

Condensation and localization of the partitioning protein ParB on the bacterial chromosome

Chase P. Broedersz^{a,b}, Xindan Wang^c, Yigal Meir^d, Joseph J. Loparo^e, David Z. Rudner^c, and Ned S. Wingreen^{a,f,1}

^aLewis-Sigler Institute for Integrative Genomics, ^bJoseph Henry Laboratories of Physics, and ^fDepartment of Molecular Biology, Princeton University, Princeton, NJ 08544; Departments of ^cMicrobiology and Immunobiology and ^eBiological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, MA 02115; and ^dDepartment of Physics, Ben-Gurion University, Beer Sheva 84105, Israel

Edited by Ken A. Dill, Stony Brook University, Stony Brook, NY, and approved May 1, 2014 (received for review February 10, 2014)

The ParABS system mediates chromosome segregation and plasmid partitioning in many bacteria. As part of the partitioning mechanism, ParB proteins form a nucleoprotein complex at *parS* sites. The biophysical basis underlying ParB–DNA complex formation and localization remains elusive. Specifically, it is unclear whether ParB spreads in 1D along DNA or assembles into a 3D protein–DNA complex. We show that a combination of 1D spreading bonds and a single 3D bridging bond between ParB proteins constitutes a minimal model for a condensed ParB–DNA complex. This model implies a scaling behavior for ParB-mediated silencing of *parS*-flanking genes, which we confirm to be satisfied by experimental data from P1 plasmids. Furthermore, this model is consistent with experiments on the effects of DNA roadblocks on ParB localization. Finally, we show experimentally that a single *parS* site is necessary and sufficient for ParB–DNA complex formation in vivo. Together with our model, this suggests that ParB binding to *parS* triggers a conformational switch in ParB that overcomes a nucleation barrier. Conceptually, the combination of spreading and bridging bonds in our model provides a surface tension ensuring the condensation of the ParB–DNA complex, with analogies to liquid-like compartments such as nucleoli in eukaryotes.

par system | protein–DNA condensate | protein localization

Chromosomal organization and segregation presents a major challenge in all organisms. Partitioning proteins of the ParABS system play a key role in chromosomal segregation and mediate plasmid partitioning in a variety of bacteria, including *Caulobacter crescentus*, *Bacillus subtilis*, and *Vibrio cholerae* (1, 2). This partitioning module includes a DNA binding protein (ParB) that forms a large nucleoprotein complex at centromere-like *parS* sites, frequently located near the origin of replication (Fig. 1) (3). These ParB–DNA complexes interact with ParA ATPases, leading to segregation of replicated origins (4–8). Despite the apparent simplicity of this segregation machinery, puzzles remain: What is the nature of interactions among DNA-bound ParB proteins, and how do these determine the organizational and functional properties of the ParB–DNA partitioning complex? A central question is whether ParB spreads along the DNA to form a 1D filamentous protein–DNA complex or assembles into a 3D complex on the DNA. Furthermore, it remains unclear how a small number of *parS* sites (typically 2–10) leads to robust formation and localization of such a large protein–DNA complex.

Live cell microscopy experiments indicate that ParB–GFP fusion proteins form a large fluorescent focus on the bacterial chromosome (6, 7, 9–12). ParB–DNA complexes do not rely on the presence of ParA or on the particular DNA sequence other than the *parS* site (12–14). Genome-wide chromatin immunoprecipitation studies (ChIP-chip) in *B. subtilis* have shown that ParB (Spo0J) binds site-specifically to eight origin-proximal *parS* sites (13) and also revealed significantly enhanced binding to DNA in the vicinity of each *parS* site up to distances of 18 kbp. The association of ParB proteins with sites surrounding *parS* is often referred to as spreading. However, for the sake of clarity, here we refer to this as the formation of ParB–DNA complexes. We reserve the word “spreading” to describe a purely 1D coating

of ParB along DNA to form a nucleoprotein filament. The actual structure of the ParB–DNA complexes remains unclear.

Experiments on P1 plasmids have shown that ParB overexpression leads to silencing of *parS*-proximal genes (15). However, in many cases gene silencing is only partial and strongly dependent on genomic distance from a *parS* site, suggesting that the ParB–DNA complex is partially accessible to the transcriptional machinery. Indeed, transcription of genes adjacent to *parS* sites in *B. subtilis* is virtually unaffected by Spo0J complex formation at native Spo0J expression levels (13).

The formation of a ParB–DNA complex at *parS* suggests the presence of ParB–ParB interactions. In vitro experiments indicate that ParB is largely dimerized in solution owing to its C-terminal dimerization domain. Moreover, the ParB crystal structure from *Thermus thermophilus* suggests an N-terminal interface that acts as a secondary dimerization domain between DNA-bound ParB proteins (16). In support of this idea, a Spo0J93 mutant that contains a single amino acid substitution near the N terminus (G77S), has wild-type nonspecific and specific DNA binding affinities in vitro but lacks the ability to form ParB–DNA complexes surrounding *parS* sites on the chromosome in vivo and fails to form discrete fluorescent foci (13). Taken together with the structural data, this suggests that in addition to a C-terminal dimerization domain, ParB (Spo0J) also has an N-terminal interaction domain, which is required to form a ParB–DNA complex.

Further insight into the nature of ParB–ParB interactions was gleaned from roadblock experiments (15, 17) in which a strong binding site for a transcriptional repressor was inserted near a *parS* site. In *B. subtilis* the bound repressor leads to a reduction in ParB binding to DNA adjacent to *parS*, but only in the direction of the roadblock (17). This was taken as evidence that ParB

Significance

The ParABS system is responsible for chromosome and plasmid segregation in many bacteria. A large, coherent ParB–DNA complex forms the partitioning module at the heart of this segregation machinery. Here we provide a simple theoretical model for interacting proteins on DNA to elucidate the structure of the ParB–DNA complex. We show that that both 3D bridging and 1D spreading interactions between DNA-bound ParB proteins are required to ensure the formation of a coherent protein–DNA complex. This combination of protein–protein interactions implies a surface tension that drives the condensation of ParB proteins on the DNA. The formation of such a condensed protein complex is essential for understanding how a single centromeric *parS* site can localize ParB on the DNA.

Author contributions: C.P.B., J.J.L., D.Z.R., and N.S.W. designed research; C.P.B. and X.W. performed research; C.P.B., X.W., Y.M., J.J.L., D.Z.R., and N.S.W. analyzed data; and C.P.B., D.Z.R., and N.S.W. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

¹To whom correspondence should be addressed. E-mail: wingreen@princeton.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1402529111/-DCSupplemental.

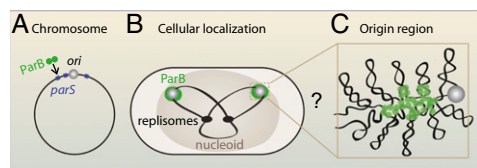


Fig. 1. Cellular localization of ParB. (A) ParB binding sites (*parS*) are frequently present near the replication origin (*ori*). (B) In cells, GFP-ParB proteins form fluorescent foci that colocalize with replication origins. The mass of DNA (the nucleoid) is shown schematically as a large oval, including a simplified view of the replicating chromosome (black lines) and the replisomes (black ovals). (C) A hypothetical magnified view of the origin region where ParB forms a large protein–DNA complex.

associates with the DNA by 1D lateral spreading from *parS* to form a nucleoprotein filament. Alternatively, ParB may form higher-order interactions resulting in a 3D protein–DNA complex.

To investigate the structure of the ParB partitioning complex, we developed a simple model for interacting proteins on DNA. We found that a combination of 1D spreading bonds and a 3D bridging bond between ParB proteins constitutes the minimal model for condensation of ParB proteins on DNA into a coherent complex. These combined interactions provide an effective surface tension, preventing fragmentation of the complex. In detail, our model predicts that ParB spreads to form multiple, short 1D domains on the DNA, connected in 3D by bridging interactions to assemble into a 3D ParB–DNA condensate (Fig. 1C). More generally, the computational model we developed here offers a simple framework to study how various interactions between DNA-binding proteins determine the structure and localization of protein–DNA complexes.

Model for Interacting Proteins on DNA

We developed a minimal model to investigate the spatial organization of interacting proteins on DNA. For simplicity, the DNA is described as a linear, self-avoiding chain on a cubic lattice in 3D; the DNA chain has a bending stiffness κ and N protein binding sites (*Supporting Information*). The addition of DNA confinement has little effect on our central results (*Supporting Information*). The DNA is coarse-grained at the scale of a protein-binding site, ℓ_0 , so that exactly one protein can bind the DNA per site of the cubic lattice.

Proteins can bind or unbind the DNA and the binding energy may vary along the DNA. Here, we considered a chain with equivalent binding sites, or a chain with just one cognate *parS* site with a binding energy $\Delta\epsilon_{\text{parS}}$ relative to all other sites. Importantly, we distinguish two types of protein–protein interactions: (i) 1D spreading interactions with strength J_S between proteins along the backbone of the DNA chain and (ii) 3D bridging interactions with strength J_B between proteins bound to nonsequential DNA sites that are nearest neighbors in 3D space. We studied the thermodynamic equilibrium behavior of this model via Monte Carlo simulations (*Supporting Information*).

Results

For many bacteria, the faithful partitioning and segregation of plasmids and the chromosome relies on ParB proteins, which form large ParB–DNA complexes localized by a few or in rare cases by one *parS* site—a 16-bp sequence that specifically binds a ParB dimer (3). We first discuss how various types of protein–protein interactions in our model would affect the structure and stability of the ParB–DNA complex in the absence of a *parS* site. We then proceed to include a *parS* site in our model to investigate how this affects the localization of the ParB cluster on the DNA, and finally we consider the role of the *parS* site in ParB cluster nucleation.

Combining 1D Spreading with 3D Bridging Interactions Is Necessary for the Formation of a Condensed Protein–DNA Complex. To serve as a benchmark, we first defined a dimer model, in which DNA-bound proteins interact through their primary dimerization domain (Fig. 2A). An additional dimerization domain on the protein (16, 18) could, in principle, engage in either 1D spreading or 3D bridging interactions leading to two additional models: a spreading model (Fig. 2B) in which each protein in the ParB dimer can form two 1D spreading interactions, one in each direction along the chain, and a bridging model (Fig. 2C), in which each dimer can engage in a single 3D bridging interaction.

The dimer, spreading, and bridging models all resulted in multiple protein clusters of varying sizes dispersed over the DNA (with no *parS* site), as shown in representative images of simulated protein–DNA complexes in Fig. 2A–C. To quantify the distribution of cluster sizes in a system with a fixed number of proteins, we evaluated the probability $P(n)$ for a randomly chosen protein to be part of a cluster of size n . We defined a cluster to be a 3D, contiguous collection of DNA-bound proteins. For all three models (dimer, spreading, and bridging) we observed a broad distribution peaked at small cluster sizes (Fig. 2F). The enhanced clustering found in the bridging model is consistent with a recent theoretical study showing that bridging can induce an effective entropic attraction between proteins (19). However, with only a single bridging bond per protein, this effective lateral interaction between proteins was found to be weak and the resulting clusters were small, consistent with our results. Changing the interaction strength in both the spreading

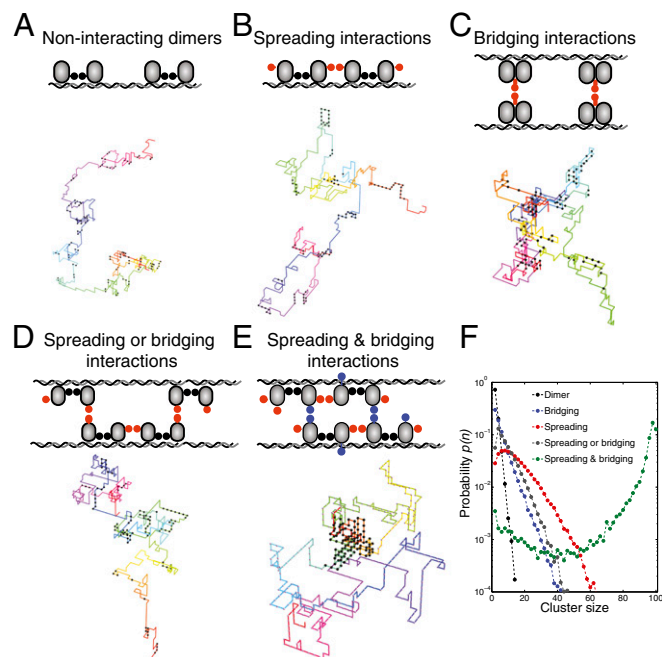


Fig. 2. Simple models for interacting DNA-bound proteins. Schematics include a representative image of a DNA–protein complex from simulations: (A) dimer model, (B) spreading model, (C) bridging model, (D) spreading or bridging model, and (E) spreading and bridging model. In the spreading or bridging model there are two interaction domains per monomer, one of which can be either in spreading or in bridging mode. In the spreading and bridging model there are two 1D spreading bonds along the DNA and one 3D bridging bond. The bond energies are $-6k_B T$ except in the spreading or bridging model, for which the bridging bond has an energy of $-7k_B T$. The bending rigidity of the DNA was set to $\kappa = k_B T \ell_0$. The total number of proteins is fixed at $m = 100$ on a DNA chain of length $N = 500$ (i.e., with 500 binding sites). (F) Probability distribution $P(n)$ for a DNA-bound protein to be part of a cluster of size n .

model and the bridging model can modify the typical cluster size; however, the cluster size distributions remain broad.

Conceptually, the fragmentation in these models can be traced to the entropy of dispersing the proteins over the DNA, which is favored over the energy of forming a single large cluster. To account for the experimental observation of coherent ParB-DNA foci (6, 7, 9–12), we expanded the model by combining 1D spreading and 3D bridging interactions to avoid entropic fragmentation. In the simplest of such models, the proteins still only have two interaction domains, as in the spreading model, but now one of the interaction domains can either function in spreading or in bridging mode, whereas the other domain can only participate in a 1D spreading bond (Fig. 2D). Importantly, this model failed to prevent fragmentation because the distribution of cluster sizes was broad and qualitatively similar to the previous models (Fig. 2F).

We next considered a model with one additional ingredient: each ParB protein is able to interact with its two neighbors via 1D spreading interactions along with an additional 3D bridging interaction. Strikingly, in this spreading and bridging model the majority of the proteins cluster together in one coherent focus on the DNA (Fig. 2E). This is reflected as a large, narrow peak centered around the maximal number of proteins in the cluster size distribution (Fig. 2F).

Importantly, for the spreading and bridging model, unsatisfied dangling ParB bonds at the surface of the cluster generate a surface tension that counters the tendency of entropy to fragment the condensate into multiple small clusters. By contrast, the spreading model cannot support a condensed phase because 1D surface tension always loses to entropy. Moreover, for the bridging model and spreading or bridging model the most stable state consists of small clusters with no surface tension because all bonds can be satisfied. (A variety of other models can be excluded for similar reasons) (*Supporting Information*). Thus, the combination of both spreading and bridging interactions constitutes the minimal requirement for a condensed phase (i.e., a large, coherent ParB-DNA cluster). These conclusions also hold when we include DNA confinement in our model (*Supporting Information*).

A *parS* Site Can Localize a Condensed ParB Complex. DNA-bound ParB proteins localize around *parS* sites. ParB binds specifically to a *parS* site, but in vitro experiments suggest that the binding affinity to a *parS* site is only roughly 10-fold higher than to nonspecific DNA (13). Thus, it is unclear how a few *parS* sites can localize the majority of ParB proteins.

To investigate the effects of a *parS* site on the organization of ParB on the DNA, we inserted a site at the center of the DNA in our model with a strong binding energy $\Delta\epsilon_{\text{parS}} = -10k_B T$ relative to the other nonspecific binding sites, where $k_B T$ is the thermal energy. Although this *parS* site binds a protein with a probability close to 1, for the spreading, bridging, and spreading or bridging models the binding probability decays rapidly to a background value ~ 0.2 , corresponding here to the average coverage of the DNA (Fig. 3A). Thus, in these models the *parS* site is capable of localizing only one of the small ParB clusters at a time. By contrast, the spreading and bridging model produces a broad binding profile peaked around the *parS* site, with a binding probability that decays nearly to zero far from the *parS* site. This approximately triangular binding profile is consistent with the formation of a single, localized ParB-DNA cluster, which can shift as a whole as long as it overlaps with the *parS* site (*Supporting Information*). Thus, the ability of one *parS* site to localize ParB on the DNA in the spreading and bridging model can be traced back to the coalescence of the vast majority of proteins into a large ParB-DNA cluster, ensuring the localization of all these ParB proteins around a sufficiently strong *parS* site (Fig. 3B and Eq. 2). Finally, we confirmed that the ParB binding profiles remain qualitatively similar when DNA confinement is included (*Supporting Information*).

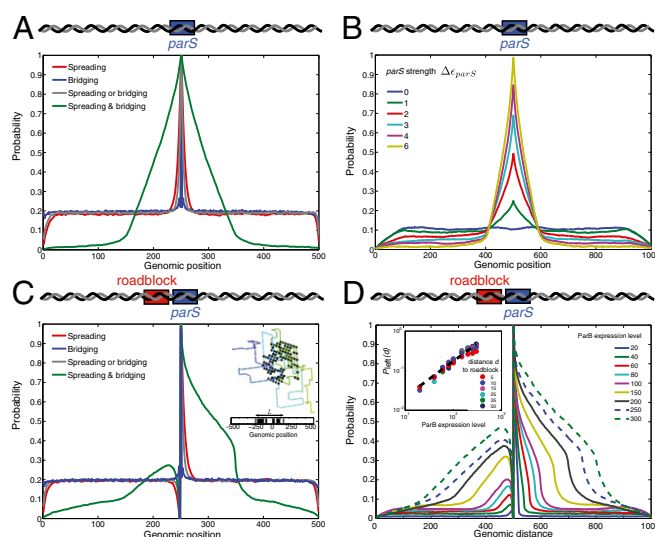


Fig. 3. ParB binding profiles, ParB localization, and the effects of a roadblock. (A) Probability of a bound protein versus genomic position with a central *parS* site (parameters as in Fig. 2 with $\Delta\epsilon_{\text{parS}} = -10k_B T$, $N = 500$ binding sites, and $m = 100$ proteins). (B) The probability of a bound ParB versus genomic position for the spreading and bridging model for different ParB-*parS* binding energies relative to nonspecific sites ($N = 1,000$). (C) A roadblock is inserted immediately to the left of the strong *parS* site, resulting in an asymmetric spreading profile ($N = 500$). (Inset) ParB-DNA complex with loops from simulations on the spreading and bridging model and the corresponding 1D binding profile on the DNA. (D) Probability of a bound protein versus genomic position for the spreading and bridging model with a roadblock inserted immediately to the left of *parS* ($N = 1,000$). The total number of ParB proteins on the DNA is varied. (Inset) Scaling of the binding probability, $p_{\text{left}}(d)$, at positions d binding sites to the left of the roadblock as a function of ParB expression level. The dashed line indicates power-law scaling with an exponent of 0.9.

Roadblocks Strongly Affect the Localization of 3D ParB-DNA Complexes. In *B. subtilis*, the insertion of a strong binding site for a transcriptional repressor in the vicinity of a *parS* site led to a reduction in ParB-DNA interaction in the direction of the repressor bound site (17). Based on this observation, it was proposed that ParB associates with the DNA by spreading one-dimensionally from a *parS* site. To investigate the effect of a roadblock on the binding profiles in our models, we inserted a blocked site, to which ParB is not allowed to bind, immediately to the left of the *parS* site (Fig. 3C). The binding profile of the bridging model is hardly affected by the roadblock, whereas for the spreading model and the spreading or bridging model the binding of ParB is obstructed by the roadblock. Strikingly, the binding profile is also strongly asymmetric for the spreading and bridging model, consistent with the observations from roadblock experiments in *B. subtilis* (17).

The small binding probability of ParB to the left of the roadblock for the spreading and bridging model indicates that there are some configurations in which the ParB cluster forms loops that bypass the roadblock. This raises the question why it is thermodynamically more favorable for the complex to stay to the *parS* side of the roadblock instead of looping around it. To investigate this, we analyzed the loops extruding from the surface of the cluster (Fig. 3C, Inset). Interestingly, we find that the average loop length $\langle \ell \rangle$ is independent of cluster size M_c and of the energy of a spreading bond J_S (*Supporting Information*). By contrast, the average number of loops behaves as $\langle N_{\text{loops}} \rangle \sim \exp(-J_S/k_B T) M_c^\alpha$, with $\alpha = 0.9$. The formation of a loop requires the breaking of a spreading bond and thus occurs with a probability $\sim \exp(-J_S/k_B T)$. The dependence of $\langle N_{\text{loops}} \rangle$ on M_c simply reflects the surface area of the cluster, because loops extend from the surface (*Supporting Information*).

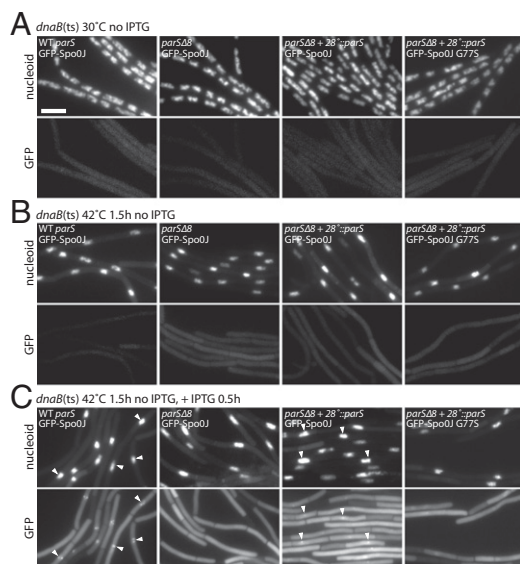


Fig. 5. A single *parS* site is necessary and sufficient to generate a GFP-Spo0J focus. (A) Representative images of *B. subtilis* cells harboring a *dnaB(ts)* allele with all eight wild-type *parS* sites (BWV2454, WT *parS*, first panel); none of the eight endogenous *parS* sites (BWV2456, *parS*Δ8, second panel); *parS*Δ8 with an ectopic *parS* site inserted at 28° (*amyE*) (BWV2458 and BWV2789, *parS*Δ8+28°::*parS*, third and fourth panels) grown at 30 °C (A) and after growth at 42 °C for 1.5 h (B) to block new rounds of initiation of replication generating a single chromosome (Supporting Information). Expression of GFP-Spo0J (first to third panels) or GFP-Spo0J (G77S) (fourth panel) was then induced by the addition of IPTG (0.5 mM final concentration) for 0.5 h at 42 °C (C) (see Supporting Information for induction control). GFP-Spo0J did not form foci in the strain lacking *parS* sites (*parS*Δ8, second panel) and formed a single focus per nucleoid in the *parS*Δ8+28°::*parS* strain (third panel). GFP-Spo0J (G77S) did not form foci in the *parS*Δ8+28°::*parS* strain (fourth panel). Based on comparison with Spo0J (G77S) (C, fourth panel) that binds *parS* but is unable to spread to neighboring sites (13), the foci observed in the *parS*Δ8+28°::*parS* strain (C, third panel) reflect GFP-Spo0J nucleoprotein complexes. Gray scale of DAPI-stained nucleoids (upper panels) and GFP (lower panels) is shown. GFP-Spo0J foci (white caret) are highlighted in the DAPI and GFP panels. (Scale bar, 4 μm.)

in vivo but does not form a ParB–DNA complex (14). When grown at the permissive temperature (30 °C) in the absence of IPTG, all four strains contained DAPI-stained chromosomes (called nucleoids) that resembled the chromosomes observed in the Spo0J null mutant (Fig. 5A). Importantly, GFP-Spo0J fluorescence was virtually undetectable (Fig. 5A). After 1.5 h of growth at the restrictive temperature (42 °C), virtually all cells in the four strains contained a single replicated chromosome (Fig. 5B and Supporting Information). At this time, IPTG was added to induce expression of GFP-Spo0J. Thirty minutes later, GFP-Spo0J localization was monitored. As expected, the cells harboring the endogenous origin-proximal *parS* sites had a single fluorescent focus (or cluster of foci) that colocalized with the nucleoid (Fig. 5C). Moreover, the cells lacking *parS* sites had no detectable GFP-Spo0J foci. By contrast, virtually all cells with a single *parS* site had a faint GFP-Spo0J focus that colocalized with the nucleoid. These foci correspond to Spo0J–DNA condensates and not simply a Spo0J dimer bound to the *parS* site because no foci were detected in the GFP-Spo0J (G77S) mutant that binds *parS* but does not form a nucleoprotein complex (Fig. 5C). Immunoblot analysis indicates that the GFP-Spo0J fusions were expressed at similar levels in the different strains and the cytoplasmic fluorescence was not due to proteolytic release of GFP (Supporting Information). Taken together, these data indicate that a single *parS* site per cell is capable of forming a Spo0J–DNA complex, whereas no complex forms in the absence of a *parS* site. We therefore favor a model in which a Spo0J dimer bound to a *parS* site is able to nucleate a

Spo0J–DNA complex much more rapidly than a Spo0J dimer bound elsewhere.

Importantly, the requisite *parS*-induced reduction of the nucleation free-energy barrier, ΔF_{parS} , would have to be dramatic because nucleation of ParB–DNA clusters does not occur at any of the nonspecific DNA binding sites on the chromosome over multiple cell lifetimes under normal conditions. Thus, *parS*-specific nucleation would have to be considerably more than $N \approx 1.3 \times 10^5$ times faster than nucleation at a nonspecific site assuming a 30-bp footprint for ParB (17); this implies $\exp(-\Delta F_{\text{parS}}/k_B T) > N$, or equivalently,

$$\Delta F_{\text{parS}} < -k_B T \ln(N). \quad [1]$$

However, the relative affinity of specific binding of ParB (Spo0J) to *parS* binding was found to be only 10-fold higher than nonspecific binding (13), corresponding to $\Delta \epsilon_{\text{parS}} \approx -2.3 k_B T$. Thus, the specificity of a *parS* site is far too low for simple nucleation to occur reliably at a single *parS* site, but not elsewhere on the chromosome. A related conundrum is that the measured specificity of *parS* sites is theoretically insufficient to localize the ParB–DNA complex; the free-energy contribution from the *parS* site is required to be larger than the entropy gained by delocalizing the cluster over the DNA,

$$\Delta F_{\text{parS}} < k_B T \ln \left(\frac{2M_c}{N - 2M_c} \right). \quad [2]$$

Here, we assumed that this cluster extends over $\sim M_c$ out of a total of N binding sites on the DNA. In a typical *B. subtilis* cell, $M_c \approx 1,000$ (21) and $N \approx 1.3 \times 10^5$, thus requiring $\Delta F_{\text{parS}} < -4.2 k_B T$, also lower than the experimental value of $\Delta \epsilon_{\text{parS}} \approx -2.3 k_B T$ (13). Thus, the measured weak specificity of ParB binding to a *parS* site cannot account for the localization of the ParB–DNA complex, nor can it explain the *parS*-specific nucleation of ParB–DNA clusters. This analysis provides additional support for the idea that binding of ParB to a *parS* site facilitates nucleation of a ParB–DNA complex, beyond simply enhancing the local ParB density on the DNA.

Discussion

In summary, we have identified a minimal model for interacting ParB proteins on DNA that produces large, coherent ParB–DNA complexes, as observed in live cells (6, 7, 9–12). This model offers a conceptual framework in which a combination of 1D spreading and 3D bridging protein–protein interactions thermodynamically stabilizes a ParB–DNA condensate. Put simply, we find that both spreading and bridging interactions are required to provide a surface tension that prevents fragmentation of condensed protein–DNA complexes. Indeed, we speculate that the ParB–DNA complex may organize like a liquid-like (non-membrane-bound) compartment analogous to, e.g., the nucleolus in eukaryotes (22, 23).

The spreading and bridging model accounts for multiple experimental observations, including ParB reorganization by DNA roadblocks and ParB-induced gene silencing. What aspects of the model are essential to explain these experiments? The roadblock experiments indicate the presence of strong spreading bonds between ParB proteins (17), and the formation and localization of a coherent ParB–DNA complex (6, 7, 9–12) requires an additional bridging bond between proteins. Crystallographic studies of ParB proteins indicate that there are at least two interaction domains: one at the C terminus that likely promotes dimerization and one at the N terminus (16). An N-terminal Spo0J mutant (G77S) binds *parS* in vivo but fails to form nucleoprotein complexes. Accordingly, this domain is likely to function in either 1D spreading or 3D bridging. Our model predicts that each ParB monomer has three interaction domains, and thus can form two spreading interactions along the DNA as well as one bridging interaction with another DNA-bound ParB

protein, resulting in the formation of clusters consisting of multiple 1D spreading domains of ParB connected in 3D by bridging interactions to form a coherent ParB–DNA complex. These structural predictions could be tested by single-molecule pulling experiments on condensed ParB–DNA complexes. Recent single-molecule experiments using flow-stretched DNA provide evidence for bridging interactions between ParB proteins in vitro, consistent with our model (24). These experiments also show that ParB is capable of condensing DNA with rates up to $\sim 1 \mu\text{m/s}$. Interestingly, polymer dynamics theory suggests that the kinetics of DNA condensation in vivo may be slow but could be enhanced in the presence of type-II topoisomerases (25).

We showed experimentally that a single *parS* site is both necessary and sufficient for nucleation of a ParB–DNA complex. These experiments together with calculations based on our model led us to conclude that binding of ParB to a *parS* site effectively lowers the nucleation barrier, leading reliably to complex nucleation at *parS* sites but not at nonspecific sites. Curiously, we find that the requisite *parS*-induced reduction of the nucleation barrier has to be much larger than the energy difference inferred from the measured binding affinity of ParB to *parS* versus nonspecific sites. What mechanism could give rise to a dramatically reduced nucleation barrier at a *parS* site? It has been speculated based on structural data that the ability of DNA-bound ParB proteins to interact may require a conformational change (16). This proposed conformational transition may involve breaking the primary C-terminal dimerization bond. Thus, there may be two binding modes: ParB could bind to the DNA as closed inert dimers as in our dimer model, or as open dimers as in our spreading and bridging model (Fig. 2*A* and *E*). Building on this idea, we conjecture that the transition to the open configuration is favored at a *parS* site, whereas the closed configuration is favored at nonspecific sites. An estimate based on this scenario is that the nucleation barrier at a *parS* site, relative to nonspecific sites, is lowered by $E_0 \approx 11.5k_B T$, using the dissociation constant $K_d \approx 10 \mu\text{M}$ for ParB dimerization in solution (16). Combining this with the contribution from binding ParB specifically to a *parS* site gives $\Delta F_{\text{parS}} \approx -13.8k_B T$, which

meets both the nucleation and localization requirements (Eqs. 1 and 2). Thus, this *parS*-specific two-state nucleation model solves the puzzle of how a single *parS* site can be necessary and sufficient for the production and localization of a ParB–DNA complex. Importantly, equilibrium behavior of this two-state model is equivalent to our one-state spreading and bridging model (Supporting Information).

The biophysical properties of the ParB–DNA complex implied by our model may have important implications for the segregation and organization of the chromosome. The surface tension of the ParB–DNA complex in the spreading and bridging model indicates that it is thermodynamically most favorable to form a single ParB–DNA complex that recruits all ori-proximal *parS* sites. Our model highlights a puzzling aspect of the observed splitting of the ParB–DNA complex during segregation (6, 7, 9–12), which would be disfavored by the surface tension of the ParB–DNA complex. This suggests that the splitting of the ParB–DNA complex may rely on active processes, in the case of *B. subtilis*, likely mediated by the ParA ATPase and the action of structural maintenance of chromosomes condensin complexes (26, 27).

Our model may also help provide insights into the organization of DNA by nucleoid-associated proteins (NAPs). The model offers a simple framework to study how the specific types of protein–protein interactions in NAPs determine the structure and function of protein–DNA complexes.

Materials and Methods

Bacteria Strains and Growth. All *B. subtilis* strains were derived from the prototrophic strain PY79. Cells were grown in defined rich casein hydrolysate medium (28) at indicated temperatures. Strains, plasmids, oligonucleotide primers, and our data analysis are described in Supporting Information.

ACKNOWLEDGMENTS. We thank B. Bratton, Z. Gitai, B. Machta, and M. Tikhonov for insightful discussions. This work was supported in part by a Lewis–Sigler fellowship (to C.P.B.), National Science Foundation Grant PHY-0957573 (to N.S.W.), National Institutes of Health Grants GM086466 and GM073831 (to D.Z.R.), National Science Foundation CAREER Award MCB-1148818 (to J.J.L.), the Human Frontier Science Program (X.W.), and the hospitality of the Aspen Center for Physics (National Science Foundation Grant PHYS-1066293).

1. Toro E, Shapiro L (2010) Bacterial chromosome organization and segregation. *Cold Spring Harb Perspect Biol* 2(2):a000349.
2. Surtees JA, Funnell BE (2003) Plasmid and chromosome traffic control: How ParA and ParB drive partition. *Curr Top Dev Biol* 56:145–180.
3. Livny J, Yamaichi Y, Waldor MK (2007) Distribution of centromere-like *parS* sites in bacteria: Insights from comparative genomics. *J Bacteriol* 189(23):8693–8703.
4. Banigan EJ, Gelbart MA, Gitai Z, Wingreen NS, Liu AJ (2011) Filament depolymerization can explain chromosome pulling during bacterial mitosis. *PLOS Comput Biol* 7(9):e1002145.
5. Gerdes K, Howard M, Szardenings F (2010) Pushing and pulling in prokaryotic DNA segregation. *Cell* 141(6):927–942.
6. Shebelut CW, Guberman JM, van Teeffelen S, Yakhnina AA, Gitai Z (2010) Caulobacter chromosome segregation is an ordered multistep process. *Proc Natl Acad Sci USA* 107(32):14194–14198.
7. Toro E, Hong SH, McAdams HH, Shapiro L (2008) Caulobacter requires a dedicated mechanism to initiate chromosome segregation. *Proc Natl Acad Sci USA* 105(40):15435–15440.
8. Vecchiarelli AG, Hwang LC, Mizuuchi K (2013) Cell-free study of F plasmid partition provides evidence for cargo transport by a diffusion-ratchet mechanism. *Proc Natl Acad Sci USA* 110(15):E1390–E1397.
9. Kusiak M, Gapczynska A, Plochocka D, Thomas CM, Jagura-Burdzy G (2011) Binding and spreading of ParB on DNA determine its biological function in *Pseudomonas aeruginosa*. *J Bacteriol* 193(13):3342–3355.
10. Lin DC, Levin PA, Grossman AD (1997) Bipolar localization of a chromosome partition protein in *Bacillus subtilis*. *Proc Natl Acad Sci USA* 94(9):4721–4726.
11. Mohl DA, Gober JW (1997) Cell cycle-dependent polar localization of chromosome partitioning proteins in *Caulobacter crescentus*. *Cell* 88(5):675–684.
12. Sullivan NL, Marquis KA, Rudner DZ (2009) Recruitment of SMC by ParB-*parS* organizes the origin region and promotes efficient chromosome segregation. *Cell* 137(4):697–707.
13. Breier AM, Grossman AD (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(3):703–718.
14. Lee PS, Lin DC, Moriya S, Grossman AD (2003) Effects of the chromosome partitioning protein Spo0J (ParB) on *oriC* positioning and replication initiation in *Bacillus subtilis*. *J Bacteriol* 185(4):1326–1337.
15. Rodionov O, Lobočka M, Yarmolinsky M (1999) Silencing of genes flanking the P1 plasmid centromere. *Science* 283(5401):546–549.
16. Leonard TA, Butler PJ, Löwe J (2004) Structural analysis of the chromosome segregation protein Spo0J from *Thermus thermophilus*. *Mol Microbiol* 53(2):419–432.
17. Murray J, Ferreira H, Errington J (2006) The bacterial chromosome segregation protein Spo0J spreads along DNA from *parS* nucleation sites. *Mol Microbiol* 61(5):1352–1361.
18. Surtees JA, Funnell BE (1999) P1 ParB domain structure includes two independent multimerization domains. *J Bacteriol* 181(19):5898–5908.
19. Brackley CA, Taylor S, Papantonis A, Cook PR, Marenduzzo D (2013) Nonspecific bridging-induced attraction drives clustering of DNA-binding proteins and genome organization. *Proc Natl Acad Sci USA* 110(38):E3605–E3611.
20. Glaser P, et al. (1997) Dynamic, mitotic-like behavior of a bacterial protein required for accurate chromosome partitioning. *Genes Dev* 11(9):1160–1168.
21. Rokop ME, Auchtung JM, Grossman AD (2004) Control of DNA replication initiation by recruitment of an essential initiation protein to the membrane of *Bacillus subtilis*. *Mol Microbiol* 52(6):1757–1767.
22. Hyman AA, Simons K (2012) Cell biology. Beyond oil and water—phase transitions in cells. *Science* 337(6098):1047–1049.
23. Weber SC, Brangwynne CP (2012) Getting RNA and protein in phase. *Cell* 149(6):1188–1191.
24. Graham TGW, et al. (2014) ParB spreading requires dna bridging. *Genes and Development*, in press.
25. Sikorav JL, Pelta J, Livolant F (1994) A liquid crystalline phase in spermidine-condensed DNA. *Biophys J* 67(4):1387–1392.
26. Wang X, Montero Llopis P, Rudner DZ (2013) Organization and segregation of bacterial chromosomes. *Nat Rev Genet* 14(3):191–203.
27. Wang X, Tang OW, Riley EP, Rudner DZ (2014) The SMC condensin complex is required for origin segregation in *Bacillus subtilis*. *Curr Biol* 24(3):287–292.
28. Youngman PJ, Perkins JB, Losick R (1983) Genetic transposition and insertional mutagenesis in *Bacillus subtilis* with *Streptococcus faecalis* transposon Tn917. *Proc Natl Acad Sci USA* 80(8):2305–2309.

Supporting Information

Broedersz et al. 10.1073/pnas.1402529111

1. Monte Carlo Simulation of a Lattice Model for DNA with Interacting DNA-Binding Proteins

We used a Monte Carlo procedure to simulate the equilibrium conformations of DNA with interacting DNA-bound proteins. In our model, the DNA is described as a linear, self-avoiding chain on a cubic lattice in 3D. For simplicity, the DNA is taken to be coarse-grained at the scale of a protein-binding site, ℓ_0 , so that exactly one protein can bind the DNA per site of the cubic lattice. Hence, a DNA chain that visits N lattice sites has exactly N potential binding sites. To account for the bending stiffness κ of the DNA chain, we assign a bending energy

$$H_{\text{DNA}} = \frac{\kappa}{2\ell_0} \sum_{i=1}^{N-1} (\Delta \mathbf{r}_i)^2, \quad [\text{S1}]$$

where $\Delta \mathbf{r}_i$ is a vector describing the change in orientation of a DNA segment between sites i and $i+1$. Because our lattice model only allows for right-angle bends, $(\Delta \mathbf{r}_i)^2 = 2$ for a bend, and $(\Delta \mathbf{r}_i)^2 = 0$ otherwise.

Proteins can bind to the DNA and move between DNA binding sites, and DNA-bound proteins can interact with each other. The DNA binding energy of a protein, ϵ_i , may vary along the DNA. In our simulations we considered a chain with equivalent binding sites, or a chain with just one cognate *parS* site with a binding energy $\Delta \epsilon_{\text{parS}}$ relative to all other sites. The occupancy of a binding site i is described by ϕ_i , where $\phi_i = 1$ if the binding site is occupied by a protein and $\phi_i = 0$ otherwise.

1.1. Protein–Protein Interactions. Our model allows for two distinct types of interactions between proteins on the DNA: 1D spreading interactions with a strength J_S between proteins bound to sequential binding sites on the DNA chain and 3D bridging interactions with strength J_B between proteins bound to nonsequential DNA sites. Such bridging interactions are only possible when nonsequential DNA binding sites at which proteins are bound are nearest neighbors in 3D. The spreading interactions are directed along the DNA and are thus independent of the 3D conformation of the DNA. By contrast, the formation of bridging bonds depends sensitively on the 3D DNA conformation.

This model encompasses the freedom to choose the number of spreading and bridging interaction domains on each protein. Because we defined the system on a cubic lattice, the general model allows at most two spreading and four bridging interactions per protein. For example, in the spreading and bridging model there are two spreading interactions, but only one bridging interaction per protein, whereas in the dimer model each protein can engage in only one spreading interaction and no bridging interactions. Thus, in general a protein can only form a limited number of interactions at a time. To account for this constraint in our model, we assign oriented interaction domains to the proteins. Each spreading domain is oriented in either of the two directions along the DNA chain, and each bridging domain is oriented in one of the four remaining directions. Furthermore, two interaction domains on the same protein cannot be oriented in the same direction simultaneously. Importantly, interactions between proteins are only possible when the interaction domains of the two proteins are oriented to face each other.

The energies of the spreading and bridging interactions, as well as the DNA binding energy of the proteins, yield the total interaction energy:

$$H_{\text{int}} = J_S \sum_{i=1}^{N-1} g_{i,i+1} \phi_i \phi_{i+1} + J_B \sum_{\langle ij \rangle_{3D}} g_{ij} \phi_i \phi_j. \quad [\text{S2}]$$

Here, the second sum only runs over nonsequential DNA sites, which are nearest-neighbor sites in 3D space. Interactions between proteins are only possible when the associated interaction domains of the proteins are oriented to face each other. If the two respective domains on neighboring proteins at sites i and j are favorably aligned such that a bond can form $g_{ij} = 1$ (proteins interact), and otherwise $g_{ij} = 0$ (no interaction).

Our Monte Carlo procedure consists of a set of DNA and protein moves, as detailed below. We used the Metropolis algorithm to accept or reject moves based on the change in energy associated with the move. Contributions to this change in energy may include the bending energy of the DNA, the binding energy of the proteins, and the protein–protein interaction energy, as described by Eqs. S1 and S2.

1.2. DNA Moves. To sample DNA configurations we use a combination of local and nonlocal moves. The local moves include the corner and crankshaft moves commonly used in lattice polymer models (Fig. S1 A and B). We supplement this local move set with a collection of nonlocal moves designed to efficiently simulate dense polymer configurations. Local moves, such as the corner and crankshaft moves, require conformation changes in which polymer sites move from occupied to unoccupied sites. However, when the polymer is highly condensed such moves will frequently be impossible because two DNA sites cannot occupy the same location in space. To address this issue, a set of moves was designed to change the polymer conformation without altering the occupancy of the polymer sites in space. Thus, these moves result in sampling new configurations by rewiring the polymer, as illustrated in Fig. S1 C–E. This long-range move set is described in detail in ref. 1. Importantly, if there is a *parS* site located in the center of the DNA with a larger ParB binding affinity, such a nonlocal polymer move can amount to a relocation of this *parS* site to maintain its genomic position in the center of the DNA. We account for this by including the change in the energies for ParB binding to the DNA that would result from the nonlocal move.

As a consistency check, we confirmed that this combination of local and nonlocal polymer moves yields the well-known equilibrium properties for self-avoiding polymers with varying degrees of self-attraction between polymer sites. Specifically, we confirmed the expected scaling of the radius of gyration with the molecular weight of the polymer and the scaling of the average 3D distance between two sites on the polymer as a function of their distance measured along the polymer chain.

1.3. Protein Moves. We used a variety of protein moves. First, we used a binding move, in which a protein can either bind to an unoccupied DNA site or unbind from an occupied DNA site (Fig. S1F). We control the protein binding properties with a chemical potential for the proteins in solution. We can fix the concentration of proteins in solution by fixing this chemical potential. Alternatively, we can fix the total number of proteins in the system and account for the titration of ParB protein out of the cytoplasm by adjusting this chemical potential when proteins bind to or unbind from the DNA. Second, we used a hop move in which we attempt to move a protein from one site on the DNA to a randomly chosen other site (Fig. S1G). Third, we used a “bond change” move. Even when the DNA configuration and

the proteins are fixed in space, there may be multiple possible bonding configurations. Thus, we also used a bond change move in which all existing protein bonds at a given DNA site are broken and a random selection is made from all possible local bond configurations (Fig. S1H).

2. Binding Isotherms

To further quantify the different behavior of the five models described in the main text, we calculated binding isotherms, with the same interaction parameters as in Fig. 2. The binding curves describe the average fraction of binding sites occupied by ParB proteins on a DNA chain (without a *parS* site) in equilibrium with a solution of ParB dimers at fixed temperature. The proteins are assumed to be in excess and the free protein concentration was varied. The first four models all exhibited sigmoidal, Hill-like binding curves (Fig. S2). By contrast, the spreading and bridging model yielded a sharp vertical jump.

One expects the steepness of the binding curve to increase with the degree of cooperativity of protein binding. A vertical jump, as we observed for the spreading and bridging model, indicates cooperativity among all proteins binding to the available sites on the DNA. This behavior corresponds to a first-order phase transition in the large- N limit, similar to the transition between vapor and water at atmospheric pressures. What are the implications of such an all-or-nothing condensation transition for the cell? At very low ParB concentration, the ParB proteins would be dissolved, that is, dispersed throughout the cytoplasm and scattered on the DNA. By contrast, under normal conditions the ParB concentration would exceed this threshold, and the majority of ParB proteins would condense to form large protein-DNA clusters, limited in size by the total number of available ParB proteins or, possibly, by constraints on the condensation of the DNA, such as tethering.

3. Triangular Binding Profiles

For a sufficiently strong *parS* site, we observed that the vast majority of ParB proteins localize around the *parS* site in the spreading and bridging model. In this case, the average ParB binding profile exhibits a triangular shape centered around the *parS* site: The ParB binding probability is ~ 1 at the *parS* site, and decays to nearly 0 a distance M_c from the *parS* site, where M_c is the total number of proteins in the cluster (Fig. 3A).

The triangular ParB binding profile is consistent with the formation of a single large ParB-DNA cluster; the cluster as a whole can shift over a range of sites along the DNA as long as it overlaps with the *parS* site. The triangular shape of the profile can be understood by noting that the probability for ParB binding at a particular site is proportional to the number of states in which the ParB-DNA cluster overlaps with that site. The number of states in which a cluster with M_c proteins overlaps with a given site at a genomic distance s from the *parS* site scales as $\sim M_c - s$ (when $s \leq M_c$), thus giving rise to a triangular binding ParB profile. Note that significant deviations from this triangular binding profile occur in parameter regimes for which the cluster includes multiple unoccupied DNA loops (Fig. 3C, *Inset*).

4. Loop Statistics

In this section we describe the statistics of DNA loops that emanate from the ParB-DNA cluster. Fig. S3 shows the loop-size distribution for the spreading and bridging model. The probability $p(s)$ to form a loop of size s exhibits a power-law decay. These results can be understood by noting that this distribution is determined by the entropic cost of forming a loop of size s : $S_{\text{loop}}(s) \sim k_B \ln(s^{3\nu})$ (2), where $\nu \approx 0.588$ is the Flory exponent (3). This dependence of the loop entropy on loop size implies a universal loop-size distribution, $p(s) \sim s^{-3\nu}$, independent of clus-

ter size M_c or the spreading bond strength J_s , consistent with our simulation results (Fig. S3).

The loop-size distribution implies that the average loop length $\langle \ell \rangle$ is insensitive to protein-protein bond strengths or ParB cluster size, in agreement with our simulation results shown in Fig. S4A. However, the average loop length does depend on the total length of the DNA:

$$\langle \ell \rangle = \frac{(3\nu - 1)(N^{3\nu} - N^2)}{(3\nu - 2)(N^{3\nu} - N)}. \quad [\text{S3}]$$

For a DNA strand with $N = 1,000$ binding sites, we expect $\langle \ell \rangle \approx 13$, consistent with our simulations (Fig. S4A).

4.1. Scaling Behavior of the Number of Loops. Because loops are formed at the cluster surface, the average number of loops $\langle N_{\text{loops}} \rangle$ should increase with the surface area of the cluster, which increases with cluster size. In particular, the number of surface sites of a cluster is expected to increase as $M_{\text{surface}} \sim M_c^\alpha$. One expects $\alpha = 2/3$ for a 3D spherical cluster and $\alpha = 1$ for, say, a 1D filamentous cluster, which is all surface. To test this prediction for the surface-area scaling, we defined as a surface site any site occupied by a protein with fewer than the maximal number of six proteins at neighboring sites and measured the average number of such surface sites for different cluster sizes M_c . Indeed, for a model with two spreading and four bridging interactions per protein we found $\alpha = 2/3$, which indicates a compact, spherical cluster (Fig. S5). By contrast, for the spreading and bridging model we found that $M_{\text{surface}} \sim M_c^\alpha$ with an exponent $\alpha \approx 0.9$, indicating a more extended cluster with fractal surface area, as shown in Fig. S5.

The scaling of the surface area with cluster size implies that $\langle N_{\text{loops}} \rangle \sim M_c^\alpha$. In addition, $\langle N_{\text{loops}} \rangle$ should also depend on the spreading bond strength J_s , because the formation of a loop from the ParB-DNA cluster requires breaking a spreading bond. Taking the corresponding Boltzmann factor together with the expected surface area scaling yields

$$\langle N_{\text{loops}} \rangle \sim \exp(-J_s/k_B T) M_c^\alpha. \quad [\text{S4}]$$

To test this prediction within the spreading and bridging model, we determined the average number of loops in a ParB cluster as a function of cluster size M_c and the strength of the spreading bond J_s , as shown in Fig. S4B. Indeed, we find that $\langle N_{\text{loops}} \rangle \sim M_c^\alpha$ with an exponent $\alpha \approx 0.9$. Furthermore, we observe that $\langle N_{\text{loops}} \rangle \sim \exp(-J_s/k_B T)$, consistent with the theoretical prediction.

4.2. Scaling Behavior of Gene-Silencing Profiles. How can we understand the scaling of the ParB binding profile (Fig. 4B and D), and how is it affected by looping? Because the formation of a loop requires breaking a spreading bond, reducing the magnitude of J_s increases the average number of loops. Indeed, weakening J_s from $-8k_B T$ to $-6k_B T$ causes the ParB binding profile to deviate more strongly from a triangular shape owing to an increase in DNA looping, as shown in Fig. 4A. Surprisingly, these binding profiles can still be collapsed by scaling with cluster mass M_c . In the absence of loops, the average footprint L of the ParB-DNA cluster (Fig. 4B) is simply set by total number of proteins in the cluster M_c . However, the presence of loops extending from the cluster will increase the average 1D range of binding sites that contribute to the 3D cluster. Thus, the ParB cluster footprint L has two contributions: a contribution from the total number of proteins in the cluster M_c and a contribution from the accumulated length in loops. However, because the total loop length only scales sublinearly with cluster mass (total loop length $\sim M_c^\alpha$ with $\alpha \approx 0.9$), the loop contribution becomes

negligible for large enough M_c . Indeed, for all our simulations we find that the DNA footprint L of the ParB cluster scales approximately linearly with the number of proteins in the cluster, despite the presence of loops. It is this property of the modeled ParB cluster that allows for the simple data collapse shown in Fig. 4B and D; this collapse demonstrates that the DNA exposure at a specific location is governed by the ratio of the footprint L of the ParB–DNA complex to the genomic distance from the *parS* site.

To further investigate the effect of looping on DNA exposure, we generated three datasets using the spreading and bridging model with the strength of a spreading bond set to $J_S = -8k_B T$, $-6k_B T$, and $-4k_B T$; we expect the largest effects due to looping for $J_S = -4k_B T$. Interestingly, the DNA exposure at *parS*-proximal sites is dramatically enhanced and exhibits a weaker decay with (ParB expression level)/(*parS* distance) for the system with abundant looping ($J_S = -4k_B T$), as shown in Fig. 4D. The results with $J_S = -8k_B T$ with sparse looping predict that the exposure depends on (ParB expression level)/(*parS* distance) with a power law exponent of -1 (Fig. 4D), consistent with the behavior observed in P1 plasmids (Fig. 4F) (4).

5. Roadblock Simulations

In the main text we investigated how the ParB binding profile for a system with a roadblock placed immediately to the left of a *parS* site depends on the size of the ParB cluster M_c (Fig. 3D). Here we provide a scaling argument leading to a prediction that the probability p_{left} to the left of the roadblock scales as $p_{\text{left}} \sim M_c^\alpha$.

In the absence of a roadblock, the ParB–DNA cluster can shift up to $\sim M_c$ binding sites to the left while still overlapping with the *parS* site. Thus, the probability for ParB binding close to the left of *parS* site is proportional to these $\sim M_c$ possible configurations. However, in the presence of a roadblock, we should only consider the fraction $\langle \ell \rangle N_{\text{loops}} / M_c \sim M_c^{\alpha-1}$ of the configurations for which a loop is actually formed at the position of roadblock. Thus, $p_{\text{left}}(d) \sim M_c M_c^{\alpha-1} = M_c^\alpha$ (for $d < M_c$), consistent with the results shown in Fig. 3D, *Inset* for the spreading and bridging model ($\alpha = 0.9$). As a consistency check, we considered a model in which each protein can form two spreading bonds and four bridging bonds. Owing to the larger number of bridging bonds in this model, we find that the clusters become more spherical, resulting in a lower value of the loop exponent $\alpha \approx 2/3$ (Figs. S5 and S6). Thus, for this model, we would expect that $p_{\text{left}} \sim M_c^{2/3}$, which is consistent with the simulation results shown in Fig. S6, *Inset*.

6. DNA Confinement

To investigate the effects of DNA confinement on the formation of protein–DNA clusters, we performed Monte Carlo simulations with confined DNA for the bridging model, the spreading or bridging model, and the spreading and bridging model. We used a simple harmonic DNA confinement potential of the following form:

$$V_{\text{conf}} = \frac{1}{2} \sum_{i=1}^N k_{\text{conf}} (r_i - R_0)^2 \theta(r_i - r_0). \quad [\text{S5}]$$

Here, k_{conf} is the strength of the confinement potential, r_i measures the distance from site i to the origin, and R_0 is the confinement radius. The Heaviside step-function $\theta(r)$ ensures that the confinement potential only contributes when $r \geq R_0$. Using $k_{\text{conf}} = 2k_B T$ and $R_0 = 7$, roughly 2.5 times less than the equilibrium radius of gyration without confinement for a chain of length $N = 500$, we performed simulations to determine the cluster-size distribution and the ParB binding profiles for the bridging model, the spreading or bridging model, and the spreading and bridging model in the presence of DNA confinement. The results of the simulations with confinement (Fig. S7A and B) are qualitatively similar to the

results for the cluster-size distribution and ParB binding profiles in the absence of DNA confinement (Figs. 2F and 3A).

7. Nucleation Kinetics

ParB foci do not form in cells where the *parS* site is absent (Fig. 5). This suggests that *parS* is crucial for the nucleation of ParB complexes on the DNA. In the main text we showed that a *parS*-induced reduction of the nucleation barrier ΔF_{parS} is required to satisfy the condition $\exp(-\Delta F_{\text{parS}}) < N$ for simple nucleation to occur reliably at a single *parS* site and not at any of the N nonspecific sites. Thus, this would require $\Delta F_{\text{parS}} \approx -12k_B T$. By contrast, if nucleation were to occur through the formation of a bond between two ParB dimers bound to distinct, but nearby, *parS* sites (as occurs during replication), this condition would become $\exp(-\Delta F_{\text{parS}}) < N^2$, requiring roughly $\Delta F_{\text{parS}} \approx -6k_B T$. However, we showed experimentally that a single *parS* site per cell is sufficient for the formation of ParB foci on the DNA (Fig. 5), ruling out the scenario that ParB foci formation requires the presence of two nearby *parS* sites.

8. Equilibrium Properties of a Two-State Model for ParB

In *Discussion* we describe a two-state model in which a ParB dimer can be in two different states (closed and inert, or open and capable of forming interactions), and we focused on the implications that such a two-state model might have on the kinetics of cluster nucleation. Interestingly, as discussed below, the equilibrium behavior of such a model is not qualitatively different from the one-state models (in which the protein is always in the open state) discussed in *Results* in the main text.

For a two-state model for ParB, suppose the energy of the open state is E_0 higher than that of the closed state, except when the protein is bound at a *parS* site, in which case the open and closed states have, say, the same energy. To form clusters, the proteins must be in the open state. Thus, in this two-state model the energy of a cluster is increased by $\Delta E = (M_c - 1)E_0$ relative to the energy of the cluster in the one-state model if the cluster is formed around a *parS* site (localized), and $\Delta E = M_c E_0$ if the cluster is formed with all of the proteins bound to nonspecific sites (delocalized). This has two implications:

- The cluster has an additional energetic contribution to localization at a *parS* site within a two-state model, because the energy of a cluster localized at a *parS* site (Eq. 2) is reduced by E_0 relative to a nonlocalized cluster, i.e., $\Delta F_{\text{parS}} = \Delta \epsilon_{\text{parS}} - E_0$.
- In the two-state model, a cluster is destabilized by an energy $M_c E_0$. Importantly, the total cohesive energy E_{coh} from the (spreading and bridging) bonds that hold the cluster together also scale with M_c in the same way, that is, $E_{\text{coh}} \sim (3/2)J M_c$. For the sake of simplicity we can assume that the spreading and bridging bonds have the same strength ($J_B = J_S = J$). Thus, the cluster energies per protein in the two-state model are reduced by an amount E_0 per protein compared with the one-state model. Put simply, the cluster energies of the two-state model with bond strength J are the same as the cluster energies of a one-state model with bond strength $J + 2E_0/3$. Thus, this essentially amounts to the same model, but with a redefined bond strength.

From this analysis, we conclude that the two-state model will have the same equilibrium properties at the one-state model, including cluster sizes and silencing profiles, and will still require three interaction domains per protein to stabilize the clusters.

9. Variety of Models with Spreading and/or Bridging Interactions Do Not Result in the Formation of ParB Condensates

One could envision other simple models with just a few interaction domains that are not discussed in *Results* in the main

text. These other options are not explicitly addressed either because they will show equivalent behavior to the models we did discuss or because one can argue a priori that they are not consistent with experiments.

Models with multiple bridging domains per protein could possibly form condensates, but if they do not also have two spreading domains such models will not be capable of explaining the roadblock experiments. We find that if the proteins have two spreading domains, only one additional bridging domain is required to form ParB condensates (as in the spreading and bridging model).

A model with a single spreading and a single bridging domain is part of a larger class of models (to which the bridging model and the spreading or bridging models also belong) that fails to result in a condensed ParB–DNA complex. At root, this type of model with only two binding sites per ParB does not allow for a 3D surface tension, which is required to drive condensation. In all such models, proteins can interact to form small clusters in which all of the bonds (or all but two) are satisfied, and, because of entropy, the formation of many small clusters is more favorable than formation of a single large cluster.

SI Materials and Methods

Live Cell Imaging and Image Analysis. Fluorescence microscopy was performed with a Nikon Ti microscope equipped with Plan Apo 100 \times /1.4 N.A. phase contrast oil objective and a CoolSnapHQ2 CCD camera (Photometrics). Cells were immobilized on 2% (vol/vol) agarose pads containing growth media. DNA was visualized with DAPI (2 μ g/mL) (Molecular Probes). Images were cropped and adjusted using MetaMorph software (Molecular Devices). Final figure preparation was performed in Adobe Illustrator.

Immunoblot Analysis. Cells were grown exponentially at 30 °C and shifted to 42 °C for 1.5 h. Then IPTG was added to a final concentration of 0.5 mM to induce the expression of GFP-Spo0J and GFP-Spo0J (G77S) at 42 °C. Samples were harvested immediately before and 0.5 h after induction. Cultures were diluted with prewarmed media when needed to make sure the OD₆₀₀ of

the cultures was below 0.6 at all times. Whole-cell lysates were prepared as described (5). Samples were heated for 5 min at 80 °C before loading. Equivalent loading was based on OD₆₀₀ at the time of harvest. Proteins were separated by SDS/PAGE on 12.5% polyacrylamide gels, electroblotted onto Immobilon-P membranes (Millipore), and blocked in 5% nonfat milk in PBS containing 0.5% Tween-20. The blocked membranes were probed with anti-GFP (1:10,000) (6) or anti-SigA (1:10,000) (7), diluted into 3% BSA in PBS–0.05% Tween-20. Primary antibodies were detected using horseradish peroxidase-conjugated goat anti-rabbit IgG (BioRad) and the Super Signal chemiluminescence reagent as described by the manufacturer (Pierce).

Plasmid Construction. pNS040 [*amyE::parS* (*kan*)] was constructed by annealing oTD40 and oTD41 and inserting this fragment into pER82 [*amyE::kan*] at BamHI site.

pWX589 [*yvbJ::Pspank-gfp-spo0J* (Δ *parS*) (*cat*)] was generated by cloning *gfp-spo0J* (Δ *parS*) [PCR-amplified from pKM256 (8) using oWX998 and oWX999 and digested with XmaI and NheI] into pER134 [*yvbJ::Pspank* (*cat*)] between XmaI and NheI.

pWX624 [*yvbJ::Pspank-gfp-spo0J* (G77S) (Δ *parS*) (*cat*)] was generated from pWX589 by site-directed mutagenesis using primers oTG049F and oTG049R.

Strain Construction. Δ *spo0J* (Δ *parS*)::*spec* was obtained by direct transformation of an isothermal assembly product (9) into the wild-type strain PY79. The isothermal assembly reaction contained three PCR fragments: (i) a 0.7-kb region upstream the *spo0J* gene (amplified from wild-type genomic DNA using primers oWX507 and oWX886); (ii) *loxP-spec-loxP* cassette (amplified from pWX466 using primers oWX438 and oWX439), and (iii) a 1.6-kb region downstream of the *spo0J* gene (amplified from wild-type genomic DNA using primers oWX887 and oWX888). pWX466 contains a *loxP-spec-loxP* cassette. The resulting construct was sequenced across the *soj-spo0J* region using primers oWX507 and oWX508. This construct deletes the *parS* site that is within the *spo0J* gene, whereas Δ *spo0J*::*spec* in AG1468 (10) retains the *parS* site at this locus.

- Deutsch JM (1997) Long range moves for high density polymer simulations. *J Chem Phys* 106(21):8849–8854.
- de Gennes P-G (1979) *Scaling Concepts in Polymer Physics* (Cornell Univ Press, Ithaca, NY).
- Le Guillou JC, Zinn-Justin J (1977) Critical exponents for the n-vector model in three dimensions from field theory. *Phys Rev Lett* 39(2):95.
- Rodionov O, Lobočka M, Yarmolinsky M (1999) Silencing of genes flanking the P1 plasmid centromere. *Science* 283(5401):546–549.
- Doan T, Rudner DZ (2007) Perturbations to engulfment trigger a degradative response that prevents cell-cell signalling during sporulation in *Bacillus subtilis*. *Mol Microbiol* 64(2):500–511.
- Rudner DZ, Fawcett P, Losick R (1999) A family of membrane-embedded metalloproteases involved in regulated proteolysis of membrane-associated transcription factors. *Proc Natl Acad Sci USA* 96(26):14765–14770.

- Fujita M (2000) Temporal and selective association of multiple sigma factors with RNA polymerase during sporulation in *Bacillus subtilis*. *Genes Cells* 5(2):79–88.
- Sullivan NL, Marquis KA, Rudner DZ (2009) Recruitment of SMC by ParB-parS organizes the origin region and promotes efficient chromosome segregation. *Cell* 137(4):697–707.
- Gibson DG, et al. (2009) Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat Methods* 6(5):343–345.
- Ireton K, Gunther NW, 4th, Grossman AD (1994) *spo0J* is required for normal chromosome segregation as well as the initiation of sporulation in *Bacillus subtilis*. *J Bacteriol* 176(17):5320–5329.

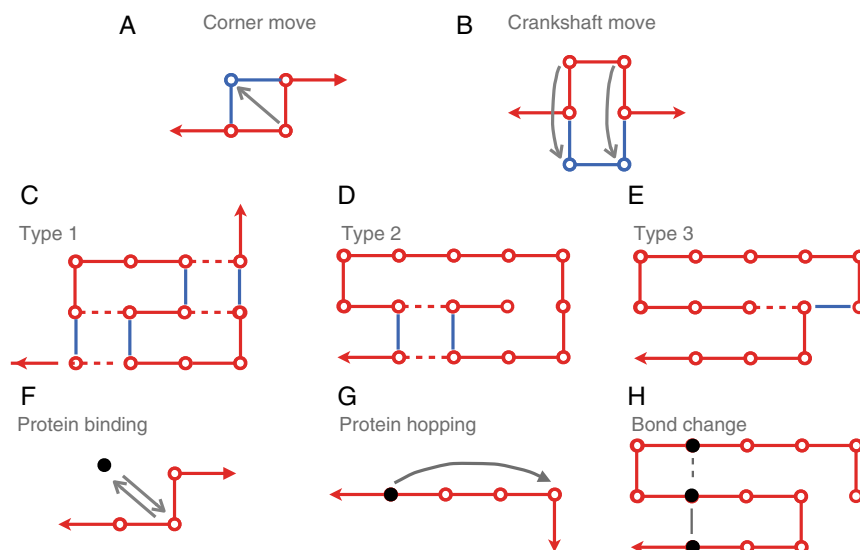


Fig. S1. Monte Carlo move set. (A and B) Local corner and crankshaft polymer moves. The nonlocal polymer moves (C–E) are described in detail in ref. 1. Dashed red polymer connections are broken and solid blue connections are formed. Arrows indicate continuation of the polymer chain. (F–H) Depiction of the protein moves, including binding moves, hopping moves, and bond change moves.

1. Deutsch JM (1997) Long range moves for high density polymer simulations. *J Chem Phys* 106(21):8849–8854.

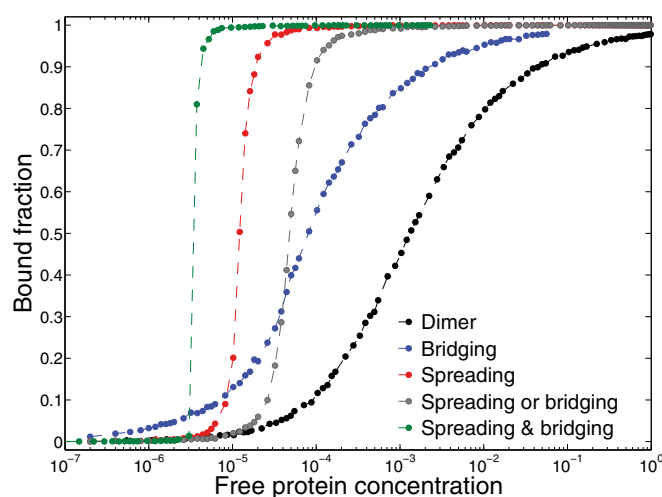


Fig. S2. Binding isotherms. The fraction of DNA sites bound by proteins as a function of the concentration of free protein dimers in solution (in arbitrary units). The first four models exhibit a sigmoidal, Hill binding curve with varying degrees of cooperativity, whereas the spreading and bridging model exhibits a sharp vertical jump indicative of a first-order phase transition. In all models, the proteins are assumed to be dimerized in solution.

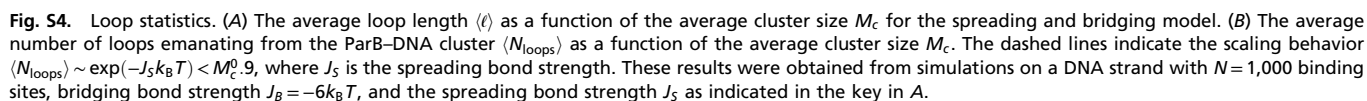
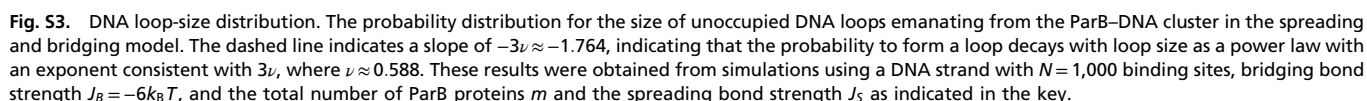


Figure 2 consists of a main plot and an inset. The main plot shows the probability distribution of genomic distance between ParB binding sites for various ParB expression levels. The x-axis is 'Genomic distance' (0 to 1000) and the y-axis is 'Probability' (0 to 1). A vertical line at 500 indicates the expected distance. The curves are color-coded by expression level: 20 (blue), 40 (green), 60 (red), 80 (cyan), 100 (magenta), 150 (yellow), 200 (black), 250 (dashed blue), and 300 (dashed green). The inset shows $A_{eff}(d)$ vs. ParB expression level on a log-log scale, with data points and a dashed line fit.

7 of 11

A 42°C 1.5h no IPTG

	WT <i>parS</i>	<i>parS</i> Δ8	<i>parS</i> Δ8 + 28°:: <i>parS</i>
nucleoid			
GFP-Spo0J			

B 42°C 1.5h no IPTG, +IPTG 0.5h

	WT <i>parS</i>	<i>parS</i> Δ8	<i>parS</i> Δ8 + 28°:: <i>parS</i>
nucleoid			
GFP-Spo0J			

42°C 1.5h no IPTG + IPTG 0.5h

1 2 3 4 1 2 3 4

GFP-Spo0J

►

σ^A

1: *dnaB(ts)*, WT *parS*, *Pspank-gfp-spo0J*
2: *dnaB(ts)*, *parSΔ8*, *Pspank-gfp-spo0J*
3: *dnaB(ts)*, *parSΔ8*, +28'':*parS*, *Pspank-gfp-spo0J*
4: *dnaB(ts)*, *parSΔ8*, +28'':*parS*, *Pspank-gfp-spo0J*(G77S)

Oligos	Sequence	Use
oTD40	gatcaatcagaatgttacacgtgaacaagaaaaac	pNS40
oTD41	gatcgttttctttgttcacgtgtaacattcgatt	pNS40
oTG049F	tatgatattgttgcgagtgaaacggcggttttc	pWX624
oTG049R	gaaaacgccgttcactcgcaacaatatcata	pWX624
oTG049R	gaaaacgccgttcactcgcaacaatatcata	pWX624
oWX438	gaccagggagcactggtcaac	<i>Δspo0J (ΔparS)::spec</i>
oWX439	tccttctgctccctcgctcag	<i>Δspo0J (ΔparS)::spec</i>
oWX507	cgctgctgaattttcaattattttcc	<i>Δspo0J (ΔparS)::spec</i>
oWX508	acccgttgcaaaggctcactggggcgc	<i>Δspo0J (ΔparS)::spec</i>
oWX886	ctgagcgagggagcagaaggatcctttcaaggccttagccattcgagc	<i>Δspo0J (ΔparS)::spec</i>
oWX887	gttgaccagtgctccttggtccgagaatcataaatgaaaaaacattcttc	<i>Δspo0J (ΔparS)::spec</i>
oWX888	agaggtaaactgaatgtctgcaggcc	<i>Δspo0J (ΔparS)::spec</i>
oWX998	aaaCCCGGAcataaggaggaactctatgagtaaagg	pWX589
oWX999	tttGCTAGCagagtgagggaagaacgccttaacc	pWX589