THE ROLE OF P1 PARA IN THE PARTITIONING OF THE P1 PLASMID

By

Thomas Dunham

A THESIS/DISSERTATION

Presented to the Department of Biochemistry and Molecular Biology
and the Oregon Health & Science University
School of Medicine
in partial fulfillment of
the requirements of the degree of

Doctor of Philosophy

April 2009

School of Medicine

Oregon Health & Science University

CERTIFICATE OF APPROVAL

This is to certify that the Ph.D. thesis of

Thomas Dunham

has been approved

_____

Mentor/Advisor

_____

Member

_____

Member

_____

Member

_____

Member

_____

Member

TABLE OF CONTENTS

LIST OF FIGURES

Chapter 1

Chapter 2

Chapter 3

Chapter 6

Chapter 7

ACKNOWLEDGEMENTS

I would first like to thank Dr. Maria Schumacher for being my mentor and providing the support necessary for conducting the research included in this thesis. I would like to thank Dr. Richard Brennan and the Schumacher and Brennan labs for providing a stimulating and friendly lab environment.

I would also like to thank Dr. Peter Rotwein for being my sponsor at OHSU and for the support of my off-site thesis research at UTMDACC.

I would like to acknowledge Dr. Schumacher for solving the ParA structures discussed in this thesis. I would also like to acknowledge our collaborator Dr. Barbara Funnell for conducting the DNAse I experiments supporting the mutational experiments on the basic region of ParA (Chapter 4) and for the work she has performed in characterizing the P1 *par* system. I would also like to acknowledge the UTMDACC HREM facility for assisting in the electron microscopy studies of the P1 ParA filaments (Chapter 4).

ABSTRACT

The partition system of the low copy number P1 plasmid ensures that plasmids are faithfully segregated between daughter cells using three components contained in a single *par* operon. The P1 *par* cassette includes the genes of two proteins (*parA* and *parB*) and a *cis* acting plasmid centromere (*parS*). P1 ParA is a deviant Walker A ATPase that plays an essential, but unknown role in plasmid partitioning as well as in the autorepression of its own operon. These two disparate functions are modulated by the nucleotide binding state of the ParA protein. The understanding of how nucleotide binding to ParA can affect its functional role has been limited by the absence of high-resolution structural information.

X-ray crystallography studies were undertaken to address the questions centering on the nature of ParA's different nucleotide bound states. A structure of P1 ParA in its apo state revealed that it was a dimer with an extended N-terminal α-helix. Supporting static light studies established that apo ParA is able to form a dimer in solution at physiologically relevant concentrations. The studies also revealed that the stability for the dimer state was found to be dependent upon an interaction between the N-terminal α-helix (α1) and its dimer partner. ADP binding potentiates ParA's ability to bind DNA and function in the auto-repression of its own operon. An ADP-bound structure of P1 ParA was solved and provided the basis for a ParA-DNA-bound model that required the DNA to be significantly bent in order to align with the recognition helix of ParA's HTH motif. A nucleotide dependent basic region revealed in an ADP-bound structure of ParA provided a mechanism to stabilize the bending of the DNA. Mutagenesis and supporting DNA binding analysis established that the basic region is important for the binding of DNA.

When ParA is in its ATP-bound state it carries out a different function providing an essential role in partitioning. Negative stain electron microscopy (EM) experiments were used to demonstrate that ParA formed filaments in an ATP dependent manner. Reconstitution experiments further demonstrated that the components of the P1 *par* cassette were localized to the filaments suggesting that the ParA filaments are important

to the segregation of the P1 plasmid. ParB has been known to modulate ParA's ATPase activity in an unknown manner through an interaction involving its N-terminus. ParA-ParB fusion constructs were used to further localize both the specificity and activity features responsible for the potent interaction involving ParB's N-terminal 28 residues. Further investigations into the source of the activity found that two arginine residues (R6 and R11) were critical for this activity possibly acting through an "arginine finger" motif mechanism.

The studies on P1 ParA show that unlike other deviant Walker A ATPases that ParA likely exists as a dimer in all its states. The activity of ParA is governed by its nucleotide bound state, which locks ParA into conformations that are preferred for its disparate cellular functions – autorepression and partitioning.

Chapter 1

Definition and Description of the Cellular Process of Partitioning

1.1 Introduction.

Partition is the deliberate movement of newly replicated chromosomes or plasmids to specific regions in daughter cells prior to cell division. It is the process that guarantees the accurate total distribution of genetic material between daughter cell offspring and as a result is a fundamental process necessary for the viability of organisms across all kingdoms. The mechanisms and machinery employed to carry out this process are diverse and vary greatly in their complexity. While the general framework of partitioning in Eukaryotes has been established with visual distinction - microtubules pulling newly replicated chromosomes back toward the cell poles - understanding partitioning in Prokaryotes is at an earlier stage with much of our current understanding arising from genetic and biochemical studies that are now being augmented by more biophysical methods of examination. Much of the progress in our understanding has been made studying the simple systems utilized by low-copy number plasmids to ensure their accurate distribution into host daughter cells. The components of this low copy plasmid system carry the designation of *par*, for partition, to specify their link to this biological process in the cell. As an introduction to the work carried out in the body of this thesis on the ParA protein from the *par* system of the P1 plasmid, I will briefly discuss the components of the low copy systems, the classification of the different plasmid partition systems, their common mechanisms and the current challenges to understanding this useful model.

Low-copy number (4-6 copies) plasmid systems employ active mechanisms to ensure that at least one newly replicated plasmid is accurately distributed to each daughter cell as opposed to high-copy (>10 copies) plasmid systems, which rely on passive diffusion to statistically ensure the same end. The composition of the low-copy plasmid *par* systems is an established paradigm characterized by three essential elements contained in a single cassette: a trans-acting DNA binding protein, a trans-acting motor protein and a cis-acting plasmid centromere-like site (Austin and Abeles, 1983; Funnell, 2005; Hayes and Barilla, 2006; Schumacher, 2008). These components are often all that is required to ensure the accurate division of genetic material within the bacterial cell. Host components, which may confer added efficiency to partitioning *in vivo* (localization and synchronization of components), were found to be fundamentally unnecessary *in vitro* (Garner et al., 2007). The protein components are usually encoded on single operon within the plasmid itself with the motor protein immediately upstream of the DNA binding protein and the centromere-like site either up or downstream of the other two components (Figure 1_1).

The general process of partitioning commences when the DNA binding protein component (later distinguished as ParB or ParR) recognizes and binds to its own specifically matched plasmid centromere-like site (termed *parS* or *parC*, respectively) within the newly replicated plasmid (Figure 1_1 (I.)). This nucleoprotein complex is termed the "partition complex" and becomes the first important superstructure of the partitioning process. At this point some protein dependent or DNA dependent pairing or clustering of the plasmids has been observed to take place (Funnell, 2005). The next

important step occurs as a result of a conformational change in the motor protein (ParA or ParM) resulting from a nucleotide-binding event (ATP or GTP) (Figure 1_1 (II.)). The motor protein in its activated ATP or GTP-bound form is able to recognize the "partition complex" and drive the segregation of the newly replicated plasmids to their respective daughter cells (Figure 1_1 (III.)). The daughter plasmids are finally localized to the cell poles or at positions ¼ and ¾ of the cell (Figure 1_1 (IV.)).

The expression of the two protein components (ParA and ParB) in the *par* systems appears to be crucial and strictly controlled balancing the requirements for the partitioning process with the deleterious consequences of over-or-missed timed expression (Abeles et al., 1985; Funnell, 1988a). In response, the plasmid partition systems exert control over their own expression levels through the auto-repression of the *par* operon (Figure 1_1 (IV.)).

1.2 Classification of the Different Partitioning Systems.

Plasmid partition systems have by convention been classified by the type of motor protein present in the cassette. These motor proteins fall under two main general types: a deviant Walker A type ATPase (type I) or an actin-like protein (type II) (Gerdes et al., 2000). Walker-type ATPases are classified on the basis of a sequence motif (A box or P-

loop) GxxGxGKS/T containing a signature lysine residue important for its ATPase function (Walker et al., 1982). Deviant Walker A ATPases have an A box motif KGGXXK[S/T] containing a second lysine residue that is important to the function of this sub-family (Koonin, 1993). Recently, two more potential classes have been identified with links to either a tubulin-like GTPase, TubZ, or a coiled-coil motif illustrating that the mechanisms at the center of these *par* systems are likely to be more diverse (Schumacher, 2008).

Type I *par* system components generally have the following designations: ParA, for the motor protein; ParB, for the DNA binding protein; and *parS*, for the centromere-like DNA site. Type I *par* systems are the largest class and contain two sub-types, Ia and Ib, distinguished by the organization of their operon, limited sequence homology and the relative sizes of the *par* proteins. The general scheme for the type I *par* operon follows the normal motor protein centromere binding protein progression, but the location of the *parS* element is downstream *parAB* in the Ia sub-type, while it is upstream in the Ib sub-type (Figure 1_2A). Type Ia ParA proteins (~ 30-40 kDa) are larger than the type Ib homolog possessing an extra ~ 110 residues in what forms an N-terminal extension (Figure 1_2B). Contained within this extended N-terminus is a putative helix-turn-helix (HTH) motif presumably enabling ParA to fulfill its role as an auto-repressor of type Ia *parAB* (Davey and Funnell, 1994; Radnedge et al., 1998). Both type Ia and Ib *par* motor proteins are Walker-type ATPases based upon an A box sequence motif GxxGxGKS/T. Both type Ia and Ib *par* motor proteins are weak ATPases and are stimulated most strongly by an interaction with their corresponding ParB protein. ATPase activity has

been demonstrated to be necessary for correct partitioning, but the nature of the mechanistic role is unknown (Davey and Funnell, 1997; Davis et al., 1996).

There is little sequence homology among the type Ia and Ib ParB proteins. Type Ia ParB proteins (~ 30-35 kDa) are also significantly larger (2-6 times) than their type Ib analogues. Type Ia ParB proteins share predicted HTH, B and C-terminal dimerization domains even though there is little sequence similarity between the Ia family members (Figure 1_2A)(Surtees and Funnell, 1999). Type Ib ParB proteins have no recognizable DNA binding motif based on sequence despite being both a centromere binding protein and an auto-repressor of the type Ib *parAB* operon.

Like the type I system, the components of the type II system carry their own general designations: ParM, for the motor protein; ParR, for the DNA binding protein; and *parC*, for the centromere-like site (Gerdes and Molin, 1986; Gerdes et al., 2000). The type II cassette is organized similarly to the type Ib cassette (Figure 1_2B). The motor protein component, ParM, is an actin-like protein. Type II centromere binding proteins, like the type Ib ParB proteins, are small and act to bind centromere-like DNA as well as functioning as auto-repressors of their own *parMR* operon (Jensen et al., 1994). The centromere-like site, *parC*, like the type Ib protein is an arrangement of repeated sequences (Dam and Gerdes, 1994)(Figure 1_2C).

1.3 Formation of the Partition Complex.

As noted previously, the first key step in partitioning is the formation of the "partition complex." The arrangement of the centromere sites are varied, but have the common feature of some form of repeated sequence covering 80-500 base pairs (bp) (Figure 1_2C). It appears that the configuration of the DNA binding centromere-like sites allows these proteins to bind cooperatively in order to promote the formation of a superstructure conducive to the plasmid segregation process (Erdmann et al., 1999). Protein loading is a common mechanism among plasmid DNA binding proteins and they often contain structural elements to promote oligomerization (Funnell and Gagnier, 1993). It is clear that this leads to spreading, but its specific role in partitioning is controversial (Rodionov et al., 1999).

One of the most complicated centromere arrangements is *parS* from the type Ia P1 *par* system. It is composed of two different binding site, a heptad (A Box) and a hextet (B Box), repeats arranged asymmetrically around a binding site for the *E. coli* derived Host Integration Factor (IHF) (Funnell, 1988b; Funnell, 1991; Funnell and Gagnier, 1994) (Figure 1_2C). IHF's role is to bend the DNA by ~180^{o} allowing the two arms of the centromere site to come closer in proximity to each other. This site has been replaced by a DNA sequence with an inherently bent structure to demonstrate that IHF's only contribution is its role in bending DNA. The proposed mechanism of partition complex formation involves multiple ParB proteins binding cooperatively to plasmid centromere

resulting in a higher order nucleoprotein complex (Bouet et al., 2000; Funnell and Gagnier, 1993; Schumacher and Funnell, 2005; Surtees and Funnell, 2001). Partition complexes on replicated plasmids are next thought to pair via ParB-ParB and ParB-DNA interactions (Funnell, 2005). Structural studies on P1 ParB showed that it utilized two different DNA binding domains to bind *parS* (Schumacher and Funnell, 2005). The C-terminal dimerization domain recognizes the DNA hextet B-box motifs, while the HTH motif binds to the heptad A-box motifs (Figure 1_3A&B)(Schumacher, 2007; Schumacher and Funnell, 2005). A model of the P1 partition complex integrating IHF and the *parS* site was made showing the circular nature of the nucleoprotein superstructure (Figure 1_3C). The radius of gyration (R_g) of the model closely corresponded to experimental R_g generated from supporting SAXS experiments (Dr. Schumacher - personal communication).

A lingering question has been how DNA binding might occur in the other *par* systems. It is only through recent structural studies on type Ib and type II DNA binding proteins that the mysteries of their DNA binding capabilities have been elucidated. Structures of type Ib TP228 ParG in the apo state and type Ib pSM19035 ω protein in the apo and DNA-bound states show that type Ib ParB proteins are dimers and form a dimer of dimers upon binding DNA (Golovanov et al., 2003; Murayama et al., 2001; Weihofen et al., 2006). Type Ib proteins fold into a ribbon-helix-helix (RHH) motif forming a dimer primarily through an anti-parallel β-strand and α-helical core. The DNA bound state of the ω protein illustrates that it forms a dimer of dimers and that the protein contacts two successive major grooves through the anti-parallel β-strand formed by each dimer. As

the ω protein is also a general transcriptional regulator, the DNA site used for the study was distinct from its *parH* centromere site. Crystal structures of type II pB171 ParR in the apo state and pSK41 (1-53) in the DNA-bound state demonstrate that the type II DNA binding proteins also contain RHH motifs and utilize a dimer of dimer configuration to bind DNA (Figure 1_4A&B)(Moller-Jensen et al., 2007; Schumacher et al., 2007). The underlying mechanism of DNA binding is through a common RHH motif that appears to be common to both type Ib and type II centromere binding proteins despite the lack of homology in their primary sequences.

Since the 20 bp tandem repeat DNA site used in the type II pSK41 ParR-DNA structure was a portion of the *parC* site, it allowed for each ParR dimer of dimers to stack and form the pseudo continuous centromere structure (Figure 1_4Aa). The resulting super-helical structure, formed from the wrapping of DNA around the mostly electropositive face of the ParR protein complex, had a pitch of ~ 24 nm and a diameter of ~ 18 nm (Figure 1_4Ab). Biochemical assays demonstrated that ParR binds cooperatively to the full-length *parC* site and that spreading to adjacent DNA likely extends the structure. In the structure the missing C-terminus would rest in the pore. It was demonstrated that ParM bound only to the full-length ParR-*parC* complex. Interestingly, it appears that the type II ParR proteins prefers this super-helical structure as the apo structure of the full-length type II pB171 formed a similar structure with a pitch of ~ 13 nm and a diameter of ~ 15 nm. The pB171 structure also illustrated how the C-terminus might fill the pore and act to support oligomerization through inter-dimer contacts (Figure 1_4B&C). Finally, a model of the full segrosome (motor protein + DNA binding protein + centromere site)

showed the feasibility of the ParM-ParR*C* segrosome complex mediated through an interaction between the C-terminus of ParR and ParM filaments (Figure 1_4D). It appears that a circular "partition complex" superstructure is key to recruiting and coordinating the motor protein's activity in partition, but it has yet to be demonstrated how this might occur in the type I *par* system.

1.4 Nucleotide Binding by *par* Motor Proteins.

One of the most important and emerging questions revolving around plasmid partition systems concerns the mechanism utilized by the motor proteins to drive plasmid partitioning. Common to all the motor proteins is the question of a "switch mechanism" that coincides with the transition from a resting state to an active state protein. Within the partition systems, this switch is believed to be nucleotide dependent with the ATP-bound state being distinct from the ADP-bound state, which is again distinct from the apo state. In all the *par* motor proteins the ATP-bound (or GTP-bound) form is the active state capable of participating in the segregation reaction. What is the role of nucleotide binding?

The answer is unclear. The two best-studied examples come from type Ib and type II motor proteins. Crystal structures of the deviant Walker A type Ib-like motor protein Soj

from *B. Subtilis*, one of the few chromosomal partition systems dissected to date, in the apo, ADP-bound and ATP-bound states show a clear and dramatic effect of nucleotide binding on the functional state of the protein (Leonard et al., 2005). In this case, the Soj protein is a monomer until it binds ATP, which converts Soj to its dimeric state (Figure 1_5A). Interestingly, Soj filaments were observed only in the presence of non-specific DNA resulting in a nucleo-protein filament. Another type Ib motor protein whose structure has been solved is the δ protein from the plasmid pSM19035 in *Streptococcus pyogenes*. The ATPγS-δ_2 protein complex was also found to be in a nucleotide-sandwich fold similar to the Soj protein, but supporting cross-linking experiments found that δ_2 is a dimer in its apo form (Figure 1_5B)(Pratto et al., 2008). This opens the possibility that there may be more variability in type I *par* motor proteins and leads to the question of what constitutes the "molecular switch" if it does not involve the act of dimerization.

Crystal structures of the type II motor protein ParM in the apo and ADP-bound state confirmed that the nucleotide binds to a cleft in between two subunits in an actin-fold changing its conformation to a more closed state (van den Ent et al., 2002). Recently, ParM was crystallized in its GDP and GMP_PNP-bound forms (Popp et al., 2008). Studies suggest that ParM may actually function by binding to guanine nucleotides and ParM can hydrolyze GTP at a faster rate than ATP (Popp et al., 2007). Analysis of the different structures revealed that the binding mechanism for GDP and GMP-PNP were similar to that of ADP and an overlay of all the structures identified only minor differences (Figure 1_5C). It is clear from studies that type II ParM is able to form

filaments readily in the presence of ATP or GTP and ParM has no DNA binding capabilities on its own.

1.5 Mechanisms Used by Motor Proteins During Partitioning.

The mechanisms employed by the motor proteins during partitioning are only now being elucidated. As mentioned previously, in terms of partitioning, it is the ATP/GTP-bound form of the motor protein that is the only active state capable of recognizing the "partition complex" and triggering the segregation of daughter plasmids. So far, there has been little consensus on whether these motor proteins work directly (alone) or indirectly (in concert with other factors) in actually trans-locating the plasmids across the cell. Recently, there have been more reports that the *par* motor proteins are capable of forming polymers in a nucleotide (ATP/GTP or analogs) dependent manner. In fact, filaments have now been observed in microscopy experiments from both Walker A type Ia (SopA) and Ib (δ, ParF, and pB171 ParA) *par* motor proteins (Barilla et al., 2005; Ebersbach et al., 2006; Ebersbach et al., 2005). These polymers have been hypothesized to form the critical link between type I *par* motor protein activity established through genetic and biochemical experiments and its biological role in plasmid segregation observed in time-averaged fluorescence microscopy experiments. Still, there is little structural information on the composition of these filaments or any consensus on the

molecular mechanisms used by the filaments to accomplish the segregation of the plasmids. It is by understanding the molecular mechanisms used by motor proteins at the structural level that we can begin to have the confidence to understand how they function in a coordinated way in plasmid segregation.

Although filaments have now been observed in all of the *par* systems, the type II system (ParM) is the only system that has been examined in detail. ParM was observed to form filaments in an ATP dependent manner that were similar to F-actin (Moller-Jensen et al., 2003). More recent efforts to examine the structure of ParM filaments combining data from a variety of sources (X-ray crystallography, electron microscopy (EM), and fiber diffraction) resulted in the generation two different models based upon either GMP_PNP or AMP_PNP negative stain EM data (Orlova et al., 2007; Popp et al., 2008). The resulting models differed from the earlier F-actin based model by suggesting ParM filaments form left handed helices (instead of right as in F-actin), but they differed in their use of subunit conformations at the heart of their respective models. The model (Popp) using a nucleotide-bound ("closed") conformation contained the inter-subunit contacts likely to support polymerization, whereas the other model (Orlova) used an apo ("open") conformation that lacked some of the contacts. The difference appears to be of critical importance to the validity of the filament model as nucleotide is required for the maintenance of filament structure and its loss could account for the dynamic instability observed in the ParM system.

The demonstration of ParM filament formation led to the development of an associated hypothesis that proposed that ParM directs the movement of a *parRC* "partition complex" to its respective pole through a pushing rather than pulling manner. Initial studies on ParM filament formation suggested the dynamics of proto-filament generation were faster than actin and did not require other nucleating factors suggesting that ParM filament formation was spontaneous in the presence of ATP (Garner et al., 2004). The model of ParM dependent partitioning was further refined in reconstitution experiments of the type II RI plasmid system where three components (ParM, ParR, and *parC*) *in vitro* were found to be sufficient to propel "plasmids" (*parC* attached to beads) being pushed apart in an ATP-dependent manner (Figure 1_6A)(Garner et al., 2007). Photo-bleaching experiments on the filaments further demonstrated that the filaments were growing from the ends (where the ParR*C*-beads were attached) confirming that the mechanism of force generation was through the insertion of additional ParM molecules that effectively "pushed" the two bead complexes apart. This was the first piece of concrete proof in support of the insertion model explaining the segregation of plasmids in the type II *par* system. In the next series of studies on a reconstituted *parMRC* partition complex (with gold labeled native *parC*), electron micrographs showed that ParM filaments were being capped by the ParR*C* complex ("partition complex") (Figure 1_6B)(Salje and Lowe, 2008). The imaged structure of the ParR*C* complex along with supporting biochemical experiments lead to the proposal of a "ratchet mechanism" to explain the movement observed in the reconstitution experiments (Figure 1_6B). In the proposed mechanism the ParR*C* complex caps the ATP-bound ParM filament through a high affinity interaction with ParR. As a result of ATP hydrolysis in this ParM subunit, the ParR*C*

complex would rotate away from this ADP-bound (lower affinity) subunit towards the higher affinity ATP-bound ParM subunit of the adjacent filament strand. This rotation would create a channel for the next ATP-bound subunit to move atop the growing filament. The cycle would repeat until some force knocked-off the ParR*C* cap exposing the filament to its inherent dynamic instability (resulting from its ADP-bound or apo state). More recently, freeze-fractured transmission EM experiments demonstrated for the first time the existence of ParM filaments *in vivo* forming bundles proportional to the number of plasmids present in the experiment (proposed to be 2 filaments/plasmid) (Figure 1_6C&D)(Salje et al., 2009).

1.6 The P1 *par* System as a Model System to Study Partitioning.

The type Ia *Escherichia coli* P1 and F plasmid partition systems are well studied and serve as models for the study of partition in type I systems (Barilla et al., 2005; Funnell, 2005; Surtees and Funnell, 2003). P1 ParA (~45 kDa) is a deviant Walker A ATPase containing the signature P-loop sequence KGGXXK[S/T] motif with a functionally important second lysine residue (Koonin, 1993). As a type Ia motor protein, P1 ParA has an extended N-terminus (~100 residues) outside the consensus ATPase domain common to all type I *par* motor proteins. Contained within the N-terminus is a predicted HTH motif believed to enable ParA to fulfill its role in the auto-repression of the P1 *par* operon

(Davey and Funnell, 1994; Davey and Funnell, 1997; Davis et al., 1992). The precise control of ParA and ParB is vital to the health of the cell as the over expression of either component leads to a disruption in partition (Abeles et al., 1985; Funnell, 1988a).

The P1 *par* operon contains all the elements necessary for the successful partition of the P1 Plasmid. Within the P1 *par* operon are the genes for the motor protein, *parA*, a centromere binding protein, *parB*, and the centromere-like site, *parS*. The partition process begins with ParB binding to the *parS* site on a newly replicated plasmid forming an important nucleo-protein complex called the "partition complex" (Schumacher and Funnell, 2005; Surtees and Funnell, 2003). ParA (in the ATP-bound state) is recruited to the "partition complex" through an interaction with ParB where it assists in the segregation of the daughter plasmids (Figure 1_7A)(Bouet and Funnell, 1999; Davey and Funnell, 1994; Fung et al., 2001). ParA's ATPase activity is also stimulated by a similar interaction with ParB (Radnedge et al., 1998). The hydrolysis reaction is essential to partition as hydrolysis deficient mutants no longer stably maintain plasmids in the cell, but the nature of its role is unclear (Fung et al., 2001).

How P1 ParA is able to function both as a transcriptional regulator and as motor protein involved in the segregation of the P1 plasmid is unclear. However, it is clear that ParA switches between these two functional roles by binding distinct adenine nucleotides. ParA, in its ADP-bound state, binds to the promoter region of the P1 *par* operon with a 20-fold increase in affinity (Davey and Funnell, 1994). While ParA, in its ATP-bound state, becomes an activated player essential to the segregation of the P1 plasmid (Davey

and Funnell, 1997). What remains unknown is how this nucleotide dependent switching mechanism relates to the underlying structural changes that allow ParA to alter its function. Being a deviant Walker A ATPase, ParA should dimerize in an ATP-dependent manner (Lutkenhaus and Sundaramoorthy, 2003). The ATP-bound dimer state would make it ideally suited for its role in partitioning. It has been reported that ParA also dimerizes in its ADP-bound state, which is the preferred state for DNA binding (Figure 1_7A)(Davey and Funnell, 1997). The underlying switching mechanism that is responsible for the change in ParA's function is unknown.

One of the most important remaining unanswered questions in the study of type I *par* systems involves the activity of the motor proteins during partitioning. While the molecular mechanism employed by the type II motor protein in segregation is understood to a level bordering on atomic resolution, the mechanisms by which the type Ia and Ib motor proteins function to drive plasmid segregation is less clear. Although recent evidence suggests that both type Ia and Ib proteins also form filaments the mechanical nature of their composition and their activity is not well understood. One point of departure from the type II system is the requirement by some of these type I filaments for a direct interaction with DNA – both chromosomal Soj and type Ib δ form nucleoprotein polymers. Evidence of filament formation in the type Ia system comes from F-plasmid SopA where fluorescence studies suggest they oscillate in spiral formations reminiscent of the cell mid-point regulator MinD (Figure 1_7B)(Hatano et al., 2007). Other reports have suggested that it also forms aster-like projections that help guide the partition complex toward the poles (Figure 1_7C) (Lim et al., 2005). Most of the observational

data comes from time-averaged fluorescence microscopy studies of cells where the motor protein has been fluorescently tagged (GFP). Being at the mercy of the small dimensions of the bacterial cell, the de-convolution of these images results in the gross modeling of motor protein movements. Studies on these type I filaments have yet to address the mechanisms employed by the motor proteins during partitioning so their exact role remains unclear.

1.7 Discussion

All the low copy number plasmid partition systems utilize three main components to ensure the faithful segregation of replicated plasmids to daughter cells. The three essential components (a motor protein, a centromere binding protein, and a *cis*-acting plasmid centromere site) are contained within a single cassette within the plasmid. These systems have been well-studied using genetics and biochemical techniques, which have shaped the understanding of partitioning in general.

Recent efforts at studying the components of these systems using structural techniques have expanded the understanding of the detailed mechanisms involved in the biological process of partitioning. Studies of the type II system have provided a detailed understanding of the partitioning process from the assembly of the "partition complex" to

the dynamic workings of the plasmid segrosome being pushed apart by elongating ParM filaments.

Despite being well studied, the steps involved in partitioning (except for the P1 partition complex) in the type Ia *par* system are less clear. We undertook structural studies (crystallography and electron microscopy) on the P1 ParA protein in an effort to identify and understand the underlying mechanisms present in type Ia *par* motor proteins that account for its activity in partitioning and the auto-repression of its operon (DNA binding).

In the body of the following thesis I will address some of the open questions encountered in the course of the introduction. First, in Chapter 2, I will address the topic of X-ray crystallography, the primary approach used to address questions of ParA structure and function in the thesis. In Chapter 3, the crystal structure of P1 ParA (apo state) will be discussed along with issues pertaining to the oligomeric state of P1 ParA in solution. In the following chapter (Chapter 4), I will address our investigations into ParA's different nucleotide-bound states (ADP and ATP) and how they pertain to ParA's function within the cell. Chapter 5 will cover our attempts to understand the P1 ParA-ParB interaction in relation to ParB's ability to stimulate ParA's ATPase function. In Chapter 6, I will discuss the DNA binding characteristics of P7 ParA. Finally in Chapter 7, I will discuss our attempts to understand the conformational mechanisms ParA may utilize to carry out its functions in the cell.

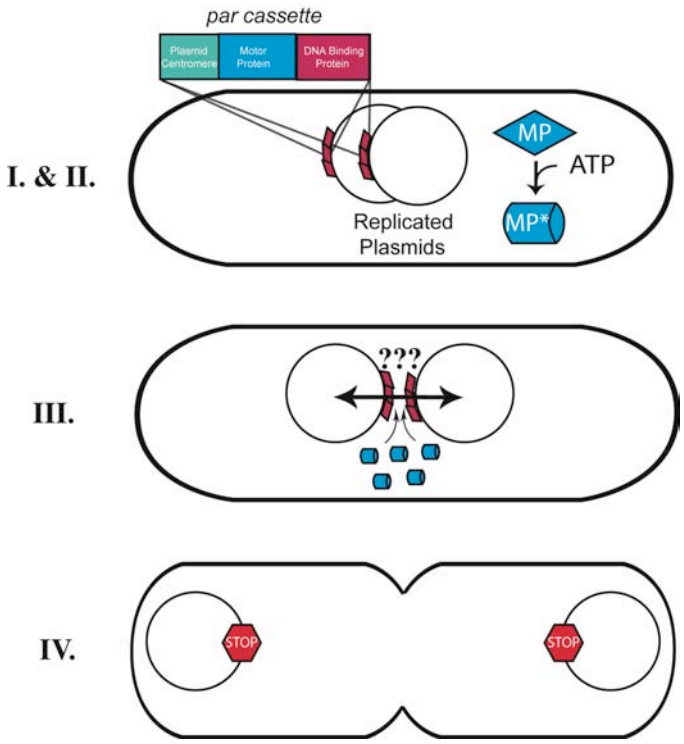


Figure 1_1: Plasmid Partitioning. Partitioning begins with plasmid replication. In low copy number systems 4-6 plasmids are localized to the mid-cell. High copy systems have many more plasmids and are also localized at the mid-cell (**I.**). At this point in low copy systems the centromere binding protein recognizes the centromere-like site on the plasmid and forms a nucleo-protein "partition complex." (**II.**) In both types of systems some pairing or clustering has been observed to take place. In the low copy system the formation of the "partition complex" mediates the recognition and binding of the corresponding motor protein (ATP-bound) to this site. The motor protein then mediates the segregation of the plasmids in a class type-specific manner. High copy systems are thought to rely on diffusion to ensure the distribution of their plasmids. (**III.**) In the end, plasmids in the low copy number systems are localized to either the cell poles or positions ¼ and ¾ of the cell prior to cytokinesis. (**IV.**)

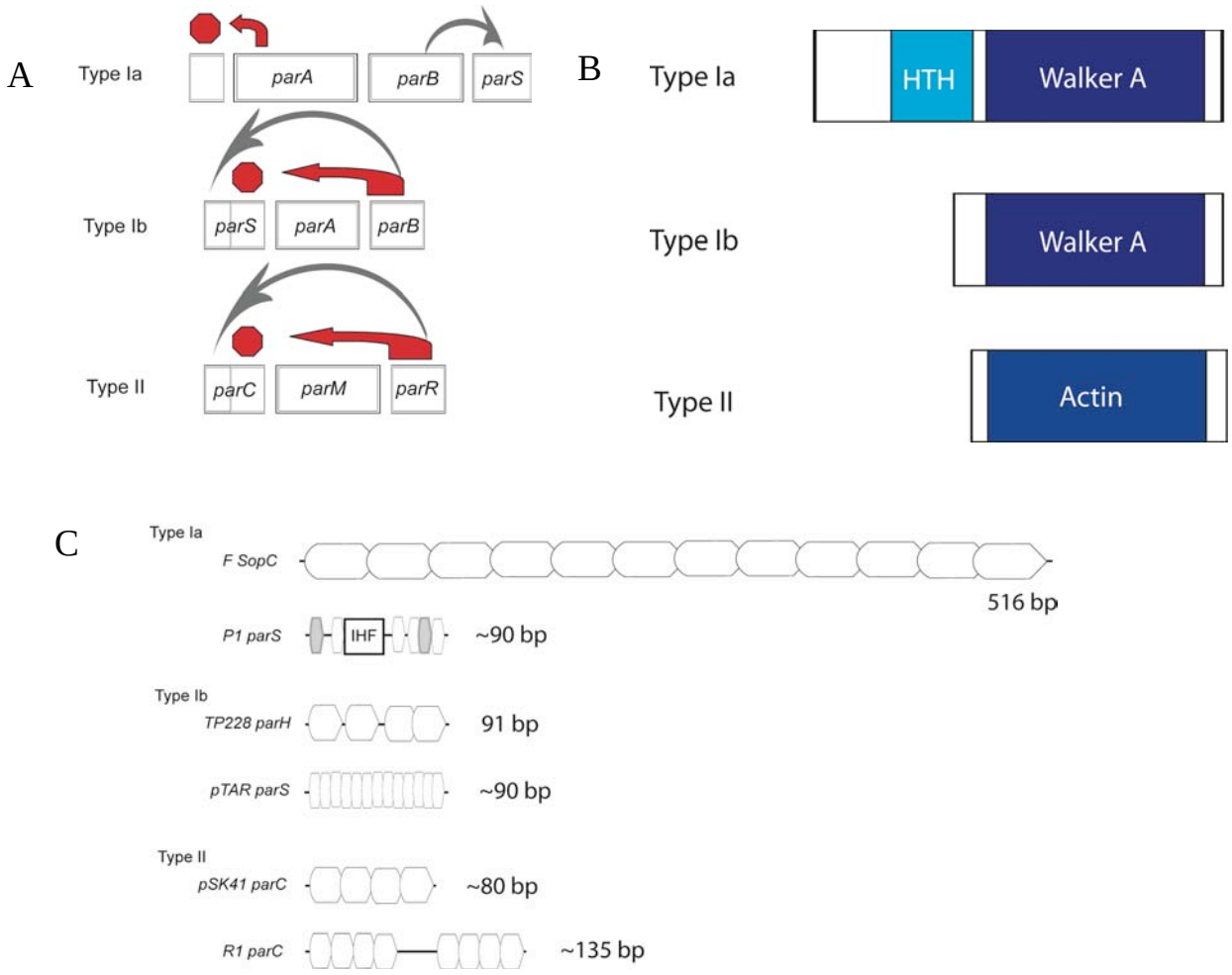


Figure 1_2: The *par* System Cassettes. (A) Schematic of the major *par* cassettes illustrating the interactions between the different components in partitioning (grey) or auto-repression (red). (B) Major domains within motor proteins from each of the major *par* systems. (C) Schematic of plasmid centromere sites from representative members of the major *par* systems (adapted from Schumacher *et al*., 2008)

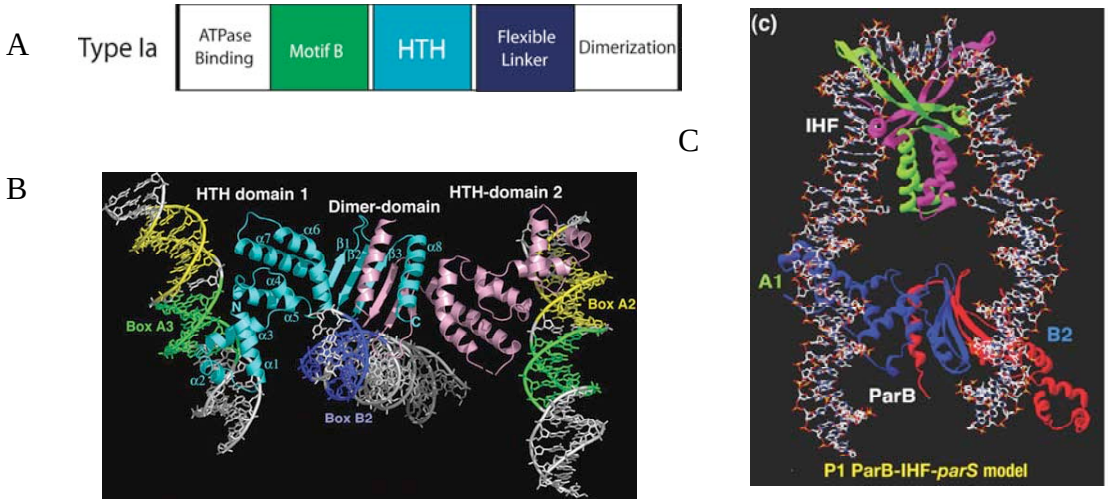


Figure 1_3: Type Ia ParB Protein DNA Binding Characteristics. (A) Schematic of the type Ia ParB protein highlighting the different known domains and characteristics. Motif B is thought to promote ParB oligomerization. (B) Structure of P1 ParB (131-333) bound to 20 bp (A2-A3-B2 box) oligonucleotide of portion of *parS*. (C) P1 ParB-IHF-*parS* "partition complex" model.

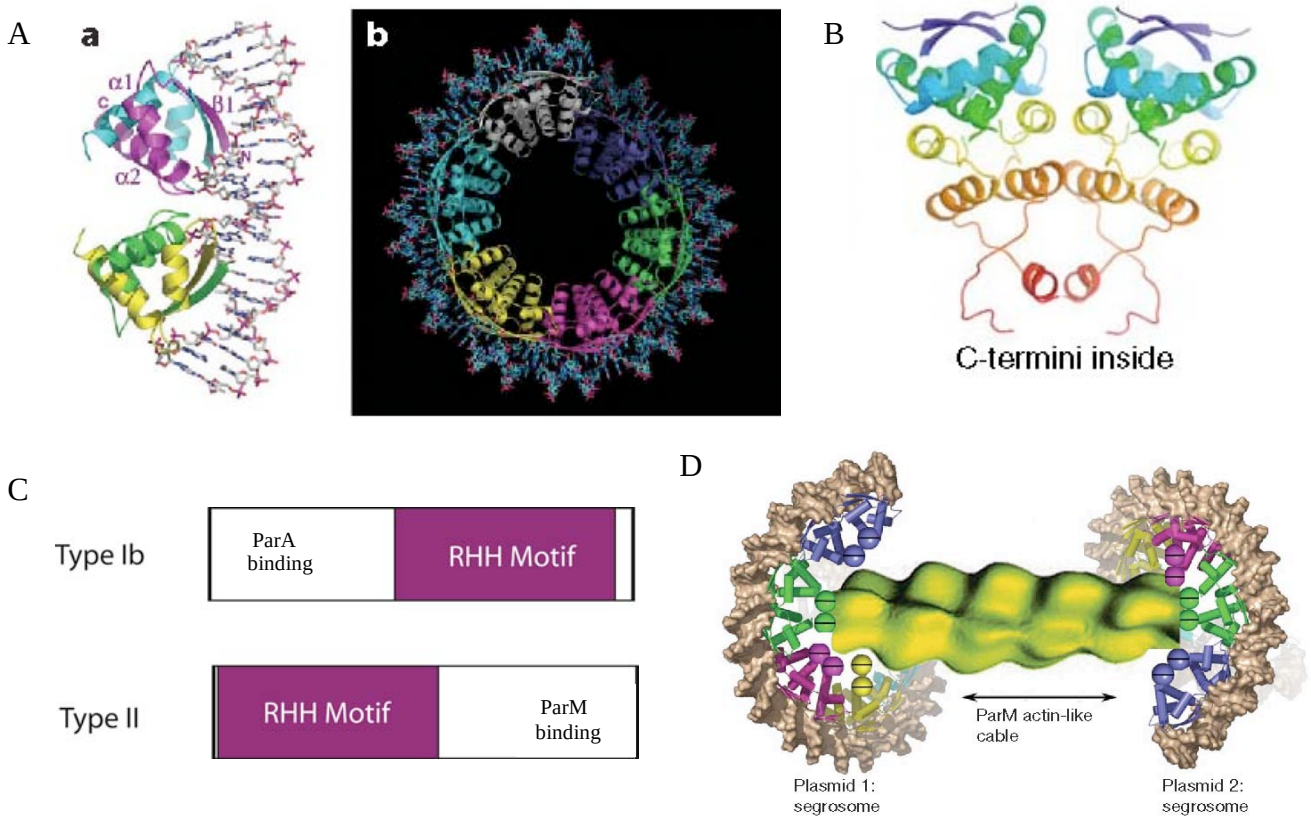
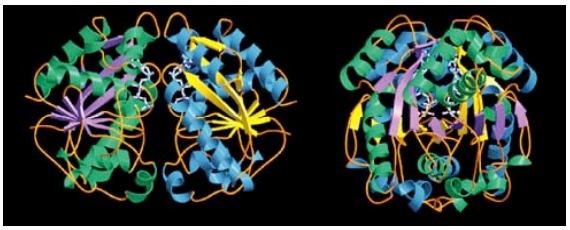
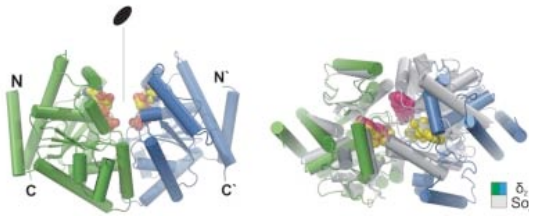


Figure 1_4: Type Ib and type II Centromere Binding Protein DNA Binding Characteristics. (A) Structure of type II pSK41 ParR (1-59) bound to 20 bp segment of *parC*. Two DNA-binding dimers (left) and the pseudo-continuous ParR-*parC* super-helical structure (right)(Schumacher and Funnell, 2005). (B) Apo structure of the type II pB171 ParR (Moller-Jensen et al., 2007). (C) Schematic of type Ib and type II centromere binding proteins illustrating the structurally determined RHH DNA binding motifs. (D) Model of the complete type II segrosome (Schumacher et al., 2007).

A



B



C

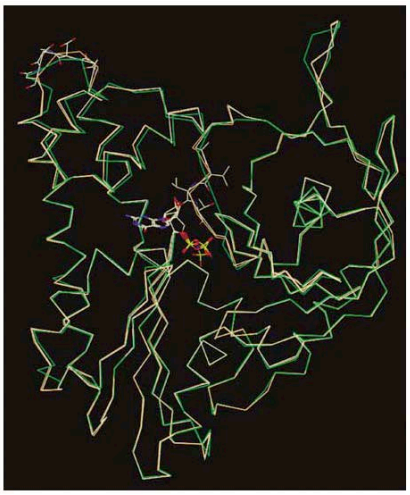


Figure 1_5: Nucleotide Bound State of *par* Motor Proteins. (A) Chromosomal type Ib-like Soj motor protein in its ATP-bound state (Leonard et al., 2004). (B) Type Ib δ motor protein bound to ATPγS (left). Comparison of nucleotide bound Soj and δ proteins (right)(Pratto et al., 2008). (C) Multiple alignment of type II motor protein ParM in its ADP, GDP and GMP-PNP-bound states (Popp et al., 2008).

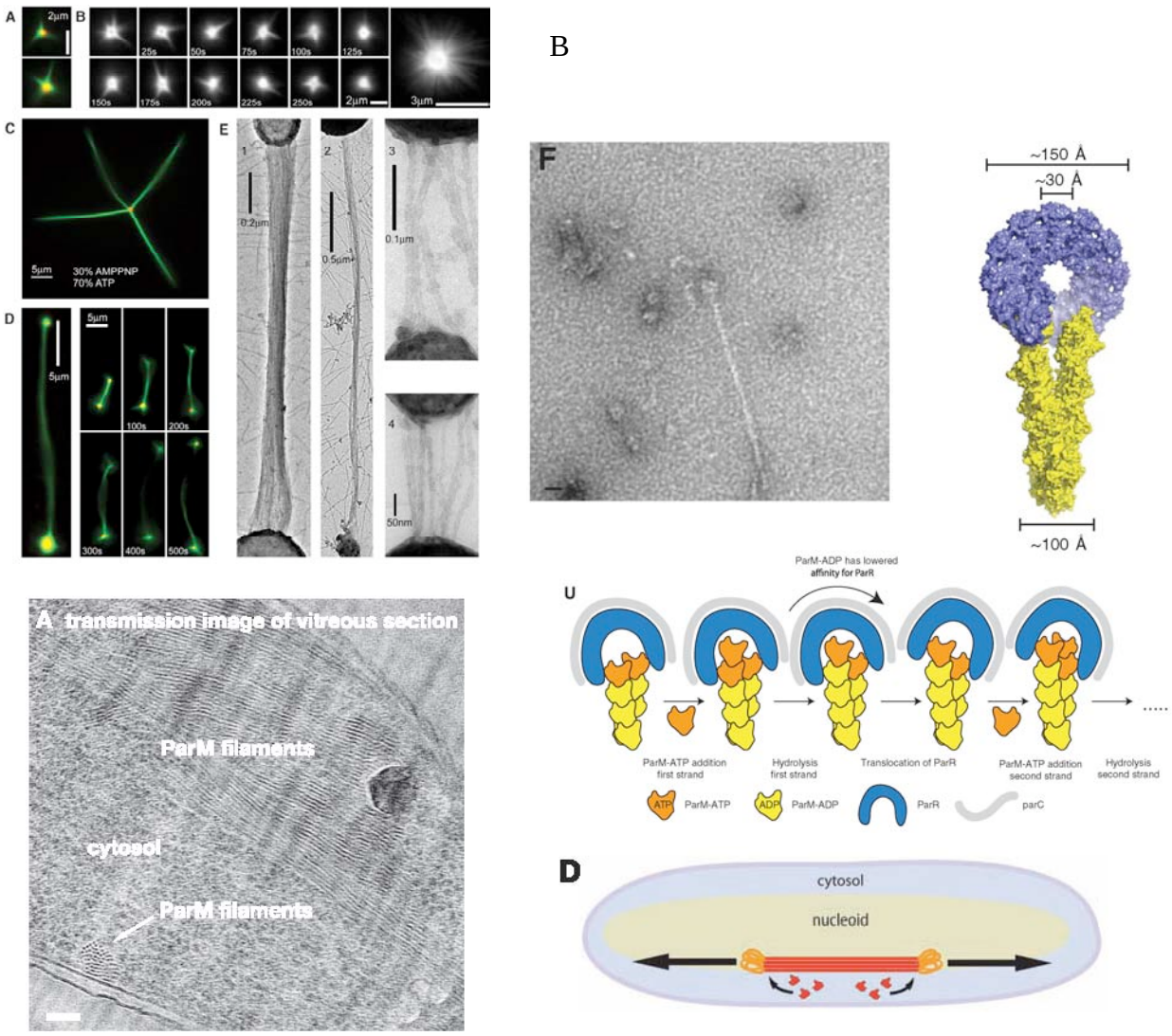
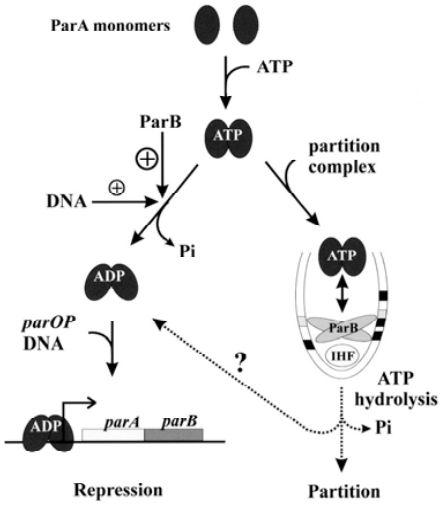


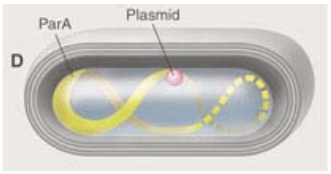
Figure 1_6: The Type II Motor Protein Filament Formation. (A) Type II ParMRC reconstitution studies. TIFM images of Alexa488 ParM forming proto-filaments. Negative stain electron micrographs of filaments forming from beads (Garner et al., 2007). (B) Negative stain electron micrographs of reconstituted system (native *parC* capping filament). Proposed model of ratchet mechanism explaining filament extension (Salje and Lowe, 2008). (C) Freeze-fracture negative stain electron micrographs of filaments *in vivo* (Salje et al., 2009). (D) Model of type II partition.

A



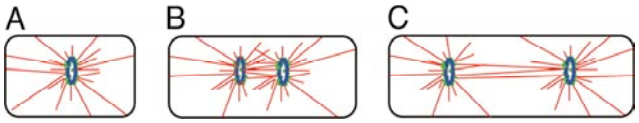
Bouet, JV. Et al. (1999) *EMBO J*. **18**, 1415-1424.

B



Moller-Jensen, J *et. al.* (2004) *Science*. 306, 987 - 989.

C



Lim, GE. *et. al.* (2005) *PNAS.* 102, 17658-17663

Figure 1_7: Type Ia ParA Activity. (A) Model of P1 ParA activity based upon experimental results. (B) Modeled movement of ParA based on time-lapse fluorescence microscopy experiments. (C) Modeled activity of type Ia SopA filaments forming asters.

Chapter 2

Crystallography Techniques used to Solve a Protein Structure

2.1 Introduction

The goal of structural biology is to provide insights into the structural elements utilized by proteins to perform their physiological function. There are a number of techniques that allow one to look at a particular aspect of a protein's structure. They provide a range of data from general landscapes to fine detail all of which contributes to the knowledge of the particular protein or system being studied. There are only a few techniques that are considered high resolution where one can extract mechanistic detail at or near atomic resolution (X-ray crystallography, NMR, electron microscopy – for complexes).

In this chapter, I will cover the basic concepts as well as pointing out the key steps in determining a protein crystal structure. Understanding how a crystal structure is solved allows one to approach a protein structure with the confidence to address questions relating to its function.

2.2 Basic Principles of Crystallography.

There are many techniques based upon the different regions of the Electro-Magnetic (EM) spectrum that can be used to provide useful information on protein structure.

Experiments with light have established that in order to observe diffraction ("see" an object) the size of the wavelength used must be no larger than object one wishes to discern. Translating this in terms of X-rays (λs from 10 to 0.01 nm) means that we may obtain information from X-ray diffraction for objects that are within the same range. As we are interested in the function of proteins, we think of useful structural information in terms of the covalent bonds that compose a protein and the hydrogen and ionic bonds that influence their tertiary structure. This information equates to distances that range from 1 to 3.5 Å apart. In practice, X-rays used for X-ray crystallography are between 0.5 and 2.25 Å. One source of X-rays comes as a consequence of electron bombardment of a metal that results in a characteristic emission spectrum for that metal. The X-rays can then be filtered further to obtain a monochromatic wavelength X-ray. A common source of experimental X-rays comes from generators containing a copper anode, which results in a signature K_α 1.54 Å X-ray. More recent developments have given crystallographers access to a source of tunable X-rays from 0.5-1.6 Å that have expanded the information available from X-ray experiments making structure determination more readily feasible in many cases (see below).

X-rays used in crystallography effectively "see" an atom's electrons. As a result, the information from an experiment represents distributions of electron density. One of the useful results of these interactions is scattering, or diffraction. Like light, scattered X-rays can be described by a periodic wave function:

$$f(x) = F \cos 2\pi (hx + \alpha) \qquad \text{(Eqn. 2.1)}$$

where $f(x)$ describes the vertical height at any position x. As a wave the resulting diffraction is also composed of an amplitude (F), a frequency (h), and a phase (α). Scattering from an atom is proportional to the number of electrons contained within the atom. In this scheme, an atom with more electrons (termed "heavy") like Sulfur will have a strong associated intensity whereas atoms with fewer electrons (termed "light") such as Carbon have a weaker associated intensity. Hydrogen atoms are effectively invisible. A single molecule, a very weak source of diffraction, is impractical for X-ray diffraction techniques and its signal must be amplified to be of use for crystallography. Regular packing of a molecule can result in a crystal, which then may be used as a source of signal amplification for the single molecule. A crystal is composed of a regularly repeating unit extending in all directions (xyz). This unit, termed the unit cell, is the simplest building block that defines the crystal's symmetry. The vectors defining a unit cell are termed **a**, **b** and **c** related by angles α (between **b** and **c**), β (between **a** and **c**), γ (between **a** and **b**)(Figure 2_1A). These six parameters in turn can be used to describe the seven basic crystal systems. As these parameters converge, the number of independent parameters increase, and the internal symmetry decreases (Figure 2_1B). The asymmetric unit is the smallest part of the crystal that can generate the unit cell using symmetry. The asymmetric unit can be a single molecule, part of a molecule or multiple molecules. Symmetry operators for proteins (rotations, translations and screws) can be applied to the asymmetric unit in order to generate the unit cell.

The corners of the unit cell can serve as equivalent positions of a primitive lattice system. The Bravais lattice system originated from the realization that there can be other

equivalent positions beyond the ones at the corners of a primitive lattice. Expanding the equivalent points (non-primitive) to those located on the A, B or C face of the unit cell (Face centered), or to the center (Inner), or to a combination (F) expanded the lattice system while still conforming to the symmetry of the original seven crystal systems. Combinations of symmetry operators (rotations, inversions, etc.) can be used to describe the relationships between equivalent positions in the crystal. These combinations result in 32 unique "point groups." Layering the 32 point groups atop the 14 Bravais lattices leads to the generation of 230 possible "space groups," but in protein (chiral L-amino acids eliminates the possibility of some symmetry elements) crystallography this number is limited to 65. By selecting a higher symmetry the space group from a lattice one can benefit from equivalent reflections, which expands the available options during data collection (collecting less data vs. higher redundancy).

The faces of the of the unit cell can be thought of as planes and these as well as any other repeating set of planes that pass through the unit cell can be thought of a sources of diffraction. These Miller indices, *hkl*, describe all parallel planes. The planes pass through the x-axis at a/*h*, the y-axis at b/*k* and the z-axis at c/*l* and define the number of times the planes cut through each axis in the unit cell and ultimately the number of planes per cell (Figure 2_1C). The entire set of planes acts as a single diffractor and all the spheres along the plane result in a single reflection. Different planes have different perpendicular inter-planar spacing. The associated index (*hkl*) increases as the inter-planar spacing decreases.

Bragg's study of diffraction found that a series of parallel planes with an index hkl and perpendicular inter-planar spacing d_{hkl} produces a diffracted beam when X-rays of wavelength λ impinge upon the planes at an angle θ only if;

$$2d_{hkl}\sin\theta = n\lambda \qquad \text{(Eqn. 2.2)}$$

(Figure 2_2). $2d_{hkl}\sin\theta$ equates to the extra distance traveled by the X-ray to remain in phase. Rays emerging from successive planes are in phase and interfere constructively resulting in a strongly diffracting beam. X-rays incident with the same planes at an angle other than theta are not in phase and no diffractive beam emerges. The intensity of the reflection is proportional to the number of electrons and their distribution along the plane.

To make a connection between the real space lattice of the crystal and the diffraction pattern observed as a result of an experiment one must make the transition to a reciprocal lattice. Where real space is defined by the coordinates x, y, and z, reciprocal space is defined by *h*, *k*, and *l*. The reciprocal space lattice is generated on the perpendiculars of the Miller indices (Figure 2_1C). The lattice spacing is inversely proportional to the inter-planar spacing within the crystal. The resulting unit cell in reciprocal space is either larger (if d_{hkl} is small) or smaller (if d_{hkl} is large) than in real space.

The number of measurable reflections depends upon the unit cell lengths in the crystal and the X-ray wavelength. The mathematician Ewald recognized a relationship between diffraction and reciprocal space could be made using the value 2θ for the total angle of

deflection and a convenient geometrical construction, Ewald's sphere (Figure 2_3). To construct Ewald's sphere an X-ray represented by a vector *XO* is drawn through a point O, the origin of the reciprocal lattice, such that a point C (serving as the center of the sphere) is $1/\lambda_{xray}$ away. The construction determines that diffraction will occur if a vector emerging from point C ("the crystal") that is 2θ from *CO* and intersects a lattice point from the reciprocal lattice at the surface of the sphere (radius $1/\lambda_{xray}$). By adapting Bragg's law to Ewald's construction:

$$\sin\theta = \tfrac{1}{2}\, OP / 1/\lambda \qquad \text{(Eqn. 2.3)}$$

diffraction will occur when the line segment between O (the reciprocal lattice origin) and P (the lattice point) is equal to $1/d_{hkl}$ and n equals 1. Diffraction in reciprocal space takes place whenever a lattice point intersects a sphere constructed with a radius of $1/\lambda_{xray}$. The construction also allows one to determine the angle of diffraction relating a reflection in reciprocal space (hkl) to real space (xyz). The reciprocal lattice rotates around the point O as the crystal is rotated such that any point on the lattice that intersects the surface of the sphere (radius $1/\lambda_{xray}$) will result in diffraction. Any point within a radius of $2//\lambda_{xray}$ from the reciprocal lattice origin, O, (the limiting sphere) has the potential to be a reflection.

2.3 Data Collection and the "Phase Problem."

Now that we have established the means for collecting the data, we need a way of finding what the data say about the molecular structure. The ultimate goal is to find a way of calculating the 3-dimensional electron density (ρ (x, y, z)) of the molecule in the form of a 3-dimensional contour map. Fourier theory tells us that any periodic function can be approximated with the inclusion of the proper amplitude, frequency and phase. As electron density is repeated in every unit cell it can be represented by a wave function. To help find where the atoms are located within the unit cell we can enlist the help of a Fourier transform. The Fourier transform provides the mathematical relationship between an object and its diffraction pattern. A reflection can be described in terms of a structure factor equation with one term for each atom or small volume of the unit cell (Eqn. 2.4).

$$F_{hkl} = \Sigma_j f_j \, e^{2\pi i (hxj + kyj + lzj)} \qquad \text{(Eqn. 2.4)}$$

The basis of the structure factor equation (Eqn 2.4) is a mathematical summation of sine and cosine wave functions. The resulting wave function is a complex number (a vector) of the form ($a + ib$) where i is an imaginary number equal to $-1^{\frac{1}{2}}$ (Figure 2_4A). The scattering factor, f_j, is a function dependent on the number of electrons in the jth atom (or on the atom type) and marks the amplitude of the contribution. The terms (x, y, z) are the fractional coordinates in the unit cell of the jth atom. The terms (h, k, l) represents both

the frequencies of the wave in the x, y, z directions and the indices of a specific reflection (reciprocal space). The resulting contribution of the *j*th atom to the structure factor, F_{hkl}, is dependent upon the elemental composition and its location in the unit cell (x_j, y_j, z_j), which establishes the phase of its contribution.

The unit cell can also be divided into very small volume elements where the electron density at (x, y, z) can be represented by the average electron distribution, $\rho(x, y, z)$, of the area. F_{hkl} can alternatively be expressed in terms of each volume element of electron density in the unit cell (Eqn. 2.5 or Eqn. 2.6);

$$F_{hkl} = \int_x \int_y \int_z \rho(x, y, z)\, e^{2\pi i (hxj + kyj + lzj)}\, dxdydz \qquad \text{(Eqn. 2.5)}$$

or

$$F_{hkl} = \int_V \rho(x, y, z)\, e^{2\pi i (hxj + kyj + lzj)}\, dV \qquad \text{(Eqn. 2.6)}$$

where V is the unit cell volume. As the Fourier transform above represents the relationship between the electron density and the diffracted X-rays, an inverse Fourier transform can be used to calculate the electron density (Eqn. 2.7);

$$\rho(x, y, z) = (1/V) \Sigma_h \Sigma_k \Sigma_l \left| F_{hkl} \right| \exp^{-2\pi i (hxj + kyj + lzj) + i \alpha_{hkl}} \qquad \text{(Eqn. 2.7)}$$

where $\rho(x, y, z)$ is composed of a sum of all structure factor amplitudes, $|F_{hkl}|$, and phases, α_{hkl}. The structure factor amplitude can be obtained from the measurements of reflection intensities (Eqn. 2.8).

$$I_{hkl} \propto |F_{hkl}|^2 \qquad \text{(Eqn. 2.8)}$$

Unfortunately the phase component is lost in diffraction. It is a critical component for determining the electron density map and overcoming its absence constitutes what is termed the "phase problem."

The three most common methods for determining phases in crystallography experiments are: Molecular Replacement (MR); Multiple Isomorphus Replacement (MIR); and Multiple Anomalous Diffraction (MAD). The goal of all three of these methods is to allow one to determine the phases of each of the reflections in the unit cell. As each reflection has its own phase and the phase component carries the most effective information for resolving a structure, it is critical that the determined phases are as accurate as possible to ensure the resulting electron density map is useful.

In MR, the initial phases are calculated from a known structure with strong sequence homology (>30%) to the protein of interest. Since the phases of the homolog are used in conjunction with the intensities from your protein, the key operation involves the overlaying the homolog onto the density of the protein of interest. The known structure serves as a search model, which is compared with your protein in Patterson space using rotation and translation operations. The advantage of this method is that it only requires a native data set (but comes at the risk of using non-experimentally derived phases).

A second method used to determine the phases is MIR. This method is based upon the ability to derivatize a crystal form for which you have a native data set. As protein crystals usually are composed of 30-70% solvent, channels within the crystal provide an avenue for the introduction of elements that have more electrons. These "heavy atoms," often Pt, Hg or Au, can be complexed to the surface of proteins based upon their chemical properties. As these "heavy atoms" will have a stronger signal than the surrounding atoms the resulting diffraction pattern from a derivatized crystal will contain differences when compared to the native data set. These differences can be used to estimate phases. To generate useful phases one must use multiple derivatives that have distinct heavy atom sites. The key caveat is that the resulting derivative must not disturb the crystal packing and thus the space group of the native crystal – it must be isomorphus. The difficulty of this method comes in finding derivatives that maintain their diffraction qualities – meaning the screening of many crystals.

A third method for determining phases is MAD (Hendrickson, 1991; Hendrickson et al., 1990). I will be spending the most time on this method as it is the method that proved successful in determining the phases of apo P1 ParA. One component of this method comes from having access to a tunable source of X-ray radiation (0.5 to 2 A), which is provided by a synchrotron source. Atoms scatter X-rays and the scattering factor, *f,* is dependent upon the atom being scattered and the wavelength of the X-ray. The other component comes from having an ability to introduce an anomalous diffractor, a "heavy-atom," into the sequence of a protein. By inhibiting the synthesis of methionine (exposing bacterial cells to Thr, Lys and Ile), seleno-methionine (selenium) is introduced

into the protein sequence in place of methionine (Doublié, 1997). The benefit of this method is that only one crystal is required, so isomorphism is perfect, but the data collection demands are extensive.

2.4 Multiple Anomalous Diffraction (MAD).

The magnitude of the anomalous scattering component is dependent upon the number of electrons in the source and the wavelength of the X-rays. The source of scattering (Thompson) in X-ray experiments normally comes from the outer electrons, which are independent of wavelength, f_o. Atoms scatter anomalously when the source of scattering involves the inner electrons and is a result of selecting a wavelength E near the atom's absorption edge. Anomalous scattering is out of phase with normal scattering and has two components f' and f'' in addition to f_o that can be related to F(*hkl*), the structure factor, by Eqn. 2.9.

$$f_o + f' + f'' = \text{F}(hkl) \qquad \text{(Eqn. 2.9)}$$

f' is the real component and is 180° out of phase with the normal and can be thought of as absorption. f'' is an imaginary component and is 90° out of phase. Both components are wavelength dependent. The most common source of the anomalous signal is a Selenium

atom (atomic number 34). f'' is a measurable quantity equal to fluorescence. f' can be calculated from Kramer-Kronig relation that allows the calculation of the imaginary out of phase component from the in phase dissipative force and visa versa (Eqn. 2.10).

$$f'(\omega) = 2/\pi \int_0^\infty \omega' f''(\omega') \, \partial\omega' / \omega^2 - \omega'^2 \qquad \text{(Eqn. 2.10)}$$

Anomalous scattering occurs near the absorption edge. A wavelength scan measuring fluorescence is used to empirically determine f'' (max fluorescence – "peak"), f' ("inflection"), and f_o (remote). The differences are very small so the collection must be very precise. The selection of wavelengths is required to maximize the signal of each component f', f'', and f_o.

According to "Friedel's Law" reflections from normal scattering (f_o) 180° apart (*hkl* and *–h-k-l*) have equal magnitude but opposite phase (Eqn. 2.11)(Figure 2_4B).

$$|F(hkl)| = |F(-h-k-l)| \qquad \alpha_H(hkl) = -\alpha_H(-h-k-l) \qquad \text{(Eqn. 2.11)}$$

In anomalous scattering Friedel's law is broken as the f'' term is 90° out of phase and is always positive (Eqn. 2.12)(Figure 2_4B).

$$|F(hkl)| = |F(-h-k-l)| \qquad \alpha_H(hkl) \neq -\alpha_H(-h-k-l) \qquad \text{(Eqn. 2.12)}$$

What is measurable in anomalous scattering is F_{PH}, the component that contains both protein and heavy atom contributions as related by Eqn. 2.13.

$$F_{PH}(hkl) = F_P(hkl) + F_H(hkl) \qquad \text{(Eqn. 2.13)}$$

The observable anomalous difference, ΔF_H, is determined from the relation in Eqn. 2.14.

$$\Delta F_H = \big|\,|F_H(hkl)| - |F_H(-h\text{-}k\text{-}l)|\,\big| \qquad \text{(Eqn. 2.14)}$$

Another source of anomalous differences in a MAD experiment are the Bragg reflections that are space group symmetry equivalents to a Friedel pair called "Bijvoet pairs." They have unequal amplitudes in the presence of anomalous scattering and the differences may also be used for phase determination. All the anomalous differences are very small – equal to a few percent – thus it is important to use all the information that is physically available. To maximize the data available in a crystal it is useful to collect data using an inverse beam strategy, or a small wedge of data 180° away (although this means collecting up to 6 data sets on one crystal - peak, inflection and remote). This strategy works best if your crystal has high symmetry and is not overly sensitive to radiation damage.

In order to use the anomalous differences to find the phases the heavy atom sites must be located. A special Fourier summation is used called a Patterson function where intensities, or $|F|^2$, are used along with coefficients without phase information (phase

angles set to 0) (Eqn. 2.15). Coordinates (u, v, w) are used to locate a point on a Patterson map. A Patterson map is a contour map equal to the size and shape of the unit cell where peaks are representations of the inter-atomic vectors between two atoms in the unit cell. To find the anomalous scattering sites a difference Patterson function is used calculated from the Bijvoet difference, ΔF_H (Eqn. 2.16). The results are displayed on a difference Patterson map where the (u, v, w) coordinates of the inter-atomic vectors between pairs of anomalous scattering atoms are represented as peaks.

$$P(u, v, w) = 1/V \, \Sigma_h \Sigma_k \Sigma_l \, | F_{hkl} |^2 \cos 2\pi (hu + kv + lw) \qquad \text{(Eqn. 2.15)}$$

$$P(u, v, w) = 1/V \, \Sigma_h \Sigma_k \Sigma_l \, \Delta F_H^2 \cos 2\pi (hu + kv + lw) \qquad \text{(Eqn. 2.16)}$$

A Patterson map can be hard to interpret as the number of peaks is determined by the relationship (n^2-n), where n is the number of atoms within the asymmetric unit. Patterson peaks generated by symmetry operations are present on Harker sections. A Harker section allows one to look at the reflection from one set of indices – like the unit cell edges. Patterson space coordinates *uvw* can be converted to real space coordinates xyz, so locating all the possible sites allows one to calculate the estimated phases.

By establishing the positions of the anomalous scattering atoms, one can calculate the vectors for F_H (*hkl*) and F_H *(-h-k-l)*. The amplitudes of both F_{PH} (*hkl*) and F_{PH} *(-h-k-l)* are available from the derivative data set. We also know the component F_P without the phase angle. To establish the phase of a reflection it is necessary to complete a Harker

construction, which begins with the placement of F_H (*hkl*) on an Argand diagram (Figure 2_4C)(Schumacher, 2004). Next, F_H (*hkl*) is translated so the tip is now at the origin. F_P is now set with its tail at the origin and becomes the radius of the first circle (grey). F_{PH} (*hkl*) is now aligned so its tail originates at the tail of F_H (*hkl*) becoming the radius of the next circle (red). Now there are at least two distinct phase angle possibilities for F_P (red lines, F_{PH}, intercepting red circle). F_H *(-h-k-l*) can now be used to complete the triangle by invoking Friedel's law to make the following relationship F_P (*hkl*) = F_P *(-h-k-l*) = F_{PH} *(-h-k-l*) - F_H *(-h-k-l*) (grey rotation). By placing F_H *(-h-k-l*) with its tip on the origin we can draw a new circle (blue) with radius F_{PH} *(-h-k-l*) and complete the Harker construction. The remaining ambiguity in the phase angle should be resolved.

With the initial phases in hand, the first electron density map can be calculated. The figure of merit (FOM), *m*, is a weighting factor (zero to 1) calculated from a probability distribution of the potential errors in the phases ($m = \overline{\cos\{\alpha - \alpha_{best}\}}$). As phases are crucial to the generation of an electron density map, which is ultimately the basis for the model they are often weighted for confidence using FOM as a coefficient that determines that phases' contribution.

This process is integrated into the software package SOLVE which uses a combination of anomalous Pattersons (Eqn. 2.17) and "free" self difference Pattersons (Eqn. 2.18) to locate heavy atom sites and compute phases (Terwilliger and Berendzen, 1999). The results are scored and ranked on the basis of a Z score.

$$\rho\,(u, v, w) = 1/V\ \Sigma_h\Sigma_k\Sigma_l\ \left|\,|F_{PH}\,(+)| - |F_{PH}\,(-)|\,\right|^2 \cos 2\pi\,(hu + kv + lw) \qquad \text{(Eqn. 2.17)}$$

$$\Delta\rho\ (x, y, z) = 1/V\ \Sigma_h\Sigma_k\Sigma_l\ \ |F_H\ (hkl) - F_P\ (hkl)|^2\ e^{\ -2\pi i\ (hu + kv + lw) + i\alpha_{hkl}} \qquad \text{(Eqn. 2.18)}$$

Once the initial phases have been determined they can be processed further in an effort to improve them (phase extension) and make the resulting first electron density map more useful. The process is termed Density Modification (DM) and entails the use of outside information ("prior knowledge") to refine the protein borders within the unit cell. One essential piece of outside information required for this process is the determination of the percent solvent of the unit cell. The percent solvent of the unit cell fixes the parameters for the density that may be disregarded, considered disordered non-protein, during the process of defining the protein's electron density, which is used to refine the phases. Empirical studies of protein crystals determined that the ratio of the unit cell volume and a protein's molecular weight falls between 1.7 and 3.5 $Å^3$/Da (Matthews, 1968). This number termed the Mathews' coefficient, V_m, can be used to estimate the percent solvent in a unit cell as well as the # of molecules in the asymmetric unit. This is derived from the relationship:

$$\begin{aligned} V_{protein} &= (\text{Vol. of the protein in the unit cell})/V_{cell} \\ &= \frac{(Z * M_r * \text{specific Vol. of the protein})/ N}{V_m * Z * M_r} \\ &= \text{specific volume } (cm^3/g) / V_m\ (Å^3/Da) \end{aligned}$$

where V_{cell} is the unit cell volume; M_r is the molecular weight of the protein; and Z is the number of molecules in the unit cell. The specific volume of a protein has been

experimentally approximated at 0.74 cm^3/g, resulting in the following relationship $V_{solvent} = 1- 1.23/V_m$.

Once the contribution of the solvent has been established the electron density can be fractionated and evaluated based upon the electron density of the fraction. In solvent flattening the fractions equated with a lower density are set to a prescribed value equating the area as solvent (Solvent flipping sets the value as a negative number sharpening the distinction). The result of the process is to establish a mask around the protein in the unit cell, which can be used to generate new phases. As the average unit cell contains between 30-70% solvent, which is a uniform density distribution any effort to subtract this from the phase calculations reduces phase angle errors. Another process involves the assessment of the electron density of an area and comparing the value with a set of established values for protein densities. This process is termed Histogram matching and is another way of defining the regions of protein within the unit cell. A third method depends upon the existence of identical protein molecules within the cell that are not related by crystallographic symmetry. In a process termed NCS averaging the electron densities of these molecules can be averaged in an effort to improve the phases calculated from known protein in the unit cell. Density modification is an iterative process where all the phase contributions are added through phase combination where each contribution is weighted (FOM coefficients) as a measure of its quality and its impact in computing the final phases.

Through efforts in Density Modification the phases become refined enough that they generate a useful electron density map that has a clear continuous density (connectivity) and where secondary structural features may be easily distinguished (Figure 2_5). At this point you reach the stage of model building (using a program such as O) where the crystallographer begins to manually arrange protein segments into the electron density (Jones, 1991). First, armed with the proteins primary sequence and the location of heavy atoms (SelMet for Met) the main chain connectivity (bond angle restrictions) prescribed by secondary structural elements is determined. In early maps it may be enough to distinguish the presence of large bulky side chains from smaller side-chains in an effort to find your way through the protein. Variation in the quality of electron density as you proceed ranging from clearly identifiable main chain and side chain electron density to areas where there is a clear lack of density suggesting the main chain may be disordered in the crystal. Ultimately, the goal is to build in protein density that improves the calculation of phases without introducing error.

2.6 Refinement.

Once a model is built into the electron density, the goal shifts to improving the resolution of the model and eliminating the errors. The process is termed refinement. Refinement is an iterative process that involves moving from real space to reciprocal space with the

goal of improving the agreement between structure factor amplitudes calculated from the model (F_{calc}) and those from the collected data (F_{obs}). An important consideration in the process is that the only experimentally measured value is the diffraction data reflected as Intensity ($I_{hkl} \propto |F_{hkl}|^2$), and therefore, it is only through structure factor amplitudes that we can compare the model with experimental observations.

In real space refinement certain resources can be used to guide the process of improving the model so that it is a better representation of the collected data. These resources are in the form of special electron density maps calculated from structure factors that represent the difference between the calculated and observed data. As phases are important and contain errors that may bias the resulting model the goal is always to attempt to emphasize the collected data, F_{obs}. One way this is accomplished is through the weighting of the respective structure factors used to calculate the difference maps. These maps designated F_o-F_c or $2F_o$-F_c guide the crystallographer in making adjustments to the model that are more representative of the observed data. In F_o-F_c maps, the most unbiased map, there is both positive and negative density – making the map harder to interpret. In $2F_o$-F_c maps there is more continuous positive density, but it introduces some phase bias back into the map. To minimize the bias of information from sections of the structure with poor corresponding electron density, the model from that section can be removed when calculating the improved phases. The resulting "omit" map may result in electron density in the section in question that is less biased and of better quality.

The goal in reciprocal space refinement is to computationally match the structure factors of the model to those of the observed data by making slight adjustments to the atom locations in the model. Least squares is one method used in reciprocal space refinement, which tries to establish a function where the differences between the observed and calculated values (structure factors) are minimized (Eqn. 2.19). In doing so, atoms in the model are adjusted (x, y, z) in order to result in a structure factor that fits the data more closely. These adjustments to an atom's position can be limited by real world constraints and restraints, which are added to the parameters in the transform.

$$\Phi = \sum_{hkl} w_{hkl} \left(\left|F_{obs}(hkl)\right| - \left|F_{calc}(hkl)\right| \right)^2 \qquad \text{(Eqn. 2.19)}$$

Besides the inclusion of a weighting factor other parameters can be added to the calculation of the structure factors of the model including a temperature factor, B_j. The temperature factor accounts both for the movement of an atom's position around a "resting point" and the resulting occupancy of that atom in the structure. B_j influences the resulting structure factor in a manner dependent upon the angle of reflection [(sin θ)/λ].

Another refinement method is Maximum likelihood (ML), which tries to arrive at the best model by making sure it is statistically (using probability) the most consistent with the observations. The algorithm used in ML is based upon the probability of the data given the model p(data; model). It operates by establishing a probability distribution of structure factors (or with the phases removed - structure factor amplitudes). The shape of

distribution is then used as a reference point for comparing models. By comparing the models to probability distribution with that of the data the likeliness of the model given the data can be calculated. In an iterative process the aspects of the model that fit the data well are retained while the other sections are allowed to sample more space eventually leading to the best outcome. Errors are assumed to have a bell shaped distribution.

R-factor, or residual index, is a measure of the convergence of the model to the collected data as reflected by the structure factor amplitudes ($I_{hkl} \propto |F_{hkl}|^2$). An R factor is a useful a tool to tell how your model is progressing during the refinement process. The

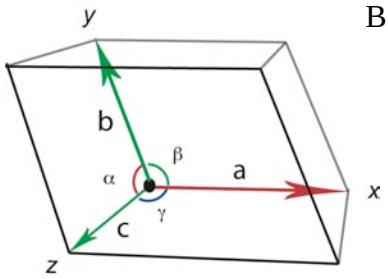
$$\text{R-factor (\%)} = \pm \frac{\sum_{(hkl)} \left| |F_{obs}(hkl)| - |F_{calc}(hkl)| \right|}{\sum_{(hkl)} |F_{obs}(hkl)|} \times 100 \qquad \text{(Eqn. 2.20)}$$

iterative refinement process is normally continued until a converging R-factor value is attained. An R-factor of 0 is perfect and 60 is random. An R factor is limited by the resolution of the data, but a value of 20-30% is generally considered acceptable. R_{free} is another more unbiased way of judging the true convergence of the model and the observed data. R_{free} is generated through the removal of a random 5-10% section of the model before using a Fourier transform (Eqn. 2.6) to calculate the improved structure factors (and phases), which are then used by an inverse Fourier transform (Eqn. 2.7) to calculate a new electron map. At the end of the refinement process the 5-10% portion of

the model that has been withheld from the calculation of improved phases is used to judge the unbiased quality of the model.

At the end of the refinement process, once the R-factor has converged to an acceptable end-point, the resulting model can be checked against established parameters to validate the structure. A Ramachandran plot is one way to judge the quality of the structure in classifying the phi-psi bond angles to make sure they are within the energetically allowable range (Laskowski et al., 1993).

A



B

Crystal Systems	Number of Independent Parameters	Parameters
Triclinic	6	$a \neq b \neq c$; $\alpha \neq \beta \neq \gamma$
Monoclinic	4	$a \neq b \neq c$; $\alpha = \gamma = 90^o$; $\beta > 90^o$
Orthorhombic	3	$a \neq b \neq c$; $\alpha = \beta = \gamma = 90^o$
Tetragonal	2	$a = b \neq c$; $\alpha = \beta = \gamma = 90^o$
Trigonal		
rhombohedral lattice	2	$a = b = c$; $\alpha = \beta = \gamma \neq 90^o$
hexagonal lattice	2	$a = b = c$; $\alpha = \beta = 90^o$; $\gamma = 120^o$
Hexagonal	2	$a = b = c$; $\alpha = \beta = 90^o$; $\gamma = 120^o$
Cubic	1	$a = b = c$; $\alpha = \beta = 90^o$

C

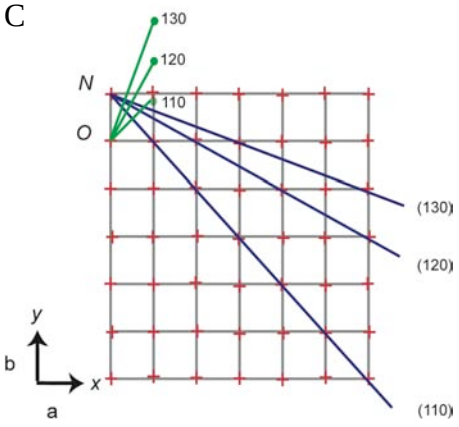


Figure 2_1: Crystal Lattices. (A) Unit cell with 6 different parameters: lengths a (x-direction), b (y-direction), c (z-direction), and angles α, β, γ. (B) Seven crystal systems. (C) Two-dimensional lattice (grey). Real lattice planes (blue). Construction of the reciprocal lattice. Origin of the reciprocal lattice, O. Reciprocal lattice planes and points (green)(adapted from Rhodes).

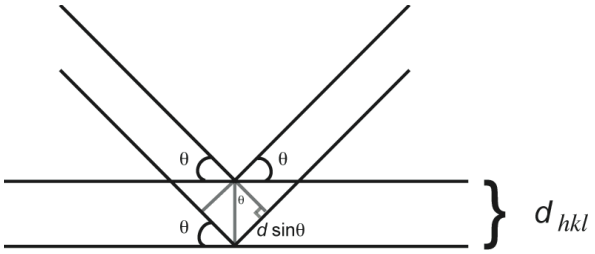


Figure 2_2: Bragg's Law. Conditions for strong diffraction and satisfaction of Bragg's law when the path length equals $2d_{hkl}\sin\theta = n\lambda$. d_{hkl} is the perpendicular inter-planar spacing of the lattice. Black lines intersecting at angle θ are incident and diffracting X-rays.

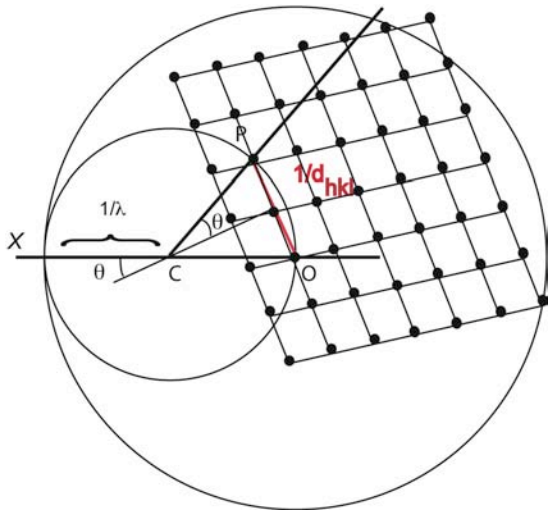


Figure 2_3: Ewald's Sphere. Two-dimensional representation of Bragg's law in reciprocal space. The radius of the sphere of reflection is 1/λ. O is the origin of the reciprocal lattice. Diffraction will occur when a lattice point intersects the sphere (when OP = $1/d_{hkl}$).

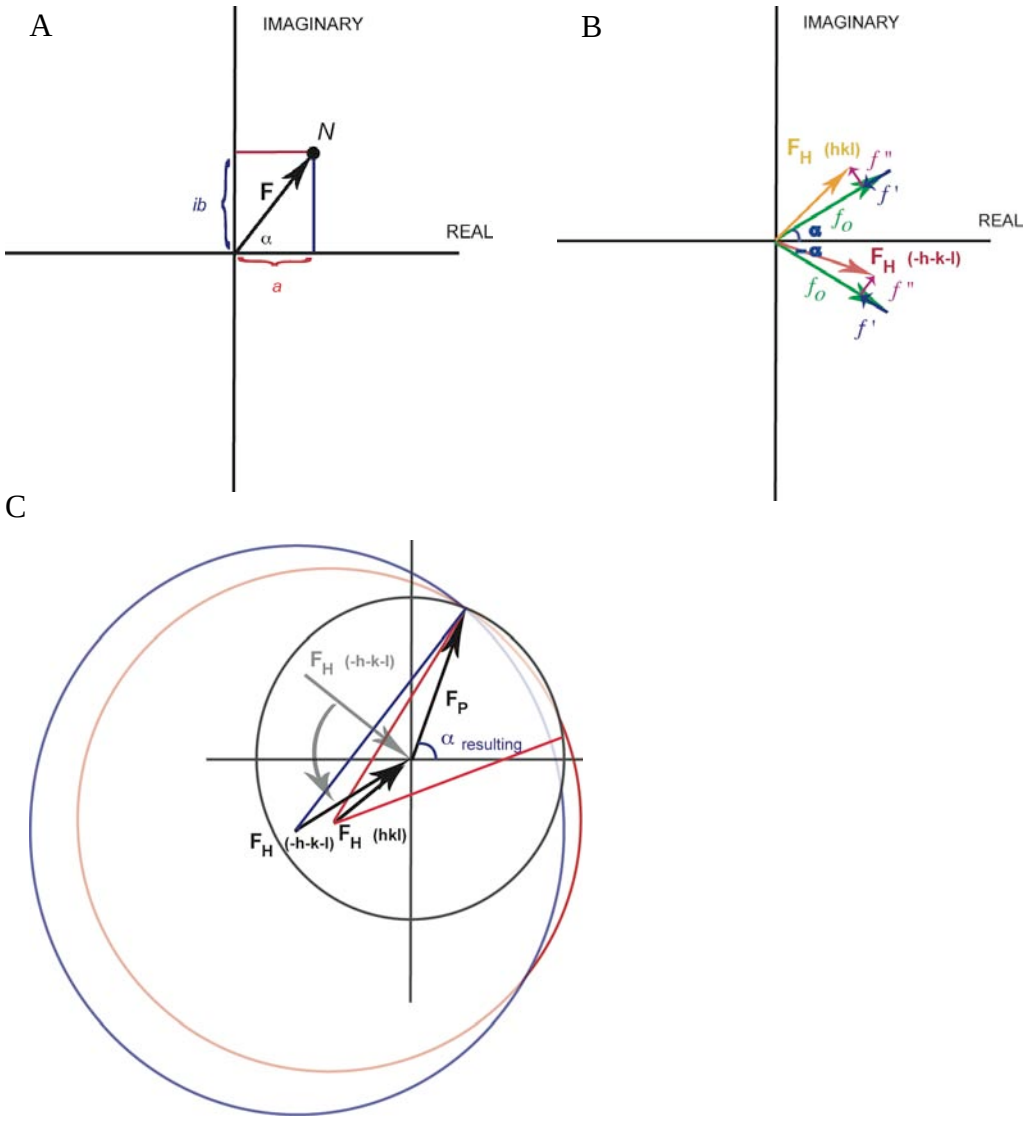


Figure 2_4: Phase Determination. (A) Argand diagram. Description of a complex number with real and imaginary components (adapted from G. Rhodes). (B) Friedel Pairs. Anomalous scattering components: f_o, Thompson scattering (green); f', real (blue); f'', imaginary (pink) (Schumacher, 2004). (C) Harker Construction. Diameter of black circle, $|F_P(hkl)|$; diameter red circle, $|F_{PH}(hkl)|$; diameter blue circle, $|F_{PH}(-h-k-l)|$. Through the relation $F_P\,(hkl) = F_P\,(-h-k-l) = F_{PH}\,(-h-k-l) - F_H\,(-h-k-l)$ the grey arrow translation of $F_H\,(-h-k-l)$ resolves phase ambiguity.

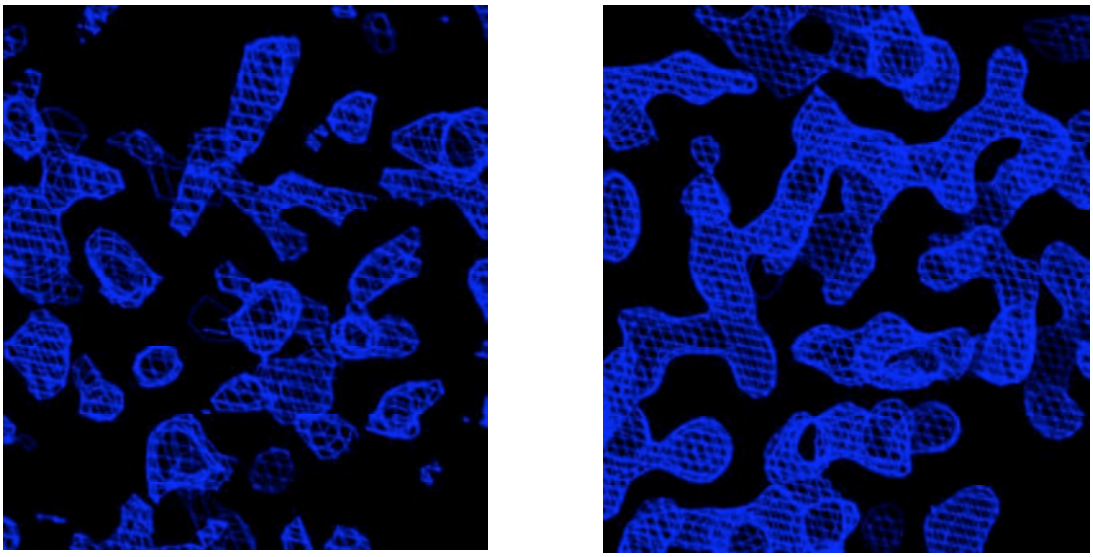


Figure 2_5: Density Modification (DM). Example of the results of DM on an electron density map. Electron density map calculated from initial phases from SIR (left). Electron density in the same map after DM (right). Example borrowed from class power point presentation (Schumacher, 2004).

Chapter 3

The Apo P1 ParA Structure and Oligomeric State

3.1 Introduction

The type I *par* system is broken into two sub-families, type Ia and type Ib, based upon the identity of their motor proteins. The type Ia *par* motor proteins are larger (~ 40 kDa) than the type Ib motor proteins (~ 30 kDa)(Gerdes et al., 2000). The type Ia *par* motor protein has an extended N-terminus of roughly 100 residues that is outside the homologous C–terminal ATPase domain. It is this extended sequence that distinguishes type Ia motor proteins from type Ib (Figure 3_1A). This N-terminus contains a predicted helix-turn-helix (HTH) motif that is believed to impart ParA with DNA binding capabilities critical to its functional role in the auto-repression of its own operon (Davey and Funnell, 1994; Radnedge et al., 1998).

Biochemical studies on the type Ia P1 ParA protein have characterized its activity within the cell as having two functions – partitioning of the P1 plasmid and the auto-repression of its operon (Bouet and Funnell, 1999; Bouet et al., 2000). ParA's participation in each of these activities is mediated by its nucleotide-bound state (Figure 3_1B). When ParA is bound to ATP it is recruited to the ParB-*parS* "partition complex" where it functions in an unknown manner to segregate the P1 plasmid to a specific location that will become the mid-point of the daughter cell (Bouet et al., 2000). On the other hand, when ParA is bound to ADP it is able to bind to a DNA site in the promoter region of its operon repressing any further transcription.

ParA is a deviant Walker A ATPase containing the A box sequence motif KGGXXK[S/T] that has a second functionally important lysine residue. Crystal structures of the deviant Walker A protein Soj, a chromosomal type Ib-like *par* motor protein, provided the structural evidence that it undergoes a monomer to dimer transition in the ATP-bound state by forming a nucleotide sandwich fold (Figure 3_1C)(Leonard et al., 2005). It is believed that the nucleotide-binding event acts as a "molecular switch" leading to the dimerization of the protein and to the alteration of its activity. It is believed that P1 ParA might possess a similar switch mechanism.

We have carried out structural studies on the type Ia ParA proteins (P1 and P7 ParA) in order to understand how ParA's activity is mediated through the binding of adenosine nucleotides and to gain a detailed account of the "molecular switch" mechanism operating within ParA.

3.2.1 Expression, Purification and Crystallization Studies of P1 ParA.

Initial crystallographic trials using "wild type" P1 ParA were unsuccessful. These "wild type" crystals were hexagonal (P6 with cell edges a=b=100 Å c=400 Å) and diffracted poorly (4-6 Å resolution). Attempts to improve the diffraction quality by refining the crystallization conditions (>0.5 mm in shortest dimension), annealing or dehydration were also unsuccessful. The lack of success with the "wild type" crystals lead to the

inclusion of other strategies notably the use of a ParA-ParB fusion construct in my crystallization experiments. A ParA-ParB fusion construct incorporating the N-terminal 28 amino acids of ParB onto the C-terminus of ParA (through a linker – GGSGGGSG) was originally designed to examine the stimulating effects of the ParB N-terminus on ParA ATPase activity as well as to serve as a platform for structural studies of the ParA-B interaction using a variety of spectroscopic techniques. Being composed of both ParA and an N-terminal portion of ParB, this fusion construct is designated ParA_fusNt28ParB, which is the nomenclature used in the remainder of the text. The ParA-ParB interaction and the nature of ParB's stimulatory effect will be covered in more detail in Chapter 5. Briefly, the basis for the fusion construct came from a genetic study of the P1 and P7 operons that identified and localized the source of the ParA-ParB interaction to the first 28 amino acids of ParB (Radnedge et al., 1998). The inclusion of the ParB peptide was also thought to provide three useful benefits; one, it might help stabilize the nucleotide-bound form; two, reveal critical information about the nature of the ParA-ParB interaction; and three, alter the shape of the ParA molecule in a way that might change the potential crystal contacts and give rise to additional crystal forms.

Assembly of the fusion construct was achieved through the use of 40 base pair (bp) oligonucleotides engineered with alternating 20 bp overlaps that made the extension of the ParB fusion piece possible with a simple PCR protocol (Stemmer et al., 1995). The amplified fusion piece was then incorporated into the reading frame of the C-terminus of ParA (inserted in a pET15b plasmid) using existing restriction sites resulting in the ParA_fusNt28ParB fusion construct (Figure 3_2A and Chapter 5 methods).

Additionally, a K122Q mutation was introduced into the ATPase domain corresponding to the signature lysine in the deviant A box KGGXX**K**[S/T] motif. This mutation had previously been found to abrogate ATP hydrolysis while maintaining ParA's ability to bind nucleotide (Fung et al., 2001). The fusion constructs were expressed in BL21 (DE3) *E. coli* cells and purified in a single step using a Ni^{2+} affinity column (Figure 3_2B). The use of an absorbance assay sensitive to the generation of inorganic phosphate (P_i) demonstrated that the fusion construct was able to hydrolyze ATP where that activity was abolished by the K122Q mutation (Figure 3_2C). Pure fractions were pooled and concentrated to 100 μM (4.5 mg/mL) and crystallization experiments were performed using hanging drop vapor diffusion.

Both ParA_fusNt28ParB and ParA(K122Q)_fusNt28ParB fusion constructs were screened in conjunction with ATP and non-hydrolyzeable ATP analogs (AMP-PCP, AMP-PNP, ATPγS, ADP-Vanadate, ADP-Al-F and ADP-Be-F). Diffracting crystals were achieved using 2 M sodium formate, 0.1 M sodium acetate (pH 4.6) as a crystallization solution. The best quality crystals were grown in the presence of AMP-PCP (0.1-3 mM) and $MgCl_2$ (2-5 mM). As the best chance at getting a well-diffracting crystal in the activated state was to maintain the integrity of the nucleotide in the binding pocket, using AMP-PCP/Mg^{2+} in addition to the K122Q mutant was thought to provide twice the protection against any hydrolysis. Crystals were cryo-preserved in a solution of the crystallization reagent doped with 20% glycerol. Native data were collected at ALS beamline 8.2.1 and processed with MOSFLM (Leslie, 1992).

3.2.2 Data Collection, Processing and Building of the ParA Structures.

The P1 ParA_fusNt28ParB crystal form - space group I422 (cell parameters; a=b=144 Å c=127 Å) - was distinct from that of the "wild type" ParA crystal and had a resolution limit beyond 3.0 Å (Figure 3_3A). A diffracting native crystal only provides data in the form of intensities that are only a part of what is required to solve a structure. Absent is the phase information, which accounts for the other critical component of a data set necessary for solving a structure. The effort of finding the phases is what is termed overcoming the "phase problem." As P1 ParA was the first type Ia *par* protein structure there were no homologous structures in the Protein Data Bank (PDB) that could serve as a search model to determine phase information using molecular replacement (MR). Two common strategies were pursued in parallel to determine the phases *de novo* in order to overcome the "phase problem" – Multiple Isomorphous Replacement (MIR)(utilizing heavy atom soaks) and Multiple Anomalous Diffraction (MAD)(incorporating the heavy atom into the polypeptide chain). Ultimately, the phase problem yielded to a MAD strategy. Using this strategy it is first necessary to express and purify seleno-methionine substituted protein (ParA contains 11 methionines) in order to use the anomalous diffraction to obtain the phasing information necessary to calculate an electron density map (Hendrickson, 1991; Hendrickson et al., 1990). The crystals of seleno-methionine substituted protein diffracted to a resolution limit of 3.5 Å. It can be a common occurrence to lose resolution in heavy atom substituted crystallization experiments, but phases can often be determined with lower resolution data. I collected several data sets

with these crystals at ALS beamlines 8.2.1 and 8.3.1. With the assistance of Dr. Schumacher data sets were collected using an inverse beam strategy at three wavelengths (remote, peak and inflection). The processed data was to 3.5 Å resolution having an I/σI (~ signal to noise) of 7 (2.0 in the highest resolution shell) and an R_{sym} of 10%. Phase information generated from 10 selenium sites identified using SOLVE (Terwilliger and Berendzen, 1999). The program "O" was used to construct a model of P1 ParA (Jones, 1991). As the resolution of these crystals were low, the model I had built contained sections that were disordered or out of register, preventing the refinement of my 3.5 Å resolution model (Figure 3_3B&C).

Several features were evident from the structure that was built; one, ParA was in a dimeric state that had an inter-dimer domain swap involving a portion of the C-terminus; two, the ParB peptide fragment was likely disordered and missing from the electron density; and three, no AMP-PCP could be found in the electron density. The resulting apo dimer was surprising as the model for deviant Walker A proteins had previously been to dimerize in the presence of ATP. The apo dimer in the crystal structure contains the notable characteristic of having a domain swap at the dimer interface through an exchange of the C-terminus (~ 80 residues). This swap of analogous structural components results in a symmetric dimer (Figure 3_3C&D). A search for the occurrence of domain swaps in crystal structures reveals many examples (30+) including a classic example: bovine RNase A in 1962 (Crestfield et al., 1962). There appears to be a strong correlation between their formation and low pH environments dating back to the RNase structure where the sample was prepared in 50% acetic acid (this historical observation

does not exclude other chemical environmental conditions from also promoting the swapping of domains). As the crystallization conditions of P1 ParA were at pH 4.6, below what is normally considered within its physiological range, it is likely the large shift in environmental conditions promoted the interactions amenable to domain swap contacts. Interest in the formation of amyloid plaques has lead to more intensive examination of domain swapping as it is thought to be one of the mechanisms by which these structures are stabilized. These examinations have found a link between domain swapping and their proximity to unstructured hinge or loop regions (flexible regions) (Bennett et al., 1994; Liu and Eisenberg, 2002; Rousseau et al., 2003). The swapped dimer has a significant amount of buried accessible surface area in its interface and is likely not a physiological dimer, but it does not exclude ParA from normally being an apo dimer in solution.

Dr. Schumacher finished the apo P1 ParA structure by refining her own model using CNS (Brünger, 1998). The P1 apoParA structure was solved to 2.92 Å resolution with R_{work}/R_{free} values of 24.6%/28.1% respectively (other statistics of the P1 apoParA structure are listed in Table 3_1). Dr. Schumacher went on to solve two structures of P7 ParA in different crystal forms ($C222_1$ and $P6_222$). Similar to the case of P1 ParA both crystal forms resulted in apo dimer structures despite being set up in the presence of ATP (or ATP analogs). Unlike P1 ParA the crystals were grown from conditions of higher pH (~ 7) and the resulting structures did not include the domain swap. Despite containing the domain swap, the P1 ParA structure is homologous to both the P7 structures (RMSD

of <1.0 Å for all corresponding C_α atoms). This points to the pH of the set-up conditions as likely being responsible for the domain swap observed in the apo P1 structure.

3.2.3 The apo P1 ParA Structure.

The completed apo P1 ParA and two apo P7 ParA structures solved by Dr. Schumacher are presented here in order to introduce further work on the physiological properties and biological role of ParA in the cell. The apo P1 and P7 ParA structures are the first examples of type Ia *par* proteins and are valuable as they reveal the structural nature of what distinguishes the type Ia from the type Ib families. The P1 and P7 ParA proteins are multi-functional, so it is not surprising that they notably contain three main domains: first, an extended alpha helix that bridges the two dimer partners (amino acids 1-43); second, a winged helix-turn-helix motif (amino acids 44-104); and third, a C-terminal deviant Walker A ATPase domain (amino acids 105-399). It is the N-terminus that distinguishes the type Ia *par* proteins from the type Ib family (Figure 3_4A). It is also the N-terminus that is structurally unique and leads the overall ParA structures to be considered a novel fold. The extreme N-terminus forms an extended alpha helix that bookends the dimer and makes contacts with the other dimer partner. Also contained within the N-terminus is a winged-HTH motif. The HTH motif had been predicted from its primary sequence (Davey and Funnell, 1994). It is likely to be the main factor

contributing to ParA's ability to bind DNA allowing it to carry out its role in repression. The final motif in the ParA structure is the C-terminus that contains the deviant Walker A ATPase domain. The deviant Walker A fold contained within the C-terminal three quarters of the protein is the basis of strong homology to other Walker A proteins including the *Bacillius subtilis* Soj protein, which is a chromosomal type Ib ParA homolog, MinD, and NifH (Cordell and Lowe, 2001; Georgiadis et al., 1992; Leonard et al., 2005).

The structures of P1 and P7 ParA in the apo state show that these proteins are structurally homologous, but are in distinct apo dimer conformations owing to the flexibility provided by the N-terminal contacts. Excluding the swapped P1 ParA structure each of the P7 ParA dimers have approximately 5000 $Å^2$ of buried accessible surface area (buried-ASA). There is great variability in the degrees of openness and in the interfaces among the different dimers. The "nucleotide-sandwich" dimer in the Soj protein (bound to ATP) has a buried-ASA of 2855 $Å^2$/dimer by comparison. In the case of the Soj protein it is the nucleotide that helps establish the dimer interface making the apo ParA dimers even more interesting as they achieve these large buried-ASAs without the need for this additional mechanism.

What is the common feature in the apo ParA structures that could account for the stability of the dimeric state? Although diverse, there is one interaction that is common to all apo ParA dimers, which is the interface created by the insertion of α1 into a surface accessible hydrophobic crevice between α14′ and α15′ of the other dimer partner (Figure

3_6B). What is even more remarkable is that this interaction accounts for 2800 $Å^2$ of buried-ASA (~2/3 the total for the non-swapped dimers), which is sufficient to account for the dimerization by itself.

3.2.4 Apo ParA Dimerizes at Physiologically Relevant Concentrations.

The prevailing model for deviant Walker A ATPases has been that nucleotide binding marks the required committing step for dimerization. In our crystallographic studies of ParA (P1 & P7) in the apo state we observed several conformational states that had the common remarkable feature of being in the dimeric state. This is interesting as the unique feature of the type Ia *par* ATPases is an additional N-terminus that appears to provide a key interaction in the formation of the apo dimer state. The N-terminus forms an extended alpha-helix in both P1 and P7 ParA structures that serves as the primary source of contact between dimer partners.

Before addressing any potential contributions to the stability of the structural dimer it is first important to verify that ParA exists as a dimer in its native solution state. While the crystal structures of ParA were obtained at concentrations of 70-100 μM (~ 4.5 mg/mL), the physiological concentration of ParA has been measured to be ~ 3 μM (Bouet and Funnell, 1999). To assess whether the observed structural dimer was also the native

solution state we undertook a variety of studies to ascertain the solution state of ParA. These studies include gel filtration chromatography, size exclusion chromatography resolved static light scattering (SEC-LS), and simple cross-linking studies using glutaraldehyde. SEC-LS is a technique that uses size-exclusion chromatography in-line with a multi-angle light scattering and refractive index instruments. The sizing column serves to resolve the sample into roughly homogeneous fractions before it enters the light scattering chamber. Laser light (mono-chromatic λ_{658}) passes through the sample fractions as they move through the instrument and the resulting scatter is detected at multiple angles. The instrument using the scatter intensity (at each angle) per unit time determines the molecular mass, M_r (g/mol), according to Brownian motion (see methods).

SEC-LS experiments with native ParA proteins in the apo state show that both P1 and P7 forms have an apparent molecular mass of roughly 90 kDa in solution (Figure 3_5B: 89 kDa for P1 and Figure 3_6A: 87 kDa for P7). This molecular mass is roughly twice the calculated weight of 45 kDa for ParA and corresponds to a dimeric state. Additionally, running samples over a gel-filtration column (Sephacryl 200) calibrated with molecular weight standards resulted in similar findings of an apparent molecular weight of $\sim$ 98 kDa for "wt" ParA. The dimer interaction also appears to be independent of concentration as sample concentration was titrated down to 5 μM (roughly the detectable limit for ParA using SEC-MALS)(Figure 3_5B). As the believed physiological concentration is just under that which is detectable using biophysical techniques, we undertook cross-linking studies using glutaraldehyde to confirm ParA's state below its

physiological concentration (500 nM and 1 μM). Glutaraldehyde, a five-carbon linear molecule, is sufficiently long to complete inter-molecular stable links through accessible amines. A mixture of ParA with 0.2% glutaraldehyde at 30 ºC for 15-30 min. showed that ParA forms a dimer at concentrations of 500 nM and 1 μM (Figure 3_5C). These studies confirmed that ParA is primarily a dimer at and below its physiological concentration. Collectively the data show that the preferred state of ParA in solution is in its dimeric form.

It is also probable that ParA spends most of its time in the dimeric state (apo, ATP or ADP-bound) *in vivo* impacting its mode of activation and biological function. Deviant Walker A proteins have been hypothesized to form a dimer in the ATP-bound state, which marks their transition to an active state. Structures of the type Ib chromosomal motor protein Soj were solved in the apo (monomer), ADP-bound (monomer) and ATP-bound (dimer) states illustrating this transition (Leonard et al., 2005). Recently, another type Ib motor protein δ was solved in its dimeric ATPγS–bound state, however accompanying cross-linking studies found that it was also a dimer in its apo form (Pratto et al., 2008). These results open the possibility that other mechanisms besides ParA's oligomeric state may be important for its transition among its different functional states.

3.2.5 The Role of α1 in ParA Dimerization.

Apo-ParA is a dimer in solution within its physiological range (~ 2 μM) as determined by light scattering and cross-linking experiments. This finding was unexpected as the model for all Walker A ATPases has been that nucleotide binding enabled dimerization through a "nucleotide sandwich" fold (Lutkenhaus and Sundaramoorthy, 2003). ParA, a type Ia *par* protein, is unique among Walker A ATPases in that it contains an extended N-terminus beyond the consensus ATPase domain. Published genetic studies pointed to the importance of the first 20 residues in ParA function and an examination of all the P1 and P7 ParA structures reveal that the N-terminal α-helix makes substantial contacts with its dimer partner using these residues (Radnedge et al., 1998). The strong structural homology of the interaction between the α1 N-terminus and the α14´-α15´ hydrophobic crevice found in all the apo structures along with its large buried accessible surface area points to the potential role of this interaction in stabilizing the apo dimer state (Figure 3_6B).

To ascertain whether these interactions observed in the structures of the apo dimer state are also important in solution, we mutated strategic residues in the dimer interface and assayed their effects using light scattering (SEC-LS). The mutations targeted residues within the α1-C-domain interface or within the C-domain-C-domain interface (observed in the domain swapped dimer). To assess the role of the α1 N-terminus in stabilizing the apo dimer, we made an N-terminal deletion mutant (Δ1-20). Using light scattering it was

found that the deletion mutant of ParA is a monomer in solution (Figure 3_6A). Next, to target the inter-dimer contacts of the N-terminus more specifically, we made a series of point mutants along the inside of the N-terminal helix (A11S, R13Q, A14S). Two of these, A11S and A14S, were found to be primarily monomeric in solution (Figure 3_6C). These studies highlight the significant role the N-terminal α-helix plays in the dimerization of ParA. Analysis of these mutants using SEC-LS showed that they have apparent molecular masses corresponding to a monomeric form in solution (Figure 3_6A: 49kDa for the Δ1-20 mutant, 63kDa for the A11S mutant and 51kDa for the A14S mutant). The apparent molecular weights determined by SEC-LS are dependent upon the column for resolution of the sample and the creation of a homogeneous data sample. Any resulting poly-dispersity is preferentially weighted for the species with higher mass and has a direct impact upon the resulting calculated molecular mass. This is the case in the A11S mutant which has a higher contribution from a fraction of the dimeric form resulting in a ~ 10 kDa shift in the apparent molecular mass from what would be predicted. Point mutations to the "domain swapped" dimer interface (A277Y and T278W) designed to introduce "steric bulk" in the same fusion construct had no affect on the dimeric state of ParA (each had an apparent molecular mass of 100 kDa, or twice the calculated molecular weight)(Figure 3_6C). These results illustrate the critical importance of the contacts made by the α1 helix in stabilizing the apo dimer state. The a1 helix contacts were present in all conformations observed and act to maintain the structural integrity of the dimer, while providing for the rotational flexibility to function in its various nucleotide-bound forms.

3.3 Discussion

Type Ia ParA contains an extra N-terminal ~100 residues compared with the smaller type Ib *par* proteins. The prevailing paradigm for most deviant Walker A ATPases has been that nucleotide binding marked the required step for dimerization resulting in a nucleotide sandwich fold. The monomer to dimer transition was believed to comprise part of the "switch mechanism" that coincided with a change in protein function. Previous structures of the type Ib-like Soj and type Ib δ Walker A *par* proteins in their ATP-bound (or ATPγS) forms resulted in nucleotide sandwich dimers. Structures of the Soj protein in both its apo and ADP-bound states are monomers consistent with previous deviant Walker A activity. Whereas, biochemical evidence in support of the δ structure found it to be a dimer in its apo state.

As there was no previous structural information of a type Ia *par* protein, solving the ParA_fusNt28ParB protein crystal structure required the *de novo* determination of phase information using a MAD strategy. The resulting P1 ParA structure solved by Dr. Schumacher was found to be a domain swapped apo dimer. Two subsequent structures of P7 ParA solved by Dr. Schumacher were also apo dimers in different conformations. The structures revealed that the unique N-terminal ~100 residues contained a winged-HTH motif as well as an extended α-helix (α1). The N-terminal α-helix (α1) makes substantial contacts with a crevice between α14´ and α15´ in the dimer partner. These

contacts appear to be the only consistent dimer interface in what appears to be a conformationally flexible apo state.

In supporting static light experiments ParA was determined to be an apo dimer in solution down to a concentration of 5 μM (the limit of detection for the instrument). Additional cross-linking experiments using glutaraldehyde determined that ParA was an apo dimer at a concentration of 500 nM well below its reported physiological concentration. It appeared that the dimer was stabilized in its nucleotide bound form and that the apo dimer transitioned to higher oligomers at higher concentrations.

A series of mutations were made to two different dimer interfaces to determine whether the observed structural dimer was the same in solution. The mutations Δ1-20, A11S and A14S made to disrupt the α1 - C-domain (α14´ - α15´) interface all successfully resolved ParA to its monomeric form. Whereas the mutations A277Y and T278W made to the interface of the swapped dimer did not alter the apo dimer state. These results are consistent with the observations in structure and point to the α1 - C-domain as being the only substantial interface in the conformationally flexible dimer state.

In summary, ParA appears to be a dimer in solution in its apo state (Figure 3_7). With its equilibrium point shifted in favor of the dimeric state, questions addressing the constitution of the underlying molecular switch mechanism must move beyond a monomer-dimer transition. As a result of these studies on the structure and oligomeric state of apo P1 ParA, portions of Chapter 3 (Figures: 3_4A&B; 3_5B&C; 3_6A&B) were

included in a recent publication entitled “Structural basis for ADP-mediated transcriptional regulation by P1 and P7 ParA” (Dunham et al., 2009).

3.4 Methods

P1 ParA constructs. The P1 ParA gene was cloned out of pBEF198 (a gift from Dr. Funnell) using oligonucleotides (5′-GGAATTCCATATGAGTGATTCCAGCCAGCTTCAC-3′ (NdeI site) and 5′-CGGGATCCTCAGTTAGATCTGATAAATTCAATACG-3′ (BamHI site)) that introduced new restriction sites (NdeI and BamHI sites) facilitating its insertion into a pET15b vector (Novagen). A ParA-ParB fusion construct incorporating the N-terminal 28 residues of ParB onto the C-terminus of ParA (through a linker – GGSGGGSG) was designed to examine the stimulating effects of the ParB N-terminus on ParA ATPase activity as well as to serve as a platform for structural studies of the ParA-ParB interaction. Assembly of the fusion construct was achieved through the use of 40 bp oligonucleotides engineered with alternating 20 bp overlaps that made the extension of the ParB fusion piece possible with a simple PCR. The amplified fusion piece was then incorporated into the reading frame of the C-terminus of ParA using existing restriction sites creating the ParA_fusNt28ParB construct (see Chapter 5 methods for details).

Expression, purification and function. N-terminal His tagged ParA and ParA-ParB fusion proteins were expressed in BL21 (DE3) *E. coli* cells. The cells were induced at OD_{600}=0.5-0.7 with 0.5 mM IPTG and grown at 37 ^{o}C while shaking at 250 rpm for another hour. Cells were lysed in 25 mM Tris (pH 7.2-7.5), 300 mM NaCl, 10 mM imidazole, 1x Complete, 3μg/mL DNase I and RNase A, and 0.5 mM TCEP using a microfluidizer. Lysate was clarified by centrifugation at 30,000 xg for 45 minutes. Lysate was loaded onto a HisTrap column (FF or Super Flow)(4^{o} C) and washed with 25

mM Tris (pH 7.2), 250 mM NaCl, 15 mM imidazole & 0.2 mM TCEP. ParA proteins were eluted using an imidazole gradient (15 mM-200 mM) in 25 mM Tris (pH 7.2), 250 mM NaCl & 0.4 mM TCEP.

The ParA_fusNt28ParB construct was assessed for function by judging its ability to hydrolyze ATP as judged by a coupled reaction assay capable of detecting the generation of P_i (described in detail in Chapter 5 methods).

Mutagenesis. The following mutants were constructed using a QuikChange (Stratagene) strategy unless otherwise specified. To verify the specific nature of ParA's ATPase domain function in the ParA_fusNt28ParB construct the signature lysine residue was mutated to a glutamine using the following oligonucleotide : 5′-GCGGTGTGTCA CAGACTGTATCGACG-3′ (K122Q). In order to probe the nature of ParA's α1 helix in stabilizing its apo dimer state, the interacting domains were altered either through deletion (re-cloned using flanking oligonucleotides: 5′-GGAATTCCATATGCT GACTGAACAAGTACAGTTGC-3′ (Δ1-20)) or by mutating important alanine residues (Ala11 and Ala14) proposed to be important in the α1-α14′/15′ interaction (Ala to Ser mutations: 5′-CACAAGGTTTCTCAAAGAGC-3′ (A11S) and 5′-GCTCAAAGATCCAACAGAATGC-3′ (A14S)). To judge whether contacts observed in the domain swapped apo dimer structure were important to the dimer state two residues at the interface (Ala277 and Thr278) were mutated to introduce bulk using the following oligonucleotides: 5′-CCTCTGCCGCCATATA CTGTCGATTTCC-3′ (A277Y) and 5′-CTGCCGCCAGCATGGGTCGATTTCCAC-3′ (T278W). The resulting mutant constructs were sequenced to confirm the proper configuration using the

Sanger method (Applied Biosystems instrument) at the Vollum Institute or MD Anderson core facilities.

Static Light Scattering. Static Light Scattering (SLS) experiments were undertaken to assess the solution states of native *par* proteins and probe the stability of the apo dimer state as judged through a relative molecular mass, M_r (g/mol). Samples were analyzed using a Wyatt technologies SEC-LS instrument (consisting of a Shodex size exclusion column coupled in-line to a multi-angled light scattering detector, Dawn-Heleos, and a refractive index detector, OptilabrEX.) Samples (5-50 μM) were injected onto the column equilibrated with buffer: 25 mM Tris-Cl (pH 7.5), 250 mM NaCl and 2 mM DTT) at 1 mL/min. and data was collected at 0.5 s intervals. In-line size exclusion column, multi-angle light scattering (MALS) instrument and refractive index (RI) instruments work in a coordinated fashion to facilitate the acquisition of both the intensity of scattered light at different angles and the concentration of the sample (refractive index) that allow for the calculation of the molecules molar mass. Astra software from Wyatt technologies uses the relationship below to relate the changes in the angular dependent scatter intensity (expressed as a Rayleigh ratio) over time of these measurements to the molecules molar mass.

$$I_{\text{scattered}}(\theta) \propto R(\theta) = K^* McP(\theta)[1 - 2A_2 McP(\theta)]$$

In the relationship R is the Rayleigh ratio; K is $4\pi^2 n_o^2 (dn/dc)^2 / N_A \lambda^4$ (n_o is the refractive index of the solvent); Mc is the molar mass; P is the scatter function that relates angular variation in scatter intensity to mean square radius r_g; and A_2 is the second virial

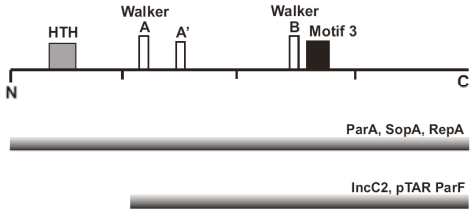
constant. Results were evaluated using a Debye plot ($R(\theta)/K^*$ vs. $\sin^2(\theta/2)$) where least squares analysis fits a polynomial in $\sin^2(\theta/2)$ to the data resulting in Mc (from the intercept) and r_g (from the slope).

ParA cross-linking. ParA samples were prepared in concentrations ranging from 500 nM to 10 μM in 30 mM HEPES (pH 7.8), 200 mM NaCl and 0.5 mM TCEP. Cross-linking was initiated with the addition of 0.2% glutaraldehyde and allowed to proceed for 15 or 30 min. The reaction was quenched with the addition of 0.25 M Tris (pH 8.2) followed by a 20 min. incubation period at 30 °C. Samples were then concentrated before being resolved by SDS-PAGE (protein bands illuminated by Coomassie staining). The resulting gels were documented using an Alpha Innotech FluorChem system and saved in digital format with AlphaEase FC software.

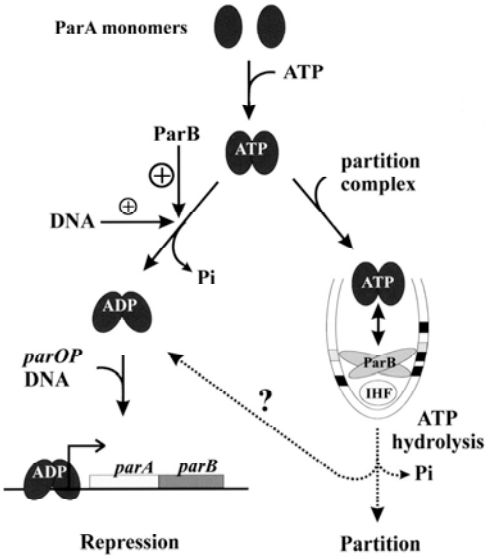
3.5 Acknowledgements

I would like to recognize Dr. Maria Schumacher for her help in the data collection and processing of data for the P1 “apo” ParA_fusNt28ParB crystals. I would also like to recognize her work in solving the structures and analysis of the apo P1 ParA_fusNt28ParB and P7 ParA proteins.

A



B



C

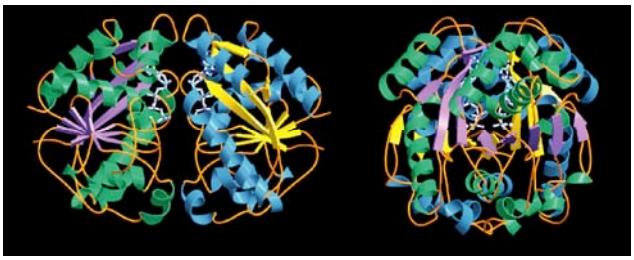


Figure 3_1: Characterization of the P1 ParA Protein. (A) Schematic of the general alignment of type Ia and Ib *par* proteins (Barbara Funnell, personal communication). (B) Schematic of P1 ParA's characterized activity (Bouet and Funnell, 1999). (C) Structure of Soj (type Ib *par* protein) emphasizing the essential role of the nucleotide in dimerization ("nucleotide sandwich")(Leonard et al., 2004).

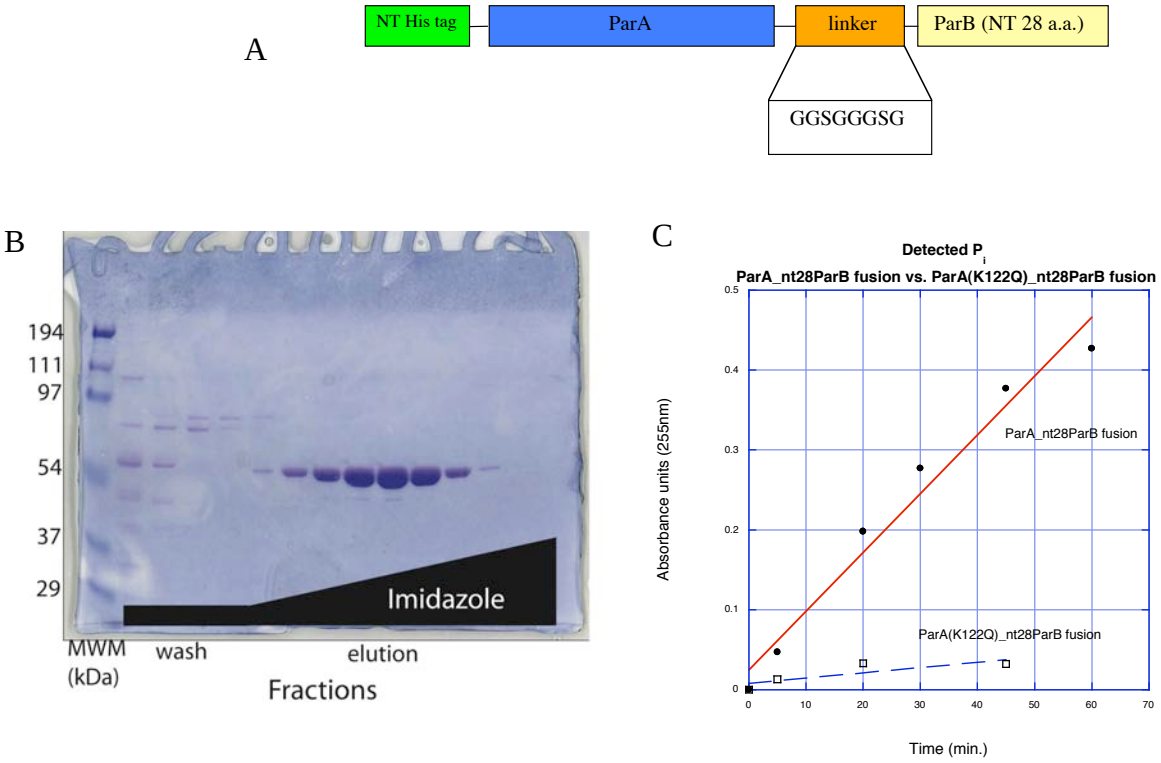
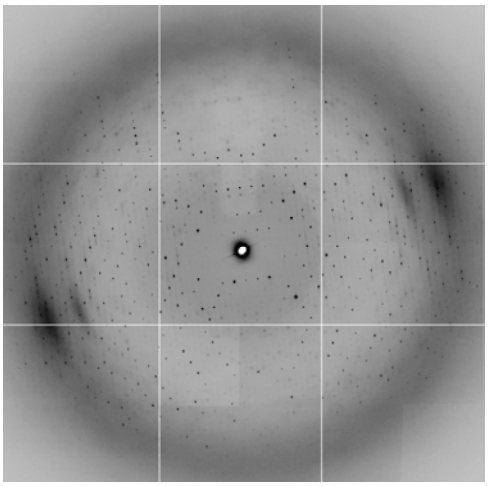


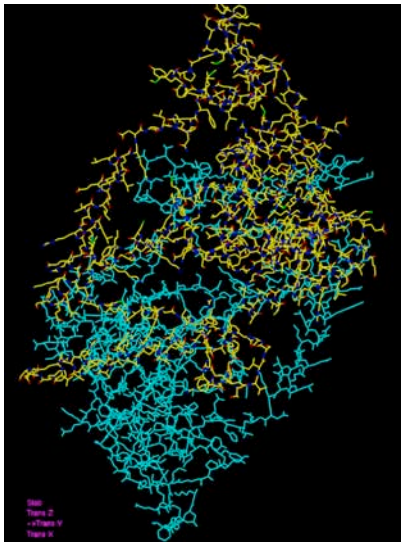
Figure 3_2: P1 ParA-ParB Fusion Construct - ParA_fusNt28ParB. (A) Schematic of the P1 ParA_fusNt28ParB construct incorporating the N-terminal 28 residues of ParB onto the C-terminus of ParA. (B) Representative purification result of P1 ParA_fusNt28ParB (SDS-PAGE-Coomassie stained). (C) ATPase activity of the purified P1 ParA_fusNt28ParB construct compared to the K122Q mutant.

A



B

C



D

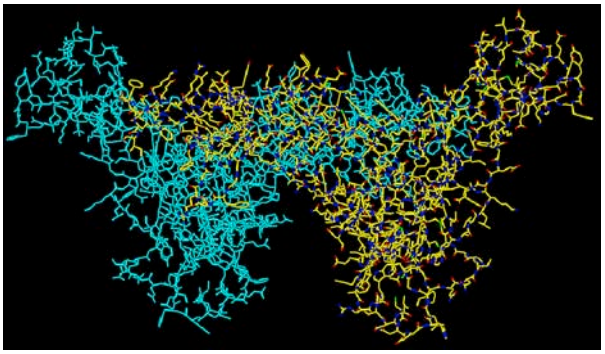


Figure 3_3: Building Process of the Apo P1 ParA Structure.

(A) X-ray diffraction (image) in P1 ParA(K122Q)_fusNt28ParB (I422) data set. (B) Early-stage model of "domain swapped" apo ParA with secondary structural elements. (C) Early-stage model "domain swapped" apo ParA dimer illustrating contributions to the swapped interface. (monomers colored in cyan and gold) (D) Same as in (C) with two 90^{o} rotations to illustrate a different perspective.

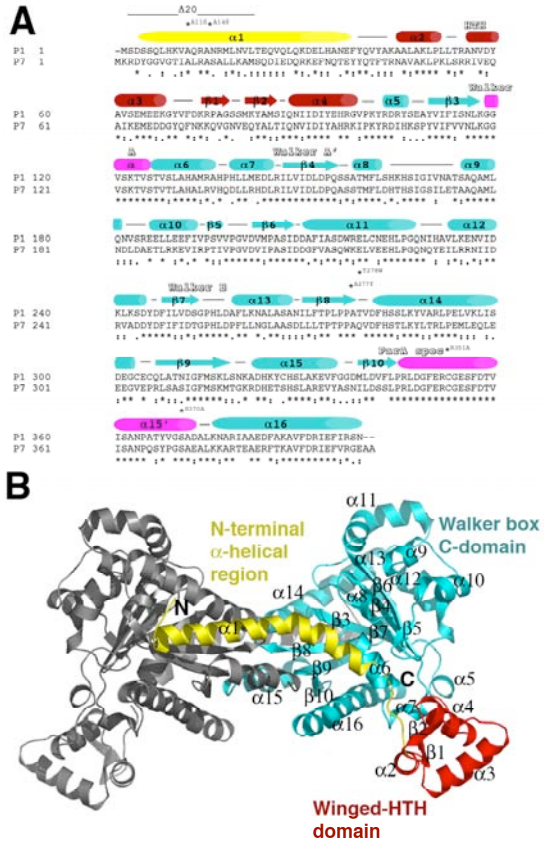
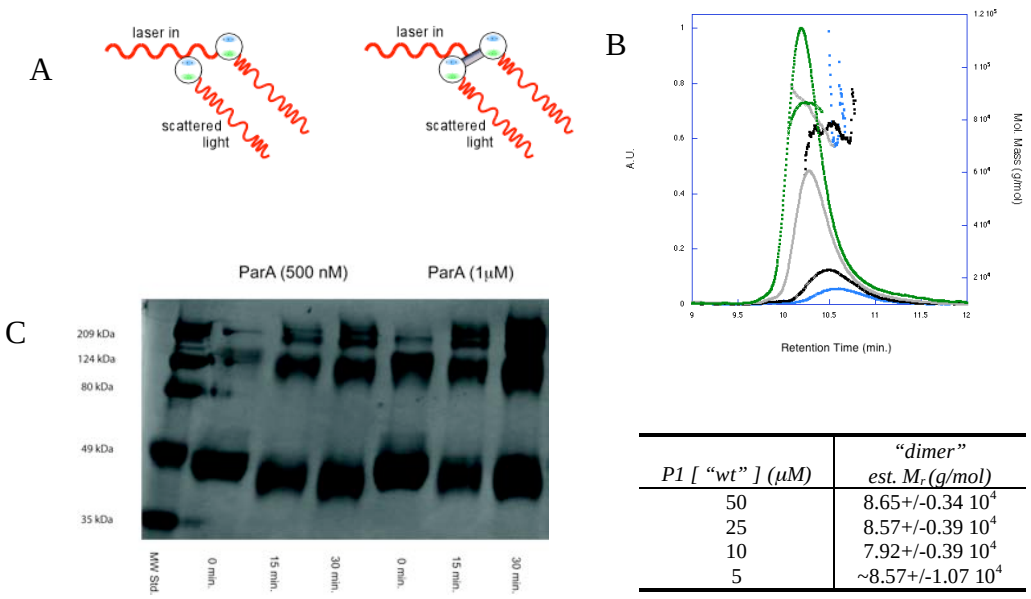
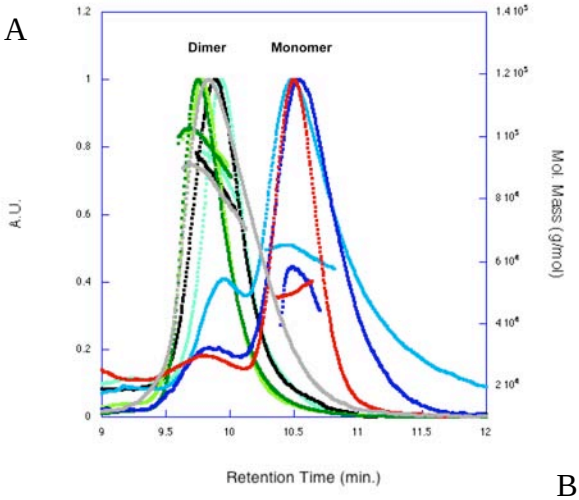


Figure 3_4: P1 ParA apo Structure. (A) Alignment of P1 and P7 ParA proteins. (B) Structure of the apo P1 ParA dimer with main domains highlighted: (1) alpha helix 1 (yellow); (2) winged – HTH (red); and (3) ATPase C-domain (teal). This figure comprised part of a recent publication on P1 and P7 ParA (Dunham et al., 2009).

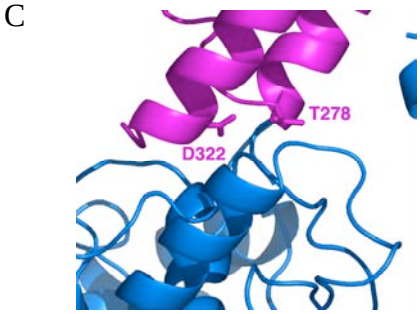


P1 ["wt"] (μM)	*"dimer" est. M_r (g/mol)*
50	8.65+/-0.34 10^4
25	8.57+/-0.39 10^4
10	7.92+/-0.39 10^4
5	~8.57+/-1.07 10^4

Figure 3_5: Characterization of the Solution State of P1 ParA. Characterization of the solution state using SEC-LS and glutaraldehyde cross-linking. For clarity the portion of the chromatogram used for the analysis for the monomer and dimer states has been highlighted. Included in the graph are the Rayleigh Ratio and molar mass for each sample. The estimates for the molecular mass of each of the ParA proteins calculated using ASTRA software are depicted graphically on the second y-axis and in the adjacent Tables. (B) Dimerization state is independent of concentration (to the lower limit of detection). P1 "wt" at 50 μM (Green), P1 "wt" at 25μM (Grey), P1 "wt" at 10 μM (Black), and P1 "wt" at 5 μM (Blue). (C) Cross-linking experiments were conducted to confirm that the solution state of apo ParA was still dimeric at low and sub-micro molar concentrations. Panels B and C were included in a recent publication on ParA (Dunham et al., 2009).



ParA construct	*"dimer" est. M_r (g/mol)*	*"monomer" est. M_r (g/mol)*
P1 "wt"	8.86+/-0.02 10^4	
P7 "wt"	8.73+/-0.07 10^4	
P1-fus28 ParB	9.16+/-0.82 10^4	
P1(A277Y)-fus28 ParB	1.00+/-0.01 10^5	
P1(T278W)-fus28 ParB	1.00+/-0.01 10^5	
P1(Δnt20)-fus28 ParB		4.99+/-0.03 10^4
P1(A11S)-fus28 ParB		6.27+/-0.05 10^4
P1(A14S)-fus28 ParB		5.14+/-0.11 10^4



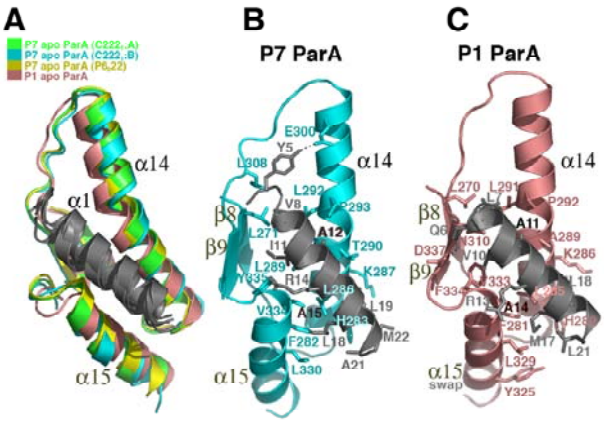


Figure 3_6: Probing the Stability of ParA's Dimeric State. For clarity the portion of the chromatogram used for the analysis for the monomer and dimer states has been highlighted. Included in the graph are the Rayleigh Ratio and molar mass for each sample. The estimates for the molecular mass of each of the ParA proteins calculated using ASTRA software are depicted graphically on the second y-axis and in the adjacent Tables. (A) Samples fall into two groups characterized by their molecular masses; a dimeric state: P1 ParA"wt"(Jade), P7 ParA"wt" (**Black**), P1 ParAB fusion (**Grey**), P1 ParA(T278W)B fusion (**Green**), ParA(A277Y)B fusion (Lt. Green); or monomeric state: ParA (Δnt20)B fusion (Red), ParA(A14S)B fusion (**Blue**), and ParA(A11S)B fusion (Aqua). Tabulated results from SEC-LS experiments (right). (B) Alignment of N-terminal α1 to C-terminal α14-α15 (dimer partner) contacts. (C) Dimer interface contacts from apo P1 ParA ("domain swapped") structure. Panels A and B of this figure comprised part of the work included in a recent publication (Dunham et al., 2009).

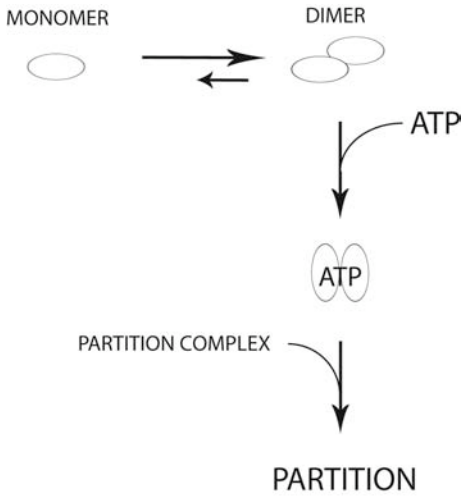


Figure 3_7: New Model of ParA's Functional Oligomeric State. P1 ParA is now believed to be an apo dimer in solution being stabilized by a pair of interface contacts established by the N-terminal α1 helix. These contacts involving a structural element that is unique to the type Ia *par* proteins may account for its altered behavior compared to other deviant Walker A proteins.

Crystal	P1 apoParA
Space group	$I4_122$
Cell dimensions	a=b=145.3 Å, c=126.2 Å
#molecules/ASU	1
Resolution (Å)	72.6-2.92
Overall R_{sym} (%)	5.5 (37.4)
Overall I/σ (I)	9.4 (2.0)
# total reflections	73,545
# unique reflections	14,700
Refinement statistics	
Resolution (Å)	72.6-2.92
R_{work}/R_{free} (%)	24.6/28.1
Rmsd	
Bond Angles (^{o})	1.59
Bond Lengths (Å)	0.009
B-values ($Å^2$)	2.8
Ramachandran values	
Most favored (%/#)	82.1/285
Add. allowed (%/#)	15.0/52
Gen. allowed (%/#)	2.9/10
Disallowed (%/#)	0/0

Table 3_1: Statistics of Dr. Schumacher's P1 apoParA crystal structure.

Chapter 4

The Roles of the Nucleotide-Bound States of P1 ParA

4.1 Introduction

We have learned from the apo P1 and P7 ParA structures and supporting biochemical studies that the type Ia *par* proteins appear to behave differently from their Walker A relatives. The additional N-terminus provides extra stability to the dimer state such that it has become the preferred form of the protein rather than the form of one particular functional state (ATP-bound form). From earlier biochemical studies it is clear that the nucleotide-bound state has a direct effect on P1 and P7 ParA's functional state from partition (ATP) to transcriptional regulation (ADP). The mechanisms by which this is achieved are unclear as there is an absence of structural information addressing this question to guide our understanding.

What is known from the extensive studies (genetic and biochemical) on the type Ia *par* motor proteins, ParA and SopA, is that the act of binding a specific type of nucleotide directs the activity of the protein. The familiar nucleotide switch concept can still serve as the basis of our conceptual understanding of what directs ParA's activity although the mechanistic connotations are likely more complicated than a simple monomer to dimer transition. Generally, it has been established that ParA in an ADP-bound state is reported to have a role in the repression of its own transcription by binding to a site within its own *parOP* (Davey and Funnell, 1994; Davey and Funnell, 1997). Juxtaposed to this, ParA in its ATP-bound state is able to recognize the "partition complex" (ParB – *parS*) and in an unknown manner drive or direct the segregation of the daughter plasmids. This binary

model is more complicated when viewed through the prism of ParA's interactions with ParB, its *parAB* partner, which clouds our understanding of ParA's activities. The results of these investigations have been synthesized into an activity model that serves as the basis for understanding ParA's role in the cell (Figure 4_1A).

The type Ia *par* systems share the characteristic of having their *parAB* operons regulated via auto-repression by their motor proteins rather than their centromere binding proteins like other *par* systems. P1 ParA is among the best characterized of these auto-repression systems. The *parOP* binding site has been mapped using a DNase I protection assay (Davey and Funnell, 1994; Davis et al., 1992). ParA binds to this site by nucleating around a ~ 42 base pair (bp) site composed of a ~ 21 bp imperfect inverted repeat (between the -10 and ribosome binding sites (RBS))(Figure 4_1B). ParA appears to load this site with multiple copies ultimately protecting a ~ 150 bp stretch that extends upstream of the -35 site to the *parA* start codon. As the nucleation site is pseudo-palindromic, it was hypothesized that ParA must bind to this site as a dimer. Additionally, it was demonstrated that binding of nucleotide increased ParA's affinity for the *parOP* site. ParA in the ATP-bound or ADP-bound state increased its affinity for the *parOP* site 10-fold or 20-fold, respectively (Davey and Funnell, 1994)(Figure 4_1B). Early glycerol gradient sedimentation experiments suggested that ADP binding stabilizes the dimeric state (Davey and Funnell, 1994).

Establishing the DNA binding characteristics used by P1 ParA to bind the ~ 42 bp pseudo-palindromic site is complicated by the fact that ParA appears to bind DNA in all

of its examined states (apo, ADP-bound and ATP-bound). A further complication arises from the fact that initial recognition is likely different from the higher nucleoprotein complex believed to be at the heart of ParA's repression of the *par* operon. The role the ATP-bound state plays in this auto-repression is also uncertain. Initial studies on P1 ParA found that binding to *parOP* occurred in the ATP-bound state (Davis et al., 1992). DNA binding was later expanded to include ParA's ADP-bound state (the "preferred state" for binding DNA), while binding to the nucleation site in the ATP-bound state remained evident. Later, ParA binding to *parOP* was shown to be stimulated by an interaction with ParB only in the ATP-bound state bringing its affinity more in line with the ADP-bound state (Bouet and Funnell, 1999). ParB was demonstrated to have no ability to bind to *parOP* on its own. As ParB was also shown to stimulate ParA's ATPase activity, it was initially hypothesized that it converted ParA to its ADP-bound state, the preferred DNA-bound state, but the measured activity could not account for the increase in binding. Studies in the F plasmid system showed that SopB (ParB) was necessary for full auto-repression also by enhancing SopA's (ParA) ability to bind to its *OP* (Ravin et al., 2003). It remains to be seen what the nucleoprotein super structure will look like and how, or even if, states (ParA(ATP-bound)-ParB) other than the ADP-bound state will contribute to the repression of the *par* operon.

While the ADP-bound state appears to be the preferred state for *parOP* binding and to be responsible for ParA's role in auto-repression, the ATP-bound state is intricately tied to ParA's role in partition. From the earliest genetic studies *parA* was determined to be essential to partitioning, but the nature of that role remained unclear (Hayes and Austin,

1993; Radnedge et al., 1998). Being a Walker A ATPase it is assumed the ATP-bound state of ParA forms a nucleotide sandwich fold and is in the dimeric state, like Soj (Leonard et al., 2005). Although this has not been demonstrated experimentally, ATP binding was shown to increase ParA's thermal stability suggesting that it changes its conformation to a more compact form. ParA was shown to be able to associate with the "partition complex" in its ATP-bound state. How it then acts to ensure that paired plasmids are resolved into separate daughter plasmids and properly segregated to each daughter cell is not known.

Our understanding of ParA's role in partitioning has been shaped by fluorescence microscopy experiments that provide temporal and positioning information on *par* motor proteins within the cell. More recently, the dissection of the type II ParM system has lead to a better understanding of the role *par* motor proteins (through the formation of filaments) can play partitioning (Garner et al., 2007; Salje and Lowe, 2008). Filaments have been observed for both type Ib and type Ia *par* proteins as well (type Ia (SopA) and Ib (δ, ParF, and pB171 ParA) (Barilla et al., 2005; Ebersbach et al., 2006; Ebersbach et al., 2005; Hatano et al., 2007)(Figure 4_1C). The role filaments might play in type I partitioning has yet to be determined, but it is likely that there will be some variety in the mechanisms used by both type Ia and Ib *par* systems.

In this chapter, I will cover our efforts at understanding the structural states of the two functionally distinct roles that ParA plays within the cell.

4.2.1 ParA's ADP-bound state: implications of Dr. Schumacher's ADP-bound structure.

I introduce Dr. Schumacher's structures of ADP-bound ParA in order to discuss the biochemical implications of the structural features important to its role in the auto-repression of the *par* operon. As the apo form of ParA already exists in a dimer state our goal was to discover the mechanism by which ParA is converted to an activated form (ADP-bound state) more suitable to binding DNA. The two P1 ParA ADP-bound structures resulted from two different crystal forms both of which grew out of conditions at pH 7.0. One crystal form ($P2_12_12_1$ -R_{work}/R_{free} of 19.0%/23.0% to 2.05 Å resolution) was grown in the presence of higher concentrations of ADP and contained a second ADP site believed to be non-physiological. The other crystal form ($P4_12_12$ - R_{work}/R_{free} of 18.5%/24.0% to 2.58 Å resolution) contained a single ADP binding site analogous to the central binding site in the other structure.

A comparison of the apo ParA and ADP-bound structures revealed that a large conformational change had occurred resulting in a structure that is more compactly organized around the bound nucleotide. The fact that the two ADP-bound crystal forms were nearly identical suggests that the nucleotide-bound form is a single conformational state. Previous studies on the thermal stability of P1 ParA found the binding of nucleotide (ADP and to a larger extent ATP) lead to the protection of its DNA binding function from exposure to heat (Davey and Funnell, 1997). Nucleotide binding also

resulted in an increase in ParA's denaturation temperature. These results lend support to the observed nucleotide dependent changes observed in the structure, namely that ADP binding results in a more rigid dimeric state eliminating the flexibility observed in the apo P1 and P7 structures. Our collaborator, Dr. Barbara Funnell, performed limited trypsin digest experiments on ParA to examine the nucleotide dependent (ADP) changes in the surface accessible area. She found that both P7 and P1 ParA were more resistant to trypsinolysis in the presence of ADP (Figure 4_2B). These observed nucleotide dependent changes also lend support to the idea that the resulting ParA structure is more conformationally rigid lowering its exposure to the surrounding environment.

The ADP-bound structure also revealed that protein segments contacting or adjacent to the bound nucleotide become more structured as a result of the binding event. Walker A ATPases are comprised of three motifs designated the Walker A (or P-loop), Walker A´ (switch 1) and Walker B (switch 2) motifs. The Walker A motif plays a critical role in the formation of the nucleotide pocket principally by making contacts with the phosphate groups of the bound nucleotide. The P-loop is a glycine rich sequence and in P1 ParA the motif contains a deviant lysine within the signature sequence KGGXXK[S/T] (residues 117-123)(Koonin, 1993). While the P-loop was mostly disordered in the apo structures, ADP binding stabilizes this motif through extensive interactions with the phosphate groups of the ADP, similar to what is observed in other structures of Walker box proteins. More specifically, the P-loop encapsulates the β phosphate allowing the backbone nitrogen of Gly119 and residues Val120, Ser121, Lys122 and Thr123 to make hydrogen bonds with the β phosphate. Lysine 122 also wraps around to contact the β

phosphate. Lys122 performs its critical role in the function of P-loop ATPases by stabilizing the negative charge that is generated in the transition state. Studies of P1 ParA protein show that mutating Lys122 to either a glutamine or glutamic acid results in a reduction or loss of ATPase activity as well as altered repressor and partition function (Fung et al., 2001). The amide groups of Thr123 and Thr124 and the side chain of Thr124 all contact the α phosphate of ADP. Also present within the ADP-bound structure was the presence of the key conserved Walker B motif residue Asp251. This aspartate residue helps to coordinate the essential Mg^{2+} ion in the binding pocket that helps to neutralize the negative charge of the transition state. The presence of these components, which are critical to the identity and function of all Walker A ATPases, lend support to the integrity of the resulting P1 ParA nucleotide binding pocket.

Besides resolving the structural elements within the nucleotide-binding pocket (like the P-loop), the ADP-bound structure resulted in the formation of new α-helix from residues 343 to 372 (designated α15´)(Figure 4_2A,C&D). This small increase in α-helical content had been noted previously in circular dichroism experiments where a ~ 5% change occurred upon ADP binding (Davey and Funnell, 1997). The potential functional ramification of this new structural element became more apparent when the DNA binding capabilities of ParA were examined. As ParA functions in the auto-repression of its own operon and the ADP binding state was believed to be the preferred state for DNA binding, we were interested in what the ADP-bound structure might reveal about the mechanism of this interaction. P1 ParA protects a wide region (~ 150 bp) around its own promoter through cooperative binding interactions, but seems to preferentially protect a ~

40 bp pseudo-palindromic section near the center of the *parOP* (Davey and Funnell, 1994). Expansive efforts to crystallize a DNA-ParA complex focusing on the 40 bp pseudo-palindromic site have been unsuccessful thus far.

As a result, Dr. Schumacher modeled the ParA-ADP-DNA complex by taking a ~40 bp DNA duplex and docking each helix-turn-helix (HTH) element in successive DNA major grooves. Previous homology sequence scanning studies revealed that residues 22-66, corresponding to the *parA* HTH, are essential for operator specific DNA binding (Davey and Funnell, 1994; Radnedge et al., 1998). Interestingly, Dr. Schumacher found that HTH elements could not be fit into successive major grooves unless the DNA was significantly bent (Figure 4_2D). This bending mechanism could be significant as it could point to how ParA might orient itself within a larger nucleoprotein complex in order to protect a 150 bp segment of its promoter (Davey and Funnell, 1994). This modeling exercise lead to the functional identification of a "basic region" (residues 340-353 and 365-380) that arises from the binding of ADP and is within contact distance of the phosphate backbone in the central region of the DNA (Figure 4_2A&D). These studies demonstrate how the binding of nucleotide (ADP) may lead to structural rearrangements and in the process provide a mechanism that enhances ParA's ability to perform its DNA binding function.

4.2.2 Validation of a new basic region and its implications to DNA binding state.

One of the potentially important structural findings was the emergence of additional secondary structure – an α-helical "basic region" – absent in the apo structures that could contribute to the binding of DNA. To test the hypothesis that residues from the "basic region" make a significant contribution to DNA binding, a series of mutations were generated (R351A, S370A and K375A/R378A) to assess the contributions of these polar residues to ParA's DNA binding function. These residues are conserved in both P1 and P7 ParA proteins. The impact of these mutations on DNA binding was first assessed using a site based around a pseudo-inverted repeat within the *par* operator as measured by fluorescence anisotropy (FA). In FA, binding (to labeled DNA) is reflected in the maintenance of polarized light, which is related to the hydrated volume (rotation correlation time) of the complex. These studies showed that the R351A and K375A/R378A had the most pronounced shift in their binding affinities for the extended pseudo-palindromic site (Figure 4_3A). Also, by comparing the deflections, ΔmP, one sees that all the mutants have a smaller deflection, but the K375A/R378A mutant is significantly smaller. As ΔmP is related to the hydrated volume, or size of the complex, this could mean that the "basic region" mutants have an impaired ability to properly bind, or have a lowered capacity to load onto the site (smaller hydrated volume). A significant piece of missing information that might provide insights into the mechanism of binding and further explain the role being played by the "basic region" mutants is the stoichiometry of the complex. Stoichiometry for the *parOP* binding site could be

established using FA (for smaller DNA) and possibly SEC-LS (for the full-length *parOP* site). The studies were performed in the presence of ATP (1 mM) rather than with ADP as binding to the shorter oligonucleotides (roughly half the site shown to be protected by ParA binding) were more easily detected experimentally in the presence of ATP. It remains unclear whether the ATP-bound state has any physiological role in binding to the operator site where as the role of the ADP-bound state has been established (Davey and Funnell, 1994). As a result our collaborator, Dr Barbara Funell, repeated these studies using a DNase I protection assay of the whole *parOP* in the presence of ADP. The DNase I protection studies confirmed that binding was altered in all the mutations and most impaired by the R351A and K375A/R378A mutants (Figure 4_3B). The functional significance of these mutations illustrates the importance of the "basic region" along with the wing and HTH in composing the P1 ParA DNA binding motif.

4.2.3 ATP-bound state: P1 ParA filament formation.

From our structure function studies of the ADP-bound state one can see how P1 ParA can be switched to a conformation that promotes its function in the auto-repression of its operon. Likewise, we know from the modeled activity of P1 ParA based upon biochemical studies that ATP must lead to a similar change allowing it to participate in partitioning (Figure 4_1A). Unfortunately, crystals of P1 or P7 ParA in the ATP-bound

state could not be obtained as the addition of ATP-Mg^{2+} resulted in the dramatic precipitation of both proteins. Closer inspection of the precipitate revealed that it was string-like in its consistency. As one of the hypothesized roles for *par* motor proteins is to function in a direct and active way in the segregation of plasmids through the formation of filaments it was possible that type Ia ParA was forming filaments relevant to its role in partitioning.

Early evidence of type II ParM filaments quickly led to a hypothesis that envisioned the direct mobilization of daughter plasmid *parRC* "partition complexes" to their respective poles occurring in a pushing rather than pulling manner. Studies of the ParM filaments found that they nucleated in the presence of ATP and proceeded to form proto-filaments that exhibited hallmark features like dynamic instability (Garner et al., 2004). In *in vitro* reconstitution experiments of the type II RI plasmid system the three canonical *par* components (ParM, ParR, and *parC*) were found to be sufficient to stabilize growing apolar ParM filaments so they could push the "partition complexes" apart in an ATP dependent manner (Garner et al., 2007). This study provided a crucial piece of evidence in support of the "insertional model," which holds that filament growth occurs through the addition of ParM subunits at the filament -"partition complex" interface. The "insertional model" was further refined with *in vitro* reconstitution experiments substituting full-length *parC* for the magnetic bead-complex used previously. This allowed the visualization of the native filament-"partition complex" interface, which lead to the proposal of a ratchet-like mechanism to explain filament growth (Salje and Lowe,

2008). These studies on the type II system show that plasmid segregation is driven by growing ParM filaments that are stabilized by interactions with the "partition complex."

Filaments have now been observed in all the *par* systems, but those originating from the type I *par* proteins have not been characterized to the extent of type II ParM. Type Ib-like chromosomal motor protein Soj and type Ib motor protein δ appear to require DNA to form nucleo-protein filaments (Leonard et al., 2005; Pratto et al., 2008). While type Ib motor proteins pB171 ParA and TP228 ParF do not require DNA to form filaments in the presence of ATP (Barilla et al., 2005; Ebersbach et al., 2006; Ebersbach et al., 2005). SopA is the only type Ia motor protein observed to form filaments in an ATP-dependent manner (Bouet et al., 2007; Hatano et al., 2007; Lim et al., 2005). The mechanisms proposed for explaining its function have mostly been observational and contradictory among the groups reporting the filaments. One group reported that SopA oscillates from pole to pole in a spiral shape reminiscent of MinD (Hatano et al., 2007). While another group reported seeing aster formations (Lim et al., 2005). There is likely to be some variability in how each type I family's motor protein participates in the segregation of its plasmid, but until more detailed structural analysis is performed it is likely that the questions about the underlying mechanisms will not be addressed.

To address the question of how P1 ParA might act in partitioning, we set out to characterize the precipitant we had observed during our crystallization trials. First, we utilized negative stain electron microscopy (EM) to ascertain the structural features of the ATP-Mg^{2+} induced precipitation at higher resolution. The resulting electron micrographs

showed that ParA (~ 4 μM) forms regular filaments in the presence of ATP-Mg^{2+} with rough dimensions (188 x 140 Å) (Figure 4_4A). To ensure that these observed filaments were the result of an interaction with ATP, the experiments were repeated in the presence of ADP instead of ATP. The resulting electron micrographs were devoid of any filamentous objects (Figure 4_ 4B). These control experiments demonstrate that the filaments are selectively generated in the presence of ATP (or ATP-analog).

As the studies on type II ParM filaments were the most extensive, they served as a roadmap for addressing the underlying mechanisms of P1 ParA filaments function. As P1 ParA appears to form filaments exclusively in the presence of ATP (or ATP-analogs), we were interested in how these filaments might act in partitioning. P1 ParA has been reported to associate with the "partition complex" (ParB-*parS* complex) while in its ATP-bound state. This link between ParA and the "partition complex" is vital for the proper segregation of the P1 plasmid and could serve as a way of demonstrating that the ParA filaments were functionally relevant. The recent studies of a reconstituted type II system using Alexa488-labeled ParM in conjunction with fluorescence microscopy opened the possibility of viewing active filaments pushing the "partition complex" (attached to beads) apart in real-time (snap shots taken at minute resolution). This was an attractive experimental model to follow as it allowed the assessment of the critical link between the ParA filaments observed in earlier experiments with the required association with the "partition complex" during partitioning. First, to determine if we could visualize the P1 filaments using fluorescence microscopy similar to reconstitution experiments in the type II system, the experiments were repeated with fluorescently labeled ParA (Alexa488)

using either ATP of ADP. Again, the addition of ATP resulted in the generation of filaments (not shown), while the addition of ADP had no effect (uniform green haze similar to the apo state)(Figure 4_4C).

4.2.4 Studies on the reconstituted "partition complex."

Are the P1 ParA filaments functionally relevant – can they associate with the partition complex (ParB + *parS*)? Experiments were undertaken in an effort to determine whether the P1 ParA filaments function similarly to the type II system or if they utilized another mechanism to partition plasmids. In order to establish the existence of P1 ParA filaments that could provide the vital link between the function of the "partition complex" and the divided plasmids in their respective daughter cells we used an *in vitro* reconstituted system. Contained within the P1 *par* operon (ParA, ParB and a *cis*-acting DNA element *parS*, the plasmid centromere) are the necessary components to drive segregation. We separately purified each component and recombined the elements to demonstrate the potential for segregation function. In the *in vitro* reconstituted partition system the "partition complex" is composed of a segment of *parS* (a sequence including either *parS* or *parS*small) that has been linked to a magnetic bead (~ 1 μm) and ParB. When ParA in its ATP-bound state is mixed with the "partition complex" we expect to see the filaments localizing to the bead-"partition complex" possibly forming connections between

multiple bead complexes (Figure 4_4D). Fluorescently labeled (Alexa488) ParA was used to visualize these interactions with a fluorescence microscope. When the ParAB*S* components were mixed in an eppendorf tube in the presence of ATP-Mg^{2+} and then applied to a microscope slide we were able to capture what appears to be a late-stage snap shot of ParA (Alexa488 labeled) in the process of partitioning ParB-*parS* bead complexes (Figure 4_4E). Upon examining the reaction on a glass microscope slide we demonstrated that ParAB*S* can form structures that not only localize the substituents to a central locus (nodes), but also are reminiscent of those observed in the ParMR*C* reconstitution experiments. We turned to negative stain electron microscopy to acquire higher resolution images of the P1 reconstituted system (ParAB*S*) in order to examine the details of the ParA-"partition complex" interaction. From the electron micrographs we observed filamentous bundles emminating from the ParB-*parS* bead complex (Figure 4_ 4G&H) that are consistent with what we would expect from a locus containing many ParB*S* complexes. Examining the filaments themselves we see they are aligned parallel to one another (Figure 4_4F). This shows that the observed filaments are consistent with ParA's known biochemical function, namely that it forms filaments in the ATP-bound state and it recognizes the "partition complex" while it is in that nucleotide-bound state.

So far, we have been able to determine that P1 ParA filaments form in an ATP dependent manner and that the filaments are able to associate with "partition complexes" *in vitro*. Reconstitution experiments of the type II system had used Alexa488 labeled ParM to demonstrate its ability to push "plasmids" (beads) apart in an ATP-dependent manner as well as demonstrating that the ParM filaments were dynamic and were stabilized by

interactions with “partition complexes” (beads)(Garner et al., 2007). They were able to accomplish this using total internal reflection microscopy (TIFM), which produces an image with a much higher signal-to-noise ratio. As a result of using TIFM, they were able visualize ParM filaments and their interactions with the “partition complex” at an earlier time-point when the complexes are just forming (and are much smaller) and the filaments can still be seen undergoing dynamic transitions (Figure 4_5). As this type of microscope was not available for our P1 reconstitution experiments, the best we could be able to resolve are the later-stage complexes that result from establishing multiple parallel connections between “partition complexes” (beads) making them much larger and visible to a standard fluorescence microscope (Figure 4_5). In effort to view the P1 ParA-ParB*C* complexes at an earlier time-point, the reaction was moved from an eppendorf tube to the microscope slide glass surface that had been washed and silanized. The limit to the sensitivity of the traditional fluorescent microscope comes from the need to use longer integration times, which results in a bleaching effect at earlier time points where small objects cannot be resolved.

From the model of the reconstituted par components shown in Figure 4_4E, we see that ParA (filaments), ParB and *parS* (“partition complex”) should all be localized to the magnetic bead. As shown in earlier reconstitution experiments of later stage partitioning Alexa488 labeled ParA is localized around both the bead complexes (nodes) and in what appeared to be filament bundles between the nodes (Figure 4_4D). At the time, we did not have the ability to localize any of the other components, so they were assumed to be present. In an effort to refine the way filaments were imaged in the past it was first

important to be able to localize all the required components of the reconstituted system. To accomplish this, the *parS* DNA used in the reconstitution experiments was labeled with a fluorophore (cy3) allowing the *parS*-beads to be localized in the complex. As ParB is the only unlabeled component, it is assumed that it is present on the *parS*-bead complex. ParB binds to *parS* forming the "partition complex" that is recognized by ParA (Bouet and Funnell, 1999). To make sure this assumption is valid, ParB was labeled with Alexa488 and reacted with the *parS*-beads. The Alexa488 labeled ParB sample appeared to contain particles of punctate fluorescence (Figure 4_6A). The *parS*(cy3)-beads were localized to these particles by changing the filter set (Figure 4_6B). The resulting merged image demonstrates that ParB localizes to the *parS*-beads (cy3-labeled)(Figure 4_6C).

Now that it has been established that ParB is localized to the *parS*-bead complex it was important to demonstrate that ParB is required for the ParA-"partition complex" interaction. ParB interacts with ParA through a domain loosely mapped to its N-terminus (Radnedge et al., 1998). It is the only component in the "partition complex" capable of interacting with ParA directly. Establishing this relationship in the reconstituted P1 *par* system would suggest that the ParA-"partition complex" interaction is specific. By removing ParB from the reaction containing ParA, ATP-Mg^{2+}, and *parS*(cy3)-beads, ParA filaments should not be observed at the *parS*-bead locus. From the resulting series of microscopy images no filaments were observed at the *parS*-beads when ATP-bound ParA was present suggesting that ParB is required for localization of ParA to the bead-"partition complex" (Figure 4_7A&C). The *parS*(cy3)-beads were localized on the slide

signifying that they were present in the reaction (Figure 4_7B). As ParA filaments were not observed during the course of the experiment it is also possible that ParB plays a role in stabilizing the P1 ParA filaments. The controls show that ParB is integral to the reconstituted P1 *par* system and that filaments are localized to the partition complex in a specific manner.

We have now established that ParB is present at the *parS*-beads forming the "partition complex" and that ParB is the central component of the "partition complex" capable of interacting with ParA. As the goal was to observe the reconstituted system at an earlier time point we repeated the earlier P1 ParAB*S* (now cy3-*parS*) reconstitution experiments (Figure 4_4E) by initiating the reaction with ATP/Mg^{2+} on a microscope slide. Monitoring the reaction with the fluorescence microscope ParA filaments became visible in ~15-20 minutes. The filaments were elongated and intersected at nodes (punctate fluorescence) with some coming out of the plane of the slide (Figure 4_8A). Switching the channel (changing the filter) in the microscope allowed the *parS*-beads to be localized (Figure 4_8B). The merged image (including the Alexa488 & cy3 channels) showed that the "partition complex" (cy3 *parS*) was localized to points on what are likely filament bundles (white arrows in Figure 4_8C). These experiments demonstrated that the observed filaments were able to interact with the "partition complex" and, in conjunction with the earlier ParB control experiments (Figure 4_7), with ParB specifically. These filaments were smaller, but were also observed at an earlier time point than the complexes in Figure 4_4E.

The partition complexes themselves were similar and are reminiscent of those observed in the type II ParMR*C* reconstitution experiments. It is possible that the ParA filaments are being stabilized by ParB*S* in a similar manner that ParM filaments are stabilized by ParR*C*. Efforts are proceeding to characterize the nature of the interaction by substituting gold-labeled full-length *parS* (intrinsically bent or with the integration host factor (IHF)) for the beads so that the native segrosome (ParA(ATP/Mg^{2+})-ParB*S*) may be imaged using negative stain EM. As each bead contains multiple "partition complexes" each capable of forming their own interactions (Figure 4_4D), their removal should simplify the evaluation of stabilizing interactions at the filament ends (primarily to see if they are capped by the gold-"partition complex").

To validate definitively the function of the P1 ParA filaments and begin to probe the underlying mechanisms by which they may operate two further investigations must be undertaken. First, the ParA filaments must be examined using TIFM in order to study the dynamics of their formation (Figure 4_5). Specifically, the TIFM experiments would provide a glimpse at the early timeframe (1-5 min.) where proto-filaments would expect to be governed by some form of dynamic instability (or reversibility), a hallmark of filament systems. Second, higher resolution studies on the composition of the filaments themselves, which might provide packing arrangements that would demonstrate that the filaments formed by specific interactions. Corresponding mutagenesis experiments (targeting the sub-unit interfaces) could provide useful candidates for crystallization trials targeting a structure of the ATP-bound form of ParA.

4.3 Discussion

One of the key questions has been to understand the structural mechanisms composing the "molecular switch" that is responsible for directing ParA's functional state. It was known that this switching mechanism was nucleotide dependent with the ADP-bound state of ParA favoring DNA binding and the ATP-bound state marking ParA's participation in partitioning (Figure 4_1A). Our studies have shown that ParA prefers to be a dimer in solution, which means that the switching mechanism is likely more nuanced than a simple monomer to dimer transition.

Dr. Schumacher's ADP-bound structure revealed that the act of binding nucleotide altered ParA's conformational flexibility locking it into a form that is more conducive to binding DNA. The emergence of a new structural element, α15, was given functional relevance when Dr. Schumacher docked a 40 bp segment of DNA onto the structure. It was discovered that the α15 "basic region" was likely a new DNA binding element that stabilizes the bending of the DNA in a manner that accommodated the participation of ParA's other DNA binding structural elements, winged-HTH. Polar residues (R351A, S370A, and R375A/R378A) of the basic region making contacts with the backbone of the DNA were neutralized through mutagenesis to assess their contributions to the DNA binding interaction. The R351A and K375A/R378A mutations in ParA resulted in a binding affinity that was significantly reduced. The primary remaining task to be accomplished in the effort to understand the mechanism of ParA's repression of its

operon is to establish the "basic regions" ability to stabilize the bending of *parOP* DNA. Altogether our studies suggest that one of the ways that ParA's functional state is influenced by the binding of ADP is through the advent of structural features that lead to its ability to form a DNA-bound complex necessary for the protection of the 150 bp segment of its promoter region.

ParA has long been known to play a vital, but unknown role in the segregation of the P1 plasmid. ParA performs this roll in an ATP-bound state where it is able to recognize and associate with a newly formed "partition complex." In our investigations the binding of ATP/Mg^{2+} resulted in the consistent precipitation of the ParA protein. While this appears to have factored into the inability to solve an ATP-bound structure, it opened investigations into functional relevance of the P1 ParA filaments that resulted. The filaments were found to form in the presence of ATP-Mg^{2+} using both negative stain electron microscopy (EM) and fluorescence microscopy, whereas the presence of ADP/Mg^{2+} had no effect. Further, we successfully reconstituted the P1 *par* system *in vitro* co-localizing the filaments with the "partition complex" (ParB + *parS*). These studies demonstrate that the filaments are the specific result of binding ATP-Mg^{2+} and in the ATP-bound state the filaments can co-localize with the "partition complex," which fulfill two of the major requirements of a relevant *par* filament. Future TIFM experiments would need to be conducted to investigate the dynamics of segrosome (ParAB*S*) assembly and to definitively address the mechanistic role of each of the components. Our observations along with recent investigations into other type I partition

systems have established that Walker A ATPase motor proteins are also capable of forming filaments in the ATP-bound state.

Our studies demonstrated that ParA filaments could be co-localized with the "partition complex" in a way reminiscent of those formed by ParM in type II *par* system reconstitution studies. If the P1 partition system is functioning in a similar manner to the type II system, the interface between the filaments and the partition complex is not only serving to stabilize the ParA filaments, but is also undergoing dramatic conformational changes leading to the elongation of the filament. Since the link between ParA filaments and the partition complex would be a common trait in most partition mechanisms, it is also possible that ParA functions by forming filaments in a similar manner to MinD. MinD is a deviant Walker A protein that establishes the bacterial cell mid-point. It functions by sampling the cell membrane in what has been observed as a spiral pattern although it has also been reported to form filaments in a ATP-dependent manner (Hu and Lutkenhaus, 2001; Lutkenhaus and Sundaramoorthy, 2003; Suefuji et al., 2002). How this activity would translate into a mechanism that would explain partitioning is unknown. Recently, Type Ia SopA has been observed to also form similar spiraling shapes in time-lapse fluorescence microscopy experiments (Hatano et al., 2007). The uncertainty in the underlying mechanism of ParA's function in partition must be resolved with higher resolution studies.

In summary, our studies show that the nucleotide bound state mediates ParA's functional states by stabilizing a specific dimeric state suited for its role in the cell – repression or

partition (Figure 4_9). Some of these findings (Figure 4_2) were included in the recent publication “Structural basis for ADP-mediated transcriptional regulation by P1 and P7 ParA” (Dunham et al., 2009).

4.4 Methods

Mutagenesis. Structural elements discovered in the P1 ParA ADP-bound crystal structure ("basic region") are proposed to play a role in ParA's DNA binding function. Selected residues in the "basic region" (Arg351, Ser370, Lys375, and Arg378) capable of forming interactions with the phosphate backbone of a bound DNA molecule were mutated to alanine residues in order to assess their contributions to the DNA interaction. Mutants were generated using a QuikChange (Stratagene) strategy with the following oligonucleotides: 5′-CGGTTTTGAAGCGTGCGGCGAGTC-3′ (R351A), 5′-CGTATGTTGGTGCG GCTGATGCATTG-3′ (S370A), and 5′-GCTGATGCATTGGCGAACGCGGCCATTGCCGCGG-3′ (K375A/R378A). The resulting mutant constructs were sequenced to confirm the substitution using the Sanger method (Applied Biosystems instrument) at the MD Anderson core facility.

Fluorescence anisotropy (FA) measurements. Binding affinities of mutant P1 ParA proteins for the P1 *parOP* site were performed using a Beacon Red Ranger steady-state fluoresence anisotropy (FA) system. Two nanomolar fluorescein labeled 5′ oligonucleotide (5′-GCTATGTACAAGCATCTACGCATACATTATTATTTTATGCAGCA TTTTTAATTAAATTCAAAAATACAGCATAAAGGATGACTTTCGATGAGTGATTCCAGCC-3′) was used as a probe to represent the P1 *parOP* site. Protein was titrated into 150 μL of buffer (25 mM Tris-Cl (pH 7.5), 150 mM KCl, 2 mM DTT) containing the DNA probe and 1 mM ATP at 25 ^{o}C. Resulting binding isotherms were generated in Kaleidagraph

using non-linear least squares analysis to fit the data to an equation describing a simple binding model: $mP = mP_i + \Delta mP^*[P]/(K_d + [P])$ (Lundblad et al., 1996).

Negative stain electron microscopy (EM). ParA filament formation and the reconstitution of the P1 ParAB*S* partition system. Polymerization of ParA filaments was initiated with the addition of ATP and $MgCl_2$. Approximately 1 mg of purified ParA (~ 40 μM) was reacted with 2 mM ATP and 5 mM $MgCl_2$ in polymerization buffer (30 mM Tris-Cl (pH 7.5), 100 mM KCl and 1 mM DTT) for 10 min. at room temperature and then subsequently diluted ten-fold in polymerization buffer before being applied to continuous carbon coated Ni^{2+} grids. Filament samples were prepared on Ni^{2+} grids with Formvar film (EMS - Hatfield, PA) that had been treated with poly-L-Lysine or Cu^{2+} grids with either continuous Carbon (C-Flat) or Formvar/Carbon film (EMS - Hatfield, PA) (glow discharged just prior to sample application). After being applied to grids samples were negatively stained with a 1% uranyl acetate/formate solution and excess stain was removed by blotting and the grids were allowed to dry before being viewed in a JEOL 1010 transmission electron microscope at the MDAnderson Microscopy Facility (Houston, TX).

Reconstitution. Reconstitution of the P1 partition system (ParA, ParB and *parS*) was accomplished using purified individual components. ParB (the gene, a gift from Dr. Funnell, contained an N-terminal Deca-His tag) was expressed in BL21(DE3) *E. coli* cells and purified using Ni^{2+} affinity followed by gel filtration chromatography (Sephacryl 200 column). *parS* components in the reconstitution experiments were

attached to magnetic beads (Bangs Laboratories, Inc) using a Biotin-Streptadvidin linkage in a manner described previously (Garner, E.C. *et al.*, 2007). The composition of *parS* consisted of either a 53 bp segment containing a ~25 bp linker followed by the A2, A3 and B2 boxes of the *parS* site or a 330 bp segment containing the entire *parS* site. The 53 bp site was prepared by annealing top (5' Biotin) and bottom strand oligonucleotides (Sigma). The *parS* segment containing the full site was assembled from individual overlapping 40 bp oligonucleotides and amplified using flanking primers (top strand 5' Biotin labeled) in a PCR method described previously (Stemmer, W.P. *et al*., 1995). The following oligonucleotides were assembled in a PCR protocol using a higher number of reaction cycles (55) to generate full-length *parS*: 5′-CGAGGTGAAGCTTGGCTGCC-3′ (prop_1), 5′-CATTGAAGCAGGAA ATCACCAAAATGATTCAGGCTACAAC-3′ (prop_3), 5′-CTGAACGTAGAAGAAATCCG CGTCCTTTATGCGTGGAGGA-3′ (prop_5), 5′-TGCCAAAGCATGTTGTGACACACTTGGCAA AGGAGTAAGC-3′ (prop_7), 5′-ATGCAGAGAATGCTATGTACAAGCATCTACGCATACATTA-3′ (prop_9), 5′-TTATTTTATGCAGCATTTTTAATTAAATTCAAAAATACAG-3′ (prop_11), 5′-CATA AAGGATGACTTTCGATGAGTGATTCCAGCCAGCTTC-3′ (prop_13), 5′-ACAAGGTTGCTCA AAGAGCAAACAGAATGCTCAATGTTCT-3′ (prop_15), 5′-GACTGAACAAGTACAGT TGCAAAAGGATGAGCTACACGCG-3′ (prop_17), 5′-AACGAGTTTTACCAGGTCTAT GCGAAAGCGGCACTGGCAA-3′ (prop_19), 5′-AATTGCCTCTACTGACTCGCTCG TGAGTCGAGTATACCGG-3′ (prop_21), 5′-GCAAAAAACGGCAAAAAACGGGCA AAAAACGGCAAAAAAC-3′ (prop_23), 5′-GGGCAAAAAACGGCAAAAAAATCGAT TGAGGATCCAAACT-3′ (prop_25), 5′-TTCGCCATTCAAATTTCACTATTAACT GACTGTTTTTAAA-3′ (prop_27nat), 5′-GTAAATTACTCTAAAATTTCAAGGTGA AATCGCCACGATT-3′ (prop_29nat), 5′-TCACCTTGGTCTAGAGTACGGCGAATCG GTAGCTACTTGC-3′ (prop_35nat), 5′-GCAAGTAGCTACCGATTCGC-3′ (prop_24nat),

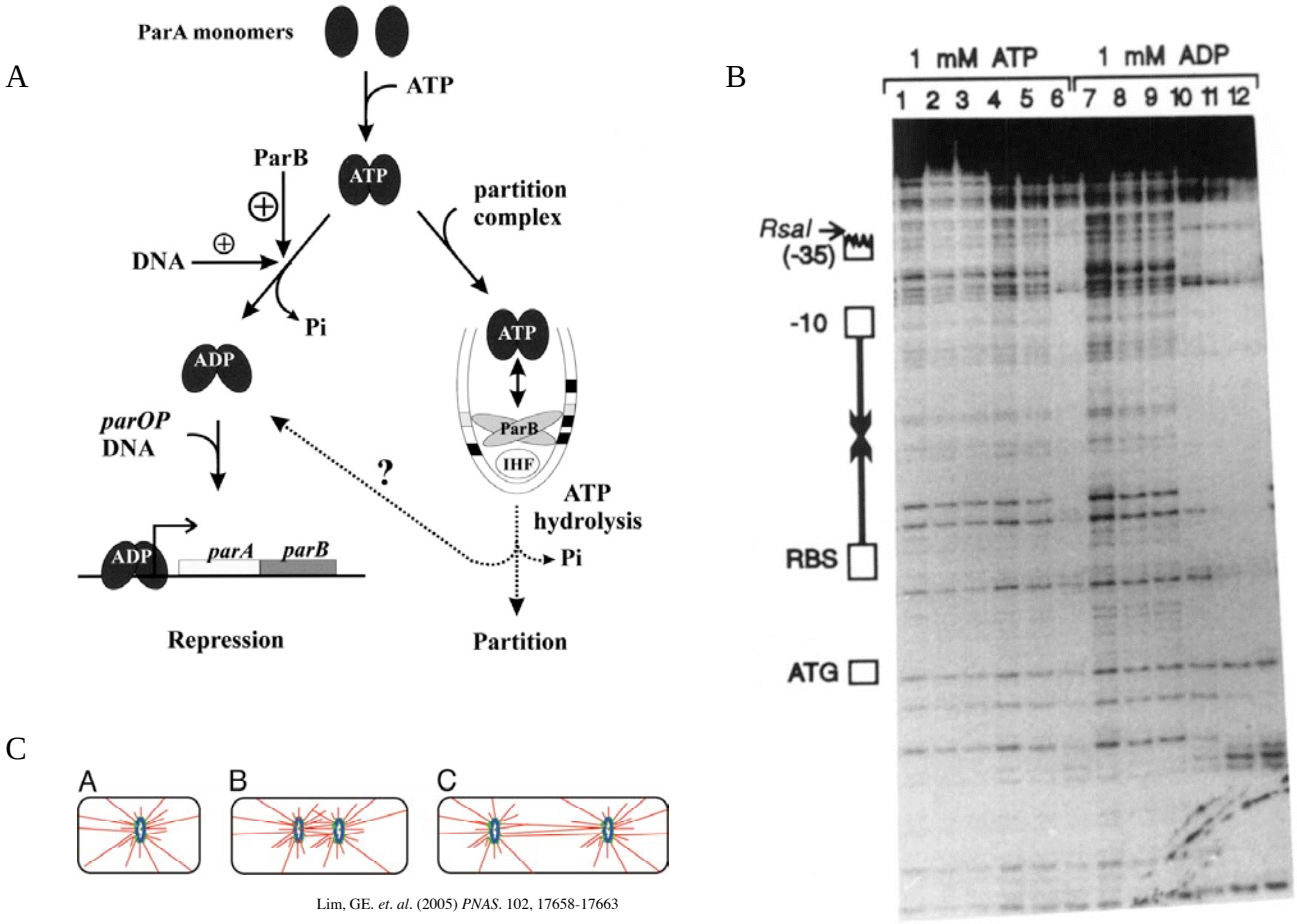
5′-CGTACTCTAGACCAAGGTGAAATCGTGGCGATTTCACCTT-3′ (prop_6nat), 5′-GAAATT TTAGAGTAATTTACTTTAAAAACAGTCAGTTAAT-3′ (prop_8nat), 5′-AGTGAAATTTGA ATGGCGAAAGTTTGGATCCTCAATCGAT-3′ (prop_10), 5′-TTTTTTGCCGTTTTTTG CCCGTTTTTTGCCGTTTTTTGCC-3′ (prop_12), 5′-CGTTTTTTGCCGTTTTTTGCCCGG TATACTCGACTCACGA-3′ (prop_14), 5′-GCGAGTCAGTAGAGGCAATTTTGCCAGTGCCG CTTTCGCA-3′ (prop_16), 5′-TAGACCTGGTAAAACTCGTTCGCGTGTAGCTCATCCTTTT-3′ (prop_18), 5′-GCAACTGTACTTGTTCAGTCAGAACATTGAGCATTCTGTT-3′ (prop_20), 5′-TG CTCTTTGAGCAACCTTGTGAAGCTGGCTGGAATCACTC-3′ (prop_22), 5′-ATCGAAAGTCA TCCTTTATGCTGTATTTTTGAATTTAATT-3′ (prop_24), 5′-AAAAATGCTGCATAA AATAATAATGTATGCGTAGATGCTT-3′ (prop_26), 5′-GTACATAGCATTCTCTG CATGCTTACTCCTTTGCCAAGTG-3′ (prop_28), 5′-TGTCACAACATGCTTTGGCAT CCTCCACGCATAAAGGACG-3′ (prop_30), 5′-CGGATTTCTTCTACGTTCAGGTTGTAGCCT GAATCATTTT-3′ (prop_32), 5′-GGTGATTTCCTGCTTCAATGGGCAGCCAAGCTTCACCTCG-3′ (prop_34). The *parS* segment was amplified using flanking primers (prop_1 (5′- Biotin) and prop_24 (3′- Cy3 molecule).

Fluorescence microscopy. For fluorescence microscopy reconstitution experiments ParA was labeled with an Alexa488-maleimide derivative (Molecular Probes) and re-purified using a Ni^{2+} affinity spin column (GE). Reconstitution experiments consisted of 40 μM ParA (25% Alexa488 labeled ParA), 2-5 μM ParB, 1 μL *parS* beads in a buffer containing 20 mM Tris-Cl (pH 7.5), 0.1 M KCl, 5 mM $MgCl_2$, 10% Glycerol, 1.5 mg/mL BSA, and 1 mM DTT. Polymerization was initiated with the addition of 2 mM ATP and allowed to proceed for 10-30 min. at room temperature before being diluted 10 fold. Fluorescence microscopy images were taken using Image Pro software on a Leica

DM4000B fluorescence microscope (100x oil-immersion objective) mated to a Leica color camera 300.

4.5 Acknowledgements

I would like to recognize Dr. Schumacher's work in solving the P1 ParA – ADP bound structures, modeling the ParA-ADP-DNA structure and analysis. I would also like to recognize the work of or collaborator Dr. Barbara Funnell in performing the limited trypsin-digest and DNase I protection experiments.



Lim, GE. *et. al.* (2005) *PNAS.* 102, 17658-17663

Figure 4_1: Nucleotide dependent functional states of P1 ParA. (A) Model of P1 ParA activity based upon accumulated data from biochemical experiments (Bouet and Funnell, 1999). (B) DNase I protection assay illustrating the ~ 150 bp region (center is a ~ 21 bp imperfect repeat) protected by ParA in its ADP-bound form (Davey and Funnell, 1994). (C) Model of type Ia SopA motor protein observed to form asters during time-lapse fluorescence microscopy experiments.

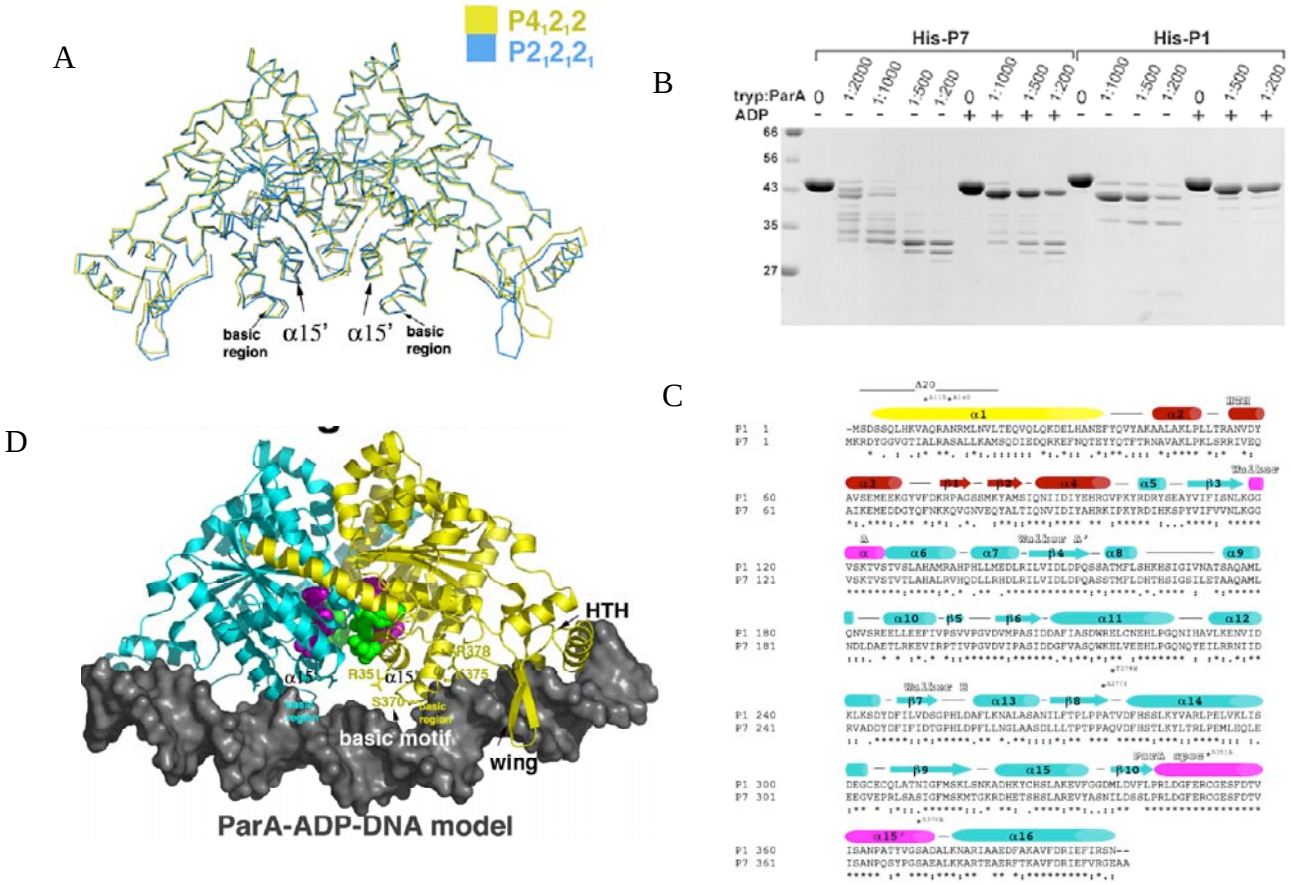
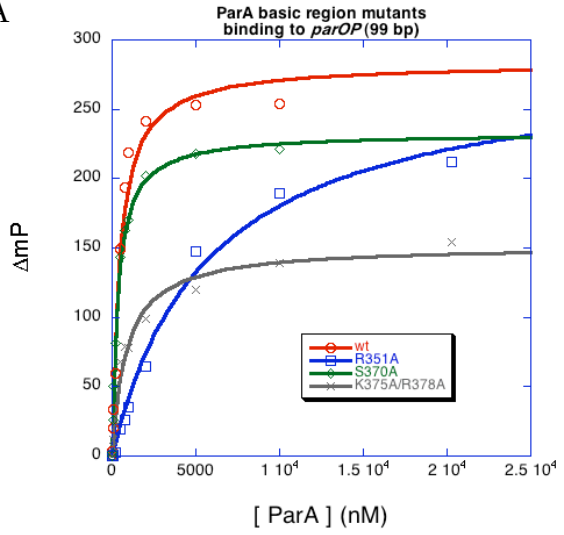


Figure 4_2: P1 ParA in the ADP-bound state. (A) Overlayed structures (main-chain) of ADP-bound P1 ParA from two crystal forms. New structural elements (basic region and α15') noted. (B) Results of trypsin-digest experiments of P1 and P7 ParA proteins +/- ADP/Mg^{2+} performed our collaborator – Dr. Barbara Funnell. (C) P1 and P7 ParA sequence alignment with structural features noted above (pink are ADP dependent changes). (D) Model of DNA-bound ParA (ADP-bound) complex illustrating the main contributions (wing, HTH and basic region motifs) to the binding interaction. This figure composed part of a recent publication (Dunham et al., 2009).

Construct	*Kd (nM)*	*[Nucleotide](mM)*	*parOP*
NtHis ParA	460.0 +/- 80.0	1mM ATP	99
R351A	5750.0 +/- 840.0	1mM ATP	99
S370A	390.0 +/- 20.0	1mM ATP	99
K375A/R378A	870.0 +/- 110.0	1mM ATP	99

A



B

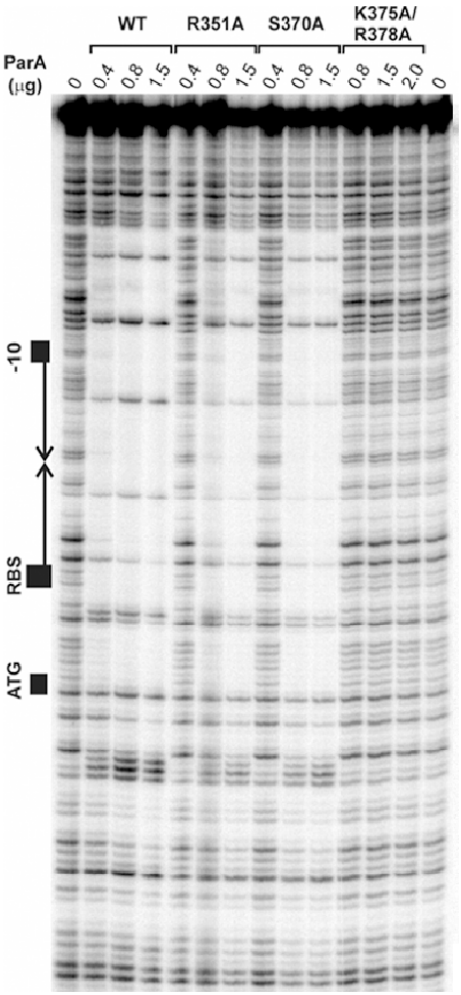


Figure 4_3: DNA binding experiments evaluating the role of the basic region. (A) Representative fluorescence anisotropy (FA) experiments measuring ParA's affinity for a 99 bp fluorescein labeled-oligonucleotide centered around the 42 bp pseudo inverted repeat. Measurements were conducted in the presence of 150 mM KCl and 1 mM ATP. Basic region mutants R351A and K375A/R378A displayed impaired ability to bind DNA. (B) DNAse I protection assay of *parOP* by ParA in ADP-bound state. Basic region mutants R351A and K375A/R378A also displayed impaired ability to protect the site.

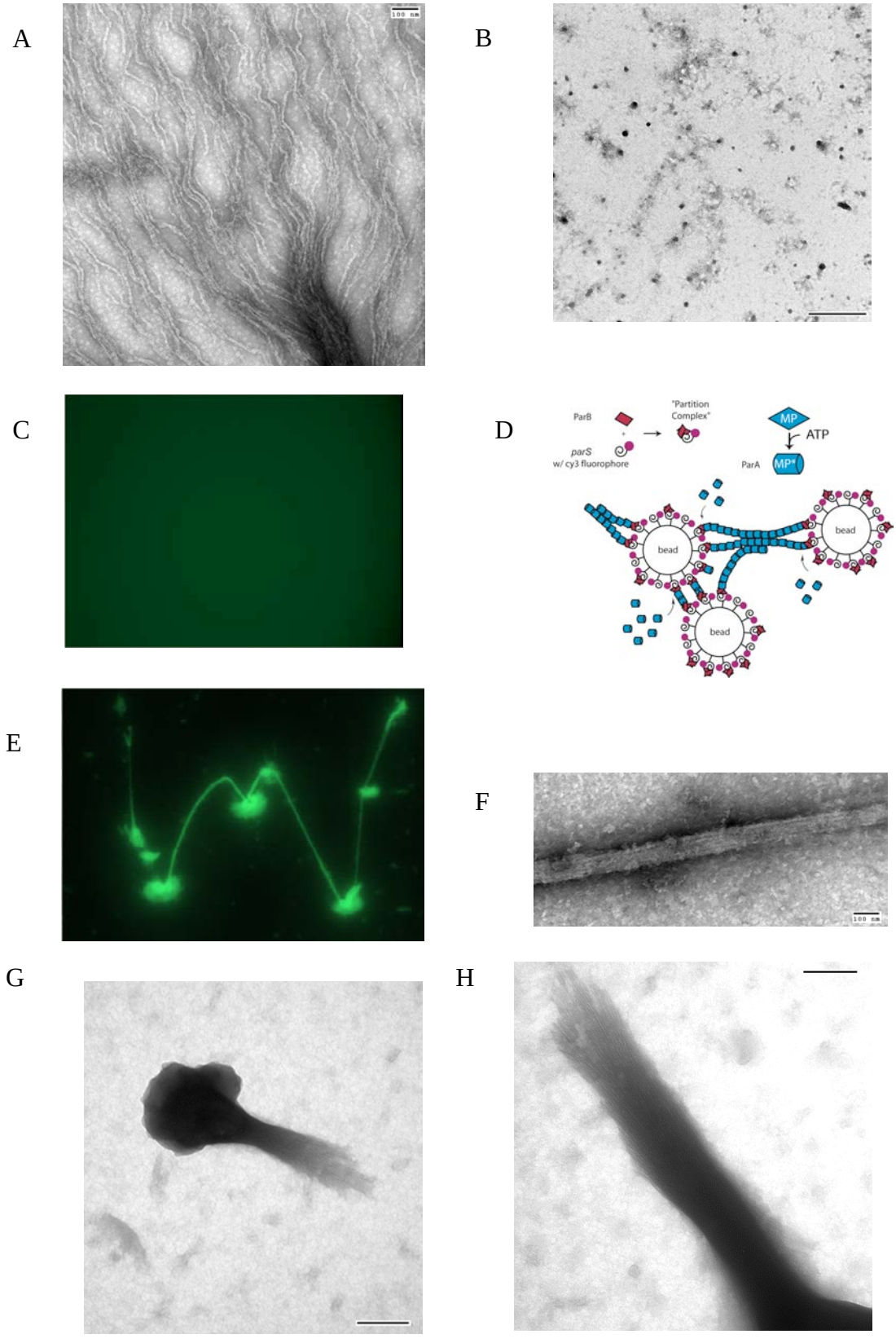


Figure 4_4: Reconstituted P1 Partition Complex. (A) Electron micrograph of negative stained ParA-(ATP/AMP-PCP)/Mg^{2+} filaments (bar = 100 nM). (B) Electron micrograph ParA-ADP/Mg^{2+}: no filaments (bar = 500 nM). (C) Fluorescence microscopy image of Alexa-488 labeled ParA-ADP/Mg^{2+}: no filaments. (D) Experimental model of type Ia reconstituted system – ParA-B*S*(magnetic beads). (E) Fluorescence microscopy image of Alexa-488 labeled ParA in an *in vitro* reconstituted system containing ParB/*parS*-complexed beads after ~ 60 min. exposure to ATP/Mg^{2+}. (F) Electron micrograph of filament bundle. (bar = 100 nM) (G) Electron micrograph of reconstituted ParAB*S* system in the presence of ATP and Mg^{2+} resulted in the formation of filament bundles when ParB-*parS* were complexed to magnetic beads via a Streptavidin-Biotin linkage. Image shown at 50,000x magnification (bar = 500 nM). (H) Same as G at 100,000x (bar = 100

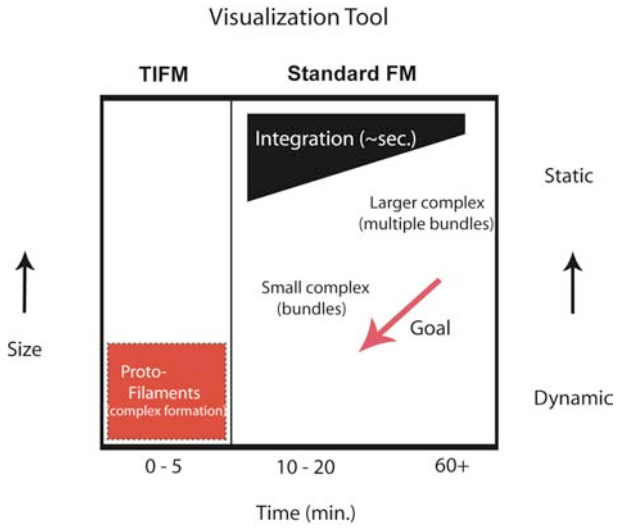


Figure 4_5: Fluorescence microscopy approach. Comparison of the resolution of total internal fluorescence microscopy (TIFM) and the experimental method standard fluorescence microscopy. Goal is to visualize at an earlier time point when proto-filaments are forming.

B

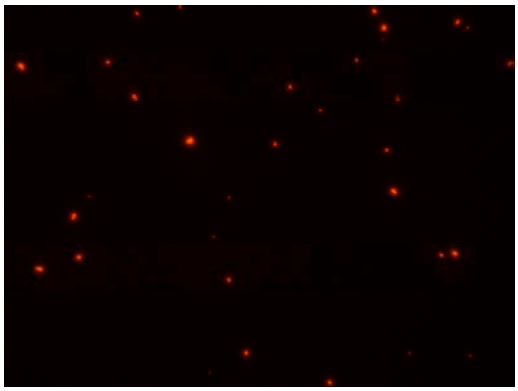
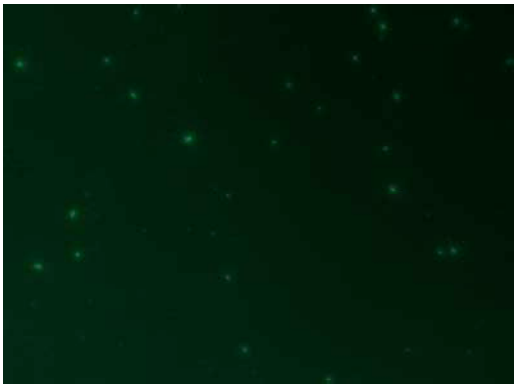
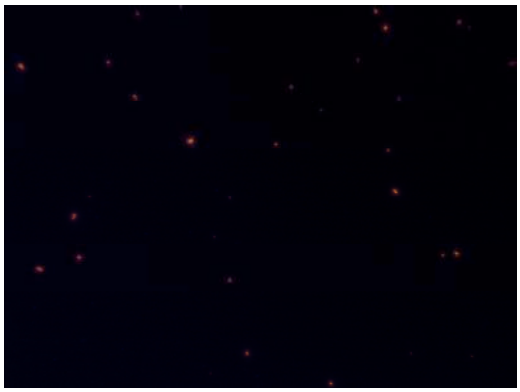


Figure 4_6: P1 *par* reconstituted system – ParB-*parS* interaction. (A) Alexa488 labeled ParB. (B) cy3 labeled *parS* magnetic beads. (C) Merged image demonstrating the co-localization of ParB and *parS*.

A

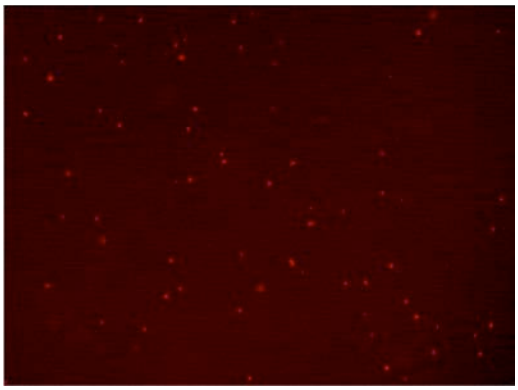


C



A

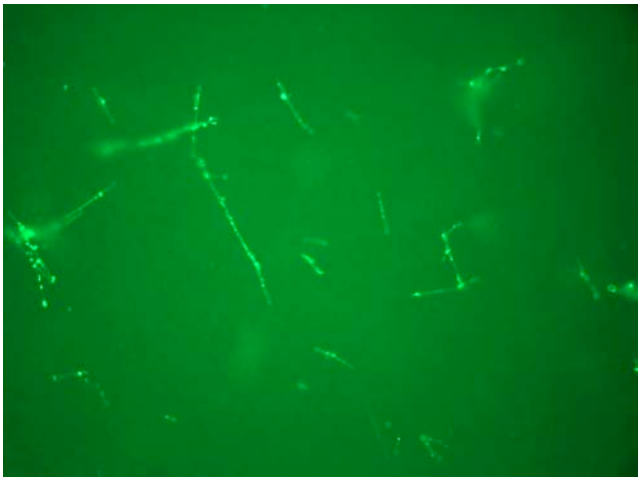
B



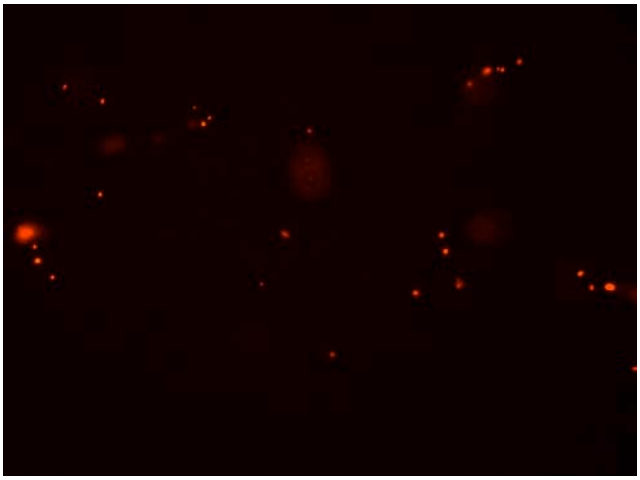
C

Figure 4_7: P1 *par* reconstituted system – ParA-*parS* interaction. (A) Alexa488 ParA-ATP/Mg^{2+}. No filaments are observed. (B) cy3 labeled *parS*. (C) Merged image. Does not appear to be a ParA-*parS* interaction.

A



B



C

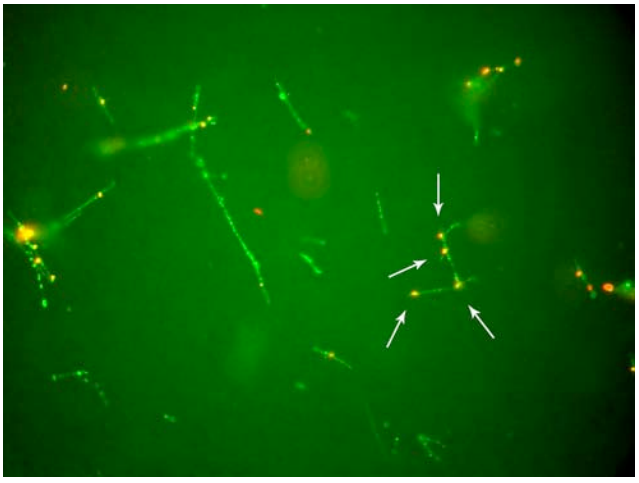


Figure 4_8: P1 *par* reconstituted system – ParAB*S* interaction. (A) Alexa488 labeled ParA-ATP/Mg^{2+}. Filaments are clearly visible. (B) *parS*-ParB bead complex – the "partition complex." (C) Merged image. ParA filaments in (A) are co-localized with "partition complex"-beads (white arrows).

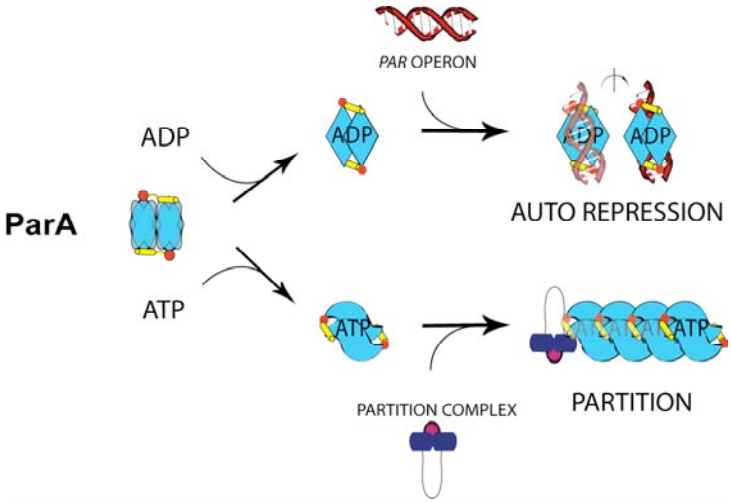


Figure 4_9: Model of P1 ParA's functional activity. A flexible apo dimer is stabilized by the binding of adenosine nucleotides in a form that is appropriate for its functional activity. The binding of ADP locks ParA into a conformation that is suitable for binding DNA – AUTO REPRESSION. The binding of ATP presumably locks ParA into a conformation where inter-dimer contacts promote its polymerization – PARTITION.

Chapter 5

Functional Implications of a ParA-ParB Fusion Protein

5.1 Introduction

Par motor proteins perform an essential role in partitioning in the cell. Whether the motor proteins are acting through filament formation or in some other capacity, the nucleotide-bound state is directly linked to its role in the cell. The binding of ATP marks the ability of these motor proteins to recognize their respective "partition complexes" (centromere binding protein bound to the plasmid centromere-like site) and participate in the segregation of newly replicated plasmid. Further, the nucleotide binding state of all *par* motor proteins being comprised of either ATPase (type Ia and Ib) or actin-like domains are modulated by their partner centromere binding proteins. In type Ia and Ib systems ParA is dependent upon an interaction with ParB to stimulate the hydrolysis of ATP. The timing and role hydrolysis plays is an essential yet understudied part of the partition process in both type Ia and type Ib systems.

The underlying mechanisms of partitioning are best understood in the type II *par* system. ParM segregates newly replicated plasmids through the formation of filaments between paired plasmids that act to drive the plasmids to the cell's poles. A new model of ParM driven partition suggest that most ParM molecules in the filaments remain in the ADP-bound state after their interaction with the partition complex (Popp et al., 2008; Salje and Lowe, 2008). This model suggests that ParM conforms to the basic model of dynamic instability and that the filaments themselves are only stabilized through their contact with the "partition complex." *In vivo* studies of ParM filament formation suggest that the

proto-filaments are stabilized and become elongated when bound to non-hydrolysable ATP analogs suggesting that hydrolysis is an essential step for de-polymerization and dynamic instability (Garner et al., 2004).

The mechanistic features of partitioning are not known in the type I *par* systems although increasing evidence suggests the motor proteins of these systems act by forming filaments as well (type Ia (SopA) and Ib (δ, ParF, and pB171 ParA) (Barilla et al., 2005; Ebersbach et al., 2006; Ebersbach et al., 2005; Hatano et al., 2007). Our evidence now suggests that P1 ParA forms filaments in the presence of ATP (or ATP analogs) as well. ATP hydrolysis is thought to be an essential step in ParA's role in partitioning as genetic studies introducing mutants (hydrolysis deficient) into *parA* on P1 *par* locus abrogates plasmid partitioning (Davis et al., 1996; Fung et al., 2001). An additional consequence of hydrolysis in the type Ia system is that the conversion of ParA to its ADP-bound state, which is the preferred form for binding DNA, enables ParA to fulfill its role in auto-repression (Bouet and Funnell, 1999).

Type I ParA proteins are Walker A ATPases distinguished by their signature GXXGXGK[S/T] P-loop motif (Saraste et al., 1990; Walker et al., 1982). The residues within the P-loop motif are essential for the correct orientation of the di/tri-phosphates of bound nucleotide moieties as well as for the hydrolysis of the gamma phosphate (Figure 5_1A). The Walker A ATPases of the *par* systems contain the hallmarks of a P-loop ATPase including the switch I (Walker A´) and switch II (Walker B) motifs. Within ParA's Walker A´ and Walker B motifs are two essential aspartate residues (D152 and

D251) that coordinate a Mg^{2+} ion, which is also necessary for the hydrolysis reaction. These interactions prime the nucleotide for a water molecule to attack the gamma phosphate leading to hydrolysis (Rayment, 1996)(Figure 5_1B). The molecule serving as the general base is uncertain and could be the gamma phosphate itself. The P1 ParA molecule is a member of a deviant Walker A motif with a signature KGGXXK[S/T] motif containing a second lysine that is the signature residue of this sub-family (Koonin, 1993)(Figure 5_1C). What has distinguished this class of ATPases is that they dimerize upon binding ATP forming a nucleotide sandwich fold. In the nucleotide sandwich fold the deviant lysine (K117 in ParA) makes a cross-dimer contact to the opposing γ-phosphate adding to the stability of the binding pocket (Georgiadis et al., 1992; Leonard et al., 2005)(Figure 5_1C). The *par* Walker A motor proteins are very weak ATPases and are stimulated by an unknown mechanism through an interaction with its partner protein, *parB*.

This interaction is common among *par* motor proteins and has been observed in other type Ia, type Ib and type II systems. The basis for the interaction between *parA* and *parB* has been mapped to the flexible N-terminus in several *par* systems using genetics and biochemical techniques. The ParB proteins are specific for their own motor protein, so the domain must contain a recognition signature. In the type Ia P1 and P7 systems this specificity was mapped to the first 28 residues in domain swapping experiments. The source of the stimulation was localized to the first 68 residues by measuring the stimulatory effects of including a GST-N-terminal 68 residue ParB poly-peptide in an ATPase reaction with ParA (Radnedge et al., 1998). In the type Ia F plasmid system this

interaction was mapped to the N-terminal 45 amino acids of ParB using a yeast 2-hybrid assay (Ravin et al., 2003). The type Ib-like system of the Soj protein was found to be stimulated by a peptide composed of the N-terminal 20 amino acids of its partner SpoOJ (Leonard et al., 2005). What remains to be determined is the underlying mechanism responsible for the stimulatory effects of the ParA-ParB interaction.

5.2.1 Design and Construction of P1 ParA-ParB Fusion Protein

In structural studies of P1 ParB, the N-terminus was found to be flexible and sensitive to proteolytic degradation and was removed in order to obtain diffraction quality crystals ((Schumacher and Funnell, 2005). The N-terminus is unstructured in the full-length structure of the chromosomal SpoOJ protein (partner to Soj) (Leonard et al., 2004). This appears to be common among *par* centromere binding proteins and as a result there is currently no structural information on the N-terminus of a type I ParB protein. Without any structural information to guide the discussion it is a difficult task to model any potential mode of interaction or propose a mechanism for the ParB dependent stimulation of ParA's ATPase activity. As the N-terminal interaction domains are believed to be unstructured it is possible that the right segment would have an induced structure through a ParA interaction. One approach to determining the structure of a functionally relevant ParA-ParB complex would be to establish the interaction/activation domain and focus

your efforts on that ParB site. Previous studies on P1 ParB determined that the source of the specificity for ParA was within the N-terminal 28 residues, while the source of its activity was within the N-terminal 68 residues (Radnedge et al., 1998). A ParA-ParB fusion could also be used to examine the biochemical nature of ParB's interaction with ParA. To avoid the complicated reconstitution efforts of working with the full length P1 ParB protein, I used a strategy of employing a fusion protein in order to expedite my efforts to study the ParA-ParB interaction.

Fusion constructs have been found to be effective tools for studying protein-protein interactions capable of probing the relationships between both ligands and effector molecules. They have been utilized extensively to study G protein coupled receptor (GPCR)-G protein interactions. G_{α}-proteins interact with their associated GPCRs when the receptors are in their active state through a sequence located in the G_{α} C-terminus (Hamm et al., 1988). Coupling the receptor to their associated G_{α}-protein in a fusion construct has been demonstrated to be an effective means of mimicking the native interaction enabling the rapid examination of different classes of ligand compounds (Milligan, 2000).

The first ParA-ParB fusion protein incorporated the N-terminal 28 residues of ParB onto the C-terminus of ParA using a flexible linker (Figure 3_2A and methods). This construct expressed and could be purified like "wt" ParA (Figure 3_2B). It appeared to be indistinguishable from "wt" in its ability to form the apo dimer as measured in static light scattering experiments. As our goal was to characterize the interaction to establish

the functional interacting/stimulating regions, we decided to use a functional assay that measured ParB's ability to stimulate ParA's ATPase activity. ParA's ATPase activity was monitored through the detection of P_i from the hydrolysis reaction using a spectroscopic detection assay based upon the P_i dependent conversion of methylthioguanosine (MESG) ribonucleoside to free MESG, which has an associated absorbance shift from 335nm to 360nm (Figure 5_2A)(Webb, 1992). The assay was used either to monitor the generation of P_i (in a cuvette) continuously in real time at 22 ^{o}C or at a single time-point after a 30 minute incubation period at 30 ^{o}C (similar to published values).

5.2.2 Functional Characterization and Validation of the Important ParB Interaction Domain.

P1 ParA is a weak ATPase. In published *in vitro* measurements, ParA (10 pmol) is able to generate about 45 pmoles of ADP during a 90 minute reaction (at 30 ^{o}C)(~ 0.0005 pmol P_i/pmol ParA $*$ min.). This reaction is stimulated by an interaction with ParB which leads to a ~ 5-10 fold increase over the basal reaction level (248 pmoles)(~ 0.003 pmol P_i/pmol ParA $*$ min.)(Fung et al., 2001). The level of stimulation by ParB is similar in other type I *par* systems (~ 20 fold) (Leonard et al., 2004; Pratto et al., 2008). As ParA's ATPase "activity" and its interaction with ParB has been documented, its

characteristics can used as a way to validate the function and behavior of the *par* proteins expressed in my studies. To ensure that the ParA-ParB fusion constructs were able to stimulate ParA's ATPase activity in a similar manner to a "wt" *par* system it was first important to establish a baseline for the reconstituted P1 ParAB*S* system.

Reconstitution of the P1 *par* system using purified components (ParA, ParB and *parS*small) demonstrated an ability to hydrolyze ATP at a level significantly higher than ParA alone establishing a baseline level of activity for comparison with other constructs (Figure 5_2B). When the ParA_fusNt28ParB construct was also assayed at 22 ºC its activity was similar to the reconstituted *par* system (ParAB*S*small) (Figure 5_2B). This demonstrates that the N-terminal 28 residues of ParB are similarly effective in stimulating ParA's ATPase activity indicating that the residues necessary for recognition as well as activation lie within the N-terminal 28 amino acids of ParB. A peptide corresponding to the N-terminal 28 amino acids of ParB was also assayed for its ability to stimulate ParA, but it was required to be present in 5-10x molar excess of ParA to see significant stimulation of ParA's ATPase activity. Attempts to assay the affinity of a fluorescently labeled version of the ParB N-terminal peptide were unsuccessful as significant precipitation prevented the interaction from saturating (extrapolating the saturation point resulted in an estimated affinity in the tens of micro molar range).

Progress curves of ParA's stimulated ATPase activity are reminiscent of pre-steady state conditions with "burst" and "steady-state" components of the reaction (Figure 5_2B). The burst displayed by the ParA_fusNt28ParB construct is higher than the ParAB*S*

reconstituted system and is likely a result of the fact that ParB, while being at the same concentration as ParA (1:1), tends to suffer from degradation (effectively lowering its concentration in this reaction). The burst component of the ParA_fusNt28ParB construct shows that the stimulatory effect of ParB leads to an increase in the rate of hydrolysis. At the end of the burst as ParA has completed one cycle it appears the rate of hydrolysis slows considerably as the release of ADP appears to be the limiting step. It is possible that ParB also functions as a nucleotide exchange factor aiding in the release of ADP and increasing the turnover rate.

The activity of the ATPase was demonstrated to be specific to ParA as a mutation to the signature Walker A lysine (K112Q) in the fusion construct abrogated its ability to hydrolyze ATP (Figure 3_2B). Interestingly, an N-terminal ParA deletion mutation (Δ1-20) also abrogated the ATPase activity (Figure 5_2B). This mutation was demonstrated to break the dimer interaction in earlier studies (Chapter 3). A P7 for P1 N-terminal 20 amino acid swap was demonstrated to prevent partitioning (Radnedge et al., 1998). It is likely that this mutation was ineffective in partitioning due to its inability to form a stable dimer and hydrolyze ATP.

Apart from an attempt to understand the mechanistic basis of ParB's stimulation of ParA, crystallizing ParA in a complex with its ParB partner remains a major research goal. As the N-terminal 28 residue ParB portion of the resulting P1 ParA apo structure was missing, attempts were made to engineer a more stable complex using an alignment of motifB (SpoOJ) and secondary structural prediction as guides. The inclusion of residues

1-69 became ParA_fusNt69ParB; the inclusion of residues 1-121 became ParA-fusNt121ParB; and the addition of full-length ParB (1-333) became ParA_fusfullParB (Figure 5_3A). Fusion constructs based upon the N-terminal 78 and 101 ParB residues resulted in proteins that were excessively proteolyzed and were not pursued further. The goals of these studies were to confirm that the domain responsible for ParB's stimulatory activity was within the N-terminal 28 residues and to monitor the functional effect of the additional structural elements.

The reaction conditions were altered for this study (30 ºC) in order to more closely match earlier published results on ParA's ATPase production of P_i. As the increase in temperature resulted in an increase in scatter that added instrumental complexity, the real-time measurements were abandoned in favor of a single time point measurement that could be used for comparison within the study and with published values. Absorbance values were converted to nmoles of P_i using a standard curve established by plotting the absorbance assay results against assays where known concentrations of phosphate (KH_2PO_4) were used (Figure 5_3B). The activity of the full ParA-ParB fusion with *parS*small was measured to be ~ 1 nmol after 30 minutes at 30 ºC (~ 0.067 pmol P_i/pmol ParA ∗ min.)(Figure 5_3C). The ParA fusion protein's productivity is roughly 4x the published results (248 pmoles P_i) for the P1 ParA-ParB-*parS* system measured at 90 min. at 30 ºC (Fung et al., 2001). The slightly higher production of the fusion construct could be due to the fact that an attached ParB protein would have fewer degrees of freedom and as a result would effectively be at a higher local concentration. Interestingly, the activity of the fusion protein containing the N-terminal 28 residues of ParB (ParA_fusNt28ParB –

2.3 nmoles P_i) was roughly twice that of the fusion incorporating the full-length ParB molecule (~ 0.153 pmol P_i/pmol ParA ∗ min.)(Figure 5_3C). This is likely due to the fact that the smaller N-terminal 28 residue construct is less constrained and may be free from additional regulatory domains. This was observed in the MinD system where $MinE_{1-31}$ was found to be a more potent stimulator of MinD activity than $MinE_{1-88}$ (Ma et al., 2003). MinD is a deviant Walker A protein involved in establishing the cell mid-point. Its partner, MinE, stimulates MinD's ATPase activity in a similar manner to ParB.

The difference in the activity of the N-terminal peptide compared to full-length ParB could also be a manifestation of an important regulatory element in the P1 *par* system. In genetic experiments it was proposed that *parB* has a regulatory element that is composed of a C-terminal to N-terminal contact that limits *parB's* ability to bind to *parA* (Radnedge et al., 1998). The internal constraint within ParB was demonstrated to be liberated when ParB bound DNA (Surtees and Funnell, 1999). ParB binds DNA through a bridging interaction utilizing both its helix-turn-helix (HTH) domain and its C-terminal dimerization domain (Schumacher and Funnell, 2005). It has been hypothesized that this release of the internally constrained N-terminus by the correct interaction between ParB and *parS* in the formation of the "partition complex" allows ParB to recruit ParA.

The two intermediate fusion constructs displayed interesting behavior as well (Figure 5_3A). The ParA_fusNt69ParB protein had a measured P_i production of ~ 600 pmol in 30 min. (~ 0.040 pmol P_i/pmol ParA ∗ min.)(Figure 5_3C). This level of activity was lower than both the full fusion and fusion composed of the N-terminal 28 residues of

ParB. The addition of ~ 40 residues has a marked impact on the fusions ability to hydrolyze ATP effectively. In previous studies the N-terminal 59 residues was found to contain an oligomerization domain thought to be important for establishing protein-protein contacts in the "partition complex" (Surtees and Funnell, 1999). If the addition of these residues has lead to the oligomerization of fellow fusion domains this could prevent their proper interaction with ParA accounting for the observed reduced activity. The addition of only ~ 50 more residues in the ParA_fusNt121ParB protein restores the ATPase activity back to the level similar to that of the N-terminal 28 residue fusion. Within that stretch lies the conserved B motif (this conserved primary sequence has an undefined function), which may help dilute the potency of the oligomerization domain along with other structural elements freeing up the ParB N-terminus to interact with ParA. From a functional standpoint it appears that the fusion incorporating the N-terminal 121 amino acids is the best candidate for future structural studies.

5.2.3 Mutational Studies – "R-Finger Hypothesis."

One of the more interesting outstanding questions in the type I *par* system is how an interaction between ParB and ParA can lead to the stimulation of ParA's ATPase activity. While the interaction domain in *parB* has been generally mapped to the N-terminal 28 to 45 residues in both type Ia and Ib systems its corresponding interaction domain in *parA*

remains less studied. More significantly, there is almost no information on the underlying molecular mechanism responsible for the stimulation of ATPase activity. One of the prime candidate mechanisms that can explain this ParB dependent activation is the "arginine-finger hypothesis" (Barilla et al., 2007). The mechanism first identified in G-protein activating proteins (GAPs) as a mechanism that explained their ability to promote hydrolysis in G-proteins. The arginine finger mechanism was later found to be important for the stimulation of other P-loop ATP/GTPases (Saraste et al., 1990). The mechanism operates by inserting an arginine residue into the nucleotide-binding pocket on the opposite side of the P-loop where it can interact with the NTP phosphates (Scheffzek et al., 1998; Scrima and Wittinghofer, 2006). It promotes hydrolysis by neutralizing negative charges occurring during phosphoryl release reaction. There is little sequence homology to identify the domain other than they are often found on flexible loops or tails.

An examination of the amino acid composition within the first 28 residues of P1 ParB for arginine residues that could potentially act as a catalytic residue uncovered two residues, Arg6 and Arg11. A sequence alignment of deviant Walker A type I ParB proteins highlighting the region where these arginines are contained shows that there are several conserved residues including an arginine at position 11 in P1 ParB (Figure 5_4). The importance of this residue was demonstrated in a mutagenesis experiment in the type Ib-like *Thermus thermophilus* SpoOJ (ParB) protein (Leonard et al., 2005). When the R10A mutation was made in a 20 mer peptide of the N-terminus of *T. therm* SpoOJ it was shown to abrogate the stimulatory activity of the "wt" peptide on its partner deviant Walker A protein Soj (ParA). The importance of arginine (basic) residues within the type

I ParB families appears to be a recurring theme as it was observed in the type Ib TP228 ParB. Here, the important arginine residue was mapped further in at position R19.

To investigate the role of these conserved residues in the N-terminus of the P1 ParB (K3, K4, R6, and R11), we made a series of point mutations to first neutralize then probe the specificity of these charged residues to see if they function either in the binding specificity or in contributing to ParA's ATPase activity. First, making substitutions of alanines at positions K3, K4, R6 and R11 and checking the ATPase activity as measured by the generation of P_i. Demonstrating that the activity of the ParA_fusNt28ParB construct is roughly equivalent to the reconstituted system (Figure 5_3C) enables these mutations to be made in the fusion construct without compromising the integrity of the results. When the K3A and K4A mutations are made in the ParA_fusNt28ParB construct they appear to have little effect on the ParA's ability to hydrolyze ATP (Figure 5_5). While the R6A and R11A mutations in the ParA_fusNt28ParB construct appear to have a dramatic effect on the ATPase activity of ParA. Next, these two residues were mutated to serines (R6S and R11S) to ensure that the resulting effects of these mutations were specific and not a result of introducing a more hydrophobic residue in place of a charged one resulting in a structural change rather than a mechanistic one. From the resulting measurements of P_i production, the R6S and R11S residues were equally effective in reducing the ATPase activity of ParA suggesting the effects were more than introducing a hydrophobic residue (Figure 5_5). Finally, to see if these sites were specific for arginines they were replaced with another basic residue to conserve charge (Arg to Lys). Again, the ATPase activity of ParA in the both mutations was abrogated (Figure 5_5). The

mutagenesis experiments suggest that the R6 and R11 sites are both playing a role in ParB's stimulatory activity. They also suggest that the identity of these sites prefer to be arginines. The alteration of either one resulted in the loss of ATPase stimulation suggesting they are not merely redundant arginines. From the alignment the R11 site is the most conserved. The R6 site may provide an essential contact for the recognition of the ParA binding site. The next step would be to screen the mutations to both the R6 and R11 within ParB for their ability to bind to ParA to see if either site is important to the recognition of the ParA binding site. This way the purpose for each site could be classified to be either important for recognition or for the mechanism leading to the stimulation of ATPase activity. As attempts have been made in the past to ascertain the affinity of an N-terminal 28 residue ParB peptide with little success the best chance may be to introduce the mutations in full length ParB and attempt to measure the ability of each to pull down ParA. Checking the R6 and R11 sites *in vivo* by introducing them into the P1 *par* cassette and assessing the mutations for their ability to stably maintain the P1 plasmid is also another essential experiment that would confirm the importance of the arginine residues.

5.3 Discussion

Par motor proteins play an active role in partitioning in their ATP-bound state. The type I Walker A motor proteins are weak ATPases whose activity is stimulated through a specific interaction with its partner ParB protein. The important ParA interaction domain in many type I ParB proteins has been mapped to its N-terminus. The ability to hydrolyze ATP in the type I *par* motor proteins is essential for the proper segregation of daughter plasmids, but its mechanistic significance is unknown. As some type I *par* motor proteins, including P1 ParA, have been observed to form filaments in an ATP dependent manner, it is possible that the hydrolysis of ATP and the bound-nucleotide conversion to ADP in some way modulates the stability of these filaments. In the type II *par* system ParM filaments are now thought to be composed of ADP-bound sub-units with the ATP to ADP conversion playing a critical role in the growth of the filament.

The critical domain responsible for the specificity of P1 ParB's interaction with P1 ParA has been mapped to the first 28 residues and the domain responsible for its stimulating activity was found to be within its N-terminal 68 residues. We used a series of ParA-ParB fusion constructs to investigate the nature of the interaction between P1 ParA and ParB that leads to the stimulation of ParA's ATPase activity. A comparison of ATPase function in the fusion construct with the "wt" ParAB*S* reconstituted system demonstrated that they had similar abilities to hydrolyze ATP that were also within range of published "wt" values. As the ParA-ParB fusion construct appeared to be an effective mimic of the

"wt" *par* system, it was used to study the nature of the ParA-ParB interaction. A comparison of the stimulatory activities of different segments of ParB within the fusion construct determined that both the specificity for and activity resulting from the ParA interaction reside in the N-terminal 28 residues of ParB. As the previous attempt to capture the ParA-ParB interaction using the ParA_fusNt28ParB construct in crystallization trials resulted in an apo ParA structure where the ParB fusion peptide was missing, the next best candidate for future crystallization trials is the ParA_fusNt121ParB construct (2^{nd} highest activity).

The N-terminal 28 residues of ParB were found to be sufficient to stimulate ParA's ATPase activity, so they were used as the basic construct to further investigate the underlying mechanism responsible for the stimulation. One of the prime candidate mechanisms to explain the *trans*-stimulation of a P-loop NTPase is the Arg-finger hypothesis. A sequence alignment of deviant Walker A *par* motor proteins identified several conserved charged residues in the first 15 amino acids of ParB's N-terminus. Two of these residues (Lys3 and Lys4) were found not to be important, as substituting either of these residues had no effect. Substitution at Arg11, a conserved arginine residue, impaired ParA's ATPase activity regardless of the residue type (non-polar, polar, or similar charge). The arginine at residue eleven remains a candidate as a possible "Arg-finger" as it fits the profile for what has been used to identify the motif in other systems. It contains three principal elements; one, it is located in a flexible region that would allow its interaction with another proteins binding pocket; two, it appears to be conserved

across a family of proteins (loosely across deviant Walker A); and three, the site is specific for arginine (even a R11K mutation could not retain any activity).

There is a second proximally located arginine, Arg6, which is not conserved, but appears to be just as sensitive to the corresponding mutations (non-polar, polar, or similar charge). Arg6 and Arg11 don't appear to be redundant residues as there is no compensation for the loss of either one. It could also be that one of the residues is important for the specificity of the ParA-ParB interaction, while the other has a stimulating function. They could also be involved in a different mechanism that leads to the stimulation of ParA's ATPase activity (either contributing to the hydrolysis of ATP or by facilitating the nucleotide exchange). The stimulating properties of another deviant Walker A protein MinD were found to be located on the N-terminus of MinE (its partner protein) at R21 or R30 (Ma et al., 2003). This is an example of how a structure of the ParA-ParB complex would be important to elucidating the underlying mechanism responsible for providing the stimulus of ParA's ATPase activity.

Apart from sorting out the mechanism underlying ParB's stimulation of ParA's ATPase activity, questions remain about the physiological relevance of ParA's ATPase activity. ParA is a weak ATPase. From the progress curves of ParA's hydrolysis of ATP (Figure 5_2B), it is clear the introduction of ParB results in a burst of activity. The "burst" represents a single turnover of the enzyme. The rate-limiting step in ParA's ATPase activity is the release of ADP making the ADP-bound form of ParA a longer lasting species. In the type II system, the resulting MinD filaments were found to be in the

ADP-bound state. It is possible that the P1 ParB is acting at the partition complex to convert ParA filaments to their ADP-bound form, a state more suitable for eventual depolymerization. Establishing the nucleotide-bound state of ParA's filaments will be critical to understanding the physiological role of ParA's ATPase activity.

5.4 Methods

Fusion Construct. Using the P1/P7 homology sequence scanning (HSS) genetic studies as a guide, we engineered a fusion construct composed of the full-length P1 ParA and the N-terminal 28 amino acids of ParB (attached to the C-terminus of ParA through a linker – GGSGGGSG)(Radnedge et al., 1998). The ParA_fusNt28ParB fusion construct was assembled using a PCR protocol pioneered for the construction of synthetic genes (Stemmer, W.P. *et al*., 1995). Briefly, oligonucleotides were designed in ~ 40 bp lengths to generate a codon optimized sequence that encoded the C-terminus of P1 ParA, a flexible linker, and the N-terminus of P1 ParB. Synthetic oligonucleotides with overlapping 20 bp sections (5′-GCGCGAATTGCCGCGGAAGA-3′ (PrTop1), 5′-TTTTGCTAAAGCGGTTTTTGACCG TATTGAATTTATCCGC-3′ (PrTop3), 5′-TCTAACGGTGGCTCCGGTGGTGGCTCCGGTC TAGACATGAGCAAGAAAA-3′ (PrTop51'), 5′-ACCGCCCGACGATTGGGCGCACCCTG AATCCGAGCATCCT-3′ (PrTop712'), 5′-GAGCGGATTTGATAGTTCTAGCGCCTCTGGCTGAG GATCCCG-3′ (PrTop912'), 5′-CAAAAACCGCTTTAGCAAAATCTTCCGCGGCAATTCGCGC-3′ (PrBttm2), 5′-CCACCGGAGCCACCGTTAGAGCGGATAAATTCAATACGGT-3′ (PrBttm4), 5′-GCGCCCAATCGTCGGGCGGTTTTTCTTGCTCATGTCTAGACCGGAGCCA-3′ (PrBttm61'), 5′-TAGAACTATCAAATC CGCTCAGGATGCTCGGATTCAGGGT-3′ (PrBttm812'), 5′-CGGGATCCTCAGCCAGAGGCGC-3′ (PrBttm1012')) were mixed to generate a template molecule (~ 200 bp in length) using a PCR protocol featuring a high number of reaction cycles (55). The resulting product was diluted and the template molecule was amplified using flanking oligonucleotides (PrTop1 and PrBttm1012' - containing SacII and BamHI

restriction sites) using a traditional PCR protocol. The resulting amplified product corresponding to the fusion segment was digested using restriction endonucleases SacII and BamHI (NEB) and ligated in-frame to the sequence encoding the N-terminus of P1 ParA (via a SacII site). The fusion construct was inserted into a pET15b (Novagen) plasmid using NdeI and BamHI sites. The resulting construct was sequenced to insure fidelity using the Sanger method (Applied Biosystems instrument) at the Vollum Institute core facility.

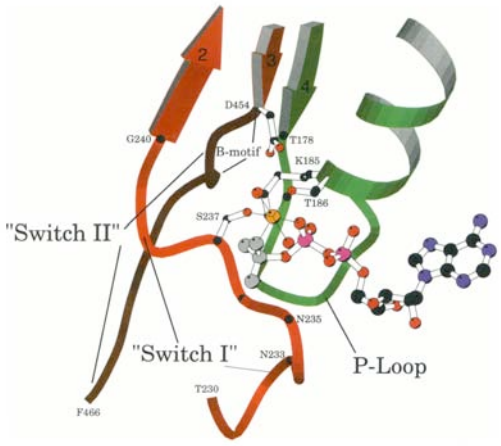
The ParA_fusNt40ParB construct was made as described above substituting the following oligonucleotides: 5′-GAGCGGATTTGATAGTTCTAGCGCCTCTGGCGATCGAGTC-3′ (PrTop92'40), 5′-GAGCAGGTATTCAAGTTATCAACCGGTTGAGGATCCCG-3′ (PrTop112'40), 5′-GATAACTTGAATACCTGCTCGACTCGATCGCCAGAGGCGC-3′ (PrBttm102'40), 5′-CGGGATCCTCAACCGGTT-3′ (PrBttm122'40). A silent AgeI restriction site was introduced into the ParA_fusNt40ParB construct at Thr39 of ParB. This site facilitated the inclusion of the full ParB gene in-frame using an amplified PCR product made with flanking oligonucleotides introducing AgeI and BamHI sites. The ParA_fusNt69ParB (5′-CAACGGGCGTTAACAGGCATCTC-3′), ParA_fusNt78ParB (5′-GCCAAAATCATAAAAAGTATCCG-3′), ParA_fusNt101ParB (5′-GACGGGCTACATAAAAATTGAAATTTTGG-3′), and ParA_fusNt121ParB (5′-GAGAACGTATAATTGCGGGTTTTAG-3′) fusion constructs were subsequently generated by introducing a termination site after the targeted residue using a QuikChange (Stratagene) protocol. The resulting constructs were sequenced to insure correct composition using the Sanger method (Applied Biosystems instrument) at the MD Anderson core facility. The P1 ParA constructs were expressed and purified as described in Chapter 3 methods.

ATPase assay. The ATPase assay is a coupled reaction linking the production of something difficult to detect, P_i, with something that is more amenable – the conversion of methylthioguanosine (MESG) ribonucleoside to free MESG, which has an associated absorbance shift from 335 nm to 360 nm (Webb, 1992). 5 μM ParA or ParA fusion construct (500 pmol) was used in a 100 μL reaction with 0.5 mM ATP. ParB (500 pmol) and *parS*small (1:1) were included with ParA to reconstitute the ParAB*S* system. The assay was used either to monitor the generation of P_i continuously (in a cuvette) in real time at 22 ºC or at a single time-point after a 30 minute incubation period at 30 ºC plus an additional 10 min. spin at 13,000xg. A standard curve relating absorbance values (A_{360}) to P_i (nmol) was generated using known concentrations of KH_2PO_4. A_{360} values resulting from the reactions were then converted to P_i (nmol) for comparison with published values or between mutant constructs.

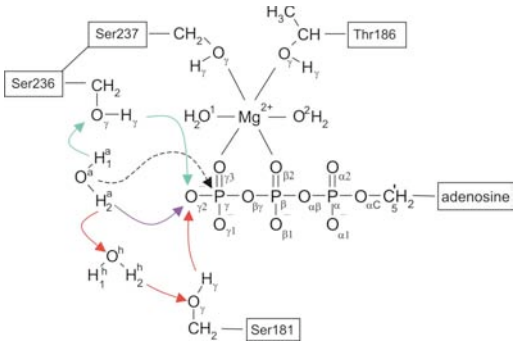
Mutagenesis. Residues on the N-terminus of P1 ParB have been identified as being potentially important for its role in the stimulation of P1 ParA's ATPase activity. Targeted charged residues (Lys3, Lys4, Arg6 and Arg11) were mutated to either neutralize or alter their potential contributions. The mutagenesis of these sites (to alanine, serine or lysine residues) was accomplished using a QuikChange (Stratagene) strategy using the following oligonucleotides: 5′-GGTGGCTCCATGAGCGCGAA AAACCGCCCGACG-3′ (K3A), 5′-GCTCCATGAGCAAGGCGAACCGCCCGACG-3′ (K4A), 5′-GCAAG AAAAACGCGCCGACGATTGGG-3′ (R6A), 5′-GCAAGAAAAACAGCCCGACGATTGGG-3′ (R6S), 5′-GCAAGAAAAACAAACCGACGATTGGG-3′ (R6K), 5′-CCGACGATTGGGGCGACCC TGAATCCG-3′ (R11A), 5′-CCGACGATTGGGAGCACCCTGAATCCG-3′ (R11S), 5′-CCGACGA TTGGGAAAACCCT GAATCCG-3′ (R11K). The resulting constructs were sequenced to insure

fidelity using the Sanger method (Applied Biosystems instrument) at the MD Anderson sequencing core facility.

A



B



C

***Par* type Ia**	*P1* ParA	LKGGVSKTVSTVS
	P7 ParA	LKGGVSKTVSTVT
	F SopA	HKGGVYKTSVSVH
***Par* type Ib**	Soj	QKGGVGKTTTSVN
Other	MinD	GKGGVGKTTTTSN
	NifH	GKGGIGKSTTSQN
		***: *: :

D

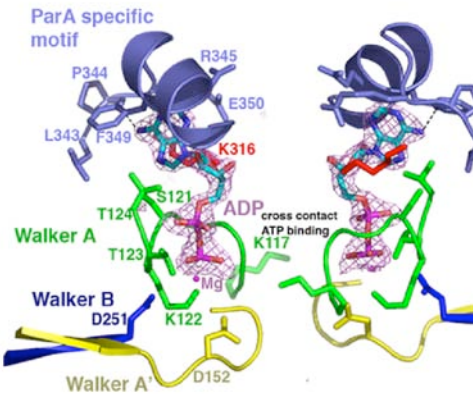
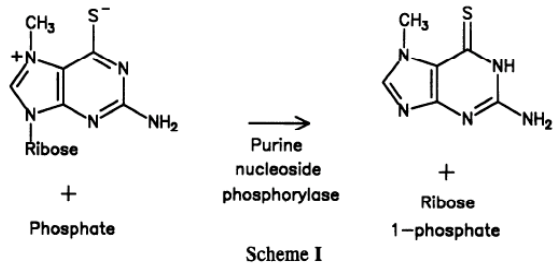


Figure 5_1: ATP Hydrolysis in Walker A P Loop Proteins. (A) Interaction scheme of a P-loop ATPase binding pocket (myosin)(Rayment et al., 1996). (B) Proposed scheme of hydrolysis mechanism in a P-loop ATPase (myosin). (C) Sequence alignment of the P-loop of deviant Walker A proteins. (D) Nucleotide binding pocket of P1 ParA in the ADP-bound state.

A



Scheme I

B

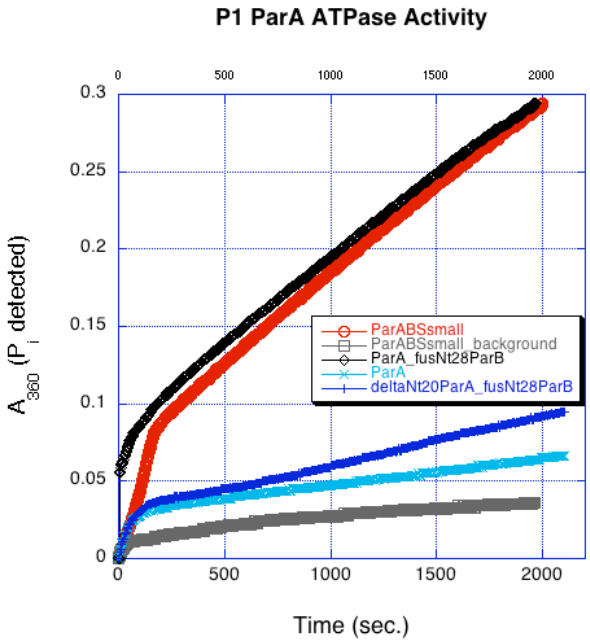
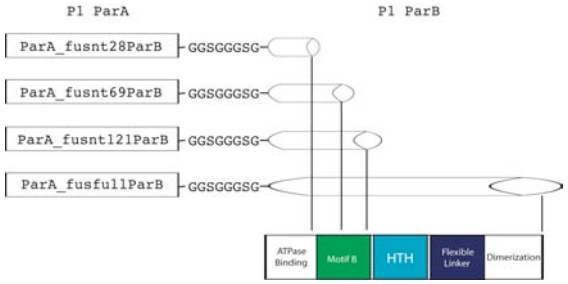
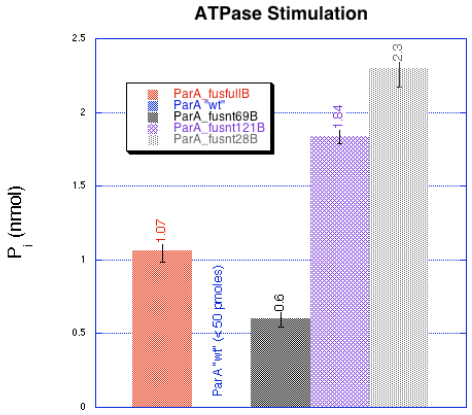


Figure 5_2: ATPase Activity of P1 ParA and P1 ParA_NtParB Fusion Constructs. (A) Reaction scheme of P_i detection assay using substrate MESG in a coupled reaction. (B) ATPase activity as measured by the detection of product P_i using a coupled reaction (Absorbance change at 360 nm) as shown in A. A comparison of ATPase activity of "wt" ParA, a reconstituted P1 *par* system (ParAB*S*small), a fusion construct (ParA_fusNt28ParB) and an N-terminal ParA deletion mutant fusion construct (Δ20ParA_fusNt28ParB).

A



C



B

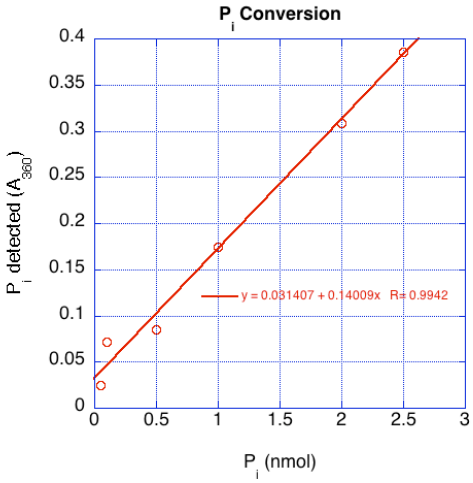


Figure 5_3: Identifying the ParB Activation Domain. (A) Schematic of the different fusion constructs. (B) Standard curve for the conversion of absorbance values to molarity of P_i. (C) Measured ATPase activity comparing the different sized fusion constructs (ParA "wt" < 50 pmoles P_i.).

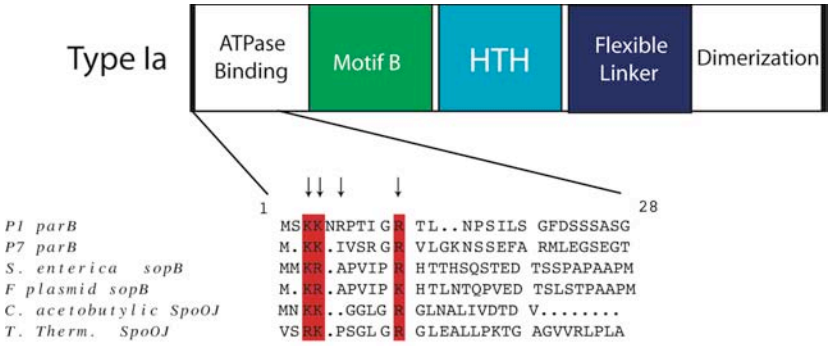


Figure 5_4: Sequence Alignment of the N-terminus of Deviant Walker A *par* Proteins. Conserved residues highlighted in red and include three strongly basic residues. The Arg (R11 in P1 ParB) was shown to be important in the stimulation of Soj (partner to listed *T.Therm* SpoOJ's).

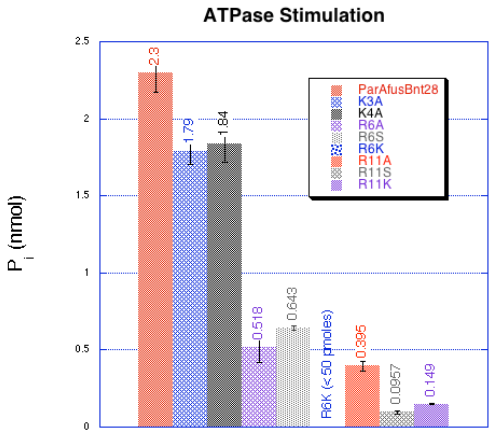


Figure 5_5: Identifying the Arg-Finger in ParB N-Terminus. Four sites containing charged residues in the N-terminus of ParB were assessed for their contribution to the stimulatory interaction with ParA. ATPase activity - P_i detected (nmoles) labeled (R6K mutant < 50 p moles P_i.).

Chapter 6

The P7 *parOP* Site

6.1 Introduction

Auto-repression of *par* operons has been demonstrated to be an essential element of a properly functioning partition system as the over-expression of either protein element in the *par* operon of many systems results in the breakdown of proper plasmid segregation. The characterization of the P1 *parOP* has demonstrated that P1 ParA protects a roughly 150 base pair (bp) (-80 to +70) segment of the promoter in an ADP (or ATP) dependent manner (Figure 6_1A&Di)(Davey and Funnell, 1994). A further truncation to a segment of the *parOP* starting at -44 functioned in a manner similar to the full site suggesting that the important recognition segment is smaller than the 150 bp fully protected site (Figure 6_1Dii)(Hayes et al., 1993). Within the site protected by P1 ParA is a ~ 40 bp site downstream of its promoter consisting of a 21 bp pseudo inverted repeat (Figure 6_1A&D)(Davey and Funnell, 1994). The P7 ParA protein also protects a large region of its own *parOP* (~ 80 bp) in a similar manner, but the hypothesized operator sites lie upstream of its promoter (Figure 6_1B)(Hayes et al., 1994). An alignment of the P1 and P7 *parOP* sites reveal that they are both very AT-rich (~ 70%), but beyond that there is little sequence similarity. An analysis of the P7 *parOP* sequence found no obvious palindromic site similar to that found in the P1 *parOP*. Based upon loose sequence similarity two potential binding site mechanisms have been proposed (Figure 6_1C)(Hayes et al., 1994). First, there is a pseudo palindromic sequence upstream of the P7 promoter site that contains 9 of the 16 symmetric bases required for perfect symmetry. Second, there are a series of four loosely similar 9 bp repeats that span the protected P7

parOP site. Identifying the site that P7 ParA uses to recognize its *parOP* could provide a better understanding of the mechanism P7 ParA uses to accomplish the auto repression of its operon.

As a goal has been to understand the structural role of type Ia motor proteins in the auto-repression of their operons, efforts to study the P1 operator site in more detail have been partially successful. Oligonucleotides designed around the 40 bp pseudo-palindromic site have been used to demonstrate specific binding (in 100-150 mM NaCl) in fluorescence anisotropy (FA) experiments (for example Figure 4_3A). Binding affinity was determined to be significantly higher (10x) with an extended site (99 bp) as compared to sites limited to the pseudo-inverted repeat (37-42 bp)(Figure 6_1D). P1 ParA was found to bind these oligonucleotides in an ATP-dependent manner. Affinity for the oligonucleotides in the presence of ADP were unsuccessful as the higher concentrations of nucleotide that were required for binding interfered with the fluorescence measurements. Crystallization trials using oligonucleotides based upon the full pseudo-inverted repeat (34-42 bp) as well as the half site (16-22 bp) resulted in crystals that diffracted poorly.

As P7 ParA had been shown to protect a similarly large portion of its promoter region when bound to nucleotide, the P7 system could also be used to study the mechanistic details of ParA DNA binding interaction that is the basis for autorepression of the type Ia *par* operon. Central to these efforts was a desire to obtain a ParA-DNA complex structure that would be invaluable in contributing to the understanding of the binding mechanism used by type Ia ParA proteins to repress their own operons. To this end, P7 ParA could be used to expand the possibilities of determining a DNA-bound structure.

The difficulty in using the P7 system is the uncertainty in the mechanism of binding makes the design of oligonucleotides for P7 ParA-DNA complex crystallography experiments difficult.

6.2.1 Approach Used to Find the P7 Operator Site

In order to identify candidate sites for the P7 ParA-DNA complex crystallography experiments, it is necessary to have a rudimentary understanding of the mechanism of binding. The P7 binding site is only known in terms of the segment of the *parOP* that ParA protects in DNase I experiments. Two hypothesized sites have been proposed based upon an examination of the sequence to explain the protection. One, based loosely on the P1 site, is a pseudo-inverted repeat covering a ~ 36 bp segment in the upstream region of the protected site (Figure 6_1C). The second, is composed of a set of four loosely similar nine bp sites that are spaced 6-10 bp apart spanning a 61 bp segment nested in the middle of the protected site (Figure 6_1C). There is overlap between the predicted recognition sites, which adds to the difficulty in distinguishing the sites.

To begin to address the differences in the possible binding mechanisms, the protected site was fractionated using separate oligonucleotide sets, so that the binding affinity of each portion of the site could be determined separately. The binding affinity of P7 ParA for

each oligonucleotide was determined in the presence of nucleotide (ADP and ATP) using FA to establish the boundaries of the high affinity binding site. With the boundaries established for what is required for high affinity binding it should then be possible to distinguish between the two proposed biniding sites. Understanding the nature of the site's archetecture will allow the design of oligonucleotides for co-crystallization experiments as well as establishing a testable model that could be used to further define the binding mechanism and mode of protection occurring at the P7 *parOP* site.

6.2.2 P7 Operator Site identification.

Analysis of the type Ia *parOP* regions to date have relied heavily on DNase I protection and gel shift assays using larger DNA segments (Davey and Funnell, 1997). Smaller DNA segements can be utilized in fluorescence anisotropy (FA) experiments to examinine the site in more detail. To establish the experimental model of using a fluorescently tagged oligonucleotide to ascertain the affinity of P7 ParA for its cognate DNA, an oligonucleotide (whole site - 75 mer) corresponding to most of the protected P7 *parOP* site was used to determine the specificity of P7 ParA binding in the apo, ATP-, and ADP-bound states (Figure 6_2A). This oligonucleotide also contains the proposed sites of both binding mechanisms. The results demonstrated that P7 ParA binds the site with high affinity in its ATP- and ADP-bound states (~ 550 and 300 nM, respectively)

(Figure 6_2B). ParA's affinity for this site is similar to the affinity estimated from DNase I protection assays using the larger DNA segment. P7 ParA's affinity for the whole site oligonucleotide was greater than 20x lower (~ 12 μM) than either of the nucleotide bound states. This is consistent with the results from DNase I protection assays and with the finding in Chapter 4 that binding of nucleotide (ADP) stabilizes ParA's conformation that is preferred for binding DNA. As ΔmP is related to the hydrated volume, or size of the complex, the higher mP (and a larger non-specific component) observed in the ATP-bound state may be indicative of additional protein-protein interactions. Establishing the stoichiometry of the complex would provide the insight into the mechanism used by ParA to bind to the P7 *parOP*. The characteristics of P7 binding to the whole site oligonucleotide illustrates that it is a valid model system to further investigate the binding mechanism of P7 ParA to its operator site.

An additional characteristic of type Ia ParA proteins is that they are specific for their own *parOP* site. To demonstrate that P7 ParA DNA binding was specific for an oligonucleotide composed of its own *parOP*, the affinity of P7 ParA for an oligonucleotide composed of the P1 *parOP* (OP99 - 99 mer) was determined in its ATP- and ADP-bound states. The affinity of P7 ParA for this site was determined to be ~ 6 μM in its ATP-bound state and little binding was detected for P7 ParA in its ADP-bound state (Figure 6_2B). This shows that P7 ParA binding is specific for its own *parOP* oligonucleotide as P7 has little affinity for the P1 *parOP* site. The fact that there was little binding to the P1 OP99 oligo, which is longer than the P7 whole site oligo, also demonstrates the interaction is likely base specific.

After establishing that the behavior of the P7 *parOP* oligonucleotide model system was similar to "wt", additional oligonucleotide sets could be used to probe the *parOP* site further to identify the core binding site. The P7 *parOP* site (whole site) was first divided into thirds creating three ~ 40 bp oligonucleotide sets. The oligonucleotide OP_1 (- 75 to – 35) corresponds to the left arm of the protected site and contains both the pseudo-inverted repeat site and sites 1 and 2 of the 9 bp repeat motif (Figure 6_2A). The oligonucleotide OP_2 (-35 to + 5) corresponds to the right arm and contains sites 3 and 4 of the 9 bp repeat motif. The oligonucleotide OP_3 (- 55 to - 15) corresponds to the middle of the protected site and contains sites 2 and 3 of the 9 bp repeat motif. P7 ParA (ATP- and ADP-bound states) bound to OP_1 with low μM affinity and bound to OP_2 with even lower affinity (data not shown). P7 ParA bound to OP_3 with the highest affinity (Kd ~ 1 μM) in its ADP-bound state (Figure 6_2C). In order to look at P7 ParA's affinity for the pseudo-inverted repeat again (in addition to OP_1) a second slightly shorter oligonucleotide (36 bp) comprised of the core inverted repeat site alone was used. P7 ParA (ADP-bound state) bound to the more refined pseudo-inverted repeat site with low μM affinity (Kd ~ 2.5 μM) similar to the oligonucleotide OP_1. The results show that the fractionation of the P7 *parOP* site into smaller oligonucleotides (thirds) leads to the loss of the high affinity interaction. Additionally, the pseudo-inverted repeat site proposed as one possible binding site used by P7 in the autorepression of its operon is not a high affinity interaction.

As the fractioned *parOP* segment with the highest affinity (OP_3) was in the middle of the protected region it was retained as a core element in the next series of

oligonucleotides. First, the OP_3 site was extended symmetrically to see if the inclusion of additional bases also increased the affinity of P7 ParA for the site. P7 ParA (ADP-bound state) bound to the new oligo (OP_middle) with slightly higher affinity (Kd ~ 800 nM) than the previous OP_3. By expanding the OP_3 site asymmetrically two additional footprints could be tested that included the 9 bp repeat sites 1, 2, and 3 in one oligonucleotide (OP_5) and sites 2, 3 and 4 in another oligonucleotide (OP_6) (Figure 6_2A). P7 ParA (ADP-bound state) bound to the OP_5 oligo with "wt"-like affinity (Kd ~ 100 nM) where P7 affinity for the OP_6 oligo similar to OP_middle (repeat sites 2 and 3)(Kd ~ 800 nM)(Figure 6_2C). The elimination of the 4th site does not appear to alter P7 ParA's affinity for the oligonucleotide as its inclusion or exclusion does not lead to a change in affinity. Site 4 is the most dissimilar of the four repeat sites as it is without a guanine base at position 6 which all the other sites have. The highest affinity appears to come from the inclusion of repeat sites 1, 2 and 3.

6.2.3 P7 Operator Site Refined

The search for the highest affinity core site within the P7 *parOP* was narrowed to an asymmetric composition of pseudo repeats (sites 1, 2 and 3) using different oligonucleotides that represented fractions of the full-protected site. The OP_5 oligonucleotide represents the smallest site found to have similar binding characteristics

to the full site. The 9 bp pseudo repeats had been hypothesized as one mechanism that explained the binding characteristics of P7 ParA for its *parOP* site. To examine the elements that may convey P7 ParA's specificity for its *parOP* site, the pseudo repeat sequences were randomized one site at a time. A monte-carlo randomized sequence of the bases contained within each site was used as the basis to test each mini site's contribution to the binding specificity of the P7 *parOP* site. The resulting pseudo-repeat sequences were altered retaining 0-3 bases out of each nine bp sequence (Figure 6_3A). P7 ParA's affinity for the altered site OP_5 (site 1′, 2′ and 3′) oligonucleotides was reduced 4-6 fold (Kd's ~ 1 μM) in its ADP-bound state (Figure 6_3B). The similar reduction in binding affinity for each of the randomized sites suggests that each pseudo-repeat site makes roughly the same contribution to the overall site specificity. P7 ParA's affinity (ADP-bound state) for the 9 bp pseudo-repeat site itself was also tested (site 3) and was found to be very low (Kd ~ 12 μM) (Figure 6_3A&B).

In an effort to establish the boundaries for the core P7 *parOP* site the OP_5 oligonucleotide was shortened by roughly 10 bp to the start of the pseudo site 1 eliminating the extended 5′ region (Figure 6_3A). ParA's affinity (ADP-bound state) for the new 45 bp site (OP_5b) was reduced roughly 3 fold (Kd ~ 600 nM)(Figure 6_3B). It is likely that a 2-3 bp 5′ extension to pseudo site 1 would rectify the reduced affinity resulting from the shortened OP_5b oligonucleotide.

As the highest affinity core-binding site is likely to resemble the OP_5b oligonucleotide (containing pseudo repeat sites 1, 2 and 3), the OP_5b site (and OP_3 site) provided the

basis for designing oligonucleotides for use in P7 ParA(ADP-bound)-DNA co-crystallization trials. Screening with these oligonucleotides resulted in crystals, but like the co-crystallization trials before they had poor diffracting characteristics.

6.3 Discussion

Both P1 and P7 ParA protect a large region of their respective promoter regions. The operator site of P1 ParA has been mapped to a roughly 42 bp pseudo-inverted repeat site. The operator site for P7 ParA has been undefined with several imperfect repeat sites being proposed as core binding sites. Efforts to refine the P7 operator site in this study resulted in the identification of a roughly 43 bp site containing three 9 bp pseudo-repeats. P7 ParA was found to bind to the site with high affinity in its ADP-bound state.

Efforts to ascertain structural information from co-crystallization studies with P7 ParA were unsuccessful. A P1 ParA-DNA bound model has been made based upon the P1 ParA structure in its ADP-bound state (Figure 6_4B). The model features the incorporation of a basic region in ParA that is proposed to facilitate the necessary bending of the DNA site required to incorporate ParA's HTH motifs. Both the P1 and P7 *parOP* regions are AT rich and an alignment of the protected sites show they share a concentrated region of AT bases near the center of their sites (Figure 6_4A). AT rich

base regions, while being insufficient alone, have been known to facilitate the bending of DNA. Acting in conjunction with site specific AT/GC transitions these regions may provide the mechanism to initiate the protection of the *parOP*. The internal DNA recognition sequence (AT/GC transition) flanking the AT rich region would have to be tested empirically to determine its importance to ParA's binding. ParA binding to its operator site is reported to lead to the additional protection of the larger promoter site most likely through the formation of a larger nucleoprotein structure.

As the functional complex leading to the autorepression of the *par* operon likely involves a larger nucleoprotein structure with multiple copies of ParA involved in the binding event, the best way to evaluate the structure may be through electron microscopy (EM) studies. Another possible method of studying the larger nucleoprotein superstructure is to evaluate ParA's effectiveness in bending the DNA around a core of composed of ParA molecules. ParA's effectiveness could be judged in terms of affinity for different length DNA molecules as well as the stoichiometry of the binding interaction. A qualitative measure of the bending of the different DNA segments is to utilize fluorescence resonance energy transfer (FRET) to evaluate the convergence of the labeled ends of the DNA in a ParA dependent manner (Figure 6_4C). These tools should provide the needed additional information that can be used in conjunction with the ParA-DNA model in order to understand how ParA carries out its functional role in the autorepression of its own operon.

6.4 Methods

P7 ParA. The P7 ParA protein was purchased as a codon optimized synthetic gene (GenScript) inserted into a pET15b vector (Novagen). The P7 ParA protein was expressed and purified in a manner analogous to P1 ParA (described in Chapter 3 methods).

Fluorescence anisotropy (FA) measurements. Binding affinities of P7 ParA protein for the P1 and P7 *parOP* sites were performed using a Beacon Red Ranger fluoresence anisotropy (FA) system. Two nanomolar fluorescein labeled 5′ oligonucleotides (Figures 6_1D, 6_2A and 6_3A) were used as probes to sample different segments of the *parOP* sites to establish the borders of the core ParA binding site. Protein was titrated into 150 μL of 25 mM Tris-Cl (pH 7.5), 150-200 mM KCl, and 2 mM DTT at 25 ^{o}C (nucleotide concentration was 1 mM when included). Resulting binding isotherms were generated in Kaleidagraph using non-linear least squares analysis to fit the data to an equation describing a simple binding model: $mP = mP_i + \Delta mP^*[P]/(K_d + [P])$ (Lundblad et al., 1996).

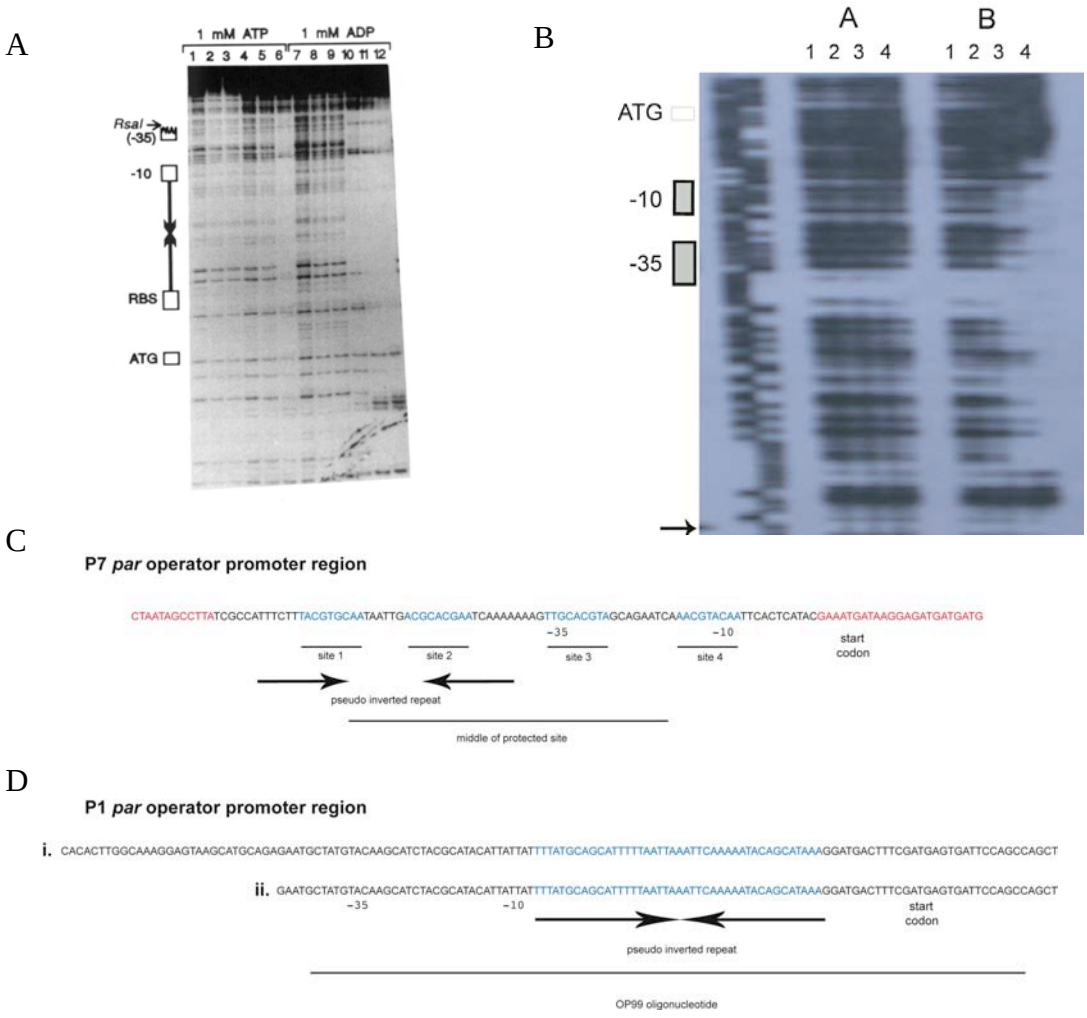
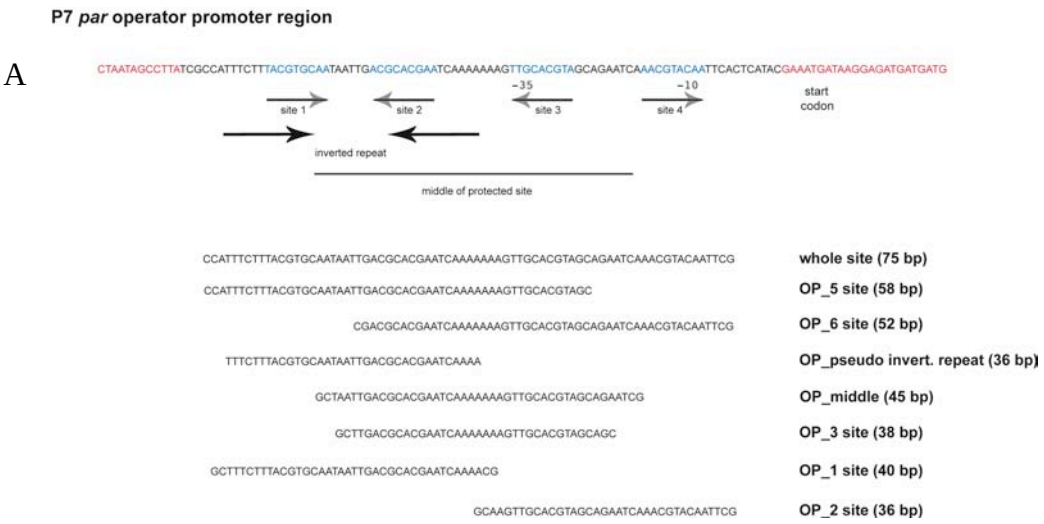


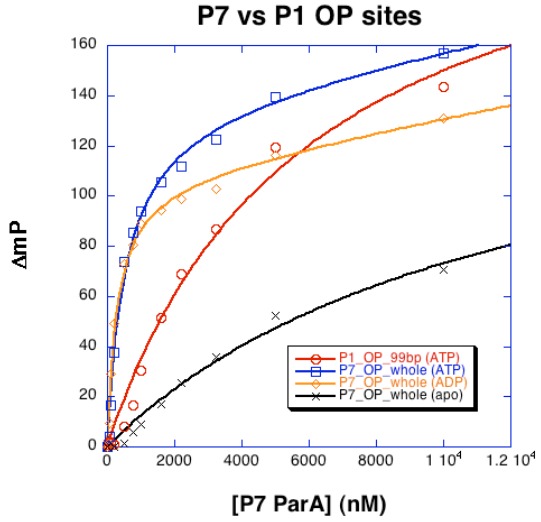
Figure 6_1: P1 and P7 *parOP* Sites. (A) DNase I protection assay of the P1 *parOP* site. P1 ParA protects a ~ 150 bp site in a nucleotide dependent manner. (Davey and Funnell, 1994) (B) DNase I protection assay of the P7 *parOP* site. P7 ParA protects a ~ 80 bp site also in a nucleotide dependent manner. (Hayes et al., 1993) (C) Sequence of the P7 *parOP* site. Black type designates the boundary of the protection. Proposed sites are denoted: pseudo inverted repeat (arrows) and pseudo 9 bp repeats (blue text). (D) Sequence of the P1 *parOP* site. i.) the full site and ii.) the truncated site used as an artificial construct. The pseudo inverted repeat site is denoted (arrows).

A

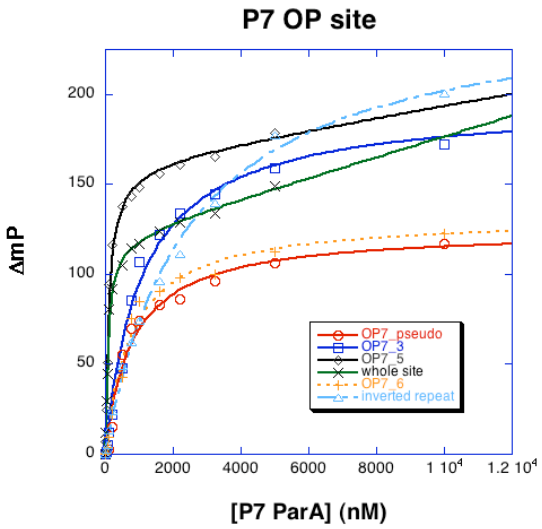


B

OP site	*Kd(nM)*	*R*
P7 OP (apo)	1.2×10^3	0.992
P7 OP (ADP)	300.0	0.996
P7 OP (ATP)	560.0	0.996
P1 OP (ATP)	6.0×10^3	0.992



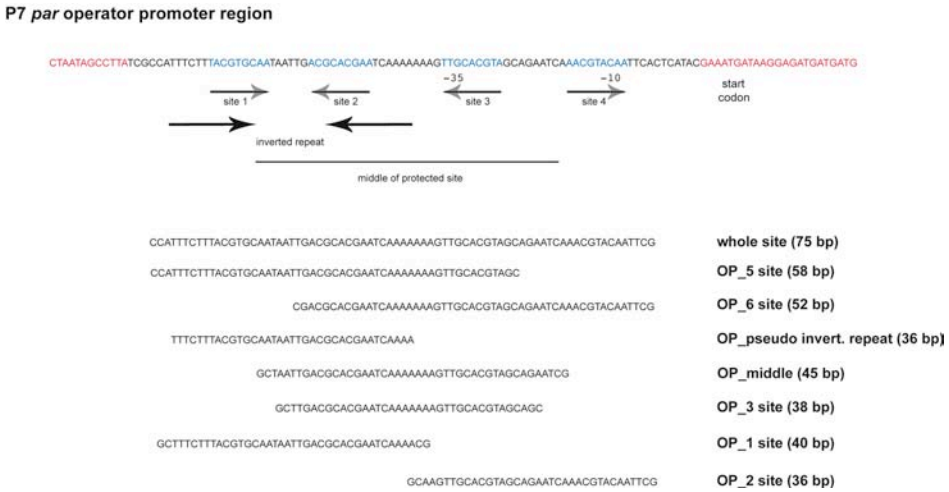
C



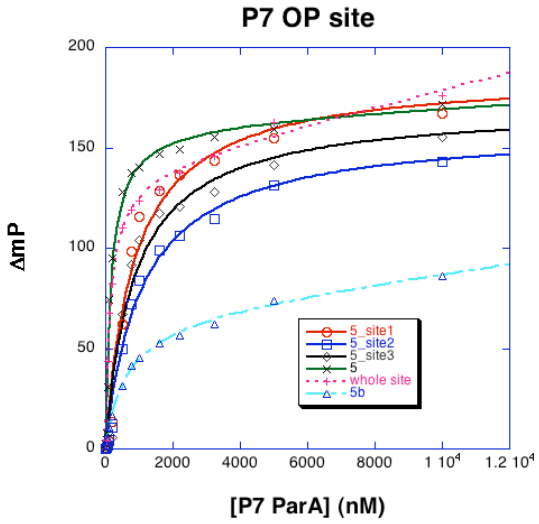
OP site	*Kd(nM)*	*R*
whole	80.0	0.997
5	100.0	0.996
6	780.0	0.992
3	1.1×10^3	0.995
pseudo	810.0	0.989
inverted	2.5×10^3	0.998

Figure 6_2: Dissection of the P7 *parOP* Site. (A) Sequences of the oligonucleotides used to find the core-binding site. (B) Representative binding isotherms of P7 ParA binding to the P7 whole site in its apo, ATP-, and ADP-bound states. Also included is P7 ParA binding to the P1 OP99 oligonucleotide in its ATP-bound state. (C) Representative binding isotherms of P7 ParA binding to different oligonucleotides representing the different segments of the P7 *parOP* site.

A



B



OP site	*Kd(nM)*	*R*
whole	220.0	0.9875
5	180.0	0.9939
5_1	870.0	0.9875
5_2	1.1 x 10^3	0.9931
5_3	870.0	0.9829
5b	620.0	0.9989
Site 3	1.2 x 10^4	0.9952

Figure 6_3: Further Dissection of the P7 *parOP* Site. (A) Sequences of the oligonucleotides used to find the core-binding site. (B) Representative binding isotherms of P7 ParA binding to altered pseudo-repeat *parOP* sites.

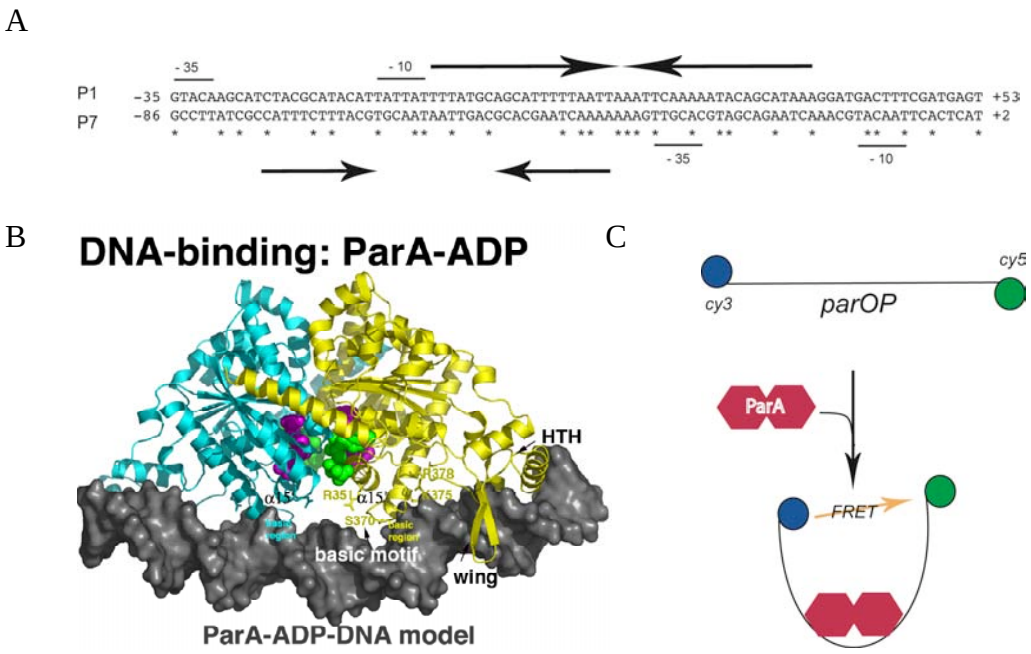


Figure 6_4: Model of type Ia Motor Protein Binding to the *parOP* Region. (A.) Alignment of the center of the P1 and P7 *parOP* sites protected by their respective motor proteins. An AT-rich sequence is located in the center of both sites. (B) Model of P1 ADP-bound ParA binding to a 40 bp sequence of DNA. (C) Method for detecting a ParA dependent bending of a *parOP* sequence of DNA.

Chapter 7

Fluorescence Studies on the N-terminal Helix of P1 ParA

7.1 Introduction

Deviant Walker A ATPases are believed to undergo an ATP-dependent dimerization step. The transition to its dimer state also coincides with the functional activation of the Walker A protein. The binding of nucleotide constitutes a molecular switch mechanism that modulates the activity of the Walker A protein based upon the properties of the bound-nucleotide. The characteristics of the bound-nucleotide lead to the stabilization of distinct conformational states that are optimal for a particular function. In the P1 *par* system the motor protein ParA has two distinct roles (functioning in partitioning and autorepression), which are mediated by the binding of two distinct adenine nucleotides, ATP and ADP respectively.

In previous work it was established that P1 ParA is a dimer in its nucleotide-bound state, ADP (Davey and Funnell, 1994). Further structural and biochemical studies on P1 ParA outlined earlier in this text (Chapter 3) found evidence that ParA can also exist as a dimer in its apo state. Thus, P1 ParA is a dimer in all forms examined suggesting that an underlying molecular switch mechanism is more nuanced than a simple monomer to dimer transition. The important contacts for dimer stability were localized to an interaction involving the N-terminal α-helix (α1) that is a unique extension found in type Ia *par* motor proteins (Chapter 3)(Figure 7_1A). Mutations (Δ1-20, A11S and A14S) in the N-terminus were found to disrupt the dimer state. The N-terminal 20 residues had

previously been shown to be important for the specific function of P1 ParA as a swap of amino acids from P7 ParA abrogated P1 ParA function (Radnedge et al., 1998)

MinD is a deviant Walker A protein (ParA homolog) and functions as a regulator of polarity within the bacterial cell. MinD is believed to dimerize in an ATP-dependent manner (Figure 7_1B)(Lutkenhaus and Sundaramoorthy, 2003). While MinD is in its ATP-bound state it serves to localize a partner protein MinC to the cell membrane covering an area from the cell pole to a point adjacent to the cell mid-point. An amphipathic C-terminal helix (α9) in MinD serves to anchor it to the membrane (in the ATP-bound state) (Hu and Lutkenhaus, 2003). Once MinD is localized to the membrane it forms polymers. Time-lapse fluorescence microscopy studies revealed that MinD polymers move in spiral formations from the cell pole to the cell mid-point. The mean concentration of MinD (and MinC) is established from the cell pole to the cell mid-point by the MinD polymer's dwell time at any point in between. MinE, localized to the cell mid-point, interacts with MinD and stimulates the hydrolysis of ATP in a manner similar to ParB's interaction with ParA (Chapter 5). The hydrolysis of ATP in MinD leads to the release of MinC and ultimately the depolymerization of MinD. In this system the concentration of MinE is critical to the localization of MinD at the cell membrane.

Type Ia motor proteins (SopA) have also been observed to form spiral-like structures in time-lapse fluorescence microscopy experiments, which may be indicative of an association with the cell membrane. An examination of the N-terminus of the P1 ParA alpha I helix shows that it, like α9 in MinD, is an amphipathic helix with two faces (one

that is hydrophobic and one that is more hydrophilic) (Figure 7_1C). It is possible that the N-terminal helix could also serve a similar purpose in P1 ParA. The following fluorescence-based studies were then initiated with the idea that α1 in P1 ParA represented a strong candidate as an amphipathic α-helix that might serve as a membrane anchor.

For the N-terminus to function as a link to the membrane in a nucleotide dependent manner (ATP) ParA must undergo a conformational change that would reposition the dimer contacts allowing the α1 helix to establish new associations with its hydrophobic face. One way to detect possible conformational changes in solution is with a fluorescence-based technique that allows one to monitor changes in a protein's structure at a specific location. A fluorescence-based technique has been developed that allows one to qualitatively judge the proximity of an introduced fluorescent label and any neighboring tryptophan residue (Mansoor and Farrens, 2004). The fluorescent label bimane contains inherent properties that make its fluorescence sensitive to interactions with quenching agents such as a tryptophan residue. The bimane label may be attached to specific locations within a protein (engineered cysteine residues) using sulfhydryl chemistry (Figure 7_2A). By pairing the introduction of a bimane fluorescent label with an engineered neighboring tryptophan residue, the qualitative proximity of the pair can be determined in a steady-state fluorescence experiment (Figure 7_2B). The proximity can be determined for the resting protein as well as in the protein in its active state (Figure 7_2C). Any apparent changes that may result within the targeted area would provide

clues to the nature of the underlying conformational change that is taking place as a protein undergoes a transition from its inactive to active states.

7.2.1 P1 ParA's Association with the Cell Membrane.

The deviant Walker A protein MinD localizes to the membrane in an ATP-dependent manner. The binding of ATP results in a nucleotide dependent dimerization step that may be stabilized by an interaction with the membrane (Zhou and Lutkenhaus, 2004). MinD's association with the membrane is dependent upon the formation of an amphipathic C-terminal helix, which anchors the protein to the lipid bi-layer. There appears to be some stabilization of the membrane association in the ADP-bound state (Figure 7_3A). The ATP-bound state must stabilize a conformation that allows MinD to establish this interaction in preference to the ADP-bound state.

To investigate whether P1 ParA also forms a nucleotide-dependent association with the membrane, ultra-centrifugation was used to study the interaction in conjunction with *E. coli* lipid extract vesicles used to mimic the cell membrane. If ParA forms an interaction with the lipid vesicles it should preferentially associate with the lipid pellet after a 100,000xg centrifugation step. In the presence of ATP ParA was preferentially found to

be in the pellet fraction (Figure 7_3B). The ADP-bound form of ParA also appears to form an association with the membrane to a lesser degree.

A candidate for a source of this interaction in ParA is the amphipathic N-terminal helix, α1. α1 forms the hydrophobic face that interacts with α14′-15′ of the dimer partner and could interact with the acyl chain of the lipid bi-layer similar to the C-terminal helix of MinD. A N-terminal deletion construct (Δ1-20) was used to assess the contribution of the N-terminus on ParA's association with the membrane. The P1 ParA(Δ1-20) construct was not able to associate with the membrane in the presence of ATP (Figure 7_3C). Titrating P1 ParA into a fixed limited amount of lipid vesicles demonstrates that full-length ParA is able to associate with the membrane in a specific manner (saturable association), while the Δ1-20 mutant does not associate to the same degree (Figure 7_3D). As the deletion mutant was found to break the apo ParA dimer (Chapter 3), it is possible that the inability or instability of the dimeric state rather than the absence of the N-terminal segment is responsible for the lack of association.

As the association with the membrane appears to be dependent upon an interaction with the N-terminus, we undertook fluorescence studies to determine if there are any nucleotide-dependent movements that would suggest that N-terminus could participate in the membrane association interaction.

7.2.2 Design of the Model Construct.

In order to create a platform for the examination of dynamic changes occurring in P1 ParA during nucleotide binding using fluorescence, a construct must be made that is absent of any native reactive cysteines. P1 ParA contains five cysteine residues (C220,C303, C305, C326 and C352) (Figure 7_4A). These cysteines were removed individually or in pairs (C220A, C303A/C305A, C326A, and C352A) to assess their reactivity to a fluorescence label PyMPO-maleimide. An examination of each of the mutant's reactivity showed that C220 in α11 was the only cysteine residue exposed to solvent that could react with the fluorescent label (Figure 7_4B). Next, to make sure the resulting reactive cysless construct was functional the C220 site was mutated to four different residues (Ala, Ser, Thr, and Val) to assess each mutant's impact on P1 ParA ATPase function. The purified mutants were tested for their ability to hydrolyze ATP in an absorbance-based assay that detects the generation of inorganic phosphate, P_i (Figure 7_4C). As they all had similar activity a comparison of the P1 and P7 sequences suggested that a valine residue was suited for the C220 site. The final reactive cysless construct P1 ParA(C220V) was used as a basis for future experiments.

In order to detect potential nucleotide-dependent movements in the N-terminal α1-helix, sites allowing for the introduction of fluorescent labels and partner tryptophan residues must be engineered into ParA. Using the P1 ParA crystal structure, the region around the N-terminal α1 helix-α14′/α15′ (′ denotes that it is contributed by the dimer partner)

interaction site was examined for potential sites for the introduction of single cysteine sites (fluorescence label incorporation) and proximal tryptophan residues (Figure 7_5A&B). As there is a link between bimane label–tryptophan pairs placing them strategically across from one another allows any movement in α1 to be represented as a change in fluorescence intensity. Two sites were selected on α14′ to introduce tryptophan residues (H282W and K286W) due to their proximity to the α1 helix. To pair with the tryptophan residues in α14′, a series of residues on α1 downstream of the important contact site (A11S and A14S) were selected to introduce single cysteine residues (N19C through V25C) to serve as sensors (bimane label) of α1 movement (Figure 7_5A&B). The end result is to match a cysteine mutation at each residue in α1 (N19C through V25C) with a tryptophan mutation at either H282 or K286 on α14′. Each set of mutant pairs can be used to assess movements in α1 in each of its nucleotide-bound states.

7.2.3 Characterization of the α1 Single Cysteine Sites in P1 ParA.

The series of cysteine mutants (N19C through V25C) were generated in a C220V/H282W background. The mutant constructs were expressed in *E. coli* Bl21(DE3) cells, purified using a Ni^{2+} affinity column and labeled with the fluorescent probe pyridyl bimane (methods). Labeling efficiency was estimated using a standard curve of bimane

label (methods) and are included in parentheses next to each mutant: N19C (40%), V20C (9%), L21C (37%), T22C (6%), E23C (70%), Q24C (55%), and V25C (44%)(free label was determined to be $\leq$ 5%). A quick examination of the labeling efficiency of each of the mutants generally reveals where they are located in the α1 helix with respect to the α14′ helix. The T22C mutant is located along the inner face of the α1 helix where it is less accessible and therefore it only has a 6% label incorporation whereas E23C located on the outer face has a 70% label incorporation. Each mutant was assayed for ATPase function to ensure it retained activity.

Each bimane label is attached to its mutant site through a disulfide linkage. This property provides a convenient way of internally normalizing the amount of quenching occurring at a site within a single sample. Measuring a sample and then reducing the label off the site with TCEP before recording a second measurement (label in solution) allows each mutant site to be assessed for fluorescence quenching, which is an indication of how close each residue is to a neighboring tryptophan residue (Figure 7_2B). A comparison of the fluorescence spectra of each ParA mutant in the apo state (red) with the spectra of the reduced label (blue) illustrates the range of quenching interactions occurring in the series of α1 mutants (Figure 7_6). An examination of the mutants individually shows the amount of quenching occurring at residue E23C-bimane is small indicating that it is located far from the tryptophan residue (H282W) located on α14′. Whereas the quenching occurring at residue V25C-bimane is much greater indicating that it is located nearer to the tryptophan residue (H282W) located on α14′. Comparing the overall pattern of quenching occurring at each α1 residue with the proximity of each residue to

H282W in the P1 ParA crystal structure shows that there is a strong correlation between the measured values and what would be predicted from the apo dimer crystal structure. The finding that the measurements are consistent with the apo dimer structure is another piece of evidence that suggests that ParA prefers the dimer state in solution. The close correlation between the measurements and the P1 ParA crystal structure supports the use of systematic fluorescent probe (cysteine mutant)-tryptophan pairings as a valid platform to examine nucleotide-dependent conformational changes.

7.2.4 Nucleotide Dependent Movement.

One of the outstanding questions on the function of P1 ParA is whether it might behave in a similar manner to MinD including being localized to the membrane in its ATP-bound state. One way this question could be addressed is by determining if a similar amphipathic helix in P1 ParA (α1) undergoes any nucleotide dependent changes that would indicate that it also could participate in anchoring the protein to the membrane. Using a fluorescence-based detection system (bimane-tryptophan pairs) that allows one to monitor changes that occur in a specific segment of the protein (P1 α1 helix), we set out to test whether ParA in its different nucleotide-bound states resulted in a significant conformational change that affected the N-terminal α1-α14′/15′ dimer interaction. A

change in the position of α1 would be indicative of the first step to a solution to membrane transition.

P1 ParA was found to be a dimer in the ADP-bound crystal structure that was locked in a single conformation (Chapter 4). The contacts between the N-terminus of the α1 helix and the α14′/α15′ helices of its dimer partner were conserved in this structure. Comparing the bimane fluorescence resulting from each of the single cysteine mutants (a series of bimane labeled mutants from N19C to V25C) in their apo (red spectrum) and their ADP-bound states (blue spectrum) shows that the position of the α1 helix changes very little as a result of the binding event (Figure 7_7). The bimane spectrums of the mutants V21C and T22C (low label incorporation) were noisy and were not included in the analysis. The analysis confirmed that the N-terminus of ParA does not move out of its α14′/α15′ pocket in the course of binding ADP.

The ATP-bound state of P1 ParA is predicted to be a dimer. Recent evidence that suggests that it forms filaments in its ATP-bound state that function in some manner to segregate newly replicated plasmids (Chapter 4). It is also possible that ParA is localized to the cell membrane in its ATP-bound state similar to other deviant Walker A *par* motor proteins through an interaction with its unique N-terminus. An examination of the bimane fluorescence from the single cysteine mutants (bimane labeled N19C-V25C) reveals that there is little change in intensity due to the binding of ATP (blue spectra) suggesting that the α1 helix does not undergo a major transition in solution (Figure 7_7). This is also supported by the lack of a significant shift in the emission characteristics of

bimane due to an ATP-binding event (Table 7_1). Bimane has an additional property of having solvent (or environmental) sensitive excitation and emission characteristics. One of the most sensitive mutants (reflected in changes to the intensity) for examining changes in the proximity of the α1 to α14′ helices is V25C as it is nearer to H282W. The fluorescence at this location is slightly quenched during the process of binding ATP indicating that there might be a slight movement of the mutant pair towards one another. As there is no apparent movement of the N-terminus as a result of binding ATP in solution there is an absence of evidence suggesting that it could act as a membrane anchor. But, this should not eliminate the possibility of the N-terminus interacting with the membrane as MinD requires an interaction with lipid in order to function (dimerize and interact with MinC).

P1 ParA in its non-hydrolyzeable ATP-analog-bound state locks ParA in a conformation that mimics the ATP-bound state. *Par* motor protein filaments are stabilized and can form extended structures in the presence of non-hydrolyzeable ATP analogs (Garner et al., 2004). Examining the bimane fluorescence (intensity) of all the single cysteine mutants in their ATP-analog-bound states shows that they are similar to the fluorescence observed in the ATP-bound state (Figure 7_7). Again, V25C appears to be the most effective sensor of N-terminal α1 movement as it exhibits the largest quenching of fluorescence indicating a slight rearrangement bringing the two mutant residues closer together. Interestingly, the addition of ParB (with *parS*small) lead to additional quenching at V25C suggesting that the ParB binding site is near enough to alter the local structural features. Overall, the ATP-analog studies show that it behaves in a similar

manner to ATP and that other than indications of slight rearrangements there is no major movement of the N-terminus.

The additional quenching of bimane observed in the V25C sample due to the addition of ParB may be a reflection of a nearby binding event. In the MinD system, MinE is responsible for the stimulation of MinD's ATPase activity in a manner similar to ParB. MinE's binding site on MinD has been mapped to a site on MinD's α7 helix, which is analogous to the α14 helix in ParA. Mutagenesis of certain residues in the α7 helix of MinD (D152 along with S148, D154, and I159) abrogates its association with MinE in 2-hybrid studies (Ma et al., 2004; Ma et al., 2003). Sequence alignment of P1 ParA and MinD showed that S283 in ParA is analogous to D152 in MinD and is located on the outer face the α14 helix adjacent to the α1 interaction domain (and H282W). The other residues found to disrupt the MinD and MinE interaction are also mapped to α14 in P1 ParA and are near the α1 interaction domain. The α14 helix had been previously shown to be important to ParA's function in homology sequence scanning experiments that found swapping sequences representing residues 282-290 in *parA* blocked partitioning (Radnedge et al., 1998). The possibility that the ParB binding site lies along ParA's α14 helix near the α1 interaction domain is worth pursuing in the future and the bimane-tryptophan partner constructs made to investigate α1 movement may be well suited to mapping the ParA-ParB interaction.

The N-terminus of ParA does not appear to move in the course of binding any nucleotide tested in solution. There is a strong possibility the α1 helix movement would require an

interaction with the cell's membrane as is the case with MinD. Attempts to assay fluorescence changes in the presence of *E.coli* phospholipids lead to excessive scatter that prevented accurate measurements, so a membrane induced movement cannot be ruled out. Additional measurements using detergent-lipid mixed micelles, detergent micelles, or less polar solvents needs to be undertaken to address questions revolving around lipid dependent movements.

7.3 Discussion

The fluorescence studies on the N-terminus of P1 ParA demonstrated that in solution the N-terminus does not move away from its interaction with the $\alpha14'$-$\alpha15'$ dimer partner interface in response to binding any of the nucleotides examined. This finding agrees with the N-terminal structural interactions found in the structures of ParA in its apo and ADP-bound states. The N-terminal $\alpha1$ interaction with its $\alpha14'$-$\alpha15'$ dimer partner interface appears to be at the core of the ATP-bound ParA dimer as well. The stability of this N-terminal interaction makes it unlikely that it would move unless it was exposed to an environment that could make compensatory interactions (the membrane). The deviant Walker A protein MinD requires an interaction with the membrane to function. The same possibility may exist with ParA, but will require further experiments to make a judgment either way.

The bimane-tryptophan constructs made for this study provides a platform that can easily be adapted for future structural studies on P1 ParA. One of the most interesting findings to come out of the present study was the possible detection of a ParB binding event by the V25C-bimane mutant. The core C220V mutant background could easily be adapted to the ParA-ParB fusion constructs in a way that would facilitate the mapping of the interaction site. Placing a tryptophan residue on the α1 helix at position Q23 and corresponding cysteine residues at positions adjacent to the ParA interaction domain in ParB (residues 2-12 from Chapter 5) should provide effective coverage of the interaction site. The placement of the tryptophan residue on the α1 helix allows it to be near the proposed α14 binding site without interfering with a binding event. A detected binding event could be quickly validated using the mutants mapped onto α14 from yeast-2-hybrid studies on the MinE-MinD binding interface. Using this construct would also allow the efficient screening of potential catalytic residues in ParB (discovered in Chapter 5) for activity.

CC-3′ (C220A), 5′-CGGATGAAGGCGCCGAGGCCCAGCTTGCGAC-3′ (C303A/C305A), 5′-CCATAAGTATGCCCATAGCCTGGC-3′ (C326A), and 5′-CGGTTTTGAACGCGCCGGCGAGTCT TTTGAC-3′ (C352A). Once the site of the reactive cysteine was determined (C220) the site was mutated conservatively to alternate residues to assess any potential impact on protein function. The following oligonucleotides were used to mutate the Cys220 site to valine, threonine and serine residues: 5′-GGAGAGAGCTGGTCAATGAGCATCTACC-3′ (C220V), 5′-GGAGAGAGCTGACCAATGAGCATCTACC-3′ (C220T), 5′-GGAGAGAGCTGAGCAATGAG CATCTACC-3′ (C220S). To undertake experiments addressing the potential movements in the α1 helix a sensor residue (tryptophan) had to be strategically placed in the α14′ helix. His282 was judged from the crystal structure to be adjacent to the area of interest in the α1 helix and was mutated to a tryptophan residue (5′-GCAACTGTCGATTTCTGGTCATC GCTTAAATACG-3′ (H282W)). Single cysteine residues serving as sites for introducing fluorescent probes were inserted into the α1 helix from Asn19 to Val25 (5′-GCAAACA GAATGCTCTGCGTTCTGACTGAAC-3′ (N19C), 5′-CAGAATGCTCAATTGCCTGACTGAACAAG-3′ (V20C), 5′-GAATGCTCAATGTTTGCACTGAACAAGTACAG-3′ (L21C), 5′-GCTCAATGTTCT GTGCGAACAAGTACAGTTGC-3′ (T22C), 5′-GCTCAATGTTCTGACTTGCCAAGTACAGTTGC-3′ (E23C), 5′-GTTCTGACTGAATGCGTACAGTTGCAAAAGG-3′ (Q24C), 5′-CTGACTGAAC AATGCCAGTTGCAAAAGGATG-3′ (V25C)). All the resulting constructs were sequenced to check for correct substitutions using the Sanger method (Applied Biosystems instrument) at the MD Anderson sequencing core facility. ParA mutant proteins were expressed in BL21 (DE3) *E. coli* cells and purified as described previously (Chapter 3 methods). The function of the ParA mutant proteins was assessed to insure the maintenance of activity using a coupled reaction assay that detects the generation of P_i (as covered in Chapter 5 methods)(Webb, 1992).

Site accessibility studies – making a reactive cysless construct. Purified cysteine mutants were assayed for their accessibility to solvent judged by their ability to react with PyMPO-maleimide label. Reactive cysteines were assessed for their inability to be labeled by PyMPO. ParA (5 μM) was reacted with 15x PyMPO-maleimide label in buffer (20 mM MES pH 6 and 200 mM NaCl) for 20 minutes at 25 ^{o}C. Sample reactions were quenched with 50x L-cysteine before being resolved on a 12% PAGE. Resulting gels were documented under UV light using an Alpha Innotech FluorChem SP system and saved in digital format with AlphaEase FC software.

Purification of Bimane labeled proteins. Single cysteine mutant ParA proteins were purified using Ni^{2+} resin chromatography as described previously (Chapter 3 methods). Purified ParA samples (diluted reducing agent) were next labeled with 10x excess Pyridyl-Bimane in buffer (20 mM MES (pH 6) and 200 mM NaCl) for 60 minutes at 4 ^{o}C. The PDT-bimane labeled samples were first diluted then reapplied to the Ni^{2+} column and were washed extensively before being re-eluted.

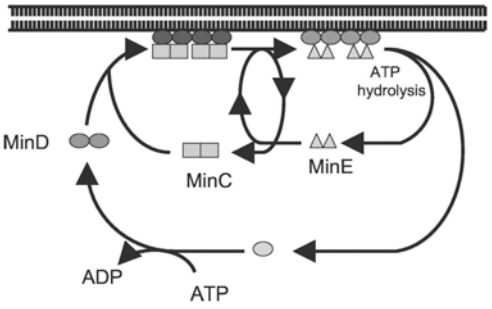
Labeling efficiency. Pyridyl bimane was reacted with DTT creating a mimetic of a bound fluorescent molecule suitable for generating a standard curve used to quantify fluorescence intensities. This method has been described previously (Janz and Farrens, 2004). Briefly, 2 μM labeled ParA protein samples were split in half (2 x 100 μL). One part was then reacted with $[TCA]_{final}$ (3%), while the other part was reduced with $[TCEP]_{final}$ (30mM). The samples were then centrifuged for 15 min. at 12,000xg. The supernatants of each sample were reacted with the opposite agent. Samples that had been

treated with TCA first were compared against a standard bimane curve to determine the amount of free label. Samples treated with TCEP first were compared against a standard bimane curve to determine the labeling efficiency.

Fluorescence measurements. Steady-state fluorescence measurements of PDT-bimane labeled ParA samples (2 μM) were conducted at 25 ^{o}C in 25 mM HEPES (pH 7) and 250 mM NaCl using a PTI fluorometer. Emission scans were recorded from 410 to 605 nm (with 2.0 nm band pass) while exciting at 380 nm (0.3 nm band pass). Nucleotide when present was at a concentration of 1 mM.

A

B



C

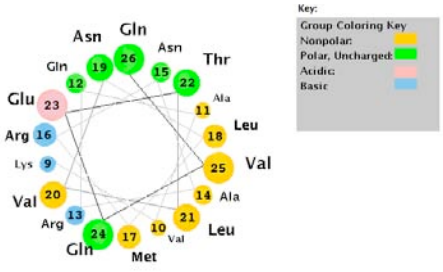


Figure 7_1: Structural Elements of Deviant Walker A Proteins. (A) P1 ParA apo structure denoting the position of the unique α1 helix. (B) Model of MinD protein activity illustrating the role of the C-terminal amphipathic helix in anchoring the protein and its partners to the membrane (Hu et al., 2003). (C) Amphipathic nature of the P1 ParA α1 helix. Adapted from cti.itc.virginia.edu/

7.4 Methods

ParA association with small unilamellar vesicles (SUVs). *E.coli* phospholipid extract (chloroform:methanol) (Avanti Polar Lipids Inc.) was evaporated and then hydrated in 15 mM Tris-Cl (pH 7) and 300 mM NaCl through sonication at 22 ^{o}C for 10-15 min. (until clarified) producing a ~1 mM mixture of unilamellar vesicles 15-50 nm size (SUVs). Prepared SUVs were stored under Argon at RT. 500 nM SUV were interacted with 200 nM P1 ParA_fusNt28ParB in 20 mM Tris-Cl (pH 7), 250 mM NaCl and 2 mM DTT in the presence of 1 mM nucleotide (ADP or ATP) at 30 ^{o}C for 10 min. before being resolved using ultra-centrifugation (100,000xg for 1 hour). The supernatant was removed and the remaining pellet was resuspended in an equal volume. Samples were then resolved by 12% SDS_PAGE (Coomassie stained). Gels were digitized using an Alpha Innotech FluorChem SP gel documentation system and analyzed using AlphaView software. Resulting protein band intensities were used to generate a plot demonstrating that P1 ParA's association with the 425 nM SUV was saturable in the ATP-bound state (10 nM – 1 μM).

Mutagenesis, purification and functional characterization. Mutagenesis was performed using a QuikChange strategy (Stratagene). In order to generate a reactive cysless P1 ParA construct each of its 5 native cysteine residues was removed piecemeal to assess the reactivity of the remaining residues. The following oligonucleotides were used mutate the cysteine residues to alanines: 5′-GGAGAGAGCTGGCCAATGAGCATCTA

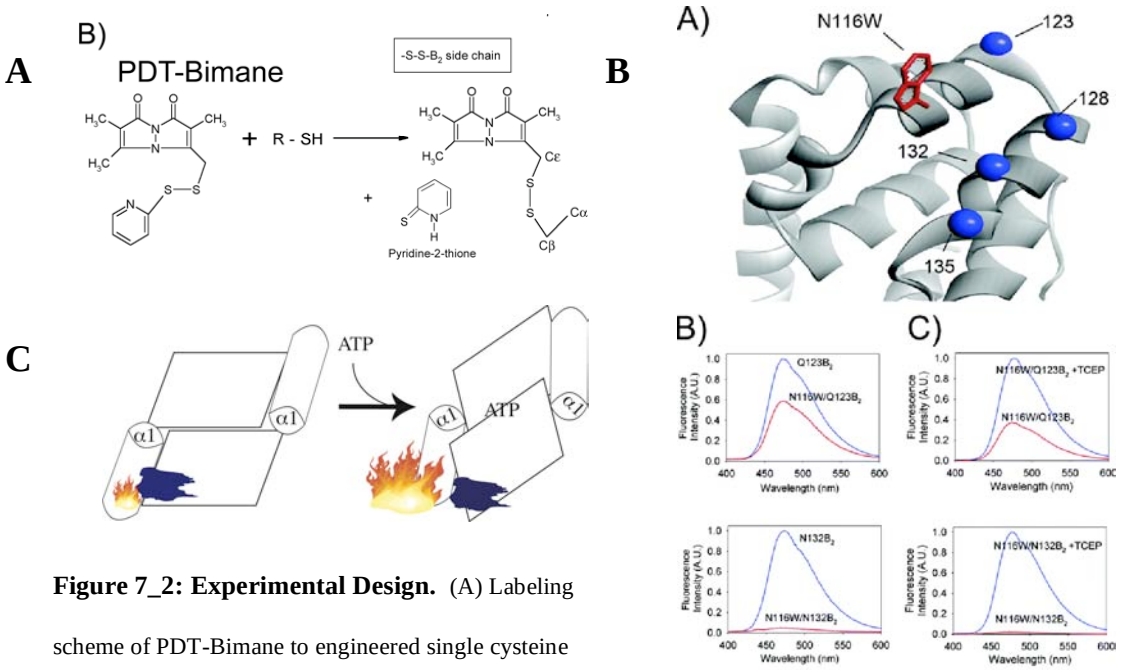


Figure 7_2: Experimental Design. (A) Labeling scheme of PDT-Bimane to engineered single cysteine residues. (B) Tryptophan dependent quenching of Bimane. Study of located sites demonstrating the specific proximity/distance relationship between a tryptophan residue and bimane sites. Residue 123 in lysozyme shows some quenching of fluorescence where residue 132 located across from the tryptophan residue is almost completely quenched. (Mansoor and Farrens, 2004) (C) In ParA this will be used to test the nucleotide dependent movement of the α1 helix.

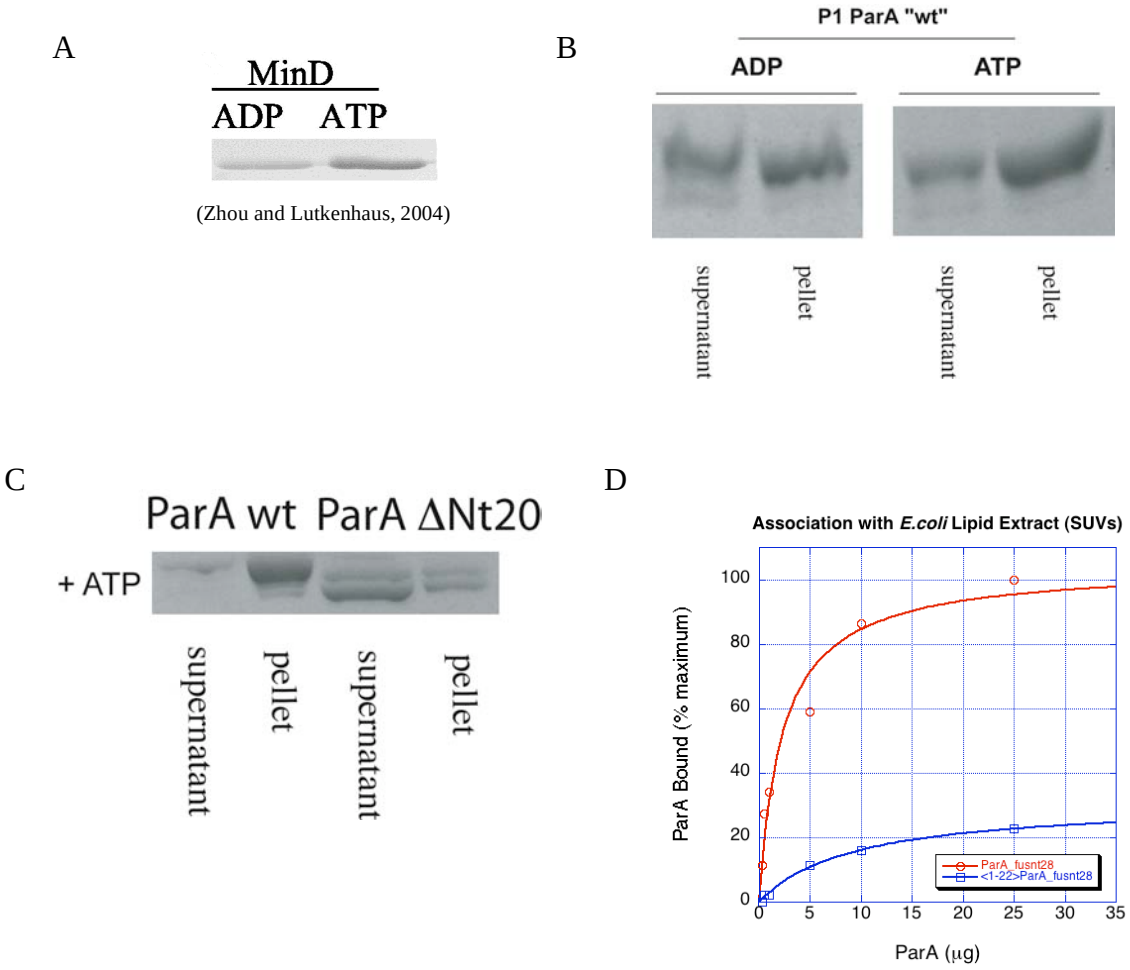
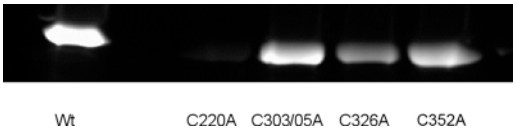


Figure 7_3: ParA Association with *E. coli* Lipid Vesicles. (A) MinD *E. coli* lipid association post centrifugation. (B) Nucleotide dependent P1 ParA association with *E. coli* lipid. (C) N-terminal α1 helix Δ1-20 mutant impaired association with lipids. (D) Plot of ParA titration into a fixed amount of lipid resulting in a saturable association.

A **Location of ParA's 5 native cysteine residues**

α11(C220)
α14/B9 (C303/05)
α15 (C326)
α15 extended (C352)

B



C

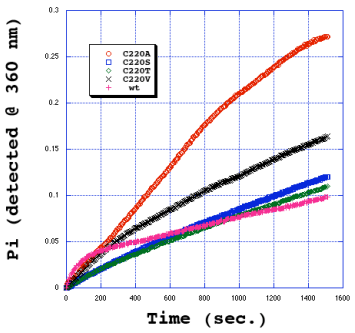


Figure 7_4: Function of the Reactive Cysless Construct. (A) Locations of P1 ParA's 5 native cyteine residues. (B) Accessibility labeling study utilizing PyMPO-maleimide to identify reactive cysteines. SDS-PAGE of fluorescently labeled mutant ParA protein. (C) ParA ATPase activity of C220 mutants (C220A, C220S, C220T and C220V) used to identify the most functional mutant.

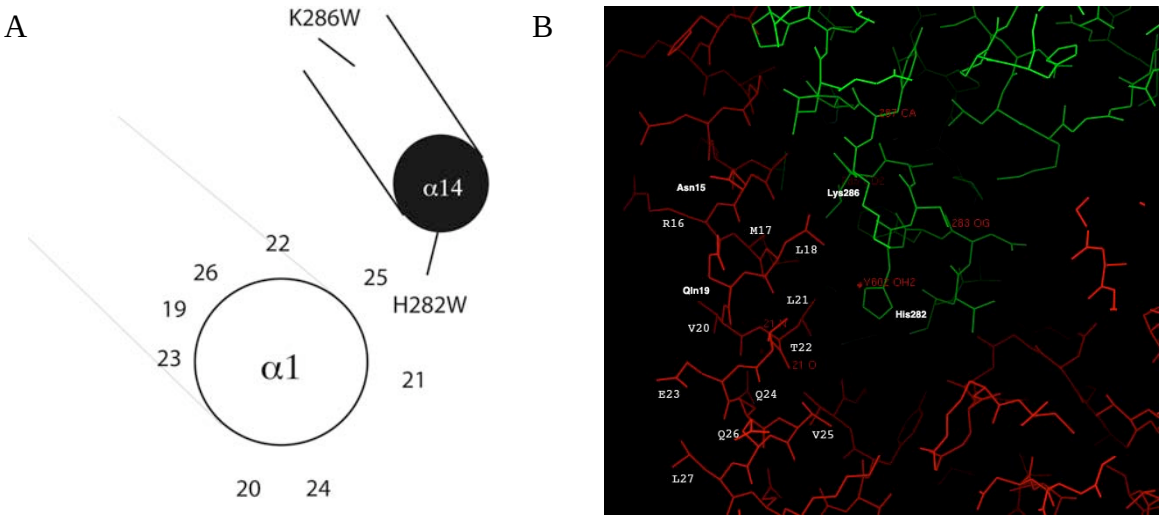


Figure 7_5: Location of the Series of Single Cysteine Residues (Bimane Sites) and Corresponding Tryptophan Residues. (A) Cartoon showing the relative positions of the single cysteine residues (helical tube) and their approximate relation to corresponding strategic tryptophan residues introduced into the α14′ helix. (B) The positions of the sites in the crystal structure of apo P1 ParA.

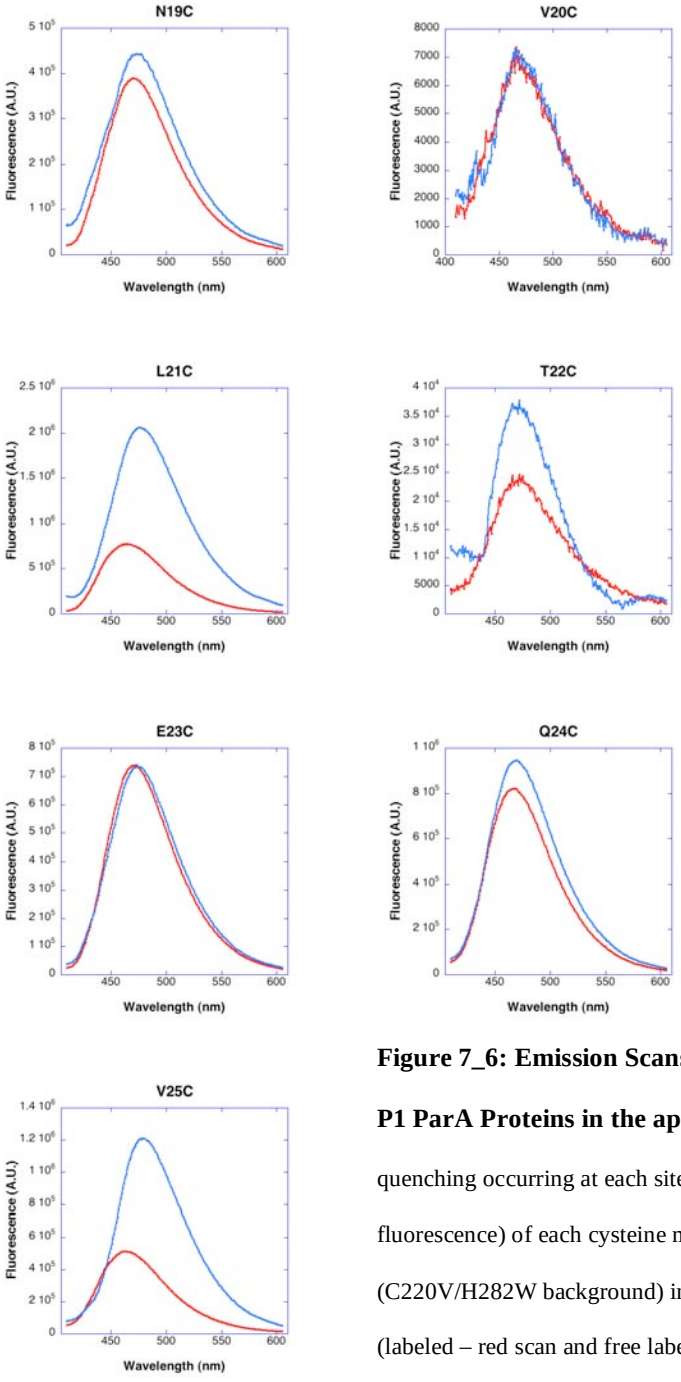


Figure 7_6: Emission Scans of Bimane Labeled P1 ParA Proteins in the apo State. Illustrating the quenching occurring at each site. Emission scans (raw fluorescence) of each cysteine mutant in the α1 helix (C220V/H282W background) in ParA's apo state (labeled – red scan and free label - blue scan). The transition from the red to blue scans illustrates the degree of quenching due to the proximity of the H282 tryptophan residue.

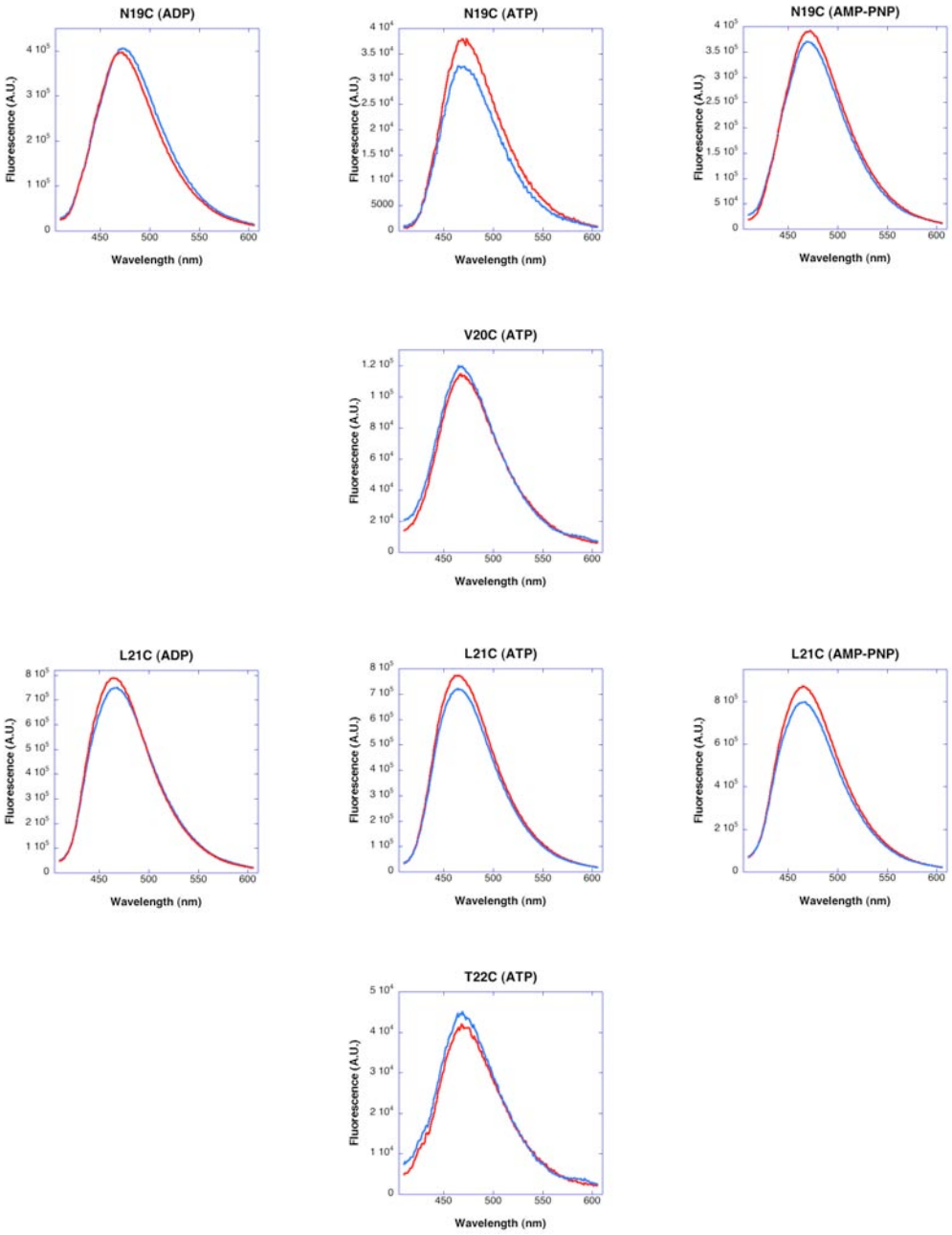


Figure 7_7 (part A): Emission Scans of Bimane Labeled ParA Cysteine Mutants in Different Nucleotide Bound States (ADP, ATP, and AMP-PNP). Emission scans (raw fluorescence) of labeled ParA in its apo (red scan) and nucleotide-bound (blue scan) states. The change in fluorescence is an illustration a change in the sites proximity to the neighboring H282 tryptophan residue (nucleotide dependent movement of the α1 helix).

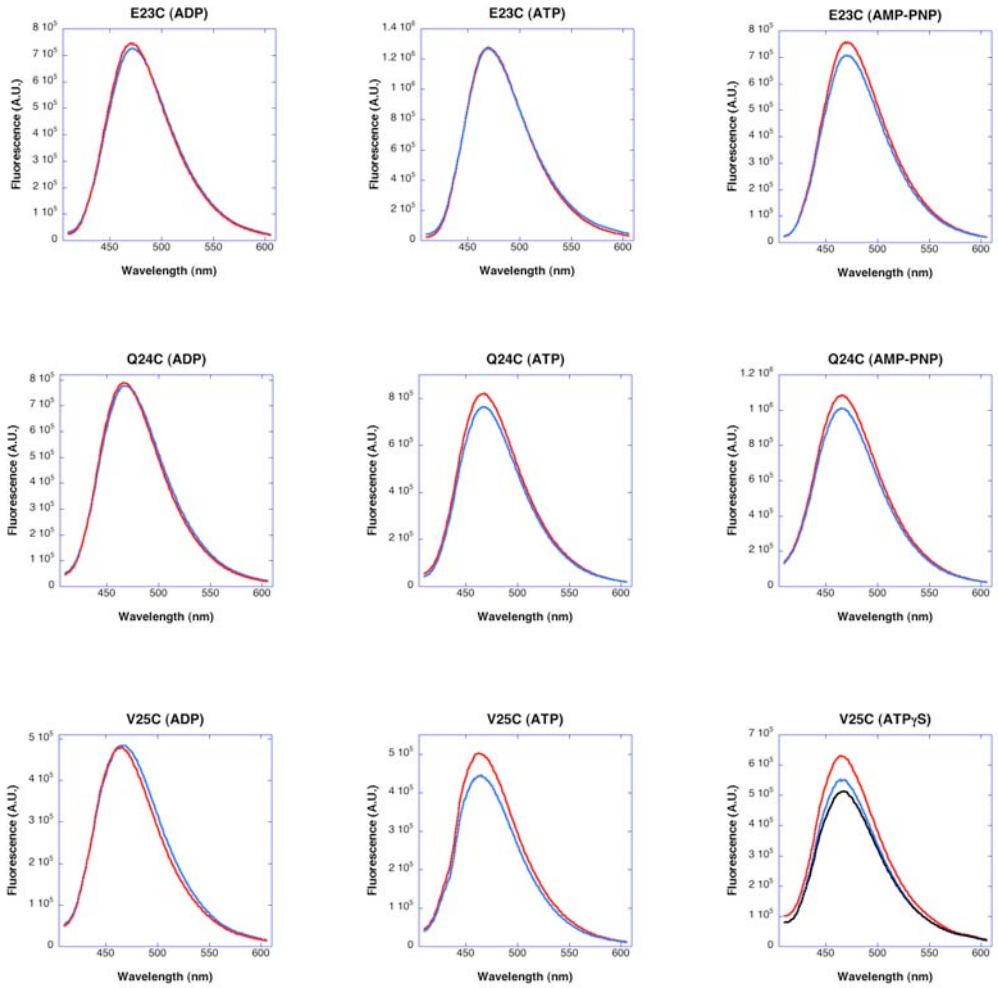


Figure 7_7 (part B): Emission Scans of Bimane Labeled ParA Cysteine Mutants in Different Nucleotide Bound States (ADP, ATP, and ATP-analog). Emission scans (raw fluorescence) of labeled ParA in its apo (red scan) and nucleotide-bound (blue scan) states. The change in fluorescence is an illustration a change in the sites proximity to the neighboring H282 tryptophan residue (nucleotide dependent movement of the α1 helix). The V25C mutant contains an extra scan (black) representing the addition of P1 ParB + *parS*small.

Mutant *(C220V/H282W background)*	λ_{max} *(apo)* *(nm)*	λ_{max} *(ATP)* *(nm)*
N19C	471	470
V20C	470	465
L21C	465	464
T22C	468	469
E23C	469	470
Q24C	466	468
V25C	465	465

Table 7_1: Wavelength maxima of emission scans.

The wavelength maxima of the emission scans from each ParA labeled mutant in both its apo and ATP-bound states. A comparison of the λ_{maxima} illustrates that there is no dramatic shift due to the binding of ATP.

SUMMARY

The work included as part of this thesis was undertaken to address the outstanding questions involving the structural features that allow ParA (type Ia *par* proteins) to function in the disparate activities of partitioning of the P1 plasmid as well as the autorepression of the *par* operon.

P1 ParA is a deviant Walker A ATPase and a type Ia *par* motor protein. Typical deviant Walker A proteins undergo a monomer to dimer transition when binding ATP, which marks the proteins transition to an activated functional state. As a type Ia motor protein, P1 ParA contains a roughly 100 residue N-terminus that extends beyond the core ATPase domain present in other Walker A proteins. In X-ray crystallography experiments the structure of P1 ParA in the apo state revealed that it surprisingly existed in the dimer state. Supporting experiments using SEC-MALS confirmed that ParA was able to form a dimer in solution at physiologically relevant concentrations. P1 ParA's unique N-terminus formed an extended α-helix (α1) that bridged the two ParA monomers resting in a pocket created by α14′ and α15′ in the dimer partner. These bridging contacts between dimer partners were found to be critical for the stability of the apo dimer in solution. Mutating key hydrophobic residues (A11 and A14) along the inner face of the α1 helix destabilized the dimer state. From the structural and biochemical experiments ParA was found to be a dimer, which has important implications for the nature of the nucleotide switching mechanism in type Ia motor proteins.

P1 ParA's function within the cell is directed by its nucleotide-bound state in an unknown manner. ParA in its ADP-bound state protects a large segment (~ 100 bp) of its own *parOP* region from DNase I activity. ParA's protection of its *parOP* region is believed to be critical to its function as an autorepressor of its own operon. X-ray crystallography studies on the ADP-bound state were undertaken to understand the underlying structural features that lead to ParA's improved DNA binding capabilities. Dr. Schumacher's ADP-bound structure revealed that ParA formed a dimer structure that was locked in a single conformation. A ParA-DNA bound model based upon the ADP-bound structure revealed

that the binding of DNA required the DNA to bend so that the recognition helices in ParA's HTH motif are able to bind to the major grooves. A basic region emerging as a structural feature of the ADP-bound state was found to be in a position to stabilize the bending of DNA. Mutagenesis studies to the basic region were subsequently found to impair binding to the *parOP* segment lending support to the DNA-bound model.

ParA in its ATP-bound state participates in the segregation of the P1 plasmid in an unknown manner. Difficulty with precipitation in X-ray crystallography studies on the ATP-bound state of ParA lead to the further investigation into the nature of the precipitation. Using negative stain electron microscopy (EM) ParA was shown to form filaments composed of regular subunits in an ATP-dependent manner. Further investigation into the nature of ParA's participation in the segregation of the P1 plasmid using a reconstituted system of the P1 *par* cassette revealed that the purified components (ParB and *parS*) were localized to the ParA filaments in fluorescence microscopy experiments. The exact mechanism of ParA's mobilization of the plasmid has yet to be elucidated, but it appears that ParA's functional role is to segregate the P1 plasmid to the daughter cells through the formation of filaments.

P1 ParA's ATPase function is modulated through an interaction with its P1 partner protein ParB. The nature of the stimulating interaction is unknown, but the activity of the interaction had been previously mapped to the N-terminal 68 residues. X-ray crystallography studies were undertaken using a ParA-ParB fusion construct to resolve the structural nature of the interaction. Although the studies were unsuccessful both the specificity and the source of the stimulating interaction were further mapped to the N-terminal 28 residues of ParB in biochemistry experiments. ATPase functional studies on the ParA-ParB fusion protein found that two arginine residues (R6 and R11) play an important role in the stimulation of ParA's ATPase activity.

In summary, from the studies of P1 ParA's different functional states our studies illustrate how nucleotide binding mediates ParA's role in the cell by stabilizing a specific dimeric state suited for its activities in either repression or partition.

REFERENCES

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. J Mol Biol *185*, 261-272.

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. J Mol Biol *169*, 353-372.

Barilla, D., Rosenberg, M. F., Nobbmann, U., and Hayes, F. (2005). Bacterial DNA segregation dynamics mediated by the polymerizing protein ParF. Embo J *24*, 1453-1464.

Barilla, 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. Proc Natl Acad Sci U S A *104*, 1811-1816.

Bennett, M. J., Choe, S., and Eisenberg, D. (1994). Domain swapping: entangling alliances between proteins. Proc Natl Acad Sci U S A *91*, 3127-3131.

Bouet, J. Y., and Funnell, B. E. (1999). P1 ParA interacts with the P1 partition complex at parS and an ATP-ADP switch controls ParA activities. Embo J *18*, 1415-1424.

Bouet, J. Y., Surtees, J. A., and Funnell, B. E. (2000). Stoichiometry of P1 plasmid partition complexes. J Biol Chem *275*, 8213-8219.

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

Brünger, A. T., Adams, P.D., Clore, G.M., DeLano, W.L., Gros, P., Crosse-Kunstleve, R.W., Jiang, J.S., Kuszewski, J., Nilges, M., Pannu, N.S., Read, R.J., Rice, L.M., Simonson, T., and Warren, G.L. (1998). Crystallography and NMR System: A new software suite for macromolecular structure determination. Acta Crystallogr D Biol Crystallogr *54*, 905-921.

Cordell, S. C., and Lowe, J. (2001). Crystal structure of the bacterial cell division regulator MinD. FEBS Lett *492*, 160-165.

Crestfield, A. M., Stein, W. H., and Moore, S. (1962). On the aggregation of bovine pancreatic ribonuclease. Arch Biochem Biophys *Suppl 1*, 217-222.

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. J Mol Biol *236*, 1289-1298.

Davey, M. J., and Funnell, B. E. (1994). The P1 plasmid partition protein ParA. A role for ATP in site-specific DNA binding. J Biol Chem *269*, 29908-29913.

Davey, M. J., and Funnell, B. E. (1997). Modulation of the P1 plasmid partition protein ParA by ATP, ADP, and P1 ParB. J Biol Chem *272*, 15286-15292.

Davis, M. A., Martin, K. A., and Austin, S. J. (1992). Biochemical activities of the parA partition protein of the P1 plasmid. Mol Microbiol *6*, 1141-1147.

Davis, M. A., Radnedge, L., Martin, K. A., Hayes, F., Youngren, B., and Austin, S. J. (1996). The P1 ParA protein and its ATPase activity play a direct role in the segregation of plasmid copies to daughter cells. Mol Microbiol *21*, 1029-1036.

Doublié, S. (1997). Preparation of selenomethionyl proteins for phase determination. Methods Enzymol *276*, 523-530.

Drenth, J. Principles of Protein X-Ray Crystallography. (New York, Springer-Verlag)

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. Embo J *28*, 1792-1802.

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. Mol Microbiol *61*, 1428-1442.

Ebersbach, G., Sherratt, D. J., and Gerdes, K. (2005). Partition-associated incompatibility caused by random assortment of pure plasmid clusters. Mol Microbiol *56*, 1430-1440.

Erdmann, N., Petroff, T., and Funnell, B. E. (1999). Intracellular localization of P1 ParB protein depends on ParA and parS. Proc Natl Acad Sci U S A *96*, 14905-14910.

Fung, E., Bouet, J. Y., and Funnell, B. E. (2001). Probing the ATP-binding site of P1 ParA: partition and repression have different requirements for ATP binding and hydrolysis. Embo J *20*, 4901-4911.

Funnell, B. E. (1988a). Mini-P1 plasmid partitioning: excess ParB protein destabilizes plasmids containing the centromere parS. J Bacteriol *170*, 954-960.

Funnell, B. E. (1988b). Participation of Escherichia coli integration host factor in the P1 plasmid partition system. Proc Natl Acad Sci U S A *85*, 6657-6661.

Funnell, B. E. (1991). The P1 plasmid partition complex at parS. The influence of Escherichia coli integration host factor and of substrate topology. J Biol Chem *266*, 14328-14337.

Funnell, B. E. (2005). Partition-mediated plasmid pairing. Plasmid *53*, 119-125.

Funnell, B. E., and Gagnier, L. (1993). The P1 plasmid partition complex at parS. II. Analysis of ParB protein binding activity and specificity. J Biol Chem *268*, 3616-3624.

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*, 924-932.

Garner, E. C., Campbell, C. S., and Mullins, R. D. (2004). Dynamic instability in a DNA-segregating prokaryotic actin homolog. Science *306*, 1021-1025.

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*, 1270-1274.

Georgiadis, M. M., Komiya, H., Chakrabarti, P., Woo, D., Kornuc, J. J., and Rees, D. C. (1992). Crystallographic structure of the nitrogenase iron protein from Azotobacter vinelandii. Science *257*, 1653-1659.

Gerdes, K., and Molin, S. (1986). Partitioning of plasmid R1. Structural and functional analysis of the parA locus. J Mol Biol *190*, 269-279.

Gerdes, K., Moller-Jensen, J., and Bugge Jensen, R. (2000). Plasmid and chromosome partitioning: surprises from phylogeny. Mol Microbiol *37*, 455-466.

Golovanov, A. P., Barilla, 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. Mol Microbiol *50*, 1141-1153.

Hamm, H. E., Deretic, D., Arendt, A., Hargrave, P. A., Koenig, B., and Hofmann, K. P. (1988). Site of G protein binding to rhodopsin mapped with synthetic peptides from the alpha subunit. Science *241*, 832-835.

Hatano, T., Yamaichi, Y., and Niki, H. (2007). Oscillating focus of SopA associated with filamentous structure guides partitioning of F plasmid. Mol Microbiol *64*, 1198-1213.

Hayes, F., Davis, M. A., and Austin, S. J. (1993). Fine-structure analysis of the P7 plasmid partition site. J Bacteriol *175*, 3443-3451.

Hayes, F., Radnedge, L., Davis, M. A., and Austin, S. J. (1994). The homologous operons for P1 and P7 plasmid partition are autoregulated from dissimilar operator sites. Mol Microbiol *11*, 249-260.

Hayes, F., and Barilla, D. (2006). The bacterial segrosome: a dynamic nucleoprotein machine for DNA trafficking and segregation. Nat Rev Microbiol *4*, 133-143.

Hendrickson, W. A. (1991). Determination of macromolecular structures from anomalous diffraction of synchrotron radiation. Science *254*, 51-58.

Hendrickson, W. A., Horton, J. R., and LeMaster, D. M. (1990). Selenomethionyl proteins produced for analysis by multiwavelength anomalous diffraction (MAD): a vehicle for direct determination of three-dimensional structure. Embo J *9*, 1665-1672.

Hong, Minsun. (2005) Structural studies on the MarR family, Crystal structures of OhrR-DNA and SlyA-inducer complexes. (Doctoral Thesis)

Hu, Z., and Lutkenhaus, J. (2001). Topological regulation of cell division in E. coli. spatiotemporal oscillation of MinD requires stimulation of its ATPase by MinE and phospholipid. Mol Cell *7*, 1337-1343.

Hu, Z., and Lutkenhaus, J. (2003). A conserved sequence at the C-terminus of MinD is required for binding to the membrane and targeting MinC to the septum. Mol Microbiol *47*, 345-355.

Janz, J. M., and Farrens, D. L. (2004). Rhodopsin activation exposes a key hydrophobic binding site for the transducin alpha-subunit C terminus. J Biol Chem *279*, 29767-29773.

Jensen, R. B., Dam, M., and Gerdes, K. (1994). Partitioning of plasmid R1. The parA operon is autoregulated by ParR and its transcription is highly stimulated by a downstream activating element. J Mol Biol *236*, 1299-1309.

Jones, T. A., Zou, J.-Y., Cowan, S.W., and Kjeldgaard, M. (1991). Improved methods for building protein models in electron density maps and the location of errors in these models. . Acta Crystallogr *A47*, 110-119.

Koonin, E. V. (1993). A superfamily of ATPases with diverse functions containing either classical or deviant ATP-binding motif. J Mol Biol *229*, 1165-1174.

Laskowski, R. A., Moss, D. S., and Thornton, J. M. (1993). Main-chain bond lengths and bond angles in protein structures. J Mol Biol *231*, 1049-1067.

Leonard, T. A., Butler, P. J., and Lowe, J. (2004). Structural analysis of the chromosome segregation protein Spo0J from Thermus thermophilus. Mol Microbiol *53*, 419-432.

Leonard, T. A., Butler, P. J., and Lowe, J. (2005). Bacterial chromosome segregation: structure and DNA binding of the Soj dimer--a conserved biological switch. Embo J *24*, 270-282.

Leslie, A. G. W. (1992). Recent changes to the MOSFLM package for processing film and image plate data. Joint CCP4 + ESF-EAMCB Newsletter on Protein Crystallography *26*.

Lim, G. E., Derman, A. I., and Pogliano, J. (2005). Bacterial DNA segregation by dynamic SopA polymers. Proc Natl Acad Sci U S A *102*, 17658-17663.

Liu, Y., and Eisenberg, D. (2002). 3D domain swapping: as domains continue to swap. Protein Sci *11*, 1285-1299.

Lundblad, J. R., Laurance, M., and Goodman, R. H. (1996). Fluorescence polarization analysis of protein-DNA and protein-protein interactions. Mol Endocrinol *10*, 607-612.

Lutkenhaus, J., and Sundaramoorthy, M. (2003). MinD and role of the deviant Walker A motif, dimerization and membrane binding in oscillation. Mol Microbiol *48*, 295-303.

Ma, L., King, G. F., and Rothfield, L. (2004). Positioning of the MinE binding site on the MinD surface suggests a plausible mechanism for activation of the Escherichia coli MinD ATPase during division site selection. Mol Microbiol *54*, 99-108.

Ma, L. Y., King, G., and Rothfield, L. (2003). Mapping the MinE site involved in interaction with the MinD division site selection protein of Escherichia coli. J Bacteriol *185*, 4948-4955.

Mansoor, S. E., and Farrens, D. L. (2004). High-throughput protein structural analysis using site-directed fluorescence labeling and the bimane derivative (2-pyridyl)dithiobimane. Biochemistry *43*, 9426-9438.

Matthews, B. W. (1968). Solvent content of protein crystals. J Mol Biol *33*, 491-497.

Milligan, G. (2000). Insights into ligand pharmacology using receptor-G-protein fusion proteins. Trends Pharmacol Sci *21*, 24-28.

Moller-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. Mol Cell *12*, 1477-1487.

Moller-Jensen, J., Ringgaard, S., Mercogliano, C. P., Gerdes, K., and Lowe, J. (2007). Structural analysis of the ParR/parC plasmid partition complex. Embo J *26*, 4413-4422.

Murayama, K., Orth, P., de la Hoz, A. B., Alonso, J. C., and Saenger, W. (2001). Crystal structure of omega transcriptional repressor encoded by Streptococcus pyogenes plasmid pSM19035 at 1.5 A resolution. J Mol Biol *314*, 789-796.

Orlova, A., Garner, E. C., Galkin, V. E., Heuser, J., Mullins, R. D., and Egelman, E. H. (2007). The structure of bacterial ParM filaments. Nat Struct Mol Biol *14*, 921-926.

Popp, D., Narita, A., Oda, T., Fujisawa, T., Matsuo, H., Nitanai, Y., Iwasa, M., Maeda, K., Onishi, H., and Maeda, Y. (2008). Molecular structure of the ParM polymer and the mechanism leading to its nucleotide-driven dynamic instability. Embo J *27*, 570-579.

Popp, D., Yamamoto, A., Iwasa, M., Narita, A., Maeda, K., and Maeda, Y. (2007). Concerning the dynamic instability of actin homolog ParM. Biochem Biophys Res Commun *353*, 109-114.

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 Res *36*, 3676-3689.

Radnedge, L., Youngren, B., Davis, M., and Austin, S. (1998). Probing the structure of complex macromolecular interactions by homolog specificity scanning: the P1 and P7 plasmid partition systems. Embo J *17*, 6076-6085.

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. J Mol Biol *329*, 875-889.

Rayment, I. (1996). The structural basis of the myosin ATPase activity. J Biol Chem *271*, 15850-15853.

Rayment, I., Smith, C., and Yount, R. G. (1996). The active site of myosin. Annu Rev Physiol *58*, 671-702.

Rhodes, G. (2006) Crystallography made crystal clear. (Burlington, MA AP)

Rodionov, O., Lobocka, M., and Yarmolinsky, M. (1999). Silencing of genes flanking the P1 plasmid centromere. Science *283*, 546-549.

Rousseau, F., Schymkowitz, J. W., and Itzhaki, L. S. (2003). The unfolding story of three-dimensional domain swapping. Structure *11*, 243-251.

Salje, J., and Lowe, J. (2008). Bacterial actin: architecture of the ParMRC plasmid DNA partitioning complex. Embo J *27*, 2230-2238.

Salje, J., Zuber, B., and Lowe, J. (2009). Electron cryomicroscopy of E. coli reveals filament bundles involved in plasmid DNA segregation. Science *323*, 509-512.

Saraste, M., Sibbald, P. R., and Wittinghofer, A. (1990). The P-loop--a common motif in ATP- and GTP-binding proteins. Trends Biochem Sci *15*, 430-434.

Scheffzek, K., Ahmadian, M. R., and Wittinghofer, A. (1998). GTPase-activating proteins: helping hands to complement an active site. Trends Biochem Sci *23*, 257-262.

Suefuji, K., Valluzzi, R., and RayChaudhuri, D. (2002). Dynamic assembly of MinD into filament bundles modulated by ATP, phospholipids, and MinE. Proc Natl Acad Sci U S A *99*, 16776-16781.

Schumacher, M. (2004) Protein Crystallography (course notes).

Schumacher, M. A. (2007). Structural biology of plasmid segregation proteins. Curr Opin Struct Biol *17*, 103-109.

Schumacher, M. A. (2008). Structural biology of plasmid partition: uncovering the molecular mechanisms of DNA segregation. Biochem J *412*, 1-18.

Schumacher, M. A., and Funnell, B. E. (2005). Structures of ParB bound to DNA reveal mechanism of partition complex formation. Nature *438*, 516-519.

Schumacher, M. A., Glover, T. C., Brzoska, A. J., Jensen, S. O., Dunham, T. D., Skurray, R. A., and Firth, N. (2007). Segrosome structure revealed by a complex of ParR with centromere DNA. Nature *450*, 1268-1271.

Scrima, A., and Wittinghofer, A. (2006). Dimerisation-dependent GTPase reaction of MnmE: how potassium acts as GTPase-activating element. Embo J *25*, 2940-2951.

Stemmer, W. P., Crameri, A., Ha, K. D., Brennan, T. M., and Heyneker, H. L. (1995). Single-step assembly of a gene and entire plasmid from large numbers of oligodeoxyribonucleotides. Gene *164*, 49-53.

Surtees, J. A., and Funnell, B. E. (1999). P1 ParB domain structure includes two independent multimerization domains. J Bacteriol *181*, 5898-5908.

Surtees, J. A., and Funnell, B. E. (2001). The DNA binding domains of P1 ParB and the architecture of the P1 plasmid partition complex. J Biol Chem *276*, 12385-12394.

Surtees, J. A., and Funnell, B. E. (2003). Plasmid and chromosome traffic control: how ParA and ParB drive partition. Curr Top Dev Biol *56*, 145-180.

Terwilliger, T. C., and Berendzen, J. (1999). Automated MAD and MIR structure solution. Acta Crystallogr D Biol Crystallogr *55*, 849-861.

van den Ent, F., Moller-Jensen, J., Amos, L. A., Gerdes, K., and Lowe, J. (2002). F-actin-like filaments formed by plasmid segregation protein ParM. Embo J *21*, 6935-6943.

Walker, J. E., Saraste, M., Runswick, M. J., and Gay, N. J. (1982). Distantly related sequences in the alpha- and beta-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. Embo J *1*, 945-951.

Webb, M. R. (1992). A continuous spectrophotometric assay for inorganic phosphate and for measuring phosphate release kinetics in biological systems. Proc Natl Acad Sci U S A *89*, 4884-4887.

Weihofen, W. A., Cicek, A., Pratto, F., Alonso, J. C., and Saenger, W. (2006). Structures of omega repressors bound to direct and inverted DNA repeats explain modulation of transcription. Nucleic Acids Res *34*, 1450-1458.

Zhou, H., and Lutkenhaus, J. (2004). The switch I and II regions of MinD are required for binding and activating MinC. J Bacteriol *186*, 1546-1555.