DNA segregation or partition is an essential process that ensures stable genome transmission. In prokaryotes, partition is best understood for plasmids, which serve as tractable model systems to study the mechanistic underpinnings of DNA segregation at a detailed atomic level owing to their simplicity. Specifically, plasmid partition requires only three elements: a centromere-like DNA site and two proteins: a motor protein, generally an ATPase, and a centromere-binding protein. In the first step of the partition process, multiple centromere-binding proteins bind cooperatively to the centromere, which typically consists of several tandem repeats, to form a higher-order nucleoprotein complex called the partition complex. The partition complex recruits the ATPase to form the segrosome and somehow activates the ATPase for DNA separation. Two major families of plasmid par systems have been delineated based on whether they utilize ATPase proteins with deviant Walker-type motifs or actin-like folds. In contrast, the centromere-binding proteins show little sequence homology even within a given family. Recent structural studies, however, have revealed that these centromere-binding proteins appear to belong to one of two major structural groups: those that employ helix–turn–helix DNA-binding motifs or those with ribbon–helix–helix DNA-binding domains. The first structure of a higher-order partition complex was recently revealed by the structure of pSK41 centromere-binding protein, ParR, bound to its centromere site. This structure showed that multiple ParR ribbon–helix–helix motifs bind symmetrically to the tandem centromere repeats to form a large superhelical structure with dimensions suitable for capture of the filaments formed by the actin-like ATPases. Surprisingly, recent data indicate that the deviant Walker ATPase proteins also form polymer-like structures, suggesting that, although the par families harbour what initially appeared to be structurally and functionally divergent proteins, they actually utilize similar mechanisms of DNA segregation. Thus, in the present review, the known Par protein and Par–protein complex structures are discussed with regard to their functions in DNA segregation in an attempt to begin to define, at a detailed atomic level, the molecular mechanisms involved in plasmid segregation.

**Key words:** deviant Walker box motif, DNA segregation, insertional polymerization model, partition, partition complex, plasmid.

**INTRODUCTION**

DNA partition or segregation, is the process whereby the genetic material is accurately moved and positioned to daughter cells during cell division. The newly replicated eukaryotic chromosomes is carried out by the formation of microtubule-based spindles, which pull chromosomes to opposite cell poles [1]. In contrast, the molecular machinery that mediates the segregation of prokaryotic DNA has been less clear. Much of our understanding of prokaryotic partition has resulted from cellular, genetic and biochemical analyses on the segregation of low-copy-number plasmids. Such plasmids play a significant role in the spread of multidrug resistance and also play a central role in DNA manipulation technology. Thus understanding the mechanism behind their maintenance has important consequences. Whereas high-copy-number plasmids primarily rely on passive diffusion for plasmid maintenance, low-copy-number plasmids utilize so-called partition (par) systems, which are carried on the plasmid DNA [2–13]. The majority of par operons or cassettes contain two genes: one encoding a motor protein, which is typically an ATPase, and the second encoding a centromere-binding protein [2–11]. In addition, the centromere-like site bound by the centromere-binding protein is located near the par cassette. These three components are all that is required to direct the partition reaction (Figure 1A). However, the amounts of each of the Par proteins are carefully regulated, as deficiency or excess of either protein can completely or seriously impede the partition process [14–17]. Fluorescence microscopy studies have provided a visual picture of the highly dynamic partition reaction. For the most common type of par system, immunofluorescence and FISH (fluorescent in situ hybridization) studies have shown that newborn cells usually have one central plasmid focus, whereas larger older cells have two foci located at the one-fourth and three-fourth positions of the cell, indicative of plasmid segregation [18–23]. Studies using GFP (green fluorescent protein)-tagged plasmids gave visualized plasmid movement in live cells and have revealed that plasmid movement and separation are relatively rapid events in the entire cell cycle. The challenge has been to relate these visual observations with the molecular steps involved in plasmid partition.

A general understanding of the steps involved in DNA partition has been obtained from biochemical, cellular and genetic studies. Specifically, the first step involves binding of the centromere-like partition site by the centromere-binding protein, ultimately leading to the creation of a large nucleoprotein complex [3–11] (Figure 1A). Next, interactions between the centromere-binding proteins and the DNA lead to pairing between partition complexes on two different plasmids. Once the plasmids are paired, the
partition complexes are recognized by the motor protein leading to the assembly of the segregation, which then actively mediates the separation of the plasmids (Figure 1A). The simplicity of the plasmid par systems marks them as extremely attractive model systems to address the fundamental biological question of how DNA is segregated to daughter cells at the atomic level. As described in the present review, structural work carried out in the last 5 years has started to elucidate the detailed molecular mechanisms utilized by these proteins in each step of the DNA segregation process.

PLASMID PARTITION CASSETTES AND FAMILIES

The first identified plasmid partition systems and their associated genetic loci are those of the Escherichia coli P1 and F plasmid par systems. These discoveries were made over 25 years ago by the Hiraga and Austin groups [11,24,25]. Since then, multiple par cassettes have been found experimentally and by bioinformatics approaches in plasmids from very diverse bacteria [7]. Most of these cassettes have very similar genetic organizations whereby they consist of three elements: a cis-acting centromere-like site and two trans-acting proteins: a centromere-binding protein and a motor or force-generating protein (Figure 1B). Although the force-generating proteins were all originally found to be ATPases, recent work shows that tubulin-like GTPases and, possibly, coiled-coil proteins can also function as the force-generating protein to drive partition. To reflect this, these proteins are referred to as motor proteins. The upstream gene in the operon encodes the motor protein, whereas the downstream gene encodes the centromere-binding protein (Figure 1B). The centromere-like site itself is typically located either upstream or downstream of the par operon [3–11]. An important property of par cassettes is their ability to stabilize heterologous replicons, which means that they can act independently of the replication system and still mediate plasmid separation, ensuring plasmid maintenance.

An analysis of all the par systems that had been identified to date by Gerdes et al. in 2000 [7] led to their classification into two main types. This categorization was based on the kind of motor protein that is present, as the centromere-binding proteins exhibit significant sequence diversity. Accordingly, the par motor proteins can readily be broken into two main groups. In these two groups, the motor protein belongs to one of two families of ATPases: one with a deviant Walker box motif (also called P-loop) and the second containing an actin-like fold. In the present review, this categorization of par systems as suggested by Gerdes et al. [7] has been used because it provides a useful framework for describing and comparing Par protein structure. More recently, additional putative par systems that are unrelated to the two major par families have been identified. However, structures of these proteins have yet to be determined. We shall begin by defining and describing the two par major families as well as two more recently discovered putative par systems, which we delineate as the type III and type IV systems, as this sets the stage for the ensuing discussion on partition biology and Par protein structure–function relationships.

Type I and II par systems

The type I are the most numerous of the par systems and contain motor proteins with Walker-type or P-loop ATPase motifs [5]. These ATPase proteins are typically called ParA, and their centromere-binding protein counterpart is called ParB. The type I par systems can be divided further into type Ia and type Ib on the basis of the genetic organization of the par operon and the limited sequence and size homologies of the Par proteins encoded by these systems (Figure 1B). In both type Ia and Ib systems, the gene for the Walker-type ATPase is located upstream of that for the centromere-binding protein. However, the location of the centromere site differs: for the type Ia systems, the centromere-like site is located downstream of the par operon, whereas, in the type Ib systems, the centromere is found upstream (Figure 1B).
The type Ia Par proteins are characteristically larger than type Ib proteins. Specifically, the type Ia ParA and ParB proteins typically consist of 251–420 residues and 182–336 residues respectively. In contrast, type Ib ParA and ParB proteins generally contain 208–227 residues and 46–113 residues respectively. Notably, the type Ia ParA proteins have a dual role in partition; they function in both partition and transcription autoregulation, whereas the type Ib Par proteins contain 236–336 residues, whereas the ParR proteins, of the type Ib systems, wherein the centromere-binding protein ParB, whereas the F par system encodes an ATPase called SopB and a centromere-binding protein named SopB. Studies on these systems revealed that the ParB and SopB proteins bind to specific centromere sites and form higher-order complexes called partition complexes, which recruit the ParA/SopA proteins for completion of the partition reaction [30–32]. More recently, Hayes and co-workers have carried out an extensive analysis of the type Ib par system from the Salmonella Newport TP228 plasmid [33–35]. These studies revealed that the TP228 par system contains functional analogues of ParA and ParB, called ParF and ParG respectively [33–35].

The second type of par system, termed type II, contains ATPase proteins called ParM (for motor) that belong to the type Ib Walker-box ATPase proteins. The TP228 par system contains consensus repeats that are 7 bp. The type II par systems are sequentially diverse compared with the type I centromere-binding proteins. The most conserved region among the type Ia ParB proteins is a predicted HTH motif. The type Ib centromere-binding proteins display even more distinct sequences than the type Ia ParB proteins and contain no identifiable DNA-binding motif(s). Indeed, it was not until recent structures were solved that the DNA-binding motifs of these proteins were revealed.

The type Ia par loci from the E. coli P1 and F plasmids are among the best characterized par systems [10, 11, 24, 25]. The P1 par system contains the prototypical ParA ATPase and the centromere-binding protein ParB, whereas the F par system encodes an ATPase called SopB and a centromere-binding protein named SopB. Studies on these systems revealed that the ParB and SopB proteins bind to specific centromere sites and form higher-order complexes called partition complexes, which recruit the ParA/SopA proteins for completion of the partition reaction [30–32]. More recently, Hayes and co-workers have carried out an extensive analysis of the type Ib par system from the Salmonella Newport TP228 plasmid [33–35]. These studies revealed that the TP228 par system contains consensus repeats that are 7 bp. The type II par systems are sequentially diverse compared with the type I centromere-binding proteins. The most conserved region among the type Ia ParB proteins is a predicted HTH motif. The type Ib centromere-binding proteins display even more distinct sequences than the type Ia ParB proteins and contain no identifiable DNA-binding motif(s). Indeed, it was not until recent structures were solved that the DNA-binding motifs of these proteins were revealed.

The type Ia par loci from the E. coli P1 and F plasmids are among the best characterized par systems [10, 11, 24, 25]. The P1 par system contains the prototypical ParA ATPase and the centromere-binding protein ParB, whereas the F par system encodes an ATPase called SopB and a centromere-binding protein named SopB. Studies on these systems revealed that the ParB and SopB proteins bind to specific centromere sites and form higher-order complexes called partition complexes, which recruit the ParA/SopA proteins for completion of the partition reaction [30–32]. More recently, Hayes and co-workers have carried out an extensive analysis of the type Ib par system from the Salmonella Newport TP228 plasmid [33–35]. These studies revealed that the TP228 par system contains consensus repeats that are 7 bp. The type II par systems are sequentially diverse compared with the type I centromere-binding proteins. The most conserved region among the type Ia ParB proteins is a predicted HTH motif. The type Ib centromere-binding proteins display even more distinct sequences than the type Ia ParB proteins and contain no identifiable DNA-binding motif(s). Indeed, it was not until recent structures were solved that the DNA-binding motifs of these proteins were revealed.
places the gene encoding the centromere-binding protein before the motor protein, is distinct from the type I and II par systems [56]. A possible centromere site for the *B. thuringiensis israelensis* subspecies has been identified and consists of four 12 bp direct repeats [57]. The TubR protein, which contains a putative HTH domain, was shown to bind specifically to these repeats. Perhaps the most interesting feature of this system, however, was the finding that the TubZ protein contains neither a deviant Walker box motif nor an actin-like fold and instead appears to be a member of the tubulin/FtsZ GTPase superfamily of motor proteins. In fact, it was clearly demonstrated that TubZ assembles into dynamic polymers that exhibit directional polymerization with plus and minus ends, indicating that plasmid partition of pBtoxis may be mediated by tubulin-like polymers [56]. How TubZ might attach to plasmids to drive segregation remains unclear. However, it appears that TubR can recruit TubZ polymers to the partition site, similarly to how the centromere-binding proteins of the type I and II par systems engage their requisite ATPase motor proteins. The recent finding that the *Bacillus anthracis* pXO1 plasmid RepX protein (now called TubZ-Ba) forms two stranded filaments identical with those formed by the *B. thuringiensis* TubZ protein (TubZ-Bt) supports the contention that there may be several type III par systems [58–60]. Indeed, at least four TubZ-like sequences have been found on different *Bacillus* plasmids [56].

Finally, a potential type IV partition system that utilizes only one protein has been identified on the *Staphylococcus aureus* plasmid pSK1 [61]. This protein, called Par, contains 245 residues and was shown to be critical for extending the segregational stability of the pSK1 plasmid, supporting the idea that it may function in partition. It seems unlikely that a second protein is present in the pSK1 par operon as only one significant open reading frame, belonging to *par*, has been identified in the coding region. How this protein may play both centromere-binding and motor functions is unknown. However, structural prediction studies suggest that Par contains an N-terminal HTH domain, which could function in centromere-binding, and a central coiled-coil domain, which might form polymeric structures similar to those formed by coiled-coil-containing proteins in the cytoskeleton of eukaryotic cells [62–64]. Therefore the pSK1 Par protein contains domains that could conceivably perform both roles in partition. Clearly, more studies are needed to elucidate the mechanism of pSK1 segregation and the involvement of the pSK1 Par protein in this process. Interestingly, Par homologues have been found on plasmids in a range of Gram-positive bacteria, including *Staphylococcus*, *Streptococcus*, *Lactococcus*, *Lactobacillus*, *Clostridium* and *Tetragenococcus*, suggesting that such systems are not uncommon [65].

**Bacterial chromosomal par systems**

It is interesting that, although the mechanism(s) of DNA segregation utilized by eukaryotes and plasmids have been delineated in general terms, the means by which most bacteria partition chromosomes is still largely a mystery. This is notably true for the model Gram-negative organism, *E. coli*. Indeed, no partition proteins with homology with plasmid Par proteins have been identified in *E. coli* genome searches, nor have any Par-like proteins been isolated from genetic screens [66]. However, DNA segregation of some bacterial chromosomes do involve Par protein homologues and these proteins are surprisingly similar to their plasmid counterparts. The best-studied examples of chromosomal par loci are from *Bacillus subtilis*, *Caulobacter crescentus*, *Pseudomonas putida* and *Pseudomonas aeruginosa* [67–70]. Although the present review focuses on studies on plasmid partition machinery, we shall briefly mention some of these examples of bacterial chromosomal DNA segregation by plasmid-like Par proteins.

Bacterial chromosomal par systems appear to be hybrids of plasmid Type Ia and Ib systems in that the chromosomal ParA proteins contain deviant Walker box motifs of the small Type Ib category, whereas the corresponding chromosomal centromere-binding proteins contain HTH motifs and are very similar to plasmid type Ia centromere-binding proteins [7]. Initial indications that Par-like proteins might be involved in chromosome segregation in bacteria came from genetic studies showing that mutation of the chromosomally encoded *B. subtilis* ParB homologue, spo0J, resulted in a 100-fold increase in the production of anucleate cells [71]. Found on the same operon as spo0J is a gene, spoJ, which encodes a deviant Walker box-containing protein.

Initial studies on Soj mutants suggested that it is not involved in DNA segregation [71,72]. However, more recent data show that Spo0J and Soj together play an important role in bringing together centromere sites that are far apart on the *B. subtilis* chromosome and organizing the origin region into a compact structure that facilitates separation of replicated origins [73]. To do this, Spo0J, like its plasmid counterparts, binds a conserved DNA site, numerous copies of which are clustered about the spo0J operon and the replicated chromosomal origins.

There are only a few cases in which chromosomal Par proteins have been shown to be essential for bacterial DNA partition under normal cellular conditions [74–76]. In *C. crescentus*, the ParB and ParA homologues are required for chromosome segregation, and inactivation of either one is lethal. Another example is chromosome II of *Vibrio cholerae*. *V. cholerae* contains two chromosomes and each has a parAB system. Recent studies have found that segregation of chromosome II is absolutely dependent upon its parAB system [77,78]. It is interesting to speculate that this chromosome may actually represent a co-opted plasmid that has taken on a role as a host chromosome. In contrast with these examples in which par systems are essential for DNA segregation, some bacterial par systems are important only under specialized conditions. For example, deletion of the *Streptomyces coelicolor* par system causes serious chromosome defects only during the formation of chain spores [79]. Consistent with this, parAB transcription is greatly stimulated during *S. coelicolor* sporulation [79]. In *P. putida*, the parAB locus plays a vital role in DNA segregation during the transition from exponential to stationary phase in cells grown in minimal medium [80,81]. Thus these combined data suggest that plasmid Par homologues play important roles in DNA segregation in some bacteria, whereas they may be important only under certain conditions in others. Consistent with this supposition, more than 50 homologues of ParA and ParB have been identified in bacteria thus far [82].

**CELLULAR AND BIOCHEMICAL STUDIES REVEAL PLASMID SEGREGATION STEPS**

Cellular and biochemical studies have provided a general outline of the steps involved in the plasmid partition process. The important findings from these studies will be summarized next to set the stage for the discussion of structural studies, which, in turn, have begun to elucidate the detailed molecular mechanisms involved in DNA segregation.

**First step: formation of the partition complex**

The first step of the partition reaction involves the co-operative binding of the centromere-binding proteins to their cognate centromere site. Just as eukaryotic centromeres serve as attachment
sites for spindle microtubules, plasmid segregation depends on the creation of a defined nucleoprotein complex, called the partition complex, which is formed by the accurate assembly of centromere-binding proteins on to the centromere site. As noted, little homology exists between centromere-binding proteins, and the centromere-like sites that they bind are also diverse. However, most plasmid centromere-like sites share an important characteristic, which is that they consist of multiple DNA repeat elements (Figure 2). The repeats can be extensive, as exemplified by the F plasmid centromere, sopC, which contains 12 tandem repeats of a 43 bp element [30,83]. Also, the parS centromere of the Agrobacterium tumefaciens pTAR plasmid contains 13 repeats that are separated by an integral number of turns of the DNA double helix [84]. The type II R1 parC centromere has a complex arrangement in that it consists of two sets of 11 bp repeats that are separated by a 39 bp region that includes the parMR promoter [13]. But the most complex centromere is the parS centromere of the P1 plasmid. Similar, albeit less well studied, centromeres are found on related plasmids such as the E. coli P7 plasmid [85–88]. These centromeres contain a central binding site for the E. coli protein IHF (integration host factor) (Figure 2). IHF is an architectural protein that bends DNA by ~180° [89]. On either side of the IHF site in the P1 centromere are asymmetrically arranged DNA-binding motifs called A-boxes and B-boxes that are recognized by the P1 ParB protein. Not only are the boxes asymmetrically arranged, they also differ in number on each arm whereby the ‘left’ arm (left of the IHF-binding site) contains one B-box followed by a single A-box, whereas the ‘right’ arm contains one B-box sandwiched between two A-boxes on the left and one A-box on the right (Figure 2). Less complicated centromeres include the plasmid TP228 parH centromere, which consists of four direct repeats and the type II pSK41 centromere, parC, which also contains four tandem DNA repeats [35,90] (Figure 2).

Despite the apparent diversity among centromeres and centromere-binding proteins, in all cases it appears that the protein–centromere interaction leads to the formation of higher-order partition complexes. The higher-order structure of the partition complex is postulated to be important in the capture and activation of the motor protein for DNA segregation [90]. The strongest evidence for higher-order protein–DNA structure, including DNA wrapping, in formation of the partition complexes comes from studies on the interactions between the centromere-binding proteins SopB and R1 ParR with their centromere sites. The dramatic effect on DNA topology caused by formation of the SopB–sopC partition complex is consistent with a structure in which the centromere DNA is wrapped in a right-handed manner around a multimeric SopB protein core [91]. The R1 parC centromere itself appears to be intrinsically bent, and ParR binding leads to further distortion of the site [92]. Similarly, the centromere of the Enterococcus faecium gentamicin-resistance plasmid pGENT is also intrinsically curved [93]. In this regard, it is notable that analogous structural features are associated with eukaryotic centromeres. Specifically, studies indicate that there is inherent curvature in Saccharomyces cerevisiae centromere DNA sites and binding of yeast centromere-binding proteins with these DNA elements promotes DNA wrapping into higher-order complexes [92–94]. These combined findings indicate that the specific architecture of the higher-order partition complex plays a critical role in its function. In the case of partition systems, the function would probably be assembly of the motor protein to form a fully active segregosome. Indeed, as will be described, the recent structure of the pSK41 ParR–centromere complex shows a highly wrapped protein–DNA superstructure, with overall dimensions ideal for ParM filament binding [90].

Recruitment of the motor protein is thought to take place after the two plasmids have paired. The pairing event is apparently mediated by interactions between centromere-binding proteins and/or centromere sites located on the partition complexes of the replicated plasmids. Pairing was initially suggested by the finding that plasmid centromeres act as incompatibility determinants. This means that plasmids that replicate or segregate using similar mechanisms cannot coexist in the same cell [95]. In contrast, plasmids with distinct partition systems are compatible, can coexist and localize to alternative subcellular positions. Although pairing has been somewhat controversial, it has been demonstrated in vitro for the E. coli plasmids R1 and pB171 and in vivo for the P1 plasmid [8,96–98].

Second step: DNA separation as mediated by motor proteins

After plasmid pairing, the motor protein is recruited to create the active segregosome and mediates plasmid separation. The mechanism(s) by which motor proteins carry out this feat was initially unclear, especially for the Walker-type ParA ATPase motor proteins. To be activated for partition, both type II ParM proteins and type I Walker-type proteins must bind ATP. As noted, for the type II ParM proteins, ATP binding was found to stimulate filamentation, thus providing an explanation for the ATP requirement and also providing a model for plasmid separation. According to this model, growth of ParM filaments anchored between two ParR/parC-containing plasmids pushes or pulls the plasmids apart. Interestingly, a recent study reported that GTP binding to ParM stimulates the formation of filamentous structures identical with those formed in the presence of ATP [99]. Whether ATP or GTP, or a mixture of these nucleotides, functions in type II partition in vivo remains to be resolved.

The Walker-type ATPase proteins also require ATP binding for partition [26–28]. How ATP binding stimulates segregation by the type I motor proteins, however, has been less clear than for the type II motor proteins. Initial models, based primarily on knowledge regarding other Walker-type proteins, suggested that type I motor proteins are monomers in the absence of ATP and form nucleotide sandwich dimers upon ATP binding [26]. Dimerization is mediated via cross-contacts from one subunit to the ATP molecule bound in the second subunit. However, it is difficult to speculate how such a simple monomer to dimer switch could cause the large-scale propulsion of paired plasmid DNA from the cell centre to opposite cell poles. More recent structural and biochemical data, which have revealed that these Walker-type proteins can form polymers upon ATP binding, suggest that they actually function in a manner similar to the ParM proteins in the type II par systems [3,6]. The ability of Walker-type proteins to form polymers is not unprecedented. Indeed, the Walker-type protein MinD forms polymers upon binding ATP. The MinD protein is part of the MinCDE system that is involved in cell division control, and ATP binding to MinD causes the release of its C-terminal region, which forms an amphipathic helix when localized to the cell membrane. Once attached to the membrane, multiple MinD proteins coalesce and form long filaments [100,101]. Along with MinC, the role of MinD is to prevent formation of the cell division septum at random locations in E. coli, particularly near the cell poles. The inhibition of septum formation by MinCD causes a rapid oscillation of this complex between the cell poles. This remarkable oscillation correlates with the polymerization of the MinD protein [101].

The theory that Par Walker-type proteins utilize filaments formation for segregation has been somewhat controversial. However, data showing that several type Ia and Ib Walker-type proteins can form polymers provide strong support for the idea that
this capability is important functionally. The pB171 ParA, which is a member of the type Ib ParA family, was shown to both oscillate over the nucleoid and form polymers [52,102]. This led to a model for pB171 partition in which polymerization by pB171 ParA pushes the plasmids apart by the insertion of ParA–ATP molecules between the ParA–ParB/parC interface, similar to the insertional polymerization model proposed for type II segregation of the R1 plasmid [102]. However, an important feature of this model, called the ‘sweeping par model’ is that the force exerted by pB171 ParA is assumed to be proportional to the number of filaments involved in forming the polymer at that point. This model differs from the type II insertional polymerization model in some ways. Perhaps the major difference is that, unlike ParM, which forms non-oscillating filaments that separate plasmids paired by ParR and pushes them to cell poles, oscillating pB171 ParA positions plasmids near quarter cell locations. As a result, whereas the type II R1 reaction is binary because the mitotic machinery can only accommodate one pair of plasmids at a time, each attached at the ends of the growing ParM filament, the pB171 ParA can position many plasmids at a time because of the oscillating nature of its polymer, and so all plasmids present at a given moment can participate in the segregation process. Notably, this model is consistent with the observed higher efficiency of type I loci compared with type II. Another type Ib motor protein that has been shown to form polymers is the TP228 ParF protein [34]. Similarly to pB171 ParA, the polymerization of ParF is dependent on ATP and is stimulated by the presence of its centromere-binding protein [34].

The *E. coli* type Ia ParA protein SopA has also been shown to form filament-like structures [18,103,104]. Several detailed mechanisms have been posited for SopA-mediated segregation based on its ability to form polymers [103,104]. One model proposes that SopA forms asters that project from the partition complex [103]. It is hypothesized that the asters not only function to push the plasmids apart, but also to position the plasmids at midcell before replication and maintain the position of daughter plasmids after separation. Thus, according to this model, the pushing and/or pulling forces that are imparted by the projecting asters within the cell would allow for the correct positioning of the plasmids in the cell. A second model, based on studies in which there was no evidence for aster formation, proposes that nucleoid DNA inhibits SopA polymerization. In this model, increased concentrations of SopB bound to sopC alleviate the DNA inhibition of SopA polymerization, allowing polymers to attach to SopB-sopC partition complexes [104]. Interestingly, in the studies used to formulate the aster hypothesis, there was no evidence for DNA inhibition of SopA polymerization. A third study, which shows evidence for SopA oscillation between cell poles, suggests that SopA filaments function as a railway track for the movement of the plasmids [18]. Despite the differences in detail, all of these models propose that F plasmid segregation is powered by SopA polymers, which are formed in a manner stimulated by ATP and SopB/sopC.

**STRUCTURAL STUDIES ON PARTITION PROTEINS**

Cellular and biochemical studies have provided a general outline of the plasmid partition process. However, multiple questions regarding the structure–function relationship of Par molecules remain and must be addressed in order to provide a complete and detailed understanding of the mechanisms employed by the Par proteins to segregate DNA. Outstanding questions regarding the first step include: what types of DNA-binding motifs do centromere-binding proteins utilize; how do these motifs recognize centromere repeats; how do multiple centromere-binding proteins bind co-operatively to multiple repeats to form a higher-order partition complex; and what kind of structures are adopted by these higher-order complexes? In terms of the second step of plasmid partition, important questions include: what structures are adopted by the putative Walker-type proteins (do they really contain canonical Walker-type folds?); how does ATP binding lead to a ‘switch’ in conformational states that ultimately results in plasmid separation; how does interaction with the partition complex stabilize and possibly stimulate this process; and what role does ATP hydrolysis play in plasmid separation? As discussed below, recent structural information has started to shed light on some of these questions.

**Centromere-binding proteins: sequence diversity belies structural homology**

Within the last year, new structures have been obtained for several centromere-binding proteins. As a result, structures are now available for each type of centromere-binding protein: type Ia, type Ib and type II. Remarkably, these combined structures indicate that, despite the significant lack of sequence homology among these proteins, they share common DNA-binding folds. In fact, it appears that there are two primary structural classes of centromere-binding proteins: those that contain HTH folds, which are of the type Ia family, and those that contain RHH (ribbon–helix–helix) DNA-binding motifs, which unexpectedly include both the type Ib and type II centromere-binding proteins.

**Structures of type Ib centromere-binding proteins**

The type Ib centromere-binding proteins are typically small, consisting of 46–113 residues. The sequences of these proteins provided no hint as to how centromere binding is mediated. However, recent structures of these proteins and their complexes with DNA have revealed the DNA-binding motif they utilize and how they bind DNA (Figure 3).

**Structure of TP228 ParG protein**

One of the best-characterized type Ib par systems is that of the *E. coli* multidrug-resistance plasmid, TP228. The TP228 centromere-binding protein, called ParG, is a 76 residue, 8.6 kDa dimer that is essential for TP228 plasmid retention. The structure of apo ParG was solved by NMR and revealed that it contains a bipartite structure with an N-terminal disordered region (residues 1–32) and a C-terminal region (residues 33–76) that has a RHH DNA-binding motif [105] (Figure 3B). The RHH topology is β1, residues 34–41; α1, residues 43–55; α2, residues 60–75. The RHH is a very common DNA-binding motif present in the Arc/MetJ superfamily of DNA-binding proteins [106]. In fact, it is predicted that there are over 2000 RHH-containing proteins in various bacterial sources, 55 in bacteriophages and 300 such proteins in archaea [106]. Thus far, the RHH fold appears to be absent from eukaryotes. However, RHH folds are difficult to predict without experimental data as sequence conservation is strikingly low among these proteins and there is a general lack of strictly conserved residues within the fold. As was noted, this was the case for the centromere-binding proteins. Further impeding the identification of RHH proteins is the fact they regulate multiple diverse cell processes such as transcription, the lytic cycle of bacteriophages and cell division [106]. However, all RHH proteins whose structures have been solved to date form dimers-of-dimers upon DNA binding, whereby each β-strand inserts into consecutive major grooves. The defining characteristic of the RHH motif is the utilization of β-strand residues for making
Structure and function of plasmid partition proteins

Figure 3  Structures of type Ib centromere-binding proteins

(A) Schematic diagram showing the domain structure of the type Ib centromere-binding proteins. The N-terminal domain (grey) is a flexible region that functions in ATPase stimulation, ATPase recruitment and pairing. In addition, a role in DNA binding has been suggested. The C-terminal domain (magenta) contains the DNA-binding RHH motif. In some cases, there is a short region following the RHH (shown in grey). The RHH motifs of the structures in (B) and (C) are shown in the same orientations for comparison purposes. (B) Structure of the ParG protein from the E. coli TP228 plasmid shown colour-coded as in (A) [105]. The protein is a dimer. The N-terminal flexible region is grey, and the RHH, which mediates all the dimer contacts, is magenta. (C) Structure of Streptococcus pyogenes plasmid pSM19035 ω protein colour-coded as in (A) [110,111]. On the left is the apo structure, and on the right is the ω–DNA structure. The DNA is shown as sticks and is coloured blue. N- and C-termini regions are labelled for one subunit in each apo structure. (B) and (C) were modelled using PyMOL (DeLano Scientific).

In addition to binding their centromere sites to form higher-order partition complexes that recruit and stabilize motor proteins, centromere-binding proteins also stimulate the ATPase activity of the motor protein. This presents a conundrum as ATP hydrolysis leads to depolymerization of ATPase motor proteins. To date, it is not understood at a molecular level how these two opposing functions are mediated and co-ordinated to achieve accurate plasmid separation and this remains one of the important questions in partition biology. In studies aimed to address this issue, Hayes and co-workers have shown that there are two separate regions within the flexible N-terminal region of ParG that mediate these activities: one acts to promote polymerization of the Walker-box protein ParF and the other to enhance its ATPase activity 30-fold [107]. Specifically, they identified residue Arg19 as part of a putative arginine finger motif within the N-terminal region of ParG that is responsible for stimulation of the ParF ATPase activity, but which is dispensable for the ParG-mediated affects on ParF polymerization. Arginine fingers are common motifs utilized by P-loop NTPases, either in cis, where it is from the same protein, or trans, where, similar to the case of ParG and ParF, it is located on another protein [108]. These motifs, which are typically found on extended loop regions on a protein, stimulate nucleotide hydrolysis by inserting into the active site and stabilizing the transition state. Because ATPase-interacting regions located on type Ia and Ib centromere-binding proteins are usually located in extended regions and contain arginine residues, an attractive hypothesis, proposed by Hayes and co-workers, is that all these proteins contain arginine finger motifs that are used in trans to enhance the ATPase activity of their requisite ATPase motor proteins [107].

Remarkably, two additional functions have been posited for the N-terminal domains of type Ib proteins: DNA binding and pairing. Specifically, studies on ParG indicate that residues in its N-terminal region can affect how it organizes the formation of higher-order ParG–DNA structures [109]. Moreover, the continued deletion of this N-terminal tail greatly perturbs its DNA-binding kinetics [109]. Additionally, studies on the ω protein from the Streptococcus pyogenes plasmid pSM1903, another type Ib centromere-binding protein with a RHH fold, revealed that removal of its entire N-terminal tail causes a 2-fold reduction in its DNA-binding affinity [110]. Indications that the N-terminal region of type Ib centromere-binding proteins impact pairing come from studies on the pB171 ParB protein, which showed that removal of its entire N-terminal tail abrogates ParB–ParB interactions between plasmids [53]. Combined, these data demonstrate that the flexible N-terminal regions of type Ib centromere-binding proteins play several pivotal roles in partition (Figure 3A). However, the precise locations of these functional regions within the N-terminal domains of these proteins have yet to be mapped, and their specific molecular mechanisms have yet to be resolved.

Structure of inc18 ω protein in the presence and absence of DNA

In addition to ParG, the structure of the ω protein from the Streptococcus pyogenes plasmid pSM19035 found in the Gram-positive broad host range inc18 family of plasmids has also been
determined [110,111]. Like most type Ib centromere-binding proteins, ω functions in transcriptional autoregulation. On the other hand, ω is distinct from other type Ib centromere-binding proteins in that it also acts as a global regulator of transcription [111]. To carry out these functions, ω binds co-operatively to DNA sites containing seven to ten consecutive non-palindromic DNA heptad repeats (5'-ATCACA/T-3'). Interestingly, these sites can be arranged as direct or inverted repeats. ω binds with low affinity (Kₘ of >500 nM) to one heptad, but it binds strongly to DNA sites containing two or more consecutive heptad repeats [110,111]. The heptad arrangement also affects the affinity of ω for the DNA site, with direct or converging repeats providing the strongest binding interaction.

Structural and biochemical studies on the Streptococcus pyogenes ω protein revealed that, similarly to ParG, it has a bipartite structure with an N-terminal flexible region and a C-terminal domain containing a RHH motif (Figure 3C). The flexibility of the N-terminal region is reflected in the structure of apo ω, whereby the N-terminal 20 amino acids were lost to proteolysis during crystallization [111]. The C-terminal RHH has the topology of β, residues 28–32; α1, residues 34–47; and α2, residues 51–66. To gain insight into the unusual DNA-binding preferences exhibited by ω, structures were solved in the presence of direct and converging 17 bp DNA repeats using a deletion mutant of the ω protein, Δ19ω, in which the first 19 residues were removed. As noted, the N-terminal region has been implicated to be important in DNA binding, and this deletion results in a 2-fold reduction in DNA binding compared with wild-type protein [110,111]. However, the role(s) played by the N-terminal tails of type Ib centromere-binding proteins in DNA binding has yet to be determined as, so far, the Δ19ω–DNA complexes are the only structures available for a type Ib centromere-binding protein bound to cognate DNA.

The ω–DNA structures show that ω binds its DNA as a dimer-of-dimers, wherein each dimer binds a consecutive DNA major groove. Unexpectedly, the crystallographic asymmetric unit contains a Δ19ω dimer-of-dimers bound to a 17 bp DNA site, which in turn is arranged pseudocontinuously with a second uncomplexed 17 bp DNA site. As a result, there is an unbound 17 bp DNA site on either side of each ω–DNA complex in the crystal lattice (Figure 3C). The DNA bound by ω is essentially straight and B-DNA in character. The presence of the uncomplexed operator in the structure permitted the direct comparison of protein-bound and protein-free DNA conformations. This analysis showed that the unbound DNA adopts the same conformation as that bound by Δ19ω. Therefore any deviations from ideal B-DNA observed in the structure of the Δ19ω-bound DNA are the likely result of nucleotide sequence-induced alterations. Comparison of the structures of ω bound to direct and converging repeats showed that the protein can bind these different sites because the contacts of ω to the two DNA sites are related by a pseudo-2-fold rotation axis, and, as a result, the specific protein–DNA interactions are comparable, despite the heptad orientation [110]. Presumably, the formation of the higher-order ω–DNA partition complex involves the co-operative interaction of multiple ω dimers-of-dimers binding to consecutive DNA heptad repeats. However, confirmation of this awaits structural studies on ω and other type Ib centromere-binding proteins bound to full-length centromere sites.

Structures of type II ParR proteins

Structures for type II ParR centromere-binding proteins have only become available within the last year. Similarly to the type Ib centromere-binding proteins, the amino acid sequences of the type II centromere-binding proteins revealed no clues about the structures these proteins adopt. However, these proteins are similar in size to the type Ib centromere-binding proteins. Thus it was intriguing when the structures of the pB171 and the pSK41 type II ParR proteins were solved and shown to belong to the same family of DNA-binding proteins as the type Ib centromere-binding proteins, namely the RHH family (Figure 4).

Structure of pB171 ParR protein

The E. coli pB171 ParR protein functions as the centromere-binding protein for the type II locus on the pB171 plasmid, which contains both type I and II par systems [48–53]. The structure of the full-length pB171 ParR protein was recently solved in the absence of DNA [112]. The C-terminal 35 residues were not resolved in this structure, neither were residues 1–5. The structure, consisting of residues 6–95, revealed a RHH motif and a C-terminal domain composed of a three helix ‘cap’ (Figure 4B). The topology is β1, residues 7–14; α1, residues 20–30; α2, residues 36–52; α3, residues 56–65; α4, residues 71–81; and α5, residues 88–94. The three-helix cap at the C-terminal region not only reinforces the tight dimerization of the RHH, but also stabilizes interactions between dimers-of-dimers. Several RHH proteins contain domains outside the RHH motif that similarly act to bolster the dimer and dimer-of-dimer; however, none of these structures is similar to the pB171 three-helix cap. A notable finding in the crystal structure was the creation in the lattice of a continuous helical structure created by the assembly of 12 ParR dimers per turn. In this arrangement, the N-terminal RHH motifs face one direction, whereas the C-termini point in the opposite direction. This was interesting because EM (electron microscopy) images of the R1 ParR protein bound to its centromere site parC, revealed that the ParR–parC partition complex forms ring-like structures on pre-linearized DNA that are between 15 and 20 nm in diameter [112]. Because R1 and pB171 ParR proteins probably form similar higher-order partition complexes, it was postulated that the EM projections of R1 ParR–parC were possibly hiding the helical nature of the arcs observed in the pB171 ParR apo structure. However, it was not possible to distinguish the exact nature of the superstructure formed by the R1 or pB171 ParR–DNA partition complexes by these data. The resolution to the issue was obtained by the crystal structure of the type II pSK41 ParR–centromere complex, which revealed the first structure of a higher-order partition complex [90].

pSK41–ParR–centromere complex reveals partition complex structure

pSK41 is a prototypical multiresistance plasmid from Staphylococcus aureus and is stably maintained by the presence of a type II partition system encoding ParR and ParM proteins [113]. The pSK41 Par proteins show little sequence homology with R1 ParR and ParM. Nonetheless, the pSK41 ParM protein is readily identifiable as an actin-like protein, thus establishing the system as a type II par system. The pSK41 centromere site was identified by a combination of footprinting and FP (fluorescence polarization)-based DNA-binding studies [90]. These analyses revealed that the centromere consists of four 20 bp elements, in which the 20 bp repeat constitutes the minimal element for binding of a ParR molecule (Figure 2). The simplicity of this centromere site made it an ideal candidate for attempts to obtain the structure of a higher-order partition complex. As data-quality crystals of the pSK41-full-length ParR–centromere complex could not be obtained, the protein was subjected to limited
Figure 4 Structures of type II centromere-binding proteins

(A) Schematic diagram showing the domain structure of the type II centromere-binding proteins. The N-terminal domain (magenta) forms a RHH domain responsible for centromere binding. The C-terminal region is a partially flexible domain that functions in ParM binding and also contributes to the stabilization of the RHH dimer and dimer-of-dimers. The RHH motifs of the structures in (B) and (C) are shown in the same orientations for comparison purposes. (B) Structure of the E. coli pB171 ParR colour-coded as in (A) [112]. The N-terminal RHH domain is coloured magenta, and the C-terminal region, which consists of a three-helix cap and additional disordered residues, is grey. N- and C-terminal regions are labelled. (C) Structure of the Staphylococcus aureus pSK41 ParR protein coloured as in (A) [90]. On the left is the protein in the absence of DNA, and on the right is the ParR–centromere complex. N- and C-termini regions of one subunit are labelled. (B) and (C) were modelled using PyMOL (DeLano Scientific). An interactive three-dimensional version of the structure shown in (C) can be seen at http://www.BiochemJ.org/412/0001/bj4120001add.htm.

proteolysis, and the minimal DNA-binding domain was found to reside in the first 53 residues of the 109 residue protein. This domain, termed ParRN, contains all of the determinants required for centromere binding as ascertained by FP-binding studies [90]. The structure of the ParRN–centromere was determined by utilizing a 20 bp repeat, which stacked pseudointerminously in the crystal to form the full-length centromere (Figure 4C).

The structure revealed that, like pB171 ParR, pSK41 ParR contains a RHH motif in its N-terminal DNA-binding domain with topology of $\beta_1$, residues 5–12; $\alpha_1$, residues 16–25; and $\alpha_2$, residues 33–47. Like all RHH proteins, the $\beta$-strands combine in an antiparallel fashion and helices interdigitate to form the extensive RHH dimer. The DNA elements bound by each ParR dimer-of-dimers stack pseudointerminously to generate the full-length centromere. Thus the most notable feature of the structure is the creation of the higher-order partition complex that mediates full segregation assembly (Figure 4C). Consistent with the cooperative nature of ParR binding, the protein dimer-of-dimers that bind the DNA form intimate protein interactions that lead to the generation of a continuous protein superstructure that has distinct positive and negative electrostatic surfaces. As would be expected, the positive face wraps the centromosomal DNA about itself to create the unique superstructure that is best described as a large superhelix. Characteristic of the superstructure is its superhelical parameters: it has a pitch of $\sim 24$ nm with six ParR dimers-of-dimers in one turn of the superhelix and a measured diameter of $\sim 18$ nm. Within this structure, the C-terminal regions of the ParR protein would face inwards, towards the pore, which suggests that it plays a role in ParM recruitment. This was confirmed by binding experiments, which showed that ParM bound only to complexes containing full-length ParR and not ParRN [90].

DNase I-protection studies carried out on the pSK41 ParR binding to the centromere revealed striking nucleosome-like periodicity in the protection pattern consistent with the crystal structure. As would be expected from binding a continuous protein superstructure, the footprint showed almost complete protection. Indeed, only eight symmetrically disposed adenine nucleotides on each strand are susceptible to DNase I attack. The basis for the susceptibility of these nucleotides is revealed by the crystal structure, which shows that these adenines reside in the only minor groove positions that are exposed outside the partition complex superhelix and between ParR molecules that bind each repeat (Figure 4C). Thus the multiple features of the DNase I footprint are readily explained by the crystal structure. In this same study, cryo-EM was also carried out whereby the partition complex formed by binding of ParR to the centromere site with no surrounding DNA regions was examined. The image revealed a circular superhelical structure with dimensions identical with those observed in the crystal structure.

The structural studies on the pSK41 ParR–centromere complex resolved the issue of whether R1 and pB171 ParR–centromere complexes formed rings or helices as, clearly, the structure is a superhelix [90,112]. Indeed, the measured diameter of the pSK41 superhelical partition complex crystal structure was $\sim 18$ nm, which closely matches the 15 and 20 nm diameter obtained for the superhelical rings of the R1 ParR–centromere complex by EM [112]. These data, which indicate a conserved structure of the type II segregosome assembly, seem to support the notion that the higher-order structure of the partition complex is likely to be crucial to its function of recruitment of ParM filaments and their stabilization. Indeed, as discussed below, recent EM data on ParM filaments combined with the structure of the pSK41 partition...
complex provide insight into the role played by the partition complex in ParM filament recruitment and final segrosome assembly.

Structure of type Ia ParB centromere-binding proteins

It had been predicted that the type Ia ParB proteins contain HTH motifs. Studies on the type Ia centromere-binding protein, SopB, also clearly showed that these proteins form higher-order segrosomal structures. The DNA-binding properties of these proteins that enable such higher-order structure formation are starting to be resolved with recent structure determinations of type Ia centromere-binding proteins.

All plasmid centromere-binding proteins share in common the fact that they are multidomain proteins that contain flexible extended regions, usually at their N- and/or C-termini. Yet, the type Ia centromere-binding proteins clearly differ from those in the type Ib and II families. The type Ia proteins are larger than the type Ib and II proteins and typically contain between 182 and 336 residues. Sequence comparisons indicate that the type Ia centromere-binding proteins have two conserved regions, which correspond to a domain called motif B and a HTH unit. Motif B is near the N-terminal region and the HTH unit is centrally located (Figure 5A). The linker domains of type Ia proteins show limited sequence and, as it turns out, structural homology (Figure 5A). Like the type Ib and II centromere-binding proteins, all the type Ia ParB proteins that have been characterized form dimers whereby the dimer domain is located in the C-terminal ∼70 residues. The C-terminal dimer domain is one of the least conserved domains among type Ia members. As will be described, this is reflected in the finding that this domain can adopt very different structures. In fact, in one case, this domain functions in DNA binding as well as dimerization.

Structure of Thermus thermophilus chromosomal protein Spo0J

The first described structure of a type Ia centromere-binding protein was that of the N-terminal domain of the T. thermophilus Spo0J, which is a chromosomally encoded type Ia protein [114]. Attempts to crystallize the Spo0J protein from B. subtilis and eight other bacterial genomes all failed. Only a truncated form consisting of residues 1–222 of the 269 residue T. thermophilus Spo0J was successfully crystallized. This fragment contains the regions of highest homology among ParB proteins, motif B and the HTH unit. Spo0J homologues typically bind DNA sites containing multiple 16 bp repeats located in the origin proximal region of the requisite bacterial chromosome [69,115]. Gel-retardation assays revealed that the Spo0J-(1–222) fragment is greatly weakened for DNA binding, despite the fact that it contains the HTH unit [114]. This is consistent with data on other type Ia ParB proteins, which indicate that the C-terminal dimer domain is essential for high-affinity DNA binding. In most cases, this domain is not directly involved in DNA binding by type Ia centromere-binding proteins and thus it appears that dimerization
is important in mediating the proper docking of the HTH units on to successive grooves of the DNA.

The structure of Spo0J-(1–222) is predominantly helical (Figure 5B). Residues 1–22 and 210–222 were not visible, consistent with data indicating that the N-terminal domains of type Ia proteins are flexible. The structure can be broken into three main regions based on homology with other type Ia proteins. The region containing residues 23–118 corresponds to the motif B region and consists of four α-helices and a two-stranded β-sheet, with topology α1–α2–α3–β1–β2–α4 (Figure 5B). The HTH unit includes residues 119–160. The linker domain, which in the full-length protein connects to the dimer domain, includes residues 161–209. Although Spo0J-(1–222) does not contain the dimer domain, a clear dimer is observed in the crystal structure, which is formed by contacts primarily involving residues from motif B.

There are two main interaction interfaces within the dimer: one between residues from α1 of one subunit and residues from α3, β1 and β2 of the other subunit and the second between α2 of one subunit and helices 4, 5, 8 and 10 of the other subunit. Despite the large surface area buried in this dimer interface, biochemical data show that Spo0J-(1–222) only dimerizes when the protein is at high concentrations. This is consistent with studies on other type Ia ParB proteins, such as P1 ParB, which suggest that motif B is involved in secondary oligomerization when bound to DNA [86].

Structures of RP4 KorB

The structures described above underscore a conserved feature of centromere-binding proteins, their flexible multidomain nature. Following this general theme, the type Ia RP4 KorB protein had to be broken into domains in order for structures to be obtained. KorB is carried on the RP4 plasmid and is a member of the E. coli incompatibility group P (IncP-1α) [116,117]. This plasmid displays a broad host range and thus can be transferred to a variety of Gram-negative bacteria [117]. Structures have been solved for the N-terminal HTH-containing domain of KorB as well as the C-terminal dimer domain [118,119]. The KorB C-terminal dimer domain (residues 297–358) consists of a five-stranded antiparallel all-β structure with striking homology with the SH3 (Src homology 3) domain found in eukaryotic signalling proteins (Figure 5D) [120,121]. However, unlike the SH3 domain, the KorB C-domain is missing a long loop between β1 and β2 that is essential for the SH3 recognition of proline-rich motifs. In contrast, the KorB C-domain contains a significantly elongated β5 compared with SH3 domain structures and is essential for KorB dimerization.

The KorB C-terminal domain is not directly involved in centromere binding. Instead, this function is fulfilled by the N-terminal region of KorB [119]. However, similarly to Spo0J, the C-terminal domain of KorB is critical for high-affinity DNA binding by KorB again probably by correctly positioning and anchoring the DNA-binding domains on cognate DNA. Insight into how KorB binds its individual centromere-like palindromic repeats was provided by the structure determination of the KorB DNA-binding domain (residues 101–294) bound to one of its 12 operator/centromere-like sites, O6 [120]. The structure of the KorB DNA-binding domain consists of eight helices, including the C-terminal portion of motif B (coloured cyan in Figure 5D), the DNA-binding HTH unit (magenta) and part of the linker region (blue) connecting the HTH unit to the dimer domain (Figure 5D). Each half site of the palindromic operator is bound by one KorB subunit in its major groove (Figure 5D). Although the KorB HTH has a canonical HTH structure, Thr211 and Arg240, which are outside the recognition helix and located on the second and fourth helices in the linker domain, are utilized for base-specific contacts.

Structures of P1 ParB–centromere complexes

As noted, the E. coli P1 and F par loci were the first partition systems to be discovered and have since served as paradigms for studies on plasmid partition [4,10,12,16–17,24–29,85–88]. The P1 par system is of particular interest owing to its unusual centromere-like site. Studies carried out largely in the Funnell and Austin laboratories have provided a detailed description of the P1 centromere site and how it is recognized by the P1 ParB protein [24,25,85–88]. These studies showed that the P1 centromere site, parS, is an unusual ~74 bp DNA centromere element that can be divided into three main regions: a centralized ~29 bp binding site for IHF, a so-called ‘rightside’ region with two heptameric A-boxes, A2 and A3 (the third A-box, A4, to the right of B2 is not essential for partition) and one hexameric B-box, and a ‘leftside’ element with one A-box and one B-box element (Figures 2 and 6A). Although not required for partition, the E. coli IHF αβ heterodimer increases partition efficiency by bending the central region of parS to juxtapose the A-box- and B-box-containing
‘arms’, which appears to facilitate ParB binding across the bend (Figure 5C) [122].

Interestingly, data revealed that both the A-boxes and B-boxes are recognized and bound by P1 ParB [86]. This finding distinguishes ParB from other centromere-binding proteins, which only bind one cognate DNA element. Although full-length parS provides maximal partition efficiency, the rightside parS site composed of A2–A3–B2 is sufficient for partition in an IHF-independent manner. P1 ParB residues 142–333, which contains the HTH unit, linker domain and dimer domain, contain all the determinants required for centromere binding. Similarly to other type Ia centromere-binding proteins, motif B functions as a secondary oligomerization domain in P1 ParB when the protein is present at high local concentrations. The extreme N-terminal region of P1 ParB and nearly all other type Ia centromere-binding proteins are used to bind to their requisite motor proteins (Figure 5A). An exception is KorB, which binds its motor protein using residues within its centromere-binding domain [86,116].

Insight into how ParB can recognize two DNA elements and bind the unusual parS centromere site were provided by recent structures of ParB-(142–333)–rightside parS complexes (Figure 5C) [123]. These structures show that ParB-(142–333) comprises two separate domains. The first corresponds to residues 147–270 and includes the HTH unit and linker region. This domain consists of seven α-helices and is connected by a short flexible linker (residues 271–274) to the C-terminal domain (residues 275–333). As in other type Ia centromere-binding proteins, the C-terminal domain mediates dimerization. Unlike KorB, which contains a C-terminal SH3-like domain, the C-terminal domain of P1 ParB contains a novel fold consisting of three antiparallel β strands and a C-terminal α helix [118,123]. These elements lock together in the dimer to form a continuous antiparallel β-sheet, flanked by a coiled-coil (Figure 5C).

Unique structural features evident in the ParB–DNA structures explain how ParB can interact with multiple arrays of box elements on the looped parS site [123]. Importantly, each box element is bound by a separate DNA-binding module: the A-boxes are bound by the HTH motif, whereas the B-boxes are recognized by the dimer domain. In addition, the structure shows that ParB acts as a bridging factor between DNA boxes, explaining how it can bind sites on looped DNA or box elements on adjacent plasmids.

importantly, the two DNA-binding modules of ParB do not interact with each other and freely rotate, enabling them to contact multiple arrangements of A- and B-boxes, as found in parS (Figure 5C). Residues from the recognition helix of the HTH contact the major groove of the A-box elements. Interestingly, only one of these base contacts appears to be specific, suggesting the possibility that the HTH domain may play a role in non-specific DNA spreading. Indeed, non-specific binding of P1 ParB to DNA enables it to bind several kilobases upstream and downstream of the parS site [124–126]. Studies on P1 ParB and SopB, which both show spreading, suggest that DNA spreading may be a shared feature of at least some of the centromere-binding proteins [104,115,124–126]. Initial data on spreading-defective ParB mutants led to the speculation that it was important for partition. However, more recent studies in which plasmid stability was determined after spreading was restricted by the introduction of roadblocks on either side of parS revealed that extensive spreading was, in fact, not required for plasmid partition [124]. One result of spreading is the silencing of genes several or many kilobases away from the ParB–DNA interaction site [125]. The exact role(s) that spreading plays in DNA partition and transcription regulation has yet to be clarified.

Unlike the HTH domain–A-box interaction, the contacts made by the novel dimer domain to the B-box are highly sequence-specific. Initial structures of ParB-(142–333)–parS rightside showed only one side of the dimer domain bound to a B-box. However, this was due to the stoichiometry of box elements to ParB dimer used in the initial structural studies. Indeed, modelling suggested that each side of the dimer domain should be capable of B-box binding. The more recent structure of ParB-(142–333) bound to a 16 bp A-box/B-box-containing site confirmed this prediction and not only revealed a completely novel type of protein–DNA interaction, but also indicated that one P1 ParB is able to interact simultaneously with four separate DNA duplexes or DNA sites on bent DNA or a mixture of these permutations [127] (Figure 6). This multi-bridging capability has important implications for the ability of P1 ParB to form wrapped nucleoprotein structures as well as mediate plasmid pairing. Indeed, the unique DNA-bridging function of ParB reveals how it can bind across the rightside and leftside arms of the bent parS site. This unusual bridging capability would also permit P1 ParB to bind between the two arms of one looped parS site while simultaneously contacting a parS site on a second plasmid, thus providing a very attractive model for plasmid pairing. Moreover, the multi-bridging capability, especially that provided by the double B-box interaction, which juxtaposes DNA close in space, suggests multiple strategies for forming higher-order partition complex superstructures by P1 ParB.

Comparison of the Spo0J, KorB and ParB structures suggest that there are larger differences among the type Ia proteins as compared with the type Ib and II centromere-binding proteins. Although a mechanism for partition complex formation similar to that demonstrated by pSK-41 ParR can be envisioned for all centromere-binding RHH proteins, a clear structural model for the partition complex by the type Ia proteins is not yet forthcoming.

Structures of motor proteins: mechanisms of plasmid separation

Once the higher-order partition complex is formed, the final step in partition is formation of the final segrosome via recruitment of the motor protein, which is typically either an actin-like protein (type II) or, more commonly, a Walker-type protein (type I). Filament growth then propels the plasmids to opposite cell poles. Currently, structural information is available for one type II plasmid motor protein, ParM from the E. coli R1 plasmid [99,128–130]. Although a structure has yet to be reported for a plasmid type I Walker-type motor protein, the structure of the chromosomal Walker-type protein, T. thermophilus Soj, which is related to the type Ib motor proteins, has been described [131]. Structures of R1 ParM and T. thermophilus Soj combined with EM studies on several type II and I motor proteins have revealed important insight into how these proteins mediate separation of replicated plasmids.

Structures of the T. thermophilus Soj protein

The T. thermophilus Soj protein collaborates with the T. thermophilus Spo0J protein to play an important, yet incompletely characterized, role in chromosome segregation. Soj is also required together with Spo0J for stable maintenance of a plasmid bearing the parS centromere site, suggesting that they mediate plasmid segregation [69]. Localization studies of Soj have revealed dynamic oscillations of the protein similarly to oscillations observed for the type Ib Walker-type protein pB171 ParA [101,102,132]. Soj has also been shown to form nucleoprotein filaments, but in a DNA-dependent manner [131]. The structure of T. thermophilus Soj was solved in its apo and ADP-bound forms. In addition, the structure of the ATPase hydrolysedeficient mutant, Soj(D44A), was solved bound to ATP [131].
The structure confirmed previous predictions, that the protein contains a canonical Walker-type fold (Figure 7).

The Soj structure shows the highest homology with the cell division regulator MinD [133]. Accordingly, Soj consists of a core of twisted β-strands surrounded by α-helices. The α-helices cluster on either side of the β-sheet formed by one antiparallel β-strand and seven parallel β-strands. The Soj apo and ADP-bound forms are monomers, whereas ATP binding to Soj(D44A) leads to formation of an ATP ‘nucleotide sandwich dimer’. Notably, the formation of such a dimer had been predicted previously, but not observed until the structure of the Soj(D44A)–ATP complex [131]. Comparison of apo and ADP-bound Soj, reveals that ADP binding is accompanied by minor structural changes in the P-loop region wherein residues GGVG (Gly-Gly-Val-Gly) become less extended as they make multiple interactions with the ADP phosphate moieties. Upon ATP binding, dimerization is induced by numerous hydrogen bonds between residues on adjacent subunits (Figure 7). However, the key interaction in mediating the formation of dimer is provided by the Walker box signature lysine residue, Lys15, which contacts the α- and γ-phosphates of the ATP bound in the other subunit.

The formation of Soj polymers requires not only ATP, but also non-specific DNA as well [131,134]. A recent study aimed at elucidating what role Soj DNA binding might play in segregation identified several surface arginine residues, conserved among chromosomal Soj proteins, that were essential for DNA binding, but appeared to have no significant role in nucleotide binding [134]. Remarkably, creation of a DNA-binding-deficient Soj protein by mutation of these residues led to a partition-deficient phenotype in a model plasmid partitioning system in E. coli [134]. A recent study showed that, in addition to ATP, ParM can bind GTP. In this study, the structure of the R1 ParM apo form filaments, but also that these filaments function in vitro. These studies have firmly established that not only does ParM form filaments, but also that these filaments function in the segregation of the two domains [129]. Structure and function of plasmid partition proteins include DnaK, FtsA and MreB [135,136]. Crystal structures of ParM in its apo and ADP-bound form confirmed that it has an actin-like fold and, accordingly, consists of a two-domain architecture [129]. These two domains, called I and II, are delineated further into Ia and Ib and IIa and IIb (Figure 8). The larger subdomains, Ia and IIa, share a common fold consisting of a five-stranded β-sheet surrounded by three α-helices. In contrast, the two smaller subdomains are much more diverse structurally. Comparison of the apo ParM and ParM–ADP structures revealed that domains I and II act as rigid bodies and undergo a 25° rotation upon binding ADP [129] (Figures 8A and 8B). Interestingly, analysis of the ADP-binding pocket of ParM revealed no specificity-determining interactions between ParM residues and the adenine base. Indeed, a recent study showed that, in addition to ATP, ParM can bind GTP. In this study, the structure of the R1 ParM was solved bound to GDP and the R1 ParM–p[NH]ppG (guanosine 5′-β,γ-imido)triphosphate) complex structure was obtained by soaking p[NH]ppG into ParM–GDP crystals [99]. These ParM–ADP, ParM–GDP and ParM–p[NH]ppG structures are essentially identical in their conformations, which suggests that nucleotide binding stabilizes the closed conformation.

Although actin is a well-known filament-forming protein, several actin family members have roles unrelated to filament or polymer formation. For example, DnaK functions in protein folding, and FtsA is involved in the regulation of cell division [137–139]. Sequence homology between ParM and actin did not necessarily dictate filament formation as a mechanism for ParM action. To address the function of ParM, numerous cellular, biochemical and EM analyses have been carried out [42–45,128,129]. More recent studies by other groups have provided additional insights into the structure–function dynamics of ParM. These studies have firmly established that not only does ParM form filaments, but also that these filaments function directly in plasmid segregation [99,140–141]. Initial hints that ParM forms filaments were obtained from immunofluorescence microscopy studies, which revealed that ParM forms highly dynamic polymeric structures that extended along the E. coli cell. In addition, ParM was shown to self-assemble in vitro when ATP is present [129]. EM studies on these assemblies of R1 ParM filaments revealed the presence of ordered filaments that were initially described as actin-like [129]. More recently, two low-resolution models of ParM filaments have been constructed based on EM data [99,130]. Specifically, Orlova et al. [130] used the high-resolution crystal structures of R1 ParM to carry out image reconstructions of cryo-EM data of R1 ParM filaments.

Other bacterial proteins that are actin superfamily members include DnaK, FtsA and MreB [135,136]. Crystal structures of ParM in its apo and ADP-bound form confirmed that it has an actin-like fold and, accordingly, consists of a two-domain architecture [129]. These two domains, called I and II, are delineated further into Ia and Ib and IIa and IIb (Figure 8). The larger subdomains, Ia and IIa, share a common fold consisting of a five-stranded β-sheet surrounded by three α-helices. In contrast, the two smaller subdomains are much more diverse structurally. Comparison of the apo ParM and ParM–ADP structures revealed that domains I and II act as rigid bodies and undergo a 25° rotation upon binding ADP [129] (Figures 8A and 8B). Interestingly, analysis of the ADP-binding pocket of ParM revealed no specificity-determining interactions between ParM residues and the adenine base. Indeed, a recent study showed that, in addition to ATP, ParM can bind GTP. In this study, the structure of the R1 ParM was solved bound to GDP and the R1 ParM–p[NH]ppG (guanosine 5′-β,γ-imido)triphosphate) complex structure was obtained by soaking p[NH]ppG into ParM–GDP crystals [99]. These ParM–ADP, ParM–GDP and ParM–p[NH]ppG structures are essentially identical in their conformations, which suggests that nucleotide binding stabilizes the closed conformation.

Structures and EM studies on the R1 ParM protein

Before high-resolution structures were available, the type II ParM protein had been predicted to belong to the actin superfamily.
the ParM molecules within the filaments adopt a conformation that is even more open than the apo ParM conformation. In contrast, in the filament model proposed by Popp et al. [99], ParM adopts a closed conformation like that found in the crystal structures of R1 ParM with bound nucleotides. Despite these differences in detail, both ParM filament models are similar in overall shape and dimension and thus both studies support the insertional polymerization model whereby ParM filaments insert between plasmids to propel them towards opposite cell poles (Figure 9).

The insertional polymerization model has also recently received strong support from both in vitro reconstitution assays and in vivo time-lapse fluorescence microscopy analyses [46,140,141]. Moreover, these studies have provided additional insight into this process. The in vitro studies, which followed R1 plasmid dynamics in real time, revealed that ParM filaments are short-lived and are stabilized only when each end of the filament is bound to a ParR–parC-associated plasmid [141]. The rates of in vivo ParM polymerization were found to be similar to the timing required for plasmid segregation [46]. The fact that, in vitro, purified ParR and ParM proteins and parC-labelled magnetic beads alone could drive bead separation and the strong similarities between the in vitro and in vivo filament-mediated separation of plasmids and parC-labelled beads also strongly supports the idea that host factors are not required for type II plasmid segregation. The in vitro reconstitution study also demonstrated the important finding that the filament grows by insertion of additional ParM–ATP molecules at the ParR–parC interface. This insertional polymerization causes the ParM filaments to extend bidirectionally, moving the two plasmids to opposite cell poles. The in vivo studies suggest that, after reaching the poles, the force of colliding into the ends of the cell lead to dissociation of the filament ends from ParR–parC, prompting the rapid depolymerization of the filament [141] (Figure 9).

Indeed, it has been demonstrated that pressure applied to a filamentous structure such as actin can lead to loss of non-covalent interactions, bound nucleotides and susceptibility to depolymerization [142]. This knowledge provides additional insight into how the filaments depolymerize at the correct time after plasmid separation. Specifically, ParR is known to stimulate the ATPase activity of ParM. Because ParR is in contact with ParM, it will hasten the formation of ParM–ADP (or ParM–GDP) within the filament. However, as long as ParR is bound to the ends of the filaments, it strongly stabilizes the polymer form of ParM. But, after separating, the plasmids will be forced against the cell poles [141]. This force will dissociate the ends of one or both ParM filaments from its interaction with the partition complex on the plasmid. The end subunit of ParM, which had previously been in contact with ParR, will probably exist in the NDP-bound state. Release of the protective partition complex contacts will allow release of the NDP molecule. According to the recent ParM–GDP and ParM–p[NH]ppG structures, the NTP-bound and NDP-bound states of ParM are in the same conformation, which is distinct from the apo conformation [99]. Thus release of the NDP and subsequent conversion into the apo form will lead to domain rotation, which in turn will cause the loss of one subunit from the end, stimulating the disruption of the entire filament.

One problem with the insertional polymerization model is that it necessitates that each end of the ParM filament interacts with an identical ParR–DNA partition complex. However, structural information indicates that ParM filaments exhibit polarity, meaning they have a plus end and a minus end. To examine how the different ends of the ParM filament may interact equally with an identical partition complex, we docked the ends of the R1 ParM filament model obtained from EM on to the pSK41 partition

Subsequently, Popp et al. [99] used EM, TIRF (total internal reflection fluorescence) microscopy, high-pressure SAXS (small-angle X-ray scattering), and X-ray fibre diffraction to construct three-dimensional images of ParM–p[NH]ppG filaments and obtain insights into filament dynamics [99] (Figure 9). The resulting models from these two studies agree that ParM filaments are left-handed rather than right-handed like F-actin. However, the two low-resolution ParM filament models differ quite markedly in detail [99,130]. For example, in the model by Orlova et al. [130],

Figure 9 EM reconstructions of R1 ParM filaments and type II insertional polymerization model

complex superstructure [90,130]. Remarkably, this modelling revealed that the filament can readily fit within the partition complex pore (Figure 10). Further modelling shows that this is true even when accounting for the addition of the ParR C-terminal domains, which were not present in the structure [90]. As can be seen in Figure 10(A), ParM filament polarity would not be a problem because each end of the filament would encounter ParR molecules in multiple orientations (Figure 10). Furthermore, the C-terminal region of ParR, which interacts with ParM, is highly flexible, thereby permitting it to dock on to ParM molecules that are arranged in any orientation. An additional feature of the partition complex structure that is likely to be important for its interaction with ParM filaments is that its pore region displays a very high local concentration of ParR C-terminal tails which are perfectly poised for ParM filament interaction and capture [90]. ParR spreading could help further encase the ParR molecule and increase the likelihood of productive ParM binding.

Could this proposed partition complex capture mechanism apply to other par systems? As noted, although ParR/ParB proteins show little sequence identity, recent structures of the type Ia and type II centromere-binding proteins have revealed a striking homology in that both contain RHH DNA-binding domains. In fact, the RHH fold is an excellent partition complex-forming motif because binding to tandem centromere sites would promote its co-operative polymerization and the formation of higher-order superstructures. The combined structural data on type II ParR proteins and partition complexes, including EM data on the R1 partition complex, the crystal structure of pB171 ParR and the crystal structure of the pSK41 partition complex suggest that type II partition complexes form highly similar superhelical structures. It might be anticipated that type Ib centromere-binding proteins, which also contain RHH motifs, form related superstructures. We speculate that the specific conformation of each of the partition complex and segrosomal superstructures is likely to be dictated by the polymer structure of the motor protein bound by each partition complex.

CONCLUSIONS

DNA segregation is an essential process for the survival of all species. The segregation of plasmid DNA requires only three components and thus is an ideal model system to study the detailed mechanisms involved in this process at the atomic level. The structures of bacterial chromosomal and plasmid partition proteins and their DNA complexes that have been obtained over the last 5 years have shed significant light on the molecular mechanisms of segregation of several types of partition systems. In fact, studies on the type II system have revealed that three components, the centromere-binding protein, ATPase and centromere site, can orchestrate DNA separation without the addition of cellular factors. It has also been established unequivocally that ParM forms filaments in an ATP- or GTP-dependent manner, and formation of these filaments drives DNA partition. Stabilization of these filaments requires interaction of filament ends with the partition complex superstructure, which is formed by the binding of ParR molecules to the centromere. The first high-resolution structure of a higher-order partition complex, that of the type II pSK41 partition complex, revealed that the ParR–centromere complex is a superhelical structure with dimensions suitable for capture of the ParM filaments [90]. The position of the ParM-binding sites within the flexible C-terminal region of ParR indicates how it can bind the dissimilar ends of the ParM filament.

Despite the substantial progress in delineating a molecular understanding of type II plasmid partition, many questions remain. For example, how do ParR and other centromere-binding proteins mediate the seemingly opposing functions of ATPase stimulation and polymer stabilization? Structural studies on centromere-binding proteins bound to their requisite motor proteins should help clarify this paradox. Although many of the steps of type II partition are understood at the molecular level, much less is known about type I segregation. Unexpected structural homology revealed between the type II and type Ib centromere-binding proteins suggests that the type Ib centromere-binding proteins may form similar superstructures when interacting with their centromere sites; however, this remains to be demonstrated. Also, the types of partition complexes formed by the type Ia proteins are, at this time, completely unknown. The presence of conserved HTH and motif B domains hints at a mechanism involving DNA spreading as mediated by the combination of these motifs. Perhaps the most unresolved issue regarding type I partition is the molecular mechanism(s) utilized by type I motor proteins in DNA separation. In fact, no structures are available for a plasmid type I motor protein and thus how they may form polymers and
bind nucleotides remains a mystery. Mounting evidence suggests that polymer formation by type I motor proteins is important in segregation. If this is the case, do these proteins utilize an insertional polymerization mechanism similar to that of type II par systems? Support for a common mode of polymer-driven plasmid separation comes from recent findings that putative type III and IV par systems employ a tubulin-like GTPase and a coiled-coil protein respectively for partition. Thus the future challenges are to fill in the missing pieces of the puzzle left in the type II partition reaction and to determine the molecular mechanism(s) used by the type Ia and Ib systems, as well as newly identified par systems, to form partition complexes and drive plasmid separation. Structural studies combined with biochemical and cellular analyses will be needed to address these issues and, hopefully, in so doing, provide atomic-level snapshots of each step in the partition reaction.

I acknowledge support from the Burroughs Welcome (Burroughs Wellcome Career Development Award 992863), the University of Texas M.D. Anderson Cancer Center Trust Fellowship and the National Institutes of Health. I thank Professor Finnar Hayes for helpful discussions and comments. I apologize to those whose work was not discussed.

REFERENCES


© The Authors Journal compilation © 2008 Biochemical Society