

Three ParA dimers cooperatively assemble on type Ia partition promoters

François Boudsocq^{1*}, Maya Salhi¹, Sophie Barbe² and Jean-Yves Bouet^{1*}

¹ Laboratoire de Microbiologie et Génétique Moléculaires, Centre de Biologie Intégrative (CBI), Centre National de la Recherche Scientifique (CNRS), Université de Toulouse, UPS, F-31062 Toulouse, France.

² Toulouse Biotechnology Institute (TBI), Université de Toulouse, CNRS, INRAE, INSA, F-31077 Toulouse, France.

*Corresponding authors francois.boudsocq@univ-tlse3.fr and jean-yves.bouet@univ-tlse3.fr

ABSTRACT

Accurate DNA segregation is essential for faithful inheritance of genetic material. In bacteria, this process is mainly ensured by a partition system (Par) composed of two proteins, ParA and ParB, and a centromere site. The auto-regulation of Par operon expression is important for efficient partitioning, and is primarily mediated by ParA for type Ia plasmid partition systems. For the plasmid F, four ParA_F monomers were proposed to bind to four repeated sequences in the promoter region. By contrast, using quantitative surface plasmon resonance, we showed that three ParA_F dimers bind to this region. We uncovered that one perfect inverted repeat (IR) motif, consisting of two hexamer sequences spaced by 28-bp, constitutes the primary ParA_F DNA binding site. A similar but degenerated motif overlaps the former. ParA_F binding to these motifs is well supported by biochemical and modeling analyses. In addition, molecular dynamics simulations predict that the winged-HTH domain displays high flexibility, which may favor the cooperative ParA binding to the promoter region. We propose that three ParA_F dimers bind cooperatively to overlapping motifs thus covering the promoter region. A similar organization is found on both closely related and distant plasmid partition systems, suggesting that such promoter organization for auto-regulated Par operons is widespread and may have evolved from a common ancestor.

INTRODUCTION

Genome stability depends on faithful segregation of replicated DNA to daughter cells. Stable inheritance of low-copy-number bacterial replicons (chromosomes and plasmids) requires an active partition mechanism. Each such replicon typically carries a self-specific partition module comprising a two-gene operon and an adjacent set of short sequence repeats that acts as a centromere. The first gene encodes an ATPase and the second a centromere-binding protein. Partition systems have been classified into three main classes depending on the type of ATPase encoded with Type I, II and III, which are characterized by deviant Walker A, actin-like and tubulin-like NTPases, respectively (for review (1) (2)). Type I systems, generically termed ParABS, are the most common type in the bacterial world, being widespread on low-copy-number plasmids and the only type present

on chromosomes. ParA ATPases act on large nucleoprotein complexes formed by ParB on *parS* sites to drive the duplicated centromere DNA to opposite sides of the cell, ensuring faithful DNA segregation.

The mechanism of DNA segregation from type I partition systems is not fully understood (2). Nevertheless, the DNA binding properties of ParA have been shown to play an essential role in segregation. In its ATP-bound form ParA binds non-specifically to DNA (3) (4) (5) and this activity was shown to be essential for DNA partition *in vivo* (4) (6). Some ParA also display a specific DNA binding activity (7), provided by an additional amino-terminal (Nter) domain. This distinction defines two subclasses, namely type Ia and type Ib (8). Only type Ia ParAs bind DNA specifically to their promoter regions and function as transcriptional regulators of their own operons (9) (10) (11). This autoregulation is enhanced by ParB proteins acting as

corepressors of transcription (9) (11) and is further enhanced by the centromere DNA (12) (13). The additional Nter domain is composed of ~100 aa and contains a winged-helix-turn-helix (wHTH) motif (14). The wHTH is involved in promoter binding activity (15) (16) when ParA is either in ADP-bound or ligand-free forms (5) (17). The role of the autoregulation of plasmid-encoded Par operons has been highlighted for a long time as both the absolute and relative levels of ParA and ParB are important for the partition process (18) (19) (20).

The partition system of *Escherichia coli* plasmid F, ParABS_F (historically called SopABC), belongs to the type Ia. It encodes ParA_F and ParB_F proteins in an operon upstream of the *parS_F* centromere site (Figure 1). ParA_F binds to the promoter region of *parAB_F* (*PparAB_F*) and ParB_F enhances this binding (7). DNase I footprint assays suggested that ParA_F binds to four copies of the hexamer motif 5'CTTTGC present in *PparAB_F*. Three hexamer motifs are arrayed in direct orientation with 13-bp intervals, and one is present in an inverted orientation 28-bp further downstream (Figure 1). ParA_F binds to *PparAB_F* cooperatively and induces DNA bending (21). It was proposed that ParA_F monomers bind to *PparAB_F* with a 4:1 stoichiometry, one on each hexamer motif (21). However, how several ParA monomers bind in a coordinate fashion to these binding motifs is not known. Furthermore, results of our own investigations were inconsistent with this proposal and, in consequence, we have re-examined the mechanism underlying the autoregulation of type Ia plasmid partition operon.

In the present work, we further characterized ParA_F site-specific DNA binding properties on its promoter region. Using quantitative biochemical studies, we revealed a stoichiometry of three ParA_F dimers assembled on *PparAB_F* DNA. By modeling and docking a ParA_F dimer on the promoter region, we proposed that two hexamer motifs form a perfect inverted repeat (IR) binding motif separated by a large spacer. We also uncovered another similar but degenerated motif that overlaps the perfect IR. Molecular dynamics simulations suggested that the wHTH is a highly flexible region supporting the binding of ParA_F dimer to the overlapping motifs. Based on the *PparAB_F* DNA sequence and on closely related

Par systems, we proposed that three ParA_F dimers bind in an overlapping manner to the promoter region. Finally, similar promoter organization was also found in other type Ia partition systems suggesting that the proposed assembly for auto-regulating partition promoter may be conserved and widespread on low-copy-number plasmids.

MATERIALS AND METHODS

Protein purification and analytical gel filtration

Native ParA_F and ParB_F proteins were purified from derivatives of strain DLT812 carrying pDAG127 (*ParaBAD::parA_F*) (20) and pDAG170 (*ParaBAD::parB_F*) (5). Protein production was induced by adding arabinose (final concentration 0.1%) to 1 L cultures grown at 37°C to OD₆₀₀ ~ 0.6 in LB supplemented with thymine (10 µg.ml⁻¹) and kanamycin (50 µg.ml⁻¹). Four hours after induction, cells were harvested by centrifugation, resuspended in cold buffer A (50 mM NaCl, 1 mM EDTA, 20 % sucrose) equilibrated at pH 6 (50 mM bis-Tris) for ParA_F or pH 8 (50 mM Tris) for ParB_F. Cells were frozen at -80°C and slowly thawed on ice. All subsequent steps were performed at 4°C. Cell free extracts were obtained by sonication followed by centrifugation for 20 min. at 27,000 g and filtration of the supernatant through a 0.2 µm hydrophilic filter (PALL). The filtrates were loaded onto 10 ml Heparin column on a FPLC (GE Healthcare) equilibrated in buffer A without sucrose. After washing with at least 100 ml of buffer A, the proteins were eluted over a 50-500 mM NaCl gradient. Fractions containing ParA_F or ParB_F, identified by SDS-PAGE, were pooled. Ammonium sulphate was added to concentrations of 0.32 mg.ml⁻¹ and 0.15 mg.ml⁻¹ for ParA_F and ParB_F, respectively, and the solutions were incubated for 30 min prior to centrifugation for 20 min at 20,000 g. The supernatants were loaded on a Superdex 200 column (GE Healthcare) equilibrated in buffer C (200 mM NaCl, 0.5 mM EDTA) equilibrated at pH 6 for ParA_F (40 mM bis-Tris) or pH 8 for ParB_F (40 mM Tris). Fractions containing ParA_F or ParB_F were pooled, diluted two-fold in buffer C without NaCl, loaded onto a mono-Q or a mono-S column (GE Healthcare), equilibrated in buffer D (buffer C containing 100 mM NaCl), for ParA_F or ParB_F, respectively. After elution over a 100-600 mM

NaCl gradient, each fraction containing ParA_F or ParB_F was mixed with cold glycerol (50 % final) and DTT (5 mM). All fractions were flash frozen in liquid nitrogen and stored at -80°C.

Protein concentrations were determined from absorbance at 280 nm using a Nanodrop spectrophotometer. The ParA_F and ParB_F preparations were pure to > 97% homogeneity as judged by SDS-PAGE stained by Instant Blue (Euromedex). An example of ParA_F purification was shown in Supplementary Figure S1A.

For analytical gel filtration, ParA_F diluted in buffer E (50 mM bis-Tris pH 6, 400 mM NaCl, 0.5 mM DTT) to 4 µM, in the presence or absence of ADP (1 mM), was injected onto a calibrated Superdex 75_10/300GL column (GE healthcare) and eluted at 0.5 ml.min⁻¹ in the same buffer.

Surface Plasmon Resonance assays

SPR experiments were performed using a Biacore 3000 (Biacore) or a Biacore X100 (GE Healthcare) apparatus. 200-530 RU (indicated in the figure legends) of biotinylated DNA probes (Supplementary Figure S2) were immobilized on a streptavidin-coated sensorchip in HBS-EP buffer (10 mM Hepes pH 7.4, 150 mM NaCl, 3 mM EDTA, 0.005 % P20). Binding analyses were performed at 24°C with multiple injections at different protein concentrations. Samples, diluted in BD buffer (20 mM Hepes-KOH pH 7.4, 100 mM KCl, 10 mM MgCl₂, 1 mM DTT), were injected at 20 µl.min⁻¹ for 180 sec. Reference sensorgrams containing non-specific DNA (Biacore 3000) or no DNA (Biacore X100) were subtracted from sensorgrams containing the tested probes to yield binding responses. No significant difference was observed between the two methods of subtraction. Kinetic constants were calculated using BIAevaluation 4.0.1 software (Biacore) taking into account survey's recommendations (22). Values of K_D , k_a and k_d were derived from local analysis fit with the 1:1 Langmuir binding model. R_{max} is the theoretical maximum response of binding to the ligand at a specific surface density of analytes. The kinetics values were variable depending on experiments and apparatus preventing their accurate determinations as indicated in the result section. Nevertheless, SPR assays allowed to accurately calculating the precise measurement of binding stoichiometry.

Electrophoretic mobility shift assays

The standard binding mixture (10 µl) contained 15 nM Cyanine 3 (Cy3) labelled DNA probes in 30 mM HEPES-KOH (pH 7.5), 150 mM KCl, 3 mM MgCl₂, and 1 mM DTT, 10 % glycerol and, where indicated, 1 mM ADP. The mixtures, assembled on ice, at the indicated protein concentration, were incubated for 15 min at 37°C, and analyzed by electrophoresis at 50 V for 4 h or 200 V for 1 h at 4°C on 6% polyacrylamide gels in TGE (25 mM Tris, 25 mM Glycine and 5 mM EDTA). Wet gels were scanned at 532 nm using a Typhoon trio imager (GE Healthcare). Percentages of bound and unbound fractions were quantified using TINA software. Each experiment was performed at least three times. Data were subjected to non-linear regression analysis using GraphPad Prism 4.0 Software.

Microscale thermophoresis assays

Microscale thermophoresis measurements were performed at 25°C using Monolith NT™ premium capillaries and the Monolith NT.115 device with a green filter (NanoTemper, Germany). Power of LED and IR-laser sources are indicated in figure legends. Binding reactions, assembled in low binding grade tubes (Eppendorf), consisted of MST buffer (10 mM Tris-HCl pH 8, 150 mM NaCl, 50 mM CaCl₂, 0.05% Tween 20, 20 mg.ml⁻¹ BSA), 25 nM Cy3-labelled, annealed oligonucleotide probes (Supplementary Figure S2) and increasing concentrations of ParA_F. Proteins and DNA were centrifuged prior use. Fluorescence response was measured once total fluorescence reached equilibrium in all capillaries (SDS denaturation was undertaken following Nanotemper recommendation). Thermophoresis data were analyzed when fluorescence signals lies within the 10% acceptance range using MO Affinity Analysis software V2.1.2 (Nanotemper). Data were plotted using GraphPad Prism 4.0 software and fitted with nonlinear equation.

Protein melting point analysis

Protein thermal stability was measured in a label-free fluorescence assay using a Prometheus NT.48 (NanoTemper, Germany). Briefly, the shift of intrinsic tryptophan fluorescence upon temperature-induced unfolding was monitored by detecting the emission fluorescence at 330 and 350 nm.

Thermal unfolding was performed in nanoDSF grade high-sensitivity glass capillaries (NanoTemper, Germany) at a heating rate of 1°C per minute. Melting points (T_m) were calculated from the first derivative of the ratio of tryptophan emission intensities at 330 and 350 nm. A screen was performed over 18 different buffers. Melting temperature varies from 34.4°C up to 49.5°C (average 44.5°C).

Modeling of ParA_F structure and docking on DNA

A three-dimensional model (3D) of ParA_F UniProt P62556 (SOPA_ECOLI) was built by homology modeling using Swiss-Model and ExPASy with the X-ray structure of ParA_{P1} as template (PDB 3ez2) (14). The resulting 3D model includes the first five aa and the last seven aa of ParA_F (Figure 4B). We also modelled ParA_F using I-TASSER (23) and Phyre2 (24). Automatic docking of DNA on the 3D model of ParA_F was carried out using HADDOCK 2.2 (25). Curved DNA models were obtained using 3D DART (26) by applying different angles between each nucleotide. The best docking was observed with a 3° angle given rise to an overall curvature of 120° over a 40-bp DNA fragment. The automated prediction of hinges in ParA_F structure was performed by HingeProt analysis (27).

Molecular dynamics simulation

From the 3D model of ParA_F, molecular dynamics simulations were carried out using the AMBER 14 suite of programs (28) and the molecular all-atom ff14SB force field (29). Parameters for ADP were derived from (30). To obtain a neutral charge of the simulated system, a number of Na⁺ counter-ions were included. Proteins together with the counter-ions were solvated with TIP3P water molecules, using the rectangular parallelepiped box with a minimum distance of 0.10 nm between the solute and the simulation box edge. Preparation of simulations consisted of initial energy minimization steps (steepest descent and conjugate gradient methods). Positions of the protein backbone and ADP atoms were first restrained using a harmonic potential during the minimization schedule. The force constant was then progressively diminished until a final unrestrained minimization step. The minimization process was followed by a slow heating under constant volume over a period of

100 ps. At the final temperature (310 K), the system was equilibrated under constant volume condition over 10 ps and then subjected to constant pressure (1 bar) condition over 90 ps. The final production phase of simulations was then carried out for a total of 40 ns at constant temperature (310 K) and pressure (1 bar). Temperature and pressure parameters were setup using a Berendsen thermostat and barostat (31) with a collision frequency of 2 ps⁻¹ and pressure relaxation time of 2 ps. Long-range electrostatic forces were handled by using the particle-mesh Ewald method (32). The time step of the simulations was 2.0 fs and the SHAKE algorithm (33) was used to constrain the lengths of all chemical bonds involving hydrogen atoms to their equilibrium values. To avoid artefacts, MD simulations were run twice, with different starting velocity distribution. The resulting trajectories were analyzed using the *cptraj* module of the AMBER14 package. The RMSD was calculated for the protein backbone atoms using least squares fitting. Positional fluctuations (Δr_i^2) of protein backbone atoms were calculated. A mass-weighted average value was then calculated for each residue. These parameters are related to the B-factors (depicting the atomic fluctuation) through the following relationship: $B_i = \frac{8\pi^2}{3} \langle \Delta r_i^2 \rangle$. The simulated B-factors were calculated using the coordinates of the 40 ns trajectory. All graphics were prepared using PyMOL (34).

RESULTS

Three ParA_F dimers bind to the *parA* promoter region

To investigate how ParA_F interacts with its promoter, we performed a combination of three techniques measuring protein-DNA interactions. We first characterized this interaction using Surface Plasmon Resonance (SPR) measurements with increasing concentrations of ParA_F injected onto immobilized 136-bp PparAB_F DNA substrate. We obtained typical dose response binding curves up to signal saturation above 500 nM (Figure 2A). Similar results were obtained in the presence of ADP (Supplementary Figure S1B), showing that the ADP-bound and apo forms of ParA_F display the same DNA binding activity, as previously observed (5). No SPR signal was detected if

PparAB_F DNA was replaced by non-specific DNA probes of same length indicating that *ParA_F* binding to the promoter region is highly specific in the absence (Figure 2B) or presence of ADP (Supplementary Figure S1C). SPR allowed direct and accurate calculation of the stoichiometry of *ParA_F* per DNA substrate as (i) the signal response is directly proportional to the molecular weight of the complexes at the surface of the sensorchip where a given and known amount of DNA is immobilized, and (ii) SPR only measures the active protein present in the protein preparations. We estimated the stoichiometry of *ParA_F* per DNA substrate by measuring the ratio of the SPR signal response relative to the amount of *PparAB_F* DNA immobilized (Figure 2C, left panel). We found that six *ParA_F* monomers bind per *PparAB_F* DNA fragment when saturation is achieved. The same stoichiometry was also observed in the presence of ADP (Supplementary Figure S1B).

In the case of plasmid P1, apo-*ParA_{P1}* was shown to be mostly a dimer in solution (35) (14). To investigate the oligomeric state of *ParA_F* we performed analytic exclusion chromatography. At nearly physiological *ParA_F* concentration (4 μ M), we observed only one peak, in the presence or absence of ADP, which corresponds to a molecular weight slightly above 80 kDa (Figure 2D). Thus, this indicates that *ParA_F* is mostly a dimer in the apo and ADP-bound forms. In the above SPR experiments, the stoichiometry is therefore three *ParA_F* dimers per *PparAB_F* promoter (Figure 2C). This result differs from a previous report using electrophoretic mobility-shift assay (EMSA) which concluded that four *ParA_F* monomers bind to the four 5'-CTTTGC sequences of *PparAB_F* (21). To discriminate between these contradictory results, we performed SPR analyses in other conditions such as in the presence of competitor DNA or non-hydrolysable ATP (ATP γ S, which conferred a similar conformation of *ParA_F* as in the presence of ADP (6)), and we observed the same stoichiometry of three dimers per DNA fragments (Figure 2C, right panel). Importantly, this 3:1 stoichiometry was also observed with a 240-bp DNA fragment (Figure 2C, right panel) indicating that the size of the DNA probe does not limit the amount of *ParA_F* bound to the promoter region.

These experiments, reproduced on different Biacore apparatus with the same stoichiometry

outcomes (Supplementary Figure S1D), displayed a large variability in the SPR binding isotherms (Supplementary Table S1), thus preventing to determine accurately the *ParA_F*-*PparAB_F* dissociation constant (K_D). By contrast, microscale thermophoresis (MST) assays and direct fluorescent measurements yielded a K_D of 140 and 130 nM, respectively, for *ParA_F* binding to a 120-bp Cy3-labelled *PparAB_F* DNA probe (Figure 3A, upper panel; Supplementary Figure S3A). MST measurements performed in the presence of BSA, PEG or ADP resulted in similar K_D ranging from 60 to 200 nM.

Lastly, we performed EMSA with the 120-bp Cy3-labelled *PparAB_F* fragment (Figure 3B). A unique band shift was obtained from 30 nM of *ParA_F*. No band shift was observed when the DNA fragment does not harbor the motif #4 whatever the combination tested (Figure 3C). However, a band shift was observed when at least motifs #3 and #4 are present. The binding curve from the 120-bp *PparAB_F* fragment, fitted with a nonlinear regression, resulted in an apparent K_D of 85 ± 20 nM (Figure 3D) with a sharp slope suggesting high cooperativity as previously observed (21). With longer incubation times, some *ParA_F* appeared stacked in the wells (judged by silver staining) at concentrations above 0.5 μ M indicating that some level of aggregation occurs at high concentrations, as also observed in the other techniques used (see figure legends). This effect is less pronounced in the presence of ADP (discussed below). Nevertheless, combining the results obtained from these different techniques, we estimated the K_D of the interaction of *PparAB_F* promoter with three *ParA_F* dimers to be in the order of 100 nM and that motifs #3 and #4 are the primary *ParA_F* binding sites.

Modeling *ParA_F* structure from *ParA_{P1}*

Understanding how three *ParA_F* dimers interact with the promoter region would benefit from the knowledge of the 3D structure. Lacking an experimentally determined structure of *ParA_F*, we built a 3D structural model by comparative modeling using the 3D coordinates of the ADP-*ParA_{P1}* structure (14). The two proteins share strong functional features and all deviant Walker-A *ParA* ATPases display important structural similarities (36) (14) (37). We first characterized the extent of their secondary structure similitudes. Despite *ParA_F*

and ParA_{P1} sharing low sequence identity and similarity, 22 % and 54 %, respectively, the predicted ParA_F secondary structure is similar to the one of ParA_{P1} obtained by X-ray crystallography (Figure 4A). ParA_F contains two main parts discerned by hydrophobic cluster analysis (HCA) and IUPred analyses (Supplementary Figure S4B-C). The first part (aa 1-110) has the lowest structuration of the protein and encompasses the dimerization and the wHTH domains. The first Nter α -helix (aa 6-27) may be important for ParA_F dimerization as shown for ParA_{P1} (14). This is further supported by our failure to purify an active variant that lacks this part (Vergne, Castaing & Bouet, unpublished). In addition, HingeProt analysis predicted that two hinges separates the wHTH domain from the first α -helix from (position 30) and the second part (position 105) (Figure 4A and Supplementary Figure S4B). The second part (aa 111-388) contains the main conserved residues (Supplementary Figure S4D) and comprises the deviant Walker-box and the ns-DNA binding domains (Figure 4A). Notably, the last α -helix (aa 367-388) is not conserved amongst ParA members (Supplementary Figure S4D).

These numerous similitudes between ParA_F and ParA_{P1} allowed us to build a ParA_F 3D model (Figure 4B) by homology modeling using three different approaches (SwissModel, Phyre2 and I-tasser). The RMSD value that compares the template and the best model backbone was ~1 Å indicating a good modeling accuracy. Besides this structural analysis, we attempted to construct several variants in the Nter domain and the wHTH motif specifically impaired for PparAB_F binding. However, all deletion variants tested were poorly expressed suggesting that these proteins were not stable. Also, two variants in the wHTH (A46P and R57A-K61A) were purified but melting point analysis using nanoDSF (advanced Differential Scanning Fluorimetry) suggests that they were significantly unstructured, thus preventing further studies.

Overall, these results indicated that ParA_F and ParA_{P1} display structural similarities and that the first 110 amino acids of ParA_F including the wHTH are important for correct folding.

Overlapping binding of three ParA_F dimers in the promoter region

Next, by careful examination of the PparAB_F sequence, we found that the 5'CTTTGC motifs #3 and #4 (Figure 1) may form an inverted repeat (IR) motif with a large 28-bp spacer that would accommodate the binding of a ParA_F dimer. We investigated this hypothesis by docking a ParA_F dimer onto this putative IR motif (Figure 5A). The distance between the centers of the two binding motifs is 34-bp, therefore they both locate nearly on the same face of the DNA molecule. Each motif could interact with each wHTH, separated by ~10 nm, of the ParA_F dimer. However, such binding could only be possible if the DNA is curved by 3° between each base pair as deduced from docking trials (Figure 5A). Interestingly, the PparAB_F region is AT rich (73 % over 100-bp), presents A-stretches, a property which is known to be prone to induce DNA curvature (38), and is bent upon ParA_F binding (21). Moreover, the modeling of ParA_{P1} dimer bound to a 40-bp DNA fragment has also suggested that DNA bending is required to allow the two wHTH to contact the DNA major groove (14). In the following, we will refer to this motif as IR3-4 (Figure 5C).

We next extended our analyses to the binding motif #2. We found the similar but slightly degenerated sequence, 5'tcTTGC, 27-bp away in an inverted orientation (Figure 5C), that we named #2' and that could form a putative IR2-2'. For binding motif #1, we could not find a complementary inverted repeat ~28-bp away since this position exactly overlaps the -35 promoter box. Previous DNase I footprint experiments performed on PparAB_F DNA (7) (21) have shown that ParA_F protection extends over 87-bp (schematically represented in Figure 5C), starting and ending in the close vicinity of binding motifs #1 and #4, respectively. Importantly, a protection was observed over the imperfect binding motif #2'. Combining these results with the stoichiometry experiments, we propose that ParA_F dimers could bind to IR3-4, IR2-2' and motif #1, the latter being bound only through one wHTH of a dimer, leading to three dimers bound to PparAB_F as observed by SPR (Figure 2C). This binding organization implies that ParA_F dimers are overlapping each other to contact interspersed binding motifs present two by two in inverse orientations on the same face of the DNA molecule.

We then examined promoter regions of three partition systems closely related to the one

of plasmid F, the *Klebsiella pneumoniae* virulence plasmid pLVPK (39), the *E. coli* linear plasmid-prophage N15 (40) and the *Klebsiella oxytoca* plasmid-prophage phiKO2 (41). In all three cases, while the sequences in between the nearly identical binding motifs vary, they display the same spacing between each other with (Figure 5C). Notably, for N15, the binding motif #2', present in inverted orientation with a 28-bp spacing from the motif #2, has only one variation from the consensus (5'GtAAAG). This strengthens the hypothesis that these two hexamer sequences form an IR to which a ParA_F dimer could bind. We performed Blastn searches with the Par promoter regions of F, LVPK, N15, and KO2, and found 12 other closely related, but not identical, partition promoters (Supplementary Figure S5). We carried out a quantitative Logo analysis using these 16 sequences (Figure 5D). The conserved pattern still emerged clearly both in term of sequence and spacing between each binding motif. This finding further argues for a binding model involving overlapping IR motifs in these partition promoters.

Is the promoter organization with overlapping IR motifs a conserved feature amongst type Ia plasmid partition operons?

To evaluate whether a promoter organization with overlapping IR binding motifs is a general or restricted feature, we extended the comparative analysis to more phylogenetically distant partition systems. The promoter regions of the Par operons of plasmids P1 (PparAB_{P1}) and P7 (PparAB_{P7}) display a different organization from the one of plasmid F. They both harbor a large degenerated palindromic sequence but with different positioning relative to their respective -35 and -10 boxes (schematically represented by black arrows in Figure 5E) and also, in PparAB_{P7}, a set of four 9-bp imperfect repeats (open arrows) (42) (10). In PparAB_{P1}, we found a 6-bp motif, 5'TTATGC, at both extremities of the large palindromic sequence, in an inverse orientation with a 28-bp spacing (IR3-3'; Figure 5E), which could accommodate the binding of a ParA_{P1} dimer as proposed above for ParA_F. Strikingly, we found two other sets of similar sequences in inverse orientation with 26- and 28-bp spacing. As for F, only one IR motif (IR3-3') is perfectly conserved, the two others (IR1-1' and IR2-2') being imperfect. The binding motif #2,

5'acAaGC, is the less conserved but overlaps the -35 promoter box. The logo analysis of the six proposed binding motifs in PparAB_{P1} indicates that four out of six nucleotides are highly conserved as for the motifs in PparAB_F (Figure 5F). Overall, the layout of these three putative IR motifs is highly reminiscent to the one of PparAB_F, with only IR3-3' being shifted downstream 8-bp respectively, in correlation with the -35 and -10 boxes being shifted upstream compare to PparAB_F. We found that the position of these proposed IR motifs matches the protection zones of previous DNase I footprinting experiments (35) (43) (44) which begins at the motif #1 and ends a few bp after motif #3' (Figure 5E, grey lines). Moreover, DNase I footprinting performed with a version of PparAB_{P1} lacking the DNA sequence upstream of the -35 promoter box still revealed a protection at the position corresponding to the missing binding motif #1, 5'gcATGC (35). Thus, only one binding motif of the IR (in this case #1') is required to allow ParA_{P1} dimers to efficiently protect the upstream DNA sequence (#1) by a non-specific DNA binding interaction with the second wHTH.

In the P7 Par promoter, PparAB_{P7}, a perfect IR motif composed of the sequences 5'ACGTGC separated by 29-bp also encompassed the imperfect palindromic sequences previously proposed (10). Two other sets of similar motifs matching a consensus 5'ACGTgc present two by two in inverted orientation with 26- and 27-bp spacers were present in PparAB_{P7} (Figure 5E). However, in contrast to PparAB of F and P1, only the proposed motifs IR1-1' and IR2-2' overlapped; the putative IR3-3' being shifted 21-bp downstream. This latter is also less conserved which leads to a slightly lower signature for a binding motif in the logo analysis (Figure 5F). The overlapping motifs IR1-1' and IR2-2' correspond perfectly to the ParA_{P7} protected zones in DNase I footprinting (Figure 5E, grey bars) further arguing for being the binding motifs. In the case of IR3-3', a DNA protection is seen close to the motif #3 but not at the proposed motif #3', which raises the question of whether this latter IR motif is involved in autoregulation of PparAB_{P7}.

These analyses support the possibility that, as for PparAB_F, the partition promoters of P1 and P7 harbor IR motifs with 26- to 29-bp internal spacing that may be involved in the

binding of their cognate ParA dimers. Interestingly, with the only exception of IR3-3' from P7, these motifs are similarly positioned relatively to each other in these three different promoter regions (Figure 5C-E). This finding suggests that some type Ia promoters of phylogenetically distant plasmid-encoded partition operons may share this overlapping binding organization.

The winged-HTH domain is highly flexible

The analysis of ParA_F and ParA_{PI} amino acid sequences and structures has revealed a conservation of secondary and tertiary structures. An important feature of the type Ia plasmid ParA proteins is their additional Nter domain comprising the wHTH. Our observation suggests that the PparAB_F region harbors inverted repeats where three ParA_F dimers could bind if the DNA is curved (Figure 5A). Another possibility would be that the protein is flexible and undergoes conformational changes enlarging its amplitude. We investigated how the wHTH region participates in the specific DNA binding by analyzing the conformational plasticity of ParA_F protein. To this end, we performed Molecular Dynamics (MD) simulations in explicit water environments. The stability of ParA_F dimers during the course of the MD simulation was evaluated by backbone atoms RMSD measurement after least square fit from the starting structure (supplementary Figure S6A). The profile obtained indicates that the backbone RMSD gradually increased by 7 Å within the first 7 ns of the simulation and then remained stable until the end of the simulation. Such RMSD variation indicates thus that the protein structure had undergone significant conformational changes (supplementary Figure S6B). The analysis of the B-factors during the course of MD simulations highlights particularly flexible regions of the protein (Figure 4C). They are predominantly located in the wHTH region involving α 2- α 3 helices and β 2- β 3 antiparallel β strands (supplementary Figure S4A). In particular, the glutamic acid and the arginine residues at positions 60 and 75, respectively, exhibit the highest simulated B-factor values and correspond to the end of the HTH (E60) and to the inflexion point of the winged (R75). It is noteworthy that ParA_{PI} also presents high simulated B-factors within the wHTH domain (centered around residue 75; Supplementary Figures S6C-D).

The backbone superposition of the starting ParA_F structure with the one obtained after 40 ns of MD simulation, shows that the arginine residue R75 can move ~30 Å away from its initial position (Figure 5B), while the core domain remains stable (Supplementary Figure S6B). Notably, the conformational flexibility enlarges the wing span and increases the width of ParA_F dimer from 10 nm to 11 nm. This increase of protein amplitude may be needed for contacting the inverted repeats motifs. This result is also in agreement with the prediction of two Hinges at positions 30 and 105 (Figure 4A). Overall, the MD simulations show that the wHTH region is highly flexible and may undergo important variation in amplitude movement increasing the wing span of ParA_F. We propose that, in addition to DNA bending, conformational changes may adjust ParA_F dimers binding to the IR motifs to support local topological constraints.

The IR3-4 motif is sufficient to nucleate the binding of three ParA_F dimers.

The minimal size of the DNA fragment required for three ParA_F dimers to bind to PparAB_F was investigated by SPR assays. We designed DNA duplexes containing at least the IR3-4 motif (Supplementary Figure S2). No response signal was obtained in SPR assays using a 58-bp DNA fragment indicating that a minimal size is required for ParA_F binding. By contrast, a SPR response was observed with an 85-bp probe. Despite the concentration of ParA_F needed to reach saturation was significantly higher (5 μ M) than for the larger DNA fragments, a stoichiometry of 2.7 was calculated (Figure 2C, right panel). This indicates that the IR3-4 motif is sufficient to nucleate a complex with three ParA_F dimers on a DNA fragment as small as 85-bp.

Strikingly, we found that the position of the IR motif on small DNA fragments is crucial for ParA_F binding. Indeed, using EMSA, we observed that if the IR3-4 motif is positioned in the middle of the 120-bp DNA fragments (probe IR2-2*), no band shift was observed (Figure 3C). The same result was obtained if the motif #4 is also present (probe IR2-2* #4). This result indicates that there is not enough space for a complex of three ParA_F dimers to bind starting from an IR motif positioned in the middle of the DNA fragment. Thus, these data confirm that the IR3-4 motif is sufficient to nucleate the

binding of three ParA_F dimers to *PparAB_F* DNA providing that the length of the DNA available on left is large enough allowing stable interaction.

ADP stabilizes ParA_F but does not increase its affinity to *PparAB_F*

ParA_{P1} was shown to interact preferentially with its promoter region in the presence of ADP (43) (17). For ParA_F, no significant difference in binding affinity to *PparAB_F* was detected in the presence or absence of ADP in SPR and EMSA experiments (compare Figure 2A-B and Supplementary Figure S1B-C; Supplementary Table S1). However, in MST experiments, we observed that ADP stabilized the thermophoresis signals and limits ParA_F aggregation. To investigate this further, we performed protein melting point analyses using nanoDSF, in different conditions (Table 1). In the presence of ADP, the ParA_F melting temperature (*T_m*) increased by 1.5°C. By contrast, in the presence of *PparAB_F* DNA, ParA_F is stabilized only moderately with an increased *T_m* by 0.5°C and this effect is not cumulative with the ADP one (Table 1). Thus, ADP stabilizes ParA_F dimers compare to ligand-free form. This may prevent the aggregation that occurs over time but ADP does not increase ParA_F affinity to *PparAB_F*.

ParB_F increases ParA_F binding affinity to the promoter region

Early experiments using DNase I footprinting assays have suggested that ParB_F increases the binding affinity of ParA_F to *PparAB_F* (7). To quantify this stimulatory effect, we performed EMSA in the presence of both ParA_F and ParB_F. ParB_F by itself has no specific affinity for *PparAB_F* DNA and presents a *K_D* of ~1 μM to non-specific DNA probes (Supplementary Figure S3D). Maximal ParB_F stimulation of ParA_F binding to *PparAB_F* was obtained when the same stoichiometry of both proteins was used (Figure 3E). It resulted in an apparent *K_D* of about 10 nM compare to 85 nM with ParA_F alone, corresponding to an eight-fold ParB_F dependent stimulation. Lastly, we also tried the addition of a duplex DNA containing *parS_F* in the presence of ParB_F and found no further stimulation of ParA_F binding.

DISCUSSION

This study provides a new picture of how the partition operon, *parAB_F*, is transcriptionally auto-regulated. It establishes that ParA_F is a dimer in solution and binds primarily to an IR motif. This first interaction allows two other dimers to bind, in a cooperative and oriented manner, to other overlapping motifs surrounding the -35 and -10 transcriptional boxes. The resulting complex composed of three dimers covers an 85-bp region. We propose that the cooperativity is mediated by protein-protein interaction through the highly flexible winged-HTH domain, and that this binding organization may be conserved for some plasmid partition promoters of the type Ia subfamily whose finely-tuned expression is auto-regulated by ParA. Our findings are summarized in Figure 6.

ParA proteins of type I partition systems are dimers in the presence of ATP (37) (35). In contrast to chromosomally-encoded ParAs (belonging to type Ib; (8)) which are converted to monomers upon ATP hydrolysis (37), type Ia ParAs remain mainly dimers, as exemplified by ParA_{P1} and ParA_{P7} (35) (14) and shown for ParA_F (Figure 2D). The first alpha-helix of the Nter region of ParA_{P1-P7} stabilizes the dimer form (14). This alpha-helix was conserved on ParA_F (Supplementary Figure S4) and 3D-modeling strongly suggests that it may exhibit the same role in dimer stabilization as for ParA_{P1-P7} (Figure 4B). The Nter domain is not present on type Ib ParAs (8), therefore explaining this difference in dimer to monomer conversion upon ATP hydrolysis between plasmid and chromosomally encoded ParAs. Also, the expression of partition operons is auto-regulated by ParAs only for type Ia plasmid members, through the wHTH DNA binding in the Nter domain (12) (10) (16).

ParA_F specifically binds the promoter region in the apo and ADP-bound forms (5) as for ParA_{P1} (35) (17). Four ParA_F monomers were initially proposed to bind to the four hexamer repeats in the promoter region of *parAB_F* (21). However, gel filtration experiments indicated that ParA_F is a dimer in solution both in the presence or absence of ADP (Figure 2D), which rather argue, along with the quantitative SPR analyses, that three ParA_F dimers bind to the promoter region (Figure 2C). We thus hypothesized that ParA_F dimers may contact binding sites that would be present in inverted orientations and uncovered that the hexamer motifs #3 and #4 defines a perfect

inverted repeat motif with a 28-bp spacer. We also identified a slightly degenerated motif (#2') in reverse orientation and spaced by 27-bp relatively to the motif #2. These proposed binding sites are all protected by ParA_F in DNase I footprinting assays (summarized in Figure 5C), including the motif #2' (21). In addition, mutations in motif #2 abolished the protection of the motif #2' (21). Therefore, these previous results together with our data strongly argue for ParA_F binding as a dimer contacting two DNA sites separated by about three helical turn (~33-bp between the centers of each repeat). We propose that binding motifs #3 and #4 constitute a perfect IR motif (IR3-4). Accordingly, motifs #2 and #2' could form an imperfect IR motif (IR2-2'). These IR motifs are highly conserved in term of sequences and spacing amongst closely related promoters (Figure 5C and Supplementary Figure S5). For the motif #1, the position expected for a cognate inverted binding motif with a ~28-bp spacing would be exactly at the position of the -35 box, which prevents a match with the consensus binding motif. Interestingly, the removal of the binding site #1 changed the footprint at the expected position for a cognate binding site (21) further arguing for ParA_F binding as a dimer even from a unique binding motif providing other IR motifs are present (see below).

Previous studies have investigated the promoter region of the Par operons of the plasmids P1 and P7 and reported that they are very different in sequence and organization (10). Nevertheless, the 3D-structure of ParA from these two plasmids is highly similar and allowed a 3D-modeling of ParA_F with high accuracy (Figure 4B), thus suggesting that ParA binding as dimers to IR motifs in the promoter regions would also apply for P1 and P7. In both cases, as for F, we found two perfect 6-bp inverted motifs spaced by 28- or 29-bp, IR3-3' and IR1-1' for P1 and P7, respectively (Figure 5E). Also, two other putative IR motifs with imperfect binding sites spaced by ~26- to 28-bp were present. These three promoter regions have different core sequences but display overlapping IR binding motifs with striking similar positioning relatively to each other (Figure 5E). Therefore, we propose that the binding of a ParA dimer to the perfect IR motif serves as nucleating the cooperative binding of two other ParA dimers to the degenerated motifs, thus stabilizing the binding of three dimers to

PparAB (see model Figure 6). This proposition also supports the unexplained observation that a truncated version of PparAB_{P1}, lacking the DNA sequence upstream of the -35 promoter box, is still protected from DNase I cleavage at the position corresponding to the missing binding motif #1 in a ParA_{P1} dependent manner (35). For PparAB_F, providing that at least one IR repeat motif is present, a ParA dimer could nucleate the binding of two other ParA efficiently to cover the promoter region (Figure 2C and 3C). Altogether, these data suggest that the promoter structure of these type Ia plasmid partition operons is conserved from a putative common ancestry and also highlights the importance of the auto-regulation of the *par* genes expression.

ParA_F covers a large DNA region (~85-bp) encompassing the -35 and -10 core promoter sequences as observed in DNase I footprinting assays (7) (21). This is in agreement with our findings that the minimal DNA fragment size for the binding of three ParA_F dimers is ~85-bp (Figure 2C, right panel). The binding affinity of ParA_F on the promoter region was measured to be in the order of ~100 nM (Figure 3). The binding curve displays a sharp slope (Figure 3D), suggesting an important binding cooperativity. Also, mutations in any of the four conserved binding motifs does not change the overall protected zone (21) and the IR3-4 motif is sufficient to nucleate ParA_F binding to PparAB_F (Figure 3C) further arguing for the high binding cooperativity. Our data suggest that dimer-dimer interaction may favor the cooperative binding of three ParA_F dimers to PparAB_F. Interestingly, wHTH motifs has been reported to be involved in protein-protein interaction (45), as in the case of the *B. subtilis* protein RacA (46), whose wHTH motif is the closest structure to that of ParA_F. Moreover, a model prediction from the fitting of the ParA_{P1} crystal dimer (ADP form) into the nucleofilament of its close homolog ParA2_{Vcho} from the chromosome 2 of *Vibrio cholera* also suggested that the wHTH makes contact with the N-terminal of other subunits (47). In light of our results and the specific organization of the promoter region with overlapping binding motifs, it is therefore tempting to propose that a dimer-dimer interaction mediated by the wHTH motifs may allow ParA_F dimers to bind cooperatively on PparAB_F promoter (see model Figure 6). The wHTH is composed of a classical

tri-helical bundle followed by two β -strand hairpins and one α -helix forming an approximately triangular outline. The wing interacts with the minor groove of DNA through charged residues in the hairpin (45). In the ParA_F dimer, the two wHTHs are present on both edges of the structure (Figure 4B). This symmetrical conformation allows binding to IR motifs. The two binding motifs of the proposed IR are separated by 33- or 34-bp (Figure 5C), corresponding to about three helical turns of 10.5-bp, and therefore are present nearly on the same face of the DNA molecule. The 3D-modeling of a ParA_F dimer on a DNA fragment indicates that the distance between the two wHTHs motifs is shorter than the distance between the inverted binding motifs. Therefore, the specific interaction could occur only if the DNA is bent (Figure 5A) as proposed for ParA_{P1} (14), and observed on PparAB_F DNA upon ParA_F binding (21). The promoter sequence contains numerous A-tracks known to promote important local bending of the DNA, thus suggesting that the DNA curvature is probably not homogeneous over the entire region. Our results based on molecular dynamics simulations suggest that the wHTH, insulated by two hinges motifs (Figure 4A), is the most flexible region of ParA_F (supplementary Figure S6B). It could move up to 3 nm away from its initial positions (Figure 5B), expanding the overall span of the dimer by 10 angstroms. Therefore, this capability to change the amplitude between the two wHTH motifs and to adapt their positioning in relation to the local DNA constraint might be important for contacting the IR binding motifs and for interacting with a wHTH motif of another dimer (see model Figure 6).

ParB_F participates in the autoregulation of *parAB* expression (11) as also observed in the P1 Par system (9) and stimulates ParA_F binding to its promoter region (7). We quantified this effect *in vitro* and observed a ~8-fold increase in binding affinity with a K_D of ~10 nM confirming that ParB_F stimulates ParA_F binding to the promoter region in the absence of ATP hydrolysis (Figure 3D). The action of ParB_F on ParA_F binding does not occur through a stable ParA-ParB interaction since no super-shift of PparAB_F-ParA_F complexes in EMSA or larger footprinting were observed upon addition of ParB_F (7). In the case of plasmid P1, a specific enhancement of repression by ParB_{P1} was also

shown to be due to a protein-protein interaction with ParA_{P1} rather than through a direct recognition of the operator by ParB_{P1} (15), in agreement with the observation that ParB_{P1} does not bind to the promoter region and does not cause any qualitative modification of the footprint of ParA_{P1} bound to it (42). In the presence of the centromere site, *parS_F*, the level of repression measured *in vivo* increased (12). This also occurs when *parS_F* is present in *trans* on a different DNA molecule, thus indicating that the assembly of ParB_F on *parS_F* and surrounding sequences including PparAB_F does not account for this additional repression level. One possible explanation for the stimulatory effect of ParB_F on ParA_F binding to its promoter region could be that a conformational change in ParA_F is induced by ParB_F. In the case of plasmid P1, ParA_{P1} super-repressor variants are no more stimulated by ParB_{P1} to further increase repression (44). It was suggested that these mutations, in the Walker A motif, might block ParA_{P1} in the conformation proficient for promoter binding (48). We propose that a conformational change upon ParB_F interaction may extend the overall span of ParA_F and/or the relative orientation of the flexible wHTH motifs favoring the promoter-specific DNA binding affinity. A similar ParB-induced conformational change acting on ParA-ATP form, releasing ParA binding to non-specific DNA, has been previously proposed (49) (48). Further experiments are needed to determine if this ParB-induced conformational change is transient or requires the release of ATP before being converted back to the non-specific DNA binding conformation.

In conclusion, our results shed some light on key structural features of ParA_F dimers leading us to propose a new organization with overlapping binding motifs in the promoter region. This organization is conserved on closely but also distantly related partition systems suggesting that it may be important for the fine-tuning of the partition operon expression. The nature of the interactions between ParA_F dimers and their orientation relative to each other await future biochemical and structural studies.

SUPPLEMENTARY DATA

Supplementary data includes six figures, one tables and references.

ACKNOWLEDGEMENTS

We thank members of the ‘Active genome segregation: mechanism and diversity’ group for discussion, J-P. Castaing, M. Vergne and S. Ait Bara for constructing *parA_F* mutants, P. Rousseau, D. Lane and C. Johnston for helpful comments on the manuscript, and F. Pasta for helpful contribution. We also thank the French Nanotemper team for their support, the Biacore facility at the ITAV and the CRCT Proteomics platforms. This work was granted access to the HPC resources on the Computing mesocenter of Région Midi-Pyrénées (CALMIP, Toulouse, France).

FUNDING

This work was supported by Agence National pour la Recherche (10-BLAN-1316-01).

REFERENCES

1. Baxter,J.C. and Funnell,B.E. (2014) Plasmid Partition Mechanisms. *Microbiol. Spectr.*, **2**.
2. Bouet,J.-Y. and Funnell,B.E. (2019) Plasmid Localization and Partition in Enterobacteriaceae. *EcoSal Plus*, **8**.
3. Leonard,T.A., Butler,P.J.G. and Löwe,J. (2004) Structural analysis of the chromosome segregation protein Spo0J from *Thermus thermophilus*. *Mol. Microbiol.*, **53**, 419–432.
4. Hester,C.M. and Lutkenhaus,J. (2007) Soj (ParA) DNA binding is mediated by conserved arginines and is essential for plasmid segregation. *Proc. Natl. Acad. Sci. U. S. A.*, **104**, 20326–31.
5. 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.
6. Castaing,J.-P., Bouet,J.-Y. and Lane,D. (2008) F plasmid partition depends on interaction of SopA with non-specific DNA. *Mol. Microbiol.*, **70**, 1000–1011.
7. Mori,H., Mori,Y., Ichinose,C., Niki,H., Ogura,T., Kato,A. and Hiraga,S. (1989) Purification and characterization of SopA and

SopB proteins essential for F plasmid partitioning. *J. Biol. Chem.*, **264**, 15535–15541.

8. Gerdes,K., Moller-Jensen,J. and Jensen,R.B. (2000) Plasmid and chromosome partitioning: surprises from phylogeny. *Mol. Microbiol.*, **37**, 455–466.
9. Friedman,S.A. and Austin,S.J. (1988) The P1 plasmid-partition system synthesizes two essential proteins from an autoregulated operon. *Plasmid*, **19**, 103–112.
10. 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.
11. Biek,D.P. and Strings,J. (1995) Partition Functions of Mini-F Affect Plasmid DNA Topology in *Escherichia coli*. *J. Mol. Biol.*, **246**, 388–400.
12. Yates,P., Lane,D. and Biek,D.P. (1999) The F plasmid centromere, sopC, is required for full repression of the sopAB operon1. *J. Mol. Biol.*, **290**, 627–638.
13. Hao,J.-J. and Yarmolinsky,M. (2002) Effects of the P1 Plasmid Centromere on Expression of P1 Partition Genes. *J. Bacteriol.*, **184**, 4857–4867.
14. 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.
15. 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.
16. 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.
17. 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.
18. 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.

19. Funnell, B.E. (1988) Mini-P1 plasmid partitioning: excess ParB protein destabilizes plasmids containing the centromere parS. *J. Bacteriol.*, **170**, 954–960.
20. Lemonnier, M., Bouet, J.-Y., Libante, V. and Lane, D. (2000) Disruption of the F plasmid partition complex in vivo by partition protein SopA. *Mol. Microbiol.*, **38**, 493–503.
21. Komai, M., Umino, M. and Hanai, R. (2011) Mode of DNA binding by SopA protein of Escherichia coli F plasmid. *J. Biochem. (Tokyo)*, **149**, 455–461.
22. Rich, R.L. and Myszka, D.G. (2008) Survey of the year 2007 commercial optical biosensor literature. *J Mol Recognit*, **21**, 355–400.
23. Zhang, Y. (2008) I-TASSER server for protein 3D structure prediction. *BMC Bioinformatics*, **9**, 40.
24. Kelley, L.A., Mezulis, S., Yates, C.M., Wass, M.N. and Sternberg, M.J.E. (2015) The Phyre2 web portal for protein modeling, prediction and analysis. *Nat. Protoc.*, **10**, 845–858.
25. van Zundert, G.C.P., Rodrigues, J.P.G.L.M., Trellet, M., Schmitz, C., Kastiris, P.L., Karaca, E., Melquiond, A.S.J., van Dijk, M., de Vries, S.J. and Bonvin, A.M.J.J. (2016) The HADDOCK2.2 Web Server: User-Friendly Integrative Modeling of Biomolecular Complexes. *J. Mol. Biol.*, **428**, 720–725.
26. Lu, X.-J. and Olson, W.K. (2003) 3DNA: a software package for the analysis, rebuilding and visualization of three-dimensional nucleic acid structures. *Nucleic Acids Res.*, **31**, 5108–5121.
27. Emekli, U., Schneidman-Duhovny, D., Wolfson, H.J., Nussinov, R. and Haliloglu, T. (2008) HingeProt: Automated prediction of hinges in protein structures. *Proteins Struct. Funct. Bioinforma.*, **70**, 1219–1227.
28. D.A. Case, V. Babin, J.T. Berryman, R.M. Betz, Q. Cai, D.S. Cerutti, T.E. Cheatham, III, .A. Darden, R.E. Duke, H. Gohlke, A.W. Goetz, S. Gusarov, N. Homeyer, P. Janowski, J. Kaus, I. Kolossváry, A. Kovalenko, T.S. Lee, S. LeGrand, T. Luchko, R. Luo, B. Madej, K.M. Merz, F. Paesani, D.R. Roe, A. Roitberg, C. Sagui, R. Salomon-Ferrer, G. Seabra, C.L. Simmerling, W. Smith, J. Swails, R.C. Walker, J. Wang, R.M. Wolf, X. Wu and P.A. Kollman (2014) AMBER.
29. Maier, J.A., Martinez, C., Kasavajhala, K., Wickstrom, L., Hauser, K.E. and Simmerling, C. (2015) ff14SB: Improving the accuracy of protein side chain and backbone parameters from ff99SB. *J. Chem. Theory Comput.*, **11**, 3696.
30. Meagher, K.L., Redman, L.T. and Carlson, H.A. (2003) Development of polyphosphate parameters for use with the AMBER force field. *J. Comput. Chem.*, **24**, 1016–1025.
31. Berendsen HJ, Postma JP, Van Gunsteren WF, Di Nola A, and Haak JR (1984) Molecular dynamic with coupling to an external bath.
32. Essmann U, Perera L, Berkowitz ML, Darden T, Lee H, and Pedersen LG (1995) A smooth particle mesh Ewald method. *J Chem Phys*.
33. Ryckaert JP, Ciccotti G, and Berendsen HJC (1977) Numerical integration of the Cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J Comput Phys*, **23**, 327–333.
34. DeLano, W.L., Ultsch, M.H., De, A.M., Vos and Wells, J.A. (2000) Convergent Solutions to Binding at a Protein-Protein Interface. *Science*, **287**, 1279–1283.
35. 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.
36. 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. *Nucl Acids Res*, **36**, 3676–3689.
37. 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.
38. Wu, H.M. and Crothers, D.M. (1984) The locus of sequence-directed and protein-induced DNA bending. *Nature*, **308**, 509–513.
39. Chen, Y.-T., Chang, H.-Y., Lai, Y.-C., Pan, C.-C., Tsai, S.-F. and Peng, H.-L. (2004) Sequencing and analysis of the large virulence plasmid pLVPK of Klebsiella pneumoniae CG43. *Gene*, **337**, 189–198.

40. Ravin, N. and Lane, D. (1999) Partition of the linear plasmid N15: interactions of N15 partition functions with the *sop* locus of the F plasmid. *J. Bacteriol.*, **181**, 6898–6906.
41. Casjens, S.R., Gilcrease, E.B., Huang, W.M., Bunny, K.L., Pedulla, M.L., Ford, M.E., Houtz, J.M., Hatfull, G.F. and Hendrix, R.W. (2004) The pKO2 Linear Plasmid Prophage of *Klebsiella oxytoca*. *J. Bacteriol.*, **186**, 1818.
42. 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.
43. 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.
44. 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.
45. Aravind, L., Anantharaman, V., Balaji, S., Babu, M.M. and Iyer, L.M. (2005) The many faces of the helix-turn-helix domain: Transcription regulation and beyond. *FEMS Microbiol. Rev.*, **29**, 231–262.
46. Schumacher, M.A., Lee, J. and Zeng, W. (2016) Molecular insights into DNA binding and anchoring by the *Bacillus subtilis* sporulation kinetochore-like RacA protein. *Nucleic Acids Res.*, **44**, 5438–5449.
47. Hui, M.P., Galkin, V.E., Yu, X., Stasiak, A.Z., Stasiak, A., Waldor, M.K. and Egelman, E.H. (2010) ParA2, a *Vibrio cholerae* chromosome partitioning protein, forms left-handed helical filaments on DNA. *Proc. Natl. Acad. Sci.*, **107**, 4590–4595.
48. Baxter, J.C., Waples, W.G. and Funnell, B.E. (2020) Nonspecific DNA binding by P1 ParA determines the distribution of plasmid partition and repressor activities. *J. Biol. Chem.*, **295**, 17298–17309.
49. Vecchiarelli, A.G., Havey, J.C., Ing, L.L., Wong, E.O.Y., Waples, W.G. and Funnell, B.E. (2013) Dissection of the ATPase Active Site of P1 ParA Reveals Multiple Active Forms Essential for Plasmid Partition. *J. Biol. Chem.*, **288**, 17823–17831.
50. Schneider, T.D. and Stephens, R.M. (1990) Sequence logos: a new way to display consensus sequences. *Nucleic Acids Res.*, **18**, 6097–6100.

	T _m (°C)
ParA _F	44.5 +/- 0.4
ParA _F + PparAB _F	45.1 +/- 0.2
ParA _F + ADP	46.1 +/- 0.1
ParA _F + PparAB _F + ADP	45.8 +/- 0.1

Table 1. Measurement of ParA_F melting temperature.

The melting temperature (T_m) of ParA_F was determined using a label-free fluorescence assay in presence or absence of ADP and/or 120-bp PparAB_F DNA. The standard deviation (+/-) was calculated from triplicate measurements. The data shown are from a typical experiment. Each assay has been reproduced three times with the same variations observed between each tested condition.

Figure 1

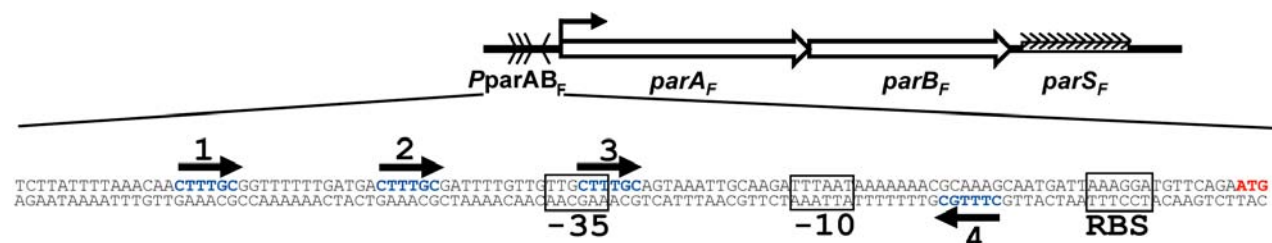


Figure 1. The *parABS_F* (*sopABC*) partition locus of the plasmid F.

Top: Schematic representation of the *parABS_F* locus. The *parA_F* and *parB_F* genes, the promoter region (*PparAB_F*) and the *cis*-acting *parS_F* site are drawn at scale over 3-kbp. The arrow represents the transcriptional start point. Arrowheads in *PparAB_F* display the positions and orientations of the hexamer 5'CTTTGC motifs. The arrowheads in the *parS_F* centromere site represent the twelve 43-bp tandem repeats containing the *ParB_F* binding sites. Bottom: Nucleotide sequence of the *PparAB_F* promoter region. The *parA_F* start codon and the hexamer motifs are labeled in red and blue, respectively. The -10 and -35 promoter motifs and the ribosome binding site (RBS) are indicated by boxes.

Figure 2

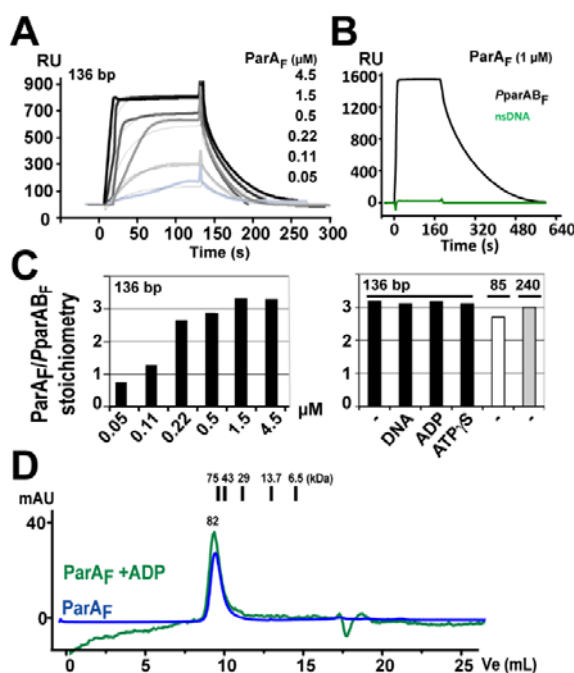


Figure 2. Three *ParA_F* dimers bind to the *PparAB_F* promoter.

A. Surface Plasmon Resonance analysis of *ParA_F* binding to *PparAB_F*. 300 RU of a 136-bp biotinylated DNA fragment were immobilized on a streptavidin chip. Increasing concentrations of *ParA_F* ranging from 0.05 to 4.5 μM in BD buffer without adenine nucleotide were injected at 20 $\mu\text{l} \cdot \text{min}^{-1}$ at time 0. Dose response sensorgrams and corresponding fits according to 1:1 Langmuir interaction model are represented by bold and light curves, respectively.

B. *ParA_F* binds specifically to the promoter region in the absence of ADP. *ParA_F* DNA binding activity was measured by SPR using 136-bp DNA fragments containing *PparAB_F* (530 RU; black line) or nsDNA (597 RU; green line) immobilized on two different channels of the same sensorchip. *ParA_F* (1 μM) was injected at 10 $\mu\text{l} \cdot \text{min}^{-1}$ at time 0 for 180 s before dissociation.

C. Stoichiometry of *ParA_F* dimers binding. *ParA_F* dimers bound per *PparAB_F* DNA molecule was calculated according to the formula $S = R_{\text{max}} /$

$[(\text{MW}_A / \text{MW}_L) * R_L]$ where R_{max} is the maximum response, MW_A is the molecular weight of *ParA_F*, and MW_L and R_L are the molecular weight and the amount of ligand immobilized, respectively. ADP and non-hydrolysable ATP analog (ATP γ S) were used at 1 mM and, where indicated, sonicated salmon sperm DNA is added at 1 $\text{mg} \cdot \text{ml}^{-1}$. Biotinylated DNA fragments of 85-, 136- and 240-bp, immobilized at 200, 300 and 530 RU, respectively, were infused with 5, 1.5 and 1.85 μM of *ParA_F*, respectively.

D. *ParA_F* is a dimer in solution in the apo- and ADP-form. Analytical gel filtration chromatography of *ParA_F* (4 μM) pre-incubated with (green line) or without (blue line) 1 mM ADP using a Superdex 75 column was monitored at 280 nm. A drift of the baseline is observed throughout the elution in the presence of ADP probably to a slight difference in the buffers used during the experiment. The elution peak (9.5 ml) corresponds to 82 kDa. Conalbumin (75 kDa), ovalbumin (43 kDa), anhydrate carbonic (29 kDa), ribonuclease A (13.7 kDa) and aprotinin (6.5 kDa) were used as molecular weight standards and eluted at 9.66, 10.55, 11.74, 13.35 and 15 ml, respectively.

Figure 3

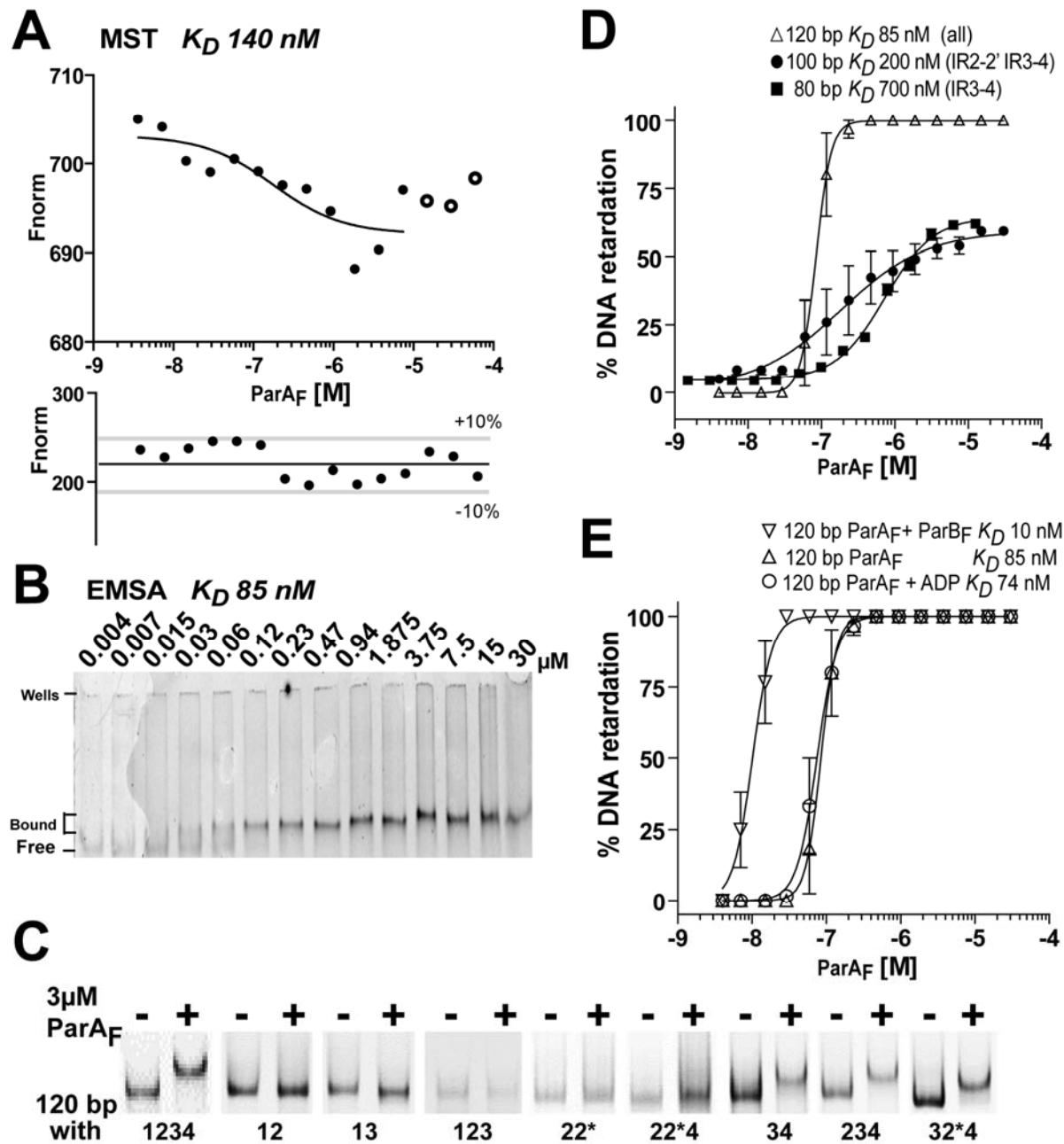


Figure 3. Measurements of ParAF binding affinity to the *PparABF* promoter.

A. MST measurement of ParAF binding to a 120-bp Cy3-labelled *PparABF* DNA fragment. DNA (25 nM) is titrated over ParAF concentrations ranging from 3.6 nM to 60 μ M. The apparatus power sources were set at 40% LED and 80 % IR-laser. At high protein concentrations (above 15 μ M) the decrease in the fluorescence signal is due to ParAF aggregation (larger dots) and is not taken into account for curve fitting. K_D was determined to be 140 ± 23 nM with a stable fluorescence signal in the 10 % acceptance range (lower panel).

B. ParAF binding to *PparABF* in EMSA. Cy3-labelled 120-bp DNA fragments (15 nM) carrying *PparABF* were incubated with increasing concentrations of ParAF ranging from 4 nM to 30 μ M. They were

analyzed by electrophoresis in a 6% TGE polyacrylamide gel. The positions of the bound and free DNA fragments, and of the wells, are indicated on the left.

C. ParA_F binding to PparAB_F DNA fragments harboring mutations in the hexamer motifs. EMSA were performed as in B with 3 μ M of ParA_F and various 120-bp probes carrying the indicated binding motifs (see Supplementary Figure S2). Note that 2* refers to the 2' motif replaced by a perfect consensus site.

D. ParA_F binding to PparAB_F in the presence or absence of ADP and ParB_F. Quantification of EMSA experiments performed as in (B) with a 120-bp Cy3-labelled PparAB_F probe (15 nM) titrated over 4 nM to 30 μ M ParA_F, and incubated when indicated with ADP (1 mM) or ParB_F (equal molar concentration as ParA_F). The apparent K_D were calculated after fitting with non-linear regression from three independent assays. Note that, in the presence of ParB_F, the apparent K_D is an estimation since it is close to the DNA probe concentration (15 nM).

E. ParA_F binding to PparAB_F in the presence of ADP or ParB_F. As in C, with Cy3-labelled 120-bp probes incubated with ParA_F concentrations ranging from 4 nM to 30 μ M in the presence of ADP (1 mM) or ParB_F (equal molar concentration as ParA_F). The apparent K_D measured in the presence of ParB_F is an estimation since it is close to DNA probe concentration (15 nM).

Figure 4

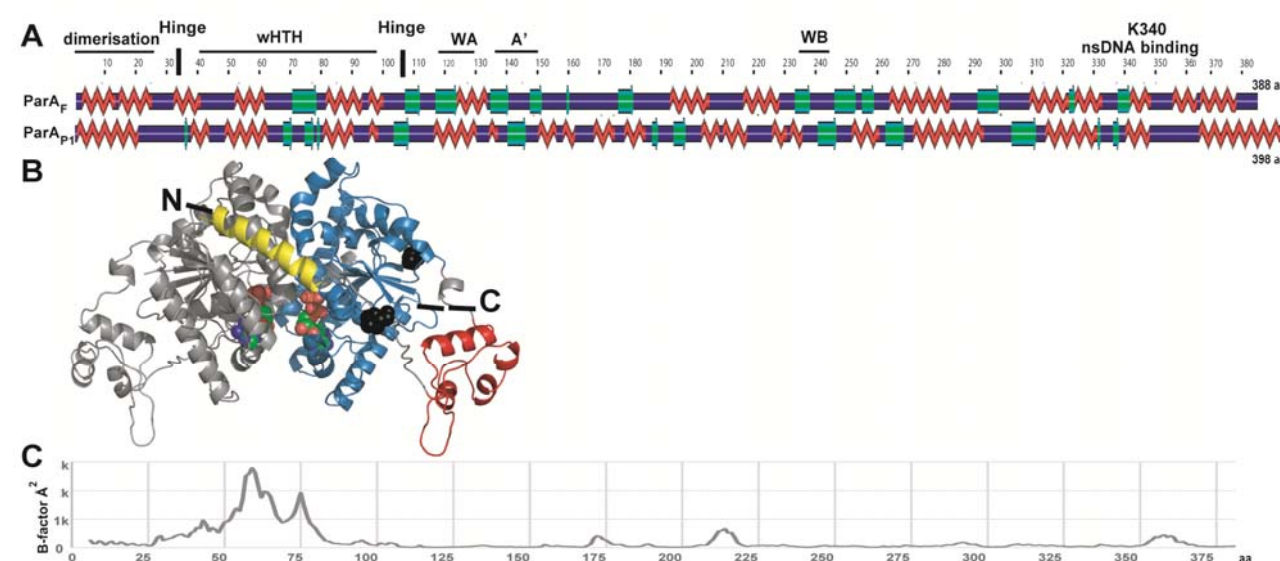


Figure 4. Structural similarity of ParA_F with ParA_{P1}.

A. Alignment of secondary structures predicted for ParA_F (388 aa) and derived from X-ray crystallography for ParA_{P1} (398 aa). The comparison was generated by POLYVIEW2D. α -helices and β -sheets are displayed in red and green, respectively. Domains and motifs for ParA_F are indicated as follows: dimerization (1-27 aa), wHTH (winged helix turn helix; 40-100 aa), WA (Walker A; 114-122 aa), WA' (Walker A'; 139-149 aa), WB (Walker B; 246-251 aa), nsDNA binding K340 (lysine 340). Vertical bars represent Hinge positions predicted by HingeProt analysis.

B. Model of ParA_F 3D structure generated by Swiss-Model using ADP-ParA_{P1} structure as a reference (see Materials and Methods). The two monomers are colored separately in grey and blue. For the right monomer, the first alpha-helix and the wHTH are displayed in yellow and red, respectively, and the Nter (N) and Cter (C) are indicated. Hinge positions are represented by black spheres. Each monomer contains an ADP represented with spheres, red for O, green for C, blue for N and orange for P.

C. The wHTH of ParA_F undergoes a large conformational change. The B-factor, estimating the atomic fluctuation and expressed in angstrom (A²), were calculated from residues 5 to 381 from Molecular Dynamic simulations. The local mobility within the molecule is mostly restricted to the region encompassing the wHTH motif.

Figure 5

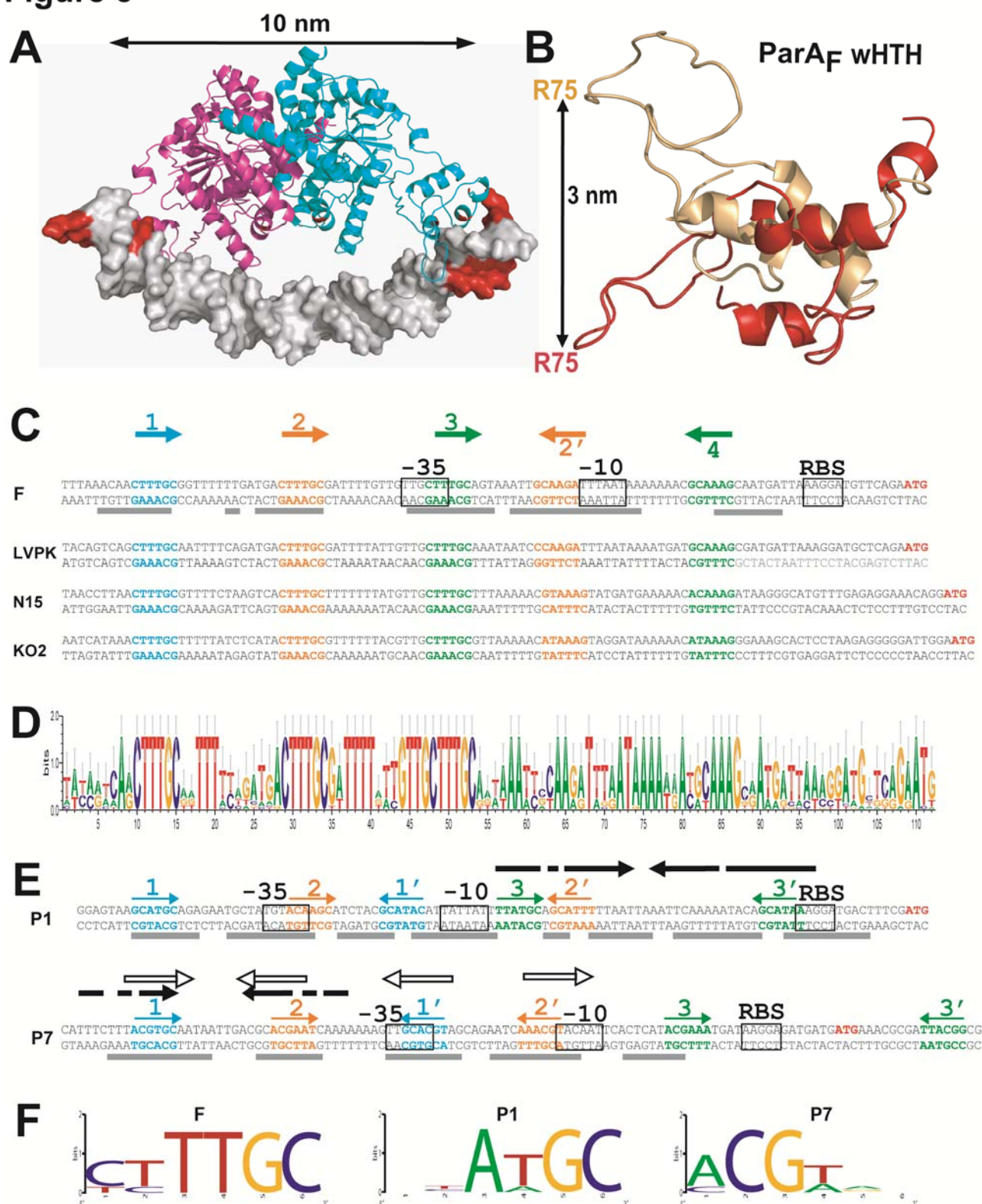


Figure 5. Type Ia plasmid ParA binds to overlapping motifs arranged as inverted repeats.

A. Docking ParA_F onto PparAB_F. ParA_F from the modeled structure was docked onto a 40-bp DNA carrying the IR 3-4 motif. Different DNA bending angles (see Material and Methods) were assayed with one possible conformation of ParA_F dimer. The best docking on the binding motifs was observed with a 120° DNA curvature. ParA_F monomers are represented in magenta and cyan. The DNA is displayed in gray with the binding motif #3 and #4 colored in red.

B. The winged-HTH domain of ParA_F undergoes large conformational rearrangement. The backbone of the wHTH is superimposed before (red) and after (salmon) 40 ns MD simulation. The distance between the C-α atoms of arginine 75 (R75) in the initial and final states is shown by a black line. The full structures are presented in Supplementary Figure S6B.

C. Par promoter regions of the plasmid F and close homologs display a strong conservation of the ParA binding motifs. The sequences of PparAB_F and of three close homologs from LVKP (large virulence plasmid pLVPK of *Klebsiella pneumoniae* CG43), N15 (linear prophage N15 of *E. coli*) and KO2 (phage phiKO2 of *Klebsiella oxytoca*) are shown up to the *parA* start codon (ATG) colored in red. The -10 and -35 transcriptional signals and the ribosome binding sites (RBS) are indicated by open boxes. Binding motifs #1 and IR 2-2' and 3-4, are written in blue, orange and green, respectively, and are represented with corresponding colored arrows. The regions protected by ParA_F in footprint experiments from (21) are indicated with grey lines below the PparAB_F sequence.

D. Partition promoters related to PparAB_F displays a high level of conservation for the inverted repeats motifs and their spacing. Sixteen different partition promoters recovered from Blastn searches using the Par promoter regions of F, LVPK, N15, and KO2 were subjected to quantitative Logo analysis (50). The sequences and the alignment of these promoter regions are presented in Supplementary Figure S5.

E. Promoter regions of the *par* operons of plasmids P1 and P7. DNA sequences and promoter core signals are displayed with the same color codes and drawings as in C. Inverted black arrows indicates the imperfect repeat sequences from a previous study (10). Open arrows over P7 indicates motifs previously identified (10). The region protected from DNase I footprinting by ParA_{P1} (35) (44) and ParA_{P7} (10) are indicated by grey lines below each sequence.

F. Logo analysis of the proposed ParA binding sites for the plasmids F, P1 and P7. The six proposed binding motifs in the PparAB regions were subjected to Logo analysis, except for the plasmid F for which only five motifs were identified (see main text).

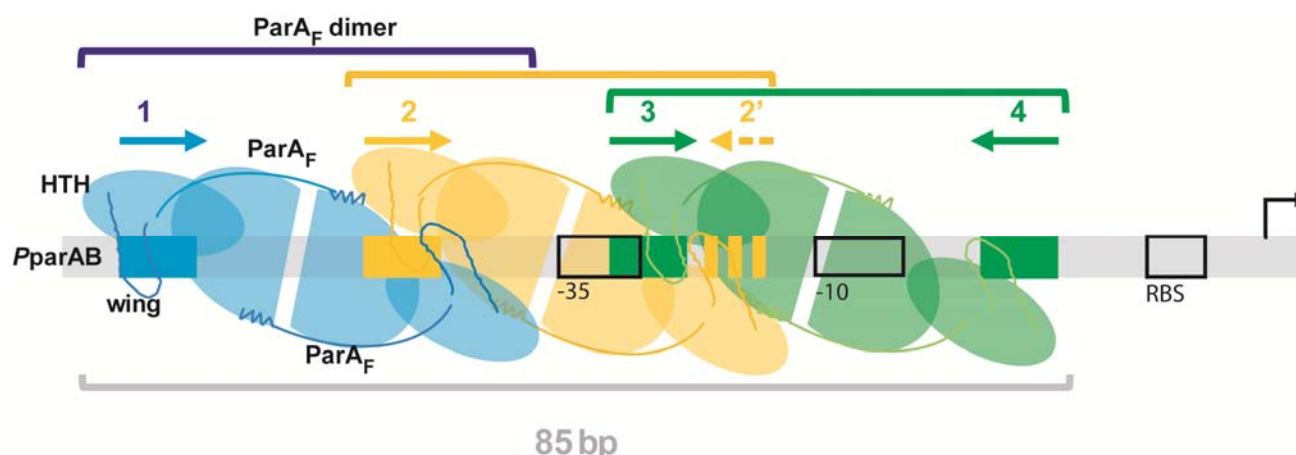


Figure 6. Model of ParA cooperative binding to plasmid partition promoters.

Schematic representation of three ParA dimers bound over the *PparAB* promoter region. DNA is represented by a grey bar with the ParA binding motifs displayed by rectangular boxes and arrows colored accordingly to Figure 5C. Each ParA monomer is represented by a semi-oval (Walker-box domain) linked to an oval (Nter domain) with the same color-code as their respective binding motifs. ParA are dimers in the apo and ADP-bound forms that are stabilized by the first Nter α -helices (represented by curved lines ending by a zigzag) which maintained the two monomers together. The wHTH motifs, symbolized by the ovals with loops (drawn with darker lines) emanating from them, contact specifically the DNA binding motifs. A ParA dimer (green) binds, through the two symmetrically oriented wHTHs, to an inverted repeat motif composed of two perfect binding sites spaced by 28-bp (for F, IR3-4). This initial binding enables the cooperative loading of a second ParA dimer (orange) on imperfect IR motif (IR2-2') composed of a slightly degenerated motif #2' in reverse orientation and spaced by 27-bp relatively to the motif #2, and of a third dimer (blue) on an imperfect IR motif (for F, composed of only 1 perfect binding site (#1)). We propose that these two secondary binding are mediated by dimer-dimer contacts through the wings of two intertwined dimers resulting in a stable complex composed of three dimers bound to the promoter region. The flexible wHTH allow each monomer to contact one hexamer motif by adapting to the local DNA constraint and to interact with an adjacent dimer bound to the overlapping binding motifs through the wing. The overall complex assembled on the promoter region covers ~85-bp. For simplicity, the schematic is drawn with a straight representation and not with bent DNA. ParB may increase ParA binding affinity by enabling an extended conformation proficient for the assembly of a stable complex without interacting directly on the complex.