



**HAL**  
open science

# Modeling the formation and positioning of intracellular macromolecular assemblies: Application to bacterial DNA segregation

Jean-Charles Walter

► **To cite this version:**

Jean-Charles Walter. Modeling the formation and positioning of intracellular macromolecular assemblies: Application to bacterial DNA segregation. Biological Physics [physics.bio-ph]. Université de Montpellier, 2020. tel-02893703

**HAL Id: tel-02893703**

**<https://hal.science/tel-02893703>**

Submitted on 5 Nov 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License



Thèse

présentée pour l'obtention de

l'Habilitation à Diriger des Recherches

par **Jean-Charles WALTER**

**Modeling the formation and positioning of  
intracellular macromolecular assemblies:  
Application to bacterial DNA segregation**

Soutenance publique le 24/06/2020

Membres du Jury :

Président :	Andrea PARMEGGIANI	PR (HDR), L2C, Montpellier Univ. & CNRS
Rapporteurs :	Walter KOB	PR (HDR), L2C, Montpellier Univ. & CNRS
	Thomas RISLER	MCF (HDR), Institut Curie, Sorbonne Univ.
	Jean-Marc VICTOR	DR CNRS (HDR), LPTMC, Sorbonne Univ. & CNRS
Examineurs:	Ludovic BERTHIER	DR CNRS (HDR), L2C, Montpellier Univ. & CNRS
	Izaak NERI	Assistant Professor, King's College, London

---

**Laboratoire Charles Coulomb (L2C)  
CNRS & Université Montpellier**







# Contents

<b>1 Protein-DNA interactions</b>	<b>11</b>
1.1 Introduction	11
1.1.1 The ParABS system on plasmid and chromosomes of <i>E. coli</i>	12
1.1.2 Characteristics of ParB foci	13
1.1.3 ChIP-sequencing experiments	15
1.1.4 Evidences of the metastability of the ParBS complexes	16
1.2 The Stochastic Binding model	19
1.2.1 Long range decay of density profiles along DNA	19
1.2.2 The effect of supercoiling on long DNA molecules: towards an estimation of supercoiling <i>in vivo</i>	22
1.3 The Looping and Clustering model	29
1.3.1 The model	30
1.3.2 Protein binding profiles from LC	34
1.3.3 Discussion	38
1.4 Spreading & Bridging model	39
1.4.1 The model LRLG	39
1.4.2 The variational approach	42
1.4.3 Application to biological systems	44
<b>2 Positioning of complexes in liquid-like phase</b>	<b>47</b>
2.1 The Proteophoresis model	49
2.1.1 Definition	49
2.1.2 Restoring proteophoresis force	51
2.2 Stability analysis	53
2.2.1 Traveling Wave (TW) <i>Ansatz</i>	53
2.2.2 TW solution	53
2.2.3 Stability criterion	54
2.3 Comparison with experiments	57
2.4 Discussion	61
<b>Bibliographie</b>	<b>77</b>

<b>3</b>	<b>Curriculum Vitæ</b>	<b>79</b>
3.1	Education & Positions . . . . .	79
3.2	Scientific activity . . . . .	79
3.2.1	Research interests in brief . . . . .	79
3.2.2	Research grants . . . . .	80
3.2.3	Main collaborations . . . . .	80
3.3	Scientific communications . . . . .	81
3.3.1	Publication list . . . . .	81
3.3.2	Recent presentations 2015–present . . . . .	84
3.4	Experience of mentoring and teaching . . . . .	86
3.4.1	PhD Students supervision . . . . .	86
3.4.2	Supervision of students . . . . .	86
3.4.3	Teaching experience . . . . .	87
3.5	Various collective activities . . . . .	87
<b>4</b>	<b>Research Interests</b>	<b>89</b>
4.1	Spin systems at and out of equilibrium (2006–2012) . . . . .	89
4.2	Description of polymers and biopolymers . . . . .	90
4.2.1	Physical behavior of DNA Microarrays (2010–2011) . . . . .	90
4.2.2	Winding and zipping properties of a polymer (2011–2016) . . . . .	91
4.3	Collective behaviour of protein assemblies in the cell (2014–present) . . . . .	92
4.3.1	Organisation of the genome during active bacterial DNA seg- regation . . . . .	92
4.3.2	Modeling the cytoplasmic active transport . . . . .	92

# Introduction

Bacteria, also called prokaryotes, are autonomously replicating unicellular organisms where all the processes take place in the same inner compartment, contrary to eukaryotes where the genome is separated from the cytoplasm by a membrane. Among the first living forms appearing on Earth, bacteria are involved in almost all habitats: in water, soil, earth's crust, even the more hostile like acidic hot springs, and deep down the Mariana Trench. They live in symbiotic relationships with plants and animals (for the latter especially in the gut), which thus depend on bacteria for their survival. On the other hand, tuberculosis is responsible each year for two millions of human death essentially in Africa.

Any kind of cell is confronted to the remarkable challenge of large information storage in the form of the DNA molecules. The bacterial chromosome, a looped molecule of linear size of the order of the millimeter, needs to fit into a cell of micrometer size, i.e., three orders of magnitude smaller. On top of this remarkable compaction, the genome is positioned precisely inside the cell and displays a hierarchical organization involving namely Nucleated Associated Proteins (NAPs) and supercoiling. Not only DNA has to be compacted, but, at the same time, it needs to be easily accessible in the dense intracellular environment to adjust to the actual needs of the cell. These constraints are put under high evolutionary stress with bacteria: these cellular functions have to be efficiently executed in the competitive environment constituted by other bacterial strains. In Fig.1A, we show a population of *E. coli* in exponential growth.

The perpetuation of a species relies on the inheritance of the genomic information to the offsprings. A few main steps of the processes of genetic inheritance are sketched in Fig.1B and described in Ref.[125] in more details. Firstly, DNA is replicated: the DNA double helix is opened, the two single strands are read by a processing RNA polymerase in order to produce two identical DNA molecules. Secondly, DNA is segregated: the replicated DNA molecules are separated and positioned at either side of the cell division plane to ensure faithful DNA inheritance. Finally, cell division consists in the constriction of a biopolymer at midcell position to literally cut the mother cell into two daughter cells. The two last examples share a same feature: the importance of the physical positioning along cell axis, whether it is chromosomes at quarter-cell position or the constricting polymer at mid-cell position.

These biological processes, among many others, need a strong, local, concentration of proteins to catalyze biochemical reactions. One way to realize a fast reorganization of intracellular proteins without compartmentalizing them with a membrane are *phase transitions*. Intracellular phase transitions have been first shown with P-granules (RNA

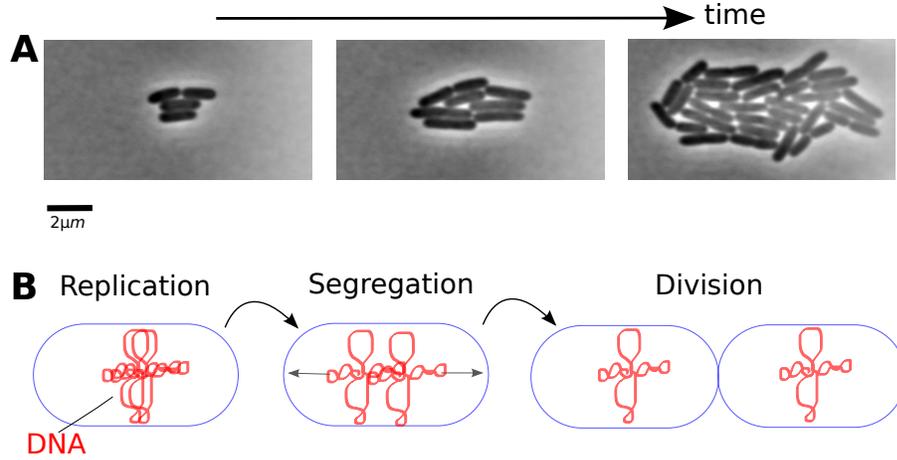


Figure 1: *Main steps of the bacterial cell cycle.* (A) Pictures of a population of bacteria *E. coli* at different times during an exponential growth (fluorescence microscopy, credit: J. Rech). (B) Some of the main biological processes during bacterial cell cycle: DNA replication, DNA segregation and cell division. This thesis is interested in the process of DNA segregation, prototypical system combining intracellular phase transition and equipositioning of replicated DNA molecules along the cell axis.

and proteins-containing bodies) in eukaryotes in *Caenorhabditis elegans* [21]. Since then, phase transitions have been shown in many other biological processes both in eukaryotes and, later on, in prokaryotes. Experimental tools need a much higher precision to probe the behaviour of protein self-assemblies in bacteria (much smaller cells than eukaryotes), namely with super-resolution microscopy at the scale of a single nanometric protein.

This thesis is mainly interested in the theoretical description of bacterial DNA segregation, which offers a prototypical system of protein phase transition, as well as positioning of macromolecular complexes in the cell cytoplasm. The prevalent bacterial partition system, so-called ParABS, is sketched in Fig. 2. ParABS is a widely conserved mechanism: it is used by most of the known bacteria species to segregate their genome (it has even been detected in some archaea). It is thus one of the most ancient liquid-phase separation system known in nature. From the biophysical modeling point of view, ParABS provides a model system constituted only of three interacting components: two proteins ParA and ParB, as well as the centromere-like sequence *parS*.

The theoretical work presented in this thesis has benefited from collaborations with experimentalists, experts in the field of molecular microbiology and super-resolution imaging at the single protein resolution. It was possible to have access to first hand accurate data and to design dedicated experiments to check the predictions of our models. We have combined two powerful experimental techniques giving access to complementary informations. Microscopy techniques (Super-resolution (PALM) and

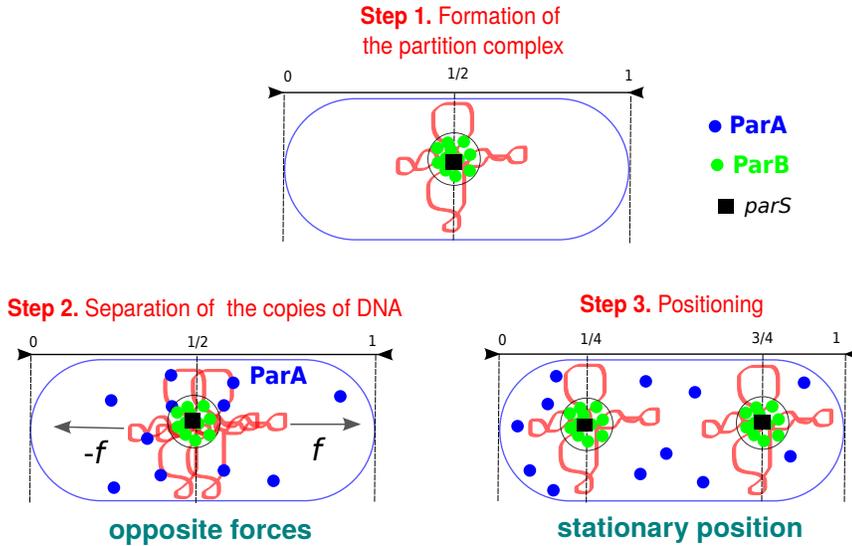


Figure 2: *The three main steps of the segregation mechanism by ParABS. 1.* A few ParB proteins bind to the specific sequence *parS*. These few proteins serve as a nucleation seed attracting a pool of a few hundreds of ParB, thus creating a region of high concentration ( $\sim 10\mu M$ ). This ParBS assembly is called the partition complex. *2.* The high concentration of ParB in the ParBS complex catalyzes the ATP hydrolysis of ParA. These interactions produce opposite forces that separate the two replicated copies of DNA. *3.* Once the two complexes are equipositioned along the cell axis, they remain at their stable position, see Fig.3 for real imaging. The goal of this thesis is to elucidate the molecular biophysical interactions involved in these three steps.

epifluorescence) give access to spatial position and size of the complex ParBS inside the cell [112, 144], at the single protein resolution [81], as shown in Fig.3. It is thus even possible to estimate the diffusion coefficient of single proteins. Molecular microbiology experiments so-called ChIP-sequencing are fragment protected experiments coupled to high-throughput sequencing. They give access to the linear protein density along DNA [112]. Thus, these techniques give access to spatial and linear density of ParB proteins along DNA, which will allow us to discriminate between different physical models.

The first part of this thesis is concerned by the description of the ParBS complex. We start with a phenomenological approach: the *Stochastic Binding model*. DNA is modeled by a freely-fluctuating polymer inside a sphere attached at *parS*, representing the ParB condensate. The model is in excellent agreement with the long range decay of ParB density along DNA in ChIP-sequencing experiments. This shows the crucial role of thermal fluctuations in the organization of bacterial DNA [112, 32]. We discuss in this context our recent work on the effect of supercoiling on DNA orga-

nization. Secondly, we present a semi-phenomenological approach, the *Looping and Clustering model*, which introduces the notion of interactions between ParB proteins with a spreading interactions. Fitting the ChIP-seq data with the model, we are able to give an estimate of this coupling. Thirdly, we use a microscopic approach where we model ParB on the DNA by a Lattice Gas on a polymer with the *Spreading and Bridging model*. We solve this system with a mapping on a 1D Long Range Lattice Gas and a variational approach: for biological parameters, the system ParBS falls in the metastability region of the phase diagram. This metastability is an efficient biological mechanism to turn on and off clusters of proteins with a small perturbation. Finally, we conclude this first part with recent experiments of our collaborators showing evidences of this phase transition scheme.

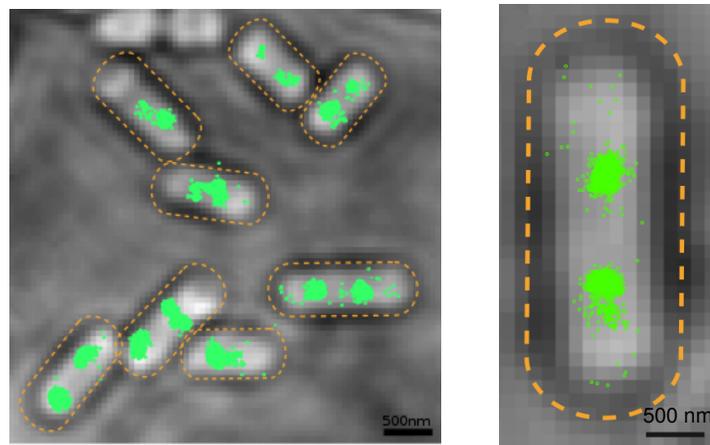


Figure 3: *Super-resolution microscopy experiments (PALM)*. Imaging of the ParB proteins at the single particle resolution in *E. coli* in slow growth condition (each green dot represents a ParB protein; credit: Marcelo Nollmann's team, CBS, Montpellier). We observe the clustering of ParBs around *parS* sequence (DNA is invisible here). These experiments give access to the spatial extent and the positioning of the clusters, as well as an approximate distribution of the diffusion coefficients of single proteins.

In the second part, we consider the dynamics of the segregation of the complexes with model of reaction-diffusion between the ParBS complex and ParA [144]. The complex displays two dynamical regimes: a regime of stability and a regime where it moves according to travelling waves. In the stability regime, we show the existence of two subregimes: a regime where the complexes are equipositioned along the cell axis. We show that the biological system falls right in the equipositioning stability regime (as observed in experiments), close to the dynamical instability. In this way, the segregation is optimized: the complexes are equipositioned, and yet they benefit from a quick segregation due to precursor fluctuations of the dynamical instability. This mechanism is called proteophoresis: a volumetric chemical force inside ParBS due to the coupling to the density field of ParA is driving the complexes to their expected positions in the cell.

# Chapter 1

## Protein-DNA interactions: application to the ParBS complex

### 1.1 Introduction

The confinement of chemical species, such as RNA or proteins, within the cytoplasm is essential for biochemical activities in the cell [60]. Cells compartmentalize the intracellular space using either membrane vesicles or membrane-less organelles. For the latter, cells may employ phase separation of chemical species in order to create localized high density regions in which specific reactions may occur [62, 23]. Such biological phase separation mechanisms often involve polymeric scaffolds like RNA or DNA to bind the chemical species [24, 80, 93, 55, 77, 115].

In the past, bacteria were often viewed as homogeneous systems lacking the complex sub-cellular spatial organization patterns observed in eukaryotes. The advent of powerful labeling and imaging methods has enabled the discovery of a plethora of mechanisms used by bacteria to specifically and precisely localize components, in space and time, within its sub-cellular volume [117, 125]. These mechanisms include pole-to-pole oscillatory systems to define the site of cell division (e.g. MinCDE), dynamic, ATP-driven polymerization to drive cell division and cell growth (e.g. FtsZ, MreB), recognition of cell curvature to localize chemotaxis complexes (e.g. DivIVA [104]), ATP-powered translocation of membrane-bound machines to power cell motility (e.g. Agl-Glt [45]), or nucleoid-bound oscillatory systems to localize chromosomal loci (e.g. ParABS [81]). More recently, it became apparent that many cellular components (e.g. ribosomes, RNA polymerases, P-granules) ([131, 97, 105, 111]) display specific sub-cellular localization patterns, leading to the spatial segregation of biochemical reactions (e.g. translation, transcription, or polyP biosynthesis). Most notably, bacteria are able to achieve this precise sub-cellular compartmentalization without resorting to membrane-enclosed organelles. Recently, important

progress has been made in understanding the formation of membrane-less organelles in eukaryotic cells [62]. A combination of *in vitro* and *in vivo* experiments demonstrated that such compartments are formed by liquid-liquid phase separation (LLPS), a mechanism similar to liquid demixing [62]. It consists of a thermodynamic process through which attractive molecular interactions counterbalance entropy-driven effects. This phenomenon promotes the self-organisation of a condensed phase, in which molecules are close enough from each other to experience their mutual attraction, interfaced with a dilute phase. This mechanism provides advantages such as rapid assembly/disassembly, absence of a breakable membrane, and has been shown to serve fundamental biological processes such as regulation of biochemical reactions, sequestration of toxic factors, or to play the roles of organisation hubs [118]. The first evidence that eukaryotic cells use LLPS came from examination of P granules in *C. elegans* [21]. In this study, Brangwynne et al. observed different key signatures of liquid droplets. They showed that P granules were spherical bodies that could fuse together, drip and wet, and that their internal organisation was dynamic. Since then, many other compartmentalizations have been shown form by LLPS [8]. More recently, it was discovered that bacterial ribonucleoprotein granules assemble into LLPS droplets [4] and that the bacterial cell division protein FtsZ forms condensates when in complex with SlmA [99].

A prominent example in the bacterial realm may be the formation of localized protein-DNA complexes during bacteria DNA segregation due to the *in vivo* ParABS system [26, 112, 81, 144]. This clustering of ParB is essential for the interaction with ParA because high concentration of ParB catalyzes the hydrolysis of ParA needed to split the complexes (see next Chapter). Although the molecular components of this widely conserved segregation machinery have been clearly identified, their dynamical interplay and the mechanism that leads to the condensation of the complexes remain elusive.

After an introduction presenting the different molecular actors (sizes, interactions etc.), the experimental methods used to confront the models, we will present three models giving an explanation of experimental observations at different level: firstly, we introduce the phenomenological approach of the Stochastic Binding model, an intermediate model that supposes a preexisting foci of ParB attached to freely fluctuating polymer (without explaining the mechanism of their formation). The predictive power of this model lies in the prediction of the DNA conformation *in vivo* and gives already a clue on the importance of the polymers loops for the DNA-protein organization. Secondly, the Looping and clustering is a semi-phenomenological model, which introduces the ParB-ParB microscopic interactions, but a strong hypothesis is performed to study the system more easily. Finally we solve the Spreading & Bridging model, the most general microscopic model of DNA-protein interactions for ParB, which offers an unconventional mechanism of phase transition for ParBS, modeled as a Lattice Gas on a fluctuating and fractal substrate.

### 1.1.1 The ParABS system on plasmid and chromosomes of *E. coli*

In Fig.1.1A, we show the ParABS operon. An operon is a genetic unit that is transcribed in one run of the RNA polymerase. It is composed by a promoter (black

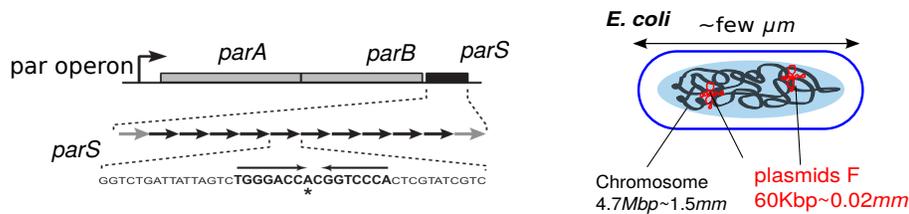


Figure 1.1: (A) The ParABS operon, the genetic unit in which the ParABS system, is encoded and regulated, is composed by a promoter (black arrow), the gene *parB*, and the gene *parA* coding for the proteins ParB and ParA, respectively. The sequence *parS* is constituted, for the F plasmid of *E. coli*, of ten identical sequences of 43bp. (B) Most of the eperiments presented in this thesis have been performed with *E. coli*, displaying a chromosome of  $\sim 4.7\text{Mb}$  fitting in a bacteria of  $\mu\text{m}$  size. Experiments have been performed on the F plasmid of length 60kb, and on modified chromosomes.

arrow) responsible for the initiation of transcription, two genes, *parA* and *parB* coding for the proteins ParB and ParA, respectively. The sequence *parS* is constituted, for the F plasmid of *E. coli*, of ten identical sequences of 43bp, in the middle of which a palindromic sequence of 16bp is specifically recognized by ParB. The affinity of ParB for these sequences is much stronger than non-specific DNA sequence. Most of the eperiments presented in this thesis have been performed with *E. coli* sketched in Fig.1.1B. *E. coli* displays a chromosome of  $\sim 4.7\text{Mb}$  fitting in a bacteria of  $\sim \mu\text{m}$  size. On top of chromosomes, bacteria have auxiliary DNA molecules, so-called plasmids, which can help the organism in case of stressful situation in order to acquire specific skill, like antibiotic resistance. Experiments shown in this thesis have been performed on the F plasmid of length 60kb, and on a modified chromosome (*E. coli* is one among few bacterial strains which do not have ParABS on their chromosomes in Wild Type (WT) conditions.). The advantage of plasmids is that, in case of genetic modifications, we can follow their loss during generations without killing the organism. Moreover, the F plasmid is linked to antibiotic resistance, thus it is possible to track the transmission among a population just by adding antibiotic to the probe.

### 1.1.2 Characteristics of ParB foci

Previous studies have revealed that the partition complex is made of  $\sim 300$  dimers of ParB [1, 16] and  $\sim 10\text{kb}$  of *parS*-proximal DNA [107]. This complex is held together by specific, high-affinity interactions between ParB and *parS*, and by low-affinity interactions of ParB dimers with themselves and surrounding DNA that are essential for the assembly of the partition complex (Fig. 1.2A) [32, 112]. However, technological limitations have thwarted investigation of the structural and dynamic properties of the partition complex. We addressed these limitations by first investigating the shape and size of the partition complex, reasoning that the former should inform us on the role

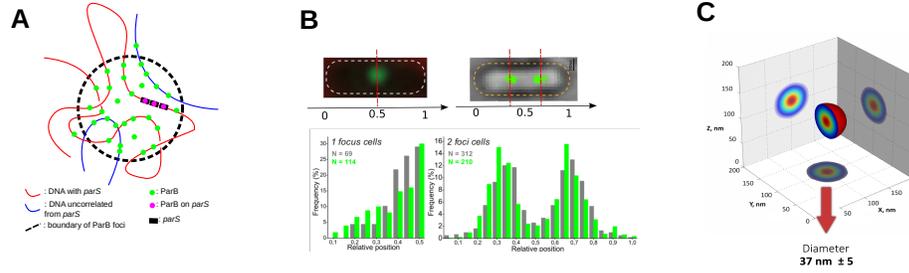


Figure 1.2: *ParB* is confined into spherical nano-condensates. (A) Schematic representation of a *ParB* partition complex [112, 32]. Three types of interactions have been reported: *ParB-parS* (high affinity, 2nM), *ParB*-non-specific DNA (low affinity, 0.3-0.5  $\mu$ M), and *ParB-ParB* dimer (low affinity, 0.3-1  $\mu$ M) [2, 112, 126]. (B) *ParBS* complexes ( $N \approx 990$  *ParBs* per cell) were visualized by PALM at the single molecule resolution. *ParBS* are equipositioned along cell axis (one complex at mid-cell, two complexes are at quarter-cell etc.) [112]. (C) A three-dimensional reconstruction of the *ParB* partition complex. Diameter was estimated with the full width at half maximum.

of the mechanisms involved in the maintenance of the complex cohesion while the latter would enable an estimation of protein concentration within the partition complex. To this aim, our collaborators from CBS, Montpellier, combined Photo-Activated Localisation Microscopy (PALM) [47, 90] with single-particle reconstruction [110]. Partition complexes are positioned near the quarter cell positions (Fig.1.2B) [112], as reported in Ref. [81]. Next, in Fig. 1.2C, we estimated the size of the partition complex by calculating the mean full width at half maximum of the reconstructions obtained from each single class average ( $43 \pm 7$  nm), which is smaller than our previous estimate  $\sim 150$ nm in Ref. [112]. Thus, from the size of the partition complex and the average number of *ParB* in each partition complex ( $\sim 300$  *ParB* dimers) [1, 16], we can estimate a local *ParB* dimer concentration of  $\sim 10$  mM. Remarkably, this extremely high concentration is comparable to that of the total protein concentration in the bacterial cytoplasm ( $\sim 10$ mM, for proteins of the same size as *ParB*) [41], suggesting that *ParB* dimers occupy a volume fraction of 25-30%, and most of the total protein volume within a partition complex. *ParB* dimers interact together with sub- $\mu$ M affinities (0.3-1  $\mu$ M) [112, 126], thus we expect virtually all *ParB* dimers within a partition complex to be interacting with another *ParB* dimer. Finally, these estimations predict a mean intermolecular distance between *ParB* dimers of  $\sim 3$  nm, comparable to the size of a *ParB* dimer itself ( $\sim 3$ nm, assuming its interaction volume to be a sphere) [114].

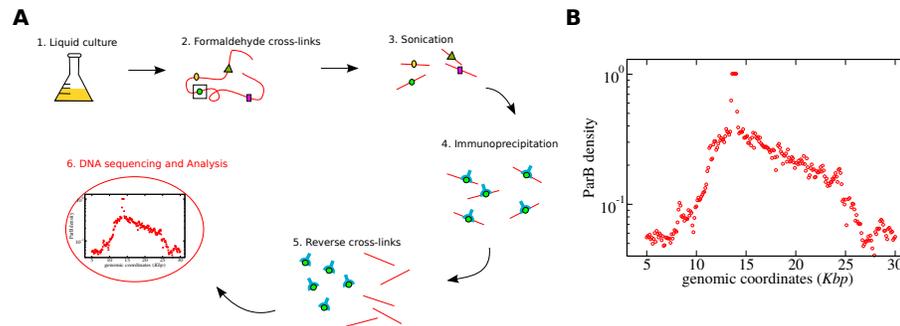


Figure 1.3: *Sketch the experimental protocol of ChIP-seq.* (A) ChIP-seq is a cross-linking experiment that gives access to the local density of protein (here ParB) on DNA. The cell is lysed after cross-linking of protein-DNA contact with formaldehyde. DNA is sonicated in fragment of  $\sim 150\text{bp}$ , which are immunoprecipitated by an antibody displaying affinity with the protein ParB. The fragment are sequenced after reverse cross-linking and a count of the fragments along the genome gives access to the local density of proteins *in vivo*. (B) Typical ChIP-seq profile for the F plasmid of *E. coli*.

### 1.1.3 ChIP-sequencing experiments

On top of superresolution microscopy data, an important set of experimental data is obtained by ChIP-sequencing experiments. These experiments are performed by my collaborators at LMGM, Toulouse. These experiments give information on the linear density along the DNA molecules. Thus we expect these density profiles to contain the information on the probability density function of the polymer. Experiments have been performed essentially on the F plasmid of *E. coli*, and its modified chromosome. Some experiments have been performed on the chromosome I of *Vibrio cholerae*. The principle of ChIP-seq experiments is shown in Fig.1.3(A). These are protected fragment experiments allowing to count the number of proteins-DNA along the genome. Eventually, these experiments give access to the linear density of proteins along the genome. These profiles are directly impacted by the conformation of intracellular DNA, and are thus complementary to the microscopy experiments. In Fig.1.3(B), we show a typical density profile of ParB along the F plasmid of *E. coli*. The *parS* sequence is saturated (density equal to one) due to the high affinity of ParB for *parS*. There is a drop in a few hundreds of bp as soon as we step out of the specific binding region. finally, we see a slower decrease over  $\sim 10 - 15\text{kb}$  on the right side of *parS*, and over  $\sim 5\text{kb}$  on the left side of *parS*. The left side displays a nucleoprotein complex hindering the binding of ParB [112, 32]. This modeling will be performed on the right side for which the decay of the signal is deeper.

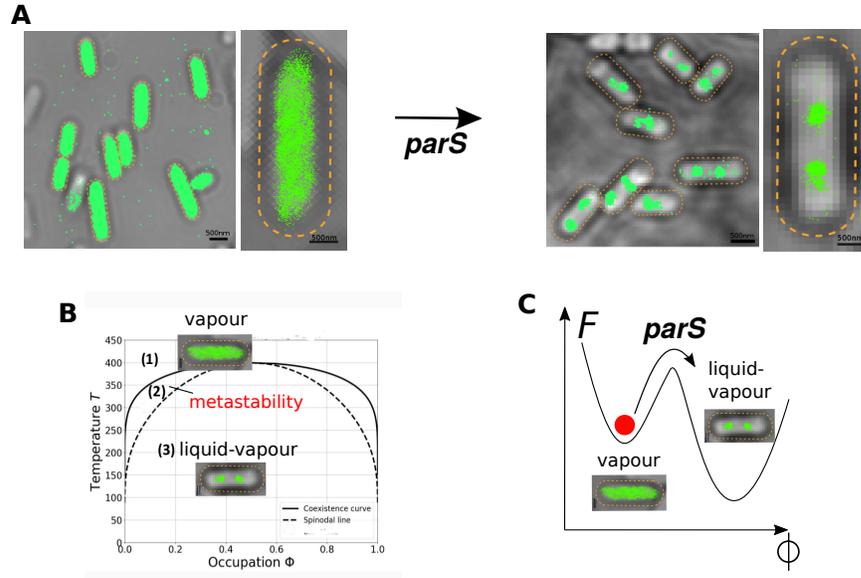


Figure 1.4: (A) Super-resolution microscopy (PALM) of a population of *E. coli* with the F plasmid. (left) No *parS* sites on plasmids: ParBs are spread inside the cell; (right) Same bacteria but with *parS* restored: ParBs condensate in droplets around *parS*. (B) A small perturbation leading to a global reorganization of proteins is a feature of metastable system. The phase diagram of the liquid-vapor transition is shown. (C) Sketch of the expected free-energy  $F$  of our scenario: in the metastable regime, without *parS*, the system is trapped in a long-lived vapor metastable state compared to the cell cycle. Adding *parS* accelerates the kinetics towards the minimum. [use fig presentation]

#### 1.1.4 Evidences of the metastability of the ParBS complexes

The formation of ParBS complexes is accelerated by *parS*. In Fig.1.4A, we show evidences from microscopy experiments PALM of the sensitivity of the formation of ParBS with respect to *parS*. On the left, a population of *E. coli* display a modified F plasmid with two functional proteins ParA and ParB, but without the *parS* sequence. Then we observe that ParB is homogeneously distributed inside the cell. On the right, the sequence *parS* is restored, and the condensation of ParBS occurs. This high local sensitivity of a system with respect to a small perturbation is reminiscent to a mechanism of metastability. In Fig.1.4B, we show the typical phase diagram of a Lattice Gas as a function of the two variables: the occupation  $\phi$  of a site and the absolute temperature  $T$ . The vapour state is governed by entropy and is characterized by a homogeneous distribution of particles inside the volume. The liquid-vapour coexistence regime is characterized by the formation of high density region governed by entropy (droplets) in equilibrium with particles in the gas state. The metastability region be-

tween the coexistence and the spinodal lines is a regime where the vapour state can exist as a metastable state, even though it is not the minimum of the free-energy. In Fig.1.4C, we plot a sketch of the free energy  $\mathcal{F}$  as a function of the occupation  $\phi$  in the metastable regime: the vapour phase is a local minimum where, without *parS*, the system can get trapped for a long time (compared to the cell cycle of the bacteria). We expect the sequence *parS* to decrease the height of the free energy barrier between the metastable state and the minimum of the free energy. Thus the kinetics is accelerated, because the smaller the free energy barrier the more likely a thermal fluctuation will lead to overcome this barrier. The sequence *parS* acts thus as a nucleation seed.

*ParS, and low- and high-affinity ParB interactions are sufficient for phase separation.* We performed a thermodynamic analysis and Monte Carlo simulations to find the minimal set of assumptions that would be necessary to reproduce an equilibrium between liquid- and gas-like phases. For this, we considered a minimalistic Lattice Gas model that shows in mean field theory a generic Liquid-Liquid Phase Separation (LLPS) diagram [54]. The ParB proteins are modeled by particles with a diameter  $a = 5\text{nm}$ , which are located on the nodes of a simple cubic lattice representing segments of the nucleoid on which ParB is able to bind to or unbind from. The distance between nearest-neighbour sites was as large as the size of a particle, i.e.,  $5\text{nm}$ . The lattice is chosen to display the dimension of a bacterial nucleoid:  $L_x = L_y = 0.5\mu\text{m}$  and  $L_z = 2\mu$  (equivalent to  $L_x = L_y = 100a$  and  $L_z = 400a$ ). The total number of binding sites in the lattice is then  $N_S = 4 \cdot 10^6$ . To match the experimentally determined number of ParB proteins per condensate, the particle number is fixed at 300. Thus, the particle density is very low ( $\rho \sim 10^{-4}$ ), placing the Lattice Gas (LG) model in a configuration where only a first-order phase separation transition can occur. Particles can interact between each other when they are at adjacent sites on the sc lattice (nearest neighbour contact interaction) with a magnitude  $J$ . For the LG model, the total energy  $\epsilon$  of the system is:

$$\epsilon = -J \sum_{\langle i,j \rangle} \phi_i \phi_j + \sum_{i \in \text{parS}} h_i \phi_i, \quad (1.1)$$

where  $\phi_i$  is the occupation variable at site  $i$  taking the value 1 if the site  $i$  is occupied or 0 if not. The sum  $\sum$  runs over the pairs of nearest-neighbour sites  $\langle i, j \rangle$  of the lattice. The F plasmid is represented by a single, static, high affinity sequence of binding sites within this lattice containing a repeat of 10 *parS* sequences (similar to the natural F plasmid). ParB interacts with high affinity  $h_i = 10\text{kT}$  with the *parS* cluster, with low-affinity to other sites in the lattice (reference energy of the system), and with low-affinity  $J$  with itself. The diffusing ParB random relocation defines the order of the time unit of the Monte Carlo simulation as approximately one second: ParB proteins diffusing at  $D \approx 1\mu\text{m}^2/\text{s}$  can move at any point on a  $\sim 1\mu\text{m}$  nucleoid after one second.

For this given density, we estimated the ParB-ParB coupling constant  $J = 4.5\text{kT}$  in order to place the system in the metastable state (cf phase diagram in function of  $J$  in Ref.[54] for details) and to match at best the number of ParB confined in condensate ( $\sim 90\%$ ). The metastable regime is obtained numerically with Monte Carlo simulations by varying  $J$  (increasing and decreasing pathways), thus giving access to the range of coupling for which metastability occurs within the hysteresis cycle of the Energy versus  $J$ .

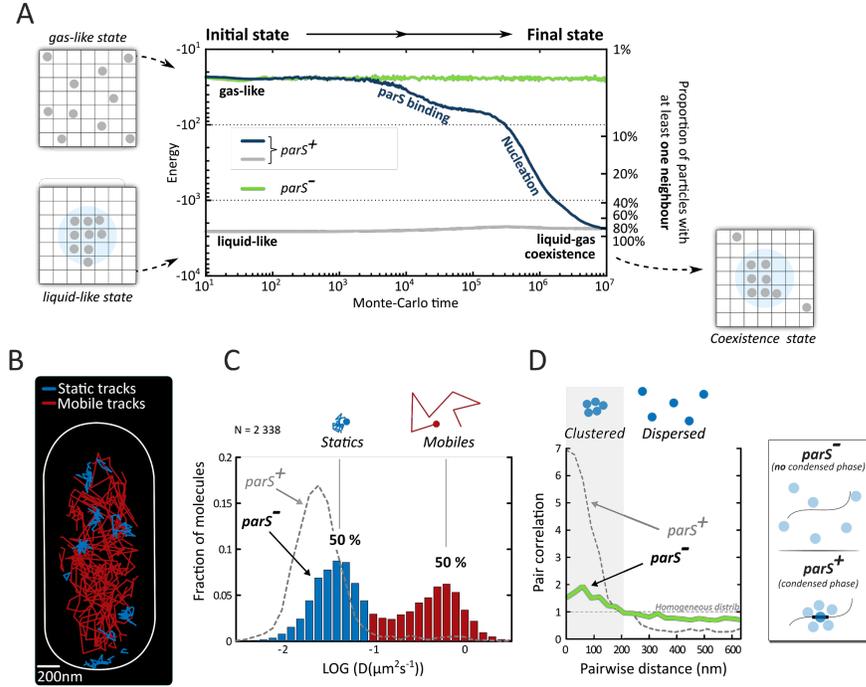


Figure 1.5: Centromeric sequences, and ParB-ParB interactions are required for the nucleation and stability of ParB condensates. (A) The dynamics of 300 ParB proteins is modeled by a LG: free diffusion ( $1 \mu\text{m}^2/\text{s}$ ), weak ParB-ParB dimer interactions ( $J = 4.5kT$ ) and high-affinity ParB-*parS* interactions. Regardless of initial conditions, ParB proteins formed a condensate in the presence of *parS* (gray, blue). In contrast, no condensate is formed in the absence of a nucleation sequence (green). For simplicity, *parS*<sup>+, -</sup> refers to a simulation with (without) *parS*, resp. (B) Representative single-molecule tracking experiment of ParB in absence of *parS*. Static trajectories are shown in blue and mobile trajectories in red. (C) Distribution of apparent diffusion coefficients for static (blue, 50%) and mobile trajectories (red, 50%). The distribution for a strain containing *parS* is shown as a dashed grey line for comparison. (D) Pairwise distance analysis of static trajectories in absence of *parS* (green curve). The expected curve for a homogeneous distribution is shown as a horizontal dashed line. The distribution for a wild-type strain containing *parS* is shown as a dashed grey line for comparison. Schematic representations of single-molecule ParB trajectories in the absence (top) and presence of *parS* (bottom) are shown in the panels on the right.

In Fig.1.5A, the system is initially set at  $t = 0^+$  in the metastable regime. The system is then left to evolve under different initial conditions at  $t = 0^-$  (pure gas-like or liquid-like states) and in the presence or absence of *parS*. The system remains in a gas-like phase in the absence of *parS* (Fig.1.5A, green curve). In contrast, in the presence of *parS*, the system displays several transitions that involved *parS* binding, nucleation, and stable co-existence of liquid- and gas-like phases (Fig.1.5A, blue curve). The system evolved towards the same endpoint when the initial condition are a pure liquid-like state (Fig.1.5A, gray curve). At this endpoint, 80% of the molecules are in a liquid-phase and the remainder in a gas-like state, mirroring the proportions observed experimentally [54]. These results suggest that the only required elements to reach a stable coexistence between liquid and gas-like ParB phases are: low-affinity interactions between ParB dimers and non-specific DNA, and high-affinity interactions between ParB and the nucleation sequence *parS*. Critically, ParB is unable to form a stable liquid-gas coexistence phase in the absence of *parS* within physiological timescales (Fig.1.5A, green curve). We experimentally tested this observation by performing sptPALM experiments in a strain lacking *parS*. In contrast to our previous results, we frequently observed mobile trajectories (Fig.1.5B-C). In fact, the proportion of this population increased from  $\sim 5\%$  in the wild-type to  $\sim 50\%$  in the absence of *parS*. The persistence of such a large proportion of static trajectories could suggest that ParB is still able to assemble into liquid-like phases, but now in transient or instable condensates (made possible by cryptic sequences for instance). To investigate this possibility, we performed pair-correlation analysis of static trajectories (Fig.1.5D). Clearly, static trajectories clustered together in the presence of *parS* (dashed curve), but were rather homogeneously distributed in the absence of *parS* (green curve). These observations are consistent with previous reports showing that ParB foci are not detectable in the absence of *parS* [43] or with *parS*-binding deficient ParB [112]. Overall, these data fully agree with our simulations and indicate that the centromeric sequence *parS* is essential for the assembly of stable liquid-like ParB condensates.

## 1.2 The Stochastic Binding model

### 1.2.1 Long range decay of density profiles along DNA

We start the description of the ParBS complex by a phenomenological approach, the Stochastic Binding model, which offers a good description of ChIP-seq data and give evidences on the importance of the loops in the bacterial DNA organization [112, 32].

ChIP-sequencing data offer an interesting tool to probe the local organization of ParB along DNA. In the Fig. 1.6A, the red dots are experimental data obtained with the F plasmid of *E. coli*. The origin of genomic coordinate is set at the rightmost *parS* binding site, and we show only the right side of the data (the signal on the left side is perturbed by proteins complexes hindering the binding of ParB). We observe a signal saturated at *parS* (probability density  $P(0) = 1$ ), a fast drop down to  $\approx 0.4$  and a slow decay until  $10kb$ . Two models have been proposed in the past to explain the enrichment of ParB around *parS*: the Spreading model sketched in Fig. 1.6B, the Spreading and Bridging in the strong coupling limit sketched in Fig. 1.6C, and our

model, the Stochastic Binding in Fig. 1.6D.

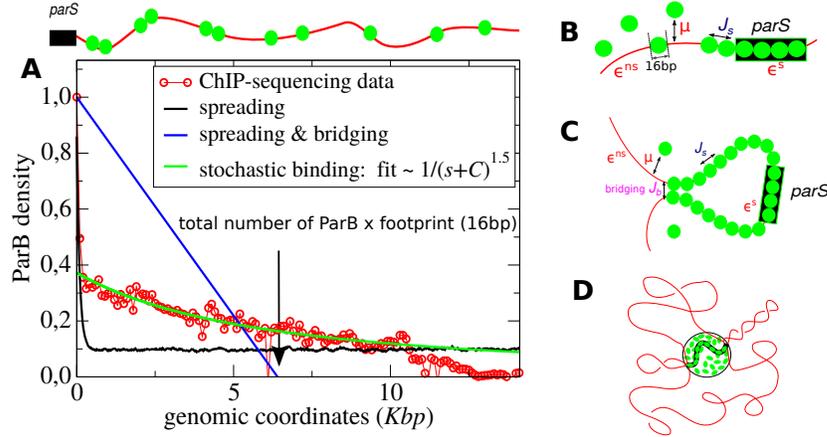


Figure 1.6: (A) ChIP-seq data can only be described by Stochastic Binding. We plot the density profiles of ParB versus the genomic length for the F plasmid of *E. coli* and compare it to the prediction of the three models proposed: (B) the Spreading model (1D Lattice Gas) (C) the Spreading and Bridging model in the strong coupling limit and (D) the Stochastic Binding model, the single model able to describe the trend of ChIP-seq data (see text for a discussion of the parameters).

(i) **The Spreading model (Fig.1.6B).** The polymerization of ParB from the centromere along the plasmid is modeled with the spreading model, or equivalently the one dimensional Lattice Gas [26, 112]. The particles can bind every 16bp, interact weakly with an interaction of magnitude  $J_S \sim 6kT$  (where  $k$  is the Boltzmann constant and  $T$  the absolute temperature) and are exchanged with a reservoir whose concentration is fixed by the chemical potential (or binding parameter)  $\mu$  [26, 112]. Due to the high affinity  $\epsilon^S$  of the particles with *parS*, we suppose that particles are always bound on these sites. The protein distribution along the plasmid from *parS* can then be computed using statistical physics models and arguments, namely by using the mapping between the one-dimensional Lattice Gas onto the Ising model in a field. In this case the probability density to find ParB bound to a generic locus  $s$  is proportional to:  $P(s) \sim \exp(-s/\xi)$  where  $\xi$  is the parameter that defines the characteristic length of the distribution (correlation length, or the typical length of a cluster in unit of bp). As expected for a 1D Lattice Gas, we obtain a fast exponential decrease of the density of ParB to an average coverage value [26, 112], see black curve in Fig. 1.6A. Therefore, the spreading effect from *parS* alone is clearly not able to describe the experimental data on a range of 10kb.

(ii) **The Spreading and Bridging model in the strong coupling limit (Fig.1.6C).** In this model, introduced in Ref. [26], the filament is a polymer embedded in space. Due to the contact of the polymer with itself, ParB dimers can form bridges at large genomic distance with a coupling constant  $J_b$  of the same order of magnitude as the

spreading interaction  $J_S$ . This bridging interaction at long distance induces in the strong interaction limit a condensation of the ParBs in a single continuous filament along the DNA. This filament can move as a whole on DNA with the constraint of overlapping with  $parS$ . Therefore, the density of ParBs is expected to decrease linearly as confirmed with numerical simulations [26]. The only parameter of the distribution is the intersect with the horizontal axis, which gives the maximal extent of the cluster from  $parS$ . Therefore, this value is fixed by the total number of ParB dimers available. This distribution is plotted in blue in Fig. 1.6A. Even if we suppose the maximal expected value of 400 ParBs dimers, we cannot describe the ChIP-seq data [112].

(iii) The **Stochastic Binding model (Fig.1.6D)**. We model the DNA molecule by a Freely Jointed Chain (FJC) characterized by  $N$  freely rotating monomers of size  $a$  (total linear length  $L = aN$ ). The probability distribution  $P(r, s)$  to have two monomers of a Gaussian polymer at a euclidean distance  $r$  and spaced by  $s$  monomers (linear distance  $as$ ) along the polymer is given by[34]:

$$P(r, s) = \left( \frac{3}{2\pi R^2(s)} \right)^{3/2} \exp \left( -\frac{3}{2} \frac{r^2}{R^2(s)} \right), \quad (1.2)$$

where  $R(s) = a\sqrt{s}$  is the averaged radius occupied by the polymer. In the same way, we define the probability distribution function to find a particle ParB with a Gaussian concentration centered at  $parS$  and with a width  $\rho$  corresponding to the averaged radius of the foci occupied by proteins:

$$C(r) = \kappa \exp \left( -\frac{r^2}{2\rho^2} \right), \quad (1.3)$$

where  $\kappa$  is an adimensional normalization constant setting the total number of ParB on the DNA. Thus the occupation rate of a protein on DNA is given by:

$$P(s) = \int_0^\infty 4\pi r^2 dr P(r, s) C(r). \quad (1.4)$$

The integration of Eq.(1.4) gives:

$$P(s) = \frac{\kappa}{\left(1 + \frac{1}{3} \frac{a^2}{\rho^2} s\right)^{3/2}}. \quad (1.5)$$

Note that  $P(0) = \kappa$ , thus  $\kappa$  is setting the height of the drop between specific and non-specific sites and can be estimated directly from ChIP-seq data. When  $R^2(s) \gg 3\rho^2$ , we recover a pure algebraic law  $P \sim s^{-3/2}$ . The theory is compared to ChIP-seq experiments in Fig.1.6A. The best match is obtained with  $\kappa \approx 0.4$ ,  $\rho = 75nm$  [112] and  $a \approx 10bp$  [112]. The good agreement suggest that we need large loop of DNA to explain the long range decay of the ChIP-seq signal, which is impossible with other model based on strong binding interactions. The success of the SB model indicates that the DNA is weakly constrain by the ParB-ParB interactions. We will see below that microscopic models confirm the finding of this phenomenological approach.

The bacterium *Vibrio cholerae* has two chromosomes, and we will be interested in the chromosome I, which uses ParABS for segregation. The ChIP-seq profile is

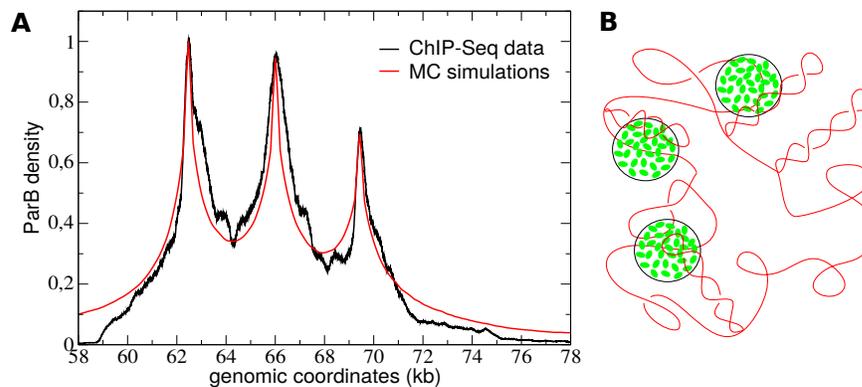


Figure 1.7: (A) The chromosome I of *Vibrio cholerae* displays 3 *parS* sites spaced by  $\sim 4kb$ . The ChIP-seq profile (black curve) displays enrichment at each *parS*. The simulations (red curve) are in good agreement with the data. (B) Simulations are performed like the F plasmid of *E. coli*, but this time with three spheres centered on each of the *parS*. The best fit is obtained with  $a = 16bp$  and  $\rho = 25nm$ .

shown in Fig.1.7A: it displays an enrichment around the three *parS* sites spaced from each other by  $\sim 4kb$ . We observe a different height of peak with different binding affinities, which is interpreted as a difference in specific affinity of ParB for each *parS*. The simulation have been performed with the Stochastic Binding, but this time we need 3 spheres centered on each of the *parS*, see Fig.1.7B. The best fit is obtained with  $a = 16bp$  and  $\rho = 25nm$ .

The Stochastic Binding model is in good qualitative agreement with ChIP-seq data compared to other models based on strong energetic coupling, and gives important information on the conformation of DNA *in vivo*, which needs to be highly unconstrained to form long loops. This model is matching well experimental data at the cost of unphysical persistence length, which we will solve by the introduction of supercoiling in the next subsection.

### 1.2.2 The effect of supercoiling on long DNA molecules: towards an estimation of supercoiling *in vivo*

Supercoiling has a strong effect in bacterial DNA organization, as sketched in Fig.1.8. In most bacteria, DNA is underwound. Despite its critical role for genome structuring [136] and coordination of gene expression [38], measurement of this negative supercoiling along chromosomes remains highly challenging, with both biological and physical difficulties. Biological difficulties stem from the complex functioning of cells. For instance, a large part of supercoiling is known to be absorbed by various histone-like proteins [128]. The remaining supercoiling, which is responsible for the formation of branched plectonemic structures [137], is usually referred to as “free” or “effec-

tive” [14]. Physical difficulties are inherent to the dual nature of supercoiling. That is, in absence of topoisomerases, a topologically constrained DNA molecule, as in the case of a circular (plasmid) molecule or of a constrained linear domain [86], is characterized by a constant linking number,  $Lk$ , equal to the sum of the twist ( $Tw$ ), the cumulative helicity of the molecule, plus the writhe ( $Wr$ ), the global intricacy of the molecule [149]. As a consequence, supercoiling, i.e. the change of  $Lk$  with respect to  $Lk_0$ , the value at rest, leads to changes in the mean values of both  $Tw$  and  $Wr$ . Having access to only  $Tw$ , when using e.g. DNA intercalating agents, is thus a priori insufficient to fully characterize the topological status associated with chromosomal loci [74]. This explains why supercoiling density,  $\sigma = (Lk - Lk_0)/Lk_0$ , has been estimated quantitatively using plasmid reporters only (see [119] for an exception, although the chromosomal measurement is global, not local). Note, in this regard, that a genetic recombination-based system sensitive to the tightness of plectonemes has been developed to address variations of supercoiling density along the chromosome [15, 108]. The quantitative estimation of  $\sigma$  yet remains problematic because the method can only be calibrated in vitro [15].

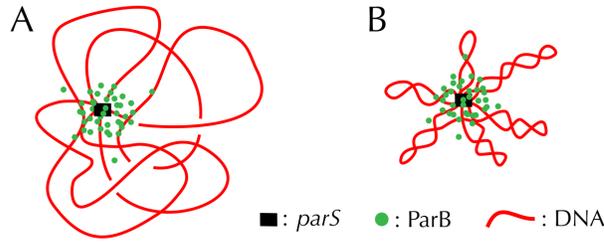


Figure 1.8: *Stochastic binding model*. When DNA enters the high concentration region of the *parS*-anchored cluster of ParB, cross-linking with ParB occurs with high probability during the ChIP-seq protocol. Compared to relaxed DNA (A), supercoiled DNA (B) tends to increase DNA compaction and, hence, cross-linking with DNA loci far from *parS*.

Here, we investigate the possibility of measuring the effective chromosomal supercoiling density using DNA binding properties of the centromere-binding protein ParB, part of the active ParABS system of DNA segregation. Specifically, it has been argued that the capture by chromatin immuno-precipitation sequencing (ChIP-seq) of the binding of ParB onto DNA in the vicinity of its specific binding site (*parS*) is driven by stochastic binding involving DNA looping properties [112] (Fig. 1.8). More precisely, ParB proteins cluster around *parS* [112, 32] through a phase separation-like mechanism [30]. In this context, it has been shown that only a process of looping, which brings DNA loci inside the cluster, can explain the long range decay of the ParB binding profile as the genomic distance to *parS* increases (black curves in Fig. 1.10) [112, 145]. Knowing that supercoiling properties strongly influence DNA looping properties, we thus assess whether a quantitative reproduction of the ParB binding profile in the vicin-

ity of *parS* is possible using a model *with no other free parameter than DNA supercoiling density* ( $\sigma$ ). To this end, we consider, on one hand, a realistic model of supercoiled DNA that has been independently calibrated using single-molecule techniques and, on the other hand, an independent estimation of the size of the ParB cluster using high-resolution microscopic experiments.

Compared to the previous stochastic binding model where a very small DNA persistence length (10bp), difficult to justify on physical grounds, was required to match experimental data [112], here we show, using numerical simulations of realistic long (i.e.  $\geq 30\text{kb}$ ) molecules, that DNA supercoiling indeed leads to a quantitative reproduction of ChIP-seq ParB binding profiles. In this context, we provide a bound for the chromosomal supercoiling density, propose new experimental protocols to further precise its exact values and demonstrate, for the first time to our knowledge, the consistency between chromosomal and plasmid measurements. In addition, we provide novel insights into the physical properties of ParB clusters. In particular, we predict a cluster shape that differs from the usual sharp boundaries of liquid droplets. Namely, we show that the cluster density profile display unconventional “leaky” boundaries, which can be explained as a perturbation induced by a source of proteins located at the edge of the cluster core. Altogether, our work thus offers insights into both bacterial DNA organization and liquid-like protein condensates. It also offers a proof of concept for measuring chromosomal supercoiling with high accuracy.

*Self-avoiding rod-like chain model of DNA.* We consider a realistic 30bp resolution polymer model of bacterial DNA, namely the self-avoiding rod-like chain (sRLC) model [136] (detailed simulation procedure in [83]). Specifically, DNA is modeled as a discrete chain of 10.2nm long (30bp of B-DNA) articulated hard-core cylinders, with radius  $r_e = 2\text{nm}$  reflecting the short-range electrostatic repulsions of DNA for *in vivo* salt conditions [84]. The chain is iteratively deformed using crankshaft elementary motions with Metropolis-Hastings transition rates, under the condition that it does not cross itself. Each articulating site is associated with bending and torsional energies such that the resulting persistence length (50nm or, equivalently, 147bp) and torsional length (86nm) are typical of B-DNA for *in vivo* salt conditions [136, 121, 84].

Here, we discuss results obtained with a 30kb long chain by varying  $\sigma$  from 0 to  $-0.08$  slowly enough so that chain statistical properties are insensitive to the associated rate of change (see simulation details in SM of Ref.[146]). Simulated conformations are thus expected to reflect thermodynamic equilibrium, even at low values of  $\sigma$  where plectonemes are tight. We further checked that our results did not depend significantly on the length of the chain by performing additional simulations of 60kb long chains (Fig. S2 of Ref.[146]). Note, here, that the motivation to work with  $\sigma \geq -0.08$  is both biological and physical: in the worst case of topoisomerase mutants, the total supercoiling density in *E. coli* has been shown to remain above  $-0.08$  [14], while recent work has revealed the existence of a transition toward a hyperbranched regime occurring at  $\sigma \simeq -0.08$  [72], which is beyond the scope of our discussion.

*Leaky vs quenched cluster.* Having in hand the corresponding  $P_s(r)$  for  $\sigma \in [-0.08, 0]$ , we now consider the spatial distribution of ParB proteins associated with the *parS*-anchored clusters. In this regard, high-resolution microscopic measurements [54] suggest that these clusters result from a phase transition-like mechanism. Theoretical models further suggest that this phase transition is unconventional as it implies a frame-

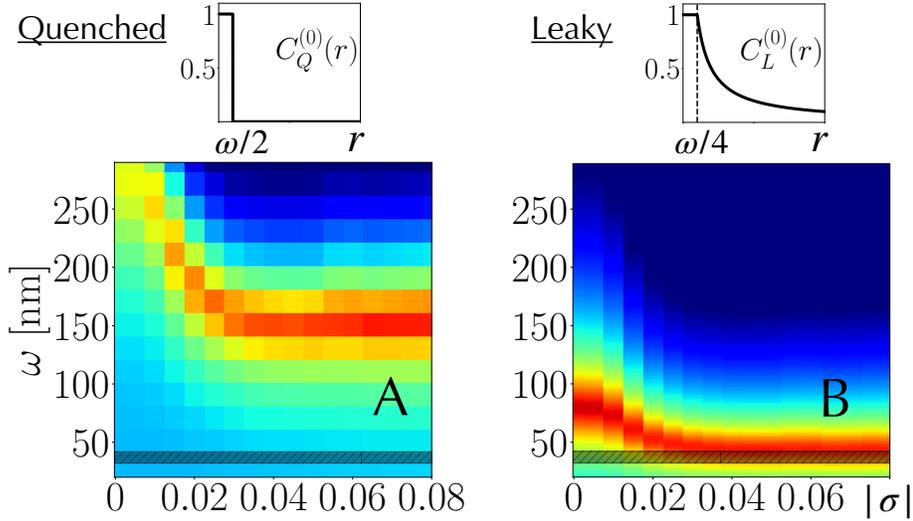


Figure 1.9: *Capturing chromosomal binding profiles*. Root mean squared deviation between modeled binding profiles and ChIP-seq chromosomal data (curves can be found in Fig. S3 of Ref. [146]); the redder the pixel, the smaller the deviation (arbitrary scale). The horizontal dark band indicates  $\omega_{exp}(37\text{nm} \pm 5\text{nm})$ . **(A)** The best models with quenched clusters imply a large cluster with  $\omega_{best} = 150\text{nm}$ . **(B)** In contrast, the best models with leaky clusters imply cluster sizes very close to microscopic data when  $\sigma \lesssim -0.04$ . In this regime, all best models indeed correspond to  $\omega_{best} = 44\text{nm}$ .

work of a lattice gas on a fluctuating polymer [30]. Moreover, the physical formation of a cluster is likely to interfere with biological processes like, e.g., the production of ParB close to the cluster, just as membrane proteins are often produced close to the membrane. In other words, the spatial distribution of ParB proteins around *parS* remains an open question.

Here, we investigate two extreme cases for the shape of these clusters, referred to as *quenched* and *leaky*. A quenched cluster (Fig. 1.9A) is defined by  $C_Q^{(0)}(r) = \theta(\omega/2 - r)$ , with  $\theta$  the Heaviside function. It corresponds to the conventional sharp interface of a droplet. A leaky cluster (Fig. 1.9B) further includes the stationary solution of a diffusion process where ParB proteins are continuously produced at the edge of the cluster core and diluted due to cell growth and division [146]. That is, the leaky cluster release proteins in excess, while  $C_L^{(0)} = 1$  for  $r \leq \frac{\omega}{4}$  (cluster core) reflects the saturation regime in which experiments are performed [32]. As a result,  $C_L^{(0)}$  includes a  $1/r$  long range decay such that  $C_L^{(0)}(r) = \theta(\frac{\omega}{4} - r) + \frac{\omega}{4r}\theta(r - \frac{\omega}{4})$ . Note that for both quenched and leaky cases, the full width at half maximum of  $C^{(0)}$  is equal to  $\omega$ .

We computed binding profiles for  $\sigma$  ranging in  $[-0.08, 0]$  and for values of  $\omega$  between 10nm and 300nm. We compared them with profiles obtained for *E. coli* by inserting *parS* along the chromosome (black curve in Fig. S3 of Ref.[146]) – only one side of the chromosome is analyzed as the other side is distorted by the presence of

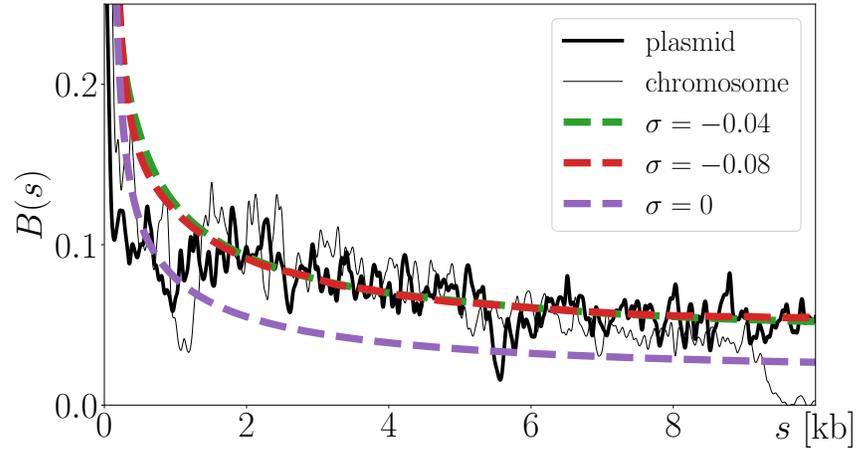


Figure 1.10: *Plasmid & chromosomal profiles vs Stochastic Binding with supercoiling.* The leaky models with  $\omega_{best} = 43\text{nm}$  (smooth curves) and  $\sigma \leq -0.04$  (green and red curves) capture even better plasmid binding profiles (thick black curve). We notice that model predictions for  $\sigma = -0.04$  and  $\sigma = -0.08$  are almost undistinguishable.

strong promoter regions [32]. In this experiment, 10 *parS* sites interspersed by 43 base pairs, as found in the natural *parS* region, were inserted at *xyIE* locus [32]. A careful analysis of the binding properties among these *parS* sites actually revealed significant variations of the ChIP-seq signal, which was thus normalized with respect to the maximum value. The origin of the curvilinear abscissa  $s$  was set right at the edge of the most extreme *parS* site.

We are interested in explaining the global shape of the binding profile as it is expected to reflect generic polymer physics principles. To that end, we quantify the explanatory power of each model by reporting the root mean square deviation with respect to the experimental binding profile for  $s \in [1.5\text{kb}, 9\text{kb}]$ . Both the lower and upper bounds at 1.5kb and 9kb, respectively, are used to avoid specific, reproducible distortions of the signal associated with the presence of gene promoters and sites for regulatory DNA proteins [32].

We find that both quenched and leaky clusters can accurately capture experimental data (Fig. S3 of Ref.[146]). However, the best quenched models are found at  $\omega_{best} = 150\text{nm}$  (Fig. 1.9A), which is much larger than  $\omega_{exp}$ . In contrast, the best leaky models are found at  $\omega_{best} = 44\text{nm}$  when  $\sigma \lesssim -0.04$  (Fig. 1.9B). That is, they explain data in the physiologically relevant plectonemic regime of bacterial DNA. They also solve the small DNA persistence length issue associated with the previous version of the stochastic binding model where DNA supercoiling was neglected [112] – a small persistence length was indeed needed to “mimic” compaction due to plectonemes. Interestingly, compared to chromosomal *parS* data, ParB binding profiles in the vicinity of a *parS* located on a plasmid (100kb long F-plasmid [32]) show less distortion (Fig. 1.10) – just as for the chromosome, only one side of the plasmid is analyzed as the other side is distorted by binding sites for a replication initiator [112]. In this

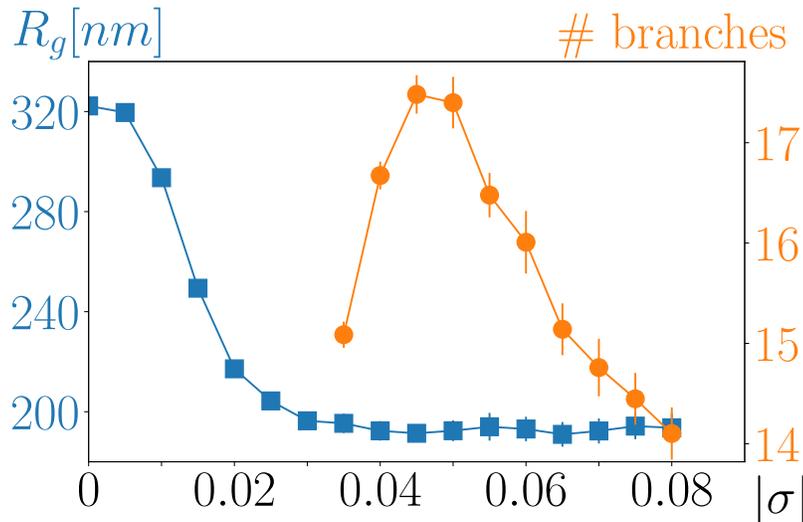


Figure 1.11: The radius of gyration of a 30kb long circular molecule plateaus at  $|\sigma| \simeq 0.04$ . The number of branches is non-monotonous, reaching a maximum at  $|\sigma| \simeq 0.05$  – see SM of Ref.[146] for how to identify plectonemic branches.

context, the best leaky models lead to similar model parameters ( $\omega_{best} = 43\text{nm}$  when  $\sigma \lesssim -0.04$ ), while providing an even better match with the data (Fig. 1.10). Compared to the chromosomal situation where the gene *parB* is located 750kb away from *parS*, this better match might reflect a phenomenology of the plasmid fitting particularly well the leaky situation, with *parB* located only 74bp away from *parS* [17]. The hypothesis of a source located on the edge of the cluster core is indeed even more relevant since the production of proteins in bacteria often occurs close to their gene.

*$\sigma$ -sensitive probes for strong supercoiling.* While leaky models with experimentally relevant  $\omega$  capture experimental data rather well, resulting binding profiles are almost indistinguishable for  $\sigma \in [-0.08, -0.04]$  (Fig. 1.10). This lack of sensitivity is concomitant with a poor variation of the radius of gyration (blue curve in Fig. 1.11) in the plectonemic regime. Note that, in contrast, branching properties can vary significantly in this regime [137, 72]. For instance, we find that the number of plectonemic branches reaches a maximum at  $\sigma \simeq -0.05$  (orange curve in Fig. 1.11), in accord with previous analyses with smaller molecules [137] and with a minimum value of the hydrodynamic radius for 10kb long plasmids [148, 137, 72].

A natural question, then, is whether it is possible to build a probe that is sensitive to variations of  $\sigma$  for strong supercoiling. Interestingly, we have found a possible solution consisting of a system that senses intertwining properties of plectonemes, in the spirit of the  $\gamma\delta$  recombination system [15]. In that respect, one would need a quenched (in-

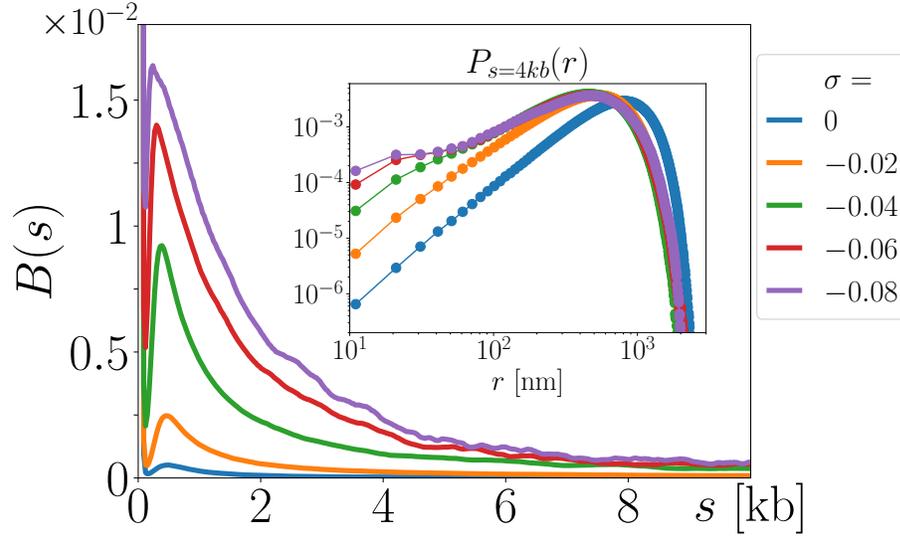


Figure 1.12: A  $\sigma$ -sensitive probe. With a small quenched cluster ( $\omega = 20\text{nm}$ ), binding profiles are well separated for values of  $\sigma \in [-0.08, 0]$  differing by 0.01, which would thus provide a reasonable precision for measuring supercoiling. Inset: in the plectonemic regime ( $\sigma \leq -0.04$ ), the spatial distribution of distances between loci differ significantly only at small distances associated with plectonemic intertwining properties.

stead of a leaky) cluster that is small enough such that the binding properties of proteins is sensitive to the diameter and pitch of plectonemes [92, 12]. For instance, our simulations reveal a strong sensitivity of  $P_s(r)$ , at the kb genomic scale for  $s$ , with respect to *all values of*  $\sigma$  for spatial distances  $r$  on the order of 10nm (inset of Fig. 1.12). One can verify, then, that a quenched cluster with  $\omega = 20\text{nm}$  provides well-distinct binding profiles for  $\sigma \in [-0.08, 0]$  (Fig. 1.12). Notice the much smaller values of  $B(s)$  in this case, compared e.g. to results in Fig. 1.10. ParB ChIP-seq experiments can nevertheless report very low binding frequencies as demonstrated by titration assays [32].

*Discussion and perspectives.* We have shown that the binding profile of ParB proteins in the vicinity of *parS* can be quantitatively explained considering a stochastic binding process between a supercoiled DNA and proteins that are issued from a saturated *parS*-anchored core cluster. To this end, we had to consider clusters from a non-equilibrium, stationary perspective, with the presence of a spatially localized source and sink. Biologically, the sink reflects protein dilution due to cell growth and division, while the source may arise from two effects: the continuous activity of genes producing new proteins in a saturated cluster and the effect of an unconventional liquid-like nature of the cluster. Namely, we predict the cluster core to result from a balance between an influx of continuously produced proteins and an outflux of proteins in excess. In the plasmid, the situation may even be more prototypical with the production of ParB occurring close to *parS*.

In this context, and for the first time to the best of our knowledge, we provide an upper bound ( $\sigma \approx -0.04$ ) for the *in vivo* supercoiling density at a chromosomal location of a bacterium (*E. coli* during its exponential growth) and we show that it also holds for plasmids. Interestingly, this value corresponds to the onset of the plectonemic regime characterized by a poor variation of the radius of gyration, on one hand, and a significant variation of branching properties, on the other hand. Importantly, we also offer a proof of concept to obtain a finer estimate of the supercoiling density. Specifically, in the spirit of existing genetic recombination-based probes [15, 108], we demonstrate that a small quenched cluster provides a supercoiling-sensitive probe as it "senses" physical properties of plectonemes.

Compared to "biological" genetic recombination-based probes, our "physical" ChIP-seq-based probe is expected to be much less invasive. It should also be less sensitive to molecular environment as it is based on generic (polymer) physics properties – for instance, recombination-based systems depend on (slow) enzymatic recombinase reactions, whose quantitative modeling has, to the best of our knowledge, remained elusive. In practice, while genetic design of quenched clusters of ParB proteins might be tricky, transcription factors could provide an efficient system. These proteins have indeed the capacity of binding both cognate DNA sites strongly and other DNA sites non-specifically with (short) millisecond residence times [40]. They could also be used in conjunction with a DNA methyltransferase to generate methylation (instead of binding) profiles without the need of crosslinking stages [106]. Finally, a sensitive system would require having the designed artificial DNA devoid as much as possible of interfering biological elements, such as gene promoters, which distort the utilizable physical signal. Along this line, one would like to have an explicit description of ParB nucleation and diffusion properties to develop a detailed model of the interactions between ParB and DNA using e.g. molecular dynamics approaches. In particular, the discrepancy between experimental and modeling profiles below  $\sim 1$ kb (Fig. 1.10) might be the result of our approximation of neglecting hard-core interactions between ParB proteins and DNA. At large scales, cellular confinement of DNA should also be included in the model. We note, nevertheless, that a complete picture would require studying the melting of a plectonemic tree-like structure at the chromosome scale, which is currently beyond the capacities of numerical simulations.

### 1.3 The Looping and Clustering model

The Looping and Clustering model [145] introduces the coupling between ParB proteins in an approximate manner, and is thus an interpolation between the phenomenological Stochastic Binding model (where the formation of the ParB cluster is not explained) and the full microscopic description given by the Spreading and Bridging model.

Physically, the Spreading and Bridging model describes the collective behavior of interacting proteins on a self-avoiding polymer chain. This model suggested that ParB assembles into a three-dimensional complex on the DNA, as illustrated in Figure 1.13(A,B). In this model, each protein can interact with nearest neighbors along the DNA through a 1D spreading interaction and, in addition, each protein can inter-

act through a 3D bridging interaction with another ParB protein bound to a site that may be distant on the DNA albeit close in 3D space. This Spreading & Bridging model is supported by single molecule experiments, which provide direct evidence for the presence of bridging interactions between two ParB proteins on DNA [53, 126]. Importantly, computational studies have shown that the Spreading & Bridging model supports phase transitions [26].

In this Section, we propose a comprehensive theory to describe the distribution of ParB proteins on the chromosome in terms of molecular interaction parameters. We expand on the ideas of the Spreading & Bridging and the Stochastic Binding models to provide a quantitative analytic approach to describe the genomic organization of ParB that are bound around *parS* sites on the DNA. To this end, we develop a simple model for protein-DNA clusters that explicitly accounts for the competition between protein-protein interactions, which tend to favour a compact cluster, and the entropy associated with the formation of loops, which favours a looser cluster configuration. This *Looping and Clustering* model represents a reduced, approximate version of the full Spreading & Bridging model that provides a clearer understanding and greatly facilitates calculations of the distribution profile of ParB or other proteins that form protein-DNA clusters.

### 1.3.1 The model

To theoretically describe the protein binding profiles of ParB on DNA, we first consider a DNA polymer of length  $L$  that can move in space on a 3D cubic lattice and is in contact with a “cytoplasm” containing a finite number of proteins. These proteins are able to bind to/unbind from the lattice sites on the DNA along which the proteins can freely diffuse. Importantly, in this model the DNA itself is also dynamic and fluctuates between different three-dimensional configurations, which are affected by the presence of interacting DNA bound proteins. When proteins are bound to the DNA, they are assumed to be able to interact attractively with each other by contact interactions in two distinct ways: (i) 1D spreading interactions with coupling strength  $J_S$ , defined as an interaction between proteins on nearest-neighbor sites along the polymer, and (ii) a 3D bridging interaction with strength  $J_B$  between two proteins bound to sites on non-nearest neighbor-sites on the DNA, but which are positioned at nearest neighbor-sites in 3D space (see Figure 1.13(A,B)). Thus, these bridging interactions couple to the 3D configuration of the DNA, while the 1D spreading interactions do not. In prior work, we introduced this model for interacting proteins on the DNA, which was termed the Spreading & Bridging model [26]. The Hamiltonian for this model is given by:

$$\mathcal{H}_{S\&B} = -J_S \sum_{i=1}^L \phi_i \phi_{i+1} - \frac{J_B}{2} \sum_{i,j} \phi_i \phi_j \delta_{|\mathbf{r}_i - \mathbf{r}_j|, 1}, \quad (1.6)$$

where the variable  $\phi_i \in \{0, 1\}$  represents the occupancy by a protein of the  $i$ -th DNA binding site, and the Kronecker delta  $\delta_{|\mathbf{r}_i - \mathbf{r}_j|, 1}$  is equal to one when binding sites  $i$  and  $j$  with respective spatial locations  $\mathbf{r}_i$  and  $\mathbf{r}_j$  are positioned on nearest-neighbor binding sites in space and it is zero otherwise. Note, this particular form of the Hamiltonian, in principle, allows a valency of up to 4 bridging bonds per protein on a 3D cubic

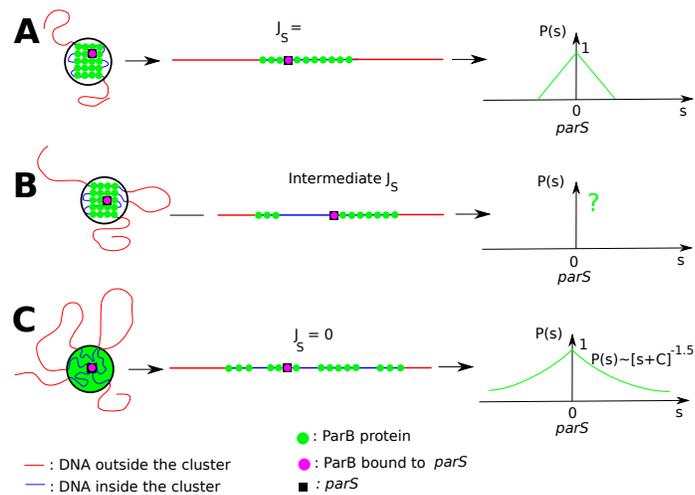


Figure 1.13: (left) Schematic illustration the models; (middle) a typical distribution of ParB on extended DNA; (right) the average distribution profile. The Spreading & Bridging model [26] is shown with (A) strong coupling  $J_S \rightarrow \infty$ , where thermal fluctuations cannot break the bonds between proteins such that all bridging and spreading interactions are satisfied, and (B) Intermediate coupling where the energetic cost of breaking a spreading bond is balanced by configurational and loop entropy. With the Looping and Clustering approach presented here, we propose a simple analytic description for this regime. (C) The Stochastic Binding model assumes presented above [112] can be seen as taking the limit of the spreading bond strength to zero ( $J_S \rightarrow 0$ ), and thus the formation of loops are not hampered by protein-protein bonds.

lattice (2 bonds out of 6 are already used for spreading interactions). Single-molecule experiments provide evidence for bridging bonds [53], but the actual bridging valency of a ParB protein may be limited to one [82, 48]. To capture this, the Hamiltonian in the Spreading & Bridging can be easily adjusted to reduce the valency for 3D bridging bonds per protein. Even in the realistic limit where each protein can form two spreading bonds and a single bridging bond, the system has been shown to exhibit a condensation transition where the majority of the proteins form a single large cluster that can be localized by a single *parS* site on the DNA [26].

While it is possible to perform Monte Carlo simulations of the Spreading & Bridging model for a lattice polymer, it remains challenging to perform analytical calculations within this framework. In this section, we are primarily concerned with describing the average binding profile of proteins along the DNA (see right panels in Figure 1.13). With this aim in mind, we can simplify our model by realizing that the configurations of ParB proteins along the DNA are more sensitive to  $J_S$  than to  $J_B$ . While both spreading and bridging bonds are necessary for the condensation of all proteins into a single cluster, loop extrusion from the cluster is controlled by  $J_S$ , and such loop extrusion will strongly impact the binding profile of proteins on the DNA. Thus, we will assume that  $J_B$  is sufficiently large that approximately all available bridging bonds between the proteins inside the 3D protein-DNA cluster are satisfied, leaving  $J_S$  as the main adjustable parameter in the model.

A contiguous 3D cluster of proteins on DNA with loops can effectively be represented by a disconnected 1D cluster along the DNA, where connections in 3D between the 1D subclusters are implied, and domains of protein-free DNA within the disconnected 1D cluster represent loops that emanate from the 3D cluster (see Figure 1.13(B,C)). We can describe this system by a reduced model for the effective 1D cluster in which we account for the entropy of the loops that originate from the protein-DNA cluster. In this model, the spreading bond energy, set by the parameter  $J_S$ , competes with the formation of loops and will therefore play a crucial role in determining the binding profile of ParB on DNA around a *parS* site.

To capture these effects, we propose the reduced Looping and Clustering (LC) model, which offers a simplified description of 3D protein-DNA clusters with spreading and bridging bonds. In this model a loop is formed whenever there is a gap between 1D clusters. We can make the connection between the gaps in the 1D cluster and the number of loops extending from the 3D cluster explicit by writing down the partition function for this model. The effective 1D cluster corresponding to a 3D cluster with  $m$  proteins and  $n$  loops has a multiplicity:

$$\Omega_{\text{cluster}} = \frac{(m-1)!}{(m-n-1)!n!}, \quad (1.7)$$

which counts the number of ways in which one can partition  $m$  proteins into  $n+1$  subclusters in 1D. In addition, we note that creating  $n$  loops will require breaking  $n$  spreading bonds, and the probability at equilibrium for this to occur will include a Boltzmann factor  $\sim \exp(-nJ_S)$ , where the interacting energy is expressed in unit of  $k_B T$ . The loops that are formed are assumed to be independent, and thus modify the

configurational entropy as [34]:

$$S_{\text{loop}} = -d\nu \sum_{i=1}^n \ln(\ell_i/\ell_0), \quad (1.8)$$

where  $d$  is the spatial dimension,  $\nu$  is the Flory exponent, and  $\ell_0$  is the lower cutoff of loop sizes and approximately represents the persistence length of DNA. This entropy is obtained by considering both the loops formed within the protein cluster and the protein-free segment of DNA outside the cluster. Indeed, the number of configurations associated with loop  $i$  for a Gaussian polymer is given by  $z^{\ell_i} \ell_i^{-d\nu}$  [34, 59], where  $z$  is the lattice coordination number. Thus, there is also an extensive contribution to the entropy given by  $k_B \ell_i \log(z)$ . However, when a loop of length  $\ell_i$  forms, the same length of polymer is removed from the DNA outside of the cluster, which also results in a reduction of the entropy by  $k_B \ell_i \log(z)$ . Thus, there is a precise cancellation between the extensive contribution to the entropy associated with the loop inside the cluster and the extensive contribution due to effectively shortening the DNA outside the cluster.

It is now straightforward to write down the partition function of the Looping and Clustering model:

$$\begin{aligned} Z_{LC} &= \sum_{n=0}^{m-1} \frac{(m-1)!}{(m-n-1)!} \exp(-nJ_S) \int_{\ell_0}^{\ell_{\max}} d\ell_1 \ell_1^{-d\nu} \dots \int_{\ell_0}^{\ell_{\max}} d\ell_n \ell_n^{-d\nu} \\ &\stackrel{\ell_{\max} \rightarrow \infty}{=} \sum_{n=0}^{m-1} \frac{(m-1)!}{(m-n-1)!} \exp(-nJ_S) \left[ \frac{\ell_0^{1-d\nu}}{d\nu-1} \right]^n \\ &= \left( 1 + \exp(-J_S) \frac{\ell_0^{1-d\nu}}{d\nu-1} \right)^{m-1}. \end{aligned} \quad (1.9)$$

where all lengths are measured in units of protein binding sites, and the bond interactions are in units of  $k_B T$ . In the partition function, it is convenient to set the upper boundary of integration to infinity. Strictly speaking, the upper boundary should be  $L - (m + \ell)$ , where  $\ell = \sum_i \ell_i$  represents the total accumulated loop length. In practice, however, for chromosomes, but arguably also for plasmids,  $L \gg m$  and the probability to have a large loop is very small. For instance, if we consider the F plasmid of *E. coli* with a length of 60 kb, it would correspond to  $L = 3750$  in units of the ParB footprint of 16 bp [19, 112]. For this system, Monte Carlo (MC) simulations of the LC model, with  $m = 100$  reveal that the average cumulated loop size is  $\ell \approx 140$  for small couplings ( $J_S = 2$ ) down to  $\ell \sim 10$  for large couplings ( $J_S = 5$ ), which in both cases is much less than the DNA length. Thus, for biologically relevant cases it is reasonable to assume that the length of the DNA polymer is much larger than the footprint of the whole protein complex on the DNA.

From the partition function, we can write down an effective 1D Hamiltonian for the LC model, which explicitly accounts for the balance between spreading bonds and loop entropy:

$$\mathcal{H}_{LC} = -J_S \sum_{i=1}^{L-1} \phi_i \phi_{i+1} + d\nu \sum_{i=1}^n \ln(\ell_i/\ell_0). \quad (1.10)$$

This Hamiltonian is useful to perform Monte Carlo simulations of the model as a benchmark for the approximations performed in the analytical approach described in the next sections.

In summary, the LC model constitutes a simple statistical mechanics approach to describe how proteins assemble into a protein-DNA cluster with multiple loops. To calculate the average distribution of proteins along the DNA, we will assume that the affinity of ParB binding to *parS* is sufficiently strong such that one of the proteins in the cluster is always bound to a *parS* site.

In Ref. [145], we describe how to compute the ParB binding profile around this *parS* site given a fixed number of loops with specified loop lengths. We then use the statistical mechanics framework provided above to perform a weighted average over all possible loop numbers and sizes to arrive at a simple predictive theory for the ParB binding profile. For the sake of clarity, we present here only the main results and conclusion.

### 1.3.2 Protein binding profiles from LC

Real protein-DNA clusters are expected to fluctuate with new loops forming and disappearing continuously. To capture such fluctuations, we will use the expressions for the binding profile of a static cluster with fixed loop length together with a statistical mechanics description of the LC model to obtain *average* binding profiles for dynamic clusters (see Ref. [145] and Appendix therein), including an ensemble average over both the number of loops and the loop lengths.

To obtain a full binding profile averaged over all realizations, it is useful to investigate the statistics of loops that extend from the protein-DNA cluster and how these statistics are determined by the underlying microscopic parameters of the model. We start by considering the number of loops that extend from the cluster. Using the partition function in Eq. (1.9), it is possible to calculate the basic features of the LC model. For instance, the moments of the distribution of the number of loops are given by:

$$\langle n^\alpha \rangle = \frac{1}{Z_{LC}} \sum_{n=0}^{m-1} n^\alpha \frac{(m-1)!}{(m-n-1)!n!} e^{-nJ_S} \left[ \frac{\ell_0^{1-d\nu}}{d\nu-1} \right]^n. \quad (1.11)$$

From this, we easily find the average loop number is:

$$\langle n \rangle = (m-1) \frac{x}{1+x} \underset{m \gg 1}{\propto} m e^{-J_S}, \quad (1.12)$$

where  $x = e^{-J_S} \ell_0^{1-d\nu} / (d\nu - 1)$ . The average loop number  $\langle n \rangle$  is depicted in Figure 1.14(A), demonstrating the exponential dependence on the spreading energy  $J_S$ . In Figure 1.14(B), we plot  $\langle n \rangle$  as a function of the total number of proteins  $m$  in the protein-DNA cluster. Over a broad range of parameters, we observe the expected linear dependence of the average loop number  $\langle n \rangle$  on  $m$ .

The linear dependence on  $m$  reflects that, in the LC model, loops can form anywhere in the cluster. However, one would expect that loops can only form at the surface of a 3D cluster. For a compact, spherical cluster, this would result in a dependence  $\langle n \rangle \sim m^{2/3}$ . However, Monte Carlo simulations of the full S&B model have revealed

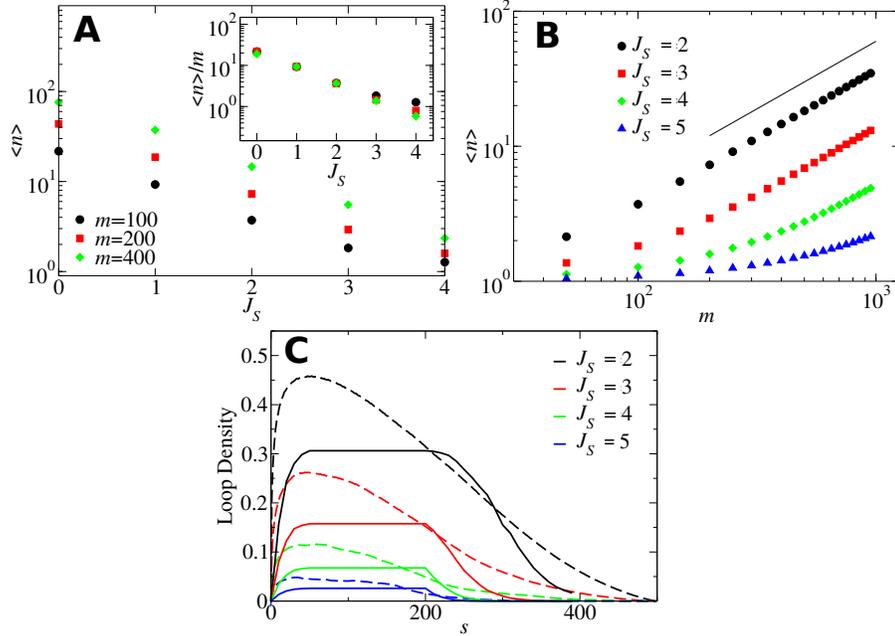


Figure 1.14: **(A)** Average number of loops,  $\langle n \rangle$ , as a function of spreading coupling strength  $J_S$  obtained by numerical integration of Eq. (1.11). The different curves correspond to protein number  $m = 100$  (black),  $m = 200$  (red), and  $m = 400$  (green), with loop-size cutoff  $\ell_0 = 7.5$ . For large enough  $\langle n \rangle$ , we observe an exponential decrease  $\langle n \rangle \propto e^{-J_S}$  in accord with Eq.(1.12). Inset: Same data replotted with expected dependence of average loop number on  $m$  scaled out. **(B)** Average number of loops  $\langle n \rangle$  as a function of  $m$  for  $J_S = 2, 3, 4$ , and  $5$ . The black solid curve is a guide to eyes with a slope 1. The prefactor that determines the vertical shift between the different curves scales with  $e^{-J_S}$ , as demonstrated in the inset of panel (A). **(C)** Average loop density,  $\rho(s, m, \ell)$ , as a function of the genomic coordinate with  $m = 200$  and  $L = 1000$  for protein-DNA clusters with fluctuating loop number and loop lengths. Different curves correspond to different spreading couplings  $J_S = 2, 3, 4$ , and  $5$ . The simple analytic approximation(explained in Ref. [145]) (solid curves) is compared to MC simulations (dashed curves).

that the protein-DNA clusters are not compact [26], but rather have a surface that scales almost linearly in  $m$ , close to the behavior of the simplified LC model presented here.

The loop statistics of protein-DNA clusters are not easily accessible in experiments. Perhaps the most relevant results for which this model can provide insight come from ChIP-Seq experiments. In the LC model, the ParB density profile along DNA can be

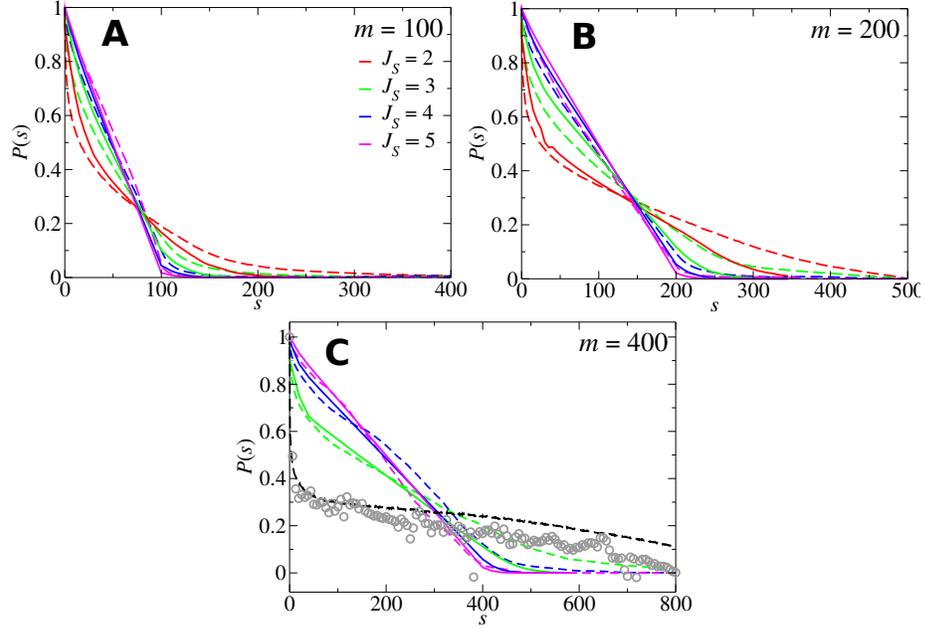


Figure 1.15: Binding profiles of ParB plotted versus the genomic distance  $s$  to  $parS$  for (A)  $m = 100$  and  $L = 1000$ , (B)  $200$  and  $L = 1000$ , and (C)  $400$  and  $L = 2000$ . Analytic results from Eq. (1.13) are shown as solid curves and numerical results from Monte Carlo simulations of the LC model are shown as dashed curves (averages over  $10^6$  configurations). In Eq. (1.13), the loop size integrals were calculated with an upper cutoff of  $7\ell_0$  and summations were truncated at  $n = 15$ . For this reason, the convergence for  $m = 400$  and  $J_S \leq 2$  was not sufficient and is not displayed here. Instead, we compare the ChIP-Seq data (grey) of the F-plasmid with MC simulations with a black dashed line ( $J_S = 1$ ,  $m = 400$ ,  $L = 2000$ ).

calculated from:

$$P_{LC}(s) = \frac{1}{Z_{LC}} \sum_{n=0}^{m-1} \frac{(m-1)!}{(m-n-1)!n!} \exp(-nJ_S) \int_{\ell_0}^{\infty} d\ell_1 \ell_1^{-d\nu} \dots \int_{\ell_0}^{\infty} d\ell_n \ell_n^{-d\nu} P_n(s, \{\ell_i\}), \quad (1.13)$$

where  $Z_{LC}$  is given in Eq. (1.9). Here,  $P_n(s, \{\ell_i\})$  represents the multiloop ParB binding profile with  $n$  loops with loop lengths  $\{\ell_i\} = \{\ell_1, \dots, \ell_n\}$ . For simplicity, we approximate this multiloop profile by the analytical 1-loop conditional probability,  $P_1(s, \ell | \overline{\text{loop@s}})$ , with the loop length equal to the accumulated loop length, i.e.  $\ell \rightarrow \sum_i \ell_i$ , weighted by the loop probability  $p(\overline{\text{loop@s}}) \approx \sum_{i=1}^n \ell_i \rho(s, m, \ell_i, \ell)$  (see Ref.[145] for the full expression and the derivation). In the expression for the loop probability,  $\rho(s, m, \ell_i, \ell)$  is defined as the contribution to the loop density of a loop of length  $\ell_i$  to a cluster of  $m$  proteins with a total accumulated loop length  $\ell$ , and we neglected correlations between contributions from different loops (see Appendix of

Ref. [145]).

In the analysis above, we aimed to capture the effects of multiple loops in a simple way by assuming statistical independence of the loops, and by using the analytical 1-loop expressions to approximate the impact of loop formation on the loop density and the ParB binding profile of the protein-DNA complex. To test the validity of these approximations, we performed MC simulations of the complete LC model. We find that the numerically obtained loop probability is in reasonable agreement with our approximate expression for the multi-loop density, as shown in Figure 1.14(C). Thus, despite the simplicity of our approach, the analytical model provided here captures the essential features of looping in protein-DNA clusters.

The full protein binding profile  $P(s)$  around a *parS* site is calculated by averaging the static binding profile for different total loop numbers and loop lengths using the Boltzmann factor (see the partition function Eq. (1.9)) from the Looping and Clustering model as the appropriate weighting factor. The resulting expression in Eq. (1.13) for the protein binding profile of a protein-DNA cluster is our the central result. We use this expression to compute binding profiles for the full LC model, which are shown in Figure 1.15 as a function of the distance  $s$  to *parS* for  $m = 100$  (A), 200 (B), and 400 (C). By construction, the site  $s = 0$  corresponding to *parS* is always occupied, and thus  $P(s = 0) = 1$  for all values of the spreading energy  $J_S$ . This feature of the LC model captures the assumed strong affinity of ParB for a *parS* binding site. For  $J_S = 5$ , the binding profile converges to a triangular profile, implying a very tight cluster of proteins on the DNA with almost no loops. The triangular profile in this case results from all the distinct configurations in which this tight cluster can bind to DNA such that one of the proteins in the cluster is bound to *parS*, and therefore the probability drops linearly to 0 at  $s \approx m$ . The same triangular binding profile was observed for the S&B model in the strong coupling limit  $J_S \rightarrow \infty$  [26]. Interestingly, as  $J_S$  becomes weaker, we observe a faster decrease of the binding profile near *parS* together with a broadening of the tail of the distribution for distances far from *parS*. This behavior results from the increase of the number of loops that extrude from the ParB-DNA cluster with decreasing spreading bond strength  $J_S$ . The insertion of loops in the cluster allows binding of ParB to occur at larger distances from *parS*. Thus, the genomic range of the ParB binding profiles is set by  $s_{\max} \approx m + \langle \ell \rangle$ , where the average cumulated loop length  $\langle \ell \rangle$  is controlled by  $J_S$  and  $m$ . These results illustrate how the full average binding profile is controlled by the spreading bond strength  $J_S$ : the weaker  $J_S$ , the looser the protein DNA-cluster becomes, which results in a much wider binding profile of proteins around *parS*.

In the limit  $J_S \rightarrow 0$ , the LC model quantitatively reduces to the statistics of non-interacting loops predicted by the Stochastic Binding model [112], See Ref. [145]. In this case, the binding profiles exhibit asymptotic behaviour  $P(s) \propto s^{-1.5}$  for large  $s$ . Interestingly, when the number of cluster proteins  $m$  increases, we observe the appearance of a second regime where  $P(s) \sim s^{-\alpha}$  with  $\alpha \approx 0.5$ . We attribute this weaker scaling  $P(s)$  with  $s$  at intermediate genomic distances to the reduced loop density near *parS* (see Figure 1.14(C)). A similar behaviour was also observed for the Stochastic Binding model: at small genomic distance from *parS*, the DNA is assumed to be always inside the region of high ParB concentration, leading to a slower decay of the binding profile of ParB near *parS*.

### 1.3.3 Discussion

In the LC model, the formation of a coherent cluster of ParB proteins is ensured by a combination of spreading and bridging bonds between DNA bound proteins, which together can drive a condensation transition in which all ParB proteins form a large protein-DNA complex localized around a *parS* site. We do not assume, however, that this protein-DNA cluster is compact. Indeed, loops of protein-free DNA may extend from the cluster, which strongly influences the average spatial configuration of proteins along the DNA. In the LC model, the formation of loops in the protein-DNA cluster is controlled by the strength of spreading bonds, i.e. the bond between proteins bound to nearest neighbor sites on the DNA. Specifically, for every protein-free loop of DNA that extends from the cluster, a single spreading bond between two proteins within the cluster must be broken. Thus, if the spreading interaction energy,  $J_S$ , is sufficiently small, thermal fluctuations will enable the transient formation and breaking of spreading bonds, thereby allowing multiple loops of DNA to emanate from the protein cluster (See Figure 1.13). The LC model predicts a profile in good quantitative agreement with binding profiles measured with ChIP-Seq on the F-plasmid of *E. coli* [112] with  $J_S = 1k_B T$  and  $m = 400$ , as shown in Fig. 1.15(c).

The Looping and Clustering model, which we introduce to calculate the binding profile of ParB-like proteins on the DNA, is a simple theoretical framework similar to the Poland-Scheraga model for DNA melting [102, 44]. An important difference in the LC model with respect to the homogeneous Poland-Scheraga model, is that translational symmetry is broken due to the presence of a *parS* site at which a protein is bound with a high affinity such that loops are effectively excluded in the vicinity of *parS*. We show that the binding profiles predicted by this model are sensitive to both the expression level of proteins and the spreading interaction strength, which directly controls the formation of loops in the protein-DNA cluster. Our model thus provides a means to use binding profiles, measured for instance in ChIP-Seq experiments, to infer molecular interaction strengths of the proteins that form large protein-DNA clusters.

Conceptually, the spreading bond interaction determines how “loose” the protein DNA cluster is, which directly impacts the ParB binding profiles. When  $J_S$  is large, loop formation is unlikely, resulting in a compact protein-DNA cluster with a corresponding triangular protein binding profile centered around *parS*. At intermediate  $J_S$ , the protein-DNA cluster becomes looser with the formation of loops, resulting in a binding profiles that are more strongly peaked around *parS* but with far-reaching tails, which is in accord with high-resolution ChIP-Seq experiments [112]. In the limit  $J_S \rightarrow 0$ , the loops can form at any position with no energetic cost. In this limit, the binding profile  $J_S \rightarrow 0$  is consistent with the Stochastic Binding model with a profile of the form [112]:  $P(s) \propto s^{-1.5}$  [145]. Thus, the Looping and Clustering model offers a description for a broad parameter regime, connecting two limits investigated in preceding studies [26, 112]. This model provides an insightful quantitative tool that could be employed to analyze and interpret ChIP-Seq data of ParB like proteins on chromosomes and plasmids.

## 1.4 Spreading & Bridging model

Despite extensive numerical studies [26, 66, 145, 116], it is still unclear theoretically how long 1D substrates like DNA polymers interact with particles to form 3D structures essential for the cellular cycle [24, 93, 69]. Interestingly, similar organizational principles may apply to the higher-order folding of chromatin and the interactions between topological domains in eukaryotic cells [69, 11, 56, 67]. A common theme is the mechanism of protein-induced polymer loop formation via bridging interactions and the role played by these loops in structuring DNA and creating localized protein-DNA complexes. Three different basic models have been studied, mainly using simple mean-field Flory-type approaches and simulations: (i) sparse but fixed interacting sites [69, 116], (ii) non-interacting mobile bound particles that can bind simultaneously to two polymer sites to form bridges [11, 116], and (iii) mobile bound particles that can interact to form both nearest-neighbor (NN) and bridging bonds [26, 129]. However, an analytical statistical mechanics framework is still needed to clarify the existence and nature of phase transitions in such systems. Here, we present an analytical Hamiltonian approach to case (iii) by introducing a basic microscopic particle-polymer model where all relevant physical parameters appear explicitly. From this model, we derive an effective 1D lattice gas model with 1D temperature-dependent long-range interactions that arise from the 3D conformational fluctuations of the polymer. We show that the existence of a phase transition in this effective model depends on the exponent describing the asymptotic power law decay of the long-range interactions. We then propose a variational method that goes beyond mean-field theory (MFT) to compute the mean occupation-temperature phase diagram. Finally, for illustration, we apply our model to the bacterial partition system ParABS and the formation of ParBS complexes. We propose a plausible explanation in terms of metastability for experiments showing the existence of high density ParB protein condensates only in the presence of specific binding sites.

### 1.4.1 The model LRLG

In our approach (see Fig. 1.16), the polymer consists of  $N$  monomers (or sites) with each monomer capable of accommodating one bound particle. The effective monomer length  $l_m$  corresponds to the footprint of one particle on the polymer, measured, for example, in terms of base pairs for DNA. Each site  $i$  is characterized by its position in 3D space  $\mathbf{X}_i$ , its occupation  $\Phi_i$  (equal to 1 if a particle is bound and 0 otherwise) and its on-site binding energy  $\varepsilon_i$ . This energy allows us to implement local specific or non-specific binding. In the particle grand-canonical ensemble, the energy of a state  $[\Phi_i, \mathbf{X}_i]$  is

$$H[\Phi_i, \mathbf{X}_i] = H_P[\mathbf{X}_i] + H_{\text{SRLG}}[\Phi_i] + H_B[\Phi_i, \mathbf{X}_i]. \quad (1.14)$$

The first term  $H_P[\mathbf{X}_i]$  describes the polymer configuration energy. The second is a 1D Short Range Lattice Gas (SRLG) Hamiltonian for bound particles,

$$H_{\text{SRLG}}[\Phi_i] = -J \sum_{i=1}^{N-1} \Phi_{i+1} \Phi_i - \sum_{i=1}^N (\mu - \varepsilon_i) \Phi_i \quad (1.15)$$

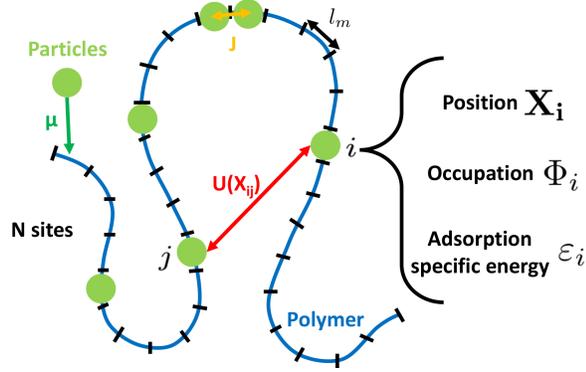


Figure 1.16: Schematic illustration of the coupled polymer-particle model. The polymer in 3D is divided into  $N$  monomers, each having three attributes: a position vector  $\mathbf{X}_i$ , an occupation  $\Phi_i$ , and a local adsorption energy  $\epsilon_i$ . Loops form when particles far apart along the polymer interact at short range in 3D.

with NN spreading interaction coupling constant  $J$  and chemical potential  $\mu$ . The contribution from 3D bridging interactions, giving the coupling between the bound particles and the fluctuating polymer, takes the form

$$H_B[\Phi_i, \mathbf{X}_i] = \frac{1}{2} \sum'_{i,j} \Phi_i U(X_{ij}) \Phi_j, \quad (1.16)$$

with  $X_{ij} = |\mathbf{X}_i - \mathbf{X}_j|$  and  $U(X_{ij})$  the potential of 3D spatial interaction between particles. The prime on the sum means that  $|i - j| \geq n_{\text{inf}}$ , where  $n_{\text{inf}}$  is the minimal internal distance in number of sites over which two particles can interact at long-range.

The polymer conformational degrees of freedom can formally be integrated out, yielding a highly non-linear 1D effective free energy for the bound particles including two and all higher body interactions along the chain. Given the complexity of this coupled model, we derive using a virial (cluster) expansion [95, 46] a more amenable 1D effective model that retains only short and two body long-range interactions:

$$\begin{aligned} \frac{\mathcal{Z}}{\mathcal{Z}_P} &= \sum_{\{\Phi_i=0,1\}} e^{-\beta(H_{\text{SRLG}}[\Phi_i] - \beta^{-1} \ln \langle e^{-\beta H_B[\Phi_i, \mathbf{X}_i]} \rangle_P)} \\ &\approx \sum_{\{\Phi_i=0,1\}} e^{-\beta \mathcal{F}_{\text{LRLG}}[\Phi_i]} \end{aligned} \quad (1.17)$$

where  $\beta = 1/(k_B T)$ ,  $\langle \cdot \rangle_P$  denotes an average over polymer conformations,  $\mathcal{Z}_P$  is the partition function of the bare polymer, and  $\mathcal{F}_{\text{LRLG}}[\Phi_i]$  is a 1D long-range Lattice Gas (LRLG) effective (temperature dependent) free energy:

$$\mathcal{F}_{\text{LRLG}}[\Phi_i] = H_{\text{SRLG}}[\Phi_i] - \frac{1}{2} \sum'_{i,j} \Phi_i G_{ij} \Phi_j \quad (1.18)$$

The second term of Eq. (1.18) is an effective 1D long-range bridging interaction between particles on the polymer that depends on the distance along the chain and arises after the chain conformational fluctuations have been integrated out, giving rise to the temperature dependence of  $\mathcal{F}_{\text{LRLG}}$ . The kernel,

$$G_{ij} = 4\pi\beta^{-1} \int_0^\infty dR R^2 \left[ e^{-\beta U(R)} - 1 \right] P_{ij}(R), \quad (1.19)$$

is obtained by performing a generalized virial expansion (assuming isotropy) with

$$P_{ij}(R) = \langle \delta(R - |\mathbf{X}_i - \mathbf{X}_j|) \rangle_{\text{P}} \quad (1.20)$$

the polymer distribution function. The effective free energy  $\mathcal{F}_{\text{LRLG}}$  is therefore completely defined by the polymer and particle parameters. The effective long-range interaction encoded by the kernel  $G_{ij}$  implicitly sums over all possible loops formed by the polymer segment bounded by the two bridging particles. This approach accounts exactly for two-body interactions and should therefore be valid for sufficiently low polymer monomer 3D spatial density (as in Flory-type approximations [33, 34, 49]). There will be no restriction, however, on the 1D occupation along the polymer.

The possibility that the LRLG model exhibit a phase separation, while the 1D SRLG model does not, is thus completely dependent on the asymptotic behavior  $|i - j| \rightarrow \infty$  of the kernel  $G_{ij}$ . The asymptotic behavior of  $P_{ij}(R)$  is [34]

$$P_{ij}(R) \xrightarrow{\frac{R}{R_{ij}} \rightarrow 0} \frac{c_0}{R_{ij}^3} \left( \frac{R}{R_{ij}} \right)^g, \quad (1.21)$$

where  $c_0$  is a constant and  $R_{ij} = \langle X_{ij}^2 \rangle_{\text{P}}^{1/2} = b|i - j|^\nu$  is the root-mean-square monomer  $i$ -to- $j$  distance with  $b$  the Kuhn length. The exponents  $\nu$  and  $g$  depend on the chosen polymer statistics. In the absence of the polymer, the monomers form an ideal gas and  $P_{ij}(R)$  is replaced by the inverse system volume  $V^{-1}$  in Eq. (1.19). The above approach then reduces to the usual non-ideal gas virial expansion. By contrast, particle-particle correlations arise from the polymer connectivity due to the presence of  $P_{ij}(R)$  in the kernel  $G_{ij}$ . Bound particles closer on the chain thus experience enhanced two-body interactions down to a lower limit imposed by polymer rigidity and self-avoidance.

By inserting Eq. (1.21) in (1.19), we obtain the asymptotic behavior of the long-range interaction,  $G_{ij} \sim |i - j|^{-\alpha}$  with  $\alpha = (3 + g)\nu$ . The effective 1D LRLG model clearly falls into the universality class of the well known 1D long-range Ising model (LRIM) [87], aside from an additional NN interaction that also appears in the effective inverse square LRIM approach to the Kondo problem [5]. The exponent  $\alpha$  is the key parameter to predict phase transitions in the LRIM [39]. Ferromagnetic-like phase transitions occur for a positive kernel and  $1 < \alpha < 2$  (Dyson criterion) and critical exponents are classical for  $1 < \alpha < 3/2$  [100]. The case  $\alpha = 2$  leads to the 1D analog of the Berezinsky-Kosterlitz-Thouless phase transition [5, 71].

Interestingly, the Dyson criterion depends here only on the polymer properties and it is straightforward to obtain the values of  $\alpha$  for the Gaussian and self-avoiding polymer (SAP) distributions. For a Gaussian polymer  $\nu = 1/2$  and  $g = 0$ , and therefore

$\alpha = 3/2$ . For a SAP  $\alpha \approx 1.92$ , since  $\nu \approx 0.588$  and  $g \approx 0.27$  [34]. Therefore, the Dyson criterion for  $\alpha$  is fulfilled and these two polymer models are expected to lead to phase separation. For an infinite compact globular polymer, we expect Gaussian behavior for interior monomers owing to internal screening of polymer self-avoidance [113]. Typical polymer conformational statistics therefore lead to a LR interaction decay exponent  $\alpha$  that ensures the existence of a 1D phase transition for bound particles.

### 1.4.2 The variational approach

Using a variational method [46], we proceed by finding the coexistence and spinodal curves to construct the entire LRLG phase diagram. Assuming homogeneous non-specific binding, a constant  $\epsilon_i$  can be absorbed into the definition of the chemical potential, and we rewrite the free energy  $\mathcal{F}_{\text{LRLG}}$  as the sum of two parts by introducing a variational parameter  $\mu_0$ :

$$\mathcal{F}_{\text{LRLG}}[\Phi_i] = H_0 + \Delta H, \quad (1.22)$$

where

$$H_0 = -J \sum_{i=1}^{N-1} \Phi_{i+1} \Phi_i - \mu_0 \sum_{i=1}^N \Phi_i \quad (1.23)$$

and

$$\Delta H = -\frac{1}{2} \sum'_{i,j} \Phi_i G_{ij} \Phi_j - (\mu - \mu_0) \sum_{i=1}^N \Phi_i. \quad (1.24)$$

$H_0$  is just the Hamiltonian of another 1D SRLG (see Eq. (1.15)) with an effective chemical potential  $\mu_0$  and therefore has the advantage of being exactly solvable. For  $J = 0$ , the variational method is equivalent to the MFT one, which consists in moving the NN interaction (term in  $J$ ) from  $H_0$  to  $\Delta H$  (see [31]). MFT, which incorrectly predicts a 1D phase in the absence of bridging, is improved by the optimal choice for  $\mu_0$  when  $J > 0$ , because correlation effects, missed entirely by MFT, are approximately accounted for in the variational  $H_0$ . This variational method is exact for the infinite range lattice gas (or Ising model [70, 101]) and therefore we expect it to lead to reasonably accurate results for the LRLG. The division in Eq. 1.22 leads to a trial grand potential  $\Omega_V = \Omega_0 + \langle \Delta H \rangle_0 \geq \Omega_{\text{LRLG}}$ , where  $\Omega_0$  is the grand potential related to  $H_0$  and  $\langle \cdot \rangle_0$  denotes an average with respect to  $H_0$ . In the thermodynamic limit ( $N \rightarrow \infty$ ),  $\Omega_0 = -Nk_B T \ln \lambda_+$ , where  $\lambda_+$  is the largest of the two eigenvalues  $\lambda_{\pm}$  which arise from the transfer matrix method applied to the SRLG model [88]:

$$\lambda_{\pm} = e^Y \left[ \cosh(Y) \pm \sqrt{\sinh^2(Y) + e^{-\beta J}} \right], \quad (1.25)$$

where  $Y = \beta(J + \mu_0)/2$ . The second term in  $\Omega_V$ ,

$$\langle \Delta H \rangle_0 = \frac{1}{2} \sum'_{i,j} G_{ij} \langle \Phi_i \Phi_j \rangle_0 - (\mu - \mu_0) \sum_{i=1}^N \langle \Phi_i \rangle_0, \quad (1.26)$$

involves the mean occupation in the ensemble  $H_0$ ,  $\Phi_0 \equiv \langle \Phi_i \rangle_0$ , where

$$\langle \Phi_i \rangle_0 = -\frac{1}{N} \frac{\partial \Omega_0}{\partial \mu_0} = \frac{1}{2} \left( 1 + \frac{\sinh(Y)}{\sqrt{\sinh^2(Y) + e^{-\beta J}}} \right), \quad (1.27)$$

and the two-site correlation function,

$$\langle \Phi_i \Phi_j \rangle_0 = \Phi_0^2 + \Phi_0 (1 - \Phi_0) e^{-|i-j|/\xi_{\text{LGG}}}, \quad (1.28)$$

in the thermodynamic limit with  $\xi_{\text{LGG}} = -1/\ln r_{\text{LGG}}$  the SRLG correlation length and  $r_{\text{LGG}} \equiv \lambda_-/\lambda_+$ . The optimization equation  $(\partial \Omega_V / \partial \mu_0)_{\mu_0 = \mu_0^*} = 0$  gives the optimal value  $\mu_0^*$  of  $\mu_0$ :

$$\begin{aligned} \mu - \mu_0^* &= 2\Phi_0^* [S' - S] - S' \\ &- \Phi_0^* (1 - \Phi_0^*) (1 - 2\Phi_0^*) S'' \beta \left( \frac{\partial \Phi_0}{\partial \mu_0} \right)_{\mu_0 = \mu_0^*}^{-1} \end{aligned} \quad (1.29)$$

with  $\Phi_0^* = \Phi_0(\mu_0^*)$  and where the sums  $S$ ,  $S'$  and  $S''$ , defined as  $S = \sum_{k=n_{\text{inf}}}^{\infty} G_k$ ,  $S' = \sum_{k=n_{\text{inf}}}^{\infty} G_k r_{\text{LGG}}^k$ , and  $S'' = \sum_{k=n_{\text{inf}}}^{\infty} G_k k r_{\text{LGG}}^k$ , depend crucially on the long-range behavior of the kernel  $G_{ij} = G_{i-j}$  (see [31]). The best variational approximation to the exact grand potential  $\Omega_{\text{LRLG}}$  is the optimal grand potential,  $\Omega_V^* = \Omega_V(\mu_0^*)$ , from which we obtain the average site occupation  $\Phi \equiv -N^{-1} \partial \Omega_V^* / \partial \mu$ . This last definition, along with the optimization condition, leads to  $\Phi = \Phi_0^*$  and since Eq. (1.27) can be inverted to obtain  $\mu_0^*$  in terms of  $\Phi_0^*$ , it is possible to write  $\Omega_V^*$  entirely in terms of  $\Phi$  (see [31]):

$$\begin{aligned} \frac{\Omega_V^*}{N} &= \frac{\Omega_0(\Phi)}{N} + \Phi^2 (S - S') \\ &+ \Phi^2 (1 - \Phi) (1 - 2\Phi) \beta S'' \left( \frac{\partial \Phi_0}{\partial \mu_0} \right)_{\mu_0 = \mu_0^*}^{-1}. \end{aligned} \quad (1.30)$$

We therefore obtain analytical variational expressions for the chemical potential  $\mu$ , the LRLG pressure  $P \approx -\Omega_V^*/(Nl_m)$  as functions of  $\Phi$  that can be used to obtain the coexistence and spinodal curves [61] (see [31]).

For simplicity, we illustrate our results for the case of an attractive square well (SW) particle interaction of depth  $u_0$ , range  $a$  and hard core  $\sigma$  [69, 116]. The asymptotic long distance behavior (for  $R_{ij}/b \gg 1$ ) is therefore given by  $G_{ij} \xrightarrow{|i-j| \rightarrow \infty} K_{\text{SW}} |i-j|^{-\alpha}$  where

$$\begin{aligned} K_{\text{SW}} &= 4\pi\beta^{-1} \frac{c_0}{3+g} \left( \frac{\sigma}{b} \right)^{3+g} \\ &\times \left\{ (e^{\beta u_0} - 1) \left[ \left( \frac{a}{\sigma} \right)^{3+g} - 1 \right] - 1 \right\}. \end{aligned} \quad (1.31)$$

This result allows us to illustrate generic behavior for potentials with short range repulsion and longer range attraction:  $K_{\text{SW}}$  is positive (attractive) at low enough  $T$  and

decreases monotonically with decreasing slope for increasing temperature, eventually becoming negative (repulsive) at high enough  $T$  due to short range repulsion. In the attractive regime of interest,  $K_{\text{SW}}$  increases with  $u_0$  and  $a$  and decreases with the Kuhn length  $b$ ,  $\sigma$ , and polymer exponent  $g$  because chain stiffness and polymer self-avoidance inhibit particle-particle bridging.

### 1.4.3 Application to biological systems

We apply our LRLG model with the SW potential to study phase separation in the ParABS partition system. This molecular machinery is composed of three components: a DNA sequence *parS*, and two protein species ParB and ParA. We focus on one of its key elements: the formation of ParB aggregates around *parS*. ParB proteins can bind to DNA non-specifically and specifically on the *parS* sequence [107]. Once bound to DNA, ParB proteins can mutually interact through both spreading and bridging interactions (see Fig. 1.16), which lead to the formation of ParBS partition complexes [26, 53]. Although we now have a better understanding of segregation dynamics [144], the conditions of complex formation are still poorly understood.

With our model we are now positioned to investigate whether or not the formation of ParBS complexes could be the result of a 1D phase separation between states of high and low ParB occupation on the DNA, qualitatively similar to conventional liquid-vapor phase separation. The available data for ParB allow us to parameterize the LRLG model at room temperature  $T_r = 300$  K (See Fig. 1.17 and [31]). Figures 1.17AB show the phase diagrams obtained using Gaussian polymer or SAP statistics. The co-existence and spinodal curves are obtained from the equality of pressure and chemical potential in the two phases and the divergence of the isothermal compressibility, respectively. The critical temperature is found in the limit  $\Phi \rightarrow \Phi_c = 1/2$  (see [31]). This leads to the variational critical temperature as a solution to the following implicit equation:

$$\frac{T_c^{\text{V}}}{T_r} = \frac{1}{2k_B T_r} \left[ (S_c - S'_c) \exp\left(\frac{J}{2k_B T_c^{\text{V}}}\right) - S''_c \right], \quad (1.32)$$

where the subscript c indicates quantities evaluated at the critical point. We observe that  $T_c^{\text{V}}$  grows with  $J$  (Figs. 1.17CD) and that this effect is severely overestimated by MFT, for which (see [31])

$$\frac{T_c^{\text{MFT}}}{T_r} = \frac{1}{2k_B T_r} [J + S(T_c^{\text{MFT}})]. \quad (1.33)$$

In the asymptotic kernel approximation adopted here

$$S(T) = \sum_{k=n_{\text{inf}}}^{\infty} G_k \approx K_{\text{SW}}(T) \left[ \zeta(\alpha) - \sum_{k=1}^{n_{\text{inf}}-1} \frac{1}{k^\alpha} \right] \quad (1.34)$$

with  $\zeta$  the Riemann zeta function. A simple approximation based on the weak temperature dependence of  $K_{\text{SW}}(T)$  for  $T > T_r$  and obtained by evaluating  $S$  in Eq. (1.33) at  $T_r$  explains the linear dependence of  $T_c^{\text{MFT}}$  on  $J$  for large  $J$  (see [31]). The temperature dependence of the kernel is, however, crucial in determining the critical temperature for

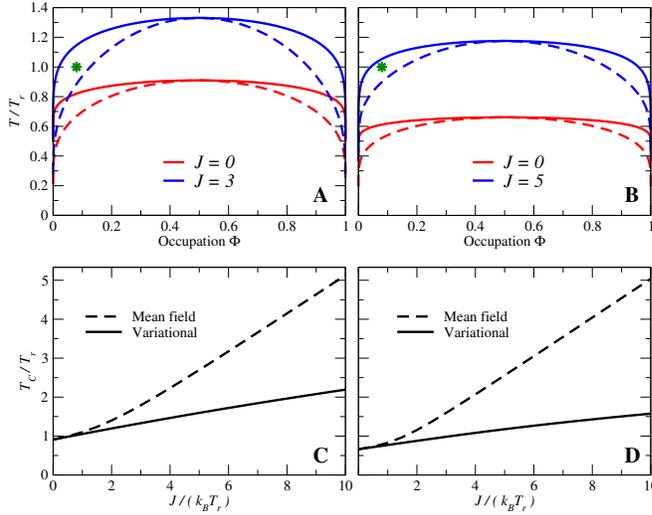


Figure 1.17: Phase diagrams for polymer-bound particles. Model parameters are (see [31]):  $l_m = 5.44$  nm,  $\sigma = l_m$ ,  $b = 23.6$  nm,  $n_{\text{inf}} = 10$ ,  $a = 2\sigma$ , and  $u_0 = 3 k_B T_r$ . Green star: biological conditions for the bacterial F-plasmid ( $\Phi = 0.08$  at room temperature  $T_r$ ). **(A)**: Gaussian polymer. Solid (dotted) line represents the coexistence (spinodal) curve for  $J = 0$  (red) and  $J = 3 k_B T_r$  (blue). **(B)**: Self avoiding polymer (SAP) with  $J = 0$  (red) and  $J = 5 k_B T_r$  (blue). **(C)**: Critical temperature  $T_c$  for the Gaussian polymer: variational approach (solid line) and MFT (dotted line). **(D)**: Same as **(C)**, but for the SAP.

small  $J$ . The variational result for the critical temperature is also close to being linear in  $J$  for large  $J$  and heuristically can be obtained from MFT by evaluating  $S$  at  $T_r$  and replacing  $J$  by  $J/3$ .

The expression (1.34) indicates how the critical temperature is crucially determined by  $n_{\text{inf}}$ , the polymer persistence length in site number, by reducing the weight of the LR interaction contribution [93, 145]. In Fig. 1.17, the lower  $T_c$  shown by the SAP compared with the Gaussian polymer at constant  $J$  is due to the faster decay of the LR interaction (larger  $\alpha$ ), despite the larger value of the SAP  $K_{\text{SW}}$  (see [31]).  $T_c$  is non-zero even for  $J = 0$ , but is far below room temperature. Therefore, the system does not exhibit phase separation without spreading interactions at this temperature. Both short range spreading with reasonable biological values for  $J$  ( $\sim 3$ - $6 k_B T_r$ ) and long-range bridging interactions are thus required at room temperature to form ParB condensates in our model, as suggested by Monte Carlo simulations [26] and experiments [32, 53].

The ParABS system ensures the segregation in *E. coli* of relatively short circular DNA strands called F-plasmids. For an F-plasmid of linear size  $\sim 60$  kbp and an average number of 300 ParB [16], the mean occupation is  $\Phi \approx 0.08$ . Its position in the phase diagram (green star in Fig. 1.17A-B) shows that for reasonable values of  $J$  the system may exist in the low occupation metastable coexistence region at room temperature, providing a plausible explanation for the experimental observations [112]:

without the *parS* sequence, experiments show a homogeneous ParB distribution in the cell, while with *parS* a ParBS complex forms. Thus, *parS* could provide the energy required to overcome the nucleation barrier and allow the system to switch from the metastable homogeneous state to the stable coexistence phase, in which ParB proteins form a stable cluster on the DNA around *parS*. Experimentally, this system should follow the conventional behavior of liquid-vapor phase transitions: (i) in the low occupation metastable region, the system can form relatively high density ParBS complexes with only a small total number of intracellular proteins, and (ii) ParB over- or under-expression will favor or repress the formation of ParBS complexes depending on the position in the phase diagram. Indeed, systems without *parS* but with sufficiently high ParB occupation would be in the unstable coexistence area and should therefore form protein (*liquid*) droplets spontaneously in a low occupation (*vapor*) background, the homogeneous state being unstable in this case. On the contrary, systems with too few ParB proteins would be in the low occupation *vapor* region, losing the ability to form complexes even in the presence of *parS*.

In this section, we proposed a general theoretical framework for the physics of particles interacting on a polymer fluctuating in 3D that leads naturally to an effective 1D LRLG model. We established a criterion for the existence of a 1D phase transition based on the exponent  $\alpha$  controlling the asymptotic decay of the LR interactions, which depends only on the polymer exponents  $\nu$  and  $g$ . Since this criterion is satisfied for standard polymer models, the conformational fluctuations of linear structures like DNA produce effective 1D long-range interactions between bound particles that lead to 1D particle phase separation along the polymer. We used our theoretical approach to construct the whole phase diagram of the ParBS bacterial DNA segregation system and concluded that the formation of ParBS complexes results from activated phase separation in the low ParB occupation metastable region. This general mechanism for triggering the formation of polymer-bound protein complexes via small nucleation sites may play an important role in membrane-less cell compartmentalization.

Our method may also be used to derive the 1D particle distribution along the polymer and the 3D particle density of the condensate that forms around a specific binding site, both of which are accessible experimentally [112, 32]. Finally, to facilitate quantitative testing of the present model, it would also be of great interest to find an *in vitro* biomimetic system of interacting polymer-bound particles that could be studied experimentally.

## Chapter 2

# Positioning of complexes in liquid-like phase

Controlled motion and positioning of colloids and macromolecular complexes in a fluid, as well as catalytic particles in active environments, are fundamental processes in physics, chemistry and biology with important implications for technological applications [150, 94]. In this Chapter, we focus on an active biological system for which precise experimental results are available. Our work is fully inspired by studies of one of the most widespread and ancient mechanisms of liquid phase macromolecular segregation and positioning known in nature: bacterial DNA segregation systems. Despite the fundamental importance of these systems in the bacterial world and intensive experimental studies extending over 30-years [50, 112, 81], no global picture encompasses fully the experimental observations.

Partition systems encode only three elements that are necessary and sufficient for active partitioning, as sketched in Fig.2.1: two proteins ParA and ParB, and a specific sequence *parS* encoded on DNA. As explained in the previous Chapters, the pool of ParB proteins is recruited as a cluster of spherical shape centered around the sequence *parS*, forming the ParBS partition complex [112]. These ParBS cargos interact with ParA bound onto chromosomal DNA (ParA-slow) [82, 18], triggering unbinding of ParA by inducing conformational changes through stimulation of adenosine triphosphate (ATP) hydrolysis and/or direct ParB-ParA contact [133], and thereby allowing ParA diffusion in the cytoplasm (ParA-fast) [81]. This process entails the oscillation of ParA from pole to pole and the separation of the ParBS partition complex into two complexes with distinct sub-cellular trajectories and long-term localization. Overall, these interactions result in an equidistant, stable positioning of the duplicated DNA molecules along the cell axis.

The specific modeling of ParABS systems falls into two categories: either “filament” (pushing/pulling the cargos, similar to eukaryotic spindle apparatus [50]) or reaction-diffusion models [132, 133, 134, 63, 85, 124, 65, 115]. Recent superresolution microscopy experiments have been unable to observe filamentous structures of

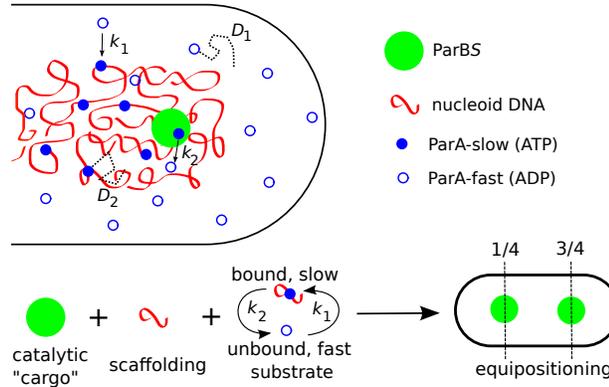


Figure 2.1: The proteophoresis model consists in interactions between solely three actors: (i) *ParBS*, the catalytic particles constituting the “cargo”, (ii) the DNA which serves as a scaffolding for *ParA* and (iii) *ParA* (fast/ADP and slow/ATP), a motor proteins constituting a substrate. *ParA*-fast and slow have respectively a diffusion coefficient  $D_1$  and  $D_2$  such that  $D_2 \ll D_1$ . *ParA*-fast can be converted in *ParA*-slow at a rate  $k_1$  anywhere in the cell by recovery of ATP, whereas the hydrolysis of *ParA*-fast occurs at a rate  $k_2$  at the contact with *ParBS*:the cargo catalyzes hydrolysis during a transient attractive interaction. These elements are sufficient to lead to the equipositioning of *ParBS* along the cell axis.

*ParA* [81, 85], disfavoring polymerization-based models [63]. Reaction-diffusion models have been mainly investigated numerically to describe experimental observations like single or multiple *ParBS* complex positioning. In most cases, these models require other assumptions - such as DNA elasticity [85, 124] - as simple reaction-diffusion mechanisms are not sufficient to predict proper positioning. Other reaction-diffusion models considered the dynamics of the partition complex on the surface of the nucleoid [132, 133, 134, 65]. Recent experiments, however, demonstrate that partition complexes and *ParA* translocate through the interior of the nucleoid, not at its surface [81].

Recently, in the context of the active colloids literature, there have been attempts to describe the *ParABS* system using models inspired by the diffusiophoresis [7, 6] of active colloidal particles in solute concentration gradients [123, 94]. These works have several important limitations for applications to *ParABS*, such as: rigid spherical particles (with surface reactions only), the steady-state approximation, only one *ParA* population, or reproducing equilibrium positioning only. The full dynamical behavior of the coupled system (*ParBS* cargo coupled to *ParA*) has thus not been elucidated.

Here we propose a general model of reaction-diffusion for *ParA* coupled to the overdamped motion of *ParBS*. Our continuum reaction-diffusion approach goes beyond the previous diffusiophoretic mechanisms [123, 63, 65, 124] by accounting for the finite diffusion of *ParA*-slow and *ParA*-fast, as well as the interaction of *ParA*-slow with the entire volume of *ParBS* partition complexes. Volumetric interactions are suggested by our recently developed “nucleation and caging” model [112, 32], which

accounts for both the formation of ParBS and the distribution of ParB in the spatial vicinity of *parS* specific DNA sites : the conformation of the plasmid is well described by a fluctuating polymer and the weak ParB-ParB interactions lead to foci of low density [112, 32]. The chromosome is thus likely to enter ParBS with bound ParA-slow thereby allowing for volumetric interactions. Such a volumetric interaction should also find useful applications in the field of porous catalytic particles. On the other hand, allowing for finite diffusion coefficients permits describing analytically the global dynamical picture of the model, contrary to previous numerical studies often restricted to a limited range of parameters. In particular this enables us to predict a dynamical transition between stable and unstable regimes. We observe that biological systems are generally close to the instability threshold. The ParABS system of the F-plasmid lies just below, enabling efficient positioning and precursor oscillations of ParA. Other ParABS systems ([124] and Refs. therein) could be just above, providing an explanation for the observed out-of-phase ParBS and ParA oscillations. Our model accounts for both these regimes.

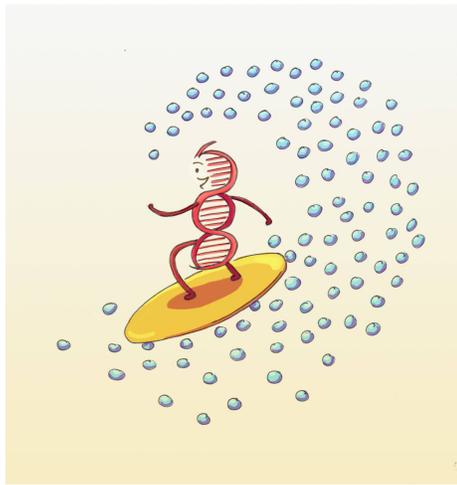


Figure 2.2: One could say that DNA (in red) is surfing on proteins (blue circles). Credit: Alexandra Jeullard.

## 2.1 The Proteophoresis model

### 2.1.1 Definition

We translate the interactions sketched in Fig.2.1 into a model of reaction-diffusion. The ParA protein population is described by two coupled density fields:  $u(\mathbf{r}, t)$  for the hydrolysed ParA-fast proteins, assumed to be unbound and diffusing rapidly within the nucleoid, and  $v(\mathbf{r}, t)$  for the non-hydrolysed ParA-slow molecules, which are bound dynamically to the nucleoid and diffuse more slowly. These two species are coupled via a system of reaction-diffusion equations: the rapid species  $u$  converts into

the slow one with a constant rate  $k_1$ , while the slow species  $v$  is hydrolysed in the presence of the ParBS partition complexes located on DNA, with a rate  $k_2$  (typically  $k_1 \approx 0.02 \text{ s}^{-1}$  [132] and  $k_2 \approx 68.5 \text{ s}^{-1}$  [63]). The ParBS assemblies form 3D-foci complexes [112, 32] and interact with the ParA-slow proteins. The interaction probability is described by the profiles  $S(\mathbf{r} - \mathbf{r}_i(t))$  centered around the ParBS positions  $\mathbf{r}_i(t)$ . These profiles play a double role: (i) they act as catalytic sources in the reaction-diffusion equations, triggering ParA-slow hydrolysis with the rate  $k_2$  and (ii) they determine a feedback “proteophoresis” (volumetric) force, in contrast with chemophoresis forces that occur in general only at the complex surface. In what follows, the function  $S(\mathbf{r})$ , representing an idealized density profile of ParBS, is assumed to be symmetric with a compact support of width  $\sigma$  and a unit value at its maximum. The dynamics of the protein population is therefore described by the coupled reaction-diffusion equations:

$$\begin{aligned}\frac{\partial u}{\partial t} &= D_1 \Delta u - k_1 u(\mathbf{r}, t) + k_2 v(\mathbf{r}, t) \sum_i S(\mathbf{r} - \mathbf{r}_i(t)), \\ \frac{\partial v}{\partial t} &= D_2 \Delta v + k_1 u(\mathbf{r}, t) - k_2 v(\mathbf{r}, t) \sum_i S(\mathbf{r} - \mathbf{r}_i(t)).\end{aligned}\quad (2.1)$$

In these equations, in which we do not invoke the steady-state approximation (cf. [94]),  $D_1$  and  $D_2$  represent the diffusion constants of the fast and slow species, respectively  $u$  and  $v$ . The sum runs over the ParBS positions  $\mathbf{r}_i(t)$ . The density fields are subjected to reflecting boundary conditions  $\nabla u \cdot \mathbf{n}|_{\partial V} = 0$  and  $\nabla v \cdot \mathbf{n}|_{\partial V} = 0$ , where  $\mathbf{n}$  is a unit vector normal to the cell boundary  $\partial V$ . The system described by Eqs.(2.1) together with these boundary conditions on  $u$  and  $v$  ensure total ParA protein number conservation. Note that ParA proteins can freely penetrate the partition complexes, which do not form barriers for diffusion.

The nonlinear coupling in the system is introduced by the forces driving the partition complexes, which are modeled as Brownian particles in an active medium. The back reaction on each complex is described by a “proteophoresis force” due to the ParA-slow concentration gradient acting on the whole volume of the complex. In the viscous medium prevailing in a cell, we do not expect inertial terms to be important. Neglecting in the first approximation the stochastic and confining forces, the dynamic equation for the  $i^{\text{th}}$  complex then read

$$m\gamma \frac{d\mathbf{r}_i}{dt}(t) = \varepsilon \int_V \nabla v(\mathbf{r}, t) S(\mathbf{r} - \mathbf{r}_i(t)) d^3\mathbf{r}.\quad (2.2)$$

Note that no direct coupling between complexes has been introduced. The constant  $\varepsilon$  represents the energy of interaction between a single ParA-slow protein and the ParBS partition complex. Hence, the order of magnitude of  $\varepsilon$  is a fraction of the energy released by the ATP hydrolysis ( $\simeq 10k_B T$ ). The drag force coefficient  $m\gamma$  is related to an effective diffusion constant of the complex  $D_{\text{pc}}$  by the Einstein relation  $m\gamma = k_B T / D_{\text{pc}}$ . Thanks to attractive protein-protein interactions (leading to hydrolysis energy consumption) the interaction energy  $\varepsilon$  in (2.2) is positive, and the corresponding proteophoresis force, and resulting motion, is in the direction of increasing

ParA density gradient. In the following, we will use the dimensionless coupling constant:  $\alpha \equiv \varepsilon/m\gamma D_2 = (\varepsilon/k_B T)(D_{pc}/D_2)$ . From numerical simulations it appears that the stochastic force does not affect crucially the main system dynamics. Superresolution microscopy [81] indicates that the partition complex motion is confined to the cell symmetry axis, i.e. within the bacterial nucleoid. Therefore, in the minimal model we limit the study of its dynamics to one dimension and denote by  $x$  the coordinate along the cell axis,  $-L \leq x \leq L$ , where  $2L$  is the cell length.

## 2.1.2 Restoring proteophoresis force

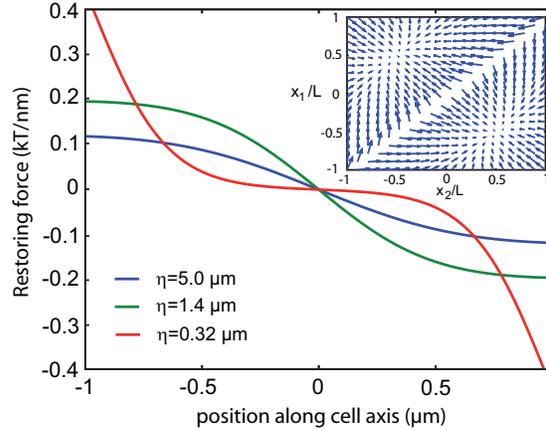


Figure 2.3: Proteophoresis force (Eq.(2.7)) for different values of the screening length  $\eta = \sqrt{D_1/k_1}$  (variable  $k_1$ ) with the other biological parameters fixed (see below). The curve in blue is plotted using physiological values ( $k_1 = 0.04 \text{ s}^{-1}$ ) and shows a marked restoring force gradient toward mid-cell positions: for  $\eta = 0.32, 1.4$  and  $5 \mu\text{m}$ , the force produces a parabolic potential well of depth  $\sim 0, 6$  and  $4 \text{ kT}$ , respectively, over a half-width of  $0.25 \mu\text{m}$  (note the non-monotonic behavior with the equilibrium position restoring force vanishing for both zero and infinite  $k_1$ , see text below). Inset: Proteophoresis force field in the phase space  $(x_1/L, x_2/L)$  of two partition complex positions leading to quarter-cell positions.

The model provides all the necessary ingredients for proper partition complex positioning. We first look for stationary solutions when a single partition complex is present within the cell at position  $x_1$ . In order to keep the algebra simple, we approximate the profile function  $S(x - x_1)$  by a Dirac-delta distribution  $\sigma\delta(x - x_1)$ , where the amplitude  $\sigma$  is the typical interaction volume of the complex. The stationary solutions

of Eqs.(2.1) with reflecting boundary conditions then reads:

$$\begin{cases} u(x) = A \frac{\cosh(q(L+x))}{\cosh(q(L+x_1))} & \text{for } -L \leq x < x_1, \\ u(x) = A \frac{\cosh(q(L-x))}{\cosh(q(L-x_1))} & \text{for } x_1 < x \leq L, \\ v(x) = C - \frac{D_1}{D_2} u(x), \end{cases} \quad (2.3)$$

where  $q \equiv \sqrt{k_1/D_1}$ . The  $x_1$  dependent constants  $A$  and  $C$  in (2.3) can be easily computed by the gradient discontinuity at  $x_1$ ,

$$D_1(\partial_x u|_{x_1^+} - \partial_x u|_{x_1^-}) = -k_2 \sigma v(x_1), \quad (2.4)$$

and by the conservation of the total number of ParA monomers:  $\partial_t \int_V (u+v) d^3\mathbf{r} = 0$ . We obtain:

$$A = \frac{C_0}{\frac{D_1}{D_2} + \left(\frac{2Lk_1}{\sigma k_2} + 1 - \frac{D_1}{D_2}\right) \frac{\tanh(q(L+x_1)) + \tanh(q(L-x_1))}{2qL}}, \quad (2.5)$$

and

$$C = C_0 \frac{\frac{D_1}{D_2} + \frac{D_1 q}{\sigma k_2} (\tanh(q(L+x_1)) + \tanh(q(L-x_1)))}{\frac{D_1}{D_2} + \left(\frac{2Lk_1}{\sigma k_2} + 1 - \frac{D_1}{D_2}\right) \frac{\tanh(q(L+x_1)) + \tanh(q(L-x_1))}{2qL}}. \quad (2.6)$$

For a delta-like complex profile, the force acting on a static partition complex located at  $x_1$  is proportional to the mean value of the ParA-slow density gradient at  $x_1$ :

$$\begin{aligned} F(x_1) &= \frac{\varepsilon\sigma}{2} (\partial_x v|_{x_1^+} + \partial_x v|_{x_1^-}), \\ &= \frac{1}{2} \alpha m \gamma \sigma D_1 q A (\tanh q(L-x_1) - \tanh q(L+x_1)). \end{aligned} \quad (2.7)$$

This result shows that the unique equilibrium position of the complex is located at the cell center, i.e.  $x_1 = 0$ . An important feature of the resulting force mediated by the ParA density distribution gradient is its finite range. Its screening length, given by  $\eta = 1/q = \sqrt{D_1/k_1}$ , is illustrated in Fig.2.3, where the force  $F(x)$  is plotted for different values of  $\eta$ . Clearly, the proteophoresis force, here estimated of the order of the picoNewton ( $\approx 0.25 kT/\text{nm}$ ), is sensed by the partition complex only if its distance to the cell boundary or to a neighboring complex is less than  $\eta$ . Note that the above quasistatic (adiabatic) analysis is valid only when the ParA distribution instantaneously adapts to the complex position (cf. [94]). The restoring character of the force, Eq.(2.7), then makes the symmetric position  $x_1 = 0$  stable.

For bacterial cells containing several partition complexes, the sum over their positions in Eqs.(2.1) generates an effective indirect interaction among them that, together with the boundary conditions and protein number conservation, brings the system to an equilibrium state with highly symmetric complex positions. For instance, when two complexes are present within the cell (as would be the case after a DNA replication event) the equilibrium positions are found to be located at  $x_1 = -L/2$  and  $x_2 = L/2$ , i.e. the “1/4” and “3/4” positions in terms of the cell axis length  $2L$ . A phase portrait of the system in the  $(x_1, x_2)$  coordinates (see inset of Fig.2.3) clearly indicates the stable nature of these positions.

## 2.2 Stability analysis

Analytical and numerical studies of Eqs. (2.1-2.2) show that stationary solutions (irrespective of the number of complexes) become unstable in cells where the ParA density profiles can develop large gradients. The concentration profiles and the partition complex start travelling together at a constant velocity  $c_{TW}$ , as if partition complexes were self-propelled by “surfing” on the ParA distribution wave they have themselves generated to eventually bounce back and forth in presence of cell boundaries. This strongly suggests the existence of travelling waves (TW) in an unbounded system or in finite-size cells whose length  $2L$  is much larger than the screening length  $\eta$ .

### 2.2.1 Traveling Wave (TW) Ansatz

For one complex, we look for solutions of Eqs. (2.1-2.2) in the TW form  $u(x, t) = u(\xi)$ ;  $v(x, t) = v(\xi)$ , where  $\xi = x - c_{TW}t$  is the wave comoving reference coordinate, with the asymptotic conditions  $u(\xi) \rightarrow 0$  and  $v(\xi) \rightarrow v_\infty = C_0$  when  $\xi \rightarrow \pm\infty$ . The resulting system of ordinary differential equations admits analytical solutions for a Dirac partition complex profile  $S(\xi) = \sigma\delta(\xi)$ , and the equation of motion of the partition complex (2.2) takes the form  $c_{TW} = \alpha D_2 \int \partial_\xi v(\xi) S(\xi) d\xi$  and provides a nonlinear relation for determining the wave celerity  $c_{TW}$ . The system to solve is:

$$D_1 u'' + c_{TW} u' - k_1 u + k_2 \sigma v \delta(\xi) = 0, \quad (2.8)$$

$$D_2 v'' + c_{TW} v' + k_1 u - k_2 \sigma v \delta(\xi) = 0, \quad (2.9)$$

$$c_{TW} = \frac{\alpha D_2 \sigma}{2} (v'(0^+) + v'(0^-)), \quad (2.10)$$

where the primes denote  $\xi$ -derivatives. By symmetry, if a TW wave exists for  $c_{TW}$ , it exists also for  $-c_{TW}$ . We therefore restrict ourselves to  $c_{TW} \geq 0$ . For more general shapes than  $S(\xi) = \sigma\delta(\xi)$ , solutions are easily obtained numerically.

### 2.2.2 TW solution

We solve Eq.(2.8) for  $u(\xi)$  requiring that  $u$  and its derivatives vanish at infinity. Integrating Eq.(2.8) over a small interval around  $\xi = 0$  shows that  $u$  is continuous at  $\xi = 0$  but that its derivative verifies the jump condition:

$$D_1 (u'(0^+) - u'(0^-)) = -k_2 \sigma v(0), \quad (2.11)$$

leading to

$$u(\xi) = A e^{r-\xi}, \quad \xi \geq 0, \quad (2.12)$$

$$u(\xi) = A e^{r+\xi}, \quad \xi \leq 0, \quad (2.13)$$

where

$$r_\pm = \frac{-c_{TW} \pm \sqrt{c_{TW}^2 + 4D_1 k_1}}{2D_1}. \quad (2.14)$$

The prefactor  $A$  is obtained from the jump condition Eq.(2.11):

$$A = \frac{k_2 \sigma v(0)}{\sqrt{c_{TW}^2 + 4D_1 k_1}}. \quad (2.15)$$

We now solve Eq.(2.9) for  $v(\xi)$ . We add Eqs.(2.8) and (2.9) and integrate once to obtain:

$$D_2 v' + c_{TW} v = c_{TW} v_\infty - D_1 u' - C u, \quad (2.16)$$

where the constant  $v_\infty = C_0$  is the value of the field  $v$  at infinity which is the (nonzero) concentration of  $v$ -particles (slow parA) away from the TW (note that this value cannot be 0 unless the solution for  $v$  and  $u$  is identically 0). It equates the total ParA concentration  $C_0$  since  $u$  vanishes at infinity. We require  $v$  to be bounded at infinity and as it is continuous through  $\xi = 0$ , we obtain:

$$v(\xi) = v_\infty + B e^{-\frac{c_{TW} \xi}{D_2}} - A \frac{D_1 r_- + c_{TW}}{D_2 r_- + c_{TW}} e^{r_- \xi}, \quad \xi \geq 0 \quad (2.17)$$

$$v(\xi) = v_\infty - A \frac{D_1 r_+ + c_{TW}}{D_2 r_+ + c_{TW}} e^{r_+ \xi}, \quad \xi \leq 0 \quad (2.18)$$

where  $B$  is given by the continuity condition of  $v$  at  $\xi = 0$ :

$$B = \frac{A c_{TW} (D_2 - D_1) \sqrt{c_{TW}^2 + 4D_1 k_1}}{c_{TW}^2 (D_1 - D_2) - D_2^2 k_1}. \quad (2.19)$$

Finally, using for instance Eqs.(2.18) and (2.15), we obtain:

$$v(0) = \frac{v_\infty}{1 + \frac{k_2 \sigma}{\sqrt{c_{TW}^2 + 4D_1 k_1}} \frac{D_1 r_+ + c_{TW}}{D_2 r_+ + c_{TW}}}. \quad (2.20)$$

Using  $v(0)$  in Eq.(2.15) yields the constant  $A$  as a function of the parameters and  $c_{TW}$  only. This expression can then be used to determine  $B$  from Eq.(2.19). We have therefore obtained an exact expression for  $u$  and  $v$  as a function of the TW velocity  $c_{TW}$  and the parameters defining the model.

For illustration, we plot typical ParA distributions calculated from Eqs.(2.12-2.13) and (2.17-2.18) in Fig.2.4A.

### 2.2.3 Stability criterion

We obtain an equation for  $c_{TW}$  with Eqs.(2.10) and (2.17-2.18):

$$c_{TW} = -c_{TW} \frac{\alpha D_2 \sigma}{2} \left( \frac{B}{D_2} + A \frac{k_1 (2D_1 - D_2)}{c_{TW}^2 (D_1 - D_2) - D_2^2 k_1} \right). \quad (2.21)$$

(i) *The static solution*  $c_{TW} = 0$

Obviously,  $c_{TW} = 0$  is one of the solution of Eq.(2.21). This corresponds to the static profiles is given by Eqs.(2.12-2.13) and (2.17-2.18) in the limit  $c_{TW} \rightarrow 0$ .

In this limit, from Eqs.(2.15), (2.19) and (2.20) for  $A$ ,  $B$  and  $v(0)$  resp., we have  $v(0) \rightarrow v_\infty/(1 + \Delta/\nu)$ ,  $A \rightarrow v_\infty/(\nu + \Delta)$  and  $B \rightarrow 0$  with  $\Delta = D_1/D_2$  and  $\nu = 2k_1\eta/(k_2\sigma)$  ( $\eta = q^{-1}$  being the screening length defined previously). The static solutions are:

$$u_s(x) = \frac{v_\infty}{\nu + \Delta} e^{-q|x|} \quad (2.22)$$

$$v_s(x) = v_\infty - \Delta u_s(x) \quad (2.23)$$

corresponding to the limit  $L \rightarrow \infty$  in the system of equations (2.3).

(ii) TW solution  $c_{TW} > 0$

We solve Eq.(2.21) for  $c_{TW} > 0$  for  $\alpha$ :

$$\tilde{\alpha} = \alpha\sigma C_0 = \tilde{c}_{TW}\nu + \frac{\Delta(\sqrt{1 + \tilde{c}_{TW}^2} + \tilde{c}_{TW}) + \nu}{(2\Delta - 1)\sqrt{1 + \tilde{c}_{TW}^2} + \tilde{c}_{TW}}, \quad (2.24)$$

with the dimensionless  $\tilde{\alpha} = \alpha\sigma C_0$  and  $\tilde{c}_{TW} = c_{TW}/(2\sqrt{k_1 D_1})$ . For  $\Delta > 1$  (biologically,  $\Delta \approx 100$ ), we can use the former expression in the limit  $c_{TW} \rightarrow 0$  to calculate the critical value  $\alpha_c$  determining the TW existence threshold. We obtain:

$$\alpha_c = \frac{\Delta + \nu}{C_0\sigma(\Delta - 1/2)}, \quad (2.25)$$

which takes the value  $\alpha_c \approx 1/(C_0\sigma)$  for the biological values  $\Delta \sim 100 \gg \nu \sim 1$ . We plot in Fig.2.4B  $\tilde{\alpha}$  versus  $\tilde{c}$  for  $\Delta = 100$  and  $\nu = 1$ .

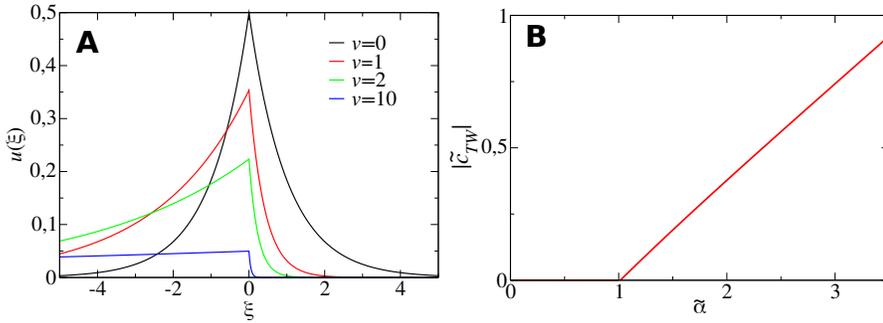


Figure 2.4: (A) Travelling wave (TW) profile  $u(\xi)$  versus the distance  $\xi$  from the complex after Eqs.(2.12-2.13) and (2.17-2.18). Different curves correspond to different TW velocities  $v = 0$  ( $K \leq 1$ ), 1, 2 and 10. (B)  $|c_{TW}|$  plotted versus  $\tilde{\alpha}$  after Eq.(2.24) for  $\nu = 1$  and  $\Delta = 100$  (same order as biological values). A critical value  $\tilde{\alpha}_c = 1$  separates two regions: (i) for  $\tilde{\alpha} < 1$ ,  $|c_{TW}| = 0$  corresponding to the stability regime and (ii) for  $\tilde{\alpha} > 1$ ,  $|c_{TW}| > 0$  and ParBS starts to move according to a traveling wave.

We have studied a vast class of similar models derived from the present one, where the analysis is made simpler by the presence of a single ParA density field. It turns out that, in these simplified models, the stability of the static solutions is directly related to the absence of travelling waves with non zero celerity. In the proteophoresis model, the expressions for the velocities of the complex  $V_S$  and of the ParA-fast distribution  $V_v$  can be written, respectively, as follows. The “escape velocity” of the complex  $V_S$  is defined as the velocity of the complex when shifted from its equilibrium position by a small quantity  $\epsilon$  (i.e., at the position  $x_1 + \epsilon$ ), the distribution of  $v$  being frozen (quasi-static hypothesis). This velocity is obtained from the distributions of  $u$  and  $v$  in Eqs.(2.3) using an unbounded system and again a Dirac profile for the density of ParB:

$$V_S = A \alpha \sigma q D_1 = \frac{\alpha \sigma q C_0 D_2}{1 + \frac{2q D_2}{k_2 \sigma}}. \quad (2.26)$$

Although it is not possible to define a velocity in a simple diffusion equation context, a quantity homogeneous to a velocity  $V_v$  can be derived in the case of reaction-diffusion equations. This quantity should be interpreted as a velocity of the ParA density distribution rearrangement after a perturbation. It is instructive to consider first a single-component ParA system. The single-field equation (outside of the complex where  $S(x - x_1)$  is negligible), replacing the two-species system of Eqs.(2.1) is now:

$$\frac{\partial v}{\partial t} = D \frac{\partial^2 v}{\partial t^2} - kv, \quad (2.27)$$

whose Green function is given by  $G(x, t) \propto \exp(-\frac{x^2}{4Dt} - kt)$ . This enables us to identify a quantity homogeneous to a speed  $V_v = 2\sqrt{kD}$ .  $V_v$  is interpreted as the spreading distance of the distribution during its life-time. In this case, the criterion of stability can be shown to be *exactly*  $V_S < V_v$  for an infinite system ( $L \rightarrow \infty$ ).

For a two-component model, the exact condition in Eq.(2.24) can be written in the same form as for the single-component model ( $V_S < V_v$ ) with the expression of  $V_S$  given by Eq.(2.26). This allows us to identify the expression for  $V_v$ :

$$V_v = \frac{\sqrt{\frac{k_1}{D_1}} D_2}{1 - \frac{1}{2} \frac{D_2}{D_1}}, \quad (2.28)$$

which reduces to  $V_v \approx qD_2$  for the biological condition  $D_2 \ll D_1$ . This produces the stability criterium found in the stability analysis above:

$$\alpha < \alpha_c = \frac{1}{\sigma C_0} \left( \frac{1 + \frac{2qD_2}{k_2\sigma}}{1 - \frac{D_2}{2D_1}} \right) \approx \frac{1}{\sigma C_0}, \quad (2.29)$$

where the last approximation has been obtained from the inequalities  $D_2 \ll D_1$  and  $2D_2q \ll k_2\sigma$ , which are satisfied for the biological values of the model parameters (see Table 2.1).

## 2.3 Comparison with experiments

A reasonable set of values for the physical constants of our model has been taken from the literature:

Table 2.1: Biological values for the parameters of the model from different methods: experiments *in vivo*, *in vitro* and *in silico*. The linear density of ParA,  $C_0$  (expressed as the number of proteins per inverse length in  $\mu\text{m}^{-1}$ ), is obtained by considering the volumetric density of 2,400 monomeric ParA proteins per  $\mu\text{m}^3$  averaged on the cell cross-section. Here the celerity of the partition complex  $V_S$  is given for the segregation phase.

Physical quantity	notation	value	methods & references
ParA-fast diffusion constant	$D_1$	$\simeq 1 \mu\text{m}^2 \cdot \text{s}^{-1}$	simulations [63] & <i>in vivo</i> [73, 81]
ParA-slow diffusion constant	$D_2$	$\simeq 0.01 - 0.05 \mu\text{m}^2 \cdot \text{s}^{-1}$	simulations [63] & <i>in vivo</i> [81]
ParA-fast to ParA-slow conversion rate	$k_1$	$\simeq 0.02 - 0.05 \text{s}^{-1}$	<i>in vitro</i> [132]
ParA-slow to ParA-fast conversion rate on the complex	$k_2$	$36 \text{s}^{-1}$ & $\simeq 68.5 \text{s}^{-1}$	<i>in vivo</i> [2] & simulations [63]
mean linear cell concentration of monomeric ParA	$C_0$	$\simeq 1,200 \mu\text{m}^{-1}$	<i>in vivo</i> [16]
celerity of ParBS	$V_S$	$\simeq 0.007 \mu\text{m} \cdot \text{s}^{-1}$	<i>in vivo</i> [52]
half width of ParBS	$\sigma$	$\leq 0.035 \mu\text{m}$	super resolution microscopy [112]

Table 2.2: Numerical values used in the figures. In Fig.2.3, the results are obtained in a quasi-static hypothesis, thus  $V_S$  and  $\alpha$  do not play a role. However, the interaction energy  $\varepsilon$  sets the scale of the force. In Figure 2.5(a), the kymograph is obtained using a single history of thermal noise with a random force of amplitude having a Gaussian distribution with a width of the order of one picoNewton. In Figure 2.5(b), the size of the cell is changing linearly with respect to time, to mimic the experiments in [81]; the density  $C_0$  is changing accordingly. Note that the quantities  $\alpha$  and  $V_S$  are not independent, but we give both values for clarity.

Quantity (units)	Fig.2.3	Fig.2.5(a) and (c)	Fig.2.6
$L$ ( $\mu\text{m}$ )	1	[0.5; 1.5] (growing cell)	1
$D_1$ ( $\mu\text{m}^2 \text{s}^{-1}$ )	1	1	1
$D_2$ ( $\mu\text{m}^2 \text{s}^{-1}$ )	0.04	0.04	0.04
$k_1$ ( $\text{s}^{-1}$ )	0.04 (biological value), 0.5 and 10	0.04	0.04
$k_2$ ( $\text{s}^{-1}$ )	40	40	40
$C_0$ ( $\mu\text{m}^{-1}$ )	1200	1200	1200
$V_S$ ( $\mu\text{m} \text{s}^{-1}$ )	–	$\approx 2.10^{-3}$	(a) $\approx 2.10^{-3}$ (b) 0.005 (c) 0.1
$\sigma$ ( $\mu\text{m}$ )	0.03 (Dirac)	0.03 (Gaussian)	0.03 (Square)
$\alpha$	undefined ( $\varepsilon = 4k_B T$ )	0.024	(a) $\alpha = 0 < \alpha_c$ (b) $0.024 \approx \alpha_c = 0.028$ (c) $0.04 > \alpha_c$

In the results of simulations presented in the Figures, we used the set of parameters in Table 2.3.

We start to plot the density profiles with a more physical rectangular complex profile in Fig.2.5. We note that the ParA distributions becomes unstable at a threshold value  $\alpha_c$  Fig.2.5(b) and (c) with  $\alpha_c \lesssim \alpha$  and  $\alpha_c \ll \alpha$ , resp.

This result is in excellent agreement with experimental observations [81, 51], and can describe even more complex experimental situations with multiple ParBS, see some examples in Fig.2.6.

*Bacterial segregation system is tuned near the instability threshold.* We now estimate the parameter  $\alpha$  which controls the stability regime. The quantity  $\alpha$  is directly related to the speed of segregation in Eq.(2.26), and thus this value is obtained from experimental data in Table 2.1:  $C_0 \sim 1, 200 \mu\text{m}^{-1}$ ,  $\sigma = 0.03 \mu\text{m}$ ,  $q = \sqrt{k_1/D_1} = \sqrt{0.04 \text{s}^{-1}/1 \mu\text{m}^2 \text{s}^{-1}} = 0.20 \mu\text{m}^{-1}$ ,  $V_S = 0.007 \mu\text{m} \text{s}^{-1}$  and  $D_2 = 0.04 \mu\text{m}^2 \text{s}^{-1}$ . We find the critical value  $\alpha_c \approx 1/(\sigma C_0) \approx 2.8 \cdot 10^{-2}$ , and the biological system displays the value  $\alpha = V_S/C_0 \sigma q D_2 \approx 2.4 \cdot 10^{-2}$ . From the stability criterion Eq.(2.29), we obtain  $\alpha \approx 0.9 \times \alpha_c$ . Note, however, that our conclusion here is semi-quantitative:  $\alpha$  appears to be of the same order of magnitude as the critical coupling  $\alpha_c$ . This leads us to suggest that the system is tuned to be near the critical dynamical threshold, permitting us to account for the main experimental observation concerning the F-plasmid: robust positioning of the partition complexes accompanied by precursor oscillations of ParA. Biological systems very often evolve near stability limits in order to efficiently switch between different regimes [62]. We can draw several conclusions from these values:

(i) The value of the screening parameter  $q \simeq 0.20 \mu\text{m}^{-1}$  and the corresponding screening length  $\eta \simeq 5 \mu\text{m}$  show that there is no screening in the volume of the cell for these values since the half-length of the cell is  $L \simeq 1 \mu\text{m}$ . Segregation and equipositioning of the complexes are thus possible.

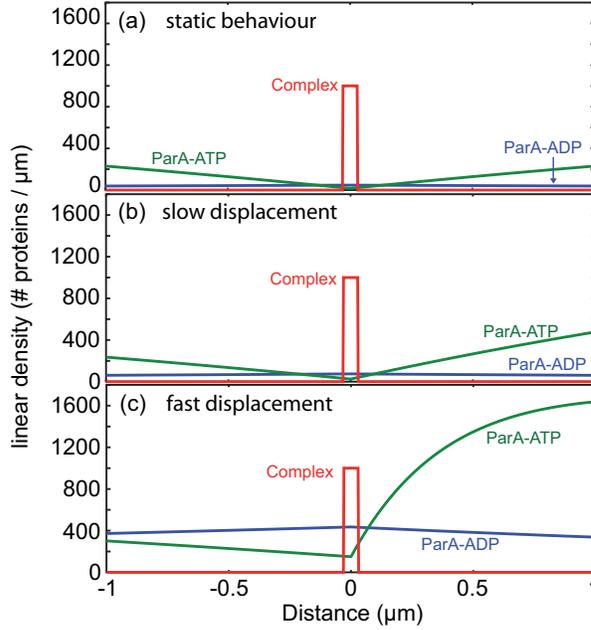


Figure 2.5: Density profile of ParA-slow  $v$  (green), ParA-fast  $u$  (blue) and ParBS (red). (a)  $\alpha < \alpha_c$ : positioning in the middle of the cell. (b) weak coupling  $\alpha_c \lesssim \alpha$ : ParBS moves as a TW and is "surfing" to the right on a protein wave. (c) strong coupling  $\alpha_c \ll \alpha$ : large asymmetry between the two sides of ParBS implying fast "surfing".

(ii) Plasmid diffusion in the relatively dense intracellular viscoelastic medium composed of the cytoplasmic fluid, DNA, and proteins, etc. is a complex phenomenon and we therefore estimate the plasmid diffusion coefficient from the above estimate for the coupling constant  $\alpha \equiv \varepsilon/m\gamma D_2 = (\varepsilon/k_B T)(D_{pc}/D_2)$ , the estimated value of  $D_2$ , and the energy  $\varepsilon$ . This energy represents a typical energy of interaction between a ParA-slow (ATPase) and a ParB protein and its order of magnitude should be in the range  $5 - 10 k_B T$  per ParA-ParB interaction. We therefore estimate the diffusion constant of the partition complex  $D_{pc}$  to be in the range  $[10^{-5} - 10^{-4}] \mu m^2 \cdot s^{-1}$ , in agreement with [63, 85]. This suggests that pure Brownian diffusion is very inefficient for partition complex segregation, and therefore justifies disregarding its effects in most of the results presented here.

(iii) The following useful relation sets the order of magnitude of the rates  $k_1$  and  $k_2$  needed in order to be in a non-trivial regime where two ParA species (slow or fast) coexist,

$$\frac{k_1 L}{k_2 \sigma} \sim \mathcal{O}(1), \quad (2.30)$$

which requires that the fluxes of ParA-fast and ParA-slow be of the same order of magnitude. It appears that this relation is fulfilled by the biological parameters of

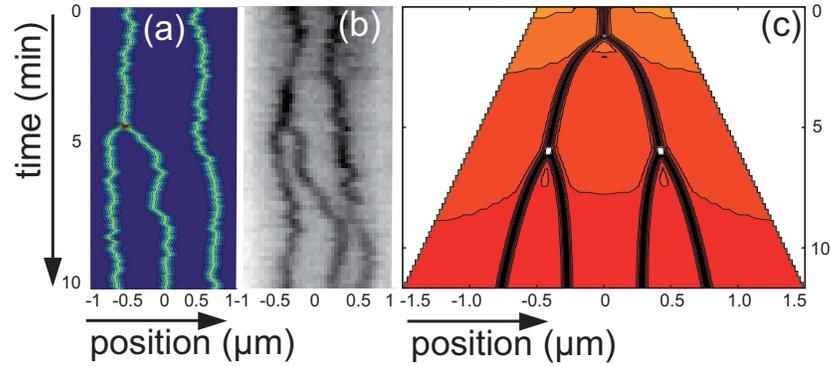


Figure 2.6: (a) Kymograph obtained from the model using an additional brownian force acting on ParBS: the model describes ParBS equilibrium, segregation and positioning. (b) Example of an experimental kymograph, obtained from 10 min. timelapse microscopy (frame every 10 sec.) of F-plasmids in *E.coli*, displaying a segregation event from two to three ParBS over the length of the nucleoid. (c) Theoretical kymograph obtained with growing cell (with an average over the stochastic noise). Trajectories are similar to experimental ones [81].

Table 2.1.

(iv) The equilibrium position restoring force for a partition complex vanishes for both zero and infinite  $k_1$  (see Fig.2.3). In the limit where  $k_1$  tends to infinity, rebinding is very fast and the ParA gradients extend only locally ( the screening length goes to zero) preventing the complex from sensing the boundaries. When  $k_1$  tends to zero, although the screening length diverges, the ParA gradients become weaker and weaker and the restoring force again tends to zero. In between the equilibrium position restoring force reaches a maximum at an optimal value of  $k_1$ .

*ParA in necessary drive ParB condensates out-of-equilibrium.* At equilibrium, a passive system undergoing phase separation displays a single condensed phase: if initially multiple condensed phases are found, the lowest energy state comprises a single liquid-separated phase that can be reached either by fusion or through Ostwald ripening [151]. Partition systems have specifically evolved to ensure faithful DNA segregation to daughter cells, which would be impeded by fusion into single, stable partition complexes. At longer time-scales (tens of minutes), however, ParB condensates are not only kept apart from each other but are also actively segregated concomitantly with the growth of the host cell (Fig.2.7A). Previous reports have shown that the ParA motor protein is necessary to ensure faithful positioning and segregation of partition complexes [81]. This function requires both ParA's ATP hydrolysis activity and its stimulation by ParB [3, 81].

Thus, we reasoned that if ParA is the factor preventing fusion of ParB condensates, then its slow removal should lead to the fusion of all ParB condensates in a cell. To test this hypothesis, we generated a strain where ParA can be selectively degraded using

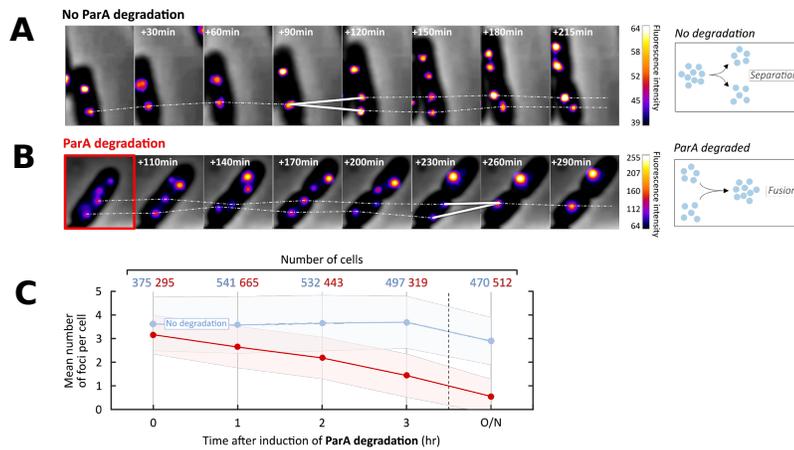


Figure 2.7: *Degradation of ParA leads to the fusion of ParBS complexes.* (A) Representative time lapse experiment of a ParA strain without induction of ParA degradation. Trajectories of ParB condensates (white dashed line) and a splitting event (white fork) are shown. Scalebar represents fluorescence intensity. (B) Representative time lapse experiment of a ParA strain upon induction of ParA degradation. Dashed lines represent the trajectories of two ParB condensates. A fork displays a fusion event between two ParB condensates. (C) Quantification of the average number of ParB condensates per cell in a ParA-ssra strain with (red circles) or without (blue circles) ParA degradation. The induction of ParA degradation was initiated at time 0. Note that the slight decrease in foci number ( $\sim 20\%$ ) without ParA degradation after overnight (O/N) culture corresponds to the decrease in cell length ( $\sim 20\%$ ) [54].

a degron system [89, 54]. In this strain, we observed the almost complete disappearance of ParA after 30 min of degron induction [54]. In the absence of ParA degradation, this strain displayed typical ParB condensate separation dynamics (Fig.2.7A, white traces). Strikingly, we observe the fusion of ParB condensates upon degradation of ParA (Fig.2.7B, white traces). Over time, the number of ParB condensates per cell collapses to  $0.6 \pm 0.7$  ParB condensates/cell when ParA is degraded, but remains at  $2.9 \pm 1.0$  ParB condensates/cell in the presence of ParA (Fig.2.7C). These data indicate that the action of the ParA motor is required to keep ParB condensates out of equilibrium by preventing their fusion.

## 2.4 Discussion

The existence of travelling waves with nonzero velocity is concomitant with the loss of stability of the equilibrium positions of the partition complexes discussed above. Thus,

we distinguish two dynamical regimes: (1) A *stable* regime without TWs ( $c_{TW} = 0$ ), with stable (equidistant, if more complexes are present) equilibrium complex positions independent of the initial conditions if the screening length  $\eta$  is large with respect to the cell size, see Fig.2.5(a). This implies a transient translocation when the initial conditions do not correspond to stable positions. This regime occurs for small values both of the coupling constant  $\alpha$  (obtained, e.g., for large values of the limiting diffusion constant  $D_2$ ) and the ParA concentration,  $C_0$ . When the screening length  $\eta$  is small, then ParBS cargos remain at their initial positions, not necessarily equidistant and without interaction between complexes. (2) A *dynamical* regime ( $c_{TW} \neq 0$ ) with unstable equilibrium positions of the complexes and ParA density oscillations in the cell corresponding to TWs in an unbounded domain, see Fig.2.5(b,c). This occurs for large values of both  $\alpha$  and the initial ParA concentration  $C_0$ . Since  $\alpha$  is large for small values of the diffusion constant  $D_2$ , there results an apparently surprising phenomenon, namely that slower ParA-slow kinetics leads to faster complex dynamics. This regime occurs because the ParA-slow distribution variation in time is not rapid enough to follow the partition complex and trails behind it. Indeed, the stability threshold corresponding to the appearance of TWs at  $c_{TW} = 0^+$  can be written as  $V_S < V_v$ , where  $V_S$  is the escape velocity of the complex and  $V_v$  the speed of spatial rearrangement of the ParA-slow distribution. When  $V_v > V_S$  the ParA distribution rapidly reequilibrates its symmetric profile with respect to the complex position and the system tends to the stable stationary regime, while in the opposite case spontaneous symmetry breaking and TW behavior occur. Using the expressions for  $V_S$  and  $V_v$ , we obtain the stability condition in the form:  $\alpha < \alpha_c \approx 1/(\sigma C_0)$ . This reveals that large complex sizes, interaction energies  $\varepsilon$ , and ParA densities, as well as low ParA-slow diffusion coefficients lead to the instability of the partition complex positioning. Importantly, a biologically reasonable choice of model parameters shows that the system is not far below the instability threshold, leading to a not only robust but also relatively fast segregation process, in agreement with experiment.

Our continuum reaction-diffusion approach significantly extends previous work [123, 63, 65, 124]. Some of these [65, 124] failed to observe a stable equi-positioning regime because ParA-slow was not allowed to diffuse ( $D_2 = 0$ ): thus  $\alpha$  diverges, setting the system in the unstable regime. In [124], relative positioning occurs only with multiple cargos as a crowding effect, whereas it is known that positioning can occur even with a single plasmid [3], as predicted by certain modeling studies [123, 63]. In line with the most recent experimental findings [81], we assume that partition complexes evolve within the nucleoid volume near the axis of the rod-shaped bacterial cells, in contrast with the translocation surface mechanism presented in [132, 133, 134, 65] performed on large surfaces coated by ParA, lacking the confinement necessary for equi-positioning. Our proposed mechanism integrates explicitly a volumetric interaction [112] with the partition complex (i.e. a length in 1D), placing the system close to the stability threshold for the biological range of parameters. In the case of a surface interaction, for which the volume is limited to the boundaries of the surface complex,  $\alpha_c$  would thus take much higher values. This argument can be easily generalized to higher dimensions  $D$ . Our approach also allows us to clarify analytically the physical mechanism at play, by going beyond the numerical simulations usually performed in a limited range of parameters, and to show explicitly that other effects like polymeriza-

tion [63] and DNA elasticity [85, 124] are not needed to account for segregation.

These elements make the active system considered in our work unprecedented, with genuine size and bulk-dependent effects, like the emergence of a critical coupling constant controlling the stability and the TW regimes. Moreover, when multiple complexes are present, they generate indirect inter-complex interactions mediated solely by the “perturbed” medium. This leads naturally to proper equilibrium partition complex positioning, as well as to spontaneous (left/right in 1D) symmetry breaking in the travelling wave regime. To our knowledge this is the first model, in the context of active bacterial segregation via ParABS systems, possessing very good qualitative and semi-quantitative agreement with all experimental observations, including segregation and position control of single and multiple partition complexes. The model robustness also suggests its application to other biological processes, like macromolecule and organelle positioning in intracellular dynamics.

Beyond its biological inspiration, this model is a novel one for active particle dynamics (accounting for “proteophoresis”) and nonlinear physics with a very rich phenomenology. Indeed, our model falls in the class of active particles (partition complexes in the present case) which locally “perturb” a medium (composed here of ParA proteins) that acts back on their dynamics and thus gives rise to particle self-propulsion. Such a behavior also provides similarities with classical polaron systems [10]. In contrast with previous works [123, 63, 94] on the subject, as well as on the self-propulsion of catalytic particles in active environments under chemical gradients [150], we do not invoke specifically the well-known mechanism of diffusiophoresis (or chemiphoresis) [36, 6, 7] or autochemotaxis, which involve surface interactions and (possibly asymmetric) catalytic surface reactions [123] coupled to surrounding hydrodynamic fluid flow relative to the particle surface (see [94, 150]). Future perspectives will include more refined comparisons with experimental observations and biological parameters and a generalization to higher dimensions.

to the community worldwide (>17,000 unique users/year).



# Conclusion

Cellular processes offer a fascinating playground for Physicists: new experimental techniques are now able to probe biological processes at the molecular scale, offering new classes of systems. This ranges from equilibrium system of binding and bridging proteins onto DNA, as well as active, out of equilibrium processes needed for the cell to create high concentration at a specific location of a given component, or simple to move efficiently objects into the cell and overcome the slowness of diffusion.

The model presented above pave the way for broader extension: we have shown application of the Spreading & Bridging in the case of ParABS, but it could in principle be used for any system of DNA interacting with proteins, like the chromatin fiber into eukaryotic cells.

We provide evidence in support of a new class of droplet-like-forming system with unique properties that ensure the stable coexistence and regulated inheritance of separate liquid-condensates. Interestingly, the three components of the ParABS partition system and the exquisite regulation of their multivalent interactions are required to observe this complex behavior: (1) High-affinity interactions of ParB and the *parS* centromeric sequence are necessary for the nucleation of the ParB condensate; (2) Weak interactions of ParB with chromosomal DNA and with itself are necessary to produce phase separation; (3) Finally, the ParA ATPase is required to counter droplet fusion and generate segregation. In passive phase-separation systems, separate liquid droplets grow by taking up material from a supersaturated environment, by Ostwald ripening, or by fusion of droplets. Over time, these processes lead to a decrease in the number of droplets and an increase in their size. This reflects the behaviour we observed for ParB condensates upon depletion of the ParA ATPase. Excitingly, in the presence of ParA, ParB condensates remain well separated. Extensive evidence suggests that this non-equilibrium state is maintained by the ability of ParA to hydrolyze ATP [81]. Recently, theoretical models have predicted mechanisms to suppress Ostwald ripening to enable the stable coexistence of multiple liquid phases [152]. These mechanisms require constituents of the phase separating liquids to be converted into each other by nonequilibrium chemical reactions [153]. Notably, the ParABS system constitutes an alternative and novel mechanism to actively control the number and size of droplets within cells, as well as their fission, segregation and stable anchoring to specific sub-cellular localizations. Interestingly, similar mechanisms – yet to be discovered – could control the number and size of other droplet-forming systems such as P-granules [22], stress granules [64], or heterochromatin domains [122].



# Bibliography

- [1] Adachi, S., Hori, K., and Hiraga, S. (2006). Subcellular positioning of F plasmid mediated by dynamic localization of SopA and SopB. *J. Mol. Biol.* 356 , 850-863.
- [2] Ah-Seng, Y Lopez, F., Pasta, F., Lane, D., and Bouet, J.Y (2009). Dual role of DNA in regulating ATP hydrolysis by the SopA partition protein. *J Biol Chem* 284, 30067-30075.
- [3] Ah-Seng, Y., Rech, J., Lane, D., and Bouet, J.-Y. (2013). Defining the role of ATP hydrolysis in mitotic segregation of bacterial plasmids. *PLoS Genet*, 9(12), e1003956.
- [4] Al-Husini, N., Tomares, D.T., Bitar, O., Childers, W.S., and Schrader, J.M. (2018).  $\alpha$ -Proteobacterial RNA Degradosomes Assemble Liquid-Liquid Phase-Separated RNP Bodies. *Mol. Cell* 71 , 1027-1039.e14.
- [5] P. W. Anderson and G. Yuval, Some numerical results on the Kondo problem and the inverse square one-dimensional Ising model, *J. Phys. C* 4, 607 (1971).
- [6] J. L. Anderson, Transport mechanisms of biological colloids., *Ann N Y Acad Sci.* 469, 166 (1986).
- [7] J. L. Anderson, Colloid transport by interfacial forces, *Annu. Rev. Fluid Mech.* 21, 61 (1989).
- [8] Banani, S.F., Lee, H.O., Hyman, A.A., and Rosen, M.K. (2017). Biomolecular condensates: organizers of cellular biochemistry. *Nat. Rev. Mol. Cell Biol.* 18 , 285-298.
- [9] Banigan, E. J., Gelbart, M. A., Gitai, Z., Wingreen, N. S., & Liu, A. J. (2011). Filament depolymerization can explain chromosome pulling during bacterial mitosis. *PLoS Comput Biol*, 7(9), e1002145.
- [10] L. Bányai, Motion of a classical polaron in a dc electric field, *Phys. Rev. Lett.* 70, 1674 (1993).
- [11] M. Barbieri, M. Chotalia, J. Fraser, L.-M. Lavitas, J. Dostie, A. Pombo, M. Nicodemi, Complexity of chromatin folding is captured by the strings and binders switch model, *Proc. Nat. Acad. Sci. U.S.A.* 109, 16173-16178 (2012).

- [12] C. Barde, N. Destainville, and M. Manghi, *Physical Review E* 97, 032412 (2018).
- [13] Baxter, J.C., and Funnell, B.E. (2014). Plasmid Partition Mechanisms. *Microbiol Spectr* 2.
- [14] J. B. Bliska and N. R. Cozzarelli, *Journal of molecular biology* 194, 205 (1987).
- [15] B. M. Booker, S. Deng, and N. P. Higgins, *Molecular Microbiology* 78, 1348 (2010).
- [16] J.-Y. Bouet, J. Rech, S. Egloff, D. P. Biek, D Lane, Probing plasmid partition centromere-based incompatibility, *Mol. Microbiol.* Vol.55, 511-525 (2005).
- [17] J.-Y. Bouet, M. Bouvier, and D. Lane, *Molecular Microbiology* 62, 1447 (2006).
- [18] Bouet J.-Y., Ah-Seng Y., Benmeradi N. and Lane D., Polymerization of SopA partition ATPase: regulation by DNA binding and SopB. *Molecular microbiology*, 63(2), pp.468-481 (2007).
- [19] Bouet, J. Y., & Lane, D. (2009). Molecular basis of the supercoil deficit induced by the mini-F plasmid partition complex. *Journal of Biological Chemistry*, 284(1), 165-173.
- [20] Bouet, J.-Y., Stouf, M., Lebailly, E., and Cornet, F. (2014). Mechanisms for chromosome segregation. *Curr. Opin. Microbiol.* 22, 60-65.
- [21] Brangwynne C. P., Eckmann C. R., Courson D. S., Rybarska A., Hoegge C., Gharakhani J., ... & Hyman, A. A. (2009). Germline P granules are liquid droplets that localize by controlled dissolution/condensation. *Science*, 324(5935), 1729-1732.
- [22] Brangwynne, C.P. (2011). Soft active aggregates: mechanics, dynamics and self-assembly of liquid-like intracellular protein bodies. *Soft Matter* 7, 3052-3059.
- [23] C. P. Brangwynne, Phase transitions and size scaling of membrane-less organelles, *J. Cell Biol.* 203, 875-881 (2013).
- [24] C. P. Brangwynne, P. Tompa, R. V. Pappu, Polymer physics of intracellular phase transitions, *Nat. Phys.* 11, 899-904 (2015).
- [25] Breier, A. M., & 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(3), 703-718.
- [26] C. P. Broedersz, X. Wang, Y. Meir, J. J. Loparo, D. Z. Rudner, N. S. Wingreen, Condensation and localization of the partitioning protein ParB on the bacterial chromosome, *Proc. Nat. Acad. Sci. U.S.A.* 111, 8809-8814 (2014).
- [27] Ciandrini L., Neri I., Walter J.-C., Dauloudet O. & Parmeggiani A. (2014) *Motor proteins traffic regulation by supply-demand balance*, *Phys. Biol.* 11, 056006.

- [28] Dame, R. T. (2005). The role of nucleoid-associated proteins in the organization and compaction of bacterial chromatin. *Molecular microbiology*, 56(4), 858-870.
- [29] Dame, R. T., Tark-Dame, M., & Schiessel, H. (2011). A physical approach to segregation and folding of the *Caulobacter crescentus* genome. *Molecular microbiology*, 82(6), 1311-1315.
- [30] David G., Walter J.-C., Brøedersz C., Dorignac J., Geniet F., Parmeggiani A., Walliser N.-O. & Palmeri J., Effective long range interactions generated by polymer fluctuations induce bound particle phase separation, submitted to *Phys. Rev. Lett.* [arXiv/1811.09234].
- [31] See Supplemental Material of [30] for details of calculations.
- [32] Debaugny R. E., Sanchez A., Rech J., Labourdette D., Dorignac J., Geniet F., Palmeri J., Parmeggiani A., Boudsocq F., Leberre V., Walter J.-C. & Bouet J.-Y., A conserved mechanism drives partition complex assembly on bacterial chromosomes and plasmids, *Molecular System Biology* **14**, e8516.
- [33] P. De Gennes, Collapse of a polymer chain in poor solvents, *J. Phys. Lett.* 36, 55 (1975).
- [34] P.-G. de Gennes, *Scaling Concepts in Polymer Physics*, Cornell University Press, Ithaca (1979).
- [35] Dekker, J., Marti-Renom, M. A., & Mirny, L. A. (2013). Exploring the three-dimensional organization of genomes: interpreting chromatin interaction data. *Nature Reviews Genetics*, 14(6), 390-403.
- [36] S. Duhkin, B.V. Derjaguin, *Electrokinetic Phenomena*, in: E. Matijevic (Ed.), *Surface and Colloid Science*, vol. 7, p. 322, Wiley-Interscience, New York, 1974.
- [37] Dillon, S. C., & Dorman, C. J. (2010). Bacterial nucleoid-associated proteins, nucleoid structure and gene expression. *Nature Reviews Microbiology*, 8(3), 185-195.
- [38] C. J. Dorman and M. J. Dorman, *Biophysical reviews* 8, 89 (2016).
- [39] F. J. Dyson, Existence of a Phase-Transition in a One-Dimensional Ising Ferromagnet, *Commun. Math. Phys.* 12, 91-107 (1969).
- [40] J. Elf, G. Li, and X. S. Xie, *Science* 316, 1191 (2007).
- [41] Elowitz, M.B., Surette, M.G., Wolf, P.E., Stock, J.B., and Leibler, S. (1999). Protein mobility in the cytoplasm of *Escherichia coli*. *J. Bacteriol.* 181, 197-203.
- [42] Emanuel, M., Radja, N. H., Henriksson, A., & Schiessel, H. (2009). The physics behind the larger scale organization of DNA in eukaryotes. *Physical biology*, 6(2), 025008.

- [43] Erdmann, N., Petroff, T., and Funnell, B.E. (1999). Intracellular localization of P1 ParB protein depends on ParA and *parS*. *Proc. Natl. Acad. Sci. U. S. A.* 96, 14905-14910.
- [44] Everaers, R., Kumar, S., & Simm, C. (2007). Unified description of poly- and oligonucleotide DNA melting: Nearest-neighbor, Poland-Sheraga, and lattice models. *Physical Review E*, 75(4), 041918.
- [45] Faure, L. M., Fiche, J. B., Espinosa, L., Ducret, A., Anantharaman, V., Luciano, J., ... & Kuru, E. (2016). The mechanism of force transmission at bacterial focal adhesion complexes. *Nature*, 539(7630), 530.
- [46] R. P. Feynman, *Statistical mechanics: A set of lectures*. Reading, Mass: W.A. Benjamin (1972).
- [47] Fiche, J.-B., Cattoni, D.I., Diekmann, N., Langerak, J.M., Clerte, C., Royer, C.A., Margeat, E., Doan, T., and Nollmann, M. (2013). Recruitment, assembly, and molecular architecture of the SpoIIIE DNA pump revealed by superresolution microscopy. *PLoS Biol.* 11, e1001557.
- [48] Fisher, G. L., Pastrana, C. L., Higman, V. A., Koh, A., Taylor, J. A., Butterer, A., & Moreno-Herrero, F. (2017). The C-Terminal Domain Of ParB Is Critical For Dynamic DNA Binding And Bridging Interactions Which Condense The Bacterial Centromere. *bioRxiv*, 122986.
- [49] P. Flory, *Principles of Polymer Chemistry*, Cornell University Press, Ithaca, NY (1971).
- [50] K. Gerdes, M. Howard and F. Szardenings. Pushing and pulling in prokaryotic DNA segregation. *Cell* 141, 927-942 (2010).
- [51] Glaser P, Sharpe ME, Raether B, Perego M, Ohlsen K, Errington J. Dynamic, mitotic-like behavior of a bacterial protein required for accurate chromosome partitioning. *Genes & Development*. 1997 May 1;11(9):1160-8.
- [52] Gordon, S., Rech, J., Lane, D. and Wright, A., 2004. Kinetics of plasmid segregation in *Escherichia coli*. *Molecular microbiology*, 51(2), pp.461-469.
- [53] T. G. W. Graham, X. Wang, D. Song, C. M. Eton, A. M. van Oijen, D. Z. Rudner, J. J. Loparo, ParB spreading requires DNA bridging, *Genes Dev.* 28, 1228-1238 (2014).
- [54] B. Guilhas, J.-C. Walter, J. Rech, G. David, N.-O. Walliser, J. Palmeri, C., Mathieu-Demaziere, A. Parmeggiani, J.Y. Bouet, A. Le Gall & M. Nollmann, ATP-driven separation of liquid phase condensates in bacteria, submitted.
- [55] M. Feric, N. Vaidya, T. S. Harmon, D. M. Mitrea, L. Zhu, T. M. Richardson, R. W. Kriwacki, R. V. Pappu, C. P. Brangwynne, Coexisting Liquid Phases Underlie Nucleolar Subcompartments, *Cell* 165, 1686-1697 (2016).

- [56] N. Haddad, D. Jost, C. Vaillant, Perspectives: using polymer modeling to understand the formation and function of nuclear compartments, *Chromosome Res.* 25, 35-50 (2017).
- [57] Hadiwikarta W. W., Walter J.-C., Hooyberghs J. & Carlon E., *Probing Hybridization parameters from microarray experiments: nearest neighbor model and beyond* (2012) *Nucl. Acids Res.* 40, e138.
- [58] Halatek, Jacob, Fridtjof Brauns, and Erwin Frey. "Self-organization principles of intracellular pattern formation." *Phil. Trans. R. Soc. B* 373.1747 (2018): 20170107.
- [59] Hanke, A., & Metzler, R. (2003). Entropy loss in long-distance DNA looping. *Biophysical journal*, 85(1), 167-173.
- [60] R. Horwitz, Cellular Biophysics, *Bio. J.* 110, 993-996 (2016).
- [61] K. Huang, Statistical Physics, Second Edition, John Wiley, New York (1987).
- [62] A. A. Hyman, C. A. Weber, F. Jülicher, Liquid-Liquid Phase Separation in Biology, *Annu. Rev. Cell. Dev. Biol.* 30, 39-58 (2014).
- [63] R. Ietswaart, F. Szardenings, K. Gerdes, M. Howard, Competing ParA Structures Space Bacterial Plasmids Equally over the Nucleoid, *PLOS Computational Biology*, 12, Vol. 10 (2014).
- [64] Jain, S., Wheeler, J.R., Walters, R.W., Agrawal, A., Barsic, A., and Parker, R. (2016). ATPase-Modulated Stress Granules Contain a Diverse Proteome and Substructure. *Cell* 164, 487-498.
- [65] Jindal L. and E. Emberly. Operational principles for the dynamics of the in vitro ParA-ParB system. *PLOS Comput. Biol.* 11.12 (2015): e1004651.
- [66] J. Johnson, C. A. Brackley, P. R. Cook, D Marenduzzo, A simple model for DNA bridging proteins and bacterial or human genomes: bridging-induced attraction and genome compaction, *J. Phys.: Condens. Matter* 27, 064119 (2015).
- [67] D. Jost, C. Vaillant, P. Meister, Coupling 1D modifications and 3D nuclear organization: data, models and function, *Cur. Op. Cell Biol.* 44, 20-27 (2017).
- [68] Jun, S., & Mulder, B. (2006). Entropy-driven spatial organization of highly confined polymers: lessons for the bacterial chromosome. *Proceedings of the National Academy of Sciences*, 103(33), 12388-12393.
- [69] I. Junier, O. Martin and F. Képès, Spatial and Topological Organization of DNA Chains Induced by Gene Co-localization, *PLoS Comput. Biol.* 6, e1000678 (2010).
- [70] L. J. Kijewski and M. P. Kawatha, One-Dimensional Ising Model with Long-Range Interaction, *Physics Letters* 31A, 479 (1970).
- [71] J. M. Kosterlitz, D. J. Thouless, Ordering, metastability and phase transitions in two-dimensional systems, *J. Phys. C: Solid State Phys.* 6, 1181-1203 (1973).

- [72] B. A. Krajina and A. J. Spakowitz, *Biophysical Journal* 111, 1339 (2016).
- [73] M. Kumar, M.S. Mommer and V. Sourjik Mobility of Cytoplasmic, Membrane, and DNA-Binding Proteins in *Escherichia coli* *Biophysical Journal*, Volume 98, February 2010, 552-559.
- [74] A. Lal, A. Dhar, A. Trostel, F. Kouzine, A. S. N. Seshasayee, and S. Adhya, *Nature communications* 7, 11055 (2016).
- [75] Laleman M., Baiesi M., Sakaue T., Belotserkovskii B., Walter J-C. & Carlon E., *Torque-induced rotational dynamics in polymers: Torsional blobs and thinning* (2016) *Macromol.* **49**, 405-414.
- [76] L.D. Landau and E.M. Lifshitz, *Statistical Physics, Third Edition, Part 1: Volume 5 (Course of Theoretical Physics, Volume 5)*, Pergamon Press, Oxford (1980).
- [77] A. G. Larson, D. Elnatan, M. M. Keenen, M. J. Trnka, J. B. Johnston, A. L. Burlingame, D. A. Agard, S. Redding, G. J. Narlikar, Liquid droplet formation by HP1 $\alpha$  suggests a role for phase separation in heterochromatin, *Nature* 547, 236-240 (2017).
- [78] Lagomarsino, M. C., Espéli, O., & Junier, I. (2015). From structure to function of bacterial chromosomes: Evolutionary perspectives and ideas for new experiments. *FEBS letters*, 589(20PartA), 2996-3004.
- [79] Le, T. B., Imakaev, M. V., Mirny, L. A., & Laub, M. T. (2013). High-resolution mapping of the spatial organization of a bacterial chromosome. *Science*, 342(6159), 731-734.
- [80] C. F. Lee, C. P. Brangwynne, J. Gharakhani, A. A. Hyman, F. Jülicher, Spatial Organization of the Cell Cytoplasm by Position-Dependent Phase Separation, *Phys. Rev. Lett.* 111, 088101 (2013).
- [81] A. Le Gall, D.I. Cattoni, B. Guilhas, C. Mathieu-Demazière, L. Oudjedi, J.B. Fiche, J. Rech, S. Abrahamsson, H. Murray, J.Y. Bouet, N. Nollmann, Bacterial partition complexes segregate within the volume of the nucleoid, *Nature Communications* 7, 12107 (2016).
- [82] Leonard, T.A., Butler, P.J. and Löwe, J., Bacterial chromosome segregation: structure and DNA binding of the Soj dimer—a conserved biological switch. *The EMBO journal*, 24(2), pp.270-282 (2005).
- [83] T. Lepage and I. Junier, *Methods in molecular biology (Clifton, N.J.)* 1624, 323 (2017).
- [84] T. Lepage, F. Képès, and I. Junier, *Biophysical Journal* 109, 135 (2015).
- [85] H.C. Lim, I.V. Surovtsev, B.G. Beltran, F. Huang, J. Bewersdorf, C. Jacobs-Wagner., Evidence for a DNA-relay mechanism in ParABS-mediated chromosome segregation. *Elife*. 2014 May 23;3:e02758.

- [86] L. F. Liu and J. C. Wang, PNAS 84, 7024 (1987).
- [87] E. Luijten, H. W. J. Blöte, Classical critical behavior of spin models with long-range interactions, Phys. Rev. B 56, 8945 (1997).
- [88] B. McCoy and T. T. Wu, The Two-Dimensional Ising Model, Cambridge, Massachusetts, Harvard UP (1973).
- [89] McGinness, K.E., Baker, T.A., and Sauer, R.T. (2006). Engineering controllable protein degradation. Mol. Cell 22, 701-707.
- [90] Marbouty, M., Le Gall, A., Cattoni, D.I., Cournac, A., Koh, A., Fiche, J.-B., Mozziconacci, J., Murray, H., Koszul, R., and Nollmann, M. (2015). Condensin- and Replication-Mediated Bacterial Chromosome Folding and Origin Condensation Revealed by Hi-C and Super-resolution Imaging. Mol. Cell 59, 588-602.
- [91] Marenduzzo, D., Micheletti, C., & Cook, P. R. (2006). Entropy-driven genome organization. Biophysical journal, 90(10), 3712-3721.
- [92] J. Marko and E. Siggia, Physical Review E 52, 2912 (1995).
- [93] John F. Marko, Biophysics of protein-DNA interactions and chromosome organization, Physica A 418, 126-153 (2015).
- [94] E.J. Banigan, J.F. Marko, Self-propulsion and interactions of catalytic particles in a chemically active medium, Physical Review E 93, 012611 (2016).
- [95] J. E. Mayer and M. G. Mayer, Statistical mechanics. 2nd ed. John Wiley & Sons, Inc., New York (1977).
- [96] Mirny, L. A. (2011). The fractal globule as a model of chromatin architecture in the cell. Chromosome research, 19(1), 37-51.
- [97] Moffitt, J. R., Pandey, S., Boettiger, A. N., Wang, S., & Zhuang, X. (2016). Spatial organization shapes the turnover of a bacterial transcriptome. Elife, 5, e13065.
- [98] Mohl, D. A., & Gober, J. W. (1997). Cell cycle-dependent polar localization of chromosome partitioning proteins in *Caulobacter crescentus*. Cell, 88(5), 675-684.
- [99] Monterroso, B., Zorrilla, S., Sobrinos-Sanguino, M., Robles-Ramos, M.A., López-Álvarez, M., Margolin, W., Keating, C.D., and Rivas, G. (2019). Bacterial FtsZ protein forms phase-separated condensates with its nucleoid-associated inhibitor SImA. EMBO Rep. 20.
- [100] D. Mukamel, Notes on the Statistical Mechanics of Systems with Long-Range Interactions, arXiv:0905.1457 (2009).
- [101] K. G. Petrosyan, C.-K. Hu, Protein-mediated loops and phase transition in non-thermal denaturation of DNA, J. Stat. Mech., P01005 (2011).
- [102] Poland, D., & Scheraga, H. A. (1970). Theory of helix-coil transitions in biopolymers.

- [103] Ptacin, J. L., Lee, S. F., Garner, E. C., Toro, E., Eckart, M., Comolli, L. R., & Shapiro, L. (2010). A spindle-like apparatus guides bacterial chromosome segregation. *Nature cell biology*, 12(8), 791-798.
- [104] Ramamurthi, K. S., & Losick, R. (2009). Negative membrane curvature as a cue for subcellular localization of a bacterial protein. *Proceedings of the National Academy of Sciences*, 106(32), 13541-13545.
- [105] Racki, L. R., Tocheva, E. I., Dieterle, M. G., Sullivan, M. C., Jensen, G. J., & Newman, D. K. (2017). Polyphosphate granule biogenesis is temporally and functionally tied to cell cycle exit during starvation in *Pseudomonas aeruginosa*. *Proceedings of the National Academy of Sciences*, 114(12), E2440-E2449.
- [106] J. Redolfi, Y. Zhan, C. Valdes, M. Kryzhanovska, I. M. Guerreiro, V. Iesmantavicius, G. Tiana, T. Pollex, J. Kind, S. Smallwood, W. d. Laats, and L. Giorgetti, bioRxiv.
- [107] O. Rodionov, M. Lobočka, M. Yarmolinsky, Silencing of Genes Flanking the P1 Plasmid Centromere, *Science* 283, 546-549 (1999).
- [108] N. S. Rovinskiy, A. A. Agbleke, O. N. Chesnokova, and N. P. Higgins, *Microorganisms* 7 (2019).
- [109] Sakaue T., Walter J.-C., Carlon E., & Vanderzande C., *Theory of the anomalous unzipping dynamics of polymers* (2017) *Soft Matter* 13, 3174-3181.
- [110] Salas, D., Le Gall, A., Fiche, J. B., Valeri, A., Ke, Y., Bron, P., ... & Nollmann, M. (2017). Angular reconstitution-based 3D reconstructions of nanomolecular structures from superresolution light-microscopy images. *Proceedings of the National Academy of Sciences*, 114(35), 9273-9278.
- [111] Sanamrad, A., Persson, F., Lundius, E. G., Fange, D., Gynna, A. H., & Elf, J. (2014). Single-particle tracking reveals that free ribosomal subunits are not excluded from the *Escherichia coli* nucleoid. *Proceedings of the National Academy of Sciences*, 111(31), 11413-11418.
- [112] A. Sanchez, D. I. Cattoni, J.-C. Walter, J. Rech, A. Parmeggiani, M. Nollman, J.-Y. Bouet, Stochastic Self-Assembly of ParB Proteins Builds the Bacterial DNA Segregation Apparatus, *Cell Syst.* 1, 163-173 (2015).
- [113] H. Schiessel, *Biophysics for Beginners: a journey through the cell nucleus*, Singapore, Pan Stanford Publishing (2014).
- [114] Schumacher, M.A., and Funnell, B.E. (2005). Structures of ParB bound to DNA reveal mechanism of partition complex formation. *Nature* 438, 516-519.
- [115] Schumacher D., Bergeler S., Harms A., Vonck J., Huneke-Vogt S., Frey E. and Sogaard-Andersen L. The PomXYZ Proteins Self-Organize on the Bacterial Nucleoid to Stimulate Cell Division. *Developmental Cell*, 41, 299-314 (2017).

- [116] V. F. Scolari, M. Cosentino Lagomarsino, Combined collapse by bridging and self-adhesion in a prototypical polymer model inspired by the bacterial nucleoid, *Soft Matter* 11, 1677 (2015).
- [117] Shapiro, L., McAdams, H. H., & Losick, R. (2009). Why and how bacteria localize proteins. *Science*, 326(5957), 1225-1228.
- [118] Shin, Y., & Brangwynne, C. P. (2017). Liquid phase condensation in cell physiology and disease. *Science*, 357(6357), eaaf4382.
- [119] R. R. Sinden, J. O. Carlson, and D. E. Pettijohn, *Cell* 21, 773 (1980). *Proc. Natl. Acad. Sci. U. S. A.* 111, E2636-E2645.
- [120] Stracy, M., Lesterlin, C., Garza de Leon, F., Uphoff, S., Zawadzki, P., and Kapanidis, A.N. (2015). Live-cell superresolution microscopy reveals the organization of RNA polymerase in the bacterial nucleoid. *Proc. Natl. Acad. Sci. U. S. A.* 112, E4390-E4399.
- [121] T. R. Strick, M.-N. Dessinges, G. Charvin, N. H. Dekker, J.-F. Allemand, D. Bensimon, and V. Croquette, *Reports on Progress in Physics* 66 (2003).
- [122] Strom, A.R., Emelyanov, A.V., Mir, M., Fyodorov, D.V., Darzacq, X., and Karpen, G.H. (2017). Phase separation drives heterochromatin domain formation. *Nature* 547, 241-245.
- [123] T. Sugawara, K. Kaneko, Chemophoresis as a driving force for intracellular organization: Theory and application to plasmid partitioning, *Biophysics*, Vol. 7, pp. 77-88 (2011).
- [124] Surovtsev IV, Campos M, Jacobs-Wagner C., DNA-relay mechanism is sufficient to explain ParA-dependent intracellular transport and patterning of single and multiple cargos. *Proceedings of the National Academy of Sciences*. 2016 Nov 15;113(46):E7268-76.
- [125] Surovtsev I. V., & Jacobs-Wagner C. (2018). Subcellular organization: a critical feature of bacterial cell replication. *Cell*, **172**(6), 1271-1293.
- [126] Taylor, J. A., Pastrana, C. L., Butterer, A., Pernstich, C., Gwynn, E. J., Sobott, F., ... & Dillingham, M. S. (2015). Specific and non-specific interactions of ParB with DNA: implications for chromosome segregation. *Nucleic acids research*, 43(2), 719-731.
- [127] Toro, E., & Shapiro, L. (2010). Bacterial chromosome organization and segregation. *Cold Spring Harbor perspectives in biology*, 2(2), a000349.
- [128] A. Travers and G. Muskhelishvili, *Current Opinion in Genetics & Development* 15, 507 (2005).
- [129] G. Le Treut, F. Képès, H. Orland, Phase Behavior of DNA in the Presence of DNA-Binding Proteins, *Bio. J.* 110, 51-62 (2016).

- [130] Umbarger, M. A., Toro, E., Wright, M. A., Porreca, G. J., Bau, D., Hong, S. H., & Shapiro, L. (2011). The three-dimensional architecture of a bacterial genome and its alteration by genetic perturbation. *Molecular cell*, 44(2), 252-264.
- [131] van Gijtenbeek, L. A., Robinson, A., van Oijen, A. M., Poolman, B., & Kok, J. (2016). On the spatial organization of mRNA, plasmids, and ribosomes in a bacterial host overexpressing membrane proteins. *PLoS genetics*, 12(12), e1006523.
- [132] A.G. Vecchiarelli, Y.-W. Han, X. Tan, M. Mizuuchi, R. Ghirlando, C. Biertümpfel, B.E. Funnell and K. Mizuuchi, ATP control of dynamics P1 ParA-DNA interactions: a key role for the nucleoid in plasmid partition, *Molecular Microbiology* 78(1), 78-91 (2010).
- [133] A.G. Vecchiarelli, L.C. Hwang, K. Mizuuchi, Cell-free study of F plasmid partition provides evidence for cargo transport by a diffusion-ratchet mechanism, *Proc. Natl. Acad. Sci. USA* 110, E1390-E1397 (2013).
- [134] A.G. Vecchiarelli, K.C. Neuman, K. Mizuuchi, A propagating ATPase gradient drives transport of surface-confined cellular cargo, *Proc. Natl. Acad. Sci. USA* 111, 4880-4885 (2014).
- [135] Viollier, P. H., Thanbichler, M., McGrath, P. T., West, L., Meewan, M., McAdams, H. H., & Shapiro, L. (2004). Rapid and sequential movement of individual chromosomal loci to specific subcellular locations during bacterial DNA replication. *Proceedings of the National Academy of Sciences of the United States of America*, 101(25), 9257-9262.
- [136] A. Vologodskii and N. Cozzarelli, *Annual Review of Biophysics and Biomolecular Structure* 23, 609 (1994).
- [137] A. V. Vologodskii, S. D. Levene, K. V. Klenin, M. Frank-Kamenetskii, and N. R. Cozzarelli, *Journal of Molecular Biology* 227, 1224 (1992).
- [138] Walter J.-C., Kroll K. M., Hooyberghs J. & Carlon E., *Nonequilibrium effects in DNA microarrays: a multiplatform study* (2011), *J. Phys. Chem. B* **115**, 6732.
- [139] Walter J.-C., Barkema G. T. & Carlon E., *The equilibrium winding angle of a polymer around a bar* (2011) *J. Stat. Mech.* P10020.
- [140] Walter J.-C., Ferrantini A., Carlon E. & Vanderzande C., *Fractional Brownian motion and the critical dynamics of zipping polymers* (2012) *Phys. Rev. E* **85**, 031120.
- [141] Walter J.-C., Baiesi M., Barkema G. T. & Carlon E., *Unwinding relaxation dynamics of polymers* (2013) *Phys. Rev. Lett.* **110**, 068301.
- [142] Walter J.-C., Laleman M., Baiesi M. & Carlon E., *Rotational dynamics of entangled polymers* (2014) *Eur. Phys. J. Special Topics* **223**, 3201–3213.
- [143] Walter J.-C., Baiesi M., Carlon E. & Schiessel H., *Unwinding dynamics of a helically wrapped polymer* (2014) *Macromolecules* **47**, 4840–4846.

- [144] J.-C. Walter, J. Dorignac, V. Lorman, J. Rech, J.-Y. Bouet, M. Nollman, J. Palmeri, A. Parmeggiani, F. Geniet, Surfing on Protein Waves: Proteophoresis as a Mechanism for Bacterial Genome Partitioning, *Phys. Rev. Lett.* **119**, 028101 (2017).
- [145] J.-C. Walter, N.-O. Walliser, G. David, J. Dorignac, F. Geniet, J. Palmeri, A. Parmeggiani, N. S. Wingreen, C. P. Broedersz, Looping and Clustering model for the organization of protein-DNA complexes on the bacterial genome, *New J. Phys.* **20**, 035002 (2018).
- [146] Walter J.-C., Lepage T., Dorignac J., Geniet F., Parmeggiani A., Palmeri J., Parmeggiani A., Bouet J.-Y. & Junier I., Modeling supercoiled DNA interacting with an anchored cluster of proteins: towards a quantitative estimation of chromosomal DNA supercoiling (2020) submitted, [arXiv:2002.00111].
- [147] Wang, X., Llopis, P. M., & Rudner, D. Z. (2013). Organization and segregation of bacterial chromosomes. *Nature Reviews Genetics*, 14(3), 191-203.
- [148] J. C. Wang, *Journal of molecular biology*87, 797 (1974).
- [149] J. H. White, *American Journal of Mathematics* 91, 693 (1969).
- [150] A. Zöttl and H. Stark, Emergent behavior in active colloids, *J. Phys.: Condens. Matter* 28, 253001 (2016).
- [151] Zwicker, D., Decker, M., Jaensch, S., Hyman, A.A., and Jülicher, F. (2014). Centrosomes are autocatalytic droplets of pericentriolar material organized by centrioles.
- [152] Zwicker, D., Hyman, A.A., and Jülicher, F. (2015). Suppression of Ostwald ripening in active emulsions. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* 92, 012317.
- [153] Zwicker, D., Seyboldt, R., Weber, C.A., Hyman, A.A., and Jülicher, F. (2016). Growth and division of active droplets provides a model for protocells. *Nat. Phys.* 13, 408.



# Chapter 3

## Curriculum Vitæ

Jean-Charles WALTER

Date of birth: January 28, 1984

Citizenship: French

Laboratoire Charles Coulomb, Université de Montpellier

Chargé de Recherche CNRS – Section 05

Address : Laboratoire Charles Coulomb, Place Bataillon, 34095, Montpellier.

Email: jean-charles.walter@umontpellier.fr

Office phone: +33467143146

### 3.1 Education & Positions

2017–present: Chargé de Recherche CNRS, Laboratoire Charles Coulomb, Université de Montpellier, team *Systèmes Complexes et Physique Non-linéaire*

2013–2017: Postdoc, Laboratoire Charles Coulomb, Université de Montpellier, team *Systèmes Complexes et Physique Non-linéaire*

2012: Postdoc, Instituut-Lorentz, University of Leiden, The Netherlands, (team of Helmut Schiessel).

(dec.)2009-2011: Postdoc, Institute for Theoretical Physics, KU Leuven, Belgium (team of Enrico Carlon).

2006-2009: PhD, University Henri Poincaré, Nancy, *Numerical study of corrections to the dominant scaling behaviour at and out of equilibrium in spin systems*, Supervisors: Christophe Chatelain & Bertrand Berche.

### 3.2 Scientific activity

#### 3.2.1 Research interests in brief

- Theoretical biophysics,

- Organisation of the genome *in vivo*,
- Physical properties of polymers and biopolymers,
- Transport phenomena along a filament (translation, transport of cargo along the cytoskeleton),
- Statistical physics,
- Numerical simulations (Monte Carlo, Molecular Dynamics).

### 3.2.2 Research grants

- **Principal Investigator** of the funding MITI-CNRS “*Défi pour le vivant*”: COILCHROM Sensing DNA supercoiling *in vivo* by combining polymer physics and molecular biology (2019-2020, 26k€).
- **Partner** of the funding 80 Prime: NUMACOILED, Nucleation mechanism for the auto-assembly of a large nucleoprotein complex on supercoiled DNA (2019-2020, 17k€ + 1 experimental PhD grant).
- **Partner** of the ANR *FlagMot* (2018-2021, 550k€).
- **Partner** of the I-site MUSE project *NMJ-on-chip* (2018-2021, 200k€).
- **Partner** of the ANR *Imaging Bacterial Mitosis* (2015–2018, 450k€).
- **Partner** of a flagship project *Gene Expression Modeling* funded by the LabEx NUMEV (2017–2020, 250k€).
- **Partner** of a CNRS Défi Inphyniti grant: (*Projet Structurant*, 2015–2016, 20k€).

#### Individual Grants:

- 2018-2023: *Prime d’Encadrement Doctoral et de Recherche* (PEDR) attributed by Section 05, CNRS.

### 3.2.3 Main collaborations

- Marco Baiesi  
Galileo Galilei Institute, Padova University.
- Gerard Barkema  
Institute for Theoretical Physics, Utrecht University.
- Jean-Yves Bouet  
LMGM, CNRS & Université Paul Sabatier, Toulouse, France.
- Chase Brøedersz  
Ludwig-Maximilian Universität, Munich, Germany.

- Enrico Carlon  
Institute for Theoretical Physics, KU Leuven, Belgium.
- Ivan Junier  
TIMC-IMAG, CNRS & Grenoble University, France.
- Izaak Neri  
King's College, London, UK.
- Marcelo Nollmann  
CBS, CNRS & Montpellier University.
- Helmut Schiessel  
Instituut-Lorentz for Theoretical Physics, Leiden University.
- Ned S. Wingreen  
Lewis-Sigler Institute of Integrative Genomics, Princeton University, USA.

### 3.3 Scientific communications

#### 3.3.1 Publication list

\*: First Author

†: Corresponding Author

2020

[26] Dauloudet\* O., Neri\*† I., **Walter\*† J.-C.**, Dorignac J., Geniet F. & Parmeggiani† A. *Modelling the effect of ribosome mobility on the rate of protein synthesis*, submitted, [arXiv:2009.14533].

[25] **Walter\*† J.-C.**, Lepage T., Dorignac J., Geniet F., Parmeggiani A., Palmeri J., Parmeggiani A., Bouet J.-Y. & Junier† I. *Modeling supercoiled DNA interacting with an anchored cluster of proteins: towards a quantitative estimation of chromosomal DNA supercoiling*, submitted, [arXiv:2002.00111].

[24] **Walter\*† J.-C.**, Rech J., Walliser N.-O., Dorignac J., Geniet F., Palmeri J., Parmeggiani A. & Bouet† J.-Y. *Physical modeling of a sliding clamp mechanism for the spreading of ParB at short genomic distance from bacterial centromere sites*, submitted [bioRxiv/2020/213413].

[23] David G., **Walter J.-C.**, Brødersz C., Dorignac J., Geniet F., Parmeggiani A., Walliser N.-O. & Palmeri J., *Effective long range interactions generated by polymer fluctuations induce bound particle phase separation*, *Physical Review Research* **2**, 033377.

[22] Guilhas B., **Walter J.-C.**, Rech J., Mathieu-Demazière C., Parmeggiani A., Bouet J.-Y., Le Gall A. & Nollmann M., *ATP-driven separation of liquid phase condensates in*

bacteria, *Molecular Cell*, **79**, 293-303.

## 2018

[21] Debaugny R. E., Sanchez A., Rech J., Labourdette D., Dorignac J., Geniet F., Palmeri J., Parmeggiani A., Boudsocq F., Leberre V., **Walter<sup>T</sup> J.-C.** & Bouet<sup>T</sup> J.-Y., *A conserved mechanism drives partition complex assembly on bacterial chromosomes and plasmids*, *Molecular System Biology* **14**, e8516.

[20] **Walter\* J.-C.**, Walliser N.-O., David G., Dorignac J., Geniet F., Palmeri J., Parmeggiani A., Wingreen N. S. & Brødersz C., *Looping and clustering model for the organization of protein-DNA complexes on the bacterial genome*, *New Journal of Physics* **20**, 035002.

## 2017

[19] **Walter\* J.-C.**, Dorignac J., Lorman V., Rech J., Bouet J.-Y., Nollmann M., Palmeri J., Parmeggiani A. & Geniet F., *Surfing on protein waves: proteophoresis as a mechanism for bacterial genome partitioning*, *Phys. Rev. Lett.* **119**, 028101.

[18] Sakaue T., **Walter J.-C.**, Carlon E., & Vanderzande C., *Theory of the anomalous unzipping dynamics of polymers*, *Soft Matter* **13**, 3174-3181.

## 2016

[17] Laleman M., Baiesi M., Sakaue T., Belotserkovskii B., **Walter J.-C.** & Carlon E., *Torque-induced rotational dynamics in polymers: Torsional blobs and thinning*, *Macromol.* **49**, 405-414.

## 2015

[16] Sanchez A., Cattoni D. I., **Walter J.-C.**, Rech J., Parmeggiani A., Nollmann M. & Bouet J.-Y., *Stochastic self-assembly of ParB proteins nucleated from the centromere builds bacterial DNA segregation apparatus*, *Cell Systems* **1**, 163–173.

[15] **Walter\* J.-C.** & Barkema G. T., *An introduction to Monte Carlo methods*, *Physica A* **418**, 78–87, Lecture Notes of the 13th International Summer School: Fundamental Problems in Statistical Physics (Leuven, Belgium, 16–29 June 2013).

## 2014

[14] **Walter\* J.-C.**, Laleman M., Baiesi M. & Carlon E., *Rotational dynamics of entangled polymers*, *Eur. Phys. J. Special Topics* **223**, 3201–3213.

[13] Ciandrini L.\*, Neri I.\*, **Walter\* J.-C.**, Dauloudet O. & Parmeggiani A., *Motor proteins traffic regulation by supply-demand balance*, *Phys. Biol.* **11**, 056006 [Fea-

tured].

[12] **Walter**<sup>\*†</sup> **J.-C.**, Baiesi M., Carlon E. & Schiessel H., *Unwinding dynamics of a helically wrapped polymer*, *Macromolecules* **47**, 4840–4846.

### 2013

[11] **Walter**<sup>\*</sup> **J.-C.**, Baiesi M., Barkema G. T. & Carlon E., *Unwinding relaxation dynamics of polymers*, *Phys. Rev. Lett.* **110**, 068301.

### 2012

[10] Berche B., Kenna R. & **Walter J.-C.**, *Hyperscaling above the upper critical dimension*, *Nucl. Phys. B* **865**, 115.

[9] Hadiwikarta W. W., **Walter J.-C.**, Hooyberghs J. & Carlon E., *Probing Hybridization parameters from microarray experiments: nearest neighbor model and beyond*, *Nucl. Acids Res.* **40**, e138.

[8] **Walter**<sup>\*</sup> **J.-C.**, Ferrantini A., Carlon E. & Vanderzande C., *Fractional Brownian motion and the critical dynamics of zipping polymers*, *Phys. Rev. E* **85**, 031120.

[7] **Walter**<sup>\*</sup> **J.-C.** & Chatelain C., *Relaxation at finite temperature in Fully-Frustrated Ising Models*, *J. Stat. Mech.* P02010.

### 2011

[6] **Walter**<sup>\*</sup> **J.-C.**, Barkema G. T. & Carlon E., *The equilibrium winding angle of a polymer around a bar*, *J. Stat. Mech.* P10020.

[5] **Walter**<sup>\*†</sup> **J.-C.**, Kroll K. M., Hooyberghs J. & Carlon E., *Nonequilibrium effects in DNA microarrays: a multiplatform study*, *J. Phys. Chem. B* **115**, 6732.

### 2009

[4] **Walter**<sup>\*</sup> **J.-C.** & Chatelain C., *Numerical investigation of the aging of the Fully-Frustrated XY Model*, *J. Stat. Mech.* P10017.

### 2008

[3] Berche B., Chatelain C., Dhall C., Kenna R., Low R. & **Walter J.-C.**, *Extended scaling in high dimensions*, *J. Stat. Mech.* P11010.

[2] **Walter**<sup>\*</sup> **J.-C.** & Chatelain C., *Logarithmic corrections in the aging of the 2d Fully-Frustrated Ising Model*, *J. Stat. Mech.* P07005.

**PhD. thesis**

[1] *Numerical study of corrections to the dominant scaling behaviour at and out of equilibrium in spin systems*, **Walter J.-C.**, PhD. thesis (2009) available online : <http://tel.archives-ouvertes.fr>

**3.3.2 Recent presentations 2015–present**• **2019**

- Life of Cancer Cells (Decembre) Montpellier, France, *Biophysical modeling of translation in cancerous human cells*.
- Quantitative Methods in Gene Regulation V (December), London, UK, *Surfing on Protein Waves: Modeling Bacterial Genome Partitioning*.
- Modeling phase separation in health and disease: from nano- to meso-scale (Septembre) Toulouse, France [invited talk] *Surfing on Protein Waves: Modeling Bacterial Genome Partitioning*.
- Physics Meets Biology (Septembre 9-11), University of Oxford, Oxford, UK *Modeling Bacterial Genome Partitioning*.
- Translation Control in Cancer (May), Montpellier, France, *Biophysical Modeling of Translation* [invited].
- Statistical Physics of Complex Systems (May), EPS, Nordita, Stockholm, Sweden, *Surfing on protein waves: proteophoresis as a mechanism for bacterial genome partitioning*.
- APS March Meeting (March), Boston, USA, *Surfing on protein waves: proteophoresis as a mechanism for bacterial genome partitioning*.

• **2018**

- DNA Transactions & Physical and Molecular Biology of Chromosomes, Egmond aan Zee, The Netherlands (Septembre) *Surfing on protein waves: proteophoresis as a mechanism for bacterial genome partitioning*
- Biophyschrom18: The Biology and Physics of Bacterial Chromosome Organisation, Leiden, The Netherlands (June) *Proteophoresis as a mechanism for bacterial genome partitioning*
- Biophysics Today and Beyond: The Physics of Unconventional Systems, Montpellier, France (April) *Surfing on protein waves: proteophoresis as a mechanism for bacterial genome partitioning*

• **2017**

- Quantitative Methods in Gene Regulation IV, Corpus Christi College, Cambridge, UK (December): *Modeling active bacterial DNA segregation*

- iPoLS, Paris (June) *Modeling active bacterial DNA segregation.*
- Seminar at the Institute for Theoretical Physics, Leuven, Belgium (April) *Physical modeling of active bacterial DNA segregation.*
- GDR ADN, Paris, Jussieu (March) *Physical modeling of active bacterial DNA segregation.*
- **2016**
- CECAM Workshop: Mesoscopic Modeling in Physics of Molecular and Cell Biology, Toulouse, France (Octobre) [invited conference] *Physical modeling of active bacterial DNA segregation.*
- Biophyschrom16 : The Biology and Physics of Bacterial Chromosome Organization, Collège de France, Paris, France (Septembre) *Physical modeling of active bacterial DNA segregation.*
- iPoLS Network Annual Meeting, Harvard University, Cambridge, USA (August) [invited conference] *Physical modeling of active bacterial DNA segregation.*
- **2015**
- Quantitative Methods in Gene Regulation III, Corpus Christi College, Cambridge, UK (December): *Stochastic self-assembly of ParB proteins nucleated from the centromere builds bacterial DNA segregation apparatus*
- Seminar at TIMC-IMAG, University of Grenoble (November): *Physical modeling of active bacterial DNA segregation.*
- Seminar at the Lewis-Sigler Institute for Integrative Genomics, Princeton University, USA (October): *Physical modeling of active bacterial DNA segregation.*
- Seminar at the LabEx NUMEV Meeting, Montpellier (July): *Physical modeling of bacterial DNA segregation.*
- Seminar at DIMNP, Montpellier (June): *Physical modeling of bacterial DNA segregation.*
- IPOLS meeting, Montpellier, France (June): *Architecture of a bacterial DNA segregation apparatus.*
- GDR Architecture & Dynamics in the Nucleus national French meeting on DNA organisation), Paris Jussieu, France (April): *Architecture of a bacterial DNA segregation apparatus.*

## 3.4 Experience of mentoring and teaching

### 3.4.1 PhD Students supervision

- Carole Chevalier (2018-2021) Spatio-temporal mechanisms of regulation during translation: Biophysical modeling of high-throughput Ribo-seq experiments (from Master *Physique pour le Vivant* at University Paul Sabatier, Toulouse).
- Paul Soudon (2018-2021) Effect of recycling and finite resources of ribosomes in the spatio-temporal mechanisms of regulation during translation (from *Magistère de Physique Théorique*, ENS, Paris).

### 3.4.2 Supervision of students

- **Master projects:**
  - M1 project, *Modeling ribosomes during the translation of messenger RNA: interactions between adjacent codons*, Margot Bellenguez, Montpellier University (2020).
  - M2 project, *Modeling translation with the TASEP model*, Carole Chevalier, Montpellier University (2018).
  - M1 project, *Study of active bacterial DNA segregation with a reaction-diffusion model*, Yvan Rousset, Montpellier University (2016).
  - M1 project: *Modeling proteins in interaction with proteins*, Thomas Lavoisier, Université Paris-Sud (2015).
  - M1 project, *Modelling of the dynamics of microtubules*, Baptiste Guilhas, Montpellier University (2013).
  - M1 project, *Simulations of zipping dynamics in lattice polymers*, Hans Vandebroeck, KULeuven, Belgium (2011).
- **Bachelor projects:**
  - Bachelor project: *Biophysical modeling of translation*, Mohamed Haddouche, Montpellier University (2019).
  - Bachelor project: *Introduction to TASEP*, Panajot Kristofori, Montpellier University (2018).
  - Bachelor project: *A model for polymers on a Lattice*, Michiel Laleman, KULeuven, Belgium (2011).
  - Bachelor project: *Finite size effects in the 2D-Ising model*, Michiel Laleman, KULeuven, Belgium (2010).
- **High School projects:**
  - Priscille Beaurain & Barry Ronca, 2019.

### 3.4.3 Teaching experience

- Computational methods for modeling biopolymers in interactions with proteins: application to the bacterial DNA segregation system, Master 2 students, Montpellier University (2017, 2019).
- Numerical methods for Physicists for Bachelor students, Montpellier University (2015–2017).
- Tutorial of mathematical methods for physicists: stochastic processes, complex analysis, group theory and non-linear dynamics & chaos for Bachelor students, KULeuven, Belgium (2010–2011).
- **During the PhD thesis** (2006–2009):
- Lecture and practical works of electrokinetics in 2nd year; Practical works of nuclear physics in 3rd year; Supervision of projects in 1st year; Lecture and practical works to prepare students for the school teacher examination; Tutorials of mathematics for physicists in 3rd year.

### 3.5 Various collective activities

- Organisation of a workshop *Supercoiling in Bacterial DNA* at Sète, France (November 2019).
- Member of the organizing team *Risques Psycho-Sociaux* of the Laboratory Charles Coulomb.
- Referee activity for J. Stat. Mech., Physica A, Physical Biology, *Agence Nationale pour la Recherche* (ANR).
- Chairman at APS March Meeting Boston, session "Active Matter in Biology" (2019).
- Membre of a PhD Committee (2016, Hasselt, Belgium, Student: Hans Vandebroek, Supervisor: Carlo Vanderzande).
- Reach out in high school to explain the career of a Researcher.



## Chapter 4

# Research Interests

I realized my PhD thesis in theoretical statistical physics on the topic of magnetic spin systems at and out of equilibrium (2006–2009). During my different postdocs (2010–2017), I transferred these skills to the field of the physical properties of polymers, biopolymers and biological assemblies of macromolecules and proteins. In KULeuven (2010–2011), I namely realised the design, analysis and modeling of experiments on microarrays, which are device using the hybridization properties of DNA to measure the concentration of unknown DNA strands in solution. In parallel, I had a long term project on the fundamental aspects of the rotational dynamics of polymers wound around each other or around a fix cylinder (2011–2016). around a cylinder, as well as the dynamics of hybridization of two complementary strands (2012–2017). I finally dived into the description of cellular processes *in vivo* (2015–present): my most recent studies concern the description of bacterial DNA active partition system during cell division and the translation of messenger RNA by ribosomes in human cells. I am using both numerical simulations (Monte Carlo, Molecular Dynamics) and analytical methods to investigate the functioning of biological systems.

### 4.1 Spin systems at and out of equilibrium (2006–2012)

I realised my PhD on the topic of phase transition in magnetic systems at and out-of-equilibrium. In a first part, I studied aging in 2D spin systems displaying frustration (fully Frustrated Ising and XY models) during a thermal quench between the high and low temperature phase. The frustration of the interactions leads to an exponential degeneracy of the ground state with respect to the system size. We have characterized the universality class of the dynamics of these systems at and out of equilibrium by considering the symmetry breaking associated to the quench in temperature. In a second part, I studied the finite size effects of the critical Ising model in high dimension ( $D > 4$ ). Finite size effects are called anomalous and are explained by the renormalization group with the existence of a scaling variable said “dangerous”. This mathematical analysis is possible only for periodic boundary conditions, and evidences of this behaviour for free boundary conditions was still lacking and debated. This issue of boundary con-

ditions was essential to promote this anomalous behaviour as an universal behaviour, i.e., independent of boundary conditions. I have shown with Monte Carlo simulations that this is also true for free boundary conditions.

## 4.2 Description of polymers and biopolymers

### 4.2.1 Physical behavior of DNA Microarrays (2010–2011)

DNA microarrays allow, many other applications, to analyse large scale gene expression, diagnosis of the state of a cell (e.g. cancer, HIV), mutation of viruses etc. In Figure 4.1, I illustrate the functioning of a DNA microarray. The physical behavior of DNA microarrays, still largely unveiled, is important in order to use them in a quantitative way. Our theoretical work is supported by experiments that we designed using statistical criteria (optimal design). We have shown that under certain circumstances, the equilibration time of the arrays exceeds the typical experimental time. We have developed a model that allows metastable states where a strand in solution can bind with many strands on the array. Experimental data are modelled successfully with our model, containing a small number of free parameters [138]. Another aspect of my work on DNA microarrays concerns the effect of the surface on the thermodynamic of hybridization[57]. We have shown the limits of the nearest-neighbor approximation (used to estimate the hybridization free-energy) in a quantitative way for DNA microarrays. We also discuss the effect of anti parallel GA mismatches displaying the signature of a structural conformational change of the double helix. See my References [138, 57] for more details.

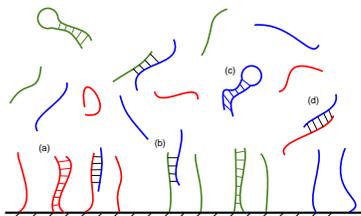


Figure 4.1: Illustration of an experiment with DNA microarrays. Known strands are synthesized and fixed to the surface. We put them into contact with an solution of labelled strands of unknown concentration. After reading the spot-dependent fluorescent intensity on the microarray, it is possible to obtain the concentration of strands in solution. However, many phenomena can occur, preventing microarray to be used into a precise quantitative tool : (a) specific hybridization with matching strands (b) non-specific hybridization (c) formation of loops (d) cross-hybridization. We have shown that cross-hybridization slows down significantly the dynamics of hybridization.

### 4.2.2 Winding and zipping properties of a polymer (2011–2016)

An important topic of research during these last years concerns the static and dynamical properties of polymers in the framework of the description of biopolymer properties, namely DNA conformation and dynamics: winding and zipping properties. The dynamics of DNA has direct implication in biological processes – the crowded environment of the cells – like replication where the double helix has to unwind and unzip. To get solid theoretical ground, we considered fundamental polymer physics questions: (a) polymer wound around each other (b) a polymer wound around a fix cylinder and (c) a polymer pulled in a rotational motion at constant angular velocity  $\Omega$ , as sketched in Figure 4.2. First, we have shown the interplay between the free-energy associated to the winding angle and the relation dynamics according to a Langevin equation description [141]. We described precisely the total relaxation time which involves multiplicative logarithmic corrections. For the short time dynamics, the released winding angle is also found to be in good agreement with the Langevin equation. We gave elements related to the shape of the rotating polymer. This study has been performed both with Monte Carlo simulations of polymers on a lattice and analytical calculation. Subsequently, we gave a theoretical framework to explain in details the dynamics of a rotating polymer at constant velocity with the use of the blob picture introduced in the context of polymer pulled by a constant force at one end. Our results apply to the case of a polymer in rotation at constant velocity where we found different conformational regime depending on the angular velocity. See References [139, 142, 143, 75] for more details.

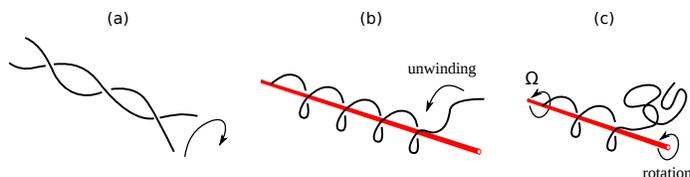


Figure 4.2: Examples of rotational dynamics of entangled polymers. (a) The unwinding dynamics of two polymer chains dissociating from each other; (b) the unwinding of a single chain from a rigid rod; (c) the stationary conformation of a polymer attached to a rod rotating with a constant angular velocity  $\Omega$ . We have given a theoretical framework for the dynamics of such system using the “blob picture” originally introduced for the case of polymer pulled by a constant force at one end.

I also studied the aspect of zipping and unzipping of two polymers in interaction with a contact potential [140, 109], see Figure 4.3. We can show that this interacting system displays a critical point. At criticality, this process shows analogies with translocation. Namely, we can map this problem on a one-dimensional Brownian particle whose coordinate corresponds to the position of the fork point where the two strands starts to unzip. I gave evidences that the dynamics is described as a fractional Brownian motion. Our conclusion is supported by the calculation of the distribution probability of the position, the autocorrelation of the speed, and the survival probability. This study has also been supported by Monte Carlo simulations on a lattice. In a

subsequent paper, we solved analytically this model with an appropriate kernel motivated by polymer physics. These analytical results are in full agreement (to the second order) with numerical simulations. See References [140, 109] for more details.

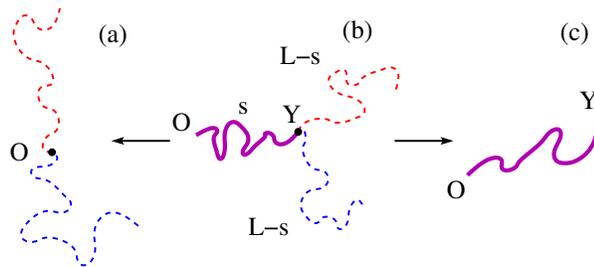


Figure 4.3: Zipping of two polymers interacting with a contact potential at complementary monomers. (a) Polymers are fully unzipped; (b) at the critical point, the polymers are in average half-zipped. Violet line is the zipped part. The reaction coordinate of the system is the linear position  $s$  of the fork point  $Y$ . We have shown numerically and analytically that it displays a Fractional Brownian Motion with characteristic exponents; (c) the polymers are fully zipped.

### 4.3 Collective behaviour of protein assemblies in the cell (2014–present)

#### 4.3.1 Organisation of the genome during active bacterial DNA segregation

This project on active bacterial segregation is the main topic of the scientific part of this thesis, thus we refer the reader to the previous Chapter. This project led to the following publications in Refs.[112, 144, 145, 32, 30, 54].

#### 4.3.2 Modeling the cytoplasmic active transport

Eukaryotic cells exploit different ways to transport matter in their interior such as passive Brownian motion or active cytoskeletal transport driven by motor proteins. The transportation of matter (such as proteins and organelles) is performed via a network of filaments, the microtubules and the microfilaments (Fig.2, left). On these filaments, motor proteins perform a walk using chemical energy from ATP (Adenosin-Tri-Phosphate) hydrolysis. In this way, matter alternates diffusive motion (inefficient in the crowded medium of cells) and active transport by motor proteins. Recent progress in experimental approaches such as fluorescent microscopies and spectroscopies have allowed a possibility to observe and quantify molecular events of transport. Exclusion processes are known to be good candidates for understanding how collective and non-equilibrium effects of large clusters of motors arise from the molecular properties of

single motors. These models consist in particles which can jump (possibly asymmetrically) on a one dimensional lattice with the constraint of volume exclusion. Generalization of these models for cytoskeletal transport include attachment and detachment of particles in contact with a reservoir and bidirectional motion, for example.

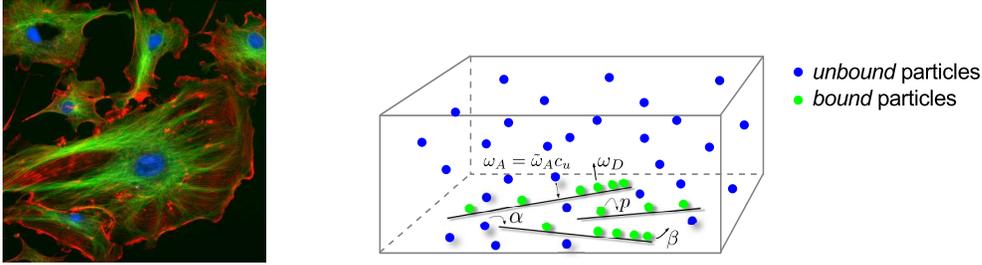


Figure 4.4: (Left) The cytoskeleton of an eukaryotic cell: the microtubules (MT) are shown in green, the actin filaments in red and the nuclei in blue. (Right) Model of directed transport (TASEP-LK) coupled to a reservoir of diffusive particles. Filaments (black segments) are immersed in a closed compartment constituting a reservoir. *Unbound* particles (in blue, at the concentration  $c_u$ ) are freely diffusing in the reservoir and can attach at the rate  $\omega_A = \tilde{\omega}_A c_u$  at any site, or at the entry at the rate  $\alpha$ . Particles *bound* can detach at the rate  $\omega_D$ , exit at the rate  $\beta$  and hop internally at a rate  $p$ .

The model we have used to model Microtubules is a TASEP filament with attachment and detachment, also known as TASEP-LK (LK stands for Langmuir Kinetics). A sketch of this model with the relevant parameters is shown in Fig.4.4 (right). In our first paper, we present a model to study the fine balance between supply and demand for molecular motors (such as kinesins, dyneins and myosins) during intracellular transport processes. In particular, we generalise the TASEP-LK: in such a model motors with finite processivity actively advance (with steric interactions) along a filament and detach from the filament into a reservoir with unlimited resources, i.e. unlimited motors. Since the amount of motors involved in active transport inside a cell is far from being infinite, here we study a model with a homogeneous but limited reservoir of motors, which explicitly takes into account the competition for particles (supply) between the filament's binding sites (demand). Our theoretical results allow to analytically study the density profiles of motors bound on a filament as a function of the concentration of binding sites (bound tubulin dimers in the case of microtubules). We have used analytical calculations and extensive Monte Carlo simulations to study this model in Ref. [27].

This work has led to a natural extension in the project part of this thesis: the transport of ribosomes on messenger RNA during translation, namely in human cancer cells.