G37 WT	pMTGmmg318:mCherry	P32ch	Gm	362142	MG_294
P32ch	pMTnCmEYFP:mg2192nd	P32ch + MG2192ndATGYFP	GmCm	347043	MG_284
mg219ch	pMTnTcmg318:EGFP	mg219ch + P32GFP	GmTet	394695	MG_316

4.2.6 Confirmation of Δmg219 based mutant strains

Mutant strains generated using the plasmids in the latter section were confirmed by direct genome sequencing for transposons insertions. This direct genomic sequence was performed with primers upstream and downstream of the antibiotic resistance gene giving information about the insertion point and confirming the genetic construct inserted in the genome. The clone selection for further studies was done using the following criteria: non-fluorescent oriented mutants were selected by insertion point, being preferred the ones not affecting ORFs (Table 1.1) and fluorescent mutants were selected for the highest signal in fluorescence analyses not for insertion site. MG_284, MG_294 and MG_316 are genes coding for proteins previously unrelated with gliding motility or adhesion and the transposon mutant inside this coding region showed a growing and gliding phenotype indistinguishable from WT (data not shown). In the generation of mg219ch strain, an antibiotic resistance confirmation approach was taken. Since recombinants by only one homology region would present both tetracycline and gentamicin resistance, only clones presenting gentamicin resistance but no tetracycline resistance were further analysed.

To confirm that each construct lead to desired MG219 form designed, a western blot approach was taken (Fig. 1.9A). MPN312 expression and stability was also tested (Fig. 1.9B). As expected due to its low identity in amino acid sequence, no cross-reaction of antisera between MPN312 and MG219 was observed.

A unique band for MG219 or MPN312 was present except for fluorescent protein fusions with MG219. MG219YFP and MG219YFP2ndATG are expected to have a Mw of 45 kDa and 44 kDa respectively. The higher molecular weight band detected around 45kDa is compatible with the expected molecular weight of 44-45 kDa for strains carrying both versions of the EYFP fusion. Unluckily, this is not the only band manifested in these strains. pMTnCmEYFP:mg219 showed a 34 kDa band and a 17 kDa band whereas that pMTnCmEYFP:mg2192nd only presents an extra band of 34 kDa. A similar event is shown in the strain carrying MG219ch when analysed by Western blot. Besides the expected 43 kDa band for MG219ch another band of 28 kDa is present.

Confirming the hypothesis that the real MG219 ORF expressed in the WT strain is smaller than computationally predicted, a 17 kDa is observed in WT and Δmg219 + MG219 strains. By contrast, a 18 kDa protein is observed in Δmg219 + MG219 strain. All further experiments were only performed on strains harbouring a copy of MG2192nd version of the protein, since all data pointed that MG_219 gene annotated in the reference genome is shorter than expected.

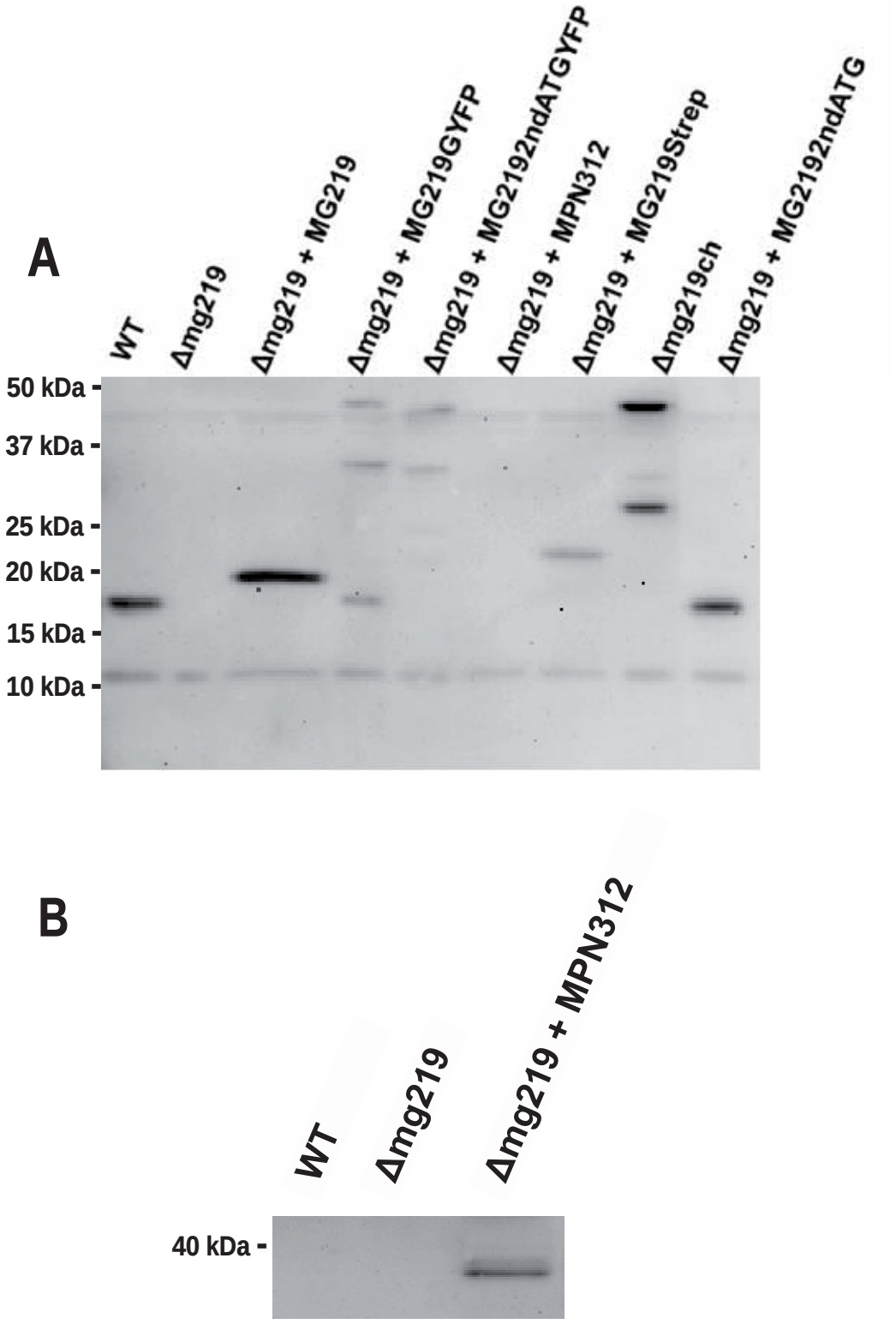


Figure 1.9. Western blot against from WT and strains generated from Δmg219. **A.** Western blot using MG219 antiserum. **B.** Western blot using MPN312 antiserum.

4.2.7 Subcellular localization of MG219 and MG318 proteins

Unlike *M. pneumoniae*, whose cells are more elongated and filamentous, *M. gentialium* cells look much more round-shaped and its polarity is difficultly assessed by visible light contrast phase microscopy. In spite of this handicap, epifluorescence analyses of P32ch+MG2192ndATGYFP and mg219ch+P32GFP strains revealed a polar localization for MG219 and P32 (Fig. 1.10). All epifluorescence studies were performed with the addition of Hoechst 33342 to the samples, resulting in the emission of blue fluorescence in the nucleoid location when excited with UV. These polar localizations are less pronounced than in *M. pneumoniae* since overlapping between part of the DNA signal and the terminal organelle tagged proteins is visible. From these polar localizations is evident that, in both strains, P32 is located farther from the nucleoid than MG219 in most of the analysed cells. Distance measurements of the brightest pixel for each signal showed that P32 is 275 ± 3 nm from the nucleoid and MG219 is at 172 ± 2 nm from P32. In some cases, MG219 fluorescence signal is not between P32 and DNA foci. Noteworthy is the presence of dual terminal organelle cells in division, corroborating the view that P32 and MG219 are members of the terminal organelle.

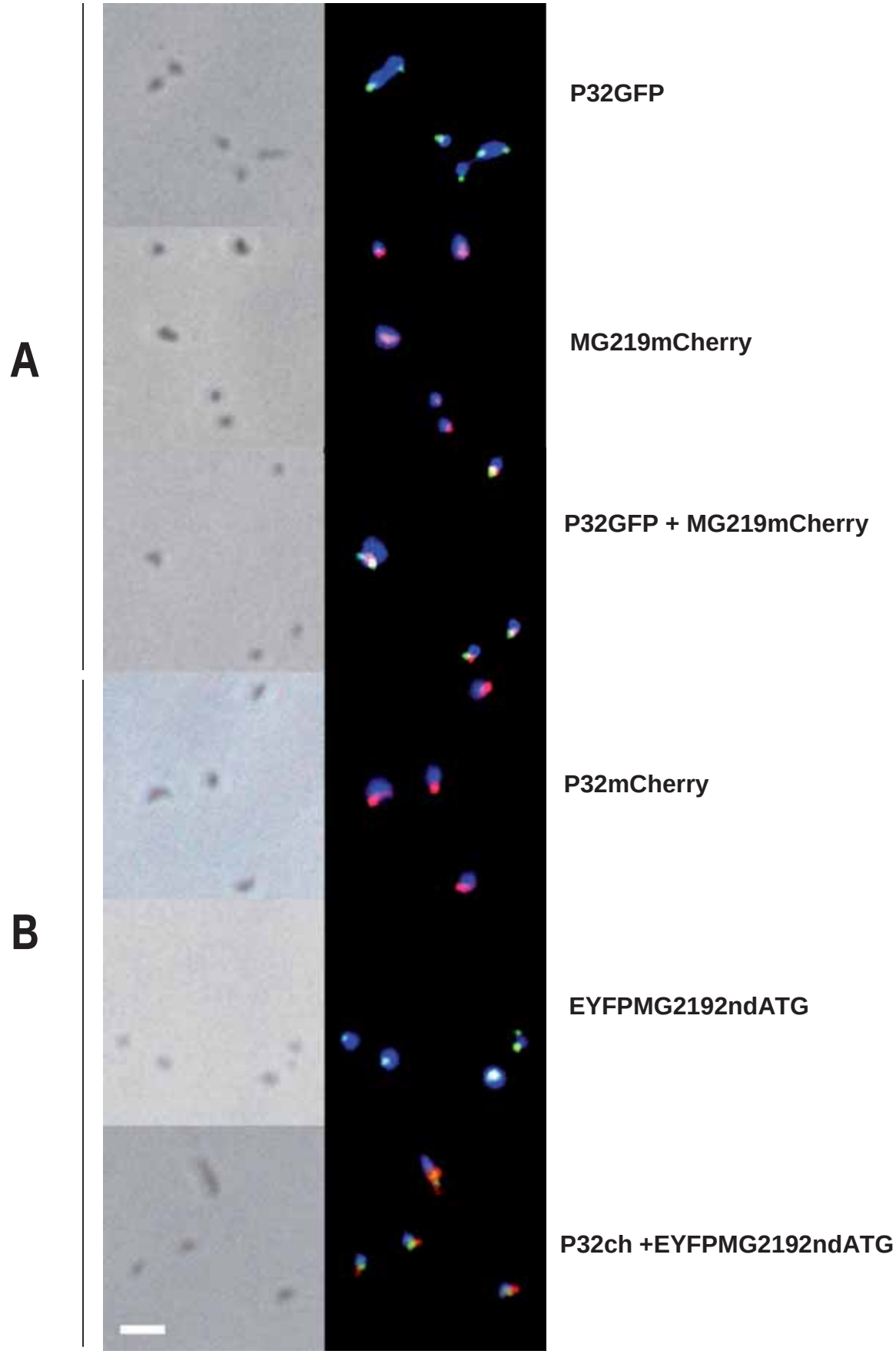


Figure 1.10. Epifluorescence microscopy for subcellular localization of MG219 and P32. **A.** mg219ch + P32GFP. **B.** P32ch + MG2192ndATGYFP Scale bar is 2.5 µm.

4.2.8 Characterization of MG219 null mutant

4.2.8.1 Gliding motility

Comparative micro-cinematographic studies of WT and Δmg219 strains showed that MG219 is involved in gliding motility (Figure 1.11A and 1.11B). A significant difference is clear in the frequency of cells depicting movement and also in the length of those showing to be much shorter. In addition, the reintroduction of MG2192ndATG, MG219ch, MG2192ndATGEYFP coding alleles did show an apparent recovery in the tracks depicted by gliding and the proportion of cells gliding (Figure 1.11C, 1.11D and 1.11E). In contrast, the introduction of MPN312 gene did not seem to restore WT levels of gliding activity (Figure 1.11F).

After carefully analysing the tracks depicted by most of the strains generated (Figure 1.12A) it seemed like the frequency of erratic tracks varies in the different analysed strains, but the Z-test statistical test indicates no significant differences ($P<0.01$). In contrast, quantitative analysis of gliding motility confirmed the notions inferred from the tracks of the cells depicted in two minutes of observation. Figure 1.12B shows that cells lacking MG219 proteins glide much less frequently than the WT (79.2% vs 18.4%). Although not completely, the gilding frequency is recovered when reintroducing any of the alleles of MG219 designed in this work. Z-test indicates there is no statistical prove that WT and complemented strains are significantly different. Distinctively, when introducing a MPN312 copy in Δmg219 strain a significant gliding frequency variation of 39% is observed.

Quantitative gliding velocity assays of the mutant strains generated revealed that Δmg219 strain glides 60% slower than the WT strain (0.045 ± 0.004 μm s^{-1} vs 0.112 ± 0.004 μm s^{-1}). No significant differences were observed in gliding speed of strains harbouring a copy of MG219 with or without fluorescent protein fusions. In contrast, mutant strain with a copy of P24 instead of MG219 showed slightly higher speed than the null mutant (0.57 ± 0.003 μm s^{-1}) but this difference was not supported by statistical significance.

The diameter of circular tracks was also analysed but since WT and Δmg219 showed no statistical differences (1.16 μm ± 0.12 vs 1.28 ± 0.07 μm, respectively) no further strains were examined.

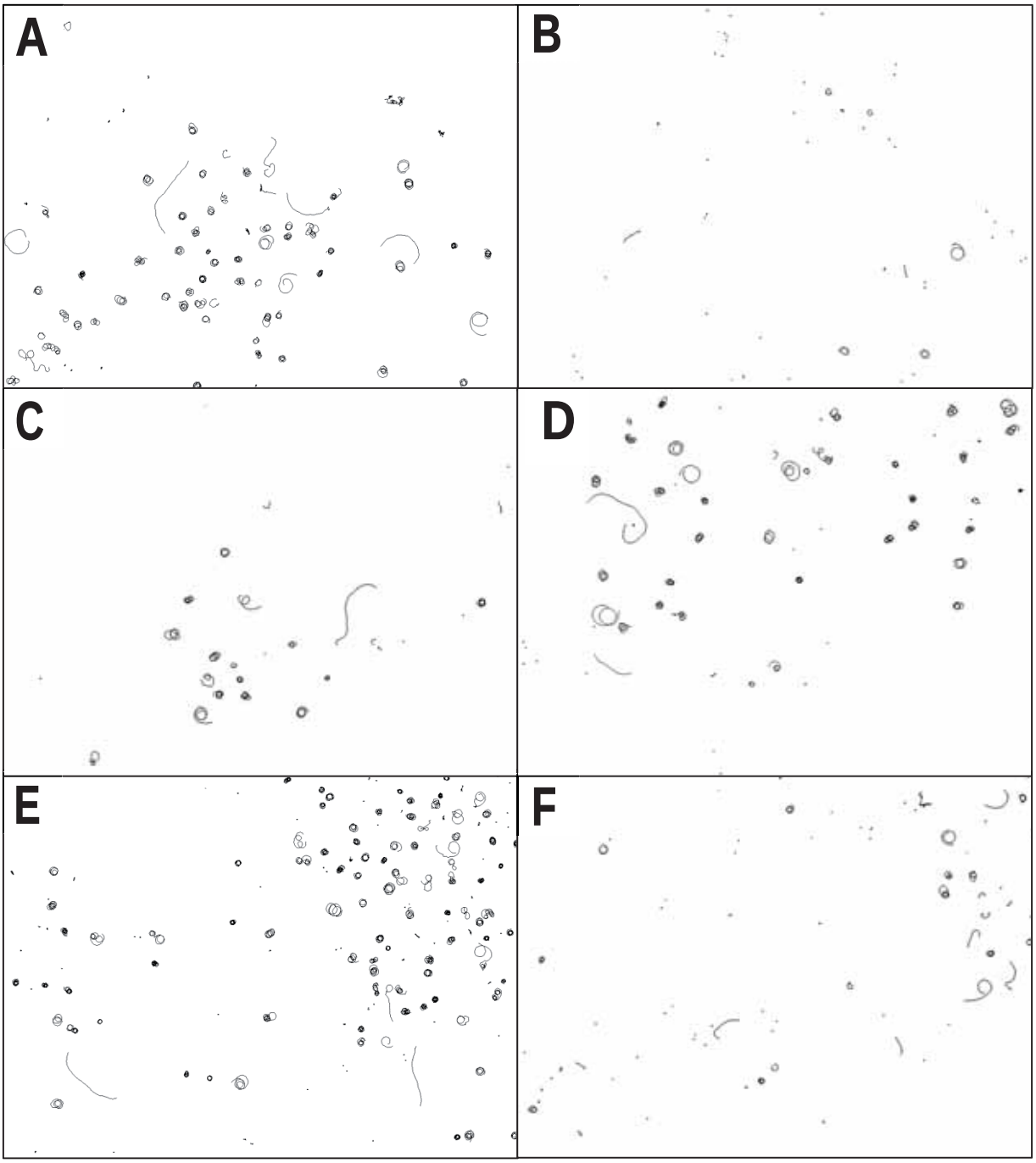


Figure 1.11. Paths depicted by MG219 mutant strains in comparison with WT strain during a period of two minutes at 37ºC in SP4 medium. **A.** WT **B.** Δmg219 **C.** Δmg219 + MG219 **D.** Δmg219 + MG2192ndATG **E.** Δmg219 + MG219GYFP **F.** Δmg219 + MG2192ndATGYFP **G.** Δmg219 + MPN312.
Each image shows a region of 86.9x65.2 µm.

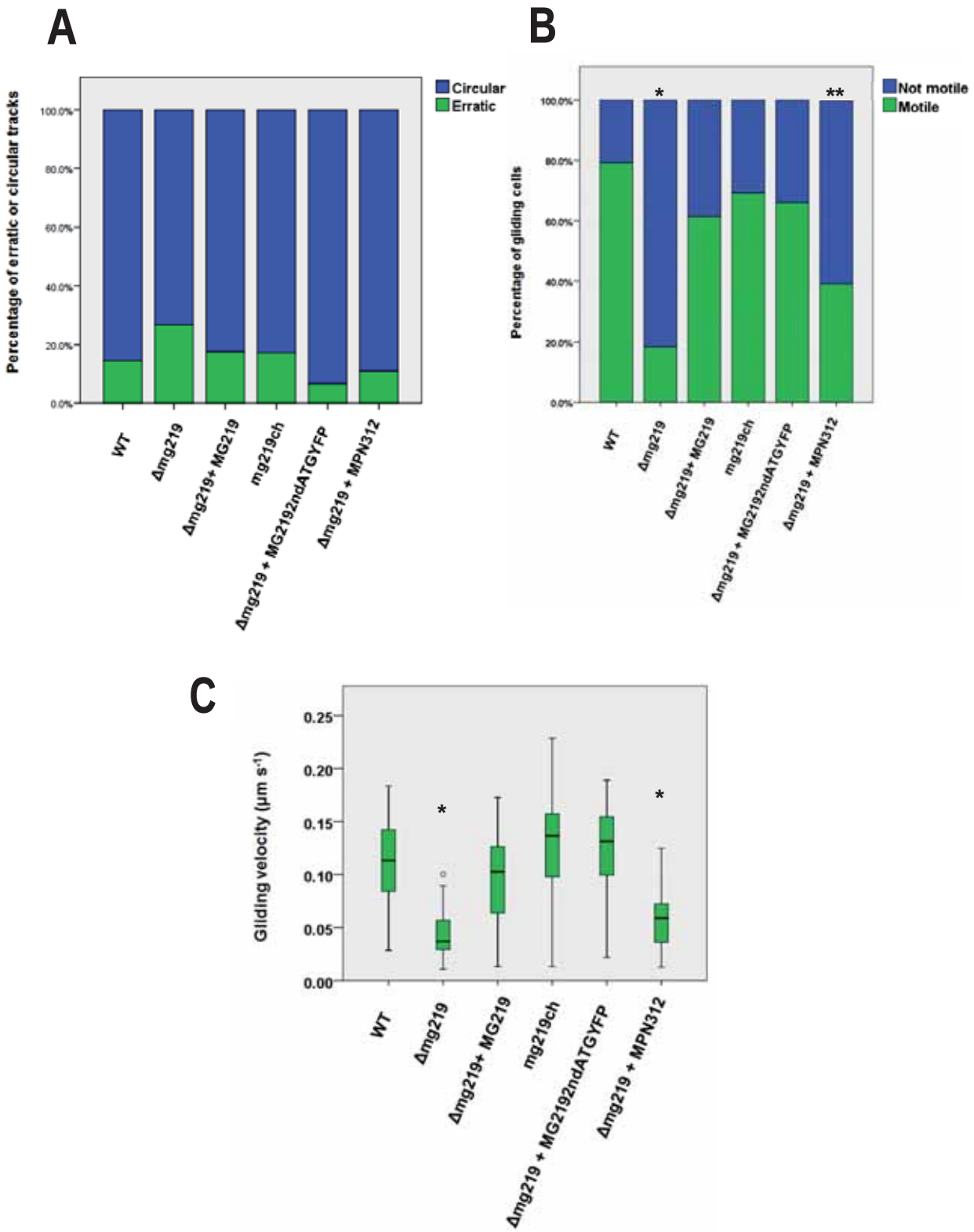


Figure 1.12. Gliding properties of WT and strains with modifications related to MG219. **A.** Frequency of erratic and circular tracks **B.** Frequency of motile and non-motile cells **C.** Boxplots of gliding velocity. The asterisk (*) indicates statistical significantly different when compared with the WT (p-value < 0.01). The double asterisk (**) indicates statistical significantly different when compared with both the WT and the null mutant Δmg219 (p-value < 0.01).

4.2.8.2 Cellular morphology

Cellular morphology of WT and mutant strains was evaluated by means of scanning electron microscopy (Figure 1.13). Although cells in all samples analysed showed flask shape morphology the frequency of cells in cell division stages is altered in mutants related to the MG219 protein. Most remarkably, the Δmg219 strain showed a high number of cells in cell division and several cells with more than one terminal organelle. After reintroducing a copy of MG2192nd in Δmg219 strain, cells presented normal cell division and no multiple terminal organelles.

For a quantitative analysis of the cell division status four groups were defined:

- Cell single status: one cell with one terminal organelle
- Division status: from one cell with two terminal organelles to two cells with one terminal organelle connected by a filament)
- Multiple cell: when more than two cells are clumped together, most likely because of inefficient cell division
- Multiple TO: when a cell have more than two terminal organelles

After the analyses (Fig. 1.14) significant differences were found in Δmg219 strain and Δmg219 + MPN312 strain. Those significant differences are in all the four cell status studied, an increase in 28% of cells in division was found in Δmg219 strain plus an increase of 10% in multiple cell status and also an increase of 7% in cells with multiple terminal organelles. As projected, Δmg219 strain after the reincorporation of MG2192nd did recover all the WT frequencies for all cell status. Interestingly, the strain carrying a MPN312 copy instead of MG_219 showed no cells with multiple terminal organelles, exactly as the WT strain.

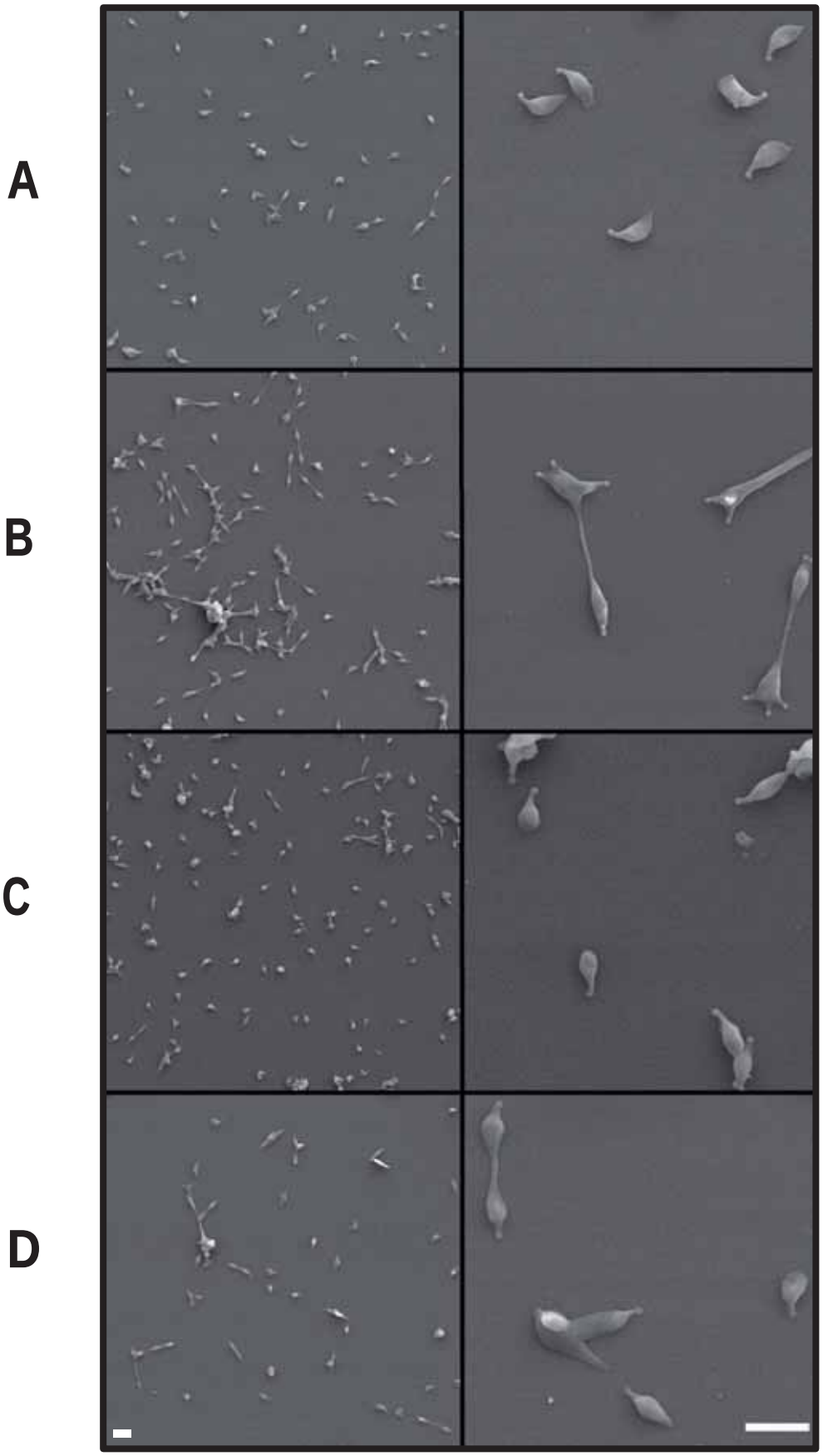


Figure 1.13. Scanning electron microscopy of WT and MG219 related strains. **A.** WT **B.** Δmg219 **C.** Δmg219 + MG219 **D.** Δmg219 + MG2192ndATG **E.** Δmg219 + MG219GYFP **F.** Δmg219 + MG2192ndATGYFP **G.** Δmg219 + MPN312. Scale bar is 1 µm.

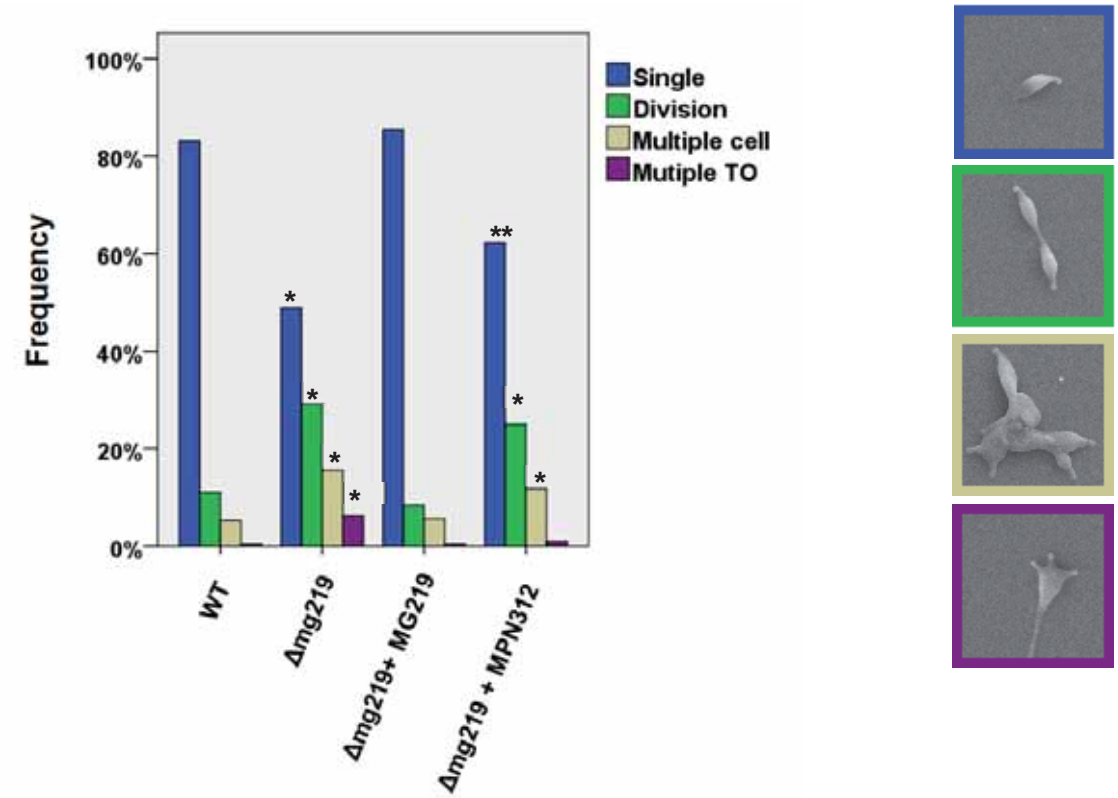


Figure 1.14. Cell division status of WT strain and MG219 mutant strains. The asterisk (*) indicates statistical significantly different when compared with the WT (p value < 0.01). The double asterisk (**) indicates statistical significantly different when compared with both the WT and the null mutant Δmg219 (p value < 0.01).

4.3 Discussion

The study of the mutants designed in this chapter revealed information about a unique protein of *M. genitalium*: MG219. MG219 protein shows no sequence homology to any other in the public repositories but it is suspected to have a similar function to the P24 protein coded by MPN312 from *M. pneumoniae*. To study if this theory is correct and to bring light to the knowledge gap that currently exists about MG219 protein, a null mutant lacking MG_219 has been successfully generated and was named Δmg219. An antiserum against this hypothetical protein has been generated using the recombinant MG219 protein produced in *E.coli* and has proved to be useful for the detection of MG219 by western blot. As expected, the null mutant for MG219 mutant showed no signal when proved with the MG219 antiserum confirming the expected phenotype.

Interestingly, primer extension analysis of MG219 showed three low intensity peaks compatible with a heterogeneous TSS within the coding region of MG219. Inspection of the coding region showed that this TSS lays within the coding region of amino acids K7 and K8. Noteworthy is the presence of a start codon (ATG) in position M12. To further analyse the possibility that the real protein is shorter than expected, a construct capable of delivering by transposition the full-length protein plus another construct bearing the MG219 protein from M12 were generated (MG2192nd). The size of the bands revealed by Western blot analyses revealed that MG219 in the wild-type strain have a lower (17 kDa) Mw than the predicted MG219 full length (18 kDa). In contrast, MG2192nd showed a MW of 17 kDa. Taken all this data together we concluded that MG2192nd is the real form expressed by MG_219 ORF and genome annotation of the WT strain should be modified in accordance. In agreement with these data, a plasmid bearing an EYFP fusion of predicted MG219 full length in its N-terminal showed a 17 kDa band while an N-terminal fusion of EYFP with MG2192nd did not. This data points that maybe in the 19 bp upstream of the heterologous TSS detected but within the full-length MG219 lies on an unidentified promoter capable of promoting the expression MG219.

Western blot analyses of strains carrying either C-ter mCherry fusion or N-ter EYFP fusion with MG219 showed that fusions of MG219 with fluorescent proteins are not entirely behaving as estimated. The expected band (45 kDa for EYFP and 43 kDa for mCherry) plus an extra band (35 kDa for EYFP and 30 kDa for mCherry) were detected. Since the decrease of the band size is detected with the same intensity that the full-length fusion and MG219 is a protein of only 17kDa it is likely that the extra band is consequence of a proteolytic cleavage of the fluorescent protein rather than a proteolytic cleavage of MG219 protein. Further

reinforcing this view, if the proteolytic cleavage is performed within the MG219 protein, a new band around 10 kDa should be observed in fluorescent fusion mutants, and this has not been observed. Considering that both fluorescent proteins present a sole degradation version, it is feasible that this band is a proteolytic cleavage than degradation by instability. Further experiments are required to demonstrate that.

As expected, fluorescent translational fusions of P32 with either EGFP or mCherry proved to be functional for the determination terminal organelle localization in *M. genitalium* cells. EYFP and mCherry translational fusions with MG219 proved that either tagged N-ter or C-ter MG219 localizes around 100 nm closer to the centre of the cell body than P32, but still polarized. Both localizations are fully compatible with the size of the terminal organelle of *M. genitalium* which is around 200 nm. In some cases, a side-to-side location of MG219 and P32 respective to the DNA signal is observed. Due to the curvature of the terminal organelle during gliding activity, it is possible that the terminal button and the wheel complex are not lined up when the cells are depicting very narrow circular tracks. Interestingly, P24 in *M. pneumoniae* (Fig. I.3B) and MG219 in *M.genitalium* showed a polar localization compatible with being part of the wheel complex, although P24 in *M. pneumoniae* showed a stronger signal also present in the cytoplasm of the cell.

Western blot using the MG219 antiserum also revealed that MG219Strep-tag is stable when introduced in a strain previously lacking MG219 protein. Using this strain bearing a copy of MG219 with a C-ter Strep-tag an experiment trying to isolate MG219 and its possible interaction partners *in vivo* was attempted. Unluckily, those experiments yielded non-detectable levels of MG219 protein either by Coomasie-blue or Western blot (data not shown). Given the insoluble nature of such an enormous macromolecular complex that is the cytoskeleton of *M. genitalium*, it is possible that MG219 remains in the insoluble fraction after cell lysis, making impossible the recovery of MG219 in the conditions tested. Even using 6 M urea in the purification protocol resulted in MG219 levels barely detectable by Western blot (data not shown). It is also possible that MG219 interaction partners, such as de C-ter of MG386 make sterically impossible the interaction of the Strep-tag with the strep-tactin sepharose. An intensive and extended protocol development was anticipated and not further developed due to time limitations. A combination of denaturation step with high concentrations of urea or guanidinium chloride with crosslinking reagents in combination with high amounts of biomass could lead to successful *in vivo* detection experiments. An affinity un-tagged purification approach have been developed with success in *M. pneumoniae* for P1 protein for the detection of interaction partners (Layh-Schmitt et al., 2000) suggesting that with the proper optimization this protocol should work.

Studies of gliding motility showed that a mutant lacking MG219 protein had a 60% decrease in gliding frequency and also a 60% in gliding velocity which is fully restored after the introduction of MG2192nd, MG219EYFP or MG219ch. This result demonstrates a specific role of MG219 in gliding activity, and that the C-ter and N-ter tags tested does not alter MG219 functionality. Interestingly, MG219 + MPN312 strain revealed that MPN312 protein is capable of partially recovering the gliding frequency (increase of 20%) in Δmg219 but not able to recover significant amount in gliding velocity.

While cells of all strains presented flask-shaped morphology after microscopy analyses, a greater number of cells in division and with multiple terminal organelles were observed in Δmg219. Reintroduction of all MG219 alleles designed in this study in a Δmg219 background resulted in the recovery of the wild-type phenotype. In contrast, when introducing a copy of MPN312 gene in Δmg219, the number of multiple terminal organelles was reduced to the levels while the number of cell in cell division were not.As previously discussed the greater number of cells in division or with multiple terminal organelles is most likely caused by the involvement of gliding motility in cell division. This view is reinforced by the findings that Δmg219 + MPN312 strain glides more frequently as the number of cells with multiple terminal organelles is lower.

In conclusion, P24 is capable of virtually reducing to zero the number of multiple terminal organelles observed in Δmg219 strain and also capable partially restoring the gliding frequency.These results suggest that P24 is capable of performing some of the functions of MG219, but there is no evidence whether this "partial functionality" is caused by the marginal sequence identity or, for example, folding problems. It is possible that MG219 and P24 had a common gliding protein ancestor but finally they ended evolving in almost completely different proteins. After taking all this data together it is possible to claim that MG219 is a crucial protein for gliding motility and it has no homologues in any of the currently sequenced organisms in public repositories.

Although the molecular mechanisms in which MG219 is involved in the generation or propagation of the movement to the cell body are unknown, interaction studies could shed light to this question. Since the known interaction partners of MG219 are MG386 and MG200, both proteins previously linked with gliding frequency and speed, it is reasonable to hypothesize that MG219 could have a role in the coordination of the gliding mechanisms where the wheel complex is involved. This view will be reinforced by experiments carried out in chapter III and chapter IV. Remarkably, the C-ter of MG219 has a theoretical pI of 11 due to its high content in lysines and its interacting partner, the C-ter of MG386, has a theoretical

pI of 4 due to its high content of glutamic acid residues. These physicochemical properties lead to hypothesize that MG386 and MG219 are interacting due electrostatic charges. Since P24 also presents a lyisine rich C-ter region with a pI higher than 10, it is possible that the partial restoration of the phenotype of MG_219 null mutant by the introduction of MPN312 gene is caused by the similar electrostatic properties of P24 protein and MG219. *In vitro* interaction studies of these proteins can be performed to enlighten that question. Structural analyses of the interaction complexes would be also very interesting to explore, since MG219 might be an intrinsically disordered protein capable of acquiring folding after interaction with its partners.