Permissive zones for the centromere-binding protein ParB on the *Caulobacter crescentus* chromosome

Ngat T. Tran[1‡], Clare E. Stevenson[2‡], Nicolle F. Som[1], Anyarat Thanapipatsiri[1], Adam S. B. Jalal[1], and Tung B. K. Le[1*]

[1]Department of Molecular Microbiology

[2]Department of Biological Chemistry

John Innes Centre, Norwich, NR4 7UH, United Kingdom

[‡]Co-first authors

[*]Correspondence: tung.le@jic.ac.uk

ABSTRACT

Proper chromosome segregation is essential in all living organisms if daughter cells are each to inherit a full copy of genetic information. In *Caulobacter crescentus*, the ParA-ParB-*parS* system is required for proper chromosome segregation and cell viability. The bacterial centromere-like *parS* DNA locus is the first to be segregated following chromosome replication. *parS* is recognized and bound by ParB protein, which in turn interacts with ParA to partition the ParB-*parS* nucleoprotein complex to each daughter cell. In this study, we investigated the genome-wide distribution of ParB on the *Caulobacter* chromosome using a combination of *in vivo* chromatin immunoprecipitation (ChIP-seq) and *in vitro* DNA affinity purification with deep sequencing (IDAP-seq). We confirmed two previously identified *parS* sites and discovered at least three more sites that cluster ~8 kb from the origin of replication. We showed that *Caulobacter* ParB nucleates at *parS* sites, then associates non-specifically with flanking DNA to form a high-order nucleoprotein complex that occupies an extensive ~10 kb DNA segment on the left chromosomal arm. Lastly, using transposon mutagenesis coupled with deep sequencing (Tn-seq), we identified a ~500 kb region surrounding the native *parS* cluster and a ~100 kb region surrounding the terminus of the *Caulobacter* chromosome that are tolerable to the insertion of a second *parS* cluster without severely affecting cell viability. Our results demonstrate that the genomic distribution of the bacterial centromere-like *parS* is highly restricted and is crucial for chromosome segregation in *Caulobacter*.

INTRODUCTION

Proper chromosome segregation is essential in all living organisms if daughter cells are each to inherit a full copy of the genome. In eukaryotes, chromosome segregation during mitosis starts with sister chromosome condensation, followed by the formation of spindle fibres that attach to the kinetochore to pull sister chromatids apart. The kinetochore is the protein structure that assembles on the centromere and links each sister chromatid to microtubules polymers from the mitotic spindle. Unlike in eukaryotes, bacterial chromosome segregation happens without a dedicated spindle-like apparatus (Lim et al., 2014; Vecchiarelli et al., 2013, 2014). Nevertheless, this process is highly organized and also involves protein-based components (reviewed in Badrinarayanan et al., 2015). The first segregated segment of the chromosome is usually proximal to the origin of replication (*ori*) (Lagage et al., 2016; Lin and Grossman, 1998; Livny et al., 2007; Toro et al., 2008). In many bacteria, this region is segregated by the tripartite ParA-ParB-*parS* partitioning system (Fogel and Waldor, 2006; Ireton et al., 1994; Lin and Grossman, 1998; Mohl et al., 2001). *parS* is a centromere-like DNA sequence that most often locates near *ori*. ParB is a DNA-binding protein that nucleates on a *parS* sequence. ParB, via its N-terminal domain, binds DNA non-specifically to spread along the chromosome from its cognate *parS* nucleation site (Graham et al., 2014; Lin and Grossman, 1998; Murray et al., 2006; Taylor et al., 2015). ParB might also bridge distal DNA together to coalesce into a large nucleoprotein complex (the “spreading and bridging” model) (Breier and Grossman, 2007; Graham et al., 2014; Murray et al., 2006; Taylor et al., 2015). Similarly, the formation of the nucleoprotein complex for plasmid-borne ParB-*parS* was proposed to happen via a “nucleation and caging” mechanism where the nucleation of ParB on *parS* creates a high local concentration of ParB, thereby caging ParB dimer-dimer together with non-specific DNA surrounding *parS* (Sanchez et al., 2015). Following ParB binding to *parS*, ParA, a Walker-box ATPase protein, interacts with ParB and powers the segregation of the ParB-DNA nucleoprotein complex to partition replicated chromosomes to each daughter cell (Easter Jr. and Gober, 2002; Leonard et al., 2005).

In *Caulobacter crescentus*, the ParA-ParB-*parS* system is essential for viability (Mohl and Gober, 1997; Mohl et al., 2001). In G1-phase *Caulobacter*, *parS*/*ori* reside at one cell pole, the terminus (*ter*) is near the opposite pole, and the two chromosomal arms run orderly in parallel down the long axis of the cell (Le et al., 2013; Viollier et al., 2004). After replication, the duplicated *parS* sites are released from the pole and separated slightly from one another before one *parS* site is translocated unidirectionally to the opposite cell pole. Toro *et al* (2008) identified two *parS* sites located ~8 kb from the *ori* on the left arm of the *Caulobacter* chromosome, while other works predicted six *parS* sites bioinformatically but did not report their sequences nor verify them experimentally (Bergé et al., 2016; Figge et al., 2003; Mohl and Gober, 1997). Furthermore, it is not yet known whether *Caulobacter* ParB spreads non-specifically on DNA, and if it does, how far it spreads along the chromosome from the *parS* nucleation site. Regarding the genome-wide distribution of *parS* sites, a comparative genomic study suggested that *parS* sites are not distributed randomly on bacterial chromosomes, rather they are found almost exclusively near the *ori* (Livny et al., 2007). Notably, in *Pseudomonas aeruginosa*, *parS* sites must be located within a ~650 kb region surrounding the *ori* for the chromosome segregation to proceed correctly (Lagage et al., 2016).

In this study, we used genome-wide techniques (ChIP-seq and IDAP-seq) together with *in vitro* biochemical characterization to clarify the number and locations of *parS* sites in *Caulobacter*. We show that there are at least five *parS* sites clustered closely near the *ori* of *Caulobacter* chromosome, and that ParB occupies ~10 kb of DNA on the left arm of the chromosome. We also show that *Caulobacter* ParB nucleates on *parS,* then spreads to flanking DNA independent of the location of *parS* on the chromosome. Moreover, using transposon mutagenesis coupled with deep sequencing (Tn-seq), we define a ~500 kb region surrounding the native *parS* cluster, and ~100 kb region surrounding *ter* of the

Caulobacter chromosome that are tolerable to the insertion of a second *parS* cluster without severely affecting cell viability. Our results demonstrate that the genomic location of *parS* is highly biased and crucial for proper chromosome segregation.

RESULTS AND DISCUSSION

ParB occupies a 10 kb DNA region near the origin of replication

To define the distribution of ParB on the chromosome, we performed chromatin immunoprecipitation with deep sequencing. We fused the *flag* tag to the ParB-encoding gene at its 5' end and placed this allele downstream of a vanillate-inducible promoter (P_{van}), at the chromosomal *vanA* locus. The vanillate-inducible *flag-parB* was then transduced to a *Caulobacter* strain where the native and untagged *parB* was under the control of a xylose-inducible promoter (P_{xyl}). *Caulobacter* cells were depleted of untagged ParB by addition of glucose for 5 hours, then vanillate was added for an additional hour before cells were fixed with 1% formadehyde for ChIP-seq (Fig. 1A). *Caulobacter* cells depleted of native ParB while producing the FLAG-tagged ParB version are viable, indicating that the tag does not interfere with ParB function (Fig. S1A). For ChIP-seq, DNA-bound to FLAG-ParB was pulled down using α-FLAG antibody coupled to sepharose beads. The immunoprecipitated DNA was deep sequenced and mapped back to the *Caulobacter* genome to reveal enriched genomic sites (Fig. 1A). As a negative control, we performed α-FLAG ChIP-seq in a *Caulobacter* strain that produces FLAG-tagged YFP, a non-DNA binding protein (Fig. 1B). The ChIP-seq profile of FLAG-ParB showed a clear enrichment in the DNA region on the left chromosomal arm, ~8 kb away from the origin of replication. No other significant enrichment was observed elsewhere on the chromosome or in the negative control (Fig. 1A-B). A closer examination of the *ori*-proximal region revealed an extended ~10 kb region with significant enrichment above background and four defined peaks (Fig. 1A). To independently verify our results, we repeated the ChIP-seq experiment using α-GFP antibody to pull down DNA from a *Caulobacter* strain that produces a CFP-ParB fusion protein from its native location as the only source of ParB in the cell or using a polyclonal α-ParB in a wild-type *Caulobacter* (Fig. S1B). For all cases, we retrieved very similar ChIP-seq profiles to that of FLAG-ParB, suggesting the extended DNA region associating with ParB is not an artefact of tagging but a property of *Caulobacter* ParB itself.

The extensive 10-kb ParB-binding DNA region cannot be explained by the length of DNA fragments that were sheared as part of a ChIP-seq protocol. We sequenced immunoprecipitated DNA from both ends to determine their exact size distribution (Table S1). Pulled-down DNA averages around 150 bp, much smaller than the size of ChIP-seq peaks in our study. However, the extended ParB-binding DNA region can be most easily explained by the non-specific binding of ParB to DNA outside of the *parS* nucleation site, either by a "spreading and bridging" or "caging" mechanism. If so, *Caulobacter* ParB mutants that are impaired in binding to non-specific DNA are predicted to spread less. To identify such mutants in *Caulobacter*, we mutated the highly-conserved N-terminal Box II motif which was shown to be important for the non-specific DNA-binding activity of *B. subtilis* ParB (Fig. S2A) (Breier and Grossman, 2007; Graham et al., 2014). Four variants were constructed *parB* (G101S), *parB* (R103A), *parB* (R104A), and *parB* (R106A). We introduced the *flag*-tagged *parB* mutant allele at the *van* locus, in the P_{xyl}-*parB* genetic background, then employed α-FLAG ChIP-seq to assess the distribution of mutated ParB on the chromosome. Two mutants, ParB (G101S) and ParB (R104A), were found to produce well-defined and symmetrical peaks (~400 bp in width) that are typical of site-specific DNA-binding proteins (Fig. 1B and Fig. S1B). On the contrary, wild-type ParB peaks are much wider and asymmetrical (Fig. 1A). These data suggest that *Caulobacter* ParB, similar to *B. subtilis* and *P. aeruginosa* ParB, also binds DNA non-specifically to spread along the chromosome from its *parS* nucleation sites. Lastly, we noted that DNA enrichment in ChIP-seq experiments with ParB (G101S) or ParB (R104A) is ~5 fold less than that of wild-type ParB (Fig. 1 A-C), despite the fact that ParB variants nucleate equally well on DNA *in vitro* (Fig. S2B). This is

most likely because ParB (G101S) and ParB (R104A) are less stable than wild-type ParB *in vivo* (Fig. S2C).

Identification of *parS* sites and correlating ParB-*parS in vitro* binding affinities to their *in vivo* ChIP-seq enrichment

Since the large width of ChIP-seq peaks obscures the exact position of *parS*, we employed *in vitro* DNA affinity purification with deep sequencing (IDAP-seq) (Belitsky and Sonenshein, 2013) to pinpoint *parS* sequence to near single-nucleotide resolution. Purified ParB-$(His)_6$ was incubated with randomly-fragmented *Caulobacter* genomic DNA, then ParB-DNA complexes were pulled-down using immobilized Ni^{2+} beads. ParB-bound DNA fragments were eluted out and sequenced *en masse*. The sequencing reads were mapped back to either the upper strand or the lower strand of the *Caulobacter* genome (Fig. 1 D and Fig. 2). Analysis of the strand-specific coverage map allows identification of seven 16 bp putative *parS* sites (see Fig. 1D and Fig. S3 for the methodology of IDAP-seq data analysis). These included the two *parS* sites (sites 3 and site 4) that were first discovered in Toro *et al* (2008) but revealed five more putative sites (sites 1, 2, 5, 6 and 7).

To correlate the sequence conservation to the binding affinity of ParB, we measured the equilibrium dissociation constant (K_d) of ParB binding to 24-bp double-stranded oligonucleotides containing individual putative *parS* sites by Surface Plasmon Resonance (SPR) (Fig. 3 and Fig. S4). The double-stranded oligonucleotides was tethered to a chip surface within an SPR flow cell. Purified ParB-$(His)_6$ was flowed over the test DNA. ParB binding was recorded by measuring the change in response units during ParB injection. After injection the chip was washed with buffer and subsequently with high salt buffer to remove any bound ParB. This cycle was repeated for an increasing concentration of ParB dimer to enable the estimation of K_d (Fig. 3 and Fig. S4). Note that the length of the double-stranded oligonucleotides was limited to 24 bp so that only the nucleating event of ParB on *parS* was observed, and not the interaction with DNA flanking *parS*. We observed that sites 2, 3, 4, 5, and 7 have low nM K_d values (Fig. 3), consistent with their high ChIP-seq peaks (Fig. 1). On the other hand, ParB binds to the putative sites 1 and 6 weakly *in vitro*, albeit more than to a scrambled *parS* control (Fig. 3), suggesting that sites 1 and 6 are perhaps unlikely to be significant *in vivo*.

ParB spreads to a maximum of 2 kb around individual *parS* site

The extended ChIP-seq peak of ParB around *parS* is consistent with ParB spreading from the *parS* nucleation site by associating with neighbouring DNA. Since *parS* sites are located within essential genes or genes that have a high fitness cost, we were not able to ablate individual *parS* site to investigate the spreading of ParB in *Caulobacter*. Instead, we investigated the spreading of ParB from individual *parS* sites by expressing the *Caulobacter* ParB/*parS* system in *Escherichia coli*. Since *E. coli* does not possess a ParB homolog nor a *Caulobacter parS*-like sequence, it serves as a suitable heterologous host for this experiment. We inserted individual *parS* sites onto the *E. coli* chromosome at the *ybbD* locus (Fig. 4). The ParB protein was expressed from an IPTG-inducible promoter as a C-terminal fusion to the T18 fragment of *Bordetella pertussis* adenylate cyclase. The T18-ParB is fully functional in *E. coli* as judged by its interactions with their known partners such as ParB itself, ParA, and MipZ in a bacterial-two hybrid assay (Fig. S5A). We induced exponentially-growing *E. coli* cells at 28°C with 500 µM IPTG for an hour before fixing with formadehyde for ChIP-seq. DNA bound to T18-ParB was immunoprecipitated using α-T18 conjugated sepharose beads. A scrambled *parS* site 3 was also inserted at the *ybbD* locus to serve as a negative control. As expected, the strong *parS* sites (sites 2, 3, 4, 5, and 7), on their own showed a high level of DNA enrichment, in agreement with their *in vitro* ParB binding affinity (Fig. 4). The weak putative *parS* sites (site 1 and 6) show little to no enrichment above background (Fig. S5B). Most importantly, we observed that ParB in an *E. coli* host spreads to a maximum of ~2 kb around each *parS* site (Fig. 4). Next, we repeated the ChIP-seq

experiment but with a spreading-defective ParB (G101S). This revealed symmetrical peaks with a ~400-bp width, confirming that *Caulobacter* ParB can spread to any neighbouring DNA and that non-specific interaction with DNA is mainly dependent on an initial ParB-*parS* nucleation event. Lastly, we noted that the spreading of wild-type ParB is not equal on both sides of *parS*. It is likely that the non-specific association of ParB with neighbouring DNA might be influenced by on-going transcription or other nearby DNA-binding proteins. This asymmetrical spreading has been observed previously with ParB homologs from other bacterial species (Attaiech et al., 2015; Breier and Grossman, 2007).

Since *Caulobacter* ParB associates maximally with ~2 kb DNA surrounding individual *parS* site, the clustering of *parS* sites might serve to enable a higher concentration of DNA-bound ParB near *ori* than is possible with a single site. A previously study estimated that ~80% of the total cellular ParB is bound at *parS* sites in *Caulobacter* (Lim et al., 2014). *Caulobacter* ParA was also found to require a high concentration of DNA-bound ParB to activate its ATPase activity, an essential step for chromosome segregation by the ParAB-*parS* system (Lim et al., 2014). Furthermore, it is known that *Caulobacter* ParB interacts with MipZ, which in turns binds PopZ to anchor the *ori*-proximal DNA to the cell pole (Bowman et al., 2008; Ebersbach et al., 2008; Thanbichler and Shapiro, 2006). A high local concentration of DNA-bound ParB would enable a robust anchorage of the *ori* DNA domain to the cell pole. Since high-affinity *parS* sites reside within essential genes or genes with a high fitness cost, we could not systematically ablate *parS* site one-by-one to test whether a reduction DNA-bound ParB affects the chromosome segregation or the anchorage of the *ori*-proximal domain. Nevertheless, we noted that the nucleation-competent but spreading-defective ParB (G101S) or ParB (R104A) variants are unable to support *Caulobacter* growth, implying that interactions with non-specific DNA is required for cell viability (Fig. S1A). In line with our study, *B. subtilis* or *P. aeruginosa* engineered with a single *parS* are defective in chromosome segregation, resulting in elevated numbers of anucleate cells (Breier and Grossman, 2007; Jecz et al., 2015; Lagage et al., 2016).

Extra copies of *parS* can reduce the fitness of *Caulobacter* depending on their genomic locations

Additional copies of *parS*, for example when is placed on a multi-copy number plasmid, can be lethal for cells because plasmid DNA can be segregated instead of the chromosome, resulting in daughter cells with either zero or two chromosomes (Toro et al., 2008). Indeed, we found the presence of a *parS*-carrying plasmid caused growth impairment in *Caulobacter*, and the fitness cost correlates well with the ParB-*parS* binding affinity (Fig. 5). Plasmid-borne sites 3 and 4, which are the strongest *parS* sites, reduced cell viability by ~1000 fold compared to a negative control (scrambled site 3). Extra copies of sites 2, 5 and 7 reduced cell viability by ~100 fold compared to a control, while the weaker *parS* sites 1 and 6 did not impact cell viability when present on a plasmid.

We reasoned that if the toxicity of a plasmid-borne *parS* site was due to the segregation of plasmids instead of the chromosome then having extra *parS* sites on the chromosome should eliminate the toxicity. Indeed, we were able to engineer a 260-bp DNA segment containing both strong *parS* site 3 and site 4 at various positions from *ori* to *ter* on both arms of *Caulobacter* chromosome. On the contrary, a plasmid containing both *parS* sites 3 and 4 are completely lethal to *Caulobacter* cells (Toro et al., 2008). Nevertheless, we noted a variation in the fitness of *Caulobacter* with extra chromosomal *parS* sites, depending on the location of the ectopic *parS* (Fig. 6). An extra $parS^{3+4}$ inserted at +200 kb (near *ori*) or at +1800 kb (near *ter*) did not impact the fitness of the cell dramatically as judged by a normal cell length distribution and a 6-fold increase in the number of anucleate cells (Fig. 6B and Fig. 6D). On the contrary, $parS^{3+4}$ inserted at +1000 kb (middle of the right arm of the chromosome) caused a more severe fitness defect. The cells were more elongated (4.74 ± 3.3 µm) compared to WT (2.97 ± 0.77 µm) (Fig. 6). Furthermore, the number of cells with no or more than two CFP-ParB foci were elevated ~ 11 fold in comparison to strains without an

ectopic *parS*$^{3+4}$ (Fig. 6C). Our data suggest that the genomic location of an extra chromosomal copy of *parS* is important for the cell fitness.

Systematic identification of a permissive zone for *parS* insertion by transposon mutagenesis with deep sequencing (Tn-seq)

Previously, a comparative genomics study surveyed and predicted the positions of *parS* sites over a wide range of bacteria and found that most *parS* sites are located close to the *ori* on the chromosome (Livny et al., 2007). Here, in *Caulobacter*, we have found that a second *parS* cluster, depending on its location on the chromosome, can affect chromosome segregation and cell fitness. To investigate this positional bias systematically, we employed a genome-wide transposon mutagenesis with deep sequencing (Tn-seq) approach. Briefly, a Tn5 transposon carrying *parS* sites 3, 4 and 5 was used to insert these strong *parS* sites randomly around the chromosome. A library of approximately half a million of single colonies were generated and the genomic locations of the inserted *parS* cluster was then determined *en masse* by deep sequencing. As a control, we generated an insertion library using a transposon that does not carry *parS*. Wild-type *Caulobacter* cells were first mutagenized with *parS*$^{+}$ or *parS*$^{-}$ transposon, and the number of insertions was binned to 10-kb segments along the *Caulobacter* chromosome. The ratio of the frequency for the *parS*$^{+}$ transposon and that of the *parS*$^{-}$ transposon was plotted as a $\log_{10}$ scale against genomic position (Fig. 7A), and used as a proxy to determine the genomic preference for an extra cluster of *parS*. We observed that a second *parS* cluster is most tolerated within ~500 kb surrounding *ori* (Fig. 7A and Fig. S6A). In contrast, an ectopic *parS* is strongly disfavoured near the middle of each chromosomal arm (Fig. 7A and Fig. S6B), consistent with our observation that *parS*$^{3+4}$ at +1000 kb caused cell elongation and chromosome segregation defects. A limited zone of *parS* enrichment was also found within ~100 kb around the *ter* (Fig. 7A and Fig. S6C). Lastly, we also note the presence of two *parS* insertion "hot spots". The first hot spot locates near the native *parS* cluster (Fig. 7B), likely strengthening the existing native ParB binding area on the left arm of the chromosome. The second hot spot encompasses the *recF*, *gyrB* and *CCNA0160* genes (Fig. 7C). One possibility is that a *parS* insertion in the vicinity of *gyrB* is preferred because it alters the global supercoiling level. However, we found that the *gyrB* transcription was unchanged compared to wild-type cells or cells with an extra *parS* elsewhere on the chromosome. The mechanism responsible for the *gyrB* "hotspot" therefore remains unknown.

We noted that *parS* insertion frequency decreases gradually from *ori* to the mid-arm without a clear boundary, suggesting that the *parS* permissive zone is perhaps dependent on the genomic distance away either from *ori* or from the native *parS* cluster. To test this hypothesis, we employed a Flip 1-5 strain where the native cluster of *parS* sites were relocated ~400 kb away from *ori* through an inversion between +3611 kb and +4038 kb (Fig. 7D) (Tran et al., 2017). The Tn5 transposon with or without the *parS* cluster was again used to randomly mutagenize the Flip 1-5 strain. As a control, we also transposon mutagenized another inversion strain (Flip 2-5) where the native *parS* cluster remains at its original location but a similar chromosome segment (between +3611 kb and +4030 kb) was inverted (Fig. 7D). Results showed that the permissive zone for insertion of an extra *parS* cluster in Flip 1-5 was now centred near the relocated *parS* site at +3611 kb, while the permissive zone remains centred at the native *parS* in the control Flip 2-5 strain (Fig. 7D) (Tran et al., 2017). Altogether, our results suggest that the genomic distance from the original *parS* cluster, not the distance from *ori*, is likely the main determinant of the permissive zone for the insertion of a second *parS* cluster.

Most bacterial species with a ParAB-*parS* system have more than one *parS* site (Livny et al., 2007), and some species such as *Streptomyces coelicolor* and *Listeria innocua* have accumulated 22 *parS* sites near their origin of replication (Jakimowicz et al., 2002; Livny et al., 2007). How the bacterial centromere-like region expands and what drives its extension over time are interesting biological questions. Our finding that new *parS* sites can locate

near the native *parS* cluster but not elsewhere could potentially explain the clustering of *parS* sites on bacterial chromosomes over time. New *parS* sites preferentially locate near the original *parS* cluster because it is the least disruptive to chromosome segregation and cell viability (Fig. 6 and 7). In *Caulobacter*, *parS,* not *ori*, is the site at which force is exerted during chromosome segregation (Toro et al., 2008). ParA forms a gradient emanating from the opposite pole to the ParB-*parS* cluster. A ParA gradient retracts upon contacting ParB-*parS* and this nucleoprotein complex moves in the retreating gradient of ParA to the opposite cell pole. ParA-ParB-*parS* are only required for the segregation of *parS*-proximal DNA, but not of the distal DNA loci (Badrinarayanan et al., 2015b). Once the *parS*-proximal DNA is properly segregated by ParA-ParB-*parS*, distal DNA regions follow suit, driven by separate molecular machinery, or more likely without the need of a dedicated system (Badrinarayanan et al., 2015b). It is, therefore, foreseeable that expanding the *parS* region by adding new *parS* sites near the native cluster is least disruptive to chromosome segregation since the *parS*-proximal DNA remains the first locus to be segregated. Similarly, in *P. aeruginosa*, *parS* is also the first segregated locus and it is preferable for cell viability that *parS* segregates soon after DNA replication (Lagage et al., 2016).

In this study, we also discovered that new *parS* sites are also tolerated near the *ter* region, albeit with less preference than near the native *parS* cluster. In *P. aeruginosa* or *B. subtilis*, insertion of *parS* near the *ter* region is strongly discouraged, presumably due to the recruitment of the Structural Maintenance of the Chromosomes (SMC) complex away from *ori* (Lagage et al., 2016; Sullivan et al., 2009). SMC is a prominent protein involved in bacterial chromosome organization and segregation (Gruber and Errington, 2009; Minnen et al., 2011; Schwartz and Shapiro, 2011; Sullivan et al., 2009; Tran et al., 2017). To test if SMC might contribute to shape the distribution of ectopic *parS* sites in *Caulobacter*, we transposon mutagenized the *Δsmc Caulobacter* strain (Fig. S6D). In *Δsmc* cells, the pattern of *parS* permissive zones does not change dramatically. New *parS* sites remain disfavoured near mid-arms, although they are less favoured near *ter* compared to wild-type cells (Fig. S6D). Our previous study showed that *Caulobacter* SMC are recruited to the *ter*-located ectopic *parS* and cohese flanking DNA together, nevertheless the global chromosome organization remained largely unchanged with *ori* and *ter* at opposite poles and two chromosomal arms running in parallel down the long axis of the cell (Tran et al., 2017). All together, we conclude that SMC contributes to the determination of *parS* permissive zones but cannot solely explain some of the preference for the *ter* region and the disfavour for mid-arm regions in *Caulobacter crescentus*. Further investigation into the molecular mechanism that gives raise to the permissive zones of *parS* will undoubtedly improve our understanding of bacterial chromosome segregation and organization.

MATERIALS AND METHODS

All experimental procedures are reported in the Supplementary Information

AUTHOR CONTRIBUTIONS

Conceptualization, N.T.T. and T.B.K.L.; Investigation and Data analysis, N.T.T., C.E.S., N.F.S., A.T., A.S.B.J., and T.B.K.L.; Writing, N.T.T., C.E.S., and T.B.K.L.; and Funding Acquisition, T.B.K.L.

ACCESSION NUMBERS

The accession number for the sequencing data reported in this paper is GEO: **GSE100233**.

ACKNOWLEDGMENTS

We thank Anjana Badrinarayanan, Matthew Bush, Mark Buttner, Michael Laub, and Susan Schlimpert for discussion and comments on the manuscript. We thank Christine Jacob-Wagner and Martin Thanbichler for materials. This study was supported by a Royal Society University Research Fellowship (UF140053), a Royal Society Research Grant (RG150448) to T.B.K.L. and a BBSRC grant-in-add (BBS/E/J/000C0683) to the John Innes Centre.

REFERENCES

- Attaiech, L., Minnen, A., Kjos, M., Gruber, S., and Veening, J.-W. (2015). The ParB-parS Chromosome Segregation System Modulates Competence Development in Streptococcus pneumoniae. MBio *6*, e00662-15.
- Badrinarayanan, A., Le, T.B.K., and Laub, M.T. (2015a). Bacterial Chromosome Organization and Segregation. Annu. Rev. Cell Dev. Biol. *31*, 171–199.
- Badrinarayanan, A., Le, T.B.K., and Laub, M.T. (2015b). Rapid pairing and resegregation of distant homologous loci enables double-strand break repair in bacteria. J. Cell Biol. *210*, 385–400.
- Belitsky, B.R., and Sonenshein, A.L. (2013). Genome-wide identification of Bacillus subtilis CodY-binding sites at single-nucleotide resolution. Proc. Natl. Acad. Sci. U. S. A. *110*, 7026–7031.
- Bergé, M., Campagne, S., Mignolet, J., Holden, S., Théraulaz, L., Manley, S., Allain, F.H.-T., and Viollier, P.H. (2016). Modularity and determinants of a (bi-)polarization control system from free-living and obligate intracellular bacteria. ELife *5*.
- Bowman, G.R., Comolli, L.R., Zhu, J., Eckart, M., Koenig, M., Downing, K.H., Moerner, W.E., Earnest, T., and Shapiro, L. (2008). A Polymeric Protein Anchors the Chromosomal Origin/ParB Complex at a Bacterial Cell Pole. Cell *134*, 945–955.
- Breier, A.M., and Grossman, A.D. (2007). Whole-genome analysis of the chromosome partitioning and sporulation protein Spo0J (ParB) reveals spreading and origin-distal sites on the Bacillus subtilis chromosome. Mol. Microbiol. *64*, 703–718.
- Easter Jr., J., and Gober, J.W. (2002). ParB-Stimulated Nucleotide Exchange Regulates a Switch in Functionally Distinct ParA Activities. Mol. Cell *10*, 427–434.
- Ebersbach, G., Briegel, A., Jensen, G.J., and Jacobs-Wagner, C. (2008). A self-associating protein critical for chromosome attachment, division, and polar organization in caulobacter. Cell *134*, 956–968.
- Figge, R.M., Easter, J., and Gober, J.W. (2003). Productive interaction between the chromosome partitioning proteins, ParA and ParB, is required for the progression of the cell cycle in Caulobacter crescentus. Mol. Microbiol. *47*, 1225–1237.
- Fogel, M.A., and Waldor, M.K. (2006). A dynamic, mitotic-like mechanism for bacterial chromosome segregation. Genes Dev. *20*, 3269–3282.
- Graham, T.G.W., Wang, X., Song, D., Etson, C.M., Oijen, A.M. van, Rudner, D.Z., and Loparo, J.J. (2014). ParB spreading requires DNA bridging. Genes Dev. *28*, 1228–1238.
- Gruber, S., and Errington, J. (2009). Recruitment of condensin to replication origin regions by ParB/SpoOJ promotes chromosome segregation in B. subtilis. Cell *137*, 685–696.
- Ireton, K., Gunther, N.W., and Grossman, A.D. (1994). spo0J is required for normal chromosome segregation as well as the initiation of sporulation in Bacillus subtilis. J. Bacteriol. *176*, 5320–5329.
- Jakimowicz, D., Chater, K., and Zakrzewska-Czerwi´nska, J. (2002). The ParB protein of Streptomyces coelicolor A3(2) recognizes a cluster of parS sequences within the origin-proximal region of the linear chromosome. Mol. Microbiol. *45*, 1365–1377.
- Jecz, P., Bartosik, A.A., Glabski, K., and Jagura-Burdzy, G. (2015). A single parS sequence from the cluster of four sites closest to oriC is necessary and sufficient for proper chromosome segregation in Pseudomonas aeruginosa. PloS One *10*, e0120867.
- Lagage, V., Boccard, F., and Vallet-Gely, I. (2016). Regional Control of Chromosome Segregation in Pseudomonas aeruginosa. PLOS Genet. *12*, e1006428.
- Le, T.B., Imakaev, M.V., Mirny, L.A., and Laub, M.T. (2013). High-resolution mapping of the spatial organization of a bacterial chromosome. Science *342*, 731–734.

- Leonard, T.A., Butler, P.J., and Löwe, J. (2005). Bacterial chromosome segregation: structure and DNA binding of the Soj dimer--a conserved biological switch. EMBO J. *24*, 270–282.
- Lim, H.C., Surovtsev, I.V., Beltran, B.G., Huang, F., Bewersdorf, J., and Jacobs-Wagner, C. (2014). Evidence for a DNA-relay mechanism in ParABS-mediated chromosome segregation. ELife *3*, e02758.
- Lin, D.C.-H., and Grossman, A.D. (1998). Identification and Characterization of a Bacterial Chromosome Partitioning Site. Cell *92*, 675–685.
- Livny, J., Yamaichi, Y., and Waldor, M.K. (2007). Distribution of Centromere-Like parS Sites in Bacteria: Insights from Comparative Genomics. J. Bacteriol. *189*, 8693–8703.
- Minnen, A., Attaiech, L., Thon, M., Gruber, S., and Veening, J.-W. (2011). SMC is recruited to oriC by ParB and promotes chromosome segregation in Streptococcus pneumoniae. Mol. Microbiol. *81*, 676–688.
- Mohl, D.A., and Gober, J.W. (1997). Cell cycle-dependent polar localization of chromosome partitioning proteins in Caulobacter crescentus. Cell *88*, 675–684.
- Mohl, D.A., Easter, J., and Gober, J.W. (2001). The chromosome partitioning protein, ParB, is required for cytokinesis in Caulobacter crescentus. Mol. Microbiol. *42*, 741–755.
- Murray, H., Ferreira, H., and Errington, J. (2006). The bacterial chromosome segregation protein Spo0J spreads along DNA from parS nucleation sites. Mol. Microbiol. *61*, 1352–1361.
- Sanchez, A., Cattoni, D.I., Walter, J.-C., Rech, J., Parmeggiani, A., Nollmann, M., and Bouet, J.-Y. (2015). Stochastic Self-Assembly of ParB Proteins Builds the Bacterial DNA Segregation Apparatus. Cell Syst. *1*, 163–173.
- Schwartz, M.A., and Shapiro, L. (2011). An SMC ATPase mutant disrupts chromosome segregation in Caulobacter. Mol Microbiol *82*, 1359–1374.
- Sullivan, N.L., Marquis, K.A., and Rudner, D.Z. (2009). Recruitment of SMC by ParB-parS Organizes the Origin Region and Promotes Efficient Chromosome Segregation. Cell *137*, 697–707.
- Taylor, J.A., Pastrana, C.L., Butterer, A., Pernstich, C., Gwynn, E.J., Sobott, F., Moreno-Herrero, F., and Dillingham, M.S. (2015). Specific and non-specific interactions of ParB with DNA: implications for chromosome segregation. Nucleic Acids Res. *43*, 719–731.
- Thanbichler, M., and Shapiro, L. (2006). MipZ, a Spatial Regulator Coordinating Chromosome Segregation with Cell Division in Caulobacter. Cell *126*, 147–162.
- Toro, E., Hong, S.-H., McAdams, H.H., and Shapiro, L. (2008). Caulobacter requires a dedicated mechanism to initiate chromosome segregation. Proc. Natl. Acad. Sci. *105*, 15435–15440.
- Tran, N.T., Laub, M.T., and Le, T.B. (2017). SMC Progressively Aligns Chromosomal Arms In Caulobacter crescentus But Is Antagonized By Convergent Transcription. BioRxiv 125344. *(Accepted in Cell Reports)*
- Vecchiarelli, A.G., Hwang, L.C., and Mizuuchi, K. (2013). Cell-free study of F plasmid partition provides evidence for cargo transport by a diffusion-ratchet mechanism. Proc. Natl. Acad. Sci. U. S. A. *110*, E1390-1397.
- Vecchiarelli, A.G., Neuman, K.C., and Mizuuchi, K. (2014). A propagating ATPase gradient drives transport of surface-confined cellular cargo. Proc. Natl. Acad. Sci. U. S. A. *111*, 4880–4885.
- Viollier, P.H., Thanbichler, M., McGrath, P.T., West, L., Meewan, M., McAdams, H.H., and Shapiro, L. (2004). Rapid and sequential movement of individual chromosomal loci to specific subcellular locations during bacterial DNA replication. Proc Natl Acad Sci U A *101*, 9257–9262.

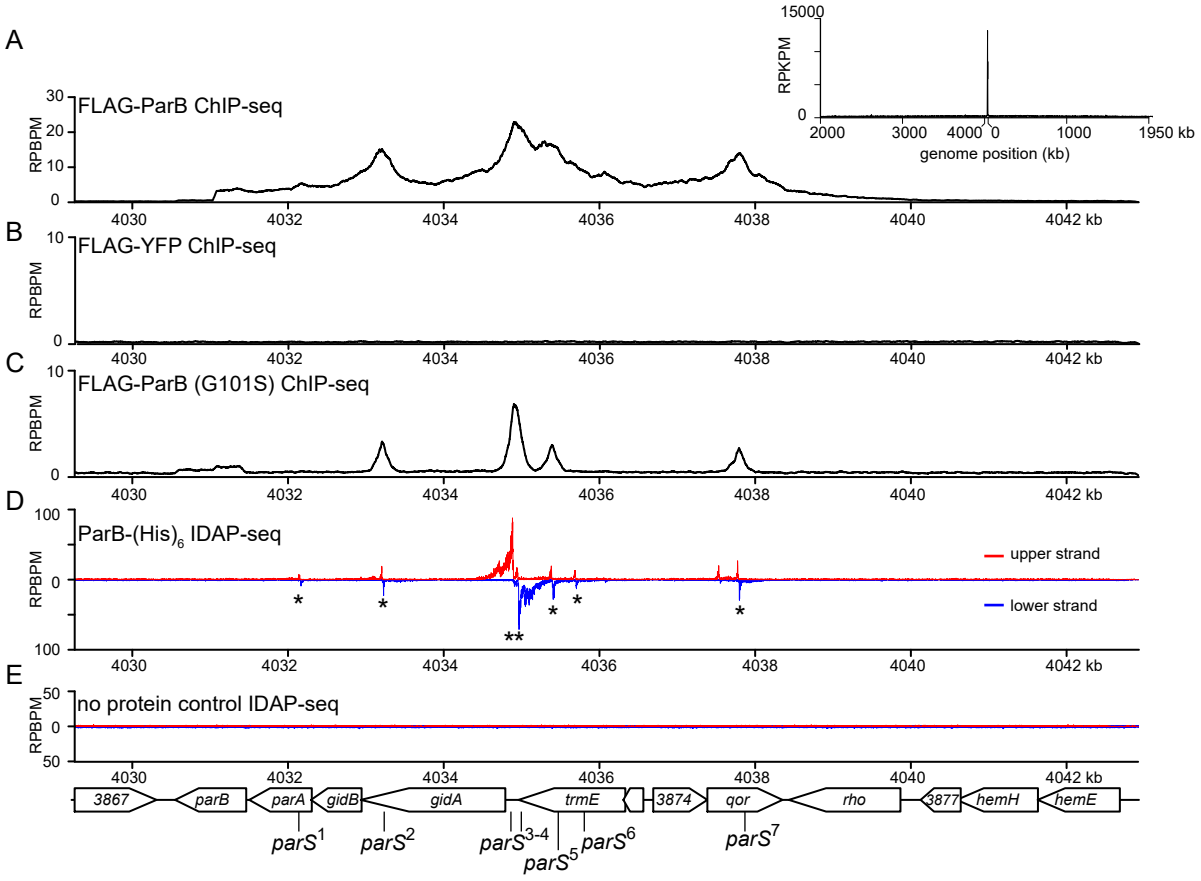

Figure 1. ParB occupies 10 kb DNA region near the origin of replication
(A) The distribution of FLAG-tagged ParB on *Caulobacter* chromosome between +4030 kb and +4042 kb. ChIP-seq signals were reported as the number of reads at every nucleotide along the genome (RPBPM value). The whole-genome ChIP-seq profile of ParB is shown in the inset. For the whole genome profile, the ChIP-seq signals were reported as the number of reads at every kb along the genome (RPKPM value). **(B)** ChIP-seq profile of FLAG-tagged YFP. **(C)** ChIP-seq profile of FLAG-tagged ParB (G101S) mutant. **(D)** IDAP-seq profile of ParB-(His)6 with sonication-fragmented genomic DNA from *Caulobacter*. IDAP-seq reads were sorted to either the upper strand (red) or to the lower strand (blue) of the reference genome to enable identification of *parS* sites (see also Fig. 2 and Fig. S3). Putative *parS* sites (1 to 7) are noted with asterisks (see also Fig. 2). **(E)** IDAP-seq profile of a negative control in which ParB-$(His)_6$ was omitted.

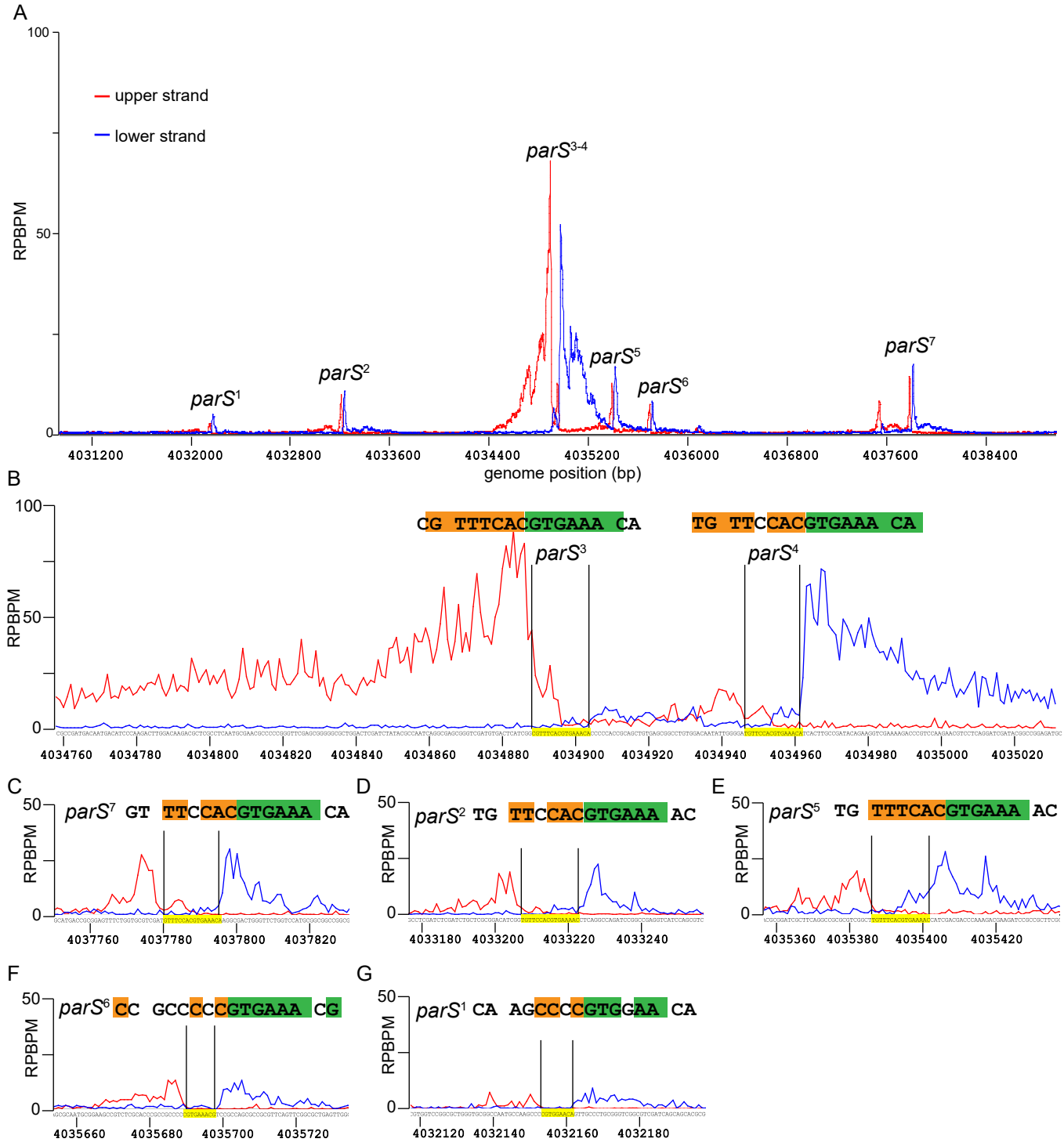

Figure 2. Identification of *parS* sequences by *in vitro* DNA purification with deep sequencing (IDAP-seq)

Sequencing reads were sorted to either the upper DNA strand (red) or to the lower strand (blue) of the *Caulobacter* reference gnome, as suggested in the original IDAP-seq publication (Belitsky and Sonenshein, 2013, Fig. S3). The sequence in between the summit of the upper strand profile and that of the lower strand profile defines the *parS* sequence required for binding to ParB *in vitro* (See also Fig. S3). **(A)** IDAP-seq profile of ParB-$(His)_6$ in the genomic region between +4031 kb and +4039 kb. **(B-G)** IDAP-seq profile of ParB-$(His)_6$ surrounding each individual *parS* site. Palindromic nucleotides within the identified *parS* site are shaded in orange and green.

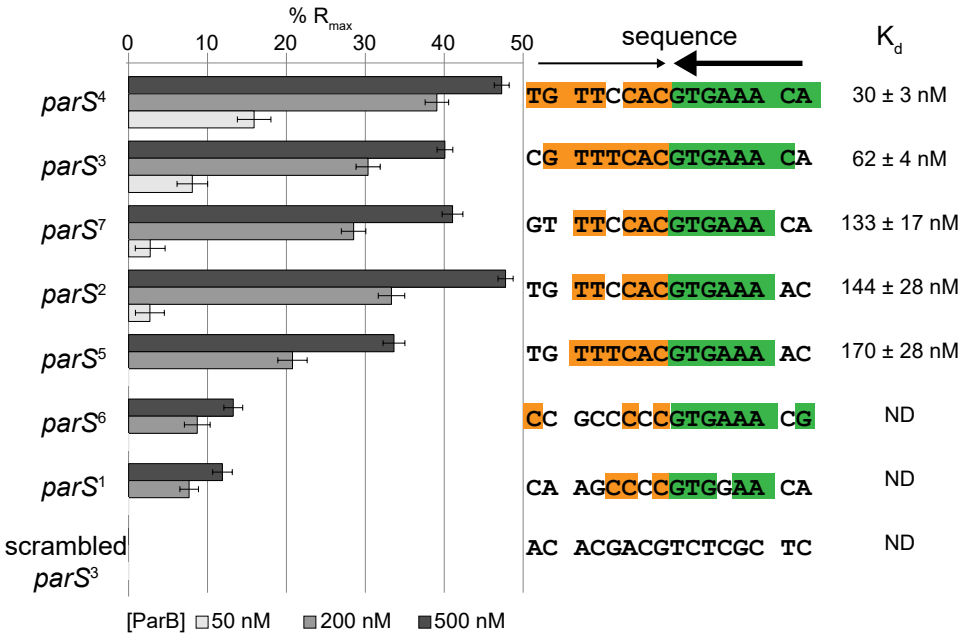

Figure 3. ParB-*parS in vitro* binding affinities correlate to their *in vivo* ChIP-seq enrichment

Surface Plasmon Resonance (SPR) was used to measure binding affinity of ParB (50 nM, 200 nM and 500 nM) to 24-bp double-stranded DNA that contains individual putative *parS* site. The level of ParB binding to DNA was expressed as a percentage of the theoretical maximum response, Rmax, assuming a single ParB dimer binding to one immobilized double-stranded DNA oligomer. This normalization process enabled the various responses to be readily compared, irrespective of the quantity length of the DNA tethered on an SPR chip surface. A wider range of ParB concentration (6.25 nM, 12.5 nM, 25 nM, 50 nM, 100 nM, 200 nM, 400 nM, 600 nM and 800 nM) was used to estimate the binding constant (K_d) of ParB to individual *parS* site (Fig. S4). The sequences of *parS* are shown with palindromic nucleotides shaded in orange and green. Convergent arrows on top of *parS* sequence indicate that *parS* sites are palindromic. Thicker arrow signifies that the second half of *parS* sequences (GTGAAA, in green) is conserved among *Caulobacter parS* sites.

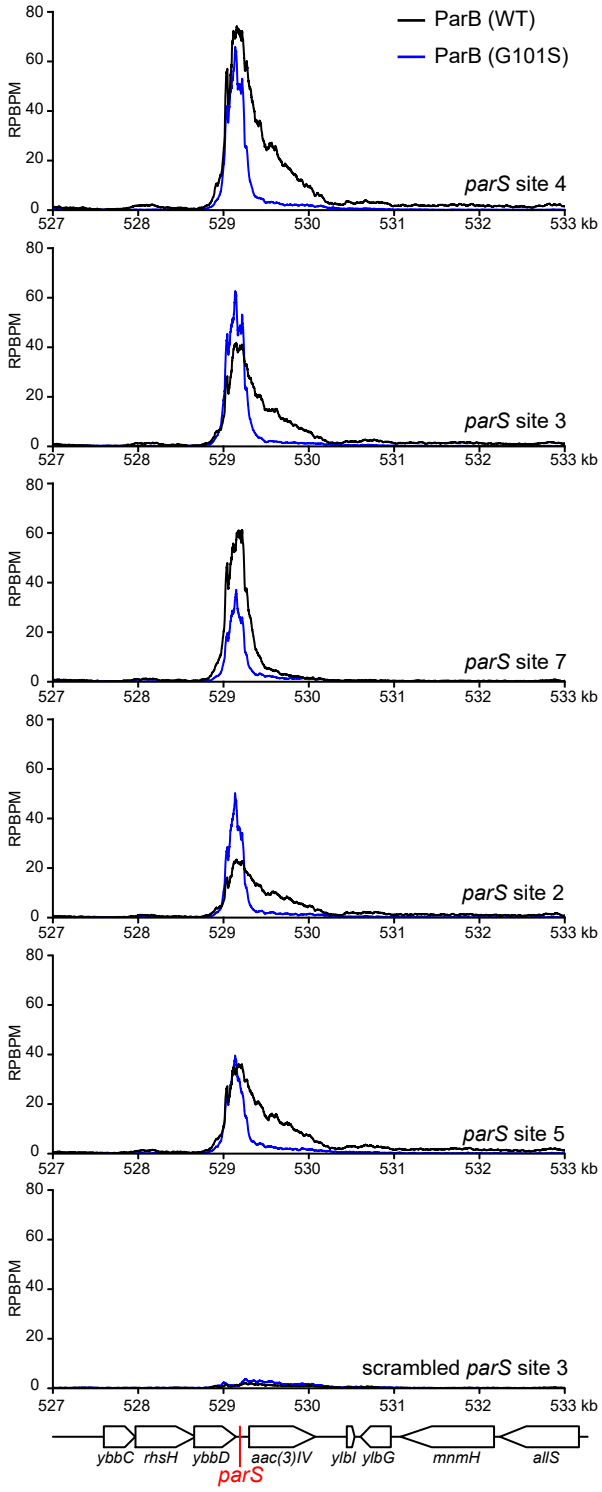

Figure 4. *Caulobacter* ParB binds to *parS* and spreads to flanking DNA in a heterologous *E. coli* host

A cassette composed of individual *parS* (red line) site and an apramycin resistance marker *aac(3)IV* was inserted at the *yybD* locus on an *E. coli* chromosome. T18-ParB (WT) (black) or T18-ParB (G101S) (blue) were expressed from an IPTG-inducible promoter, and their distribution on the *E. coli* chromosome were determined by α-T18 ChIP-seq. ChIP-seq signals were reported as the number of reads at every nucleotide along the genome (RPBPM value). A cassette composed of a scrambled *parS* site 3 and an apramycin resistance marker was also inserted at the *yybD* locus and serves as a negative control.

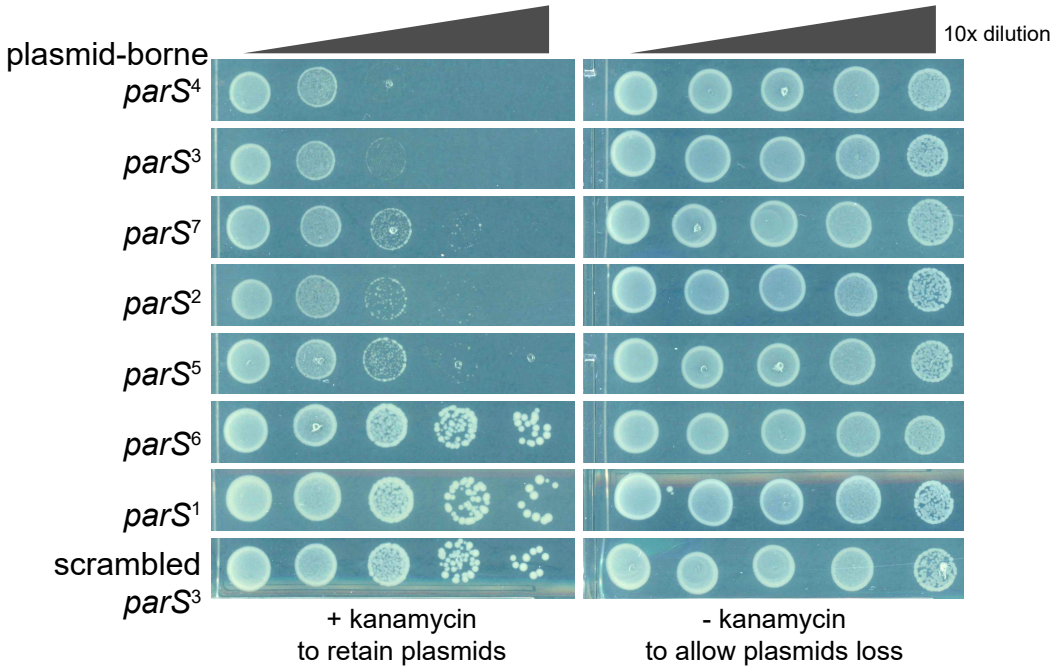

Figure 5. Plasmid-borne *parS* reduces the fitness of *Caulobacter*
Low-copy number plasmid harbouring individual *parS* site was conjugated from *E. coli* S17-1 to wild-type *Caulobacter.* The same number of *E. coli* and *Caulobacter* cells were used for each conjugation. A ten-fold serial dilution was performed and spotted on PYE plates supplemented with both nalidixic acid and kanamycin or just with nalidixic acid. Addition of kanamycin enforces the retention of *parS* plasmid, while omitting kanamycin allows plasmid loss. All cells were spotted on the same +kanamycin or -kanamycin plates, and pictures were taken after 3 day incubation at 30°C.

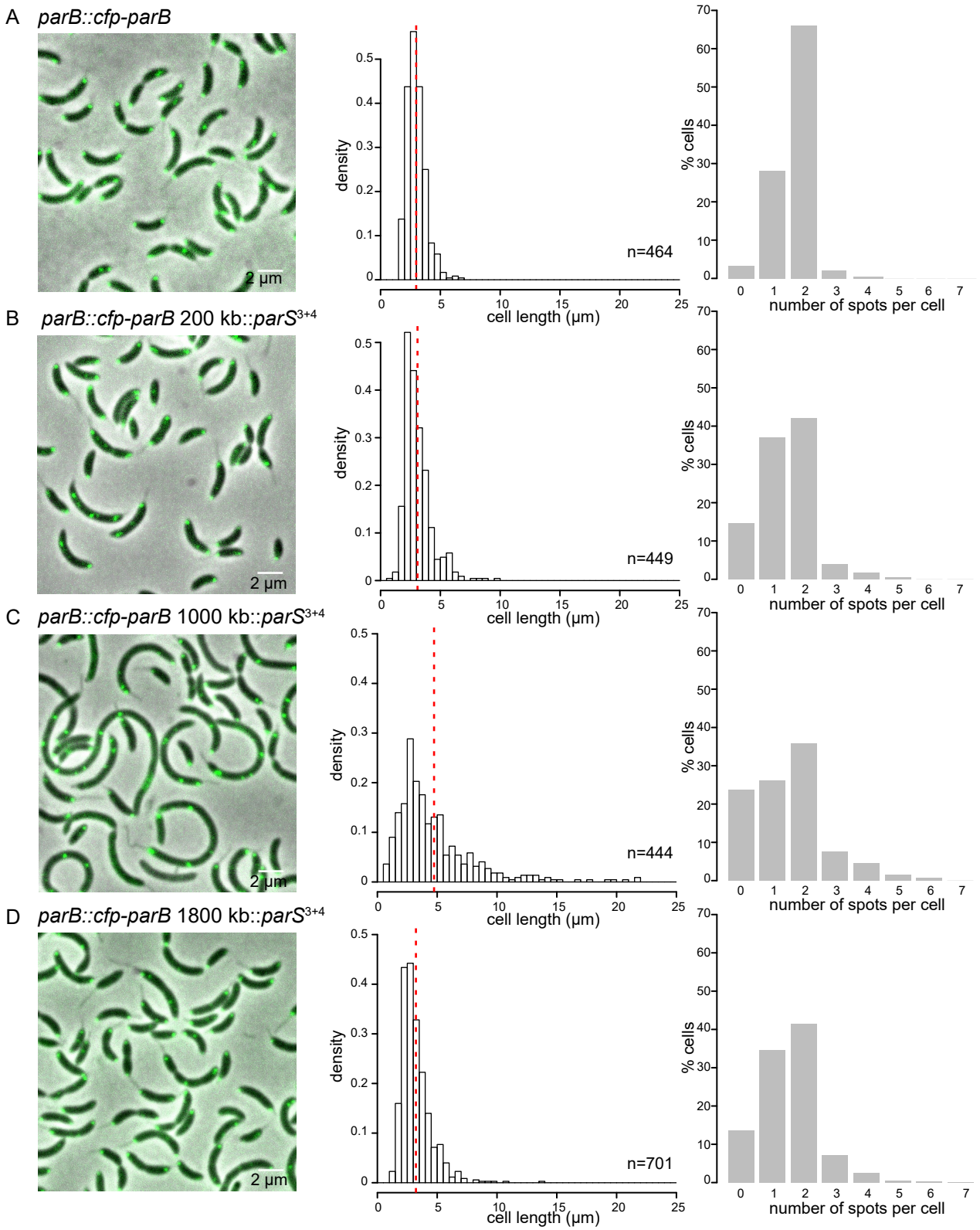

Figure 6. The position of an ectopic *parS* on the chromosome is critical for the fitness of *Caulobacter*

Micrograph of *parB::cfp-parB Caulobacter* cells **(A)** without an extra ectopic $parS^{3+4}$, **(B)** with an extra ectopic $parS^{3+4}$ at +200 kb, **(C)** at +1000 kb, or **(D)** at +1800 kb. Cell length of an exponentially-growing cells were quantified and presented as histograms. Vertical dotted red lines indicate the mean cell length. The number of CFP-ParB foci (green) per cell was also quantified and plotted as histograms. Note that we could not observe foci corresponding to an extra ectopic $parS^{3+4}$ perhaps due to the limited numbers of ParB bound to this shorter cluster. Most observable foci are likely due to the original $parS^{1-7}$ cluster that reside ~8 kb near ori.

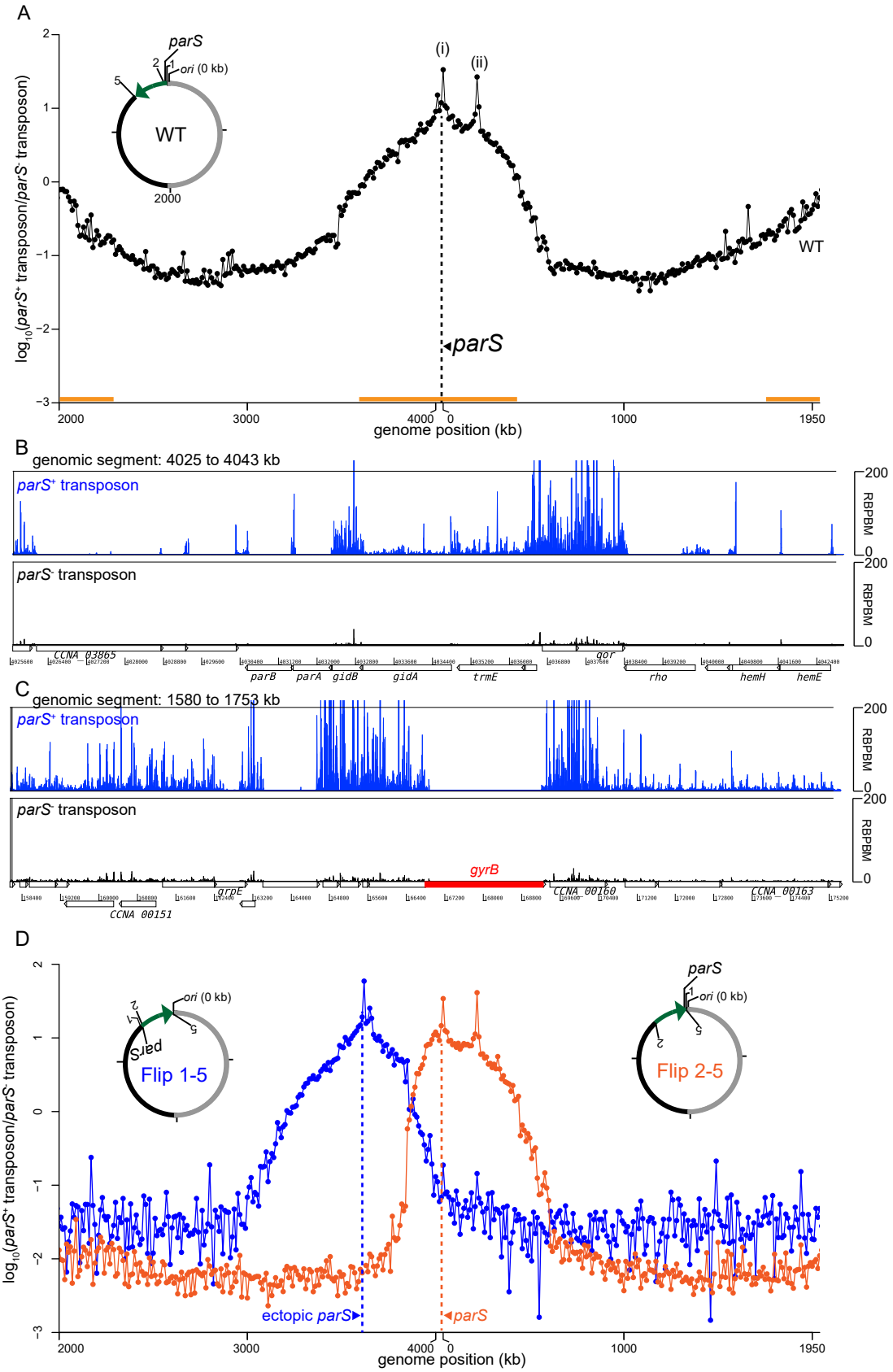

Figure 7. Tn5-seq reveals the positional bias of the centromeric *parS* site on *Caulobacter* chromosome
(A) Wild-type *Caulobacter* cells were mutagenized with the *parS*$^+$ or *parS*$^-$ transposon, and the number of insertions was binned to 10-kb segments along the genome. The ratio between insertion frequency for the *parS*$^+$ transposon and that of the *parS*$^-$ transposon was calculated and plotted as a log10 scale against genomic position. Two hotspots for insertion of the *parS*$^+$ transposon are marked with asterisks (*). The vertical dotted line (black) shows the position of the native *parS* cluster. The horizontal bar (orange) indicates the permissive zone for extra *parS* insertions. **(B-C)** Comparison between *parS*$^+$ (blue) and *parS*$^-$ (black) transposon insertions for the genomic segment between +4025 kb and +4043 kb, and between +158 kb and +175 kb. **(D)** *parS*$^+$/*parS*$^-$ Tn5-seq profiles for Flip1-5 (blue) and Flip 2-5 (orange) strains. The horizontal axis represents genome position in kilobases for each strain. A schematic genomic map of Caulobacter showing the position of *parS* and *ori* are presented in the inset. The inverted DNA segment (green arrow) is indicated together with the end points of the inversion (1, 2, and 5).

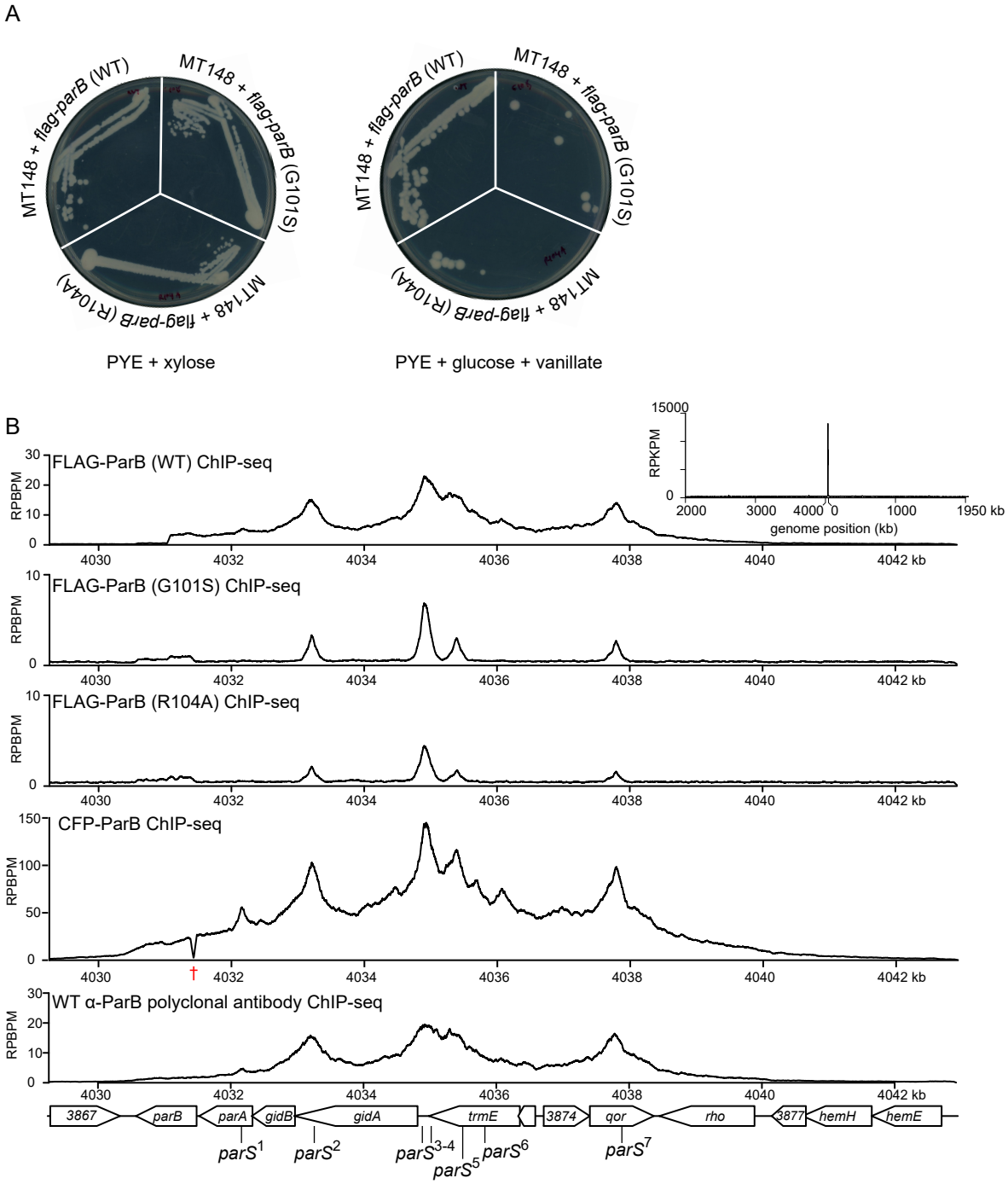

Figure S1. Genomic distributions of wild-type *Caulobacter* ParB and spreading-defective ParB (G101S) and ParB (R104A) variants. (A) *Caulobacter* strains *parB::*P$_{xyl}$*-parB van::*P$_{van}$*-flag-parB* WT, G101S, or R104A were restruck on PYE + xylose to induce the expression of the wild-type untagged ParB, or on PYE + glucose + vanillate to repress the expression of the wild-type untagged ParB while expressing the FLAG-tagged ParB WT, G101S, or R104A. The FLAG-tagged version of wild-type ParB is functional and can complement the depletion of wild-type untagged ParB while the spreading mutant ParB (G101S) or ParB (R1014A) cannot. **(B)** ChIP-seq profiles of FLAG-ParB (WT), FLAG-ParB (G101S), and (R104A) (using α-FLAG antibody), of CFP-ParB (using α-GFP antibody), and of ParB (using polyclonal α-ParB antibody). Note: the red dagger (†) symbol on the CFP-ParB ChIP-seq profile indicates the genomic region where sequencing reads were missing. This is because CFP-ParB ChIP-seq reads were mapped to the wild-type *Caulobacter* reference genome instead of to the genome of *parB::cfp-parB* strain.

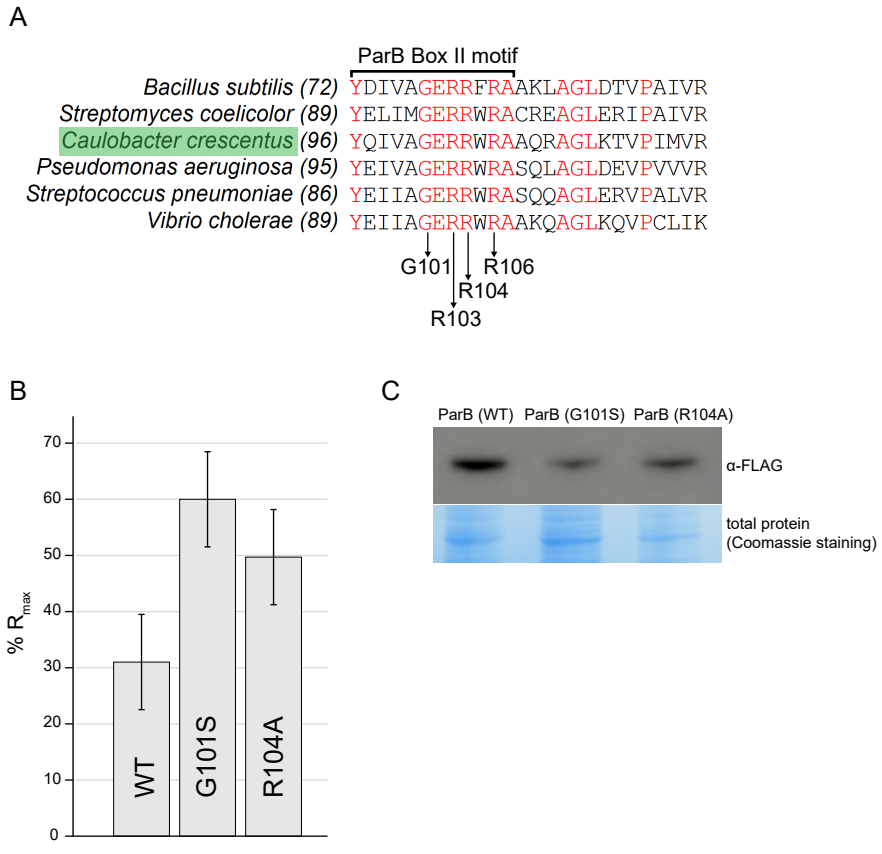

Figure S2. Identification of potential spreading-defective ParB mutants in *Caulobacter crescentus*. (A) A sequence alignment of *Caulobacter* ParB and homologs to highlight the conservation of the Box II motif (square bracket). Identical amino acids residues are shown in red. Vertical arrows indicate the position of G101, R103, R104, and R106 residues. **(B)** Surface Plasmon Resonance (SPR) was used to measure binding affinity of ParB WT, ParB (G101S) and ParB (R104A) at 200 nM to a 24-bp double-stranded DNA that contains *parS* site 4. The level of ParB variants binding to DNA was expressed as a percentage of the theoretical maximum response, R_{max}, assuming a single ParB dimer binding to one immobilized double-stranded DNA oligonucleotides. This normalization process enabled the various responses to be readily compared, irrespective of the quantity and length of the DNA tethered on an SPR chip surface. **(C)** Immunoblot analysis of FLAG-tagged ParB WT vs. G101S and R104A. Cells were depleted of wild-type untagged ParB for 5 hours, then vanillate was added for an additional hour to allow for expression of FLAG-tagged ParB. Equal amount of total protein was loaded on each well of the SDS-PAGE.

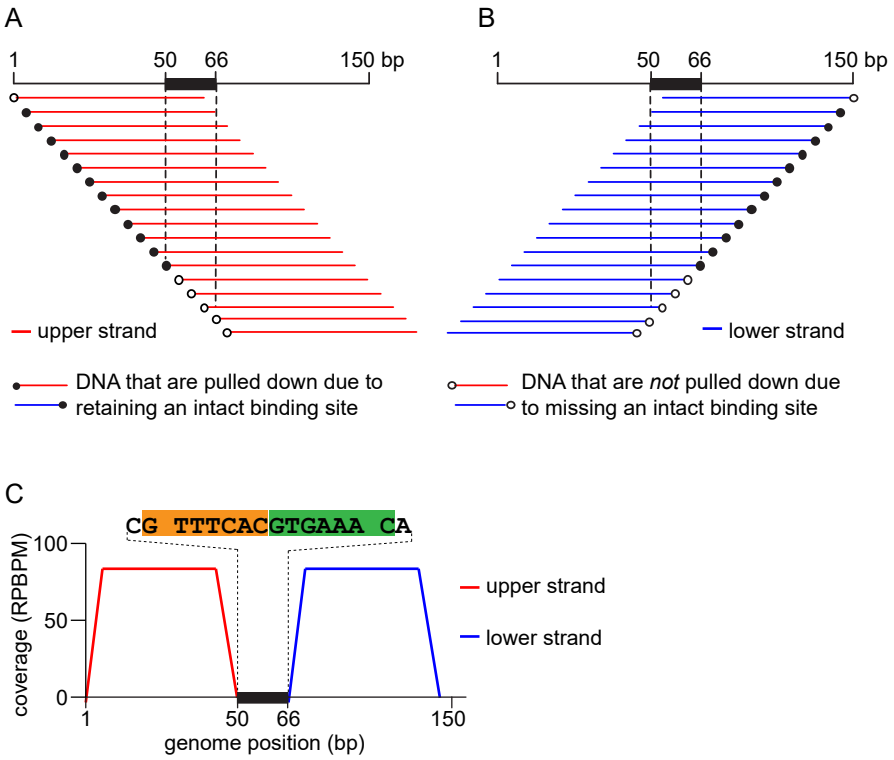

Figure S3. Methodology for the analysis of *in vitro* affinity purification with deep sequencing (IDAP-seq) data. Sequencing reads were sorted to either the upper DNA strand **(A)** or to the lower strand **(B)** of the *Caulobacter* reference genome (Belitsky and Sonenshein, 2013). The 5' nucleotides are shown as circles. Solid circles are for DNA fragments with an intact ParB binding site (black rectangle), and open circles are for DNA fragments with a partial or no ParB binding site. Only DNA fragments with an intact ParB binding site will be pulled down during affinity purification and contribute to the sequencing coverage. **(C)** A schematic strand-specific coverage map of IDAP-seq. A footprint of ParB can be identified in between the two edges of the upper-strand peak (red) and the lower-strand peak (blue). The schematic picture was adapted from Belitsky and Sonenshein (2013) with permission from A. Sonenshein.

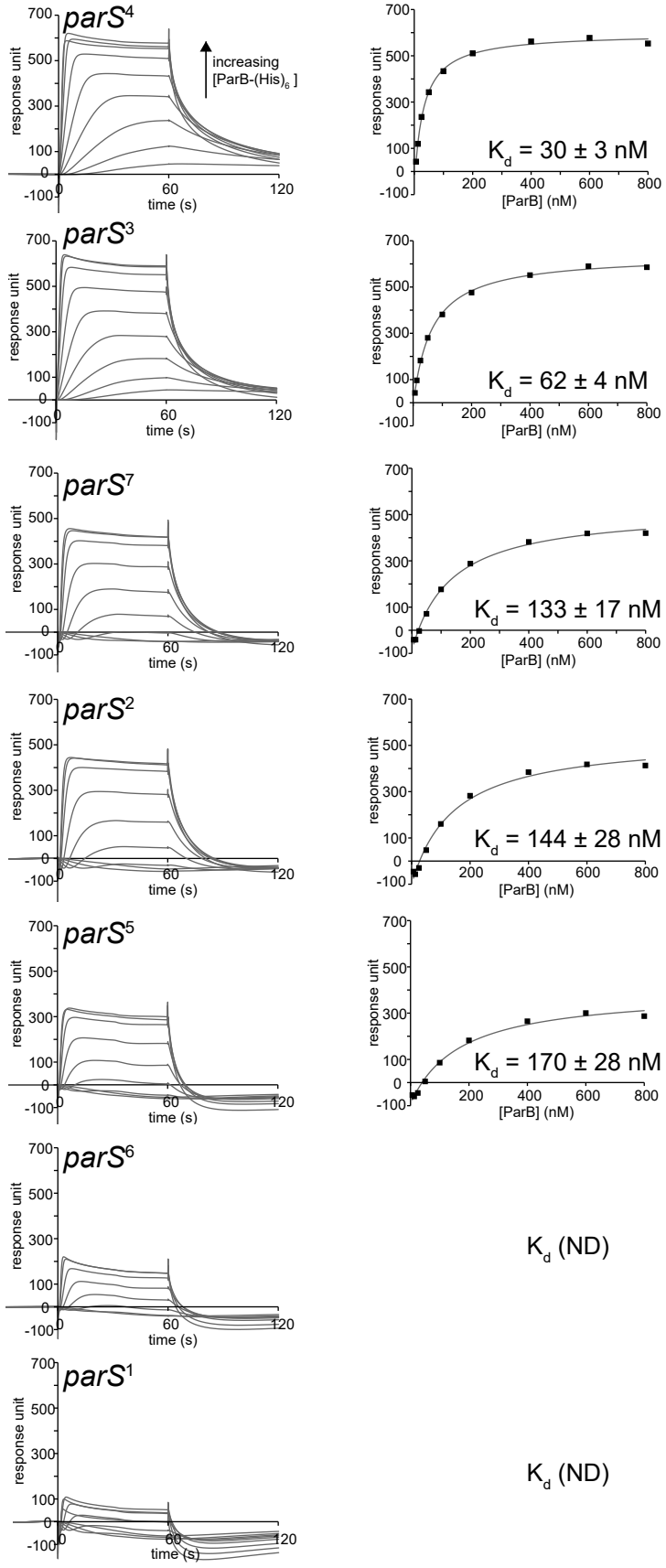

Figure S4. Determination of binding affinity constants (K_d) of ParB-*parS* interactions. 24-bp duplex DNA containing each individual *parS* was tethered on an SPR chip surface. Increasing concentrations of ParB (6.25 nM, 12.5 nM, 25 nM, 50 nM, 100 nM, 200 nM, 400 nM, 600 nM and 800 nM) were flown through the SPR chip surface. Binding of ParB to *parS* site was recorded and expressed as response unit (RU). Response units were plotted against ParB concentration and curve fitted to estimate K_d value ± standard deviation. K_d for *parS* site 1 and site 6 were not determined due to very little specific binding of ParB to the tested DNA.

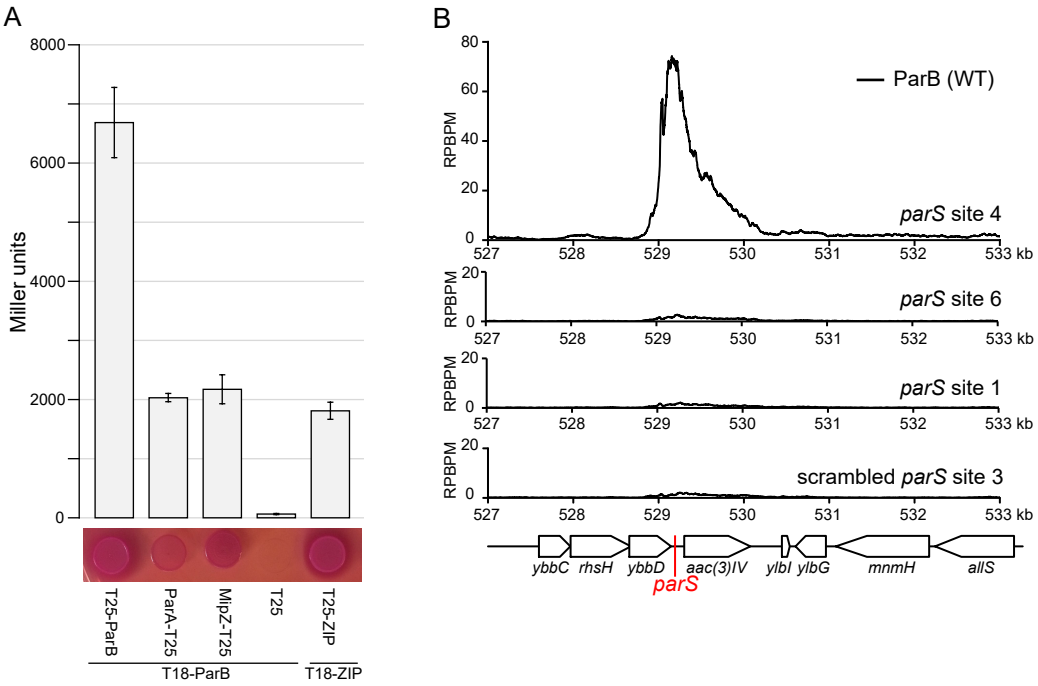

Figure S5. T18-ParB is functional in *Escherichia coli.* (A) A ParB protein was expressed from an IPTG-inducible promoter as a C-terminal fusion to the T18 fragment of *Bordetella pertussis* adenylate cyclase. Known interacting partners of ParB were expressed as fusion proteins to the T25 fragment of *B. pertussis*: T25-ParB, ParA-T25 and MipZ-T25. Interactions between ParB and partners were assessed on a solid MacConkey agar or by β-galactosidase assay. Three biological replicates were performed for each pair of interacting partners. A negative control (T25 fragment alone) and a positive control: T25-ZIP and T18-ZIP were also included. **(B)** ChIP-seq profiles of T18-ParB at *parS* site 4, site 6 and site 1 in an *E. coli* heterologous host. T18-ParB protein was expressed by addition of 500 µM IPTG for an hour before fixing with formadehyde for ChIP-seq. DNA bounds to T18-ParB was immunoprecipitated using α-T18 conjugated sepharose beads. A scrambled *parS* site 3 was also inserted at the *ybbD* locus to serve as a negative control. ChIP-seq signals were reported as the number of reads at every nucleotide along the genome (RPBPM value).

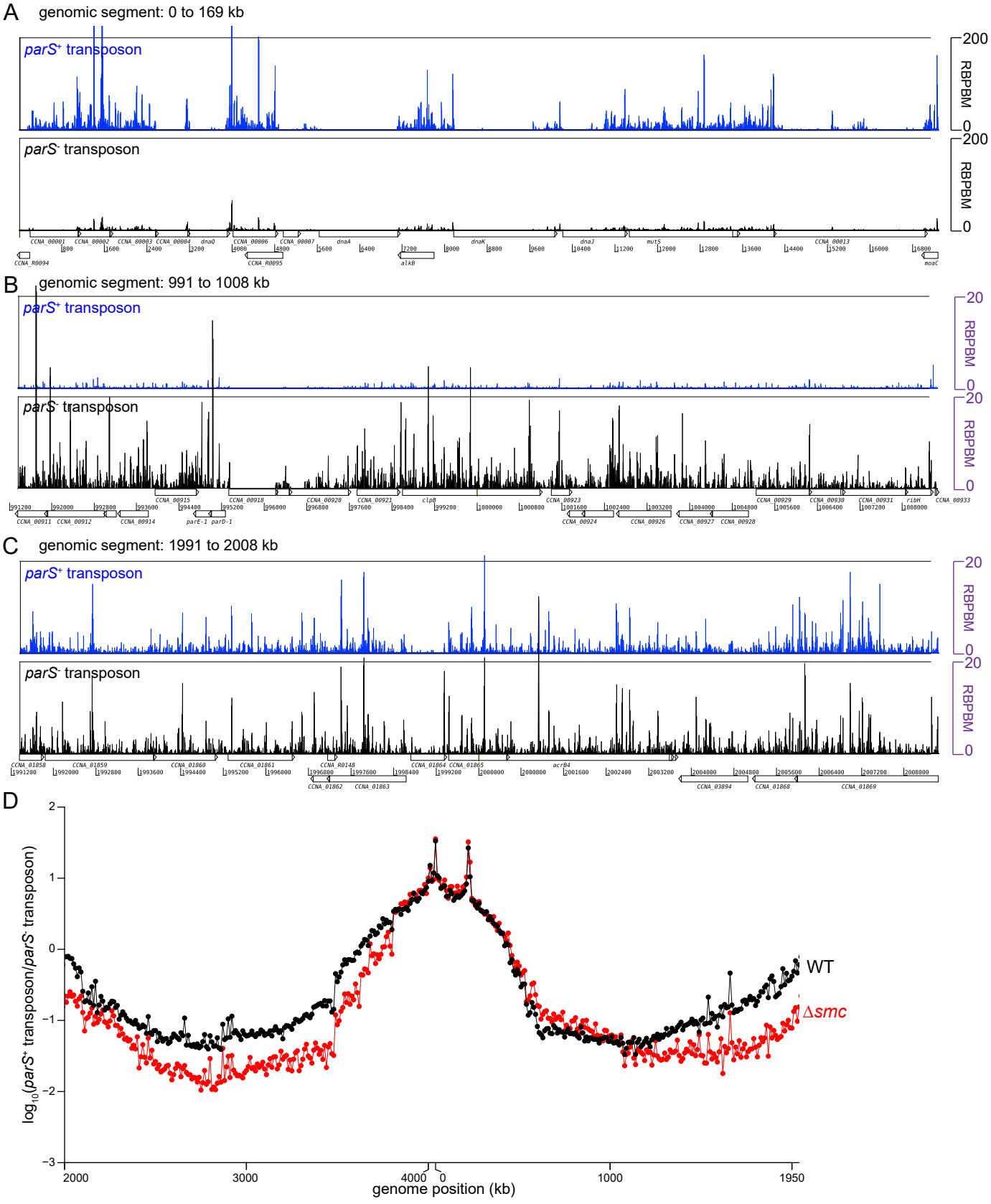

Figure S6. Tn5-seq reveals the positional bias of the centromeric *parS* site on *Caulobacter* chromosome. (A) A comparison of *parS*$^+$ (blue) and *parS*$^-$ (black) transposon insertions for the genomic segment between +0 kb and +168 kb (*ori* proximal). **(B)** A comparison of *parS*$^+$ (blue) and *parS*$^-$ (black) transposon insertions for the genomic segment between +9912 kb and +1008 kb (mid-arm). **(C)** A comparison of *parS*$^+$ (blue) and *parS*$^-$ (black) transposon insertions for the genomic segment between +1991 kb and +2008 kb (*ter* proximal). Axes labeled in purple color highlights different scaling. **(D)** Wild-type (black) or *Δsmc* (red) *Caulobacter* cells were mutagenized with *parS*$^+$ or *parS*$^-$ transposon, and the number of insertions was binned to 10-kb segments along the *Caulobacter* chromosome. The ratio between insertion frequency for *parS*$^+$ transposon and that of *parS*$^-$ transposon was calculated and plotted as a log10 scale against genomic position.