Molecular interactions underpinning the segregation of TP228 multidrug resistance plasmid in *Escherichia coli*

Cecilia Pennica

PhD

University of York

Biology

September 2018

Abstract

Accurate segregation of newly replicated plasmids is essential for genetic stability throughout generations. The conjugative plasmid TP228, identified in *Salmonella newport*, confers resistance to several antibiotics and it is segregated during cell division thanks to its self-encoded *parFGH* partition system. This is composed of two *trans*-acting elements, ParF and ParG, and a centromeric region, *parH*. Interactions between the three components are essential for plasmid maintenance and are the main focus of this work.

ParG is a dimeric protein composed of a folded domain and two unstructured tails. Alanine-scanning mutagenesis of the ParG tail has highlighted a number of residues, whose change impairs plasmid partitioning *in vivo*. These flexible regions were shown to play a fundamental role in binding to the partner protein ParF as well as conferring specificity to the interaction with the DNA. A combination of biochemical and biophysical techniques have been used to address which regions in the ParG tails are essential for the interaction with ParF and which play different roles. Surface plasmon resonance, together with microscale thermophoresis allowed quantification of the change in binding when residues are replaced in ParG. These results also allowed identification of positions in the tail that are supposedly required for binding to the DNA and structuring the tail at the initial stage of segrosome formation. Circular dichroism and electrophoretic mobility shift assays showed potential interaction of the amino acids in the middle of the tail with the site-specific DNA. On the other hand, cross-linking between ParF and ParG, in the absence and presence of *parH,* coupled to mass-spectrometry revealed a more complicated picture of the interaction among the segrosome components. Interestingly, interaction with the partition site greatly enhances ParF-ParG association, probably by structuring the ParG N-terminal tail. Finally, a tandem affinity purification protocol gave new insights into the interaction network of the ParFGH system. A number of cell division, membrane and DNA-binding proteins were identified and tested for interaction with ParF or involvement in plasmid segregation. Based on these findings, a more detailed mechanism of segrosome assembly and plasmid segregation is proposed.

Table of contents

List of Tables

List of Figures

Acknowledgments

I would like to thank my supervisors Dr. Daniela Barillà and Dr. Andrew Leech for the opportunity to work on such an amazing project and for all the help, support, guidance and stimulating conversations during the years.

I would also to extend my thanks to my TAP members, Prof. James Chong and Prof. Gavin Thomas and to Dr. Adam Dowle and Dr. Chris Taylor for the help with mass spectrometry experiments.

A big thank you goes to the past and current members of the Barillà group and to the L1 lab, particularly to Iman Alnaqshabandy and, Dr. Azhar Kabli for being the best colleagues and friends I could have asked for. You made my days much brighter and helped me more than you can imagine.
I would also like to extend my acknowledgment to Dr. Constantinos Drousiotis, Banushan Balansethupathy, and Daniel Yee for sharing this journey with me.

In addition, I would like to thank Prof. Cees Dekker for the opportunity to spend three months in his lab and Dr. Aleksandre Japaridze for being a great teacher and friend.

A massive thank you goes to my family for their support. I would like to thank my parents, Angela e Michele, for all the opportunities they gave me growing up and for letting me “dream big”. A huge thank you goes to my siblings, Isabella and Corrado for being always there for me and for their constant love and endless support.
A massive thank you goes to Fabrizio, Maria, Anastasia and Angelica without whom completing this work would have been much more difficult.
Finally, I would like to thank my friends Alice, Chiara, Laura, Micol, Sara, Sofia and Floriana for their love, patience and support during this journey.

This project was funded by Biotechnology and Biological Sciences Research Council (BBSRC).

Author’s declaration

I declare that this thesis is a presentation of original work and I am the sole author. Work concerning mass spectroscopy was performed by Dr. Adam Dowle and Dr. Chris Taylor. This work has not previously been presented for an award at this, or any other, University. All sources are acknowledged as References.

Chapter 1
Introduction

1.1 Plasmids

Plasmids are extra-chromosomal genetic elements present in most bacterial species as well as in archaea, fungi and yeasts (del Solar *et al.*, 1998; Thomas, 2000). They contain additional, non-essential genes, often beneficial to the host in niche environments, as in the presence of antimicrobial agents or in limiting nutrients conditions. Plasmids have been found to be the source of bacterial resistance to antibiotics, heavy metals and toxic anions. They can also encode genes involved in accessory metabolic pathways, for example nitrogen fixation, molecular degradation, antibiotic and bacteriocin production. In addition, they can contain genes that modify the host life-style, e.g. through the production of toxins or virulence factors (Summers, 1996; del Solar *et al.*, 1998; Thomas, 2000; Hayes, 2003a). Importantly, plasmids can be transferred within (vertically) and across (horizontally by conjugation and transformation) species, enhancing the spreading of genetic traits in the environment (Summers, 1996; Thomas, 2000; Hayes, 2003a). Due to the medical and evolutionary relevance, as well as the possibility of easy manipulation for biotechnological applications, plasmids have been the object of scientific interest since the 1950s, when Joshua Lederberg introduced the term plasmid for the first time (Lederberg, 1952).

Plasmids can vary in size and number, from as small as a couple of kilobases to large megaplasmids, of many hundreds of kilobases (Hayes, 2003a). Large plasmids are generally present in the cell in a low number of copies (1 to 10), while small plasmids are generally high copy number (more than 20 copies per cell) (Zhong *et al*., 2011).

An important feature of plasmids is that they replicate independently from the chromosome and have developed sophisticated strategies to be stably maintained and passed to the progeny. In fact, despite being often advantageous, plasmids are energetically costly to the organism and they can be lost, especially in the absence of selective pressure (del Solar *et al.*, 1998; Thomas, 2000; del Solar and Espinosa, 2002; Silva *et al.*, 2012).

In the case of high copy number plasmids, random distribution is sufficient to ensure at least one copy of the plasmid per cell after cellular division. Low copy number plasmids, instead are present in a few copies per cell and require active segregation mechanisms to be properly distributed to the offspring (Hayes and Barillà, 2006a; Salje, 2010; Brooks and Hwang, 2017). Plasmid stability and, more specifically, plasmid segregation will be the focus of this work.

1.2 Plasmid Replication

In order to be stably maintained, plasmids need to coordinate their replication with the cell cycle (Summers, 1996; del Solar *et al.*, 1998; Del Solar and Espinosa, 2002).As mentioned above, plasmids replicate independently from the host, however they often rely on the replication proteins encoded by the chromosome (del Solar *et al.*, 1998). In most plasmids, genes essential for replication are clustered and confined within a region of 1-3 kb, named the basic replicon. The basic replicon of most plasmids is characterised by an *ori* region where replication starts, a gene encoding a Rep protein, a plasmid-specific replication initiation protein and other genes involved in replication regulation (Summers, 1996; del Solar *et al.*, 1998; Thomas, 2000). Control of plasmid replication initiation is essential. In fact, plasmid copy number needs to be maintained to a level that is not a burden to the host (Thomas, 2000). Large plasmids can contain more than one replicon, although usually only one is active (Hayes, 2003). Three replication mechanisms have been identified: the theta-type, the strand displacement and the rolling-circle mechanisms (del Solar *et al.*, 1998; Thomas, 2000).

1.2.1 Theta-type mechanism

The theta-type is one of the most studied replication mechanisms, found in plasmids isolated from both Gram-negative and Gram-positive bacteria (del Solar *et al.*, 1998). Plasmids that use this replication mechanism are for example P1, R1 and pS10 (Thomas, 2000). Their origin of replication is similar to the *Escherichia coli oriC* (del Solar *et al.*, 1998; Thomas, 2000) and is characterised by short repeated DNA sequences called iterons, that are recognised specifically by the plasmid-encoded Rep protein, one or more *dnaA* boxes, that are binding sites for the DnaA initiator protein, and an A/T-rich sequence adjacent to the iterons. In some cases, Dam (DNA adenine methylase) methylation regions are also present (Thomas, 2000; del Solar and Espinosa, 2002; Hayes, 2003a). These can be essential, as for replication initiation of plasmid P1 (Abeles and Austin, 1987).

The general mechanism is the following: the Rep protein recognises and binds to the iterons. DnaA is also recruited and, together with Rep, changes the conformation of the DNA, causing DNA melting in the surrounding A/T-rich regions (del Solar *et al*., 1998a; Thomas, 2000; Hayes, 2003a). Other host-encoded factors can also be involved in the replication initiation process, as Factor for inversion stimulation (FIS) and Integration host factor (IHF), for instance. Once the replication bubble is formed, DnaB (helicase), DnaC (helicase-loader protein) and DnaG (primase) are recruited to start replication. Replication is continuous in the leading strand and discontinuous in the lagging strand (del Solar *et al.*, 1998; Thomas, 2000). Other host-factors are also involved in the

process, as DNA polymerase III HE (holoenzyme), SSB (single strand binding protein), DNA gyrase and topoisomerase IV, which resolve the DNA topology during the process (del Solar *et al.*, 1998; Thomas, 2000).

In some cases, plasmid replication depends completely on the host-encoded proteins, as for the plasmid ColE1. The replication of this plasmid does not require the presence of a Rep protein. This replication mechanism, however, requires the presence of DNA polymerase I to start replication (del Solar *et al.*, 1998).

1.2.2 Strand displacement mechanism

Strand-displacement mechanism has been observed for broad-host range plasmids from the IncQ family, such as the plasmid RSF1010. It requires the presence of three plasmid-encoded proteins, RepA (helicase), RepB (primase) and RepC (initiator) (del Solar *et al.*, 1998; Thomas, 2000; Hayes, 2003a). The replicon is characterised by three regions: the iterons; iteron-adjacent A/T-rich sequences and two small symmetrical palindromic sequences, *ssiA* and *ssiB,* located on opposite strands. In order to start replication, RepC recognises and binds to the iterons next to the *ssi* regions. RepA is then recruited to the nucleoprotein complex and unwinds the DNA in the 5'→ 3'direction (del Solar *et al.*, 1998). This causes the melting of the double stranded DNA and the exposure of the *ssi* sites on the DNA single strands. In this way, the primase RepB can be recruited to the *ssi* regions and replication starts. Replication is continuous and proceeds in both directions (del Solar *et al.*, 1998; Thomas, 2000). While proceeding along the DNA, RepA facilitates the replication process by displacing the complementary strand (Thomas, 2000). The fact that this mechanism does not require host factors may explain the broad-host range of the plasmids of the IncQ family (del Solar *et al.*, 1998).

1.2.3 Rolling-circle mechanism

The rolling-circle mechanism has been identified for the replication of small plasmids like pT181, pC221, pUB110, and pC194 (del Solar *et al.*, 1998). In this mechanism, the replication of the leading and the lagging strands are uncoupled. Replication starts with the recognition of the *dso* (double stranded origin) site by the plasmid-encoded Rep protein. The protein introduces a nick in the leading strand, leaving a 3'-OH end which works as a primer for the recruitment of host proteins, involved in the replication of the plasmid (DNA polymerase III, SSD and the helicase) (del Solar *et al*., 1998a). While the helicase (UvrD in Gram-negative and PcrA in Gram-positive bacteria (Ruiz-Masó *el al.*, 2015)) unwinds the DNA, the replication around the intact strand of DNA proceeds and the plus strand is displaced. When the replication of the leading strand is concluded, a ligase seals the nick of the double stranded molecule made of the parental minus strand

and the newly replicated plus strand. The parental single stranded plus strand is also repaired and converted into a duplex DNA molecule by the action of host proteins. First, RNA polymerase synthesises RNA primers, which are extended by the action of DNA polymerase I. Finally, DNA polymerase III synthesises the complementary strand of the parental leading strand and DNA ligase circularises the plasmid. (del Solar *et al.*, 1998; Thomas, 2000; Hayes, 2003a, Khan, 2005)

1.3 Plasmid Maintenance

1.3.1 High copy number plasmids

In order for plasmids to be stably maintained in the cell, replication must be followed by the distribution of the plasmids to the offspring. Mechanisms of plasmid maintenance vary according to the abundance of the plasmid in the cell. In the case of high copy number plasmids, random distribution is sufficient to ensure that at least one copy of the plasmid is inherited by both daughter cells. Replication control mechanisms will then ensure that the normal copy number of the plasmid is restored (Ghosh et al., 2006).

For any cell harbouring at least one plasmid copy, the probability of formation of plasmid-free cells is 50% (0.5). If the cell contains *n* number of plasmids, since each cell produces two daughter cells at cell division, the probability of plasmid loss (i.e. the probability of producing a plasmid-free daughter cell) (P_0) will be the following:

$$P_0 = 2(0.5)^n = 2^{(1-n)}$$ (Summers, 1991)

Therefore, the higher the number of copies of the plasmid, the lower is the probability of plasmid loss at cell division.

A major cause of plasmid instability is the formation of plasmid dimers or multimers due to homologous recombination during or after replication (Summers, 1996). Formation of multimers eventually leads to plasmid loss because in the cell multimers are maintained in low copy number compared to monomers (Summers and Sherratt, 1984). To avoid the propagation of multimers, plasmids have developed “resolution systems”, which use chromosomal or plasmid-encoded recombinases, to convert multimers into monomers (Thomas, 2000). For instance, the high copy number plasmid ColE1 harbours a 240 bp *cer* region where recombination takes place (Summers and Sherratt, 1984). This multimer resolution system requires the presence of four host-encoded proteins: ArgR, PepA and the recombinases XerC and XerD. Interestingly, the *cer* region contains a promoter regione P_{cer} that directs the expression of a short untranslated RNA region

called RCD (Regulator of Cell Division). RCD regulates the cell cycle in order to delay cell division until multimers have been resolved (Chant and Summers, 2007; Thomas, 2000).

Although random distribution and multimer resolution have been considered for a long time the only mechanisms that mediate high copy number plasmid maintenance, it was recently shown that other factors may contribute.

Pogliano *et al.* showed that high copy number plasmids (such as pUC19) are seen in clusters at midcell or at cell quarter positions, rather than diffusing freely in the cell. In addition, they observed that the plasmid foci migrate in opposite directions as the cell grows. Therefore, they proposed that the plasmid clusters could tether to receptors in the cells, possibly associated with the replication machinery (Pogliano *et al.,* 2001). It was later shown that the plasmid ColE1 also forms clusters, although they localise at cell poles. Again, no evidence of factors involved in the positioning of the plasmids was identified (Yao *et al.*, 2007).

In 2014, Reyes-Lamothe *et al.* demonstrated that ColE1 preferentially localises at the poles because the molecules are excluded from the nucleoid. They suggested that the plasmid clusters are not tethered to any cellular structure, instead they diffuse freely in the nucleoid-free area (Reyes-Lamothe et al., 2014).

Nucleoid exclusion has also been observed for the positioning at the cell poles of large protein complexes as well as protein aggregates (Vecchiarelli *et al.*, 2012; Neeli-Venkata *et al.*, 2016). The nucleoid works as a barrier for the passive diffusion of large molecules, which are confined at the cell poles. This confinement would therefore prevent plasmids from being lost at midcell positions during cell division (Reyes-Lamothe et al., 2014).

In general, high copy number plasmid segregation seems to be regulated somehow and it cannot be excluded that chromosomal factors may be involved in active partitioning (Million-Weaver and Camps, 2014).

1.3.2 Low copy number plasmids

Low copy number plasmids are generally larger in size and encode genes essential for their persistence. A very important feature of these plasmids is that they often harbour virulence factors or drug resistance genes (Pilla and Tang, 2018), hence understanding the mechanisms of maintenance is of fundamental priority.

Since stochastic diffusion would eventually cause plasmid loss, low copy number plasmids evolved strategies to be stably maintained at cell division. Multimer resolution, post-segregation killing and active segregation are some of the mechanisms that ensure plasmid stability. Multimer resolution has already been discussed above for high copy

number plasmids. Post-segregational killing will be described briefly, whereas active segregation will be the focus of the next sections.

1.3.2.1 Toxin-antitoxin systems

In the 1980s post-segregational killing through toxin-antitoxin (TA) or addiction systems was identified as a mechanism used by bacteria to ensure plasmid maintenance in a growing population (Gerdes *et al.*, 1986; Hiraga *et al.*, 1986).TA systems are composed of two elements: a toxin and an antitoxin, which neutralises the toxin by various mechanisms (Hayes, 2003b). By selectively killing plasmid-free cells, bacteria reduce the competition for nutrients in the environment and retain the plasmid within the population.

Toxin-antitoxin systems rely on the higher level of stability of the toxin, compared to the antitoxin. If a cell loses the plasmid, the antitoxin is quickly degraded, leaving the toxin free to target structures essential for the bacterium viability (Hayes and Van Melderen, 2011). Toxins are usually enzymes that often affect translation or DNA replication. Antitoxins can either be proteins or RNAs that bind directly to the toxin molecule or indirectly, repressing its gene transcription or translation. Four main types of TA systems have been identified and categorised according to the antitoxin mode of action (Harms *et al*., 2018).

In type I and type III, the antitoxin molecule is a RNA molecule that works by either targeting the toxin mRNA and inhibiting its translation (type I) or by direct binding to the toxin molecule, followed by inactivation (type III) (Harms *et al*., 2018).

In the *hok*/*sok* system of the R1 plasmid (type I), the *hok* gene encodes for a toxin molecule that interferes with the cell membrane and causes cell death (Hayes, 2003b). Translation of the *hok* gene is regulated by the antisense RNA *sok* product, which binds to the 5’ end of the *hok* mRNA and promotes its degradation by RNase III. Since Hok mRNA is much more stable than Sok RNA, if the cell loses the plasmid, Hok mRNA can be processed and translated, eventually leading to cell death (Harms *et al*., 2018) (Figure 1.1, top).

In type II and type IV the antitoxin is instead a protein which either forms a complex with the toxin and neutralises it (type II) or acts on the cell and counteracts the toxin effect (e.g. protecting the toxin’s target) (type IV) (Harms *et al*., 2018; Hayes and Van Melderen, 2011).

Many of the type II antitoxin proteins are composed of two domains, a DNA-binding and a toxin-binding region. Besides neutralising the toxin molecule by binding to it, the antitoxin protein also acts as a TA operon repressor, often aided by the toxin that works

as co-repressor (Page and Peti, 2016). An example of type II TA systems is represented by the *ccdA*/*ccdB* cassette of the F plasmid (Figure 1.1, bottom). In this system both CcdA (antitoxin) and CcdB (toxin) are proteins. CcdB kills the cell by acting on the topoisomerase II DNA gyrase. The short lived CcdA product can impair CcdB binding to the GyrA subunit, thus counteracting the toxin's activity. However, since CcdA is more susceptible to degradation than CcdB, when the plasmid is lost the toxin can affect the cell's vital function and kill it (Hayes, 2003b).

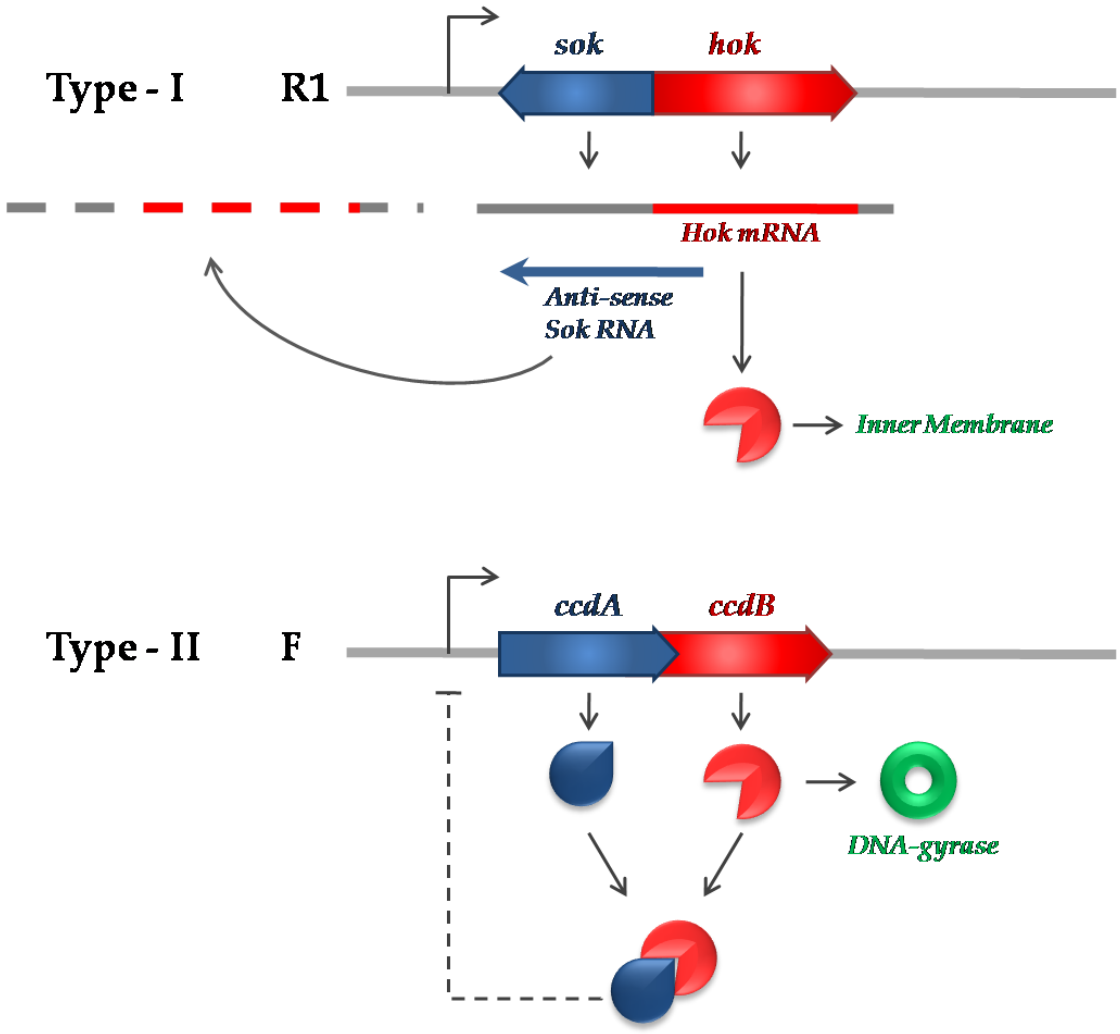

Figure 1.1 - Type I and Type II toxin-antitoxin systems

In the type I TA system, the toxin molecule is a protein while the antitoxin is a short antisense RNA. Sok RNA inhibits toxin's deleterious activity by targeting Hok mRNA leader region and enhancing its degradation by RNase III. In the type II TA system, both toxin and antitoxin molecules are proteins. In the ccdA/ccdB system, CcdA targets CcdB toxin preventing it from binding DNA-gyrase and impairing DNA replication.

1.3.2.1 Active segregation systems

Low copy number plasmids require active segregation systems to be stably maintained throughout generations. Partition cassettes (*par*) are ubiquitous in nature and are employed by bacteria and archaea to position plasmids as well as the chromosome on opposite poles of the cell prior to cell division (Vecchiarelli *et al.*, 2012; Baxter and Funnell, 2014).

Segregation or partition systems are encoded by the plasmid itself and are auto-regulated. In bacteria, they are characterised by three elements: a *cis*-acting or centromere-like site on the plasmid (*parS*), a DNA binding protein that binds specifically to the partition site, ParB, and a cytoskeletal NTPase protein that acts as a motor, ParA (Hayes and Barillà, 2006b; Gerdes *et al.*, 2010; Baxter and Funnell, 2014). The CBP (centromere-binding protein) ParB binds specifically to repeated sequences on the partition site, located upstream or downstream of the partition cassette, forming a nucleoprotein complex. ParA cannot directly interact with the plasmid. Instead it makes transient interaction with ParB which also stimulates its weak NTPase activity (Hayes and Barillà, 2006b; Gerdes *et al.*, 2010; Baxter and Funnell, 2014). Using the nucleotide (ATP/GTP) energy ParA drives the plasmids towards opposite poles of the dividing cell.

Despite sharing many common characteristics, different plasmids encode different partition cassettes. Segregation and replication machineries determine plasmid incompatibility, the inability to coexist in a cell in the absence of external pressure. The existence of multiple segregation systems allows cells to harbour more than one plasmid simultaneously, as well as the chromosome itself (Novick, 1987).

Segregation systems have been categorised according to the NTPase protein encoded by the partition cassette (Figure 1.2), since little or no sequence homology has been observed among different ParB proteins (Fothergill *et al.*, 2005).
The type I partition system is the most common. It has been found on both plasmids and chromosomes of some bacterial species (such as *Bacillus subtilis* and *Caulobacter crescentus*) (Ghosh *et al*., 2006) and it is characterised by a Walker-type ATPase. It can be further divided into two groups, type Ia (e.g. on F and P1 plasmids) and type Ib (e.g. on TP228 and pTAR plasmids), according to the size and characteristics of the ATPase proteins (Hayes and Barillà, 2006b).
Type II and type III partition systems are less abundant and encoded only by plasmids. They are characterised respectively by an actin-like ATPase (ParM), as on the plasmid R1, and by a tubulin-like GTPase (TubZ), as on the plasmid pBtoxis (Brooks and Hwang,

2017). The type IV partition system is instead composed of only one Par protein, as the one found on pSK1 plasmid (Simpson *et al.*,2003).

Recent studies also allowed ParB proteins to be categorised into two groups according to structural similarities of the DNA binding domain: Helix-turn-helix (HTH) (type Ia) or Ribon-Helix-Helix (RHH) (type Ib and type II) (Schumacher, 2008; Baxter and Funnell, 2014).

An important feature common to all partition cassettes is auto-regulation (Friedman and Austin, 1988; Jensen *et al.*, 1994; Carmelo *et al.*, 2005; Larsen *et al.*, 2007). Expression of the genes encoding the partition proteins needs to be tightly regulated as overexpression of one of the components causes plasmid instability (Friedman and Austin, 1988). Therefore, one of the Par proteins also works as a transcription repressor of the partition cassette. In type Ia, the regulatory role is played by the ATPase protein, ParA, while in type Ib, II and III by the centromere binding protein (Figure 1.3) (Friedman and Austin, 1988; Jensen*et al*.,1994; Carmelo *et al.*, 2005; Larsen *et al.*, 2007).

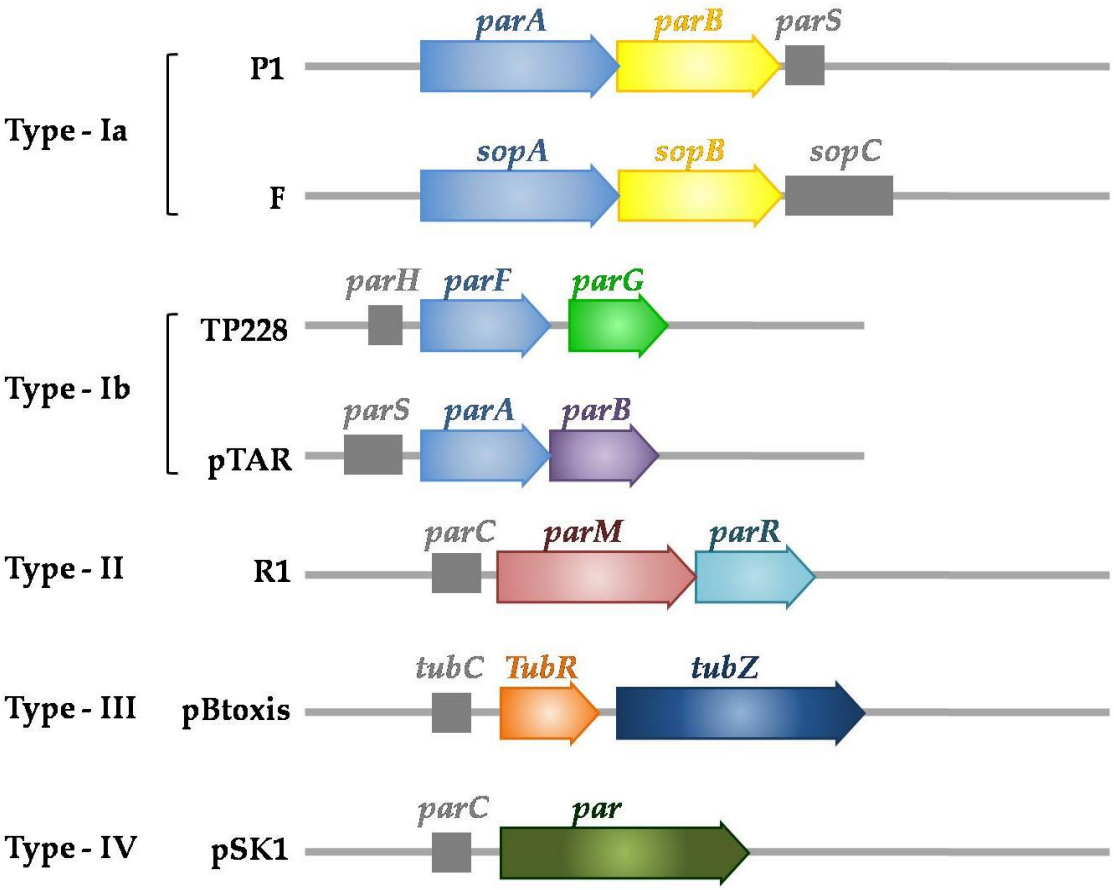

Figure 1.2 - Classification of plasmid partition cassettes

Plasmid segregation systems are categorised according to the NTPase protein encoded by the partition operon. For type I systems, the Walker type ATPase gene is shown in cyan and it is found upstream of the CBP gene (yellow in type Ia, green and violet in type Ib). In type Ia partition cassettes, the centromeric site is downstream of the partition operon, while in type Ib systems it is found upstream of the parA gene. ParM ATPase is shown in red, whereas TubZ GTPase is shown in blue. The plasmid pSK1 encodes one partition protein only, Par (dark green).

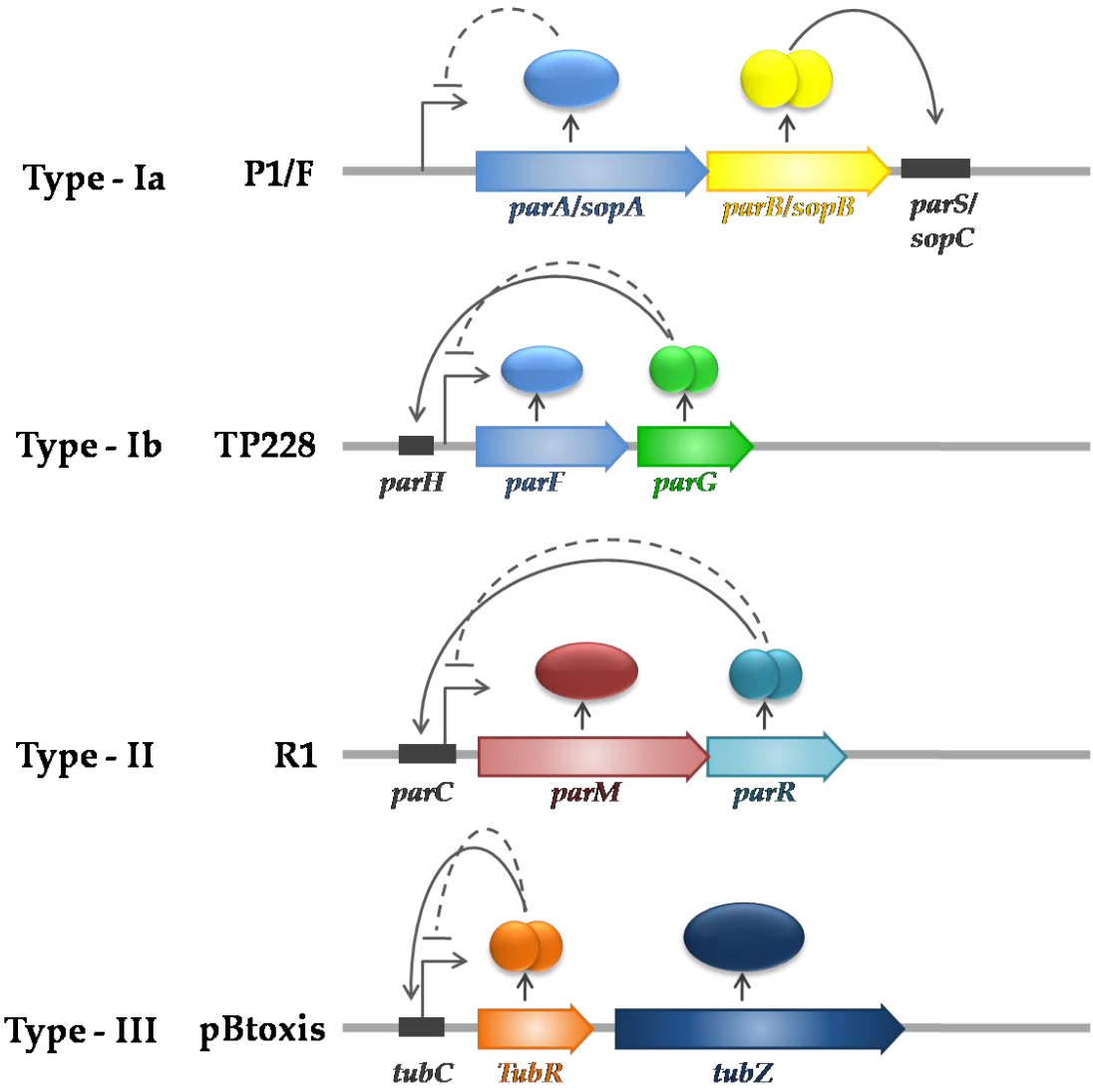

Figure 1.3 - Partition operon regulation by Par proteins

All the partition operons are tightly auto-regulated, because overexpression of one of the partition genes causes plasmid instability. In type Ia, the ATPase protein ParA works as repressor of the operon. In type Ib, II and III the CBP works as transcriptional regulator.

1.4 Type I partition systems

Type I partition systems are characterised by an ATPase protein with a deviant Walker-type motif or P-loop (phosphate-binding loop). The deviant Walker-type motif, KGGXXGKT, is characterised by two conserved lysine residues involved in the interaction with ATP. The second lysine residue is essential for binding and hydrolysing the ATP molecule, whereas the first lysine residue mediates protein dimerisation by interacting with the ATP bound to a second subunit. The cell division protein MinD also belongs to the Walker-type ATPase family, and shares many characteristics with type Ib ParA proteins (Lutkenhaus, 2012).

The gene encoding the ATPase protein is generally upstream of the CBP gene. Instead, the centromeric region is located either downstream (type Ia) or upstream (type Ib) of the partition cassette.
Type Ia partition proteins are usually larger than those of type Ib. ParB size ranges from 182-336 residues in type Ia to 46-113 residues in type Ib. As for ParA, its size ranges from 251-420 residues in type Ia to 208-227 residues in type Ib (Schumacher, 2007).

Recent structural information revealed that ParB proteins belonging to type Ia have a different fold compared to type Ib CBPs (Schumacher, 2011). Structures of type Ib CBP ParG (TP228), ω (pSM19035) and ParB (pCXC100) shows that they all share a similar RHH fold (Huang *et al.*, 2011; Golovanov *et al.*, 2003; De La Hoz *et al.*, 2004). Instead, type Ia CBPs such as P1 ParB and RP4 KorB, have a more complex structure with three distinct domains: a flexible N-terminal domain that interacts with the partner protein, a conserved central HTH DNA binding domain and a C-terminal dimerisation domain (Schumacher*et al*., 2010; Schumacher, 2011).

As far as ParA is concerned, the difference in size between type Ia and Ib is mainly due to the presence of a N-terminal HTH motif which is absent in type Ib ATPases. This region is important for DNA recognition, as type Ia ParA proteins also work as transcription regulators of the partition operon. This role is performed by the centromere binding protein ParB in type Ib systems (Schumacher, 2008, 2011).

Although Type Ib ParA proteins lack the HTH motif, they can interact non-specifically with the DNA. Non-specific DNA binding is an essential feature shared by all type I ATPases, as the nucleoid is used as scaffold for plasmid movement inside the cell (Vecchiarelli *et al.*, 2012).

1.4.1 Type Ia partition system of P1 plasmid

One of the best characterised segregation systems is that of the *E. coli* prophage P1. In 1983 Austin and Abeles discovered that the *par* site is a 2.5 kb region located next to the replicon of the plasmid. They argued that, although adjacent, the *par* and *rep* regions are separable, since *rep* could be replaced without affecting plasmid stability (Austin and Abeles, 1983).

The partition cassette of the P1 plasmid is composed of two *trans*-acting elements, *parA* and *parB*, and a *cis*-acting site, *parS* (Austin and Abeles, 1983; Martin *et al*, 1987). In addition to the essential *par* components, IHF was also identified as an important, although dispensable, element for P1 plasmid segregation (Funnell, 1988b).

The centromeric site *parS* is an 84 bp region located downstream of the *parAB* operon and is composed of six repeated sequences, four heptameric motifs called Box A and two hexameric motifs named Box B, both recognised specifically by ParB dimers (Davis *et al.*, 1990; Funnells, 1991; Funnells and Gagnier, 1993). A 29 bp IHF-binding region is located in the middle of the *parS* site, dividing the centromere into two asymmetrical arms (Funnell, 1988b). It was shown that ParB binds with higher affinity to the right side. However, when IHF is present, both sites are equally occupied by ParB. The presence of IHF improves ParB binding to *parS* 10,000 fold (Funnell, 1991). The mechanism by which IHF enhances ParB assembly onto the *parS* site is by bending the DNA and bringing the two arms into close proximity (Figure 1.4A) (Funnell and Gagnier, 1994; Funnell, 1991).In this way, ParB can interact with distant regions of the DNA, forming a loop structure where ParB spans the arms of the *parS* site (Figure 1.4B) (Hayes and Austin, 1994; Bouet *et al.*, 2000).Together with IHF binding, plasmid intrinsic supercoiling was also shown to improve ParB affinity for *parS* and to promote accurate positioning of the protein onto the centromere (Funnell, 1991). First, ParB binds to the IHF-bound *parS* as a dimer. Then, other dimers are cooperatively recruited to the site forming a larger complex (Bouet *et al.*, 2000). Based on the above observations, it was hypothesised that *parS* wraps around the IHF-ParB complex forming a nucleoprotein superstructure (Funnell and Gagnier, 1994; Hayes and Austin, 1994, 1993).

ParB is a DNA binding protein that binds to *parS* as an asymmetrical dimer (Funnell, 1991). ParB monomers are characterised by an all-helical N-terminal domain, comprising the HTH motif, and a C-terminal dimerisation domain, consisting of three anti-parallel β-strands and a final α-helix. Both domains can interact with the DNA. The HTH motif was shown to bind specifically to A-boxes, while the dimerisation domain interacts specifically

with the B-boxes. The two domains are connected by a flexible loop that allows ParB to bind to *parS* in different orientations and bridge the *parS* arms together (Schumacher and Funnell, 2005). In addition, ParB can also interact with non-specific DNA, thus spreading to neighbouring regions (up to 500 bp upstream and downstream of *parS*) and condensing the DNA (Schumacher *et al.*, 2007b).

Besides binding to the DNA, the flexible N-terminal domain of ParB is also involved in the interaction with the partner protein ParA (Surtees and Funnell, 1999). ParA is a Walker-type ATPase, whose weak ATP-hydrolysing activity is enhanced by the interaction with ParB and non-specific DNA (Davis *et al.*, 1992). The protein structure can be divided into three regions: an N-terminal α-helix dimerisation domain, a central HTH DNA-binding motif, and a large C-terminal domain containing the deviant Walker-type motif that is essential for nucleotide and ParB binding (Dunham *et al.*, 2009; Vecchiarelli *et al.,* 2013). From the structure it is clear that ParA performs multiple roles in the segregation process. Besides being the main motor protein driving plasmids apart, it works as a transcriptional repressor of the *parAB* operon, specifically binding to a 40 bp sequence within the operator site in a ADP-dependent manner (Davis *et al.*, 1992). ParB can enhance ParA regulatory activity and work as co-repressor (Friedman and Austin, 1988). Both ATP and ADP promote ParA dimerisation and interaction with the DNA, while ATP hydrolysis counteracts DNA binding (Davey and Funnell, 1994, 1997).In addition, ATP was shown to enhance ParA polymer formation *in vitro* (Dunham et al., 2009).

ParA switches between regulatory and motor function by binding to different adenine nucleotides. While ADP binding greatly enhances ParA regulatory activity, ATP binding activates ParA for plasmid segregation, promoting its binding to ParB (Dunham *et al*., 2009; Bouet *et al*., 1999). ATP binding causes a conformational change in ParA structure, essential for the protein to make non-specific interaction with the DNA *in vitro* and the nucleoid *in vivo* (Vecchiarelli *et al.*, 2010).It is now well established that ParA drives plasmids apart by using the nucleoid as a scaffold (Vecchiarelli *et al*., 2012).

Although P1 plasmid segregation mechanism is still under investigation, several models have been proposed over the years. Fluorescence microscopy experiments showed that early after replication, sister plasmids are visible at midcell position and then driven apart by the segregation machinery and located at one quarter and three quarter of the cell length (Gordon *et al*., 1997). Hatano and Niki further proved this hypothesis by fluorescently labelling ParA. They saw that ParA-YFP (yellow fluorescent protein) co-

localises with the nucleoid, with intense foci at midcell and cell quarter positions. In addition, they saw that in *parB* and *parS* deletion mutants ParA foci are not visible, suggesting that ParA localisation depends on the presence of the *parS*-ParB complex. The dependence of ParA localisation on ATP hydrolysis was also tested. It was shown that when ParA ATPase activity is impaired, foci do not form. Furthermore, differently from other Walker-type ATPases, no filament structure was observed *in vivo* for P1 ParA. Filaments were only observed *in vitro* when high concentration of ParA were used (Dunham *et al*., 2009). Based on these observations, Hatano and Niki suggested that P1 plasmids move by following a ParA-ATP gradient on the nucleoid, without the formation of polymers that would push or pull the plasmids at opposite sides of the cell (Hatano and Niki, 2010). Microscopical investigation, combined with further biochemical characterisation led to the proposal of a diffusion-ratchet mechanism for the segregation of P1 and type I plasmids. This will be discussed in the next sections.

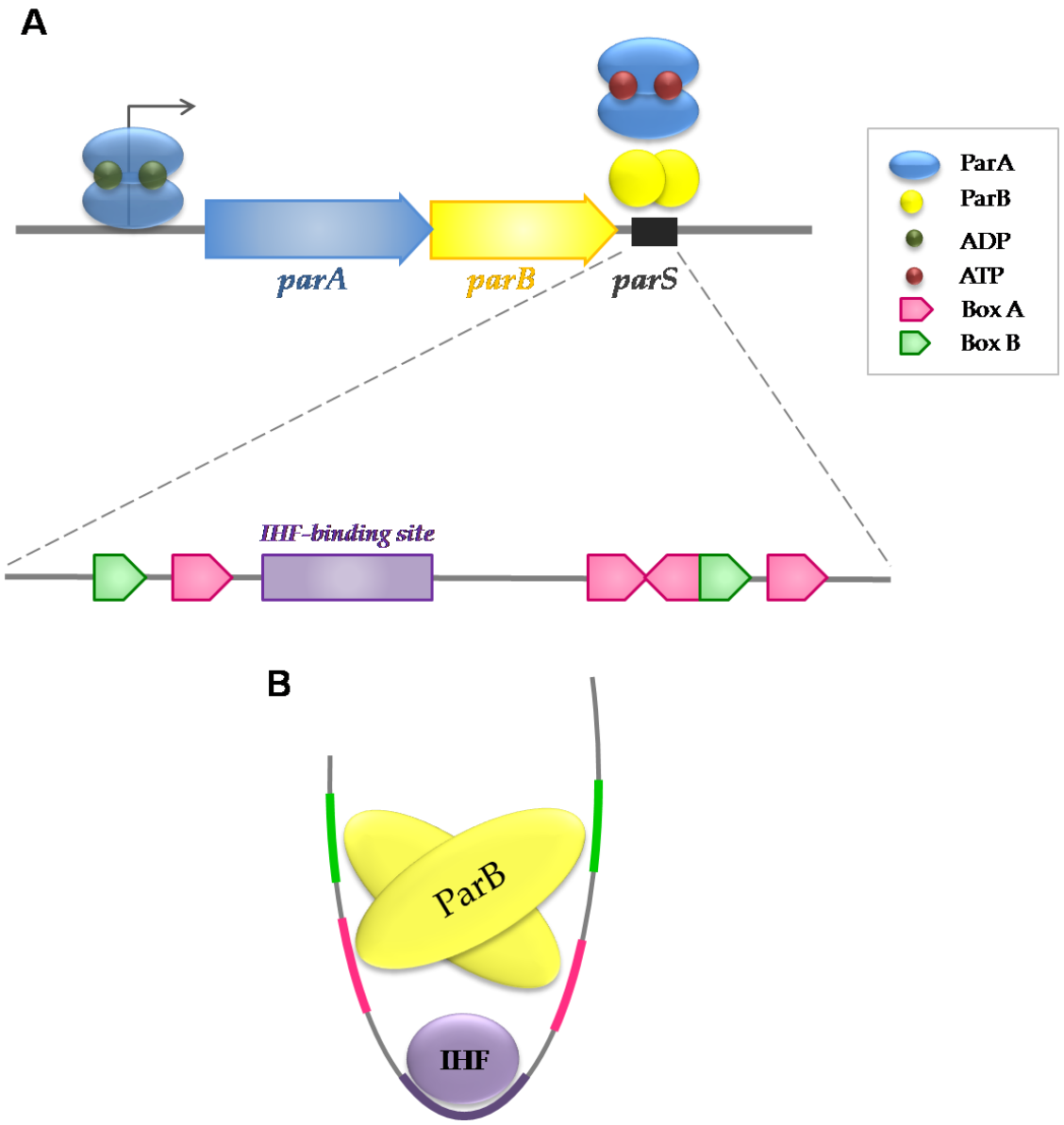

Figure 1.4 - Organisation of the P1 plasmid ParABS partition system and IHF-mediated *parS* recognition by ParB

A) The *parABS* partition cassette encodes for the Walker-type ATPase ParA (blue), the CBP ParB (yellow) and contains the *parS* site downstream of the partition operon (dark grey). ParA can bind ATP (red) or ADP (green). When bound to ADP, ParA works as transcriptional repressor of the partition operon by binding to the *parAB* operator site. When bound to ATP, ParA can recognise ParB, as well as bind to the nucleoid. ParB binds to *parS* in a sequence-specific manner, recognising both 7 bp Box-A motifs (pink) and 6bp Box-B motifs (green).

B) IHF (purple) enhances ParB binding to the centromere by bending the DNA and bringing in close proximity the right and the left arms of the *parS* site. Once a ParB dimer recognises the cognate site, other ParB dimers are recruited forming a nucleoprotein superstructure.

1.4.2 Type Ia partition system of F plasmid

A second archetype of the type Ia segregation system is that of the F plasmid, which in *E. coli* is maintained in one or two copies per cell. The conjugative plasmid was shown to encode a 2.5 kb *par* region next to the replicon, responsible for plasmid stability (Austin and Wierzbickj, 1983; Austin and Abeles, 1983; Ogura and Hiraga, 1983). Similarly to the P1 plasmid, the partition cassette of the F plasmid is characterised by three components: two *trans*-acting elements, *sopA,* encoding a Walker-type ATPase and *sopB*, encoding a DNA-binding protein, and a *cis*-acting element, *sopC*, downstream of the partition operon *(Ogura and Hiraga, 1983; Lane et al.*, 1987) (Figure 1.5).

The centromeric site *sopC* is composed of twelve 43 bp direct repeats, which are recognised specifically by SopB (Hayakawa *et al.*, 1985; Mori *et al.*, 1989). SopB interacts as a dimer with a 16 bp palindromic sequence within the 43 bp region (Mori *et al*., 1989; Hanai *et al.*, 1996). Different studies showed that a single repeat is sufficient to ensure plasmid segregation (Lane *et al.*, 1987; Biek and Shi, 1994). It was proposed that *sopC* repeats can work as nucleation sites, from which SopB can spread on the surrounding DNA, independently of the sequence (Lynch and Wang, 1995). Preliminary studies on SopB-*sopC* complex formation also suggested that the protein could introduce positive supercoiling in the DNA (Biek and Strings, 1995). It was hypothesised that the centromeric site could wrap around the SopB protein core, in a right-handed helix (Lynch and Wang, 1994, 1995). This hypothesis was later disproved by further biochemical and structural investigation on the nucleoprotein complex. Bouet and Lane demonstrated that while SopB binding does not affect DNA topology, the protein does nucleate on the *sopC* site and spread upstream and downstream of the partition site, coating the DNA (Bouet and Lane, 2009). Furthermore, the crystal structure of SopB in complex with its specific binding site showed that SopB does not distort or bend the DNA, however it makes both specific and non-specific contacts with it. While specific binding allows SopB to recognise its cognate DNA site, non-specific interactions are essential for the protein to coat the DNA as well as to spread *in trans* on distant DNA strands (Schumacher *et al.*, 2010; Schumacher, 2012).

Similarly to P1 ParB, the SopB structure is characterised by three distinct domains connected by flexible linkers. The N-terminal domain is involved in the interaction with the partner protein SopA, while the C-terminal domain is essential for protein dimerisation. The central domain is characterised by an HTH motif, which makes the primary interactions with the specific DNA site. The SopB crystal structure in complex with the centromeric site showed that a second dimerisation domain is also present

within the central region. This would allow the protein to bridge distant DNA duplexes (Schumacher *et al.*, 2010).

SopA shares many characteristics with P1 ParA. It can interact specifically with the DNA through the conserved HTH N-terminal domain and it works as a transcriptional repressor of the *sopAB* operon by binding to four 6 bp repeated sequences in the operator site. SopA regulatory activity is further enhanced by the presence of SopB, which acts as co-repressor of the partition operon (Mori *et al*., 1989).

In addition to its regulatory role, SopA works as a motor protein for the segregation of the F plasmid. The protein does not bind directly to its cargo. Instead, it interacts with SopB, which works as an adaptor between the two (Ravin *et al.*, 2003). SopB can stimulate SopA's weak ATPase activity by inserting an arginine finger motif (Arg-36) into the ATP-binding domain of the partner protein (Ah-Seng *et al*., 2009). Stimulation of ATPase activity, however, requires the presence of non-specific DNA. Interaction with non-specific DNA is extremely important for SopA activity. Non-specific DNA alone stimulates SopA ATPase activity very weakly, while SopB alone lacks this ability (Watanabe *et al*., 1992). Maximum SopA ATPase activity stimulation was recorded when both non-specific DNA and *sopC*-bound SopB were present (Castaing *et al.*, 2008; Ah-Seng *et al.*, 2009).

SopA can bind to ATP and ADP through the Walker motif. Binding of different adenine nucleotides causes conformational changes in the protein structure, important for SopA activity (Libante *et al.*, 2001). Indeed, SopA was shown to interact with the operator site in a ADP-bound form, and to non-specific DNA in a ATP-bound form (Bouet *et al.*, 2007). In addition, the presence of ATP was shown to cause SopA to polymerise into higher order structures, while ADP and ATPγS binding did not seem to have the same effect. ATP hydrolysis was shown not to be important for polymer formation. Instead, *in vitro* experiments showed that the presence of *sopC* and non-specific DNA counteracts SopA polymerisation, while the presence of SopB can reverse this effect by coating the DNA (Bouet *et al.,* 2007).

Microscopy experiments by Gordon *et al*. showed that, as for the P1 plasmid, F localises either at midcell or at cell quarter positions (Gordon *et al.*, 1997). In addition, it was shown that SopA oscillates between the cell poles and forms filaments both *in vivo* and *in vitro*, while SopB forms foci where SopA co-localises. Initially, the dynamic polymerisation and

depolymerisation of SopA filaments was identified as responsible for plasmid movement inside the cell (Lim *et al.,* 2005; Hatano *et al.*, 2007).

Although originally a pushing/pulling mechanism based on polymer formation and dissociation through the ATP/ADP cycle was proposed for the segregation of the F plasmid, it was later refuted. As for the P1 plasmid, rather than polymers, a dynamic SopA gradient over the nucleoid is believed to be the driving force moving sister plasmids apart. This will be discussed in the next sections.

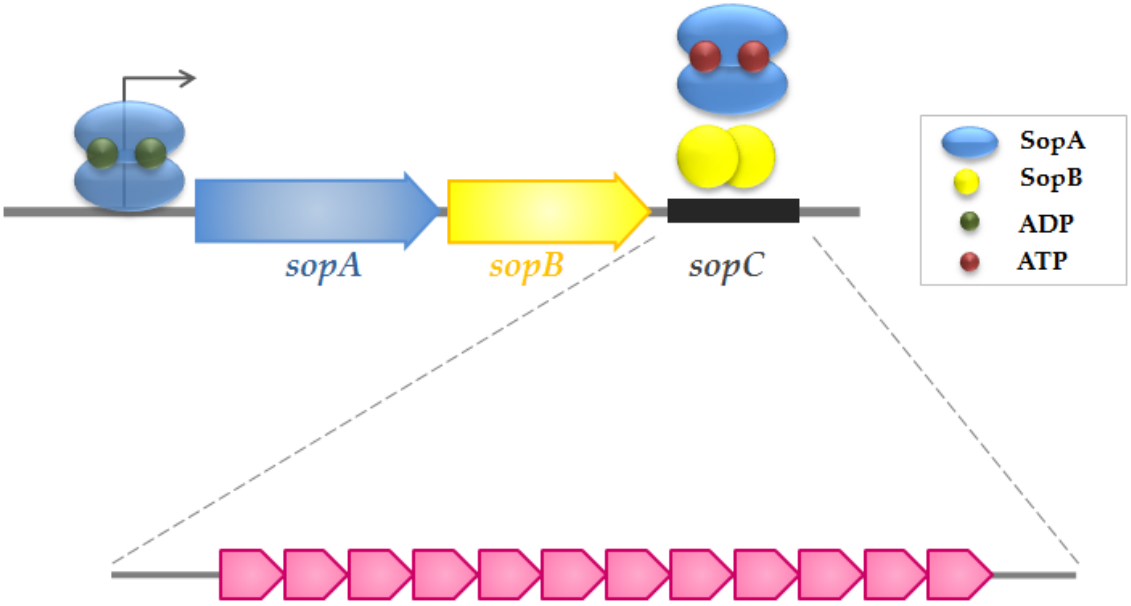

Figure 1.5 - Organisation of F plasmid SopABC partition system

The *sopABC* partition cassette encodes for the Walker-type ATPase SopA (blue), SopB (yellow) and includes the centromeric sitesopC (dark grey). SopA can bind ATP (red) or ADP (green). When bound to ADP, SopA works as transcriptional repressor of the partition operon by binding to the *sopAB* operator site. When bound to ATP, SopA can recognise SopB, as well as bind to the nucleoid.*sopC* is composed of twelve 43 bp repeated sequences (pink). SopB dimers bind to the centromeric site by recognising a 16 bp palindromic region within the 43 bp repeated sequences.

1.4.3 Type Ib segregation system (*par2*) of pB171 plasmid

The *E. coli* virulence plasmid pB171 is characterised by two adjacent partition cassettes, *par1* (type II) and *par2* (type Ib). Although each of the partition systems is sufficient for plasmid segregation, optimal plasmid stability is achieved when both cassettes are present. *par1* encodes for two proteins homologous to ParM and ParR of the plasmid R, while *par2* encodes for ParA, a Walker type ATPase, and ParB, a small DNA binding protein (Ebersbach and Gerdes, 2001). Two centromeric sites are present on the plasmid: one between *par1* and *par2* (*parC1*) and one downstream of *par2 (parC2)*. Both of them are characterised by 6bp degenerate repeats of two types: class I and class II. *parC1* is composed of 17 repeats organised into two clusters, while *parC2* is composed of 18 repeats organised in three clusters (Figure 1.6). The centromeric site *parC1* is shared by the two partition systems (Ebersbach and Gerdes, 2001; Ringgaard *et al.*, 2007).

Dimeric ParB binds in a sequence-specific manner to both *parC1* and *parC2* (Ebersbach and Gerdes, 2001; Fothergill *et al.*, 2005). In addition, ParB works as a transcriptional repressor of both *par1* and *par2*, demonstrating cross-talk between the two partition systems. Furthermore, it is possible that IHF could be involved in pB171 plasmid segregation. Two IHF-binding sites were identified within *par2*, one inside the *parA* sequence and one at the end of *parB*. *In vitro,* binding was observed between IHF and both the *ihf* sites (Ringgaard *et al.*, 2007). The interaction between ParB and its cognate sites could therefore produce a large nucleoprotein structure involving the host factor.

ParA belongs to type Ib class of ParA Walker type ATPases, as it shows similarities to TP228 ParF and pTAR ParA (Ebersbach and Gerdes, 2001). Similarly to other ParA proteins, ParA also oscillates over the nucleoid. The oscillatory pattern was shown to be dependent on the presence of the *parC*-ParB complex and on ATP binding, as mutations in the Walker-A motif impair ParA oscillation (Ebersbach and Gerdes, 2001). Plasmids harbouring a *par2* site were seen to localise either at midcell or at quarter cell positions. ParA was shown to form dynamic spiral-shaped structures on the nucleoid, that do not relocate in the absence of the rest of the partition complex (Ebersbach and Gerdes, 2004). Formation of filaments was observed both *in vivo* and *in vitro* and was shown to be enhanced by ATP binding (Ebersbach *et al.*, 2006). Indeed, ParA was shown to polymerise on the nucleoid and to shrink upon interaction with the ParB-bound plasmid. ParB interaction with ParA promotes ATP hydrolysis and causes ParA polymers to dissociate into monomers and detach from the nucleoid. ParB was hypothesised to enhance ParA ATP hydrolysis by inserting the Arg-26 into the ATP-binding pocket, in a

similar way as ParG from TP228 plasmid inserts Arg-19 into ParF's ATP-binding pocket (Barillà *et al.* 2007; Ringgaard *et al.*, 2009). In fact, mutation of ParB Arg-26 was seen to dramatically affect plasmid stability and to impair ParA ability to dynamically relocate on the nucleoid. A pulling mechanism based on polymerisation and depolymerisation of ParA over the nucleoid was proposed. This will be discussed in the next sections.

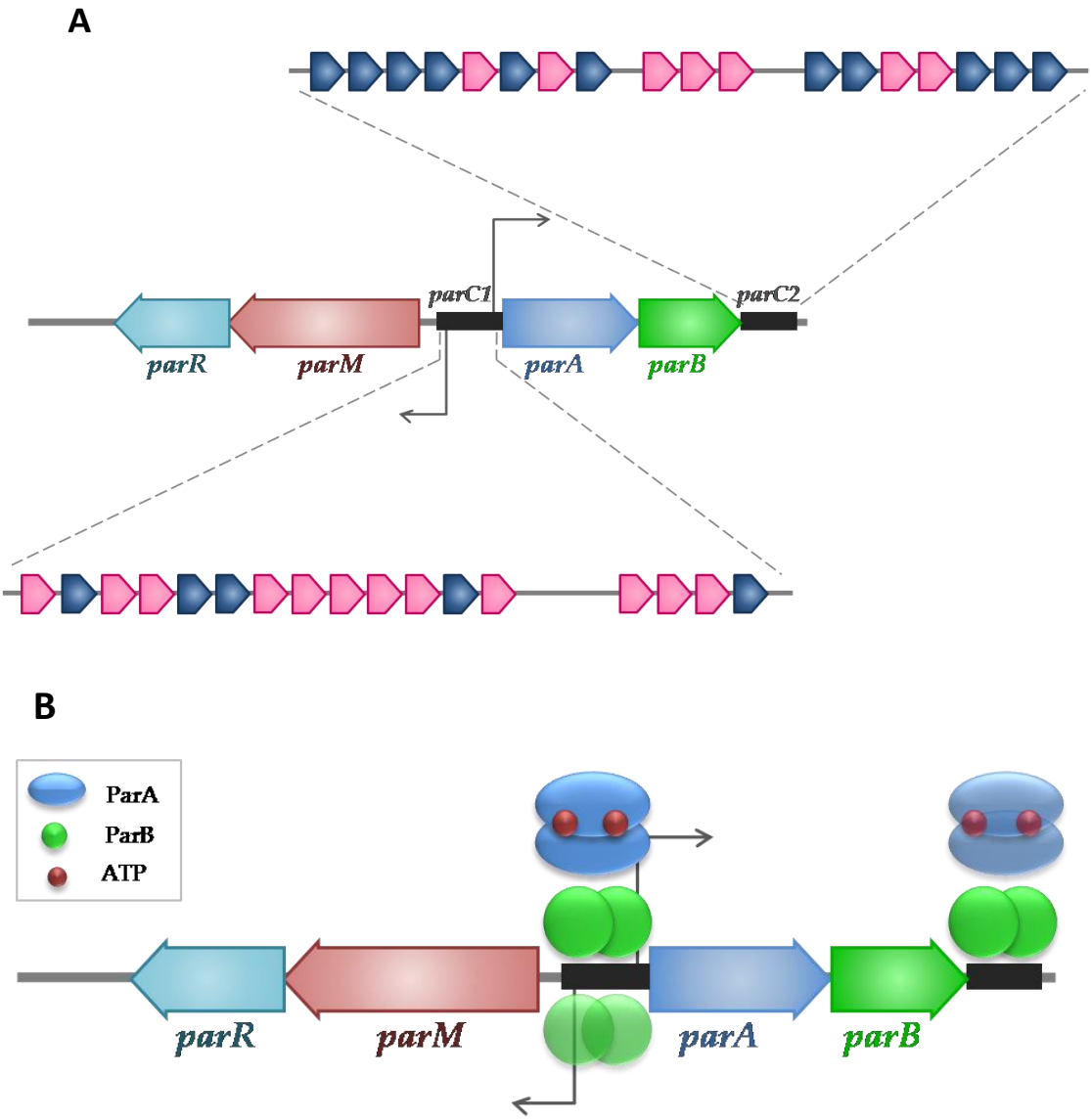

Figure 1.6 - Organisation of pB171 *par2* ParABC1C2 partition systems

A) pB171 encodes two partition cassettes: *par1* (*parRMC1*) and *par2* (*parABC1-2*). *par2* encodes for the Walker-type ATPase ParA (blue) and the DNA-binding protein ParB. The partition site *parC1* and *parC2* are composed of degenerate repeats of two types: class I (pink) and class II (blue)

B) *parC1* and *parC2* are recognised by dimeric ParB as centromeric regions for assembly of the partition complex. ParB also works as transcriptional repressor of both *par1* and *par2* operons. ParA-ATP on the nucleoid binds to ParB on the partition sites and drives plasmid segregation.

1.4.4 Type Ib segregation system of TP228 plasmid

The TP228 conjugative plasmid is a member of a group of plasmids that are evolutionarily distinct from the P1 prototype (Jones *et al.*, 1993). It was first identified in *Salmonella newport* and confers resistance to several antibiotics, such as kanamycin, neomycin, spectinomycin, streptomycin, sulphonamides and tetracycline as well as mercuric ions (Hayes, 2002).

The plasmid is very stable *in vivo* and no detectable loss was recorded after 25 generations of unselected growth in *E. coli*. Segregation stability was shown to be conferred by the type Ib *parFGH* partition cassette, which encodes a Walker-type ATPase, ParF, a DNA-binding protein, ParG, and harbours a centromeric site upstream of the partition operon, *parH*. The interactions among the partition system components are the focus of this work.

Although still under investigation, the segregation system of the plasmid TP228 shares many characteristics with other type I systems. The ATPase protein ParF is responsible for positioning the plasmids at opposite poles of the cell, using the nucleoid as scaffold and ATP hydrolysis energy (Vecchiarelli *et al.*, 2012;McLeod *et al.*, 2017). ParF does not interact directly with the plasmid. Instead it binds to its partner ParG, which recognises specifically *parH* and thus works as an adaptor between the plasmid cargo and the motor protein. A Venus-fly trap model has been recently proposed for the segregation of the plasmid(McLeod *et al.*, 2017). This will be discussed further in the next sections.

1.4.4.1 Centromeric site *parH*

The centromeric site *parH* is located upstream of the *parFG* operator site (O_F). O_F is composed of eight variant 5'– ACTC –3' 4bp motifs, either in direct or inverted orientation, interspaced by 4 bp AT-rich sequences (Zampini *et al.,* 2009a). Similarly to the operator site, *parH* is composed of twelve 4bp degenerate sequences of the 5'– ACTC –3' motif. The repeats are interspaced by 4 bp AT-rich sequences and follow a direct orientation, apart from one in the middle of the sequence (Figure 1.7A) (Wu *et al.*, 2011).

ParG works as a partition protein and a transcriptional repressor of the *parFG* operon (Carmelo *et al*., 2005). The protein binds to both O_F and *parH* as a dimer (Barillà and Hayes, 2003) (Figure 1.7B). Each ParG dimer recognises a single 5'– ACTC –3' motif. Therefore, eight ParG dimers coat the operator site for transcription regulation (Zampini *et al.,* 2009). In a similar way, twelve ParG dimers cooperatively bind to *parH* for the assembly of the segrosome. Interestingly, no binding is detected when ParG reacts with

DNA harbouring a single repeat. Two repeats is the minimum required to detect ParG binding *in vitro* (Wu *et al.*, 2011; Zampini *et al.*, 2009).

Deletion of the operator site showed no effect on plasmid stability and nor did replacing *parH* with O_F. This suggests that the operator region could also be part of the plasmid centromere. Furthermore, progressive deletion of *parH* repeats causes gradual reduction in plasmid stability. However, change in orientation and insertion of bases affected plasmid retention only slightly (Wu *et al*., 2011). It was shown that eight repeats are the minimum required for plasmid stability, although the orientation is not extremely important. The partition site seems to lack intrinsic curvature. However, the fact that rearrangment of the repeats had only minor effects on plasmid stability suggests that *parH* is highly elastic. This suggests that the DNA could form different structures during partition (Wu *et al.*, 2011).

1.4.4.2 Centromere binding protein ParG

ParG is a homodimeric DNA-binding protein, that interacts with the region upstream of the *parFG* operon in a sequence specific manner (Barillà and Hayes, 2003). During plasmid partitioning, ParG works as an adaptor between the plasmid and the motor protein ParF, directly binding to it and enhancing its weak ATPase activity (up to 30 fold) (Barillà and Hayes, 2003; Barillà *et al.*, 2005) .

ParG is small in size, 8.6 kDa (76 amino acids), and is mostly found as a dimer (Barillà and Hayes, 2003). ParG monomers are characterised by a ribbon-helix-helix C-terminal domain (residues 33-76) and an unstructured N-terminal domain (residues 1-32) (Figure1.8). Structure alignment showed that ParG belongs to the Arc/MetJ superfamily of transcriptional repressors, as it shares a similar RHH fold (Golovanov *et al.,* 2003).

In the dimeric structure, C-terminal domains are intertwined to form a double-stranded antiparallel β-structure with four α-helices on one site. The protein is predicted to interact with the DNA by inserting the double stranded β-sheet into the major groove of the DNA (Golovanov *et al.,* 2003). ParG C-terminal domain is essential for DNA interaction and dimerisation (Carmelo *et al.*, 2005; Golovanov *et al.,* 2003). The core structure of the dimer is held together by hydrophobic interactions. Many residues have been shown to be fundamental to maintain the dimeric state of the protein, such as Lys-45, Arg-48, Phe-49, Val-64, Leu-67, Val-68, Trp-71 Leu-72. Among those, mutations of Phe-49 in Trp-71 and Leu-72 were shown to be the most deleterious for plasmid segregation (Golovanov *et al.*, 2003; Saeed *et al.*, 2015).

NMR investigation showed that the ParG N-terminal tails are unstructured and flexible. The first 6 to 10 amino acids show the highest level of flexibility, while the residues 17-23 show limited flexibility, possibly because of the formation of a transient secondary structure or contacts with the folded domain (Golovanov *et al*., 2003). It is very likely that amino acids 20-30 interact with the folded C-terminal domain, as full N-terminal tail deletion affects a significant number of residues in the β-strand and the first α-helix of the folded domain. Indeed, NMR studies showed that residues 23-29 could form an additional anti-parallel β-strand (Carmelo *et al*., 2005).

Comparison of the HSQC spectra of ParG in the absence and presence of part of the operator site allowed the identification of residues that are or could be involved in the interaction with site-specific DNA. Upon O_F addition, the ParG spectrum was seen to change. The C-terminal domain, known to be involved in DNA binding, completely disappeared because of association with the DNA. On the other hand, the chemical shift of the amino acids of the N-terminal domain changed to different extents, according to the residue position. The most affected residues were the following: His-7, Lys-11, Lys-12, Met- 13, Asn-18, Arg-19, Arg-23, Thr-26, Ala-27, and Val-29. Lys-11 and Lys-12 show the largest changes in chemical shift. In this case, the change in chemical shift was hypothesised to be due to weak and transient interaction with the DNA. Interestingly, addition of non-specific DNA caused no change in the chemical shifts of the N-terminal domain (Carmelo *et al.*, 2005).

Partial and full deletion of the N-terminal tail affected ParG binding to the cognate DNA site as well as the ability to enhance ParF ATPase activity (Barillà *et al.*, 2007; Wu *et al.*, 2011). On the other hand, deletions did not affect ParG binding to ParF or dimerisation *in vivo* (Carmelo *et al.*, 2005; Barillà *et al.*, 2007).
Analogously to arginine finger residues in RasGAPs, Arg-19 in ParG was identified as the arginine finger responsible for promoting ATP hydrolysis by ParF. The arginine finger was predicted to stimulate ATP hydrolysis by inserting its side chain into the catalytic cleft of ParF. The residue stabilizes the transition state through the neutralization of the negative charges developed during the phosphoryl transfer reactions. This catalytic mechanism is commonly used by P-loop ATPases, which often require *trans*-acting elements to neutralise the negative charges developed during catalysis (Golovanov *et al.*, 2003; Barillà *et al.*, 2007). Changing Arg-19 into either alanine or lysine drastically reduced plasmid retention (Barillà *et al.*, 2007).

Progressive deletion of the N-terminal tail was seen to increase ParG interaction with site-specific DNA, while it also improved the non-specific component of the binding. The tail was shown to modulate the interaction with *parH* and reduce the non-specific binding (Wu *et al.*, 2011). In addition, the N-terminus is involved in ParG transcription regulation role, since deletion of the domain severely impairs ParG repressor activity *(Carmelo et al.,* 2005).

Besides enhancing ParF ATPase activity, ParG is also involved in ParF higher-order structures formation. *In vitro* experiments showed that ParF forms polymers in the presence of ATP and its polymerisation rate increases when ParG is added to ParF-ATP (Barillà and Hayes, 2003). It was hypothesised that the tips of the ParG flexible tails could act as a sticky tentacles and bundle ParF oligomers together, forming fibres (Barge, 2015).

Furthermore, ParG oligomerises in the presence of both site-specific and non-specific DNA fragments (Carmelo *et al.*, 2005; Golovanov *et al.,* 2003). It was hypothesised that ParG oligomerisation on the DNA is physiologically relevant, as it could provide the strength necessary to hold the plasmid during transport inside the cell (Carmelo *et al.*, 2005).

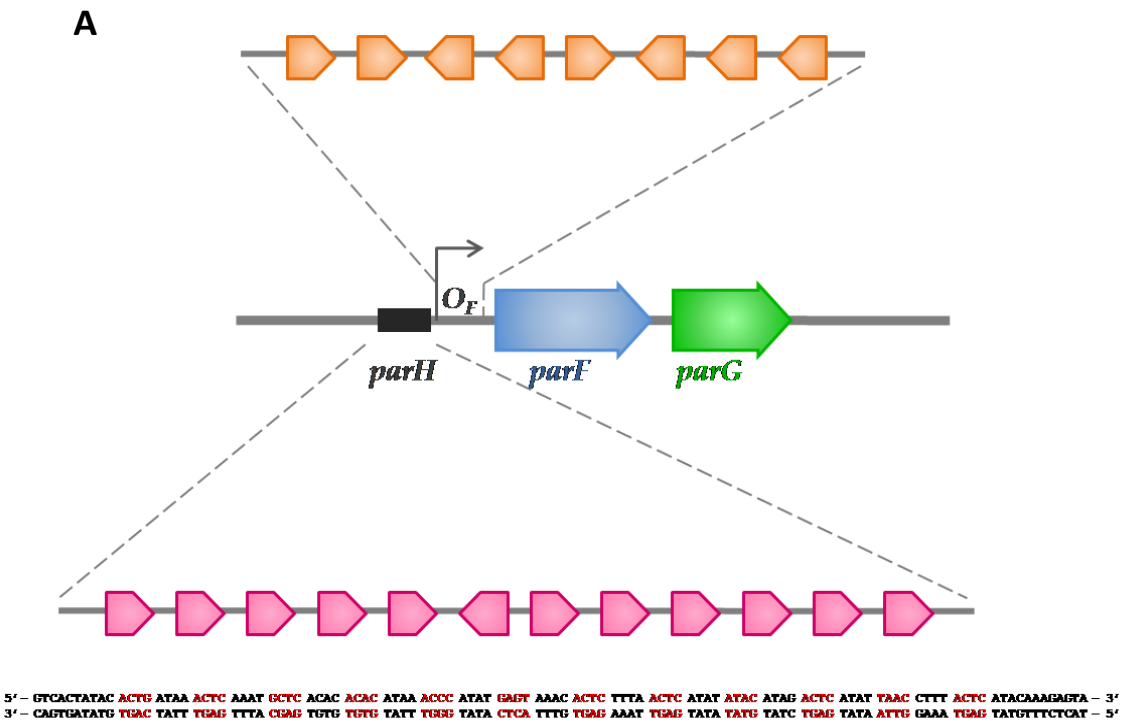

Figure 1.7 - Organisation of TP228 ParFGH partition system

A) *parFGH* partition cassette encodes for a Walker-type ATPase, ParF (blue), and a DNA binding protein, ParG (green). The centromeric site, *parH* is located downstream from the *parFG* operon and is composed of twelve degenerate 4bp repeats (pink) interspaced with 4bp AT-rich sequences. The *parFG* operator site (O_F) has a similar organisation, although it is composed of only eight repeated sequences (orange). The sequence shown represents *parH*. The nucleotides in red represent the 4bp motif recognised by ParG. Upon saturation, twelve ParG dimers bind to *parH*.

B) ParG works as transcriptional repressor of the *parFG* operon, by binding to the O_F operator site. ParG also works as an adaptor between the plasmid and ParF by binding to the centromeric site *parH.* ParF binds to the partition complex in the ATP-bound form. ParF-ATP also binds non-specifically to the DNA and drives plasmids apart using the nucleoid as a scaffold.

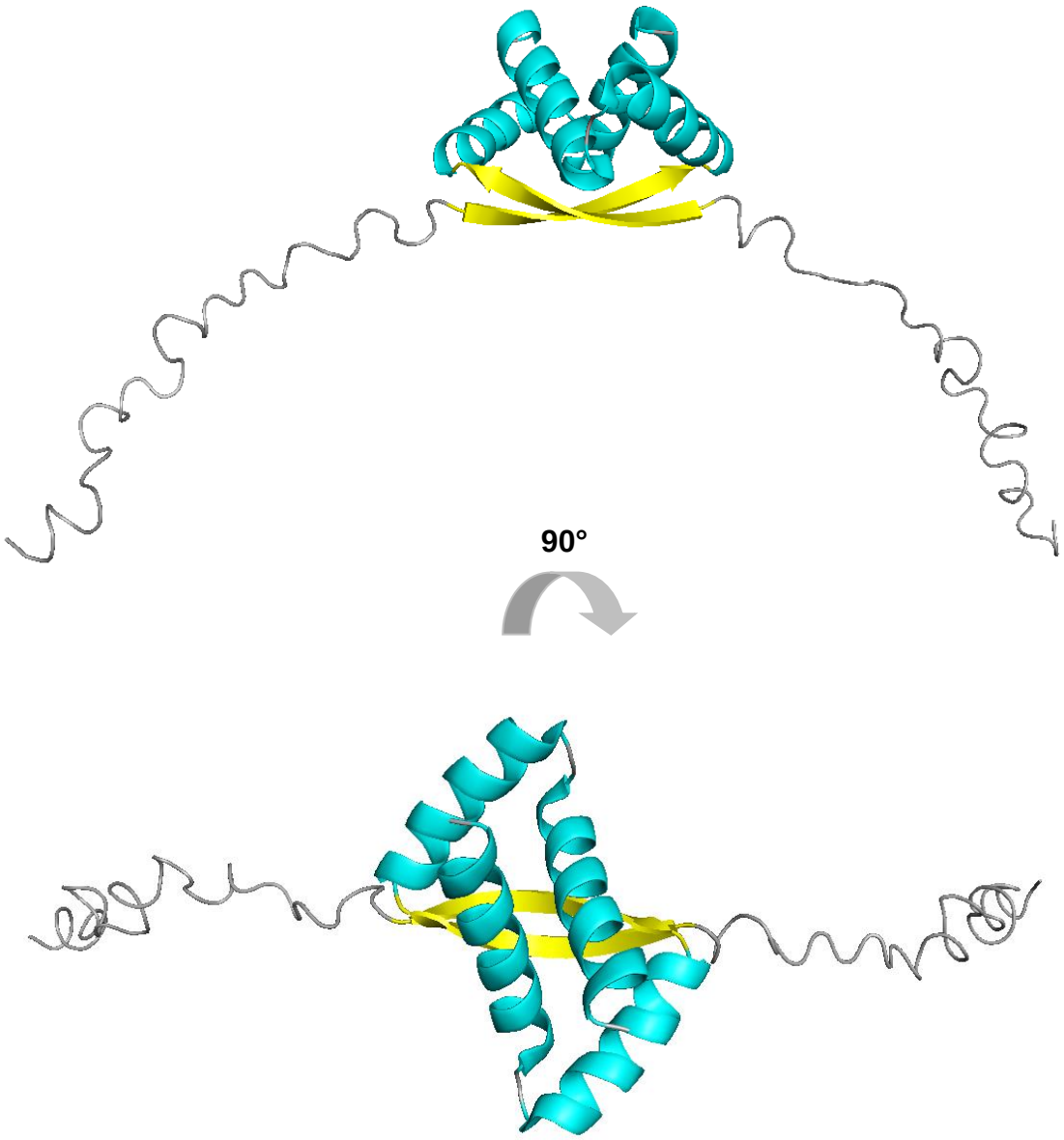

Figure 1.8 - NMR structure of ParG dimer

ParG is a homodimeric DNA binding protein. ParG monomers have a ribbon-helix-helix structure, which symmetrically intertwined together forming a dimer. The dimer is predicted to bind to *parH* by inserting the antiparallel β-sheet (yellow) in the major groove of the DNA. The C-terminal domain (residues 33-76) is involved in DNA binding and dimerisation, while the unstructured N-terminal domains (grey) (residues 1-32) perform different roles in plasmid segregation. The image was generated using PyMOL v.1.8.6.0 (The PyMOL Molecular Graphics System, Version 1.8.6.0, Schrödinger, LLC), PDB: 1P94.

1.4.4.3 Walker-type ATPase ParF

The motor protein ParF belongs to the ParA superfamily of Walker-type ATPases. However, the similarities between ParF and ParA are limited to the ATP binding motif, since structure and sequence alignment showed that ParF is more closely related to *E. coli* MinD and to *Thermus thermophilus* Soj than to P1 ParA. Therefore, ParF and its homologues have been categorized into a discrete subgroup of the ParA superfamily that is evolutionarily related to MinD. The members of this group are generally half the size of ParA and lack the characteristic helix-turn-helix motif (Hayes and Barillà, 2006b; Schumacher *et al.,* 2012).

The ParF crystal structure revealed that the protein is composed of a single domain with a central seven-stranded twisted β-sheet surrounded by four α-helices on each side (Figure1.9) (Schumacher *et al.*, 2012). ParF is a monomer when bound to ADP and a dimer when bound to ATP. The ParF-ADP structure shows that the ADP molecule is tightly wedged in a surfaced exposed cavity and interacts with the Walker-A motif. ParF Walker-A motif is located at the N-terminus of the α1-helix,residues 9-16 (PKGGSGKT), and is involved in the interaction with the β-phosphate group of the ATP/ADP molecule (Schumacher et al., 2012). A second Walker motif, Walker-B motif (residues 73-83), is also present and it is involved in magnesium binding and in the ATP hydrolysis reaction (Figure 1.9) (Schumacher *et al.*, 2012).

The structure of the ParF-AMPPCP (phosphomethylphosphonic acid adenylate ester, non-hydrolysable ATP analogue) complex shows that ParF dimerises in the presence of ATP. Each monomer binds to an AMPPCP molecule through the Walker-A motif. In the dimer, two AMPPCP molecules are sandwiched between the two ParF monomers, with the γ-phosphate interacting with the signature Lys-10 residue belonging to the Walker-A motif of the opposite monomer. This interaction and others, such as the insertion of the proline rich motif into the ATP-binding pocket of the adjacent subunit, greatly stabilize the dimer structure (Schumacher *et al.,* 2012) (Figure1.9).

Similarly to other ParA proteins, ParF is a dimer in the ATP-bound form. However, differently from its functional analogues, the ADP-bound form of ParF is monomeric (Schumacher et al., 2012). P1 ParA was found to be a dimer even in the apo state and ATP/ADP binding locks the proteins in a conformation optimal for ParA function as transcriptional repressor or partition protein (Bouet *et al.,* 1999; Dunham *et al.,* 2009). Another type Ib ParA protein, δ from *Streptococcus pyogenes* pSM19035 plasmid, was also found to be a dimer in the apo as well as in the ATP-bound form (Pratto *et al.,* 2008).

It is important to mention that in the ParF-AMPPCP complex, the sugar moiety of the ATP molecule adopts a C3-endo conformation, in contrast to the C2-endo conformation of the ADP binding form. In fact, in the ParF-ADP structure, the sugar molecule cannot bend in a C3-endo conformation due to steric hindrance. However, the C2 to C3 transition seems to be important for ParF function, since dimerisation is an essential step for ParF higher-order structure formation (Schumacher *et al*., 2012).
When the oligomers in the crystal structure are analysed, ParF always shows the same pattern of organization. ParF filaments are formed through the interaction between units composed of dimers of dimers. The dimer of dimer units have two types of surfaces which are complementary to each other geometrically and electrostatically. Each dimer potentially possesses interacting surfaces on all sides, explaining why ParF easily associates into higher order structures (Schumacher *et al.*, 2012).

In vitro, ParF polymerizes and forms long filaments in a ATP-dependent fashion. Barillà *et al.* demonstrated that ParF spontaneously self-associates *in vitro* and that ATP binding increases its association rate. In fact, mutations of the residues invoved in ATP binding abolish ParF filament formation and eventually prevent plasmid segregation *in vivo.* However, ATP hydrolysis is not necessary for polymerization, since ADP counteracts and reverses polymer assemby (Barillà *et al.*, 2005).

Electron microscopy experiments showed that ParF filaments have a polarity. One end of the filament tends to be more compact while the opposite end is irregular, suggesting that a “growth and shrinkage” process could take place at that end. On the basis of these observations, it was hypothesised that polymerisation and depolymerisation of ParF, through the ATP/ADP cycle, could drive the movement of the plasmid inside the cell (Barillà *et al.*, 2005).
In addition, the presence of ParG was shown to enhance ParF polymer formation. It has been speculated that ParG could crosslink adjacent polymers, forming ParF fibres, or being involved in polymers nucleation (Barillà *et al*., 2005; Barge, 2015).

The interplay between the components of the segrosome is fundamental for the process of plasmid segregation. ParF and ParG interact in the presence or absence of the ATP, although this was shown to stabilise the complex and improve the strength of the binding (Barillà and Hayes, 2003). The resence of *parH* is not required for binding between the two protein components. In addition the interaction between ParG and *parH* is enhanced by the presence of ParF (Barillà and Hayes, 2003).

Similarly to other ParA proteins, ParF interacts with non-specific DNA *in vitro* in a ATP-dependent manner and the nucleoid *in vivo*. Live-cell imaging showed that the protein oscillate over the nucleoid and this oscillation depends on the presence of the ParG-*parH* complex (McLeod *et al.*, 2017; Allison-Gamble, 2016). A Venus flytrap model was proposed for the segregation of the TP228 plasmid. This is based on the formation of a ParF meshwork within the nucleoid, where the ParG-bound plasmid becomes trapped and transported to opposite poles (McLeod *et al.*, 2017).

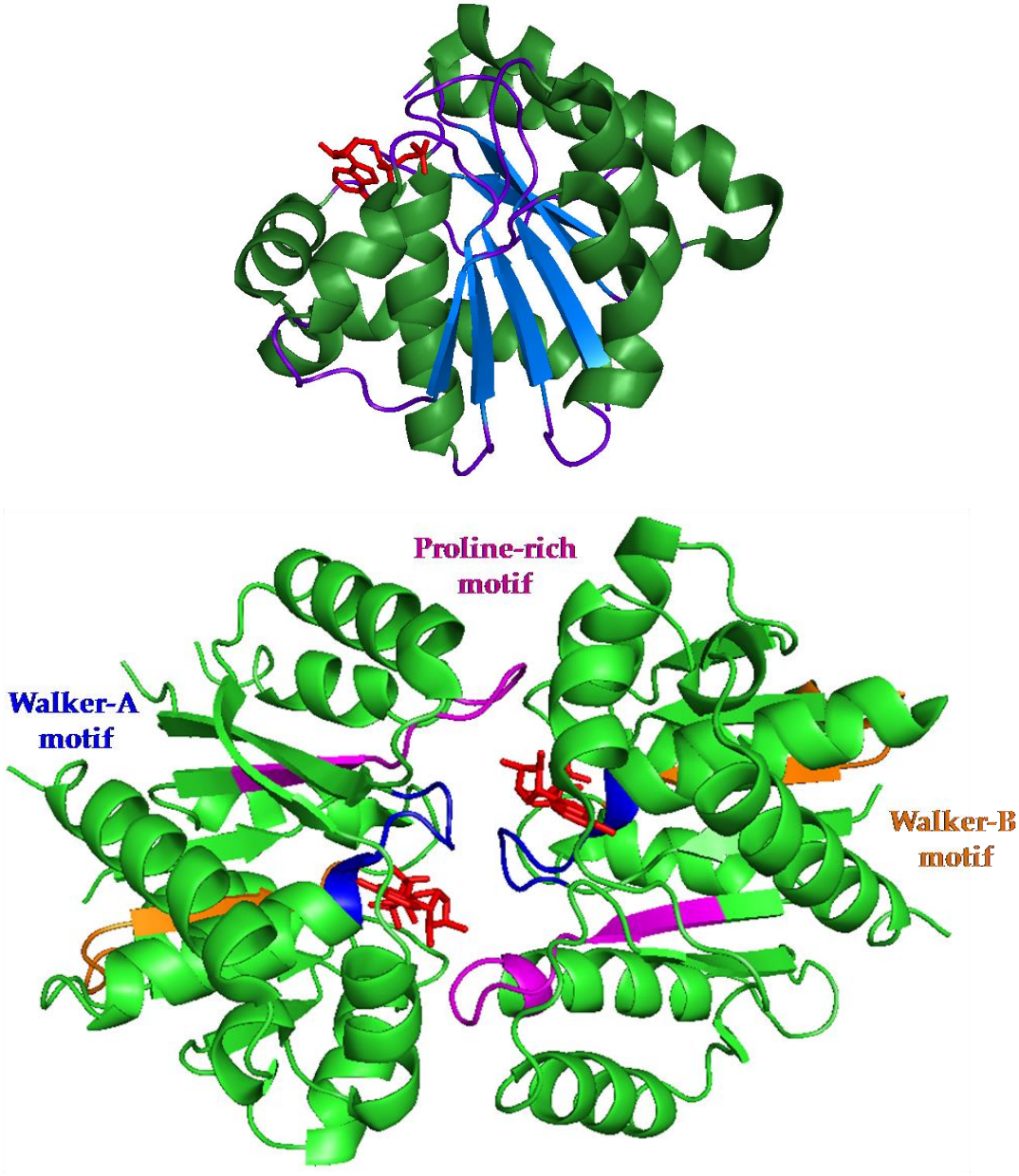

Figure 1.9 - Crystal structure of ParF-ADP and ParF-AMPPCP

The top structure represents ParF monomer bound to ADP. ParF shows seven central β-strands (blue), with four α-helices on each side. The ADP molecules shown in red binds to the α1 helix and to the P-loop (PDB: 4E03).

The bottom structure represents ParF dimers bound to the non-hydrolysable ATP analogue AMPPCP (red). In the dimer the AMPPCP molecules are sandwiched between the ParF monomers. The Walker-A motif (blue) and the Walker-B motif (orange) make interaction with the nucleotide. The proline-rich motif (magenta) of each monomer extends to make contacts with the other monomer and stabilises the dimeric structure (PDB: 4E07).

The images were generated using PyMOL v.1.8.6.0 (The PyMOL Molecular Graphics System, Version 1.8.6.0, Schrödinger, LLC).

1.4.5 Type Ib segregation system of pSM19035 plasmid

B.subtilis pSM19035 plasmid has evolved to encode a number of systems to ensure its stable maintenance throughout generations (Dmowski *et al.*, 2006). One of the stability mechanisms involves the type Ib segregation system composed of the proteins δ and ω. δ is a Walker-type ATPase, which lacks the type Ia signature HTH N-terminal domain (Dmowski *et al.*, 2006). Instead, ω is a DNA binding protein that binds to different combinations of direct and inverted heptameric motifs (5'-WATCACW-3') upstream of its own gene, as well as in the promoter region of the δ gene and *copS* (involved in copy number control). The δ promoter overlaps with the identified *parS* region, which is made of seven direct and two inverted heptameric repeats (→→→→→→←←) (De La Hoz *et al.,* 2004). Therefore, besides being involved in plasmid partition, the protein has also a transcriptional regulatory role (De La Hoz *et al.,* 1999; Dmowski *et al.*, 2006).

ω is an asymmetrical homo-dimer with a ribbon-helix-helix C-terminal domain. The N-terminal domain is supposed to be unstructured. The first twenty residues were proteolysed during crystallography, while residues 21 to 23 could not be identified because they were disordered. The protein belongs to the MetJ/Arc superfamily, although it shares very little sequence similarities with the other members (Murayama *et al*., 2001). The protein binds cooperatively to repeated sequences upstream of the *δ* and *copS* genes, as well as its own. The interaction with one repeat is very weak, while it becomes much stronger when at least three or four heptads are present. It was also shown that ω has higher affinity for repeated sequences in the same direction, rather than in opposite orientations (De La Hoz *et al.,* 2004). The protein can recognise both palindromic and non-palindromic sequences, which is uncommon for RHH proteins (Weihofen *et al*., 2006).
The crystal structure of ω in complex with two repeats, either direct (→→) or opposite orientation (→←), showed that the protein recognises the DNA through the double stranded β-sheet of the dimer and that although the proteins affects slightly the DNA conformation, overall it does not bend the DNA (Weihofen *et al.,* 2006). Instead, the protein wraps linearly around the centromeric site in a left-handed helix.

The truncated version of ω, Δ19ω, where the first 19 N-terminal amino acids were deleted, shows binding to the DNA two-fold weaker than the wild type protein. This suggests that the N-terminal domain is not essential for DNA binding. Instead the N-terminal domain is essential for δ activation, as observed for the N-terminal domain of TP228 ParG with its partner ParF (Weihofen *et al.,* 2006; Carmelo *et al.,* 2005; Barillà *et al.*, 2007).

The δ protein is a Walker-type ATPase with eight central β-strands surrounded by twelve α-helices. The protein presents structural similarities to MinD, Soj and ParF. δ, however, is a dimer in the apo form, as well as in the ATP and ADP-bound state. The nucleotide-binding pocket forms a U-shape cleft, where the two arms of the U belong to two δ units bound to one nucleotide molecule (Pratto *et al*., 2008; Schumacher *et al.,* 2012). ATP hydrolysis does not cause dissociation of δ dimers, as in the case of its homologues that dimerise only in the presence of ATP (Pratto *et al.*, 2008).
δ ATPase activity is enhanced by the ω-*parS* complex, when the ω:δ ratio is 1.4:1. Above this value, stimulation of δ ATPase activity decreases. Furthermore, δ-GFP was seen to co-localise with the nucleoid in an ATP-dependent manner and when the full partition system was present (δ, ω, *parS*, ATP), δ-GFP (green fluorescent protein) seemed to form helical structures within the nucleoid. Biochemical assays showed that the protein can polymerise on the DNA at high ω:δ ratios. Instead, ATP hydrolysis causes polymer disassembly (Pratto *et al.*, 2008).

Pratto *et al.* proposed a model for the segregation of the plasmid pSM19035 according to which the process of segregation depends on the concentration of ω. They suggested a polymers-mediated pulling mechanism on the nucleoid, whose dynamics depends on the abundance of ω inside the cell (Pratto *et al.*, 2008).

1.4.6 Plasmid segregation mechanisms for type I partition systems

Type I segregation systems are the most widespread in nature and have been found on plasmid and chromosomes of bacteria and archaea. Although the object of extensive studies for over three decades, the precise mechanism through which these systems mediate DNA segregation remains elusive. Among the alternatives, two models have been proposed over the years. The first was put forward by Ringgaard *et al.* and involves the formation of extensive ParA filaments over the nucleoid (Ringgaard *et al*., 2009). The second was first proposed by Vecchiarelli *et al.* and involves the formation of a ParA gradient within the nucleoid (Vecchiarelli *et al.,* 2013b). In both mechanisms, ParA tethers the plasmid to the nucleoid and ATP-hydrolysis is essential for plasmid movement.

1.4.6.1 Filament-pulling mechanism

The hypothesis of a polymer-mediated mechanism first arose from the observation that type I ParA proteins form filamentous structures *in vitro* and *in vivo* (Ebersbach and Gerdes, 2004; Barillà *et al*., 2005; Bouet *et al.,* 2007; Pratto *et al.*, 2008; Dunham *et al.*, 2009; Schumacher *et al.*, 2012). Therefore, it was originally believed that ParA polymers

could push or pull plasmids apart in an ATP-dependent manner. It was later discovered that ParA-ATP could interact non-specifically with the DNA and could oscillate over the nucleoid (Castaing *et al.*, 2008; Pratto *et al.,* 2008; Vecchiarelli *et al.,* 2010; McLeod *et al*., 2017). Based on this evidence, the nucleoid was added to the system as a scaffold for polymer assembly. It was also seen that ParA dynamics over the nucleoid depends on the presence of ParB-*parS* nucleoprotein complex, which is essential to stimulate ParA ATPase activity (Erdmann *et al.*, 1999; Ebersbach *et al.,* 2006; Barillà *et al.*, 2007; Bouet *et al.,* 2007; Ringgaard *et al.,* 2009; McLeod *et al.,* 2017).

Work on the pB171 plasmid gave new insights into this mechanism. Fluorescence microscopy on the plasmid harbouring the *par2* partition system showed that ParA-GFP formed elongated structures over the nucleoid, which retracted upon interaction with the plasmid, directing its movement (Ringgaard *et al.,* 2009). The model suggested by these authors, which found supporting evidence in live-cell imaging and biochemical assays, was a pulling mechanism, based on the dynamic polymerisation and depolymerisation of ParA over the nucleoid through the ATP/ADP cycle. According to this model ParA-ATP dimers associate with the nucleoid and form extensive filament structures. When the protein encounters the plasmid, the interaction with ParB, concentrated in the *parS* region, triggers ParA ATPase activity causing dissociation of the polymers at the interface with the plasmid. ParA-ATP is converted into ParA-ADP and dissociates from the nucleoid. The plasmid can move towards the ParA-ATP concentrated area, following the direction of the polymer's retraction. Once the polymers are completely disassembled, the plasmid is released. ParA-ADP is then regenerated into ParA-ATP dimers, which can associate elsewhere on the nucleoid and re-start the process (Ebersbach *et al.,* 2006; Ringgaard *et al.*, 2009; Gerdes *et al.,* 2010; Szardenings *et al.*, 2011) (Figure1.10).

This model was further supported by comparison of *in vivo* results and computational modelling. The formation of ParA structures, compared to a ParA gradient, would ensure a localised high concentration of ParA, preventing plasmids from diffusing away. Furthermore, a polymer-mediated mechanism would allow plasmid to migrate at a speed comparable to the one observed through time-lapse microscopy (Ietswaart *et al.*, 2014). According to this model, the perpetual association and dissociation of ParA structures on the nucleoid would allow accurately positioning of plasmids at cell quarters before cell division.

1.4.6.2 Diffusion-ratchet mechanism

A diffusion-ratchet mechanism based on a ParA gradient, rather than ParA filaments was proposed by Vecchiarelli *et al.* The model is based on the premise that ParA-ATP dimers bind non-specifically to the nucleoid (Castaing *et al.*, 2008; Pratto *et al.*, 2008; Vecchiarelli *et al.*, 2010; McLeod *et al.*, 2017). However, before becoming competent to bind to the DNA, ParA-ATP undergoes a series of conformational changes that delay its interaction with the nucleoid. Conversely, ADP binding prevents ParA from interacting with the chromosome. Involvement of polymers in the segregation mechanism is still under debate (Vecchiarelli *et al.*, 2010). However, Vecchiarelli *et al.* pointed out that no filament structures were observed *in vivo* for P1 ParA and that high concentrations of protein were required to see polymerisation *in vitro* (Dunham *et al*., 2009; Vecchiarelli *et al*., 2010). This was not the case for several other ParA proteins that were seen to form extensive filaments both *in vivo* and *in vitro* (Ebersbach and Gerdes, 2004; Bouet *et al.,* 2007; Pratto *et al.*, 2008; Dunham *et al.*, 2009; Schumacher *et al.,* 2012).
The diffusion-ratchet model suggests that rather than following polymer retraction, the plasmid moves in the direction of a ParA-ATP gradient.

Plasmid segregation starts with the recognition of the centromeric site on the plasmid by ParB. ParA-ATP dimers, competent to bind to the DNA, associate unevenly on the nucleoid. Since ParB interacts directly with ParA, it becomes tethered to the nucleoid in the region where ParA is more concentrated. Interaction with ParB enhances ParA ATPase activity, leading to ATP hydrolysis and ParA-ADP dissociation from the nucleoid. Before ParA is able to associate again with the DNA, a series of conformational changes are required, from ParA-ADP to inactive ParA-ATP and finally active ParA-ATP. The time delay prevents ParA from re-associating in the same position. Instead, it allows ParA to diffuse away and bind again in a new random position on the nucleoid. As ATP is hydrolysed through the catalytic action of ParB, a ParA depleted area is left around the partition complex. The plasmid partially dissociates and re-associates to the nucleoid following ParA concentration. Once at the nucleoid pole, the plasmid changes direction following the newly formed ParA gradient on the nucleoid (Vecchiarelli *et al.*, 2010, 2012; 2013; Hwang *et al*., 2013b) (Figure 1.11).

This model was further supported by cell-free reconstructions of the segregation dynamics of the P1 and F plasmids. These studies used purified components, which were made to interact with a DNA carpeted flow-cell, which mimicked the nucleoid. The system was visualised by total internal reflection microscopy (TIRFM) (Hwang *et al*., 2013; Vecchiarelli *et al*., 2013b, 2014b).

Experiments with P1 and F plasmid partition proteins showed similar results. Both ParA and SopA do not seem to form polymers on the DNA and can tether the plasmid to the flow cell surface. According to the model proposed originally by Vecchiarelli *et al.* plasmids are tethered to the nucleoid surface and segregation takes place in the narrow space between the nucleoid and the cell membrane. While these experiments provided support to the diffusion ratchet mechanism, they also showed that spatial confinement is an essential prerequisite for this mechanism (Hwang *et al*., 2013; Vecchiarelli *et al.*, 2013b). Space confinement was addressed by using a magnetic field to constrain the movement of a *parS*/*sopC*-coated magnetic bead. The combination of space constraint, slow plasmid diffusion (caused by the viscosity of the cytoplasm and the size of the cargo) and the large number of ParB-ParA interacting units allow the plasmid to move directionally, without losing contact with the nucleoid (Vecchiarelli *et al.*, 2014a).

The diffusion-ratchet mechanism was further investigated and amplified by two important discoveries. Lim *et al.* proposed a "DNA-relay" mechanism for *C. cresencentus* chromosome segregation, where the elastic force of the chromosome was reported to be essential for the directed movement of the ParABS system (Lim *et al.,* 2014).
In addition, it was recently determined that the partition elements move within the volume of the nucleoid, rather than on the nucleoid surface (Le Gall *et al*., 2016; McLeod *et al*., 2017) and that ParA could associate with high density regions of the chromosome (Le Gall *et al.,* 2016).

1.4.6.3 Venus flytrap mechanism for the segregation of the TP228 plasmid

In vivo investigation of the ParFGH partition system, using a combination of confocal and super-resolution microscopy techniques, allowed McLeod and colleagues to propose a novel mechanism for the segregation of the TP228 plasmid, which can be applied to all type I systems (McLeod *et al.,* 2017).

Fluorescent tag fused to ParF and ParG allowed visualisation of the protein dynamics. ParF was fused to an enhanced green fluorescent protein called Emerald and ParG was fused to mCherry. ParG interacts with the plasmid by binding to the *parH* site, visible as a bright focus. When cells contained only one focus, this was usually positioned at midcell. Two foci, instead, were visible at cell quarters positions (1/4 or 3/4 on the long axis). The majority of the cells where the *parF* gene harboured mutation in the Walker-A motif or where *parF* was deleted (*ΔparF*) showed only one focus, which often did not co-localised with the nucleoid. In cells where the whole partition system was present, ParF-Emerald formed a comet-like pattern on the nucleoid where it was more concentrated on

one side and diffuse on the opposite site. In contrast when the ParGH complex was not present, the ParF protein diffused evenly on the chromosomal DNA (McLeod *et al.,* 2017). In the presence of the full partition system, ParF was seen to oscillate over the nucleoid, as observed for other type I ParAs (Ebersbach and Gerdes, 2004; Hatano *et al.*, 2007; Pratto *et al.*, 2008). ATP binding was also proved to be essential, as ParF Walker-A motif mutants were unable to dynamically relocate. Furthermore, enhancement of ParF ATPase activity by ParG was another prerequisite for ParF oscillation. In fact, when ParG was changed into ParG-R19K, where the arginine finger was changed into lysine, ParF co-localised with the ParG-R19K-*parH* complex and was unable to oscillate. It was also shown that ParF oscillation alone is not enough to ensure plasmid segregation, instead the interplay between the components of the system is essential (McLeod *et al*., 2017).

Biochemical assays showed that ParF forms higher order structures in presence of ATP (Barillà *et al.,* 2005; Machón *et al.,* 2007; Schumacher *et al.*, 2012) and 3D-SIM (Structured Illumination Microscopy) reinforced these findings by showing that ParF forms a three-dimensional meshwork within the nucleoid (McLeod *et al.*, 2017). It was proposed that ParF could define plasmid partition area within the nucleoid and direct plasmid movement by the continuous reorganisation of the ParF meshwork. However, ParF meshwork dynamics is dependent on ParG, which enhances ParF ATPase activity. Therefore, plasmid segregation dynamics are determined by ParG.

Based on the evidence collected through live-cell imaging the authors suggest a Venus flytrap model. In this model ParF assembles into a three-dimensional meshwork, a combination of polymers and small oligomers that extend in all directions inside and across the nucleoid. The meshwork traps the newly replicated sister plasmids by interacting with the *parH*-bound ParG. Sister plasmids are transported through the meshwork to the nucleoid pole. The meshwork has a leading, compact, end and a lagging, more loose, end. As the meshwork moves over the nucleoid, polymers grow between the sister plasmids. ParG enhancement of ParF ATPase activity causes polymers dissociation in the vicinity, causing the plasmid to be eventually released from the lagging end of the meshwork. The other plasmid is maintained in the meshwork and transported to the opposite pole. As the migration over the nucleoid proceeds, ParG triggers ATP hydrolysis by ParF and the disassembly of the meshwork at the interface with ParG. This causes the plasmid to be finally dropped at the opposite end. The meshwork reorganisation leads to plasmid position finer adjustments at cell quarters (Figure 1.12) (McLeod *et al.*, 2017).

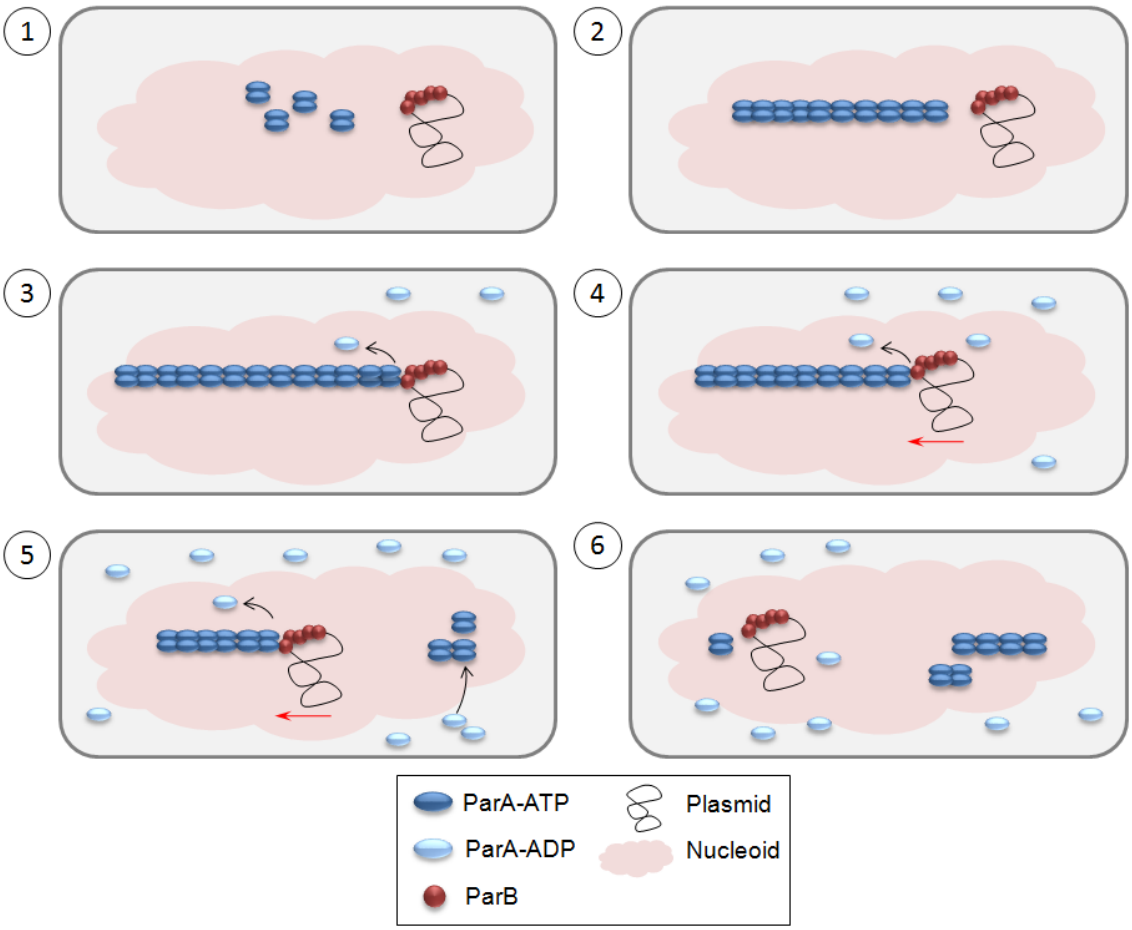

Figure 1.10 - Filament pulling model for plasmid segregation mediated by Type I partition systems

(1) ParB associates with *parS* on the plasmid. ParA-ATP dimers interact non-specifically with the nucleoid and start polymerising. (2) ParA-ATP polymers extend over the nucleoid forming filaments, that eventually encounter the plasmid. (3) ParA polymers interact with the plasmid through the ParB-*parS* complex. When ParA binds to ParB, ParB enhances ParA ATPase activity, promoting the conversion of ATP into ADP. ParA-ADP is released from the polymer and from the nucleoid. (4) The plasmid moves toward ParA-ATP following the direction of the polymer retraction, leaving behind a ParA depleted zone. (5) The depolymerisation of ParA continues as the plasmid moves toward the nucleoid pole. ParA-ADP is converted into ParA-ATP and can associate with the nucleoid elsewhere to re-start the process. (6) Once at the pole, ParA polymer is completely dissociated and the plasmid is dropped. Diagram adapted from Ringgaard *et al.*, 2009.

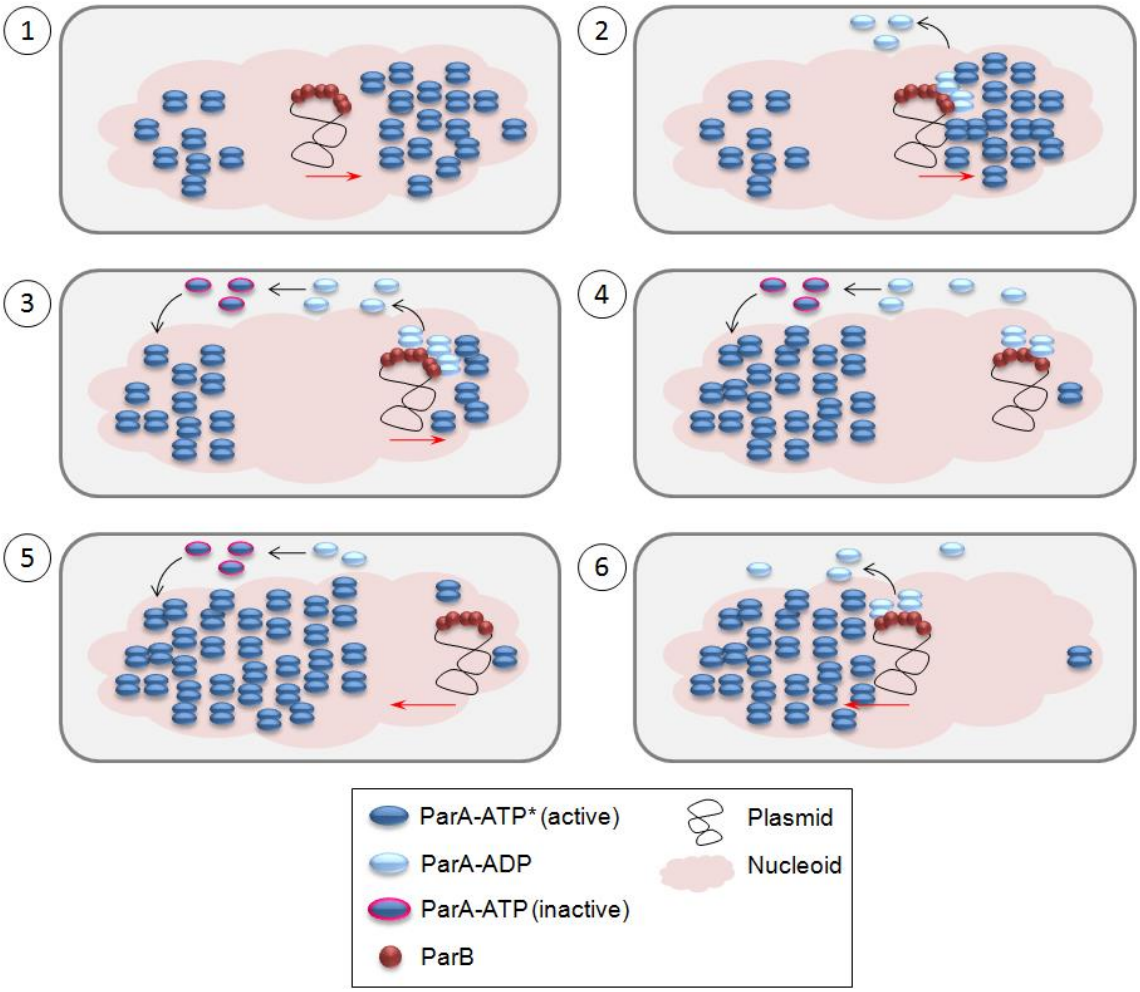

Figure 1.11 - Diffusion-ratchet model for plasmid segregation mediated by Type I partition systems

(1) ParB associates with *parS* on the plasmid. Active ParA-ATP dimers interact non-specifically with the nucleoid. (2) The plasmid-ParB is tethered to the nucleoid through the interaction with ParA. (3) The interaction with ParB enhances ParA ATPase activity, promoting the conversion of ATP into ADP. ParA-ADP can no longer bind to the DNA and it is released from the nucleoid. ParA-ADP needs to undergo a series of conformational changes before it becomes active and able to interact with the DNA again. This delay allows ParA to randomly diffuse and to re-associate with the nucleoid elsewhere. (4) The plasmid moves towards ParA-ATP concentration, leaving a ParA depleted area behind. (5) Once at the nucleoid pole, the partition complex changes direction moving towards the ParA gradient. (6) ParB interaction with ParA enhances ATP-hydrolysis into ADP and the cycle re-starts. Diagram adapted from Vecchiarelli *et al*., 2010.

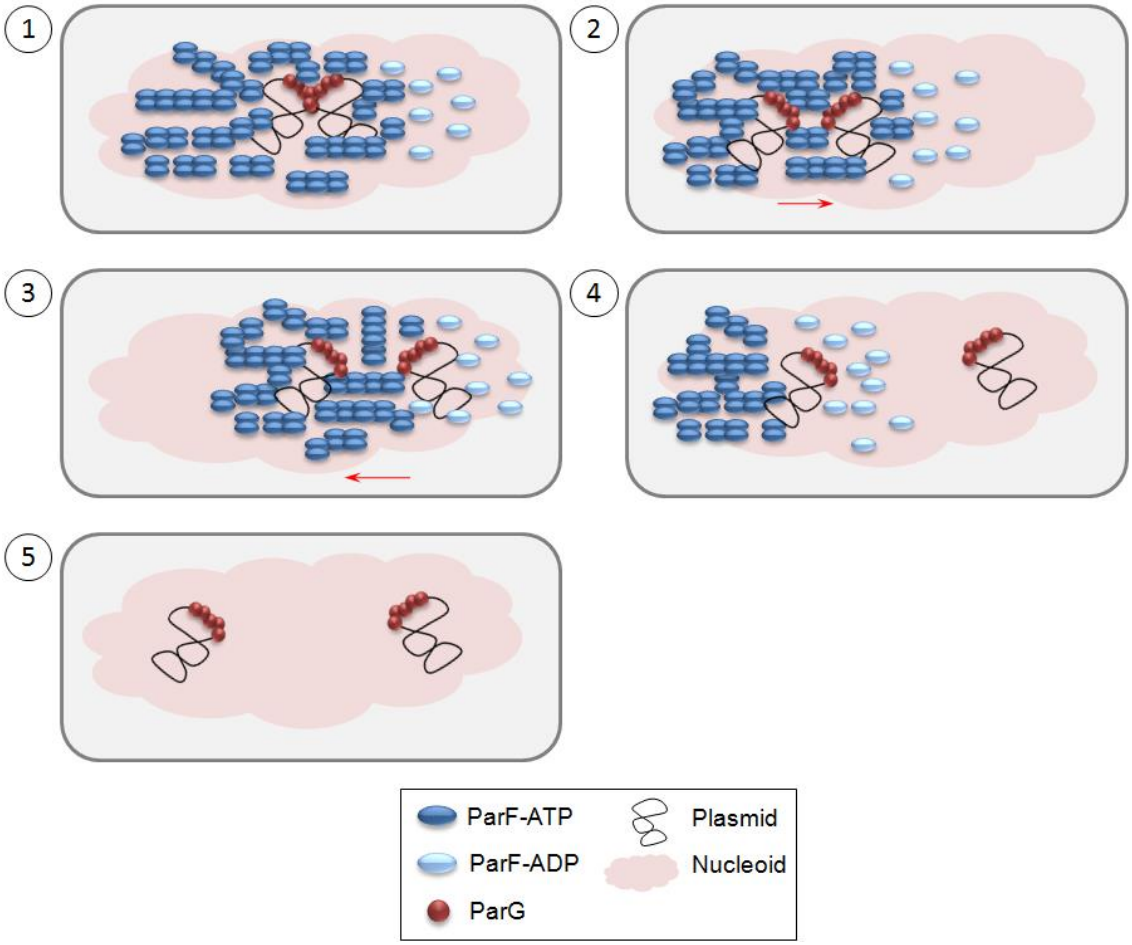

Figure 1.12 - A Venus flytrap model for plasmid segregation mediated by TP228 and Type I partition systems

(1) ParF-ATP forms a meshwork protruding within the nucleoid. ParG interacts with the plasmid by binding to *parH*. Replicated sister plasmids are trapped within the meshwork through the interaction between ParF and ParG. The meshwork has a leading front and a lagging tail. (2) The plasmids are transported to the nucleoid pole by the meshwork. As the meshwork moves, polymers grow between the sister plasmids. (3) The interaction with ParG promotes ParF polymer dissociation into monomers, leading one of the plasmids to be released through the less compact end of the meshwork. (4) The other plasmid remains tethered to the ParF lattice and moves towards the opposite nucleoid pole, till it is released. (5) Meshwork rearrangements adjust plasmid location to quarter cell positions. Diagram adapted from McLeod *et al*., 2017.

1.5 Type II partition systems

Type II segregation systems are characterised by a motor ATPase protein with an actin-like fold that segregates plasmids by pushing them apart. Type II partition cassettes have been identified on plasmids isolated from Gram-positive and Gram-negative bacteria. The best characterised segregation system is the one of the multidrug-resistance *E. coli* plasmid R1, however studies on the plasmids pSK1 and pB171 enhanced our knowledge on the organisation of the segrosome (Salje *et al.*, 2010).

1.5.1 Type II partition system of R1 plasmid

The partition cassette of the R1 plasmid is composed of two *trans*-acting elements, ParM, an actin-like ATPase, and ParR, a DNA binding protein, and a *cis*-acting region, *parC*, upstream of the partition operon (Gerdes *et al.*, 1985; Gerdest and Molin, 1986). The centromeric site *parC* is composed of two clusters of five 11 bp direct repeats, which also include the *parMR* operator site (Dam and Gerdes, 1994) (Figure 1.13 A). ParR binds cooperatively to *parC* to promote plasmid segregation and to repress transcription of the partition operon. ParR is also involved in pairing plasmids after replication, through *parC* sites (Jensen *et al.*, 1998).

In vitro investigations showed that each *parC* repeat is bound by a ParR dimer and that at least two repeats are required to achieve initial binding (Møller-Jensen *et al.,* 2003). Atomic force microscopy showed that ParR binding changes the conformation of linear DNA harbouring *parC*. The ParR-*parC* complex was seen to have a U shape, although bridging between the two clusters of repeats that compose *parC* was not observed (Hoischen *et al.*, 2008).

Structural studies on pSK41 ParR showed that the protein has a RHH fold like type Ib CBPs. The protein forms a continuous structure on the DNA, which wraps around ParR forming a solenoid structure. The protein cooperatively binds to itself and to the DNA, giving rise to a super-helical structure, in the middle of which ParM can insert (Schumacher *et al.*, 2007a) (Figure 1.13 B). Structural studies on pB171 ParR showed a similar arrangement of ParR dimers, with the DNA binding domain facing outward. This suggests the formation of a similar structure, where the *parC* site wraps around the ParR core (van den Ent *et al*., 2002). Although structural information on R1 ParR is lacking, the protein is likely to behave like its homologues.

ParM is an actin-like ATPase, belonging to the same family of ATPases as the cytoskeletal protein MreB. ParM crystal structure showed that the protein forms left-

handed double helical filaments, similar to actin in eukaryotic cells. Binding to ATP promotes filament formation, while ADP favours the formation of monomers. Therefore, ATP hydrolysis causes polymer instability (van den Ent *et al.*, 2002).
The protein was also shown to form filament structures *in vivo*. Filaments grow between the separating plasmids and localise them close to the cell poles (Møller-Jensen *et al.*, 2002). Besides requiring ATP, filament formation was seen to depend on the presence of ParR and *parC* (Møller-Jensen *et al*., 2002, 2003). Binding experiments showed that ParM interacts with ParR only when the protein is in complex with *parC* and when ATP is present (Møller-Jensen *et al.*, 2003). Furthermore, ParM ATPase activity is enhanced by ParR, when bound to *parC* (van den Ent *et al.*, 2002).

Evidence suggests that ParM filaments easily dissociate, while they are stabilised by the interaction with the partition complex. Polymers were seen to grow at the interface between ParM and ParR, suggesting an insertional polymerisation mechanism (Møller-Jensen *et al.*, 2003).

1.5.2 Insertional polymerisation or "pushing" mechanism

R1 plasmid segregation mechanism was reconstructed by Garner *et al.* using purified components. These authors determined that ParM forms filaments stabilised at one end by the ParR-*parC* complex on the plasmid. When two filaments capped at one end with the partition complex meet, they form an active spindle that pushes plasmids apart. The dynamic instability of the filaments would allow ParM to search for the plasmids and dissociate if not needed. They also saw that in micro-fabricated channels, ParM filaments always found the long axis, suggesting that no other cellular components is involved (Garner *et al.,* 2007).

The ParMRC system was further characterised to near atomic resolution. Cryo electron microscopy showed that ParM forms left-handed double helical filaments, that pair into anti-parallel double helical structures to form a bipolar spindle apparatus. In this way, the partition system can push plasmids originally paired together towards opposite poles (Gayathri *et al.,* 2012; Salje*et al.*, 2009; Bharat *et al.*, 2015).

The mechanism of plasmid segregation starts with the recognition of *parC* by ParR, shortly after replication. ParM searches for the plasmids by forming double-helical structures in an ATP-dependent manner. These structures easily disassemble if they are not capped at both sides by the ParR-*parC* complex. Therefore, when ParM filaments encounter the plasmid, they become stabilised and start to elongate by inserting ParM-

ATP units at the interface with ParR. The spindle apparatus is formed by two anti-parallel double helical structures that grow in both directions and push the plasmids to the poles, beyond the nucleoid limits. Once plasmids reach the poles ParM polymers disassemble (Salje *et al.*, 2010) (Figure 1.14).

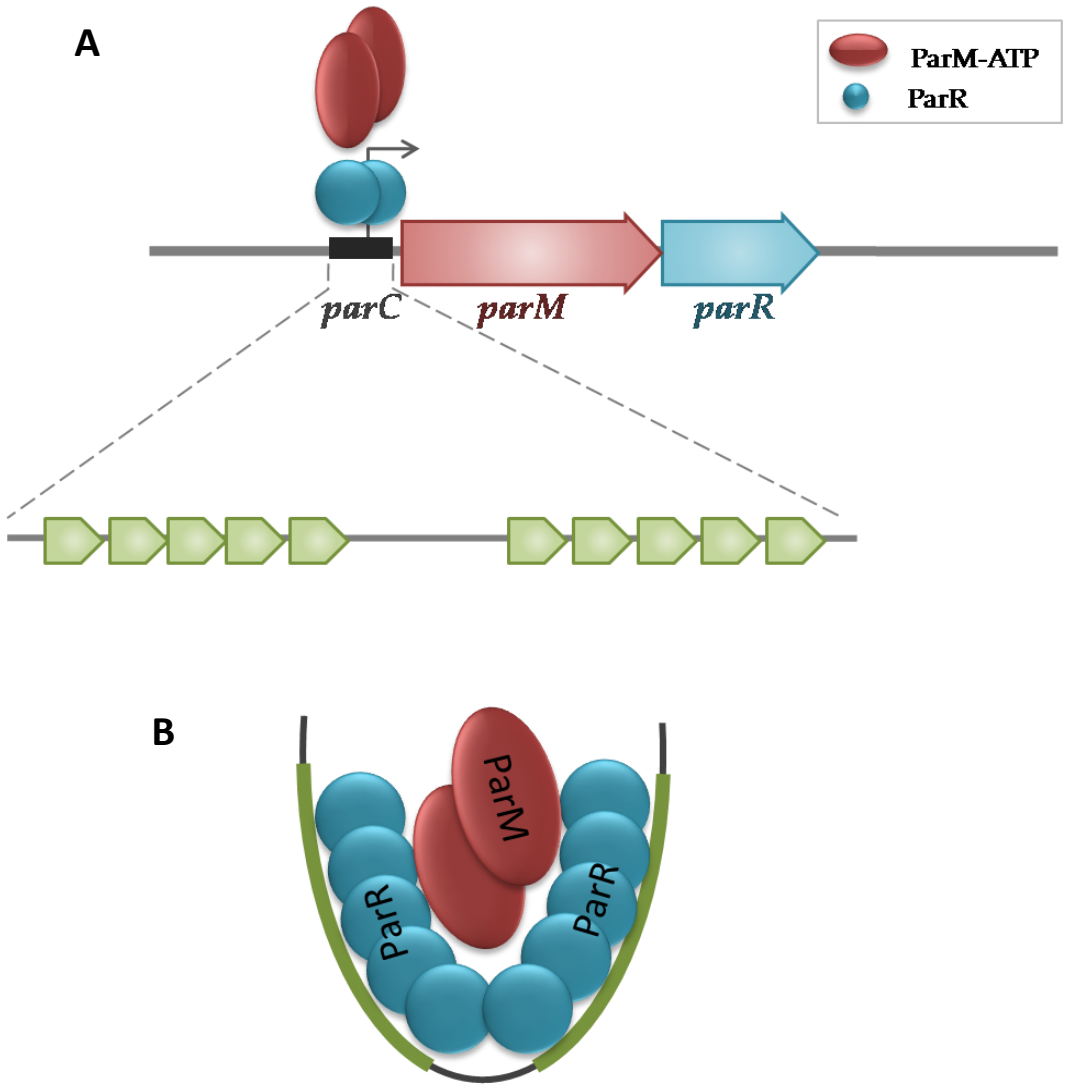

Figure 1.13 - Organisation of R1 ParMRC partition system and structure of the segrosome
A) The *parMRC* partition cassette encodes for the actin-like ATPase ParM (red) and the DNA binding protein, ParR (green). *parC* is the centromeric region (dark grey). ParR works as centromere binding protein and repressor of the partition cassette by interacting with the region upstream the partition operon. *parC* is composed of two clusters of five repeated sequences. Each dimer of ParR recognise one of the repeats.
B) ParR dimers cooperatively bind to themselves and to cognate site, *parC*. This causes ParR-bound DNA to bend and wrap around the ParR complex forming a superhelical structure in the middle of which ParM can bind.

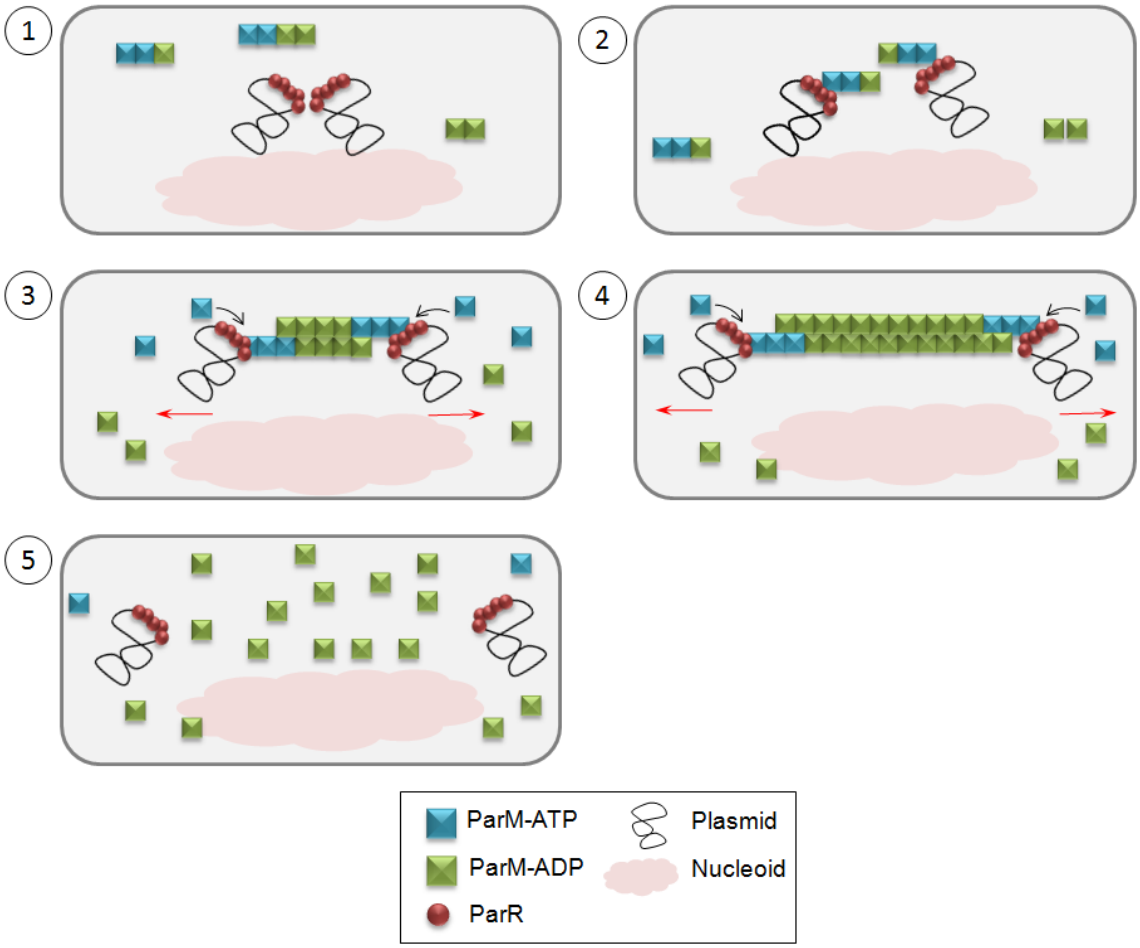

Figure 1.14 - A pushing mechanism for plasmid segregation mediated by Type II partition systems

(1) After replication, ParR binds to the plasmid through the recognition of the *parC* site. (2) ParM searches for the plasmids by forming polymeric structures in a ATP-dependent manner. If the polymers remain unbound, they quickly dissociate. When ParM filaments encounter the plasmid, they get stabilised and start to elongate by inserting ParM-ATP units at the interface with ParR. (3)(4) The spindle apparatus keeps growing and pushes the plasmids to the poles. (5) When the plasmids reach the poles ParM polymers disassemble.

1.6 Type III partition systems

Recently a new class of segregation systems has been identified on plasmids isolated from Gram-positive bacteria such as *B. thuringensis* pBtoxis and *B. anthracis* pDSW208 and pXO1. The system is characterised by a GTPase motor protein with a tubulin-like fold. The TubZR*C* system on pBtoxis represents the archetype of type III segregation systems (Larsen *et al.*, 2007; Brooks and Hwang, 2017).

1.6.1 Type III segregation system of plasmid pBToxis

The *pBtoxis* partition cassette is composed of two *trans*-acting elements, TubZ, and TubR, and a *cis*-acting site upstream of the cassette, *tubC* (Figure 1.15). TubZ is a GTPase protein distant homologue of tubulin and FtsZ, while TubR is a small dimeric DNA binding protein (Larsen *et al.*, 2007).

TubR has a HTH domain which makes interactions with the DNA in a peculiar way, different from the canonical HTH proteins mode of binding. The protein recognises *tubC,* composed of seven iterons localised upstream of the partition cassette, and is involved in plasmid segregation, as well as transcription regulation of the system (Fink and Löwe, 2015).

TubZ is a cytoskeletal protein that forms right-handed double helical structures, which are stabilised by GTP (Aylett *et al.*, 2010). The protein shows similarities with its functional analogue ParM, which also form double helical structures and can associate into bundles (van den Ent *et al.*, 2002; Aylett *et al.*, 2010). TubZ can interact with TubR though a flexible C-terminal tail (Ni *et al.*, 2010). TubZ filaments were shown to have a polarity, with a plus end where the filament grows and a minus end where it disassembles. TubZ polymers can, therefore, move inside the cell using a treadmilling mechanism (Larsen *et al.,* 2007).

1.6.2 “Pulling” treadmilling mechanism

A pulling mechanism based on treadmilling was proposed for the segregation of type III plasmids. Fink and Löwe showed that TubR-*tubC* complex binds preferentially to the minus end of the filament. Binding of the nucleoprotein complex causes reduction in filament shrinkage. In addition TubR*C* is able to nucleate TubZ filaments, suggesting that the nucleoprotein complex may recruit TubZ to start plasmid segregation, rather than the filaments searching for their cargo as for ParM (Fink and Löwe, 2015).

The segregation mechanism, therefore, involves the association of the TubR-bound plasmid to the minus end of the TubZ filament, reducing the depolymerisation rate. The

filament keeps growing on its plus side, while shortening at the interface with TubR*C*. TubZ filaments interact with TubR through the C-terminal extension. However, since several TubR-TubZ units interact with each other, the plasmid remains tethered to the minus end, while the filament shrinks and grows on the opposite end (Fink and Löwe, 2015; Ni *et al*., 2010). The way in which the plasmid is dropped at cell pole remains unclear, but it was suggested that the natural curvature of the cell could make the plasmid spontaneously dissociate (Larsen *et al*., 2007) (Figure 1.16). There are still many unanswered questions regarding this mechanism, among which is whether other factors contribute to the system. It was suggested that the complex could be tethered to the nucleoid or to the membrane (Fink and Löwe, 2015).

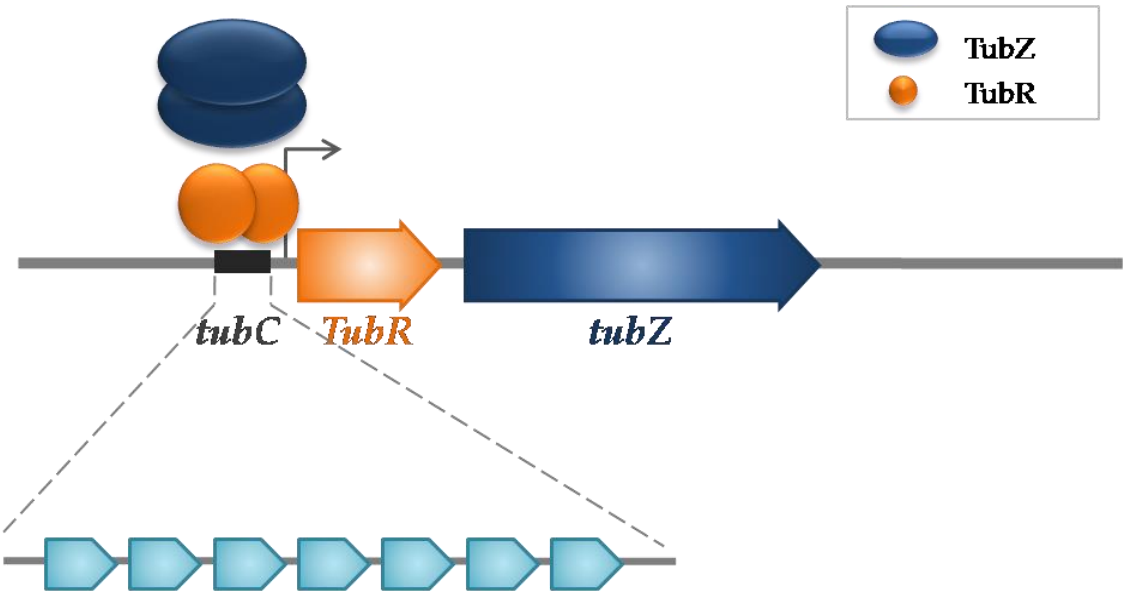

Figure 1.15 - Organisation of pBtoxis TubZRC partition system

A) The *tubZRC* partition cassette encodes for the tubulin-like GTPase TubZ (blue) and the DNA binding protein, TubR (orange). *tubC* is the centromeric region (dark grey). TubR works as centromere binding protein and repressor of the partition cassette by interacting with the region upstream the partition operon. *tubC* is composed of seven repeated sequences.

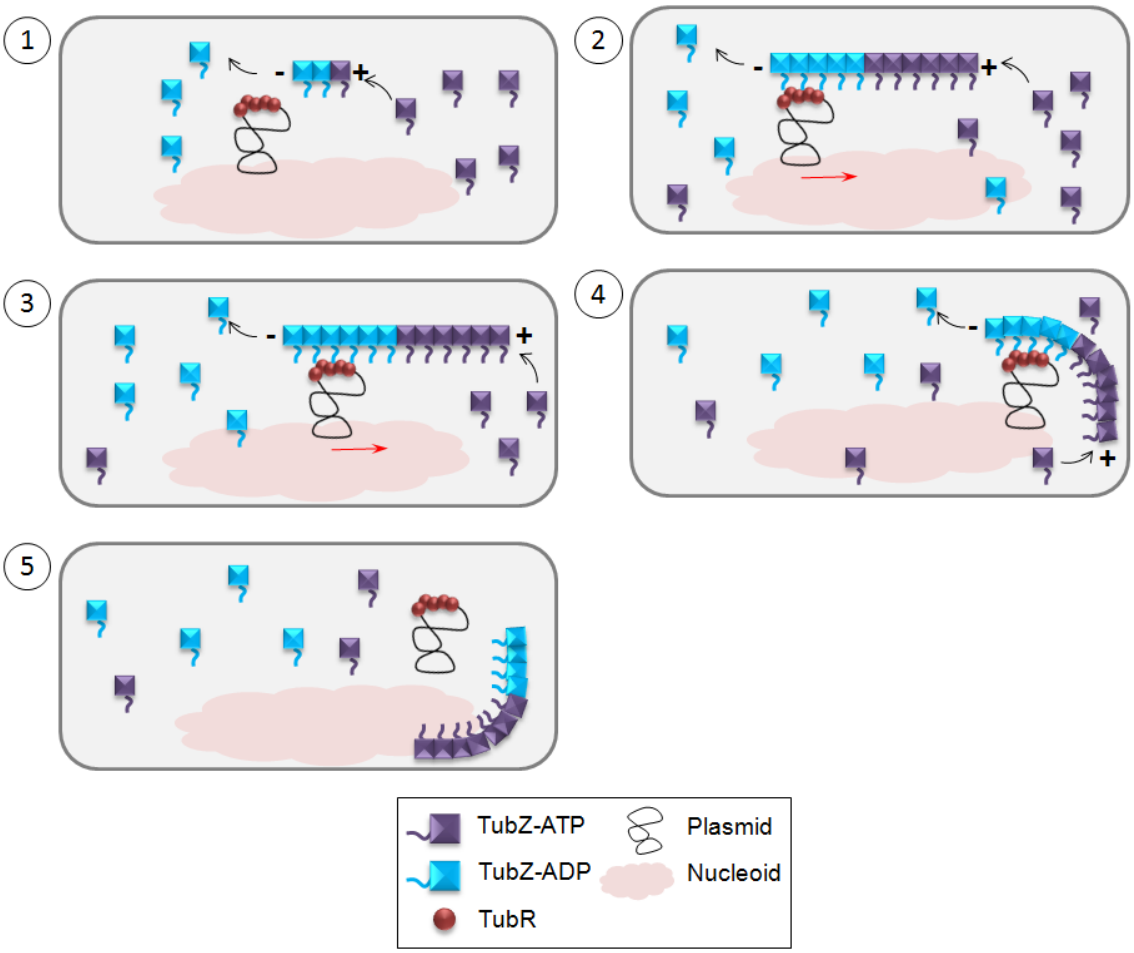

Figure 1.16 - A pulling-treadmilling mechanism for plasmid segregation mediated by Type III partition systems

(1) TubZ filaments have a polarity, they grow on the plus end and shrink on the minus end (2) TubZ filament grows and associates with the TubR-bound cargo via the C-terminal extension. TubRC binds preferentially to the minus pole of the treadmilling filament (3) TubZ filaments keep growing on the plus pole and shrink on the minus pole, moving by treadmilling. In this way the cargo is pulled to the pole. (4) Once at the pole, the filament bends. (5) The natural curvature of the membrane causes filament bending and dropping of the cargo at the pole.

1.7 Type IV partition systems

A type IV segregation system was first identified on *Staphylococcus aureus* multi-drug resistance plasmid pSK1. The system differs from the others because it is composed of only one protein that works as DNA-binding as well as motor protein (Simpson *et al.*, 2003). The protein does not show similarities to any other Par protein, therefore it has been categorised as belonging to a new group of plasmid segregation systems.

Although no structural information is available for this protein, structure prediction showed that the protein has a N-terminal HTH motif, which could recognise the centromeric site, and a central coiled-coil that could be involved in polymerisation (Schumacher, 2008). More investigation on the system needs to be done in order to understand whether this protein works alone or requires the presence of other factors and to determine a segregation mechanism.

1.8 Chromosomal ParABS systems

parABS cassettes were more recently identified also on bacterial chromosomes and were seen to play a role in chromosome segregation. Phylogeny of the chromosome-encoded Walker-type ATPase shows that all ParA proteins encoded on Gram-positive and Gram-negative bacteria chromosomes, bar a few exceptions, fall in the same subgroup that contains only one plasmid ParA from *Methanococcus jannaschii* pURB800. Therefore, chromosomal ParAs differ from plasmid ParA proteins and some of their functions could be chromosome-specific (Gerdes *et al.*, 2002). ParAB*S* systems are responsible for positioning the origin region of the newly replicated chromosomes at opposite poles of the cell before cell division, in a process similar to plasmid segregation. *C. crescentus* and *B. subtilis* are two of the bacterial species that encode chromosomal ParAB*S* systems, ParA/ParB in *C. crescentus* and Soj/Spo0J in *B. subtilis*. In contrast, both *E. coli* and *Haemophilus Influenzae* lack ParAB systems and the precise mechanism of their chromosome partitioning is still under investigation (Gerdes *et al.*, 2002). In *E. coli*, the interplay between different proteins, including the structural maintenance of the chromosome (SMC) MukBEF complex, MatP and Topoisomerase IV seems to drive sister chromosomes apart (Danilova *et al.,* 2007; Zawadzki *et al*., 2015; Nolivos *et al.*, 2016; Kleckner *et al.*, 2018).

1.8.1 *B. subtilis* chromosome segregation

Genes encoding homologues of plasmid-encoded ParA and ParB proteins were identified on the *B. subtilis* chromosome. Soj is a Walker type ATPase, while Spo0J is a DNA binding protein. Both proteins were initially characterised because they are involved in *B. subtilis* sporulation, as Soj inhibits the activation of Spo0A (transcription factor of many sporulation proteins) and Spo0J counteracts Soj activity leading to sporulation initiation (Ireton *et al.*, 1994). Conversely, during vegetative growth the proteins are involved in chromosome segregation. After replication, chromosomes are separated, with the origin (*oriC*) moved quickly towards the pole and the terminus remaining at midcell (Webb *et al.,* 1998). Both *spo0J* and *soj* genes are located close to the origin. *spo0J* null mutation causes a dramatic effect on chromosome segregation and increases 100 fold the number of anucleate cells (Ireton *et al.*, 1994). On the contrary, Soj deletion increases the number of anucleate cells only in *smc* null mutant strains (Ireton *et al.*, 1994).

Spo0J binds to ten 16-bp *parS* sites across the chromosome, eight in close proximity to the *oriC* and two further away (>1 Mb away). The protein can also spread around the centromeric sites, covering up to 18 kb (Breier and Grossman, 2007). After recognition

and binding to *parS*, Spo0J recruits SMCs to the partition sites, leading to chromosome structuring around the origin (Sullivan *et al.*, 2009).

Similarly to its ParB homologous, Spo0J is composed of an N-terminal domain, involved in interaction with its partner ParA, a central HTH domain, required for DNA binding and a C-terminal dimerization domain (Schumacher, 2008).
Mutation of the N-terminal domain of Spo0J affects the ability of the protein to spread, probably by interfering with protein oligomerisation (Breier and Grossman, 2007).
Recently, a C-terminal truncated version of Spo0J from *Helicobacter pylori* (Ct-*Hp*Spo0J), crystallised in complex with *parS,* shows that each Spo0J dimer interacts with one *parS* site through the central domain, while the N-terminal domains extend toward opposite directions. The same work also revealed that the protein oligomerises forming a tetramer, where each Spo0J monomer binds to a single *parS* site. Interaction between the monomers occur through the N-terminal domain, supporting the hypothesis that this region is involved in Spo0J spreading onto the DNA (Chen *et al.*, 2015). Mutagenesis analysis on B. subtilis Spo0J (*Bs*Spo0J) further supported the hypothesis that both *cis* and *trans* interactions between Spo0J dimers are required for ParB spreading and the formation of a higher order nucleoprotein complex (Song *et al.*, 2017).

In the past few years, investigation on how Spo0J binds and spread on the chromosome have provided new insight into the mechanism of chromosome segregation. Spo0J can interact specifically as well as non-specifically with the DNA, through two different regions of the C-terminal domain (Fisher *et al.,* 2017). Besides spreading, the protein can bridge distant sites on the chromosome, eventually leading to a higher nucleoprotein complex and, in cooperation with SMCs, overall DNA condensation (Broedersz *et al.*, 2014; Graham *et al.*, 2014; Taylor *et al*., 2015).

The Spo0J/*parS* nucleoprotein complex represents the unit recognised by the motor protein Soj. Fluorescence microscopy showed that Soj could oscillate over the nucleoid. When Spo0J was not expressed, Soj did not oscillate but it distributed evenly on the DNA (Marston and Errington, 1999). In a different study, Murray and Errington showed that, in the presence of Spo0J, Soj localised on the septa or as small foci in the cytoplasm. When Spo0J was not present, Soj localised non-specifically on the nucleoid. However, under no conditions Soj was seen to dynamically relocate on the nucleoid (Murray and Errington, 2008).

Analysis of the *Thermus thermophilus* Soj homologue showed that the protein changes its oligomerisation state upon interaction with the nucleotide. In the apo and ADP-bound form Soj is monomeric, while in the ATP-bound form the protein is dimeric (Figure 1.17). Soj-ATP dimers can interact with the DNA, forming extensive structures, and with the N-terminal domain of Spo0J. Spo0J can also enhance Soj weak ATPase activity, as its plasmid counterparts (Leonard *et al.*, 2005).
Soj performs multiple activities during cell division. The protein is also involved in initiation of DNA replication, since it can act as an activator or repressor of DnaA activity according to its oligomerisation state. Dimeric Soj-ATP can interact with DnaA and activate DNA replication, while Soj-ADP inhibits the process (Murray and Errington, 2008; Scholefield *et al*., 2011).

Despite the differences, the Soj-Spo0J chromosome segregation machinery presents many similarities with type I plasmid segregation systems. Recently, Le Gall *et al.* suggested a similar Brownian-ratchet mechanism within the nucleoid volume for *B. subtilis* chromosome and for the F plasmid (Le Gall *et al.,* 2016). Once at the poles, however, Spo0J anchors the origin to the cell membrane via the proteins RacA (**R**emodeling and **A**nchoring of the **C**hromosome) and DivIVA (Wu and Errington, 2003; van Baarle *et al.*, 2013; Ben-Yehuda *et al.*, 2005).
Chromosome and plasmid segregation differ in term of complexity. While a minimalistic three components system is sufficient to ensure plasmid segregation, chromosome partition involves additional proteins which ensure chromosome condensation and firm localisation at cell poles.

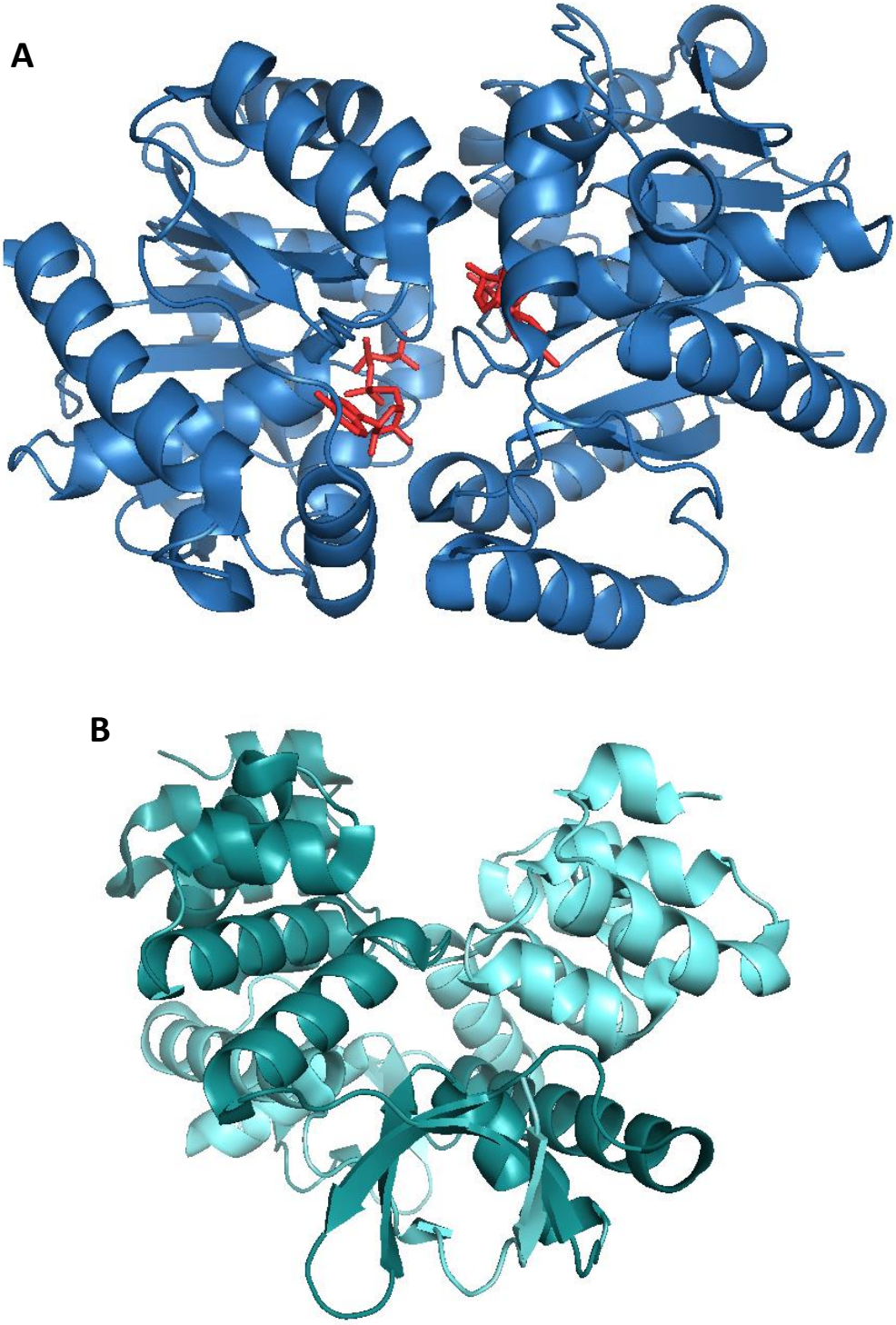

Figure 1.17 – Crystal structure of *T.thermophilus* Soj (A) and Spo0J (B)

(A) Crystal structure of Soj dimer. The ATP molecule is shown in red and is sandwiched between Soj monomers (PDB: 2BEK). (B) Crystal structure of Spo0J dimer. The two monomers are shown in different shades of cyan (PDB: 1VZ0). Images were generated using PyMOL v.1.8.6.0.

1.8.2 *C. crescentus* chromosome segregation

A partition system similar to Soj-Spo0J, was discovered also on the *C.Crescentus* chromosome. The partition system is composed of the two partition protein genes, *parA* and *parB*, as well as *parS* centromeric sites around the origin. *parA* and *parB* share around 50% sequence identity with *soj* and *spo0J* and were seen to be constitutive genes, as deletion impairs cell viability (Mohl and Gober, 1997).
Chromatin immunoprecipitation with deep sequencing showed that ParB recognises at least five *parS* sites clustered close to the origin, covering around 10 kb. ParB nucleates on the centromeric sites and spreads on the adjacent DNA (Tran *et al.,* 2018). The formation of a large nucleoprotein complex is essential to drive the origins to opposite poles of the dividing cell (Toro *et al.,* 2008). ParB interacts with a number of proteins. MipZ binds to ParB and prevents FtsZ ring formation by enhancing its ATPase activity (Thanbichler and Shapiro, 2006). At the poles ParB also interacts with PopZ, which tethers the nucleoprotein complex to the membrane (Bowman *et al*., 2010).

Movement of *oriC/parS*-ParB complex to the pole is mediated by the Walker-type ATPase ParA. The apo and ADP-bound form of the protein are monomeric and inactive. ATP binding promotes the formation of ParA-dimers, which are proficient in DNA and ParB binding (Ptacin *et al*., 2010). ParB interacts with ParA dimers and enhances its weak ATPase activity, leading to ParA reorganisation on the nucleoid (Figge *et al*., Gober, 2003; Ptacin *et al.,* 2010). ParA can also interact with TipN at the cell pole. Although the details of the interaction are lacking, TipN may direct ParA migration (Ptacin *et al.,* 2010).

ParA can form polymers in an ATP-dependent manner *in vivo* and *in vitro* (Ptacin *et al*., 2010). Originally this evidence lead to the hypothesis that the chromosome origin moves following retracting ParA polymers on the nucleoid, tethered to the membrane by TipN (Ptacin and Shapiro, 2010). However, this model was later replaced by a DNA-relay mechanism based on the diffusion-ratchet model. ParA-ATP dimers associate unevenly to the nucleoid and trap the ParB-*parS* complex around the chromosome origin. Interaction of ParA with ParB causes ATP conversion into ADP and ParA dimers dissociation and release from the DNA. The nucleoprotein complex at the origin diffuses away and interacts with the next available ParA-ATP dimers on the nucleoid, following a ParA gradient (Vecchiarelli *et al.*, 2010, 2014b). However, the protein gradient alone is not sufficient to ensure chromosome movement. Elasticity of the DNA was proved to be essential to allow proper orientation of the chromosome migration (Lim *et al*., 2014).

1.9 Involvement of host factors in plasmid segregation

Low copy number plasmids are self-replicating extra-chromosomal elements that encode their own segregation and stability systems. Despite being mostly independent units, the question of whether chromosome-encoded factors are involved in the process of plasmid partition has always been an object of interest. IHF was seen to play a role in the segregation of P1 plasmid, as it bends the *parS* site for ParB binding during the initial stage of segrosome assembly (Funnell, 1988b). IHF may also play a role in pB171 plasmid segregation, since the protein can bind to *parC2* on the *par2* site (Ringgaard *et al*., 2007). Nevertheless, to our knowledge, no other host-encoded factors have been seen to participate in plasmid segregation.

Based on the observation that plasmids are always positioned either at midcell or at cell quarters, Ebersbach *et al.* screened the binding of a number of cell division proteins with ParA and ParB from pB171 *in vivo*, using a bacterial two-hybrid system. None of the proteins tested, FtsZ, MreB, FtsA, FtsB, FtsI, FtsL, FtsN, FtsQ, FtsW and FtsX, showed binding to $ParA_{pB171}$ and $ParB_{pB171}$, even in the presence of the plasmid. The interaction of P1 ParA and ParB with FtsZ and MreB was also checked but no binding was recorded (Ebersbach *et al.,* 2006). Further testing of the interaction between the partition proteins and FtsZ *in vitro* also showed no binding.

MukB is a bacterial condensin involved in chromosome segregation in *E. coli*. Therefore, it was hypothesised that the protein could also be involved in plasmid segregation. Ezaki *et al.* showed that F plasmid is stably maintained in a MukB null mutant *E. coli* strain and that the plasmid can be properly segregated into anucleate cells (Ezaki *et al.*, 1991). Funnel and Gagnier performed a similar set of experiments to look at the segregation of the P1 plasmid in MukB mutant *E. coli* cells. As for F, P1 plasmid does not require MukB to be properly segregated, suggesting that plasmid segregation is independent from chromosome segregation (Funnell and Gagnier, 1995). MukB involvement was also tested for the plasmid R1, and again the condensin protein was seen to play no role in plasmid segregation (Weitao *et al.*, 2000).
Work by Yamaichi and Niki showed that the *B. subtilis* chromosome Soj-Spo0J-*parS* (ParA-ParB-*parS*) partition system can stabilise a plasmid in *E. coli*. When the *B.subtilis parABS* cassette was cloned into an unstable miniF plasmid where the *sopABC* cassette was deleted, plasmid stability increased. The plasmid was less stable compared to miniF harbouring the native *sopABC* cassette but could still be properly segregated. The authors attempted to increase plasmid stability by inserting multiple *parS* sites as on the *B. subtilis* chromosome, but this did not improve plasmid retention (Yamaichi and Niki,

2000). In addition, when plasmid position was studied by FISH (**F**luorescence **I**n **S**itu **H**ybridisation), the plasmid harbouring the Soj-Spo0J-*parS* system was seen to localise either at midcell or around cell quarter positions. Plasmid positioning was not as reproducible as for the miniF plasmid encoding the *SopABC* partition system, but a similar localisation pattern was observed. Yamaichi and Niki inferred that the plasmid encoding the *B. subtilis* chromosome segregation system is actively segregated in *E. coli*, however a host-encoded factor could be missing. They suggested that the segregation system may tether to membrane components at cell quarters, that could have different structures in *B. subtilis* and *E.coli*, explaining the lower efficiency of the Soj-Spo0J-*parS* partition system, when compared to *sopABC* (Yamaichi and Niki, 2000).

A similar experiment was performed to test the activity of the *P. putida* chromosome-encoded ParAB proteins in a miniF unstable plasmid, lacking its own partition system and encoding conserved sequences from *B. subtilis parS* site, also identified in *P. putida* chromosome close to *oriC*. Stability of the plasmid indeed increased, when *P. putida* partition proteins were included in the system (Godfrin-Estevenon *et al.*, 2002). Plasmids can therefore be stabilised by heterologous partition systems, although partition is not as efficient as with the native segregation apparatus. This suggests that host factors may not exist, or may be very conserved elements, found in both Gram-positive and Gram-negative bacteria.

1.10 Aims of the project

Segregation of the genetic material is an essential step for all living organisms and understanding how low copy number plasmids are partitioned at cell division is of fundamental importance. Not only are plasmids implicated in a number of diseases and confer multiple drug resistance to the host, but they also represent a model to understand chromosome segregation in bacteria.

The overarching aim of this project was to understand the interactions and interplay between the components of the TP228 plasmid segrosome. The TP228 partition cassette is characterised by genes for the proteins ParF and ParG, as well as the centromeric site *parH*. It is known that the N-terminal domain of ParG performs different roles in the segregation of the TP228 plasmid. Among those, the interaction with ParF is of particular interest. Previous work by Barge allowed the identification of a number of residues, whose changed into alanine, greatly affected plasmid segregation (Barge, 2015). The interaction between these mutant proteins and ParF was studied by **S**urface **P**lasmon **R**esonance (SPR) and **M**icro**S**cale **T**hermophoresis (MST), to ultimately understand the role played by these amino acids in plasmid segregation.

During the development of the project, a cluster of amino acids that was believed to interact with the centromeric site was identified. This led to the hypothesis that the N-terminal domain of ParG could become structured upon the interaction of these amino acids with *parH*. A combination of biochemical and biophysical assays was employed to test this hypothesis.

Finally, for years researchers have speculated that host-encoded factors could be involved in plasmid segregation. We investigated the interaction network of ParF and ParG by using a tandem affinity purification, combined to mass-spectrometry to identify putative interacting partners. Identified proteins were further investigated with *in vivo* approaches.

Overall, the project has provided further insights into the interactions required for the segregation of the TP228 plasmid.

Chapter 2
Materials and Methods

2.1 Bacterial strains and plasmids

2.1.1 Bacterial strains

The *E. coli* strains used in this work were obtained from glycerol stocks in the laboratory collection. Deletion mutants from the Keio collection were kindly provided by Prof. Peter McGlynn (University of York). Glycerol stocks were streaked on Luria-Bertani (LB) agar (with antibiotic selection when needed) and grown overnight at 37°C. One colony was selected, inoculated into 5 ml sterile LB and grown overnight at 37 °C. Competent cells were prepared as described in Section 2.3.9 and stored at -80 °C. Table 2.1 includes the strains used in this work and their specifications.

***E. coli* strain**	**Genotype**	**Application**
DH5α	F–Φ80*lac*ZΔM15 Δ(*lac*ZYA-*arg*F) U169 *rec*A1 *end*A1 *hsd*R17 (rK–, mK+) *pho*A *sup*E44 λ– *thi*-1 *gyr*A96 *rel*A1	Cloning and plasmid mini-prep
BL21(DE3)	*fhu*A2 [*lon*] *omp*T *gal* (λ DE3) [*dcm*] Δ*hsd*S	Protein overproduction
BW25113	Δ(*ara*D-*ara*B)*567*Δ(*rha*D-*rha*B)*568* Δ*lacZ4787*(::rrnB-3) *hsd*R514 *rph*-1]	Tandem affinity purification
BR825	*polA* In this strain the DNA polymerase gene was inactivated to support only low copy number replication (Ludtke,Eichorn and Austin, 1989)	Partition assay
SP850	*rel*A1 *spo*T(*cya*-1400)::Km *thi*-1 e14-λ-	Bacterial two-hybrid assay
BW25113 ΔdamX	Δ(*ara*D-*ara*B)*567*Δ(*rha*D-*rha*B)*568* Δ*lacZ4787*(::rrnB-3) *hsd*R514 *rph*-1], *damX::*Km^R (Baba *et al*., 2006)	Partition Assay
BW25113 ΔompC	Δ(*ara*D-*ara*B)*567*Δ(*rha*D-*rha*B)*568* Δ*lacZ4787*(::rrnB-3) *hsd*R514 *rph*-1], *ompC::*Km^R (Baba *et al*., 2006)	Partition Assay
BW25113 ΔompX	Δ(*ara*D-*ara*B)*567*Δ(*rha*D-*rha*B)*568* Δ*lacZ4787*(::rrnB-3) *hsd*R514 *rph*-1], *ompX::*Km^R (Baba *et al*., 2006)	Partition Assay
BW25113 ΔppiD	Δ(*ara*D-*ara*B)*567*Δ(*rha*D-*rha*B)*568* Δ*lacZ4787*(::rrnB-3) *hsd*R514 *rph*-1], *ppiD::*Km^R (Baba *et al*., 2006)	Partition Assay
BW25113 ΔydgA	Δ(*ara*D-*ara*B)*567*Δ(*rha*D-*rha*B)*568* Δ*lacZ4787*(::rrnB-3) *hsd*R514 *rph*-1], *ydgA::*Km^R (Baba *et al*., 2006)	Partition Assay

Table 2.1 - List of *E. coli* strains used in this work

2.1.2 Plasmids

Plasmids used in this work are described in Table 2.2. The plasmids were either available in the laboratory collection or constructed in the course of this work. The C-SF-TAPpcDNA3 plasmid was obtained from Dr. Christian Gloeckner (DZNE, University of Tübingen).

Plasmid	**Construction**	**Antibiotic selection**
pFH450	A pBR322 derivative having P1 and ColE1 origins of replication without any partition elements (Hayes, 2000)	Chloramphenicol
pFH547	A pBR322 derivative having P1 and ColE1 origins of replication with the wild type partition cassette *parFGH*(Hayes, 2000)	Chloramphenicol
pFH553	A pBR322 derivative having P1 and ColE1 origins of replication with the wild type partition cassette *parFGH*(Hayes, 2000)	Chloramphenicol
pET-22b(+)	An overexpression vector with the bacteriophage T7 promoter and (His)6-tag at the C-terminal end (Novagen)	Ampicillin
pDB-ParF	The plasmid partition gene *parF* cloned into vector pET-22b(+) (Barillà and Hayes, 2003b)	Ampicillin
pDB-ParG	The plasmid partition gene parG cloned into vector pET-22b(+) (Barillà and Hayes, 2003b)	Ampicillin
pET-ParGΔ9	N-terminal truncated ParG cloned into the vector pET22b(+) (Carmelo *et al*., 2005)	Ampicillin
pET-ParGΔ30	N-terminal truncated ParG cloned into the vector pET22b(+)(Carmelo *et al*., 2005)	Ampicillin

pET-MBL3A	*parGL3A* allele cloned into pET-22b(+) vector by using *NdeI* and *XhoI* sites (Barge, 2015)	Ampicillin
pET-MBK5A	*parGK5A* allele cloned into pET-22b(+) vector by using *NdeI* and *XhoI* sites (Barge, 2015)	Ampicillin
pET-MBK11A	*ParGK11A* allele cloned into pET-22b(+) vector by using *NdeI* and *XhoI* sites (Barge, 2015)	Ampicillin
pET-MBK12A	*ParGK12A* allele cloned into pET-22b(+) vector by using *NdeI* and *XhoI* sites (Barge, 2015)	Ampicillin
pET-MBM13A	*ParGM13A* allele cloned into pET-22b(+) vector by using *NdeI* and *XhoI* sites (Barge, 2015)	Ampicillin
pET-MBN18A	*parGN18A* allele cloned into pET-22b(+) vector by using *NdeI* and *XhoI* sites (Barge, 2015)	Ampicillin
pET-MBR19A	*parGR19A* allele cloned into pET-22b(+) vector by using *NdeI* and *XhoI* sites (Barge, 2015)	Ampicillin
pET-MBL21A	*parGL21A* allele cloned into pET-22b(+) vector by using *NdeI* and *XhoI* sites (Barge, 2015)	Ampicillin
pET-ParF-K160ER163E	*parFK160E-R163E* allele cloned into pET-22b(+) vector by using *NdeI* and *XhoI* sites (Allison-Gamble, 2016)	Ampicillin
pMBK11A	*parGK11A* allele cloned into pFH547 vector by using *ClaI* and *HpaI* sites (Barge, 2015)	Chloramphenicol
pT25-ParF	The plasmid partition gene *parF* cloned into pT25 (Barillà and Hayes, 2003b)	Ampicillin

pT18	A derivative of pBluescript II KS having a T18 fragment corresponding to the amino acids 225–399 of the catalytic domain of adenylate cyclase, CyaA (Karimova *et al*., 1998)	Chloramphenicol
pT18-ParG	The plasmid partition gene *parG* cloned into pT18 (Barillà and Hayes, 2003)	Chloramphenicol
C-SF-TAP pcDNA3	pcDNA3 mammalian expression vector containing the C-terminal *SF-TAP-tag* (Gloeckner *et al.,* 2007)	Ampicillin
pFH554	pFH547 derivate, where the ColE1 origin of replication has been deleted (Hayes, 2000)	Chloramphenicol
pFH450-LC	pFH547 derivate, where the ColE1 origin of replication has been deleted (McLeod, unpublished)	Chloramphenicol
pET-aspA	*aspA* gene cloned into the vector pET-22b(+) (Schumacher *et al.,* 2015)	Ampicillin
pBAD-parF	*parF* gene cloned into the pBAD24 vector by using *XbaI* and *PstI* sites (this work)	Ampicillin
pBAD-parF-SF	*SF-tag* cloned into the pBAD-parF vector by using *PstI* and *HindIII* sites (this work)	Ampicillin
pBAD-parG-SF	*parG* gene cloned into the pBAD-parF-SF vector by replacing *parF* with *parG* by using *XbaI* and *PstI* sites (this work)	Ampicillin
pBAD-aspA-SF	*aspA* gene cloned into the pBAD-parF-SF vector by replacing *parF* with *aspA* by using *XbaI* and *PstI* sites (this work)	Ampicillin

pBAD-SF	*SF-tag* cloned into the pBAD vector by using *PstI* and *HindIII* sites (this work)	Ampicillin
pCP-K11AK12A	*parGK11AK12A* allele cloned into pFH547 vector by using *ClaI* and *HpaI* sites (this work)	Chloramphenicol
pCP-K11Ak12AM13A	*parGK11AK12AM13A* allele cloned into pFH547 vector by using *ClaI* and *HpaI* sites (this work)	Chloramphenicol
pET-K11AK12A	*parGK11AK12A* allele cloned into pET-22b(+) vector by using *NdeI* and *XhoI* sites (this work)	Ampicillin
pET-K11Ak12AM13A	*parGK11AK12A* allele cloned into pET-22b(+) vector by using *NdeI* and *XhoI* sites (this work)	Ampicillin
pT18-FtsZ	*ftsZ* gene cloned into pT18 by using *XhoI* and *HindIII* restriction sites (this work)	Chloramphenicol
pT18-MreB	*mreB* gene cloned into pT18 by using *XhoI* and *HindIII* (this work)	Chloramphenicol
pT18-MukB	*mukB* gene cloned into pT18 by using *ClaI* and *ApaI* (this work)	Chloramphenicol
pT18-GyrB	*gyrB* gene cloned into pT18 by using *XhoI* and *HindIII* (this work)	Chloramphenicol

Table 2.2 - List of plasmids used in this work

2.2 Media and antibiotics

2.2.1 Media

2.2.1.1 Luria-Bertani (LB)

E. coli cells were grown either in liquid LB Broth (Fisher Scientific) or on solid LB agar (Formedium). Both media were prepared following the manufacturer's instruction by dissolving a defined amount of powder into an appropriate volume of distilled water. Table 2.3 shows a list of the components of LB media. Media were sterilised by

autoclaving at 121 °C for 20 minutes and then supplemented with antibiotics or specific inducer when required.

Component	Concentration g/L
Tryptone	10
Yeast extract	5
Sodium Chloride	10
Agar (solid medium only)	12

Table 2.3 – Components of Luria-Bertani Media

2.2.1.2 MacConkey

MacConkey medium (Oxoid) was used for the bacterial two-hybrid assay. The medium was prepared as per manufacturer instructions, by dissolving 40 g of powder into 1L of distilled water. To make the medium more solid, 6 g of Agar (Sigma Aldrich) were added to the mix. Table 2.4 shows a list of the components. MacConkey agar was sterilised by autoclaving at 121 °C for 20 minutes. MacConkey agar was prepared and used on the same day.

Component	Concentration g/L
Peptone	20
Lactose	10
Bile salts	5
Sodium chloride	5
Neutral red	0.075
Agar	12
pH 7.4 ± 0.2	

Table 2.4 – Components of MacConkey agar

2.2.2 Antibiotics

Stock solutions of antibiotics were prepared by dissolving an appropriate amount of powder in 10 ml Milli-Q water (Table 2.5). All antibiotics were filter sterilised using 0.22 µm syringe filters (Millex GP, Millipore) and stored at -20 °C.

Antibiotic	Stock concentration	Working concentration
Ampicillin (Amp)	100 mg/ml	100 µg/ml
Chloramphenicol (Cm)	30 mg/ml	30 µg/ml (cloning) 10 µg/ml (partition assay)
Kanamycin (Kan)	50 mg/ml	25 µg/ml

Table 2.5 – List of antibiotics and concentrations

2.3 Recombinant DNA techniques

2.3.1 Plasmid DNA isolation- small scale

One colony of DH5α *E. coli* cells transformed with the plasmid to be amplified was inoculated into 5 ml sterile LB, containing the appropriate antibiotic. The liquid culture was grown overnight at 37 °C with shaking. Cells were harvested by centrifugation of the culture at 11,000 x g for 2-3 minutes. The supernatant was discarded and plasmid was extracted using Macherey-Nagel Nucleospin Plasmid miniprep kit, following the manufacturer's instructions. To lyse the cells, pellets were resuspended in 250 µl of buffer A1. Afterwards, 200 µl of buffer A2 were added, the reaction was mixed by inversion and incubated at room temperature for five minutes. To the reaction, 300 µl of buffer A3 were added and mixed by inverting the tube 6-8 times. The reaction was centrifuged for 5-10 minutes at 11,000 x g. The supernatant containing the plasmid was loaded onto the column provided, to bind the DNA. The column was centrifuged for 1 minute at 11,000 x g and the flow through was discarded. The silica membrane of the column was then washed with 500 µl buffer AW and 600 µl buffer A4. In each step the column was centrifuged at 11,000 x g for 1 minute. At the end of the washes, the column was inserted into a new collection tube and the silica was dried by centrifugation at 11,000 x g for two minutes. The column was inserted into a fresh Eppendorf tube and the plasmid was eluted by applying 30-50 µl Buffer AE or Milli-Q water to the column and centrifuging it at 11,000 x g for 1 minute. This protocol was used for high copy number plasmids isolation. For low copy number plasmids, all the volumes in the protocol, including the liquid culture volume, were doubled. Purified plasmids were stored at -20 °C.

2.3.2 Polymerase chain reaction (PCR)

The DNA of interest was amplified by polymerase chain reaction (PCR) using Eppendorf "Mastercycler personal" PCR machine. The PCR reaction was prepared on ice following Table 2.6, with a final volume of 60 µl. 5 mM dNTP (deoxynucleotide triphosphate) solutions were prepared by diluting the 100 mM stock solutions (Roche) in Milli-Q water. Primers were designed according to the purpose and synthesised by

Sigma Aldrich (Table 2.8). Primers were resuspended in Milli-Q water and diluted to a final working concentration of 5 pmol/µl. The PCR program used is shown in Table 2.7. Annealing temperature was chosen according to the primers used, while extension time was selected according to the length of the DNA to amplify. Pfu enzyme was either purchased from New England Biolabs (NEB) or from Promega.

PCR reaction components	**Volume (µl)**
Template DNA (1:10 dilution in water)	1 µl
dNTPs (dATP, dTTP, dCTP, dGTP) (5 mM)	2.4 µl each
Primers (5pmol/µl)	3 µl forward 3 µl reverse
10 x Pfu Enzyme Buffer (+$MgCl_2$)	6 µl
Pfu Polymerase (2 U)	1 µl
Sterile milli-Q water	36.4 µl

Table 2.6 - PCR reaction components

PCR step	**Temperature**	**Time**
1. Initial denaturation	93 °C	3 min
2. Denaturation	93 °C	1 min
3. Annealing	40 - 70 °C (depending on the primers used)	1 min
4. Extension	72 °C	30 sec to 4 min (depending on the size of the DNA to amplify)
Repeat step 2 to 4 for 30 cycles		
5. Final extension	72 °C	6 min
6. Hold	10 °C	Indefinite

Table 2.7 – PCR thermocycler programme

Primer name	**Primer Sequence (5'→ 3')**	**Application**
ParF-TAP-Forward	CTCTCTTCTAGAATGAAAGTGATCTCATTT	Amplify *parF* for cloning in pBAD24
ParF-TAP-Reverse	CTCTCTCTGCAGCTCAAATATTCTAACTAT	Amplify *parF* for cloning in pBAD24
ParF-pBAD(XbaI)-Forward	TCTAGAATGAAAGTGATCTCATTTCTGAATCCG	Whole plasmid PCR to exclude *NcoI* site
ParF-pBAD(XbaI)-Reverse	ATATATTCTAGATTCCTCCTGCTAGCCCAAAAA	Whole plasmid PCR to exclude *NcoI* site
ParG-TAP-Forward	CTCTCTTCTAGAATGTCACTTGAAAAAGCGCAT	Amplify *parG* for cloning in pBAD24
ParG-TAP-Reverse	CTCTCTCTGCAGTTCGTTCTCTTTGAGCCAGTT	1. Amplify *parG* for cloning in pBAD24 2. PCR2 site directed mutagenesis
AspA-TAP-Forward	ATATATTCTAGAATGGGAAAAATATCCACAGACAAA	Amplify *aspA* for cloning in pBAD24
AspA-TAP-Reverse	ATATATCTGCAGTTGAACCACCTCTACTAATTTTCTAAT	Amplify *aspA* for cloning in pBAD24
SF-TAP-Forward	TATATACTGCAGATGTGGAGCCACCCTCAGTTCGAG	Amplify *SF* for cloning in pBAD24
SF-TAP-Reverse-	CTCTCTAAGCTTTCATTTATCATCATCATC	Amplify *SF* for cloning in pBAD24
ParF(middle)-Forward	CTTATCACCCGTAAGATAGAAATGGCAACCATG	PCR1 for site directed mutagenesis
K11AK12A-Forward	CATACGTCAGTAGCAGCAATGACCTTTGGTGAA	PCR2 for site directed mutagenesis
K11AK12A-Reverse	TTCACCAAAGGTCATTGCTGCTACTGACGTATG	PCR1 for site directed mutagenesis
K11AK12AM13A-Forward	CATACGTCAGTAGCAGCAGCGACCTTTGGTGAA	PCR2 site directed mutagenesis
K11AK12AM13A-Reverse	TTCACCAAAGGTCGCTGCTGCTACTGACGTATG	PCR1 siste directed mutagenesis
ParG1	GAGGAAACCATATGTCACTTGAAAAAGCG	Cloning in the pET vector
ParG2	TTCTTTCTCGAGTTCGTTCTCTTTGAG	Cloning in the pET vector
parH(biotin)-Forward	[Btn]GTCACTATACACTCATAAACT	Amplify parH for EMSA
parH(EMSA)-Reverse	TACTCTTTGTATGAGTAAAGG	Amplify parH for EMSA
parH(CD)-Forward	ACTCATAAACTCAAATGCTCACACACACATAAACCCATATGAGTAAACACTC	Produce dsDNA by annealing for CD and cross-linking

parH(CD)-Reverse	TGAGTATTTGAGTTTACGAGTGTGTGTGTATTTGGGTATACTCATTTGTGAG	Produce dsDNA by annealing for CD and cross-linking
FtsZ_forw	TATATACTCGAGATGTTTGAACCAATGGAACTT	Amplify *ftsZ* to clone in pT18
FtsZ_rev	GAGAGAAAGCTTCCATCAGCTTGCTTACGCAG	Amplify *ftsZ* to clone in pT18
GyrB_forw	ATATATCTCGAGATGTCGAATTCTTATGACTCC	Amplify *gyrB* to clone in pT18
GyrB_rev	ATATATAAGCTTCCAATATCGATATTCGCCGC	Amplify *ftsZ* to clone in pT18
MreB_forw	TATATCTCGAGATGTTGAAAAAATTTCGTGGC	Amplify *mreB* to clone in pT18
MreB_rev	TATATAAAGCTTCCCTCTTCGCTGAACAGGTC	Amplify *mreB* to clone in pT18
MukB_forw	ATATATGGGCCCATGATTGAACGCGGTAAATTT	Amplify *mukB* to clone in pT18
MukB_rev	ATATATTCGATCCACTCGCCTGAGAAGGCGC	Amplify *mukB* to clone in pT18
T18_forw	ATATCGAATTCCGCCGCCAGC	Sequencing pT18 clones
GyrB_1681bp_FW	CTGGACGGCGCAACGCTGCAC	Sequencing
GyrB_865bp_FW	ATGACCCGTACCCTGAACGCC	Sequencing
MreB_721bp_FW	CTGGCAGAAGGTGTTCCACGC	Sequencing
MreB_3394081	TTCAACCTGGGTCGCGCCAAC	Sequencing
FtsZ_102481	GAGCGGTGAAGACCGTGCGGA	Sequencing
MukB_972661	GAGCTACATACTTCGCGTCAG	Sequencing
MukB_973321	CAGGTTCAGCCGCTGCGGATG	Sequencing
MukB_974221	CTGCATCAGGCGTTCAGC	Sequencing
MukB_975061	CGCAAACTGCGCAAGCTGGAG	Sequencing
MukB_975721	AAACTGTATCAACGTCTTAAC	Sequencing

Table 2.8 - Primers list

2.3.3 Restriction enzyme digestion

Restriction enzyme digestion was used to construct recombinant plasmids and to check positive clones. For cloning purposes, restriction sites were added to the primers used to amplify the gene of interest. Plasmid vector and PCR insert were then digested with the same restriction enzymes. Double digestion reactions contained the DNA substrate (usually around 0.5-0.8 µg), 1-2 µl of each of the enzyme, 1 x enzyme buffer (suitable for both enzymes, as per manufacturer's instructions) and appropriate amount of sterile Milli-Q water to make 30-60 µl final volume. Digestion reactions were usually carried out at 37 °C for 2 hours and stopped by heat inactivating the enzymes at 65°C or 80°C for 20 minutes. All enzymes used were purchased from NEB, Promega or Thermo Fisher Scientific.

2.3.4 Ethanol precipitation

Ethanol precipitation was used as a purification step to clean the DNA from proteins and salts contained in the solution after PCR and restriction enzyme digestion. The volume of the DNA solution was adjusted to 200 µl with Milli-Q water. 20 µl (1/10 of the volume) of 3M Sodium acetate (pH 5.3) was added, together with 440 µl (2x the volume) of ice cold 100% ethanol. The solution was gently mixed and incubated at -20 °C for a few hours (from 2 hours to overnight). The solution was centrifuged for 30 minutes at 4 °C at 13,000 rpm. The supernatant was removed carefully without touching the DNA pellet at the bottom of the tube. The pellet was resuspended with 500 µl of ice cold 70% ethanol and centrifuged 10 more minutes at 13,000 rpm at 4°C. The supernatant was removed and residual ethanol was left to dry for either 20-30 minutes at room temperature or 5 minutes at 37 °C. The DNA pellet was resuspended in 30 µl Milli-Q water and stored at -20 °C.

2.3.5 Alkaline phosphatase treatment of DNA

Alkaline phosphatase treatment was used to dephosphorylate the plasmid digested with restriction enzymes. This step was necessary to prevent re-circularisation of the vector during ligation. The dephosphorylation reaction was set up in 100 µl final volume. To the restriction digestion reaction, 2 µl of alkaline phosphatise (NEB), 10 µl of 10X alkaline phosphatasebuffer and an appropriate volume of sterile Milli-Q water, to make up a 100 µl total volume, were added. The reaction was incubated at 37°C for 1 hour in total. After the first 30 minutes, 2 µl of alkaline phosphatise were added to the reaction and incubated for 30 more minutes. After incubation with the enzyme, 10 µl of 200 mM ethylene glycol tetraacetic acid (EGTA) was added to the reaction and incubated at 75°C for 10 minutes. The DNA was then purified using Macherey-Nagel PCR clean-up and gel

extraction kit, following manufacturer's instructions. To the 100 µl dephosphorylation reaction, 200 µl of NTI buffer were added. The reaction was mixed, loaded onto a column provided with the kit and centrifuged at 11,000 x g for 1 minute. The silica membrane was washed two times by adding 700 µl NT3 buffer and centrifuging for 1 minute at 11,000 x g. To remove residual buffer, the column was centrifuged for 2 more minutes at 11,000 x g. The DNA was eluted by adding 30 µl NE buffer to the column and centrifuging it for 1 minute at 11,000 x g. The DNA was then stored at -20°C if not used on the day.

2.3.6 Agarose gel electrophoresis

DNA samples were analysed on 1% (w/v) agarose gel electrophoresis. When smaller fragments of DNA were visualised, 1.2-1.5% (w/v) agarose concentration was used to obtain a sharper resolution of the bands. 1% agarose gel was prepared by dissolving 0.5 g of agarose in 50 ml 1X TAE buffer (40 mM Tris, 20 mM acetic acid, 1 mM Ethylenediaminetetraacetic acid (EDTA)). SYBR safe DNA stain was added to the molten agarose at a 10,000-fold dilution to allow visualisation of the DNA bands under UV light. 6X DNA-loading buffer (10 mM Tris-HCl (pH 7.6) 0.03% bromophenol blue, 0.03% xylene cyanol FF, 60% glycerol 60 mM EDTA) was added to the samples (10-20 µl) to a final 1X concentration. A DNA ladder was run in parallel to the samples to determine the size of the DNA under investigation. GeneRuler 1kb DNA ladder (Thermo Fisher Scientific) was used for larger DNA pieces, while PCR marker (New England Biolabs) was used for smaller DNA fragments. Agarose gels were run for 50 minutes at 90-100 V in TAE buffer at room temperature. DNA bands were visualised using a BIORAD Gel doc EZ imager. Image lab 4.0.1 software was used to analyse the images.

2.3.7 Gel extraction

DNA samples were run and separated on 1% (w/v) agarose gel. The DNA was visualised using a UV transilluminator and the band of interest was excised using a scalpel. DNA was extracted using the Macherey-Nagel PCR clean-up and gel extraction kit, following the manufacturer's instructions.

The gel piece was weighed in an Eppendorf tube and two volumes of buffer NTI were added to it (200 µl for 100 mg of gel). The sample was then incubated for 10 minutes or longer at 50°C to dissolve the agarose. Once the agarose was completely dissolved, the sample was transferred into a NucleoSpin column and centrifuged for 30 seconds at 11,000 x g. The DNA on the silica membrane was then washed twice with 700 µl NT3 buffer (30 seconds, 11,000 x g). To remove residual buffer, the silica was dried by centrifugation at 11,000 x g for 1 minute. The sample was eluted by adding 30 µl NE

buffer and centrifuging the column for 1 minute at 11,000 x g. If not used on the day, the sample was stored at -20 °C.

2.3.8 DNA ligation

After restriction digestion of the vector and the insert (section 2.3.3), the vector was dephosphorylated as described in Section 2.3.5 to prevent re-circularisation. For the ligation reaction, vector and insert were added in a 1:9 ratio for small inserts, 1:1 or 1:3 ratio for larger inserts. DNA concentration was quantified using a Nanodrop 1000 (Thermo Fisher Scientific). To an appropriate volume of insert and vector, 1 µl T4 ligase (NEB) (2.5 U) and 3 µl 10X T4 ligase buffer were added. The reaction volume was adjusted to 30 µl with Milli-Q water and incubated at room temperature for 3 hours. After incubation, DH5α *E. coli* competent cells were transformed with the ligation reaction and grown overnight at 37 °C (Section 2.3.10).

2.3.9 Preparation of competent cells

E. coli cells from the glycerol stock were streaked in aseptic conditions on a LB plate (containing antibiotics if necessary) and grown overnight at 37°C. One colony from the plate was selected and inoculated into 10 ml sterile LB (with antibiotics if needed) and grown overnight at 37°C. An aliquot of 300 µl of the liquid culture was inoculated into 60 ml liquid LB (with antibiotics if required) and grown at 37 °C until it reached the appropriate growth level. Growth of the liquid culture was checked by measuring the OD_{600} (optical density at λ=600 nm) against a blank sample containing LB only. When the OD_{600} reached a value between 0.4 and 0.6, the liquid culture was incubated on ice for ten minutes. Cells were then harvested by centrifugation at 2,000 x g for 5 minutes at 4°C and the supernatant was discarded. Cells were resuspended in 20 ml of chilled RF1 (Table 2.9) and incubated on ice for 30 minutes. The cells were then harvested again by centrifuging the culture at 2,000 x g for 5 minutes and discarding the supernatant. Cells were resuspended in 4.8 ml of chilled RF2 (Table 2.9) and incubated on ice for 15 minutes. Chemically competent cells were then divided into 500 µl aliquots and stored at -80 °C.

Chemical	**RF1**	**RF2**
Glycerol	15%	15%
RbCl	100 mM	10 mM
$MnCl_2$	50 mM	---
KCH_3CO_2	30 mM	--

CaCl2	10 mM	75 mM
MOPS	---	10 mM
pH	5.8	6.8

Table 2.9 - Composition and pH of RF1 and RF2 solutions for chemically competent cells preparation

All buffers were filter sterilised and kept in aseptic conditions.

2.3.10 Transformation of competent cells

For transformation of *E. coli* cells, 100 µl of chemically competent cells were incubated with 1 µl of purified plasmid or 30 µl of the ligation product for 40 minutes on ice. Cells were then heat shocked at 42°C for 90 seconds. Before incubation at 37°C for 1 hour, 400 µl of sterile liquid LB were added to the cells. A 100 µl aliquot of the cells was spread on LB agar and grown at 37°C overnight.

2.3.11 Screening of recombinant plasmid

After cloning, recombinant plasmids were tested by colony PCR and by restriction digestion. The final step of cloning involved the transformation of *E. coli* DH5α cells with the ligation product. Ten clones were isolated from the transformation plate and checked by colony PCR as described in Section 2.3.2. Bacterial colonies were used as template in the PCR reaction, rather than purified DNA. Primers included in the PCR reaction were the same as the ones used for cloning. If the recombinant plasmid was correctly constructed, the insert was amplified by PCR. Colony PCR was checked on 1% (w/v) agarose gel. After colony PCR screening, the recombinant plasmid was extracted as per Section 2.3.1 and digested with the same restriction enzymes used for cloning, as per Section 2.3.3. Restriction digestion of the recombinant plasmid was checked on 1% (w/v) agarose gel, to confirm the presence of the correct size insert. Plasmids showing the insert of interest were sequenced so that the integrity and the accuracy of the construct sequence was verified.

2.3.12 DNA sequencing

Sanger Sequencing was performed by GATC Biotech (currently Eurofins Genomics), Ebersberg, Germany. The reactions to be sequenced contained 5 µl plasmid (80-100 ng/µl) and 5 µl of primer (5 pmol/µl). Table 2.8 shows the list of primers used. Sequences were analysed using Chromas Software (Technelysium).

2.4 Mutagenesis and cloning

2.4.1 Site-directed mutagenesis by overlapping extension PCR for ParG mutagenesis

Site-directed mutagenesis was used to insert double and triple mutation in the codons that translate into the middle cluster of ParG N-terminal domain. Mutagenesis aimed to create recombinant *parG* sequences where the K11-K12 and K11-K12-M13 codons were replaced by alanine codons.

Mutagenesis by overlapping extension PCR involved three PCR steps and four primers (Figure 2.1). Primers a and b were used in PCR1, while primers c and d in PCR2. Primers a and d are external and anneal in the middle of *parF* and end of *parG* respectively. Primers b and c are internal, harbour the mutation to insert and anneal at the beginning of *parG*. PCR1 and PCR2 were carried out as described in Section 2.3.2. The plasmid pMBK11A was used as template (Barge, 2015).

After PCR, the products were run on 1% (w/v) agarose gel and extracted from the gel as described in Section 2.3.7. PCR1 and PCR2 products were mixed together in a 1:1 ratio and a pre-cycle reaction was run to allow the overlapping sequences to anneal. Pre-cycle reaction included equal amount of PCR1 and PCR2 products, 2.4 µl of each of the dNTPs, 10 µl of 5X GoTaq G2 buffer, 3.6 µl of 25 mM $MgCl_2$ and water to make up to 50 µl reaction. After initial denaturation at 93°C for 3 minutes, 1 µl of GoTaq G2 polymerase (2.5 U) was added and the pre-cycle started. Pre-cycle consisted of 20 cycles of 40 seconds at 94 °C and 40 seconds at 72 °C. At the end of the pre-cycle, 5 µl of each of the forward and reverse external primers (5 pmol/µl) were added and PCR3 was performed as per Section 2.3.2. After PCR, the DNA product was gel extracted (Section 2.3.7) and purified by ethanol precipitation (section 2.3.4). PCR3 product was then digested with the restriction enzymes ClaI and HpaI which cut inside *parF* sequence and at the end of *parG*. The vector pFH547 was also digested with the same enzymes and dephosphorylated as per Section 2.3.5. Vector and insert were then run on 1% agarose gel and extracted from the gel using Macherey-Nagel PCR and gel extraction kit. Ligation of the two was then performed as per Section 2.3.8 and checked as per Section 2.3.11. Mutation was then confirmed by Sanger sequencing by GATC biotech. Once the sequence of *parG-K11A-K12A* was confirmed, the newly constructed plasmid was used as template for the construction of *parG-K11A-K12A-M13A* in the pFH547 vector, following the same protocol.

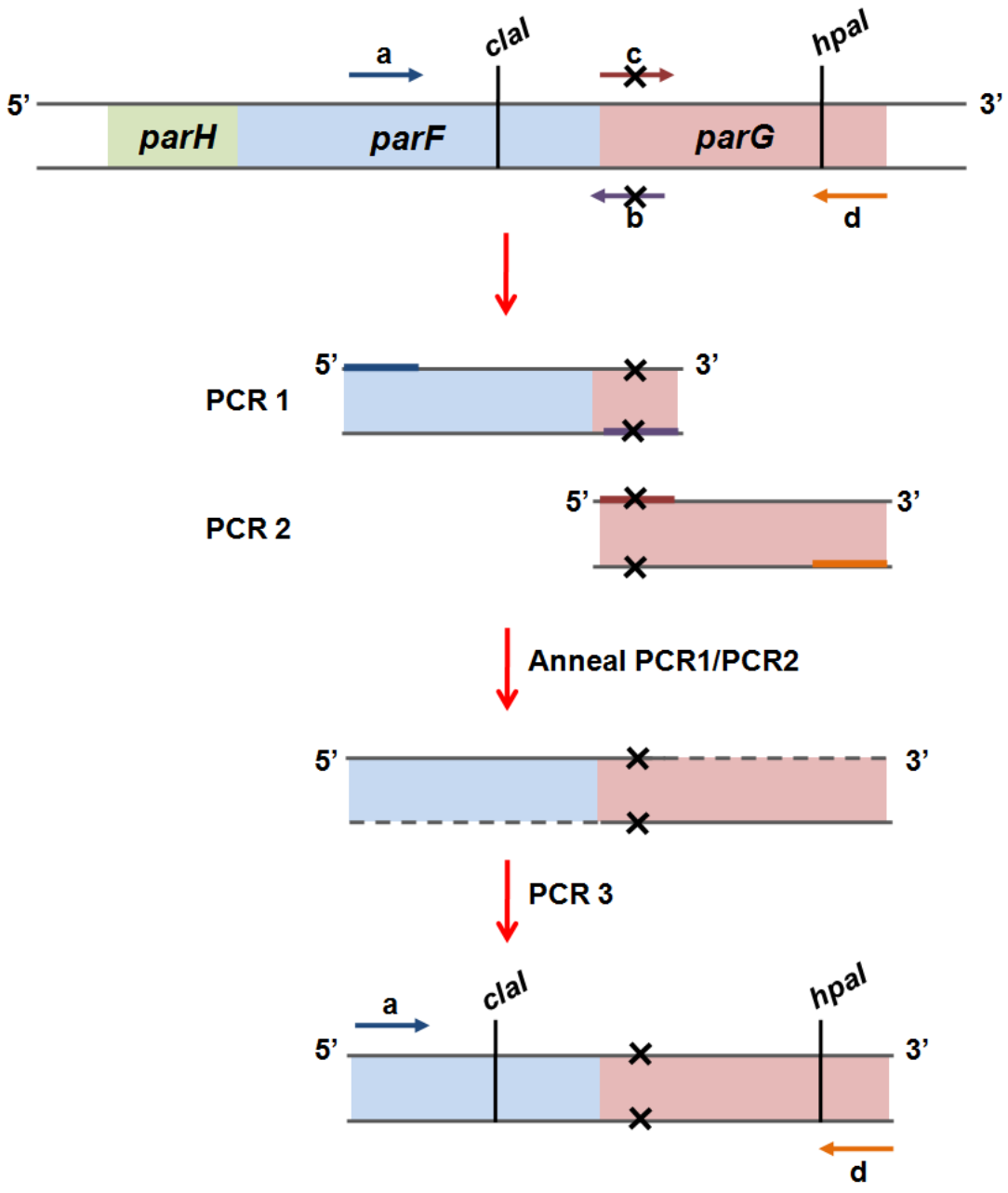

Figure 2.1 - Outline of site-directed mutagenesis by overlapping PCR

Site-directed mutagenesis by overlapping PCR involves three steps. Primers are designed so that the internal primers harbour the mutation to insert into the gene (b and c). Two PCR runs are carried out: PCR1 with the primers a (forward external-blue) and b (reverse internal-violet) and PCR2 with the primers c (forward internal-red) and d (reverse external-orange). PCR1 and PCR2 products are annealed together shortly, before amplification through PCR3 using the two external primers a and d.

2.4.1 Cloning ParG mutant in pET-22b(+) vector

The *parG* mutant alleles constructed in the pFH547 vector were cloned into the pET-22b(+) vector for gene over-expression and protein production.

Forward and reverse primers were designed so that they annealed at the beginning and at the end of *parG* sequence and harboured the NdeI and XhoI restriction site respectively. *parG* mutant alleles were amplified from the pFH547 plasmid derivates pCPK11AK12A and pCPK11AK12AM13A (constructed in this work). The PCR product was purified by ethanol precipitation. Both the vector pET-22b(+) and PCR inserts were then digested with the restriction enzymes NdeI and XhoI. After digestion, the vector was dephosphorylated. Both vector and inserts were run on 1% (w/v) agarose gel and were gel extracted following the protocol in Section 2.3.7. Ligation was carried out as per Section 2.3.8 and clones were screened as per Section 2.3.11.

2.4.2 Cloning fusion proteins for Tandem Affinity Purification

parF, *parG* and *aspA* genes were cloned into pBAD24 vector fused to a sequence encoding a C-terminal SF-tag for tandem affinity purification (TAP). Cloning of *parF* in frame with the SF-tag was performed in two steps. *parF* was amplified by PCR from the plasmid pFH553 (Section 2.3.2). Primers were designed containing the XbaI restriction site upstream of the gene and the PstI restriction site downstream. After amplification, the PCR product was purified by ethanol precipitation as described in Section 2.3.4. Both PCR product and pBAD24 vector were digested with XbaI and PstI. The digested PCR product was purified by ethanol precipitation. The vector was subjected to alkaline phosphatase treatment and purified using Macherey-Nagel (MN) PCR clean up and gel extraction kit (Section 2.3.5). The digested vector and PCR insert were ligated together as per Section 2.3.8. After ligation, DH5α competent cells were transformed with the ligation reaction and grown at 37°C overnight. Clones were screened as per Section 2.3.11. pBAD-parF-SF was constructed by cloning the sequence encoding the SF-tag in frame with *parF* in the pBAD-parF plasmid. The SF-tag sequence was amplified from the C-SF-TAP pcDNA3 plasmid obtained from the Gloeckner lab (Gloeckner *et al.*, 2007), using primers which harboured the PstI (forward) and the HindIII (reverse) restriction sites. The PCR product was purified by ethanol precipitation, then digested with PstI and HindIII restriction enzymes. The pBAD-parF vector was digested using the same enzymes and dephosphorylated. Both the vector and the PCR insert were then purified using either ethanol precipitation or Macherey-Nagel PCR clean up and gel extraction kit. The SF-tag insert was then ligated into the pBAD-parF vector as described in Section 2.3.8. DH5α cells were transformed (section 2.3.10) with the second ligation reaction and grown overnight at 37°C. Clones were screened as per Section 2.3.11. After plasmid

construction, the NcoI restriction site was excluded from the plasmids' multicloning site (MCS) to prevent transcription from the ATG inside its sequence. This was done by performing whole plasmid PCR and excluding the NcoI restriction site. Primers were designed to exclude the *ncoI* sequence (e.g. forward primer annealed downstream and reverse primer annealed upstream of *ncoI*) and harbouring the XbaI restriction site. The plasmid was amplified by PCR as per Section 2.3.2. The product was run on 1% (w/v) agarose gel and extracted as per Section 2.3.7. Ethanol precipitation (Section 2.3.4) was performed on the sample to further purify it. The plasmid was then digested with XbaI and purified by ethanol precipitation. The plasmid was then re-ligated to itself following the protocol in Section 2.3.8. Cloning was checked as per Section 2.3.11 and sequence was confirmed by Sanger Sequencing by GATC Biotech.

pBAD-*parG-SF* and pBAD-*aspA-SF* were constructed by swapping *parF* with *parG* and *aspA* respectively. *parG* and *aspA* were amplified by PCR from pFH547 and pET-aspA templates using primers harbouring XbaI (forward) and HindIII (reverse) restriction sites. PCR products were purified by ethanol precipitation (Section 2.3.4). pBAD-parF-SF was digested with the same restriction enzymes and run on 1% (w/v) agarose gel before being extracted and purified. Ligation of the vector and the insert was carried out as described in Section 2.3.8 and plasmids were screened as per Section 2.3.11, before being sequenced.

2.4.5 Cloning genes in the pT18 vector

A few of the proteins identified through tandem affinity purification were tested for binding *in vivo* with ParF using a bacterial two-hybrid assay. The genes *ftsZ*, *mreB*, *gyrB* and *mukB* were cloned into the pT18 plasmid upstream of the T18 moiety. To insert *ftsZ*, *mreB* and *gyrB* into the pT18 vector, the genes were amplified by PCR using genomic DNA as a template and primers harbouring XhoI (forward) and HindIII (reverse) restriction sites. The *mukB* gene was amplified from genomic DNA with plasmids harbouring the ClaI (forward) and ApaI (reverse) restriction sites. The plasmid and the inserts were either digested with XhoI and HindIII or with ClaI and ApaI. Both insert and vector were analysed on 1% (w/v) agarose gel and extracted from the gel (Section 2.3.7). After extraction samples were also purified by ethanol precipitation. Ligation was carried out as described in Section 2.3.8 and clones screened as per Section 2.3.11.

2.5 Proteins preparation

2.5.1 ParF overproduction and purification

To produce His-tagged ParF, BL21(DE3) cells were transformed with the pET-parF vector and grown overnight at 37 °C on selective LB agar (100 µg/ml ampicillin). A few

isolated colonies (8-10 colonies) were picked with a sterile loop and inoculated into 10 ml sterile LB containing ampicillin. The culture was first grown for 1-2 hours at 37 °C in a 250 ml flask and then inoculated into 300 ml sterile selective LB (100 µg/ml ampicillin) in a 2 L flask. The culture was grown at 37 °C until the optical density of the culture (OD_{600}) reached a value between 0.8-0.9 and it was then induced with isopropyl β-D-1-thiogalactopyranoside (IPTG) at a final concentration of 1 mM. Overexpression and protein overproduction was carried out at 30°C for 3 hours, since this temperature was shown to reduce the amount of ParF in the inclusion bodies. Cells were then harvested for 30 minutes at 11,000 rpm at 4 °C, using Sorvall Evolution Superspeed Centrifuge. The supernatant was discarded and pellet were stored at -20°C to be used the day after or after a short period of time.

Protein overproduction was checked by SDS-PAGE (as per Section 2.5.4) on 12% acrylamide gel. A sample of 100 µl was collected from the liquid culture before and after 3 hours of IPTG induction. Samples were centrifuged shortly at 13,000 rpm, the supernatant was removed and the pellets were resuspended in 50 µl 1X SDS-Loading Buffer. 25 µl of the sample were loaded on SDS-PAGE. The gel was run for 1 hour at 150-190 V, then stained with Coomassie staining and destained with Methanol-Acetic acid destaining solution.

The cell pellets were resuspended in 15 ml Binding Buffer (Table 2.10). One tablet of protease inhibitor cocktail (cOmplete™, Mini, EDTA-free Protease Inhibitor Cocktail, Roche) and 150 µl of 10 mg/ml lysozyme were added to the resuspension. The reaction was incubated at 30 °C for 15 min. Before further 15 minutes incubation at 30 °C, 150 µl of lysozyme were added to the cells. Cell resuspension was then cooled down on ice before sonication. To lyse the cells, the cell resuspension was sonicated 12 times for 15 seconds (power 60%) with one minute interval using a micro-tip. Cells were kept in iced water during the sonication process.

The sonication product was then centrifuged at 4 °C for 1h at 11,000 rpm. This low speed was used to keep ParF in solution. The crude extract was loaded on a Ni^{2+}-charged 2.5 ml His-bind resin (Merck) column. The resin was washed with water, charged with 50 mM $NiSO_4$ and equilibrated with binding buffer before being used. The cell extract was flowed over the column in a loop for about 1.5 hour to allow all the protein to bind. The column was then washed with 30 ml binding buffer, and with 70 ml wash buffer (Table 2.10) to remove non-specifically bound proteins. ParF was then eluted with 12 x 1 ml elution buffer (Table 2.10). The elution fractions were collected separately and quantified by Bradford assay as per section 2.5.3. Since plasmid stability assay confirmed that the activity of ParF-His was identical to that of native ParF, the His-tag was not removed (Barillà and Hayes, 2003).

Elution fractions were buffer exchanged using a 5 ml Hi-trap desalting column (GE Healthcare). The column was first washed with water and equilibrated with 8 CVs (column volume) of storage buffer (Table 2.10). Subsequently, 1.5 ml of protein was applied onto the column and eluted with 2 ml storage buffer in 2 x 1 ml aliquots. The process was repeated for all the high concentration elution aliquots. Protein concentration was measured by Bradford assay (Section 2.5.3) before storage. The column was washed and stored in ethanol to be reused. The protein was divided in 200 µl aliquots, flash frozen in liquid nitrogen and stored at – 80 °C.

2.5.2 ParG overproduction and purification

To produce His-tagged ParG, BL21 (DE3) cells were transformed with the pET-parG vector and grown overnight at 37°C on selective LB Agar (100 µg/ml). Colonies were isolated from the transformation plate and inoculated into sterile LB containing ampicillin (100 µg/ml). Cells were grown at 37 °C for 1-2 hours in a 250 ml flask and then transferred into 2 L flask containing 300 ml selective sterile LB (100 µg/ml ampicillin). The cell culture was grown until the OD_{600} reached a value around 0.8 – 0.9 and was then induced with IPTG at a final concentration of 1 mM. Protein induction was carried out for 3 hours at 37 °C. 100 µl samples were collected before and after induction to test protein production on 15% SDS-PAGE, before purification (as in section 2.5.1).
Cells were then harvested by centrifugation at 11,000 rpm for 30 minutes at 4°C. The cell pellets were used the day after or stored at -20°C for a short period of time.

To extract the crude extract, the cell pellets were resuspended in 10 ml binding buffer (Table 2.10). One tablet of protease inhibitor cocktail (cOmplete™, Mini, EDTA-free Protease Inhibitor Cocktail, Roche) and 100 µl of lysozyme (10 mg/ml) were added to the resuspension. The mixture was incubated for 15 minutes at 30 °C. Further 100 µl of lysozyme were then added to the cells, before 15 more minutes of incubation at 30°C. Prior to sonication, the cell-resuspension was cooled down on ice. Cells were then kept in iced water during the sonication process. Cells were sonicated 12 times for 15 seconds (power 60%), with 1 minute interval between sonication steps. The sonication product was then centrifuged for 50 minutes at 12,000 rpm to remove cell debris and obtain the cleared lysate.
A Ni^{2+}-charged column was prepared as per 3.5.1. The crude extract was loaded on the column and flowed over it in a closed loop for 1.5 hours to allow all the protein to bind. To remove non-specifically bound proteins, the column was washed with 30 ml binding buffer and 50 ml wash buffer (Table 2.10). The protein was eluted with 12 x 1 ml elution buffer (Table 2.10). Protein concentration was assessed by Bradford assay (Section

3.5.3). As for ParF, ParG-His activity was shown to be similar to that of native ParG, therefore the tag was not removed (Barillà and Hayes, 2003).

High concentration aliquots were buffer exchanged using a pre-packed 5 ml Hi-trap desalting column (GE Healthcare). The column was first washed with water and equilibrated with 8 CVs storage buffer (Table 2.10). 1.5 ml of protein were added to the column and eluted with 2 x 1 ml of storage buffer. The two steps were repeated for the other elution aliquots. Protein concentration was assessed by Bradford assay (Section 2.5.3). The protein was divided in 200 µl aliquots, flash frozen and stored at – 80 °C.

Protein Purification Buffer	**ParF proteins**	**ParG proteins**
Binding Buffer	50 mM TRIS-HCl pH 7.5-8.0 500 mM NaCl 15 mM Imidazole 10% Glycerol	20 mM TRIS-HCl pH 7.5-8.0 500 mM NaCl 15 mM Imidazole 10% Glycerol
Wash Buffer	50 mM TRIS-HCl pH 8.0 1M NaCl 85 mM Imidazole 10% Glycerol	20 mM TRIS-HCl pH 8.0 1M NaCl 90 mM Imidazole 10% Glycerol
Elution Buffer	50 mM TRIS-HCl pH 8.0 150 mM NaCl 300 mM Imidazole 10% Glycerol	20 mM TRIS-HCl pH 8.0 500 mM NaCl 400 mM Imidazole 10% Glycerol
Storage Buffer	30 mM TRIS-HCl pH 7.5-8.0 100 mM KCl 10% Glycerol 2 mM DTT	50 mM HEPES pH 8.0 50 mM KCl 10% Glycerol 2 mM DTT

Table 2.10 - Protein purification buffers

All buffers were pH adjusted and filter sterilised using 0.22 µm filter (Millipore)

2.5.3 Protein quantification by Bradford Assay

Protein concentration was quantified by Bradford Assay. This colorimetric assay is based on the change in colour of Coomassie G-250 reagent upon binding to protein. In acidic conditions, as in the Bradford Reagent solution, Coomassie G-250 is red. When the Coomassie dye binds to protein in solution, it becomes deprotonated and turns into the blue anionic form with a maximum absorbance (Abs_{max}) at 595 nm. This value can be measured to calculate protein concentration.

Pierce™ Bovine Gamma Globulin Standard (2mg/ml) was used as protein standard for Bradford Assay. The protein standard was diluted 1:10 in sterile water and used as in Table 2.11. The standard concentrations were made in triplicate and the Abs_{595} recorded for each point and averaged. Absorbance was measured using a UV/vis spectrophotometer (Jenway 6315). The average absorbance values were plotted against the protein amount and the points were fitted linearly. The coordinates of the line were used to calculate the concentration of the protein of interest.

BGG (0.2 mg/ml) Volume (µl)	Bio-rad Bradford reagent Volume (µl)	Water Volume (µl)	Protein Amount (µg)
0 µl	200 µl	800 µl	0 µg
2 µl	200 µl	798 µl	0.4 µg
5 µl	200 µl	795 µl	1 µg
10 µl	200 µl	790 µl	2 µg
20 µl	200 µl	780 µl	4 µg
30 µl	200 µl	770 µl	6 µg
50 µl	200 µl	750 µl	10 µg

Table 2.11 - Bovine Gamma Globulin (BGG) standard curve for Bradford protein quantification assay

BGG (2mg/ml, ThermoFisher) was diluted in water to 0.2 mg/ml and used to make the different calibration points of the curve.

2.5.4 Sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE)

2.5.4.1 Preparation of gels

Polyacrylamide gels were casted as per manufacturer instructions (Bio-rad). Glass plates were assembled in a sandwich and fixed on the support. The resolving gel was prepared as described in Table 2.12. APS and TAMED were added last and just before pouring the gel. The resolving gel solution was added in the space between the glass plates using a Pasteur pipette, leaving enough space to add the stacking gel (1-2 cm from the top). A thin layer of pure isopropanol was added on top of the resolving gel to prevent any exposure to air and the gel was left to solidify for 20 minutes. Once the resolving part was solid, isopropanol was removed using filter paper and the stacking gel (Table 2.13) was added to fill the remaining space between the plates. Bio-rad 10 well comb

was inserted into the stacking gel before it solidified to form the wells. The gel was left to solidify for another 20 minutes before being ready to be used.

Components	12% resolving gel solution (10 ml)	15% resolving gel solution (10 ml)
Deionised water	4.9 ml	3.4 ml
30% (w/V) Acrylamide mix	6 ml	7.5 ml
1.5 M Tris-HCl, pH 8.8	3.8 ml	3.8 ml
10% (w/V) Sodium dodecyl sulphate (SDS)	0.15 ml	0.15 ml
10% (w/V) Ammonium persulfate (APS)	0.15 ml	0.15 ml
Tetramethylenediamine (TEMED)	0.006 ml	0.006 ml

Table 2.12 – Composition of 12% and 15% resolving gels

Components	Stacking gel solution (5ml)
Deionised water	2.7 ml
30% (w/v) Acrylamide mix	0.67 ml
1.5 M Tris-HCl, pH 6.8	0.5 ml
10% (w/v) Sodium dodecyl sulphate (SDS)	0.04 ml
10% (w/v) Ammonium persulfate (APS)	0.04 ml
Tetramethylenediamine (TEMED)	0.004 ml

Table 2.13 – Composition of stacking gel

2.5.4.2 Samples preparation

Samples to be run on SDS-PAGE were prepared by adding 4 µl of 5X SDS loading buffer (250 mM Tris-HCl pH 6.8, 10% SDS, 0.5% Bromophenol blue, 50% Glycerol, 500 mM DTT) to 16 µl of sample. The sample was heated at 95°C for 5 minutes.

2.5.4.3 Electrophoresis

Gels were run on Mini-PROTEAN (Bio-Rad) system. The gel was inserted inside the tank together with 1X SDS running buffer (5X SDS running buffer: 125 mM Tris, 1.25 M

Glycine and 0.5% SDS). The comb was removed and wells were washed quickly with buffer. The samples were loaded inside the wells and run in parallel to molecular weight markers, PageRuler Unstained or Prestained Protein Ladder (Thermo Fisher Scientific). The gel was run at 150 V until the samples entered the resolving gel and at 190 V until the dye front reached the bottom of the gel.

2.5.4.4 Staining and destaining of SDS-PAG

At the end of electrophoresis, the gels were carefully removed from the glass plates and stained with Coomassie Brilliant Blue dye (1.25 g Coomassie Brilliant Blue dye, 250 ml methanol, 50 ml acetic acid and 250 ml deionised water) for 40 minutes with gentle shaking. After staining, the gel was rinsed briefly with water and destained with destaining solution (700 ml deionised water, 200 ml methanol and 100 ml of acetic acid) for a few hours or overnight, until the bands were visible and the background was clear. Gels images were taken using BIORAD Gel doc EZ imager. Image lab 4.0.1 software was used to analyse the images.

2.6 Protein-protein interaction assays

2.6.1 MicroScale Thermophoresis (MST)

Microscale Thermophoresis is a biophysical technique used to investigate and quantify molecular interactions in close-to-physiological conditions.

This technique is based on the thermophoresis principle, according to which molecules in solution move directionally under a temperature gradient, away or towards a heated spot.

The thermophoresis effect (or Ludwig-Soret effect) is sensitive to change in molecular weight, charge and solvation shell. The Soret coefficient describes the relationship between the parameters as follows:

$$S_T = \frac{A}{kT}\left(-S_{hyd} + \frac{\beta\sigma^2{}_{eff}}{4\varepsilon\varepsilon_0 T} \times \lambda_{DH}\right)$$

where A is the particle surface area, λ_{DH} is the Debye length (measure of the net electrostatic effect of a charge carrier), $\sigma^2{}_{eff}$ is the effective surface charge density, S_{hyd} is the hydration entropy, ε is the dielectric coefficient, T the temperature and kT is the kinetic energy of the particle (Duhr and Braun, 2006).

Therefore, interactions between molecules can in principle be observed using thermophoresis, as at least one of these parameters may changes upon binding (Wienken *et al.*, 2010).

In order to follow thermophoresis, one of the molecules under investigation needs to be fluorescently labelled. This molecule is usually kept at a constant low concentration, in the nanomolar range, while the binding partner (unlabelled) is titrated in at increasing concentration. Usually 10 to 16 concentrations are necessary to build a saturation binding curve. Concentration range is important and selected based on the expected K_D value. In fact, to obtain a sigmoidal saturation binding curve the concentration should range from 10 fold above to 10 fold below the K_D value.

The Monolith NT.115 instrument from Nanotemper Technology is designed with an infrared (IR) laser source combined with a fluorescence excitation/emission detector using an IR dichroic mirror (Jerabek-Willemsen *et al*., 2011) (Fig. 2.2).
The IR laser induces a small temperature gradient in the mK range into the glass capillaries containing the binding reaction. When the laser is turned on, the applied temperature gradient (ΔT) causes a depletion or enhancement of the molecules from or towards the hot spot, as described by the following equation:

$$c_{hot}/c_{cold} = \exp(S_T \Delta T)$$

Thermophoresis is recorded over a 30 seconds period, until the laser is turned off and the molecules diffuse back into the depleted area (Figure 2.2).

2.6.1.1 ParG protein fluorescence labelling

In order to track the thermophoresis change, one of the proteins under investigation was labelled with Alexa Fuor-647 dye (Thermo Fisher Scientific). In this case, ParG and ParG mutants have been successfully labelled using the following protocol.

150 µl aliquot of the ParG protein was buffer exchanged into 1X PBS buffer using 0.5 ml Zeba Spin desalting column (Thermo Fisher). The column was prepared as per the manufacturer's instructions. First the storage buffer was removed by centrifugation (1500 x g centrifugation, for 1 minute at 4°C) and then equilibrated three times with 300 µl Phosphate-buffered saline (PBS) (Table 2.14). 150 µl of protein was loaded onto the column without touching the resin. After the protein was absorbed by the resin, the column was centrifuged for 2 minutes at 1500 x g, at 4°C. After buffer exchange, 100 mM Sodium bicarbonate $NaHCO_3$ (pH 8.3) was added to the protein sample to increase the pH for labelling. Protein concentration was then checked by Bradford assay (Section 2.5.3). The Bradford standard curve was constructed on the day to improve accuracy.

If, after buffer exchange, the protein concentration was above 20 µM, this was diluted to 20 µM final concentration with labelling buffer (Table 2.14).
The fluorescent dye Alexa Fuor-647(λex = 650 nm; λem = 670 nm) used in this experiment contained NHS-ester chemistry, which reacts efficiently with primary amines of proteins to form highly stable dye-protein conjugates. Although ParG storage buffer does not contain primary amines (e.g. Tris buffer) which would compete with the protein for the labelling reaction, the protein was buffer exchanged because labelling proved to be more efficient in PBS-$NaHCO_3$ buffer.

The dye was dissolved in Dimethyl Sulfoxide (DMSO) to a final concentration of 500 µM. It was then aliquoted in 30 µl aliquots and stored at – 80 °C if not used. On the day, an aliquot was thawed at room temperature and diluted with labelling buffer to a concentration 2-3 fold the concentration of the protein. 150 µl of protein was mixed to 150 µl diluted Alexa-647-fluor dye in a 1:1 (v/v) ratio and incubated for one hour at room temperature in the dark.

2.6.1.2 Removal of unreactive dye

Since the fluorescence dye was added in great excess, the unreactive dye was removed using 2 ml Zeba Spin desalting columns (Thermo Fisher). As per manufacturer's instruction, the column storage buffer was removed by centrifugation at 1000 x g for 2 minutes. The column was then equilibrated three times with 1 ml interaction buffer (Table 2.14). After equilibration, 300 µl protein-dye labelling reaction was applied to the centre of the compact resin bed. To ensure maximal protein recovery a 40µl stacker of interaction buffer was applied onto the resin bed after the sample was fully absorbed. The column was centrifuged at 1000 x g for 2 minutes. Sample was collected and protein concentration was measured by Bradford assay.

2.6.1.3 Preparation of the binding reactions

Labelled ParG protein was diluted using interaction buffer to a concentration of 400 nM. Fluorescence of the protein was checked using the MST machine. Since optimal fluorescence values for the binding experiments are between 200 and 800 units, optimal fluorescence value for ParG alone was around 1500 units. In fact, when the protein is mixed with ParF ligand solutions, its concentration is halved from the stock value as is the fluorescence recorded.

ParF (stored at -80°C) was thawed at room temperature and protein concentration was measured by Bradford Assay, which was usually around 33 µM. 0.05% Tween-20 was added to ParF to improve stability and reduce sticking to the capillaries.
Sixteen serial dilution of ParF were prepared as follows: 16 PCR tube were prepared. 10 µl of interaction buffer were added to 15 tubes (excluding the first). 10 µl of ParF were added to the first and the second tube. The solution in the second tube was mixed and 10 µl were transferred into the third tube. Again, the solution was mixed and 10 µl transferred into the following tube. Serial dilution was repeated to the sixteenth tube, where the final 10 µl were discarded.
10 µl of labelled ParG was then added to each of the tubes. In this way, the interaction reactions contained a constant concentration of ParG (200 nM final concentration) and exponentially decreasing concentrations of ParF (usually 16.5 µM in the first tube). The reactions were loaded onto standard or hydrophilic MST capillaries (Nanotemper Technologies) (depending on the ParG mutant behaviour in solution) simply by dipping the capillary in the tube. The reaction was introduced into the glass capillary by capillary action.

Once loaded, the 16 capillaries were inserted into the Monolith NT.115 MST machine (Nanotemper Technologies). The power of the LED (red for excitation of Alexa Fluor-647) was adjusted according to the protein concentration, usually to 20%, and two measurement were recorded, one with 20% and one with 40% IR laser power. Most interactions show better results when 40% IR laser power was used.
At the end of the measurement, experiments were opened with the MO.Control Software (Nanotemper technologies). The normalised fluorescence values were plotted against the protein concentration to calculate the K_D of the interaction. For comparison between different binding curves, the normalised fluorescence values were converted into percentage of binding by the MO.Control Software. All experiments were repeated in triplicate and curves were fitted using GraphPad Prism 7.04. The binding model used was one-site specific binding equation:

$$Y = \frac{B_{max}X}{K_D + X}$$

where *Y* is the fraction bound, *X* is the concentration of the ligand protein, B_{max} is the maximum binding and K_D is the equilibrium dissociation constant.

MST Buffer	Composition
1 x PBS	137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4 (pH 7.4)
Labelling Buffer	137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4 (pH 7.4), 100 mM $NaHCO_3$ (pH 8.3)
Interaction Buffer	50 mM Tris-HCl, pH 7.5, 150 mM NaCl, 5 mM $MgCl_2$, 0.05% Tween

Table 2.14 - Microscale thermophoresis buffers composition

All buffers were pH adjusted and 0.22 µm filter sterilised.

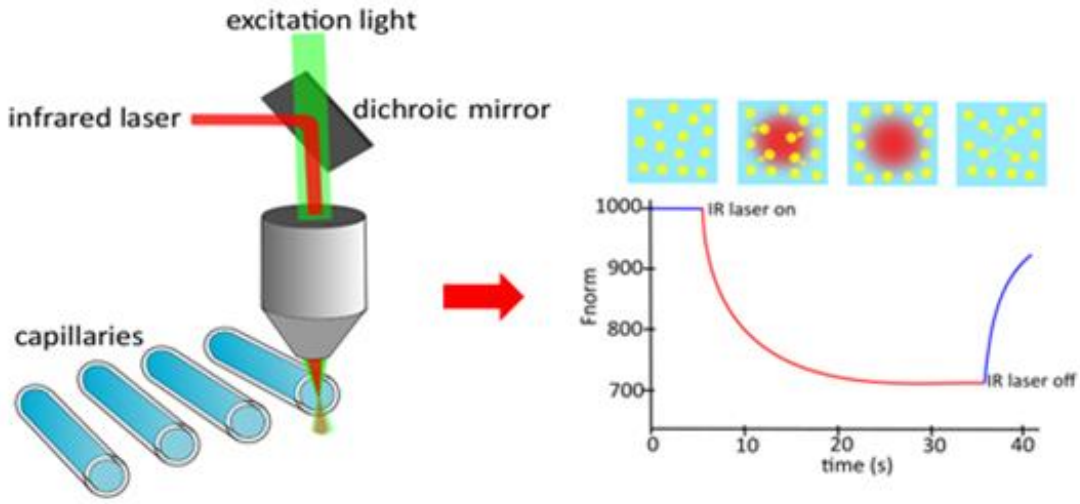

Figure 2.2 - Microscale thermophoresis set up

The binding reactions between labelled ligand and increasing concentrations of analyte are loaded into thin glass capillaries. When the laser is turned on, the molecules move away or towards the heated spot. The change in fluorescence recorded is proportional to the thermophoretic effect and can be measured. When the laser is turned off, the molecules diffuse back into the molecules depleted area. Diagram adapted from Jerabek-Willemsen *et. al*, 2011.

2.6.2 Surface Plasmon Resonance (SPR)

Surface plasmon resonance (SPR) is a biophysical technique used to investigate kinetics of molecular interactions. This label-free technique relies on the immobilisation of one of the binding partners on a thin metal surface, usually gold, present on the top of a half circular prism (Patching, 2014; de Mol and Fischer, 2010; Pattnaik, 2005).

The principle this technique is based on is known as surface plasmon resonance, a physical phenomenon that occurs when plane-polarised light hits a metal surface at a certain angle under conditions of total internal reflection. When the light beam hits the gold surface, the energy from the photons is transferred to the free electrons of the metal surface, which converts into surface plasmons (an electron density wave) (de Mol and Fischer, 2010; Arnoud Marquart, 2016). Resonance is achieved when the energy of the photons is identical to the quantum energy level of the plasmons. At this angle a dip occurs in the intensity of the reflection. The SPR angle depends on the metal surface used, as well as the temperature, the wavelength of the light and the refractive index of the media on both sides of the metal film (de Mol and Fischer, 2010; Arnoud Marquart, 2016). Refractive index is sensitive to the accumulation of mass on the metal immobilisation surface. Therefore binding of molecules can be measured by monitoring the change in the resonance angle (as in the Biacore instruments) (Arnoud Marquart, 2016).

In the following experiments, ParF was immobilised through amine chemistry to a gold carboxymethylated dextran-coated sensor chip (GE healthcare, CM5 sensor chip), while ParG was flowed over the surface at increasing concentration (Figure 2.3). All measurements were recorded using a T200 Biacore SPR instrument (GE Healthcare).

2.6.2.1 ParF immobilisation

SPR requires the immobilisation of one of the binding partners under investigation. ParF substitute, ParF-K160E-R163E, was immobilised to the CM5 sensor chip (GE Healthcare) through amine coupling. The CM5 chip is a glass slide, coated with a thin layer of gold, to which carboxymethylated dextran is covalently attached. The sensor chip is made of four flow cells (Fc) arranged in two pairs Fc-1-2 and Fc-3-4. In the first experiment ParF was immobilised on Fc-2 and Fc-1 was used as reference cell, while in the second experiment Fc-4 was used for ParF immobilisation and Fc-3 was used as reference. Using a reference cell is important to exclude non-specific binding of the analyte to the chip surface. The same procedure, described below, was used for the immobilisation of ParF in both sets of experiments.

After docking the sensor chip in the T200 Biacore system, the machine was primed with water and immobilisation buffer (Table 4). The carboxymethylated dextran surface in the Fc-2 was activated for ParF immobilisation by flowing a 1:1 (v/v) solution of N-Hydroxysuccinimide (NHS, 11.5 mg/ml) : 1-Ethyl-3-(3-dimethylaminopropyl) carbodiimide hydrochloride (EDC, 75 mg/ml) (GE Healthcare) over the chip to convert carboxyl groups to reactive NHS-ester. The NHS-EDC mixture was flowed for 7 minutes at 10 µl/min flow rate. After activation of the carboxyl groups on the dextran, immobilisation buffer was flowed over the cell for 10 minutes.

Since ParF storage buffer contained Tris, whose primary amine could react with the chip surface instead of ParF, the protein was buffer exchanged into HEPES buffer (Immobilisation buffer, Table 2.15) using ZebaSpin desalting columns, as described in section 2.6.1.1. The protein was then diluted to a final concentration of 1 ng/µl in 10 mM sodium acetate, pH 5.0. Dilution in the low pH buffer was done just before ParF injection into the sensor chip. ParF was flowed over the activated Fc-2 for 10 minutes at a 10 µl/min flow rate. The NHS-ester on the dextran reacts spontaneously with the primary amines on the ligand (lysine residues and N-terminus) to form covalent bonds. Loosely associated protein was washed away by flowing immobilisation buffer for 25 minutes at a flow rate of 20 µl/min. Unreactive NHS-ester groups were then deactivated with 1 M ethanolamine hydrochloride-NaOH pH 8.5, flowed for 7 minutes at 10 µl/min flow rate. More buffer was flowed over the chip to remove any remaining chemicals (Figure 2.4).

The same procedure was repeated for the reference cell (Fc-1). In this case, instead of ParF, buffer was flowed over the surface. In this way, both binding and reference cells were treated in the same way.

2.6.2.2 SPR binding experiment

After ParF immobilisation in HEPES buffer, the machine was equilibrated in running buffer. ParG and ParG mutants were buffer exchanged into running buffer (Table 2.15) using ZebaSpin Desalting columns (as per 2.6.1.1). Concentrations were checked by Bradford assay (section 2.5.3) and dilutions of ParG/ParG mutants proteins were prepared. Nine serial dilutions in running buffer from 15 µM to 0.058 µM were made.

In the binding experiments, ParG and ParG mutants were flowed over the ParF-coated chip at increasing concentrations. Buffer alone was injected as negative control, as first and last run for each of the mutants. The concentrations 15 µM and 0.058 µM were repeated twice, to check binding reproducibility.

Each sample was injected for 2 minutes at 30 µl/min flow rate (association phase). Running buffer was then injected for 10 minutes at 30 µl/min flow rate (dissociation phase). At the end of each run, the surface of the flow cell was regenerated with high

salt regeneration buffer (Table 2.15). Regeneration is essential to remove any analyte left on the surface before the next binding reaction. Different buffers were tested and 2M NaCl proved to be the most efficient. Regeneration buffer was injected twice for 1 minute at 30 µl/min flow rate. The surface was then stabilised with 300 µl running buffer (10 minute at 30 µl/min flow rate), before the next sample injection. Table 2.15 summarise the buffers used for the SPR experiments.

The full set of experiment was performed in four days. ParF was immobilised two times on the same sensor chip (Fc-2 and Fc-4). Quality of ParF immobilisations was tested by flowing wild type ParG over the ParF-coated surfaces and obtaining identical binding curves.

Binding curve analysis was done using the Biacore T200 analysis software. The curves were fitted using a 1:1 binding model algorithm. Further analysis of the curves was performed using Prism GraphPad 7.04.

SPR Buffer	Composition
Immobilisation Buffer	10 mM HEPES, pH 7.4, 150 mM NaCl, 0.05% Tween-20
Running Buffer	50 mM Tris-HCl, pH 7.4, 300 mM NaCl, 0.05% Tween-20
Regeneration Buffer	2 M NaCl, 0.05% Tween-20

Table 2.15 - Surface plasmon resonance buffers

All buffers were pH checked and sterilised with 0.22 µm cut-off filters (Millipore).

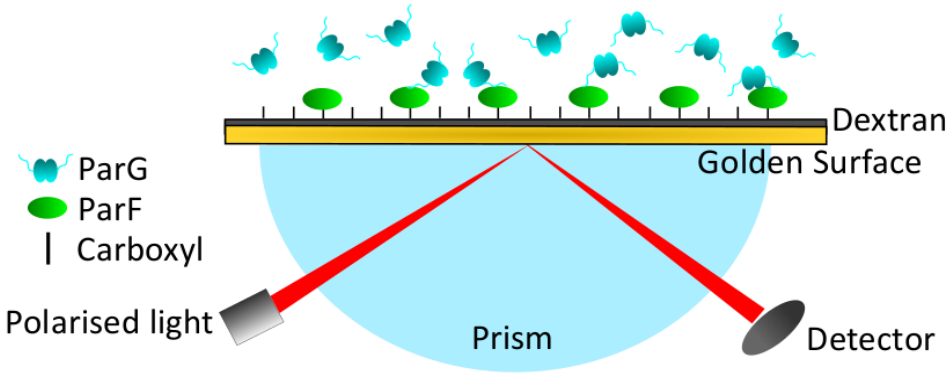

Figure 2.3 - Surface plasmon resonance set up

SPR relies on the immobilisation of the molecule of interest on a metal surface. The gold surface is coated by carboxymethylated dextran, a linear sugar chain made of 1,6-linked glucose, where each glucose unit is modified with a carboxyl group. The carboxyl group in the CM5 chip is used to covalently bind the ligand through amine chemistry. The half circular prism can be found behind the gold surface to achieve condition of total reflection of the polarised light. Binding is detected by monitoring changes in the resonance angle.

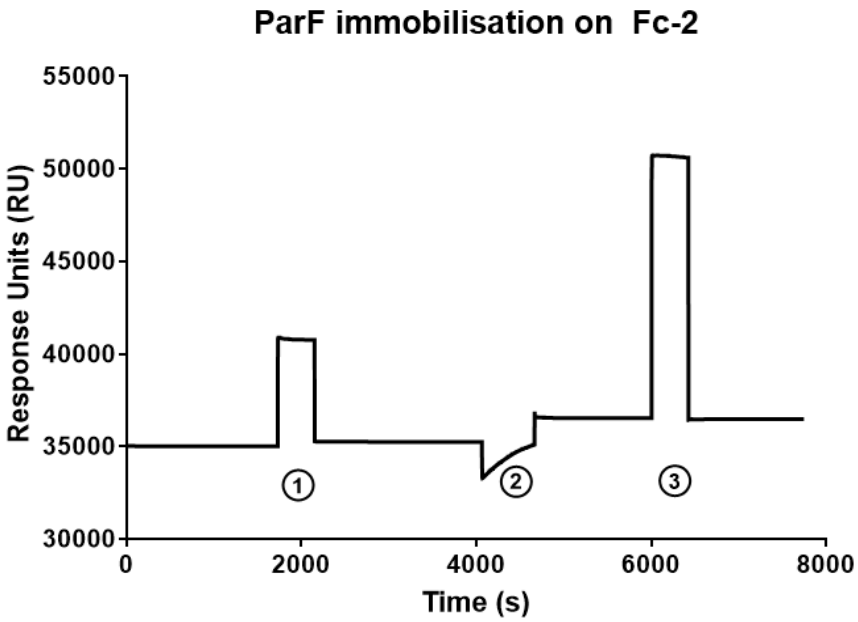

Figure 2.4 - ParF immobilisation on Fc-2 of CM5 sensor chip by amine coupling

1. NHS:EDC mixture is injected on the flow cell for 7 minutes to activate the carboxyl groups on the dextran. EDC cross-links the carboxyl groups to the NHS. 2. ParF in low pH buffer is injected for 10 minutes in the flow cell. The amine groups on the protein react spontaneously with the reactive NHS ester groups on the dextran, forming covalent bonds. 3. Ethanolamine hydrochloride deactivate NHS groups which did not react with ParF. At the end of the process, 902 RU of ParF were present on the chip.

2.6.3 Chemical cross-linking

Chemical cross-linking was performed to analyse the interaction between ParF and ParG *in vitro*, when amino acids where replaced in the ParG protein or when specific and non-specific DNA were added to the mix. Two cross-linker molecules were used in different experiments: dimethyl pimelimidate (DMP) and bis(sulfosuccinimidyl) suberate (BS3). DMP is a imidoester cross-linker (spacer arm length: 9.2 Å), while BS3 is a NHS (N-hydroxysuccinimide) ester cross-linker (spacer arm length: 11.4 Å). Both molecules can cross-link primary amine groups, present on lysine residues or at the N-terminus of a polypeptide chains, and react best at alkaline pH between 8.0 and 9.0.

2.6.3.1 DMP cross-linking

DMP cross-linking was used to investigate the interaction between ParF and ParF-K160E-R163E with ParG, ParG-L3A and ParG-Δ9. ParF and ParG/ParG mutants were buffer exchanged into cross-linking buffer (50 mM HEPES (pH 8.5), 50 mM KCl and 5 mM $MgCl_2$) using ZebaSpin Desalting columns (as in Section 2.6.1.1) and protein concentration was measured by Bradford assay. ParF/ParF-K160E-R163E and ParG/ParG-L3A/ParGΔ9 were added to the reaction to a final concentration of 0.4 mg/ml (8 µg in 20 µl volume reaction). DMP was diluted in water just before being used and added to the reactions in a concentration of 10 mM. Cross-linking was carried out for 1 hour at room temperature. The reaction was then quenched by adding 50 mM Tris-HCl, pH 6.8. 5X SDS-loading buffer was added to the reactions and SDS-PAGE was performed as per 2.5.4.3. Gels were visualised using BIORAD Gel doc EZ imager. Bands were quantified using ImageJ (National Institutes of Health).

DMP cross-linking was also used to analyse the binding between ParF and ParG/ParG mutants when *parH* was present in the reaction. For each of the ParG mutants under investigation, two cross-linking reactions were prepared: one containing ParF and ParG/ParG mutant and one containing ParF, ParG/ParG mutant and *parH*.

The first 52 bp of the centromeric site *parH* (harbouring 7 ParG dimer binding sites) were produced by annealing complementary DNA strands. Primers (Table 2.8) were purchased from Sigma Aldrich and resuspended in annealing buffer (10 mM Tris, pH 8.0, 50 mM NaCl, 1 mM EDTA). Equal concentration and equal volume of the two primers (50 µl of 100 µM forward and reverse primers) were mixed together and incubated for 5 minutes at 95 °C. The reaction was then allowed to cool down at room temperature and concentration of the final product was measured by a Nanodrop.

In 30 µl volume reactions, 10 µg of each protein was added (0.3 mg/ml concentration, 17 µM dimeric ParG and 14.3 µM ParF), together with 2.4 µM *parH*, cross-linking buffer

and 10mM of freshly prepared DMP. The same reactions, without DNA were also prepared for comparison. Cross-linking reactions of ParF-parH, ParG-parH and ParG alone were used as controls. Cross-linking reactions were made in triplicate in order to be loaded on three different 12% SDS-PAGE. Reactions were incubated 1 hour at room temperature and quenched by adding 50 mM TRIS-HCl, pH 6.8. At the end of the cross-linking SDS-PAGE was performed as per Section 2.5.4.3. Of the three gels resolved, one was stained by Coomassie Staining and two were visualised through Western Blot with antibodies specific for ParF and ParG as per Section 2.7. Gels were visualised using BIORAD Gel doc EZ imager. Image lab 4.0.1 software was used to analyse the images and quantify the protein bands.

2.6.3.2 BS3 cross-linking coupled to Liquid chromatography - tandem mass spectrometry (LC-MS/MS) and data analysis

BS3 cross-linking coupled to liquid chromatography– tandem mass spectrometry (LC-MS/MS) was used to identify which amino acids interacted in the ParF-ParG complex. Cross-linking was analysed in different conditions: when ParF and ParG interacted alone, when *parH* was present and when *parH* and chromosomal DNA were present in the reaction. ParF was buffer exchanged into cross-linking buffer (50 mM HEPES (pH 8.5), 50 mM KCl and 5 mM $MgCl_2$) as per Section 2.6.1.1 and centrifuged at 13,000 rpm for 10 minutes to remove aggregates. ParF supernatant was transferred in a fresh tube and protein concentration was quantified by Bradford assay. The cross-linking reactions were prepared in 70 µl volume and 10 µg of each protein were used. 50 µl of ParF at a concentration of 0.2 mg/ml (6.2 µM) (buffer exchanged caused protein precipitation, hence the low concentration) and 8.3 µl of ParG (7.4 µM) at a concentration of 1.2 mg/ml were added to each cross-linking reaction. When *parH* was present, this was added with a final concentration of 1.2 µM (1.7 µl of 50 µM stock). When chromosomal DNA was present, this was gently sonicated to obtain small fragments and 2µl of 100 ng/µl stock were added to the reaction. 2mg of BS3 were dissolved in 70 µl of Mili-Q water to obtain 50 mM concentration. The cross-linker was added just before incubation of the reactions, with a final concentration of 1.4 mM. The reactions were incubated at room temperature for 30 minutes and quenched for 30 min by adding 3.5 µl of 1M ammonium bicarbonate, pH. 8.5 (50 mM final concentration). 30 µl of each sample were run on 12% SDS-PAGE as per Section 2.5.4.3. The gel was stained with fresh Coomassie staining solution and destained for a few hours. Bands identified as the product of ParF-ParG cross-linking were excised with a scalpel and transferred in a Eppendorf tube for LC-MS/MS. Samples digestion was performed as per section 2.9.6.1.

Samples were loaded onto anUltiMate 3000 RSLCnano HPLC system (Thermo) equipped with a PepMap 100 Å C18, 5 µm trap column (300 µm x 5 mm Thermo) and a PepMap, 2 µm, 100 Å, C18 EasyNano nanocapillary column (75 µm x 500 mm, Thermo). The trap wash solvent was aqueous 0.05% (v:v) trifluoroacetic acid and the trapping flow rate was 15 µL/min. The trap was washed for 3 minutes before switching flow to the capillary column. Separation used gradient elution of two solvents: solvent A, aqueous 1% (v:v) formic acid; solvent B, aqueous 80% (v:v) acetonitrile containing 1% (v:v) formic acid. The flow rate for the capillary column was 300 nL/min and the column temperature was 40°C.The linear multi-step gradient profile was: 3-10% B over 7 minutes, 10-35% B over 30 minutes, 35-99% B over 5 mins and then proceeded to wash with 99% solvent B for 4 minutes. The column was returned to initial conditions and re-equilibrated for 15 minutes before subsequent injections.

Mass spectra were acquired in an Orbitrap Fusion Tribrid mass spectrometer supplied by Thermo Fisher scientific. The spray voltage was set to 1.9 kV and the ion transfer tube temperature was held at 275°C. Precursor scans were acquired in profile in the Orbitrap at a resolution of 120K and with a scan range of between 375 to 1500 m/z. Injection times were limited to 100 milliseconds and an AGC (automatic gain control) target of 1e6 was set.

Peptides with a charge state of greater than 3 were selected for fragmentation. Dynamic exclusion settings allowed 2 repeats of the same precursor with a mass tolerance of +/- 10ppm followed by an exclusion duration of 50 seconds. A minimum intensity of 4500 counts was required for selection with priority given to the highest charge state present. To improve sequence coverage for cross-linked peptide candidates, consecutive centroid CID and EThcD1 MS2 scans were acquired in the Orbitrap at a resolution of 30K and a maximum injection time of 54 milliseconds. Isolation was carried out via the quadrupole and an AGC target of 5e4 was set for both fragmentation modes. A NCE (Normalised Collision Energy) of 35% was applied for CID acquisitions.

Thermo .RAW files were converted to .MGF files using RawConverter version 1.1.0.212. Acquired mass spectra were analysed using the Stavrox3 software package version 3.6.6.6. Searches were configured to identify cross-links present at lysine residues and to exclude dead-end identifications. A dead-end link may be defined as where the BS3 reagent has formed a covalent bond with one amino acid residue. Searches allowed for three tryptic missed cleavages using cysteine carboxymethylation as a static modification and the oxidation of methionine as a variable. Precursor ion precision was set to 5 ppm and 10 ppm for fragments. Ion types included in the MS2 searches were b,y, c and z due to the use of Electron Transfer Dissociate (ETD) in the MS2 acquisitions described. A FDR (false discovery rate) was calculated by randomizing the original sequences while

preserving the positions of the arginine and lysine residues. The resulting decoys were then reformed to mixed target decoys where each species contains a decoy peptide combined with one from the original species.

2.7 Western Blot

2.7.1 Samples preparation

Western blot was used to check the integrity of the SF-tagged proteins used in the tandem affinity purification experiments. Samples of 100 µl from the overproduction cultures were collected and cells were harvested by centrifugation. Cell pellets were resuspended in 40 µl Buffer W (Table 2.19) with 10 µl 5X SDS-loading buffer. The samples were heated at 95 °C for 10 minutes and 25 µl were run on 12% SDS-PAGE as per 2.5.4.3. Parallel to the cell samples, 5 µl of a Prestained Page Ruler (Thermo Fisher Scientific) were run to estimate the size of the proteins.

2.7.2 Protein transfer onto Hybond ECL nitrocellulose membrane

The SDS- polyacrylamide gel was transferred onto a Hybond ECL nitrocellulose membrane (GE Healthcare). Before assembling the transfer cassette, 4 pieces of 3MM Whatman paper (cut according to the size of gel glass plate) and two sponges were soaked in 1X transfer buffer (Table 2.16) for 10 minutes. The membrane was also equilibrated shortly in transfer buffer. The transfer cassette was assembled in wet environment, keeping the components in buffer. The gel was positioned on top two pieces of filter paper and a sponge. The membrane was carefully positioned on top of the wet gel without forming bubbles. Two more pieces of filter paper and the second sponge were then placed on top of the membrane. The sandwich was inserted into the transfer cassette and then into the tank. The tank was filled with 1X transfer buffer and an ice block was added to keep the system cold. Transfer was performed at either 30 mA overnight or 60V for 60 minutes. At the end of the transfer, the cassette was opened and the membrane carefully isolated from the rest. Wells were labelled with a pen for easy identification.

2.7.3 Proteins detection

The membrane was rinsed with 1X PBS (Table 2.14), then blocked with 50 ml of blocking buffer (Table 2.16) for 1 hour at room temperature, with gentle shaking. The membrane was then rinsed with 1X PBS-T (Table 2.16) before incubation with the primary antibody. α-ParF and α-ParG antibodies (Barillà and Hayes, 2003) were used at 1:250 and 1:270 dilutions respectively. Commercially available α-Flag antibody (Sigma) was used at 1:3000 dilution. Each antibody was added to 10 ml of blocking buffer and poured onto

the membrane. The membrane was incubated with a primary antibody for 2 hours at room temperature or overnight at 4°C on a shaker. The membrane was then washed twice with 1X PBS-T for 10 minutes with shaking at room temperature. For detection of α-ParF and α-ParG antibody a 1:30,000 diluted α-rabbit antibody (Life technologies) was used, while for the α-Flag a 1:10,000 diluted α-mouse antibody was used. Secondary antibodies were added to 10 ml of blocking buffer and added onto the membrane. The membrane was incubated at room temperature with the secondary antibody for 1 hour, with gentle shaking. The membrane was washed three times in 1X PBS-T for 10 minutes with shaking, at room temperature. The membrane was then rinsed briefly with 1X PBS.

For the detection step, Pierce ECL Western blotting substrate (Thermo Fisher Scientific) was used. To a tray, 2 ml of luminol enhancer solution and 2 ml of peroxide solution were added. The membrane was dried on a filter paper and gently positioned on top of the detection solution (incubated for 1 minute at room temperature). Excess of the solution was removed by tapping the membrane gently on a filter paper. The membrane was then placed in an autoradiography cassette and covered with an acetate sheet. An X-ray film was immediately exposed to the membrane for a discrete amount of time, according to the signal obtained after developing the film.

Buffer	**Composition**
1X Transfer Buffer	48 mM Tris base, 39 mM Glycine, 0.037% SDS, 20% methanol
Blocking Buffer	1X PBS, 0.1% Tween-20 (Sigma), 5% milk (Marvel original dried skimmed milk)
PBS-T	1X PBS, 0.1% Tween-20 (Sigma)

Table 2.16 Composition of Western blot buffers

2.8 Electrophoretic mobility shift assay (EMSA)

2.8.1 Production of biotinylated DNA

EMSA was used to measure the interaction between ParG and ParG mutants with *parH*. The centromeric site *parH* was amplified by PCR as described in section 2.3.2, using the plasmid pFH547 as template and primers which annealed at the beginning and at the end of the *parH* site (more details in Section 1.4.4.1). The forward primer was designed with a biotin label at the 5’ end, while the reverse primer was not modified. The product of PCR amplification was 114 bp long. PCR product was analysed on 1.5% (w/v) agarose gel and extracted following the protocol in 2.3.7. DNA concentration was assessed using

a Nanodrop 1000 (Thermo Fisher). The DNA was stored at 4°C and used within a few days from production.

2.8.2 Samples preparation and electrophoresis

20 µl binding reactions were prepared by adding: 2 µl of *parH* (1 ng/µl), 2 µl of 10X binding buffer (100mM Tris, 500mM KCl, 10mM DTT; pH 7.5), 1 µl of 50% Glycerol, 1 µl of 100 mM $MgCl_2$, 1 µl Poly(deoxyinosinic-deoxycytidylic) acid (Poly (dI-dC)) (1µg/µL) and increasing concentrations of ParG/ParG mutant (50, 75, 100, 200, 30, 400, 500, 750 and 1000 nM). A control sample without protein was also prepared. Reactions were incubated at room temperature for 20 minutes before being loaded on 6% acrylamide gel. The gel was prepared as described in Table 2.17 and pre-run at 100 V for 30 minutes before loading the samples. DNA-loading buffer was added only to the control sample to follow the electrophoresis and to avoid disturbing the protein-DNA complex. Binding reactions were carefully loaded on the acrylamide gel and run at 100 V in 0.5X TBE buffer (Table 2.18) in the cold. The gel was stopped when the dye front reached 3/4 of the gel length.

Components	**Volume (12 ml)**
Milli-Q water	8.4 ml
5x TBE	1.2 ml
30% Acrylamide	2.4 ml
APS 10%	70 µl
TEMED	10 µl

Table 2.17 - Composition of 6% acrylamide native gel

Components	**Quantity**
Tris base	54 g
Boric acid	27.5 g
EDTA	20 ml
Distilled water	To make 1L volume

Table 2.18 - 5X TBE buffer composition

2.8.3 DNA transfer on positively charged membrane

At the end of electrophoresis, the DNA was transferred from the acrylamide gel onto a positively charged nylon membrane (Roche) for detection. Four pieces of Whatman 3MM filter paper (cut to the size of the gel glass plates) and two sponges were soaked in 0.5X TBE buffer (Table 2.17) for 10 minutes. The transfer cassette (Bio-rad) was assembled with the sponge and two pieces of filter papers underneath the gel, carefully removed from the glass plates and placed on top of the Whatman paper. The nylon membrane, soaked briefly in 0.5X TBE buffer, was then positioned on top of the gel, carefully avoiding bubbles formation. Two more pieces of filter papers and the second sponge were used to cover the membrane. The cassette was inserted inside the tank and 0.5X TBE buffer was added, together with an ice block to keep the system cold during the run. The transfer was carried out for 30 minutes at 380 mA. Once the transfer was completed, the cassette was opened and the membrane labelled. The membrane was wrapped in cling film and exposed for 5 minutes to the UV light to cross-link the DNA to the membrane.

2.8.4 DNA detection

The Detection of Biotinylated DNA was carried out using the LightShift™ Chemiluminescent EMSA kit (Thermo Fisher Scientific). Blocking buffer and 4X Wash buffer were warmed at 40 °C to solubilise the components. The membrane was transferred in a clean tray and blocked with 20 ml Blocking buffer for 15 minutes with gentle shaking. The Blocking buffer was discarded and the membrane was incubated with 10 ml Blocking buffer containing 1 µl streptavidin-horseradish peroxidase conjugate (Sigma-Aldrich) for 15 minutes on a shaker. Again, the buffer was discarded and the membrane was washed four times for 5 minutes with 20 ml of 1X Wash buffer (diluted from the 4X Wash buffer) with shaking. The membrane was then transferred into a clean tray and incubated with 30 ml of Equilibration buffer for 5 minutes. For the detection, the membrane was dried on a piece of Whatman filter paper, then carefully positioned face down on the substrate solution containing 1 ml of luminol/enhancer solution and 1 ml stable peroxide solution, mixed together in a clean tray. The membrane was incubated with the substrate solution for 5 minutes with gentle shaking. Excess of solution was removed by tapping the membrane on a piece of Whatman paper. The membrane was then inserted into an autoradiographic cassette with the DNA side up and covered with an acetate sheet. After 5 minutes, the membrane was exposed to an X-ray film in the dark, first for 30 seconds and then for a longer or shorter time according to the intensity of the bands detected after developing the film.

2.9 Tandem Affinity Purification

2.9.1 *parF-SF* and *parG-SF* expression test

Expression tests of *parF*-SF and *parG*-SF were performed to select the concentration of arabinose to use for large scale protein overproduction. BW25113 *E. coli* competent cells were transformed as per section 2.3.10 and grown at 37°C overnight. Colonies from the transformation plates were inoculated into three flasks containing 20 ml selective sterile liquid LB with (100 µg/ml ampicillin). Cells were grown until the OD_{600} reached a value around 0.8. At that point, cells were induced with three different concentrations of arabinose, 0.001%, 0.01% and 0.025%, for 3 hours at 37°C with shaking. Samples were collected from the liquid cultures before induction and every hour from arabinose induction. The samples were centrifuged at 13,000 rpm for 2 minutes and the supernatant was discarded. Pellets were resuspended in 1X SDS loading buffer (Section 2.5.4.2) and analysed on 12% SDS-PAGE (Section 2.5.4.3) to check the expression level.

2.9.2 Production of the StrepII/Flag-tagged proteins

BW25113 *E. coli* competent cells were transformed with pBAD-parF-SF, pBAD-parG-SF and pBAD-aspA-SF plasmids, encoding the SF-tagged proteins to be overproduced. After transformation the cells were spread on selective LB agar (100 µg/ml ampicillin) and grown overnight. A few colonies (8-10 colonies) from the transformation plate were inoculated into 20 ml sterile selective LB broth (100 µg/ml ampicillin) and grown at 37°C with shaking for 2 hours. The liquid culture was transferred to 1 L sterile LB containing 100 µg/ml ampicillin and grown at 37°C on a shaker until an OD_{600} equal to 0.7-0.8 was reached. The cells were then induced with 0.01% arabinose at 37°C for 3 hours with shaking. At the end of the induction process, cells were harvested by centrifugation at 11,000 rpm for 30 minutes at 4°C and then stored at -20°C until utilised.

2.9.3 Crude extract preparation

Pellets were resuspended in 25 ml of cold buffer W (Table 2.19). To the cells, two tablets of protease inhibitor cocktail (cOmplete™, Mini, EDTA-free Protease Inhibitor Cocktail, Roche) and 250 µl of lysozyme (10 mg/ml) were added. The cells were the incubated at 30°C for 15 minutes before further addition of 250 µl of lysozyme and a further 15 minutes incubation at 30°C. After incubation at 30°C, the cell solution was cooled down on ice for ten minutes. Cells were sonicated 15 times for 15 seconds (power 70%) with 1 minute and 15 seconds interval in between sonication steps. Cells were kept on iced water throughout the process and checked at regular interval to prevent heating. Cellular debris was removed by centrifuging the sonication product at 14,000 rpm for 1 hour at 4°C. The

cleared lysate was poured into a fresh falcon tube and the pellet discarded. Since the proteins under investigation can bind to the DNA, the crude extract was incubated with 50 µl deoxyribonuclease I (DNAaseI) (20 U/ml) for 20 minutes at 30 °C. The solution was then filtered using 0.22 µm filters. The crude extract was kept at 4°C throughout.

2.9.4 Affinity purification through StrepII - Strep-Tactin interaction

The column used for the first step of purification was prepared by adding 5 ml of 50% suspension of Strep-Tactin Sepharose (IBA GmbH) into a 10 ml Pierce centrifuge column (Thermo Fisher). The storage buffer was allowed to flow through by gravity and the resin was equilibrated with 2 CVs of buffer W.

Purification was performed at 4°C. The crude extract was loaded onto the column and was let flow through by gravity. It was crucial that protein binding took place on the column, therefore, the crude extract was loaded five subsequent times onto the Strep-Tactin sepharose and it was allowed to flow through without centrifuging to enhance binding. The flow through of the column was collected and used for analysis. The resin was washed 5 times with 1 CV buffer WS (Table 2.19) and its flow through was also collected for analysis. Bradford assay was used to assess whether wash fractions still contained protein. If protein was still present, the resin was washed longer. To elute the protein, 6 X 0.5 CVs 1X Buffer E (containing desthiobiotin, IBA GmbH) were loaded onto the Strep-Tactin resin and allowed to flow through. Elution fractions were collected and protein concentration was measured by Bradford assay (Section 2.5.3). Aliquots containing protein were pooled together and loaded onto the second column.

2.9.5 Affinity purification through Flag – α-Flag interaction

The column for the second step of purification was prepared by adding 1.5 ml of 50% α-Flag M2 affinity gel suspension (Sigma Aldrich) into a 10 ml Pierce centrifuge column (Thermo Fisher). The storage buffer was allowed to flow through by gravity and the gel was washed by loading three sequential column volumes of 0.1 M glycine-HCl (pH 3.5). Each wash was let drain completely before adding the next. The resin was then equilibrated with 5 CVs of TBS buffer (Table 2.19). The elution fractions from the Strep-Tactin were loaded on the α-FLAG affinity gel and incubated for 2 hours at 4°C while gently rotating. When the binding step was completed, the tube was positioned in vertical position and the supernatant was allowed to flow through. The column was washed four times with 2.5 ml TBS buffer (12 CVs). After the wash, Bradford assay was used to quantify the amount of protein in the wash fractions. If protein was still present the column was washed longer. The protein and putative binding partners were then eluted by competition with the FLAG peptide (Sigma Aldrich). Six subsequent 1CV (6X 750 µl)

aliquots of 100 µg/ml Flag peptide in TBS buffer were loaded onto the α-FLAG gel and allowed to flow through. The aliquots were collected separately and quantified by Bradford assay. Aliquots containing protein were pooled together and concentrated forty times using Sartorius Vivaspin centrifugal concentrator (3,000 Da cut off). An aliquot of 100 µg was loaded on 4–20% Mini-PROTEAN® TGX™ Precast Protein Gels (Bio-rad) and run for 5 minutes to enter the gel (Section 2.5.4.2 for sample preparation). The gel was stained with fresh Coomassie staining (2.5.4.4) and destained, before the non-resolved band containing the pulled-down proteins was excised for Mass-Spectrometry (section 2.9.3.1). The same amount was loaded on a separate 4–20% Precast Protein Gel together with samples from the two purification steps for analysis and electrophoresis run as per 2.5.4.3. Figure 2.5 summarises the steps of tandem affinity purification.

Buffer	**Composition**
Buffer W	100 mM Tris-HCl (pH 8.0), 150 mM NaCl
Buffer WS	100 mM Tris-HCl (pH 8.0), 200 mMNaCl
Buffer TBS	50 mM Tris-HCl (pH 7.6), 150 mMNaCl

Table 2.19-Composition of the buffers used for TAP

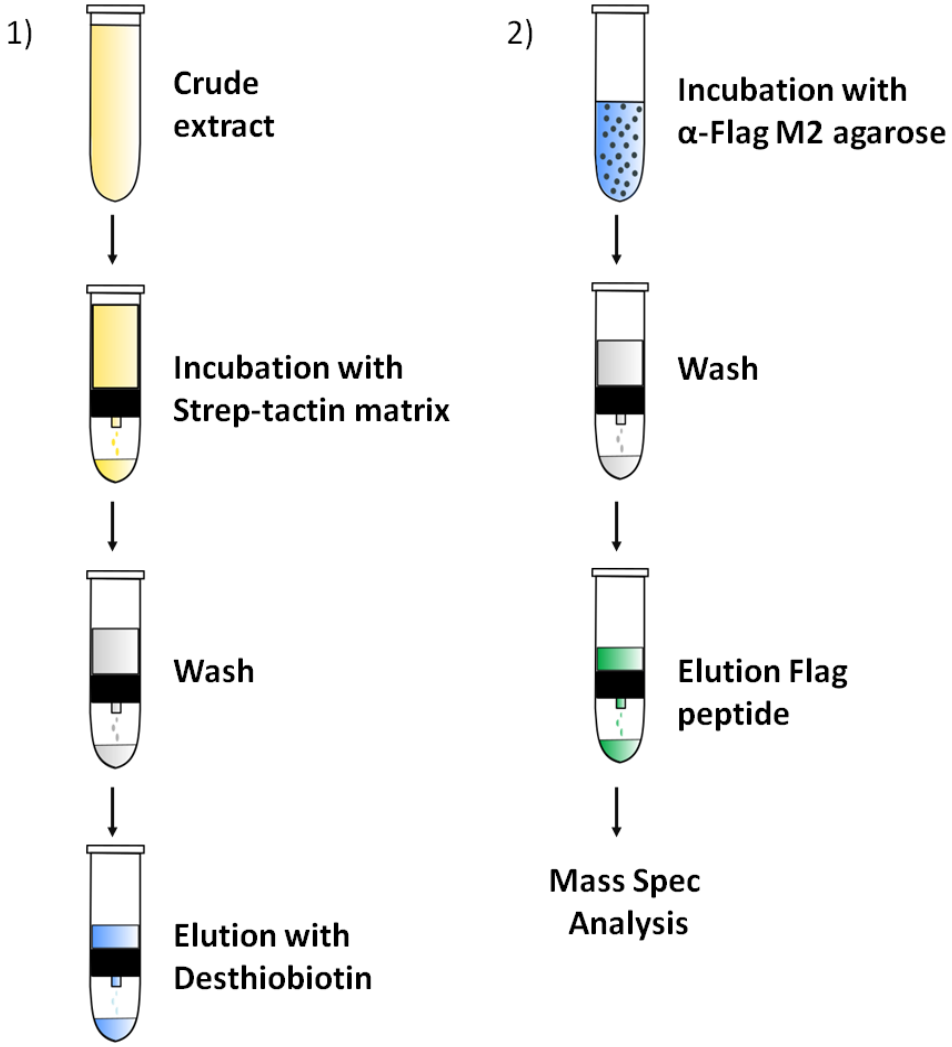

Figure 2.5 – Overview of the TAP protocol

TAP involves two steps of purification of the SF-tagged proteins. 1) The crude extract is incubated with Strep-Tactin matrix that associates with the StrepII peptides. The column is then washed and protein is eluted with Desthiobiotin. 2) The elution from the first step is incubated with α-Flag M2 agarose, that binds to the Flag peptide. The column is washed and the protein is eluted with Flag peptide. Image adapted from Gloeckner *et al*., 2007.

2.9.6 Mass Spectrometry and Protein identification

2.9.6.1 Digestion

In-gel tryptic digestion was performed after reduction with dithioerythritol and *S*-carbamidomethylation with iodoacetamide. Gel pieces were washed two times with 50% (v:v) aqueous acetonitrile containing 25 mM ammonium bicarbonate, then once with acetonitrile and dried in a vacuum concentrator for 20 min. Sequencing-grade, modified porcine trypsin (Promega) was dissolved in 50 mM acetic acid, then diluted 5-fold with 25 mM ammonium bicarbonate to give a final trypsin concentration of 0.02 µg/µl. Gel pieces were rehydrated by adding 25 µl of trypsin solution, and after 10 minutes enough 25 mM ammonium bicarbonate solution was added to cover the gel pieces. Digests were incubated overnight at 37 ^{o}C. Peptides were extracted by washing three times with aqueous 50% (v:v) acetonitrile containing 0.1% (v:v) trifluoroacetic acid, before drying in a vacuum concentrator and reconstituting in aqueous 0.1% (v:v) trifluoroacetic acid.

2.9.3.2 LC-MS/MS

Samples were loaded onto anUltiMate 3000 RSLCnanoHPLC system (Thermo) equipped with a PepMap 100 Å C_{18}, 5 µm trap column (300 µm x 5 mm Thermo) and aPepMap, 2 µm, 100 Å, C_{18}EasyNanonanocapillary column (75µm x 500 mm Thermo). The trap wash solvent was aqueous 0.05% (v:v) trifluoroacetic acid and the trapping flow rate was 15 µl/min. The trap was washed for 3 minutes before switching flow to the capillary column. Separation used gradient elution of two solvents: solvent A, aqueous 1% (v:v) formic acid; solvent B, aqueous 80% (v:v) acetonitrile containing 1% (v:v) formic acid. The flow rate for the capillary column was 300 nl/min and the column temperature was 40°C.The linear multi-step gradient profile was: 3-10% B over 7 mins, 10-35% B over 30 mins, 35-99% B over 5 minutes and then proceeded to wash with 99% solvent B for 4 minutes. The column was returned to initial conditions and re-equilibrated for 15 min before subsequent injections.

The nanoLC system was interfaced with an Orbitrap Fusion hybrid mass spectrometer (Thermo) with an Easy Nano ionisation source (Thermo). Positive ESI-MS andMS^2 spectra were acquired using Xcalibur software (version 4.0, Thermo). Instrument source settings were: ion spray voltage,1,900 V; sweep gas, 0 Arb; ion transfer tube temperature; 275°C. MS^1spectra were acquired in the Orbitrap with: 120,000 resolution, scan range: *m/z* 375-1,500; AGC target, 4e5; max fill time, 100 ms. Data dependant acquisition was performed in top speed mode using a 1 s cycle, selecting the most intense precursors with charge states >1. Easy-IC was used for internal calibration. Dynamic exclusion was performed for 50 s post precursor selection and a minimum threshold for fragmentation was set at 5e3.MS^2spectra were acquired in the linear ion

trap with: scan rate, turbo; quadrupole isolation, 1.6*m/z*; activation type, HCD; activation energy: 32%; AGC target, 5e3; first mass, 110 *m/z*; max fill time, 100 ms. Acquisitions were arranged by Xcalibur to inject ions for all available parallelizable time.

2.9.3.3 Data Analysis

Peak lists were converted from centroided .raw to .mgf format using MSConvert (ProteoWizard 3.0.9967). Mascot Daemon (version 2.6.1, Matrix Science) was used to search against the *E. coli* subset of the UniProt database (23,036 sequences; 7,069,511 residues)using a locally-running copy of the Mascot program (Matrix Science Ltd., version 2.6.1). Search criteria specified: Enzyme, trypsin; Fixed modifications, Carbamidomethyl (C); Variable modifications, Oxidation (M), Deamidation (N,Q), Acetylation (N-term); Peptide tolerance, 3 ppm; MS/MS tolerance, 0.5 Da; Instrument, ESI-TRAP. The Mascot .dat result file was imported into Scaffold (version 4.8.4, Proteome Software) and a second search run against the same databases using X!Tandem, version CYCLONE (2010.12.01.1). Protein identifications were filtered to require a maximum protein and peptide false discovery rate of 1% with a minimum of two unique peptide identifications per protein. Protein probabilities were assigned by the Protein Prophet algorithm (Nesvizhskii, 2003). Proteins that contained similar peptides and could not be differentiated based on MS/MS analysis alone were grouped to satisfy the principles of parsimony. Proteins sharing significant peptide evidence were grouped into clusters. Relative protein quantification between sample groups was calculated in Scaffold using normalised spectral counts. Statistical differences were assessed using Student's t-test with Benjamini-Hochberg multiple test false discovery rate correction applied.

2.10 Circular Dichroism

Circular dichroism (CD) was used to monitor the change in ParG and ParG mutant proteins secondary structure spectra upon addition of specific DNA. ParG and ParG mutant proteins were buffer exchanged into CD buffer (10 mM KPO_4, 10 mM KCl, 20 mM $MgCl_2$) using ZebaSpin desalting columns (Section 2.6.1.1) and protein concentration was measured by Bradford assay. Proteins were carefully diluted to 0.2 mg/ml (10.3 µM) in CD buffer and left on ice. A region of the centromeric site parH, 52 bp sequence harbouring 7 ParG dimers binding sites, was produced by annealing complementary strands. Primers were purchased from Sigma Aldrich and resuspended in annealing buffer (10 mM Tris-HCl, pH 8.0, 50 mM NaCl, 1 mM EDTA). Equal concentration and equal volume of the two primers (50 µl of 100 µM forward and reverse primers) were mixed together and incubated for 5 minutes at 95 °C. The reaction was then allowed to

slowly cool down to room temperature. Final concentration of the primers was measured using the Nanodrop. A volume of 250 µl of ParG/ParG mutant protein was loaded into a quartz cuvette with a path length of 1 mm. CD spectra were collected from 190 nm to 350 nm. Increasing concentration of DNA were then added to the protein inside the cuvette. Very small volumes of DNA were added, therefore, spectra were adjusted during the analysis to take into account the slight change in protein concentration. DNA was titrated into the protein at the following concentrations: 0.05, 0.1, 0.2, 0.4, 0.6, 0.8, 1, 1.2, 1.4 µM. Since the *parH* site selected contained 7 ParG binding sites, 1.4 µM was the concentration at which each site would have been occupied by a ParG dimer, as ParG concentration inside the cuvette was 10.3 µM. After each addition, protein and DNA were incubated together at room temperature for 5 minutes before recording the CD spectrum. Measurements were collected using a Jasco J810 circular dichromator. The values were recorded in m° (millidegrees) and converted into residual molar ellipticity ([θ]) using the following formula:

$$[\theta] = \frac{\mathrm{m}^{\circ}\mathrm{MW}_{\mathrm{R}}}{10\ cl}$$

where *m*° is the value recorded by the Jasco circular dichromator, MW_R is the residual molecular weight of the protein (MW divided by the number of amino acids), *c* is the concentration of the protein (in mg/ml) and *l* is the cuvette path length (0.1 cm). Analysis of the CD spectra was performed using GraphPad Prism 7.04.

2.11 Dynamic Light Scattering (DLS)

Dynamic light scattering (DLS) was used to measure ParG mutants ability to stimulate ParF polymerisation. A Malvern Zetasizer Nano system was used for this experiment. Proteins were centrifuged for 30 minutes at 13,000 rpm at 4°C, to exclude protein aggregates. The supernatant was transferred into a fresh tube and quantified by Bradford assay (section 2.5.3). ParF was diluted to a concentration of 50 ng/µl (2.16 µM) using the protein storage buffer. The buffers used for DLS were filtered with 0.22 µm Millex-GP syringe filters. A volume of 45.15 µl of diluted ParF was loaded onto a 50 µl quartz cuvette and placed in the Zetasizer chamber at 30 °C. Initially 30 readings were taken to obtain a baseline. Then, 2.5 µl of 100 mM $MgCl_2$ (5 mM final concentration) and 1.25 µl of 20 mM ATP (0.5 mM final concentration) were added. $MgCl_2$ was dissolved in water, while ATP was dissolved in 20 mM Tris-HCl, pH 7.5. The cuvette was placed back into the machine and 30 more readings were recorded. A volume of 1.1 µl of 100 µM ParG/ParG mutant stock solution was added to the ParF-$MgCl_2$-ATP reactioninside the cuvette to a final concentration of 2.16 µM and mixed. The cuvette was inserted into the

chamber and 30 more readings were taken. The values recorded were transferred to GraphPad 7.04 and graphs were generated.

2.12 Partition Assay

Partition assay was used to assess plasmid stability when the *parG* gene was mutated. BR825 *E. coli* cells were transformed with the pFH547 plasmid derivates which harboured the *parG* mutation. As positive and negative controls of the experiment, BR825 cells were also transformed with the wild type pFH547 (positive control) and the empty vector pFH450 (negative control). Cells were transformed as per Section 2.3.10 and grown overnight at 37 °C. Ten colonies were streaked on LB containing 10 µg/ml chloramphenicol (Cm). To perform this step, two LB-Cm plates were divided into 5 sections and ten colonies were streaked on the ten sections individually. The cells were then grown overnight at 37 °C. One colony from each of the sections was then streaked on LB agar with no selection and incubated at 37 °C overnight. The next day this step was repeated so that cells were allowed to grow for 25 generation without selective pressure. The next day, ten colonies from each section were stabbed on two different plates: one containing chloramphenicol and one without antibiotics. A total of 100 colonies was analysed, allowing to determine the percentage of cells that retained the plasmid.

2.12.1 Partition Assay to screen Keio collection deletion mutants

Partition assay was also used to determine the involvement of the proteins identified through TAP in plasmid segregation. Plasmid stability was measured in cells where the gene of the proteins identified through TAP was replaced by the kanamycin resistance gene (Baba *et al.*, 2006). Deletion mutants were obtained from Prof. Peter McGlynn's Keio collection and made chemically competent as described in 2.3.9. Deletion mutant cells were then transformed with the low copy number plasmids pFH554, harbouring the entire partition cassette, and pFH450-LC with no partition gene (negative control). As a control, plasmid stability in wild type BW25113 cells was also analysed. Partition assay was performed as described in 2.12 and percentage of plasmid retention measured.

2.13 Bacterial two-hybrid assay

The bacterial two-hybrid protocol was adapter from Barillà and Hayes, 2003 (Barillà and Hayes, 2003b). *E. coli* SP850 cells were co-transformed with the pT25-parF plasmid and the pT18 plasmid, where the *T18* moiety was fused to one of the following genes: *parG*, *ftsZ*, *mreB*, *gyrB* and *mukB*. The aim of the bacterial two-hybrid assay was to measure

the interaction between ParF and the proteins identified through the TAP experiment (FtsZ, MreB, GyrB, MukB) *in vivo.* ParF-ParG binding was used as a positive control of the interaction. Transformations were grown on LB agar containing both ampicillin (100 µg/ml) and chloramphenicol (30µg/ml) to select for cell which contained both the plasmids. Plates were incubated at 30 °C for 36 hours. McConkey plates were prepared fresh just before their utilisation as described in section 2.2.1.2. To 200 ml of liquid McConkey agar, 200 µl of ampicillin (100 mg/ml) and 200 µl of chloramphenicol (30 mg/ml) were added, together with 10 ml of 20% (w/v) maltose (Sigma). The plates were then poured and allowed to dry completely. Colonies from the transformation plates were streaked on the MacConkey-maltose plates using a sterile loop. Plates were incubated at 30 °C overnight or longer, until red colonies appeared.

2.13.1 β-galactosidase assay

Co-transformation of SP850 cells was carried out as in 2.9.1. Cells were grown on LB agar containing ampicillin (100 µg/ml) and chloramphenicol (30 µg/ml) and grown for 36 hours at 30 °C. Single colonies from the transformation plates were inoculated into 5 ml broth containing both antibiotics and 0.5 mM IPTG and the cultures were incubated at 30 °C overnight, on a shaker. Liquid cultures were then incubated on ice for 20 minutes. To 1L M63 medium 1ml of 1M $MgSO_4$and 10 ml of 20% maltose were added. An aliquot of 1 ml was taken from the liquid cultures and diluted with 4 ml of supplemented M63 medium (Table 2.20). OD_{600} of the diluted liquid cultures was recorded before 3 x 500 µl aliquots were take and centrifuged at 11,000 rpm for 10 minutes. The supernatant was removed and cells were resuspended in 700 µl of buffer Z containing 50 mM β-mercaptoethanol (Table 2.21). To the same tube, 20 µl of trichloromethane ($CHCl_3$) and 20 µl of 0.1% SDS were added and the tubes were vortexed to mix. The samples were incubated at 28 °C for 5 minutes. The reaction was then started after 200 µl of O-nitrophenyl-β-D-galactopyranoside (ONPG) (4 mg/ml) (Sigma) were added. After a few minutes a yellow colour developed and the reaction was stopped by adding 500 µl of 1M sodium carbonate (Na_2CO_3). The time from the start of the reaction to when the reaction was stopped was recorded. The samples were then centrifuged for 1 minute at 11,000 x g and the supernatant was transferred into a UV-Vis cuvette. A_{420} and A_{550} were measured using a UV/is spectrophotometer. Miller units (MU) were calculated by usingthe following equation:

$$MU = 1000 \times \frac{A_{420} - (1.75 \times A_{550})}{T \times V \times OD_{600}}$$

where *T* is the time needed for the reaction, *V* is the volume of the sample, *A* is the absorbance at the different wavelengths and OD_{600} is the cell density.

Component	Amount (g) in 1L
$(NH_4)2SO_4$	10
KH_2PO_4	68
$FeSO_4.7H_2O$	0.0025
Dissolved in 1L distilled water and pH adjusted to 7.0	

Table 2.20– Composition of 5X M63 buffer

Component	Concentration	Amount (g) in 1L
$Na_2HPO_4.7H_2O$	0.06M	16.1
$NaH_2PO_4.H_2O$	0.04 M	5.5
KCl	0.01 M	0.75
$MgSO_4$	0.001 M	0.246
Dissolved in 1L and pH adjusted to 7.0		

Table 2.21– Composition of Z-buffer

Chapter 3
Interaction of ParG N-terminal mutant proteins with the partner ParF

3.1 Introduction

The segregation system of the plasmid TP228 is composed of three elements: ParF, a Walker-type ATPase, ParG, a DNA binding protein, and *parH,* the centromeric site on the plasmid (Wu *et al.*, 2011). The interaction among these three elements, that together form the segrosome, is essential for plasmid stability and the focus of this work.

ParG is a homodimeric protein that binds specifically to its cognate site *parH* and mediates the interaction between the plasmid and the motor protein ParF (Barillà and Hayes, 2003). The protein has a folded C-terminal domain and an unstructured N-terminal tail (Golovanov *et al.*, 2003). While the C-terminal region is essential for dimerisation and DNA recognition, the N-terminus is involved in other functions, such as binding to ParF and enhancing its weak ATPase activity through an arginine finger motif (Barillà and Hayes, 2003; Barillà *et al.,* 2007). ParG can also promote ParF polymerisation (Barillà and Hayes, 2003). It was suggested that the N-termini of ParG dimers could work as sticky tentacles to hold ParF units together (Barge, 2015). ParG also works as a repressor of the *parFG* operon (Wu *et al.*, 2011) and deletion of its N-terminus has a detrimental effect on ParG transcriptional regulatory activity (Carmelo *et al.*, 2005). In fact, the N-terminal domain improves the specificity of ParG interaction with *parH* and the O_F operator site (Wu *et al.*, 2011: Carmelo *et al.*, 2005).

Alanine scanning mutagenesis of the thirty amino acids that compose the ParG N-terminal tail allowed the identification of a number of residues whose change into alanine either impairs or abolishes plasmid segregation (Barge, 2015). The following residues changes in the ParG N-terminus have been characterised as detrimental for plasmid partitioning *in vivo*: L3A, K5A, K11A, K12A, M13A, N18A, R19A, L21A (Figures 3.1 and 3.2). The ParG proteins harbouring these changes will be referred as ParG mutants in this work. They have been extensively characterised in their ability to dimerise, to bind to the *parFGH* cassette operator site, to bind to ParF *in vivo* and to stimulate ParF ATPase activity (Barge, 2015). From the data acquired by Barge it was possible to infer that many ParG characteristics do not depend on the activity of single amino acids in the N-terminal tail. ParG's ability to stimulate ParF ATPase activity is greatly affected by all the changes in the N-terminal domain mentioned above. ParG binding to ParF is also mildly affected by these single amino acid changes, while dimerisation, transcriptional repressor activity and the ability to bundle ParF filaments are almost unaffected, apart from a few exceptions (Table 3.1) (Barge, 2015).

The NMR study on ParG showed that the N-terminal region is very flexible and unstructured, with different levels of flexibility along the tail. The first six to ten amino

acids are the most flexible, while residues 17 to 23 are the least flexible, possibly due to either transient secondary structures or contact with the folded C-terminal domain (Golovanov *et al.*, 2003).
Importantly, Arg-19 in ParG was identified as the prime residue of an arginine finger motif, responsible for stabilising the transition state during ATP hydrolysis by ParF, thus enhancing ParF weak ATPase activity (Barillà *et al.*, 2007). Therefore, Asn-18 and Leu-21 could be part of a scaffold for Arg-19 positioning.
The middle cluster of amino acids, represented by Lys-11, Lys-12 and Met-13, was seen to make transient interactions with the cognate DNA sequence recognised by ParG (Carmelo *et al*., 2005). In fact, heteronuclear single quantum coherence (HSQC) spectroscopy showed a large change in chemical shift in this region when a 48bp operator site (O_F) was added to ParG. Interestingly no change in chemical shift was observed when non-specific DNA was used. In addition, residues His-7, Asn-18, Arg-19, Arg-23, Thr-26, Ala-27, and Val-29 were also affected by the addition of the cognate site, suggesting that they could also be involved in recognition and interaction with DNA.

In 2017, fragments of ParG were co-crystallised with AMPPCP-bound ParF (Zhang and Schumacher, 2017). Although only a small part of the ParG N-terminus was detectable, because of protein degradation caused by the long crystallization time, amino acids 13 to 22 (comprising the arginine finger motif) were reported to form a helical structure that inserts at the interface between ParF monomers. According to Schumacher, insertion of the α-helix in that position stabilises the ATP-bound ParF dimer (Figure 3.3).

From these results it is possible to speculate that three different clusters of amino acids are present in the N-terminal domain of ParG. These are likely to perform different activities based on their location in the N-terminus, as well as the nature of the specific side chain. The aim of this chapter is to investigate the interaction between ParF and ParG N-terminal mutant proteins *in vitro*. ParF - ParG interaction was previously characterised *in vivo* by Barge using a bacterial two-hybrid system (Barge, 2015). In this assay, all the mutants showed reduced interaction with ParF. ParG-K11A and ParG-K12A were identified as the mutants with the lowest affinity for ParF. However, this method is semi-quantitative and the interaction required further *in vitro* characterisation. Surface Plasmon Resonance (SPR) and MicroScale Thermophoresis (MST) were used to quantify the K_D of the interaction between ParF, wild type ParG and ParG mutants in order to identify which amino acids are essential for the protein-protein interaction.

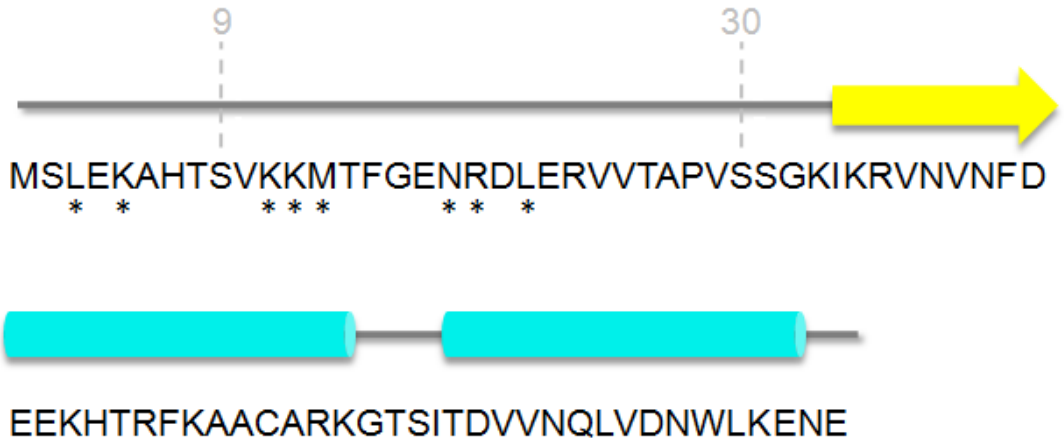

Figure 3.1 – ParG sequence and secondary structure

ParG secondary structure, showing the corresponding amino acid underneath. The asterisk highlights the amino acids which have been changed into alanine and identified as essential for plasmid segregation. The amino acids 9 and 30, end of the N-terminal tail, are shown.

.

Figure 3.2 - NMR structure of ParG dimer

The ParG monomer has a ribbon-helix-helix folded C-terminal domain and an unstructured N-terminal domain (PDB 1P94). The amino acids shown as sticks represent the residues that when changed into alanine impair plasmid segregation. The same colour for each amino acid position is used throughout this chapter. The image was generated using PyMOL v1.8.6.0 (The PyMOL Molecular Graphics System, Version 1.8.6.0, Schrödinger, LLC).

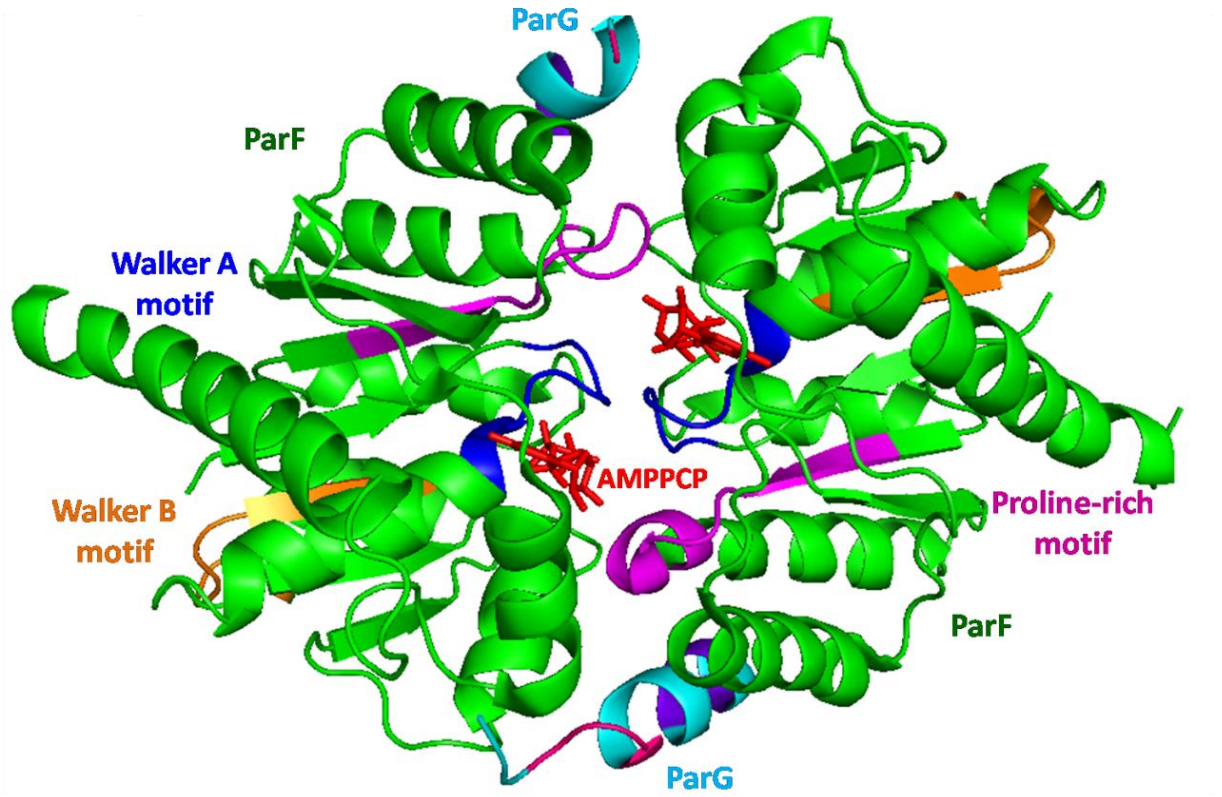

Figure 3.3- Crystal structure of the ParF - ParG (residues 8-22) complex

ParF dimer is shown in green, with residues in blue representing the Walker A motif (9-16), in orange the Walker B motif (73-83) and in purple, the proline-rich motif (102-112). AMPPCP (red) is sandwiched between the ParF monomers. A small region of ParG is visible (cyan), with the residues Lys-11, Lys-12 and Met-13 in pink and the arginine finger motif in violet (PDB 5U1G). The image was generated using PyMOL v1.8.6.0 (The PyMOL Molecular Graphics System, Version 1.8.6.0, Schrödinger, LLC).

N-terminal mutant	Plasmid stability	Dimerization	Site-specific DNA binding	Transcriptional repression	ParF bundling	Interaction with ParF *in vivo*	ParF ATPase activity stimulation	ParF oscillation
ParG WT	+	+	+	+	+	+	+	+
ParG L3A	--	+	+	+	+	-	--	--
ParG K5A	--	+	+	+	+	-	--	--
ParG K11A	--	+	-	+	+	-	--	--
ParG K12A	--	+	+	+	+	--	--	--
ParG M13A	--	+	+	+	+	-	--	--
ParG N18A	--	+	+	+	+	-	--	--
ParG R19A	--	-	+	+	+	-	--	--
ParG L21A	--	+	--	+	+	-	--	--

Table 3.1 - Summary of the characteristics identified for the ParG N-terminal mutant proteins as compared to WT ParG

The green colour (+) was used to represent similar behaviour to the wild type protein. The yellow colour (-) represents slightly affected function. The red colour (--) represents severely affected function. Adapted from Barge, 2015.

3.2 Results

3.2.1 ParF and ParF-K160E-R163E overproduction and purification

In vitro investigation of the interaction among ParF, ParG and ParG mutants, using SPR and MST required the purification of the proteins of interest. Besides wild type ParF, ParF-K160E-R163E (Allison-Gamble, 2016) was used as ParF substitute in the SPR experiments. Wild type ParF easily self-associates and precipitates at high concentrations. Insoluble protein was usually removed by centrifugation, however, polydispersity of the sample was often a problem in quantification of the interaction with ParG by biophysical techniques. For this reason ParF-K160E-R163E was used instead. ParF-K160E-R163E was constructed in an attempt to disrupt ParF binding to the nucleoid. It was shown that the mutant protein did not interact with non-specific DNA *in vitro*, however the amino acid changes did not affect the protein's ability to bind and to hydrolyse ATP and to dimerise. Two-hybrid experiments also confirmed that the mutant protein could interact with ParG as well as wild type ParF (Allison-Gamble, 2016). However, the mutant ParF was unable to self-associate into higher-order structures in the absence of nucleotide (Allison-Gamble, 2016), making it a perfect candidate for the SPR binding experiments.

The pET expression vectors harbouring wild type and mutant *parF* in frame with the codons encoding for a C-terminal hexahistidine tag, were available in the laboratory plasmids collection. The construct was transformed into BL21(DE3) *E. coli* cells, which contain a copy of the T7 RNA polymerase gene in the chromosome, under the control of the *lac* promoter. Therefore, when cells are treated with IPTG, T7 RNA polymerase is produced, triggering the transcription of the gene cloned in the pET vector downstream of the T7 promoter.

Protein overproduction was carried out as described in Section 2.5.1. Induction for three hours at 30 °C, was sufficient to obtain high levels of protein production for both ParF proteins (Figure 3.4 A). Overexpression of ParF at 30 °C decreases the amount of protein that goes in the inclusion bodies, leading to higher amount of soluble protein. Nevertheless around 40% of the protein produced is insoluble and is lost in the pellet during extraction of the cleared lysate.

The proteins were purified by Ni^{2+} affinity chromatography, as described in the Materials and Methods (Section 2.5.1). Both proteins were purified with concentrations ranging from 0.1 to 1.3 mg/ml for wild type ParF (Figure 3.4 B) and from 0.5 to 2 mg/ml for ParF-K160E-R163E (Figure 3.5). The proteins have a molecular weight of 23.1 kDa, including the His-tag.

3.2.2 ParG and ParG N-terminal variants overproduction and purification

The pET vectors encoding *parG* and *parG* mutant alleles under the control of the T7 promoter were available in the laboratory plasmid collection. The proteins were overproduced as described in Section 2.5.2. Three hours induction with IPTG at 37°C generated a high level of protein overproduction (Figure 3.6). Identical conditions were used for wild type ParG and ParG mutants-producing cells. Previous experiments had shown that the proteins are soluble (Barge, 2015). Examples of ParG mutant protein overproduction are represented by ParG-R19A and ParG-L21A (Figure 3.6 B).

All the proteins were purified by one step Ni^{2+} affinity chromatography as described in Section 2.5.2. Purification of wild type ParG, ParG-L3A and ParG-M13A (Figures 3.7 and 3.8) exemplify the results obtained for the rest of ParG mutant proteins. All the proteins were pure after chromatography, with concentrations ranging from 0.5 to 2 mg/ml.

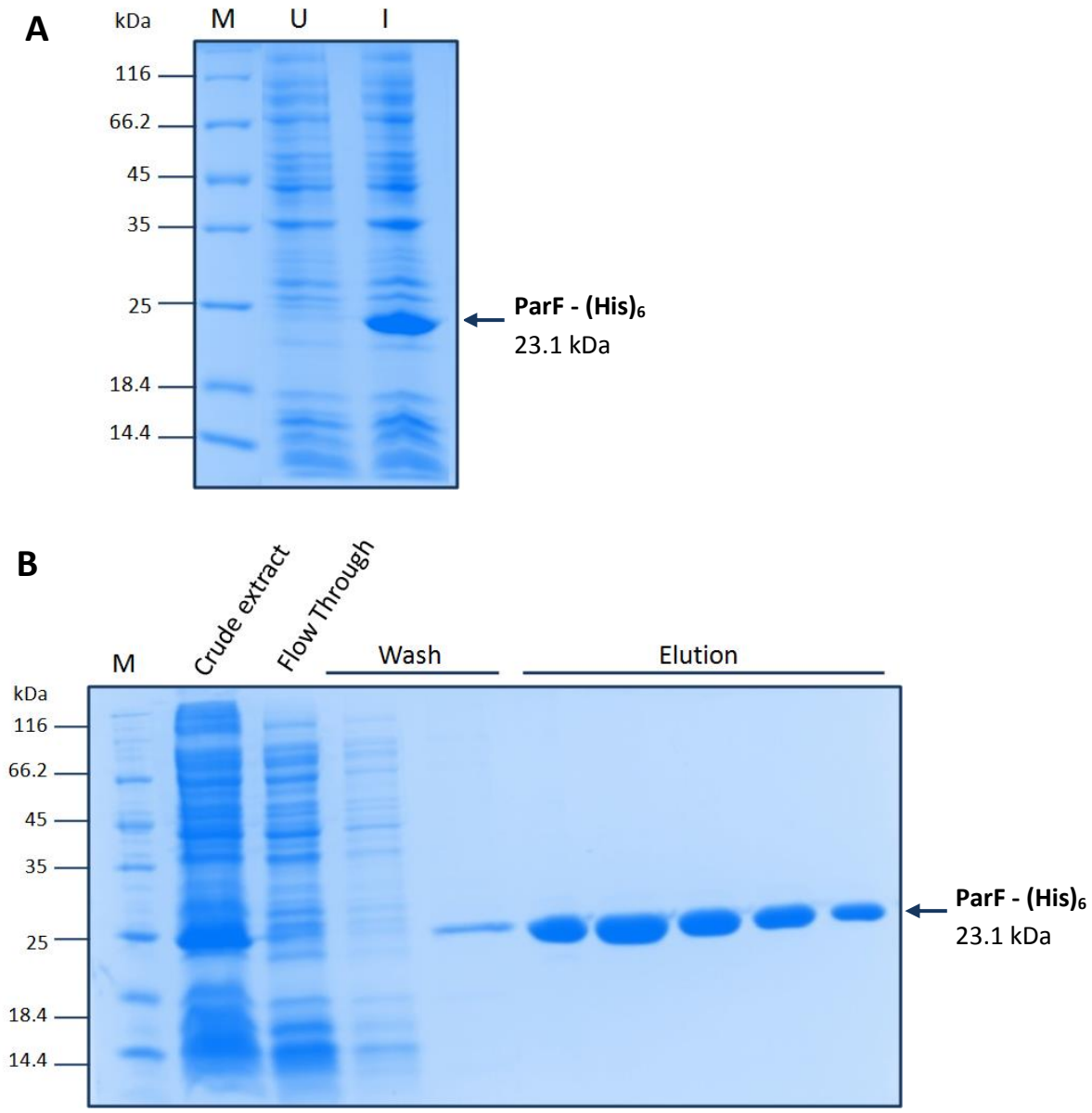

Figure 3.4 - SDS polyacrylamide gel of (A) ParF overproduction, (B) ParF Ni^{2+} affinity chromatography

(A) ParF overproduction test comparing uninduced cell (U) and induced cell (I), using 1 mM IPTG. M represents the MW marker (Pierce™ Unstained Protein MW marker).

(B) Fractions collected during ParF purification. The crude extract represents the cleared lysate of ParF-producing cells before nickel affinity chromatography. Non-binding proteins are shown in the "flow through" fraction. The column was washed with 15 mM (lane 4) and 85 mM (lane 5) imidazole to remove proteins binding non-specifically to the column (wash fractions). Lanes 5 to 10 represent ParF elution fractions after buffer exchange into storage buffer. M represents the MW marker (Pierce™ Unstained Protein MW marker).

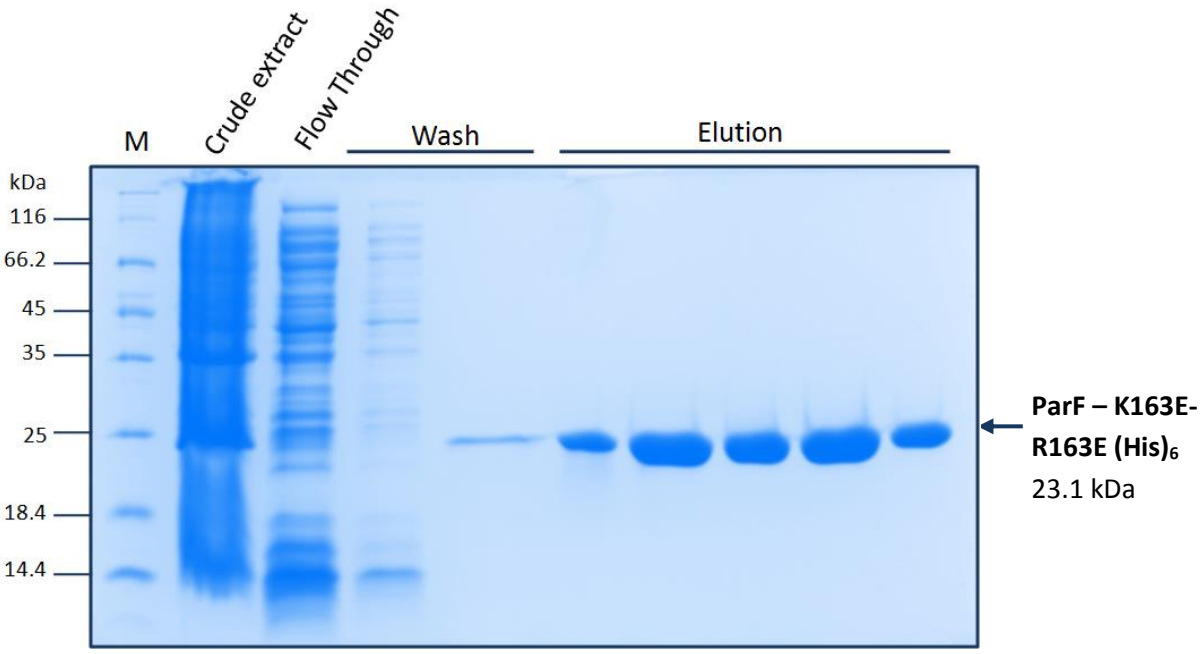

Figure 3.5 - SDS polyacrylamide gel of ParF-K160E-R163E Ni^{2+} affinity chromatography
(B) Fractions collected during ParF-K160E-R163E purification. The crude extract represents the cleared lysate of ParF-K160E-R163E-producing cells before nickel affinity chromatography. Non-binding proteins are shown in the "flow through" fraction. The column was washed with 15 mM (lane 4) and 85 mM (lane 5) imidazole to remove proteins binding non-specifically to the column (wash fractions). Lanes 5 to 10 represent ParF-K160E-R163E elution fractions after buffer exchange into storage buffer. M represents the MW marker (Pierce™ Unstained Protein MW marker).

.

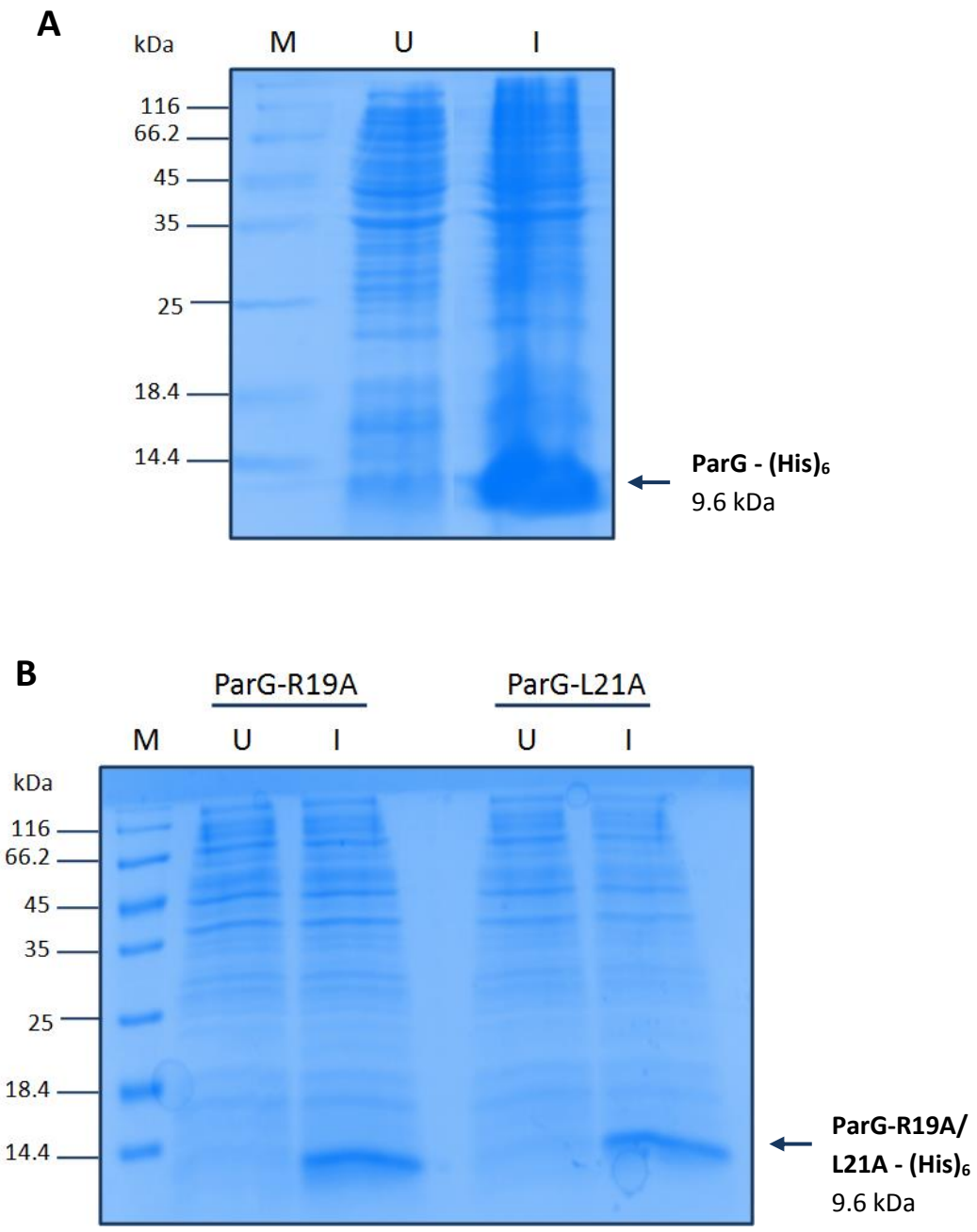

Figure 3.6 - SDS polyacrylamide gel of (A) wild type ParG, (B) ParG-R19A and ParG-L21A overproduction

(A) ParG overproduction test comparing uninduced (U) and induced (I) cells, using 1 mM IPTG. M represents the MW marker (Pierce™ Unstained Protein MW marker).

(B) ParG-R19A (lane 2 and 3) and ParG-L21A (lane 4 and 5) overproduction test comparing uninduced (U) and induced (I) cells, using 1 mM IPTG. M represents the MW marker.

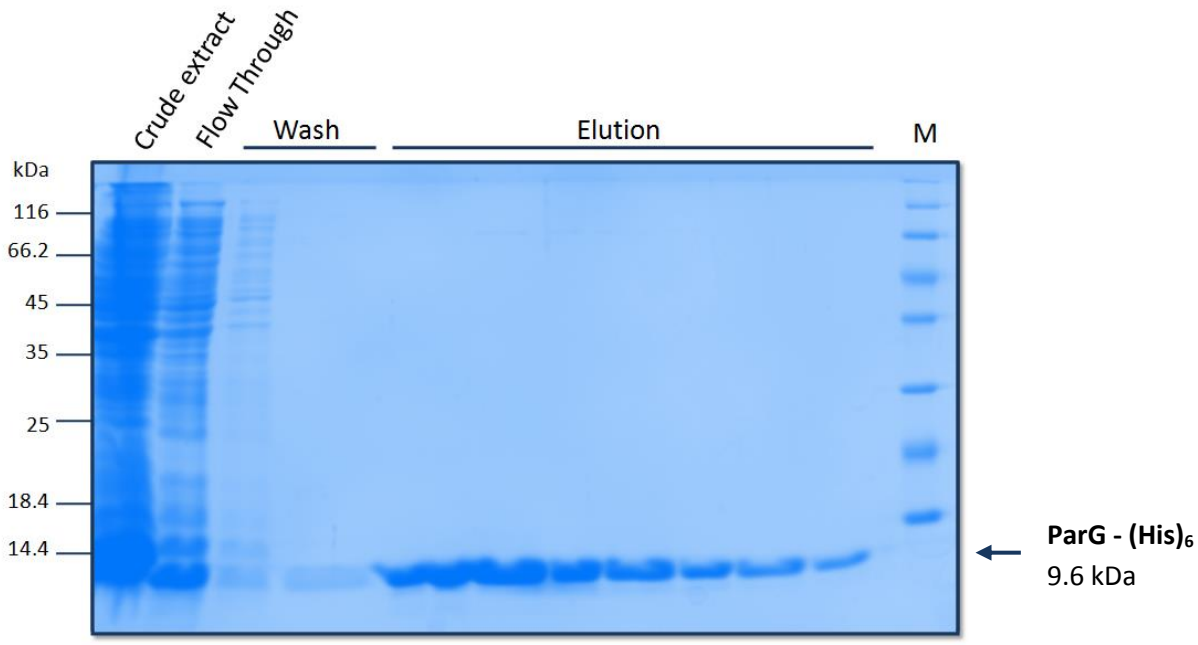

Figure 3.7 - SDS polyacrylamide gel of ParG Ni^{2+} affinity chromatography.
Fractions collected during ParG purification. The crude extract represents the cleared lysate of ParG-producing cells before nickel affinity chromatography. Non-binding proteins are shown in the “flow through” fraction. The column was washed with a buffer containing 15 mM (lane 3) and 90 mM (lane 4) imidazole to remove proteins binding non-specifically to the column (wash fractions). Lanes 5 to 12 show ParG elution fractions after buffer exchange into storage buffer. M represents the MW marker (Pierce™ Unstained Protein MW marker).

.

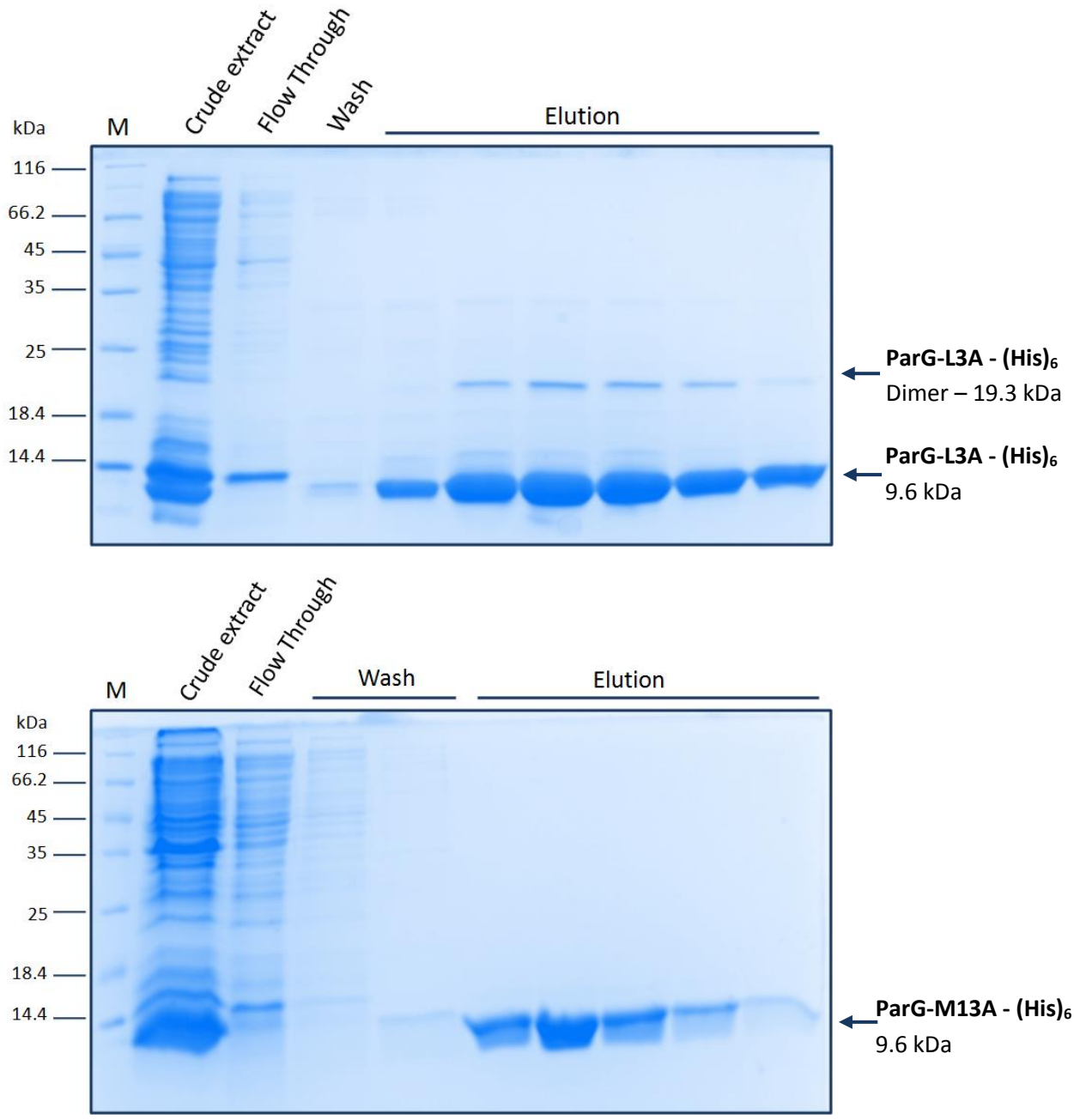

Figure 3.8 - SDS polyacrylamide gel of (A) ParG-L3A and (B) ParG-M13A Ni^{2+} affinity chromatography

Fractions collected during ParG mutant protein purification. In both (A) and (B), the crude extract represents the cleared lysate of ParG mutant protein-producing cells before nickel affinity chromatography. Non-binding proteins are shown in the “flow through” fraction. The column was washed with a buffer containing 15 mM (lane 4) and 90 mM (lane 5) imidazole to remove proteins binding non-specifically to the column (wash fractions). Lanes 6 to 10 represent ParG-L3A/M13A elution fractions after buffer exchange into storage buffer. M represents the MW marker (Pierce™ Unstained Protein MW marker).

3.2.3 Using MST to quantify the interaction between ParF and ParG mutants

3.2.3.1 Optimisation of the MST protocol

MST was used as a quantitative approach to assess the interaction between ParF and ParG mutants. MST is a technique based on thermophoresis, the directed movement of molecules caused by a temperature gradient. This behaviour depends on the molecular weight, charge and solvation shell of the molecule. Therefore, if a protein associates with a ligand, this will cause a change in the thermophoresis of the protein that can be used to obtain information regarding the interaction, since at least of one of the above parameters will change (Wienken *et al*., 2010). Thermophoresis can be monitored by fluorescently labelling the protein of interest, while the ligand is titrated at increasing concentration over a constant, low concentration of the first.

Although both interacting species are in solution, this technique presents numerous limitations for the system under investigation. Therefore, several steps of optimisation were required to obtain consistent and reproducible results.

In order to screen systematically the interaction between ParF and ParG mutants, first ParF was fluorescently labelled. However, while labelling was easily optimised, the self-associating nature of ParF made it impossible to distinguish whether the change in thermophoresis was caused by ParG binding or ParF self-association. Low concentrations of Tween-20 (0.05 %) and bovine serum albumin (BSA) (0.5 mg/ml) were added to improve protein solubility and stability. However, this did not improve the quality of the results. Buffer conditions were also tested, with no improvement.

As a consequence, ParG and ParG mutants were fluorescently labelled instead of ParF. ParG is generally more soluble than ParF and when the cognate DNA binding site *parH* is not present, it does not oligomerise. Therefore, ParG represented a better candidate for thermophoresis. ParG and ParG mutants were labelled following the protocol in 2.6.1.1 and MST experiments were carried out as described in the Materials and Methods.

3.2.3.2 ParG mutants bind to ParF with lower affinity compared to wild type ParG

MST experiments were carried out in triplicate for all the interactions between ParF and ParG/ParG mutants.

The outcome from the MST instrument was expressed in Normalised Fluorescence values and converted into fraction bound using the MO.Control software from Nanotemper Technologies. The difference between bound and unbound Normalised

Fluorescence values recorded by MST is referred as amplitude. When the amplitude is at least three times larger than the fluctuation of the points of the baseline, binding is considered authentic. Instead, if the amplitude value is smaller, the change in fluorescence can be considered caused by random fluctuation rather than binding. Conversion of the Normalised Fluorescence values into fraction bound was done to allow easy comparison between ParG mutants. However, the amplitude value was used to check binding authenticity before conversion. If the mutant showed no binding, as in the case of ParG-K5A, the values recorded were left as Normalised Fluorescence.

Binding curves were obtained fitting the points using the one-site specific binding non-linear regression equation according to the formula:

$$Y = \frac{B_{max}X}{K_D + X}$$

where *Y* is the fraction bound, *X* is the concentration of the free ligand protein, B_{max} is the maximum binding and K_D is the equilibrium dissociation constant (GraphPad Prism 7.04).

MST showed that the interaction between ParF and wild type ParG has an equilibrium dissociation constant of ~ 0.9 µM (Figure 3.9).
Most of the ParG mutants showed lower binding affinity for ParF, compared to wild type ParG (Figures 3.10 and 3.11). The only two mutants that behave differently are ParG-L21A that shows an affinity for ParF similar to that of the wild type protein and ParG-K5A that does not seem to interact with ParF (Figure 3.10). A summary of the K_D recorded using MST is shown in Table 2 and in Figure 3.12. For some mutant (e.g. K11A and M13A), the magnitude of the standard error is substantial and the K_D value calculated is therefore less accurate.

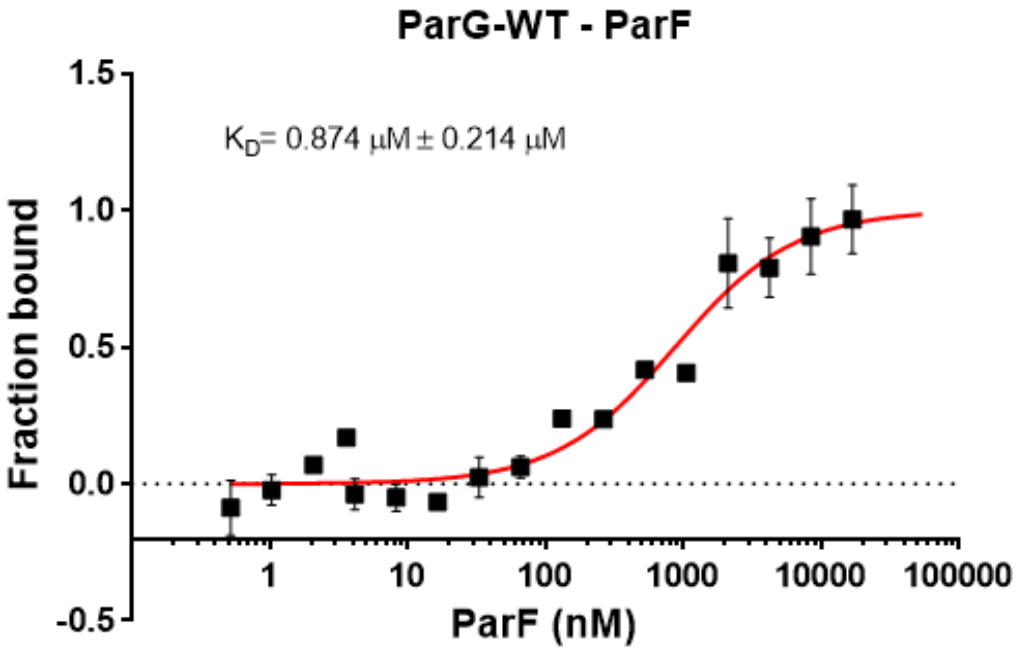

Figure 3.9 - MST binding curve of the interaction between ParF and wild type ParG

Sixteen ParF concentrations (0.52 nM to 16.95 µM) were titrated over 200 nM ParG. The experiment was done in triplicate. The values were averaged and error bar plotted calculating the standard error of the mean for each point. The curve fitting follows the one-site specific binding equation (Prism GraphPad 7.04).

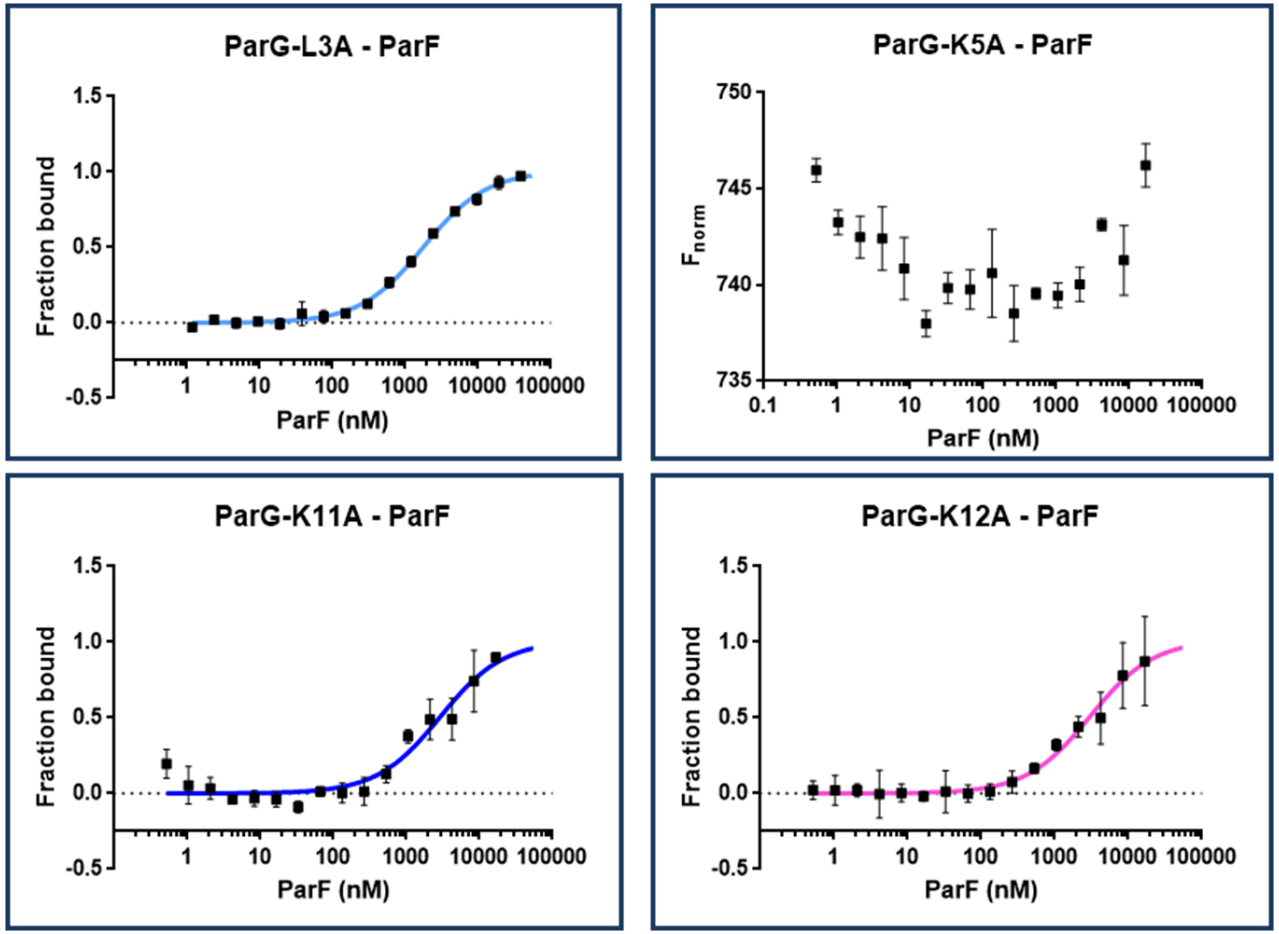

Figure 3.10 - MST binding curve of the interaction between ParF and ParG-L3A/K5A/K11A/K12A

Sixteen ParF concentrations (from 0.52 nM to 16.95 µM for the interaction with ParG-K5A/K11A/K12A and from 1.19 nM to 39 µM for the interaction with ParG-L3A) were titrated over 200 nM ParG. The experiment was done in triplicate. The values were averaged and error bar plotted calculating the standard error of the mean for each point. The curve fitting follows the one-site specific binding equation (Prism GraphPad 7.04). No curve fitting was possible for the interaction between ParG-K5A and ParF because of low fluorescence change amplitude (no binding). Values are therefore shown as normalised fluorescence rather than fraction bound

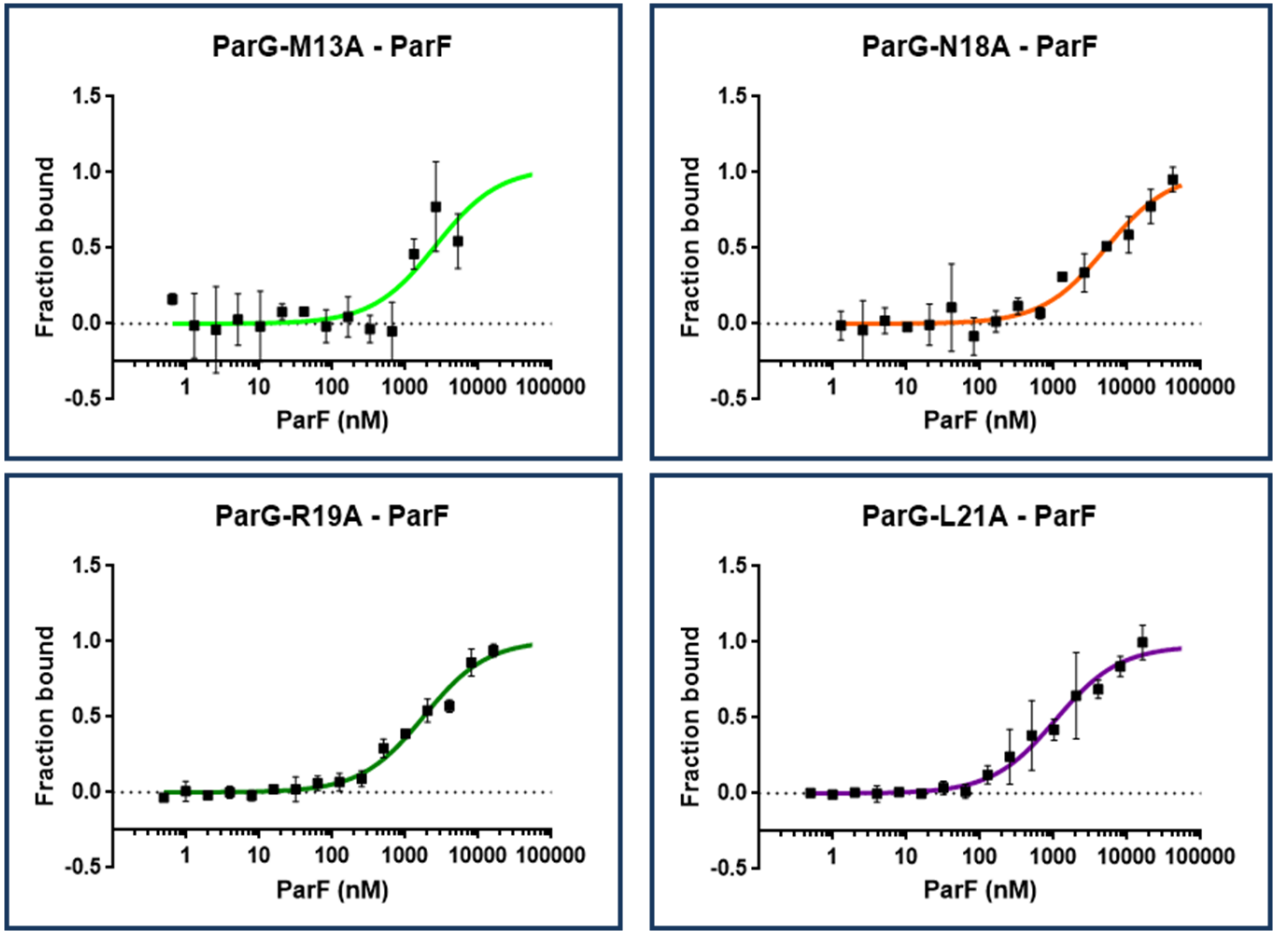

Figure 3.11 - MST binding curve of the interaction between ParF and ParG-M13A/N18A/R19A/L21A

Sixteen ParF concentrations (from 0.5 nM to 16.2 µM for the interaction with ParG-M13A/R19A/L21A and from 1.2 nM to 42 µM for the interaction with ParG-N18A) were titrated over 200 nM ParG. The experiment was done in triplicate. The values were averaged and error bar plotted calculating the standard error of the mean for each point. The curve fitting follows the one-site specific binding equation (Prism GraphPad 7.04).

ParG protein	K_D (μM)	Standard Error (μM)
ParG – WT	0.874	0.214
ParG – L3A	1.791	0.103
ParG – K5A	N/A	N/A
ParG –K11A	2.893	0.956
ParG – K12A	3.023	0.434
ParG – M13A	2.595	2.418
ParG – N18A	5.140	1.139
ParG – R19A	1.797	0.285
ParG – L21A	0.648	0.152

Table 3.2 - K_D values and standard errors of the interaction between ParF and ParG/ParG mutants

K_D values were calculated by fitting the curve with a non-linear one-site specific binding equation. GraphPad Prism 7.04 was used to do the fittings. The colour represents the effect of the residue changehad on the ParG binding with ParF. Green represents binding not affected; yellow, binding slightly affected; orange, binding moderately affected and red, no binding being recorded.

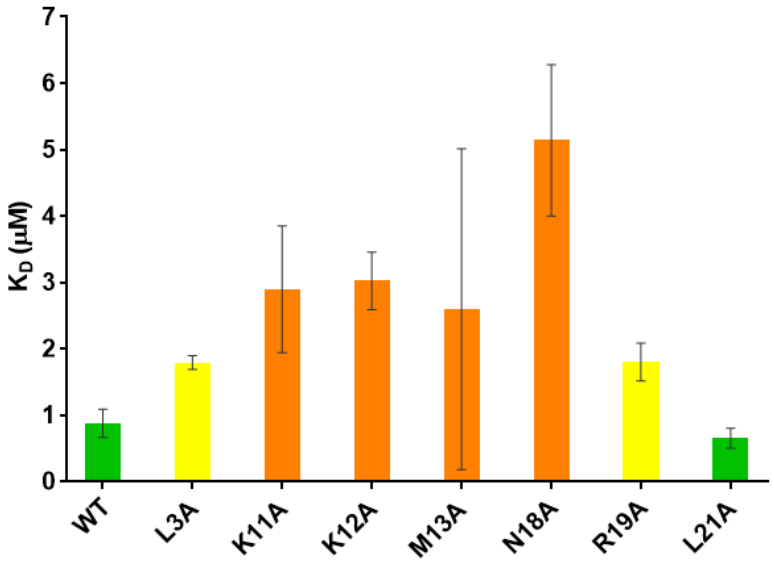

Figure 3.12 – Histogram showing the K_D values and standard errors of the interaction between ParF and ParG/ParG mutants

K_D values of the interaction with ParF are represented for each ParG mutant, including the standard error calculated from the three repeats of the experiment. The colour coding follows the one used in Table 3.2, to represent the effect that single amino acid change had on ParG binding to ParF. Green represents binding not affected; yellow, binding slightly affected; orange, binding moderately affected and red, no binding being recorded.

3.2.4 SPR investigation of the kinetics of the interaction between ParF and ParG/ParG mutants

3.2.4.1 Optimisation of the SPR protocol

SPR was used to validate the results obtained by MST. In fact, MST was poorly reproducible due to the heterogeneity of the sample, as ParF self-associates even in the absence of ATP. As a consequence the reaction between ParF and ParG likely contained ParF monomers, as well as dimers and small oligomers (Barillà *et al.*, 2005). On the other hand, ParG is mostly found as a dimer (Saeed *et al*., 2015). Due to these limitations, SPR also required several optimisation steps.

The SPR set up requires the immobilisation of one of the binding partners. Since the aim of the experiment was to screen systematically the binding of ParG mutants to ParF, the ParF protein was selected as the partner to be immobilised. The self-associating nature of ParF represented a challenge, because upon immobilisation, different ParF species were present on the chip. To make ParF more soluble and mono-disperse, the protein was immobilised in presence of ADP and $MgCl_2$. After immobilisation ADP was removed by flowing ethylenediaminetetraacetic acid (EDTA) over the ParF-ADP coated chip. EDTA is a chelating agent and sequesters magnesium, thus enhancing ADP release from the protein. However, while ADP improved the mono-dispersity of the sample, the EDTA wash affected ParF ability to bind to ParG.

To overcome the limitations caused by ParF's sticky nature, a ParF mutant was used for these experiments. The ParF-K160E-R163E mutant was originally constructed to abolish ParF interaction with non-specific DNA (Allison-Gamble, 2016). The lysine and the arginine residues that were converted into glutamic acid are surface exposed (Figure 3.13). Allison demonstrated that while the mutant was unable to self-associate into higher order structures in the absence of nucleotide, it could slowly polymerise in presence of ATP (Allison-Gamble, 2016). Bacterial two hybrid assays showed that the mutant ParF protein interacts with ParG as the wild type ParF (Allison-Gamble, 2016).

ParF-K160E-R163E interaction with ParG was further investigated using MST. MST supported the two hybrid *in vivo* results, since the estimated K_D of the interaction between ParF and ParF-K160E-R163E was comparable to that observed for wild type ParG (Figure 3.14). The interaction between ParG and ParF-K160E-R163E showed a K_D= 0.927 ± 0.264 µM, while the interaction between ParG and wild type ParF showed a K_D= 0.874 ± 0.214 µM.

As the ParF double mutant's ability to interact with ParG was similar to the that of the wild type protein, ParF-K160E-R163E was used as a substitute to investigate the interaction of ParF with ParG mutants.

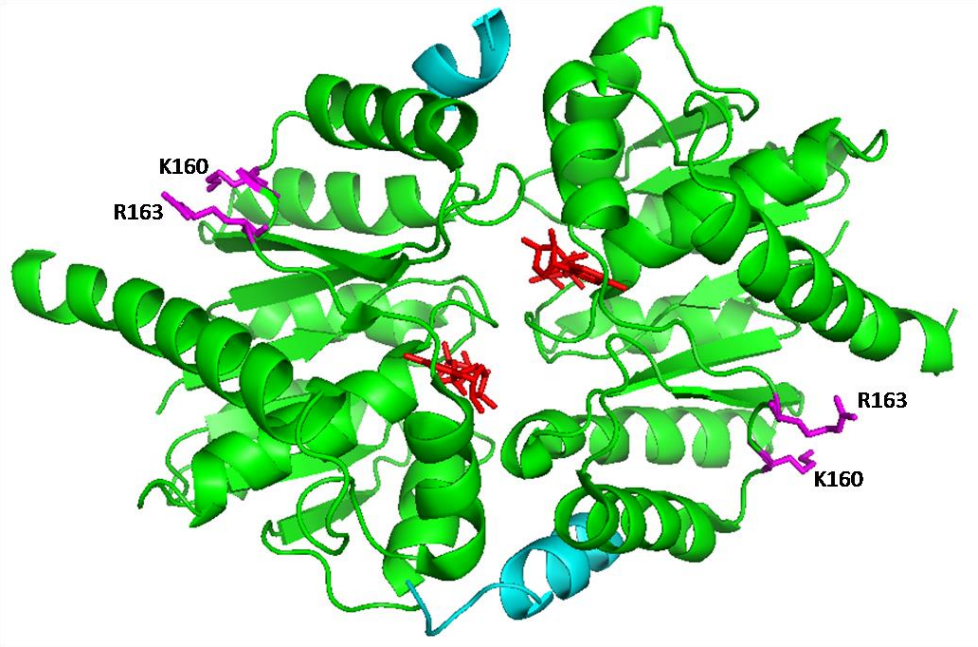

Figure 3.13 - Crystal structure of ParF bound to a small fragment of the ParG N-terminus
The green structure represents the ParF dimer, while the small helix represents the ParG arginine finger motif. AMPPCP is sandwiched between ParF monomers and shown in red. The two amino acids shown in purple are the surface exposed Lys-160 and Arg-163 (PDB 5U1G). The image was generated using PyMOL v1.8.6.0 (The PyMOL Molecular Graphics System, Version 1.8.6.0, Schrödinger, LLC).

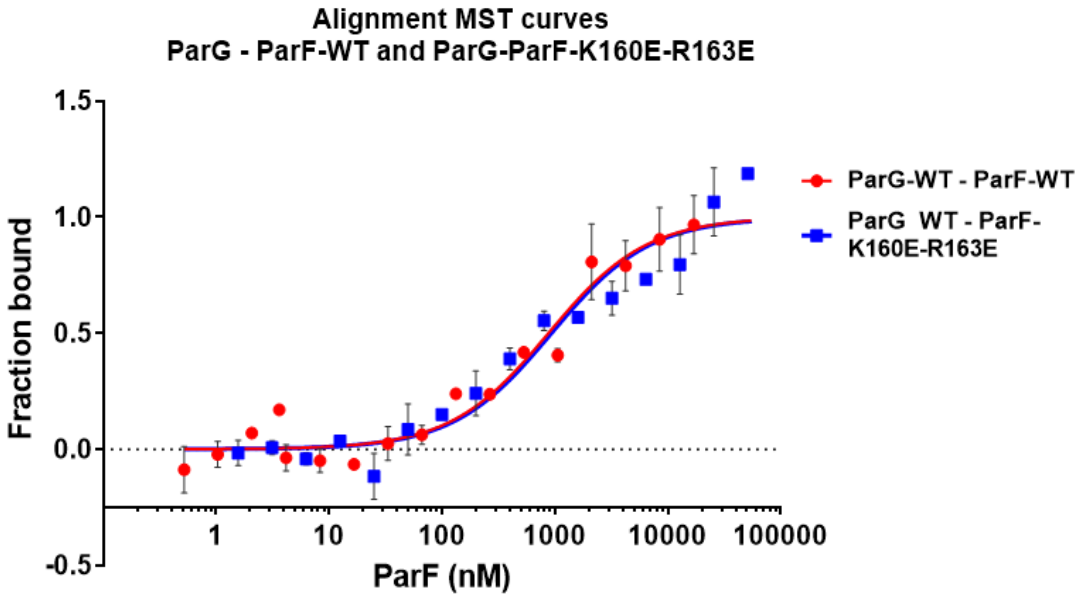

Figure 3.14 - ParG-ParF and ParG-ParF-K160E-R163E MST binding curves alignment
The curve in blue represents the interaction between wild type ParG and ParF-K160E-R163E. Sixteen ParF concentration (from 1.5 nM to 51.1 µM) were titrated over 200 nM wild type ParG. The curve in red represents ParG-ParF-WT (Prism GraphPad 7.04).

3.2.4.2 SPR shows that ParG mutants bind ParF with decreased affinity as compared to wild type ParG

ParF-K160E-R163E was immobilised on CM5 sensor chip by amine coupling to the carboxymethylated dextran covalently attached to the gold chip surface, as described in 2.6.2.1.

Two sets of experiments were performed in order to analyse all the ParG mutants. As ParF was stable on the chip for up to 24 hours, the protein was immobilised twice on different flow cells of the same chip, following the protocol described in the Materials and Methods section. The CM5 chip contains four flow cells (fc). In the first set of experiments, fc-1 was used as a reference cell and ParF was immobilised on fc-2. In the second set of experiments, fc-3 was used as reference cell and ParF was immobilised on fc-4. For both immobilisations, wild type ParG was used as a binding control (Figure 3.15). Experiments were performed over four days. ParG/ParG mutants were flowed over the ParF-coated chip at increasing concentration from 58.6 nM to 15 µM. ParG dilutions were performed by serial dilution to improve accuracy. The first and the last concentration were injected twice to check binding reproducibility.

SPR showed that most ParG mutants interact with lower affinity with ParF, as compared to wild type ParG. Figure 1.14 shows the binding curves for the interaction between ParF and ParG-WT. Conversion of residues Leu-3 and Lys-5 into alanine abolishes ParG binding to ParF almost entirely (Figure 3.16). Flowing high concentration of these mutants showed very little response, suggesting that these residues may play an important role in anchoring ParG to ParF. The other ParG mutants showed decreased interaction with ParF, although a good level of binding was still recorded (Figure 3.17 and 3.19).

Importantly, the immobilisation of ParF on the chip is random. The protein sequence harbours fourteen lysine residues. Some of them are surface exposed and probably more reactive than others, however the orientation of the molecule on the surface of the chip is unknown. For this reason, wild type ParG has been used as binding control for both rounds of ParF immobilisation.

From the binding curves (Figures 1.16 1.17, 1.18 and 1.19), it was possible to extrapolate the association rate constant (k_a) and the dissociation rate constant (k_d) of the interaction between ParF and the different ParG mutants. From those values, it was possible to obtain an estimated equilibrium dissociation constant (K_D).

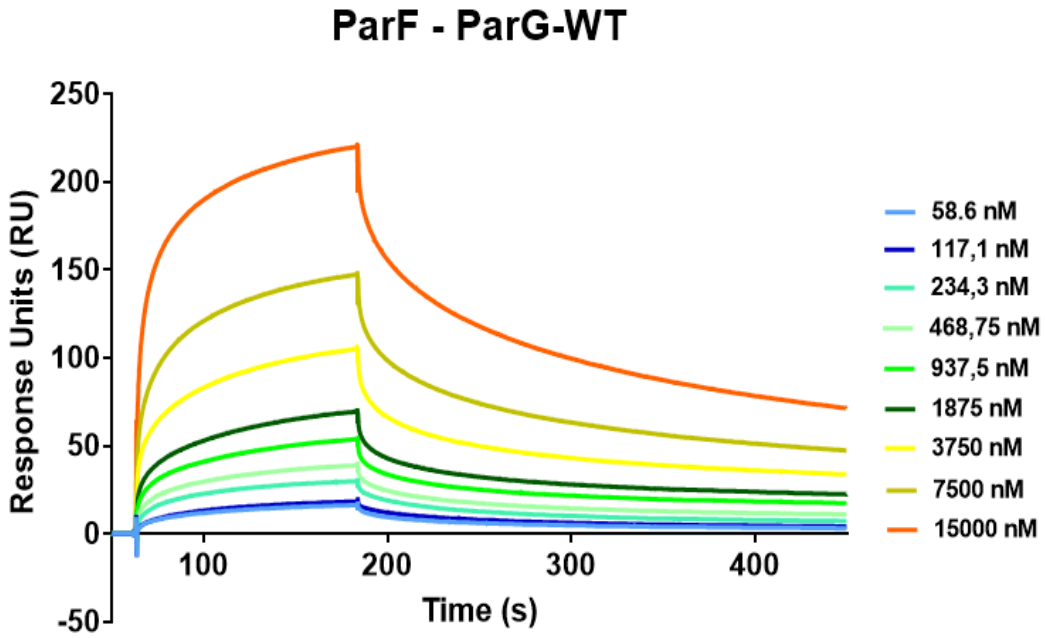

Figure 3.15 - SPR binding curves of the interaction between ParF and wild type ParG
Nine ParG concentrations were flowed over the ParF-coated CM5 chip and the response was recorded as Response Units (RU) over time.

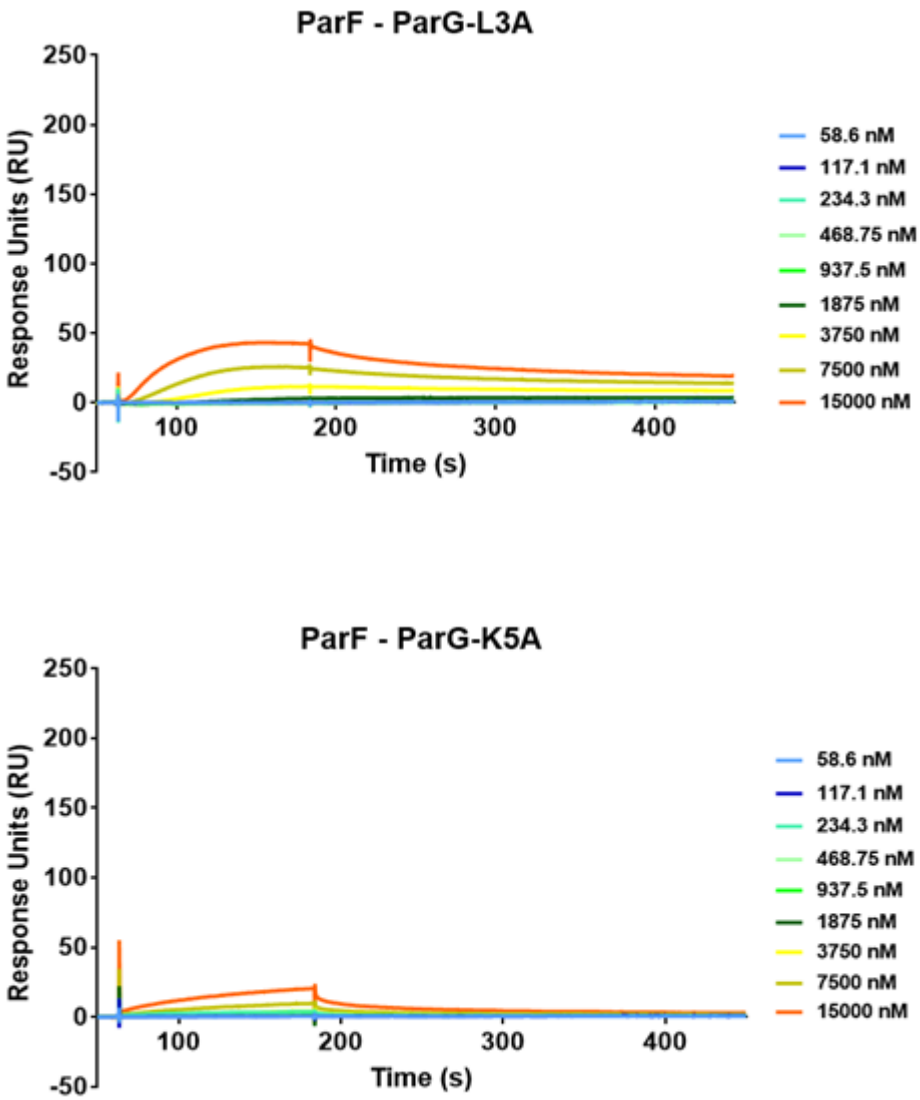

Figure 3.16 - SPR binding curves of the interaction ParF – ParG-L3A and ParF - ParG-K5A
Nine ParG-L3A/K5A concentration were flowed over the ParF-coated CM5 chip and the response was recorded as Response Units (RU) over time.

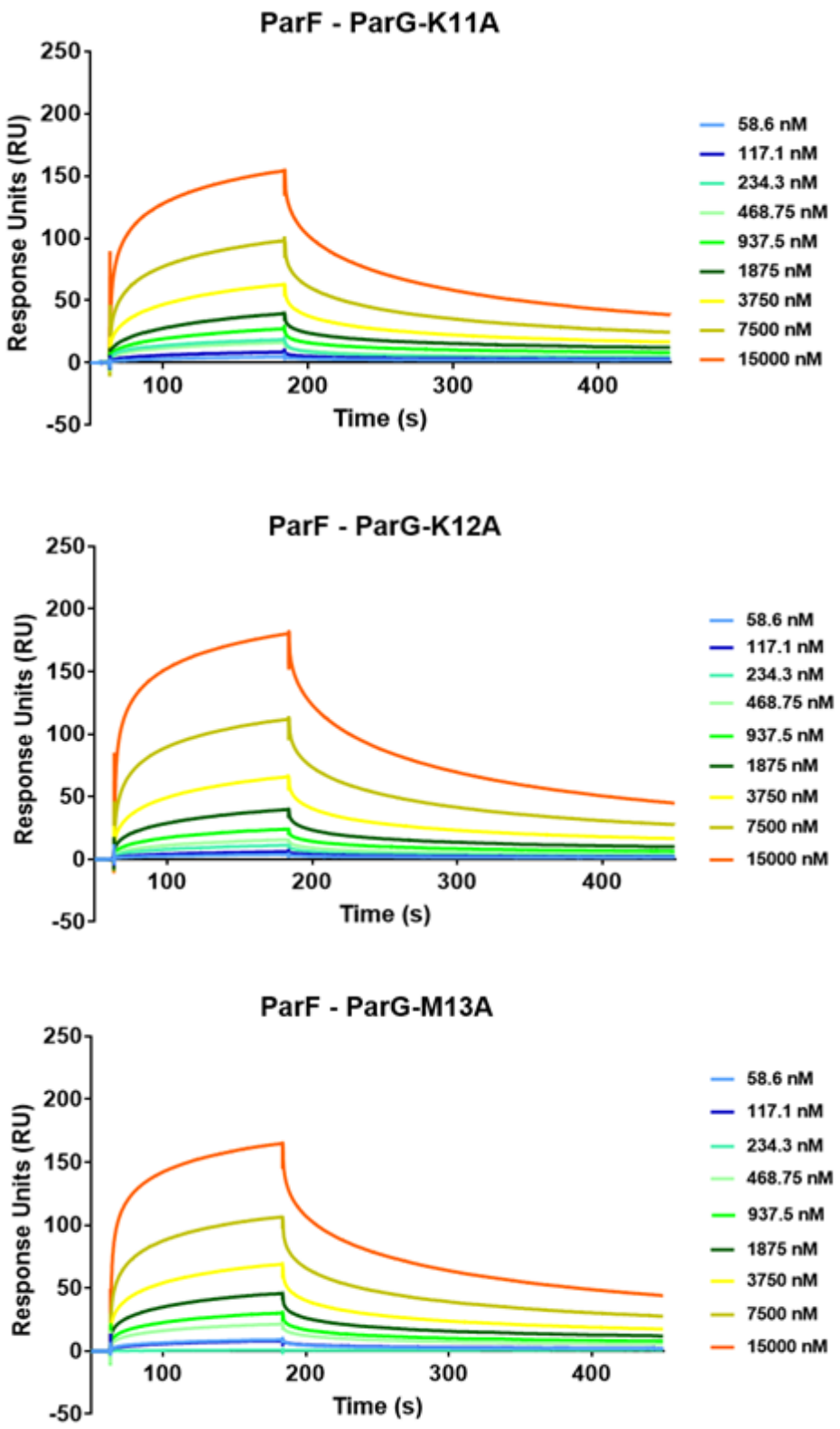

Figure 3.17 - SPR binding curves of the interaction between ParF – ParG N-terminal mutants harbouring changes in the middle cluster of amino acids

Nine ParG-K11A/K12A/M13A concentration were flowed over the ParF-coated chip and response was recorded as Response Units (RU) over time.

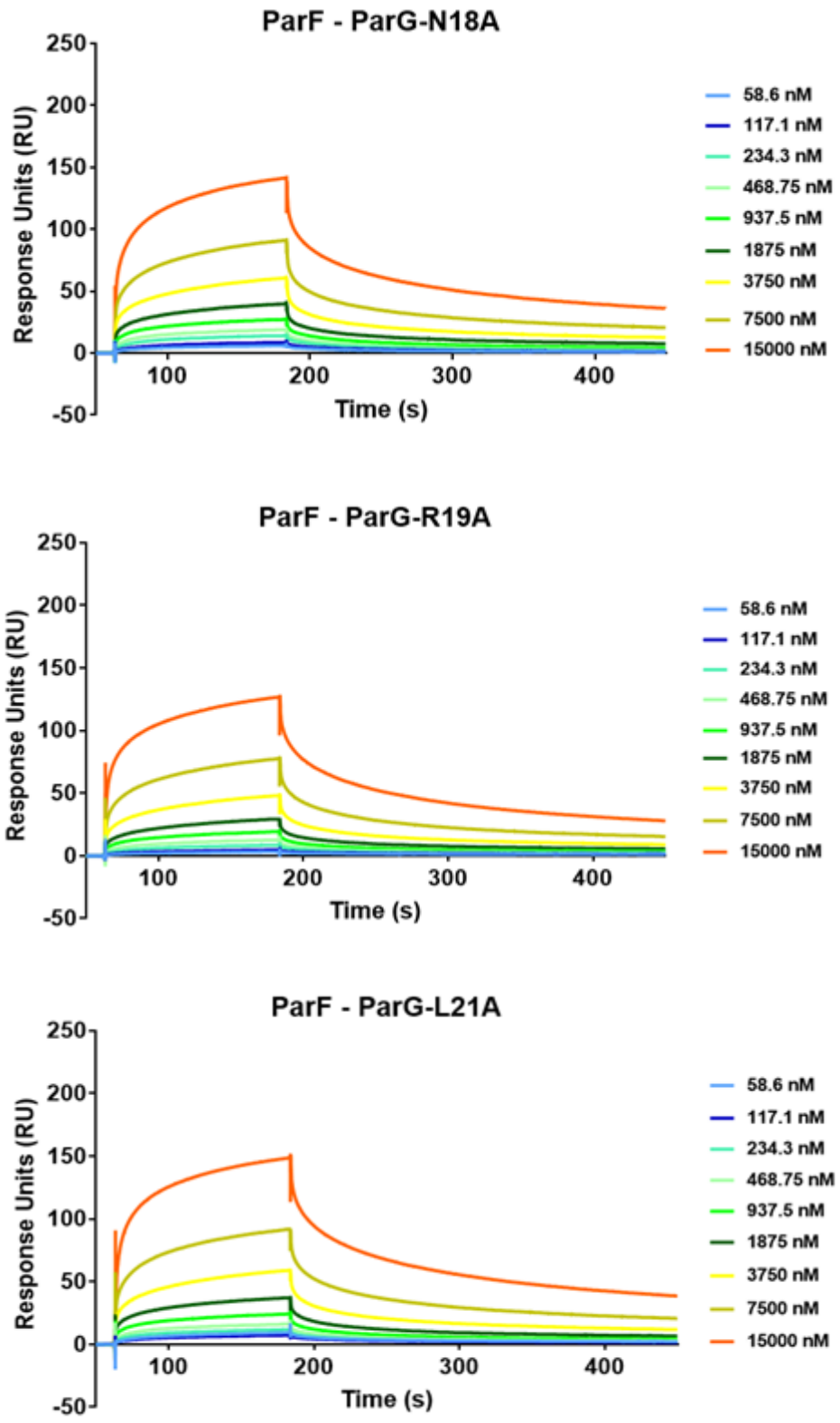

Figure 3.18 - SPR binding curves of the interaction between ParF and ParG N-terminal mutants harbouring changes in the arginine finger motif

Nine ParG-N18A/R19A/L21A concentration were flowed over the ParF-coated CM5 chip and the response was recorded as Response Units (RU) over time.

The SPR curves do not appear to reach steady state completely, as after a fast association, they grow slowly before dissociation. This may be due to ParG interacting specifically as well as non-specifically with ParF, depending on the orientation of the protein on the chip. Therefore, ParG could interact with its specific binding site and transiently and weakly with different regions of ParF. It is also possible that ParG monomers were present in solution in addition to the most abundant ParG dimers. Therefore, two types of binding would take place, ParF-ParG dimer and ParF-ParG monomer. It cannot be ruled out that ParF immobilisation was not completely homogenous and that different ParF species were present on the chip.

Although not perfectly fitting, a 1:1 binding model was used to estimate the k_a and k_d values from the binding curves (Table 3.3). According to this model, the reaction between the ligand ParF and the analyte ParG follows a pseudo first order kinetics. The ParF-ParG complex forms as a function of time and when the protein flow stops, the complex dissociates in a time dependent manner. The integrated rate equation for this model of binding is the following:

$$R_{eq} = \frac{k_a C}{k_a C + k_d} \times R_{max}$$

Where R_{eq} is the response when the binding between ligand and analyte is at the equilibrium, R_{max} is the maximum response and C is the concentration of the analyte. The K_D can then be calculated from:

$$K_D = \frac{k_d}{k_a}$$

Figure 3.19 shows an alignment of the SPR binding curves of the interaction between immobilised ParF and ParG and ParG mutants (15 µM). The SPR response is proportional to the strength of the binding between the immobilised ligand protein and the analyte, that is flowed over the chip.

ParG mutant	k_a (1/Ms)	k_d (1/s)	K_D (μM)
ParG – WT	1567	0.0 02105	1.34
ParG – L3A	N/A	N/A	N/A
ParG – K5A	N/A	N/A	N/A
ParG –K11A	1222	0.002749	2.25
ParG – K12A	1849	0.005901	3.19
ParG – M13A	1127	0.002440	2.16
ParG – N18A	1905	0.005001	2.62
ParG – R19A	673.6	0.002969	4.40
ParG – L21A	613.1	0.002546	4.15

Table 3.3 - ParF – ParG/ParG mutant K_D of the interaction

The k_a and k_d values were obtained using a 1:1 binding model for the curves fitting (Prism GraphPad 7.04).The green colour representswild type ParG binding affinity for ParF. Orange represents lower interaction than wild type ParG. Red represents no binding or severely affected interaction.

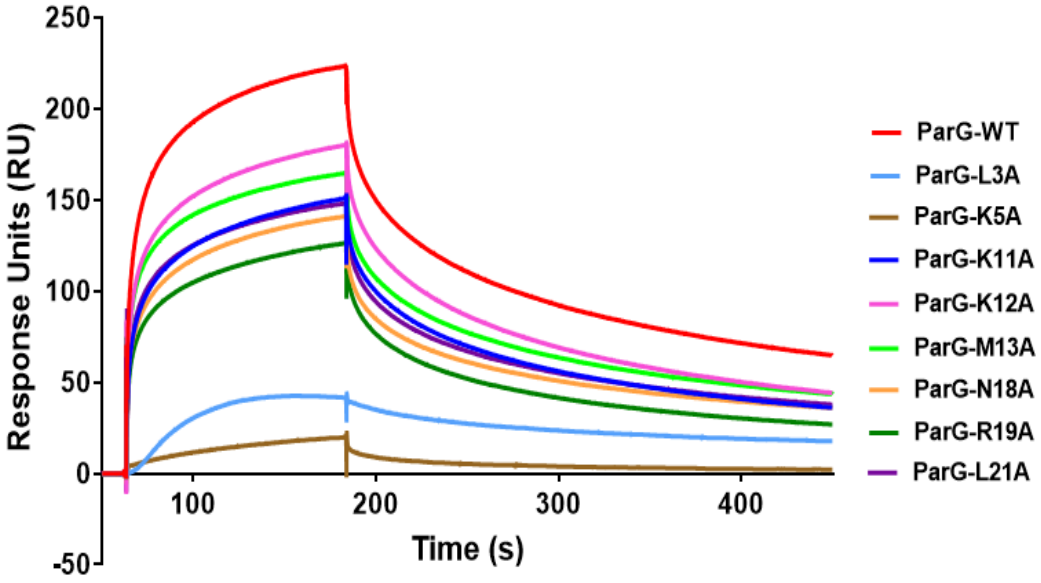

Figure 3.19 - Alignment of the SPR binding curves obtained for the interaction between ParF and ParG/ParG N-terminal mutants

After recording the ParF baseline signal, ParG/ParG mutants (15 μM) were flowed over the ParF-coated CM5 chip for two minutes at a 30 μl/min flow rate. After two minutes, buffer was injected and ParG slowly dissociated. The chip surface was regenerated with 2M NaCl.

The curve alignment shows that residue changes in the same cluster of amino acids cause a similar response in the binding to ParF (Figure 3.21). Residue changes in the tip of the tail, e.g. L3A and K5A, have the most detrimental effect on ParF binding. Perturbations in the cluster of the arginine finger motif cause a decrease in ParF binding, since the K_D recorded is three-four times higher than the one measured for the wild type protein. Finally, mutations in the middle cluster of amino acids (Lys-11, Lys-12, Met-13) reduce ParF-ParG binding by two-three fold.

Steady-state binding was investigated by plotting the SPR response at the end of the protein injection (time= 182 sec), before dissociation, versus the protein concentration (Figure 3.20). K_D values were calculated by fitting the points with a non-linear equation representing a one-site total binding model that takes into account a specific-binding hyperbolic component and a linear non-specific component.

$$Y = \frac{B_{max}X}{K_D + X} + (nsX)$$

Where X is the protein concentration, B_{max} is the maximum binding, the K_D is the equilibrium dissociation constant and *ns* is the slope of the linear non-specific component of the binding.
The fitting of the data was done with Prism GraphPad 7.04 (Table 3.4).

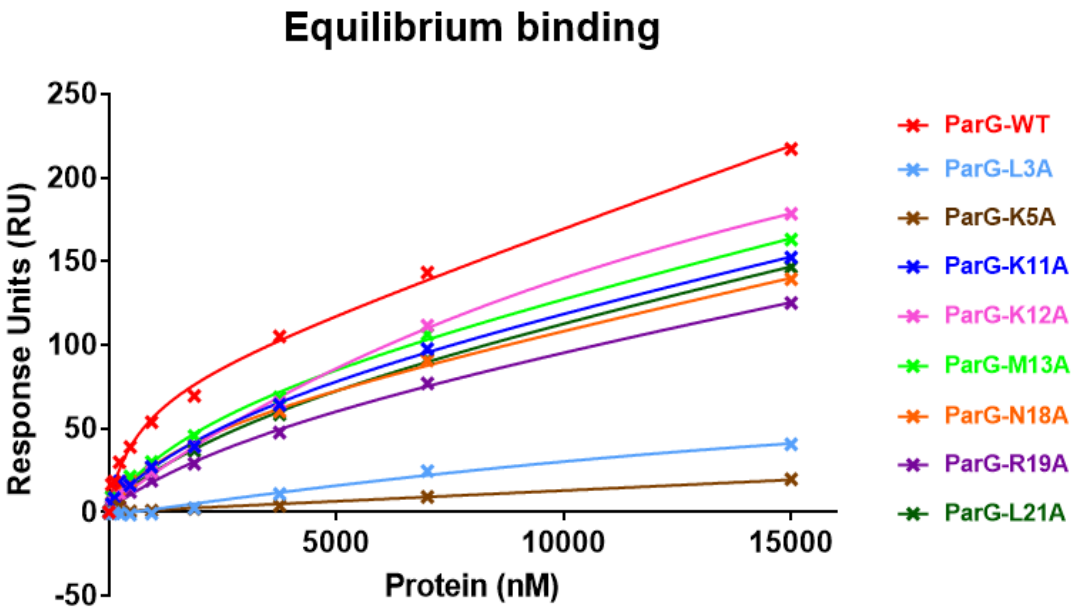

Figure 3.20 - Alignment of the steady-state binding curves of the interaction between ParF and ParG/ParG mutants

The curves fitting follows a one-site total binding model (Prism GraphPad 7.04).

Protein	B_{max} (RU)	K_D (nM)	*ns* (RU/nM)
ParG-WT	72.04	724.2	0.009615
ParG-L3A	75009	1006003	-0.07059
ParG-K5A	~ 182,2	~ 38956070	~ 0,001277
ParG-K11A	75.11	3470	0.005794
ParG-K12A	~ 373,4	18365	0.0004653
ParG-M13A	75.48	2743	0.00637
ParG-N18A	55.24	1918	0.005845
ParG-R19A	60.18	4199	0.005091
ParG-L21A	71.89	4176	0.005728

Table 3.4 - K_D of the steady-state binding curves obtained from the interaction between ParF and ParG/ParG mutants

The curve fitting follows a one-site total binding model and the K_D has been extrapolated from this fitting. The values of the binding curves of the mutants shown in grey could not be calculated accurately by the software (Prism GraphPad 7.04).

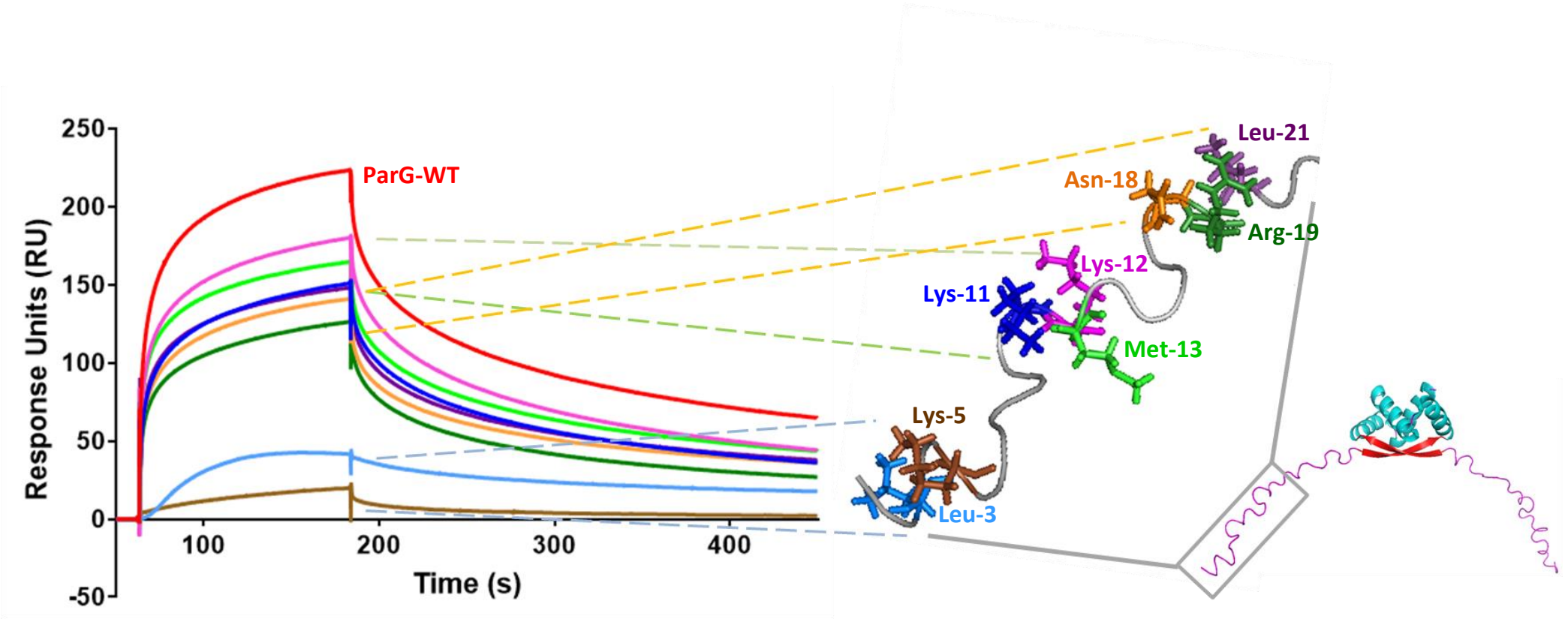

Figure 3.21 - Alignment of the SPR binding curves for the interaction between immobilised ParF and ParG/ParG N-terminal mutants

The amino acids shown on the ParG structure (PDB 1P94) were changed into alanine. The dashed lines connect the SPR binding curves to the amino acid that was changed in the corresponding ParG protein mutant. The line in red represents the control ParF - ParG-WT binding curve.

3.2.6 Chemical cross-linking shows that the ParG-L3A mutant interacts weakly with both wild type ParF and ParF- K160E-R163E

MST and SPR have been used as complementary approaches to measure the interaction between ParG N-terminal mutants and ParF in order to overcome the limitations of a single technique. However, for some of the mutants it was impossible to obtain an unambiguous result. ParG-L3A shows conflicting results in MST and SPR experiments. To prove or disprove the results obtained through the biophysical investigation, chemical cross-linking was used as qualitative evidence. DMP (dimethyl pimelimidate) cross-linking of ParF and ParF-K160E-R163E with ParG-L3A was carried out as detailed in 2.6.3.1, to assess whether or not binding was taking place. In addition, ParG-Δ9N was also cross-linked with both ParF proteins to investigate whether removal of the N-terminal region abolished ParG interaction with the partner protein. SPR and MST showed conflicting results also for the Δ9N deletion mutant (Figure 3.22).

The cross-linking experiment suggests that ParG-L3A and ParG-Δ9N interact with both ParF and ParF-K160E-R163E (Figure 3.23). However, quantification of the complex bands by ImageJ (National Institutes of Health) showed that ParF – ParG-L3A complex band is around seven times less intense than ParF - ParG-WT complex band, suggesting that the interaction is greatly affected by the change in position 3. Interestingly, the deletion of the nine N-terminal amino acids of ParG does not affect the binding as much as the change of Leu-3 alone. In fact, quantification shows that the band corresponding to ParF - ParG-Δ9N complex is between three to four fold less intense than the ParF – ParG band.

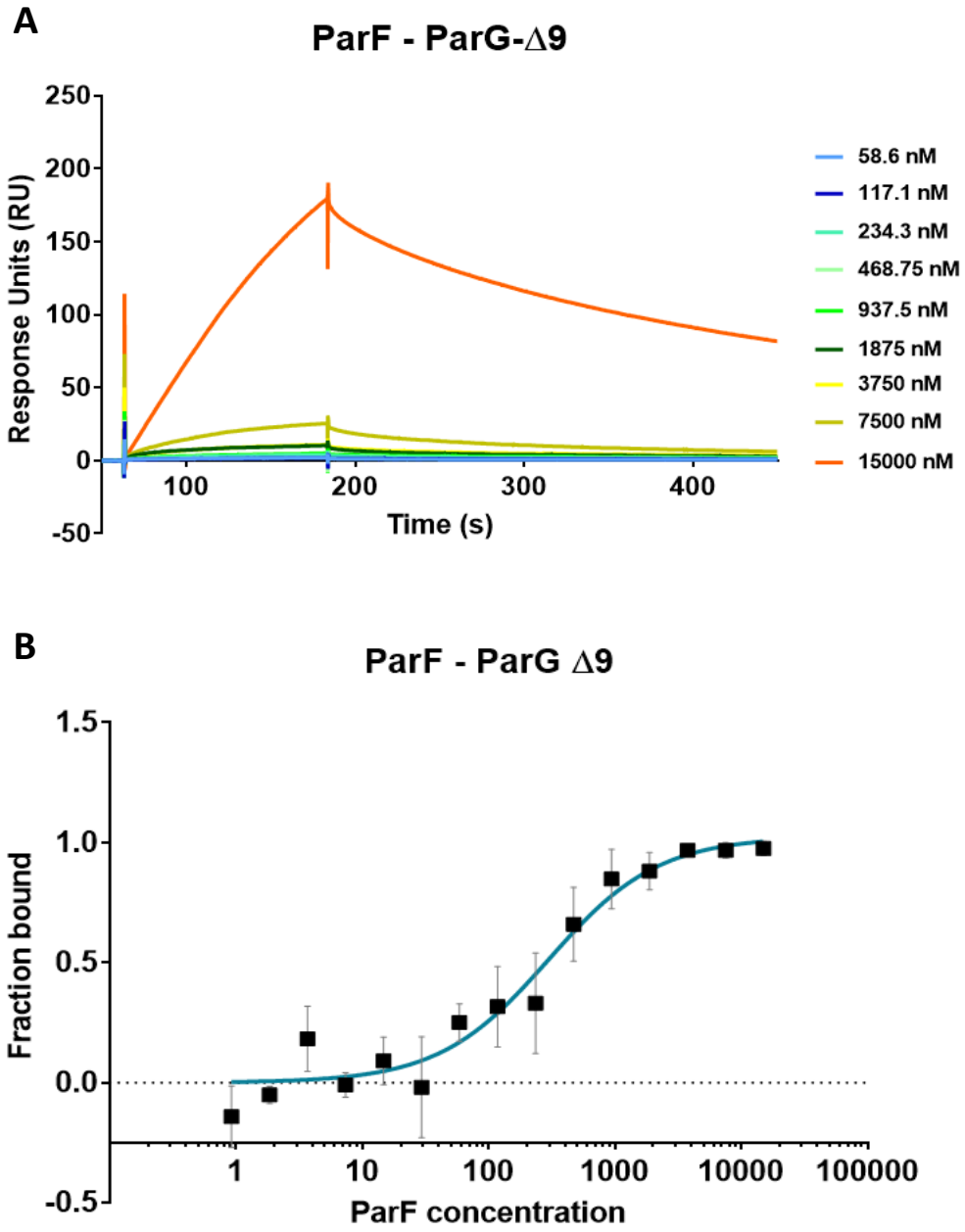

Figure 3.22 - SPR (A) and MST (B) binding curves of the interaction ParF – ParG-Δ9N

(A) Nine ParG-Δ9N concentrations were flowed over the ParF-coated CM5 chip and the response was recorded as Response Units (RU) over time. The high response observed for ParG-Δ9 (15 μM) is due to non-specific binding to the sensor chip surface. Overall, the protein binds very weakly to ParF.

(B) Sixteen ParF concentrations (0.9 nM to 15 μM) were titrated over ParG-Δ9N (200 nM). The experiment was done in triplicate. The values were averaged and error bar plotted calculating the standard error of the mean for each point. The curve fitting follows the one-site specific binding equation. The K_D of the interaction calculated by the one-site saturation binding equation is 0.296 μM ± 0.068 μM.

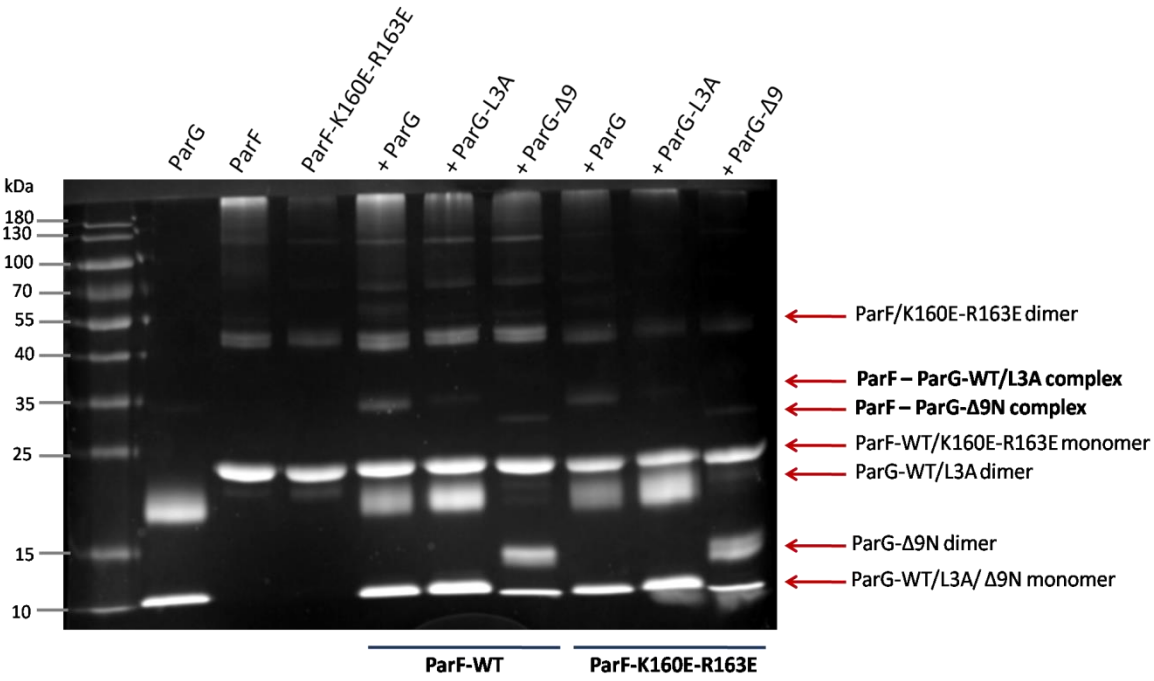

Figure 3.23 - Chemical cross-linking between ParF/ParF-K160E-R163E and ParG/ParG-L3A/ParG-Δ9N

Lane 2, 3 and 4 contain ParG, ParF-WT and ParF-K160E-R163E cross-linking controls. DMP (10 mM) was used in all the cross-linking reactions. Lanes 5, 6 and 7 show the cross-linking reactions between ParF-WT and ParG-WT/L3A/Δ9N. Lanes 8, 9 and 10 show the cross-linking reactions between ParF-K160E-R163E and ParG-WT/L3A/Δ9N. Lane 1 contains the MW marker PageRuler™ Plus Prestained Protein Ladder.

3.3 Conclusions

Analysis of the interaction between ParF and ParG N-terminal mutants allowed the identification of three distinct regions in the ParG tail that interact with ParF to a different extent.

MST showed that the interaction between ParG and ParF has an equilibrium dissociation constant K_D ~ 0.9 μM. Most of the ParG mutants interact with ParF with lower affinity, besides ParG-L21A which shows similar binding as that of wild type ParG. ParG-L3A and ParG-R19A could still interact with ParF quite strongly, since the K_D of the interaction was only less than two-fold higher compared to the ParF-ParG equilibrium dissociation constant. This suggests that these amino acids may not be essential for ParG binding to the partner protein and might have other roles.

ParG proteins harbouring changes in the middle cluster of amino acids (e.g. K11A, K12A and M13A) could also still interact with ParF quite strongly, with a two-to-three-fold increase in the K_D compared to the one recorded for the wild type protein. This cluster also does not seem to be essential for the interaction with ParF and, based on previous observations (Golovanov *et al.*, 2003), the decrease in the binding might be due to the effect that these changes have on the overall structure of ParG N-terminal domain.

On the other hand, the change of Asn-18 seems to affect more strongly the interaction of ParG with ParF, since a more dramatic increase in the K_D is recorded for this protein (five-fold increase). Interestingly, the only ParG mutant showing no association based on MST results is ParG-K5A, for which almost no change in fluorescence is recorded upon incubation with ParF. This strongly suggests that Lys-5 is key residue to enable the recognition of ParF by ParG.

The results obtained by MST were partially supported by SPR. In fact, binding experiments between ParF and ParG mutants showed contrasting results in some instances. ParG mutants K11A, K12A and M13A show very similar results with both techniques. Therefore, while changes in this cluster affects the interaction with ParF, the binding is not abolished, suggesting that the role of these residues could be to structure the tail, rather than making essential interaction with ParF.

On the other hand, the mutant ParG-K5A shows almost no binding based on the results obtained with both techniques, supporting the hypothesis that the tip of the tail might be important to grab and stick to ParF during plasmid segregation. As for ParG-L3A, a very similar behaviour was expected, since it belongs to the same cluster of amino acids as ParG-K5A. However, while SPR and cross-linking show little interaction with ParF, MST shows an interaction with ParF comparable to that of the wild type protein. Difference in

binding could be due to the specific set up of the two techniques. While SPR is usually more accurate and reliable, it has the major limitation of requiring one protein to be immobilised on a surface. Immobilisation of ParF might somehow affect the interaction with ParG-L3A, which would otherwise interact in a free-in-solution context. Conversely, it is also possible that ParG-L3A labelling for MST could affect the protein structure and make it more prone to bind to ParF than the unlabelled protein. Cross-linking proved that binding is still happening, although much less complex is formed, suggesting involvement of Leu-3 in the interaction with ParF.
When ParF was cross-linked to ParG-Δ9N deletion mutant, lacking the nine N-terminal amino acids, binding with ParF was observed. Interestingly, in the deletion mutants, the ParG N-terminus sequence is the following: Val-10, Lys-11, Lys-12. In this case, a change of register might occur and the two lysine residues might replace Lys-5 in the interaction with ParF, explaining the cross-linking results and further supporting the hypothesis that a lysine residue at the tip of the tail is essential for ParG interaction with ParF.

Finally, SPR shows that perturbation of the amino acids in the arginine finger motif causes a decrease in ParG binding to ParF of around three-to-four-fold. The results obtained from SRP do not completely agree with the ones obtained from MST. In particular, the mutant ParG-L21A shows binding to ParF comparable to wild type ParG in MST and four-fold weaker in SPR. Leucine and alanine are both amino acids with aliphatic and hydrophobic side chains. Leucine contains an isobutyl group, while alanine is smaller and contains a methyl group. Due to the similarities between the side chains, change of amino acid may not affect the binding when the proteins are in solution due to complete freedom of movement, as demonstrated through MST. Instead, when one of the interacting partners is immobilised, as in SPR, the leucine of the wild type protein may be essential to reach ParF, which is not free to move. The co-crystal structure of ParF and a ParG fragment (Zhang and Schumacher, 2017) shows that the arginine finger motif inserts at the interface between the ParF monomers. Crystal structure and binding analysis support the hypothesis that the arginine finger motif inserts in the ParF ATP-binding region to stimulate ParF ATPase activity. However, the fact that binding is not abolished in arginine-finger mutants, suggests that a different region of the tail is involved in holding ParG and ParF together.

Regardless of the discrepancies observed with the two biophysical techniques, overall MST and SPR experiments showed that the ParG N-terminal domain is composed of three distinct regions. The first is the tip of the tail, characterised by Leu-3 and Lys-5 which might work as sticky tentacles to grab ParF. The second is the arginine finger

motif, consisting of residues Asn-18, Arg-19 and Leu-21, which is positioned at the ParF monomer-monomer interface, close to the ATP binding pocket, and stimulates ParF weak ATPase activity. The third is the cluster of amino acids in the middle of the tail, composed of Lys-11, Lys-12 and Met-13, which might work as a structural domain or hinge, for the positioning of the remaining amino acids of the tail.

Chapter 4
Defining the role played by the Lys-11, Lys-12 and Met-13 amino acid cluster in ParG structure and function

4.1 Introduction

ParG's role in plasmid segregation is to work as an adaptor between the plasmid DNA and the motor protein ParF, as well as acting as a transcriptional repressor of the *parFG* operon (Wu *et al*., 2011; Zampini *et al.*, 2009; Barillà and Hayes, 2003). The protein has a folded C-terminal domain with a ribbon-helix-helix structure and an unstructured N-terminal tail (Golovanov *et al.*, 2003). ParG forms symmetrical dimers and oligomerises in presence of the DNA (Golovanov *et al.*, 2003; Barillà and Hayes, 2003; Carmelo *et al.*, 2005). Based on similarities with the Arc-repressor, a model for ParG binding to the DNA was proposed. According to this model, the protein interacts with the cognate DNA region as a dimer of dimers (tetramer) by inserting the anti-parallel β-sheet into the major groove of the DNA (Golovanov *et al.*, 2003). Biophysical and biochemical analysis showed that the folded domain is involved in DNA binding and dimerisation (Golovanov *et al.*, 2003; Barillà and Hayes, 2003; Carmelo *et al.*, 2005).

By contrast, the N-terminal domain is composed of thirty amino acids, plus a small linker composed of the amino acids 31-33 that connects the flexible region to the β-strand of the folded domain. ParG tail can be divided into three regions characterised by different levels of flexibility (Golovanov *et al.*, 2003). The first six to ten amino acids are the most flexible and include Leu-3 and Lys-5. Both amino acids were identified as essential for ParG function, since single amino acid change into alanine impairs plasmid segregation (Barge, 2015). In particular, mutation of Lys-5 seems to abolish ParG binding to ParF (Chapter 3, this work). On the other hand, amino acids in position 17 to 23 are the least flexible. This region includes the arginine finger motif which was reported to form a helix structure that inserts at the interface between ParF monomers (Zhang and Schumacher, 2017). The region 23-29 was instead hypothesised to form a β-strand, making transient interaction with the folded C-terminal domain. However, a single amino acid change into alanine, in this region, did not affect plasmid stability (Barge, 2015).

NMR studies on ParG in complex with part of the operator site (O_F) and with non-specific DNA showed that the N-terminal domain confers specificity to ParG interaction with the cognate DNA site. The HSQC spectrum of the protein changed when the cognate specific DNA is added (Carmelo *et al.,* 2005). The signal of the C-terminal domain residues disappeared, as bound to O_F, and the chemical shift of many of the amino acids composing the N-terminal region was affected. Interestingly, larger changes in chemical shift were recorded for the amino acids Lys-11 and Lys-12. On

the contrary, addition of non-specific DNA did not cause a change in the chemical shifts of the N-terminal residues (Carmelo *et al.,* 2005).
Furthermore, upon addition of DNaseI to the ParG-DNA complex, the C-terminal domain was seen to remain tightly bound to the DNA, while the N-terminus was seen to dissociate from it (Carmelo *et al.*, 2005). Carmelo *et. al* suggested that the binding of ParG N-terminal tail to the specific DNA is transient and cooperative with the C-terminal domain.

Although unstructured, the ParG N-terminal domain is essential for the protein's function. Complete (30 N-terminal amino acids) or partial (19 N-terminal amino acids) deletion of the N-teminus greatly affects ParG's ability to work as a transcriptional repressor, while deletion of the first nine amino acids did not affect ParG regulatory activity (Carmelo *et al.*, 2005).
ParG tails are also involved in the binding to ParF and in enhancing its weak ATPase activity. Furthermore, ParG was also seen to promote ParF polymerisation and to bundle ParF oligomers (Barillà *et al.*, 2007). Deletion of the thirty amino acids that composed the N-terminal tail were seen to abolish both catalytic and bundling activities (Barillà *et al.*, 2007). However, *in vivo* experiments showed that binding to ParF was partially maintained (Barillà *et al.*, 2007). This suggests that the C-terminal domain is also involved in ParG binding to the partner protein.

In the past, a protein's function was believed to be strictly connected to its structure. However, disordered regions are quite common in proteins from all domains of life and several proteins have been identified, containing unstructured regions that fold upon binding to their target (Dyson and Wright, 2002; Ward *et al.,* 2004). As an example, both the λ cI repressor protein and Antp homeodomain harbour flexible N-terminal regions that become ordered when bound to the DNA (Otting *et al.*, 1990; Clarke *et al.*, 1991; Spolar and Record, 1994).

Similarly to ParG, its analogue proteins type Ia KorB (RK2 plasmid) and type Ib ω (pSM19035 plasmid) also contain an unstructured N-terminal tail. KorB is composed of a central DNA-binding domain and a C-terminal dimerisation domain, in addition to a N-terminal domain, homologous to Spo0J. Two unstructured regions are also present, one between the DNA-binding domain and the dimerisation domain, and one at the N-terminus of the protein (residues 1-54) (Hyde *et al.*, 2017). KorB can interact with KorA for gene regulation and with IncC for plasmid segregation (Rosche *et al.*, 2000; Bingle *et al.*, 2008). IncC interacts with the central domain of KorB. In fact, deletion of the N-

terminal and the C-terminal domains still showed binding *in vivo* (Lukaszewicz *et al.*, 2002). KorA binds to KorB through the dimerisation domain as well as through the DNA-binding domain (Bingle *et al.*, 2008; Hyde *et al.*, 2017). The role played by the N-terminal region is unknown.
KorB is a rather dynamic protein, hence flexible and disordered regions are believed to allow the protein to acquire different conformations to allow for the variety of partners KorB binds to (DNA, itself, IncC and KorA) (Rajasekar *et al.,* 2010). Studies on KorB in complex with KorA and their cognate site showed that complex formation retains KorB flexibility and that the presence of KorA and DNA does not affect KorB overall structure (Hyde *et al.*, 2017).

The ω protein is a small protein sharing structure similarities with ParG. It has a RHH C-terminal domain and an unstructured and flexible N-terminal tail (residues 1-23) (Murayama *et al.*, 2001). The protein interacts with the ParF homologue δ through the N-terminal domain and deletions of 4, 9 and 17 N-terminal amino acids were seen to greatly affect the interaction with δ to a similar extent. Similarly to what was observed for ParF-ParG interaction (Barillà and Hayes, 2003), the presence of ATP improves the binding of ω to δ (Dmowski and Jagura-Burdzy, 2011).

In Chapter 3 the role of the ParG N-terminal tail in the interaction with ParF was dissected. It was shown that specific regions of the N-terminal domain are involved in the interaction with ParF, while others perform other tasks. The amino acids at the tip of the tail were identified as important for binding to ParF. The region around Arg-19 was characterised as the arginine finger motif by Barillà *et al* (Barillà *et al.*, 2007). The role played by the amino acids in the middle of the tail instead, e.g. Lys-11, Lys-12 and Met-13, remains elusive. These amino acids were identified as essential for plasmid partition, since change into alanine severely affected plasmid stability (Barge, 2015). Carmelo *et al.* argued that these amino acids could interact with *parH* and O_F and confer specificity to DNA binding. Single amino acids changes in the Lys-11, Lys-12 and Met-13 triad, affected the binding to ParF, with a three-fold decrease in binding affinity (Chapter 3). However, we speculated that decrease in the interaction between ParF and ParG was caused by a change in the tail's structure, rather than deletion of important non-covalent bonds. The aim of this chapter is to study how the Lys-11, Lys-12 and Met-13 cluster contributes to ParG function. In particular, we put forward the hypothesis that the central triad could structure the ParG tail upon binding to the DNA. Structuring of the tail would then facilitate the binding to ParF.

To prove this hypothesis, double and triple mutants of ParG (ParG-K11A-K12A and ParG-K11A-K12A-M13A) were constructed. Interaction between these mutants and ParF was tested, as well as binding to the centromeric site *parH*. Circular dichroism was then used to investigate whether the secondary structure of the ParG and ParG double and triple mutants changed, when in complex with the cognate DNA site. The hypothesis that *parH* binding by ParG enhances ParG-ParF interaction was then tested by using chemical cross-linking. Chemical cross-linking coupled to mass-spectrometry shedded light into which regions of ParF and ParG interact together, as well as whether the protein-protein interaction is aided by the centromeric site p*arH*.

4.2 Results

4.2.1 Cloning *parG-K11A-K12A* and *parG-K11A-K12A-M13A* in the pFH547 vector

ParG double and triple mutants, ParG-K11A-K12A and ParG-K11A-K12A-M13A, were constructed to investigate the role played by the triad of amino acids, present in the middle of the N-terminal tail, in plasmid segregation. Although single amino acid changes of Lys-11, Lys-12 and Met-13 into alanine, were shown to reduce plasmid stability (Barge, 2015), testing the effect that double and triple mutation had *in vivo* on plasmid segregation was important to understand to which extent multiple mutation affected the system. Therefore, *parG* mutants were constructed in the vector pFH547 (Figure 4.1). The plasmid is a derivate of pALA136, which harbours both the medium copy number ColE1 and the single copy number P1 origin of replication, as well as the chloramphenicol resistance gene (Martin *et al.*, 1987). In wild type cells, the plasmid replicates through the ColE1 replicon. However, in *polA* strains (such as BR825 *E. coli* cells), inactivation of the DNA polymerase I causes ColE ori to be non-functional (section 1.2.1), leading the plasmid to replicate at low copy number through the P1 replicon. The pFH547 plasmid harbours the full *parFGH* partition cassette, which confers stability to the plasmid, that otherwise would be lost in the absence of chloramphenicol (Hayes, 2000). The plasmid pFH450 is also a pALA136 derivate but it does not contain any partition gene, making it unstable under selective pressure. This plasmid was, therefore, used as a negative control for the experiment.

The *parG-K11A-K12A* allele was constructed using site-directed mutagenesis by overlapping extension PCR as described in 2.4.1. The plasmid pMBK11A was used as a template, as it already harboured the mutation causing the K11A change. The second mutation was introduced by designing internal primers leading to change of the Lys-12 codon into an alanine codon. The external primers used annealed in the middle of the *parF* sequence and at the end of the *parG* sequence. The product of PCR1 and PCR2

were respectively 317 bp and 214 bp long (Figure 4.2 A). The PCR products were then purified before being annealed together for PCR3. PCR3 was performed by using the external primers of PCR1 and PCR2 (Figure 2.1). The product obtained was around 530 bp long (Figure 4.2 B). Both PCR product and pFH547 vector were then digested with ClaI and HpaI. The enzymes cut unique restriction sites within *parF* and *parG* sequences (Figure 4.1, bottom). The two restriction sites are 260 bp apart and include the mutated region of *parG*. In this way the end of *parF* and part of wild type *parG* can be replaced by the mutated sequence. After restriction digestion, the insert and vector were purified and ligation was then carried out as described in 2.3.8. Clones were screened by colony PCR and restriction digestion (Figure 4.2 C). Putative positiveclones were then sent to be sequenced. Once the sequence of the *parG* double mutant was confirmed, the same procedure was repeated to construct the triple mutant *parG-K11A-K12A-M13A* in the pFH547 vector. The template used for PCR1 and PCR2 was pFH547-parG-K11AK12A (pCP-K11AK12A).

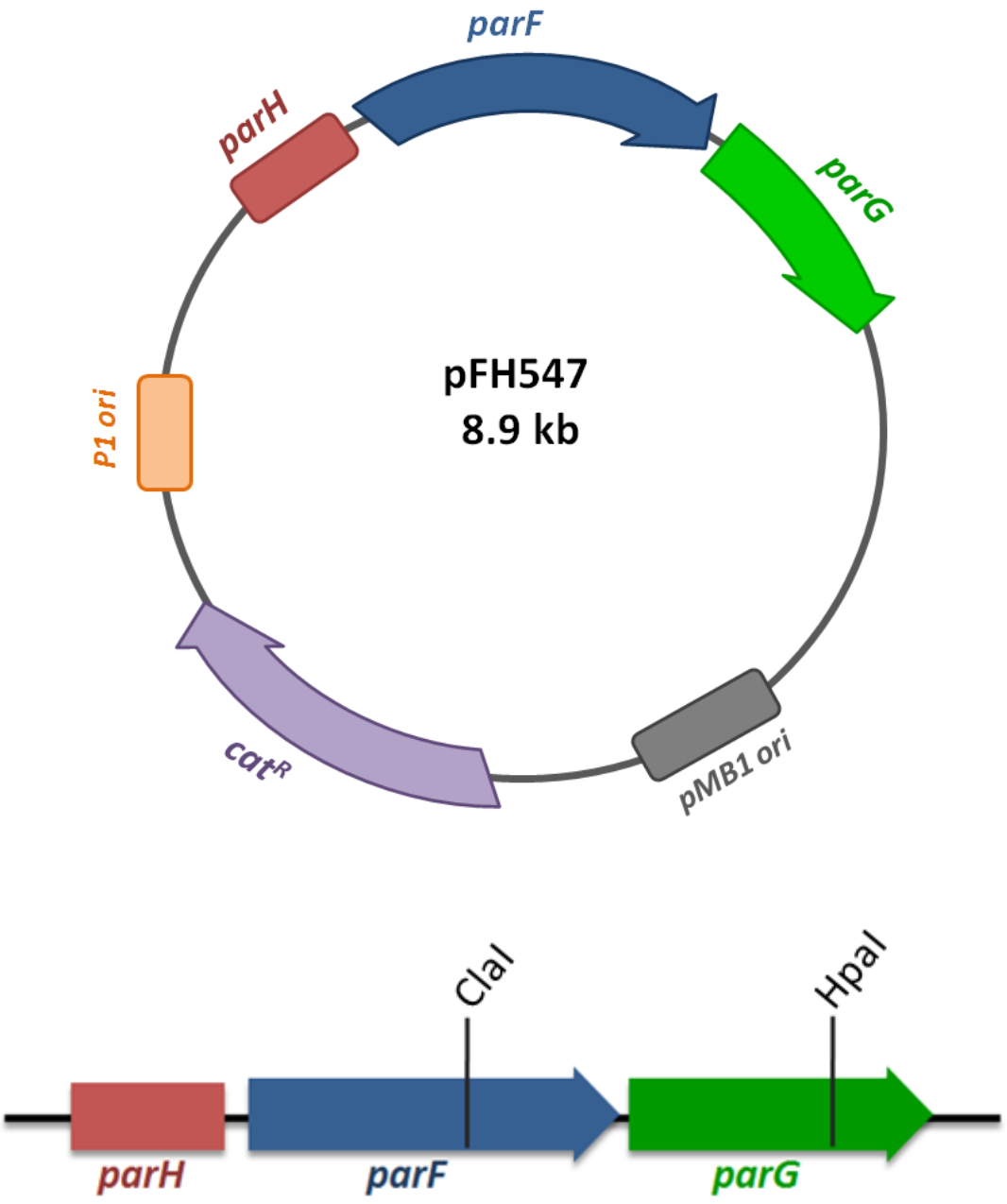

Figure 4.1 Map of pFH547 plasmid

The plasmid pFH547 is a derivate of pALA136. The plasmid harbours the full *parFGH* partition cassette, P1 *ori*, pMB1 (ColE1 *ori*) and the chloramphenicol resistance gene (cat^R). ClaI and HpaI restriction sites cut unique sites within *parF* and *parG* respectively.

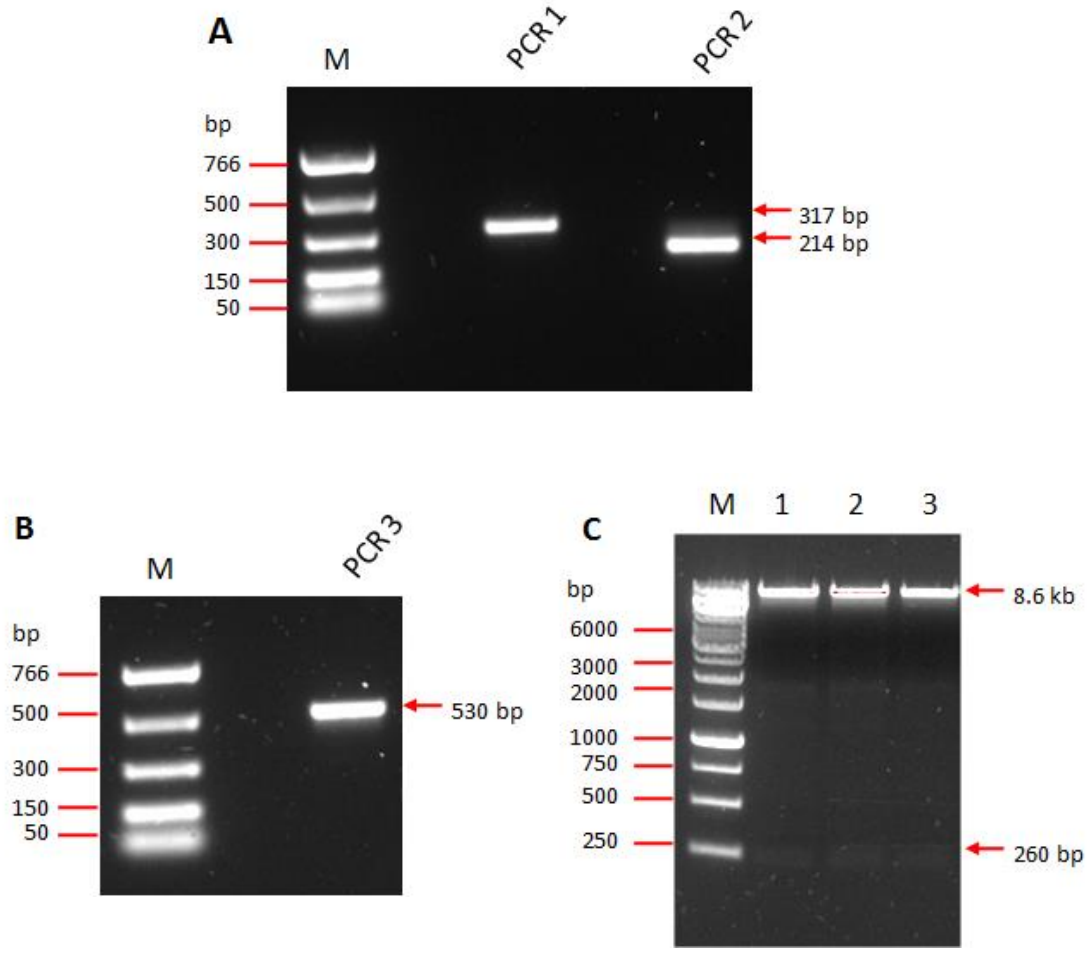

Figure 4.2 Agarose gels showing the construction of *parG-K11A-K12A* in the pFH547 vector

A) PCR1 (317 bp) and PCR2 (214 bp) products harbouring the changes leading to K11A-K12A in the *parG* gene. M represents the PCR marker (NEB). B) PCR3 amplification after annealing of PCR1 and PCR2 (530 bp). M represents the PCR marker (NEB). C) Cloning screening by restriction digestion with ClaI and HpaI enzymes. The three plasmids selected show, after digestion, a 8.5 kb band (vector) and a 260 bp band (insert harbouring the end of *parF* and part of mutant *parG*). M represents GeneRuler 1kb MW marker (Thermo Fisher Scientific).

4.2.2 Investigating the stability of the low copy number plasmid harbouring the cassette with *parG* double and triple mutants

The stability of a low copy number plasmid harbouring the full partition cassette, where wild type *parG* was substituted by *parG-K11A-K12A* and *parG-K11A-K12A-M13A*, was measured through partition assays (section 2.12). The pFH547 plasmids, containing the *parG* double and triple mutants, were transformed into BR825 *E. coli* cells and grown overnight on chloramphenicol-containing plates to select cells harbouring the plasmid. The same strain of *E. coli* cells was also transformed with the pFH547 plasmid, containing the wild type partition cassette, and the empty vector pFH450. These were used as positive and negative controls respectively for the experiment.

After selection on chloramphenicol, cells were grown for 25 generations on LB plates with no selection, so that only plasmids containing a functional partition cassette would be retained. Chloramphenicol was then added to the growth medium for the final step of the partition assay, to measure the percentage of plasmid retention. Plasmid harbouring the wild type partition cassette showed on average 65% plasmid retention. The empty vector was, instead, lost and plasmid retention was on average only 8%. ParG double and triple mutant were greatly affected in plasmid segregation and plasmid retention was on average 2% for the plasmid harbouring *parG* double mutant and 16% for the plasmid containing *parG* triple mutant. Interestingly, ParG-K11A-K12A was less efficient in plasmid retention than ParG-K11A-K12A-M13A, although both proteins were unable to support plasmid segregation efficiently.

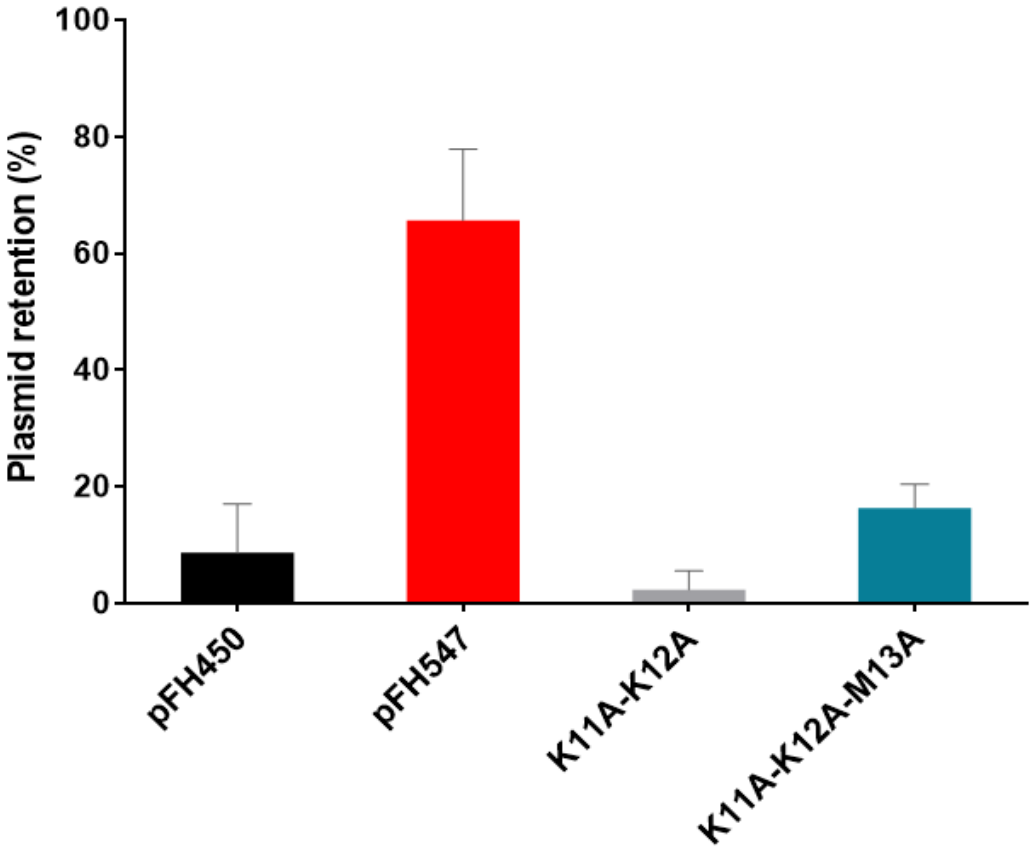

Figure 4.3 Percentage of plasmid retention in *E. coli* cells expressing *parG* double and triple mutant

Retention of a low copy number plasmid, where the *parG* gene in the partion cassette was replaced by *parG-K11A-K12A* and *parG-K11A-K12A-M13A*, was measured by partition assays. The plasmid pFH450 does not contain any partition gene and represents the negative control of the experiment, while the plasmid pFH547 harbours the wild type partition cassette and represents the positive control. Each experiment was repeated in triplicate. The average of plasmid retention is shown, with the error bars plotted by calculating the standard error of the mean.

4.2.3 Cloning *parG-K11A-K12A* and *parG-K11A-K12A-M13A* into the pET-22b(+) vector

Having observed an *in vivo* effect, the mutant alleles were cloned into the pET-22b(+) vector for over-expression and protein over-production to perform further *in vitro* investigations.

The alleles *parG-K11A-K12A* and *parG-K11A-K12A-M13A* were amplified from the corresponding pFH547 construct and cloned in the pET-22b(+) vector in frame with the sequence of a C-terminal His-tag (Figure 4.4).

Further *in vitro* characterisation of ParG-K11A-K12A and ParG-K11A-K12A-M13A required purification of the proteins. Over-expression from the pET vector constructs and protein overproduction was carried out as described in 2.5.2. Purification was performed by a single step of nickel affinity chromatography, as for the other ParG N-terminal mutants. Both proteins showed a level of purity comparable to the wild type protein and to the single mutants. Examples of ParG mutants purifications are shown in Chapter 3 and are not shown in this chapter. Protein concentration was measured after chromatography and ranged between 0.5 and 2 mg/ml.

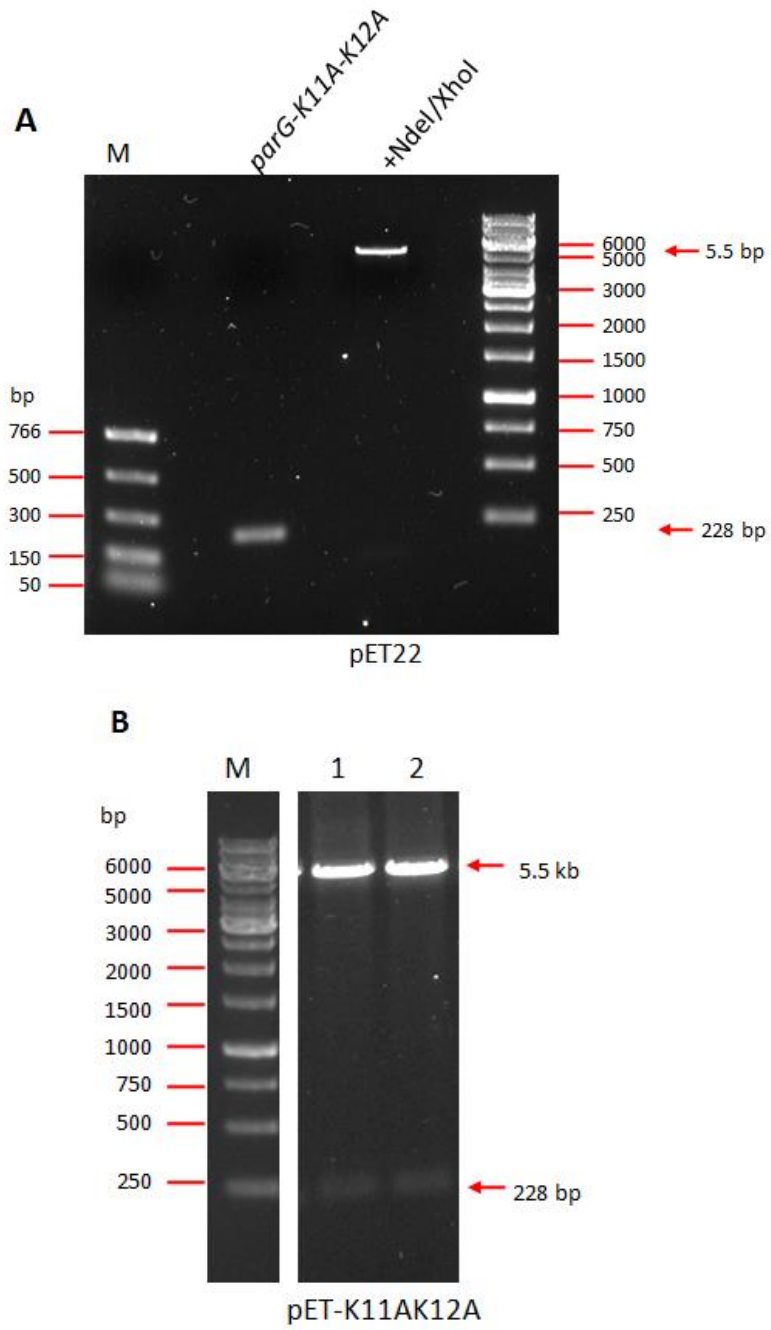

Figure 4.4 Agarose gels showning the construction *of parG-K11A-K12A* in the pET-22b(+) vector

A) Agarose gel showing PCR amplification of *parG-K11A-K12A* (228 bp) and the pET22 vector digested with NdeI and XhoI endonucleases. The first lane represents the PCR marker (NEB), while the last lane represents the GeneRuler 1kb MW marker (Thermo Fisher Scientific).

B) Cloning screening by restriction digestion with NdeI and XhoI. The two plasmids selected show a 5.5 kb band (vector) and a 228 bp band (*parG*) after digestion. M represents GeneRuler 1kb MW marker (Thermo Fisher Scientific).

4.2.4 Double and triple mutations affect ParG ability to bind to ParF

ParG-K11A-K12A and ParG-K11A-K12A-M13A ability to bind to ParF was tested using MST and SPR. MST and SPR experiments were carried out as detailed in 2.6.1 and 2.6.2. Both techniques showed that binding was severely affected by the residue changes. According to MST, no binding could be detected between the double and triple ParG mutants and ParF (Figure 4.5). For these experiments the amplitude was very small compared to the noise of the baseline. Therefore, the fluctuation of the fluorescence was considered random, rather than caused by binding, and no fitting, following a saturation binding equation, was possible. On the other hand, SPR showed that ParG mutants could bind to ParF very weakly, as a change in RU could be detected when a high concentration of ParG was flowed over the ParF-coated chip (Figure 4.6). Nevertheless, binding was too low to be measured accurately. It is interesting to notice that ParG-K11A-K12A-M13A caused a higher response than ParG-K11A-K12A, when the same concentration of protein was flowed over ParF (Figure 4.7). This could partially explain why partition assay showed that plasmid retention was higher for ParG triple mutant, compared to the double mutant.

As an additional control of the binding experiment, a truncated version of ParG, ParGΔ30N, where the 30 N-terminal amino acids of the tail were deleted, was used for the interaction with ParF. The truncated ParG showed binding to ParF *in vivo*, in a bacterial two-hybrid system (Barillà *et al.*, 2007). However, both MST and SPR showed no binding between the protein and ParF (Figure 4.5 and 4.6). Therefore, it is likely that although weak, ParG double and triple mutant, as well as ParGΔ30, could still interact with ParF. Both MST and SPR present limitations. It is possible that immobilisation of ParF on the SPR surface prevents the ParG C-terminal domain from reaching its binding site. Therefore, when the ParG tail is deleted, no binding can be observed. As far as MST is concerned, lysine labelling could have involved amino acids essential for ParG binding to ParF, thus causing abrogation of the interaction. All things considered, multiple amino acids change in the Lys-11, Lys-12 and Met-13 triad severely affected the interaction of ParG with ParF. Nevertheless, weak binding could still be observed, suggesting that these amino acids are not strictly essential for the interaction and supporting the hypothesis that their role is to structure the tail, rather than binding to ParF.

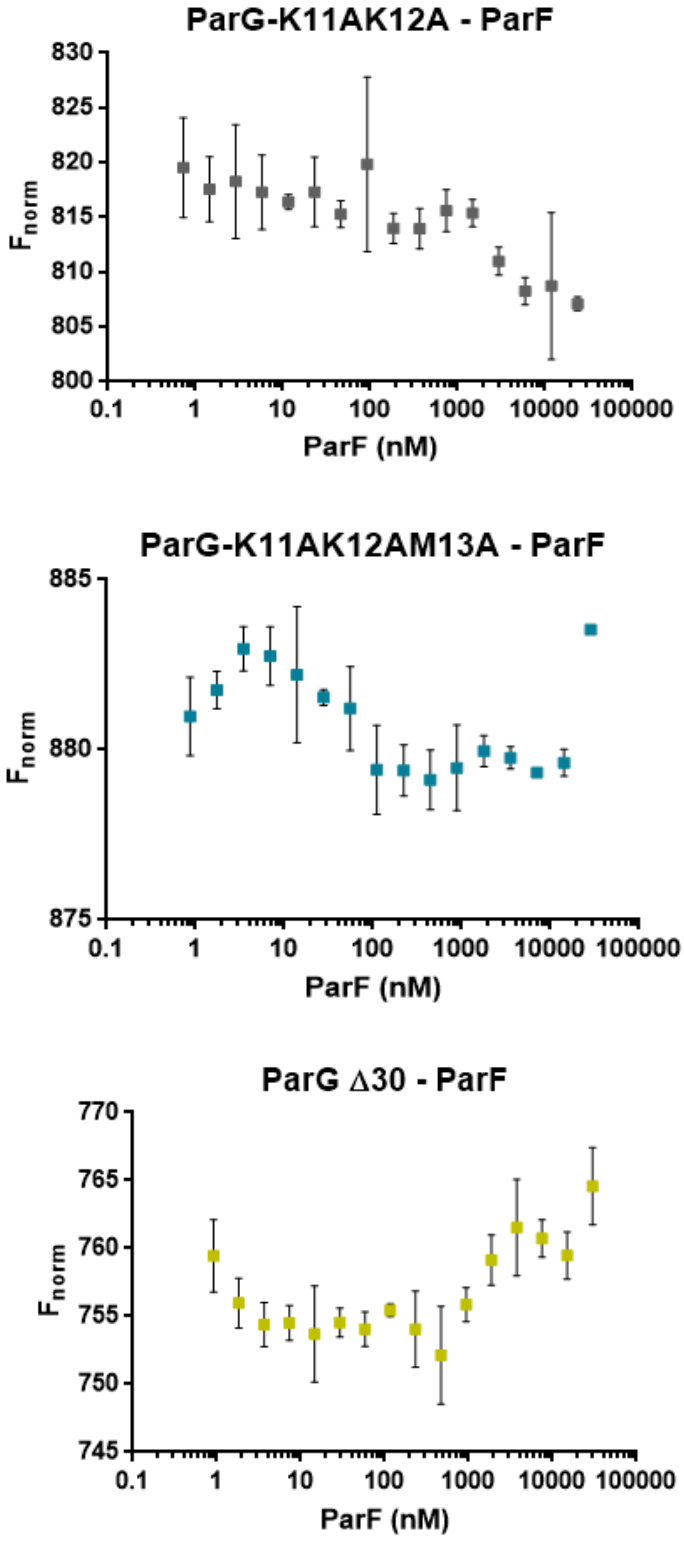

Figure 4.5 - MST binding curve of the interaction between ParF and ParG-K11A-K12A/K11A-K12A-M13A and ParGΔ30

Sixteen ParF concentrations (from 0.86 nM to 28.5 µM) were titrated over 200 nM ParG. The experiment was performed in triplicate. The values were averaged and error bar plotted calculating the standard error of the mean for each point. The curve could not be fitted with a one-site specific binding equation.

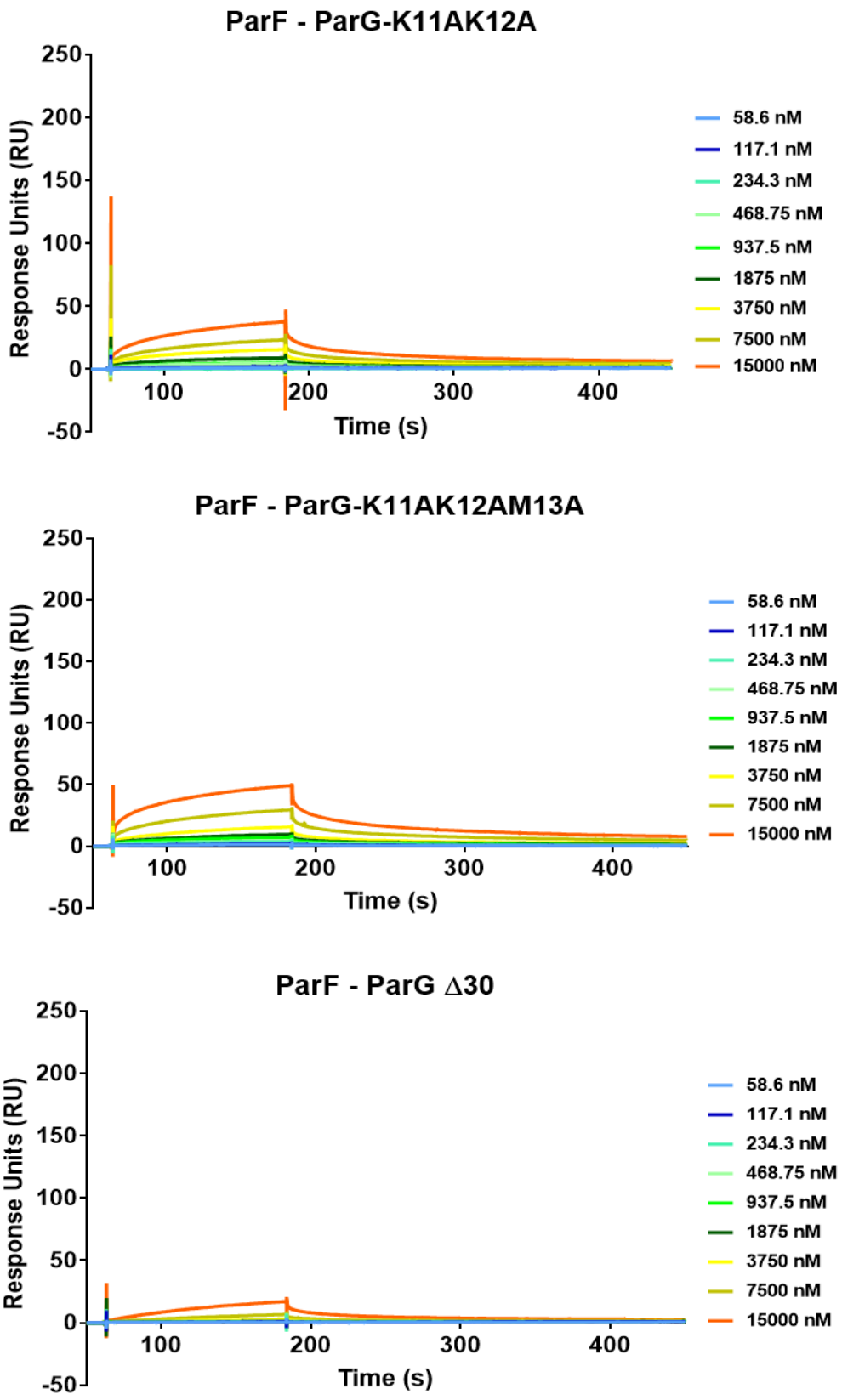

Figure 4.6 - SPR binding curves of the interaction ParF – ParG-K11A-K12A, ParF - ParG-K11A-K12A-M13A and ParF – ParGΔ30

Nine ParG concentration were flowed over the ParF-coated CM5 chip and the response was recorded as Response Units (RU) over time.

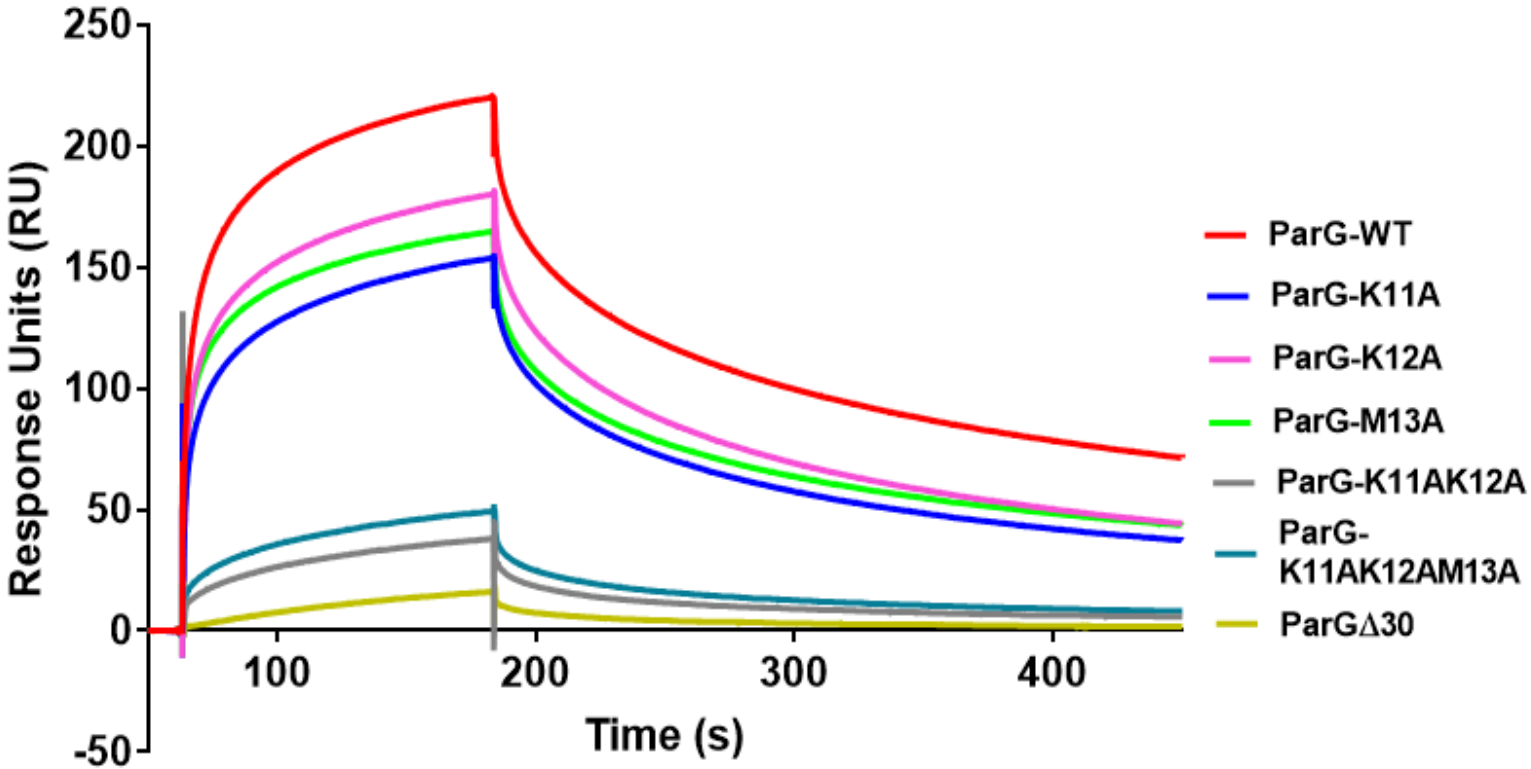

Figure 4.7 - Alignment of the SPR binding curves obtained for the interaction between ParF and ParG/ParG N-terminal mutants

After recording the ParF baseline signal, ParG/ParG mutants (15 μM) were flowed over the ParF-coated CM5 chip for two minutes at a 30 μl/min flow rate. After two minutes, buffer was injected and ParG slowly dissociated. The chip surface was regenerated with 2M NaCl.

4.2.5 Double and triple mutations in the tail affect ParG's ability to enhance the formation of ParF higher-order structures

An important characteristic of ParF is that the protein can polymerise in the presence of ATP. ParG can further enhance the formation of higher-order structures by cross-linking and bundling ParF oligomers through its N-terminal domain (Barillà *et al*., 2005, 2007). Barge demonstrated the ParG N-terminal mutants indentified through alanine-scanning mutagenesis could enhance ParF oligomerisation as efficiently as wild type ParG (Barge, 2015). This suggested that cross-linking of ParF polymers was performed synergistically by the entire N-terminal domain. However, if the ParG N-terminal tail structure was disrupted by the residue changes in the middle amino acid cluster, the protein's ability to bundle ParF structures would be affected. Dynamic light scattering (DLS) was used to investigate in real-time the formation of ParF structures by wild type ParG and ParG mutants. This was compared to the result obtained for ParGΔ30, that was shown to fail to enhance ParF oligomerisation (Barillà *et al.*, 2007). DLS allows the assessment of particle size by looking at the intensity of the light scattered by the particles in solution. For the DLS experiments, ParF was first centrifuged to remove any polymer or large particles present in solution. A baseline of ParF (2.16 μM) was recorded before addition of ATP (0.5 mM), showing a scattering intensity of ~100 kct/s. Upon addition of the nucleotide, ParF started polymerising and the light scattering intensity recorded increased to ~1500 kct/s. When wild type ParG (2.16 μM) was added, particles size increased even further and the light scattering intensity reached ~ 8000 kct/s (Figure 4.8). Instead, addition of ParG double and triple mutants, as well as ParGΔ30 could not promote polymer formation and bundling as efficiently as wild type ParG. ParG-K11A-K12A could enhance the formation of ParF higher-order structure and the intensity of light scattering recorded was around 6000 kct/s (Figure 4.8). However, the increase in intensity recorded was slower compared to the wild type protein, suggesting that change in the structure of the N-terminal region was affecting the property of ParG to mediate ParF assembly into higher-order structures.

Addition of ParG-K11A-K12A-M13A to the ParF-ATP complex did not completely fail in promoting ParF polymerisation and bundling, since intensity increased up to ~ 3000 kct/s (Figure 4.8). However, ParGΔ30 showed exactly the same pattern as ParG triple mutant, demonstrating that the structure of the tail was severely affected in the triple mutant and that the slight increase in particles size recorded was probably due to the presence of ParG C-terminal domain. The experiment was repeated three times and results were confirmed.

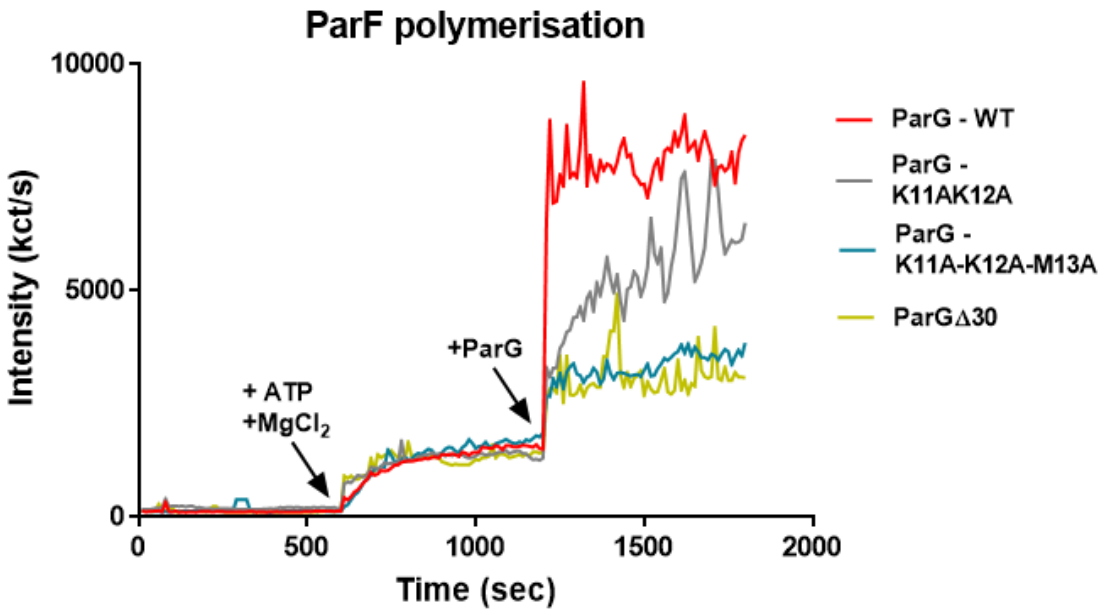

Figure 4.8 – DLS showing the effect of ParG N-terminal mutants on ParF polymerisation

ParF polymerisation was measured by DLS. ParF baseline was recorded for 10 minutes. ATP (0.5 mM) and $MgCl_2$ (5 mM) were then added to the cuvette and measurement were recorded for 10 more minutes. At that point, ParG or ParG mutants were added and increase in light scattering was measured. Light scattering intensity is expressed in kct/s and plotted against time. Each experiment was repeated three times. Data were plotted using GraphPad Prism 7.04.

4.2.6 Double and triple mutation slightly reduce ParG affinity for *parH*

NMR studies showed that Lys-11, Lys-12 and Met-13 could make transient interactions with the site-specific DNA (Carmelo *et al.*, 2005). Barge showed that the ParG mutant harbouring K11A, K12A and M13A single amino acid change could bind DNA and work as transcriptional repressors proficiently. However, K11A showed slightly reduced affinity for the O_F operator site, while K12A and M13A showed slightly lower transcription repressor activity compared to the wild type protein (Barge, 2015). Therefore, it was interesting to understand whether double or triple amino acid change in this region would affect the binding to the centromeric site *parH*. Interaction with the DNA was investigated using electrophoretic mobility shift assays (EMSA). Primers were designed to amplify the full *parH* sequence (114 bp) (Figure 4.9). The 5' end of the forward primer was biotinylated for DNA detection using LightShift Chemiluminescent EMSA kit (Thermo Fisher Scientific). After PCR amplification (section 2.3.2), the product was run on a 1.5% agarose gel and extracted as described in Section 2.3.7, to obtain pure DNA product. Nine increasing concentrations of protein (50, 75, 100, 200, 300, 400, 500, 750, 1000 nM) were incubated with 1.35 nM of *parH* (0.1 ng/µl), as detailed in Section 2.8.2. Poly-dI-dC (1 µg) was included in the reactions to reduce non-specific binding of ParG to the DNA. The complex was then run on 6% acrylamide gel to separate the protein-bound from the unbound DNA. The gel was transferred onto a nylon membrane and band detection was performed as described in Section 2.8.4. Each experiment was repeated three times and a rough estimation of the strength of the interaction was obtained from quantification of the bands identified through EMSA. Band quantification was performed by using Image lab 4.0.1 software. The values from the quantification of the "unbound DNA" bands were plotted against the protein concentrations. Fitting of the points was done using the specific binding with Hill slope equation (GraphPad Prism 7.04), stated below:

$$Y = \frac{B_{max}X^h}{({K_D}^h + X^h)}$$

where Y is the percentage of binding, X is the protein concentration, B_{max} is the maximum binding, K_D is the equilibrium dissociation constant and h is the Hill coefficient.

The specific DNA site, *parH*, harbours twelve ParG binding sites, each of which is bound by one ParG dimer. The Hill coefficient describes the cooperativity of the binding. If h is greater than one, the binding is cooperative, hence initial association of protein to the DNA enhances the binding of more molecules.

Different binding models were compared using GraphPad Prism 7.04 and the Hill slope equation showed the best fitting of the data.
Fitting of the values obtained from EMSA allowed an estimation of the K_D of the interaction between ParG/ParG mutants and *parH*. Wild type ParG showed a K_D= 144.6 ± 21.4 nM, with a Hill coefficient h= 1.305 ± 0.240 (Figure 4.10). ParG-K11A-K12A showed a K_D= 335.7 ± 21.66 nM, with a Hill coefficient h= 4.894 ± 1.451 (Figure 4.11). Finally, ParG-K11A-K12A-M13A showed a K_D= 248.6 ± 24.8 nM, with a Hill coefficient h= 2.775 ± 0.683 (Figure 4.12).

From K_D and Hill coefficient estimation, it appeared that the interaction between wild type ParG and the DNA is cooperative and quite strong. Double and triple mutation slightly reduced the strength of the interaction between the protein and the DNA, while cooperativity seemed to increase. Interaction with the DNA was, as expected, only mildly affected, since the protein makes essential interaction with the DNA through the RHH fold in the C-terminal domain, that was unchanged. Furthermore, from this quantification it appears that triple mutation affects ParG interaction with the DNA less than double mutation.
Since EMSA is a semi-quantitative technique, further quantification was attempted by using MST. In the MST experiments the DNA was labelled with Cy5 Cyanine dye at the 5' end. However, it appeared that the DNA labelling was greatly affecting the interaction with the DNA and it was impossible to obtain reliable and reproducible results. Due to time constraint, this was not explored further.

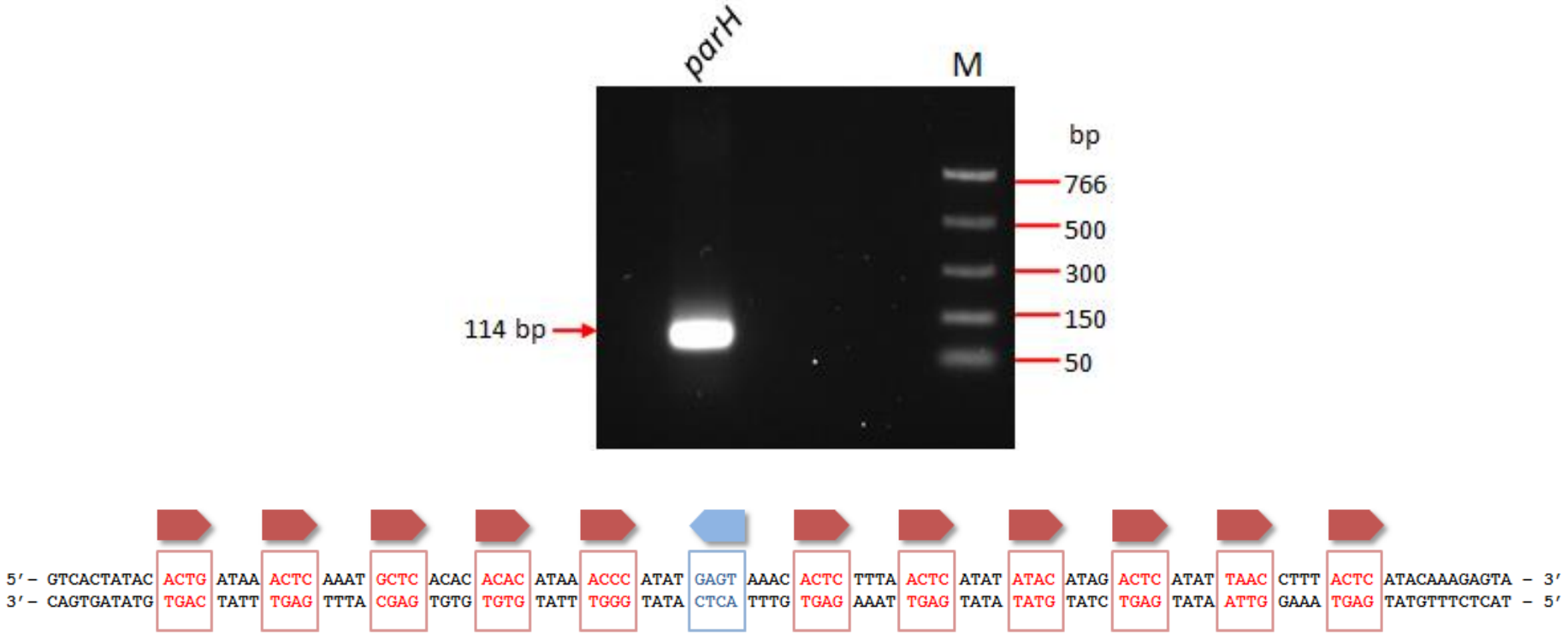

Figure 4.9 – Agarose gel showing PCR amplification of the centromeric site *parH* and *parH* sequence

The region upstream ParF, *parH*, was amplified from the plasmid pFH547. PCR product was run on 1.5% agarose gel for extraction and purification, before EMSA experiments. M represents the PCR marker (NEB).

parH sequence is shown. The repeated sequences are represented by the arrows. The direct repeats are shown in red and the inverted repeat in blue.

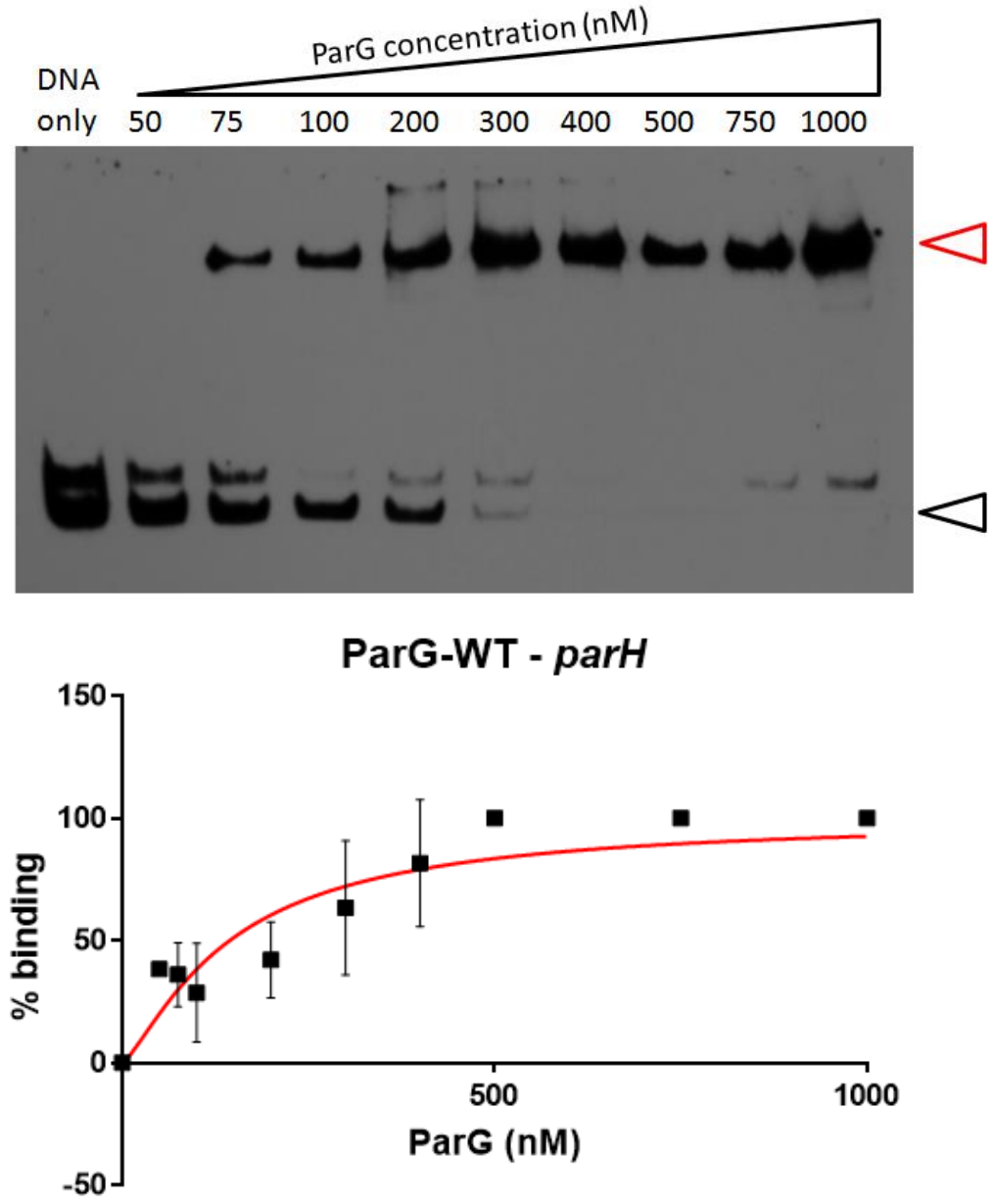

Figure 4.10 - Quantification of the interaction between wild type ParG and *parH* by EMSA
The top image shows an example of EMSA experiment, which was repeated in triplicate. Nine increasing ParG concentration were incubated with *parH* (lanes 2 to 10). The first lane represents the DNA alone, used as a control. The black arrow points at the unbound DNA, while the red arrow points at the ParG-bound DNA. The second band, visible just above the DNA band, is probably due to the formation of a secondary structure of the biotinylated DNA, which does not react with the protein. Although the same DNA was used in all the reactions, the band is visible in some lanes but not in others. The bottom image represents the quantification of the interaction. The values obtained from EMSA were averaged and error bar plotted calculating the standard error of the mean for each point. The curve was fitted with a specific binding with Hill slope equation.

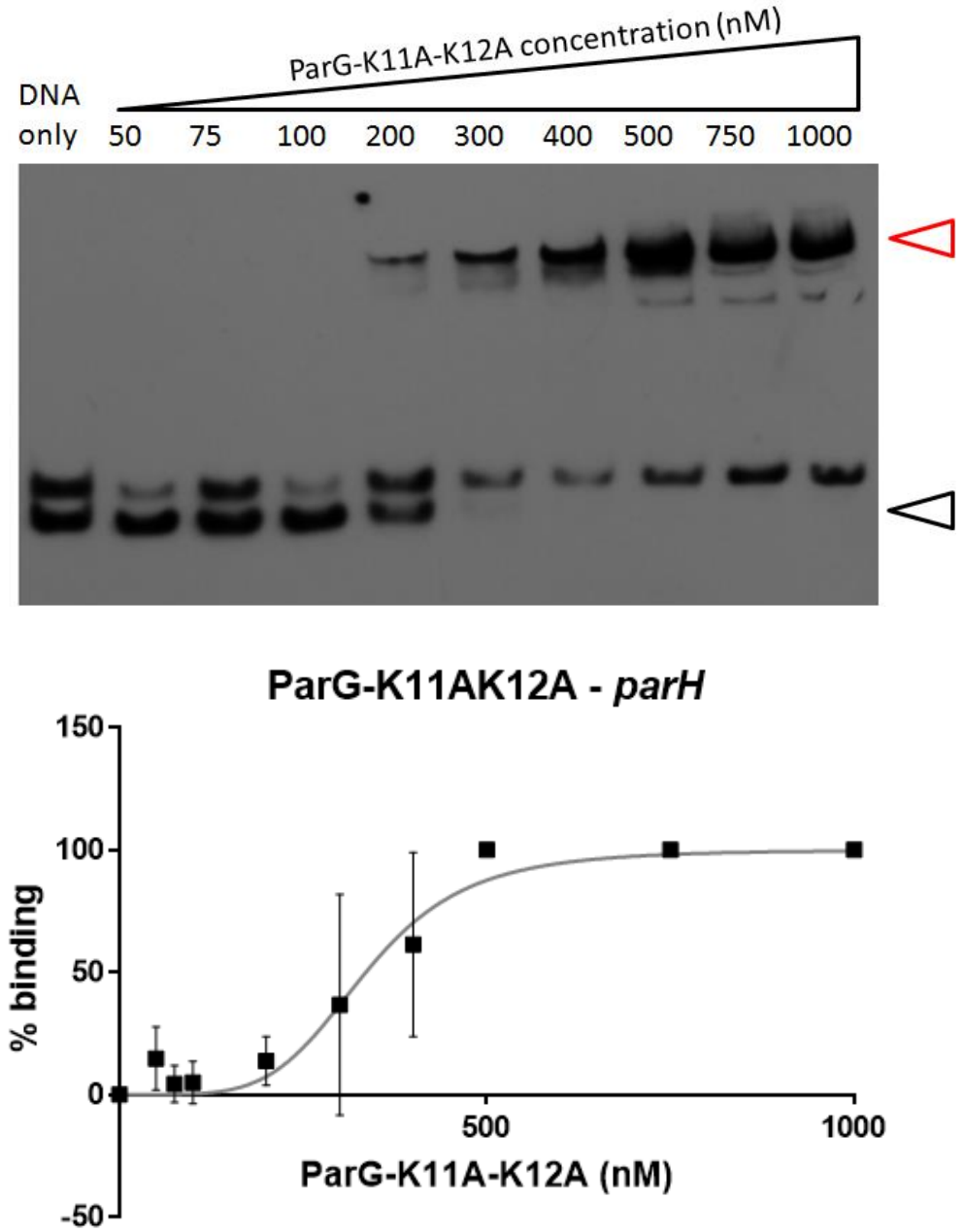

Figure 4.11 - Quantification of the interaction between ParG-K11A-K12A and *parH* by EMSA
The top image shows an example of EMSA experiment, which was repeated in triplicate. Nine increasing ParG-K11A-K12A concentration were incubated with *parH* (lanes 2 to 10). The first lane represents the DNA alone, used as a control. The black arrow points at the unbound DNA, while the red arrow points at the protein-bound DNA. The second band, visible just above the DNA band, is probably due to the formation of a secondary structure of the biotinylated DNA, which does not react with the protein. The bottom image represents the quantification of the interaction. The values obtained from EMSA were averaged and error bar plotted calculating the standard error of the mean for each point. The curve was fitted with a specific binding with Hill slope equation.

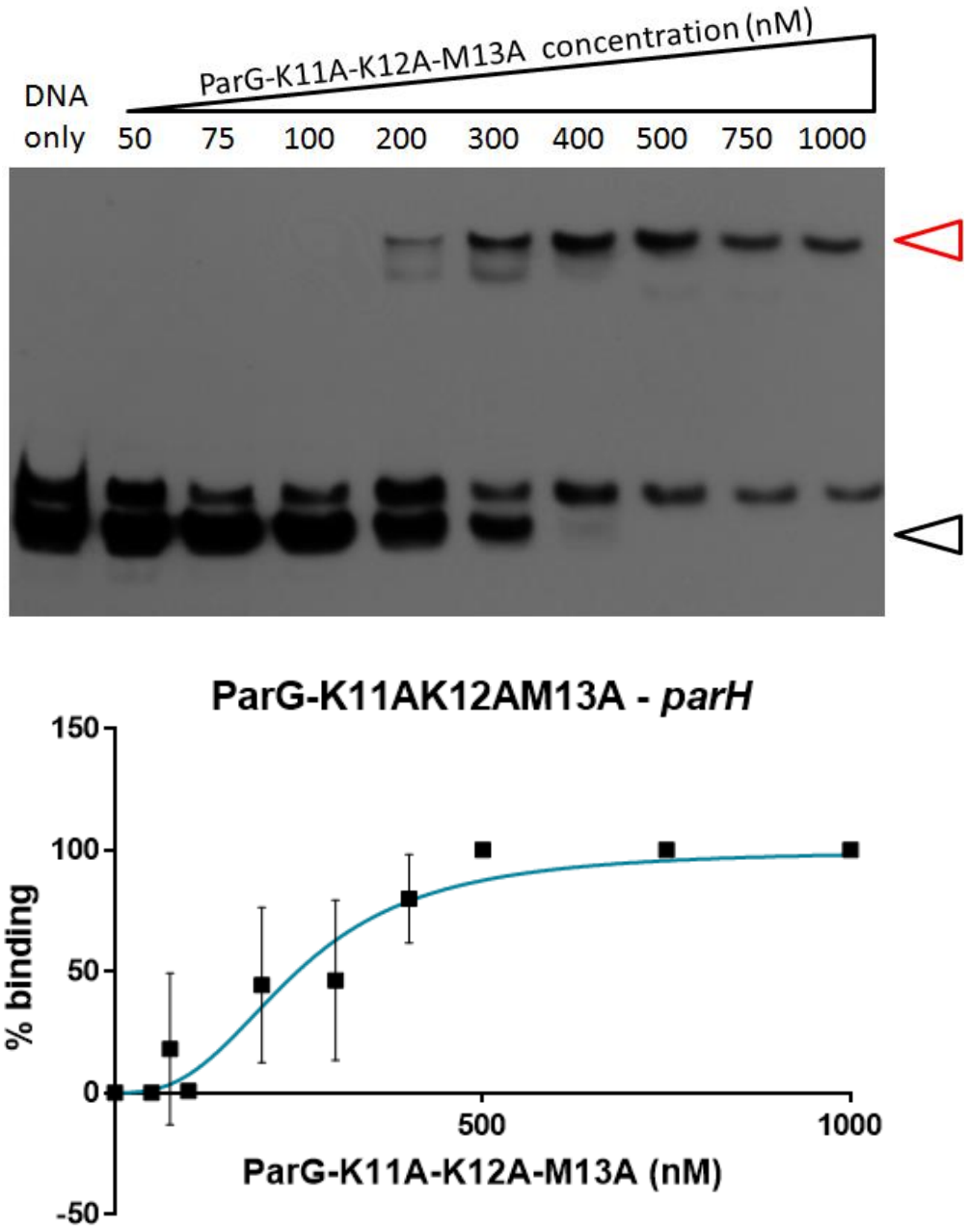

Figure 4.12 - Quantification of the interaction between ParG-K11A-K12A-M13A and *parH* by EMSA

The top image shows an example of EMSA experiment, which was repeated in triplicate. Nine increasing ParG-K11A-K12A-M13A concentration were incubated with *parH* (lanes 2 to 10). The first lane represents the DNA alone, used as a control. The black arrow points at the unbound DNA, while the red arrow points at the protein-bound DNA. The second band, visible just above the DNA band, is probably due to the formation of a secondary structure of the biotinylated DNA, which does not react with the protein. The bottom image represents the quantification of the interaction. The values obtained from EMSA were averaged and error bar plotted calculating the standard error of the mean for each point. The curve was fitted with a specific binding with Hill slope equation.

4.2.7 ParG-WT and ParG mutants structure changes to different extent upon interaction with *parH*

Circular Dichroism (CD) spectroscopy was used to study the conformational change in ParG secondary structure, caused by the interaction with *parH*. CD is the difference between left-handed and right-handed circularly polarised light, absorbed by an optically active sample ($\Delta A = A_L - A_R$). At certain wavelengths, chiral molecules, such as proteins, absorb to different extent right and left-handed circularly polarised light. Therefore, each protein secondary structure (α-helix, β-strand and random coil) has a signature spectrum, which can be used to predict the general structure of the protein (Greenfield, 2007). When the conditions change, e.g. change in pH, temperature or the interaction with another molecule, the protein can assume a different conformation that can be monitored through CD.

The original hypothesis was that the triad of amino acids, Lys-11, Lys-12 and Met-13 could structure the tail upon interaction with the ParG cognate DNA site. Binding of ParG to its cognate site *parH* should, therefore, cause a change in the conformation of the protein. DNA is also a chiral molecule and has a specific CD spectrum. If the interaction between protein and DNA does not trigger any change in ParG secondary structure, the spectrum recorded would simply be the sum of the two spectra. However, if ParG binding to the DNA causes a change in the protein conformation, the spectrum of the protein would change when in complex with the DNA. The aim of this experiment was to understand whether double and triple amino acid change in ParG would prevent the tail from structuring upon interaction with *parH*. ParG and ParG mutants spectra were recorded upon addition of increasing concentrations of a region of *parH*. The DNA was produced by annealing complementary primers, containing the first 52 bp of the *parH* site (section 2.10). This region contains 7 repeated sequences, each of which interacts with one ParG dimer. Since the concentration of ParG used for the experiment was 0.2 mg/ml (10.3 μM), the maximum concentration of DNA used was 1.4 μM (1/7 of ParG concentration). A volume of 250 μl of protein was added to a quartz cuvette with 0.1 cm path length and CD spectrum recorded using Jasco J810 circular dichromator. Increasing concentrations of DNA (0.05, 0.1, 0.2, 0.4, 0.6, 0.8, 1, 1.2 and 1.4 μM) from a 50 μM stock solution were added to the protein and the reaction was incubated five minutes at room temperature before recording the CD spectrum. The same procedure was repeated for the two ParG mutants. A “blank” CD spectrum of the cuvette containing buffer was also recorded to subtract the background from the protein measurements. CD spectra were adjusted taking into account the concentration

change caused by DNA addition. Values were then converted from m° (millidegrees) into residual molar ellipticity ([θ]) as described in 2.10.

Figure 4.13 shows the spectra of wild type ParG, ParG-K11A-K12A, ParG-K11A-K12A-M13A and *parH*. When the three proteins are not in complex with the DNA, the spectra are almost overlapping. The secondary structure of ParG triple mutant slightly differs from the other two ParG proteins, suggesting that mutation of Met-13 into alanine affects ParG secondary structure. For analysis of protein secondary structure the concentration is very important, because small changes affect the spectrum recorded. However, the dilution and the measurement were repeated and the result was confirmed, excluding the possibility of an artefact. Secondary structure analysis using the CONTIN algorithm (Sreerama and Woody, 2000) from Dichroweb (Birkbeck, University of London), with the SP175 database of soluble proteins (Lees *et al.*, 2006), showed that the secondary structure content changed a little for the three proteins. Since the algorithm simply compares the CD spectra of known protein structures to the one of interest, the percentage of secondary structures obtained from the software may not be accurate. However, this shows that there is a difference between the proteins, even in the absence of DNA (Figure 4.14). When DNA is added the structure of the proteins studied changes a little. For recording the percentage of secondary structure in the protein-DNA (1.4 μM) samples, after buffer subtraction, the DNA spectrum was subtracted from the spectrum of the complex and run through the CONTIN algorithm. It is interesting to see that the percentage of β-strand increases in the wild type sample, while it decreases in the mutant proteins. Again, this type of quantification is often unreliable, therefore change in conformation was analysed in a different way.

As mentioned above, if the formation of the protein-DNA complex does not cause a change in protein conformation, the spectrum recorded would simply be the sum of the protein and the DNA spectra. The spectrum obtained from the sum of the protein and the DNA spectra will be referred as “expected” spectrum. However, the values recorded for each protein-DNA complex, show that there is a difference between the “expected” spectrum and the “recorded” spectrum, for each of the proteins investigated at increasing concentrations of DNA (Figures 4.15, 16 and 17). In particular, we looked at the spectrum obtained from the subtraction of the “recorded” spectrum from the “expected” spectrum of the three ParG proteins at increasing concentrations of DNA. The new spectra should only show the change in conformation. Plots were made of the absolute value obtained for the wavelengths 208 and 220 nm against the concentrations of *parH* used. The titration graphs show that the conformation of wild

type ParG is the one changing the most. ParG double and triple mutants also change, but to a lower extent. In particular, ParG-K11A-K12A is the protein changing the least. The results do not contradict the EMSA results and again show that mutation of the Met-13 into alanine partially restores ParG characteristics. All things considered, the structure of ParG mutants does not change to the same extent as wild type ParG, supporting the hypothesis that the middle cluster of residues confers structure to the ParG N-terminal tail.

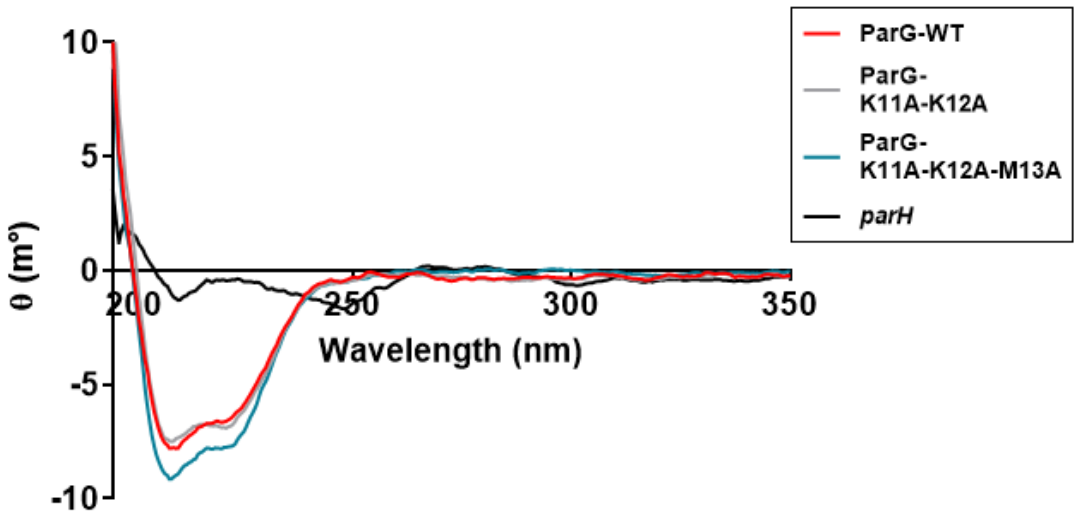

Figure 4.13 – Circular Dichroism spectra of ParG/ParG mutant proteins and *parH*

CD spectra of wild type ParG, ParG-K11A-K12A and ParG-K11A-K12A-M13A before addition of DNA and *parH*. The buffer spectrum was subtracted from the spectra shown. The CD signal is shown in millidegrees (m°).

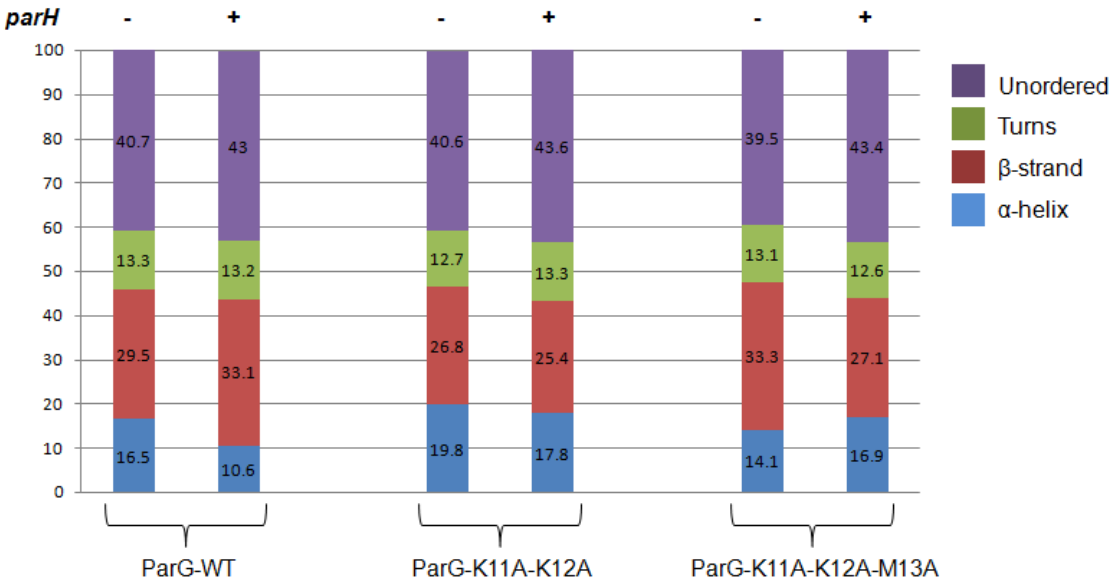

Figure 4.14 - Secondary structure prediction from the CD spectra using CONTIN algorithm and SP175 protein database (Dichroweb)

Spectra obtained from JASCO J180 circular dichromator were adjusted by subtracting the buffer spectrum and run through the CONTIN algorithm with SP175 database of soluble proteins, using Dichroweb software (Birkbeck University of London). The protein-DNA complex spectra were adjusted by first subtracting the DNA spectrum and then the buffer spectrum. The values were then analysed as above. Percentage graphs were plotted using Microsoft Excel.

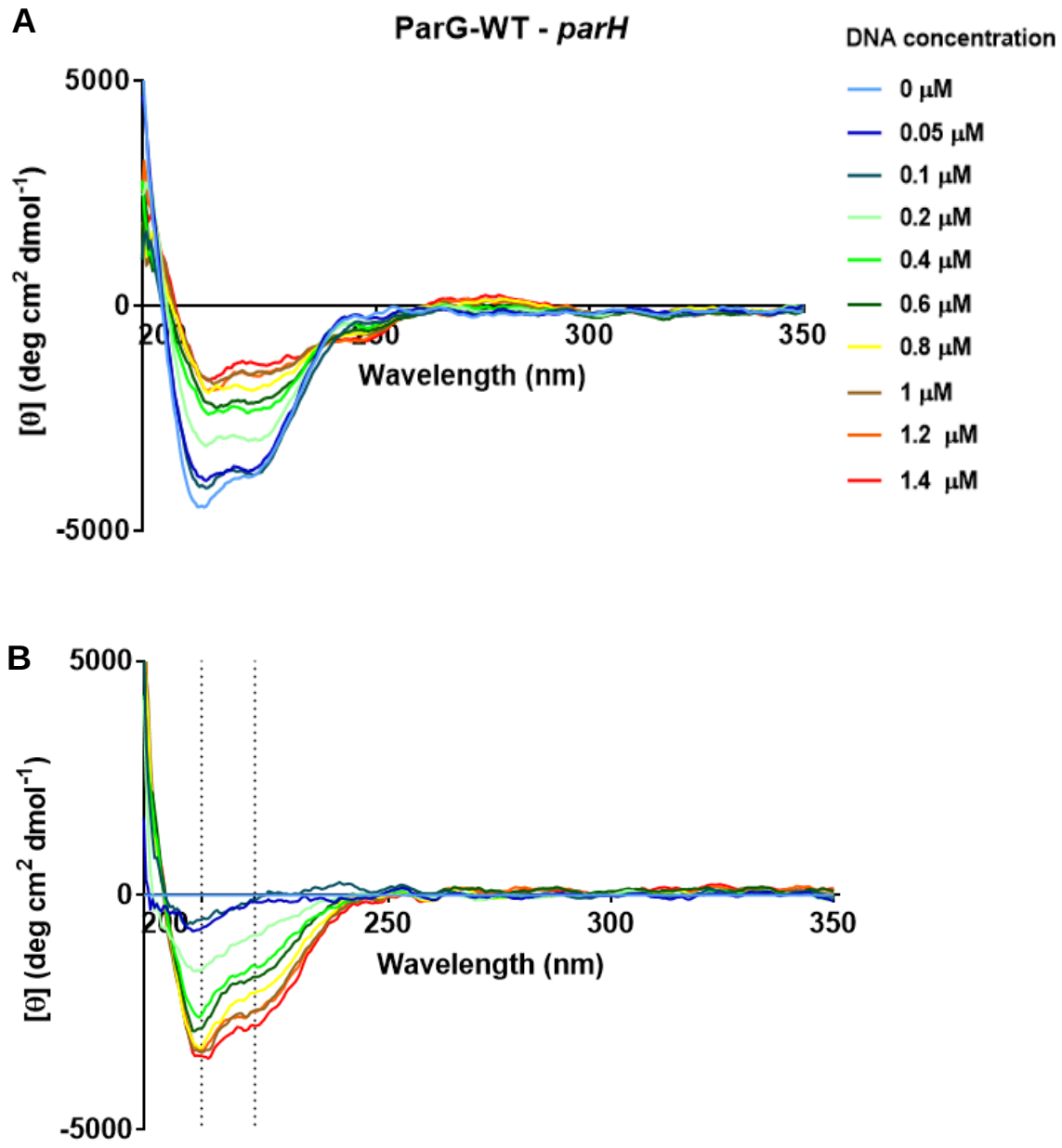

Figure 4.15 – Circular Dichroism spectra of ParG-WT with increasing concentration of *parH*

A) Circular dichroism spectra of ParG-WT in complex with increasing concentration of a region of *parH.* The values obtained from Jasco J810 circular dichromator, m° (millidegrees), were converted into Residual Molar Ellipticity and plotted using GraphPad Prism 7.04. The concentration of ParG used was 0.2 mg/ml (10.3 µM), with DNA concentration from 0 to 1.4 µM. Concentration changed slightly upon DNA addition, which was considered when plotting the graphs. Buffer CD spectrum was subtracted from each spectrum.

B) CD spectra obtained by subtracting the CD spectrum recorded from the expected CD spectrum if protein and DNA were not interacting (ParG CD spectrum + *parH* CD spectrum) for each DNA concentration. The dotted lines represent 208 and 220 nm wavelengths. The graph was plotted using GraphPad Prism 7.04.

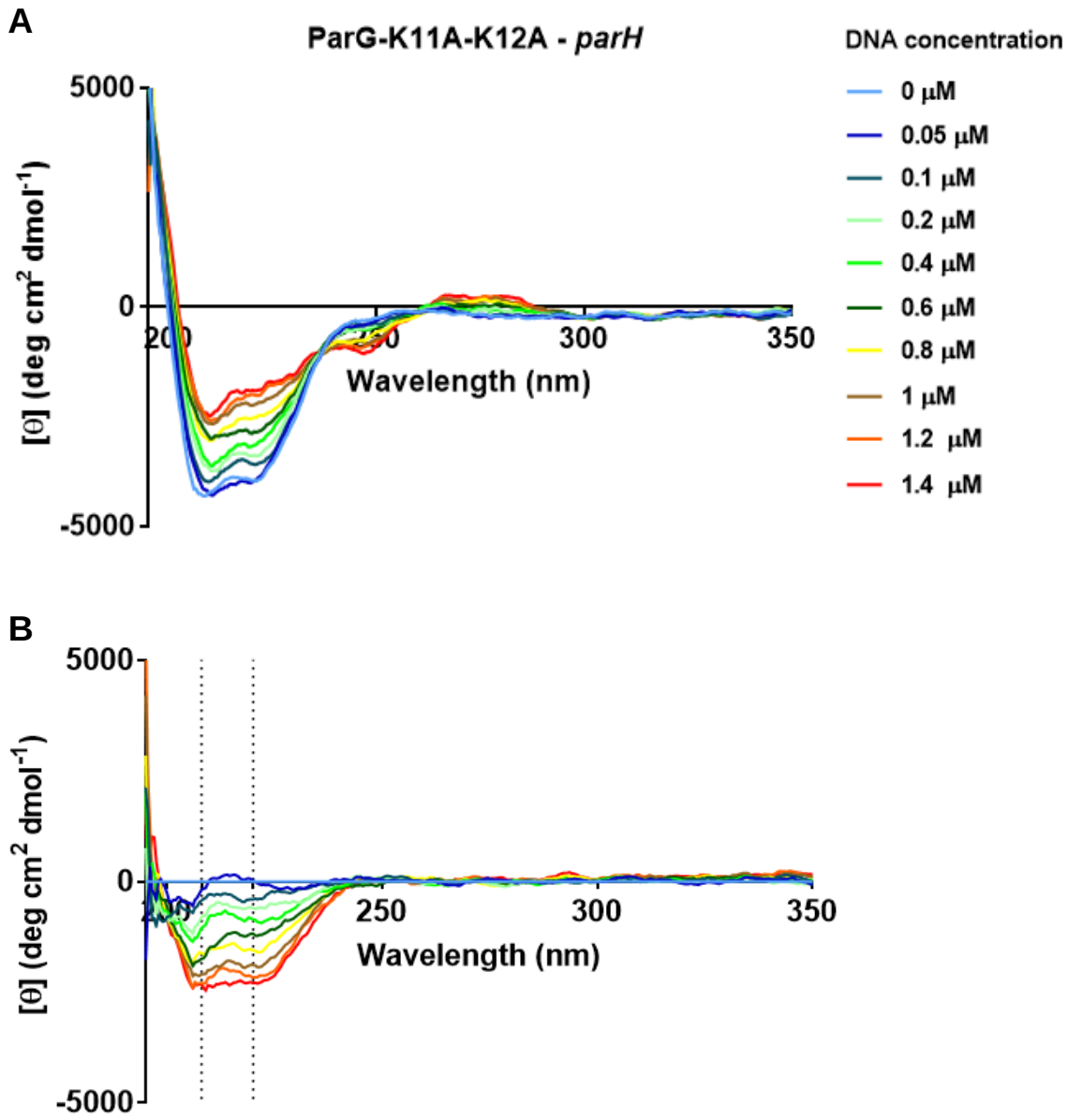

Figure 4.16 – Circular Dichroism spectra of ParG-K11A-K12A with increasing concentration of *parH*

A) Circular dichroism spectra of ParG-K11A-K12A in complex with increasing concentration of a region of *parH.* The values obtained from Jasco J810 circular dichromator, m° (millidegrees), were converted into Residual Molar Ellipticity and plotted using GraphPad Prism 7.04. The concentration of ParG double mutant used was 0.2 mg/ml (10.3 µM), with DNA concentration from 0 to 1.4 µM. Concentration changed slightly upon DNA addition, which was considered when plotting the graphs. Buffer CD spectrum was subtracted from each spectrum.

B) CD spectra obtained by subtracting the CD spectrum recorded from the expected CD spectrum if protein and DNA were not interacting (ParG double mutant CD spectrum + *parH* CD spectrum) for each DNA concentration. The dotted lines represent 208 and 220 nm wavelengths. The graph was plotted using GraphPad Prism 7.04.

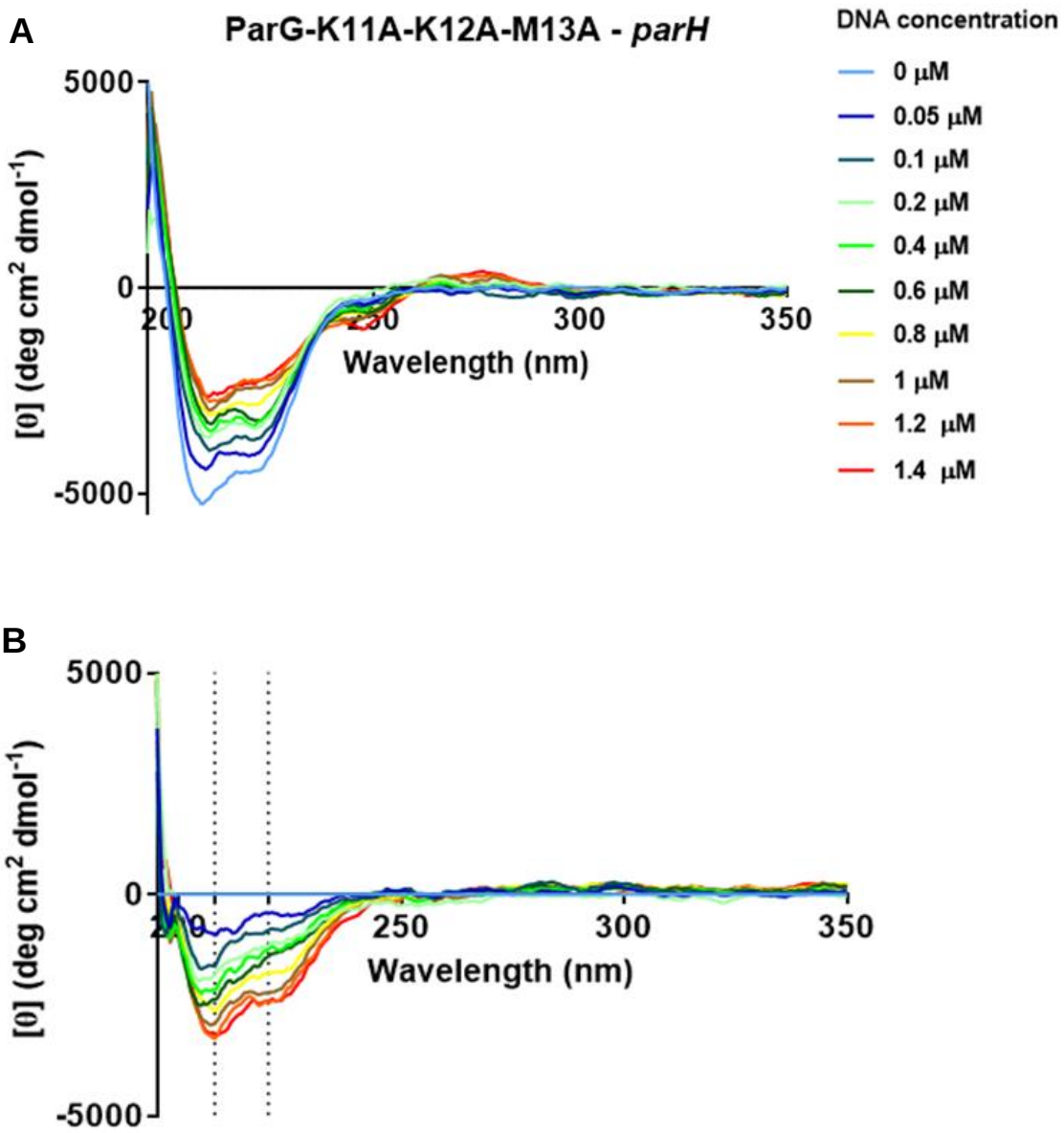

Figure 4.17 – Circular Dichroism spectra of ParG-K11A-K12A-M13A with increasing concentration of *parH*

A) Circular dichroism spectra of ParG-K11A-K12A-M13A in complex with increasing concentration of a region of *parH.* The values obtained from Jasco J810 circular dichromator, m° (millidegrees), were converted into Residual Molar Ellipticity and plotted using GraphPad Prism 7.04. The concentration of ParG triple mutant used was 0.2 mg/ml (10.3 µM), with DNA concentration from 0 to 1.4 µM. Concentration changed slightly upon DNA addition, which was considered when plotting the graphs. Buffer CD spectrum was subtracted from each spectrum.

B) CD spectra obtained by subtracting the CD spectrum recorded from the expected CD spectrum if protein and DNA were not interacting (ParG triple mutant CD spectrum + *parH* CD spectrum) for each DNA concentration. The dotted lines represent 208 and 220 nm wavelengths. The graph was plotted using GraphPad Prism 7.04.

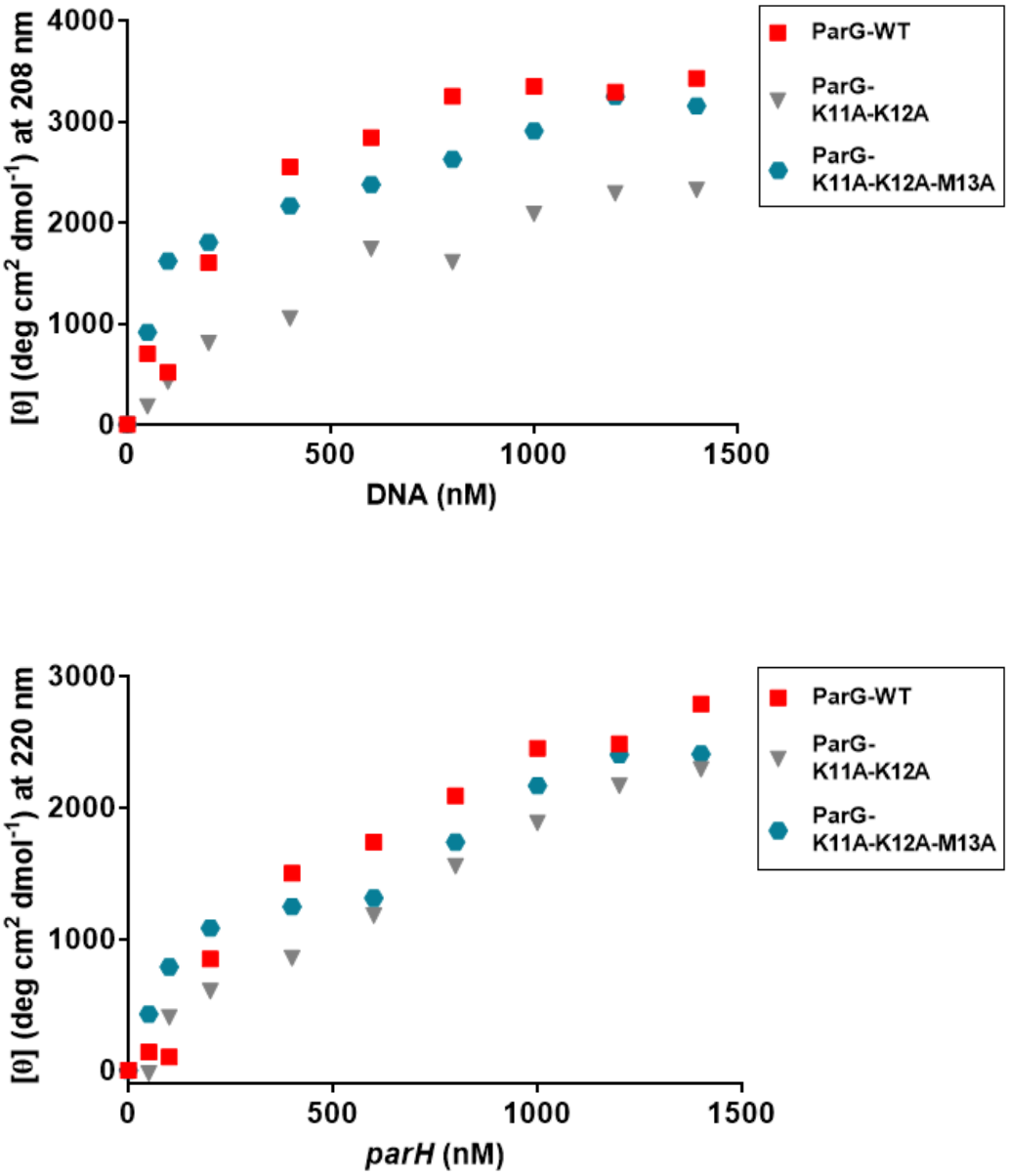

Figure 4.18 – Distribution of the Residual Molar Ellipticity values at 208 and 220 nm for wild type ParG and ParG mutants with increasing concentrations of *parH*.

The values plotted were obtained by subtracting the "recorded" ParG and ParG-DNA spectra from the "expected" corresponding spectra. The same was done for ParG mutants. The values were converted in absolute values to simplify visualisation of structural changes. Graphs were plotted using GraphPad Prism 7.04.

4.2.7 *parH* enhances ParG binding to ParF

Mechanistically, association between ParG and ParF is preceded by the formation of the ParG-*parH* nucleoprotein complex. Therefore, it is possible that the interaction between ParG and the site-specific DNA is required to structure ParG in an optimal conformation for ParF binding. According to our hypothesis the Lys-11, Lys-12 and Met-13 triad of amino acids structures the tail upon interaction with the DNA, enhancing ParG binding to ParF. This was tested by chemically cross-linking the two proteins in presence of *parH*. The same experiment was repeated with ParG double and triple mutants, as well as with ParGΔ30, to investigate whether amino acid changes would affect the cross-linking result.

Cross-linking reactions were carried out as detailed in Section 2.6.3.1. The same ParF-ParG reaction were prepared with and without *parH* for comparison. The DNA used for this experiment was the same as used for the circular dichroism experiments. In this case, the concentration of *parH* used was 1/7 of the ParG concentration (2.4 μM). ParF and ParG were present in a 1:1 ratio, with concentrations of 0.3 mg/ml (17 μM dimeric ParG, 14.3 μM ParF). Cross-linking reactions were incubated for 1 hour at room temperature and quenched with 50 mM Tris-HCl, pH 6.8, for 30 minutes before being loaded on 12% SDS- polyacrylamide gels. The reactions were prepared in triplicate to be loaded on three different gels. One of the gels was visualised by Coomassie staining, while the other two were transferred onto a nitrocellulose membrane for Western blotting. Band detection was then possible using α-ParF and α-ParG antibodies. Western blots with α-ParG were confusing because ParG trimers (29.1 kDa) migrate very closely to ParF-ParG complex (32.8 kDa). For this reason, those Western blots are not shown.

Both gels, Coomassie stained and immune-detected with α-ParF antibodies, show that cross-linking between wild type ParG and ParF led to the formation of a 32.8 kDa complex, that became more abundant when *parH* was added to the reaction (Figure 4.19). Similar results were obtained when ParG double mutant was cross-linked to ParF (Figure 4.19). Quantification of bands intensity showed that when *parH* is present, the amount of ParF-ParG complex formed is two-fold higher than when the two proteins react alone (Figure 4.21). On the other hand, cross-linking of ParG triple mutant with ParF showed very similar results to ParGΔ30-ParF cross-linking (Figure 4.20). Although both complexes are visible, *parH* does not enhance ParF-ParG-K11A-K12A-M13A or ParF-ParGΔ30 complexes assembly. Therefore, a similar amount of protein complex is visible in the presence and absence of *parH* in solution. This suggests that triple mutation affects the tail in such a way that the interaction with the DNA cannot enhance ParG binding to ParF.

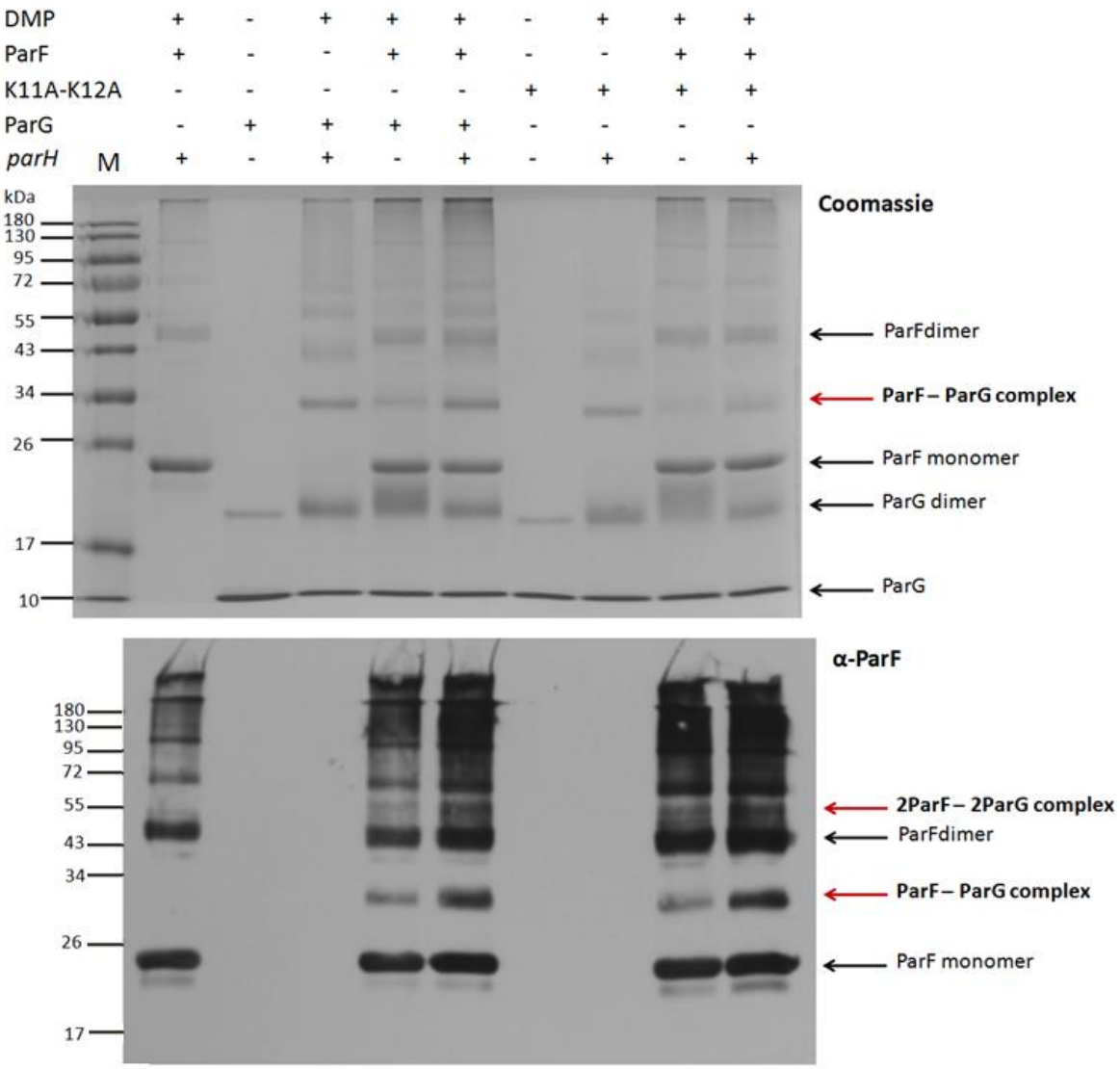

Figure 4.19 – SDS-PAG and Western blot of the DMP cross-linking between ParF and ParG/ParG-K11A-K12A in presence and absence of *parH*

Top and bottom images represent the same 12% SDS- polyacrylamide gel, visualised with different techniques. The top was Coomassie stained, while the second was detected using Western blot with α-ParF antibodies. For all the cross-linking reactions, DMP was used at a final concentration of 10 mM. The first lane represents the PageRuler prestained protein ladder (M) (Thermo Fisher Scientific). Lane 2 shows ParF cross-linked in the presence of *parH*. Lane 3 and 4 represent ParG cross-linked alone or in presence of *parH*. Lane 5 and 6 represent ParF-ParG cross-linking without and with *parH*. Lane 7 and 8 represents ParG double mutant cross-linked alone and with *parH*. Lane 9 and 10 represent ParF-ParG double mutant cross-linking reaction without and with *parH*.

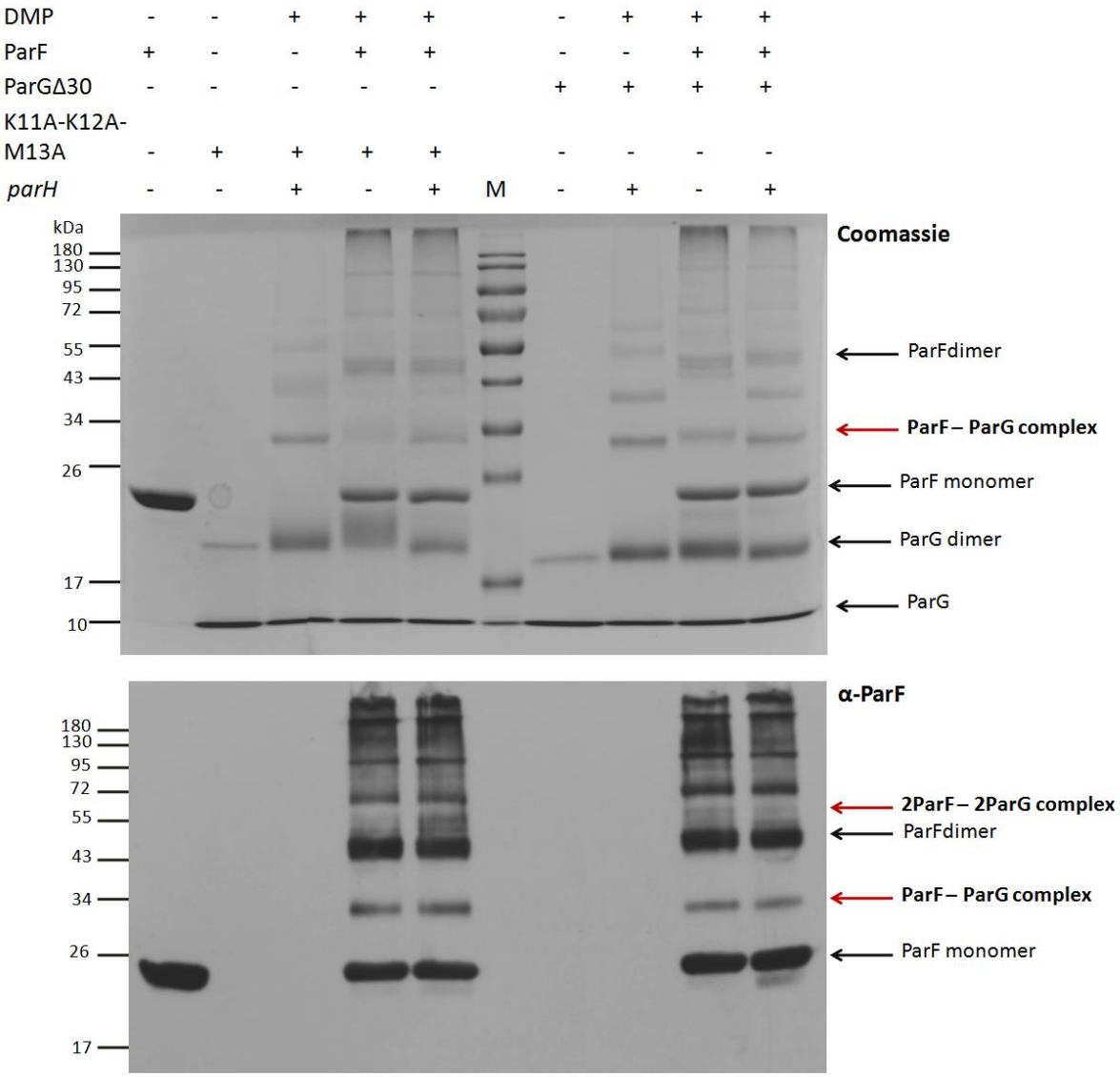

Figure 4.20 – SDS-PAG and Western blot of the DMP cross-linking between ParF and ParG-K11A-K12A-M13A/ParGΔ30 in presence and absence of *parH*

Top and bottom images represent the same 12% SDS- polyacrylamide gel, visualised with different techniques. The top was Coomassie stained, while the second was detected using Western blot with α-ParF antibodies. For all the cross-linking reactions, DMP was used at a final concentration of 10 mM. Lane 1 shows ParF non-crosslinked. Lane 2 and 3 represent ParG triple mutant cross-linked alone or in presence of *parH*. Lane 4 and 5 represent ParF-ParG triple mutant cross-linking reaction without and with *parH*. Lane 6 represents the PageRuler prestained protein ladder (M) (Thermo Fisher Scientific). Lane 7 and 8 represents ParGΔ30 cross-linked alone and with *parH*. Lane 9 and 10 represent ParF-ParGΔ30 cross-linking reaction without and with *parH*.

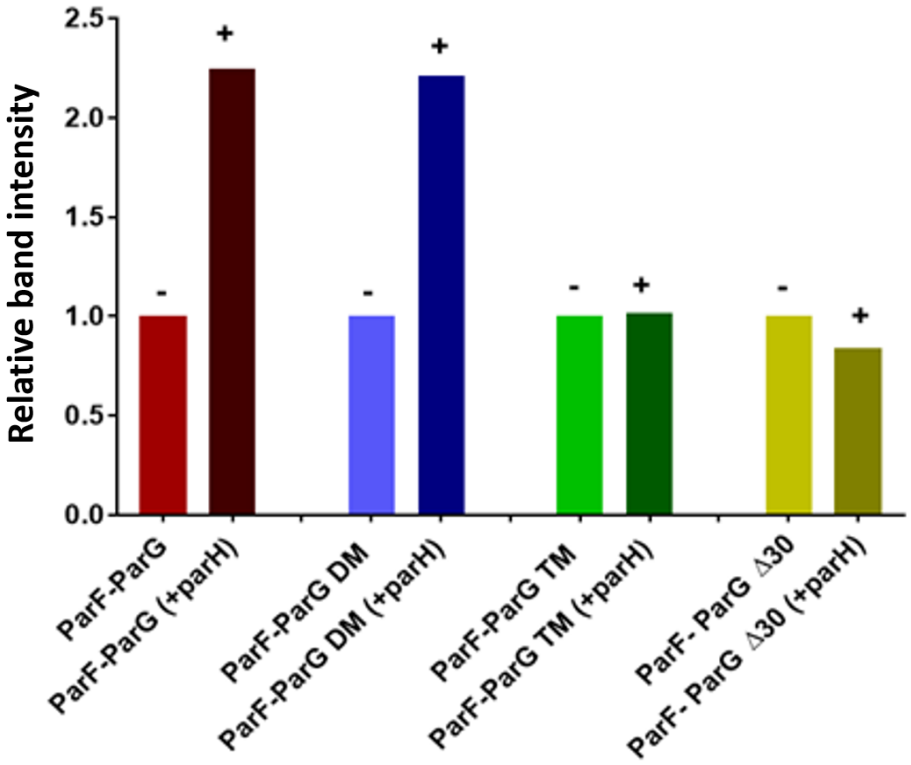

Figure 4.21 – Quantification of the intensity of ParF-ParG bands identified through α-ParF Western blot

ParF-ParG complex bands were quantified from the Western blot images using Image lab 4.0.1 software (Bio-rad). Band intensities were then plotted using GraphPad prism 7.04. The symbols (+) and (-) indicate the presence or absence of *parH*. DM represents ParG double mutant, ParG-K11A-K12A, TM represents the triple mutant, ParG-K11A-K12A-M13A.

4.2.8 Identification of amino acids interacting in the ParF-ParG complex using BS3 cross-linking and Liquid chromatography–mass spectrometry (LC-MS)

The hypothesis of ParG tail being structured upon interaction with *parH* and acquiring an optimal conformation for the interaction with ParF was further investigated using BS3 cross-linking, coupled to mass-spectrometry for identification of reacting amino acids and quantification of the cross-linking. In this experiment, ParF and ParG were incubated in four different conditions:

1. Alone in solution
2. In presence of *parH*
3. In presence of *parH* and non-specific chromosomal DNA (nsDNA)
4. In presence of *parH*, chromosomal DNA and ATP

In the segrosome, ParF and ParG do not interact alone. Instead, ParG is associated to *parH* and ParF to non-specific chromosomal DNA. ATP is also important in the system, as it mediates ParF dimerisation and oligomerisation, as well as enhancing ParG-ParF binding (Barillà and Hayes, 2003; Barillà *et al.*, 2005; Schumacher *et al.*, 2012). Cross-linking, followed by LC-MS/MS is a powerful tool to identify regions of interaction, or close proximity in protein complexes. BS3 is a 11.4 Å cross-linker that forms covalent bonds with amino groups (lysine residues), often used in this type of study (Yuan *et al.*, 2017; Sheppard *et al.,* 2016). In addition, quantitative cross-linking comparison between the different conditions allowed the identification of whether or not addition of specific and non-specific DNA enhanced the formation of the ParF-ParG complex.

Cross-linking reactions were prepared in triplicate as described in Section 2.6.3.2. Cross-linked species were then separated on gradient SDS-polyacrylamide gels and the bands corresponding to the ParF-ParG complex (1:1) were excised, digested with trypsin and analysed by LC-MS/MS. An example of SDS-PAGE showing the cross-linking reactions and a peptide spectrum with cross-linking identification are shown in Figure 4.22.

All cross-linked peptide identifications were filtered to 1% false discovery rate as assessed against a reversed sequence database. With the list of accepted identified cross-linked peptides, precursor ion areas were mapped across all samples to measure the relative response of each cross-linked peptide in each sample. The absolute response of peptides is sequence-specific, so values were normalised to relative percent for each peptide (i.e. for each identified cross-linked peptide the MS response summed across all 12 samples was fixed to 100) (Table 4.1). Plotting the relative MS1 response in this way allows for a visualization of relative abundance between samples

and also calculation of fold change and significant difference between treatments (Table 4.2). Although the ParF-ParG band was excised from the gel for analysis, ParG-ParG and ParF-ParF cross-linked species were also identified besides the most abundant ParF-ParG complex. Therefore, the cross-linked peptides identified were classified into two groups: inter-protein, if peptides from ParF and ParG were cross-linked together, or intra-protein, if cross-linking was within the same protein. Several ParF-ParF cross-linked peptides were identified, while only one ParG-ParG cross-linked peptide could be found. In this case the cross-linking was intra-peptide, suggesting that the cross-linker may have bridged adjacent lysine residues (Lys-33 and Lys-35), which could be an artefact. Although no other ParG-ParG cross-linkings were identified, ParG monomers could interact in regions which are not accessible to the BS3 cross-linker, thus not visible using this technique. ParF-ParF complex formation was clearly enhanced by the presence of ATP, as many of the cross-links identified were present in very low amount in the reactions containing no nucleotide. The lysine residues involved in the cross-linking are mostly surface exposed and covalent binding between ParF monomer was probably due to protein polymerisation either laterally or longitudinally (Schumacher *et al.*, 2012).

As for the ParF-ParG complex, many cross-links were identified between the two proteins. The lysine residues are all surface exposed, suggesting that it was hard for the cross-linker to reach regions embedded within the protein complex. Nevertheless, it is interesting to notice that Lys-11 and Lys-12 were among the residues involved in the cross-linking reaction with ParF. The other two amino acids identified within ParG were the two lysine residues at position 33 and 35, which belong to the β-strand of the folded C-terminal domain. ParF, made instead, interactions through the lysine residues in position 49, 52, 130, 140, 155, 160, 174 and 199 (Figure 4.23).
Interestingly, the cross-linking between Lys11/Lys-12 in ParG and Lys155 in ParF increased significantly when *parH* was present in solution. Since many amino acids interacted with more than one partner, it was hard to map the conformation of the complex. However, the number of cross-links identified in the samples where *parH* was added, was around 2.8 fold higher than the one identified in sample were the proteins were reacting alone, suggesting that specific DNA may somehow affect ParG conformation and enhance its interaction with ParF. In contrast, addition of chromosomal DNA only slightly improved the interaction, and values of cross-linking obtained were very similar to the ones of the ParFGH sample. The amount of nsDNA used was quite low (around 200 ng in the reaction) and it is possible that at this concentration its effect was negligible. As expected, addition of ATP was crucial, since

it improved complex formation almost 5.5 fold. As cross-linking data were not normally distributed, the median was used to describe central tendency, rather than the mean (Table 4.4). The violin and the box plots in Figure 4.24 show the distribution of the cross-linking identified among the three repeats of the same sample and across the four different conditions tested. Statistical analysis is summarised in Table 4.3.

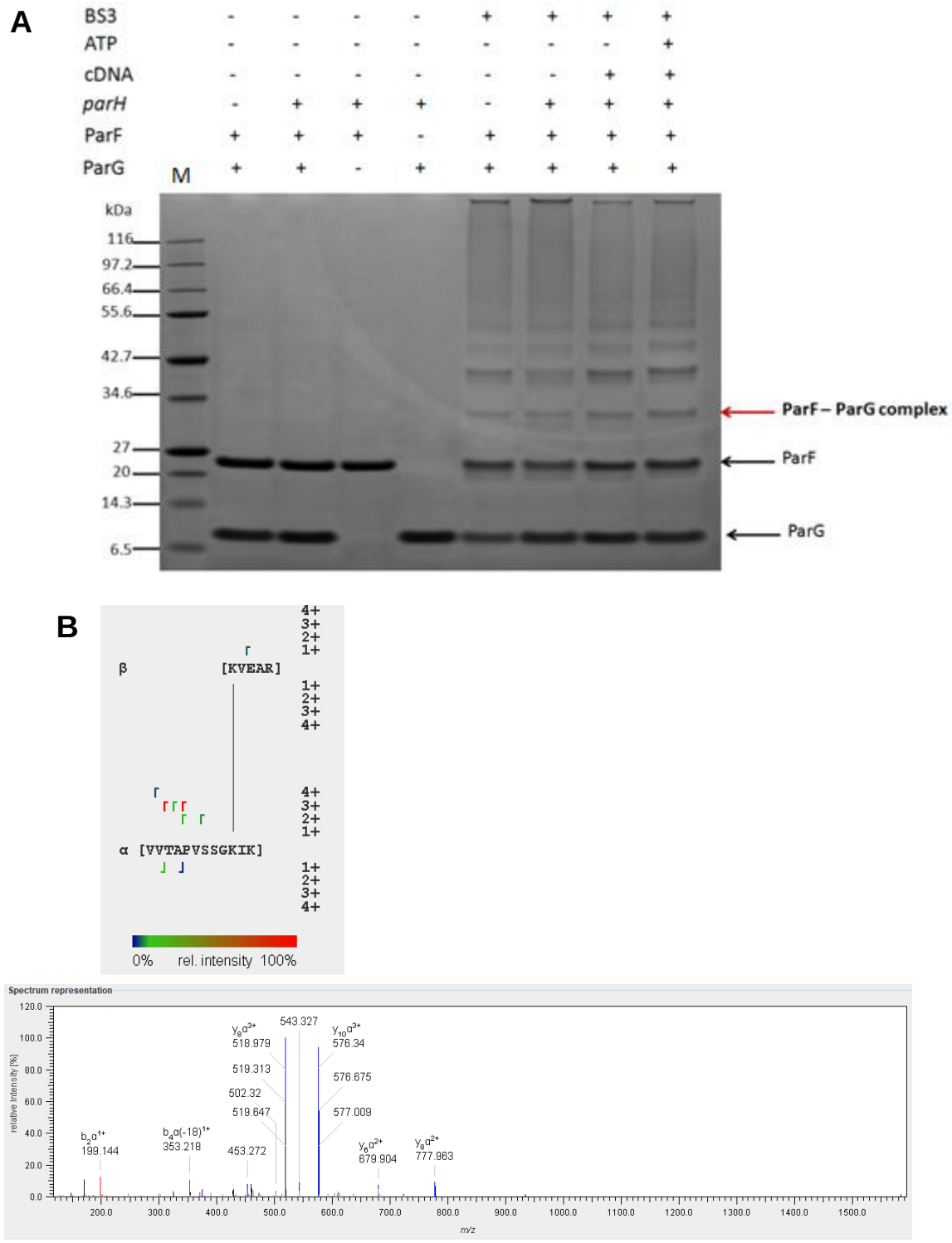

Figure 4.22 – Example of cross-linking reactions and cross-linked peptides identification after LC-MS analysis

A) Example of SDS-PAG with BS3 cross-linking between ParF and ParG reacting in different conditions. Lane 1: MW marker (NEB); lane 2-5: non-crosslinked controls; lane 6-9: ParF-ParG BS3 cross-linking. Lane 6: alone; lane 7: in presence of *parH*; lane 8: in presence of *parH* and chromosomal DNA; lane 9: in presence of *parH*, chromosomal DNA and ATP.

B) Example of cross-linking identification and annotation of the predicted fragment ions observed in the Tandem mass spec (MS/MS) spectrum. Intensity of the peaks is represented by the gradient line where blue-green is less intense and red is more intense.

m/z	FG1	FG2	FG3	FGH1	FGH2	FGH3	FGH D1	FGH D2	FGH D3	FGH DA1	FGH DA2	FGH DA3	XL Position F [1]	XL Position G	XL-Type
419.229	6	5	8	8	7	6	7	8	7	17	11	11	K155	K11	Inter-protein
431.225	3	0	5	12	5	11	1	12	13	19	6	12	K130	K12	Inter-protein
451.004	2	1	8	11	9	5	2	10	0	38	13	0	K160	K11	Inter-protein
461.933	2	2	5	7	10	5	9	12	8	26	15	0	K155	K11	Inter-protein
471.766	2	2	8	8	6	6	3	12	10	20	11	12	K155	K11	Inter-protein
482.041	8	4	12	5	4	4	2	4	6	26	11	14	K130	K33	Inter-protein
532.608	1	1	2	11	13	7	2	13	10	17	8	14	K155	K12	Inter-protein
575.584	8	5	13	7	6	4	1	6	8	22	10	10	K155	K33	Inter-protein
583.326	1	1	6	3	4	3	0	23	14	32	0	14	K140	K11	Inter-protein
600.355	1	1	6	6	10	5	4	6	6	27	16	12	K155	K35	Inter-protein
619.718	0	0	4	0	2	7	0	15	9	43	0	21	K174	K11	Inter-protein
645.338	1	2	25	3	3	9	0	12	8	24	2	12	K174	K11	Inter-protein
662.832	1	1	13	9	8	10	2	11	13	22	3	8	K52	K12	Inter-protein
681.405	0	0	7	0	0	4	0	2	3	58	0	25	K199	K35	Inter-protein
704.693	2	1	6	9	6	7	2	13	11	21	11	11	K155	K12	Inter-protein
709.644	4	4	15	10	8	6	1	8	8	22	8	7	K52	K33	Inter-protein
845.788	0	0	8	0	0	8	0	20	10	45	0	8	K199	K12	Inter-protein
934.203	0	0	15	3	4	9	1	19	17	20	0	12	K49	K12	Inter-protein

502.944	*2*	*1*	*4*	*4*	*4*	*3*	*1*	*5*	*4*	*30*	*19*	*22*	***K71-K160***		Intra-Protein
408.976	*3*	*1*	*5*	*4*	*7*	*2*	*2*	*3*	*3*	*38*	*15*	*19*	***K130-K160***		Intra-protein
523.296	*1*	*0*	*3*	*3*	*3*	*2*	*0*	*5*	*4*	*41*	*12*	*25*	***K130-K160***		Intra-protein
602.674	*3*	*2*	*7*	*14*	*6*	*13*	*0*	*20*	*16*	*9*	*5*	*5*	***K1-K155***		Intra-protein
671.722	*0*	*0*	*5*	*1*	*1*	*11*	*2*	*18*	*9*	*39*	*1*	*12*	***K155-K199***		Intra-protein
676.107	*0*	*0*	*6*	*1*	*1*	*6*	*1*	*45*	*12*	*16*	*1*	*12*	***K155-K64***		Intra-protein
709.373	*2*	*2*	*5*	*13*	*9*	*11*	*0*	*12*	*11*	*18*	*8*	*9*	***K155-K52***		Intra-protein
976.83	*0*	*1*	*6*	*4*	*3*	*9*	*0*	*13*	*12*	*37*	*4*	*11*	***K155-K174***		Intra-protein
1065.867	*0*	*1*	*5*	*3*	*1*	*8*	*4*	*14*	*7*	*41*	*4*	*11*	***K49-K155***		Intra-protein
1071.2	*0*	*1*	*4*	*5*	*3*	*7*	*4*	*9*	*7*	*29*	*12*	*19*	***K49-K155***		Intra-protein
493.968	*15*	*12*	*25*	*3*	*4*	*2*	*3*	*3*	*2*	*13*	*9*	*9*		***K33-K35***	Intra-peptide
496.941	*2*	*1*	*4*	*4*	*4*	*3*	*1*	*4*	*4*	*30*	*19*	*24*	***K155-K160***		Intra-peptide

Table 4.1 – Relative percentage of peptides identified across the ParFG (FG), ParFGH (FGH), ParFGH-nsDNA (FGHD), ParFGH-nsDNA-ATP (FGHDA) samples

Peptide identified through LC-MS/MS were filtered to 1% FDR and relative percentage of peptide abundance was calculated by fixing the MSresponse (AUC, area under the curve) across the twelve samples to 100%. Cross-links are categorised as inter-protein (ParF-ParG) or intra-protein (ParF-ParF) or intra-peptide (within the same peptide). m/z represent the mass over charge value, FG the cross-linking between ParF and ParG with no additional component, FGH the cross-linking between ParF and ParG in presence of *parH*, FGHD the cross-linking between ParF and ParG in presence of *parH* and chromosomal DNA and FGHDA the cross-linking between ParF and ParG in presence of *parH*, chromosomal DNA and ATP. All experiments were repeated in triplicate and the single repeats are shown (1-2-3).

XL Position F [1]	XL Position G	XL-Type	FGH/FG	FGH/FG p-value	FGHD/FG	FGHD/FG p-value	FGHDA/FG	FGHDA/FG p-value
K155	***K11***	Inter-protein	1.18	0.41	1.23	0.26	2.17	0.06
K130	***K12***	Inter-protein	3.37	0.08	3.17	0.26	4.64	0.11
K160	***K11***	Inter-protein	2.39	0.15	1.12	0.92	4.80	0.35
K155	***K11***	Inter-protein	2.32	0.11	2.96	0.02	4.30	0.30
K155	***K11***	Inter-protein	1.75	0.27	2.14	0.28	3.67	0.05
K130	***K33***	Inter-protein	0.52	0.24	0.47	0.21	2.12	0.18
K155	***K12***	Inter-protein	8.61	0.04	6.80	0.16	10.83	0.05
K155	***K33***	Inter-protein	0.65	0.33	0.60	0.33	1.57	0.36
K140	***K11***	Inter-protein	1.37	0.66	5.37	0.27	6.59	0.30
K155	***K35***	Inter-protein	2.76	0.10	2.18	0.19	7.13	0.06
K174	***K11***	Inter-protein	2.23	0.57	6.36	0.26	17.23	0.25
K174	***K11***	Inter-protein	0.57	0.68	0.72	0.79	1.38	0.75
K52	***K12***	Inter-protein	1.88	0.42	1.74	0.55	2.24	0.44
K199	***K35***	Inter-protein	0.61	0.75	0.81	0.88	12.05	0.27
K155	***K12***	Inter-protein	2.35	0.11	2.78	0.22	4.55	0.06
K52	***K33***	Inter-protein	1.01	0.98	0.71	0.64	1.60	0.49
K199	***K12***	Inter-protein	1.05	0.98	3.70	0.33	6.54	0.40
K49	***K12***	Inter-protein	1.11	0.93	2.44	0.39	2.15	0.49

K71-K160		Intra-Protein	1.20	0.58	1.29	0.61	8.69	**0.02**
K130-K160		Intra-protein	1.51	0.44	0.85	0.72	8.55	0.09
K130-K160		Intra-protein	1.68	0.39	1.84	0.47	16.28	0.10
K1-K155		Intra-protein	2.82	0.09	3.08	0.30	1.58	0.34
K155-K199		Intra-protein	2.66	0.49	5.72	0.23	10.13	0.30
K155-K64		Intra-protein	1.22	0.88	9.47	0.32	4.79	0.24
K155-K52		Intra-protein	3.44	**0.01**	2.43	0.36	3.71	0.10
K155-K174		Intra-protein	2.18	0.37	3.40	0.30	7.13	0.27
K49-K155		Intra-protein	2.12	0.46	4.58	0.13	9.76	0.27
K49-K155		Intra-protein	3.09	0.12	4.22	0.06	12.29	0.06
	K33-K35	Intra-peptide	0.16	0.06	0.15	0.06	0.59	0.20
K155-K160		Intra-peptide	1.38	0.41	1.24	0.67	9.51	**0.01**

Table 4.2 – Statistical difference between samples and fold increase between different conditions

Peptide abundance was compared across the sample conditions and significant difference was calculated by Student's t-test. Relative amount is shown, together with the p-value. Significant difference is described by $p<0.05$.

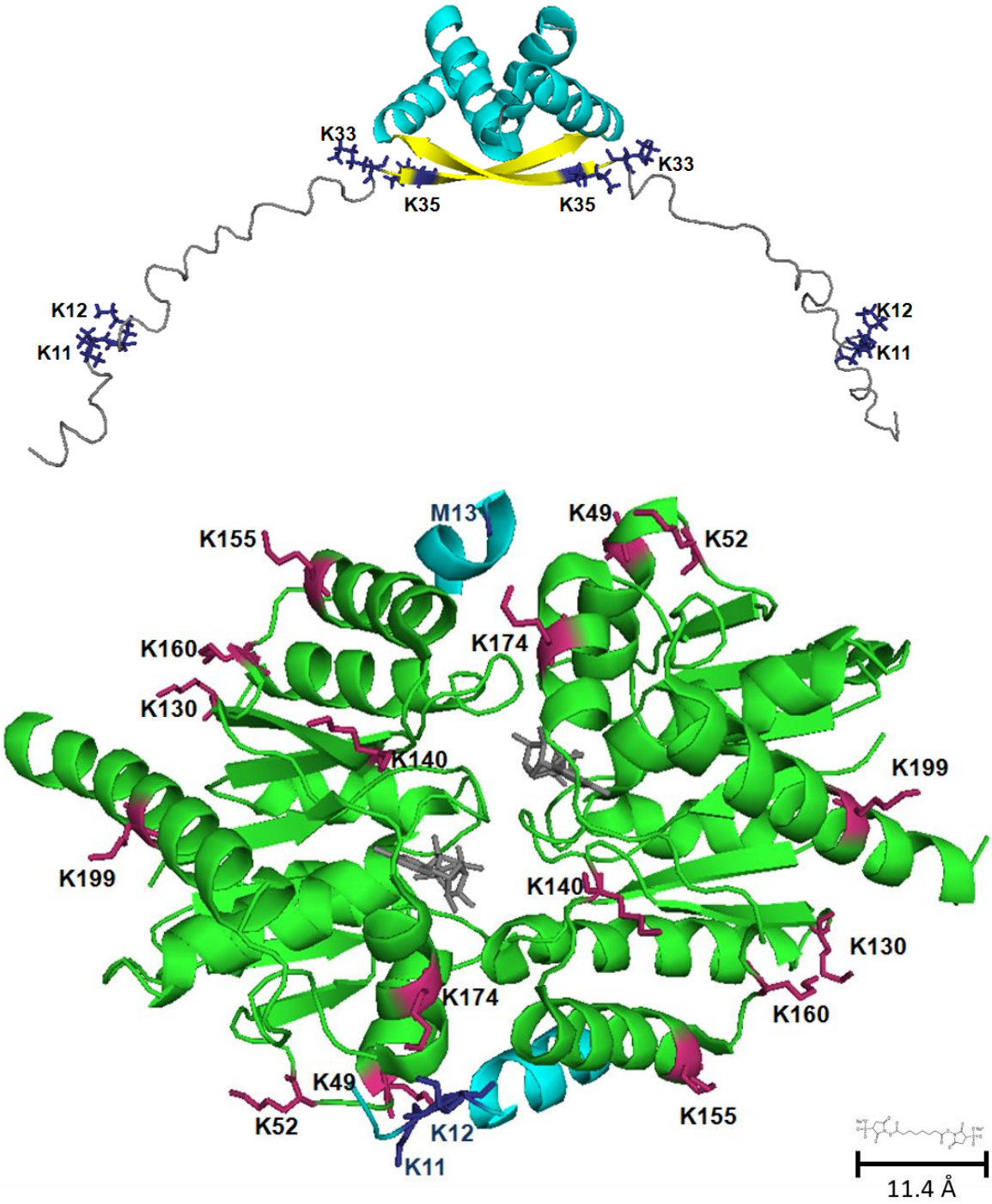

Figure 4.23 – ParG and ParF structures showing the interacting the amino acids identified through cross-linking and LC-MS

The top image shows ParG dimer NMR structure (PDB: 1P94). Amino acids shown as sticks and represented in blue are the ones identified in the cross-linking reaction with ParF. The bottom image shows ParF dimer co-crystallised with a small region of ParG arginine finger motif (PDB: 5U1G). The amino acids shown in blue belong to ParG (shown in cyan), while the one shown in hot pink belong to ParF (shown in green). The residues highlighted here were identified as involved in the cross-linking with the partner protein. The scale bar represents the length of the BS3 cross-linker (11.4 Å). The structure of the molecule is shown above.

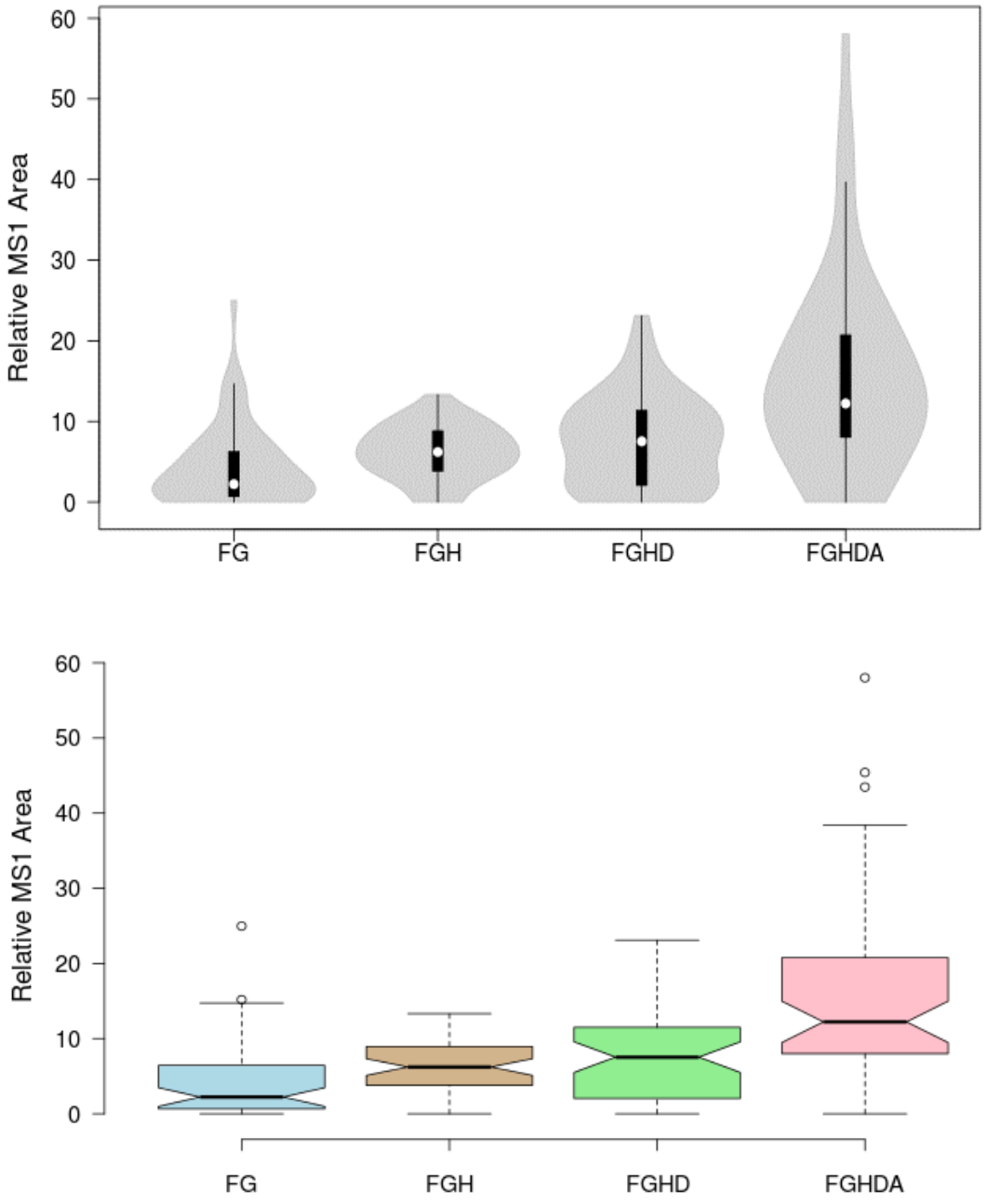

Figure 4.24 – Violin-plot (Top) and Box-plot (bottom) representing the distribution of the cross-linking data across the four conditions investigated

The violin plot represents the distribution of the data as a density curve, while the box-plot summarises the statistical information for each condition. In the violin plot, the white dot represents the median of the values, the black rectangle represents the inter-quartile range, the black line represents the 95% confidence interval and the width of the density plot represent the frequency. The box plot is a schematic representation of the same values, summarised in Table 4.3. The labels read as follows: FG-ParFG complex, FGH-ParFG*H* complex, FGHD- ParFG*H* in presence of non-specific DNA and FGHDA- ParFG*H* in presence of non-specific DNA and ATP.

Metric	FG	FGH	FGHD	FGHDA
Upper whisker	14.74	13.31	23.1	38.4
3rd quartile	6.44	8.96	11.5	20.77
Median	**2.23**	**6.22**	**7.53**	**12.23**
1st quartile	0.66	3.79	2.05	8.01
Lower whisker	0	0	0	0
Nr. of data points	54	54	54	54
IQR	5.78	5.17	9.45	12.76
IQR/Sqrt(n)	0.79	0.70	1.29	1.74
Upper 95% C.I. +1.58*IQR/sqrt(n)	**3.41**	**7.28**	**9.46**	**14.83**
Lower 95% C.I. +1.58*IQR/sqrt(n)	**0.99**	**5.11**	**5.50**	**9.49**

Table 4.3 - Summary of the statistical analysis of the three repeats of the cross-linking data for each condition

	FG	FGH	FGHD	FGHDA
Median	2.23	6.22	7.53	12.23
	Fold change vs. FG	2.79	3.38	5.48

Table 4.4 – Median of the cross-linking data among the three repeats for each condition and fold change compared to ParF-ParG cross-linking alone

4.3 Conclusions

In the segrosome, ParG interacts with the centromeric site *parH*, as well as the partner protein ParF. In the course of the interaction, the formation of the ParG-*parH* nucleoprotein complex precedes the association between the two protein components. This led us to hypothesise that ParG is required to associate with the specific-DNA in order to acquire an optimal conformation for association with ParF. Biophysical techniques such as SPR and MST monitor changes caused by molecular interactions. However, they cannot discriminate the nature of the interaction, whether it is protein-protein, protein-DNA or protein-ligand for example. Therefore, in a three components system (four components, if we consider ATP), such as the ParFGH system, at least one component must be neglected. Hence, on some occasion, ambiguous results were obtained and understanding the physiological behaviour proved challenging.

In this Chapter we tried to address an unanswered question from Chapter 1, the role played by the triad of amino acids Lys-11, Lys-12 and Met-13 in the formation of the partition complex. Single amino acid changes in this region did not abolish the interaction between ParG and ParF, leading us to believe that this cluster of amino acids could be involved in other activities during segrosome assembly. HSQC experiments comparing the chemical shift of the ParG residues in absence and presence of part of the operator site, showed that Lys-11 and Lys-12 were particularly affected by the addition of the cognate DNA. Other amino acids whose chemical shift was seen to change were: His-7, Met-13, Asn-18, Arg-19, Arg-23, Thr-26, Ala-27, and Val-29 (Carmelo *et al.,* 2005). One hypothesis was that these amino acids make transient interaction with the DNA. Another hypothesis is that the association of the folded domain with the DNA may alter the structure of the unfolded N-terminal region. Carmelo *et al.* also proposed that the transient β-strand (β^T) formed by the least flexible amino acids in position 20 to 30 could temporarily associate with the folded domain, covering the β-structure in the C-terminal domain, required for ParG interaction with the DNA. The β^T-strand would then be displaced upon interaction with the cognate DNA, allowing the C-terminal β-structure to insert into the major groove of the DNA (Carmelo *et al.*, 2005). Evidence suggests that the N-terminal domain makes interaction with the DNA, however the mechanism of binding is unknown.

Biophysical techniques such as SPR and MST showed that when multiple amino acid changes occurred in the K11-K12-M13 cluster, ParG interaction with ParF was greatly affected. MST showed that both double (K11A-K12A) and triple (K11A-K12A-M13) amino acid changes abolished ParF-ParG interaction. However, SPR showed that

flowing ParG mutants over the ParF-coated chip caused a low response, suggesting that, although very weak, ParG interaction with ParF was still happening. Differences between the binding experiments were probably caused by differences in the assay set up, as well as technique sensitivity. ParGΔ30, which was seen to interact with ParF in a two-hybrid assay (Barillà *et al.*, 2007), was unable to bind to its partner protein *in vitro* (both with SPR and MST). Binding between ParF and ParG proteins was, however, confirmed by chemical cross-linking, which showed that all three ParG variants could still weakly associate with the partner protein (Figure 4.19 and 4.20).

We were also interested in understanding whether double and triple mutations could affect the structure of the tail. ParG N-terminal tails are normally disordered but were hypothesised to become structured when in complex with ParF or *parH*. ParF forms extensive oligomers in presence of ATP, and ParG enhances the formation of ParF higher order structures by bundling or stabilising ParF oligomers (Barillà *et al.*, 2005). Barge showed that single amino acid changes in the ParG N-terminal domain did not affect the ability of ParG to enhance the formation of ParF higher-order structures. However, when double and triple mutations were inserted in the middle region of the tail, bundling by ParG was greatly affected. ParG-K11A-K12A could still enhance ParF oligomerisation, although more slowly compared to wild type ParG. On the other hand, ParG triple mutant behaved similarly to ParGΔ30. In the DLS experiment, addition of both proteins to ParF-ATP caused an increase in scattering intensity, although much lower compared to the one caused by the wild type protein. This suggests that ParG action on ParF oligomers was limited and weak. A plausible hypothesis is that triple mutation has a detrimental effect on ParG tail structuring, preventing the protein from assuming a correct conformation for the interaction with ParF and bridging ParF monomers. Thus, Met-13 could be a key player in structuring the ParG N-terminal domain.

This statement was further supported by the results obtained from the chemical cross-linking experiment with DMP. Since precise biophysical analysis is very complicated when more than two components are involved in the binding, we used semi-quantitative cross-linking to demonstrate that *parH* enhances the association between ParF and ParG. The amount of ParF-ParG complex formed when *parH* was present in solution was two-fold higher than the amount of complex obtained when the two proteins were reacting alone. Double mutation of the residues Lys-11 and Lys-12 did not affect complex formation, as *parH* caused a two-fold increase also in the interaction between ParG-K11A-K12A and ParF. However, similar enhancement was not recorded when *parH* was added to the cross-linking reaction between ParF and ParG triple mutant. Indeed, as for

ParGΔ30, addition of the cognate DNA did not improve complex formation and same amount of ParF-ParG mutant complex was obtained, whether *parH* was present or not in solution. This experiment showed that Met-13 plays an essential role in enhancing ParG-ParF association, either through the interaction with the DNA or by structuring the tail.

Interaction between the mutant proteins and the specific DNA was also tested by using EMSA. Both ParG-K11A-K12A and ParG-K11A-K12A-M13A showed slightly lower affinity for full length *parH*, compared to the wild type protein. Due to time constraint and difficulties in optimising the protocol, the interaction could not be better quantified by biophysical methods. However, these data would have helped elucidating whether or not residues 11, 12 and 13 make interaction with the DNA.
From the semi-quantitative results, it appears that the double mutation affects the interaction with the specific DNA more than the triple mutation. This could be due to the fact that methionine is a relatively large amino acid and could cause steric hindrance in the interaction with the DNA, especially if the positive charge of Lys-11 and Lys-12 has been removed. In the triple mutant, instead, the methionine is changed into a small alanine residue, facilitating the interaction with the DNA.

A similar argument could be used to interpret the CD results. The technique was used to understand whether ParG conformation changed upon interaction with the DNA and whether double and triple mutation would affect this. CD showed that ParG secondary structure changed when increasing concentration of *parH* were titrated over a constant concentration of the protein. A similar change in conformation was also recorded for ParG mutants. It is important to note that the change in the protein CD spectrum was not caused by the addition of the DNA. In fact, if the protein structure was not changing upon addition of the DNA, the spectrum recorded would have been the sum of protein and the DNA spectra. Instead, the “expected”-“recorded” spectra for each DNA concentration showed that a change in conformation was indeed visible. Comparison with double and triple mutants was attempted by plotting the absolute CD values at specific wavelengths against the DNA concentration. This plot allowed the visualisation that ParG spectra changed and reached steady state more rapidly than the ones of the mutant proteins. In particular ParG double mutant proved to be the protein least affected by the interaction with the DNA. These results supported the hypothesis that the presence of the methionine residues in position 13, when the two positively charged lysine residues have been changed into alanine, is actually detrimental for the protein, particularly for its association with the DNA. Nevertheless, this experiment successfully proved that when

amino acid changes occur in the middle cluster of the ParG N-terminal tail, protein structuring is affected.

Although these experiments did not confute the initial hypothesis, they only partially supported it. The residue Met-13 does seem to have a structural role, as demonstrated by the DMP cross-linking experiment and the DLS. The fact that triple mutation did not weaken the interaction with the DNA, instead, supported the idea that this residue is not essential for interaction with *parH*. Further understanding of the role played by Lys-11 and Lys-12 was instead possible through the BS3 cross-linking/LC-MS/MS experiment. Firstly, the experiment showed that the presence of *parH* enhanced ParG binding to ParF (around 2.8 fold) and that addition of chromosomal DNA enhanced binding slightly more. As expected, addition of ATP enhanced the binding dramatically. In fact, when ParF and ParG were cross-linked in presence of *parH*, nsDNA and ATP, the number of cross-linked peptides was 5.5 fold higher than when the proteins were cross-linked alone.
Interestingly, the cross-linking experiments showed that Lys-11 and Lys-12 were found among the amino acids involved in interaction with ParF. The two residues were also found to be interchangeable, as ParF Lys-155 was cross-linked to both ParG Lys-11 and Lys-12. The two adjacent lysine residues could, therefore, replace each other, explaining why single mutation did not affect ParF-ParG interaction to a greater extent (Chapter 3). In addition, $Lys12_{ParG}$ - $Lys155_{ParF}$ cross-linking was seen to be significantly increased when *parH* was present in the mix. $Lys\text{-}12_{ParG}$ was also seen to make interaction with ParF Lys-52, Lys-130, Lys-49 and Lys-199. However, $Lys\text{-}11_{ParG}$ was also cross-linked to ParF Lys-174, Lys-140 and Lys-160. These amino acids are scattered around ParF surface, suggesting that ParG N-terminal domain must assume different conformations to interact with all of them.
In Chapter 3, ParF-K160E-R163E mutant was used as ParF substitute in the SPR experiments, as mutations increased protein solubility without affecting the interaction with ParG (Allison, 2016). Identification of the cross-linking between ParG Lys-11 and ParF Lys-160 does not necessarily mean that the two amino acids interact, it may also simply suggest that they are in close proximity. Since biophysical analysis demonstrated that the binding between ParG and ParF double mutant is identical to the interaction with the wild type protein, it is possible that the interaction between $Lys\text{-}11_{ParG}$ and $Lys\text{-}160_{ParF}$ is non-specific.

It was interesting to observe that Lys-33 and Lys-35 in ParG C-terminal domain were also cross-linked to ParF. $Lys\text{-}33_{ParG}$ was cross-linked to Lys-52, Lys-130 and Lys-155 in ParF. Instead, $Lys\text{-}35_{ParG}$ was cross-linked to Lys-199 and Lys-155 in ParF.

In some instances, the same ParG residue were cross-linked to ParF residues locate on totally opposite regions of the protein, making impossible simultaneous interaction. Therefore, it is possible that besides the specific association for segrosome formation, ParG could make non-specific interaction with ParF. ParG was seen to bundle ParF polymers and could therefore be involved in cross-linking ParF monomers through transient interactions of the N-terminal tail with ParF surface-exposed amino acids. This would also explain why Lys-11 and Lys-12 in ParG were cross-linked to amino acids on both sides of the ParF dimer.

From this experiment it emerged that Lys-11 and Lys-12 in ParG could be involved in the interaction with both *parH* and ParF. Specific DNA improves the binding between ParG and ParF, probably by changing the conformation of the N-terminal domain. The initial hypothesis was that K11-K12-M13 cluster was performing this activity. Since Lys-11 and Lys-12 seem to interact with ParF, either other regions of the tail or, more likely, the C-terminal domain could trigger the change in conformation of the protein upon DNA recognition. Nevertheless, Met-13 showed to have a structural role, since triple mutation prevented *parH* from enhancing ParG binding to ParF, suggesting that the tail structuring could be impaired.

All things considered, it appears that *parH* recognition by ParG triggers a change in the protein conformation, required for optimal binding to ParF. Since HSQC showed that the chemical shift of Lys-11 and Lys-12 changes when the protein is in complex with the DNA, this part of the tail could become structured for the interaction with ParF and structuring could be mediated by the K11-K12-M13 cluster of amino acids.

Chapter 5
Characterisation of ParF interaction with potential *E. coli* host factors

5.1 Introduction

Plasmids are self-replicating units that encode their own segregation systems. *In vitro* reconstruction of the partition apparatus showed that the three basic elements, ParA, ParB and *parS,* that characterise all partition cassettes, can alone drive plasmid segregation (Hwang *et al.*, 2013; Vecchiarelli *et al.*, 2013; Hu *et al.*, 2015). Nevertheless, the possibility that host-encoded factors may participate in the process has always been a fascinating question that so far has remained unanswered. IHF was seen to play a role in P1 and possibly pB171 plasmid segregation, bending and structuring the DNA for ParB binding (Funnell, 1988b; Ringgaard *et al.*, 2007). However, no other host-encoded accessory protein has ever been identified. Gerdes and co-workers tested the interaction of the pB171 *par2* partition proteins with a number of cell division elements (e.g FtsZ, FtsK, MreB and others) and using a bacterial two hybrid system they observed that none of them was interacting *in vivo* with ParA or ParB (Ebersbach *et al.*, 2006).
Stability of different plasmids was also tested in Δ*mukB* strains. However, experiments showed that the condensin protein is not essential for plasmid segregation (Ezaki *et al.*, 1991; Funnell and Gagnier, 1995; Weitao *et al.*, 2000). This evidence suggests that host factors do not participate in plasmid segregation, but instead the minimalist tripartite ParAB*S* system is sufficient to ensure proper partition.

On the other hand, studies on chromosome segregation in different Gram-positive and Gram-negative bacteria showed that chromosomal ParAB*S* systems work in cooperation with other proteins. SMCs were identified as essential for chromosome compaction and segregation in different organisms, as *C. crescentus*, *B. subtilis* and *Streptococcus pneumoniae* (Jensen and Shapiro, 1999; Gruber and Errington, 2009; Sullivan *et al.*, 2009; Minnen *et al.*, 2011). During the cell cycle, SMC is recruited by ParB to the *oriC* site, through indirect interaction with *parS,* to condense the DNA for chromosome partition (Gruber and Errington, 2009; Sullivan *et al.*, 2009; Minnen *et al.*, 2011).

In addition, ParB interacts with membrane proteins to tether the partition complex at the cell poles, before cell division. In *C.crescentus* the ParB-*parS* complex is anchored to the cell membrane through the protein PopZ, which accumulates at the cell poles through a mechanism that involves the actin-homologue MreB (Bowman *et al*., 2008). ParB directly binds also to the ATPase protein MipZ. The protein follows ParB throughout the cell cycle and prevents the formation of the Z ring by promoting FtsZ polymers dissociation. ParA, instead, interacts with TipN protein, which initially was identified as the factor that attach ParA polymers to the pole (Ptacin *et al*., 2010).

In *B. subtilis*, during sporulation, the chromosome is anchored to the cell pole by a system that involves the partition proteins Spo0J and Soj, as well as RacA and DivIVA (Wu and Errington, 2003). In *Vibrio cholerae* ParA1 tethers the partition complex to the pole by interacting with the protein HubP (hub of the pole) (Yamaichi *et al*., 2012). In *Streptomyces coelicolor* the protein Scy interacts with ParA and controls its polymerisation state (Ditkowski *et al*., 2013).
In general, chromosome-encoded ParA and ParB show a wider interaction network, compared to plasmid-encoded partition proteins. Chromosome segregation needs to be tightly coordinated to cell division, while plasmids can be segregated more than once during the cell cycle. However, there is still little information regarding the existence of host-factors taking part in plasmid segregation and this is what we tried to address in this chapter.

Tandem affinity purification (TAP) is a purification technique used to study protein-protein interactions in near-to-physiological conditions. TAP has proved to be a successful technique to identify protein-protein interactions in yeasts, mammalian cells, plants and bacteria (Gloeckner *et al*., 2007; Gully *et al*., 2003; Rohila *et al*., 2004; Puig *et al.*, 2001). We tried to identify binding partners of ParF and ParG, in near to native conditions, by combining TAP and liquid chromatography-tandem mass spectrometry (LC-MS/MS). The proteins identified were then further characterised *in vivo*, using a bacterial two-hybrid system and plasmid stability assay in *E. coli* deletion strains of the proteins identified through TAP-mass spectrometry (TAP-MS).

The original TAP tag was designed with a N-terminal calmodulin-binding peptide and a C-terminal IgG-binding protein A (ProtA from *Staphylococcus aureus*), separated by a TEV (Tobacco Etch Virus) protease and fused to the C-terminus of the protein of interest (Rigaut *et al.*, 1999). Since the TAP-tag comprises two different tags, which bind to two different matrices, tandem affinity purification has the advantage of reducing the noise caused by non-specific binding (Gloeckner *et al*., 2007; Collins and Choudhary, 2008; Gully *et al*., 2003). Although tagged proteins were shown to be functional in yeast, the large molecular weight of the tag (21 kDa) could represent a limitation in bacterial systems, where proteins are in general smaller. In fact, a large TAP tag could impair the protein function as well as lead to false positives or negatives. Since the original tag was designed, different TAP-tags have been engineered and used in different applications (e.g. GS-tag (Bürckstümmer *et al*., 2006), TAPi-tag (Rohila *et al.*, 2004)).
For this work, we used a StrepII/FLAG tag (SF-tag), constructed by the Gloeckner group (Gloeckner *et al.*, 2007, 2009). The SF-tag contains two StrepII moieties and a C-terminal

FLAG peptide (Figure 5.2). Both StrepII and FLAG moieties have medium affinity for their immobilized binding partner and can be recovered by competition under mild conditions and without needing proteolytic cleavage. Furthermore, the tag is quite small (4.6 kDa) and has a hydrophilic terminus, preventing the tag from folding inside the protein (Gloeckner *et al*., 2007). In our experiments, the SF-tag was fused to the C-terminus of ParF, ParG and AspA. AspA was used as a negative binding control, to identify non-specific binding events. AspA is a small (11 kDa) DNA binding protein implicated in plasmid segregation in Archaea, with no homology to bacterial proteins (Schumacher *et al*., 2015). Due to its features, similar size and DNA binding ability to the proteins under investigation, the protein was considered a suitable control for the experiment. TAP-MS gave new insights into the interaction networks of ParF and ParG.

5.2 Results

5.2.1 Construction of pBAD-parF-SF, pBAD-parG-SF and pBAD-aspA-SF for tandem affinity purification

In order to produce recombinant proteins for TAP, *parF*, *parG* and *aspA* were cloned in a pBAD24 vector fused to a DNA sequence encoding the SF-tag at the 3' end of the genes. The pBAD vector contains the arabinose-inducible promoter upstream of the multicloning site (MCS) and has been selected because expression can be tightly regulated according to the arabinose concentration (Guzman *et al*., 1995). Cloning of *parF* in frame with the sequence encoding the SF-tag was performed in two steps as described in Section 2.4.2. First *parF* was amplified by PCR from the plasmid pFH553 (Hayes, 2000). Primers were designed containing the XbaI restriction site upstream of the gene and the PstI restriction site downstream (Table 2.8). PCR amplification produced a fragment of 621 bp (Figure 5.3 A). Both PCR product and pBAD24 vector were digested with XbaI and PstI (Figure 5.3 B). The digested vector and PCR insert were then ligated together as described in Section 2.3.8. Putative positive clones were analysed by extracting the plasmid and subjecting it to restriction digestion (Figure 5.3 C). Plasmid inserts were verified by DNA sequencing.

Once the pBAD-parF plasmid was constructed and checked by Sanger sequencing, the sequence encoding the SF-tag was cloned downstream of *parF* (Figure 5.4). The sequence was amplified from the C-SF-TAP pcDNA3 plasmid obtained from the Gloeckner lab (Gloeckner *et al*., 2007), using primers which harboured the PstI (forward) and the HindIII (reverse) restriction sites (Figure 5.4 A). The PCR product was purified and digested with PstI and HindIII restriction enzymes. The pBAD-parF vector was digested using the same enzymes and dephosphorylated (Figure 5.4 B). The region encoding the tag was then ligated into the pBAD-parF vector. Clones were screened as

described above and positive clones were sent for sequencing (Figure 5.4 C). After plasmid construction, the NcoI restriction site was excluded from the plasmids MCS to prevent transcription from the ATG inside its sequence. The plasmids harbouring *parG* and *aspA* in frame with the region encoding the SF-tag were constructed by swapping *parF* with either *parG* or *aspA* (Figure 5.5). These genes were amplified by PCR, using primers harbouring XbaI (forward) and PstI (reverse) restriction site, producing fragments of 228 and 282 bp respectively (Figure 5.5 A and B). pBAD-parF-SF and PCR products were digested with the restriction enzymes XbaI and PstI. Vector and PCR inserts were ligated as described in Section 2.3.8. Clones were screened by restriction digestion (Figure 5.5 C and D) and positive clones were sequenced. Maps of the plasmids are shown in Figure 5.1.

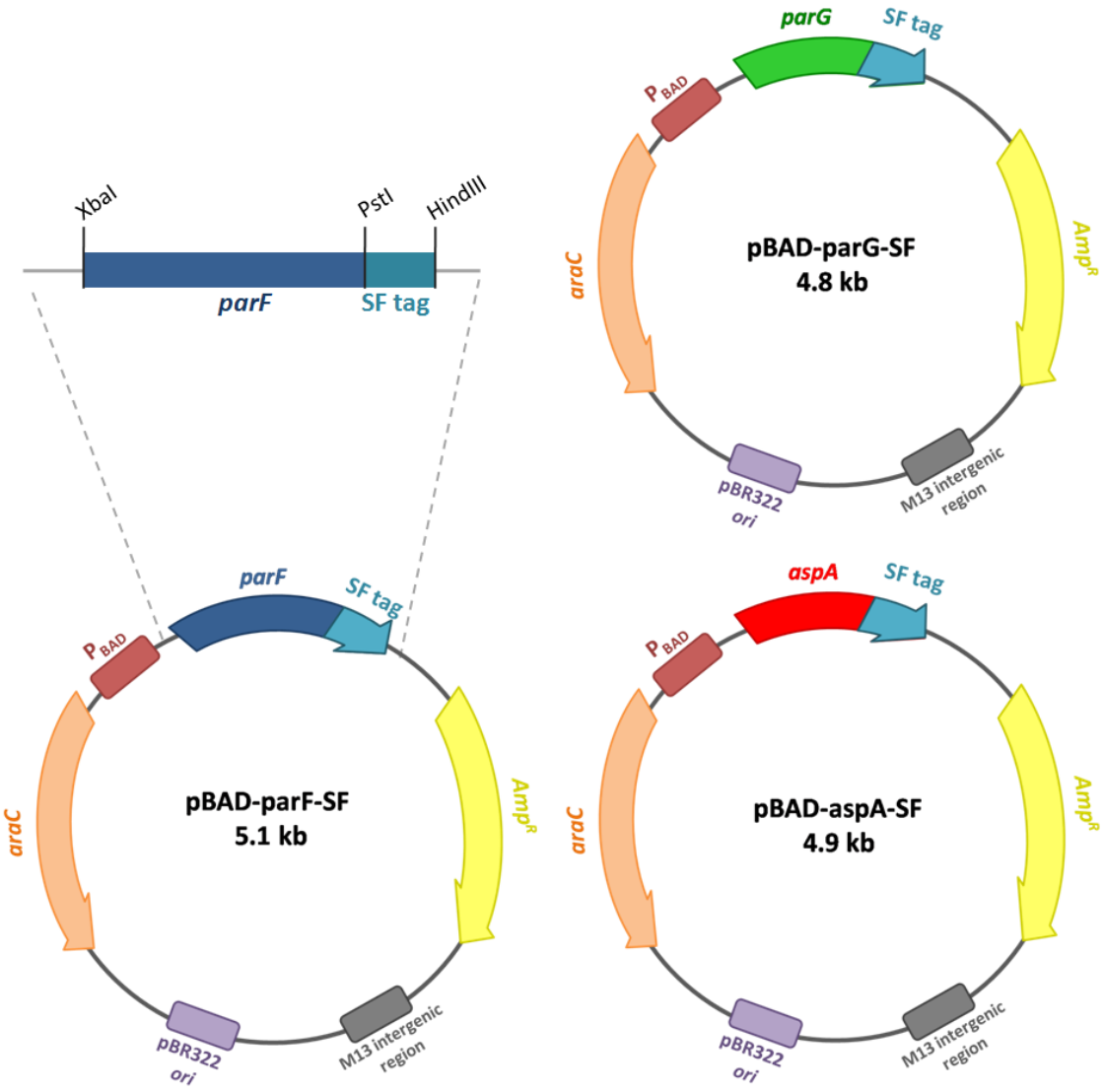

Figure 5.1 –Maps of plasmid pBAD-parF-SF, pBAD-parG-SF and pBAD-aspA-SF

The pBAD-parF-SF plasmid harbours *parF* in frame with the region encoding the SF-tag, downstream of the P_{BAD} promoter of the *araBAD* operon. The plasmid also contains the ampicillin resistance gene (*Amp*^{R}) and pBR322 origin of replication for medium copy number (15-20 copies per cell). The plasmids pBAD-parG-SF and pBAD-aspA-SF were constructed by swapping *parF* with *parG* and *aspA.*

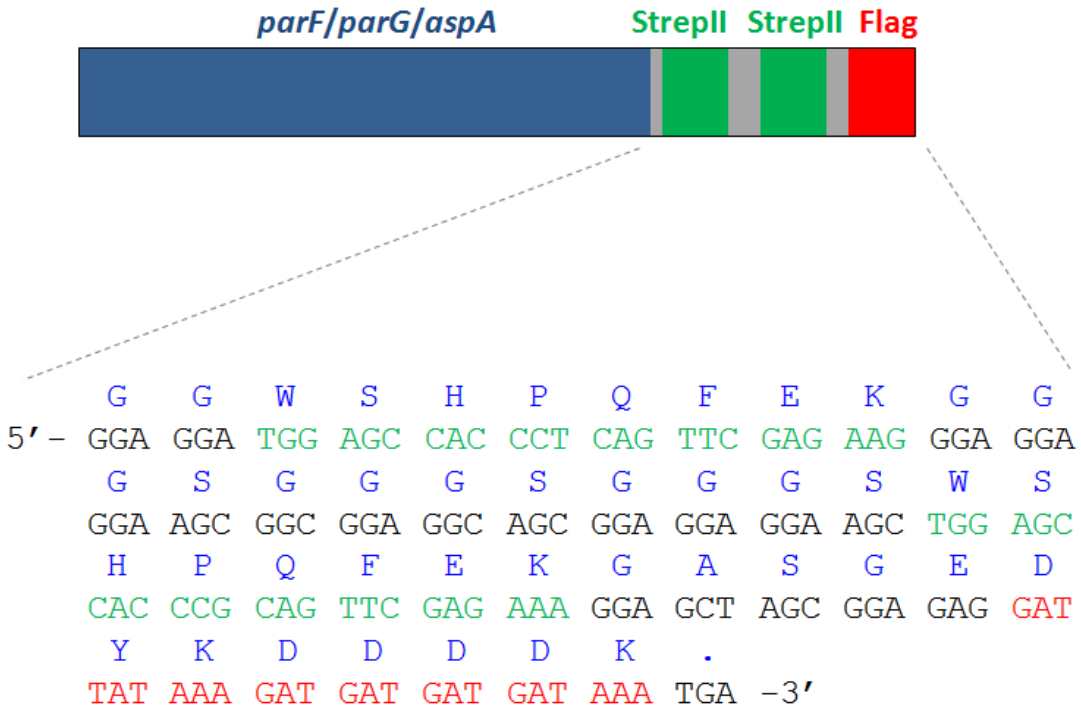

Figure 5.2 – Map of the SF-tagged gene constructs and SF-tag sequence

The genes *parF*, *parG* and *aspA* are cloned in frame with a sequence encoding a C-terminal StrepII/Flag Tag. The SF-tag is composed of two StrepII moieties (green) and one C-terminal Flag peptide (red). The sequence in blue represents the amino acid sequence of the SF-tag. Two glycine codons were added between the last codon of the protein and the starting codon of the StrepII sequence.

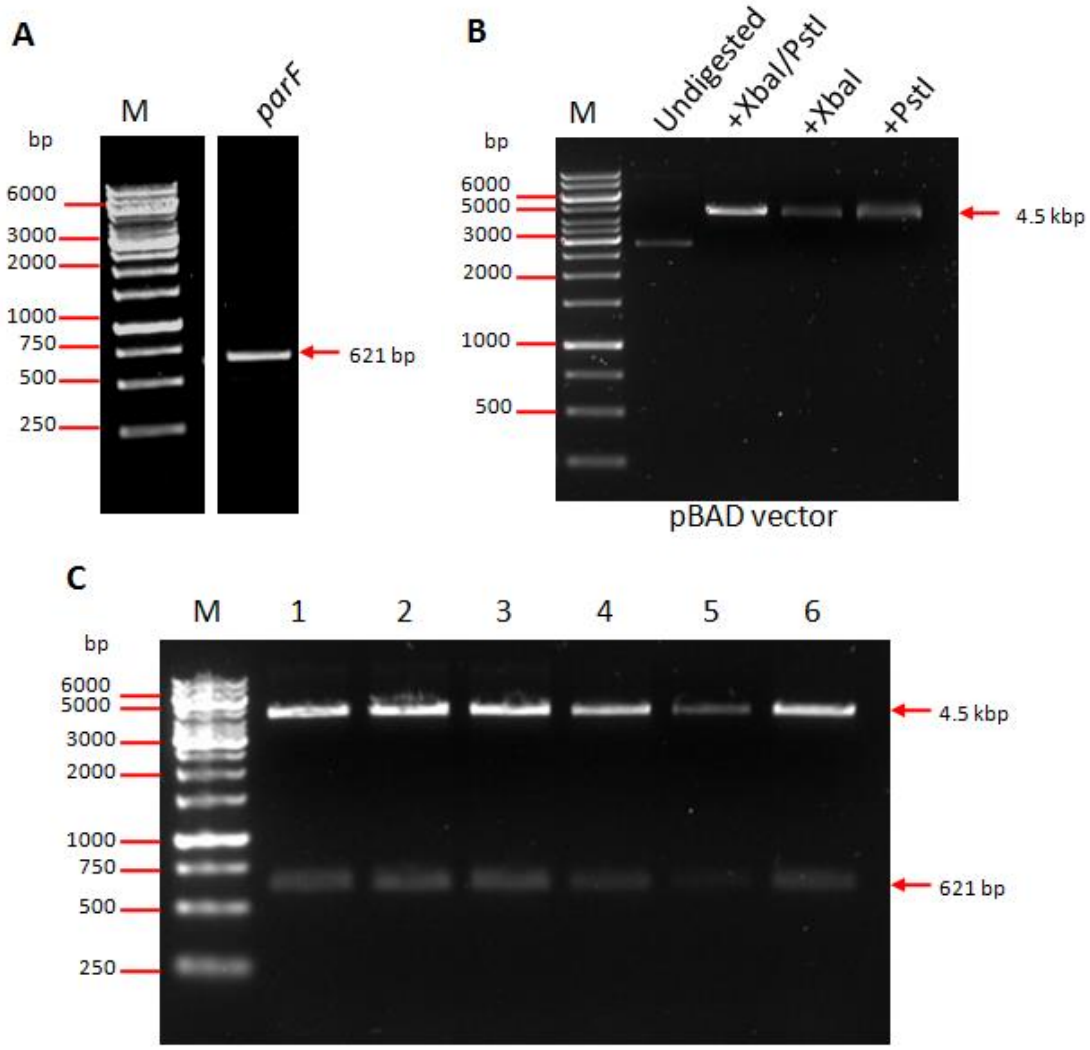

Figure 5.3 – Construction of pBAD-parF plasmid

A) *parF* amplification by PCR. Lane 1: GeneRuler 1 kb MW marker (Thermo Fisher Scientific), lane 2: parF PCR product (621 bp). B) pBAD24 restriction digestion. Lane 1: GeneRuler 1 kb MW marker (Thermo Fisher Scientific), lane 2: undigested pBAD24 plasmid (4.5 kb), lane 3: pBAD24 vector double digestion by restriction endonucleases XbaI and PstI, lane 4-5: XbaI digested and PstI digested vectors respectively C) pBAD-parF plasmid screening by restriction digestion. Lane 1: GeneRuler 1 kb MW marker (Thermo Fisher Scientific), lane 1-6: pBAD-parF recombinant plasmids were digested with XbaI and PstI.

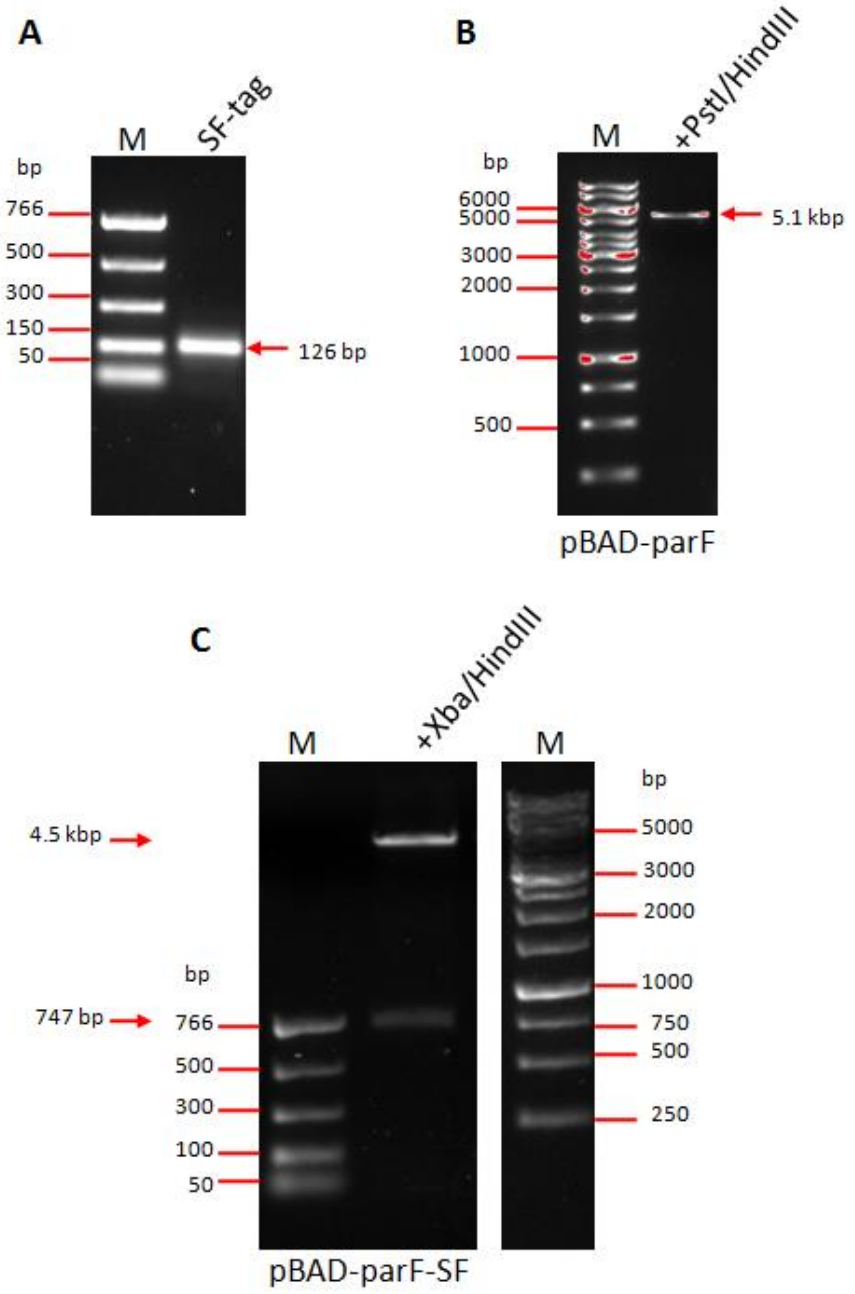

Figure 5.4 – Construction of pBAD-parF-SF plasmid

A) PCR amplification of the region encoding the SF tag. Lane 1: PCR marker (NEB), lane 2: SF-tag (126 bp) PCR product. B) pBAD-parF restriction digestion. Lane 1: GeneRuler 1 kb MW marker (Thermo Fisher Scientific), lane 2: double digestion of pBAD24-parF vector (5.1 kb) by restriction endonucleases PstI and HindIII. C) pBAD-parF-SF plasmid screening by restriction digestion. Lane 1: PCR marker (NEB), lane 3: GeneRuler 1 kb MW marker (Thermo Fisher Scientific). Lane 2: pBAD-parF-SF recombinant plasmid digested with XbaI and HindIII to check for the presence of the insert (*parF*-SF).

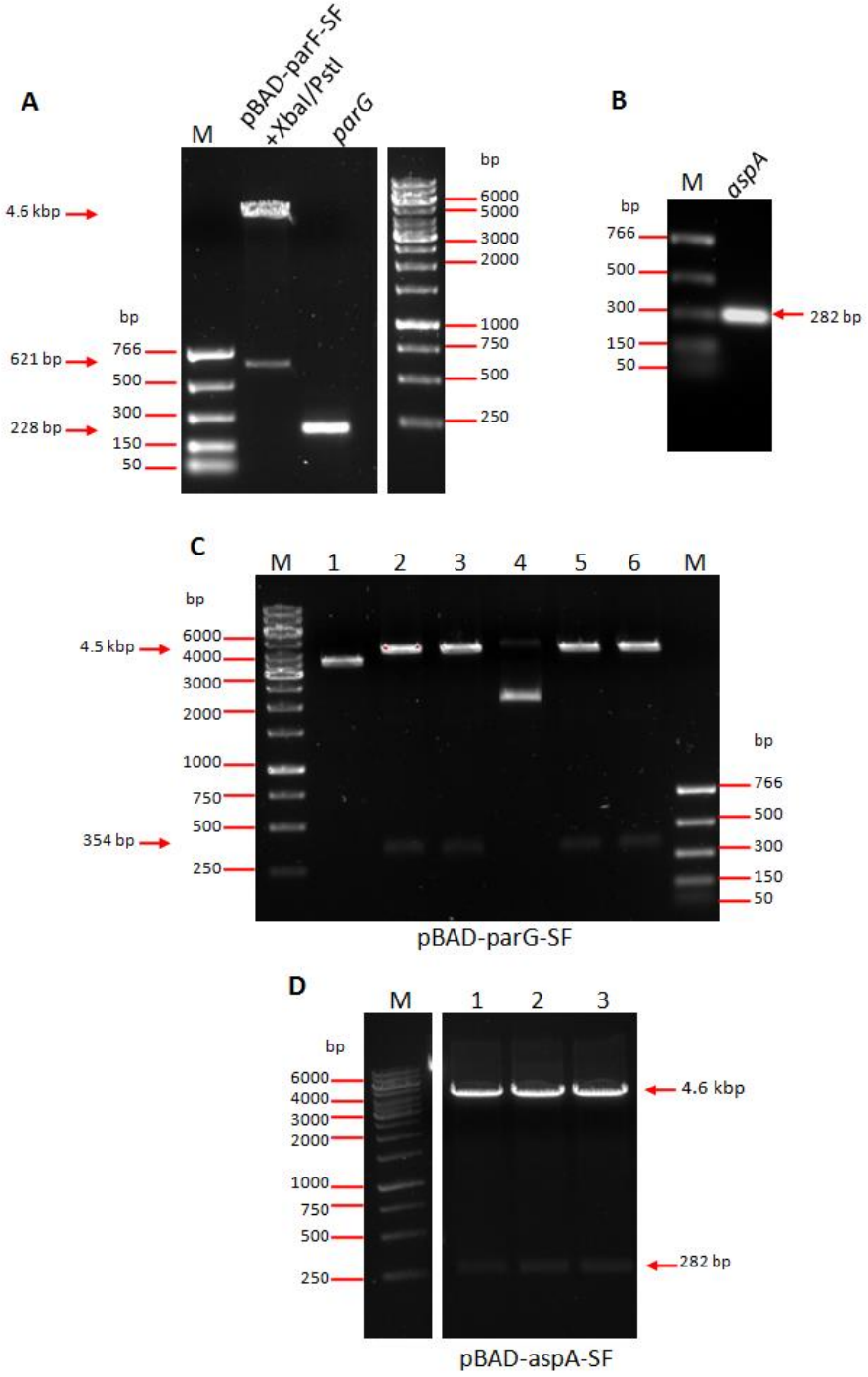

Figure 5.5 – Construction of pBAD-parG-SF and pBAD-aspA-SF plasmids

A) pBAD-parF-SF restriction digestion and *parG* PCR amplification. M: PCR marker (NEB), lane 2: pBAD-parF-SF double digestion by XbaI and PstI, lane 3: *parG* PCR product, lane 4: 1kb MW marker. B) *aspA* PCR amplification. Lane 1: PCR marker (NEB), lane 2: *aspA* PCR product. C) pBAD-parG-SF plasmid screening by restriction digestion with XbaI and HindIII. Lane 1: GeneRuler 1 kb MW marker (Thermo Fisher Scientific), lane 8: PCR marker (NEB). Lane 2-7: pBAD-parG-SF plasmid digested with XbaI and HindIII. Lane 3, 4, 6 and 7 show two bands at 4.5 kb (vector) and at 354 bp (*parG-SF*). D) pBAD-aspA-SF plasmid screening by restriction digestion with XbaI and PstI.

5.2.2 Expression test of *parF-SF, parG-SF* and *aspA-SF*

Expression of the genes fused to the region encoding the *SF-tag* was tested using increasing concentration of L-arabinose. *E. coli* BW25113 cells were transformed with the recombinant plasmids pBAD-parF-SF, pBAD-parG-SF and pBAD-aspA-SF. Transformants were inoculated into selective LB medium and induced with three concentrations of arabinose (Materials and Methods, section 2.9.1). For the genes encoding ParF-SF and ParG-SF, expression was induced at 37 °C, using 0.001%, 0.01% and 0.025% (w/v) arabinose concentration and checked after 1, 2 and 3 hours (Figure 5.6). Samples were run on 12% SDS-PAGE. Expression level did not change drastically with increasing concentration of the inducer and seemed to remain stable after the second hour from addition of arabinose. Based on these results, 0.01% arabinose concentration, 37 °C and 3 hours induction time were selected as suitable over-expression conditions. The expression of the gene encoding AspA-SF was tested with conditions selected for the production of ParF-SF and ParG-SF (Figure 5.7) and comparable expression level was observed.

As tags can be proteolysed, integrity of the overproduced proteins was tested by Western blot. *E. coli* cells overexpressing the tagged genes were run on 12% SDS-PAG and transferred onto a nitrocellulose membrane to be immunodetected using antibody specific for the proteins (α-ParF, α-ParG and α-AspA) and for the Flag peptide (α-Flag) (Figure 5.8). An aliquot of purified ParF and ParG was run in parallel to the ParF-SF and ParG-SF overproduction test to prove that the tagged proteins were not proteolysed. While the antibodies specific for the proteins mostly recognised the protein of interest, the α-Flag antibody could recognise several proteins larger than 26 kDa. This is likely due to the fact that Flag is a small peptide and may be present in other cellular proteins. Detection by α-ParF and α-ParG shows that ParF-SF and ParG-SF are, as expected, larger than purified ParF-His and ParG-His. Furthermore, α-Flag cannot recognise untagged ParF and ParG, suggesting that the overexpressed SF-tagged proteins are intact.

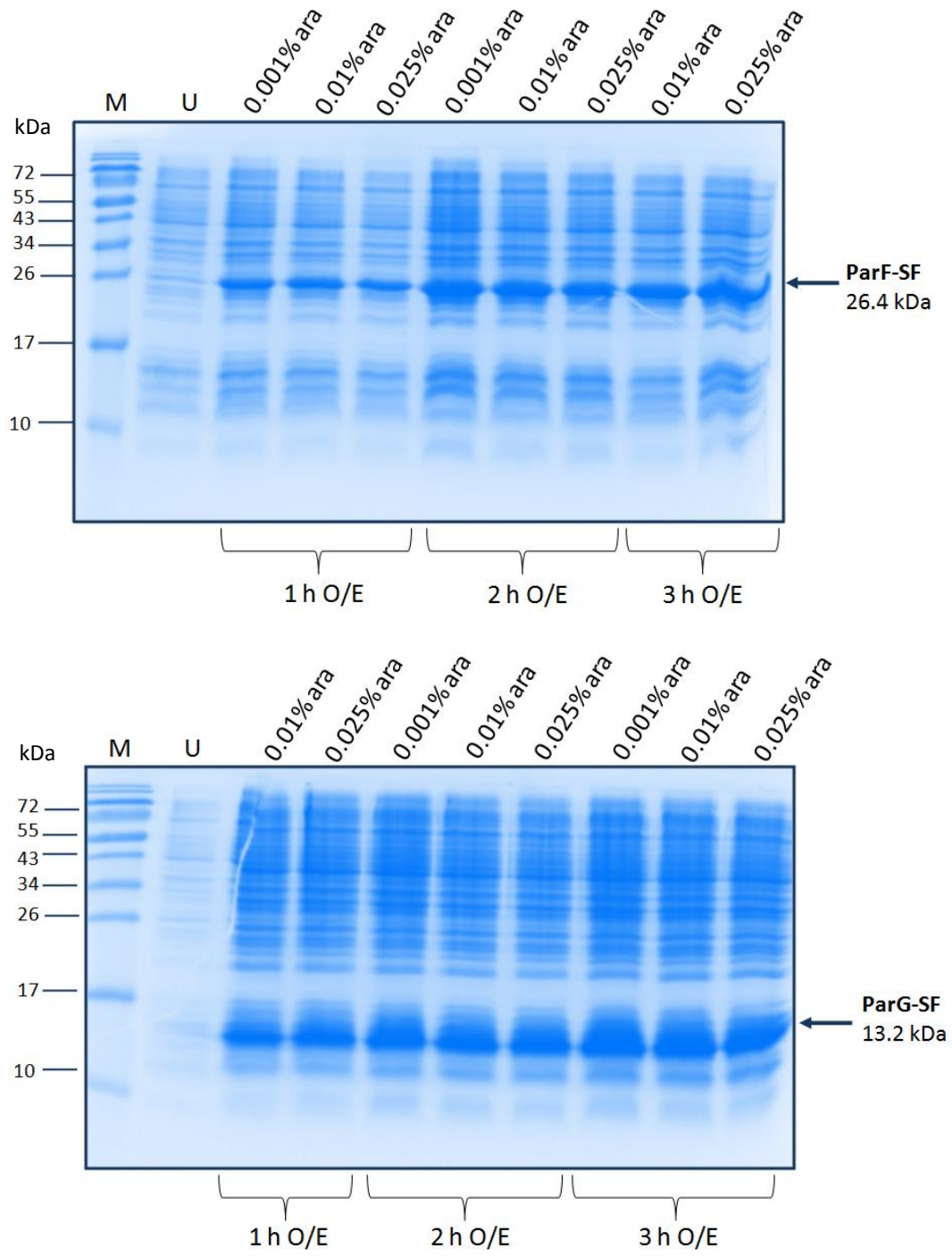

Figure 5.6 – Overproduction test of ParF-SF and ParG-SF

Overproduction of ParF-SF (top) and ParG-SF (bottom) was tested by induction with 0.001, 0.01 and 0.025% of L-arabinose for 1, 2 and 3 hours. Lane 1 represents the PageRuler Prestained protein ladder (Thermo Fisher). The arrow indicates ParF (26.4 kDa) and ParG (13.2 kDa) band.

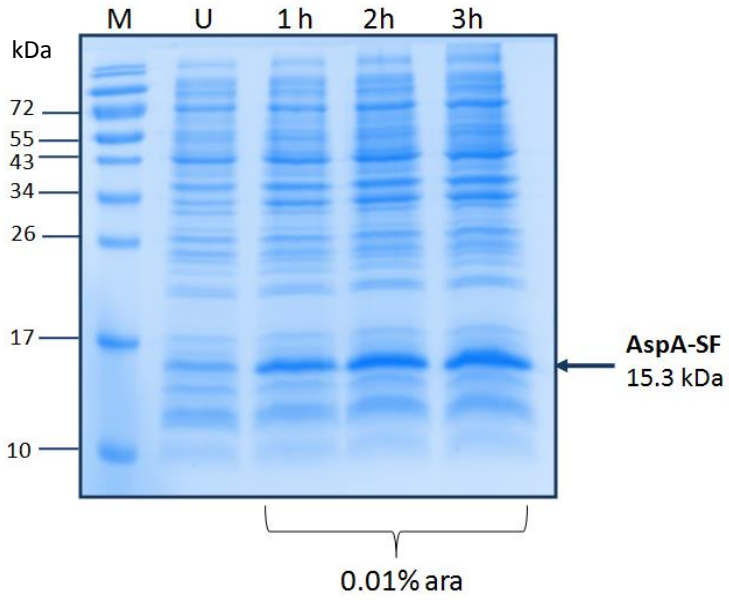

Figure 5.7 – Overproduction test of AspA-SF

Overproduction of AspA was tested by induction with 0.01% L-arabinose for 1, 2 and 3 hours. Lane 1 represents the PageRuler Prestained protein ladder (Thermo Fisher). The arrow indicates AspA (15.3 kDa) band.

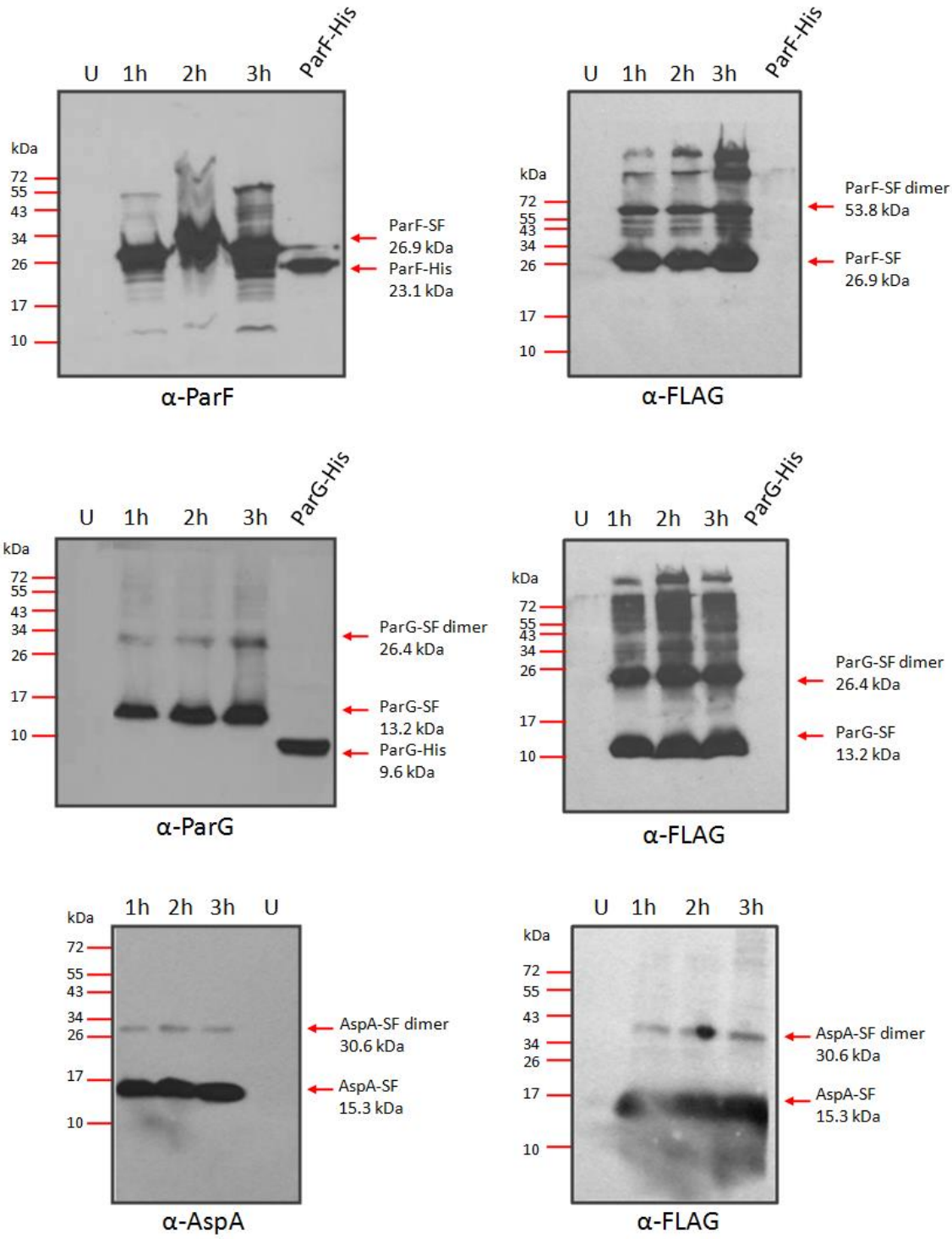

Figure 5.8 – Western blot with α-protein and α-FLAG to show the integrity of the SF-tagged proteins in *E. coli* cells

Overproduction tests for ParF-SF, ParG-SF and AspA-SF run on 12% SDS-PAG and detected by Western blot with α-ParF, α-ParG, α-AspA and α-Flag antibodies.

5.2.3 Tandem affinity purification coupled to Liquid chromatography-Tandem mass spectrometry

Tandem affinity purification aimed at identifying proteins interacting with ParF and ParG, by co-purification of the bait protein (ParF and ParG) with the binding partners in a two-steps process. To exclude factors pulled-down randomly by non-specific interaction with the SF-tag or the resins used for the purification steps, two controls were run in parallel to the experiments. AspA is a small protein found in archaea species, which has no homology to bacterial proteins. Therefore, TAP experiment on AspA-SF was performed to exclude non-specific binding events. A second control consisted in performing TAP on the crude extract of *E. coli* cells that did not contain any expression vector. This control allowed to discard proteins pulled-down by non-specific interaction with the resins. The initial idea was to perform TAP on the crude extract of *E. coli* cells producing the SF-tag peptide alone (4.6 kDa). However, it was impossible to test its overexpression in *E. coli* cells. Several attempts were made, including running the over-expression product on tricine SDS-PAGE for better separation of low molecular weight peptides. Nevertheless, all attempts failed and the peptide was never visible by Coomassie gel staining or Western blot.

In addition, to prove that the bait proteins could reproducibly pull-down the same partners, each experiment was repeated at least in triplicate (ParF TAP was repeated four times).

TAP was performed as described in section 2.9 (Materials and Methods). BW25113 cells were transformed with the recombinant pBAD-parF-SF, pBAD-parG-SF and pBAD-aspA-SF plasmids. Transformant cells were inoculated into 1L selective LB, and gene expression was induced with 0.01% arabinose for 3 hours at 37°C. Importantly, the level of expression from the pBAD24 vector is lower compared to other expression plasmids (e.g. pET-22). This vector was chosen to obtain an expression level that would not completely imbalance the native conditions in the cell.

Cells were harvested by centrifugation and the crude extract was prepared as detailed in section 2.9.3. Since the bait proteins are DNA-binding proteins, the crude extract was treated with DNaseI to prevent pulling down proteins interacting with the DNA, rather than the proteins of interest. The crude extract was then loaded onto the Strep-Tactin column and flowed over it several times for optimal protein binding. The column was then washed to remove non-specific binding partners before being eluted by competition with desthiobiotin (Sections 2.9.4). Elution fractions were pooled together and loaded onto a second column containing the α-Flag M2 agarose. The resin was incubated with the proteins for two hours, then washed and the bait protein was eluted by competition with the Flag peptide (as described in Section 2.9.5, Material and Methods). The elution

fractions were concentrated and an aliquot of 100 µg was shortly run on a precast SDS-polyacrylamide gel before excising the non-resolved band for mass-spec analysis. A second 100 µg aliquot was loaded on a separate precast gel and electrophoresis was run to visualise the TAP product.

5.2.3.1 TAP-MS/MS with ParF-SF bait allowed the identification of a number of putative binding partners

SDS-PAGE of the samples collected during the two steps of purification (Figure 5.9) showed that a high amount of ParF-SF protein was lost either in the flow through of the columns or in the washes. Different volumes of the resins and buffer conditions were tested without improving protein binding to Strep-Tactin and to the α-Flag agarose. This could be due to the fact that both StrepII and Flag moieties have medium affinity for their matrix, therefore the tagged protein did not bind to the resins tightly and was washed away. Another possibility is that the capacity of the columns was not sufficient for the volume of culture used (1 L). However, when a smaller culture was used (0.5 L), a high amount of protein was still lost during the purification steps and a low amount of protein could be recovered together with the bait protein (data not shown).

The conditions described in the methods section allowed the co-purification of ParF with several putative binding partners. SDS-PAGE shows that the elution fraction from the second purification step presents multiple proteins, visible in all four repeats (Figure 5.10 and pointed by the red arrows in Figure 5.9 B). The sample analysed by mass spectrometry was not resolved on the gel and due to the number of proteins identified, it is not possible to know which protein the bands correspond to.

LC-MS/MS analysis led to the identification of a total of 537 proteins. Many of those were excluded because they were pulled-down randomly. This means that they were either identified in the control samples or not consistently recovered in the four repeats of the experiment. In addition, many of the proteins identified were ribosomal subunits, chaperons, elongation factors or proteases. Those were also excluded because they are involved in the process of protein production and degradation.

Among the proteins consistently identified only in ParF samples, it was possible to classify many cell division proteins, outer membrane proteins, several enzymes involved in metabolic processes, outer membrane proteins and transporters (Table 5.1). It is very likely that many of these elements were pulled down by the interaction with other proteins, possibly interacting specifically with the bait ParF. However, it was impossible to identify which of those factors interacted directly with ParF and which indirectly. Since ParF is a cytoplasmic protein, it is probable that outer membrane proteins (Omp)

identified did not bind directly to ParF but were rather pulled down by other factors present in the pool. A number of proteins identified exclusively in ParF samples were further investigated *in vivo*. These were selected according to consistence and abundance, as well as their function in the cell. Among ParF's putative binding partner, it was interesting to find FtsZ, MukB and MreB, which were tested in the past for involvement in plasmid segregation. GyrB was also selected for additional investigation, as we speculated it could change plasmid supercoiling state and somehow take part in plasmid segregation. Other proteins selected were: DamX, OmpC, OmpX, YdgA and PppiD. Further experiments will be discussed in the next sections.

Similarly to the TAP experiments with ParF-SF, TAP using ParG-SF as bait protein showed that a high amount of the protein was lost in the flow through of the columns or in the washing steps. From the SDS-polyacrylamide gels it appeared that the protein was in general cleaner than ParF-SF, suggesting that a limited number of proteins were interacting with ParG. Elution fractions from the three repeats show that only a few proteins are constantly retained at the end of the purification (pointed by the red arrows in figure 5.11 and shown in 5.12). Mass-spectrometry analysis showed that none of the proteins pulled down with ParG-SF were relevant. Many of the proteins identified were chaperones, ribosomal subunits, elongation factors or were found in the control samples. Proteins identified in both ParF and ParG samples were not found consistently in all the repeats, were involved in protein metabolism or did not seem relevant.

TAP experiment using AspA-SF as bait protein allowed the identification of proteins interacting non-specifically with ParF and ParG. The protein pulled-down a number of abundant cytosolic proteins that were also identified in the other experiments, together with ribosomal subunits, chaperones, elongation factors and proteases. Although many proteins were pulled down together with AspA, they were not abundant and almost no protein band can be seen in the elution fraction of the second step of TAP on SDS-PAGE (Figure 5.13).

TAP on the crude extract of BW25113 *E. coli* cells, with no bait protein being produced, showed that most of the proteins are lost in the flow through and the wash of the first column. Little protein was retained at the end of the two steps of purification in all three repeats of the experiment and no protein can be seen on the SDS-PAGE of the TAPs (Figure 5.14). LC-MS/MS allowed identification of a few proteins that were clearly interacting with the resins, rather than the proteins of interest.

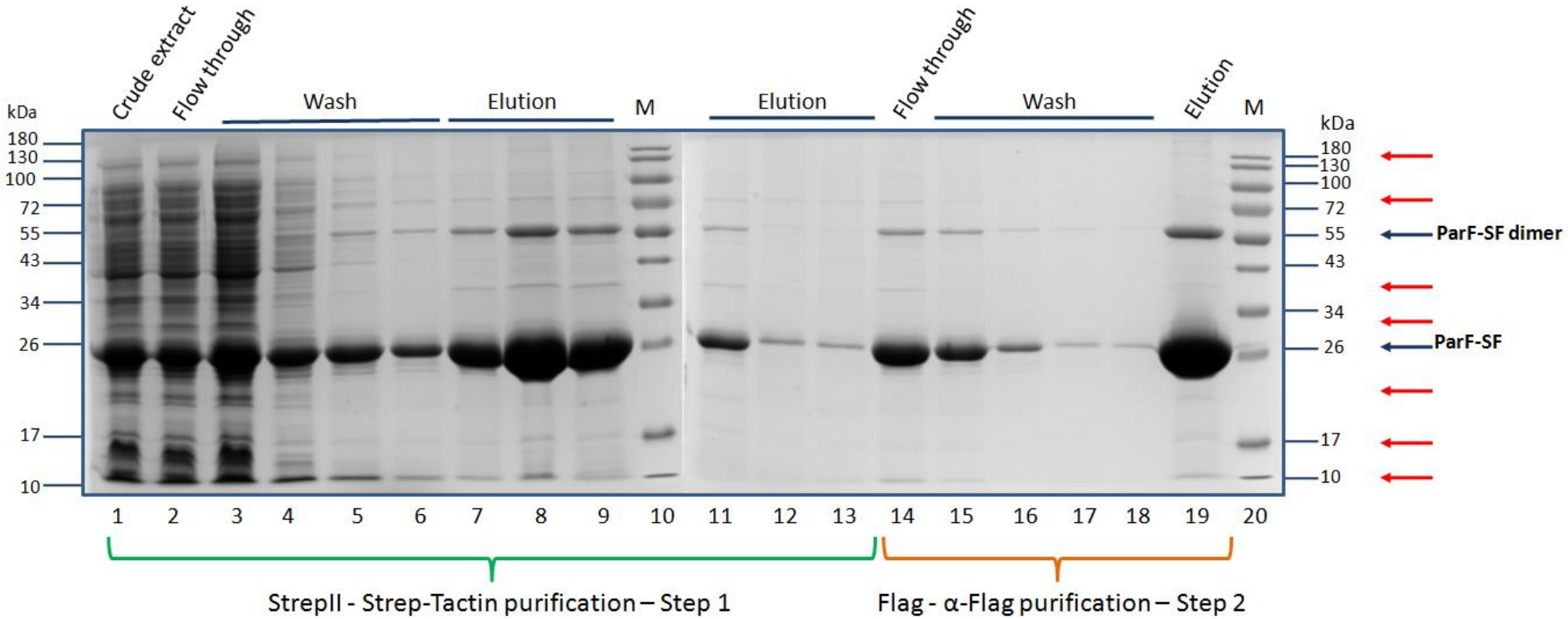

Figure 5.9 – SDS-polyacrylamide gels showing the tandem affinity purification of ParF-SF

Lanes 1 to 13 show fractions from the first step, StrepII - Strep-Tactin chromatography, of ParF-SF TAP. Lane 1: crude extract of ParF-SF-overproducing culture; lane 2: Strep-Tactin column flow through; lanes 3-6: Fractions from the wash step with Buffer WS; lanes 7-9 and 11-13: Fractions of the elution of the Strep-Tactin column. Lanes 14 to 19 show the second step, Flag – α-Flag chromatography, of ParF-SF TAP. Lane 14: Flow through of the α-Flag agarose column; lanes 15-18: fractions from the wash with TBS Buffer; lane 19: concentrated elution fraction of the α-Flag column. Lanes 10 and 20 show the PageRuler Prestained protein ladder (Thermo Fisher). The blue arrows represent ParF and ParF dimer, while the red arrows indicate bands identified in the final elution step of this and the other three repeats of TAP (Figure 5.10).

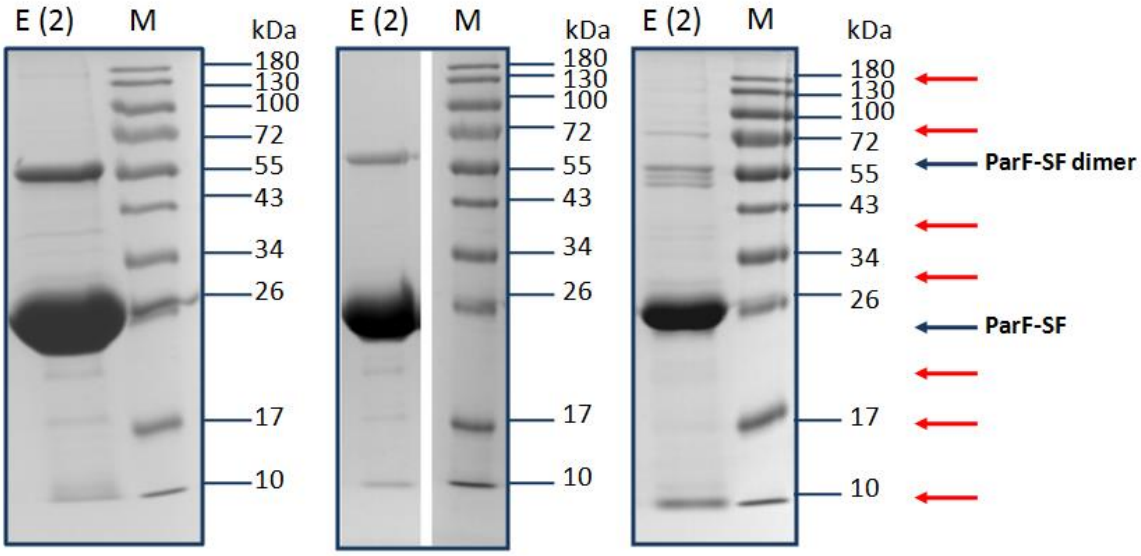

Figure 5.10 – SDS-polyacrylamide gels showing the elution fraction from the second chromatography of ParF-SF TAP for three repeats

Elution fractions from the second step of purification of ParF-SF TAP (three repeats). The elution fractions were 40 times concentrated. The fractions shown here contain 100 µg of protein. The thick band at around 26 kDa is ParF-SF. The marker used is PageRuler Prestained protein ladder (Thermo Fisher). The blue arrows represent ParF and ParF dimer, while the red arrows indicate bands identified in the final elution step the three repeats of TAP.

Protein ID	Protein Characteristics	Accession Number	Molecular Weight	Total unique peptide count			
Cell Division proteins				ParF-1	ParF-2	ParF-3	ParF-4
Ftsz	**Cell division protein**	**FTSZ_ECO57 (+2)**	**40 kDa**	**25**	**19**	**23**	**21**
Mreb	**Rod shape-determining protein**	**MREB_ECOL6 (+1)**	**37 kDa**	**13**	**11**	**13**	**16**
Damx	**Cell division protein**	**DAMX_ECOLI**	**46 kDa**	**10**	**8**	**8**	**12**
Mukb	**Chromosome partition protein**	**MUKB_ECO27 [16]**	**170 kDa**	**8**	**4**	**10**	**13**
Mind	Septum site-determining protein	MIND_ECO57 (+2)	30 kDa	7	2	5	9
Ftsa	Cell division protein	FTSA_ECO57 (+2)	45 kDa	4	2	5	9
Zipa	Cell division protein	ZIPA_ECO24	37 kDa	2	1	1	1
Outer membrane proteins							
Ompc	**Outer membrane protein**	**OMPC_ECOLI [4]**	**40 kDa**	**27**	**22**	**26**	**25**
Seca	Protein translocase subunit	SECA_ECO24 (+8)	102 kDa	20	8	19	23
Tolc	Outer membrane protein	TOLC_ECOLI	54 kDa	16	7	16	16
Bama	Outer membrane protein assembly factor	BAMA_ECO24 (+19)	91 kDa	18	5	14	14
Ompf	Outer membrane protein	OMPF_ECOLI	39 kDa	9	7	9	13
Ompx	**Outer membrane protein**	**OMPX_ECO57 (+2)**	**19 kDa**	**7**	**6**	**6**	**9**
Bamd	Outer membrane protein assembly factor	BAMD_ECO57 (+2)	28 kDa	6	2	6	9
Bamc	Outer membrane protein assembly factor	BAMC_ECOLI	37 kDa	4	2	3	5
Tolq	Outer membrane protein	TOLQ_ECO57 (+1)	26 kDa	2	1	1	2
Inner membrane proteins							

Yjiy	Inner membrane protein	YJIY_ECOLI	77 kDa	2	2	3	3
DNA-associated proteins							
Reca	DNA repair and mantainance	RECA_ECO24 (+18)	38 kDa	16	12	17	19
Gyrb	**DNA gyrase subunit**	**GYRB_ECO57 (+1)**	**90 kDa**	**20**	**8**	**15**	**18**
Ybab	Nucleoid-associated protein	YBAB_ECO24 (+19)	12 kDa	3	3	5	4
Transcription factors							
Rcsb	Transcriptional regulatory protein	RCSB_ECO57 (+3)	24 kDa	3	1	4	4
Ompr	Transcriptional regulatory protein	OMPR_ECO57 (+2)	27 kDa	1	1	2	2
Hexr	HTH-type transcriptional regulator	HEXR_ECOLI	32 kDa	2	1	2	3
Uncharacterised proteins							
Ydga		**YDGA_ECOLI**	**55 kDa**	**11**	**3**	**10**	**15**
Ydhq		YDHQ_ECOLI	43 kDa	10	4	7	7
Ykge		YKGE_ECOLI	26 kDa	2	1	2	2
Miscellaneous							
Lpda	Dihydrolipoyl dehydrogenase	DLDH_ECO57 (+2)	51 kDa	28	21	23	24
Ptsi	Phosphoenolpyruvate-protein phosphotransferase	PT1_ECOLI	64 kDa	25	17	21	27
Sdha	Succinate dehydrogenase flavoprotein subunit	SDHA_ECO57 (+2)	64 kDa	19	12	20	26
Glpd	Aerobic glycerol-3-phosphate dehydrogenase	GLPD_ECOLI	57 kDa	26	15	22	27
Frda	Fumarate reductase flavoprotein subunit	FRDA_ECOLI	66 kDa	19	13	16	20
Rne	Ribonuclease E	RNE_ECOLI	118 kDa	25	7	15	23

Ppsa	Phosphoenolpyruvate synthase	PPSA_ECOLI	87 kDa	16	12	17	20
Fdog	Cluster of Formate dehydrogenase-O major subunit	FDOG_ECOLI [2]	113 kDa	16	5	13	16
Tdce	PFL-like enzyme	TDCE_ECOLI	86 kDa	15	6	16	17
Ppid	Peptidyl-prolyl cis-trans isomerase	PPID_ECOL6 (+1)	68 kDa	12	8	13	17
Atpf	ATP synthase subunit	ATPF_ECO24 (+11)	17 kDa	13	7	10	14
Degp	Periplasmic serine endoprotease	DEGP_ECO57 (+1)	49 kDa	12	8	11	14
Nuoc	NADH-quinone oxidoreductase subunit	NUOCD_ECOLI	68 kDa	17	7	11	14
Sdhb	Succinate dehydrogenase iron-sulfur subunit	SDHB_ECOLI	27 kDa	10	7	9	10
Narg	Cluster of Respiratory nitrate reductase 1 alpha chain	NARG_ECOLI [2]	140 kDa	13	5	12	15
Acra	Multidrug efflux pump subunit	ACRA_ECO57 (+1)	42 kDa	10	7	9	11
Ftsh	ATP-dependent zinc metalloprotease	FTSH_ECO57 (+1)	71 kDa	10	5	4	9

Table 5.1 – Proteins identified through TAP – mass spectrometry with ParF bait protein

Proteins identified through TAP-mass spec were filtered by comparison with the controls (AspA-SF and *E. coli* crude extract) and ParG-SF TAP experiments. Relative protein quantification between sample groups was calculated in Scaffold 4.0 using normalised spectral counts. Statistical differences were assessed using Student's t-test with Benjamini-Hochberg multiple test false discovery rate correction applied. Protein identified in low amount were not shown. Proteins shown in bold were further characterised *in vivo*.

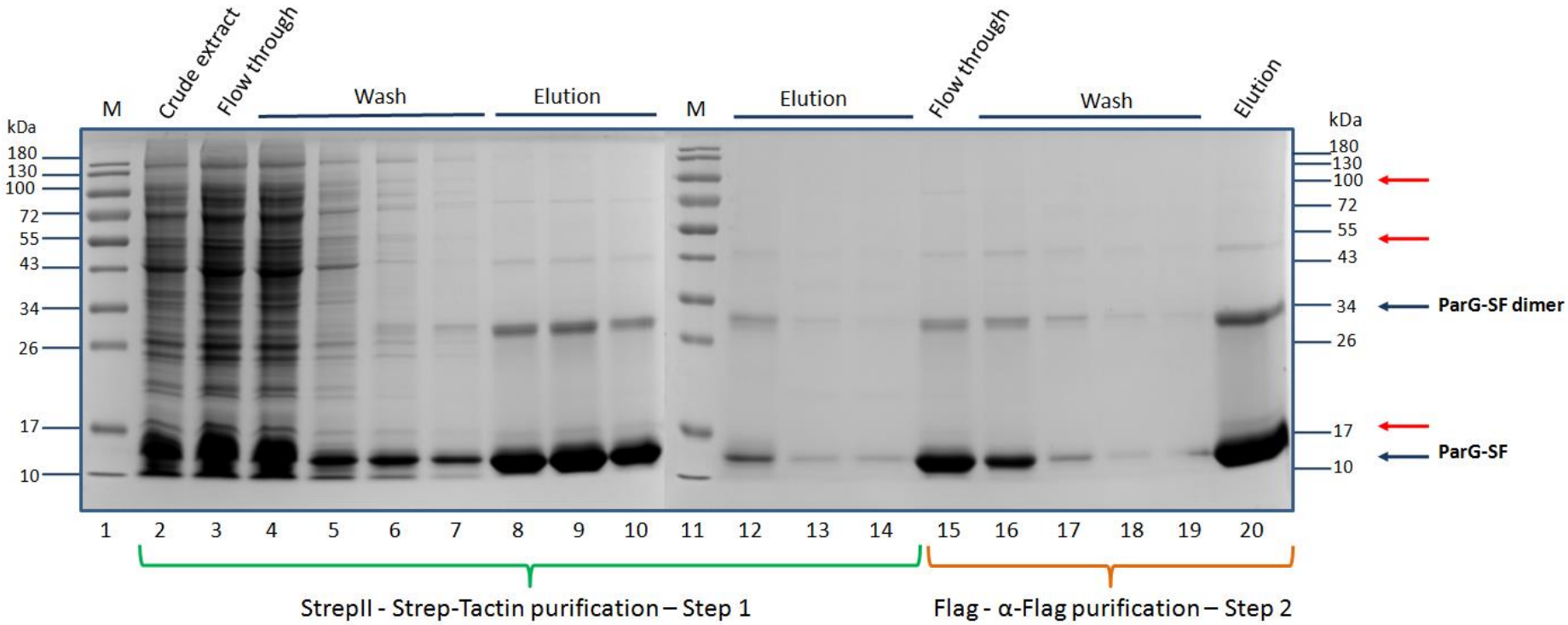

Figure 5.11 – SDS-polyacrylamide gels showing the tandem affinity purification of ParG-SF

Lanes 2 to 14 show fractions from the first step, StrepII - Strep-Tactin chromatography, of ParG-SF TAP. Lane 2: crude extract of ParG-SF-overproducing culture; lane 3: Strep-Tactin column flow through; lanes 4-7: Fractions from the wash step with Buffer WS; lanes 8-10 and 12-14: Fractions of the elution of the Strep-Tactin column. Lanes 15 to 20 show the second step, Flag – α-Flag chromatography, of ParG-SF TAP. Lane 15: Flow through of the α-Flag agarose column; lanes 16-19: fractions from the wash with TBS Buffer; lane 20: concentrated elution fraction of the α-Flag column. Lanes 1 and 11 show the PageRuler Prestained protein ladder (Thermo Fisher). The blue arrows represent ParG and ParG dimer, while the red arrows indicate bands identified in the final elution step of this and the other three repeats of TAP (Figure 5.12).

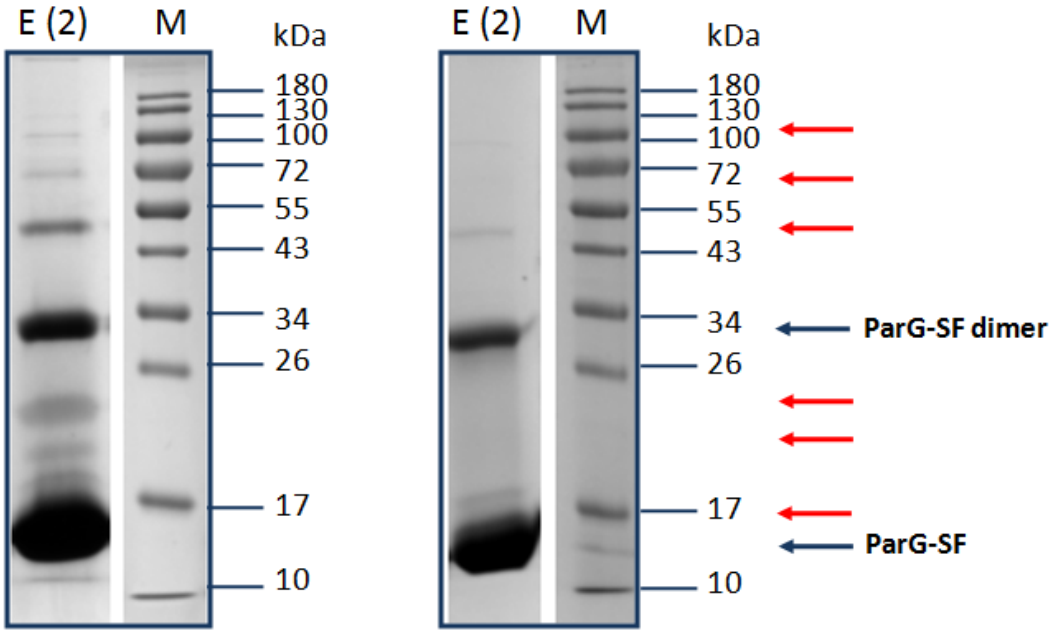

Figure 5.12 – SDS-polyacrylamide gels showing the elution fractions from the second chromatography of ParG-SF TAP for two repeats

Elution fractions from the second step of purification of ParG-SF TAP (two repeats). The elution fractions were 40 times concentrated. The fractions shown here contain 100 µg of protein. The thick band at around 13 kDa is ParG-SF. The marker used is PageRuler Prestained protein ladder (Thermo Fisher). The blue arrows represent ParG and ParG dimer, while the red arrows indicate bands identified in the final elution step the three repeats of TAP.

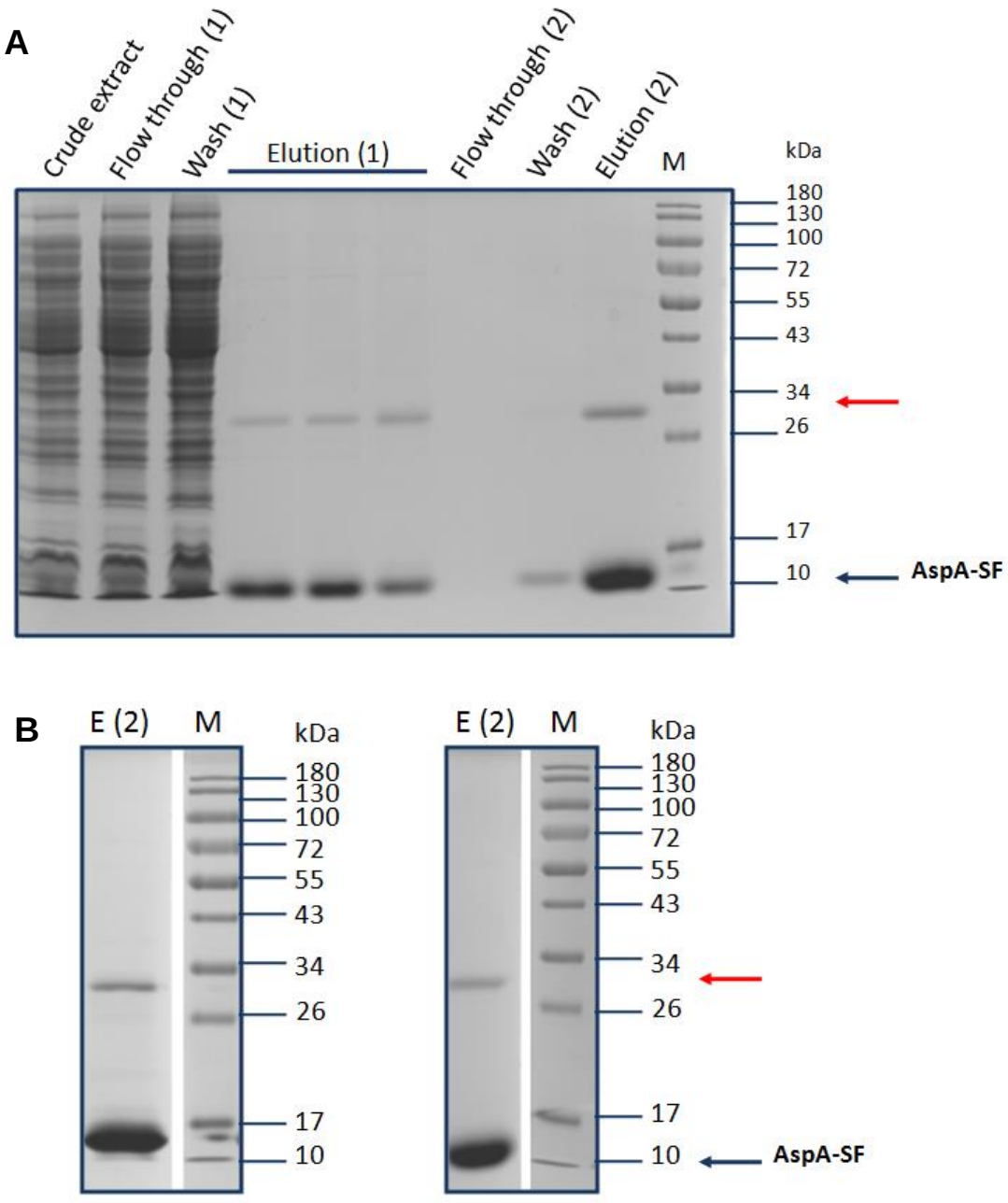

Figure 5.13 – SDS-polyacrylamide gel showing tandem affinity purification of AspA-SF

A) AspA-SF TAP. Lane 1: crude extract of AspA-SF-overproducing culture; lane 2:Strep-Tactin column flow through; lanes 3: First fraction of the wash with Buffer WS (step 1); lanes 4-6: Fractions of the elution with desthiobiotin (step 1); lane 7: flow through of the α-Flag M2 agarose column (step 2); lane 8: Fractionsfrom the wash with TBS Buffer; lane 9: times concentrated elution fraction obtained by competition with the flag peptide; lane 10: PageRuler Prestained protein ladder (Thermo Fisher).

B) Elution fractions from the second step of AspA-SF TAP for two repeats. The elution fractions were concentrated. The fractions shown here contain 100 µg of protein. The thick band at around 15 kDa is AspA-SF. The marker used is PageRuler Prestained protein ladder (Thermo Fisher). The red arrows show proteins identified in the elution of the second chromatography.

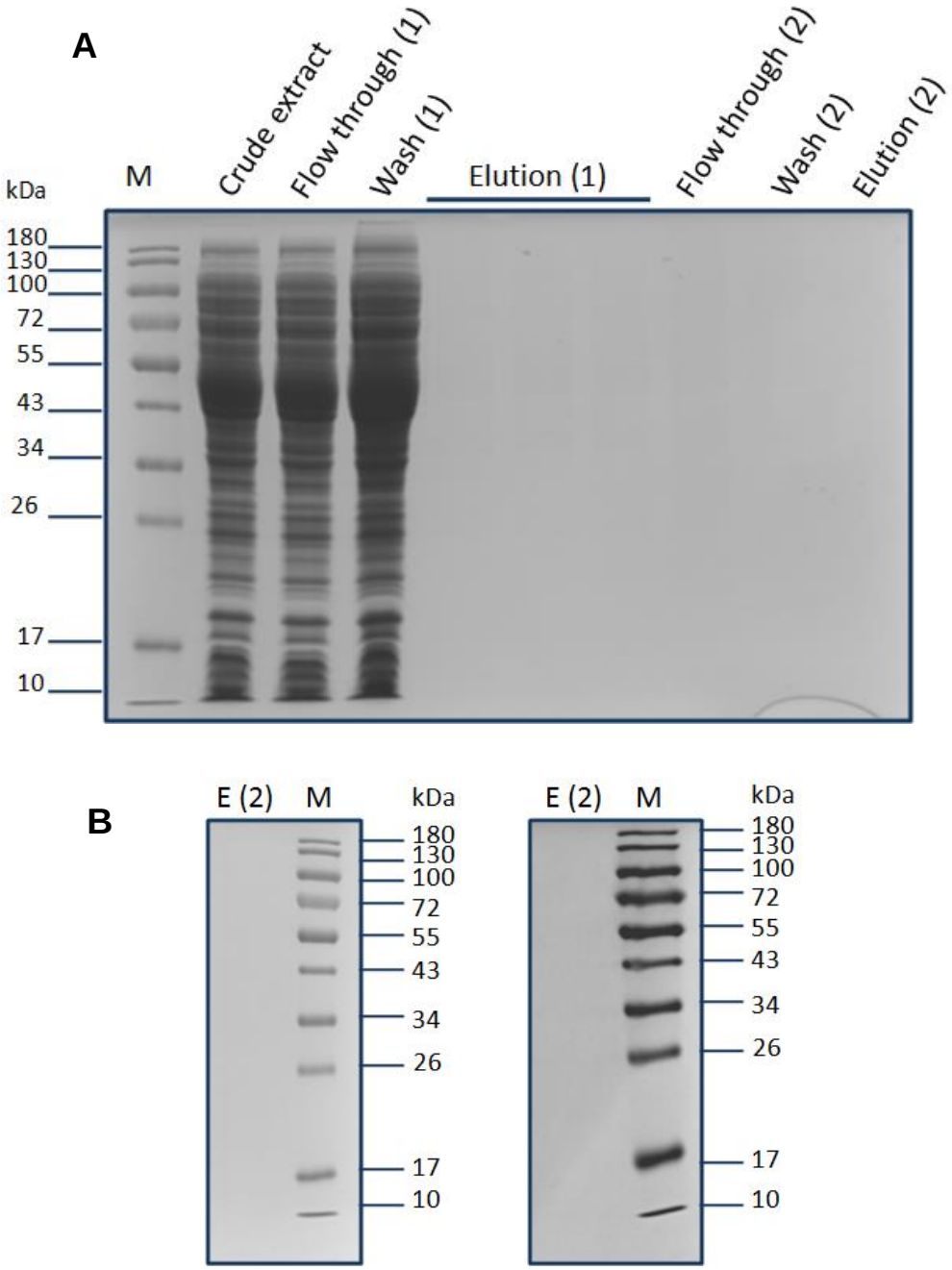

Figure 5.14 – SDS-polyacrylamide gel showing tandem affinity purification of *E. coli* crude extract

A) BW25113 *E.coli* crude extract TAP. Lane 1: crude extract of non-overproducing culture; lane 2: Strep-Tactin column flow through; lanes 3: First fraction of the wash with Buffer WS (step 1); lanes 4-6: Fractions of the elution with desthiobiotin (step 1); lane 7: flow through of the α-Flag M2 agarose column (step 2); lane 8: Fractions from the wash with TBS Buffer; lane 9: concentrated elution fraction obtained by competition with the flag peptide; lane 10: PageRuler Prestained protein ladder (Thermo Fisher). B) Elution fractions from the second step of TAP on the BW25113 *E. coli* crude extract for two repeats. The elution fractions were 40 fold concentrated. The marker used is PageRuler Prestained protein ladder (Thermo Fisher).

5.2.4 Involvement of the proteins identified through TAP-LC-MS/MS in plasmid segregation

Following TAP and LC-MS protein identification, further *in vivo* characterisation of the putative interacting partners was carried out to understand their involvement in plasmid segregation. Because of time constraints, it was essential to select a number of proteins that seemed potentially promising. This was done based on background information of plasmid partition systems and abundance of the protein in the sample analysed by mass-spectrometry. The proteins selected could be grouped into essential proteins and non-essential proteins. The housekeeping proteins selected to be further analysed were: FtsZ, MreB, MukB and GyrB.

FtsZ is a cell division protein, involved in the formation of the Z-ring at midcell for peptidoglycan synthesis and septum formation (Bisson-Filho *et al*., 2017). The protein interacts with FtsA and ZipA (Ma *et al.*, 1996; Stricker *et al.*, 2002), as well as other proteins. Therefore, it is extremely likely that the the proteins FtsA and ZipA were pulled-down by ParF because of the binding to FtsZ. FtsZ, however, was identified in high amount in the mass-spec sample, hence it was selected for further investigation. The other three proteins were also found in relative high amount, as well as being interesting for the role played in the cell. MreB is an actin-like protein essential for shape determination in rod-shaped bacteria and was reported to play a role in chromosome segregation (Kruse *et al.*, 2003). MukB is, instead, a condensing protein involved in chromosome condensation and segregation (Niki *et al.*, 1991; Cui *et al.*, 2008). Finally GyrB is an enzyme than introduces negative supercoiling in circular DNA (Gellert *et al.*, 1976).

The non-essential proteins DamX, OmpC, OmpX, YdgA and PpiD were also further investigated. OmpX and OmpC are two outer membrane proteins (Baslé *et al.*, 2006; Vogt and Schulz, 1999) that could be indirectly involved in plasmid segregation. DamX is a cell division protein, which binds peptidoglycan and is involved in septum formation (Arends *et al.*, 2010). PpiD is a peptidyl–prolyl isomerise (Dartigalongue and Raina, 1998) which could potentially be involved in post-translational modification of the proline-rich motif in ParF. Finally, YdgA is a poorly characterised protein that may tether the partition complex to the cell membrane (Stenberg *et al.*, 2005).
As this set of proteins is not essential for cell viability, plasmid stability was tested in *E. coli* cells presenting deletion in the genes encoding for these proteins. BW25113 *E. coli* deletion mutants were obtained from the Keio collection and plasmid retention was measured using plasmid partition assay.

Proteins essential for bacteria growth and development could not be tested in a similar fashion. Instead their genes were cloned in a bacterial two-hybrid vector and binding to ParF was characterised *in vivo* by a bacterial two-hybrid assay.

5.2.4.1 Plasmid stability in E. coli deletion mutants shows that DamX, OmpC, OmpX, YdgA and PpiD do not appear to be involved in plasmid segregation

The Keio collection consists of gene deletion mutants that are constructed replacing the open reading frame regions with the kanamycin resistance gene (Baba *et al*., 2006). These mutants were constructed in *E. coli* BW25113 cells, the same strain used for TAP. The deletion mutants were first checked by PCR or sequencing. For the plasmid stability assay two plasmids were used: pFH554 and pFH450-LC. Both pFH554 and pFH450-LC vectors are a derivative of pFH547 plasmid, where the ColE1 origin of replication was deleted and low copy number replication is supported by the P1 origin. The plasmid pFH554 harbours the entire *parFGH* partition cassette, while pFH450-LC does not contain any partition gene and was used as negative control (Figure 5.15). Deletion mutants were transformed with the two plasmids. Wild type BW25113 *E. coli* cells were used as control and transformed with the same plasmids. Cells were grown in presence of chloramphenicol (and kanamycin for the deletion mutant) in order to exclude cells lacking the plasmids. They were then grown for 25 generation without antibiotic selection, before growing them again on chloramphenicol. This way, only colonies containing a full working partition system would be able to retain the plasmid and grow in the presence of antibiotic. Since the pFH450-LC plasmid does not contain any partition gene, it is unstable and lost at higher rates compared to the plasmid that harbours the full partition cassette. Plasmid partition assays were carried out in triplicates to obtain the average of plasmid retention. Plasmid pFH554 retention in wild type *E. coli* cells was on average 65.5%, while pFH450-LC retention was 19.5%. Plasmid partition assays in the *E. coli* deletion mutants showed that deletion of the genes identified through TAP-MS did not affect plasmid stability, suggesting that these proteins are either uninvolved or non-essential for plasmid segregation (Figure 5.16)

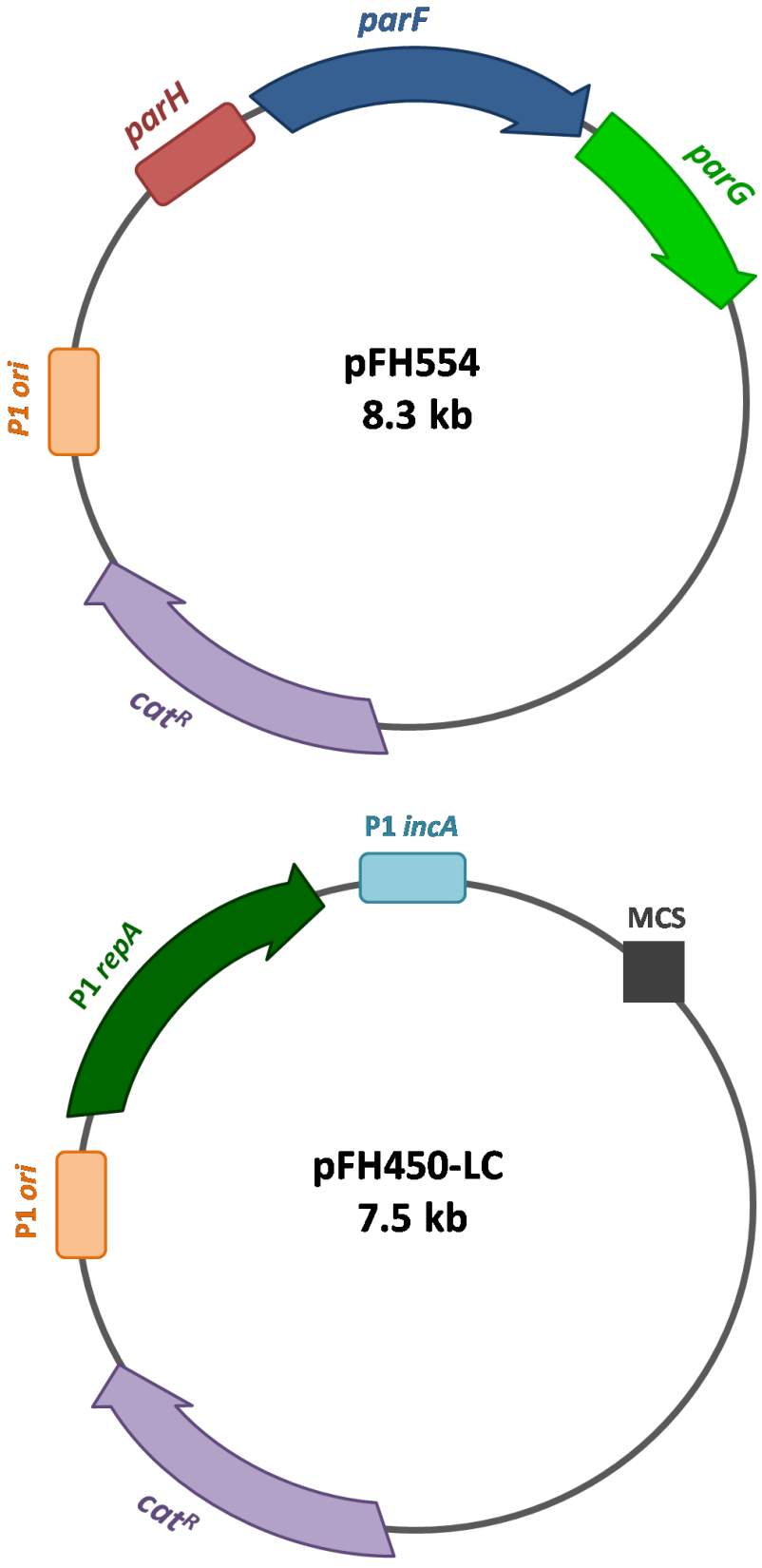

Figure 5.15 –Maps of plasmid pFH554 and pFH450

pFH554 plasmid harbours the low copy number P1 origin, the chloramphenicol resistance gene (*cat*^{R}) and the full parFGH partition cassette. pFH450 harbours a multiple cloning site (MCS) the low copy number P1 origin and the *cat*^{R} gene, but it does not harbour any partition gene.

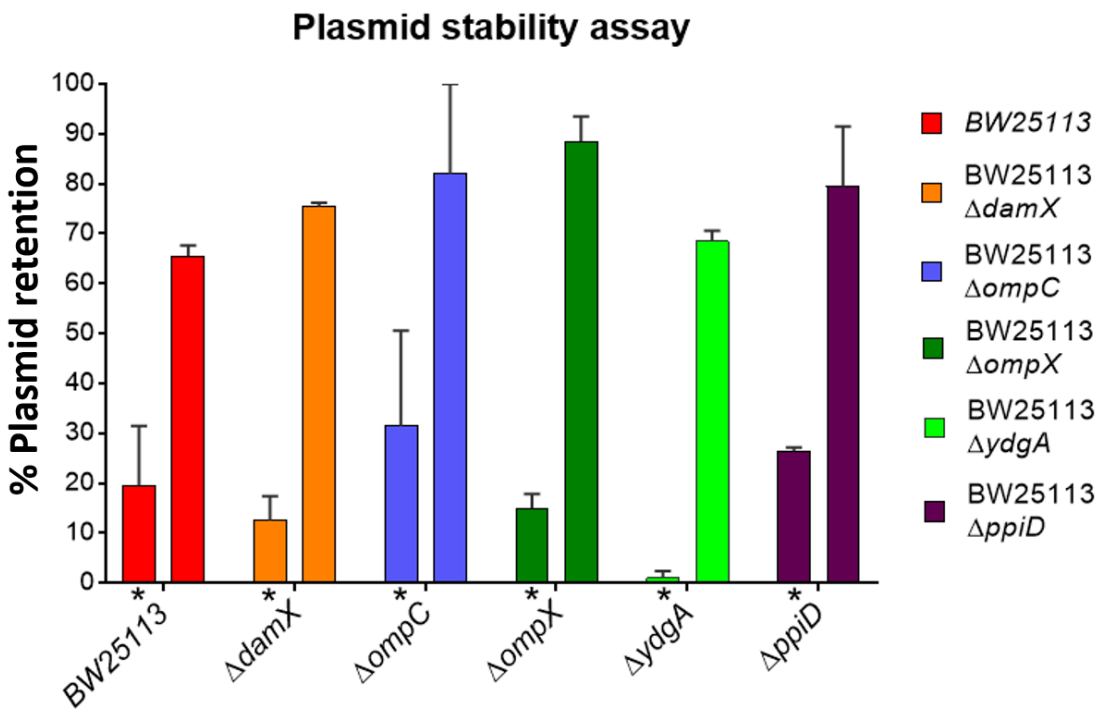

Figure 5.16 – Plasmid stability assay in *damX*, *ompC*, *ompX*, *ydgA* and *ppiD* deletion mutants

Percentage of pFH450-LC and pFH554 plasmid retation is plotted for each of the *E.coli* deletion mutants under investigation. Each colour represents a different mutant. The first column (labelled with the asterisk, *) represents the retention percentage of pFH450-LC (negative control), while the second column represents the retention value for pFH554. All experiments were carried out in triplicates and standard deviation calculated and showed by the error bars.

5.2.4.2 Bacterial two-hybrid shows that ParF does not interact with FtsZ, MreB, MukB or GyrB

A bacterial two-hybrid assay, coupled to interaction quantification by β-galactosidase assay, was used to investigate the *in vivo* interaction between ParF and FtsZ, MreB, MukB and GyrB. The bacterial two-hybrid system was first developed by Karimova *et al.* and is based on the signalling pathway triggered by the reconstruction of adenylate cyclase (CyaA) from *Bordetella Pertussis* (Karimova *et al.*, 1998). The catalytic domain of the enzyme is composed of two subunits, T18 and T25. Binding of the two subunits reconstitutes the catalytic activity of the enzyme which, in a *cya* deletion strain, can convert AMP in cAMP. Cyclic AMP binds to CAP (catabolite gene activator protein) forming the cAMP/CAP complex that triggers the expression of a number of genes, including those for maltose and lactose utilisation. In the two-hybrid assay, the change in phenotype (red colonies) cause by maltose degradation is monitored.

β-galactosidase quantification is a colorimetric assay based on the transcription of the *lacZ* gene, possible because of the formation of the cAMP/CAP complex. *lacZ* encodes for the β-galactosidase enzyme that degrades the substrate ortho-nitrophenyl-β-galactoside (ONPG), added to the reaction, causing the development of yellow colour. This can be measured spectrophotometrically.

For the experiment, pT25-parF was present in the lab collection and contained *parF* in frame with the *T25* moiety. *ftsZ*, *mreB*, *mukB* and *gyrB* genes were cloned in the pT18 plasmid upstream of the T18 moiety, as described in Section 2.4.5 (Materials and Methods).

For cloning, *ftsZ*, *mreB* and *gyrB* were amplified from chromosomal DNA using primers harbouring the XhoI (forward) and HindIII (reverse) restriction sites. The *mukB* gene was amplified using primers containing ClaI (forward) and ApaI (reverse) restriction sites (Figure 5.17 A). After cloning, plasmids were checked as described in 2.3.11 and sent to be sequenced (Figure 5.17 B).

For the bacterial two-hybrid assay, SP850 *E. coli* cells were co-transformed with the pT25-parF plasmid and one of the newly constructed pT18-ftsZ, pT18-mreB, pT18-gyrB or pT18-mukB plasmids. In addition, *E. coli* cells were co-transformed with pT25-parF and pT18-parG as a positive binding control, as well as with pT25-parF and empty pT18 as negative control. As T18 is not fused to any protein, it should not interact with ParF and no binding should be detected. Co-transformants were streaked first on LB and then on MacConkey agar supplied with maltose. When the proteins interact, the colonies turn dark pink, while if they do not interact the colonies display of a pale-pink colour. The two-hybrid assay showed that none of the proteins identified through TAP-MS interacted with

ParF *in vivo*, under the conditions used (Figure 5.18). The set-up of the assay can be considered correct, as the positive control showed development of dark-pink colonies. These results were further confirmed by semi-quantification through the β-galactosidase assay. The colorimetric assay was performed as described in Section 2.13.1 (Materials and Methods) and the interaction between the proteins quantified in Miller Units (MU). The ParF-ParG control showed an average value of ~1300 MU, while the negative control of ~50 MU. Interaction between ParF and FtsZ, MreB GyrB or MukB showed average values between 55-60 MU, suggesting that the proteins were not binding to ParF (Figure 5.19).

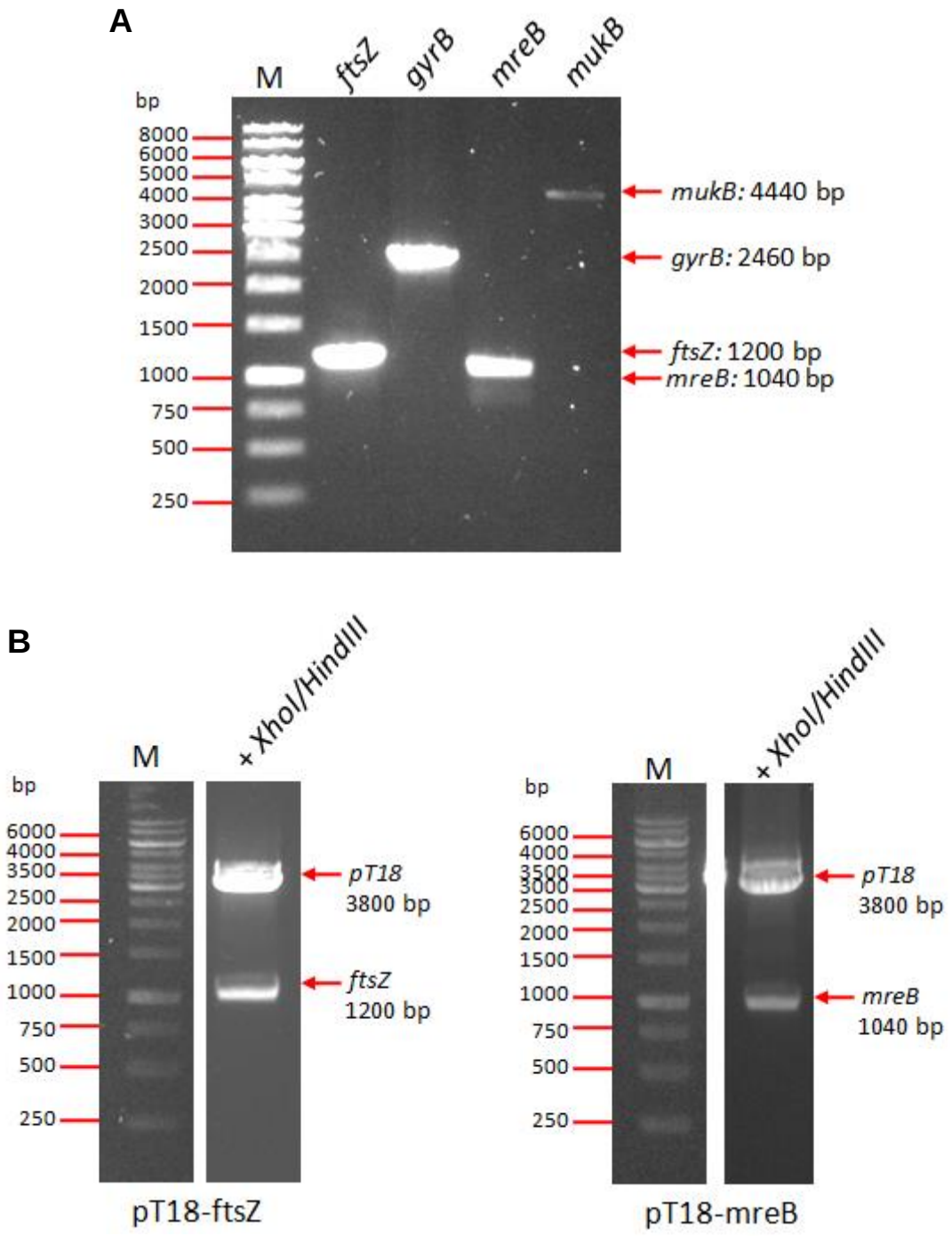

Figure 5.17 – Construction of pT18-ftsZ/mreB/gyrB/mukB

A) Amplification by PCR of *ftsZ* (1200 bp), *mreB* (1040 bp), *gyrB* (2460 bp) and *mukB* (4440 bp). B) Examples of cloning check by restriction digestion with the enzymes XhoI and HindIII of the plasmids pT18-ftsZ and pT18-mreB. The vector size is 3.8 kb.

Figure 5.18 – Bacterial two-hybrid assay for the interactions ParF-FtsZ, ParF-MreB, ParF-GyrB and ParF-MukB

E. coli cells were co-transformed with pT25-parF and pT18 containing one of the following genes: *ftsZ, mreB, gyrB, mukB* or *parG* for positive control of the binding. Co-transformation of pT25 with pT18 harbouring no gene in frame with T18 was used as negative control. Development of dark pink colour shows binding, while pale pink colour shows no binding.

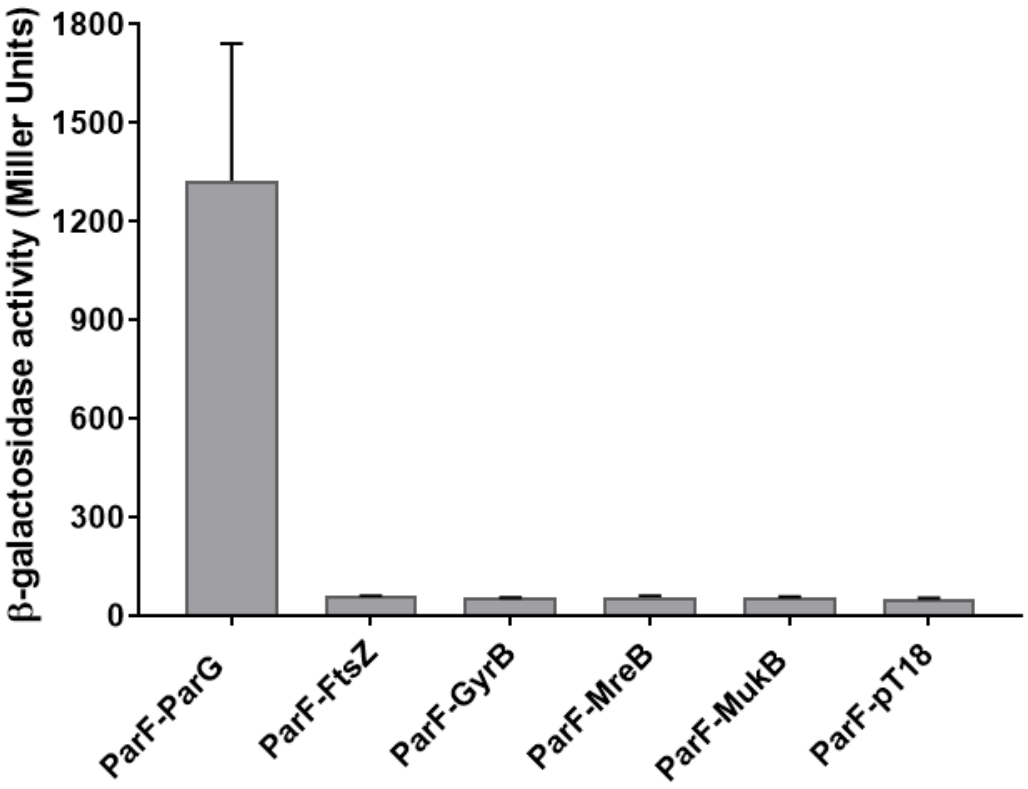

Figure 5.19 – β-galactosidase assays testing the interactions ParF-FtsZ, ParF-MreB, ParF-GyrB and ParF-MukB

β-galactosidase activity is shown for each of the proteins selected (FtsZ, GyrB, MreB and MukB) with ParF. Each experiment was repeated in triplicate. Standard deviation was calculated and shown by the error bars.

5.3 Discussion

Although promising, TAP-MS did not allow the identification of any binding partner of ParF or ParG. Based on this experiment only, it appears that ParG does not interact with any other protein, apart from ParF. Instead, TAP-MS on ParF-SF allowed the identification of a high number of proteins, among which over 100 were only found in ParF samples. Many of those were found in very low amount and were excluded from the initial analysis. In fact, due to time limitation it was impossible to test all the proteins identified, and a selection was made. Proteins were chosen mostly according to abundance in the mass-spectrometry samples, therefore it is possible that interacting partners weakly expressed and present in low amount in the cell could have been overlooked. Some of the proteins identified are enzymes whose activity does not seem to be related to plasmid segregation, therefore they were left aside in this initial analysis. Interestingly, several cell division proteins were pulled-down with ParF, suggesting a potential correlation between cell division and plasmid segregation.
An important note to make is that TAP was performed in cells that did not contain the full partition cassette, instead only one partition protein was expressed. Overexpression of one of the partition genes causes plasmid loss (Abeles *et al.*, 1985; Funnell, 1988a), since a balance between the components is required for proper plasmid segregation. However, it is possible that assembly of the full segrosome is required for unknown interacting partners to bind to the partition proteins. Therefore, we cannot rule out that a similar experiment, including the full partition system (*parH*/ParG/ParF), would give a different outcome. It was also deliberately decided not to include ParG in the cells, to avoid the protein shielding any other interaction with ParF.

Proteins identified through mass spectrometry were grouped into essential and non-essential proteins. Stability of a low copy number plasmid harbouring the *parFGH* partition cassette was measured in *E. coli* mutants, where one of the non-essential genes, expressing the proteins identified through TAP-MS, was deleted. Plasmid partition assay showed that none of the proteins selected, DamX, OmpC, OmpX, PpiD and YdgA, is involved in plasmid segregation. OmpC and OmpX are both outer membrane proteins. It was initially hypothesised that they could be indirectly involved in plasmid segregation. In the past, a gene suppressor screen was used to identify proteins interacting with ParF. In the assay, ParF was overproduced in cells transformed with a *Salmonella* library, causing plasmid instability. However, if among the genes of the library one was interacting with ParF, this would have restored the balance and prevented the plasmid from being lost. One of the proteins identified through this screen was OmpX, hence the interest in understanding whether outer membrane proteins could be relevant

in plasmid segregation. However, the results obtained strongly suggest that they do not take part in the process. Although the proteins were not identified in the control samples, it is possible that OmpC and OmpX were pulled down because of membrane debris present in the crude extract. PpiD is a prolyl isomerase which catalyses the isomerisation of proline residues and it was initially hypothesised that it could play a role in post-translational modification of ParF proline-rich motif. The protein is actually involved in the folding of Omp proteins and may have been pulled down by them (Dartigalongue and Raina, 1998; Matern *et al.*, 2010). DamX, however, is an inner membrane protein that binds peptidoglycan. The protein is involved in remodelling peptidoglycan during cell division, however its function has not been completely elucidated (Arends *et al.*, 2010; Yahashiri *et al.*, 2017). It is possible that the protein interacts with ParF, but it is not essential for plasmid segregation. It would be interesting to further characterise the interaction between ParF and DamX *in vitro*. Very little is known about YdgA, nevertheless it was interesting to investigate whether this protein is involved in plasmid segregation. Indeed, the protein does not appear to take part in the process and has probably been pulled down by the interaction with other unknown factors. A plasmid partition assay on deletion mutants from the Keio collection proved to be a powerful tool to exclude non-specific or non-essential proteins identified through TAP-MS.

The bacterial two-hybrid, used to test the *in vivo* binding between ParF and FtsZ, MreB, GyrB and MukB, also showed negative results. None of the proteins under investigation seemed to interact directly with the partition protein. However, it is important to note that false negatives are relatively common in this kind of assay, especially for constitutive proteins. Tagging essential genes can often impair their activity and lead to misinterpretation of the results. Both MreB and FtsZ form polymers, therefore activity of the tagged protein could be tested by cloning *mreB* and *ftsZ* genes in the complementary pT25 vector and looking at the binding. However, because of time limitation, this could not be checked. Since tagging the gene could represent a problem for protein activity, *in vitro* chemical cross-linking of the purified proteins (FtsZ, MreB, GyrB and MukB) with ParF could help understanding whether binding is taking place.
It is again important to specify that binding between ParF and the selected proteins was tested in *E. coli* cells that did not harbour the full partition cassette. Therefore, the negative results obtained could be due to the fact that this element is missing. It is possible that repeating the experiments including the partition system may lead to different results.

Although more work must be carried out to further prove or disprove the bacterial two-hybrid results, based on the information acquired so far, the *parFGH* partition cassette seems to act independently from host factors. Similarly to what is shown by Ebersbach *et al.* for the pB171 plasmid, FtsZ and MreB do not seem to interact with the partition proteins (Ebersbach *et al.*, 2006). Also MukB, which was reported to be involved in chromosome segregation, does not play a role in TP228 plasmid partition, as previously observed for the plasmids F and P1 (Ezaki *et al.*, 1991; Funnell and Gagnier, 1995; Weitao *et al.*, 2000). Finally, we speculated that GyrB could have been involved in supercoiling of the plasmid or of the chromosomal DNA to enhance plasmid segregation. The hypothesis was, however, disproved, as the protein shows no binding to ParF and was not identified in the ParG samples. Further *in vivo* and *in vitro* characterisation of these and other putative interacting partners will give a better understanding of ParF interaction network.

Chapter 6
Discussion and Future work

6.1 Discussion

In the last three decades, plasmids and chromosome segregation mechanisms have been extensively studied, in the attempt to understand how bacteria pass their genetic information to the offspring. In particular, work on plasmids helped elucidating mechanisms which were seen to be conserved in chromosome dynamics. The plasmid TP228, focus of this work, is one of the best characterised systems to date and the work presented in this thesis aimed at drawing a more complete picture of the mechanism of segregation, investigating the molecular interactions between the segrosome components and with putative host-encoded factors.

The segregation system of the plasmid TP228 is composed of two proteins, ParG and ParF, and a centromeric site upstream the partition cassette, *parH*. ParG has a dual role, it works as centromere-binding protein and as transcriptional repressor of the *parFG* operon (Wu *et al.*, 2011; Hayes, 2000; Golovanov *et al*., 2003; Zampin*i et al.*, 2009). ParF is a Walker-type ATPase which works as a motor protein, moving plasmids apart before cell division (McLeod *et al*., 2017; Schumacher *et al*., 2012). ParF does not directly interact with *parH*. Instead, the protein interacts with its plasmid cargo by binding to the adaptor protein ParG (Barillà and Hayes, 2003b). Synergy between the segrosome components is key and unbalance of proteins' amounts lead to plasmid loss (Abeles *et al.*, 1985).

In this Chapter, the interaction between ParG and ParF will be first dissected further, with a focus on the techniques used in the investigation. The formation of the ParG-*parH* nucleoprotein complex will be then discussed in light of the information acquired through the study on ParG double and triple mutants. A mechanism for segrosome formation will be drawn based on this and previous work. The question of whether host-encoded factors are involved in the system will be then addressed at the end of the Chapter. Future work will be discussed.

6.1.2 Dissecting the role played by the ParG N-terminus in the interaction with ParF

Interaction between ParG and ParF is a prerequisite for plasmid segregation and amino acids changes abolishing essential non-covalent interactions between the two proteins cause plasmid instability. ParF dimerises in presence of ATP, while ADP binding causes dimers' dissociation (Schumacher *et al.*, 2012). The ParF-AMPPCP crystal structure showed that the nucleotide is sandwiched between two ParF monomers (Schumacher *et al.*, 2012). Additional contacts between monomers are made by the proline-rich motif,

which is also involved in the interaction with ParG (Schumacher *et al.*, 2012, Allison-Gamble, 2016). In particular, Allison identified Ser-108, Pro-109, Leu-110 and Phe-112, belonging to the proline-rich motif, and Val-149 as essential for the interaction with the DNA-binding protein (Allison-Gamble, 2016). The interface between ParF monomers is therefore of particular importance for binding to the partner protein, explaining why ATP binding and consequent dimerisation enhances ParF-ParG association (Barillà and Hayes, 2003).

ParG was shown to interact with ParF through both C-terminal and N-terminal domains (Barillà *et al.*, 2007). Although the N-terminal domain was identified as the region essential for ParG-ParF binding (Barillà *et al.*, 2007), deletion of this region did not abolish the interaction. Both bacterial two hybrid assay (Barillà *et al.*, 2007) and DMP cross-linking (this work) showed that the ParGΔ30 deletion mutant could still bind to the partner ParF. However, biophysical techniques such as MST and SPR showed that the interaction is barely detectable, suggesting that binding through the C-terminal domain is weak and transient.

On the other hand, single amino acid changes in ParG N-terminal region were shown to have a detrimental effect on plasmid segregation for different reasons. Two of the primary activities of ParG are binding to ParF and enhancing its weak ATPase activity. Arg-19 was identified as the prime residue of an arginine finger motif, which is inserted into the ATP-binding domain and stabilises the negative charges developed during ATP hydrolysis (Barillà *et. al*, 2007). This was further confirmed by the co-crystallisation of fragments of ParG with ParF-AMPPNP dimer. The structure showed that the arginine finger motif forms a helix that is inserted at the monomer-monomer interface to attend to its function (Zhang and Schumacher, 2017). In this work we identified other residues that are required for ParG-ParF recognition and association. Lys-5 and, to a lesser extent, Leu-3 at the tip of the tail were identified as residues whose primary activity is to bind to ParF. Change of Lys-5 into alanine, in particular, was shown to completely abolish the interaction. Cross-linking/mass-spectrometry (CL/MS) allowed the identification of Lys-11 and Lys-12 as amino acids also involved in binding to ParF, while Lys-5 was not identified using this technique. However, since the tail should insert at the interface between ParF monomers, it is possible that the region containing Lys-5 was buried within the complex and not reachable by the cross-linker. On the other hand, Lys-11 and Lys-12 were cross-linked to residues of ParF (e.g. Lys-155, Lys-130, Lys-160, Lys-174, Lys-52 and Lys-49) which are surface exposed and easily accessible. Furthermore, Lys-11 and Lys-12 could be interchangeable in the binding to ParF, explaining why single amino

acid mutation of K11 and K12 had only a moderate effect on ParG interaction with ParF. Instead double mutation affected binding to a much greater extent. The methionine residue, adjacent to Lys-12 is instead hypothesised to have an important role in structuring ParG tail and is not supposed to be involved in ParF recognition by ParG. In fact, although single amino acid change of Met-13 into alanine decreases ParG affinity for ParF, adding the M13A mutation to ParG-K11A-K12A mutant did not reduce the binding further. Instead, better level of binding was recorded for ParF-ParG triple mutant binding, compared to the double mutant.

Regarding the residues that compose the arginine finger motif (N18-R19-L21), results obtained by SPR and MST were conflicting and it was hard to draw conclusions. MST showed that when Arg-19 was changed into alanine, ParG binding to ParF was slightly affected, while L21A ParG mutant had identical behaviour as the wild type protein. Replacement of Asn-18 with an alanine residue, instead, greatly affected the interaction, weakening ParG-ParF binding affinity by 5 fold. On the other hand, SPR showed that the three single amino acid changes in this cluster moderately affected ParG-ParF binding to a similar extent. Differences in the results lie in differences in the techniques set up. In MST, ParG was labelled by reaction of the succinimidyl ester moiety of Alexa-647 dye with the protein's primary amines. In SPR, instead, ParF was covalently attached to a CM5 chip by amine coupling. Therefore, both techniques rely on the modification of the lysine residues. Both ParF and ParG harbour multiple lysine residues and it is unlikely that modification of these amino acids by labelling or immobilisation could have impaired the binding with the partner protein. However, every technique has limitations, hence the importance of confirming the results by a different method. In SPR, ParF was immobilised on a surface and this could have affected the interaction with ParG because the protein could not reach its binding site. In MST, instead, both proteins are in solution and free to move. However, MST is a much less explored technique and not many papers are available on its applications and limitations. It is possible that labelling may have caused non-specific binding or that modification of certain amino acids had slightly modified protein's characteristics, leading to overestimation of the binding. Furthermore, MST was poorly reproducible, adding uncertainty to the interpretation of the results. Zhang and Schumacher showed that the arginine finger motif inserts at the interface between ParF monomers and that this interaction stabilises the ParF-ATP dimer (Zhang and Schumacher, 2017). This suggests that mutation of these residues would affect the interaction of ParG with ParF. The co-crystal structure, however was low resolution and it was impossible to precisely map the interaction between amino acids. According to the authors, the hydrophobic residues of the arginine finger motif are positioned facing the

ParF dimer interface. In particular, Gly-16 in ParG would be juxtaposed to Val-149 and Pro-109 in ParF and any amino acid other than glycine would clash with the ParF residues (Zhang and Schumacher, 2017). Barge showed that mutation of Gly-16 into alanine did not affect plasmid segregation (Barge, 2015). Furthermore, in this work we identified cross-linking between ParG Lys-11 and Lys-12 with ParF' Lys-155. Although cross-linking is artificial, this suggests that the two regions of the proteins must be in close proximity (less than 11.4 Å apart, the length of the cross-linker). According to the conformation of the complex identified by Zhang and Schumacher, instead, the amino acids are more than 20 Å apart, hence they would be too far to be cross-linked. However, cross-linking between Lys-11 in ParG and Lys-174 in ParF was also identified during this project. In this case, the structure suggested by the authors would match our finding. Therefore, it is possible that the tail could structure differently from what observed in the partial co-crystal structure or assume different conformations.

In addition, both *parH* and chromosomal DNA were seen to be involved in the ParF-ParG complex formation, probably affecting the conformation of the proteins. CL/MS showed that presence of *parH* in ParF-ParG cross-linking reaction improved binding by 2.8 fold. Addition of chromosomal DNA enhanced binding even further (1.2 fold increase compared to the ParF-ParG-*parH* complex). Specific and non-specific DNA were seen to promote cross-linking at some positions in particular. As an example Lys-12_{ParG}-Lys-155_{ParF} cross-linking was 8 fold greater when *parH* was present in the reaction and 11 fold greater when all the components were present (ParG, ParF, *parH*, nsDNA and ATP). On the other hand, cross-linking between Lys-11_{ParG} and Lys-174_{ParF} was enhanced two-fold by the presence of *parH* and up to 17 fold by the presence of all components in the reactions.
Biophysical measurements were all carried out in reactions including only ParF and ParG, because additional components as ATP or DNA would have made quantifications very challenging. Therefore, *in vitro* characterisation underestimated the strength of the binding that would be recorded in an *in vivo* setting.

Alanine-scanning mutagenesis, binding measurement by biophysical techniques and cross-linking/mass-spec analysis presented in this work suggest that ParF recognition by ParG is mediated by the amino acids at the tip of the tails, while the central domain makes superficial interactions and structures the tail. The residues composing the arginine finger motif make transient interactions at the interface between ParF monomers and localise close to the ATP-binding motif to stimulate ParF ATPase activity, as previously identified (Barillà *et al.*, 2007; Zhang and Schumacher, 2017a) (Figure 6.1).

Besides K11 and K12, CL/MS also allowed the identification of two more residues in ParG, which localise in close proximity to ParF when the two proteins are bound: K33 and K35. According to ParG DNA-binding model, ParG binds to its cognate site as a dimer, by inserting the anti-parallel β-sheet formed by the intertwining of ParG monomers (with a RHH structure), into the major groove of the DNA(Golovanov *et al.*, 2003). The position of the two residues is within or just outside the β-strand structure, which should make contact with the specific DNA. As a consequence, these residues should not be available for interaction with ParF. *In vitro* analysis showed that ParG can bundle and cross-link ParF oligomers into higher-order structures, by using the N-terminal regions as sticky tentacles (Barillà *et al.*, 2007; Barge, 2015). The hypothesis of ParG cross-linking ParF units, independently of the DNA, would explain why cross-linking was identified between four lysine residues in the ParG N-terminus (K11, K12, K33 and K35) and lysine residues scattered around ParF surface. In addition, cross-linking between ParG K33/K35 and ParF did not seem to be affected by the presence of the DNA, while ATP addition greatly enhanced the binding, further supporting this hypothesis. Therefore, ParG could assume different conformations according to the function it is performing. When it is bound to *parH*, association with the DNA causes a conformational change in the protein that affects the N-terminal domain, enhancing binding to ParF and consequently, probably, its catalytic activity, by positioning the arginine finger motif. When, instead, it is not bound to the DNA, it can extend the N-terminal regions and cross-link ParF units, exploiting the β-strand regions which are free from DNA association (Figure 6.2).

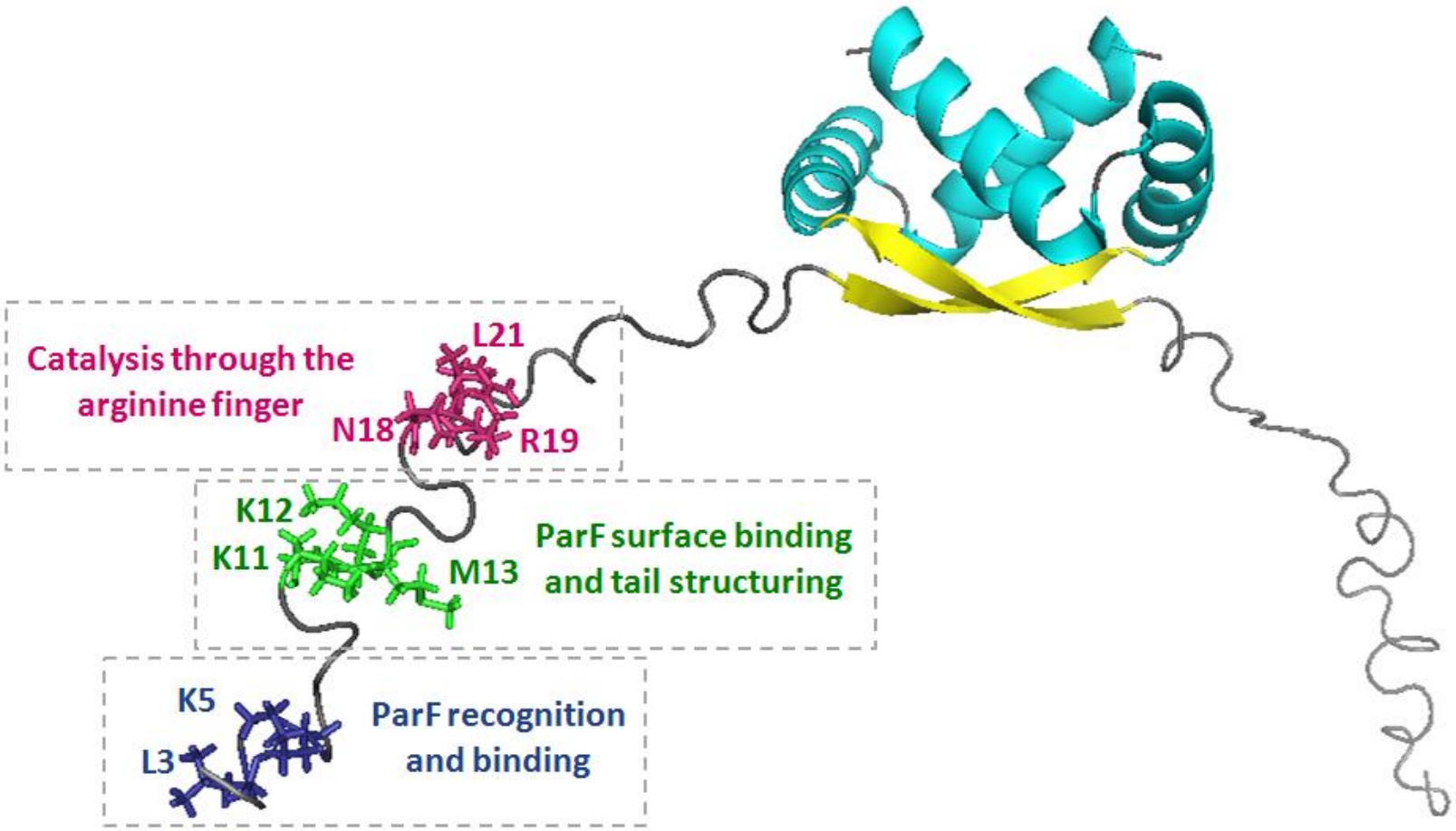

Figure 6.1 – ParG structure and N-terminal clusters description

ParG has a ribbon-helix-helix C-terminal domain and an unstructured N-terminal tail. Clusters of the N-terminal domain identified through alanine scanning mutagenesis have been highlighted in different colours according to the function. Image generated by PyMOL, PDB file: 1P94.

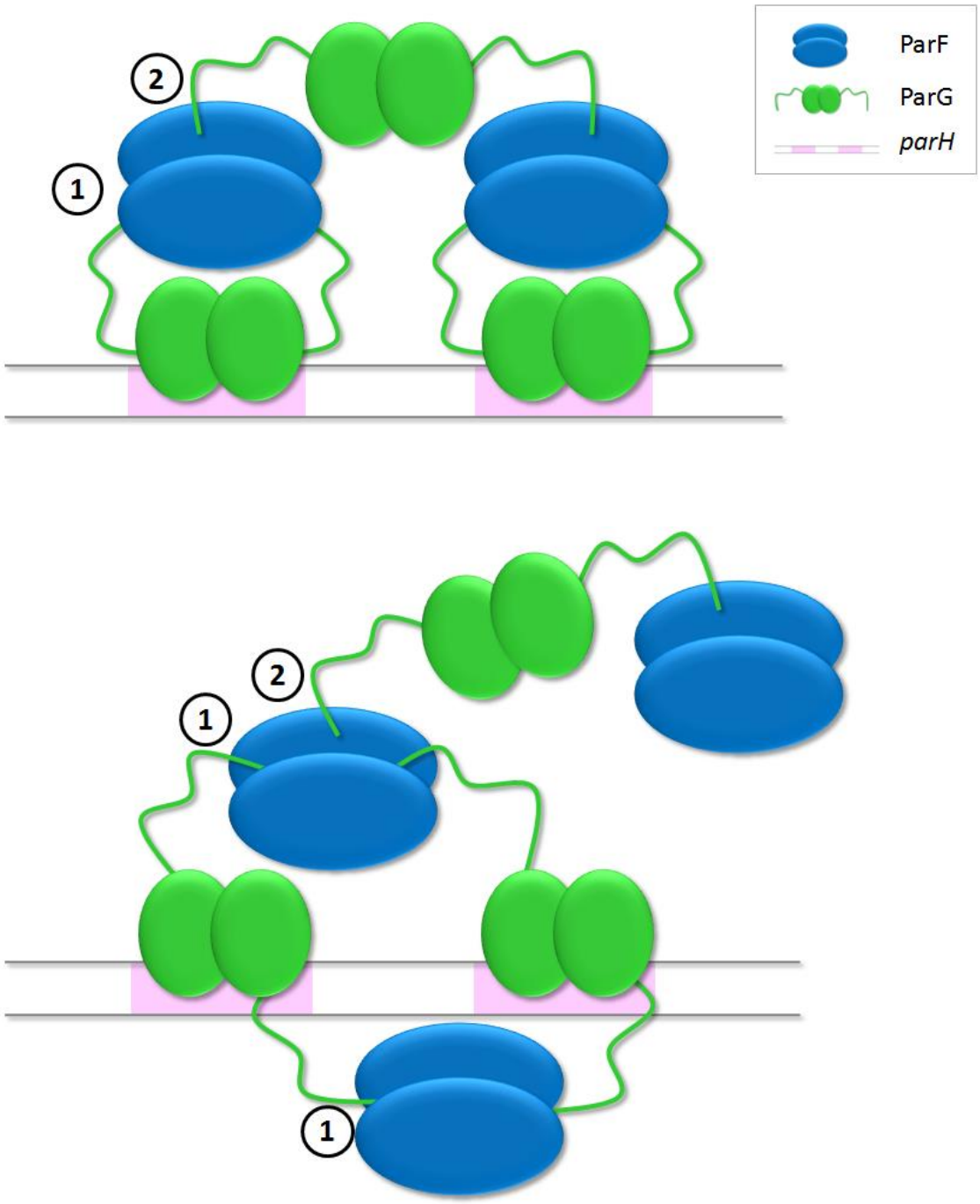

Figure 6.2 – Alternative models of ParG-ParF recognition and binding

Binding models of the interaction between ParG and ParF. The top image shows the interaction between ParG dimer and ParF dimer. The bottom image represent a model of the interaction between ParF dimers bound by ParG molecules belonging to two different dimers. (1) represents the specific interaction between ParG and ParF. ParG inserts the N-terminal tails and the arginine finger motif at the interface between ParF monomers. (2) represents the non-specific interaction between ParF and ParG, where ParG recognises regions on the surface of ParF through the N-terminal tails and cross-links ParF dimers.

6.1.2 Elucidating the role played by ParG N-terminus in the interaction with *parH*

ParG mediates the interaction between ParF and the plasmid through the centromeric site *parH*. However, it also acts as a transcriptional repressor, by binding to the operator site O_F, upstream of the *parFG* operon (Zampini *et al.*, 2009; Wu *et al.,* 2011). The two regions are adjacent to each other and consist of the same repeated ACTC motif, interspaced by 4 bp AT-rich sequences (Zampini *et al*., 2009; Wu *et al.,* 2011). The only difference between *parH* and O_F is in the orientation of the degenerate sequences, direct repeats in *parH* (apart from one in the centre) and either direct or invert repeats in O_F.

As mentioned previously, although no structure is available for ParG-*parH* complex, structural information obtained from other RHH proteins suggest that ParG could recognise its binding site as dimer of dimers, inserting the double stranded anti-parallel β-sheet of each dimer into consecutive major grooves of the DNA (Golovanov *et al.*, 2003). NMR analysis showed that while the C-terminal domain makes strong and specific interaction with the DNA, the N-terminal domain helps improving binding specificity, possibly by making transient interactions with the DNA (Carmelo *et al.,* 2005; Golovanov *et al.*, 2003). Several RHH harbour a N-terminal unstructured region (Schreiter and Drennan, 2007). The only partition protein co-crystallised with its cognate site is the ω Δ19 protein (truncated at the N-terminus) from the plasmid pSM19035 (Murayama *et al.,* 2001). Most of the tail is absent, causing a slight reduction in the binding to the DNA (Murayama *et al.*, 2001; Dostál *et al.*, 2007). How the tail enhances DNA binding is still unknown. However, it was demonstrated that the 20 amino acids at the N-terminus of the protein are not involved in the interaction, suggesting that the residues closer to the folded domain must be involved instead.

Arc and MetJ repressors are also ParG structural homologues and harbour disordered N-terminal regions of different lengths (Schreiter and Drennan, 2007). In both cases, this region was shown to interact with the DNA and improve binding specificity (Brown *et al.*, 1990; Knight and Sauer, 1989; Somers and Phillips, 1992; Augustus *et al.*, 2008). In particular, the flexible tails of the MetJ repressor were seen to change conformation completely when the protein was in complex with the DNA and to wrap around the DNA phosphate backbone (Somers and Phillips, 1992). HSQC NMR spectroscopy on MetJ showed that when DNA was present, the spectrum of the protein completely disappeared (Augustus *et al.*, 2008). Addition of the DNA to ParG, instead, was seen to cause the C-terminal domain to disappear, while the chemical shift of the N-terminus changed to different extent according to the position of the residue (Carmelo *et al.*, 2005), with Lys-

11 and Lys-12 being the most affected. Information on the involvement of the N-terminal domain of Type Ib CBPs in DNA binding are still missing and was an objective that we tried addressed in this work.

As detailed in the previous section, ParG N-terminal domain is essential for binding to ParF and enhancing its ATPase activity. HSQC spectroscopy and biochemical assays using N-terminal truncated ParG showed that this region is also involved in DNA binding. In particular, deletion of the ParG tail greatly affects ParG ability to work as a transcriptional repressor (Carmelo *et al.*, 2005; Barillà *et al.*, 2007). In this work it was hypothesised that ParG N-terminal tails interacts with the DNA through the residues Lys-11, Lys-12 and Met-13. These residues could, then, help the tail to change conformation, leading to an increase in ParG binding affinity for the partner protein, ParF. Double (K11A-K12A) and triple (K11A-K12A-M13A) ParG mutants were constructed to investigate this hypothesis during this project.

Firstly, plasmid retention following insertion of double and triple mutation in *parG* was measured *in vivo*. Replacement of ParG with the two mutants caused a decrease in plasmid stability from 65% to 2% for the double mutant and 16% for the triple mutant. Both double and triple mutations, therefore, greatly reduced plasmid retention, although replacement of wild type ParG with the double mutant had a more detrimental effect on plasmid segregation.
Since one of the main roles of the tail is ParF binding, interaction between double and triple ParG mutants and ParF was measured. Results showed that multiple amino acid changes in this region greatly reduced ParG-ParF binding. Because of the low response, it was impossible to measure the K_D of the interaction. On a related note, ParG's ability to stimulate ParF oligomerisation was also affected by the double and triple mutations. In particular, ParG triple mutant could enhance ParF oligomerisation to a similar extent as ParGΔ30, which lacks the tail, suggesting that the M13A substitution damages the tail significantly.

Insertion of single residue changes in the K11-K12-M13 cluster was seen to have almost no effect on *parH* binding and transcriptional repressor activity by ParG (Barge, 2015). Insertion of multiple amino acid changes, instead, caused a slight decrease in ParG affinity for its cognate site. Results showed that ParG-K11A-K12A mutant had a two-fold weaker affinity for *parH*, compared to the wild type protein, while ParG-K11A-K12A-M13A was less affected. One explanation for this counterintuitive result is that methionine is larger than alanine and causes steric hindrance in the interaction with the

DNA, especially when the two positive lysine residues are not available for interaction with the negatively charged phosphate backbone. Investigation of ParG-*parH* complex by CD spectroscopy showed that the protein changes conformation upon binding to the DNA. Similarly to the wild type protein, ParG-K11A-K12A and ParG-K11A-K12A-M13A structure was also seen to change, although to a lesser extent. The structure of the double mutant was seen to be the one changing the least, probably because of the reduced interaction with the DNA. Secondary structure analysis using the CONTIN algorithm, based on the CD spectra recorded, showed that the three proteins remain flexible when in complex with the DNA. Instead, α-helix and β-structure content was seen to change between the free and DNA-bound state. In particular, wild type ParG showed an increase in the β-strand content, while the double and triple mutant showed a decrease. It must be considered, however, that the algorithm simply compares known secondary structures to the one under investigation and may, therefore, be unreliable.

The hypothesis that structuring the tail through the K11-K12-M13 cluster can enhance ParG-ParF binding, after the *parH*-ParG nucleoprotein complex has formed, was further tested by chemical cross-linking of the two proteins in the absence and presence of *parH*. DMP cross-linking showed that the amount of ParF-ParG complex formed when DNA was in the reaction was two-fold more abundant than when the proteins reacted alone. The same improvement was visible for the ParG-K11A-K12A-ParF complex. Conversely, ParG-K11A-K12A-M13A cross-linking to ParF did not improved upon DNA addition. An explanation for this behaviour may be connected to the inability of the triple mutant to structure the tail. This result, together with the polymerisation assay showing that the triple mutant failed to stimulate ParF polymerisation, supports the hypothesis that Met-13 is needed to structure ParG tails. K11-K12-M13 could therefore work as a structural cluster to orientate the arginine-finger motif in the most appropriate position for its catalytic activity, anchoring the residues K11 and K12 on ParF surface.

In light of these and previous results, it appears that ParG conformation changes upon interaction with the DNA. Most likely, a change in the C-terminal domain conformation translates into a change in the N-terminus structure, which, in turn, enhances ParG interaction with ParF. The Met-13 residue seems to be key in structuring the tail. Another plausible scenario is that triple amino acid change affects the tail to such an extent that it can no longer become structured. All things considered, the initial stage of segrosome formation involves *parH* recognition by ParG. Formation of the nucleoprotein complex leads to a change in ParG overall conformation and structuring of the N-terminal domain.

Binding to *parH* improves ParG-ParF interaction and leads to stronger association between ParF and its plasmid cargo.

However, the results presented in Chapter 4 raised the question of whether ParG N-terminal tail could assume different conformations when the protein is bound to the operator site compared to when it is bound to the partition site. Mutation of Lys-11 and Lys-12 weakened ParG binding to its cognate site, suggesting that the protein may be involved in making transient interactions with the DNA. HSQC spectroscopy shows that the chemical shift of the two residues changes considerably when the protein is in complex with the DNA. Other amino acids, whose chemical shift was seen to change are: Asn-18, Arg-19, Arg-23, Thr-26, Ala-27, Va-29. Change in chemical shift could either be due to the weak and transient interaction with the DNA or to a change in conformation or environment of the residues. Based on similarities with other RHH proteins, it is tempting to hypothesise that when ParG acts as a transcriptional repressor and binds to the operator site, the tails wrap around the DNA backbone (Figure 6.3). This could also explain why the degenerate repeats composing O_F and *parH* have different orientations. It is possible that the presence of directed and inverted repeats may facilitate binding of the tails to the DNA backbone.

Since the tail retains its flexibility even when the protein is bound to the DNA, it is possible that the tails make weak and transient interactions with the DNA until a high concentration of ParF in the environment competes with the DNA for tails binding, causing the tails to change conformation and to bind to ParF instead (Figure 6.4). This hypothesis represents speculation that would be interesting to investigate further by FRET for example (described in Section 6.2).

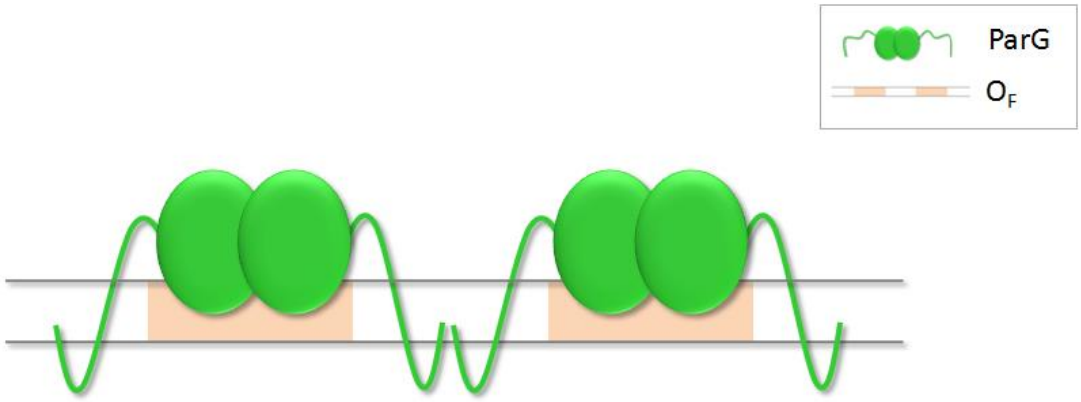

Figure 6.3 – Possible mechanism of ParG binding to O_F

Putative binding model where ParG recognise the operator site (O_F) through the C-terminal domain and uses the N-terminal tails to wrap around the phosphate backbone of the DNA

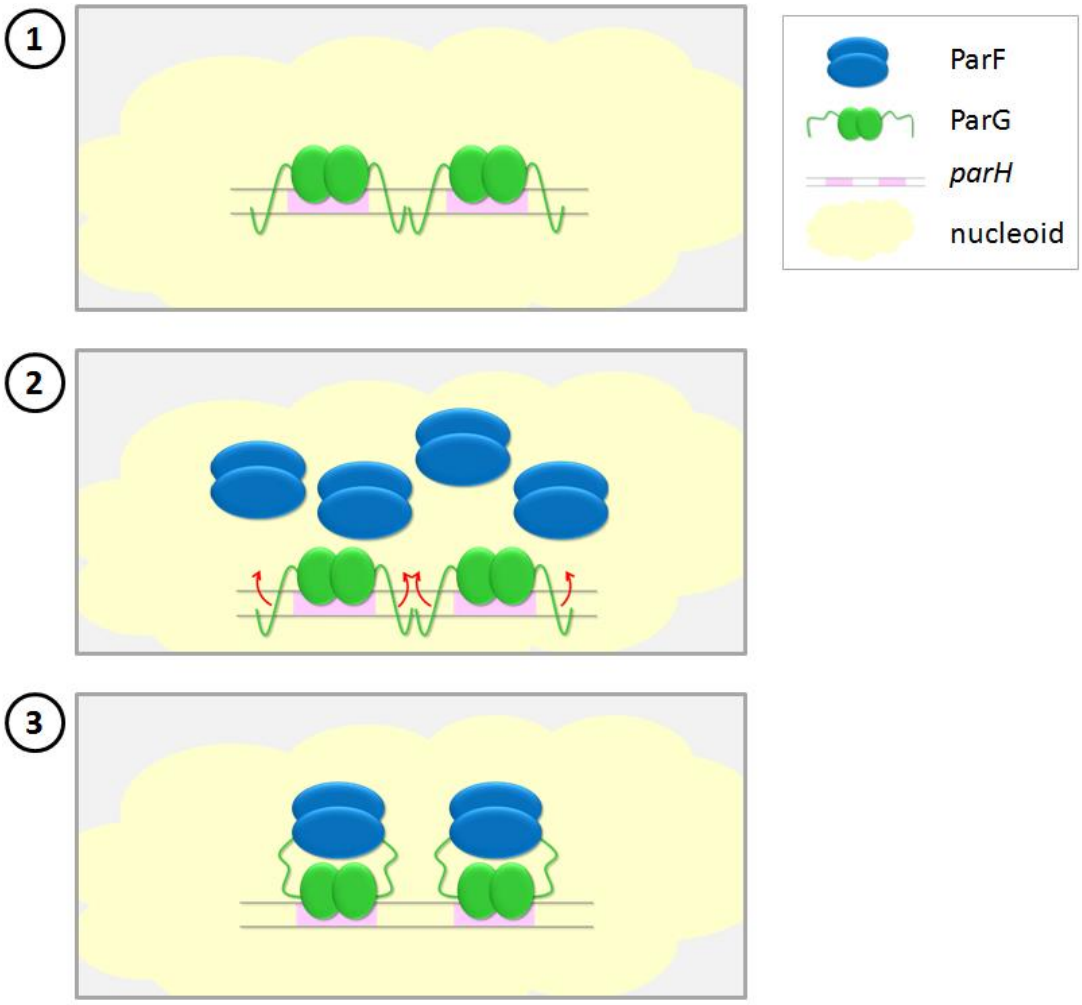

Figure 6.4 – Possible mechanism of segrosome assembly

Putative mechanism of *parH*-ParG and then *parH*-ParG-ParF complex formation. (1) ParG recognises the *parH* site through the C-terminal domain. The N-terminal tails improve binding specificity and may wrap around the phosphate backbone. (2) During plasmid segregation high concentration of ParF compete with the DNA for ParG N-terminus binding. The tails dissociate from the DNA and bind to ParF. (3) The segrosome forms thanks to the conformational rearrangement of the tails.

6.1.3 Investigating the involvement of host-factors in plasmid segregation

TP228 is a multidrug resistance conjugative plasmid, originally indentified in *Salmonella newport*. All the experiments reported in this work were carried out in *E. coli* cells, where a low copy number plasmids harbouring the *parFGH* partition cassette is stably maintained. Low copy number plasmids encode their own replication and segregation systems. However, plasmid often rely on their host for function such as replication, transcription and translation. Therefore, although the ParABS system seems autonomous and self-sustained, host-encoded factors may play a role in plasmid segregation and cooperate with the partition proteins.

In P1, and possibly pB171, plasmid segregation, IHF assists ParB loading onto the centromeric site by bending the DNA (Funnell and Gagnier, 1994; Ringgaard *et al.*, 2007). To date, there is no evidence that any other host-factor participates in plasmid partition or interacts with partition proteins.

In the past, it was proposed that plasmids could be tethered at midcell and cell quarters positions by host-factors that coordinated plasmid partition with cell division (Niki and Hiraga, 1997; Li and Austin, 2002). Ebersbach *et al.* also screened the interaction between pB171 *par2* proteins with a number of cell division proteins to test this hypothesis, finding no proof of binding (Ebersbach *et al*., 2006). P1 and F plasmid stability was also tested in *E. coli* cells, where the *mukB* gene was deleted. MukB is a condensin protein involved in chromosomes condensation and segregation, which could have also played a role in plasmid partition. However, deletion of the gene was shown to have no effect on plasmid stability, suggesting that the protein does not take part in the process (Ezaki *et al.,* 1991; Funnell and Gagnier, 1995). A systematic screen, however, has never been performed and the proteins tested were selected only on the basis of their function.

Furthermore, chromosomal ParABS systems from different organisms were seen to cooperate with proteins in the cell to faithfully segregate sister chromosomes at cell division, implying that similar interactions could also be involved in plasmid segregation (Yamaichi and Niki, 2000b). For instance, ParB from *B. subtilis* (Spo0J), *C.crescentus, S.pneumoniae* were shown to interact with SMC and recruit it to the origin of replication by indirect association with *parS* (Sullivan *et al.,* 2009; Gruber and Errington, 2009; Minnen *et al*., 2011; Jensen and Shapiro, 1999). In addition, both ParA and ParB were shown to interact with proteins involved in tethering the chromosome to the membrane. PopZ in *C. crescentus*, for example, binds to ParB and connects the chromosome origin to the inner membrane (Bowman *et al*., 2008). In *V. cholera*, instead, the protein HubP

tethers the partition complex to the cell membrane by interacting with ParA1(Yamaichi *et al*., 2012).

In this work, tandem affinity purification coupled to mass-spectrometry was used to investigate the existence of factors in *E. coli* interacting with ParF and ParG. Since ParG is a small protein, interacting with *parH* and ParF and, therefore, less likely to bind to any other protein, more effort was dedicated in understanding ParF interaction network. In the past, a gene suppressor screen with a *Salmonella* library was used to identify putative binding partners of ParF in ParF overexpressing cells (unpublished results). Overproduction of one of the partition proteins causes plasmid instability. The aim of the experiment was to identify factors, among the genes expressed by the *Salmonella* library, that could interact with ParF and suppress the unbalance caused by *parF* overexpression. However, the gene suppressor screen proved to be hard to interpret, since many of the genes identified expressed enzymes involved in metabolic pathways apparently unrelated to plasmid segregation.

On the other hand, TAP/MS allowed the identification of numerous proteins that would be worth investigating further. Among the proteins pulled down with ParF and, therefore, putative ParF-binding partners, it was possible to find: cell division proteins, inner and outer membrane proteins, DNA-associated proteins and different enzymes involved in metabolic pathways. Amongst them, it was particularly interesting to find proteins such as FtsZ, MreB and MukB, which were objects of interest in the past (Ebersbach *et al*., 2006; Funnell and Gagnier, 1995; Ezaki *et al*., 1991). Bacterial two hybrid assays showed that these proteins, as well as the DNA gyrase GyrB, do not directly interact with ParF. However, these proteins are encoded by constitutive genes and tagging may have impaired their activity. Therefore, it would be important to investigate these interactions further *in vitro*, by chemical cross-linking for example, before ruling out the involvement of these proteins in plasmid segregation.

Among the non-constitutive proteins identified, DamX, OmpC, OmpX, YdgA, PpiD were tested further. Plasmid stability was measured in *E. coli* cells where the genes encoding the proteins of interest were deleted. None of the deletions seemed to affect plasmid segregation, suggesting that these proteins may either not be involved or not be essential for plasmid segregation. Interestingly, OmpX was identified also in the gene suppressor screen (unpublished data). Nevertheless, the relevance, if it exists, of OmpX in plasmid segregation still needs to be confirmed. OmpC, YdgA and PpiD, instead, were probably pulled down by indirect interaction or by cell debris left in the crude extract. In particular,

PpiD is a prolyl isomerase (Stymest and Klappa, 2008; Dartigalongue and Raina, 1998; Matern *et al.*, 2010) that was initially hypothesised to isomerise the proline residues in ParF proline-rich motif. However, the protein participates in proteins' folding and may have been pulled down by any other factor.
DamX, instead, is a factor that would be interesting to investigate further. DamX is a membrane protein which binds to peptidoglycan and participates in cell division (Yahashiri *et al.*, 2017). The protein is still poorly characterised and may represents a factor that connects plasmid segregation to cell division, although it is not essential in the process. However, DamX also weakly interact with FtsZ and other Fts proteins (Arends *et al*., 2010) and could have been pulled down indirectly.

Although TAP/MS is a useful technique to obtain a general idea regarding the interaction network of a protein of interest, it is hard to say whether the proteins identified are pulled down by direct or indirect interaction with the bait protein. The fact that the experiment was performed in triplicate allowed ignoring proteins pulled down randomly. Instead, the negative controls were useful to exclude proteins that are pulled down by the resin or the tag, rather than the bait protein. All the proteins selected for further investigation were only found in ParF pull-downs, in relatively high amounts. However, several proteins have been overlooked because of time constraint. As an example, MinD was found among the proteins pulled down by ParF, although in lower amount compared to other proteins. MinD is a Walker type ATPase involved in the positioning of the Z-ring at mid cell, inhibiting the assembly of FtsZ at the cell poles through the interaction with MinE and MinC (Lutkenhaus, 2012b). Since MinD tagging was proved to be challenging in the past (unpublished), additional experiments were not attempted. MinD and ParF, however, belong to the same family of ATPases and share structural similarities, suggesting a possible interaction between the two proteins.

All things considered, further investigation is required to assess which of the proteins identified through TAP/MS interacts with ParF and takes part in plasmid segregation. However, evidence suggests that the ParFGH system works independently from any host-encoded factor.

6.1.4 Model for plasmid segregation

Plasmids, as well as chromosomes of different bacterial species, rely on ParABS systems for proper distributionat cell division. In the past years evidence allowed to formulate two models, both of which rely on the use of chromosome as a scaffold for plasmid movement. The chromosome is therefore the main host factor aiding the

partition systems during segregation. The first model hypothesises the formation of polymers on the nucleoid which extend stochastically and depolymerise upon interaction with ParB bound to the *parS* site on the plasmid. ParB stimulates ParA ATPase activity, leading to polymers dissociation and pulling of the plasmids to opposite poles of the cell (Ringgaard *et al*., 2009). The second model suggests a diffusion-ratchet mechanism where ParA forms a gradient over the nucleoid that recruits and directs the movement of the plasmid cargo towards the cell pole (Vecchiarelli *et al.*, 2013b; 2014a; Hwang *et al.*, 2013). This model does not include the formation of polymers, instead ParA-ATP associates with the nucleoid in the form of dimers. ParB, bound to *parS* on the plasmid, enhancing ParA ATPase activity, dictates the organisation of the gradient for plasmid movement (Vecchiarelli *et al.*, 2013b; 2014a; Hwang *et al.*, 2013). Recently a DNA-relay mechanism has been suggested to describe chromosome segregation in *C. crescentus*. This model amplifies the diffusion-ratchet mechanism and includes the elastic force of the chromosome as a prerequisite for directed movement of the ParB-*parS* cargo (Lim *et al.*, 2014).
In the past few years the idea that plasmids moved on the nucleoid surface and were excluded from its volume was revolutionised by the evidence that plasmids are segregated within the nucloid volume (Le Gall *et al*., 2016; McLeod *et al*., 2017).

The most recent model suggested for the segregation of the TP228 plasmid was put forward by McLeod *et al*. The Venus-flytrap model combines and amplify the previously identified polymerisation and diffusion-ratchet mechanisms (McLeod *et al.*, 2017). According to the Venus flytrap model, ParF forms a meshwork, composed of polymers and small oligomers, within the nucleoid volume. The sister plasmids, bound to ParG through the *parH* site, become trapped within the nucleoid by the ParF meshwork and are transported to the nucleoid pole. The meshwork has a more compact (leading) and a less dense (lagging) end. As the meshwork moves inside the nucleoid and the plasmid cargo with it, ParF oligomers grow between the sister plasmids. ParG stimulates ParF ATPase activity leading to the dissociation of oligomers in the vicinity of the ParG-*parH* site. Eventually one of the plasmid is release at one nucleoid pole, while the other is transported and dropped at the opposite pole by a similar mechanism (McLeod *et al*., 2017).

Although no new knowledge on the mechanism of segregation was achieved through this work, the study on the molecular interaction allowed plotting a more precise picture regarding segrosome formation.

After replication, sister plasmids align in the middle of the cell, held together by ParG bound to the centromeric site *parH*. Interaction between ParG and the DNA causes a change in ParG conformation and activate the tails for interaction with ParF. In the cell, ParG is probably present in two distinct forms: bound to *parH* or dispersed within the nucleoid, associated with ParF polymers. The activity of the protein on ParF was seen to be dependent on ParG concentration. At high concentration, hence on the centromeric site, ParG stimulates ParF ATPase activity. At low concentration, around the nucleoid volume, ParG can instead promotes ParF polymerisation and bundling (Barillà *et al*., 2005). The structure of the tail could also contribute to this behaviour. Binding of ParG to the *parH* site and ParF to the nucleoid make both proteins proficient for interaction with each other. In this way, the plasmid cargo associated to ParG becomes trapped within the nucleoid volume and can be transported to the nucleoid pole by the ParF meshwork.

ParG dictates timing of segregation and ParF oscillation through the N-terminal domain. Binding of the protein to the DNA triggers the conformational change in the N-terminal region required for proficient association with ParF. The plasmid can be tethered to ParF thanks to the tip of tails which inserts within the ParF-ATP dimers. Interaction between ParF and the cargo is much stronger when the protein is in ATP-bound state, allowing the cargo to discriminate which region to anchor to. Once bound together, ParG enhances ParF ATPase activity through the arginine finger motif which localises at the ParF monomer-monomer interface, close to the ATP-binding pocket. ATP hydrolysis causes the ParG bound cargo to be temporarily released and to associate again with ParF-ATP in the meshwork. As the ParF meshwork grows and remodels thanks to ParG stimulation of ParF catalytic activity, the ParF platform on the nucleoid can oscillate and release the plasmid at opposite nucleoid poles.

6.2 Future work

Although the segregation of the TP228 plasmid has been object of research for the past two decades and extensive knowledge has been built on the system, there are still mechanistic questions that remain to be answered.

The segregation mechanism starts with the replication of the plasmids at midcell and recognition of the centromeric sites on the sister plasmids by ParG dimers. In this work, it was demonstrated that interaction with the cognate DNA causes a conformational change in the protein that allows ParG to interact with ParF with higher affinity. One aspect that would be interesting to explore further is ParG conformational change by using Förster resonance energy transfer (FRET). The principle FRET is based on is the transfer of energy between two fluorescent dye molecules, a donor and an acceptor. The

energy transfer allows then to measure the distance between the two fluorophores (Roy *et al.*, 2008). FRET would be useful to investigate structural changes within the ParG dimer by labelling different regions of the protein (Truong and Ikura, 2001). It would also be useful to demonstrate whether the tails could wrap the nucleic acid, by labelling ParG N-terminus and the DNA (Blouin *et al.*, 2009). FRET could even be used to investigate the assembly of the segrosome by three colour spectral FRET (3s-FRET) (Sun *et al.*, 2010).

A parallel way to investigate ParG conformational change is by using NMR. HSQC NMR spectroscopy has been a useful tool to investigate the interaction between ParG and the operator site O_F. Although *parH* and O_F are composed by the same repeated motif, different in orientation may be important in discriminating between ParG role as transcriptional repressor and partition protein. It would be interesting to use HSQC NMR to see whether ParG behaves in the same fashion in presence of the two different regions of the DNA. It would also informative to use ParG triple mutant to investigate the conformational change caused by the interaction with the DNA.

Furthermore, it would be interesting to study plasmid segregation by single cell imaging. Super-resolution microscopy was instrumental to draw the Venus fly trap model. However, single molecule resolution microscopy, such as SLIM-fied (**S**patial **L**ight **I**nterference **M**icroscopy) microscopy would allow to map plasmid segregation process with high precision.

Other interesting questions to answer concern the loading of ParG and ParF on specific and non-specific DNA. *In vivo* NMR analysis on MetJ revealed that the protein recognise non-specific DNA and seeks for its binding site (Augustus *et al.*, 2009). Similar *in vivo* NMR approach could be used to investigate how ParG recognise its specific site. In addition, magnetic tweezers represent a great single molecule tool to investigate how the proteins interact with specific and non-specific DNA.
Beside the fact that ParA proteins stochastically associate with higher density regions of the chromosome (Hu *et al*., 2017), very little is known about non-specific binding of the motor protein to the DNA. ChIP sequencing would be an informative technique to understand whether the ParF associates preferentially to specific regions of the chromosome and whether ParG enhances this association.

Finally, it would be interesting to investigate further the proteins identified through TAP/MS starting from the purification of FtsZ, MreB, MukB as well as MinD to test the

interaction *in vitro*. Among the over 100 proteins identified with ParF, only a small portion has been investigated. Measurement of plasmid stability in Keio collection mutants was useful to exclude proteins which do not take part in the process. The same approach could be applied for other identified proteins.

Abbreviations

3D - three dimentional

3D-SIM - 3 dimensional structured illumination microscopy

°C - Degrees Celsius

µg - Microgram

µl - Microlitres

µM - Micromolar

A - Absorbance

ADP - Adenosine diphosphate

Amp – Ampicillin

AMPPCP - phosphomethylphosphonic acid adenylate ester

APS - Ammonium persulphate

ATP - adenosine 5-triphosphate

ATPase - Adenosine triphosphate hydrolase

Bp - Base pair

BSA - Bovine serum albumin

Cm - Chloramphenicol

CBP - Centromere binding protein

CD – Circular Dichroism

CL-MS – Cross-linking – Mass Spectrometry

CV – Column volume

DLS - Dynamic light scattering

DMP - Dimethyl pimelimidate

DNA - Dexoyribonuclic acid

dNTP - Dexoyribonucleoside-5’-triphosphate

dsDNA - Double stranded dexoyribonuclic acid

nsDNA – Non-specific DNA

EDC - 1-Ethyl-3-(3-dimethylaminopropyl) carbodiimide hydrochlorid

EDTA - Ethylenediaminetetraacetic acid

EGTA – Ethylene glycol tetraacetic acid

EM - Electron microscopy

EMSA - Electrophoretic mobility shift assay

Fc – Flow cell

FIS - Factor for inversion stimulation

Fnorm – normalised fluorescence

FRET - Förster resonance energy transfer

GTP - Guanosine triphosphate

GTPase - Guanosine triphosphate hydrolase

His - Histidine

HTH - Helix-turn-helix motif

IHF - Integration host factor

HSQC - heteronuclear single quantum coherence

IPTG - Isopropyl beta-D-1-thiogalactopyranoside

IR – infra red

L - Litre

LB - Luria-Bertani medium

LC-MS – Liquid chromatography-mass spectrometry

LC-MS/MS – Liquid chromatography-tandem mass spectrometry

k_a - association rate constant

Kan - Kanamycin

kbp - kilobase pairs

K_D – Equilibrium dissociation constant

k_d - dissociation rate constant

kDa - Kilodaltons

GFP - Green Fluorescent Protein

M - Molar

m^o - millidegree

MCS – Multicloning site

mg - Milligrams

Min – Minutes

mK – Millikelvin

MST – MicroScale Thermophoresis

NHS - N-Hydroxysuccinimide

NMR – Nuclear Magnetic Resonance

PCR – Polymerase chain reaction

Poly (dI-dC) - Poly(deoxyinosinic-deoxycytidylic) acid

RacA- Remodeling and Anchoring of the Chromosome

RCD - Regulator of Cell Division

RU - Response Units

SDS-PAGE - Sodium dodecyl sulphate-polyacrylamide gel electrophoresis

SLIM - Spatial Light Interference Microscopy

SMC - Structural maintenance of the chromosome

SPR – Surface Plasmon Resonance

SSB - Single strand binding protein

TA - Toxin-antitoxin

TAP – Tandem Affinity Purification

TAP-MS – Tandem affinity purification – mass spectrometry

TIRFM - Total internal reflection microscopy

YFP- Yellow Fluorescent Protein

References

Abeles, A. L. and Austin, S. J. (1987). P1 plasmid replication requires methylated DNA. *The EMBO Journal*, 6 (10), pp.3185–3189.

Abeles, A. L., Friedman, S. A. and Austin, S. J. (1985). Partition of unit-copy miniplasmids to daughter cells. III. The DNA sequence and functional organization of the P1 partition region. *Journal of Molecular Biology*, 185 (2). 261-72.

Ah-Seng, Y., Lopez, F., Pasta, F., Lane, D. and Bouet, J.-Y. (2009). Dual role of DNA in regulating ATP hydrolysis by the SopA partition protein. *The Journal of Biological Chemistry*, 284 (44), pp.30067–30075.

Allison-Gamble, G. (2016). Dissecting the role of the partition factor ParF in TP228 plasmid segregation. *PhD Thesis.*

Arends, S. J. R., Williams, K., Scott, R. J., Rolong, S., Popham, D. L. and Weiss, D. S. (2010). Discovery and characterization of three new Escherichia coli septal ring proteins that contain a SPOR domain: DamX, DedD, and RlpA. *Journal of Bacteriology*, 192 (1), pp.242–255.

Arends, S. J., Williams, K., Scott, R. J., Rolong, S., Popham, D. L. and Weiss, D. S. (2010). Discovery and characterization of three new Escherichia coli septal ring proteins that contain a SPOR domain: DamX, DedD, and RlpA. *Journal of Bacteriology*, 192 (1), pp.242–255.

Arnoud Marquart. (2016). *Surface Plasmon Resonance and Biomolecular Interaction Analysis*. 4th ed.

Augustus, A. M., Reardon, P. N. and Spicer, L. D. (2009). MetJ repressor interactions with DNA probed by in-cell NMR. Proceedings of the National Academy of Sciences, 106 (13), 5065-5069.

Austin, S. and Abeles, A. (1983). Partition of unit-copy miniplasmids to daughter cells: I. P1 and F miniplasmids contain discrete, interchangeable sequences sufficient to promote equipartition. *Journal of Molecular Biology*, 169 (2), pp.353–372.

Austin, S. and Wierzbickj, A. (1983). Two Mini-F-Encoded Proteins Are Essential for Equipartition. *Plasmid*, 10 (1), pp.73–81.

Aylett, C. H. S., Wang, Q., Michie, K. A., Amos, L. A. and Löwe, J. (2010). Filament structure of bacterial tubulin homologue TubZ. *Proceedings of the National Academy of Sciences*, 107 (46), pp.19766–19771.

van Baarle, S., Celik, I. N., Kaval, K. G., Bramkamp, M., Hamoen, L. W. and Halbedel, S. (2013). Protein-protein interaction domains of Bacillus subtilis DivIVA. *Journal of Bacteriology*, 195 (5), pp.1012–1021.

Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L. and Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: The Keio collection. *Molecular Systems Biology*, 2 (1), p.2006.0008.

Barge, M. (2015). Role of the unstructured N-terminus of the centromere binding protein ParG in mediating segregation of the multidrug resistance plasmid TP228. *PhD Thesis*.

Barillà, D., Carmelo, E. and Hayes, F. (2007). The tail of the ParG DNA segregation protein remodels ParF polymers and enhances ATP hydrolysis via an arginine finger-like motif. *Proceedings of the National Academy of Sciences*, 104 (6), pp.1811–1816.

Barillà, D. and Hayes, F. (2003). Architecture of the ParF*ParG protein complex involved in prokaryotic DNA segregation. *Molecular Microbiology*, 49 (2), pp.487–499.

Barillà, D., Rosenberg, M. F., Nobbmann, U. and Hayes, F. (2005). Bacterial DNA segregation dynamics mediated by the polymerizing protein ParF. *The EMBO Journal*, 24 (7), pp.1453–1464.

Baslé, A., Rummel, G., Storici, P., Rosenbusch, J. P. and Schirmer, T. (2006). Crystal Structure of Osmoporin OmpC from E. coli at 2.0 Å. *Journal of Molecular Biology*, 362 (5), pp.933–942.

Baxter, J. C. and Funnell, B. E. (2014). Plasmid Partition Mechanisms. *Microbiology Spectrum*, 2 (6).

Ben-Yehuda, S., Fujita, M., Liu, X. S., Gorbatyuk, B., Skoko, D., Yan, J., Marko, J. F., Liu, J. S., Eichenberger, P., Rudner, D. Z., Losick R. (2005). Defining a Centromere-like Element in Bacillus subtilis by Identifying the Binding Sites for the Chromosome-Anchoring Protein RacA. *Molecular Cell*, 17 (6), pp.773–782.

Bharat, T. A. M., Murshudov, G. N., Sachse, C. and Löwe, J. (2015). Structures of actin-like ParM filaments show architecture of plasmid-segregating spindles. *Nature*, 523 (7558), pp.106–110.

Biek, D. P. and Shi, J. (1994). A single 43-bp sopC repeat of plasmid mini-F is sufficient to allow assembly of a functional nucleoprotein partition complex. *Proceedings of the National Academy of Sciences*, 91 (17), pp.8027–8031.

Biek, D. P. and Strings, J. (1995). Partition functions of mini-F affect plasmid DNA topology in Escherichia coli. *Journal of Molecular Biology*, 246 (3), pp.388–400.

Bingle, L. E. H., Rajasekar, K. V., Muntaha, S. tul, Nadella, V., Hyde, E. I. and Thomas, C. M. (2008). A single aromatic residue in transcriptional repressor protein KorA is critical for cooperativity with its co-regulator KorB. *Molecular Microbiology*, 70 (6), pp.1502–1514.

Bisson-Filho, A. W., Hsu, Y.-P., Squyres, G. R., Kuru, E., Wu, F., Jukes, C., Sun, Y., Dekker, C., Holden, S., VanNieuwenhze, M. S., Brun Y. V., Garner, E. C. (2017). Treadmilling by FtsZ filaments drives peptidoglycan synthesis and bacterial cell division. *Science*, 355 (6326), pp.739–743.

Blouin, S., Craggs, T. D., Lafontaine, D. A. and Penedo, J. C. (2009). Functional Studies of DNA-Protein Interactions Using FRET Techniques. *Methods in Molecular Biology.* 543, pp.475–502.

Bouet, J.-Y., Ah-Seng, Y., Benmeradi, N. and Lane, D. (2007). Polymerization of SopA partition ATPase: regulation by DNA binding and SopB. *Molecular Microbiology*, 63 (2), pp.468–481.

Bouet, J. Y., Funnell, B. E., Buc, H. and Busby, S. (1999). P1 ParA interacts with the P1 partition complex at parS and an ATP-ADP switch controls ParA activities. *The EMBO Journal*, 18 (5), pp.1415–1424.

Bouet, J. Y. and Lane, D. (2009). Molecular basis of the supercoil deficit induced by the mini-F plasmid partition complex. *The Journal of Biological Chemistry*, 284 (1), pp.165–173.

Bouet, J. Y., Surtees, J. A. and Funnell, B. E. (2000). Stoichiometry of P1 plasmid partition complexes. *The Journal of Biological Chemistry*, 275 (11), pp.8213–8219.

Bowman, G. R., Comolli, L. R., Gaietta, G. M., Fero, M., Hong, S. H., Jones, Y., Lee, J. H., Downing, K. H., Ellisman, M. H., McAdams, H. H., Shapiro L. (2010). Caulobacter PopZ forms a polar subdomain dictating sequential changes in pole composition and function. *Molecular Microbiology*, 76 (1), pp.173–189.

Bowman, G. R., Comolli, L. R., Zhu, J., Eckart, M., Koenig, M., Downing, K. H., Moerner, W. E., Earnest, T. and Shapiro, L. (2008). A polymeric protein anchors the chromosomal origin/ParB complex at a bacterial cell pole. *Cell*, 134 (6), pp.945–955.

Breier, A. M. and Grossman, A. D. (2007). Whole-genome analysis of the chromosome partitioning and sporulation protein Spo0J (ParB) reveals spreading and origin-distal sites on the Bacillus subtilis chromosome. *Molecular Microbiology*, 64 (3), pp.703–718.

Broedersz, C. P., Wang, X., Meir, Y., Loparo, J. J., Rudner, D. Z. and Wingreen, N. S. (2014). Condensation and localization of the partitioning protein ParB on the bacterial chromosome. *Proceedings of the National Academy of Sciences*, 111 (24), pp.8809–8814.

Brooks, A. C. and Hwang, L. C. (2017). Reconstitutions of plasmid partition systems and their mechanisms. *Plasmid*, 91, pp.37–41.

Brown, B. M., Bowie, J. U. and Sauer, R. T. (1990). Arc repressor is tetrameric when bound to operator DNA. *Biochemistry*, 29 (51), pp.11189–11195.

Bürckstümmer, T., Bennett, K. L., Preradovic, A., Schütze, G., Hantschel, O., Superti-Furga, G. and Bauch, A. (2006). An efficient tandem affinity purification procedure for interaction proteomics in mammalian cells. *Nature Methods*, 3 (12), pp.1013–1019.

Carmelo, E., Barillà, D., Golovanov, A. P., Lian, L.-Y., Derome, A. and Hayes, F. (2005). The unstructured N-terminal tail of ParG modulates assembly of a quaternary nucleoprotein complex in transcription repression. *The Journal of Biological Chemistry*, 280 (31), pp.28683–28691.

Castaing, J.-P., Bouet, J.-Y. and Lane, D. (2008). F plasmid partition depends on interaction of SopA with non-specific DNA. *Molecular Microbiology*, 70 (4), pp.1000–1011.

Chant, E. L. and Summers, D. K. (2007). Indole signalling contributes to the stable maintenance of Escherichia coli multicopy plasmids. *Molecular Microbiology*, 63 (1), pp.35–43.

Chen, B. W., Lin, M. H., Chu, C. H., Hsu, C. E. and Sun, Y. J. (2015). Insights into ParB spreading from the complex structure of Spo0J and parS. *Proceedings of the National Academy of Sciences*, 112(21), pp.6613–6618.

Clarke, N. D., Beamer, L. J., Goldberg, H. R., Berkower, C. and Pabo, C. O. (1991). The DNA Binding Arm of A Repressor: Critical Contacts from a Flexible Region. *Science*, 254 (5029), pp.267–270.

Collins, M. O. and Choudhary, J. S. (2008). Mapping multiprotein complexes by affinity purification and mass spectrometry. *Current Opinion in Biotechnology*, 19 (4), pp.324–330.

Cui, Y., Petrushenko, Z. M. and Rybenkov, V. V. (2008). MukB acts as a macromolecular

clamp in DNA condensation. *Nature Structural & Molecular Biology*, 15 (4), pp.411–418.

Dam, M. and Gerdes, K. (1994). Partitioning of plasmid R1 Ten direct repeats flanking the parA promoter constitute a centromere-like partition site parC, that expresses incompatibility. *Journal of Molecular Biology*, 236 (5), pp.1289–1298.

Danilova, O., Reyes-Lamothe, R., Pinskaya, M., Sherratt, D. and Possoz, C. (2007). MukB colocalizes with the oriC region and is required for organization of the two Escherichia coli chromosome arms into separate cell halves. *Molecular Microbiology*, 65 (6), pp.1485–1492.

Dartigalongue, C. and Raina, S. (1998). A new heat-shock gene, ppiD, encodes a peptidyl-prolyl isomerase required for folding of outer membrane proteins in Escherichia coli. EMBO Journal, 17 (14), 3968-80.

Davey, M. J. and Funnell, B. E. (1997). Modulation of the P1 plasmid partition protein ParA by ATP, ADP, and P1 ParB. *The Journal of Biological Chemistry*, 272 (24), pp.15286–15292.

Davey, M. J. and Funnells, B. E. (1994). The P1 Plasmid Partition Protein ParA. A role for ATP in site-specific DNA binding. *The Journal of Biological Chemistry*, 269 (47), pp.29908–29913.

Davis, M. A., Martin, K. A. and Austin, S. J. (1990). Specificity switching of the P1 plasmid centromere-like site. *The EMBO Journal*, 9 (4), pp.991–998.

Davis, M. A., Martin, K. A. and Austin, S. J. (1992). Biochemical activities of the ParA partition protein of the P1 plasmid. *Molecular Microbiology*, 6 (9), pp.1141–1147.

Ditkowski, B., Holmes, N., Rydzak, J., Donczew, M., Bezulska, M., Ginda, K., Kedzierski, P., Zakrzewska-Czerwińska, J., Kelemen, G. H. and Jakimowicz, D. (2013). Dynamic interplay of ParA with the polarity protein, Scy, coordinates the growth with chromosome segregation in Streptomyces coelicolor. *Open Biology*, 3 (3), p.130006.

Dmowski, M. and Jagura-Burdzy, G. (2011). Mapping of the interactions between partition proteins Delta and Omega of plasmid pSM19035 from Streptococcus pyogenes. *Microbiology*, 157 (4), pp.1009–1020.

Dmowski, M., Sitkiewicz, I. and Ceglowski, P. (2006). Characterization of a novel partition system encoded by the delta and omega genes from the streptococcal plasmid pSM19035. *Journal of Bacteriology*, 188 (12), pp.4362–4372.

Dostál, L., Pratto, F., Alonso, J. C. and Welfle, H. (2007). Binding of regulatory protein omega fromStreptococcus pyogenes plasmid pSM19035 to direct and inverted 7-base pair repeats of operator DNA. *Journal of Raman Spectroscopy*, 38 (2), pp.166–175.

Duhr, S. and Braun, D. (2006). Why molecules move along a temperature gradient. *Proceedings of the National Academy of S*ciences, 103 (52), pp.9678-82.

Dunham, T. D., Xu, W., Funnell, B. E. and Schumacher, M. A. (2009). Structural basis for ADP-mediated transcriptional regulation by P1 and P7 ParA. *The EMBO Journal*, 28, pp.1792–1802.

Dyson, H. J. and Wright, P. E. (2002). Coupling of folding and binding for unstructured proteins. *Current Opinion in Structural Biology*, 12 (1), pp.54-60.

Ebersbach, G. and Gerdes, K. (2001). The double par locus of virulence factor pB171: DNA segregation is correlated with oscillation of ParA. *Proceedings of the National Academy of*

Sciences, 98 (26), pp.15078–15083.

Ebersbach, G. and Gerdes, K. (2004). Bacterial mitosis: partitioning protein ParA oscillates in spiral-shaped structures and positions plasmids at mid-cell. *Molecular Microbiology*, 52 (2), pp.385–398.

Ebersbach, G., Ringgaard, S., Moller-Jensen, J., Wang, Q., Sherratt, D. J. and Gerdes, K. (2006). Regular cellular distribution of plasmids by oscillating and filament-forming ParA ATPase of plasmid pB171. *Molecular Microbiology*, 61 (6), pp.1428–1442.

van den Ent, F., Møller-Jensen, J., Amos, L. A., Gerdes, K. and Löwe, J. (2002). F-actin-like filaments formed by plasmid segregation protein ParM. *The EMBO journal*, 21 (24), pp.6935–6943.

Erdmann, N., Petroff, T. and Funnell, B. E. (1999). Intracellular localization of P1 ParB protein depends on ParA and parS. *Proceedings of the National Academy of Sciences*, 96(26), pp.14905-10.

Ezaki, B., Ogura, T., Niki, H. and Hiraga, S. (1991). Partitioning of a mini-F plasmid into anucleate cells of the mukB null mutant. *Journal of Bacteriology*, 173 (20), pp.6643–6646.

Figge, R. M., Easter, J. and Gober, J. W. (2003). Productive interaction between the chromosome partitioning proteins, ParA and ParB, is required for the progression of the cell cycle in Caulobacter crescentus. *Molecular Microbiology*, 47 (5), pp.1225–1237.

Fink, G. and Löwe, J. (2015). Reconstitution of a prokaryotic minus end-tracking system using TubRC centromeric complexes and tubulin-like protein TubZ filaments. *Proceedings of the National Academy of Sciences*, 112 (15), pp.1845-50.

Fisher, G. L., Pastrana, C. L., Higman, V. A., Koh, A., Taylor, J. A., Butterer, A., Craggs, T., Sobott, F., Murray, H., Crump, M. P., Moreno-Herrero F., Dillingham M. S. (2017). The structural basis for dynamic DNA binding and bridging interactions which condense the bacterial centromere. eLife, 6, e28086.

Fothergill, T. J. G., Barillà, D. and Hayes, F. (2005). Protein diversity confers specificity in plasmid segregation. *Journal of Bacteriology*, 187 (8), pp.2651–2661.

Friedman, S. A. and Austin, S. J. (1988). The Pl Plasmid-Partition System Synthesizes Two Essential Proteins from an Autoregulated Operon'. *Plasmid*, 19, pp.103–112.

Funnell, B. E. (1988a). Mini-P1 plasmid partitioning: excess ParB protein destabilizes plasmids containing the centromere parS. *Journal of Bacteriology*, 170 (2), pp.954–960.

Funnell, B. E. (1988b). Participation of Escherichia coli integration host factor in the P1 plasmid partition system. *Proceedings of the National Academy of Sciences*, 85 (18), pp.6657–6661.

Funnell, B. E. and Gagnier, L. (1994). P1 plasmid partition: binding of P1 ParB protein and Escherichia coli integration host factor to altered parS sites. *Biochimie*, 76 (10–11), pp.924–932.

Funnell, B. E. and Gagnier, L. (1995). Partition of P1 plasmids in Escherichia coli mukB chromosomal partition mutants. *Journal of Bacteriology*, 177 (9), pp.2381–2386.

Funnell, B. E. (1991). The P1 Plasmid Partition Complex at pars. The influence of Escherichia coli integration host factor and of substrate topology. *The Journal of Biological Chemistry*, 266 (22), pp.14328–14337.

Funnell, B. E. and Gagnier, L. (1993). The P1 Plasmid Partition Complex at parS: Analysis of ParB protein binding activity and specificity. *The Journal of Biological Chemistry*, 268 (5), pp.3616–3624.

Le Gall, A., Cattoni, D. I., Guilhas, B., Mathieu-Demazière, C., Oudjedi, L., Fiche, J. B., Rech, J., Abrahamsson, S., Murray, H., Bouet, J. Y., Nollmann M. (2016). Bacterial partition complexes segregate within the volume of the nucleoid. Nature Communications, 7, 12107

Garner, E. C., Campbell, C. S., Weibel, D. B. and Mullins, R. D. (2007). Reconstitution of DNA segregation driven by assembly of a prokaryotic actin homolog. *Science*, 315 (5816), pp.1270–1274.

Gayathri, P., Fujii, T., Møller-Jensen, J., Van Den Ent, F., Namba, K. and Löwe, J. (2012). A bipolar spindle of antiparallel ParM filaments drives bacterial plasmid segregation. *Science*, 338 (6112), pp.1334–1337.

Gellert, M., Mizuuchi, K., O'Dea, M. H. and Nash, H. A. (1976). DNA gyrase: an enzyme that introduces superhelical turns into DNA. *Proceedings of the National Academy of Sciences*, 73 (11), pp.3872–3876.

Gerdes, K., Howard, M. and Szardenings, F. (2010). Pushing and pulling in prokaryotic DNA segregation. *Cell*, 141 (6), pp.927–942.

Gerdes, K., Larsen, J. E. and Molin, S. (1985). Stable inheritance of plasmid R1 requires two different loci. *Journal of Bacteriology*, 161 (1), American Society for Microbiology., pp.292–298.

Gerdes, K., Møller-Jensen, J. and Jensen, R. B. (2002). Plasmid and chromosome partitioning: surprises from phylogeny. *Molecular Microbiology*, 37 (3), pp.455–466.

Gerdes, K., Rasmussen, P. B. and Molin, S. (1986). Unique type of plasmid maintenance function: postsegregational killing of plasmid-free cells. *Proceedings of the National Academy of Sciences*, 83 (10), pp.3116–3120.

Gerdest, K. and Molin, S. (1986). Partitioning of Plasmid RI Structural and Functional Analysis of the parA Locus. *Journal of Molecular Biology*, 190(3), 269-279.

Ghosh, S. K., Hajra, S., Paek, A. and Jayaram, M. (2006). Mechanisms for Chromosome and Plasmid Segregation. *Annual Review of Biochemistry*, 75 (1), pp.211–241.

Gloeckner, C. J., Boldt, K., Schumacher, A., Roepman, R. and Ueffing, M. (2007). A novel tandem affinity purification strategy for the efficient isolation and characterisation of native protein complexes. *Proteomics*, 7 (23), pp.4228–4234.

Gloeckner, C. J., Boldt, K., Schumacher, A. and Ueffing, M. (2009). Tandem Affinity Purification of Protein Complexes from Mammalian Cells by the Strep/FLAG (SF)-TAP Tag. *Methods in Molecular Biology,* 564, pp.359-72.

Godfrin-Estevenon, A.-M., Pasta, F. and Lane, D. (2002). The parAB gene products of Pseudomonas putida exhibit partition activity in both P. putida and Escherichia coli. *Molecular Microbiology*, 43 (1), pp.39–49.

Golovanov, A. P., Barillà, D., Golovanova, M., Hayes, F. and Lian, L.-Y. (2003). ParG, a protein required for active partition of bacterial plasmids, has a dimeric ribbon-helix-helix structure. *Molecular Microbiology*, 50 (4), pp.1141–1153.

Gordon, G. S., Sitnikov, D., Webb, C. D., Teleman, A., Straight, A., Losick, R., Murray, A.

W. and Wright, A. (1997). Chromosome and Low Copy Plasmid Segregation in E. coli: Visual Evidence for Distinct Mechanisms. *Cell*, 90 (6), pp.1113–1121.

Graham, T. G. W., Wang, X., Song, D., Etson, C. M., van Oijen, A. M., Rudner, D. Z. and Loparo, J. J. (2014). ParB spreading requires DNA bridging. *Genes & Development*, 28 (11), pp.1228–1238.

Greenfield, N. J. (2007). Using circular dichroism spectra to estimate protein secondary structure. *Nature Protocols*, 1 (6), pp.2876–2890.

Gruber, S. and Errington, J. (2009). Recruitment of Condensin to Replication Origin Regions by ParB/SpoOJ Promotes Chromosome Segregation in B. subtilis. *Cell*, 137 (4), pp.685–696.

Gully, D., Moinier, D., Loiseau, L. and Bouveret, E. (2003). New partners of acyl carrier protein detected in Escherichia coli by tandem affinity purification. *FEBS letters*, 548 (1–3), pp.90–96.

Guzman, L. M., Belin, D., Carson, M. J. and Beckwith, J. (1995). Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter. *Journal of Bacteriology*, 177 (14), pp.4121–4130.

Hanai, R., Liu, R., Benedetti, P., Caron, P. R., Lynch, A. S. and Wang, J. C. (1996). Molecular dissection of a protein SopB essential for Escherichia coli F plasmid partition. *The Journal of Biological Chemistry*, 271 (29), pp.17469–17475.

Harms, A., Brodersen, D. E., Mitarai, N. and Gerdes, K. (2018). Toxins, Targets, and Triggers: An Overview of Toxin-Antitoxin Biology. *Molecular Cell*, 70 (5), pp.768–784.

Hatano, T. and Niki, H. (2010). Partitioning of P1 plasmids by gradual distribution of the ATPase ParA. *Molecular Microbiology*, 78 (5), pp.1182–1198.

Hatano, T., Yamaichi, Y. and Niki, H. (2007). Oscillating focus of SopA associated with filamentous structure guides partitioning of F plasmid. *Molecular Microbiology*, 64 (5), pp.1198–1213.

Hayakawa, Y., Murotsu, T. and Matsubara, K. (1985). Mini-F protein that binds to a unique region for partition of mini-F plasmid DNA. *Journal of Bacteriology*, 163 (1), pp.349–354.

Hayes, F. (2000). The partition system of multidrug resistance plasmid TP228 includes a novel protein that epitomizes an evolutionarily distinct subgroup of the ParA superfamily. *Molecular Microbiology*, 37 (3), pp.528–541.

Hayes, F. (2002). The partition system of multidrug resistance plasmid TP228 includes a novel protein that epitomizes an evolutionarily distinct subgroup of the ParA superfamily. *Molecular Microbiology*, 37 (3), pp.528–541.

Hayes, F. (2003a). The Function and Organization of Plasmids. In: *E. coli Plasmid Vectors*. pp.1–18.

Hayes, F. (2003b). Toxins-Antitoxins: Plasmid Maintenance, Programmed Cell Death, Cell Cycle Arrest. *Science*, 301 (5639), pp.1496–1499.

Hayes, F. and Austin, S. (1993). Specificity determinants of the P1 and P7 plasmid centromere analogs. *Proceedings of the National Academy of Sciences,* 90 (19), pp.9228–9232.

Hayes, F. and Austin, S. (1994). Topological Scanning of the P1 Plasmid Partition Site.

Journal of Molecular Biology, 243, pp.190–198.

Hayes, F. and Barillà, D. (2006a). Assembling the bacterial segrosome. *Trends in Biochemical Sciences*, 31 (5), pp.247–250.

Hayes, F. and Barillà, D. (2006b). The bacterial segrosome: a dynamic nucleoprotein machine for DNA trafficking and segregation. *Nature Reviews Microbiology*, 4 (2), pp.133–143.

Hayes, F. and Van Melderen, L. (2011). Toxins-antitoxins: diversity, evolution and function. *Critical Reviews in Biochemistry and Molecular Biology*, 46 (5), pp.386–408.

Hiraga, S., Jaffé, A., Ogura, T., Mori, H. and Takahashi, H. (1986). F plasmid ccd mechanism in Escherichia coli. *Journal of Bacteriology*, 166 (1), pp.100–104.

Hoischen, C., Bussiek, M., Langowski, J. and Diekmann, S. (2008). Escherichia coli low-copy-number plasmid R1 centromere parC forms a U-shaped complex with its binding protein ParR. *Nucleic Acids Research*, 36 (2), pp.607–615.

Hu, L., Vecchiarelli, A. G., Mizuuchi, K., Neuman, K. C. and Liu, J. (2015). Directed and persistent movement arises from mechanochemistry of the ParA/ParB system. *Proceedings of the National Academy of Sciences*, 112 (51), p.201505147.

Hu, L., Vecchiarelli, A. G., Mizuuchi, K., Neuman, K. C. and Liu, J. (2017). Brownian Ratchet Mechanism for Faithful Segregation of Low-Copy-Number Plasmids. *Biophysical Journal*, 112 (7), pp.1489–1502.

Huang, L., Yin, P., Zhu, X., Zhang, Y. and Ye, K. (2011). Crystal structure and centromere binding of the plasmid segregation protein ParB from pCXC100. *Nucleic Acids Research*, 39 (7), pp.2954–2968.

Hwang, L. C., Vecchiarelli, A. G., Han, Y.-W., Mizuuchi, M., Harada, Y., Funnells, B. E. and Mizuuchi, K. (2013). ParA-mediated plasmid partition driven by protein pattern self-organization. *The EMBO Journal*, 32, pp.1238–1249.

Hyde, E. I., Callow, P., Rajasekar, K. V., Timmins, P., Patel, T. R., Siligardi, G., Hussain, R., White, S. A., Thomas, C. M. and Scott, D. J. (2017). Intrinsic disorder in the partitioning protein KorB persists after co-operative complex formation with operator DNA and KorA. *Biochemical Journal*, 474 (18), pp.3121–3135.

Ietswaart, R., Szardenings, F., Gerdes, K. and Howard, M. (2014). Competing ParA Structures Space Bacterial Plasmids Equally over the Nucleoid. Rao, C. V. (Ed). *PLoS Computational Biology*, 10 (12), e1004009.

Ireton, K., Gunther, N. W. and Grossman, A. D. (1994). Spo0J is required for normal chromosome segregation as well as the initiation of sporulation in Bacillus subtilis. *Journal of Bacteriology*, 176 (17), pp.5320–5329.

Jensen, R. B., Dam, M. and Gerdest, K. (1994). Partitioning of Plasmid Rl. *J. Mol. Biol*, 236, pp.1299–1309.

Jensen, R. B., Lurz, R. and Gerdes, K. (1998). Mechanism of DNA segregation in prokaryotes: replicon pairing by parC of plasmid R1. *Proceedings of the National Academy of Sciences*, 95 (15), pp.8550–8555.

Jensen, R. B. and Shapiro, L. (1999). The Caulobacter crescentus smc gene is required for cell cycle progression and chromosome segregation. *Proceedings of the National Academy of*

Sciences, 96 (19), pp.10661–10666.

Jerabek-Willemsen, M., Wienken, C. J., Braun, D., Baaske, P. and Duhr, S. (2011). Molecular interaction studies using microscale thermophoresis. *Assay and Drug Development Technologies*, 9 (4), pp.342–353.

Jones, C. S., Osborne, D. J. and Stanley, J. (1993). Molecular comparison of the IncX plasmids allows division into IncX1 and IncX2 subgroups. *Journal of General Microbiology*, 139 (4), pp.735–741.

Karimova, G., Pidoux, J., Ullmann, A. and Ladant, D. (1998). A bacterial two-hybrid system based on a reconstituted signal transduction pathway. *Proceedings of the National Academy of Sciences*, 95 (10), pp.5752–5756.

Khan, S. A. (2005). Plasmid rolling-circle replication: highlights of two decades of research. *Plasmid*, 53(2), pp.126–136.

Kleckner, N. E., Chatzi, K., White, M. A., Fisher, J. K. and Stouf, M. (2018). Coordination of Growth, Chromosome Replication/Segregation, and Cell Division in E. coli. *Frontiers in Microbiology*, 9, p.1469.

Knight, K. L. and Sauer, R. T. (1989). DNA binding specificity of the Arc and Mnt repressors is determined by a short region of N-terminal residues. *Proceedings of the National Academy of Sciences*, 86 (3), pp.797–801.

Kruse, T., Møller-Jensen, J., Løbner-Olesen, A. and Gerdes, K. (2003). Dysfunctional MreB inhibits chromosome segregation in Escherichia coli. *The EMBO Journal*, 22 (19), pp.5283–5292.

De La Hoz, A. B., Ayora, S., Sitkiewicz, I., Ferná Ndez, S., Pankiewicz, R., Alonso, J. C. and Ceglow*ski, P. (1999).* *Plasmid copy-number control and better-than-random segregation genes of pSM19035 share a common regulator. Proceedings of the National Academy of Sciences, 97(2), 728-33.*

De La Hoz, A. B., Pratto, F., Misselwitz, R., Speck, C., Weihofen, W., Wel¯e, K., Saenger, W., Wel¯e, H. and Alonso, J. C. (2004). Recognition of DNA by w protein from the broad-host range Streptococcus pyogenes plasmid pSM19035: analysis of binding to operator DNA with one to four heptad repeats. *Nucleic Acids Research*, 32 (10), pp.3136–3147.

Lane, D., Rothenbuehler, R., Merrillat, A. M. and Aiken, C. (1987). Analysis of the F plasmid centromere. *Molecular & General Genetics*, 207 (2–3), pp.406–412.

Larsen, R. A., Cusumano, C., Fujioka, A., Lim-Fong, G., Patterson, P. and Pogliano, J. (2007). Treadmilling of a prokaryotic tubulin-like protein, TubZ, required for plasmid stability in Bacillus thuringiensis. *Genes & Development*, 21 (11), pp.1340–1352.

Lederberg, J. (1952). Cell genetics and hereditary symbiosis. *Physiological Reviews*, 32, pp. 403–430.

Lees, J. G., Miles, A. J., Wien, F. and Wallace, B. A. (2006). A reference database for circular dichroism spectroscopy covering fold and secondary structure space. *Bioinformatics*, 22 (16), pp.1955–1962.

Leonard, T. A., Butler, J. and Lö, J. (2005). Bacterial chromosome segregation: structure and DNA binding of the Soj dimer-a conserved biological switch. *The EMBO Journal*, 24, pp.270–

282.

Li, Y. and Austin, S. (2002). The P1 plasmid is segregated to daughter cells by a 'capture and ejection' mechanism coordinated with Escherichia coli cell division. *Molecular Microbiology*, 46 (1), 63-74.

Libante, V., Thion, L. and Lane, D. (2001). Role of the ATP-binding site of SopA protein in partition of the F plasmid. *Journal of Molecular Biology*, 314 (3), pp.387–399.

Lim, G. E., Derman, A. I. and Pogliano, J. (2005). Bacterial DNA segregation by dynamic SopA polymers. *Proceedings of the National Academy of Sciences*, 102 (49), pp.17658–17663.

Lim, H. C., Surovtsev, I. V., Beltran, B. G., Huang, F., Bewersdorf, J. and Jacobs-Wagner, C. (2014). Evidence for a DNA-relay mechanism in ParABS-mediated chromosome segregation. *eLife*, 2014 (3), p.2758.

Ludtke, D. N., Eichorn, B. G. and Austin, S. J. (1989). Plasmid-partition functions of the P7 prophage. *Journal of Molecular Biology*, 209 (3), pp.393–406.

Lukaszewicz, M., Kostelidou, K., Bartosik, A. A., Cooke, G. D., Thomas, C. M. and Jagura-Burdzy, G. (2002). Functional dissection of the ParB homologue (KorB) from IncP-1 plasmid RK2. *Nucleic Acids Research*, 30 (4), pp.1046–1055.

Lutkenhaus, J. (2012). The ParA/MinD family puts things in their place. *Trends in Microbiology*, 20, pp.411–418.

Lynch, A. S. and Wang, J. C. (1994). Use of an inducible site-specific recombinase to probe the structure of protein-DNA complexes involved in F plasmid partition in Escherichia coli. *Journal of Molecular Biology*, 236 (3), pp.679–684.

Lynch, A. S. and Wang, J. C. (1995). SopB protein-mediated silencing of genes linked to the sopC locus of Escherichia coli F plasmid. *Proceedings of the National Academy of Sciences*, 92 (6), pp.1896–1900.

Ma, X., Ehrhardt, D. W. and Margolin, W. (1996). Colocalization of cell division proteins FtsZ and FtsA to cytoskeletal structures in living *Escherichia coli* cells by using green fluorescent protein. *Proceedings of the National Academy of Sciences*, 93 (23), pp.12998–13003.

Machón, C., Fothergill, T. J. G., Barillà, D. and Hayes, F. (2007). Promiscuous stimulation of ParF protein polymerization by heterogeneous centromere binding factors. *Journal of Molecular Biology*, 374 (1), pp.1–8.

Marston, A. L. and Errington, J. (1999). Dynamic Movement of the ParA-like Soj Protein of B. subtilis and Its Dual Role in Nucleoid Organization and Developmental Regulation. *Molecular Cell*, 4 (5), pp.673–682.

Martin, K. A., Friedman, S. A. and Austin, S. J. (1987). Partition site of the P1 plasmid. *Proceedings of the National Academy of Sciences,* 84 (23), pp.8544–8547.

Matern, Y., Barion, B. and Behrens-Kneip, S. (2010). PpiD is a player in the network of periplasmic chaperones in Escherichia coli. *BMC Microbiology*, 10 (1), p.251.

McLeod, B. N., Allison-Gamble, G. E., Barge, M. T., Tonthat, N. K., Schumacher, M. A., Hayes, F. and Barillà, D. (2017). A three-dimensional ParF meshwork assembles through the nucleoid to mediate plasmid segregation. *Nucleic Acids Research*, 45 (6), pp.3158–3171.

Million-Weaver, S. and Camps, M. (2014). Mechanisms of plasmid segregation: Have

multicopy plasmids been overlooked? *Plasmid*, 75, pp.27–36.

Minnen, A., Attaiech, L., Thon, M., Gruber, S. and Veening, J.-W. (2011). SMC is recruited to oriC by ParB and promotes chromosome segregation in Streptococcus pneumoniae. *Molecular Microbiology*, 81 (3), pp.676–688.

Mohl, D. A. and Gober, J. W. (1997). Cell Cycle–Dependent Polar Localization of Chromosome Partitioning Proteins in Caulobacter crescentus. *Cell*, 88 (5), pp.675–684.

de Mol, N. J. and Fischer, M. J. E. (2010). Surface Plasmon Resonance: Methods and Protocols. *Life Sciences*, p.255.

Møller-Jensen, J., Borch, J., Dam, M., Jensen, R. B., Roepstorff, P. and Gerdes, K. (2003). Bacterial Mitosis: ParM of Plasmid R1 Moves Plasmid DNA by an Actin-like Insertional Polymerization Mechanism. *Molecular Cell*, 12 (6), pp.1477–1487.

Møller-Jensen, J., Jensen, R. B., Löwe, J. and Gerdes, K. (2002). Prokaryotic DNA segregation by an actin-like filament. *The EMBO Journal*, 21 (12), pp.3119–3127.

Mori, H., Mori, Y., Ichinose, C., Niki, H., Ogura, T., Kato, A. and Hiraga, S. (1989). Purification and characterization of SopA and SopB proteins essential for F plasmid partitioning. *Journal of Biological Chemistry*, 264 (26), pp.15535–15541.

Murayama, K., Orth, P., de la Hoz, A. B., Alonso, J. C. and Saenger, W. (2001). Crystal structure of ω transcriptional repressor encoded by Streptococcus pyogenes plasmid pSM19035 at 1.5 Å resolution. *Journal of Molecular Biology*, 314 (4), pp.789–796.

Murray, H. and Errington, J. (2008). Dynamic Control of the DNA Replication Initiation Protein DnaA by Soj/ParA. *Cell*, 135 (1), pp.74–84.

Neeli-Venkata, R., Martikainen, A., Gupta, A., Gonçalves, N., Fonseca, J. and Ribeiro, A. S. (2016). Robustness of the Process of Nucleoid Exclusion of Protein Aggregates in Escherichia coli. *Journal of Bacteriology*, 198 (6), pp.898–906.

Nesvizhskii A. I., Keller A., Kolker E., Aebersold R. (2003). A statistical model for identifying proteins by tandem mass spectrometry. Analitical Chemistry, 75(17), pp. 4646-58.

Ni, L., Xu, W., Kumaraswami, M. and Schumacher, M. A. (2010). Plasmid protein TubR uses a distinct mode of HTH-DNA binding and recruits the prokaryotic tubulin homolog TubZ to effect DNA partition. *Proceedings of the National Academy of Sciences*, 107 (26), pp.11763–11768.

Niki, H. and Hiraga, S. (1997). Subcellular Distribution of Actively Partitioning F Plasmid during the Cell Division Cycle in E. coli. *Cell*, 90 (5), pp.951–957.

Niki, H., Jaffé, A., Imamura, R., Ogura, T. and Hiraga, S. (1991). The new gene mukB codes for a 177 kd protein with coiled-coil domains involved in chromosome partitioning of E. coli. *The EMBO Journal*, 10 (1), pp.183–193.

Nolivos, S., Upton, A. L., Badrinarayanan, A., Müller, J., Zawadzka, K., Wiktor, J., Gill, A., Arciszewska, L., Nicolas, E. and Sherratt, D. (2016). MatP regulates the coordinated action of topoisomerase IV and MukBEF in chromosome segregation. *Nature Communications*, 7, p.10466

Novick, R. P. (1987). Plasmid Incompatibility. *Microbiological Reviews*, 51 (4), pp.381–395.

Ogura, T. and Hiraga, S. (1983). Partition Mechanism of F Plasmid: Two Plasmid Gene-Encoded Products and a Cis-Acting Region Are Involved in Partition. *Cell*, 32 (0), pp.351–360.

Otting, G., Qian, Y. Q., Billeter, M., Müller, M., Affolter, M., Gehring, W. J. and Wüthrich, K. (1990). Protein--DNA contacts in the structure of a homeodomain--DNA complex determined by nuclear magnetic resonance spectroscopy in solution. *The EMBO journal*, 9 (10), pp.3085–3092.

Page, R. and Peti, W. (2016). Toxin-antitoxin systems in bacterial growth arrest and persistence. *Nature Chemical Biology*, 12 (4), pp.208–214.

Patching, S. G. (2014). Surface plasmon resonance spectroscopy for characterisation of membrane protein–ligand interactions and its potential for drug discovery. *Biochimica et Biophysica Acta - Biomembranes*, 1838 (1), pp.43–55.

Pattnaik, P. (2005). Surface Plasmon Resonance: Applications in Understanding Receptor–Ligand Interaction. *Applied Biochemistry and Biotechnology*, 126 (2), pp.079–092.

Pilla, G. and Tang, C. M. (2018). Going around in circles: virulence plasmids in enteric pathogens. *Nature Reviews Microbiology*, 16(8), pp.484-495.

Pogliano, J., Ho, T. Q., Zhong, Z. and Helinski, D. R. (2001). Multicopy plasmids are clustered and localized in Escherichia coli. *Proceedings of the National Academy of Sciences,* 98 (8), pp.4486–4491.

Pratto, F., Cicek, A., Weihofen, W. A., Lurz, R., Saenger, W. and Alonso, J. C. (2008). Streptococcus pyogenes pSM19035 requires dynamic assembly of ATP-bound ParA and ParB on parS DNA during plasmid segregation. *Nucleic Acids Research*, 36 (11), pp.3676–3689.

Ptacin, J. L., Lee, S. F., Garner, E. C., Toro, E., Eckart, M., Comolli, L. R., Moerner, W. E. and Shapiro, L. (2010). A spindle-like apparatus guides bacterial chromosome segregation. *Nature Cell Biology*, 12 (8), pp.791–798.

Ptacin, J. L. and Shapiro, L. (2010). Cell Cycle Initiating bacterial mitosis: Understanding the mechanism of ParA-mediated chromosome segregation. *Cell Cycle*, 9(20), pp.4033-4

Puig, O., Caspary, F., Rigaut, G., Rutz, B., Bouveret, E., Bragado-Nilsson, E., Wilm, M. and Séraphin, B. (2001). The tandem affinity purification (TAP) method: a general procedure of protein complex purification. *Methods*, 24 (3), pp.218–229.

Rajasekar, K., Muntaha, S. T., Tame, J. R. H., Kommareddy, S., Morris, G., Wharton, C. W., Thomas, C. M., White, S. A., Hyde, E. I. and Scott, D. J. (2010). Order and disorder in the domain organization of the plasmid partition protein KorB. *The Journal of Biological Chemistry*, 285 (20), pp.15440–15449.

Ravin, N. V., Rech, J. and Lane, D. (2003). Mapping of functional domains in F plasmid partition proteins reveals a bipartite SopB-recognition domain in SopA. *Journal of Molecular Biology*, 329 (5), pp.875–889.

Reyes-Lamothe, R., Tran, T., Meas, D., Lee, L., Li, A. M., Sherratt, D. J. and Tolmasky, M. E. (2014). High-copy bacterial plasmids diffuse in the nucleoid-free space, replicate stochastically and are randomly partitioned at cell division. *Nucleic Acids Research*, 42 (2), pp.1042–1051.

Rigaut, G., Shevchenko, A., Rutz, B., Wilm, M., Mann, M. and Séraphin, B. (1999). A generic protein purification method for protein complex characterization and proteome exploration. *Nature biotechnology*, 17 (10), pp.1030–1032.

Ringgaard, S., Ebersbach, G., Borch, J. and Gerdes, K. (2007). Regulatory cross-talk in the double par locus of plasmid pB171. *Journal of Biological Chemistry*, 282 (5), pp.3134–3145.

Ringgaard, S., van Zon, J., Howard, M. and Gerdes, K. (2009). Movement and equipositioning of plasmids by ParA filament disassembly. *Proceedings of the National Academy of Sciences*, 106 (46), pp.19369–19374.

Rohila, J. S., Chen, M., Cerny, R. and Fromm, M. E. (2004). Improved tandem affinity purification tag and methods for isolation of protein heterocomplexes from plants. *The Plant Journal*, 38 (1), pp.172–181.

Rosche, T. M., Siddique, A., Larsen, M. H. and Figurski, D. H. (2000). Incompatibility Protein IncC and Global Regulator KorB Interact in Active Partition of Promiscuous Plasmid RK2. *Journal of Bacteriology*, 182 (21), pp.6014-26.

Roy, R., Hohng, S. and Ha, T. (2008). A practical guide to single-molecule FRET. *Nature Methods*, 5 (6), pp.507–516.

Ruiz-Masó J. A., Machón C., Bordanaba-Ruiseco L., Espinosa M., Coll M. and del Solar G. (2015). Chapter 4: Plasmid Rolling-circle replication. In Tolmasky & Alonso (Eds.), *Plasmids: Biology and Impact in Biotechnology and Discovery*. American Society of Microbiology, pp. 45-69

Saeed, S., Jowitt, T. A., Warwicker, J. and Hayes, F. (2015). Breaking and Restoring the Hydrophobic Core of a Centromere Binding Protein. *Journal of Biological Chemistry*, 290 (14), pp.9273-83.

Salje, J. (2010). Plasmid segregation: how to survive as an extra piece of DNA. *Critical Reviews in Biochemistry and Molecular Biology*, 45 (4), pp.296–317.

Salje, J., Gayathri, P. and Löwe, J. (2010). The ParMRC system: molecular mechanisms of plasmid segregation by actin-like filaments. *Nature reviews. Microbiology*, 8 (10), pp.683–692.

Salje, J., Zuber, B. and Lowe, J. (2009). Electron Cryomicroscopy of E. coli Reveals Filament Bundles Involved in Plasmid DNA Segregation. *Science*, 323 (5913), pp.505–509.

Scholefield, G., Whiting, R., Errington, J. and Murray, H. (2011). Spo0J regulates the oligomeric state of Soj to trigger its switch from an activator to an inhibitor of DNA replication initiation. *Molecular Microbiology*, 79 (4), pp.1089–1100.

Schreiter, E. R. and Drennan, C. L. (2007). Ribbon–helix–helix transcription factors: variations on a theme. *Nature Reviews Microbiology*, 5 (9), pp.710–720.

Schumacher, M. A. (2007). Structural biology of plasmid segregation proteins. *Current Opinion in Structural Biology*, 17 (1), pp.103–109.

Schumacher, M. A. (2008). Structural biology of plasmid partition: uncovering the molecular mechanisms of DNA segregation. *The Biochemical Journal*, 412 (1), pp.1–18.

Schumacher, M. A. (2012). Bacterial plasmid partition machinery: a minimalist approach to survival. *Current Opinion in Structural Biology*, 22 (1), pp.72–79.

Schumacher, M. A. and Funnell, B. E. (2005). Structures of ParB bound to DNA reveal mechanism of partition complex formation. *Nature*, 438(7067), pp.516-9.

Schumacher, M. A., Glover, T. C., Brzoska, A. J., Jensen, S. O., Dunham, T. D., Skurray, R. A. and Firth, N. (2007a). Segrosome structure revealed by a complex of ParR with

centromere DNA. *Nature*, 450 (7173), pp.1268–1271.

Schumacher, M. A., Mansoor, A. and Funnell, B. E. (2007b). Structure of a four-way bridged ParB-DNA complex provides insight into P1 segrosome assembly. *The Journal of Biological Chemistry*, 282 (14), pp.10456–10464.

Schumacher, M. A., Piro, K. M. and Xu, W. (2010). Insight into F plasmid DNA segregation revealed by structures of SopB and SopB–DNA complexes. *Nucleic Acids Research*, 38 (13), pp.4514–4526.

Schumacher, M. A., Tonthat, N. K., Lee, J., Rodriguez-Castañeda, F. A., Chinnam, N. B., Kalliomaa-Sanford, A. K., Ng, I. W., Barge, M. T., Shaw, P. L. R. and Barillà, D. (2015). Structures of archaeal DNA segregation machinery reveal bacterial and eukaryotic linkages. *Science*, 349 (6252), pp.1120–1124.

Schumacher, M. A., Ye, Q., Barge, M. T., Zampini, M., Barillà, D. and Hayes, F. (2012). Structural mechanism of ATP-induced polymerization of the partition factor ParF: implications for DNA segregation. *The Journal of Biological Chemistry*, 287 (31), pp.26146–26154.

Sheppard, C., Blombach, F., Belsom, A., Schulz, S., Daviter, T., Smollett, K., Mahieu, E., Erdmann, S., Tinnefeld, P., Garrett, R., Grohmann D., Rappsilber J., Werner F. (2016). Repression of RNA polymerase by the archaeo-viral regulator ORF145/RIP. *Nature Communications*, 7, p.13595.

Silva, F., Queiroz, J. A. and Domingues, F. C. (2012). Evaluating metabolic stress and plasmid stability in plasmid DNA production by Escherichia coli. *Biotechnology Advances*, 30 (3), pp.691–708.

Simpson, A. E., Skurray, R. A. and Firth, N. (2003). A single gene on the staphylococcal multiresistance plasmid pSK1 encodes a novel partitioning system. *Journal of Bacteriology*, 185 (7), pp.2143–2152.

del Solar, G. and Espinosa, M. (2002). Plasmid copy number control: an ever-growing story. *Molecular Microbiology*, 37 (3), pp.492–500.

del Solar, G., Giraldo, R., Ruiz-Echevarría, M. J., Espinosa, M. and Díaz-Orejas, R. (1998). Replication and control of circular bacterial plasmids. *Microbiology and Molecular Biology Reviews,* 62 (2), pp.434–464.

Somers, W. S. and Phillips, S. E. V. (1992). Crystal structure of the met represser–operator complex at 2.8 Å resolution reveals DNA recognition by β-strands. *Nature*, 359 (6394), pp.387–393.

Song, D., Rodrigues, K., Graham, T. G. W. and Loparo, J. J. (2017). A network of cis and trans interactions is required for ParB spreading. *Nucleic Acids Research*, 45(12), pp.7106–7117.

Spolar, R. and Record MT Jr. (1994). Coupling of local folding to site-specific binding of proteins to DNA. *Science*, 263 (5148), pp.777–784.

Sreerama, N. and Woody, R. W. (2000). Estimation of Protein Secondary Structure from Circular Dichroism Spectra: Comparison of CONTIN, SELCON, and CDSSTR Methods with an Expanded Reference Set. *Analytical Biochemistry*, 287 (2), pp.252–260.

Stenberg, F., Chovanec, P., Maslen, S. L., Robinson, C. V, Ilag, L. L., von Heijne, G. and

Daley, D. O. (2005). Protein complexes of the Escherichia coli cell envelope. *The Journal of Biological Chemistry*, 280 (41), pp.34409–34419.

Stock, A. M. and Zhulin, I. B. (2017). Two-Component Signal Transduction: a Special Issue in the Journal of Bacteriology. *Journal of Bacteriology,* 199 (18), pp.443–17.

Stricker, J., Maddox, P., Salmon, E. D. and Erickson, H. P. (2002). Rapid assembly dynamics of the Escherichia coli FtsZ-ring demonstrated by fluorescence recovery after photobleaching. *Proceedings of the National Academy of Sciences*, 99 (5), pp. 3171-5.

Stymest, K. H. and Klappa, P. (2008). The periplasmic peptidyl prolyl cis-trans isomerases PpiD and SurA have partially overlapping substrate specificities. *FEBS Journal*, 275 (13), pp.3470–3479.

Sullivan, N. L., Marquis, K. A. and Rudner, D. Z. (2009). Recruitment of SMC by ParB-parS Organizes the Origin Region and Promotes Efficient Chromosome Segregation. *Cell*, 137 (4), pp.697–707.

Summers, D. K. (1991). The kinetics of plasmid loss. *Trends in Biotechnology*, 9 (1), pp.273–278.

Summers, D. K. (1996). *The Biology of Plasmids*. Oxford, UK : Blackwell Publishing Ltd.

Summers, D. K. and Sherratt, D. J. (1984). Multimerization of high copy number plasmids causes instability: CoIE1 encodes a determinant essential for plasmid monomerization and stability. *Cell*, 36 (4), pp.1097–1103.

Sun, Y., Wallrabe, H., Booker, C. F., Day, R. N. and Periasamy, A. (2010). Three-Color Spectral FRET Microscopy Localizes Three Interacting Proteins in Living Cells. *Biophysical Journal*, 99 (4), pp.1274–1283.

Surtees, J. A. and Funnell, B. E. (1999). P1 ParB domain structure includes two independent multimerization domains. *Journal of bacteriology*, 181 (19), pp.5898–5908.

Szardenings, F., Guymer, D. and Gerdes, K. (2011). ParA ATPases can move and position DNA and subcellular structures. *Current Opinion in Microbiology*, 14(6), pp.712-8.

Taylor, J. A., Pastrana, C. L., Butterer, A., Pernstich, C., Gwynn, E. J., Sobott, F., Moreno-Herrero, F. and Dillingham, M. S. (2015). Specific and non-specific interactions of ParB with DNA: implications for chromosome segregation. *Nucleic acids research*, 43 (2), pp.719–731.

Thanbichler, M. and Shapiro, L. (2006). MipZ, a Spatial Regulator Coordinating Chromosome Segregation with Cell Division in Caulobacter. *Cell*, 126 (1), pp.147–162.

Thomas, C. M. (2000). *Horizontal Gene Pool : Bacterial Plasmids and Gene Spread.* Harwood Academic Publisher

Toro, E., Hong, S.-H., Mcadams, H. H. and Shapiro, L. (2008). Caulobacter requires a dedicated mechanism to initiate chromosome segregation. *Proceedings of the National Academy of Sciences*, 105(40), pp.15435-40.

Tran, N. T., Stevenson, C. E., Som, N. F., Thanapipatsiri, A., Jalal, A. S. B. and Le, T. B. K. (2018). Permissive zones for the centromere-binding protein ParB on the Caulobacter crescentus chromosome. *Nucleic Acids Research*, 46 (3), pp.1196–1209.

Truong, K. and Ikura, M. (2001). The use of FRET imaging microscopy to detect protein-protein interactions and protein conformational changes in vivo. *Current Opinion in Structural*

Biology, 11 (5).

Vecchiarelli, A. G., Han, Y.-W., Tan, X., Mizuuchi, M., Ghirlando, R., Biertümpfel, C., Funnell, B. E. and Mizuuchi, K. (2010). ATP control of dynamic P1 ParA-DNA interactions: a key role for the nucleoid in plasmid partition. *Molecular Microbiology*, 78 (1), pp.78–91.

Vecchiarelli, A. G., Havey, J. C., Ing, L. L., Wong, E. O. Y., Waples, W. G. and Funnell, B. E. (2013a). Dissection of the ATPase Active Site of P1 ParA Reveals Multiple Active Forms Essential for Plasmid Partition. *Journal of Biological,* 288(24), pp. 17823-31.

Vecchiarelli, A. G., Hwang, L. C. and Mizuuchi, K. (2013b). Cell-free study of F plasmid partition provides evidence for cargo transport by a diffusion-ratchet mechanism. *Proceedings of the National Academy of Sciences,* 110 (15), pp.E1390-7.

Vecchiarelli, A. G., Mizuuchi, K. and Funnell, B. E. (2012). Surfing biological surfaces: exploiting the nucleoid for partition and transport in bacteria. *Molecular Microbiology*, 86 (3), pp.513–523.

Vecchiarelli, A. G., Neuman, K. C. and Mizuuchi, K. (2014a). A propagating ATPase gradient drives transport of surface-confined cellular cargo. *Proceedings of the National Academy of Sciences,* 111 (13), pp.4880–4885.

Vecchiarelli, A. G., Seol, Y., Neuman, K. C. and Mizuuchi, K. (2014b). A moving ParA gradient on the nucleoid directs subcellular cargo transport via a chemophoresis force. *Bioarchitecture*, 4 (4–5), pp.154–159.

Vogt, J. and Schulz, G. E. (1999). The structure of the outer membrane protein OmpX from Escherichia coli reveals possible mechanisms of virulence. *Structure*, 7 (10), pp.1301–1309.

Ward, J. J., Sodhi, J. S., McGuffin, L. J., Buxton, B. F. and Jones, D. T. (2004). Prediction and Functional Analysis of Native Disorder in Proteins from the Three Kingdoms of Life. *Journal of Molecular Biology*, 337 (3), pp.635–645.

Watanabe, E., Wachi, M., Yamasaki, M. and Nagai, K. (1992). ATPase activity of SopA, a protein essential for active partitioning of F plasmid. *Molecular & General Genetics*, 234 (3), pp.346–352.

Webb, C. D., Graumann, P. L., Kahana, J. A., Teleman, A. A., Silver, P. A. and Losick, R. (1998). Use of time-lapse microscopy to visualize rapid movement of the replication origin region of the chromosome during the cell cycle in Bacillus subtilis. *Molecular Microbiology*, 28 (5), pp.883–892.

Weihofen, W. A., Cicek, A., Pratto, F., Alonso, J. C. and Saenger, W. (2006). Structures of repressors bound to direct and inverted DNA repeats explain modulation of transcription. *Nucleic Acids Research*, 34 (5), pp.1450–1458.

Weitao, T., Dasgupta, S. and Nordstrom, K. (2000). Role of the mukB gene in chromosome and plasmid partition in Escherichia coli. *Molecular Microbiology*, 38 (2), pp.392–400.

Wienken, C. J., Baaske, P., Rothbauer, U., Braun, D. and Duhr, S. (2010). Protein-binding assays in biological liquids using microscale thermophoresis. *Nature communications*, 1, p.100.

Wu, L. J. and Errington, J. (2003). RacA and the Soj-Spo0J system combine to effect polar chromosome segregation in sporulating Bacillus subtilis. *Molecular Microbiology*, 49 (6),

pp.1463–1475.

Wu, M., Zampini, M., Bussiek, M., Hoischen, C., Diekmann, S. and Hayes, F. (2011). Segrosome assembly at the pliable parH centromere. *Nucleic Acids Research*, 39 (12), pp.5082–5097.

Yahashiri, A., Jorgenson, M. A. and Weiss, D. S. (2017). The SPOR domain, a widely conserved peptidoglycan binding domain that targets proteins to the site of cell division. *Journal of Bacteriology*, 199 (14), p.00118-17.

Yamaichi, Y., Bruckner, R., Ringgaard, S., Moll, A., Cameron, D. E., Briegel, A., Jensen, G. J., Davis, B. M. and Waldor, M. K. (2012). A multidomain hub anchors the chromosome segregation and chemotactic machinery to the bacterial pole. *Genes & Development*, 26 (20), pp.2348–2360.

Yamaichi, Y. and Niki, H. (2000). Active segregation by the Bacillus subtilis partitioning system in Escherichia coli. *Proceedings of the National Academy of Sciences*, 97 (26), pp.14656–14661.

Yao, S., Helinski, D. R. and Toukdarian, A. (2007). Localization of the naturally occurring plasmid ColE1 at the cell pole. *Journal of bacteriology*, 189 (5), pp.1946–1953.

Yuan, Z., Riera, A., Bai, L., Sun, J., Nandi, S., Spanos, C., Chen, Z. A., Barbon, M., Rappsilber, J., Stillman, B., Speck C. Li H. (2017). Structural basis of Mcm2–7 replicative helicase loading by ORC–Cdc6 and Cdt1. *Nature Structural & Molecular Biology*, 24 (3), pp.316–324.

Zampini, M., Derome, A., Bailey, S. E. S., Barillà, D. and Hayes, F. (2009). Recruitment of the ParG segregation protein to different affinity DNA sites. *Journal of Bacteriology*, 191 (12), pp.3832–3841.

Zawadzki, P., Stracy, M., Ginda, K., Zawadzka, K., Lesterlin, C., Kapanidis, A. N. and Sherratt, D. J. (2015). The Localization and Action of Topoisomerase IV in Escherichia coli Chromosome Segregation Is Coordinated by the SMC Complex, MukBEF. *Cell Reports*, 13 (11), pp.2587–2596.

Zhang, H. and Schumacher, M. A. (2017). Structures of partition protein ParA with nonspecific DNA and ParB effector reveal molecular insights into principles governing Walker-box DNA segregation. *Genes & Development*, 31 (5), pp.481–492.

Zhong, C., Peng, D., Ye, W., Chai, L., Qi, J., Yu, Z., Ruan, L. and Sun, M. (2011). Determination of plasmid copy number reveals the total plasmid DNA amount is greater than the chromosomal DNA amount in Bacillus thuringiensis YBT-1520. *PloS one*, 6 (1), p.e16025.