Structural and biochemical analyses of *Caulobacter crescentus* ParB reveal the role of its N-terminal domain in chromosome segregation

Adam S. B. Jalal[1], César L. Pastrana[2], Ngat T. Tran[1], Clare. E. Stevenson[3], David M. Lawson[3], Fernando Moreno-Herrero[2], and Tung B. K. Le[1*]

[1]Department of Molecular Microbiology
John Innes Centre, Norwich, NR4 7UH, United Kingdom
[2]Department of Macromolecular Structures
CSIC-Centro Nacional de Biotechnologia, Madrid, 28049, Spain
[3]Department of Biological Chemistry
John Innes Centre, Norwich, NR4 7UH, United Kingdom

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

ABSTRACT
The tripartite ParA-ParB-*parS* complex ensures faithful chromosome segregation in the majority of bacterial species. ParB nucleates on a centromere-like *parS* site and spreads to neighboring DNA to form a network of protein-DNA complexes. This nucleoprotein network interacts with ParA to partition the *parS* locus, hence the chromosome to each daughter cell. Here, we determine the co-crystal structure of a C-terminal domain truncated ParB-*parS* complex from *Caulobacter crescentus*, and show that its N-terminal domain adopts alternate conformations. The multiple conformations of the N-terminal domain might facilitate the spreading of ParB on the chromosome. Next, using ChIP-seq we show that ParBs from different bacterial species exhibit variation in their intrinsic capability for spreading, and that the N-terminal domain is a determinant of this variability. Finally, we show that the C-terminal domain of *Caulobacter* ParB possesses no or weak non-specific DNA-binding activity. Engineered ParB variants with enhanced non-specific DNA-binding activity condense DNA *in vitro* but do not spread further than wild-type *in vivo*. Taken all together, our results emphasize the role of the N-terminal domain in ParB spreading and faithful chromosome segregation in *Caulobacter crescentus*.

INTRODUCTION

Proper chromosome segregation is essential in all domains of life. In two-thirds of known bacterial species, faithful chromosome segregation is mediated by the conserved ParA-ParB-*parS* system (1–10). This tripartite complex consists of a Walker-box ATPase ParA, a centromere-binding protein ParB, and a centromere-like DNA sequence *parS*. The *parS* site is the first DNA locus to be segregated after chromosome replication (2, 5, 11, 12). ParB, a DNA-binding protein, nucleates on *parS* before binding to adjacent non-specific DNA to form a network of protein-DNA complexes. This nucleoprotein network interacts with ParA to partition the chromosome to each daughter cell. In *Caulobacter crescentus*, ParA forms a protein gradient emanating from the opposite pole of the cell to the ParB-*parS* complex (13–15). The DNA-bound ParB complexes stimulate the ATPase activity of ParA, causing the ParA gradient to retract, bringing the ParB-DNA complex to the opposite pole of the cell in a retreating gradient of ParA (15–19). In *Caulobacter crescentus*, ParA and ParB are essential for chromosome segregation and cell viability (3, 13). In other bacterial species, engineered strains lacking ParB are still viable but have elevated numbers of anucleate cells due to defects in chromosome segregation (2, 7, 9, 12, 16, 20–22).

The binding of multiple ParB molecules onto non-specific DNA after nucleation at *parS* (i.e. spreading) is a crucial event; bacterial cells harboring a nucleation-competent but spreading-defective *parB* allele are impaired in plasmid/chromosome segregation (23–25). Spreading was first discovered for the F-plasmid-encoded SopB protein and the P1 plasmid-encoded ParB protein (26, 27), and was subsequently found to be a general feature of many plasmid and chromosomal ParB proteins in bacteria (10, 23–25, 28, 29). Recently, spreading was also reported for a ParB-unrelated AspA protein that is responsible for plasmid segregation in an archaea *Sulfolobus* (30). In addition to a linear spreading on DNA, *Bacillus subtillis* ParB can also bridge distal DNA together to coalesce into a large nucleoprotein network in a process known as “spreading and bridging” (24, 31–33). Recently, a broadly similar “nucleation and caging” mechanism was also proposed to explain the ability of F1-plasmid and *Vibrio cholerae* chromosomal ParB to form a large ParB-DNA network (34, 35). In this model, the nucleation of ParB on *parS* creates a high local concentration of ParB, thereby caging ParB molecules together with non-specific DNA surrounding *parS* to create a loose but dynamic nucleoprotein network (34, 35).

Chromosomal ParB protein consists of an N-terminal domain (NTD), a middle *parS*-specific DNA-binding domain (DBD), and a C-terminal domain (CTD) (Fig. 1). The NTD and DBD are generally conserved among ParB orthologs, while the CTD exhibits high sequence variability, except for conserved leucine-zipper residues that are crucial for dimerization of ParB monomers (36, 37) (Fig. 1). A recent co-crystal structure of a CTD-truncated *Helicobacter pylori* ParB (Ct-ParB) with *parS* provided a possible structural basis for spreading (29). In this structure, four *Helicobacter* Ct-ParB monomers bind to four individual half *parS* sites and tetramerize to bring distal DNA closer together (29). In comparison to the structure of an apo- Ct-ParB from *Thermus thermophilus*, the NTD of *Helicobacter* Ct-ParB adopts a more open conformation to position the highly conserved arginine-rich patch (GERRxR, Fig. 1) outwards to mediate ParB-ParB oligomerization (29, 38). Based on this pairwise structural comparison, Chen *et al* (2015) proposed that the nucleation on *parS* induces a transition (at the NTD of ParB) from a spreading-incompetent closed conformation to a spreading-competent open conformation (29). In addition to the NTD, the CTD of *Bacillus* ParB was also shown to contribute to the formation of the nucleoprotein network via a positively charged lysine-rich surface that binds and condenses DNA non-specifically (36, 39). However, the sequence of the CTD diverges more rapidly than other domains of ParB so it is not yet clear whether ParBs from other bacterial species also possess a functionally equivalent CTD with non-specific DNA-binding and condensation activities.

Here, we determine the co-crystal structure of a Ct-ParB-*parS* complex from *Caulobacter crescentus* to better understand the function of this protein family. By comparing our co-crystal structure to that of apo- *Thermus* Ct-ParB and *Helicobacter* Ct-ParB-*parS* complex, we show that the NTD can adopt alternate conformations. Using ChIP-seq, we then show that ParBs from different bacterial species

exhibit variation in their intrinsic capability for spreading. We discover “maxi-spreaders” (e.g. ParB from *Moorella thermoacetica*) that spread over ~50 kb, while “mini-spreaders” (e.g. ParB from *Caulobacter crescentus*) spread only ~5 kb from a single *parS* site. We construct a series of chimeric proteins and find that the NTD is a determinant for the inter-species variation in spreading, at least in the case of *Caulobacter* and *Moorella* ParBs. In addition, we show that the CTD of *Caulobacter* ParB does not display non-specific DNA-binding and DNA condensation activities *in vitro*. Engineered *Caulobacter* ParB variants with an enhanced non-specific DNA-binding activity can condense DNA *in vitro* but do not spread further than wild-type protein *in vivo*. Overall, our results emphasize the key role of the NTD in ParB spreading in *Caulobacter* and highlights the inter-species variation that exists within the chromosomal ParB family.

RESULTS AND DISCUSSION

Co-crystal structure of the C-terminal domain truncated ParB-*parS* complex from *Caulobacter crescentus* revealed the multiple conformations of the NTD

We sought to determine a co-crystal structure of a ParB-*parS* complex from *Caulobacter crescentus*. After screening several constructs with different length of ParB and *parS*, we obtained crystals of a 50 amino acid C-terminally truncated ParB (Ct-ParB) in complex with a 22-bp *parS* duplex DNA (Fig. 1 and Fig. 2A). In solution, *Caulobacter* Ct-ParB also binds to *parS*, albeit weaker than a full-length protein (Fig. S1). Diffraction data for the *Caulobacter* Ct-ParB-*parS* complex were collected to a resolution of 2.9 Å, and the structure was solved by molecular replacement using the 3.1 Å structure of the *Helicobacter* Ct-ParB-*parS* complex and the 2.3 Å structure of apo- *Thermus* Ct-ParB as search templates. X-ray crystallographic data are summarized in Table 1.

The asymmetric unit of our co-crystal contains four copies of the Ct-ParB monomer and two copies of the full-size *parS* DNA (Fig. S2A). Each Ct-ParB monomer binds to a half *parS* site via the DNA-binding domain (Fig. 2A-B). Since chain A and B are very similar to chain C and D, respectively (RMSD= 1.59 Å, Fig. S2B), we used chain C-D-*parS* complex for subsequent analysis (Fig. 2A). Each Ct-ParB monomer consists of two domains: an N-terminal domain (NTD) (helices α1-α4 and sheets β1- β3) and a *parS* DNA-binding domain (DBD) (helices α5-α10) (Fig. 1 and Fig. 2B). We previously reported a 2.4 Å co-crystal structure of a *Caulobacter* ParB (DBD only) in complex with *parS* (40), here we discuss the structure of the NTD in depth. We observed that helices α3 and α4 of the NTD are packed against the DBD and are connected to the rest of the NTD via a loop in between α3 and β4 (Fig.1 and Fig. 2B). The rest of the NTD is comprised of a four-stranded β-sheet (β1-β4) and two surrounding helices (α1-α2) (Fig. 2B). The highly conserved arginine-rich patch (G^{101}ERRWR), crucial for *Caulobacter* ParB spreading (10), resides on helix α2 (Fig. 1). We observed that while the DBD and the NTD α3-α4 are near identical between chain C and D (RMSD=0.19 Å, Fig. 2C) the rest of the NTD (α1-β4) adopts completely different arrangement (Fig. 2C-D). The NTD (α1-β4) of chain C and D are oriented ~80° apart from each other (Fig. 2D); this is due to a loop (hereafter, called the elbow) that connects α3 and β4 together (Fig. 2C-D). The role of this elbow in orientating the NTD became clearer upon comparing the *Caulobacter* Ct-ParB-*parS* structure to two other available structures of chromosomal ParBs from *Helicobacter pylori* and *Thermus thermophilus*.

Structural comparisons of the *Caulobacter* Ct-ParB-*parS* complex to other ParB family members

In the co-crystal structure of the *Helicobacter* Ct-ParB-*parS* complex, ParB adopts an open conformation in which its NTD projects outwards to contact a nearby ParB monomer (Fig. 3A, Fig. S3A) (29). In contrast, no such interaction was seen between the NTD of the two adjacent *Caulobacter* ParB monomers (Fig. 3A). By superimposing the structure of *Helicobacter* Ct-ParB onto the *Caulobacter* one, we observed that each NTD has a different orientation (Fig. 3B-C). The *Helicobacter* ParB NTD extends outwards (an open conformation), while the *Caulobacter* ParB NTD points either inwards (chain D) or side-way (chain C) (a closed conformation) (Fig. 3B-C). Superimposition of three chains showed that the elbow (residues 121-125) swivels around the α3 axis, allowing the NTD to adopt three distinct conformations (Fig. 3C). Sequence alignment of ~1800

ParB orthologs showed an enrichment for charged and polar uncharged residues in the elbow region (Fig. 3C). This amino acid preference is typically found in intrinsically disordered proteins (41, 42) and might confer flexibility to the elbow region of ParB. A further structure superimposition showed that the NTD of an apo- *Thermus* Ct-ParB also adopts a closed conformation, most similar to chain D in the *Caulobacter* Ct-ParB-*parS* structure (Fig. S3C-D). Altogether, our three-way structural comparison suggests that the NTD can adopt multiple open or closed conformations regardless of whether ParB is on or off DNA. Our finding contrasts with Chen *et al* (2015) study which proposed that a *parS*-binding event induces a transition (at the NTD) from a spreading-incompetent closed conformation to a spreading-competent open conformation (29). While all currently available structures of chromosomal ParB lack the CTD, it is reasonable to assume that the NTD is also flexible in a full-length protein on/off DNA. Indeed, a co-crystal structure of a full-length SopB (a Type-I ParB protein for F-plasmid segregation) with DNA was previously solved, but only the density for the central DBD was observed (43). Schumacher *et al* (2010) showed that the absence of density for the NTD and CTD of SopB was due to their extreme flexibility rather than proteolysis during crystallization (43). The multiple orientations of the NTD of ParB (on/off DNA) might allow a dynamic ParB-DNA network to form inside cells.

We observed the second level of flexibility at the N-terminal-most peptide (residues 1-64) of ParB. This amino acid region is extended in the *Thermus* Ct-ParB structure (pink dashed line, Fig. 3D) but folds back in the *Caulobacter* Ct-ParB to contribute the fourth strand to the core β-sheet at the NTD (green dashed line, Fig. 3D). The equivalent region was not observed in the *Helicobacter* Ct-ParB structure. Due to the alternate conformations of this N-terminal-most peptide and of the NTD as a whole, the ParA-interacting region (residues 1-30, Fig. 1 and Fig. 3D) can potentially explore a very large space surrounding ParB. This flexibility might be beneficial for the network of ParB-DNA complexes to “fly-fishing” for ParA molecules *in vivo* (15–18).

The inter-species variation in spreading among ParB orthologs is dependent on the NTD

ParB orthologs are divergent in sequence, especially at their C-terminal domain (CTD) (Fig. 1). Therefore, we wondered whether ParBs from different bacterial species have distinct capacities for spreading and if it is dependent on the variable CTD. Exploiting the conservation of the *parS* sequence in bacteria (5), we constructed an *E. coli* heterologous system that allowed us to compare the spreading ability of ten chromosomal ParBs by ChIP-seq (Fig. 4A). *E. coli* does not possess a native ParA-ParB-*parS* system. We inserted a single *parS* site at the *ygcE* locus on the *E. coli* chromosome (Fig. 4A). Genes encoding N-terminally FLAG-tagged ParBs were codon optimized and expressed individually in *E. coli* (Fig. 4B). FLAG-ParBs were produced to a similar level (Fig. S4) before bound DNA was immunoprecipitated using an α-FLAG antibody and deep sequenced to reveal the extent of spreading from a single *parS* to the flanking DNA. From the ten ChIP-seq profiles, we observed that the majority of ParB (seven out of ten, including *Caulobacter* ParB) spread ~5 kb surrounding a single *parS* (Fig. 4B). On the other hand, three ParBs are “maxi-spreaders” that spread between ~20 kb to ~50 kb (Fig. 4B). In particular, *Moorella thermoacetica* ParB, despite its lower expression in *E. coli* (Fig. S4), spread ten times more extensively on DNA than *Caulobacter* ParB (Fig. 4B). Also, by considering only the shape of ChIP-seq profiles, we noted that ChIP signals reduced to the background more gradually for *Moorella* ParB than *Caulobacter* one. We also noted that *Bacillus* ParB spread only ~5 kb surrounding a single *parS;* this is more restrictive than a previous reported ~10 kb spreading distance when ChIP-chip of *Bacillus* ParB was performed in the native bacterium (25). The reason behind this discrepancy is unknown. Due to the caveat that ParB spreading has been taken out of the context of the native organism, the biological significance of the inter-species variation in ParB spreading is unclear. Nevertheless, we utilized this inter-species variation to determine the domain responsible for spreading. To do so, we constructed a series of chimeric proteins in which different regions of a “mini-spreader” *Caulobacter* ParB were replaced with the corresponding regions of a “maxi-spreader” *Moorella* ParB (Fig. 5). These chimeric proteins were produced to the same level in the *E. coli ygcE::parS* host (Fig. S4), and α-FLAG ChIP-seq experiments were performed to determine the extent of spreading (Fig. 5). Replacing a ParA-interacting region of *Caulobacter* ParB with the corresponding region from *Moorella* ParB produced

Chimera A that spread to the same extent as a wild-type *Caulobacter* ParB (Fig. 5). Similarly, Chimera B that had the CTD of *Caulobacter* ParB replaced by an equivalent region from *Moorella* ParB spread to the same extent as a wild-type *Caulobacter* ParB (Fig. 5). However, swapping the NTD (Chimera C) or both the NTD and the DBD (Chimera D) between *Caulobacter* ParB and *Moorella* ParB produced variants that are "maxi-spreaders" i.e. having a similar extensive spreading to the wild-type *Moorella* ParB (Fig. 5). Taken together, our ChIP-seq profiles suggested that the NTD, at least in the case of *Caulobacter* and *Moorella* ParBs, dictates their variation in spreading.

Engineering a lysine-rich surface into the *Caulobacter* ParB CTD resulted in variants with non-specific DNA-binding and condensation activities *in vitro*

In addition to the NTD, the CTD of ParB from *Bacillus subtilis* also contributes to the formation of the ParB-DNA network (32, 36, 39). *Bacillus* ParB was reported to bind non-specific DNA to condense both *parS* and non-*parS* DNA *in vitro*; these activities were mediated by a positively charged lysine-rich surface on the CTD (36, 39). Whether the non-specific DNA-binding activity of the *Bacillus* ParB CTD is a shared feature among ParB orthologs is currently unknown, furthermore, the relationship between the *in vitro* DNA condensation and the *in vivo* spreading is not fully understood. To better understand this relationship, we sought to generate variants of *Caulobacter* ParB with an enhanced non-specific DNA-binding activity. Unlike *Bacillus* ParB, the CTD of *Caulobacter* ParB lacks a lysine-rich patch (Fig. 6A), and the wild-type protein does not bind or binds very weakly to non-specific DNA *in vitro* (10) (Fig. 6B). To engineer a non-specific DNA-binding activity into *Caulobacter* ParB, we introduced additional lysine residues into its CTD. We systematically introduced a single (1K), double (2K), triple (3K), quadruple (4K), and quintuple (5K) lysine substitutions from the *Bacillus* ParB CTD into equivalent positions on the CTD of *Caulobacter* ParB (Fig. 6A). Ten variants were purified to homogeneity (Fig. S5A) and analyzed by a quantitative bio-layer interferometry assay that directly assessed their binding to a *parS* and a scrambled *parS* DNA (i.e. non-specific DNA) (Fig. 6B). All ten tested ParB variants retained their binding activities to *parS* (Fig 6B). We were unable to detect any noticeable non-specific DNA-binding activity for the 1K and 2K variants (Fig. 6B). However, a further introduction of lysine residues created 3K, 4K, and 5K variants that interacted with non-specific DNA similarly to that of *Bacillus* ParB (Fig. 6B).

The non-specific DNA-binding property of *Bacillus* ParB CTD was previously shown to condense DNA *in vitro* by magnetic tweezer assay (32, 36, 39) (Fig. 7A). To test if the engineered non-specific DNA-binding activity of *Caulobacter* ParB (3K-5K) variants (Fig. 6B) also leads to DNA condensation *in vitro*, we performed magnetic tweezer experiments on these variants and compared their activities to that of wild-type *Caulobacter* and *Bacillus* ParBs (Fig. 7B). We performed experiments at 1 µM concentration of proteins and different forces using an identical setup and conditions described in experiments with *Bacillus* ParB (32, 36). The extension of a tethered DNA was tracked, and any observation of a decrease in extension that was substantially larger than by the applied force alone is an indication of DNA condensation (Fig. 7A). *Bacillus* ParB condensed both non-*parS* DNA (Fig. 7B) and *parS* DNA (32, 36). On the contrary, *Caulobacter* ParB (WT) did not display any noticeable *in vitro* DNA condensation activity with either *parS* or non-*parS* DNA substrate at the tested concentration (Fig. 7B). However, upon incubating the 3K, 4K, or 5K ParB variants with tethered DNA resulted in a decrease in the DNA extension that was much greater than that attributable to the decrease in force alone (Fig. 7B). These results indicated that introduction of three to five lysine residues to the *Caulobacter* ParB CTD resulted in DNA condensation *in vitro*.

Caulobacter* ParB variants with an *in vitro* DNA condensation activity did not spread more extensively *in vivo

We then wondered whether the enhanced DNA-condensation activity of the ParB (3K-5K) variants leads to an increase in spreading *in vivo*. To test this, we performed α-FLAG ChIP-seq experiments on *Caulobacter* strains expressing individual FLAG-tagged ParB variant in a ParB (WT)-depletable background (44). *Caulobacter* cells that were completely depleted of a native ParB (WT) while producing the FLAG-tagged ParB (3K-5K) variants were viable (Fig. S5B), suggesting that the additional lysine residues at the CTD did not impair chromosome segregation in *Caulobacter*. As

controls for ChIP-seq experiments, strains expressing FLAG-tagged versions of ParB (WT), a non-spreading FLAG-ParB (R104A) mutant (10), and a non-DNA-binding protein FLAG-YFP were included (Fig. 7C). Consistent with the previous report (10), the ChIP-seq profile of a FLAG-ParB (WT) showed a clear enrichment above the background in the ~10 kb region from 4030 to 4040 kb on the chromosome (Fig. 7C). The extensive ChIP-seq profile is consistent with ParB (WT) spreading on the chromosome *in vivo*. This contrasts with the ChIP-seq profile of a non-spreader FLAG-ParB (R104A) in which the enrichment was confined to just ~500 bp immediately surrounding *parS* sites (10) (Fig. 7C). The profiles of the FLAG-ParB (3K-5K) variants were less extended than the FLAG-ParB (WT). We also noted that the overall heights of the ChIP-seq profiles of ParB (3K-5K) are lower than that of ParB (WT). It is possible that ParB (3K-5K) might bind DNA non-specifically along the chromosome, thereby titrating ParB molecules away from the *parS* cluster, resulting in a lower concentration of DNA-bound ParB near *parS*. Another possibility is ParB (3K-5K) are defective at the *parS* nucleation step, however this scenario is less likely since ParB (3K-5K) retained their *parS*-binding activities *in vitro* (Fig. 6B) and expressed to a comparable level to wild-type protein *in vivo* (Fig. S5).

At first, we were surprised to find that *Caulobacter* ParB has no or a very weak non-specific DNA-binding activity *in vitro* since both current models for ParB-DNA network formation ("spreading and bridging" and "nucleation and caging") require some degree of interaction between ParB and non-specific DNA (31–34, 36). However, while we did not observe non-specific DNA-binding activity at 1 µM *Caulobacter* ParB *in vitro*, the local concentration of ParB near *parS* has been estimated to reach ~500 µM (five times higher than in a crystallization drop) inside *Caulobacter* cells (15). At this extreme concentration, it is entirely possible that the central DNA-binding domain can provide some non-specific DNA-binding activity (Fig 8). We also noted that the five strongest *Caulobacter parS* sites cluster more closely (within a ~5 kb DNA segment (10)), while the four strongest *Bacillus parS* sites are dispersed within a ~57 kb region on the chromosome (30). A closer clustering of tightly bound ParB-*parS* complexes might be more effective in increasing the local concentration of *Caulobacter* ParB, despite its much weaker non-specific DNA-binding activity. On the contrary, the concentration of ParB in *Bacillus* cells is lower than in *Caulobacter* (~140 dimers compared to ~360 dimers per origin of replication (15)) and *parS* sites are more dispersed in this bacterium, in this case an added non-specific DNA-binding activity might enhance the formation of a ParB-DNA network in *Bacillus*.

Final perspectives

In this study, we characterize *Caulobacter* ParB biochemically and structurally to compare to orthologous proteins from different bacterial species. The availability of the *Caulobacter* Ct-ParB-*parS* structure, together with the structures of apo- *Thermus* Ct-ParB and *Helicobacter* Ct-ParB-*parS*, allows us to propose that the NTD can adopt multiple alternate conformations with respect to the DBD regardless of whether ParB is on/off DNA. The multiple conformations of the NTD might be beneficial in promoting the formation of a loose but dynamic ParB-DNA network. This is consistent with both "spreading and bridging" and "nucleation and caging" models. We further show that the NTD, at least in *Caulobacter* and *Moorella* ParB, determines how far the protein spreads on the chromosome from a single *parS* site. Our results emphasize the key role of the NTD in the formation of the ParB-DNA network in *Caulobacter* cells (Fig. 8). Our co-crystal structure lacks the CTD, hence the role of this domain is less clear in *Caulobacter*. In *Bacillus* ParB, the CTD acts both as a dimerization and DNA-binding and bridging interface via its non-specific DNA binding and condensation activities, thereby contributing to the formation of the nucleoprotein network. Here, we discover that the *Caulobacter* ParB CTD lacks these activities, *Caulobacter* CTD might mainly function as a monomer-monomer dimerization interface (37) (Fig. 8B). Taken all together, we suggest that different bacteria might fine-tune the properties of their chromosomal ParBs and there is a noticeable inter-species variation in how each domain contributes to the optimal function of ParB inside cells.

ACCESSION NUMBER

The accession number for the sequencing data reported in this paper is GSE134665. Atomic coordinates for a protein crystal structure reported in this paper were deposited in the RCSB Protein Data Bank with the accession number 6T1F.

ACKNOWLEDGMENTS

This study was supported by a Royal Society University Research Fellowship (UF140053) and a BBSRC grant (BB/P018165/1) to T.B.K.L. A.S.B.J's PhD studentship was funded by the Royal Society (RG150448), and N.T.T was funded by the BBSRC grant-in-add (BBS/E/J/000C0683 to the John Innes Centre). We thank Diamond Light Source for access to beamline I04-1 under proposal MX13467 with support from the European Community's Seventh Framework Program (FP7/2007–2013) under Grant Agreement 283570 (BioStruct-X). F.M.H. thanks the financial support from the Spanish MINECO (project BFU2017-83794-P (AEI/FEDER, UE)) and from European Research Council (ERC) under the European Union Horizon 2020 research and innovation (grant agreement No 681299).

Conflict of interest statement. None declared.

TABLE 1. X-RAY DATA COLLECTION AND PROCESSING STATISTICS

Structure	***Caulobacter* Ct-ParB-*parS* complex**
Data collection	
Diamond Light Source beamline	I04-1
Wavelength (Å)	0.916
Detector	Pilatus 6M-F
Resolution range (Å)	72.96 – 2.90 (3.08 – 2.90)
Space Group	$P2_1$
Cell parameters (Å/°)	a = 54.25, b = 172.93, c = 72.85 Å, β = 90.54°
Total no. of measured intensities	198135 (33888)
Unique reflections	29654 (4775)
Multiplicity	6.7 (7.1)
Mean $I/\sigma(I)$	8.7 (1.4)
Completeness (%)	99.7 (100.0)
R_{merge}[a]	0.135 (1.526)
R_{meas}[b]	0.146 (1.646)
$CC_{½}$[c]	0.997 (0.677)
Wilson B value (Å^2)	81.6
Refinement	
Resolution range (Å)	72.96 – 2.90 (2.98 – 2.90)
Reflections: working/free[d]	28155/1466
R_{work}[e]	0.240 (0.366)

R_{free}[e]	0.263 (0.369)
Ramachandran plot: favored/allowed/disallowed[f] (%)	95.2/4.8/0
R.m.s. bond distance deviation (Å)	0.005
R.m.s. bond angle deviation (°)	1.045
No. of protein residues per chain	191/190/195/187
No. of DNA bases per chain	22/22/22/22
Mean *B* factors: protein/DNA/ overall (Å^2)	98/74/92
PDB accession code	**6T1F**

Values in parentheses are for the outer resolution shell.

[a] $R_{merge} = \sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl} \sum_i I_i(hkl)$.

[b] $R_{meas} = \sum_{hkl} [N/(N-1)]^{1/2} \times \sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl} \sum_i I_i(hkl)$, where $I_i(hkl)$ is the *i*th observation of reflection *hkl*, $\langle I(hkl)\rangle$ is the weighted average intensity for all observations *i* of reflection *hkl* and *N* is the number of observations of reflection *hkl*.

[c] $CC_{½}$ is the correlation coefficient between symmetry equivalent intensities from random halves of the dataset.

[d] The dataset was split into "working" and "free" sets consisting of 95 and 5% of the data respectively. The free set was not used for refinement.

[e] The R-factors R_{work} and R_{free} are calculated as follows: $R = \sum(|F_{obs} - F_{calc}|)/\sum|F_{obs}|$, where F_{obs} and F_{calc} are the observed and calculated structure factor amplitudes, respectively.

[f] As calculated using MolProbity (45).

MATERIALS AND METHODS

Strains, media and growth conditions

Escherichia coli and *Caulobacter crescentus* were grown in LB and PYE, respectively. When appropriate, media were supplemented with antibiotics at the following concentrations (liquid/solid media for *Caulobacter*; liquid/solid media for *E. coli* (µg/mL)): carbenicillin (*E. coli* only: 50/100), chloramphenicol (1/2; 20/30), kanamycin (5/25; 30/50), oxytetracycline (1/2; 12/12), and apramycin (*E. coli* only: 25/50).

Plasmids and strains construction

All strains used are listed in Supplementary Table S1. All plasmids and primers used in strain and plasmid construction are listed in Supplementary Table S2. For details on plasmids and strain constructions, see the Supplementary Materials and Methods.

Protein overexpression and purification

Plasmid pET21b::*Caulobacter crescentus* Ct-ParB-(His)$_6$ (Table S1) was introduced into *E. coli* Rosetta pRARE competent cells (Novagen) by heat-shock transformation. 10 mL overnight culture was used to inoculate 4 L of LB medium + carbenicillin + chloramphenicol. Cells were grown at 37˚C with shaking at 210 rpm to an OD_{600} of ~0.4. The culture was then left to cool to 28˚C before isopropyl-β-D-thiogalactopyranoside (IPTG) was added at a final concentration of 1 mM. The culture was left shaking for an additional 3 hours at 30ºC before cells were harvested by centrifugation. Pelleted cells

were resuspended in a buffer containing 100 mM Tris-HCl pH 8.0, 300 mM NaCl, 10 mM Imidazole, 5% (v/v) glycerol, 1 µL of Benzonase nuclease (Sigma Aldrich), 0.1 g of lysozyme (Sigma Aldrich), and an EDTA-free protease inhibitor tablet (Roche). The pelleted cells were then lyzed by sonification (10 cycles of 15 s with 10 s resting on ice in between each cycle). The cell debris was removed through centrifugation at 28,000 g for 30 min and the supernatant was filtered through a 0.45 µm sterile filter (Sartorius Stedim). The protein was then loaded into a 1-mL HiTrap column (GE Healthcare) that had been equilibrated with buffer A (100 mM Tris-HCl pH 8.0, 300 mM NaCl, 10 mM Imidazole, and 5% glycerol). Protein was eluted from the column using an increasing (10 mM to 500 mM) Imidazole gradient in the same buffer. Ct-ParB-containing fractions were pooled and diluted to a conductivity of 16 mS/cm before being loaded onto a Heparin HP column (GE Healthcare) that had been equilibrated with 100 mM Tris-HCl pH 8.0, 25 mM NaCl, and 5% glycerol. Protein was eluted from the Heparin column using an increasing (25 mM to 1 M NaCl) salt gradient in the same buffer. Ct-ParB fractions were pooled and analyzed for purity by SDS-PAGE. Glycerol was then added to ParB fractions to a final volume of 10%, followed by 10 mM EDTA and 1 mM DDT. The purified Ct-ParB was subsequently aliquoted, snap frozen in liquid nitrogen, and stored at -80˚C. Ct-ParB that was used for X-ray crystallography was further polished via a gel-filtration column. To do so, purified Ct-ParB was concentrated by centrifugation in an Amicon Ultra-15 3-kDa cut-off spin filters (Merck) before being loaded into a Superdex 200 gel filtration column (GE Healthcare). The gel filtration column was pre-equilibrated with 10 mM Tris-HCl pH 8.0, 250 mM NaCl. Ct-ParB fractions were then pooled and analyzed for purity by SDS-PAGE (Fig. S1A). Full-length *Caulobacter* ParB-$(His)_6$ and other ParB variants were also purified using the same procedure (Fig. S1A and Fig. S5A).

Reconstitution of *parS* DNA

A 22-bp palindromic single-stranded DNA fragment (5'-GGA<u>TGTTTCACGTGAAACA</u>TCC-3') (360 µM in 10 mM Tris-HCl pH 8.0, 250 mM NaCl buffer) was heated at 95°C for 5 min before being left to cool down to room temperature (RT) overnight to form a double stranded *parS* DNA (final concentration: 180 µM). The core *parS* site sequence is underlined.

Protein crystallization, structure determination, and refinement

Crystallization screens for the *Caulobacter* Ct-ParB-*parS* complex crystal were set up in sitting-drop vapour diffusion format in MRC2 96-well crystallization plates with drops comprised of 0.3 µL precipitant solution and 0.3 µL of protein-DNA complex, and incubated at 293 K. His-tagged Ct-ParB (10 mg/mL) was mixed with a 22-bp *parS* duplex DNA at a molar ratio of 2:1.2 (protein monomer:DNA) in the gel filtration elution buffer (10 mM Tris-HCl pH 8.0, 250 mM NaCl). The Ct-ParB-*parS* crystals grew in a solution containing 20.5% (w/v) PEG 3350, 263 mM magnesium formate, and 10% (v/v) glycerol. After optimization of an initial hit, suitable crystals were cryoprotected with 20% (v/v) glycerol and mounted in Litholoops (Molecular Dimensions) before flash-cooling by plunging into liquid nitrogen. X-ray data were recorded on beamline I04-1 at the Diamond Light Source (Oxfordshire, UK) using a Pilatus 6M-F hybrid photon counting detector (Dectris), with crystals maintained at 100 K by a Cryojet cryocooler (Oxford Instruments). Diffraction data were integrated and scaled using XDS (46) via the XIA2 expert system (47) then merged using AIMLESS (48). Data collection statistics are summarized in Table 1. The majority of the downstream analysis was performed through the CCP4i2 graphical user interface (49).

The Ct-ParB-*parS* complex crystallized in space group $P2_1$ with cell parameters of a = 54.25, b = 172.93, c = 72.85 Å and β = 90.54° (Table 1). Analysis of the likely composition of the asymmetric unit (ASU) suggested that it would contain four copies of the Ct-ParB monomers and two copies of the 22-bp *parS* DNA duplex (Fig. S2B), giving an estimated solvent content of ~46.6%.

Interrogation of the Protein Data Bank with the sequence of the *Caulobacter* Ct-ParB revealed two suitable template structures for molecular replacement: apo-ParB from *Thermus thermophilus* (38) (PDB accession code: 1VZ0; 46% identity over 82% of the sequence) and *Helicobacter pylori* ParB bound to *parS* DNA (29) (PDB accession code: 4UMK; 42% identity over 75% of the sequence).

First, single subunits taken from these two entries were trimmed using SCULPTOR (50) to retain the parts of the structure that aligned with the *Caulobacter* Ct-ParB sequence, and then all side chains were truncated to Cβ atoms using CHAINSAW (51). Comparison of these templates revealed a completely different relationship between the N-terminal domain and the DNA-binding domain. Thus, we prepared search templates based on the individual domains rather than the subunits. The pairs of templates for each domain were then aligned and used as ensemble search models in PHASER (52). For the DNA component, an ideal B-form DNA duplex was generated in COOT (53) from a 22-bp palindromic sequence of *parS*. A variety of protocols were attempted in PHASER (52), the best result was obtained by searching for the two DNA duplexes first, followed by four copies of the DNA-binding domain, giving a TFZ score of 10.5 at 4.5 Å resolution. We found that the placement of the DNA-binding domains with respect to the DNA duplexes was analogous to that seen in the *Helicobacter* Ct-ParB-*parS* complex. After several iterations of rebuilding in COOT and refining the model in REFMAC5 (54), it was possible to manually dock one copy of the N-terminal domain template (from 1VZ0) into weak and fragmented electron density such that it could be joined to one of the DNA-binding domains. A superposition of this more complete subunit onto the other three copies revealed that in only one of these did the N-terminal domain agree with the electron density. Inspection of the remaining unfilled electron density showed evidence for the last two missing N-terminal domains, which were also added by manual docking of the domain template (from 1VZ0). For the final stages, TLS refinement was used with a single TLS domain defined for each protein chain and for each DNA strand. The statistics of the final refined model, including validation output from MolProbity (45), are summarized in Table 1.

Chromatin immunoprecipitation with deep sequencing (ChIP-seq)

Caulobacter cell cultures (25 mL) were grown in PYE (in the presence of appropriate antibiotics, 0.3% glucose, and 0.5 mM vanillate) before fixation with formaldehyde to a final concentration of 1%. Fixed cells were incubated at RT for 30 minutes, then quenched with 0.125 M glycine for 15 minutes at RT. All subsequent steps were performed exactly as described in Tran *et al* (2018) (10). A detailed protocol was described in the Supplementary Materials and Methods. For the list of ChIP-seq experiments and their replicates in this study, see Supplementary Table S3.

Generation and analysis of ChIP-seq profiles

For analysis of ChIP-seq data, Hiseq 2500 Illumina short reads (50 bp) were mapped back to the *Caulobacter* NA1000 reference genome (NCBI Reference Sequence: NC_011916.1) using Bowtie 1 (29) and the following command: bowtie -m 1 -n 1 --best --strata -p 4 --chunkmbs 512 NA1000-bowtie --sam *.fastq > output.sam. Subsequently, the sequencing coverage at each nucleotide position was computed using BEDTools (30) using the following command: bedtools genomecov -d -ibam output.sorted.bam -g NA1000.fna > coverage_output.txt. For analysis of *E. coli* ChIP-seq data, the same procedure as above was applied, except that short reads were map to the reference genome of the *E. coli* MG1655 (NCBI Reference Sequence: NC_000913.3). Finally, ChIP-seq profiles were plotted with the x-axis representing genomic positions and the y-axis is the number of reads per base pair per million mapped reads (RPBPM) using custom R scripts.

Measurement of protein-DNA binding affinity by bio-layer interferometry (BLI)

Bio-layer interferometry experiments were conducted using a BLItz system equipped with Dip-and-Read© Streptavidin (SA) Biosensors (ForteBio). BLItz measures the wavelength shifts (binding signal or response (R), unit: nm) resulting from changes in the optical thickness of the sensor surface during association or dissociation of the analyte. The streptavidin biosensor (ForteBio) was hydrated in a low salt binding buffer (100 mM Tris-HCl pH 7.4, 100 mM NaCl, 1 mM EDTA, and 0.005% Tween 20) for 10 min. Biotinylated double-stranded DNA was immobilized onto the surface of the SA biosensor through a cycle of baseline (30 s), association (120 s), and dissociation (120 s). Briefly, the tip of the biosensor was dipped into a low salt buffer for 30 s to establish the baseline, then to 1 μM biotinylated double-stranded DNA for 120 s, and finally to a low salt binding buffer for 120 s to allow for dissociation. Biotinylated double-stranded DNA harboring *parS* or a scrambled *parS* site (i.e. non-specific DNA) were prepared by annealing a 20-bp biotinylated oligo with its unmodified

complementary strand in an annealing buffer (1 mM Tris-HCl pH 8.0, 5 mM NaCl). The sequences of oligos are listed in Supplementary Table S2. The oligos mixture was heated to 98ºC for 2 min and allowed to cool down to RT overnight. After the immobilization of DNA on the sensor, association reactions were monitored at 250 nM, 500 nM, and 1 µM dimer concentration of ParB (WT) or ParB variants for 120 s. At the end of each binding step, the sensor was transferred into a protein-free low salt buffer to follow the dissociation kinetics for 120 s. The sensor was recycled by dipping in a high-salt buffer (100 mM Tris-HCl pH 7.4, 1000 mM NaCl, 1 mM EDTA, and 0.005% Tween 20) for at least 1 min to remove bound proteins. All interaction kinetics profiles (sensorgrams) recorded during BLItz experiments were analyzed using the BLItz analysis software (ForteBio). Reactions were run in triplicate for each concentration of ParB used, and the equilibrium responses were recorded and averaged. The extent of non-specific binding was assessed by monitoring the interaction of proteins with unmodified sensors and was deemed to be negligible.

Magnetic tweezer assays

Magnetic tweezer experiments were performed using a home-made setup as described previously (32, 36). Briefly, images of micro meter-sized superparamagnetic beads tethered to the surface of a glass slide by DNA constructs are acquired with a 100x oil immersion objective and a CCD camera. Real-time image analysis was used to determine the spatial coordinates of beads with nm accuracy in x, y and z. A step-by-step motor located above the sample moves a pair of magnets allowing the application of stretching forces to the bead-DNA system. We used horizontally-aligned magnets coupled to an iron holder. Applied forces can be quantified from the Brownian excursions of the bead and the extension of the DNA tether. Data were acquired at 150 Hz to minimize sampling artifacts in force determination. We used horizontally-aligned magnets coupled to an iron holder to achieve force up to 15 pN.

Fabrication of DNA substrates for magnetic tweezer experiments containing a single *parS* sequence with biotins and digoxigenins at the tails was described previously (32).The DNA molecules were incubated with 2.8 µm streptavidin-coated beads (MyOne, Invitrogen) for 10 min. Then, the DNA-bead complex was injected in a liquid cell functionalized with anti-digoxigenin antibodies (Roche) and incubated for 10 min before applying force. Torsionally constrained molecules and beads with more than a single DNA molecule were identified from its distinct rotation-extension curves and discarded for further analysis. All the experiments were performed in a reaction buffer composed of 10 mM HEPES pH 7.5, 150 mM NaCl, 3 mM EDTA, 0.1% (v/v) Tween-20 and 100 µg/ml BSA.

Force-extension curves were obtained by decreasing the applied force in steps from 15 pN to ~0.02 pN for a total measuring time of 15 min. First, we measured the force-extension response for bare DNA molecules. Then, the force was reset to 15 pN and ParB variants were flown and incubated for 2 min before starting the measurement of a force-extension curve at the same magnet positions in absence of proteins. The force applied to each bead was determined based on the force-extension data of bare DNA molecules. Bare DNA curves were fitted to the worm-like chain model and fitted values of persistence length and contour length were used as a quality control. Molecules with a large discrepancy for contour length or persistence with respect to expected parameters (45 nm persistence length, 2.1 µm contour length) were discarded from the analysis.

REFERENCES

1. 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.
2. Lin,D.C. and Grossman,A.D. (1998) Identification and characterization of a bacterial chromosome partitioning site. *Cell*, **92**, 675–685.
3. 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.
4. Fogel,M.A. and Waldor,M.K. (2006) A dynamic, mitotic-like mechanism for bacterial chromosome segregation. *Genes Dev.*, **20**, 3269–3282.
5. 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.
6. Donczew,M., Mackiewicz,P., Wróbel,A., Flärdh,K., Zakrzewska-Czerwińska,J. and Jakimowicz,D. (2016) ParA and ParB coordinate chromosome segregation with cell elongation and division during Streptomyces sporulation. *Open Biology*, **6**, 150263.
7. Kawalek,A., Bartosik,A.A., Glabski,K. and Jagura-Burdzy,G. (2018) Pseudomonas aeruginosa partitioning protein ParB acts as a nucleoid-associated protein binding to multiple copies of a parS-related motif. *Nucleic Acids Res.*, **46**, 4592–4606.
8. 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. *Molecular Microbiology*, **45**, 1365–1377.
9. Harms,A., Treuner-Lange,A., Schumacher,D. and Sogaard-Andersen,L. (2013) Tracking of chromosome and replisome dynamics in Myxococcus xanthus reveals a novel chromosome arrangement. *PLoS Genet*, **9**, e1003802.
10. Tran,N.T., Stevenson,C.E., Som,N.F., Thanapipatsiri,A., Jalal,A.S.B. and Le,T.B.K. (2018) Permissive zones for the centromere-binding protein ParB on the Caulobacter crescentus chromosome. *Nucleic Acids Res*, **46**, 1196–1209.
11. Toro,E., Hong,S.-H., McAdams,H.H. and Shapiro,L. (2008) Caulobacter requires a dedicated mechanism to initiate chromosome segregation. *PNAS*, **105**, 15435–15440.
12. Lagage,V., Boccard,F. and Vallet-Gely,I. (2016) Regional Control of Chromosome Segregation in Pseudomonas aeruginosa. *PLOS Genetics*, **12**, e1006428.
13. Mohl,D.A. and Gober,J.W. (1997) Cell cycle-dependent polar localization of chromosome partitioning proteins in Caulobacter crescentus. *Cell*, **88**, 675–684.
14. Ptacin,J.L., Lee,S.F., Garner,E.C., Toro,E., Eckart,M., Comolli,L.R., Moerner,W.E. and Shapiro,L. (2010) A spindle-like apparatus guides bacterial chromosome segregation. *Nat Cell Biol*, **12**, 791–798.
15. 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.
16. Vecchiarelli,A.G., Mizuuchi,K. and Funnell,B.E. (2012) Surfing biological surfaces: exploiting the nucleoid for partition and transport in bacteria. *Molecular Microbiology*, **86**, 513–523.
17. Vecchiarelli,A.G., Neuman,K.C. and Mizuuchi,K. (2014) A propagating ATPase gradient drives transport of surface-confined cellular cargo. *PNAS*, **111**, 4880–4885.
18. Hwang,L.C., Vecchiarelli,A.G., Han,Y.-W., Mizuuchi,M., Harada,Y., Funnell,B.E. and Mizuuchi,K. (2013) ParA-mediated plasmid partition driven by protein pattern self-organization. *The EMBO Journal*, **32**, 1238–1249.
19. 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.
20. 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.

21. 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.
22. Yu,W., Herbert,S., Graumann,P.L. and Götz,F. (2010) Contribution of SMC (Structural Maintenance of Chromosomes) and SpoIIIE to Chromosome Segregation in Staphylococci. *J Bacteriol*, **192**, 4067–4073.
23. Murray,H., Ferreira,H. and Errington,J. (2006) The bacterial chromosome segregation protein Spo0J spreads along DNA from parS nucleation sites. *Molecular Microbiology*, **61**, 1352–1361.
24. Song,D., Rodrigues,K., Graham,T.G.W. and Loparo,J.J. (2017) A network of cis and trans interactions is required for ParB spreading. *Nucleic Acids Res.*, **45**, 7106–7117.
25. Breier,A.M. and Grossman,A.D. (2007) Whole-genome analysis of the chromosome partitioning and sporulation protein Spo0J (ParB) reveals spreading and origin-distal sites on the Bacillus subtilis chromosome. *Molecular Microbiology*, **64**, 703–718.
26. Rodionov,O., Lobocka,M. and Yarmolinsky,M. (1999) Silencing of genes flanking the P1 plasmid centromere. *Science*, **283**, 546–549.
27. Lynch,A.S. and Wang,J.C. (1995) SopB protein-mediated silencing of genes linked to the sopC locus of Escherichia coli F plasmid. *PNAS*, **92**, 1896–1900.
28. Kusiak,M., Gapczyńska,A., Płochocka,D., Thomas,C.M. and Jagura-Burdzy,G. (2011) Binding and Spreading of ParB on DNA Determine Its Biological Function in Pseudomonas aeruginosa. *J. Bacteriol.*, **193**, 3342–3355.
29. Chen,B.-W., Lin,M.-H., Chu,C.-H., Hsu,C.-E. and Sun,Y.-J. (2015) Insights into ParB spreading from the complex structure of Spo0J and parS. *Proc. Natl. Acad. Sci. U.S.A.*, **112**, 6613–6618.
30. Schumacher,M.A., Tonthat,N.K., Lee,J., Rodriguez-Castañeda,F.A., Chinnam,N.B., Kalliomaa-Sanford,A.K., Ng,I.W., Barge,M.T., Shaw,P.L.R. and Barillà,D. (2015) Structures of archaeal DNA segregation machinery reveal bacterial and eukaryotic linkages. *Science*, **349**, 1120–1124.
31. 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.
32. 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.
33. Broedersz,C.P., Wang,X., Meir,Y., Loparo,J.J., Rudner,D.Z. and Wingreen,N.S. (2014) Condensation and localization of the partitioning protein ParB on the bacterial chromosome. *PNAS*, **111**, 8809–8814.
34. 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.
35. Debaugny,R.E., Sanchez,A., Rech,J., Labourdette,D., Dorignac,J., Geniet,F., Palmeri,J., Parmeggiani,A., Boudsocq,F., Anton Leberre,V., *et al.* (2018) A conserved mechanism drives partition complex assembly on bacterial chromosomes and plasmids. *Mol. Syst. Biol.*, **14**, e8516.
36. Fisher,G.L., Pastrana,C.L., Higman,V.A., Koh,A., Taylor,J.A., Butterer,A., Craggs,T., Sobott,F., Murray,H., Crump,M.P., *et al.* (2017) The structural basis for dynamic DNA binding and bridging interactions which condense the bacterial centromere. *Elife*, **6**.
37. 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.
38. 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.
39. Madariaga-Marcos,J., Pastrana,C.L., Fisher,G.L., Dillingham,M.S. and Moreno-Herrero,F. (2019) ParB dynamics and the critical role of the CTD in DNA condensation unveiled by combined force-fluorescence measurements. *Elife*, **8**.

40. Jalal,A.S.B., Tran,N.T., Stevenson,C.E., Tan,X., Lawson,D.M. and Le,T.B.K. (2019) Evolving a new protein-DNA interface via sequential introduction of permissive and specificity-switching mutations. *bioRxiv*, 10.1101/724823.
41. Romero,P., Obradovic,Z., Li,X., Garner,E.C., Brown,C.J. and Dunker,A.K. (2001) Sequence complexity of disordered protein. *Proteins*, **42**, 38–48.
42. Vucetic,S., Brown,C.J., Dunker,A.K. and Obradovic,Z. (2003) Flavors of protein disorder. *Proteins*, **52**, 573–584.
43. Schumacher,M.A., Piro,K.M. and Xu,W. (2010) Insight into F plasmid DNA segregation revealed by structures of SopB and SopB–DNA complexes. *Nucleic Acids Res*, **38**, 4514–4526.
44. Thanbichler,M. and Shapiro,L. (2006) MipZ, a Spatial Regulator Coordinating Chromosome Segregation with Cell Division in Caulobacter. *Cell*, **126**, 147–162.
45. Chen,V.B., Arendall,W.B., Headd,J.J., Keedy,D.A., Immormino,R.M., Kapral,G.J., Murray,L.W., Richardson,J.S. and Richardson,D.C. (2010) MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D Biol. Crystallogr.*, **66**, 12–21.
46. Kabsch,W. (2010) XDS. *Acta Crystallogr D Biol Crystallogr*, **66**, 125–132.
47. Winter,G. (2010) xia2: an expert system for macromolecular crystallography data reduction. *J Appl Cryst*, **43**, 186–190.
48. Evans,P.R. and Murshudov,G.N. (2013) How good are my data and what is the resolution? *Acta Cryst D*, **69**, 1204–1214.
49. Potterton,L., Agirre,J., Ballard,C., Cowtan,K., Dodson,E., Evans,P.R., Jenkins,H.T., Keegan,R., Krissinel,E., Stevenson,K., *et al.* (2018) CCP4i2: the new graphical user interface to the CCP4 program suite. *Acta Crystallogr D Struct Biol*, **74**, 68–84.
50. Bunkóczi,G. and Read,R.J. (2011) Improvement of molecular-replacement models with Sculptor. *Acta Crystallogr D Biol Crystallogr*, **67**, 303–312.
51. Stein,N. (2008) CHAINSAW: a program for mutating pdb files used as templates in molecular replacement. *J Appl Cryst*, **41**, 641–643.
52. McCoy,A.J., Grosse-Kunstleve,R.W., Adams,P.D., Winn,M.D., Storoni,L.C. and Read,R.J. (2007) Phaser crystallographic software. *J Appl Crystallogr*, **40**, 658–674.
53. Emsley,P. and Cowtan,K. (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.*, **60**, 2126–2132.
54. Murshudov,G.N., Vagin,A.A. and Dodson,E.J. (1997) Refinement of Macromolecular Structures by the Maximum-Likelihood Method. *Acta Cryst D*, **53**, 240–255.

FIG. 1

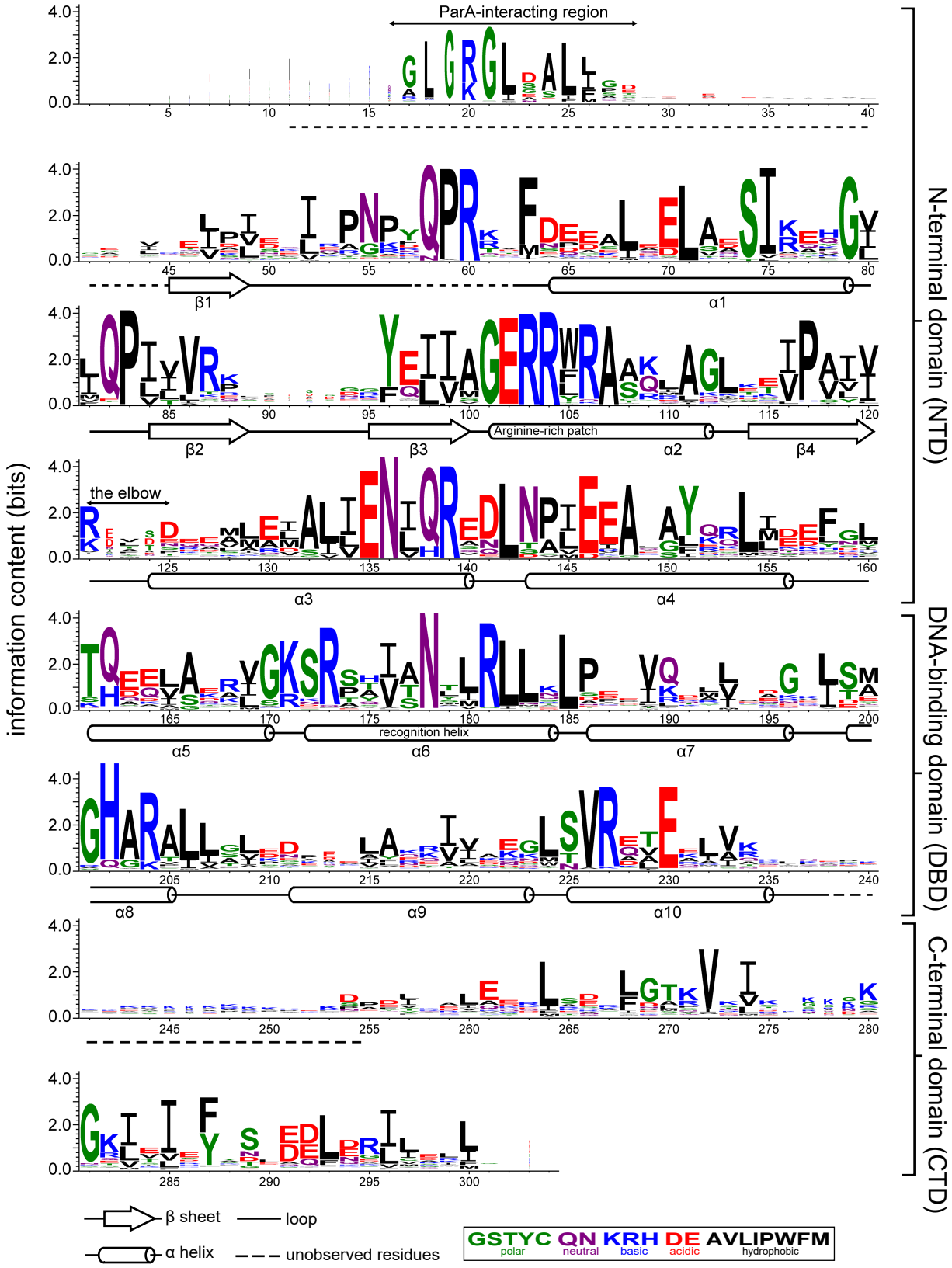

Figure 1. Sequence alignment of the chromosomal ParB protein family. An alignment of ~1800 chromosomal ParB proteins was constructed and presented as a sequence logo. The height of the stack indicates the sequence conservation, while the height of symbols within the stack indicates the relative frequency of each amino acid residue at that position. Amino acids are colored based on their chemical properties. Secondary-structure elements for *Caulobacter* Ct-ParB are shown below the alignment. Dashed lines indicate unmodelled residues due to poor electron density in the *Caulobacter* Ct-ParB-*parS* co-crystal structure.

FIG. 2

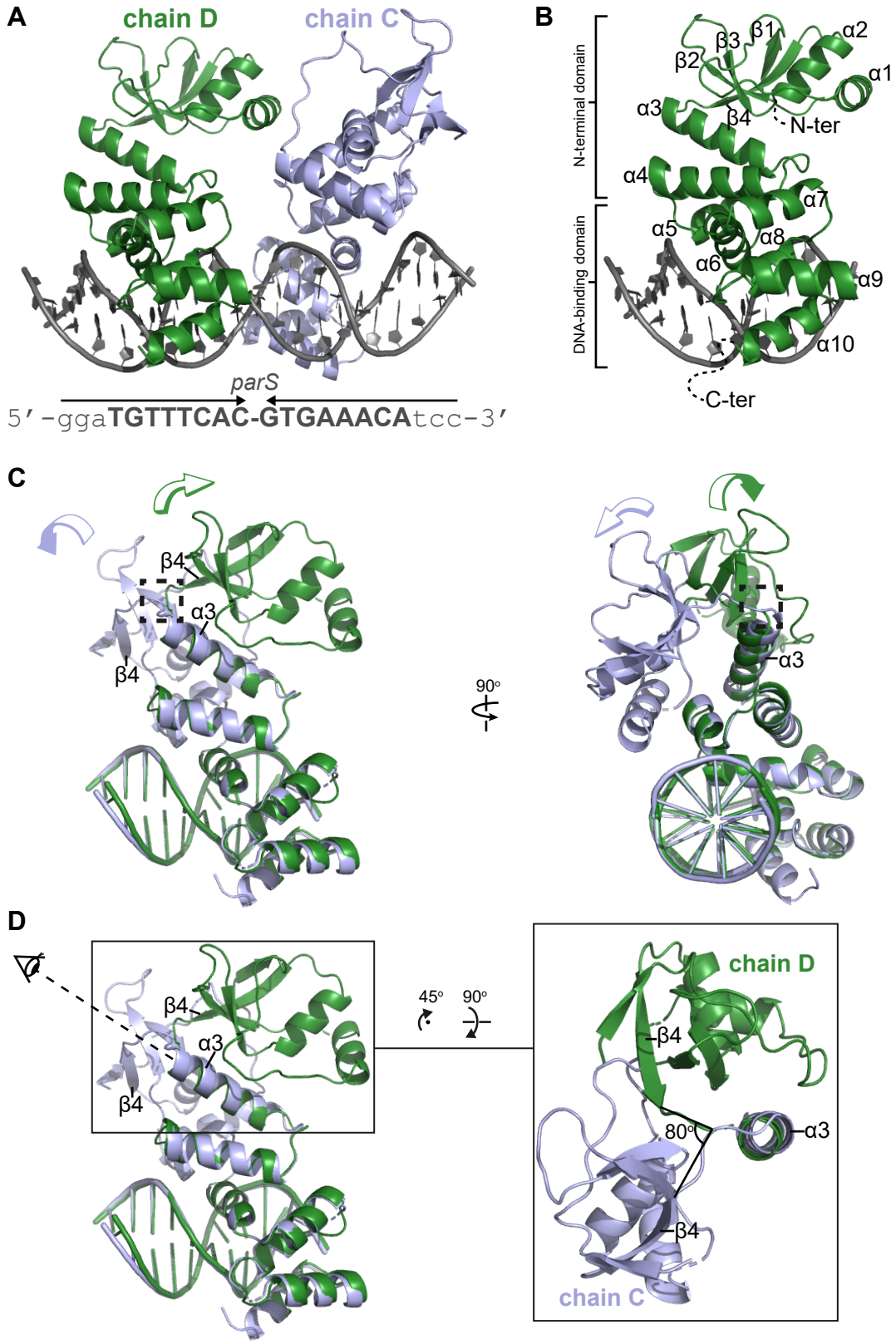

Figure 2. Co-crystal structure of a C-terminal domain truncated Ct-ParB-*parS* complex from *Caulobacter crescentus* reveals alternate conformations of the N-terminal domain. (A) Co-crystal structure of two *Caulobacter* Ct-ParB monomers (dark green and light blue) bound on a 22-bp *parS* DNA. The nucleotide sequence of the 22-bp *parS* is shown below the co-crystal structure, the core *parS* sequence is highlighted in bold. **(B)** The structure of Ct-ParB chain D bound to a *parS* half site with key features highlighted. **(C)** Superimposition of *Caulobacter* Ct-ParB chain C and D (light blue and dark green, respectively) shows two different orientations of the NTD (α1-β4). The arrow above each chain shows the direction each NTD projects towards. The dashed box shows the loop (the elbow) that connects helix α3 and sheet β4 together. **(D)** A top-down view of the superimposition of chain C and D shows that the NTDs (α1-β4) of chain C and D are oriented ~80° apart from each other.

FIG. 3

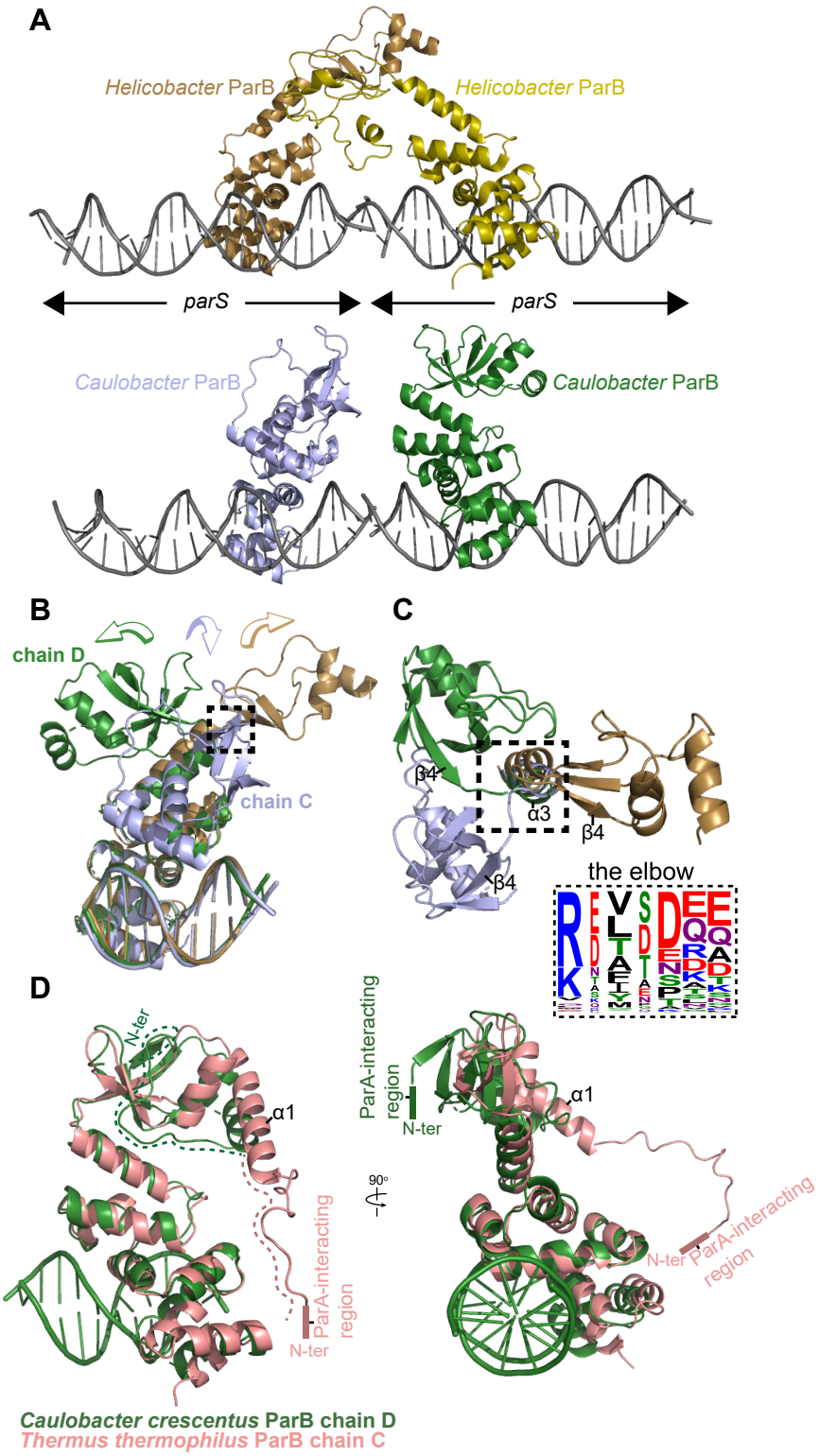

Figure 3. Structural comparisons of the *Caulobacter* Ct-ParB-*parS* complex to other ParB family members. (A) Structures of two adjacent symmetry complexes of *Helicobacter pylori* (upper panel) and *Caulobacter crescentus* Ct-ParB-*parS* (lower panel). In the *Helicobacter* Ct-ParB-*parS* complex, ParB adopts an open conformation in which its NTD projects outwards to contact a neighboring monomer. However, no such interaction was seen between the NTD of the two adjacent *Caulobacter* ParB monomers. **(B)** A side-view of the superimposition between *Caulobacter* Ct-ParB chain C (light blue), chain D (dark green), and *Helicobacter* Ct-ParB (golden) shows the three distinct orientations of the NTD (see also Fig. S3). **(C)** A top-view of the superimposition between *Caulobacter* Ct-ParB chain C, chain D, and *Helicobacter* Ct-ParB. The dashed box shows the loop (the elbow) that connects helix α3 and sheet β4 together. The conservation of amino acids at the flexible elbow is presented as sequence logos. Amino acids are colored based on their chemical properties (GSTYC: polar; QN: neutral; KRH: basic; DE: acidic; and AVLIPWFM: hydrophobic). **(D)** Superimposition of *Caulobacter* Ct-ParB chain D (dark green) and *Thermus* Ct-ParB chain C (pink) shows the two different conformations of the N-terminal-most peptide (dashed dark green and pink lines).

FIG. 4

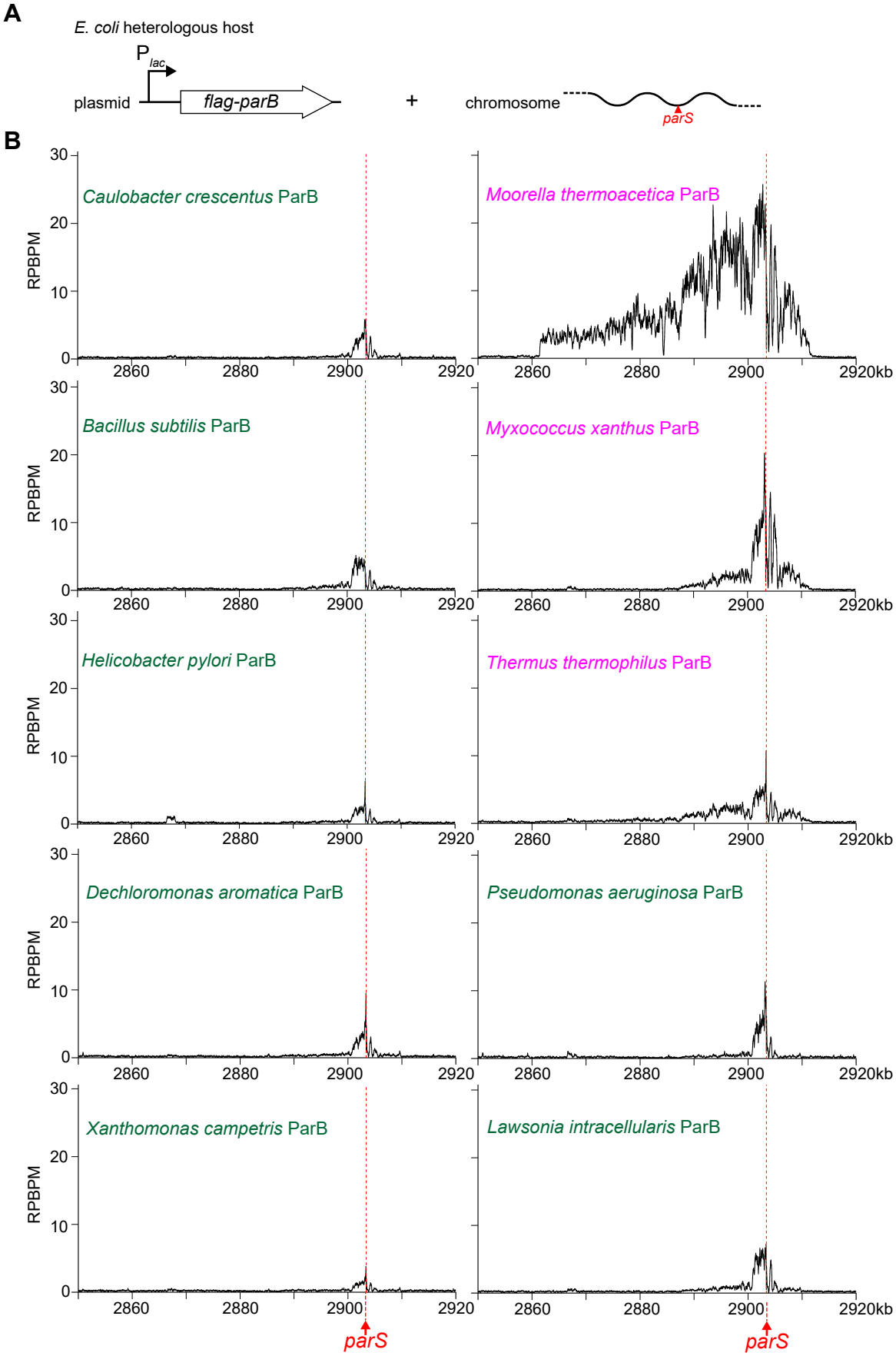

Figure 4. ChIP-seq revealed an inter-species variation in spreading among ParB orthologs. (A) The design of an *E. coli* heterologous host that allowed quantification of ParB spreading by ChIP-seq. A single *parS* site was inserted at the *ygcE* locus on the *E. coli* chromosome. Individual N-terminally FLAG-tagged ParB was produced from an IPTG-inducible P_{lac} promoter on a plasmid. **(B)** ChIP-seq profiles of FLAG-tagged ParBs from a collection of ten bacterial species (mini-spreader ParB: dark green; maxi-spreader ParB: magenta). ChIP-seq signals were reported as the number of reads at every nucleotide along the genome (RPBPM value).

FIG. 5

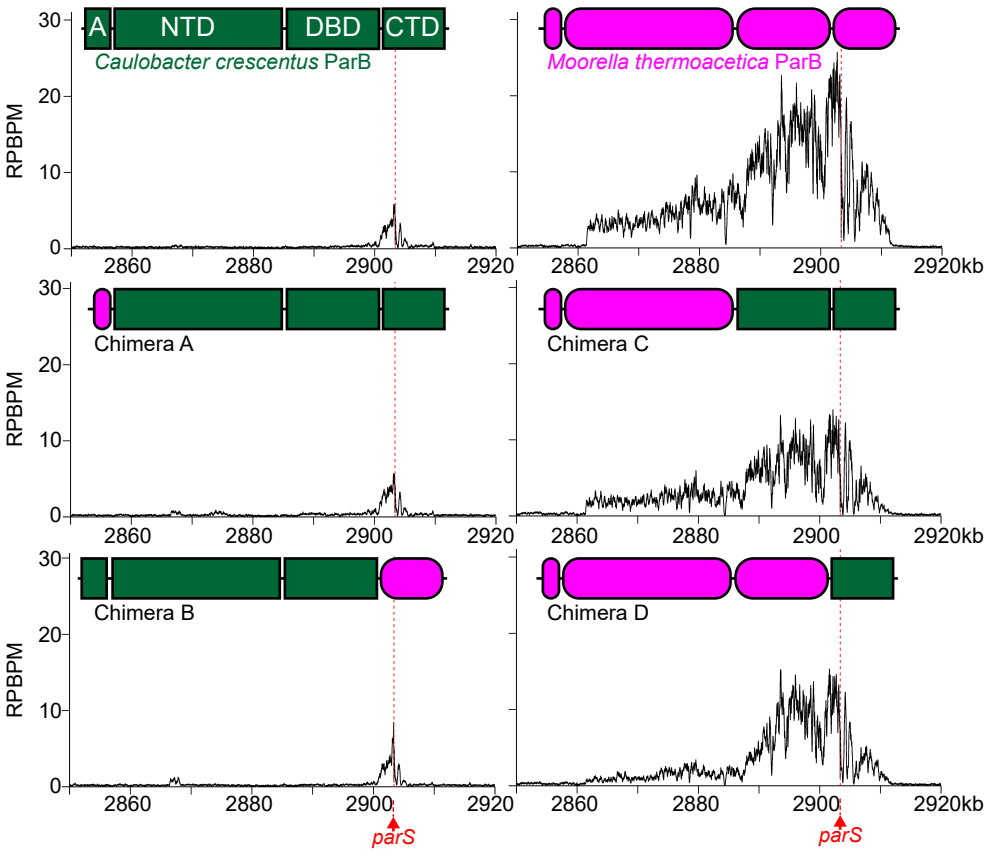

Figure 5. The N-terminal domain determines the inter-species variation in spreading among ParB orthologs. ChIP-seq profiles of a series of chimeric proteins in which different regions of a mini-spreader *Caulobacter* ParB (dark green) were replaced with the corresponding regions of a super-spreader *Moorella* ParB (magenta). N-terminally FLAG-tagged chimeras were expressed in an *E. coli ygcE*::*parS* heterologous host (See Fig. 4A). A: ParA-interacting region (res.1-32, *Caulobacter* ParB numbering); NTD: N-terminal domain (res.33-158); DBD: the *parS* DNA-binding domain (res.159-251); CTD: C-terminal domain (res. 252-end). ChIP-seq signals were reported as the number of reads at every nucleotide along the genome (RPBPM value).

FIG. 6

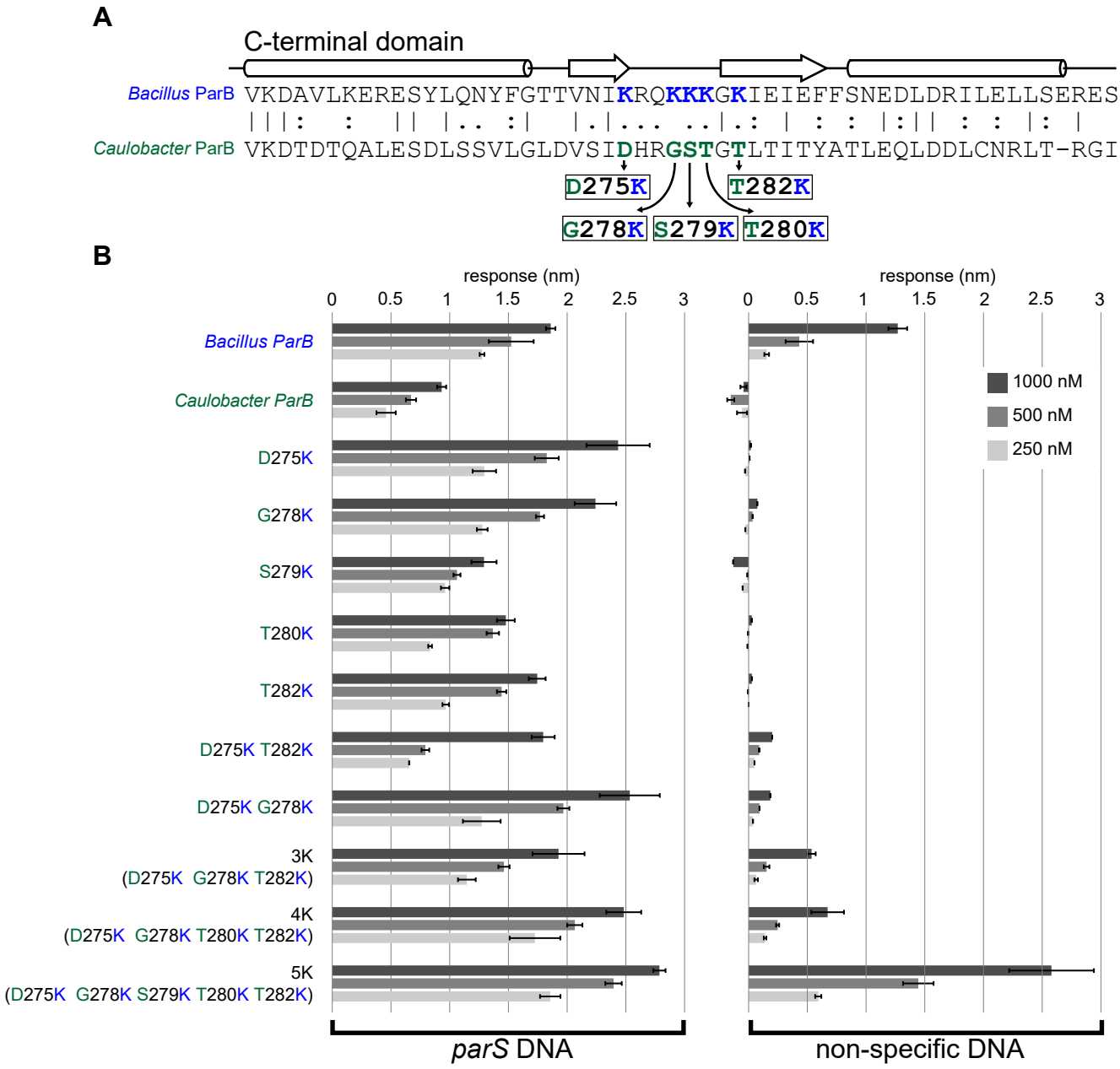

Figure 6. Engineering a lysine-rich surface into the *Caulobacter* ParB C-terminal domain resulted in variants with enhanced non-specific DNA-binding activity. (A) Sequence alignment between *Caulobacter* (dark green) and *Bacillus* ParB (blue) shows that *Caulobacter* ParB lacks the equivalent lysine-rich amino acid patch at its CTD. Lysine residues that are important for the non-specific DNA-binding and DNA condensation activities in *Bacillus* ParB (36), and positional equivalent residues in *Caulobacter* ParB are highlighted in blue and green, respectively. Secondary-structure elements for *Bacillus* ParB CTD (PDB accession number: 5NOC) are shown above the sequence alignment. **(B)** Introducing three or more lysine residues to the *Caulobacter* CTD resulted in ParB variants with enhanced non-specific DNA-binding activity. *In vitro* binding affinities between ParB variants and *parS* DNA (left panel), and non-specific scrambled *parS* DNA (right panel). Bio-layer interferometry assays were used to measure the binding affinity of ParB (250, 500, and 1000 nM) to 20-bp double-stranded DNA that contains a *parS* site or a scrambled *parS* site. The level of ParB binding to DNA was expressed as response units (nm in shifted wavelength). Error bars represent standard deviation (SD) from three replicates.

FIG. 7

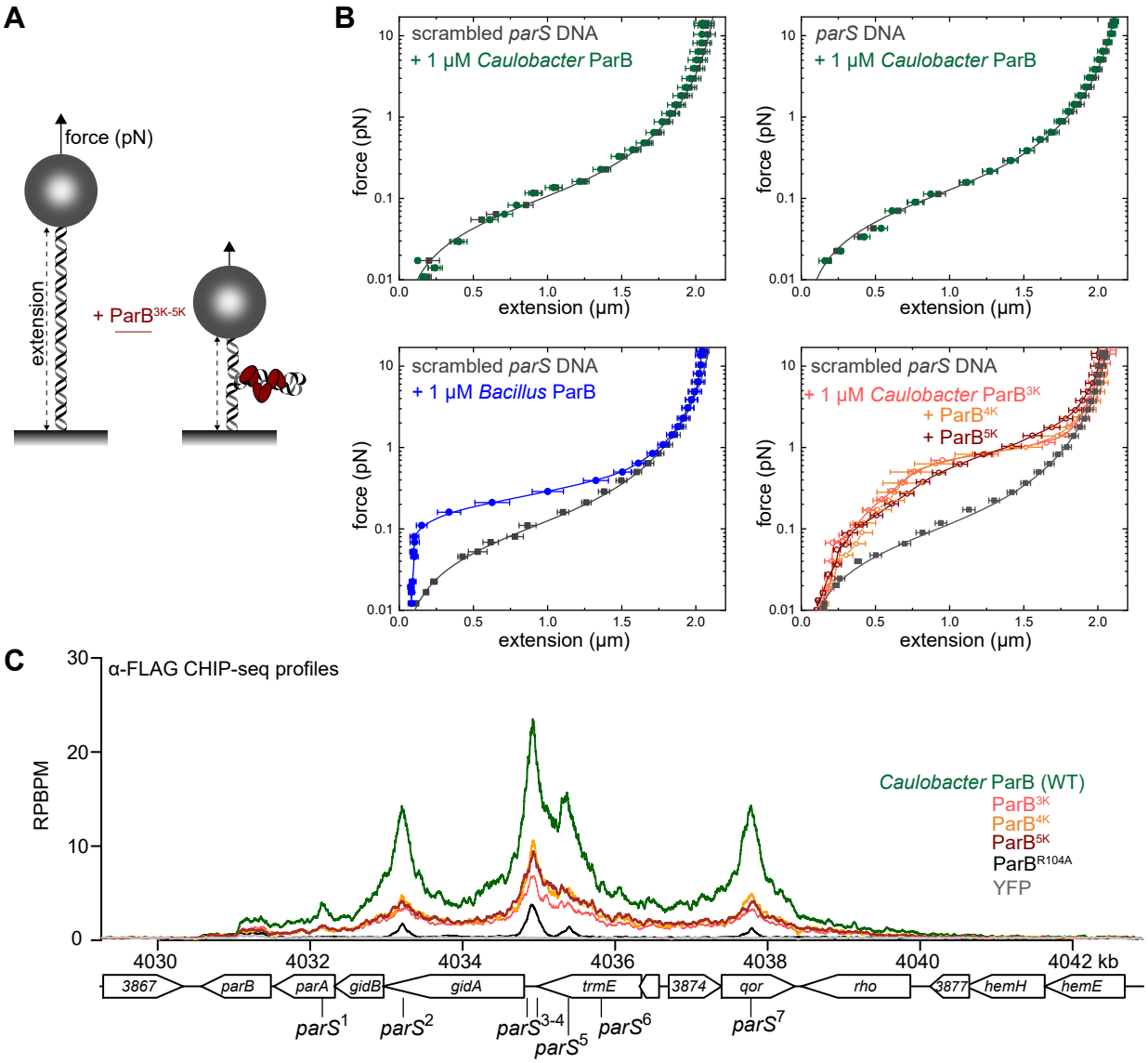

Figure 7. Engineered *Caulobacter* ParB variants with an enhanced non-specific DNA-binding activity condensed DNA *in vitro* but did not spread further than the wild-type protein *in vivo*. **(A)** Schematic of the magnetic tweezer assay that monitored ParB-dependent DNA condensation (see also the Materials and Methods). **(B)** Mean force-extension curves for *parS* DNA or scrambled *parS* DNA in the presence of 2 μM *Caulobacter* ParB (WT) (dark green), *Bacillus* ParB (WT) (blue), and *Caulobacter* ParB (3K-5K) variants (pink, orange, and brown). Data for bare DNA were fitted to the worm-like chain model. Solid lines for data in the presence of ParB variants serve as a guide to the eye. Data for non-condensed DNA were fitted to the worm-like chain model. Errors are the standard error of measurements on different molecules (N ≥ 15 molecules). **(C)** ChIP-seq profiles of a FLAG-tagged *Caulobacter* ParB (WT) (dark green), a spreading-incompetent FLAG-ParB (R104A), a non-DNA-binding protein FLAG-YFP (grey), and FLAG-ParB (3K-5K) variants (pink, orange, and brown). ChIP-seq signals were reported as the number of reads at every nucleotide along the genome (RPBPM value). The genomic context (4030 kb to 4042 kb) and *parS* sites are shown below the ChIP-seq profiles.

FIG. 8

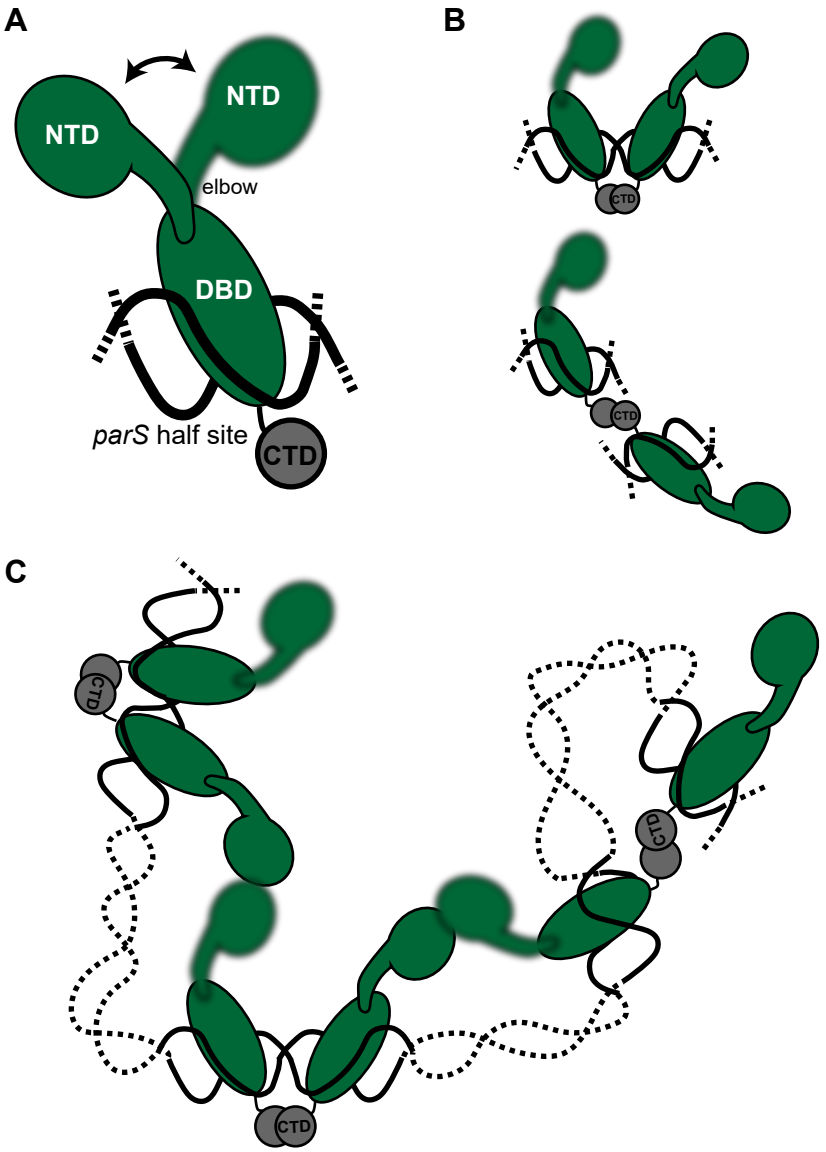

Figure 8. A model for *Caulobacter* ParB-ParB interactions via the N-terminal (NTD) and C-terminal domain (CTD). (A) The flexible elbow that connects α3 and β4 together enables the NTD of ParB to adopt multiple conformations. **(B)** The CTD might mediate the interactions between two ParB monomers bound on two adjacent half-*parS* sites (upper panel) or between two ParB monomers bound on two spatially separated half-*parS* sites (lower panel). **(C)** The coordinated actions of the CTD and the flexible NTD allows a loose but fluid network of ParB-DNA interactions to form inside the cell.

FIG. S1

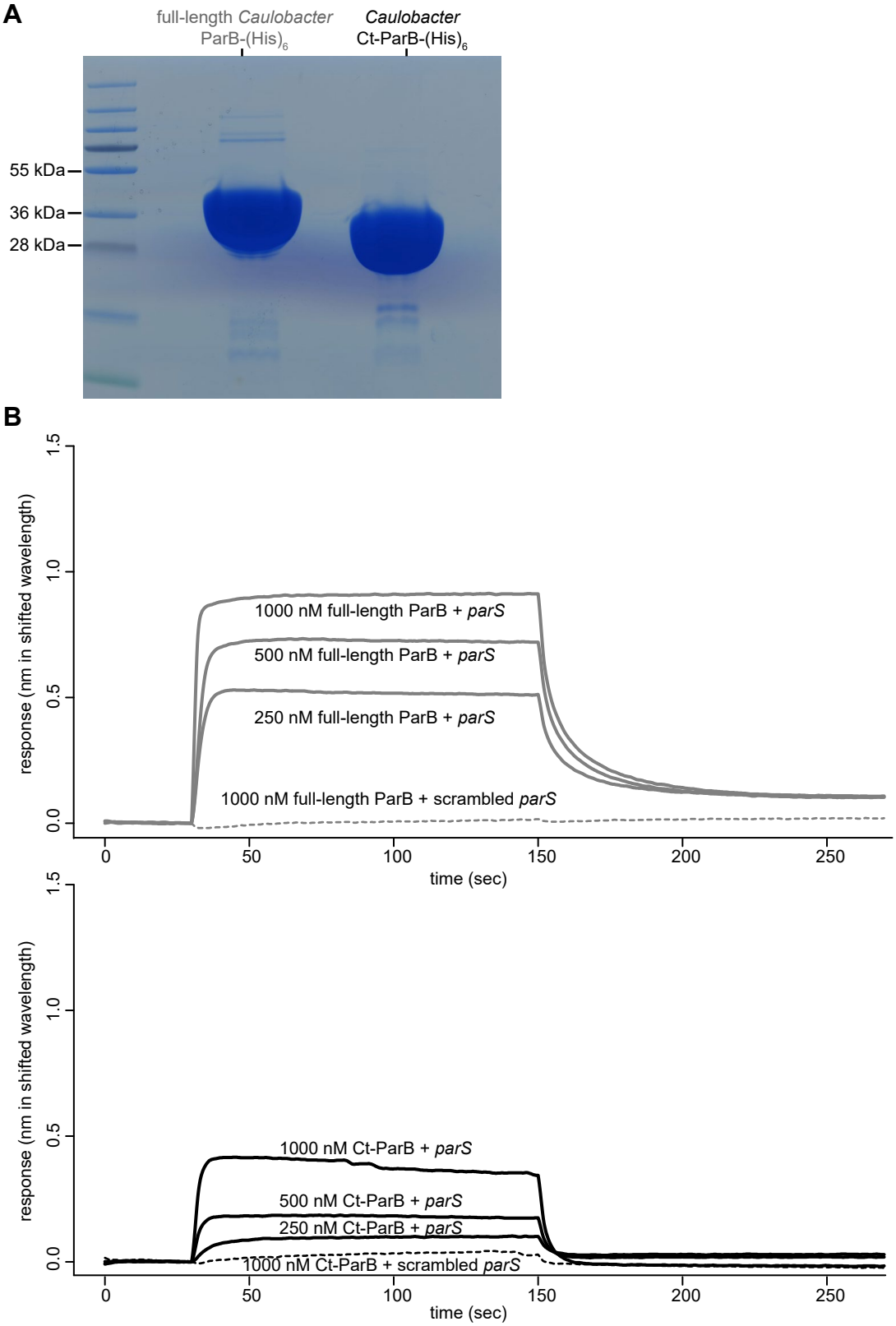

Figure S1. Purified *Caulobacter* ParB variants and their *parS* DNA-binding activities. (A) Full-length *Caulobacter* ParB and the C-terminally truncated variant (Ct-ParB) were expressed in *E. coli* and purified to near homogeneity. All variants were C-terminally His-tagged (KLAAALEHHHHHH). **(B)** Ct-ParB retains its ability to bind *parS* DNA, albeit weaker than a full-length protein. Bio-layer interferometric analysis of the interaction between a full-length *Caulobacter* ParB (grey) or a C-terminally truncated Ct-ParB (black) and a 20-bp *parS*-containing duplex DNA or the scrambled *parS* DNA. Sensors loaded with biotinylated *parS* DNA were probed with 5000 nM, 1000 nM, and 500 nM dimer concentrations of proteins. The interaction kinetics were followed by monitoring the wavelength shifts (response unit, nm) resulting from changes in the optical thickness of the sensor surface during association or dissociation of the proteins.

FIG. S2

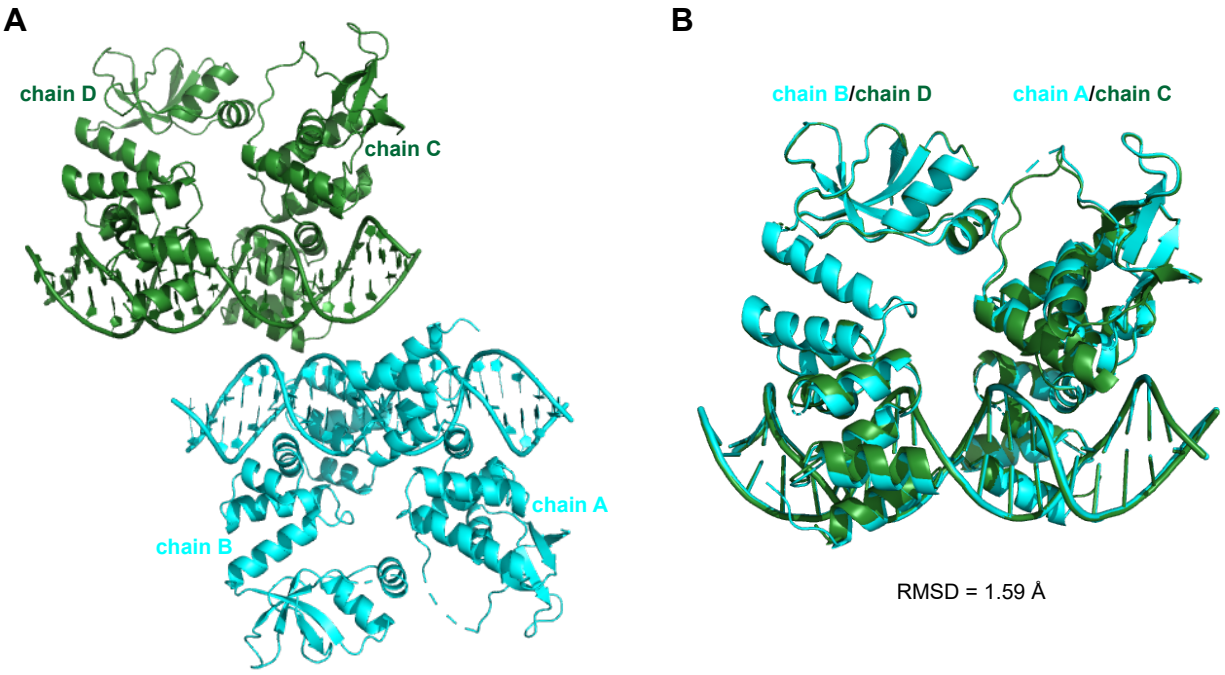

Figure S2. Composition of the asymmetric unit (ASU) of the *Caulobacter* Ct-ParB-*parS* co-crystal. (A) The ASU contains four copies of the *Caulobacter* Ct-ParB monomers (chain A, B, C, and D) and two copies of the full-size *parS* DNA. **(B)** Superimposition of the chain C-D-*parS* complex (dark green) to the chain A-B-*parS* complex (cyan) showed that the two complexes in the ASU are near identical (RMSD = 1.59 Å).

FIG. S3

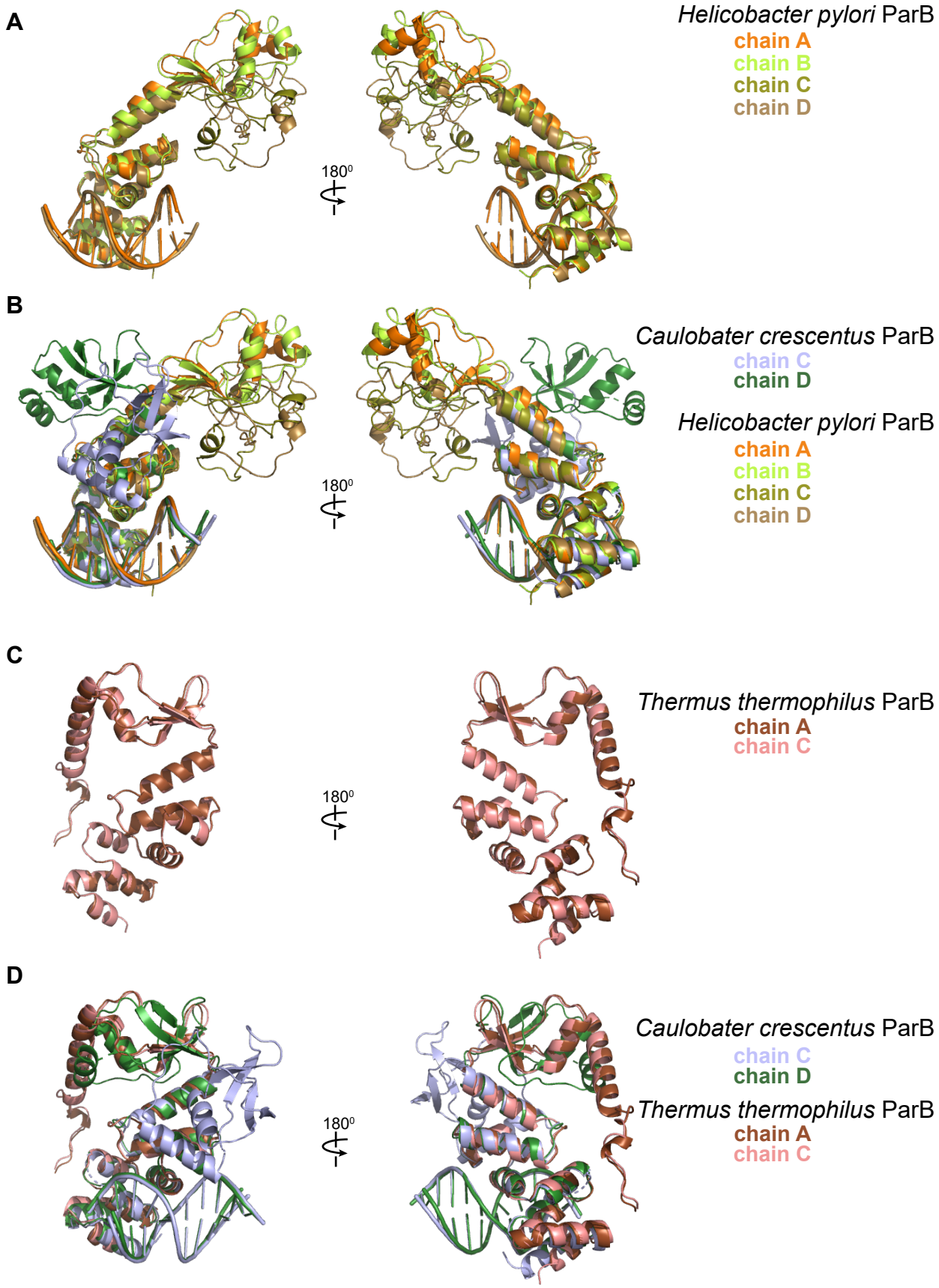

Figure S3. Structural comparisons of the *Caulobacter* Ct-ParB-*parS* complex to other ParB family members. Superposition of all crystallographically independent subunits from **(A)** *Helicobacter* Ct-ParB-*parS*; **(B)** *Helicobacter* Ct-ParB-*parS* and chain C and D of *Caulobacter* Ct-ParB-*parS*; **(C)** *Thermus* apo-Ct-ParB; **(D)** *Thermus* apo-Ct-ParB and chain C and D of *Caulobacter* Ct-ParB-*parS*. Superimposition was based on the DNA-binding domain and helices α3-α4 of the N-terminal domain.

FIG. S4

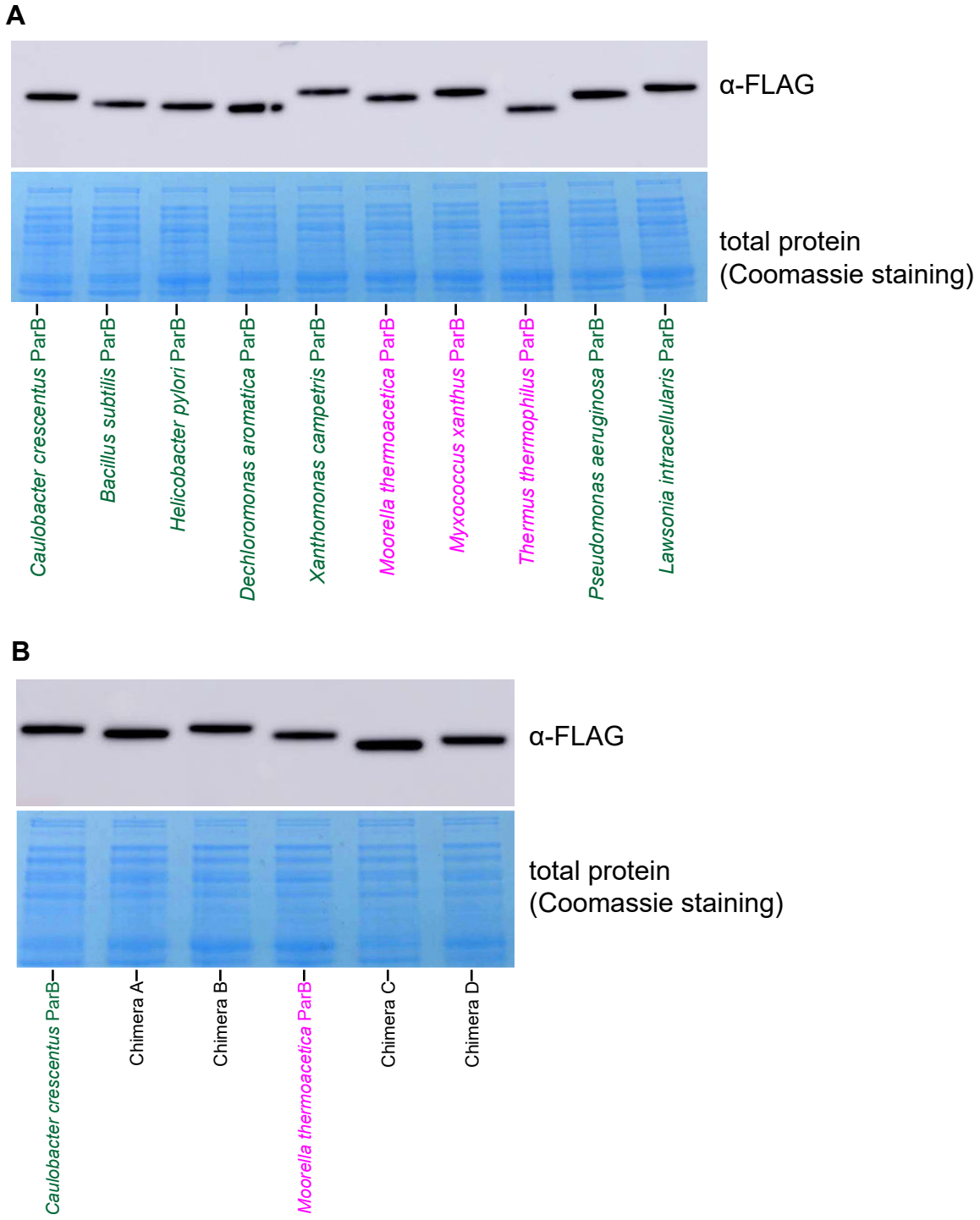

Figure S4. Production of FLAG-tagged ParB proteins from the *E. coli* heterologous host. **(A)** α-FLAG immunoblot analysis of ten FLAG-tagged ParB proteins individually expressed in the *E. coli ygcE::parS* heterologous host. **(B)** α-FLAG immunoblot analysis of FLAG-tagged chimeric ParB proteins individually expressed in the *E. coli ygcE::parS* heterologous host.

FIG. S5

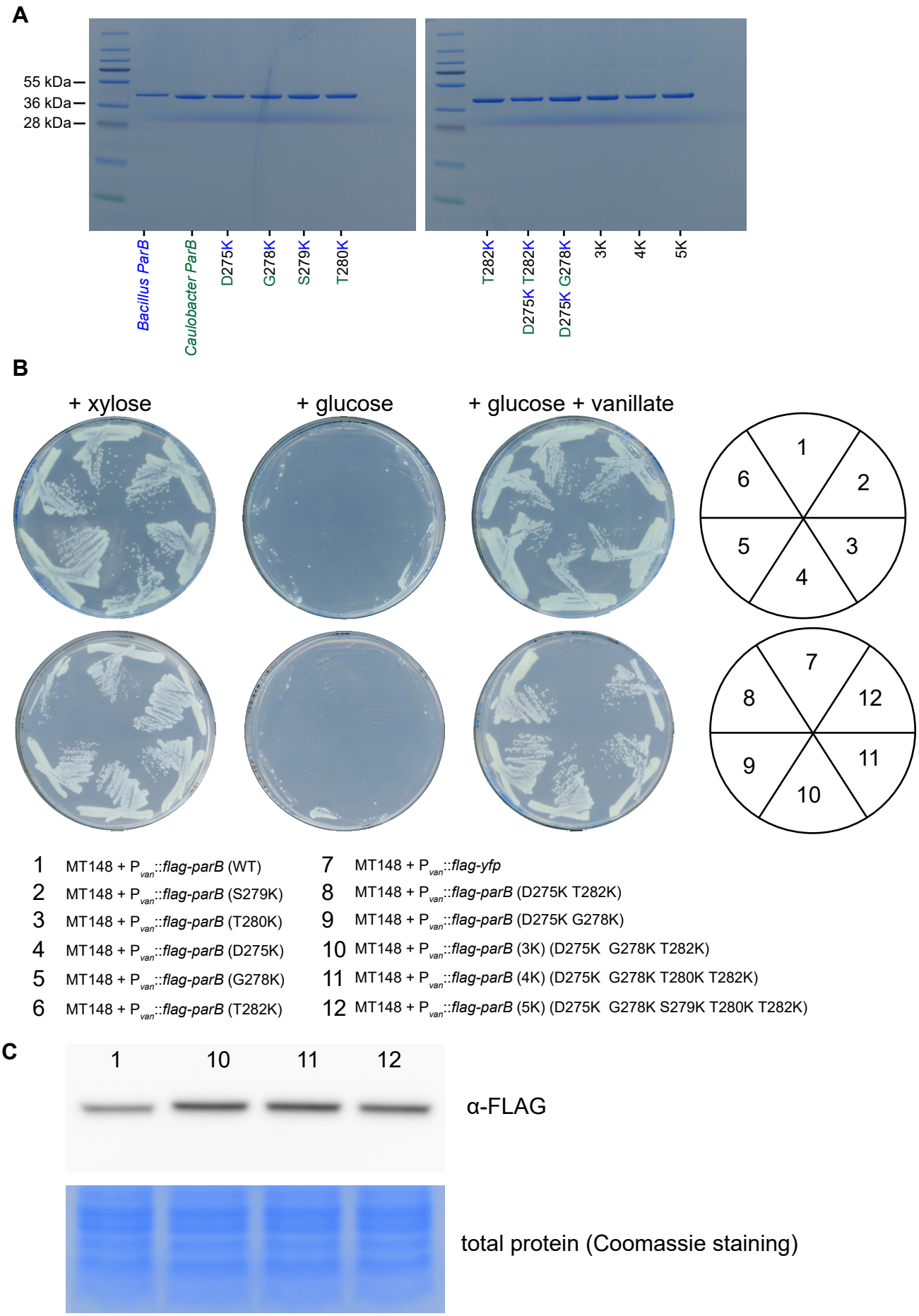

figure legend on the next page

FIG. S5

Figure S5. Expression and purification of *Caulobacter* ParB variants with additional lysine residues at their C-terminal domains. (A) *Bacillus* ParB (WT), *Caulobacter* ParB (WT) and its variants were expressed in *E. coli* and purified to near homogeneity. All variants were C-terminally His-tagged (KLAAALEHHHHHH). The identity of each protein variant (i.e. the engineered substitutions at the C-terminal domain) is shown below each lane. **(B)** *Caulobacter* ParB (1K to 5K) variants were expressed from the *van* locus (vanillate inducible) in the ParB (WT)-depletable background (*Caulobacter* MT148: *parB::*P_{xyl}*-parB*). In the presence of xylose, ParB (WT) was produced and all strains were viable, including strain 7 that expressed *yfp* from the *van* locus. On the other hand, the addition of glucose repressed the production of ParB (WT) while other ParB variants were not produced, leading to loss of viability. In the presence of both glucose and vanillate, only ParB (1K to 5K) variants were produced. Cell growth for strains (2 to 6, and 8 to 12) indicated that ParB (1K to 5K) variants can complement the lack of ParB (WT) in *Caulobacter*. **(C)** α-FLAG immunoblot analysis of *Caulobacter* strains 1, 10, 11, 12 (see panel **B**) grown in medium supplemented with glucose and vanillate.