Mechanisms of Spatial Organisation Within Bacterial Cells

Dissertation

zur Erlangung des Doktorgrades
der Naturwissenschaften
(Dr. rer. nat.)

dem Fachbereich Mathematik und Informatik
der Phillipps-Universität Marburg
vorgelegt von

Lara Connolley
aus High Wycombe, Vereinigtes Königreich

Marburg an der Lahn, Deutschland, Oktober 2023

Die vorliegende Dissertation wurde von Juni 2020 bis Oktober 2023 am Max-Planck-Institut für terrestrische Mikrobiologie unter Leitung von Dr. Seán Murray angefertigt.

Vom Fachbereich Mathematik und Informatik der Philipps-Universität Marburg (Hochschulkennziffer 1180) als Dissertation angenommen am ________________

Erstgutachter: Dr. Seán Murray
Zweitgutachter: Prof. Dr. Stephan Dahlke

Weitere Mitglieder der Prüfungskommission:
Prof. Dr. Christian Rieger
Prof. Dr. Martin Thanbichler

Tag der Disputation: 14/12/2023

ERKLÄRUNG

Ich versichere, dass ich meine Dissertation mit dem Titel „Mechanisms of spatial organisation within bacterial cells“ selbstständig ohne unerlaubte Hilfe angefertigt und mich dabei keiner anderen als der von mir ausdrücklich bezeichneten Quellen und Hilfsmittel bedient habe.

Diese Dissertation wurde in der jetzigen oder einer ähnlichen Form noch bei keiner anderen Hochschule eingereicht und hat noch keinen sonstigen Prüfungszwecken gedient.

Marburg an der Lahn, Deutschland, Oktober 2023

Lara Connolley

PUBLICATIONS

Some ideas and figures have appeared previously in the following publications:

1. **Lara Connolley**, Joanna Szczepaniak, Colin Kleanthous and Seán M. Murray. The quantitative basis for the redistribution of immobile bacterial lipoproteins to division septa. *PLoS Computational Biology* **17**(12), e1009756 (2021).
2. Manuel Osorio-Valeriano, Florian Altegoer, Chandan K. Das, Wieland Steinchen, Gaël Panis, **Lara Connolley**, Giacomo Giacomelli et al. The CTPase activity of ParB determines the size and dynamics of prokaryotic DNA partition complexes. *Molecular Cell* **81**(19), 3992-4007 (2021).
3. **Lara Connolley**, Lucas Schnabel, Martin Thanbichler and Seán M. Murray. Partition complex structure can arise from sliding and bridging of ParB dimers. *Nature Communications* **14**(1), 4567 (2023).

In addition, the following is in preparation:

4. **Lara Connolley** and Seán M. Murray. ParB-ParB tethers: An alternative model for ParABS segregation. *To be submitted*.

ZUSAMMENFASSUNG

Aufgrund ihrer geringen Größe glaubte man früher, Bakterien seien nicht mehr als Proteinsäcke und würden sich lediglich auf einfache Diffusion verlassen, um alle für ihr Überleben notwendigen chemischen Reaktionen auszuführen. Heute weiß man jedoch, dass sie eine komplexe interne Organisation besitzen, auch wenn ihnen die durch Membranen abgegrenzten Organellen fehlen, die in eukaryotischen Zellen zu finden sind. Bakterien sind in der Lage, Proteine präzise zu positionieren, Protein-Oszillationen zu koordinieren und Chromosomen sowie Plasmide effektiv zu segregieren. Obwohl moderne Mikroskopietechniken es uns erlauben, die innere Organisation zu visualisieren, sind die genauen molekularen Interaktionen, die diesen Phänomenen zugrunde liegen, nach wie vor schwer zu fassen. Die Anwendung mathematischer und physikalischer Modelle ist daher ein nützlicher Ansatz, um die grundlegenden Mechanismen der bakteriellen Organisation zu verstehen.

In dieser Arbeit setzen wir mathematische und physikalische Modelle ein, um die Feinheiten der räumlichen Organisation von bakteriellen Zellen zu untersuchen. Im Mittelpunkt unserer Untersuchung stehen zwei unterschiedliche Phänomene: Die Umverteilung eines Proteins des Tol-Pal-Systems (wie im Teil i erläutert), sowie der Mechanismus, der im Par*ABS* -System die Segregation von genetischem Material und Plasmiden mit geringer Kopienzahl steuert (wie im Teil ii beschrieben). Basierend auf experimentellen Daten entwickeln wir Modelle, um die Funktionsweise und Interaktionen innerhalb dieser Systeme zu entschlüsseln.

Der erste Teil dieser Arbeit verwendet einen deterministischen Modellierungsansatz zur Beschreibung der Proteinrelokalisierung durch ein System von Differentialgleichungen. Wir beginnen mit der analytischen Lösung eines vereinfachten Modells, das die Essenz des biologischen Problems erfasst. Im Anschluss erweitern wir dieses Modell, um eine detailliertere Abbildung des Systems zu erhalten, welche wir dann numerisch lösen. Wir zeigen, dass Proteine durch einen "Mobilisierungs- und Einfangmechanismus"lokalisiert werden können.

Im zweiten Teil wenden wir die stochastische Modellierung an und zeigen zunächst, dass es theoretisch möglich ist, das beobachtete Ausbreitungsmuster auf der DNA durch einen auf Diffusion basierenden Gleitmechanismus des Proteins zu erklären. Anschließend zeigen wir, dass Protein-Protein-Brücken aufgrund ihrer Lebensdauer Haarnadel- und Helixstrukturen in der DNA erzeugen können. Wir vereinen Gleiten und Brückenbildung in einem einheitlichen Modell und beobachten, dass kurzlebige Brücken das Gleiten nicht beeinträchtigen, jedoch sowohl das Protein-Bindungsprofil als auch die erwartete DNA-Kondensation reproduzieren können. Schließlich entwickeln wir ein Modell für einen alternativen Mechanismus der DNA-Segregation durch das Par*ABS* -System, das das experimentell beobachtete Verhalten widerspiegelt.

ABSTRACT

Due to their small size bacteria were once thought to be bags of proteins relying solely on simple diffusion to carry out all the chemical reactions needed for their survival. However, it is now clear that they have a complex internal organisation, despite lacking the membrane-bound organelles found in eukaryotic cells. Bacteria are able to position proteins precisely, orchestrate protein oscillations, and segregate the chromosome and plasmids successfully. Although contemporary microscopy techniques have illuminated the presence of spatial organisation, the precise molecular interactions underlying these phenomena have remained elusive. To address this, the application of mathematical and physical models emerges as a useful approach to understand the fundamental mechanisms behind bacterial organisation.

In this thesis, we leverage mathematical and physical modelling to explore the intricacies of bacterial cell spatial organisation. Our investigation will center on two distinct phenomena: the redistribution of a protein within the Tol-Pal system (as detailed in Part i), and the underlying mechanism governing the segregation of genetic material and low-copy number plasmids by the ParAB*S* system (as detailed in Part ii). Utilising experimental data, we construct models allowing us to decipher the operational mechanisms and interactions between components within these systems.

The first part of this thesis employs a deterministic modelling approach to describe protein relocalisation through a system of differential equations. We begin by analytically solving a simplified model that encapsulates the essence of the biological problem. Subsequently, we expand this model to provide a more comprehensive representation of the system, which we solve numerically. We show that proteins can localise using a 'mobilisation and capture' mechanism.

In the second part we pivot to utilise stochastic modelling. Initially, we show that it is theoretically possible for a diffusion based mechanism of protein sliding to explain the observed pattern of spreading on DNA. Then we look at the effects of protein-protein bridges on the DNA and demonstrate how these can organise the DNA into globular states or hairpin and helical structures, depending on bridge lifetime. Combining sliding and bridging into a unified model, we find that short-lived bridges do not impede sliding and can reproduce both the protein binding profile and the expected DNA condensation. Finally, we develop a model for an alternative mechanism of DNA segregation by the ParAB*S* system that can reproduce behaviour observed experimentally.

CONTENTS

Introduction

I Redistribution of immobile bacterial lipoproteins to division septa

II Segregation of bacterial DNA by the ParABS system

Summary

INTRODUCTION

INTRODUCTION

The primary distinction between the two major domains of life, eukaryotes and prokaryotes, is that prokaryotes largely lack membrane-bound organelles. Initially, the absence of internal organelles in prokaryotes led to the assumption that these organisms also lacked intricate internal structure. It was widely held that all the proteins needed for their survival were randomly distributed within the cell, and that all cellular processes could be accomplished via simple diffusion alone. Advancements in imaging capabilities have since enabled better visualisation of bacterial cells and it has become clear that they have a complex internal structure, in which spatial organisation arises in the absence of membrane boundaries.

The first insights into bacterial intracellular organisation were the discoveries that bacterial cells can cluster chemoreceptors at specific locations [1, 2], possess cytoskeletal structures [3, 4], produce protein oscillations [5, 6], and spatially organise and actively segregate chromosomal regions [7–9]. These works have since been followed by many others illustrating the ubiquitous nature of spatial order within different areas of bacterial life. In fact, spatiotemporal organisation is vital for essential cellular processes such as cell growth and division, DNA segregation, cell cycle regulation, cell differentiation, and motility. The list of bacterial macromolecules known to have a distinct subcellular localisation is rapidly increasing.

Whilst we are now aware of the vast array of internal bacterial organisation present, the mechanisms behind this organisation have often remained evasive. The molecules involved in these processes are exceedingly small, and their interactions are frequently beyond the reach of current microscopy techniques. In this thesis, we will leverage the power of mathematical and physical modelling to overcome these limitations. Modelling serves as a simplified representation of reality, allowing us to extract valuable insights into the real system. A successful mathematical model strives not only to replicate experimental findings but, ideally, to go beyond them by offering predictions about the behaviour of the modelled system. While it is important to acknowledge that a mathematical model may never fully encapsulate the intricacies of the corresponding biological system, this limitation does not diminish the utility of modelling as a valuable tool. In our context, models facilitate our understanding of the underlying dynamics and molecular interactions governing spatial organisation within bacterial cells.

The choice of modelling approach is intricately linked to the specific biological problem at hand and the questions being asked. In this thesis, we will employ both deterministic and stochastic modelling techniques. There is one key difference between these two types, in deterministic models the stochasticity within the system is ignored. A convenient standard approach in deterministic modelling is the use of differential equations. However, this relies on the simplifying ansatz that there is no noise, in contrast to biological systems where noise is inherent. In systems where there are a large number of molecules this stochasticity may be effectively averaged out, creating the appearance of a deterministic process. On the other hand, stochastic models can be used where the effects of this stochasticity are important for understanding the system. The analytical solution to the chemical master equation, a frequently used method, is often intractable. Instead the Gillespie algorithm can be used to provide exact simulations of the master equation [10].

In the following, we aim to provide a mechanistic understanding of two biological processes occurring within bacterial cells. Firstly, we will address the mechanism behind the redistribution of an immobile outer membrane protein as part of the Tol-Pal system (as described in Part i). Within this part we will rely on deterministic modelling utilising systems of differential equations. Secondly, we will investigate the mechanism behind the segregation of genetic material and low-copy number plasmids by the ParAB*S* system (as described in Part ii). Here we will mainly use stochastic modelling to describe protein sliding, to model the DNA in terms of a polymer model, and to model segregation. Within each part we will first introduce the underlying biological phenomena of interest and provide the requisite contextual background. Subsequently, we will present the mathematical foundations that underlie the models we will construct. The succeeding chapters will then navigate through the research conducted and the conclusions derived, with the overarching aim of providing a mechanistic description of each respective phenomena.

Part I

REDISTRIBUTION OF IMMOBILE BACTERIAL LIPOPROTEINS TO DIVISION SEPTA

1

BACKGROUND: HOW AND WHY BACTERIA LOCALISE PROTEINS

In order to produce spatial organisation in the absence of membrane bound compartments bacteria have evolved a multitude of different mechanisms. The ability of bacteria to perform dynamic subcellular protein localisation, synchronised with the cell cycle, is critical for cells to survive. In this part, we will focus on the task of precisely localising proteins to specific cellular regions.

1.1 BIOLOGICAL IMPORTANCE OF PROTEIN LOCALISATION

One of the earliest observations of protein localisation in bacterial cells was the discovery that chemotaxis receptors exhibited clustering at the cell poles in *Escherichia coli* and *Caulobacter crescentus* [1, 2, 11]. Chemotaxis, which enables bacteria to navigate their surroundings in response to chemical gradients, critically relies on the formation of these receptor clusters. Such clustering serves to amplify the input signal, allowing cells to detect and respond to even subtle nutrient gradients [12, 13]. Since these first discoveries, subsequent research has shown chemotactic clustering to be a universal occurrence, observed in all bacterial species studied to date [14]. This underlines its fundamental significance and exemplifies the pivotal role of protein localisation within bacterial cells.

Beyond chemotaxis, the phenomenon of protein localisation within bacterial cells extends to a diverse range of proteins. For instance, the cell division protein FtsZ assembles into a cytokinetic ring precisely at the mid cell of *E. coli* [3]. In *Bacillus subtilis*, the membrane phosphatase SpollE localises to the sporulation septum [15], while the actin polymerisation protein IcsA in *Shigella flexneri* targets the older of the two cell poles [16].

Whilst in some cases accurate protein localisation is not essential for cellular survival, such as for the flagellum, it plays a vital role in numerous other cases. Proteins involved in critical processes like DNA replication, chromosome segregation, or cell division rely on precise localisation within the bacterial cell to fulfill their functions. Consequently, protein localisation has emerged as an indispensable and critical process for cells to function [17].

1.2 COMMON MECHANISM OF PROTEIN LOCALISATION

In bacterial cells, the localisation of proteins is predominantly accomplished through passive mechanisms such as recruitment by an existing landmark (or inhibition by an anti-landmark), or the sensing of membrane curvature or cell geometry [18]. The actual transport of the proteins is typically governed by pure diffusion, relying on the protein coming into spatial contact with its target factor. This process is aptly termed 'diffusion and capture'. In order for this process to be effective, the protein to be localised must possess sufficient mobility such that it can reach its designated location within an appropriate timeframe.

This leads to an intriguing question: What if a protein is not sufficiently diffusive to locate its target within a reasonable timeframe?

1.3 AN EXAMPLE OF PROTEIN LOCALISATION: TOL-PAL SYSTEM

Membrane proteins are typically slowly diffusive [19]. Although they must only diffuse within two dimensions (in contrast to the three dimensions for cytoplasmic proteins), this slower rate of diffusion can pose a challenge for their localisation through conventional diffusion and capture mechanisms. In this case, active processes must be required for their localisation.

Bacterial cell membrane. Unlike cells of higher organisms, bacteria are frequently faced with unpredictable and often harsh environments. To survive in such conditions, they have evolved a sophisticated and complex cell envelope. This envelope must serve the dual purpose of shielding the bacteria from the external environment while facilitating the exchange of nutrients from the outside and the removal of waste products from the interior.

One way bacteria can be classified is based on their response to a particular staining method known as the Gram stain. This results in two distinct groups: Gram positive, where the stain is retained, and Gram negative, where it is not. This disparity in staining is attributed to differences in the composition of the bacterial cell envelope. Gram-negative bacteria possess a membrane envelope comprising three primary layers: the outer membrane, the peptidoglycan cell wall, and the cytoplasmic (inner) membrane. These two membrane layers enclose an aqueous cellular compartment referred to as the periplasm. In contrast, Gram-positive bacteria lack the outer membrane and have a thicker cell wall [20].

Of note, the outer membrane in Gram-negative bacteria is not energised, and there is no ATP in the periplasm. Consequently, any active processes occurring in the periplasm must be powered by the proton motive force (PMF) across the inner membrane or ATP hydrolysis within the cytoplasm.

An example case of a membrane-bound protein with a notably slow rate of diffusion is Pal (Figure 1.1). Pal is a widely-conserved outer membrane lipoprotein found in Gram-negative bacteria, tethering the outer membrane to the cell wall via its peptidoglycan binding domain [22]. During cell division, the outer membrane must constrict in coordination with the invagination of the cell wall. This coordina-

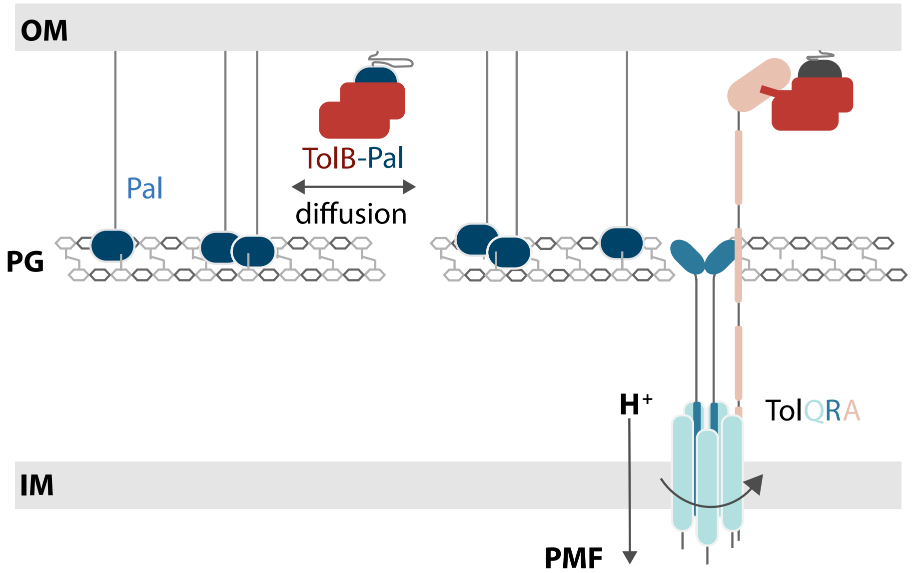


Figure 1.1: Major components of the Tol-Pal system. Pal is an outer membrane (OM) lipoprotein (blue) that can bind to peptidoglycan (PG) or to the periplasmic protein TolB (red) in a mutually exclusive manner. The inner membrane (IM) protein TolA spans the periplasm and is coupled to the PMF via its interaction with TolQ and TolR. TolA is also able to interact with TolB and can form a tripartite complex of TolA-TolB-Pal [21].

tion is believed to be orchestrated by the recruitment of Pal to the division septum by the multi-protein Tol system. The Tol system localises to the constriction site independently of Pal as part of the division apparatus [23–25].

The Tol-Pal system comprises TolQ, TolR, and TolA, which form an inner membrane stator complex, as well as the periplasmic protein TolB, and the outer membrane protein Pal [23, 25]. The binding of Pal to TolB prevents Pal from binding to the peptidoglycan layer. However, TolB can still interact with TolA even when in complex with Pal (Figure 1.1). Previous molecular dynamic simulations have suggested that this interaction, though relatively weak, has the ability to dissociate TolB from its complex with Pal by applying a pulling force.

In recent work on *E. coli*, it was found that Pal diffuses extremely slowly, with a measured diffusion coefficient through single particle tracking of $\leq 0.004\,\mu m^2 s^{-1}$ (likely an upper bound due to the limits of the technique) [26]. Such slow diffusion was attributed to Pal binding to the cell wall, rather than the embedding of its lipoylated domain in the inner leaflet of the outer membrane. Indeed, lipoylated PAmCherry, displayed a diffusion coefficient more in line with typical membrane proteins of $\sim 0.02\,\mu m^2 s^{-1}$. In light of these observations, we focus on how Pal achieves redistribution to the septum given its exceptionally slow diffusive characteristics.

2

BACKGROUND MATHEMATICS FOR MODELLING PROTEIN LOCALISATION

In order to gain insights into how the relocalisation of a slowly diffusive protein can take place we turn to mathematical modelling. The standard approach to study such spatial organisation is through reaction-diffusion systems.

2.1 REACTION-DIFFUSION SYSTEMS

A classical diffusion equation is a parabolic partial differential equation which has a general form obtained via conservation of mass and Fick's law:

$$\partial_t u = D\nabla^2 u, \quad t \geq 0, \quad x \in \Omega \subseteq \mathbb{R}^d, \quad u(x,t) \in \mathbb{R}_{\geq 0}, \tag{2.1}$$

where $u(x,t)$ is a variable usually describing the concentration of the diffusive particles, and D is the rate of diffusion. The general reaction-diffusion equation, introduced by Fisher, Kolmogorov, Petrovsky and Piskunov [27], is then a semi-linear parabolic equation given by,

$$\partial_t u = D\nabla^2 u + f(u), \quad t \geq 0, \quad x \in \Omega \subseteq \mathbb{R}^d, \quad u(x,t) \in \mathbb{R}_{\geq 0}, \tag{2.2}$$

with an initial condition $u(0,x) = u_0(x)$ and boundary conditions if Ω is a bounded domain.

Whilst in general $u(x,t) \in \mathbb{R}^n_{\geq 0}$, we will consider $n = 2$ in equations 3.8 and 3.9, and we will take $d = 1$ throughout, i.e. we will work in one spatial dimension. Furthermore, in this thesis we will only consider the case of homogeneous Neumann boundary conditions, $\frac{\partial}{\partial v} u(x,t) = 0$ on $\partial\Omega$, where v is the outward normal, since these will be the most biologically relevant.

One powerful method used for solving partial differential equations is that of Green's functions. In the next section we will first introduce some preliminary ideas needed to understand these.

2.2 THE DIRAC DELTA DISTRIBUTION

The Dirac delta distribution, $\delta(x)$, is one example of a *generalised function* or a *distribution*. It is an object that arises as the limit of a function and whilst it behaves like a function under integration it is not, strictly speaking, a function in itself.

Definition 1. *The* ***Dirac delta distribution*** *can be defined by the following two properties,*

$$\int_{-\infty}^{\infty} \delta(x)dx = 1, \tag{2.3}$$

$$\int_{-\infty}^{\infty} \phi(x)\delta(x-\xi)dx = \phi(\xi). \tag{2.4}$$

where ϕ is continuous at $x = \xi$.

The last property is sometimes called the *sifting property* of the δ distribution.

2.3 GREEN'S FUNCTIONS

In applied mathematics, Green's functions are auxiliary functions in the solution of linear partial differential equations.

Consider a general linear partial differential equation of the form,

$$L(x)u(x) = f(x) \quad \text{in } \Omega, \tag{2.5}$$

where $L(x)$ is a linear differential operator, $u(x)$ is the unknown and $f(x)$ is a known homogeneous term. The solution to the equation can be formally written as,

$$u(x) = L^{-1}f(x), \tag{2.6}$$

where L^{-1}, the inverse of L , is some integral operator. We define the inverse L^{-1} using a *Green's function*. Let,

$$u(x) = L^{-1}f(x) = \int_{\Omega} G(x,\xi)f(\xi)d\xi, \tag{2.7}$$

where the kernel of the integral operator $G(x,\xi)$, is the Green's function associated with L. Now, applying L to equation 2.7 we get,

$$Lu(x) = f(x) = \int_{\Omega} LG(x,\xi)f(\xi)d\xi. \tag{2.8}$$

Definition 2. *The* ***Green's function*** *is a kernel $G(x,\xi)$ satisfying*

$$u(x) = \int_{\Omega} G(x,\xi)f(\xi)d\xi \quad \textit{with} \quad LG(x,\xi) = \delta(x-\xi) \quad \textit{and} \quad x,\xi \in \Omega. \tag{2.9}$$

3

THE QUANTITATIVE BASIS FOR THE REDISTRIBUTION OF IMMOBILE BACTERIAL LIPOPROTEINS TO DIVISION SEPTA

As described in the background biological information (Chapter 1), the Tol-Pal system has been proposed to facilitate outer membrane constriction by populating the division septum with Pal [24]. To investigate this further, Szczepaniak et al. [26] developed a new method, known as SpatialFRAP, for the analysis of Fluorescence Recovery After Photobleaching (FRAP) experiments. This leverages the Fokker-Planck equation for diffusion in homogeneous media, to evaluate the mobility of a protein across the length of the cell (Figure 3.1a, b). The output of this analysis is $D_{\text{eff}}(x)$, a spatially-varying effective diffusion coefficient. This technique was used to measure Pal mobility in dividing (defined as having a visible constriction) and non-dividing (no visible constriction) cells.

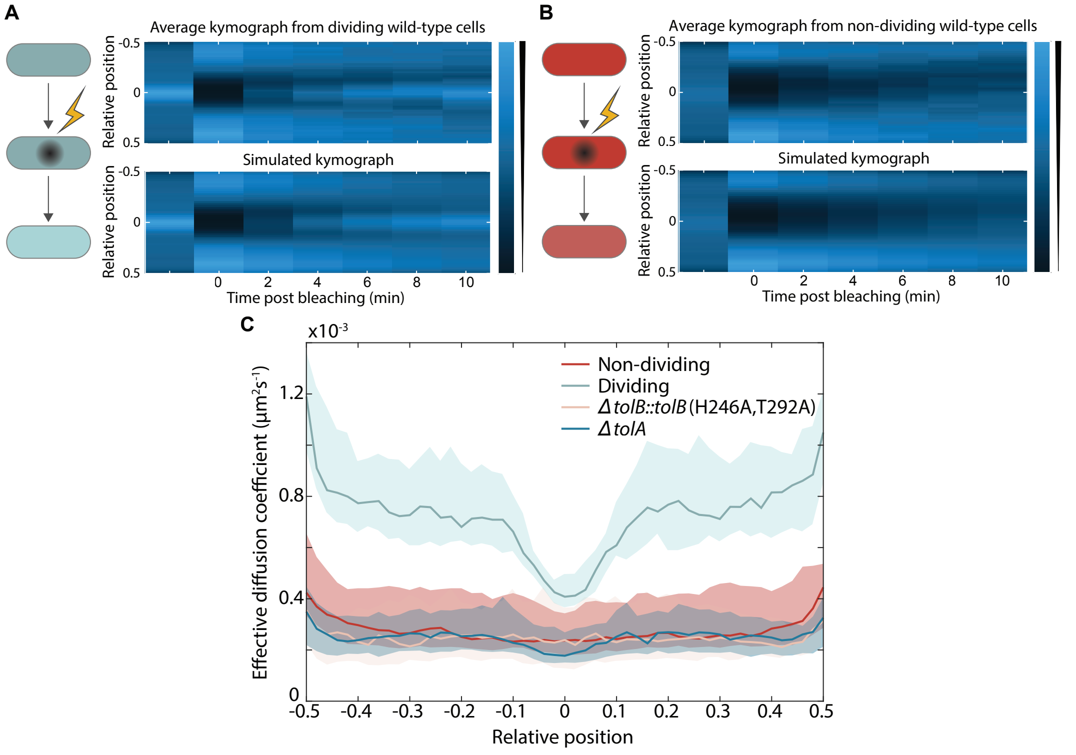


Figure 3.1: Pal in dividing cells is mobilised by TolQRA. (a) Top, average kymograph of Fluorescence Recovery After Photobleaching (FRAP) of Pal-mCherry over 30 dividing wild type cells where the color scale indicates normalised fluorescence. Bottom, simulated kymograph obtained using the SpatialFRAP method by fitting to the data above. (b) Same as in (a) for 30 non-dividing wild type cells. (c) The effective diffusion coefficient of Pal as a function of the relative cellular location in individual non-dividing, dividing, Δ*tolA*, and Δ*tolB::tolB*(H246A, T292A) cells. TolB H246A T292A is a mutant that is unable to bind Pal. Shown is the median as a function of relative long axis position with 95% confidence intervals for approximately 30 cells. Data reproduced from Szczepaniak et al. [26].

It was discovered that, whilst still slow, Pal mobility in dividing cells is increased throughout the cell, except at the septum, where it accumulates (Figure 3.1c). Note we expect the further increase at the edges to be due to edge effects as cells are narrower at their poles. Thus, instead of Pal being made less mobile at the septum, which is the commonly found diffusion and capture scenario, it is made more mobile everywhere else. Notably, this behaviour was contingent upon a functional Tol system. Specifically, the observed increase in mobility and subsequent accumulation at the septum requires TolA, TolA coupling to the PMF, and the binding interactions between TolB and Pal and TolB and TolA [21, 26].

To provide an explanation for these results, it was posited that Pal mobility arises from its binding to TolB in the periplasm (Figure 1.1). Since this binding is mutually exclusive with Pal binding to peptidoglycan, TolB-Pal complexes would be expected to diffuse much more rapidly. Additionally, it was proposed that TolQRA machines, powered by the PMF, exert a force on TolB pulling it away from Pal which is released to rebind to the cell wall. This hypothesis gains support from the structural similarities between the TolA-TolB complex and that of another PMF-driven machine, TonB, which binds to TonB-dependent transporters [26, 28]. TonB, powered by its inner membrane ExbB-ExbD stator complex, displaces the plug domains of outer membrane transporters in order to import bound nutrient substrates [29]. Since in dividing cells TolQRA machines are localised to the septum as part of the division machinery [23], Pal would be deposited at the septum to fulfill its role in outer membrane constriction.

SpatialFRAP. In Szczepaniak et al. [26] a SpatialFRAP method was developed to take advanatage of the information provided by FRAP across the length of the cell rather than exclusively looking within a particular region of interest. This method is first presented here and then generalised.

Considering a molecule of interest (Pal) which exchanges between mobile and immobile states according to its binding to peptidoglycan we have the following equations,

$$\frac{\partial u}{\partial t} = D\frac{\partial^2 u}{\partial x^2} - k_{\text{on}}(x)u + k_{\text{off}}(x)v \; , \tag{3.1}$$

$$\frac{\partial v}{\partial t} = +k_{\text{on}}(x)u - k_{\text{off}}(x)v \; , \tag{3.2}$$

where we denote the concentration of free and bound Pal by $u(x,t)$ and $v(x,t)$ respectively and binding to, and unbinding from, peptidoglycan occurs at spatially varying rates $k_{\text{on}}(x)$ and $k_{\text{off}}(x)$. Then further assuming that the timescale of binding and unbinding is much shorter than that of diffusion such that we are in the effective diffusion regime [30, 31] allows us to write $v(x,t) = \frac{k_{\text{on}}(x)}{k_{\text{off}}(x)}$. A single equation for the total concentration of Pal, $c = u + v$ can then be derived:

$$\frac{\partial c}{\partial t} = \frac{\partial^2}{\partial x^2}(D_{\text{eff}}(x)c), \tag{3.3}$$

where

$$D_{\text{eff}}(x) = \frac{k_{\text{off}}(x)}{k_{\text{on}}(x) + k_{\text{off}}(x)} D \tag{3.4}$$

is the effective diffusion coefficient. Note that this can be spatially varying according to the spatially dependent on and off rates. However, it has a simple interpretation; at each spatial location it is the fraction of free molecules multiplied by the free diffusion coefficient.

Equation 3.3 describing fluorescence recovery is the Fokker-Planck diffusion equation rather than the canonical Fickian diffusion equation. We can use this to directly model the inhomogeneity in a fluorescence profile as being due to an inhomogeneous diffusion coefficient. This allows us to move directly to the Fokker-Planck equation (3.3) describing the diffusion of the total concentration, c, of the population throughout the cell.

This conceptual model, termed 'mobilisation and capture', provides a rational for aspects of Pals behaviour in the outer membrane, such as its accumulation at the septum of dividing cells, however it is not clear that it can explain the observed mobility changes relative to non-dividing cells. Moreover, prior investigations into Pal mobility in *tol* mutants show this to be essentially identical to that in non-dividing cells, implying that the Tol-Pal system is effectively inactive in non-dividing cells.

In this chapter, we address these apparent discrepancies by developing a mathematical model of the Tol-Pal system. This model is constructed based on the physical properties of diffusion and spatially-localised dissociation. Importantly, it quantitatively elucidates how TolB facilitates the mobilisation of Pal exclusively during cell division, transporting it to the division septa. This model not only accounts for Pal mobility in both wild type and *tol* mutants but also successfully predicts variations in TolB mobility, as well as the impact of varying TolA levels.

Overall, this model provides a comprehensive explanation for the redistribution of Pal to division septa, a crucial step in coordinating outer membrane invagination during Gram-negative cell division. Furthermore, as the fundamental principles underlying the model are rooted in the physical consequences of localised dissociation, the insights gained may hold relevance for other systems enclosed by double membranes, such as mitochondria and chloroplasts.

3.1 LOCALISED TRANSPORT IS LESS EFFICIENT THAN HOMOGENEOUS TRANSPORT

Previously, SpatialFRAP was used to measure the mobility of Pal across the length of the cell, finding that in dividing cells Pal mobility away from the septum is greater than in non-dividing cells [26]. This mobility was found to be dependent on the Tol system. However, while the spatial localisation of TolQRA complexes to the division septum and their hypothesised deposition of Pal at the location may explain the spatial variation in D_{eff}, it is not clear why overall Pal mobility would be higher in dividing cells (the light blue curve in Figure 3.1c is entirely above the red curve), especially as no evidence has been found that any of the components of the Tol system are relatively more abundant in dividing cells.

However, the overall increase in Pal mobility suggests that there is more TolB in contact with Pal in dividing cells than in non-dividing. This leads us to propose that the cell wall separates the periplasm into inner (inner-membrane proximal) and outer (outer-membrane proximal) compartments and that TolA, as part of the TolQRA complex, pulls TolB from the outer to inner periplasm through holes in the peptidoglycan, powered by the proton motive force. Holes in the peptidoglycan are large enough to accommodate a protein the size of TolB [32]. Once released, TolB diffuses in the inner periplasmic region but can migrate back to the outer later through the same holes in the peptidoglycan.

To demonstrate how this scheme can lead to more TolB in the outer periplasm of dividing cells, and therefore in contact with Pal, we first developed a simple mathematical model of TolB and its transport across the peptidoglycan layer (Figure 3.2a). Let $B_{\text{out}}(x,t)$ and $B_{\text{in}}(x,t)$ denote the concentrations of TolB in the outer and inner periplasm respectively. We assume that TolB is diffuse within these compartments with different diffusion coefficients D_{out} and D_{in} respectively. The underlying reason for the difference in diffusion coefficients (binding to Pal) is not important for the moment. While transport from the inner to the outer compartment occurs spatially uniformly with rate α, transport from the outer to the inner occurs with a potentially spatially varying rate $\beta(x)$ that represents the transport by the TolQRA complexes. In one dimension, we have the following reaction-diffusion equations

$$\frac{\partial B_{\text{out}}}{\partial t} = D_{\text{out}} \frac{\partial^2 B_{\text{out}}}{\partial x^2} + \alpha B_{\text{in}} - \beta(x) B_{\text{out}} \,, \tag{3.5}$$

$$\frac{\partial B_{\text{in}}}{\partial t} = D_{\text{in}} \frac{\partial^2 B_{\text{in}}}{\partial x^2} - \alpha B_{\text{in}} + \beta(x) B_{\text{out}} \tag{3.6}$$

where we use Neumann (reflective) boundary conditions at $x = \pm\frac{L}{2}$. It is easy to see that the above equations are mass conserved i.e. we do not consider production or degradation, such that we can set

$$\frac{1}{L}\int_{-\frac{L}{2}}^{\frac{L}{2}} (B_{\text{out}} + B_{\text{in}}) dx = T. \tag{3.7}$$

We consider two different cases (Figure 3.2a): homogeneous transport where $\beta(x) = \frac{\beta_0}{L}$ is a spatial constant representing non-dividing cells and localised transport at the centre position $x = 0$ where $\beta(x) = \beta_0 \delta(x)$ for dividing cells, where $\delta(x)$ is the Dirac delta function. Note that the total concentration of TolQRA complexes is the same in both cases i.e. $\int \beta(x)dx = \beta_0$. The parameters $\alpha, \beta_0, D_{\text{out}}, D_{\text{in}}$ are taken to be strictly positive.

We implement the following non-dimensionalisation,

$$B_{\text{out}} \to \frac{B_{\text{out}}}{T}, \quad B_{\text{in}} \to \frac{B_{\text{in}}}{T}, \quad x \to \frac{x}{L}, \quad t \to \frac{D_{\text{in}}t}{L^2}$$

to obtain

$$\frac{\partial B_{\text{out}}}{\partial t} = d\frac{\partial^2 B_{\text{out}}}{\partial x^2} + aB_{\text{in}} - ab'(x)B_{\text{out}} \tag{3.8}$$

$$\frac{\partial B_{\text{in}}}{\partial t} = \frac{\partial^2 B_{\text{in}}}{\partial x^2} - aB_{\text{in}} + ab'(x)B_{\text{out}} \tag{3.9}$$

in terms of the function $b'(x)$, specified below, and the dimensionless variables,

$$d = \frac{D_{\text{out}}}{D_{\text{in}}}, \quad a = \frac{L^2\alpha}{D_{\text{in}}}, \quad b = \frac{\beta_0}{\alpha L} . \tag{3.10}$$

We next find the steady state solutions to equations 3.8 and 3.9 for both dividing and non-dividing cells.

3.1.1 *Solving the toy model for non-dividing cells*

For non-dividing cells $b'(x) = b$. This can be solved at steady state relatively easily and applying the reflective boundary conditions and mass conservation rule it can be seen that,

$$B_{\text{in}} = \frac{b}{1+b} , \tag{3.11}$$

and

$$B_{\text{out}} = \frac{1}{1+b} . \tag{3.12}$$

From this it is obvious that the total concentration of B_{in} and B_{out} across the length of the domain is given by

$$\int_{-\frac{1}{2}}^{\frac{1}{2}} B_{\text{in}} dx = \frac{b}{1+b} , \tag{3.13}$$

and

$$\int_{-\frac{1}{2}}^{\frac{1}{2}} B_{\text{out}} dx = \frac{1}{1+b} . \tag{3.14}$$

Observe that when the transport rate $b \gg 1$, almost all the TolB will be in the inner periplasm.

3.1.2 *Solving the toy model for dividing cells*

For dividing cells, $b'(x) = b\delta(x)$, we can find the solution to equation 3.9 at steady state as

$$B_{\text{in}} = b\bar{B}G(x), \tag{3.15}$$

where $G(x)$ is the Green's function defined by

$$G(x) = \frac{\kappa}{2}\frac{\cosh(\kappa x) + \cosh(\kappa(\mid x \mid -1))}{\sinh(\kappa)}, \tag{3.16}$$

where $\kappa = \sqrt{a}$, and $\bar{B} = B_{\text{out}}(0)$. The derivative of $G(x)$ with respect to x is discontinuous at $x = 0$:

$$G_x(x) = \begin{cases} \frac{1}{2}\frac{\sinh(\kappa x) - \sinh(\kappa(-x-1))}{\sinh(\kappa)}, & -\frac{1}{2} \leq x < 0 \\ \frac{1}{2}\frac{\sinh(\kappa x) + \sinh(\kappa(x-1))}{\sinh(\kappa)}, & 0 < x \leq \frac{1}{2} \end{cases} \tag{3.17}$$

Putting the solution for B_{in} in equation 3.8 we are then able to solve for B_{out} away from the sink,

$$\frac{d^2 B_{\text{out}}}{dx^2} = -\frac{a}{d}B_{\text{in}}, \tag{3.18}$$

to find that

$$B_{\text{out}}(x) = -\frac{ab}{d}\bar{B}\left(\frac{1}{\kappa^2}G(x) + Cx + D\right). \tag{3.19}$$

It can easily be seen, using the boundary conditions, that $C = 0$. Evaluating the equation at $x = 0$, we obtain

$$D = -\frac{d}{ab} - \frac{1}{2\kappa}\frac{1 + \cosh(\kappa)}{\sinh(\kappa)}. \tag{3.20}$$

Finally, we use conservation of mass $\int_{-1/2}^{1/2}(B_{\text{in}} + B_{\text{out}})\, dx = 1$ to find

$$\bar{B} = \frac{1}{b + 1 + \frac{ab}{d\kappa^2}\bar{\kappa}} \tag{3.21}$$

where $\bar{\kappa}$ is the monotonically increasing function of κ

$$\bar{\kappa} = \frac{\kappa}{2}\frac{1 + \cosh(\kappa)}{\sinh(\kappa)} - 1. \tag{3.22}$$

The total concentration of B_{in} and B_{out} across the length of the domain are

$$\int_{-\frac{1}{2}}^{\frac{1}{2}} B_{\text{in}}\, dx = b\bar{B} = \frac{b}{b + 1 + \frac{b\bar{\kappa}}{d}}, \tag{3.23}$$

$$\int_{-\frac{1}{2}}^{\frac{1}{2}} B_{\text{out}}\, dx = \frac{1 + \frac{b\bar{\kappa}}{d}}{b + 1 + \frac{b\bar{\kappa}}{d}}. \tag{3.24}$$

We immediately see that for any parameter values there is less TolB in the inner periplasm in dividing compared to non-dividing cells i.e localised transport is less efficient than homogeneous transport. Therefore, there is more TolB in the outer periplasm of dividing cells to mobilise Pal and increase its effective diffusion coefficient. In the limit of $1 + \frac{b\bar{\kappa}}{d} \gg b$ all TolB of dividing cells is located in the outer periplasm. We have seen above that when $b \gg 1$ all TolB of non-dividing cells is in the inner periplasm of non-dividing cells. In this case, we must also have $\frac{\bar{\kappa}}{d}$ sufficiently large to have a substantial amount of TolB in the outer periplasm of dividing cells.

Limits of D_{in} and D_{out}. Here, returning to our dimensionfull equations, we will consider B_{out} as we take the limits $D_{\text{in}}, D_{\text{out}} \to 0$ and $D_{\text{in}}, D_{\text{out}} \to \infty$. The limits for D_{out} are relatively trivial and we find that,

$$\lim_{D_{\text{out}} \to 0} B_{\text{out}} = 1. \tag{3.25}$$

and

$$\lim_{D_{\text{out}} \to \infty} B_{\text{out}} = \frac{1}{1 + \frac{\beta_0}{\alpha L}} = \frac{1}{1+b}, \tag{3.26}$$

where we simplify slightly by returning to the dimensionless variable b. Next we can find these limits for D_{in}, such that

$$\lim_{D_{\text{in}} \to 0} B_{\text{out}} = \frac{1}{1 + \frac{\beta_0}{\alpha L}} = \frac{1}{1+b}. \tag{3.27}$$

The limit of B_{out} as $D_{\text{in}} \to \infty$ is not so trivial, we will first find the limit of $\bar{\kappa}d$. Taylor expanding $\sinh(\kappa)$ and $\cosh(\kappa)$ and considering the dimensionfull variables we find that

$$\lim_{D_{\text{in}} \to \infty} \bar{\kappa}d = \frac{L^2\alpha}{12D_{\text{out}}}, \tag{3.28}$$

substituting this back in we can then see that

$$\lim_{D_{\text{in}} \to \infty} B_{\text{out}} = \frac{1 + \frac{\beta_0 L}{12D_{\text{out}}}}{\frac{\beta_0}{\alpha L} + 1 + \frac{\beta_0 L}{12D_{\text{out}}}} = \frac{1 + \frac{\beta_0 L}{12D_{\text{out}}}}{b + 1 + \frac{\beta_0 L}{12D_{\text{out}}}}. \tag{3.29}$$

From these limits it can be seen that in order for the majority of the TolB to be in the outer periplasm i.e. $B_{\text{out}} > 0.5$, that for any chosen D_{in} it is possible to choose a D_{out} small enough to ensure that this is true. However, for a chosen D_{out} which is too large it is not possible to choose a D_{in} which will rescue this. This is illustrated in Figure 3.2f.

For the homogeneous case, we find, as expected, uniform steady-state profiles of TolB in both compartments (Figure 3.2b, solid lines). For localised transport however, the solution is hyperbolic in shape on either side of the centre position (Figure 3.2b,

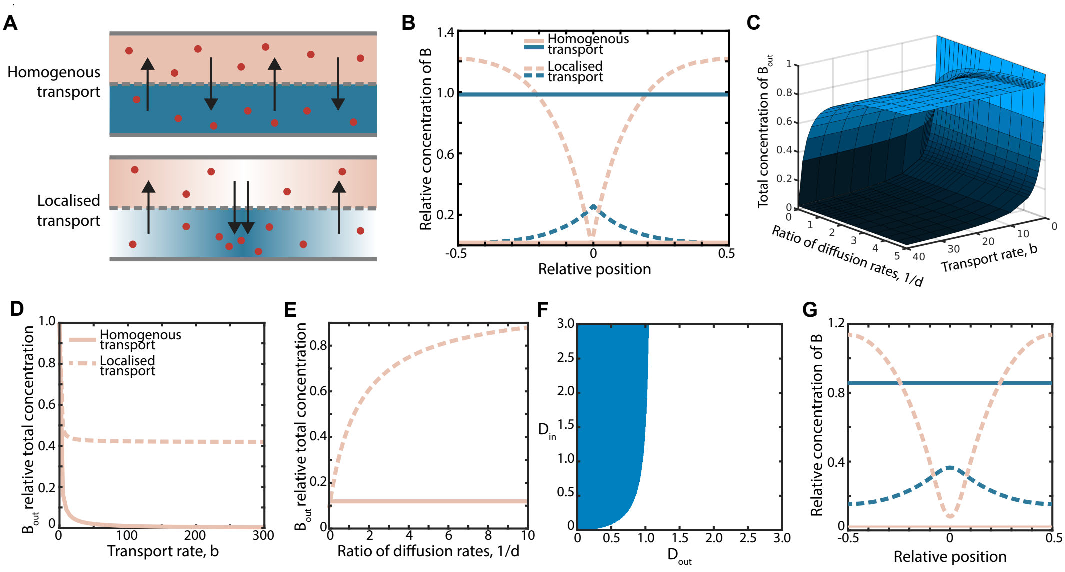


Figure 3.2: Localised transport is less efficient than homogeneous transport. (a) We compare two schemes for protein transport from the outer to inner periplasm: spatially uniformly across the length of the cell (homogeneous transport; top) and only at the centre of the cell (localised transport; bottom). In both cases the transport in the opposite direction, from inner to outer, is spatially uniform. Here, the pink and blue distinguish the outer and inner periplasm respectively. (b) Example line profiles of the concentration in the inner (blue) and outer (pink) periplasm for homogeneous transport (solid lines) and localised (dashed lines) transport. The average transport rate is the same in both cases. Only the localisation of the transport differs. Here, $a = 50$, $b = 100$, and $d = 0.2$. (c) A surface plot showing the total concentration of $B_{\text{out}}(x,t)$, in the case of homogeneous transport (lower surface) and localised transport (upper surface), as the dimensionless parameters $d = D_{\text{out}}/D_{\text{in}}$ and $b = \beta_0/\alpha L$ are varied. Here $a = \alpha L^2/D_{\text{in}} = 10$. (d) Total concentration of B_{out} as the dimensionless transport rate b is varied. For localised transport, the total concentration in the outer periplasm approaches a non-zero value as the transport rate is increased. In this regime, diffusion becomes the limiting factor. Here, $a = 10$, $d = 1$. (e) Total concentration of B_{out} as the ratio of the diffusion rates, d is varied. In terms of the dimensionfull parameters increasing $1/d$ can be achieved by increasing D_{out}. Here, $a = 10$, $b = 10$. (f) D_{in}-D_{out} space, the blue shaded region indicates where more than 50% of the TolB is in the outer periplasm, while at the same time, almost all TolB is in the inner periplasm of non-dividing cells ($\alpha = 10$, $\beta_0 = 40$). Note the upper bound on D_{out}. (g) Example concentration profiles in the outer (pink) and inner (blue) periplasm for localised transport using a transport function having the shape of a truncated normal distribution (dashed lines) compared to homogeneous transport (solid lines). This gives a similar result to that of a point sink: B_{out} is greater for localised transport.

dashed lines) due to the combination of diffusion and localised transport.

Interestingly, we find that the total amount of TolB in the outer periplasm is greater for localised transport than for homogeneous transport (compare pink lines in Figure 3.2b) i.e. TolQRA complexes are overall less efficient at transporting TolB when they are localised than when they are uniformly distributed. This can be understood in terms of the time it takes for a TolB molecule that has just migrated from the inner to outer compartments to be transported back by a TolQRA complex. In the localised case, the molecule may have to diffuse as much as half a cell length, whereas in the homogeneous case it is more likely to find a TolQRA complex in its vicinity. This result holds for any choice of parameters, but notably is greater the slower the diffusion of TolB in the outer compartment (Figure 3.2c, e). Whilst we only show this result analytically for a point sink we note that this result holds for the use of a different function to represent localised transport such as a truncated normal distribution, which we solve numerically (Figure 3.2g).

In a previous study, it was found that preventing TolB from interacting with Pal (via two point mutations) abrogates the increased mobility effect, with dividing cells having the same uniform effective diffusion coefficient as wild-type non-dividing cells (Figure 3.1c) [26]. This suggests that TolB does not interact with Pal in non-dividing cells. In the context of our inner/outer periplasm hypothesis, the implication is that in the homogeneous case most TolB is in the inner periplasm. In the toy model introduced here, this occurs when $b = \frac{\beta_0}{\alpha L} \gg 1$ i.e. when transport of TolB from the outer to the inner compartment (by TolQRA) is sufficiently greater than its transport in the opposite direction. Importantly, having the majority of TolB in the inner compartment in non-dividing cells is not mutually exclusive with having most of the TolB in the outer compartment in dividing cells (in which transport is localised), which requires D_{out} to be sufficiently small (Figure 3.2c, d), i.e. these two criteria can be independently satisfied. We will address biologically relevant parameter values below.

3.2 EXCHANGE BETWEEN DIFFERENTLY DIFFUSING STATES AFFECTS OVERALL MOBILITY

Another challenging result to explain was the finding that in the absence of TolA (Figure 3.1c), or under disruption of its coupling to the PMF (using TolA H22A) [26], Pal mobility was found to be indistinguishable from that in non-dividing cells, just as for the TolB disruption. The naive expectation is that in the absence of TolA, TolB would not be efficiently transported from the outer to the inner periplasm and therefore there should be more TolB available in the outer periplasm to mobilise Pal. However, this negates the effect of exchange on global measures of mobility.

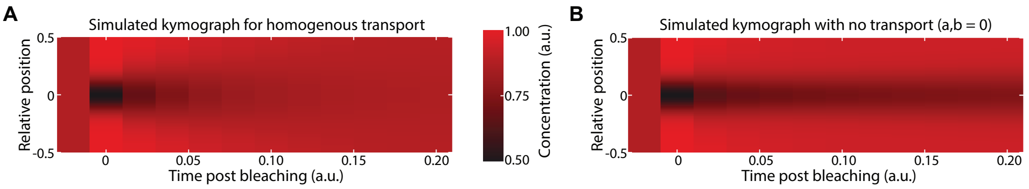


Figure 3.3: Exchange between differently diffusing states affects overall mobility. Kymographs of simulated FRAP experiments with (a) and without (b) homogeneous transport. Both cases have the same steady state with the same concentration of molecules in the inner and outer periplasm. Note that the kymograph with transport returns to the initial steady state faster, showing how the transport-induced reshuffling of the slower sub-population increases the effective rate of diffusion of the entire population. Note that the system with no transport will eventually return to the initial steady state. In both kymographs $d = 0.02$, and with homogeneous transport $a = 50$, and $b = 1$.

Consider again the simple model of TolB transport introduced in Section 3.1. Suppose we were to measure TolB mobility using a population-based technique such as FRAP. Using the model, we could simulate this in both the presence and absence of homogeneous transport, keeping the same steady-state concentration in each compartment across the two cases. Interestingly, we found that the recovery after simulated bleaching is faster in the presence of transport than without (Figure 3.3a, b), despite the fact that the average diffusion coefficient of the entire population is the same in both cases. This is because transport between the two compartments results in a faster effective recovery of the more slowly diffusing population. This is a general result and holds for any system having exchange between differently diffusing states, a very common scenario in biology e.g. DNA or membrane binding proteins. While somewhat obvious in hindsight, we believe that this is an underappreciated observation.

What is the relevance to Pal mobility in the *tolA* deletion strain? While there may be more TolB in the outer periplasm in this mutant, the resulting TolB-Pal complexes are not being continuously disassembled by the action of TolQRA. In wild-type cells the continuous turnover of TolB ensures that the particular subset of Pal molecules mobilised by TolB is continuously changing. In the absence of TolA, the exchange between the immobile (peptidoglycan-bound) and mobile (in Pal-TolB complexes) states of Pal is greatly reduced. As a result, any increase in TolB levels in the outer periplasm in a *tolA* mutant may be ineffective, on the population level, at mobilising Pal, which is ten times more abundant [33]. We will return to this point in Section 3.3 after first developing a more complete model of the Tol-Pal system.

3.3 A MINIMAL MODEL OF THE TOL-PAL SYSTEM

Our minimal model of the Tol-Pal system is built on the fundamental premise, introduced in Section 3.1, that the periplasm is separated into two compartments, inner and outer. Further, we propose that TolA, which, as part of TolQRA inner membrane complexes, extends through holes in the peptidoglycan layer, binds to TolB in the outer periplasm and pulls it into the inner periplasm, disassembling TolB-Pal complexes and releasing Pal in the process (Figure 3.4a). The only difference between dividing and non-dividing cells is in the distribution of TolQRA complexes within the cell: homogeneously distributed in non-dividing cells; localised to the septum in dividing cells (Figure 3.4b). It should be noted that the molecular basis for TolQRA localisation to the septum is not currently understood but recent work has identified a dependence on the FtsWI synthase [34].

Mathematically, the model is represented by a system of reaction-diffusion equations,

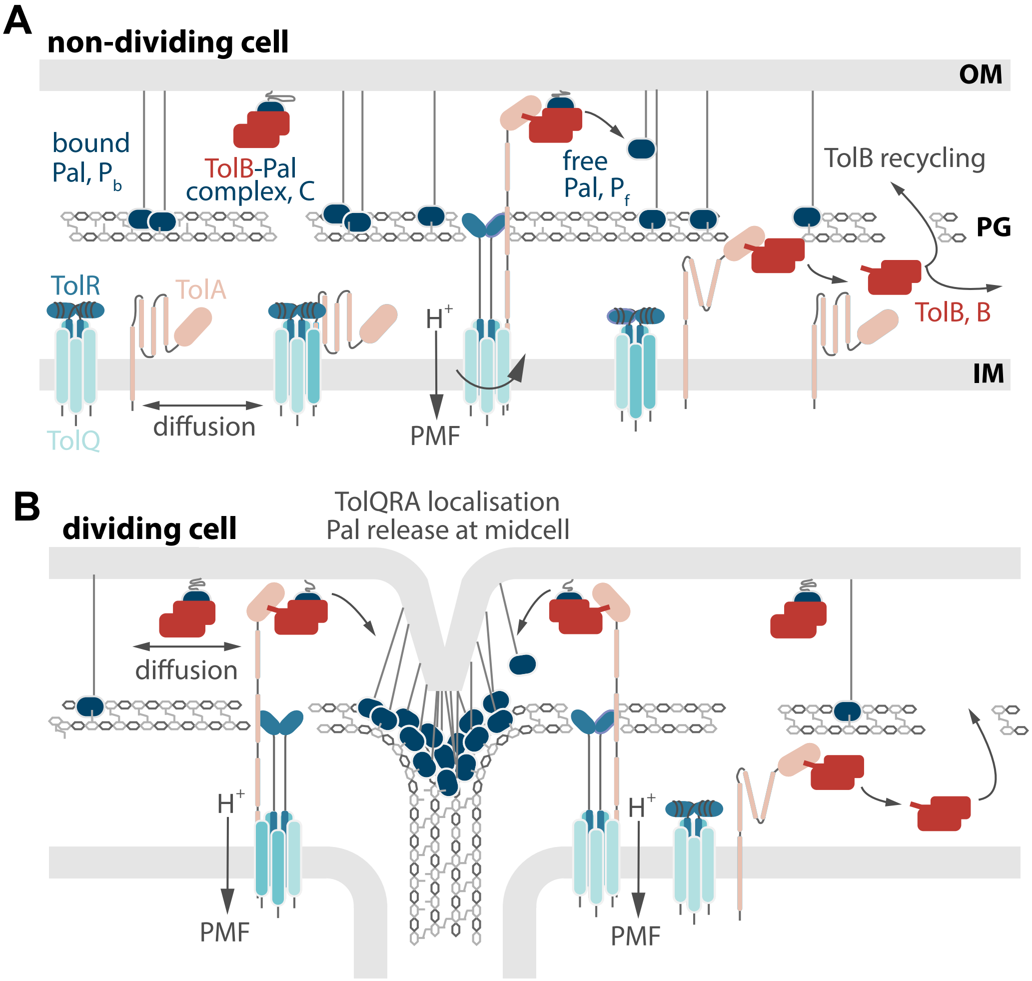


Figure 3.4: A minimal model of the Tol-Pal system. (a) Non-dividing cell: Here TolQRA is homogeneously distributed throughout the inner membrane. As a result, TolB is efficiently captured by the TolQRA machine and moved back into the inner periplasm. Pal is predominately free of TolB and hence binds to the peptidoglycan layer slowing its diffusion. (b) Dividing cell: TolQRA complexes become localised to the cell septum [23, 24] such that TolB transport by TolQRA is localised to the centre of the cell. This leads to less overall transport of TolB, resulting in more TolB in the outer periplasm, where it binds Pal, increasing Pal mobility everywhere except at the septum. In this way TolB can mobilise Pal in dividing cells.

$$\frac{\partial C}{\partial t} = D_c \frac{\partial^2 C}{\partial x^2} + \alpha B P_f - \beta(x) C - \gamma C \, , \tag{3.30}$$

$$\frac{\partial B}{\partial t} = D_b \frac{\partial^2 B}{\partial x^2} - \alpha B P_f + \beta(x) C + \gamma C \, , \tag{3.31}$$

$$\begin{aligned} \frac{\partial P_f}{\partial t} = {} & D_f \frac{\partial^2 P_f}{\partial x^2} - \alpha B P_f + \beta(x) C + \gamma C \\ & - k_{\mathrm{on}} P_f (T - P_B) + k_{\mathrm{off}} P_b \, , \end{aligned} \tag{3.32}$$

$$\frac{\partial P_b}{\partial t} = + k_{\mathrm{on}} P_f (T - P_B) - k_{\mathrm{off}} P_b \, , \tag{3.33}$$

this model contains four variables, three of which represent the concentration of one of the different states of Pal. Firstly, equation 3.30 describes the population of Pal in complex with TolB, $C(x,t)$. These complexes can either be actively dissociated by the TolQRA machine at a potentially spatially-dependent rate $\beta(x)$ or they can dissociate independently with rate γ. In either case the complex is broken down into free TolB, $B(x,t)$, and a transient population of free Pal $P_f(x,t)$. We assume that the transient state of TolB bound to TolA dissociates rapidly, consistent with their low affinity for each other [21, 26]. Therefore, we do not explicitly model this state and instead assume that TolB is immediately returned to the inner periplasm, where it is free to diffuse. TolB can then migrate through holes in the peptidoglycan layer to the outer periplasm where it can bind to free Pal to again form a TolB-Pal complex. We assume this to occur in one step with rate α. This assumption is justified by the relative abundance of Pal molecules (there are $\sim 60,000$ Pal and $\sim 6,000$ TolB molecules in the periplasm [34]). This is described in equation 3.31 and 3.32. The transient free population of Pal is immobilised by binding to the peptidoglycan of the cell wall. This occurs with rate k_{on} but this is limited by the abundance of peptidoglycan binding sites T, which we take to be spatially uniform. This limitation prevents the over-accumulation of Pal at the septum. The resulting bound population, $P_b(x,t)$, can dissociate at a rate k_{off} to return to the transient free population $P_f(x,t)$ (equation 3.33). The three mobile species, TolB, TolB-Pal complex and free Pal diffuse with diffusion coefficients D_b, D_c and D_f respectively. We model the system in one spatial dimension with reflective (Neumann) boundary conditions. This is consistent with our microscopy (FRAP) approach in which the signal along the long axis of the cell is analysed. The difference between dividing and non-dividing cells is encoded in the function $\beta(x)$, which represents the distribution of TolQRA complexes. For non-dividing cells, we set it to be a constant $\frac{\beta_0}{L}$, while in dividing cells we use $\beta(x) = \beta_0 N(0,\sigma^2)$, where $N(0,\sigma^2)$ is the (truncated) normal distribution centered at $x = 0$, the middle of the cell. Using previous experimental results and estimates we are able to obtain reasonable values for many of the model parameters. This leaves three unknown parameters, D_b, D_c and β_0 (see Table 3.1).

To proceed, we looked again at the effects of bleaching Pal in dividing and non-dividing cells, in order to gain a higher time resolution to previous results [26] (Figure 3.5a). These experiments were carried out in collaboration with Oxford University. Subsequently, we analysed this data using the SpatialFRAP methodology to obtain a spatially varying effective diffusion coefficient along the length of the cell (Figure 3.5b). As expected, this gave results similar to those obtained previously. Next, we simulated the experimental FRAP procedure using our mathematical model and fitted the results to the corresponding kymographs and effective diffusion coefficients.

Fitting the model to Pal FRAP data. In order to find the values for the unknown parameters D_b, D_c and β_0 the model is fitted to both the kymographs and the effective diffusion coefficients for Pal in dividing and non-dividing

Parameter	**Brief description**	**Value**	**Source**
D_c	Diffusion constant for the TolB-Pal complex	$0.0068\,\mu\mathrm{m}^2\mathrm{s}^{-1}$	Our fitting.
D_b	Diffusion constant for free TolB	$0.0036\,\mu\mathrm{m}^2\mathrm{s}^{-1}$	Our fitting.
D_f	Diffusion constant for free Pal	Equal to D_c	We assume that the limiting factor for diffusion is the embedding of Pal's lipoylated domain in the outer membrane, similar to how the mobility of membrane proteins is determined by the number of transmembrane domains [35].
α	Rate of binding of TolB and Pal	$0.054\,\mu\mathrm{M}^{-1}\mathrm{s}^{-1}$	[36]
β_0	Rate the TolB-Pal complex is pulled apart by TolQRA	$0.006\,\mathrm{s}^{-1}$	Our fitting.
γ	Rate the TolB-Pal complex dissociates	$0.006\,\mathrm{s}^{-1}$	[36]
k_{on}	Rate Pal binds to the peptidoglycan	$0.1\,\mu\mathrm{M}^{-1}\mathrm{s}^{-1}$	Estimate.
k_{off}	Rate Pal unbinds from the peptidoglycan	$1.0\,\mathrm{s}^{-1}$	Estimate.
T	Concentration of peptidoglycan binding sites	$320\,\mu\mathrm{M}$	Found from the height difference of the valley in the effective diffusion coefficient.
σ	Standard deviation of the truncated normal distribution for the shape of TolQRA in dividing cells	0.08	Found from fitting to the shape of the valley in the effective diffusion coefficient.

Table 3.1: Model parameters.

cells. D_c and D_b are fitted for using the parameters a and b which are related to D_c and D_b by,

$$a = D_c - D_b \quad \text{and} \quad b = \frac{D_c}{D_b} \,. \tag{3.34}$$

Then enforcing the bounds $a > 0$ and $b > 1$ ensures that $D_c > D_b$ which we have observed is necessary to return a peak in TolB as is expected experimentally [24]. During the fitting procedure D_f is chosen as being equal to D_c and all the remaining parameters are fixed as: $\alpha = 0.054\,\mu\mathrm{M}^{-1}\mathrm{s}^{-1}$, $\gamma = 0.006\,\mathrm{s}^{-1}$, $k_{\mathrm{on}} = 0.1\,\mu\mathrm{M}^{-1}\mathrm{s}^{-1}$, $k_{\mathrm{off}} = 1.0\,\mathrm{s}^{-1}$, $T = 320\,\mu\mathrm{M}$, and $\sigma = 0.08$. All simulations are run using reflective (Neumann) boundary conditions and over the domain $x \in [-L/2, L/2]$ where the values of L is taken as the average length of the cells from the respective experimental data. For dividing cells $L = 4.0\,\mu\mathrm{m}$ and for non-dividing cells $L = 2.8\,\mu\mathrm{m}$.

First, the kymographs are simulated computationally. This is performed using the MATLAB solver *pdepe*. For every parameter choice, the model is first run for a sufficiently long time such that the steady state is reached. This gives the first pre-bleach frame of the kymographs. Then a second simulation is run for both a total population and a bleached population such that we have the set of equations,

$$\begin{aligned}
\frac{\partial C_T}{\partial t} &= D_c \frac{\partial^2 C_T}{\partial x^2} + \alpha B_T P_{fT} - \beta(x) C_T - \gamma C_T \,, && (3.35)\\
\frac{\partial B_T}{\partial t} &= D_b \frac{\partial^2 B_T}{\partial x^2} - \alpha B_T P_{fT} + \beta(x) C_T + \gamma C_T \,, && (3.36)\\
\frac{\partial P_{fT}}{\partial t} &= D_f \frac{\partial^2 P_{fT}}{\partial x^2} - \alpha B_T P_{fT} + \beta(x) C_T + \gamma C_T \\
&\quad - k_{\text{on}} P_{fT}(T - P_{bT}) + k_{\text{off}} P_{bT} \,, && (3.37)\\
\frac{\partial P_{bT}}{\partial t} &= +k_{\text{on}} P_{fT}(T - P_{bT}) - k_{\text{off}} P_{bT} \,, && (3.38)
\end{aligned}$$

$$\begin{aligned}
\frac{\partial C_v}{\partial t} &= D_c \frac{\partial^2 C_v}{\partial x^2} + \alpha B_T P_{fv} - \beta(x) C_v - \gamma C_v \,, && (3.39)\\
\frac{\partial P_{fv}}{\partial t} &= D_f \frac{\partial^2 P_{fv}}{\partial x^2} - \alpha B_T P_{fv} + \beta(x) C_v + \gamma C_v \\
&\quad - k_{\text{on}} P_{fv}(T - P_{bv}) + k_{\text{off}} P_{bv} \,, && (3.40)\\
\frac{\partial P_{bv}}{\partial t} &= +k_{\text{on}} P_{fv}(T - P_{bv}) - k_{\text{off}} P_{bv} \,, && (3.41)
\end{aligned}$$

where the subscript T represents the total population and the subscript v corresponds to only the visible population, such that the first four equations are for the total population and the last three for the visible population that remains after bleaching. Note that here there is no equation for visible TolB (B_v) since this is equivalent to the equation for total TolB as only Pal is bleached and measured. The initial conditions for the total population are taken as the steady state obtained previously. We obtain a bleaching function from the experimental data by dividing the pre-bleach and the post-bleach line profiles. This bleaching function is then applied to the simulated steady state to obtain the concentration profile of the visible population (after bleaching). The coupled system is then solved numerically for the same total time as used in the experimental kymograph procedure. This gives the remainder of the simulated kymograph.

The effective diffusion coefficient is then found from the simulated kymograph using the SpatialFRAP method, in exactly the same way as for the experimental data.

The MATLAB function *immse* is used to calculate the mean square error (the cost function) between the simulated and experimental data for both the kymographs and the effective diffusion coefficient. The resultant values are added with weights of 10^5 (kymographs) and 10^7 (effective diffusion coefficients) in order to avoid numerical issues with small numbers and to evenly weight the two procedures. This cost function was minimised across the described parameters using the *patternsearch* function from MATLAB's Global Optimization toolbox.

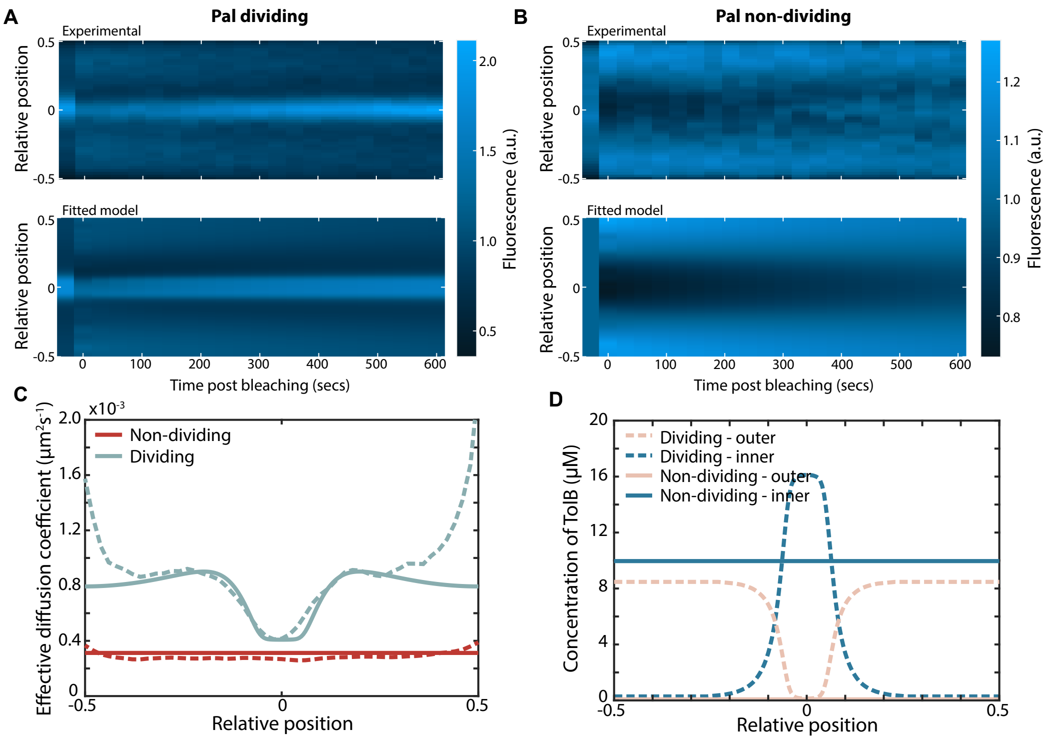


Figure 3.5: Fitting the model to experimental Pal SpatialFRAP. (a) Kymographs of (top) FRAP of Pal-mCherry averaged over 30 dividing cells and (bottom) simulations from the mathematical model showing the results for the best fit to the experimental data. Color scale is for both kymographs and shows either the normalised fluorescence (experimental data) or normalised concentration (simulations). (b) Same as in (a) but for non-dividing cells. (c) Comparison between the effective diffusion coefficient computed using the SpatialFRAP method from the average experimental kymographs in (a) and (b) (dashed lines) and the best fit of the model (solid lines) for Pal in non-dividing and dividing cells. (d) The distribution of TolB in the inner (blue lines) and outer periplasm (pink lines) for dividing (dashed lines) and non-dividing (solid lines) cells for the best fit shown in (c). Note the greater concentration of TolB in the outer periplasm of dividing cells (compare pink lines).

We found that the model could indeed accurately reproduce the experimental FRAP kymographs and the effective diffusion coefficients (Figure 3.5a,b,c). Furthermore, we found that the fitted values resulted in TolB being mostly in the inner periplasm in non-dividing cells and mostly in complex with Pal in the outer periplasm in dividing cells (Figure 3.5d), consistent with the conclusions drawn above using our simple model of TolB transport. These results demonstrate that the transport of TolB across the peptidoglycan layer, powered by the proton motive force across the inner membrane, can, at least in principle, explain the observed mobilisation of Pal by TolB and its re-distribution across the outer membrane. To test the model further, we next assessed its quantitative predictions.

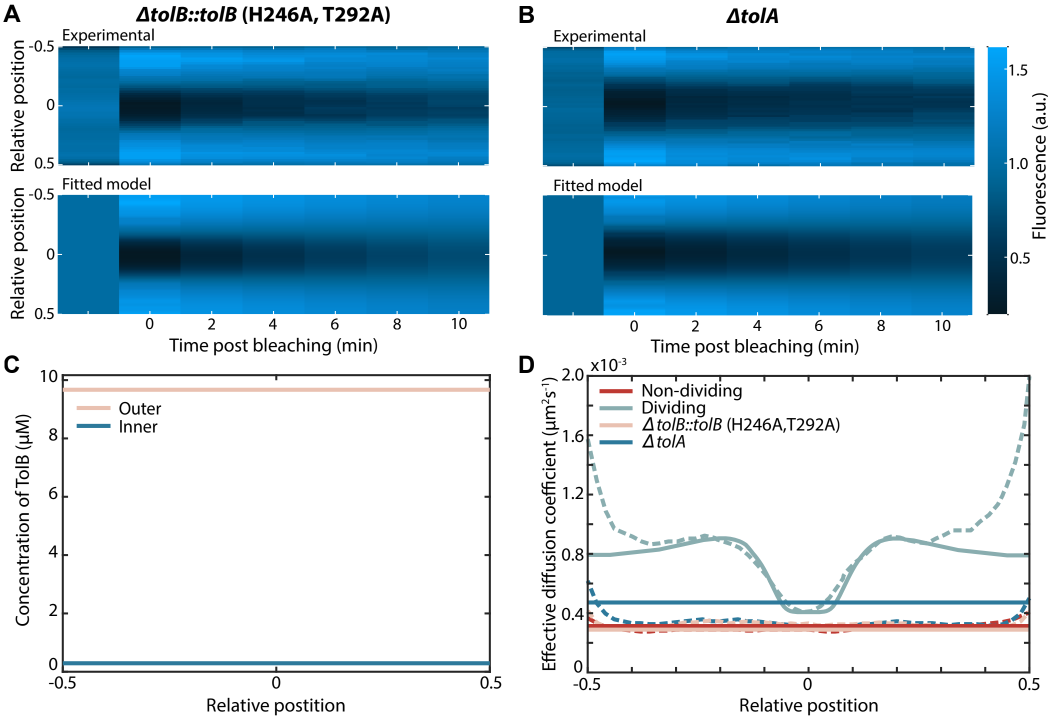


Figure 3.6: The model for the Tol-Pal system can predict the effects of mutants. (a) Kymographs of (top) FRAP of Pal-mCherry averaged over 30 Δ*tolB*::*tolB*(H246A, T292A) mutant cells and (bottom) predicted kymographs from the model using the fitting in figure 3.5. Color scale is for both kymographs and shows either the normalised fluorescence (experimental data) or normalised concentration (simulations). (b) Same as in (a) for Δ*tolA* mutant cells. (c) Predicted model concentration profile for TolB in Δ*tolA* cells. The majority of TolB is in the outer periplasm. Despite this the Δ*tolA* mutant still has a low effective diffusion coefficient (blue line in (d)). This is due to the lack of transport and the resultant recycling of TolB (see figure 3.2). (d) Comparison between the effective diffusion coefficient computed using the SpatialFRAP method from the average experimental kymographs (dashed lines) and the best fit of the model (solid lines) for Pal. Wild-type curves are as in figure 3.5. Note that the model correctly predicts both mutants have Pal mobilities close to wild type non-dividing cells.

We have previously seen that Pal mobility in the absence of either functional TolA or TolB is very similar to that of wild-type non-dividing cells (Figure 3.1c). To test if this is reproduced in the model, we simulated the FRAP procedure in the absence of these proteins but otherwise using the same parameters as above to fit to the Pal kymographs (Figure 3.5). Consistent with the experimental measurements, we found the recovery in both cases to be slow (Figure 3.6a, b) with an effective diffusion coefficient very similar to that in non-dividing wild-type cells. The reasons for these results are clear. In the absence of TolB, Pal can no longer form mobile complexes with TolB and is not deposited at the septum. This results in Pal mobility being similar to that in non-dividing cells since in that cell type TolB is almost entirely in the inner periplasm where it cannot interact with Pal. In Δ*tolA* cells, TolB is present in the outer periplasm (Figure 3.6c) but due to the lack of TolA, the turnover of Pal-TolB complexes is reduced. As we have seen above (Figure

3.3a, b), this lack of exchange results in slower recovery of the Pal signal than would otherwise be expected and we again see recovery more similar to that of non-dividing than of dividing cells.

3.4 PREDICTION OF TOLB MOBILITY

It has previously been shown that TolB exhibits mild accumulation at the septum of dividing cells (Figure 3.7b) [24]. In order to reproduce this result in our simulations we found it was necessary to have $D_c > D_b$, that is the diffusion coefficient of TolB in the outer periplasm (which is in complex with Pal) is higher than that of TolB in the inner periplasm. This is unexpected from a biochemical viewpoint since TolB-Pal complexes are tethered to the outer membrane via the lipoylated domain of Pal and would therefore be expected to diffuse more slowly than free TolB.

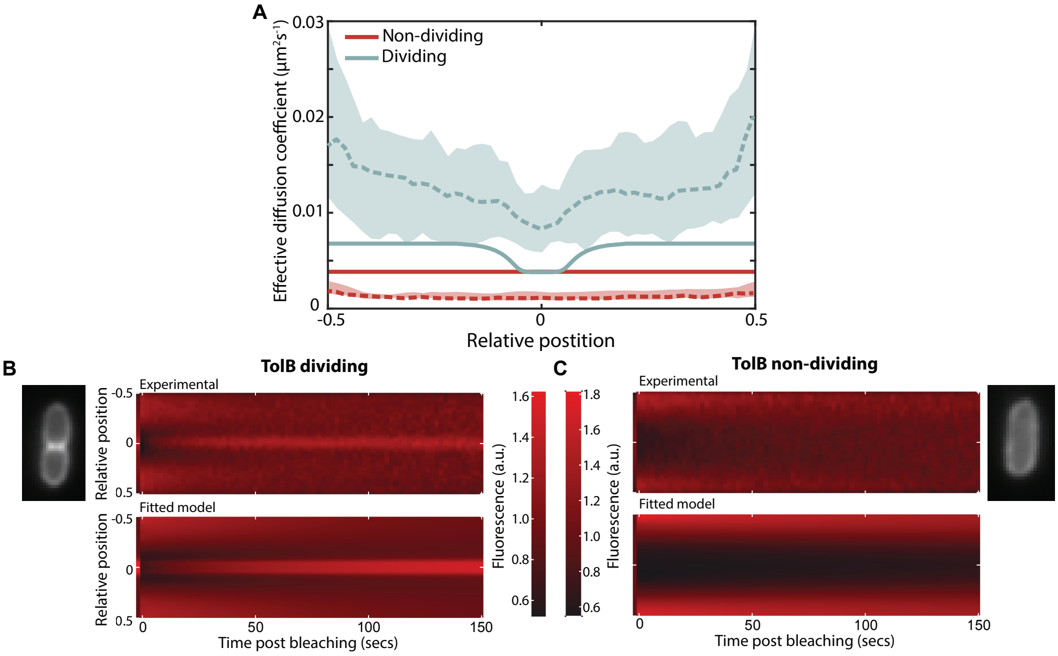


Figure 3.7: The model for the Tol-Pal system can predict the effects of TolB motility. (a) Comparison between the effective diffusion coefficient computed using the SpatialFRAP method for TolB-mCherry from experiments (dashed lines) and simulations (solid lines) in non-dividing and dividing cells. It can be seen that the average dividing effective diffusion coefficient is above that of non-dividing as we see experimentally. (b) Insert shows an example cell showing the localisation of TolB in dividing cells. Kymographs of (top) FRAP of TolB-mCherry averaged over approximately 30 cells and (bottom) predicted kymographs from the model using the fitting in Figure 3.5. (c) Same as in (b) for non-dividing cells.

To test this model prediction, we performed FRAP on cells carrying a TolB-mCherry fusion and used the SpatialFRAP method to access its mobility in both dividing and non-dividing cells. We found that TolB-mCherry in dividing cells displays a significantly higher mobility than in non-dividing cells consistent with the output of the model (Figure 3.7a). Since the model predicts that the majority of TolB is in the inner periplasm in non-dividing cells but in the outer periplasm in dividing cells, the effective diffusion coefficient in each cell type reflects the diffusion coefficient of the free and Pal-bound states respectively. Therefore, the increased D_{eff} of dividing cells is consistent with the model prediction that the mobility of TolB-Pal complexes is greater than that of free TolB. We hypothesise that this may be due to an interaction between TolB and TolA in the inner periplasm following PMF activation. This is a reasonable assumption given that in the absence of the PMF TolB forms a higher affinity complex with TolA than it does when bound to Pal [21].

3.5 SEQUESTRATION OF TOLB ABROGATES PAL MOBILISATION IN DIVIDING CELLS

Our model is based on the mobilisation of Pal due to its binding to TolB. We therefore reasoned that decreasing the amount of TolB available to bind Pal would result in a corresponding decrease of Pal mobility. Such a decrease could be achieved by over-expressing TolA, which is inserted into the inner membrane independently of TolA and TolR; and is recruited to the septum separately from the stator [23]. The resulting excess of TolA in the inner membrane would compete for TolB and sequester it away from Pal. Since TolQ and TolR are not under inducible control, the number of functional TolQRA machines should be largely unaffected.

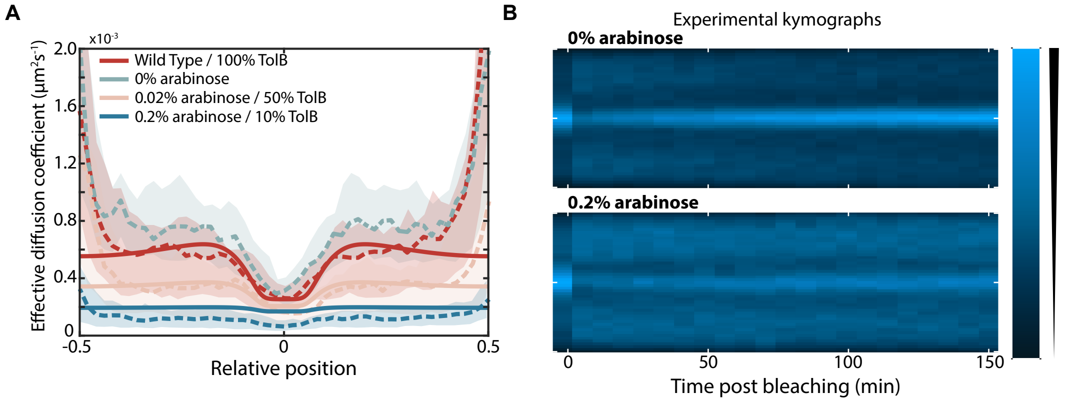


Figure 3.8: Increasing the concentration of TolA lowers the effective diffusion coefficient of Pal. (a) The effective diffusion coefficient of Pal-mCherry in dividing cells for wild type and at different levels of *tolA* expression (dashed lines). As the concentration of arabinose increases the effective diffusion coefficient of Pal decreases. Shown is the median and 95% confidence intervals from approximately 30 cells for each case. And for the model with varying amounts of TolB (solid lines). We hypothesise that excess TolA sequesters TolB, titrating it away from Pal. Decreasing the amount of TolB in the model replicates the result seen experimentally. (b) Averaged kymographs of FRAP on Pal-mCherry (data as in (a)) for 0% (top) and 0.2% (bottom) arabinose induction. While the septal signal recovers fully by 10 min at 0% arabinose, it only partially recovers for 0.2% arabinose.

To test this prediction we used an arabinose inducible promoter to over-express TolA and measured Pal mobility changes using our SpatialFRAP technique. We found that Pal mobility was reduced by TolA induction (Figure 3.8a). At the highest level measured (0.2% arabinose), Pal recovery after bleaching was significantly slower than with no inducer, with the focus of Pal at the septum still not having recovered after 10 minutes (Figure 3.8b), resulting in a much lower D_{eff} (Figure 3.8a). Importantly, our model was able to reproduce this result. We found that reducing TolB levels in the model resulted in a similar reduction of Pal mobility as observed experimentally (Figure 3.8a). Notwithstanding the molecular basis for the inhibitory effect of TolA over-expression, this data reinforces the supposition that modulating levels of TolA, which does not interact with Pal [21], nevertheless influences Pal dynamics via its effect on TolB. Overall, our quantitative model is able to explain how TolQRA, in conjunction with TolB, induces the relocalisation of Pal to the division septum and successfully recapitulates the behaviour of the system under several perturbations.

4

DISCUSSION: A NEW METHOD FOR PROTEIN LOCALISATION

In this part we have shown how the active transport of a protein *across* the bacterial cell wall can be leveraged to power the redistribution of another protein *along* the outer membrane. We focused on Pal, a slowly diffusing bacterial lipoprotein found in many Gram-negative bacteria. The role of Pal is to tether the outer membrane to the peptidoglycan of the cell wall, particularly at the division site, where it accumulates late in the division process due to the action of the Tol system [23]. This facilitates the invagination of the outer membrane during cell division, likely by a combination of tethering and cell wall remodelling [37]. However, the low mobility of Pal means that its recruitment cannot occur through a canonical 'diffusion and capture' mechanism. Furthermore, the periplasm is not energised so the energy for an active mechanism must come from the proton motive force across the inner membrane.

The slow mobility of Pal is attributable to its tethering between the outer membrane (via its lipoylated domain) and the cell wall (via binding to peptidoglycan). However, binding to peptidoglycan is prevented by complex formation with the periplasmic protein TolB. It was previously proposed that Pal exists in two states: an essentially immobile state bound to peptidoglycan and as part of mobile TolB-Pal complexes. How are these mobile complexes used to bring Pal to the division septum? Our model is that the inner membrane TolQRA complexes, which are recruited independently to the septum as part of the divisome, pull TolB through holes in the peptidoglycan layer from the outer to inner periplasm, dissociating it from Pal in the process, and eventually returning through the same holes to repeat the cycle in a manner akin to a conveyor belt. In this way, Pal is mobilised from across the cell and transported to the septum to play its role in invaginating the outer membrane. However, it was not at all clear that this scheme could explain why Pal is mobilised only in dividing cells and why its mobilisation requires functioning TolQRA machines.

Here, we have used mathematical modelling to show that a 'mobilisation and capture' model can indeed explain all the measurements of Pal mobility. In dividing cells, the TolB 'conveyor belt' results in greater Pal mobility away from, rather than at, the septum (Figure 3.5c). This mobility is greater than in non-dividing cells because localised transport is less efficient than homogeneous transport (Figure 3.2). As a result, dividing cells have, in total, more TolB in the outer periplasm than in non-dividing cells and therefore greater Pal mobility. Indeed, the fact that Pal mobility in cells lacking TolB is indistinguishable from that in wild-type non-dividing cells indicates that the majority of TolB in non-dividing cells is in the inner periplasm where it cannot interact with Pal. This is also consistent with active transport across the cell wall in one direction only. In cells lacking TolA, the active transport of TolB is disrupted. However, this does not lead to an increase in Pal mobility because without recycling by the TolQRA machine, TolB-Pal complexes are relatively stable (with an *in vitro* dissociation half-time of 2 min [36]) and therefore, within the time frame of the FRAP experiment, mobilise a smaller fraction of Pal than in wild-type dividing cells.

Our mathematical formulation of the model was able to quantitatively reproduce these results using only the wild-type data for parameter estimation (Figure 3.5, 3.6a,b). It also predicted that TolB mobility would be greater in dividing relative to non-dividing cells, which was confirmed experimentally (Figure 3.7). We also examined the effect of TolA over-expression. We hypothesised that this would titrate TolB away from Pal and therefore reduce Pal mobility in dividing cells. We confirmed this to be the case, finding a clear negative relationship between increasing TolA levels and a lower effective diffusion coefficient of Pal. In the future, the model could be tested further by modulating TolB levels. While our collaborators have found that plasmid-expressed TolB is toxic for cells, this might be surmountable by inducing TolB expression from the chromosome. This would allow us to test the model prediction of a positive relationship between TolB levels and Pal mobility.

Collectively, our data and modelling support a 'mobilisation and capture' mechanism for the mobilisation and redistribution of Pal to the division site by the action of the Tol system. The mechanism is efficient in that it naturally leads to Pal mobilisation only in dividing cells (in which TolQRA becomes localised to the septum). Furthermore, while the capture of TolB-Pal complexes by TolQRA machines is via 'diffusion-and-capture', the regulation of mobilisation makes it a distinct mechanism. More generally, our model demonstrates how an immobile protein can be actively translocated across the outer membrane despite the latter's non-energised nature.

A core component of our model of the Tol-Pal system is the active dissociation of TolB-Pal complexes by the TolQRA machines. This was based on the structural homologies to TonB dependent transporters, molecular dynamics simulations of the dissociation and the fact that TolA coupling to the PMF is required for the redistribution of Pal [26]. Secondly, our model assumes separation of the periplasm into inner and outer compartments (at least from the perspective of TolB). As well as the homologies to TonB dependent transporters, this is supported by our previous result that TolA undergoes a PMF-dependent extension through the periplasm [38]. However, this separation is not a strict requirement of the model. The critical ingredient is simply that the dissociation of TolB-Pal complexes by TolQRA machines is spatially localised to the septum. If the cell wall is relatively porous for TolB or if TolB is not pulled through the cell wall as it is dissociated from Pal, adapting the model would only require the re-interpretation of all the model variables as being in the same compartment. The equations would remain unchanged. As a result, the mechanism that we propose here may be applicable to any organelle with two membranes, such as mitochondria and chloroplasts, in which outer membrane protein redistribution needs to be coordinated with inner membrane invagination.

From a mathematical point of view there is always some uncertainty in models. Therefore, it may be prudent to further test our model by performing a sensitivity analysis and stability analysis. Sensitivity analysis would allow us to determine the parameters that have the greatest influence. This information is important when considering how accurately we know parameters from prior experiments or knowledge of the system. Furthermore, all mathematical models have variables which will be uncertain, and in biological systems there will always be an underlying amount of noise within the system which we choose to assume is effectively averaged out when taking a deterministic approach. Stability analysis is important for determining that the system remains robust under noisy conditions.

Part II

SEGREGATION OF BACTERIAL DNA BY THE PARAB*S* SYSTEM

5

BACKGROUND: THE IMPORTANCE OF DNA SEGREGATION

Segregation of genetic material is a pivotal process in cell division, enabling cells to replicate and transmit their genetic material to subsequent generations. This fundamental process is, therefore, indispensable for the survival of all life forms, from the smallest bacteria to complex eukaryotic organisms, such as plants and animals. Whilst the overarching objective of DNA segregation is shared across different organisms, the strategies employed by eukaryotes and prokaryotes are significantly different.

The mechanism by which eukaryotes segregate their chromosome, mitosis, is relatively well known. In brief, this process hinges on the formation of the spindle, a self-organising microtubule structure that attaches to the centromeres and pulls the replicated chromosomes apart [39]. In contrast, prokaryotic segregation is a continuous process where DNA replication and segregation occur simultaneously. The physical separation of the DNA must be achieved rapidly, accurately, and robustly within the tight confinements of the cell. Various methods have been proposed to play a role in accomplishing this task, including entropic forces [40], loop-extrusion facilitated by structural maintenance of chromosomes (SMC) complexes [41–44], and active partitioning systems [45–48].

In addition to segregating the chromosomal DNA, bacteria often harbour extra-chromosomal circular DNA molecules known as plasmids. Plasmids are typically smaller than the chromosome, but more abundant. Their abundance can vary, ranging from high-copy number plasmids with over 50 copies per cell [49] to low-copy number plasmids with as few as two copies [50–52]. The methods employed for plasmid segregation depend on their abundance. High-copy number plasmids can rely on Brownian diffusion to ensure an equitable distribution to daughter cells. On the other hand, low-copy number plasmids necessitate active partition systems, mirroring the strategies used by the chromosome.

Entropic segregation. The notion of entropic segregation contributing to chromosome segregation initially emerged from the field of polymer physics. When two spatially confined polymers (here representing the chromosomes) interact with excluded volumes, it becomes energetically more favourable for them to separate. This arises due to the greater number of configurations that they can adopt when they do not overlap, leading to an increased entropy in this arrangement [40, 53, 54].

Structural Maintenance of Chromosomes (SMC). The bulk of chromosome condensation is orchestrated by SMC. SMC dimers are recruited near to the origin of replication, they then translocate concurrently along the two DNA arms, facilitating the bringing together of distal loci. This process, referred to as loop extrusion, results in the folding of the DNA into a loop or multiple loops [55]. This mechanism helps to organise the replicated DNA into distinct sister chromosomes [56].

Active partitioning systems. The currently identified active partitioning systems have been classified into three types [52]. All these types share a common structure consisting of three essential components: one or more copies of a partition site (the centromere), a centromere-binding protein, and an NTPase [50, 52].

The most widespread among these systems is the type I system, which employs a Walker-type ATPase to facilitate segregation by forming dynamic patterns on the bacterial nucleoid. Notably, chromosomally encoded partitioning systems exclusively belong to the type I category. For clarity and ease of reference, we will generically use the term ParABS systems to denote type I systems, acknowledging that different bacterial species and plasmids employ varying nomenclature for this system. A more comprehensive discussion of the ParABS system is reserved for Section 5.1; for now, it is pertinent to recognise that this system promotes segregation by moving the DNA up a protein gradient. In contrast, type II and III systems rely on an actin-like ATPase and tubulin-like GTPase respectively. These components polymerise into dynamic filaments, which effectively push plasmids apart [52].

In this part of the thesis, we will focus on understanding the type I ParAB*S* active partitioning systems, which are found both on chromosomes and low-copy number plasmids.

5.1 TYPE I PARTITIONING SYSTEM: PARAB*S* SYSTEM

In the majority of bacterial species, faithful chromosome segregation is directed by the ParAB*S* system [51, 57, 58]. Additionally, this system is also employed to segregate low copy number plasmids such as F-plasmid and pB171. The ParAB*S* system is composed of three main components: the centromeric-like *parS* sequence, and two proteins, ParB a CTPase which forms dimers that bind specifically to the *parS* sequence, and ParA, a Walker-type ATPase, the activity of which is stimulated by ParB [45, 59]. ParB undergoes dimerisation and binds to *parS* sites. It is then able to spread out from these *parS* sites coating the neighbouring DNA, covering a region of a few kilobases. This process leads to the assembly of a nucleoprotein complex, known as the partition complex. Concurrently, ParB dimers can also stimulate the ATPase activity of ParA. ParA, in its dimeric ATP-bound state, binds non-specifically to DNA, effectively coating the nucleoid [60]. Subsequent ATP hydrolysis leads to the dissociation of ParA from the DNA which then undergoes an extended cytosolic conformational transition before rebinding DNA [61]. In chromosomal systems, this process acts to segregate the chromosomes to the opposite cell poles. Whereas, in the case of low-copy number plasmids, it leads to the precise positioning of plasmids at equidistant intervals along the long cell axis [62]. This ensures that each daughter cell inherits an equal share of the total number of plasmids.

Since its discovery in 1982 by Austin et al. [63], significant strides have been made in understanding the ParAB*S* system, shedding light on its functionality and the mechanism by which it operates. Nevertheless, critical questions still remain unanswered. In this part, we aim to tackle two main questions: firstly, we seek to elucidate the mechanism governing the formation and structure of the ParB-DNA partition complex. Subsequently, armed with insights gained from this investigation, we look deeper into the mechanism by which segregation of the partition complex across the nucleoid occurs. To provide a comprehensive foundation to understand these questions we will commence by first presenting an overview of the existing body of research into the ParAB*S* system.

5.1.1 *Partition complex formation*

As briefly described above, ParB spreads across several kilobases of DNA surrounding the *parS* sites [64], ultimately culminating in the formation of the partition complex. This is a nucleoprotein complex, which is clearly visible using fluorescence microscopy. The purpose of this higher-order complex, whether to strengthen the physical link between DNA and ParA, or to provide a specific DNA topology to facilitate DNA segregation is not clear. However, since bacterial strains harbouring *parS* -binding competent but spreading defective mutants of *parB* are either unviable or have an elevated number of anucleate cells, it is clear that a higher-order nucleoprotein complex is a prerequisite for faithful DNA segregation.

Whilst this spreading is essential in order for these systems to function, the degree of spreading varies substantially between systems [65–71]. Originally, spreading was proposed to be due to the formation of a nucleoprotein filament extending out from the *parS* site [65, 66, 72]. However, it was subsequently shown that there are too few ParB proteins to form such large structures [68]. Instead, ParB was found *in vitro* to condense DNA through non-specific DNA binding and the formation of protein bridges [69, 73–77].

These results motivated modelling studies of partition complex formation.

Spreading and bridging model. The spreading and bridging model is an equilibrium model that proposed that the partition complex is formed through a combination of long-range (three-dimensional) bridging interactions and short-range (one-dimensional) nearest-neighbour interactions [78]. In this way, a limited number of ParB dimers could bring together regions of DNA several kilobases long, compacting it in space to form a three-dimensional nucleoprotein complex. While this model nominally accounts for the stiffness of the DNA, the polymer model used only allowed 0° or 90° bond angles.

However, this model was subsequently argued to be incompatible with the binding profile of ParB from F plasmid [69].

Nucleation and caging model. The nucleation and caging model [69, 79] is based on similar concepts to the spreading and bridging model and can be understood as the weak-spreading limit of the spreading and bridging model [80]. Here ParB is not modelled explicitly but rather treated as a fixed spatial distribution centered on the *parS* site. This is proposed to be the result of ParB self-assembly due to protein-protein and protein-DNA interactions which cage ParB spatially into a confined volume. These interactions are also described as liquid-liquid phase separation [81]. This model initially required an unrealistically short persistence length in order to explain the distribution of ParB on the DNA. This was subsequently resolved by including the effects of DNA supercoiling [82].

In 2019, studies uncovered a new cofactor of ParB [71, 83–87]. ParB was found to bind to and hydrolyse cytidine triphosphate (CTP). This CTPase activity is required for correct partition complex formation and dynamics. CTP-bound ParB dimers were shown to load onto the DNA at *parS* sites. They then undergo a conformational change allowing them to encompass the DNA strand and subsequently slide along the DNA strand before eventually dissociating. It was also shown *in vitro* that CTP binding allows ParB bridging to occur at physiological concentrations (much lower than in the absence of CTP [68, 73]) and leads to efficient DNA condensation [88, 89]. These results fundamentally change our understanding of how the partition complex is formed and suggest that previous models need to be reevaluated. In particular, no modelling study has yet provided a unified framework for ParB dimer sliding and bridging, a question we look to answer in Chapter 9.

5.1.2 *Partition complex interaction with ParA and segregation*

The third and final component of this system left to discuss is ParA. ParA is a deviant Walker A ATPase protein [90] that enables a directional movement of ParB-bound DNA. Initial studies saw evidence of ParA polymerisation *in vitro*. This lead to the suggestion that there was an active filament-based system pulling towards the cell poles [91–97]. However, no such filaments were found *in vivo*. Instead, observations of ParA localisation within cells showed it to appear more cloud-like.

The current prevailing mechanism is based on the generation of a dynamic concentration gradient [62, 98–101]. ParA binds to DNA non-specifically in its ATP-dependent dimer state, covering the whole nucleoid [102–104]. These dimers undergo elastic fluctuations due to the nature of the DNA, and can also bind to ParB dimers as part of the partition complex. This leads to the formation of ParA-ParB tethers which exert an elastic force on the partition complex causing it to move in the direction of this force. Concurrently, ParB dimers stimulate the ATPase activity of ParA dimers, triggering ATP hydrolysis, which converts ParA back into its monomeric form in which it is unable to bind DNA [91, 105, 106]. This creates a local gradient of ParA-ATP and the partition complex (followed by the rest of the chromosome/plasmid) diffuses up this gradient following a retracting gradient of ParA-ATP [57, 99, 100]. This functions to segregate the DNA.

Several different models have been developed based on the mechanism described above.

Chemophoresis. The first model developed to explain positioning by a chemical concentration gradient was that of chemophoresis [107]. This model proposed a mechanism by which a gradient generates a chemophoresis force on a macromolecular object. Numerical simulations then showed that this force can provide the energy and information to regularly position the object. This model did not specify a particular molecular mechanism but rather offered a general methodology for the intracellular positioning of macromolecules and organelles.

Ietswaart et al. model. Ietswaart et al. [108] modelled the nucleoid as a simplified 1d system and the concentration of ParA using partial differential equations, where ParA could bind to the nucleoid, and then diffuse along it. Plasmids were initially modelled by applying homogeneous Dirichlet boundary equations at their respective positions, assumed to be due to a high ParA-ATP hydrolysis rate by ParB at the plasmid. Cell edges were modelled by applying homogeneous Neumann boundary conditions. Analysing this model to compute the diffusive flux of ParA made clear the important finding that flux balance was the key to produce regular positioning. Only at the regular positions will the ParA concentration on either side of the plasmid be equal, and thus the plasmid movement would have no directional preference and remain, on average, stationary.

Using this model as a general framework, Ietswaart et al. then built upon this to specify a particular means of plasmid movement up the ParA gradient. They developed a stochastic diffusion/immobilisation model. This model assumed that nucleoid bound ParA could immobilise the freely diffusing partition complexes and that ParA hydrolysis was stimulated at the partition complex. Cytosolic ParA dimers could then rebind the nucleoid at a slow rate. This model was able to produce equal plasmid spacing on simulated growing nucleoids. It should be noted that they also developed a model based on the formation of ParA polymeric filaments, although that will not be discussed here due to the lack of evidence of ParA polymerisation *in vivo*.

DNA relay model. The DNA relay model is a diffusion-reaction mechanism whereby the translocating force is derived from the fluctuations and elasticity of the nucleoid [99, 109]. In this model the partition complex harnesses the elastic force of the DNA by associating with ParA dimers. Specifically this model assumes that ParA binds to the nucleoid in its ATP-bound dimerised state. These DNA bound ParA dimers then undergo elastic fluctuations around an equilibrium position. When in close proximity to the partition complex, ParA and ParB are assumed to instantaneously form a tether, anchoring the partition complex to the nucleoid. The underlying DNA then exerts a force on the plasmid through this tether, pulling the partition complex towards the equilibrium position of the tether. When in a tether ParB stimulates the ATPase activity of ParA causing hydrolysis and the dissassembly of the tether, releasing ParA into the cytosol. These ParA can then rebind to the nucleoid after a waiting period, introduced due to ATP rebinding and ParA undergoing a slow conformational change to a nucleoid binding competent state.

This model results in the partition complex moving in the direction of the highest ParA gradient and generates a ParA depleted zone behind the partition complex. This results in directed motion of a partition complex until it comes into contact with the edge of the nucleoid, at which point the ParA gradient reestablishes on the other side of the partition complex and it moves in the opposite direction. This generates oscillations of a partition complex from pole to pole. By adding more plasmids (and hence more partition complexes) into the model, these can act similarly to the edge of the nucleoid repelling each other such that each plasmid is confined to a narrow region of the cell.

Brownian ratchet model. The initial proposal of the Brownian ratchet model was based on cell-free reconstruction experiments [110]. This model was then updated to specify partition complex positioning *in vivo* [111, 112]. A Brownian ratchet, as imagined in this context, refers to a situation where a particle is freely diffusion along a space divided into intervals. Each of these intervals then act as a 'ratchet' such that the particle can enter from one side but not from the other. Due to this condition of irreversibility the particle will then move in a given direction. In this way direction is provided not by powering the movement but by powering the generation of ratchets. In the case of the ParAB*S* system, this ratchet feature is generated by the formation of a depletion zone behind the partition complex.

Whilst the Brownian ratchet model is very similar to the DNA relay model, there are some notable features which separate the two. The Brownian ratchet model defines the molecular mechanism in more detail. In addition to the DNA relay model, in the Brownian ratchet model, tethers between ParA and ParB in the partition complex do not form instantly but instead at a fast rate. If stretched too far from their equilibrium position tethers break. After ParA ATP hydrolysis, there is an additional transient nucleoid-bound ADP state, nucleoid bound ParA dimers undergo diffusion, and there is a finite number of ParB on the plasmid. This model was shown by the authors to produce either diffusive behaviour or local excursions around a regular position, depending on the parameter choice.

Hopping and relay model. The hopping and relay model builds upon the DNA-relay model, additionally incorporating ParA basal hydrolysis and ParA diffusion on the nucleoid [62]. This model was able to reproduce a wide variety of different behaviours: static, diffusion, oscillations, and importantly regular positioning.

An important factor in this currently proposed mechanism and the main models describing it is the formation of ParA-ParB tethers. These are critical in order to direct the movement of the plasmid. Despite their significance, there has not yet been any experimental evidence for the presence of such tethers. The individual components of these tethers have been shown to interact [88, 113], however a continuous bridge (DNA-ParA-ParB-DNA), that is stable enough to tether the plasmid, has not been.

In contrast ParB-ParB tethers have been shown to exist through magnetic tweezer experiments, where they have been demonstrated to compact the DNA [88, 89]. This shows that ParB bridges can undergo the mechanical stress required to tether the plasmid. Furthermore, a study has shown that DNA-bound ParB dimers are able to load cytosolic ParB dimers onto non-specific DNA in close proximity [114]. It was shown that this process could either take place in cis (where a ParB dimer was loaded onto genomically neighbouring DNA) or in trans (where a ParB dimer was loaded onto spatially neighbouring but genomically distal DNA). Whilst this process was proposed as a potential mechanism to overcome roadblocks on the DNA allowing the ParB to spread over a large region, there is no apparent reason why it would not also take place to allow ParB dimers to be recruited to the nucleoid by the partition complex. These two facts suggest an alternative mechanism to that of ParA-ParB tethers, indicating instead that ParB-ParB tethers may anchor the partition complex to the nucleoid. We will explore this question in Chapter 10.

6

BACKGROUND MATHEMATICS FOR STOCHASTIC SIMULATIONS

The underlying mathematical theory behind stochastic modelling is stochastic processes, the theory of which is based on probability theory. Therefore, before we delve into stochastic modelling we will first begin with a brief overview of some important ideas from probability theory. The definitions provided here are equivalent to those in [115] and [116].

6.1 INTRODUCTION TO PROBABILITY

Here we will briefly outline ideas from probability theory necessary for defining and interpreting stochastic processes.

A random variable, X, can be described as a variable that takes on its value by chance, where this value is determined by the outcome of a random phenomenon. Let $\mathcal{S}$ be the set of values X can assume, then the probability that X takes a value less than or equal to a real number, $x \in \mathcal{S}$, defines the *cummulative distribution function* $F_X(x)$, that is,

$$F_X(x) = \Pr(X \leq x), \ x \in \mathcal{S}. \tag{6.1}$$

If the number of possible values that X can assume is finite, the random variable is discrete and the *probability mass function (pmf)*, $f(x)$, defines the probability that X assumes a specific value

$$f(x) = \Pr(X = x), \quad x \in \mathcal{S}. \tag{6.2}$$

If $\mathcal{S}$ is uncountable, X is a continuous random variable. In this case, if there exists a non-negative function $f_X(x)$ such that

$$\Pr(a < X \leq b) = \int_a^b f_X(x)dx, \quad a, b \in \mathcal{S}, \quad a < b, \tag{6.3}$$

then $f_X(x)$ is called the *probability density function (pdf)* for the random variable X. A well-known example of a continuous random variable that has particular importance for stochastic processes is described by the exponential distribution.

Definition 3. *A non-negative random variable X is said to be **exponentially distributed** with parameter $\lambda > 0$ if the probability density function is*

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0, \\ 0 & x < 0. \end{cases} \tag{6.4}$$

The uniform distribution is the basis for a random number generator, which is used extensively in numerical simulations of stochastic models.

Definition 4. *A random variable X is said to be* ***uniformly distributed*** *if the probability density function is given by,*

$$f_X(x) = \begin{cases} \frac{1}{b-a} & x \leq x \leq b, \\ 0, & \textit{otherwise}, \end{cases} \tag{6.5}$$

where $a < b$ are constants.

The uniform distribution is denoted as $U(a, b)$.

6.2 STOCHASTIC PROCESSES

Stochastic processes are simply a collection of random variables. More formally this can be defined as,

Definition 5. *A* ***stochastic process*** *is a collection of random variables $\mathcal{X} = \{X(t) : t \in T\}$ where T is some index set and $X(t)$ denotes a single random variable defined on a state space $\mathcal{S}_{\mathcal{X}}$. A stochastic process may also be a collection of n random vectors, $\mathcal{X} = \{X_1(t), X_2(T), ..., X_n(t) : t \in T\}$.*

The index set, T, is often used to denote time, here time is treated as a continuous entity, such that $T = \{t : t \in [0, +\infty)\}$. The majority of the stochastic processes described in this thesis have a discrete state space, $\mathcal{S}_{\mathcal{X}}$. Furthermore, all stochastic processes looked at here satisfy the memoryless Markov property. These processes with a discrete state space are referred to as continuous time Markov chain models. Additionally, we will also discuss models where both the index set and the state space are continuous. These are referred to as diffusion processes, and the stochastic realisation $X(t)$ satisfies a stochastic differential equation.

6.2.1 *Continuous time Markov processes*

The Markov property intuitively states that the future of a stochastic process depends only on its current state and not on any previous states.

Definition 6. *A stochastic process $\mathcal{X} = \{X(t) : t \in [0, +\infty)\}$ is called a* ***continuous time Markov process*** *if the following condition holds true, for any sequence of real numbers satisfying $0 \leq t_0 < t_1 < ... < t_n$,*

$$\begin{aligned} Pr(X(t_{n+1}) = i_{n+1} | X(t_0) = i_0, ..., X(t_n) = i_n) \\ = Pr(X(t_{n+1}) = i_{n+1} | X(t_n) = i_n). \end{aligned} \tag{6.6}$$

6.2.2 *Transition probabilities*

Transition probabilities provide a way of relating the state of a stochastic process at different time points and are defined as,

$$p_{i,j}(s,t) = \Pr(X(t) = j | X(s) = i) \quad s < t. \tag{6.7}$$

If the transition probabilities depend only on the length of the interval, $t-s$, rather than s and t explicitly, that is,

$$\begin{aligned} p_{i,j}(s,t) &= \Pr(X(t) = j | X(s) = i) \\ &= \Pr(X(t-s) = j | X(0) = i) = p_{i,j}(0, t-s), \end{aligned} \tag{6.8}$$

then they are referred to as *homogeneous* and are instead denoted by $p_{i,j}(t-s)$. The most natural way to represent homogeneous transition probabilities is in a matrix, $\mathbf{P}(t)$, known as the *probability transition matrix*. The matrix $\mathbf{P}(t)$ is a square matrix of order $|\mathcal{S}_{\mathcal{X}}|$ whose $(i,j)^{\text{th}}$ entry is equal to the transition probability $p_{i,j}(t)$. Since a stochastic process only transitions between states within its state space, it follow that each row of $\mathbf{P}(t)$ sums to one.

6.2.3 *Generator matrix*

Closely related to the transition probabilities are the transition rates, which are often specified when initially defining a stochastic process. Let $q_{i,j}$ denote the transition rate from state $i \in \mathcal{S}_{\mathcal{X}}$ to state $j \in \mathcal{S}_{\mathcal{X}}$. Provided that the transition probabilities $p_{i,j}(t)$ are continuous and differentiable for $t \geq 0$, and satisfy

$$p_{i,j}(0) = 0, \quad i \neq j, \qquad\qquad p_{i,i}(0) = 1, \tag{6.9}$$

the transition rates are defined as

$$q_{i,j} = \begin{cases} \lim_{\Delta t \to 0^+} \frac{p_{i,j}(\Delta t) - p_{i,j}(0)}{\Delta t} = \lim_{\Delta t \to 0^+} \frac{p_{i,j}(\Delta t)}{\Delta t} & \text{for } i \neq j, \\ \lim_{\Delta t \to 0^+} \frac{p_{i,i}(\Delta t) - p_{i,i}(0)}{\Delta t} = \lim_{\Delta t \to 0^+} \frac{p_{i,i}(\Delta t) - 1}{\Delta t} & \text{for } i = j. \end{cases} \tag{6.10}$$

Since each row of the transition probability matrix sums to one, it can be shown that

$$q_{i,i} = -\sum_{j=0, j \neq i} q_{i,j}. \tag{6.11}$$

As with transition probabilities, transition rates can be represented by a matrix $\mathbf{Q} = (q_{i,j})$, referred to as the *generator matrix*. From Equation 6.11 it follows that the diagonal entries of $\mathbf{Q}$ are equal to the negative of the sum of all remaining entries in the corresponding row, and thus the sum of each row of $\mathbf{Q}$ is zero.

6.2.4 *Kolmogorov differential equations*

The forward and backward Kolmogorov differential equations (or master equation) describe the rate of change of the transition probabilities. If $\mathbf{P}(t) = (p_{i,j}(t))$ is the probability transition matrix, then the backward Kolmogorov differential equation is given in matrix form by

$$\frac{d\mathbf{P}(t)}{dt} = \mathbf{QP}(t) \qquad \mathbf{P}(0) = \mathbf{I}, \tag{6.12}$$

where $\mathbf{I}$ is the identity matrix. The forward Kolmogorov differential equations can also be expressed in matrix form as

$$\frac{d\mathbf{P}(t)}{dt} = \mathbf{P}(t)\mathbf{Q} \qquad \mathbf{P}(0) = \mathbf{I}. \tag{6.13}$$

These matrix equations define a system of linear equations, and thus their solution may be written as,

$$\mathbf{P}(t) = \mathbf{P}(0)\exp(\mathbf{Q}t), \tag{6.14}$$

where the matrix exponential is defined as,

$$\exp(\mathbf{Q}t) = \sum_{k=0}^{+\infty} \frac{(\mathbf{Q}t)^k}{k!}\ . \tag{6.15}$$

6.3 STOCHASTIC SIMULATION ALGORITHMS

Consider a multivariate Markov process $\mathcal{X} = \{\mathbf{X}(t) : t \geq 0\}$, the complexity of $\mathcal{X}$ does not have to increase greatly before solving the Kolmogorov differential equations becomes too difficult. For this reason, it is often simpler to construct numerical realisations of the process $\mathcal{X}$, noting that these numerical realisations are not equivalent to solving the Kolmogorov differential equations numerically, as that would give the probability density function of $\mathbf{X}(t)$ instead of a random sample of $\mathbf{X}(t)$ [117]. Here we will describe an important approach for simulating Markov processes, the Gillespie algorithm, which is an exact method that constructs realisations through determining the next reaction to occur at each time step.

6.3.1 *Gillespie algorithm*

Rather than numerically solving the master equation, the *Gillespie algorithm* (also known as the kinetic Monte Carlo or Stochastic Simulation Algorithm) is a method that can be used to simulate intrinsic fluctuations by generating one sample time course from the master equation. By doing many simulations and averaging, the mean and variance can be calculated as a function of time. Even though the master equation itself is never explicitly used, the Gillespie algorithm is fully equivalent to the master equation.

Suppose that the Markov process $\mathcal{X}$, currently in state $\mathbf{x}$, is described by a series of reactions, R_j, $1 \leq j \leq M$, such that $a_j(\mathbf{x})\Delta t$ is the probability that reaction R_j occurs in a short time interval $(t, t + \Delta t)$. If the current state is i then a_j is equivalent to $q_{i,j}$ where $j \neq i$ in the notation of transition rates previously defined. Let J be the set of indices of these reactions, and I be the set of species for which $\mathcal{X}$ describes the dynamics. The vector $\mathbf{v}_j$ is the state change vector and describes how each species changes following the occurrence of reaction R_j. The Gillespie algorithm for obtaining a single realisation of $\mathcal{X}$ is then described in algorithm 1.

Algorithm 1 A Gillespie algorithm to obtain a single realisation of a continuous time Markov process.

```
Initialise the time $t = t_0$ and $\mathbf{x} = \mathbf{x}_0$, a vector describing the initial state of $\mathcal{X}$. Select $t_{\max}$, the maximum time point for which a realisation of $\mathcal{X}$ will be constructed.
while $t < t_{\max}$ do
    1. Compute $a_j(\mathbf{x})$ for $j = 1, ..., M$ and the sum $a_0(\mathbf{x}) = \sum_{j=1}^{M} a_j(\mathbf{x})$.
    2. Draw two random numbers $u_1, u_2 \sim U(0,1)$. The next reaction to occur is the reaction, $R_k$, satisfying $\sum_{j=1}^{k-1} \frac{a_j(\mathbf{x})}{a_0(\mathbf{x})} < u_2 \leq \sum_{j=1}^{k} \frac{a_j(\mathbf{x})}{a_0(\mathbf{x})}$.
    3. Update the state of the process $\mathcal{X}$ by replacing $\mathbf{x} \to \mathbf{x} + \mathbf{v_k}$ and $t \to t + \frac{1}{a_0(\mathbf{x})} \ln\left(\frac{1}{u_1}\right)$
end while
```

6.3.2 *Optimisation of the Gillespie algorithm*

The Gillespie system is exact, however this makes it computationally expensive. For every event several potentially expensive computational steps, such as the simulation of random numbers, the recalculation of propensities, or the identification of the next occurring reaction need to be performed. Since many trajectories are required to understand the dynamics of a system, the performance of the Gillespie algorithm is important.

In his original work, Gillespie suggested two versions of the algorithm: the direct method and the first-reaction method [10]. The direct method was shown in Algorithm 1. This is based on the generation of two random numbers u_1 and u_2 in $U(0,1)$. The time for the next reaction to occur is given by $t + \tau$, where τ is given by,

$$\tau = \frac{1}{a_0(\mathbf{x})} \ln\left(\frac{1}{u_1}\right) , \tag{6.16}$$

where a_0 is the sum of all reaction rates in the current step. The index μ of the reaction that occurs is given by the smallest integer for which

$$\sum_{j=i}^{\mu} \frac{a_j(\mathbf{x})}{a_0(\mathbf{x})} > u_2. \tag{6.17}$$

The first-reaction method generates a τ_k for each reaction R_k according to

$$\tau_k = \frac{1}{a_k(x)} \ln\left(\frac{1}{u_k}\right), \qquad k = 1, ..., M, \tag{6.18}$$

where $u_1, ..., u_M$ are M statistically independent samplings of $U(0,1)$. Then τ and μ are chosen as

$$\tau = \min(\tau_1, ..., \tau_M), \tag{6.19}$$

and

$$\mu = \text{index of } \min(\tau_1, ..., \tau_M). \tag{6.20}$$

In each case the system states are updated by $X(t+\tau) = X(t) + \mathbf{v}_\mu$ where $\mathbf{v}_\mu$ is the state change vector for the occurrence of reaction R_μ.

Evidently, the first-reaction method requires the generation of M separate random numbers, whereas the direct method always requires only two. Therefore for $M \geq 3$, the direct method is probably the method of choice.

The sampling step for the reaction event has been identified as one of the largest bottlenecks within the Gillespie algorithm. This is usually carried out, following the direct method, by putting together a list of all the reactions and calculating the cumulative sum of the probabilities. Then a uniform random number is generated whose maximum value is the sum of all the probabilities. Finally, a search method must locate in each list the reaction whose cumulative probability is smaller or equal to the random number sampled. If there are M different reactions within the model then the standard Gillespie direct method with a linear search uses $O(M)$ computational steps to select each event that takes place [118]. Each event then changes the populations of the model and the reaction rates need to be updated. This requires a further $O(M)$ computational steps. There are different methods to modify this to try and lower the computational cost. For example, a binary decision tree can be used. Here an average of $O(\log(M))$ steps are required to generate each reaction, but if the decision tree is not carefully chosen, a significant number of computational steps will still be required as a consequence of the reaction rates needing to be updated. Other different structures can be used to try and optimise this search depending on the exact nature of the problem. In this thesis we will use a matrix.

A further, trivial optimisation which can be made is to keep track of which reactions are affected by any other reactions. This means that only the necessary reactions are updated. Reactions that are unaffected by another reaction takes place do not need to be updated, and thus computational costs can be saved by not recalculating the rates for these reactions.

7

SLIDING CAN DESCRIBE THE GENOMIC DISTRIBUTION OF PARB

Since the discovery that the ParB family of proteins is a class of molecular switches that require CTP to properly function [83, 84], fundamental new insights have been found about the underlying mechanism behind partition complex formation. Namely, it has been demonstrated that ParB is a CTP-dependent DNA-sliding clamp [71, 83–87]. Here, we ignore any potential effects of ParB bridging and simply consider how ParB spreads along the DNA.

It has been shown that ParB dimers can entrap DNA at *parS* sites before sliding away in either direction. Dissociation is believed to be primarily due to CTP hydrolysis. We develop a physical model based on this proposed sliding mechanism. We test this model against the ParB DNA binding pattern from high resolution ChIP-seq data reported for the ParABS system of two different bacteria. We first study *Myxococcus xanthus*, although the binding profile is relatively noisy and consists largely of a single peak centered on a cluster of all but one of its *parS* sites. Therefore, we also look at the multi-peaked and less noisy profile of *C. crescentus*. We will continue with *C. crescentus* as our model organism when studying the formation of the partition complex in Chapter 9. The data for ParB sliding in *M. xanthus* is presented here to demonstrate the ubiquity of this mechanism, and our model in describing the distribution of ParB found on the DNA.

7.1 PARB SLIDING MODEL

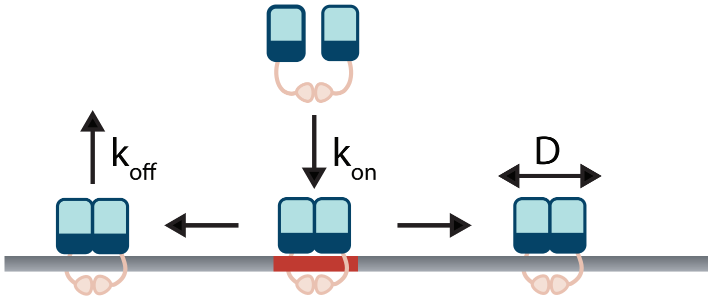


Figure 7.1: ParB sliding model. Diagram representing the sliding model, showing ParB dimers binding at a *parS* site, diffusing along the lattice and unbinding.

We developed a stochastic simulation, using the Gillespie algorithm [10] for the loading, sliding (diffusion), and dissociation of ParB dimers on the DNA (Figure 7.1). The DNA of the centromeric region is modelled as an infinite 1d lattice. Previous experimental studies have shown that upon loading at *parS* sites ParB dimers form a protein clamp that completely encompasses and subsequently slides along the DNA strand [71, 83–87]. Consistent with this, it has been shown that ParB dimers are unable to move past DNA-bound roadblocks [84, 86]. Note that due to the relatively tight entrapment of the DNA strand, we assume in our model that sliding ParB dimers act as obstacles for one another (based on the structures one dimer is not expected to be able pass through the loop of another [71, 84, 85]). Therefore, we take the model to be single occupancy, loading and sliding can only occur if the target lattice site is free. ParB dimers can load at some number, n, of special lattice sites, corresponding to the *parS* sites, with an overall loading rate k_{on},

$$B_c \xrightarrow{k_{\text{on}} \cdot k_i} B_{d,i} \;, \qquad i = s_1, s_2, ...s_n. \tag{7.1}$$

Here B_c is an unbound cytosolic ParB dimer, $B_{d,i}$ is a DNA bound ParB dimer at a *parS* site, and k_i is a relative loading rate for each *parS* site. Dimers can diffuse to an unoccupied neighbouring lattice site with a rate $d = \frac{D}{h^2}$, where D is the effective diffusion coefficient and h is the lattice spacing,

$$B_{d,i} \xrightarrow{d} B_{d,i\pm1}. \tag{7.2}$$

Dissociation occurs randomly at a rate k_{off},

$$B_{d,i} \xrightarrow{k_{\text{off}}} B_c. \tag{7.3}$$

The total number of ParB dimers is fixed at 360. This is motivated by the number estimated for *C. crescentus* [99]. Any unbound dimers are assumed to be in the cytoplasm, which we take to be well-mixed. This is a simplifying assumption and we make no claims regarding the mechanism of ParB targeting to the *parS* sites.

For each parameter set the simulation is first run until steady state is reached and then the distribution of ParB is recorded at regular time intervals, sufficiently separated to be independent samples of the steady-state distribution.

7.1.1 *Specifications for sliding model in M. xanthus*

For our simulations in *M. xanthus* we take each lattice site to be 10 bp, the approximate footprint of a ParB dimer [87, 119].

All but one of the *parS* sites in *M. xanthus* are located close together in a cluster. While the main peak of the ChIP-seq profile is centered on this cluster, there is a small peak coincident with the isolated *parS* site. This is detectable even after smoothening (Figure 7.2). However, when we took each *parS* site to have the same loading rate, this small peak was not accurately captured. We therefore increased the relative strength (three-fold) of this site compared to the others such that we obtain the same relative height of this small peak compared to the main peak.

To compare the results of our simulations to the experimental ChIP-seq profiles, we first binned the latter at a 10 bp resolution to match the simulations. As the signal is still very noisy, we also smoothen it using the *smooth* function of MATLAB with a moving window width of 151. As well as inherent noise and sequence-specific effects, this helps to mitigate the effects of transcription and other road-blocks which are not implemented in this model.

7.1.2 *Specifications for sliding model in C. crescentus*

For our simulations in *M. xanthus* we take each lattice site to be 20 bp, the approximate footprint of a ParB dimer [87, 119]. This larger lattice size is chosen so our sliding simulations correspond to our polymer simulations. These will be discussed later in Chapter 9.

In *C. crescentus* ParB dimers can load onto the DNA at any of the five *parS* sites. The relative loading rate at each *parS* site is determined by $\frac{1}{K_d}$ where K_d is the measured dissociation constant [70]. The loading rate at each site is then determined by multiplying by the overall factor k_{on}. To compare our simulations to the experimental results, we again bin the ChIP-seq data, this time at 20 bp, to match our simulated data. Here, we do not smoothen the data.

7.2 PARB SLIDING MODEL CAN REPRODUCE THE M. XANTHUS CHIP-SEQ PROFILE

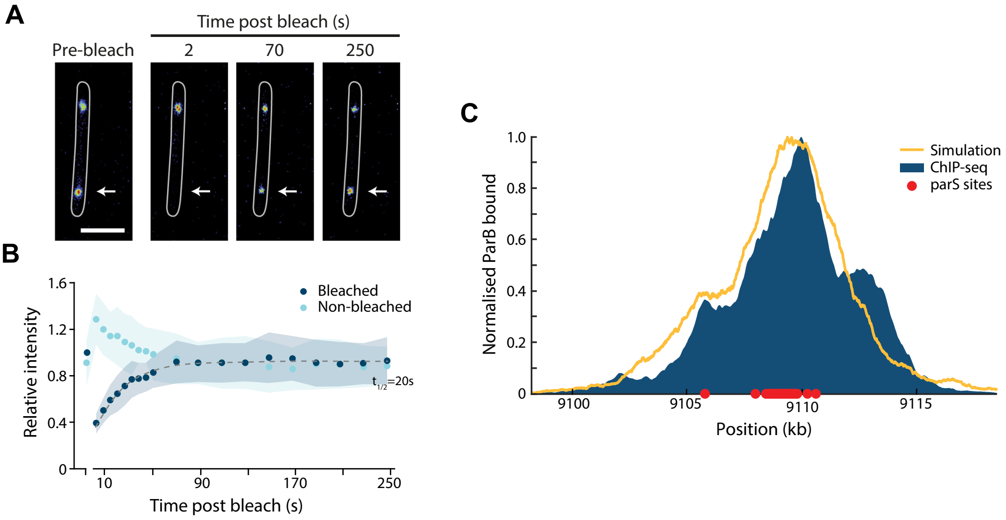


Figure 7.2: ParB sliding can reproduce the *M. xanthus* profile. (a) Representative images of fluorescence recovery after photobleaching (FRAP) experiment. A single mNeonGreen-ParB focus (arrow) was photobleached in a cell containing two foci. Scale bar is 3 µm. Reproduced from [71]. (b) Analysis of FRAP data. The average relative intensity of the bleached (blue) and unbleached (green) partition complex are shown as a function of time (n=38 cells). Dashed lines represent the behaviour from a fitted model, which found a half time of 20 s. Shading represents the SD. Reproduced from [71]. (c) Simulation of ParB sliding compared to ChIP-seq data from [71], both normalised by maximum height.

In order to compare our simulations to the experimental data, we determine the effective diffusion coefficient D, the loading rate k_{on}, and the dissociation rate k_{off}. To estimate the latter we collaborated with the Thanbichler group at the Max

Planck Institute for Terrestrial Microbiology. They performed fluorescence recovery after photobleaching (FRAP) of mNeonGreen-ParB in pre-divisional cells containing two foci (partition complexes) (Figure 7.2a). After bleaching one of the two foci, the fluorescence signal recovered with a half life of 20 s (Figure 7.2b). This provides an estimate for the dissociation rate k_{off}. We determine the loading rate k_{on} such that approximately 300 ParB dimers are bound to the DNA. This is based on previous measurements which have estimated that approximately 80% of ParB dimers in the cells are in ParB foci [99]. To determine the diffusion coefficient we find the best fit of the simulations to the ChIP-seq data. This analysis resulted in a value of $0.018\,\mu\text{m}^2\text{s}^{-1}$. This results in good agreement between the model and the ChIP-seq profile (Figure 7.2c). Furthermore, this predicted 1d diffusion coefficient corresponds to that found via single molecule microscopy measurements [71].

7.3 PARB SLIDING MODEL CAN REPRODUCE THE MULTI-PEAKED CHIP-SEQ PROFILE OF C. CRESCENTUS

Whilst we were able to fit our sliding model relatively successfully to the ParB binding profile of *M. xanthus*, this binding profile is relatively noisy and consists largely of a single peak centred on a cluster of all but one of its 24 *parS* sites. Therefore, the multi-peaked and less noisy profile of *C. crescentus* may serve as a better test of the *in vivo* relevance of the loading and sliding model. While other multi-peaked ParB binding profiles are available [79, 120–122], the binding affinity of each *parS* site of *C. crescentus* have been determined thus reducing the number of free variables within our model, the ParB loading can be described with a single parameter rather than determining the rate at each *parS* site.

As before, we determine the effective diffusion coefficient D and the dissociation rate k_{off}. To estimate the latter, we again collaborate with the Thanbichler group and follow the same procedure as before to find the half time of the recovery of the eGFP-ParB foci using FRAP (Figure 7.3b). This gives a half time of 64 s which we take as an estimate for the dissociation rate k_{off} (Figure 7.3c, d).

FRAP analysis. To calculate the residence time of ParB dimers from the photo-bleaching we perform a simple manipulation of the data. Following the standard calculation used in [81] the FRAP experiments can be described by a simple kinetic model for the ParB proteins in the partition complex and the ParB in the rest of the cytoplasm. Considering $B_1(t)$ and $B_2(t)$ as the average number of ParB proteins in each partition complex after photo-bleaching, B_{tot} the total number of ParB dimers, and k_{in} and k_{out} the rate to enter and exit the partition complexes respectively, the system can be written as,

$$\frac{dB_1(t)}{dt} = k_{\text{in}}B_{\text{tot}} - (k_{\text{in}} + k_{\text{out}})B_1(t) - k_{\text{in}}B_2(t)\ , \tag{7.4}$$

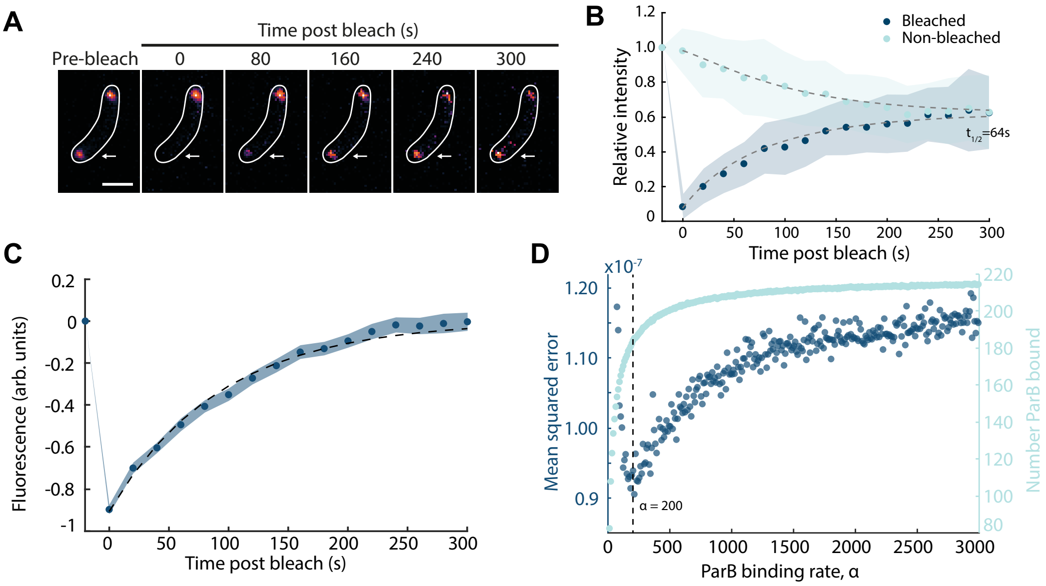


Figure 7.3: ParB sliding can reproduce the multi-peaked *C. crescentus* profile. (a) Representative images of fluorescence recovery after photobleaching (FRAP) experiment. A single eGFP-ParB focus (arrow) was photobleached in a cell containing two foci. Scale bar is 1 µm. (b) Analysis of FRAP data. The average relative intensity of the bleached (blue) and unbleached (green) partition complex are shown as a function of time (n=51 cells). Dashed lines represent the behaviour from a fitted model, which found a half time of 64 s. Shading represents the SD. (c) Difference curve $B_-(t) = B_1(t) - B_2(t)$ (blue) as described in FRAP analysis methods with corresponding fitted curve (grey) finding a half life of 64 s for ParB dimers. (d) Mean squared error for the result of simulations compared to ChIP-seq data and the number of ParB bound as the ParB binding rate is varied. Grey line marks the selected ParB binding rate, $k_{\text{on}} = 200\,\text{s}^{-1}$.

$$\frac{dB_2(t)}{dt} = k_{\text{in}}B_{\text{tot}} - (k_{\text{in}} + k_{\text{out}})B_2(t) - k_{\text{in}}B_1(t)\ . \tag{7.5}$$

In order to fit to the data more easily we consider the sum and difference, $B_\pm = B_1(t) \pm B_2(t)$,

$$\frac{dB_+(t)}{dt} = 2k_{\text{in}}B_{\text{tot}} - (2k_{\text{in}} + k_{\text{out}})B_+(t)\ , \tag{7.6}$$

$$\frac{dB_-(t)}{dt} = -k_{\text{out}}B_-(t)\ . \tag{7.7}$$

The general solution to these equations is given by,

$$S_+ = 2S_\infty - 2\left[S_\infty - \frac{1}{2}S_+(0)\right]e^{-(2k_{\text{in}}+k_{\text{out}})t}\ , \tag{7.8}$$

$$B_-(t) = -B_-(0)e^{-k_{\text{out}}t}\ . \tag{7.9}$$

A simple exponential fit of our data to the difference curve finds $k_{\text{out}} = 0.011\,\text{s}^{-1}$, or a half time in the focus of 64 s and $B_-(0) = 0.91$. Then fitting to the sum, taking B_∞ to be equal to 0.62, we find $B_+(0) = 1.06$ and $k_{\text{in}} = 0.0035\,\text{s}^{-1}$. Using these fitted values we can plot $B_1(t)$ and $B_2(t)$ using the simple transformation $S_1 = \frac{1}{2}(S_+ + S_-)$ and $S_2 = \frac{1}{2}(S_+ - S_-)$.

To determine the diffusion coefficient, we take advantage of the clear ParB profile of *C. crescentus* and fit the outer part of the third peak to an exponential $e^{-x/\lambda}$ with $\lambda = \sqrt{\frac{D}{k_{\text{off}}}}$ (Figure 7.3d), the predicted continuum distribution under this model for an isolated *parS* site. The fitted value of $\lambda = 710\,\text{bp}$, then gives $D = 5600\,\text{bp}^2\text{s}^{-1} = 6.1 \times 10^{-4}\,\mu\text{m}^2\text{s}^{-1}$. This is lower than previously reported diffusion coefficients for proteins diffusing along bacterial DNA [123] including previous measurements of ParB [71, 81], and the value we found for *M. xanthus*, which are on the order of $10^{-2}\,\mu\text{m}^2\text{s}^{-1}$. We will see in Chapter 9 that a larger value is required when we incorporate the effects of bridging into our model.

Diffusion from a point source. We provide an analytical description for the simplified case of a single *parS* site. Consider ParB dimers diffusing on an infinite single-occupancy lattice (lattice spacing h). Dimers can move to any unoccupied neighbouring site at a rate d. Dimers load onto the lattice at a site $i = 0$ with rate k_{on} and unbind with rate k_{off}. We denote the probability of there being a dimer at the n-th site by $p(n,t)$. The chemical master equation which corresponds to this system of reactions is,

$$\begin{aligned}\frac{\partial p(n,t)}{\partial t} &= dp(n-1,t) - 2dp(n,t) + dp(n+1,t) \\ &\quad - k_{\text{off}}p(n,t) + \delta_{n_0}k_{\text{on}}(1-p(0,t)). \end{aligned} \tag{7.10}$$

To obtain the solution, we will Fourier transform the master equation according to the Fourier transform convention,

$$\tilde{p}(k,t) = \sum_{n=-\infty}^{\infty} e^{ikn}p(n,t), \quad \text{and} \quad p(n,t) = \frac{1}{2\pi}\int_{2\pi}\tilde{p}(k,t)e^{ikn}dk \tag{7.11}$$

Using these definitions, we can Fourier transform both sides of the master equation to obtain the ordinary differential equation,

$$\begin{aligned}\frac{d}{dt}\tilde{p}(k,t) &= de^{ik}\tilde{p}(k,t) - 2d\tilde{p}(k,t) + de^{-ik}\tilde{p}(k,t) \\ &\quad - k_{\text{off}}\tilde{p}(k,t) + \sum_{n=-\infty}^{\infty} e^{ikn}\delta_{n_0}k_{\text{on}}(1-p(n,t)) \qquad (7.12) \\ &= \omega(k)\tilde{p}(k,t) + k_{\text{on}}(1-p(0,t)), \qquad (7.13)\end{aligned}$$

where $\omega(k) = d(e^{ik} - 2 + e^{-ik}) - k_{\text{off}}$. Integrating we find,

$$\tilde{p}(k,t) = ce^{\omega(k)t} + e^{\omega(k)t}\int_0^t e^{-\omega(k)t'}k_{\text{on}}(1-p(0,t'))dt'. \tag{7.14}$$

At steady state, as $t \to \infty$,

$$\tilde{p}(k) = \frac{-k_{\text{on}}(1-p(0))}{\omega(k)}. \tag{7.15}$$

Using the Fourier transform convention defined above (Equation 7.11), we find $p(n)$ as,

$$p(n) = \frac{1}{2\pi} \int_{2\pi} \tilde{p}(k) e^{ikn} dk \tag{7.16}$$

$$= \frac{-k_{\text{on}}(1 - p(0))}{2\pi} \int_{2\pi} \frac{e^{ikn}}{\omega(k)} dk. \tag{7.17}$$

Using Euler's formula and since $\int_{2\pi} \sin(x)dx = 0$, we can write

$$d \int_{2\pi} \frac{e^{ikn}}{\omega(k)} = \int_{2\pi} \frac{cos(kn)}{2(\cos(k) - 1) - a} dk, \tag{7.18}$$

where $a = \frac{k_{\text{off}}}{d}$. Rearranging, we can then solve this integral, making use of integral (3.613.1) from Gradshteyn and Ryzhik (GR) [124],

$$d \int_{2\pi} \frac{e^{ikn}}{\omega(k)} = 2 \int_0^{\pi} \frac{\cos(kn)}{1 + \left(\frac{-2}{2+a}\right) \cos(k)} \frac{-1}{2+a} \tag{7.19}$$

$$= \frac{-2}{2+a} \frac{\pi}{\sqrt{1 - \left(\frac{-2}{2+a}\right)^2}} \left(\frac{\sqrt{1 - \left(\frac{-2}{2+a}\right)^2}}{\frac{-2}{2+a}} \right)^n \tag{7.20}$$

$$= \frac{-2\pi}{\sqrt{a^2 + 4a}} \left(\frac{2 + a - \sqrt{a^2 + 4a}}{2} \right)^n. \tag{7.21}$$

Setting $n = 0$ in the integral we have solved above allows us to use Equation 7.17 to find $p(0)$ such that,

$$p(0) = \frac{-k_{\text{on}}(1 - p(0))}{2\pi} \frac{-2\pi}{\sqrt{a^2 + 4a}} \tag{7.22}$$

$$= \frac{k_{\text{on}} - k_{\text{on}} p(0)}{d\sqrt{a^2 + 4a}}. \tag{7.23}$$

Rearranging then gives us,

$$p(0) = \frac{k_{\text{on}}}{k_{\text{on}} + \sqrt{k_{\text{off}}^2 + 4k_{\text{off}} d}}. \tag{7.24}$$

With $p(0)$ and the solution to the integral found above (Equation 7.21) we can now solve Equation 7.17 for $p(n)$,

$$
\begin{aligned}
p(n) &= \frac{-k_{\text{on}}}{2\pi} \frac{\sqrt{k_{\text{off}}^2 + 4k_{\text{off}}d}}{k_{\text{on}} + \sqrt{k_{\text{off}}^2 + 4k_{\text{off}}d}} \frac{1}{d} \int_{2\pi} \frac{\cos(kn)}{2(\cos(k)-1) - \frac{k_{\text{off}}}{d}} dk \qquad (7.25)\\
&= \frac{k_{\text{on}}}{k_{\text{on}} + \sqrt{k_{\text{off}}^2 + 4k_{\text{off}}d}} \left(\frac{2 + \frac{k_{\text{off}}}{d} - \sqrt{\left(\frac{k_{\text{off}}}{d}\right)^2 + 4\frac{k_{\text{off}}}{d}}}{2} \right)^n \qquad (7.26)\\
&= \alpha\beta^n. \qquad (7.27)\\
& \qquad (7.28)
\end{aligned}
$$

Therefore,

$$
p(n) = \frac{k_{\text{on}}}{k_{\text{on}} + \sqrt{k_{\text{off}}^2 + 4k_{\text{off}}d}} e^{n\ln(\beta)}. \tag{7.29}
$$

We rewrite β as,

$$
\begin{aligned}
\beta &= \frac{2 + \frac{k_{\text{off}}}{d} - 2\sqrt{\frac{k_{\text{off}}}{d}}\sqrt{\frac{1}{4}\frac{k_{\text{off}}}{d} + 1}}{2} \qquad (7.30)\\
&= 1 - \sqrt{\frac{k_{\text{off}}}{d}} + O\left(\frac{k_{\text{off}}}{d}\right). \qquad (7.31)
\end{aligned}
$$

Such that,

$$
\ln(\beta) \approx -\sqrt{\frac{k_{\text{off}}}{d}} \qquad \text{if} \quad \frac{k_{\text{off}}}{d} << 1. \tag{7.32}
$$

This allows us to write $p(n)$ as,

$$
\begin{aligned}
p(n) &\approx \frac{\frac{k_{\text{on}}}{k_{\text{off}}}}{\frac{k_{\text{on}}}{k_{\text{off}}} + \sqrt{1 + 4\frac{d}{k_{\text{off}}}}} e^{-n/\sqrt{\frac{d}{k_{\text{off}}}}} \qquad (7.33)\\
&\approx \frac{\frac{k_{\text{on}}}{k_{\text{off}}}}{\frac{k_{\text{on}}}{k_{\text{off}}} + 2\sqrt{\frac{d}{k_{\text{off}}}}} e^{-n/\sqrt{\frac{d}{k_{\text{off}}}}}. \qquad (7.34)
\end{aligned}
$$

In the continuum limit ($h \to 0, d \to \infty, k_{\text{on}} \to \infty$, keeping $D = dh^2$ and $\bar{k}_{\text{on}} = k_{\text{on}}h$), the shape of the profile will be defined by $e^{-x/\lambda}$ where $\lambda = \sqrt{\frac{D}{k_{\text{off}}}}$ is the associated diffusive length scale.

One model parameter remains to be determined - the overall loading rate of ParB k_{on}. Previous measurements have estimated that approximately 80% (290) of ParB dimers in the cell are in ParB foci [99]. In contrast, we find that even at high loading rates less than 220 ParB dimers are associated with the DNA (Figure 7.3e). Increasing k_{on} further does not substantially increase the number of ParB bound as the *parS* sites are almost continuously occupied. The disparity in the number of DNA-associated dimers may be due to several factors. Firstly, the maximum possible number of associated dimers in our simulations is dependent on the chosen discretisation since each lattice site/monomer can be bound by a single ParB dimer. Thus if the footprint of ParB is smaller than our discretisation size of 20 bp, we would be underestimating the achievable ParB occupancy. Secondly the *in vivo* estimate of the cellular ParB concentration is based on quantitative Western blotting, which has a substantial margin of error [125]. ParB foci may also contain a cytosolic or non-specifically bound population that is not accounted for in our model.

Given the above, we choose the loading rate for our model by finding the best fit of the simulations to the ChIP-seq data (Figure 7.3e), obtaining $k_{\text{on}} = 200 \cdot k_{\text{off}}$. This results in remarkably good agreement between the model and the ChIP-Seq profile (Figure 7.4a), indicating that loading and diffusive sliding of ParB dimers can indeed explain the observed binding profile. It also suggests that dimers are largely unaffected by transcription and other processes that could hinder ParB spreading since we have not accounted for these effects in our model. This may not be the case for other systems such as F plasmid that show changes in the binding profile coincident with promoters [79, 126]. Indeed, *in vitro* experiments have shown that high-affinity DNA-binding proteins, such as EcoRI (with the catalytically-inactive E11Q mutation) and TetR, can block the sliding of ParB dimers along the DNA [84, 86, 89].

7.4 RESIDENCE TIME AND PERCENTAGE OCCUPANCY OF ROADBLOCKS IMPACTS THEIR EFFECT ON PARB SLIDING

To better understand how roadblocks can impact the spreading of ParB dimers, we used our sliding simulation to examine the effect on spreading from a single *parS* site. Representative of the biological situation, we do not consider a permanent roadblock but rather a dynamic one, which we specify in terms of its average lifetime and occupancy i.e. the overall fraction of time that the roadblock is bound to the DNA. These two measures are independent of each other, for instance a roadblock that is present and absent for 1 s at a time has the same 50% occupancy as a roadblock present and absent for 10 s at a time. We found that at a lifetime of 1 s, the roadblock had a surprisingly mild effect on spreading, only becoming noticeable from an occupancy of about 75%. Even at 95% occupancy, roughly half the number of dimers slide past the site of the roadblock as in its absence (Figure 7.4c). Similarly, at 75% occupancy, a negative effect on spreading was only observed for roadblock lifetimes greater than about 1 s (Figure 7.4d).

We can understand these results as follows. When the roadblock is present for a time much shorter than the time interval between dimer crossing attempts then a backlog of dimers does not develop. Even for longer times, the backlog of dimers can be cleared if there is enough time between roadblock events i.e. if the average roadblock occupancy is sufficiently low (Figure 7.4b). These results may explain

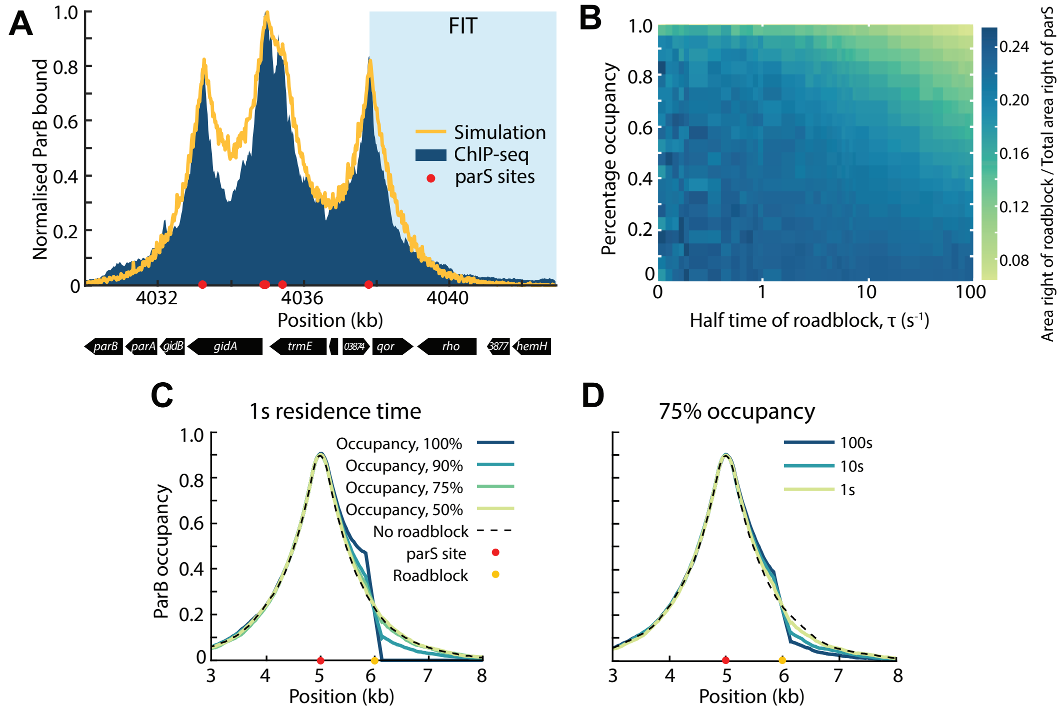


Figure 7.4: Residence time and percentage occupancy of roadblocks impacts their effect on ParB sliding. (a) Simulation of ParB sliding compared to ChIP-seq data from [70], both normalised by maximum height. Shaded area indicates the part of the ChIP-seq profile that was fitted to an exponential to find the effective diffusion coefficient. (b) Phase diagram displaying the difference between simulations with and without a roadblock. The colour indicates the ratio of the area of the distribution to the right of the roadblock to the total area on the right hand side of the *parS* site. The locations of the roadblock and the *parS* site are shown in (c) or (d). (c) Simulations of ParB sliding at 1 s residence time for different roadblock occupancies. The roadblock is indicated by the yellow dot. (d) Same as in (c) but for 75% occupancy with a varying roadblock lifetime. In both (c) and (d) the dotted line shows the profile when there is no roadblock.

why we observe no significant deviation of the ParB binding profile from that expected from our simple loading and sliding model - the *in vivo* occupancy and residence times of proteins binding to the centromeric region of *C. crescentus* and *M. xanthus* may simply not be large enough to substantially affect ParB spreading. We will apply these insights when combining these sliding simulations with polymer simulations of the centromeric DNA in Chapter 9.

8

POLYMER THEORY

DNA molecules are polymers and therefore we can make use of the large well of polymer theory developed by physicists to form models, and understand their properties. One of the deep insights developed by Pierre-Gilles de Gennes is that polymers typically show universal behaviour independent of their underlying chemical composition [127]. This result is far from trivial but we will not discuss it in detail here. It allows us to use the same theory to describe polymers ranging from polyester and nylon to collagen and DNA. Before discussing our polymer simulation framework we will briefly outline a few simple polymer models and some well-known results from polymer theory mainly following [128] and [129].

8.1 IDEAL AND REAL CHAINS

8.1.1 *Ideal chains*

We consider a flexible polymer of $n+1$ backbone atoms $a_i, 0 \leq i \leq n$. The bond vector $\mathbf{r}_i$ connects atom a_{i-1} to atom a_i, giving N total bond vectors. The polymer is then an *ideal chain* if there are no interactions between sufficiently distant atoms i.e. overlap of the chain is allowed. A simple result that one can measure is the *end-to-end vector*, defined as

$$\mathbf{R} = \mathbf{a}_N - \mathbf{a}_0 = \sum_{i=1}^{N} \mathbf{r}_i, \tag{8.1}$$

the sum of all the bond vectors in the chain. It is often useful to discuss the average properties of polymers. For an ideal chain, it can be shown that the average end-to-end vector is equal to zero,

$$\langle \mathbf{R} \rangle = \mathbf{0}, \tag{8.2}$$

where the angled brackets denote the ensemble average across different conformations of the polymer. This result arises since there is no preferred direction of the bond orientation in this model. A more interesting often measured quantity to characterise the size of the polymer chain is the *mean-square end-to-end distance*,

$$\begin{aligned} \langle \mathbf{R}^2 \rangle = \langle \mathbf{R} \cdot \mathbf{R} \rangle &= \left\langle \left(\sum_{i=1}^{N} \mathbf{r}_i \right) \cdot \left(\sum_{j=1}^{N} \mathbf{r}_j \right) \right\rangle \\ &= \sum_{i=1}^{N} \sum_{j=1}^{N} \langle \mathbf{r}_i \cdot \mathbf{r}_j \rangle. \end{aligned} \tag{8.3}$$

If we consider bonds of equal length $l = |\mathbf{r}_i|$, the scalar product can be represented in terms of the angle θ_{ij} between bond vectors $\mathbf{r}_i$ and $\mathbf{r}_j$,

$$\mathbf{r}_i \cdot \mathbf{r}_j = l^2 \cos\theta_{ij}. \tag{8.4}$$

The mean-square end-to-end distance then becomes,

$$\langle \mathbf{R}^2 \rangle = l^2 \sum_{i=1}^{N}\sum_{j=1}^{N} \langle \cos(\theta_{ij}) \rangle. \tag{8.5}$$

One of the simplest models of an ideal polymer is the *freely jointed chain* where there is no correlation between the directions of different bond vectors, $\langle \cos(\theta_{ij}) \rangle = 0$ for $i \neq j$ and $\langle \cos(\theta_{ij}) \rangle = 1$ for $i = j$. The mean-squre end-to-end distance for a freely jointed chain can then simply be given as,

$$\langle \mathbf{R}^2 \rangle = Nl^2 \tag{8.6}$$

All ideal chains can be approximated by an *equivalent freely jointed chain*, as long as the chain is long compared to the scale of short range interactions. We choose an effective bond length b sufficiently long so that neighbouring segments are non-correlated. Here b is known as the *Kuhn length*. This can be related to the stiffness of a chain. The Kuhn length can be defined by,

$$b = \sqrt{\frac{\langle \mathbf{R}^2 \rangle}{N}}. \tag{8.7}$$

An additional measure of polymer size is the *square radius of gyration*. This is defined as the average squared distance between monomers, $\mathbf{r}_i$, in a conformation and the centre of mass, $\mathbf{r}_c$,

$$R_g^2 = \frac{1}{N} \sum_{i=1}^{N} (\mathbf{r}_i - \mathbf{r}_c)^2, \tag{8.8}$$

where the centre of mass of the polymer is the number-average of all the monomer vectors,

$$\mathbf{r}_c = \frac{1}{N} \sum_{i=1}^{N} \mathbf{r}_i. \tag{8.9}$$

For a linear ideal chain, the radius of gyration can be related to the end-to-end distance as,

$$\langle R_g^2 \rangle = \frac{Nl^2}{6} = \frac{\langle R^2 \rangle}{6}. \tag{8.10}$$

The probability distribution of the end-to-end vector of an ideal chain is well described by the Gaussian function,

$$P_{3d}(N, \mathbf{R}) = \left(\frac{3}{2\pi N b^2} \right)^{3/2} \exp\left(-\frac{3\mathbf{R}^2}{2Nb^2} \right). \tag{8.11}$$

8.1.2 *Gaussian chain*

Due to the fixed-bond constraint, the freely jointed chain model is of little practical interest. A more advantageous alternative is the *Gaussian chain* model. In the Gaussian chain model, a polymer chain is represented as a linear array of beads connected by harmonic springs. Here each spring represents a Kuhn segment where each segment is Gaussian distributed,

$$P(\mathbf{r}) = \left(\frac{3}{2\pi b^2}\right)^{3/2} \exp\left(-\frac{3\mathbf{r}^2}{2b^2}\right), \tag{8.12}$$

where $\mathbf{r}$ is the bond vector. Since the bond orientation is uncorrelated from the orientation of the other bonds, the average squared end-to-end distance and the radius of gyration of a Gaussian chain are given by the same expressions as for a freely jointed chain.

We will return to the Gaussian chain later as this is used as a starting point for the Rouse model.

8.1.3 *Real chains*

In a real polymer we cannot ignore the interactions between different monomers along the chain or that we have excluded volume i.e. the polymer cannot overlap with itself. Real polymer chains have been quite successfully described by *Flory theory*. We will first consider the case of a polymer chain in a good solvent. In this case, the polymer chain experiences an effective repulsion between monomers that tend to swell the polymer chain, which is balanced by the entropy loss caused by such a deformation. Flory theory makes the assumption that we can consider a real polymer chain to be composed of two components: an ideal chain (entropic contribution) and a dilute 'gas' of monomers (energetic contribution), this avoids the need to count every possible chain conformation and pairwise energetic contribution.

The Flory approximation of the entropic contribution to the free energy of a real chain is the energy required to stretch an ideal chain to end-to-end distance R,

$$F_{\text{ent}} \approx k_B T \frac{R^2}{Nb^2}, \tag{8.13}$$

where k_B is Boltzmann's constant, and T is the temperature. To approximate the energetic contribution to the free energy of a real chain, we consider the sum of the energies associated with all pairwise monomer-monomer interactions in the chain. Each pairwise interaction carries an energetic cost of being excluded from this volume, which is $k_B T$ per exclusion. The number of interactions in the mean-field estimate is the product of N, the excluded volume V, and the number density of monomers in the pervaded volume N/r^3. Therefore, for all N monomers in the chain the energetic contribution to the free energy of the chain is,

$$F_{\text{int}} \approx k_B T v \frac{N^2}{R^3}. \tag{8.14}$$

Thus, the total free energy for the real polymer chain given by Flory theory is the sum of the entropic and energetic contributions,

$$F = F_{\text{ent}} + F_{\text{int}} \approx k_B T \left(\frac{R^2}{Nb^2} + v\frac{N^2}{R^3} \right). \tag{8.15}$$

The optimum end-to-end distance in Flory theory, R_F, can be determined from the minimum free energy of the chain with respect to R,

$$\frac{\partial F}{\partial R} = 0 = k_B T \left(-3v\frac{N^2}{R^4} + 2\frac{R}{Nb^2} \right), \tag{8.16}$$

$$R_F \approx v^{\frac{1}{5}} b^{\frac{2}{5}} N^{\frac{3}{5}}. \tag{8.17}$$

This states that the end-to-end distance of a real polymer chain in a good solvent scales as,

$$R \propto N^{3/5}. \tag{8.18}$$

The scaling exponent of $3/5$ from this model is remarkably close to that found from more complicated theory and experimentation giving a scaling exponent of ≈ 0.588. It should be noted though that this is due to a fortuitous cancellation of errors, Flory theory both overestimates the repulsion energy and the entropic elasticity.

The scaling result for a polymer in a poor solvent can also be found using Flory theory. We will present this here without derivation. In a poor solvent, there is a net attraction and the polymer collapses into a state described as a globule. Considering the Flory free energy for a polymer chain as before (Equation 8.15) and adding a term for the entropic contribution of compression it can be found that the end-to-end distance of a real polymer chain in a poor solvent scales as,

$$R \propto N^{1/3}. \tag{8.19}$$

8.2 FROM STATICS TO DYNAMICS

A small particle in any liquid will diffuse due to the fluctuations of the number of molecules hitting it randomly from different directions. The trajectory of the particle is a random walk, this motion was first observed by Robert Brown and is called *Brownian motion*. The three-dimensional mean square displacement of the particle during time t is proportional to t, with the coefficient of proportionality related to the *diffusion coefficient* D,

$$\langle [\mathbf{r}(t) - \mathbf{r}(0)] \rangle^2 = 6Dt. \tag{8.20}$$

If a constant force $\mathbf{f}$ is applied to a small particle, pulling it through a liquid, the particle will achieve a constant velocity $\mathbf{v}$ in the same direction as the applied force.

For a given particle and a given liquid, the *friction coefficient* relates the force and the velocity,

$$\mathbf{f} = \xi \mathbf{v}. \tag{8.21}$$

The diffusion coefficient D and the friction coefficient D are related through the Einstein relation,

$$D = \frac{k_B T}{\xi}. \tag{8.22}$$

8.2.1 *Rouse model*

In Section 8.1 we briefly introduced the Gaussian chain as a model for the equilibrium properties of polymers. Here we will introduce the *Rouse model* as a model for the dynamic properties of a polymer. As in the Gaussian model, we consider a polymer chain of beads connected by Gaussian springs. When a polymer chain moves through a solvent every bead will continuously collide with the solvent molecules. Besides this systematic friction force, the bead will experience random forces, resulting in Brownian motion. Each bead will experience friction with the friction coefficient, ξ. The total friction coefficient of the whole Rouse chain is the sum of the contributions of each of the N beads,

$$\xi_R = N\xi. \tag{8.23}$$

The centre of mass diffusion coefficient for a Rouse polymer can be obtained from the Einstein relation 8.22,

$$D_R = \frac{k_B T}{\xi_R} = \frac{k_B T}{N\xi}. \tag{8.24}$$

The polymer diffuses a distance of the order of its size during a characteristic time, called the *Rouse time*, τ_R,

$$\tau_R \approx \frac{R^2}{D_R} \approx \frac{R^2}{k_B T/(N\xi)} = \frac{\xi}{k_B T} N R^2. \tag{8.25}$$

The Rouse time is of special significance. On time scales shorter than the Rouse time, the polymer exhibits viscoelastic modes, and on time scales longer than the Rouse time, the motion of the polymer is simply diffusive.

8.3 BOND FLUCTUATION METHOD

Computer simulations are very useful in simulating polymers. The *Bond Fluctuation Method* (BFM) is one such model that can be used to simulate polymer dynamics. This is an ergodic model that reproduces Rouse polymer dynamics first introduced by I. Carmesin and Kurt Kremer in 1988 [130]. This is a dynamic algorithm which moves a polymeric object on a lattice via local jumps of the monomers. The number of monomers and hence the number of bonds is fixed. In order to conserve bond number, local moves are made by allowing for a variable bond length.

The polymer consists of N monomers, implemented on a square lattice. In 3d each monomer is represented by a unit cube and occupies 8 lattice sites, one at each vertex. Each lattice site can only be part of one monomer, modelling excluded volume. Bond vectors are taken from a set of 108 vectors, the resultant allowed bond lengths are $2, \sqrt{5}, \sqrt{6}, 3, \sqrt{10}$, in order to fulfill the self-avoiding walk condition. To move the chain, a monomer is selected at random, it then tries to move at random the distance of one lattice unit into one of the eight lattice directions. If this move does not lead to overlapping monomers and complies with the bond vector conditions, it is accepted.

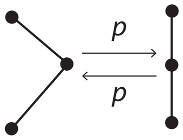
8.3.1 *Volume occupied*

It should be noted that since bond lengths can vary in the Bond Fluctuation Model, the volume occupied by the polymer depends on its specific conformation - the excluded volume associated to each monomer can overlap.

We can measure the volume occupied 'experimentally' by dilating the three-dimensional binary image describing the occupancy of each lattice site using a cubic structuring element of width 3. This gives precisely the excluded volume of the entire polymer. We will use this later in our analysis as a measure of the polymer.

8.3.2 *Kinetic implementation of the BFM*

We base our kinetic implementation of the Bond Fluctuation Method on the Gillespie method [10]. Allowable monomer moves will be attempted at a rate p, these are moves of one lattice unit in any of the six basic directions which do not break any of the applied conditions. Here applied conditions refers only to that of volume exclusion and the new bond vectors being in the allowed set, but this can be expanded if additional constraints are applied, as we will see later. This can be represented schematically as,



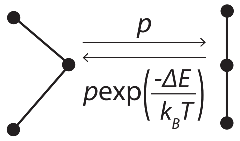
Allowable moves are calculated individually for every monomer and must be updated after every event. The rate p defines the fundamental polymeric timescale $\tau = 1/p$.

Semi-flexible polymer

Many polymers, including DNA, are semi-flexible. In order to account for this stiffness, we follow the approach of Zhang et al. [131] and introduce a squared-cosine bending potential E between successive bonds,

$$\frac{E}{k_B T} = k_s (1 - \cos\theta)^2, \tag{8.26}$$

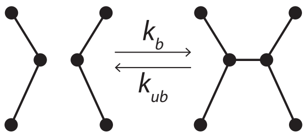
where θ is the change in angle between successive bonds and k_s is a parameter controlling the stiffness. Note that a monomer move affects three bond angles: the angle at the monomer and at its two neighbours. Applying stiffness, monomer moves are accepted at a rate p multiplied by the probability $P = \min(1, \exp(-\Delta E/k_B T))$, where $\Delta E/k_B T$ is the change in energy due to the move. Updating the schematic to include the effects of stiffness gives,



where moves towards a state with a greater bond angle between monomers require energy and those in the opposite direction do not.

Bridging

Diffusion-driven bridging between two monomers on a polymer can be implemented in a similar manner to that of Bohn and Heerman [132]. Any two monomers that are within (strictly less than) a spatial distance of three lattice sites are allowed to bridge. The rate of bridging for each monomer is dependent on the position of the monomers within the polymer and each monomer can only bridge with at most one other monomer at a time. In the simplest implementation the bridging rate for a monomer will either be k_b is there is another monomer in proximity and the monomer is unbridged, or 0 if not. Bridges break randomly with a mean lifetime $1/k_{ub}$, and therefore have exponentially distributed lifetimes. For two monomers within three lattice sites of each other we can represent this as,



It should be noted that bridged monomers are still able to move on the lattice but must maintain a bridge length strictly less than 3, this may limit the allowable monomer moves.

For any given parameter set, simulations are first run until steady state as determined by the volume occupied by the polymer reaching an approximate constant value. We use the volume occupied rather than the more classical square radius of gyration as we find this to be a much less noisy measure. The conformation of the polymer is then recorded. We repeat this process for 1000 random initial configurations.

9

PARTITION COMPLEX STRUCTURE CAN ARISE FROM SLIDING AND BRIDGING OF PARB DIMERS

As described in the background biological information (Chapter 5) it still remains an open question as to how the partition complex forms and what the structure of it is. Here, we utilise the discovery that CTP binding enables ParB dimers to slide along DNA and condense the centromeric region through the formation of DNA bridges to develop a new model of partition complex formation. We make use of both semi-flexible polymer and reaction-diffusion simulations. We first show that different ParB bridge lifetimes lead to distinctly different polymer conformations. We then study the short-lifetime regime in which distinct DNA structures (hairpins and helices) form and show that these structures result in the condensation of ParB-coated DNA. Finally, we combine ParB bridging and our previous sliding simulations in out-of-equilibrium polymer/reaction-diffusion simulations and show that bridging does not inhibit ParB sliding for sufficiently short bridge lifetimes. Overall, this supports a new model of partition complex formation in which ParB dimers load onto the DNA at *parS* sites before sliding diffusively along the DNA. Genomically distant, but spatially proximal, ParB dimers interact to form transient bridges that organise the DNA through the formation of hairpin and helical structures. We speculate that these structures facilitate the additional function of ParB to load SMC complexes onto the chromosome.

9.1 SEMI-FLEXIBLE POLYMER MODEL OF PARB BRIDGING

In order to obtain a realistic model of partition complex formation, we use a semi-flexible polymer model of the 10 kb centromeric region of *C. crescentus* in which every monomer corresponds to 20 bp, the approximate footprint of a ParB dimer [87, 119] (Figure 9.1a). The DNA is modelled as a linear chain using a kinetic implementation of the Bond Fluctuation Method polymer model [130] introduced in Section 8.3. Since the DNA is stiff at this scale, we introduce an energetic cost for bending ($k_s = 14$) to obtain the experimentally measured persistence length of $l_p \sim 120$ bp (Figure 9.1b) [133]. The BFM is well suited for this as it allowed for a large set of bond angles [130] and can therefore implement stiffness more realistically than models that only allow 0° or 90° bond angles. The lattice spacing corresponds to 2.2 nm.

Similar to other bacteria, chromosomal DNA in *C. crescentus* constitutes a volume fraction of about $1-2\%$. We obtain this volume ratio in the simulations by setting the size of the lattice appropriately. As described previously (Section 8.3), the volume occupied is not a fixed quantity in the Bond Fluctuation Method. Measuring the volume occupied from simulations we find that a $90 \times 90 \times 90$ cubic lattice with the stiffness parameter, k_s, chosen above results in an excluded volume fraction of

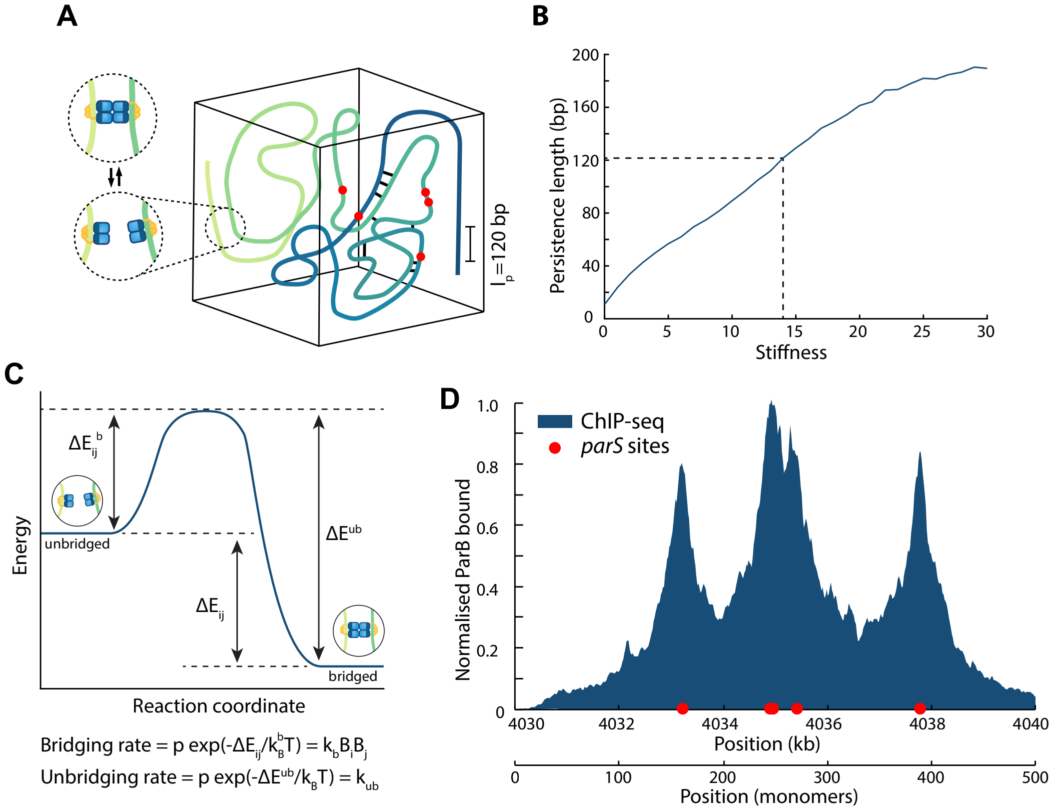


Figure 9.1: Semi-flexible model of ParB bridging. (a) Schematic of the polymer model. Bridges can form between genomically distant, spatially proximal monomers at a rate proportional to the ParB occupancy at each monomer. The ParB distribution is shown explicitly in (d). (b) The stiffness of the polymer versus the persistence length calculated. We choose a stiffness of 14 in order to give a persistence length of ~120 bp. (c) Energy landscape of a bridging-unbridging interaction between two spatially close monomers i and j. There is an overall change in energy ΔE_{ij} upon bridging. The relationship between the activation energies for bridging (ΔE_{ij}^{b}) and unbridging (ΔE^{ub}) and the reaction rates used in the model are shown. B_i is the ParB occupancy at monomer i and p is the polymeric move attempt rate defining the polymeric timescale $\tau = 1/p$. (d) The normalised ParB ChIP-Seq profile [70] specifies the ParB distribution. *parS* sites are represented by red dots. Note that two parS sites are very close together.

1.7% (an excluded volume per monomer of ~ 22 lattice sites).

The polymer is constrained by ParB induced bridging between monomers. We adapt the diffusion-driven bridging implemented by Bohn and Heerman [132], such that the bridging rate is dependent on the ParB occupancy at each monomer B_i, B_j. The rate of bridge formation between two monomers that are in proximity is then equal to the overall bridging rate k_b multiplied by the product of the ParB occupancy at each site, B_iB_j. Bridges dissociate randomly with rate k_{ub} and therefore have exponentially distributed lifetimes. The rates of bridging and unbridging are related to the activation energies for bridging and unbridging, the difference of which, ΔE_{ij} is the associated binding energy of the interaction (Figure 9.1c). The timescale of monomer dynamics $\tau = 1/p$ (where p is the move rate) is not experimentally known at the short length-scales simulated here. In the first part only the ratio of these two timescales is relevant such that we can leave k_{ub} fixed and vary the move rate p of the polymer.

Since ParB dimers can slide along the DNA, the spreading of ParB throughout the centromeric region can occur, at least in principle, independently of any 3d structure. We therefore initially model ParB dimers implicitly, using the relative probability of ParB occupancy, obtained from the experimental binding (ChIP-Seq) profile at 20 bp resolution (Figure 9.1d) [70], to specify the probability of a bridge forming when two given monomers come into proximity. This will allow us to first investigate how the observed ParB genomic distribution can, through bridging, affect the structure of the centromeric region, separately from the question of ParB spreading. We will examine the effect of both the processes of sliding and bridging later, here the ParB occupancy will depend explicitly on the presence (or not) of a ParB dimer on the polymer.

9.2 PARB BRIDGE LIFETIME RESULTS IN DISTINCTLY DIFFERENT POLYMER CONFORMATIONS

The multi-peaked ParB binding profile of *C. crescentus* consists of three clear peaks centred on five *parS* sites (note that two of these *parS* sites are only separated by 42 bp and so are not typically distinguishable in our figures) (Figure 9.1d). While two other putative *parS* sites have been identified [70], they are not associated with significant enrichment of ParB. This profile is used in our polymer model to specify the ParB dimer occupancies B_i along the polymer and thus, up to an overall parameter k_b, the bridging rate between proximal monomer pairs. Simulating the system, we found a surprising result: ParB-induced bridging leads to two distinct phases for the partition complex. Long bridge lifetimes cause the polymer to collapse into a globule-like structure (Figure 9.2c), whereas at shorter bridge lifetimes the polymer is more structured with long extended localised regions of bridging (Figure 9.2e). Note that 'long' and 'short' are relative to the polymeric timescale τ. Since we do not have an experimental estimate of this timescale at the length scale (20 bp) considered here, we cannot provide specific values.

The effect was also apparent in maps showing the location of the ParB bridges. Whereas bridge maps of the structured conformations show distinctive ±45° lines, those of the globular regime display a more random distribution. However, despite the clear differences in their conformations, both regimes exhibit very similar contact maps at the population level, which display a checker-board pattern centred on the *parS* sites and have no ±45° lines detectable (Figure 9.2d,f). Such a pattern is consistent with an overall preference for contacts within and between regions associated to peaks in the ParB binding profile, the globular regime has more contacts for the same number of bridges, as expected from its greater level of compaction. This highlights how the population-average view of DNA organisation may not be informative of the structure of individual conformations.

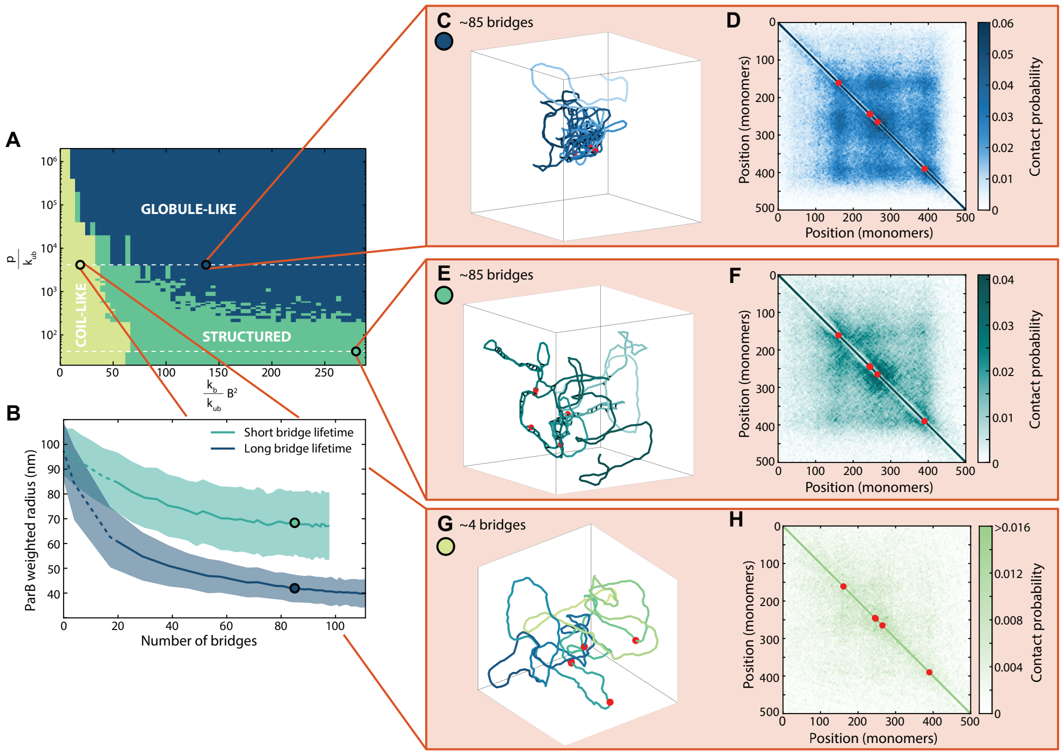


Figure 9.2: ParB bridge lifetime results in distinctly different polymer conformations. (a) Phase diagram of the system in terms of the effective binding constant $\frac{k_b}{k_{ub}}B^2 = \left\langle \exp\left(\Delta E_{ij}/k_BT\right)\right\rangle$, where B^2 is defined as the mean $\left\langle B_iB_j\right\rangle$ taken over all i, j with $|i-j| > 1$, and $\frac{p}{k_{ub}} = \exp\left(\Delta E^{ub}/k_BT\right)$, the relative bridge lifetime. (b) The mean ParB weighted radius for short and long bridge lifetimes (± SD)(indicated by the dashed lines in (a)) as a function of the number of bridges. Data from 1000 conformations for each parameter set. Circles indicate the respective locations of (c) and (d), (e) and (f), and (g) and (h). (c) An example conformation of the polymer in the globular state. (d) Average contact map at the same location based on 1000 conformations. A contact is defined as two monomers being within 5 lattice sites of one another. (e) An example conformation of the polymer in the structured state. (f) Average contact map at the same location, otherwise as in (d). The locations studied in (c) and (d), and (e) and (f) both have an average of ~ 85 bridges. (g) An example conformation of the polymer in the free coil-like state. (h) Average contact map at the same location. Based on 1000 conformations.

To better characterise these different regimes we constructed the phase diagram of the system (Figure 9.2a). Three regions could be identified: a free coil-like regime in which there are very few bridges (less than 20, a value chosen by inspection) and the polymer behaviour is dictated simply by its stiffness and volume-exclusion (Figure 9.2g, h), and the structured and globular (unstructured) regimes. We defined the transition between the structured and globular regimes using the ParB weighted radius, i.e. the radius of the *spatial* ParB distribution due to the polymer conformation. The globular state has a much smaller ParB weighted radius compared to the structured state with with the same number of ParB bridges (Figure 9.2b). This radius plateaus as the system goes further into the globular regime. We therefore chose a threshold of 55 nm to distinguish the two regimes based on two standard deviations above the plateaued mean value (Figure 9.2b).

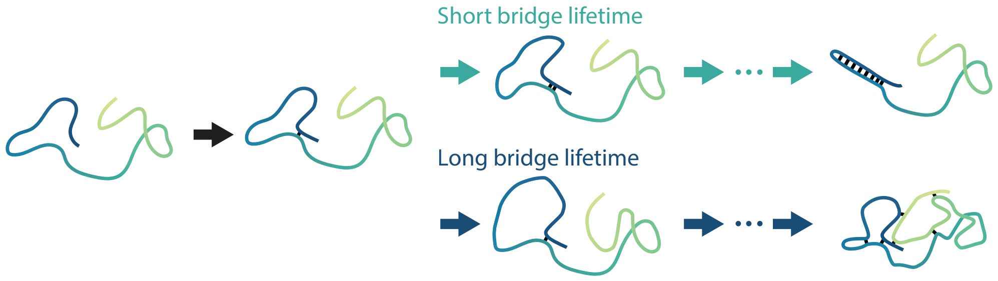


Figure 9.3: ParB bridge lifetime results in distinctly different polymer conformations. Cartoon illustrating bridging at both short and long ParB bridge lifetimes. At short bridge lifetimes the polymer conformation remains mostly unchanged and therefore bridges form close to each other. This reduces the entropic cost from the distortions caused by bridges. At long bridge lifetimes the polymer conformation can change significantly between bridging events, other genomically distant regions of the polymer may come into close contact and bridges can form between these regions. Here long and short bridge lifetimes refers to the move rate of the polymer.

ParB weighted radius. The ParB radius is calculated by combining the genomic distribution of ParB on the DNA (either based on the ChIP-seq profile for the uncoupled polymer simulations or the simulated position of ParB dimers in the coupled simulations) with the simulated conformations of the DNA polymer to obtain a spatial distribution of DNA bound ParB. We take an average across all 1000 conformations, aligning them by their centroids, to obtain a 3d density. We then determine the radius within which 95% of ParB dimers are found. We convert this value from lattice units to nanometers as follows. In our (stiff) polymer simulations, the bond length between monomers varies but has an average value of 3.0 lattice units. Since every bond/monomer corresponds to 20 bp and the length of a base pair is 0.33 nm [134], a lattice unit corresponds to 2.2 nm.

We propose that these two regimes arise due to the degree of movement that the polymer makes between bridging events. Bridging can be either kinetically limited (limited by the intrinsic bridging/unbridging rates) or diffusion limited (monomers coming into proximity is the limiting factor and the kinetics are so fast that bridge breaking and forming becomes correlated because newly broken bridges tend to recombine before the polymer can explore the conformational space) [135]. Consider a bridged polymer conformation (Figure 9.3). In the diffusion-limited region, recombination effectively increases the bridge lifetime (the activation energy of unbridging) [136, 137]. This strengthens a cooperative effect in which new bridges are more likely to form close to an existing bridge because adjacent monomers have a higher likelihood to also be in, or come into, proximity and the time needed for a new bridge to form is much less than than the (lengthened) bridge lifetime. Repetition of this process leads to the extended regions of bridges we observe (Figure 9.3). The conformational cost of bridging is also reduced by having bridges clustered together. A similar effect has previously been seen in simulations of the bridging protein H-NS [138]. On the other hand, at long bridge lifetimes bridging events are kinetically limited and the polymer is able to reorganise and explore the conformational space between bridging events. As a result, there are many more potential bridging events (monomers coming into proximity) away from existing bridges than in the short lifetime diffusion-limited regime. This results in both more bridges and a more random distribution of bridges and hence a globular polymer configuration. (Figure 9.3) The increase in the number of bridges overcomes the additional conformational cost of having the bridges dispersed rather than localised as in the structured regime.

Since the globular regime is reminiscent of previous proposals for partition complex organisation [69, 78, 79], we will focus next on examining the structured state. We will return to the globular state in the final section.

9.3 SHORT-LIVED PARB BRIDGING LEADS TO THE FORMATION OF HAIRPINS AND HELICES

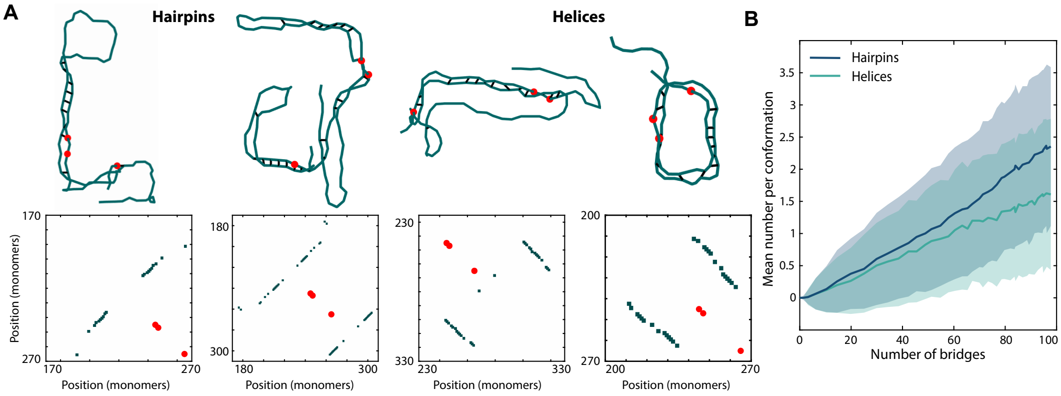


Figure 9.4: Short-lived ParB bridges results in the formation of hairpins and helices. (a) Example structures with corresponding bridge maps (bridge maps have been dilated to make lines clearer) for hairpins and helices with an average of 30 bridges. Red dots indicate the *parS* sites. Note that as two of the sites are only separated by 42 bp, they are not always distinguishable. (b) Mean number of hairpins and helices per conformation. Shading represents the SEM.

The structured regime found at short ParB bridge lifetimes is defined by the presence of two distinct structures, hairpins and helices. Hairpins form by the polymer bending 180° back on itself to form bridges between anti-parallel segments, whereas in helices, the polymer turns a full 360° with bridges between parallel segments (Figure 9.4a). These two structures are visually different but also have different underlying bridging patterns which allows them to be clearly identified in bridge maps. Hairpins correspond to +45° lines whereas helices correspond to −45° lines. The location of the line relative to the main diagonal indicates the length of the loop of the hairpin or the period of the helix. Unsurprisingly, these structures generally form near to the *parS* sites. However, we observed substantial variation: the tip of a hairpin (indicated by where the 45° line in the bridge map intersects the main diagonal) was often reasonably far from the nearest *parS* site (Figure 9.4a). At lower levels of bridging, these structures most frequently form within the region covered by the central peak containing three *parS* sites.

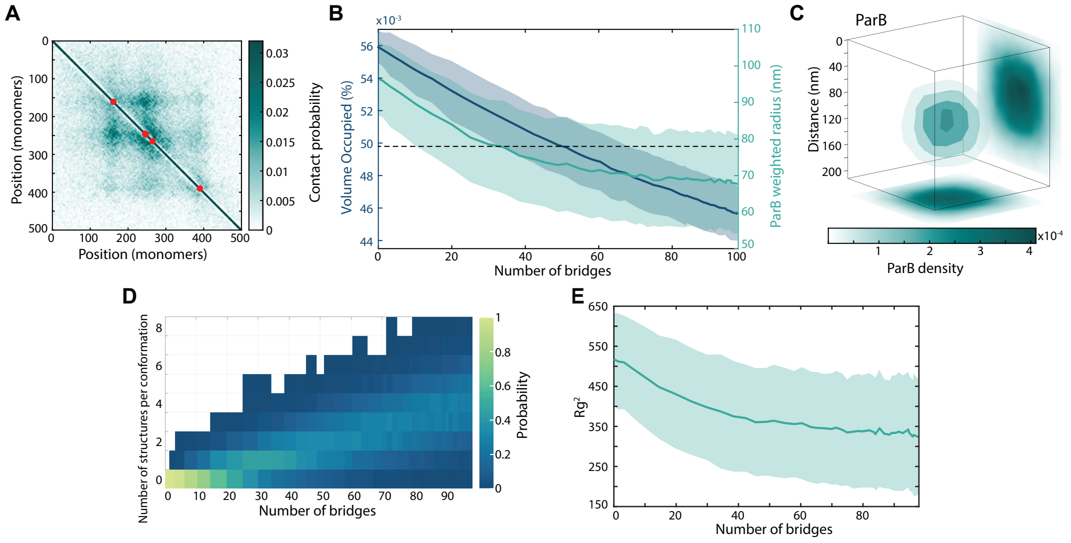


Figure 9.5: Short-lived ParB bridges results in the formation of hairpins and helices. (a) Average contact map for the polymer at an average of 30 short-lived bridges. (b) Mean volume occupied by the polymer and the mean ParB weighted radius. Shading represents the SD. The dotted line at 78 nm shows the experimentally determined ParB radius for *C. crescentus*. (c) Three-dimensional ParB density from partition complexes with an average of 30 bridges showing a radius of 78 nm. (d) Histogram showing the probability of a certain number of structures being present in any given conformation. After ~ 30 bridges it becomes most likely that a given conformation will have at least one structure. (e) Mean squared radius of gyration for the polymer as the number of bridges increases. Shading represents the SD.

We made use of the distinctive ±45° lines to quantify the occurrence of hairpins and helices as a function of the degree of bridging in the system (Figure 9.4b). We found that the frequency of both structures increased approximately linearly with the number of bridges, with hairpins being the most common. From about ~ 30 bridges every conformation contained at least one structure (Figure 9.5d). At the highest levels of bridging studied (~ 100 bridges) each conformation contained 3-4 structures, which could be of either type and involve multiple and distant *parS* sites. Nevertheless the different constituent structures could still be identified from the ±45° lines in the bridge maps. However, as discussed in the previous section, the ±45° lines are not apparent in the ensemble average contact map which displays a checkered pattern centered on the *parS* sites (Figure 9.5a).

Consistent with *in vitro* experiments, ParB bridging led to the condensation of the DNA polymer. Both the volume occupied and the squared radius of gyration decreased with the number of bridges (Figure 9.5b, e). *In vivo* the nucleoprotein complex is visualised through the spatial distribution of a fluorescently tagged variant of ParB, which forms distinct foci within cells. To connect with these observations, we combined the genomic distribution of ParB on the DNA (based on the ChIP-seq profile), with our simulated conformations of the DNA polymer to obtain the resulting spatial distribution of DNA bound ParB (Figure 9.5c). The resultant spherical density was reminiscent of that observed experimentally using single molecule microscopy. The radius of the partition complex of *C. crescentus* has been measured experimentally using super-resolution PALM microscopy to be ~78 nm

[99]. This could be achieved in our simulations with just 30 ParB bridges. This corresponds to a 20% decrease compared to the value in the absence of bridging (Figure 9.5b).

9.4 COUPLED SIMULATIONS OF SLIDING AND BRIDGING

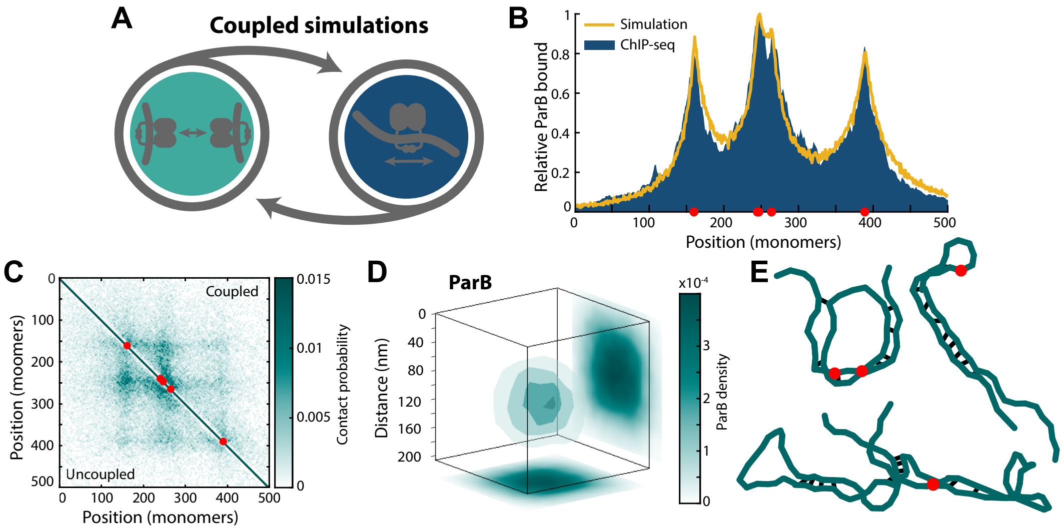


Figure 9.6: Sliding is not inhibited by short-lived ParB bridges. (a) Representation of coupled polymer simulations in which we combine bridging and sliding. (b) Profile of ParB as generated from sliding and bridging simulations with a bridge lifetime of 1 s compared to ChIP-seq data from a previous study [70], both normalised by maximum height. (c) Average contact map for coupled simulations (top right) and uncoupled simulations (bottom left) with a radius of ~78 nm. Each contact map is made from 1000 simulations. (d) Three-dimensional average of the ParB partition complex for 25 bridges with a radius of 78 nm. (e) Examples of hairpin and helical structures found in coupled simulations, with an average of 25 bridges.

We next investigate whether ParB bridging is compatible with the ParB binding profile i.e. would the spontaneous formation of ParB bridges between spatially-proximal but genomically distant ParB dimers limit overall ParB spreading and produce a fundamentally different binding profile? To answer this question we combine our polymer simulations and sliding simulations from Chapter 7 together (Figure 9.6a). Unlike the previous simulations, the bridging of proximal monomers is now explicitly dependent on the presence of a ParB dimer at each site rather than on a pre-specified ParB distribution. Note that since ParB sliding is a non-equilibrium process, this coupled model is therefore necessarily out-of-equilibrium. We assume that bridged dimers are not able to slide along the DNA, due to the entrapment of genomically distant regions, so that they act as roadblocks for unbridged sliding dimers. The simulations are ran until steady state and the ParB distributions and polymer conformations recorded.

The same values determined in Chapter 7 are used for ParB dimer loading ($k_{\text{on}} = 200 \cdot k_{\text{off}}$) and dissociation ($k_{\text{off}} = \frac{\log(2)}{64}\text{s}^{-1}$). There are currently no estimates

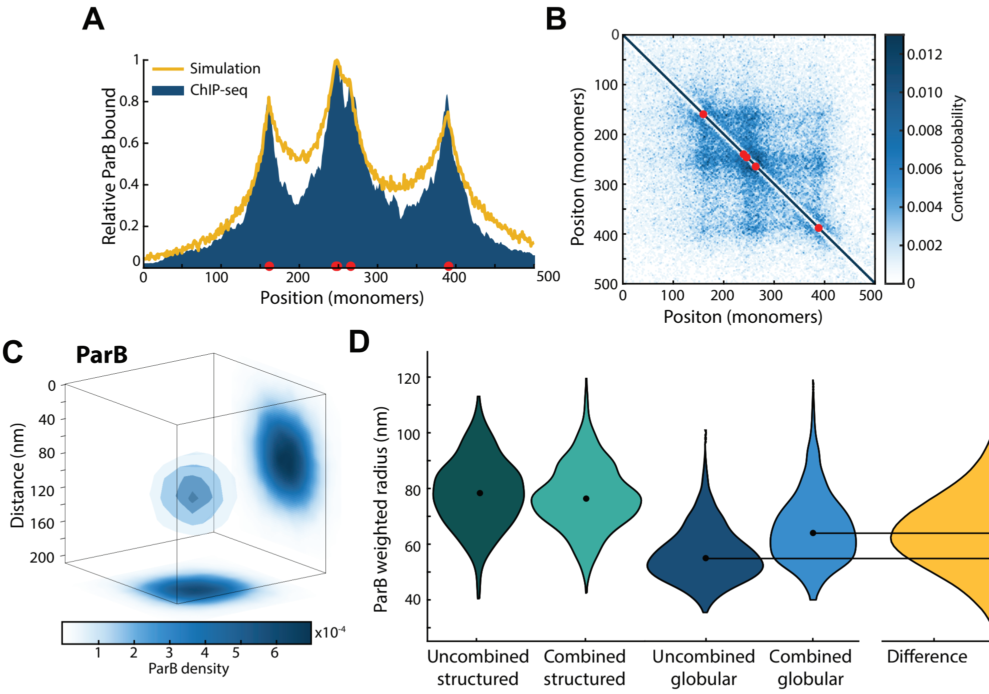


Figure 9.7: Sliding is restricted in the globular regime. (a) Profile of ParB as generated from the sliding and bridging simulations in the globular regime compared to the ChIP-seq data from a previous study [99]. (b) Average contact map for coupled simulations in the globular regime with an average of 28 bridges. (c) Three-dimensional average of the ParB partition complex in the globular state. (d) Violin difference plots for the ParB weighted radius in uncoupled and coupled (with explicit ParB dimers) simulations where there is an average of 28 bridges for both the globular and structured regimes.

for the bridge lifetime. We expect bridges to have a significantly shorter lifetime than that of ParB dimers on the DNA and therefore a nominal value of $k_{ub} = 1\,\mathrm{s}^{-1}$ is chosen. With too high a value (of the order of the ParB lifetime on the DNA) sliding ParB dimers would not have time to move past roadblocks (ParB bridges) before unbinding. We access the two regimes discussed in the first section through the mobility of the polymer. We arbitrarily choose two values of p to represent the globule-like and structured regimes (based on the sweep of the simple bridging model). This leaves the sliding diffusion coefficient and overall bridging rate as free parameters. These are chosen such that we can reproduce both a 78 nm ParB radius and the expected genomic distribution. We are unable to use the value found for the diffusion coefficient in Chapter 7 due to the introduction of ParB bridges resulting in ParB dimers sliding a shorter distance creating sharper peaks. Thus this value must be tuned based on the number of ParB-ParB bridges. For the structured regime discussed below a value of $4.4 \times 10^{-3}\,\mu\mathrm{m}^2\mathrm{s}^{-1}$ is used to resolve the ChIP-seq profile for 25 bridges. This is a lower bound. Larger values have very little effect on the ParB profile recovered as sliding ParB dimers reach equilibrium between bridging events. This lower bound is within an order of magnitude of the diffusion coefficient of ParB measured *in vivo* [71, 81].

For the structured regime, we found that the combined simulations could reproduce the binding profile measured by ChIP-seq (Figure 9.6b), with an even better fit than we obtained from the non-polymeric sliding simulation (Figure 7.4a). Importantly, we also observed the same hairpin and helical structures as in the non-combined polymer simulations that had the ParB binding profile given as a input (Figure 9.6e) and obtained very similar average contact maps (Figure 9.6c). These structures again compact the polymer and we could achieve the measured radius of 78 nm (Figure 9.6d).

In the globular regime we were able to broadly reproduce the ChIP-seq profile although the simulations could not accurately capture the depth of the valleys (Figure 9.7a). Similar contact maps were found (Figure 9.7b). Interestingly, the partition complex is less condensed in the combined simulations than in the non-combined simulations at the same mean number of bridges, whereas no significant difference was detected for the structured regime (Figure 9.7c, d).

Recent *in vitro* studies have shown that DNA-loaded ParB dimers of *B. subtilis* can load additional dimers independently of *parS* ('ParB-ParB recruitment') [114], potentially explaining the cooperative non-specific DNA binding observed previously [73, 76, 139] and consistent with interactions between dimers through their N-terminal domains [74, 140]. To explore whether such recruitment could be relevant *in vivo*, we added cis ParB-ParB recruitment to our model (Figure 9.7a). Although trans recruitment was also shown by the same authors this would be substantially more challenging and computationally intensive to implement.

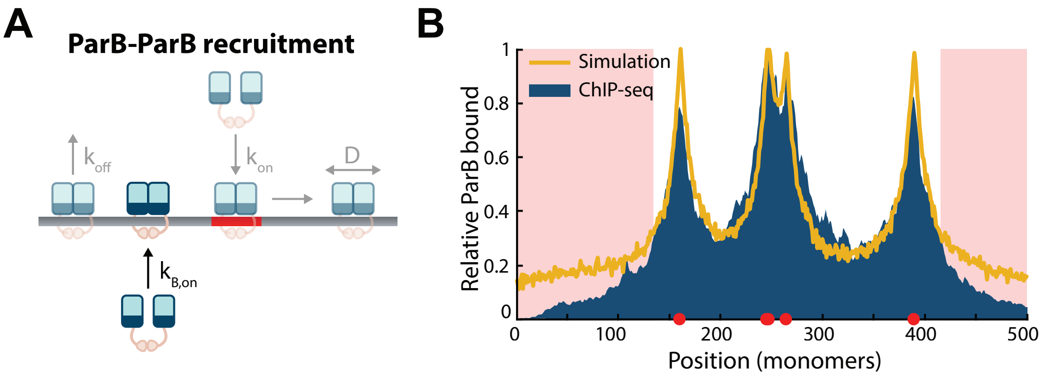


Figure 9.8: ParB-ParB recruitment fundamentally changes the shape of the ParB profile on DNA. (a) Diagram displaying the addition of ParB-ParB *in cis* recruitment to the coupled model. (b) Profile of ParB along the polymer with additional ParB-ParB recruitment, areas where the profile is substantially different to ChIP-seq data are highlighted in red. Simulations are in the structured regime.

We found that even a relatively low ParB-ParB recruitment rate, for which the total number of bound dimers increased by less than 20%, results in a fundamentally different binding profile. ParB spreading was increased through the appearance of slowly decaying 'shoulders' at the extremes of the distribution. As a result the distinctive exponential decay seen in the experimental ChIP-seq profile is no longer reproduced and this could not be remedied by changing the model parameters (Figure 9.7b). This result suggets that ParB-ParB recruitment does not play a significant role *in vivo* in ParB spreading, consistent with the finding of Tišma *et al.* that ParB-ParB recruitment accounts for only 10% of ParB loading events *in vitro* [114].

10

PARB-PARB TETHERS: A NEW MODEL FOR PARABS SEGREGATION

In the previous chapter, we have seen how ParB-ParB bridges play an important role in the formation of the partition complex leading to the emergence of different DNA structures. Coupling this knowledge with the ability of ParB dimers to recruit cytosolic ParB dimers to neighbouring DNA strands [114] leads us to wonder if ParB-ParB interactions could also play a role in the segregation of the partition complex.

The current paradigm for DNA segregation by the ParAB*S* system is that ParA-ParB tethers form between ParB dimers as part of the partition complex and ParA dimers non-specifically bound to the nucleoid. Here, we instead present a new model in which tethers are formed via ParB-ParB interactions. We take an experimentally driven modelling approach, using the F plasmid of *E. coli* as our model system, in order to compare this new model of ParB-ParB tethers to the model of ParA-ParB tethers. We propose that instead of forming half of the tether, ParA takes the role of promoting tether formation, which instills directionality into the system. We support this model experimentally via the constant diffusion rate of the partition complex independent of the ParA concentration and the lack of visible ParA colocalisation with the partition complex.

10.1 PLASMID MOBILITY IS UNAFFECTED BY PARA LEVELS

One commonly employed method to understand how a biological system functions is to examine the effects of mutants or knockouts. This can allow one to judge the functionality of the individual components. Whilst in the past this approach has provided insights into the mechanism behind the ParAB*S* system; in F plasmid, currently published mutations are highly destructive [101]. Due to the importance of the ParAB*S* system for plasmid retention and the crucial contribution of each individual component, mutations leads to plasmid mispositioning and hence a high rate of plasmid loss. Crucially, in wild type, plasmids are found within the nucleoid, whereas mutations result in the plasmid being excluded from the nucleoid, and mostly found in the cytoplasm at the cell poles [101]. This significantly alters the environment of the plasmid, and consequently would be expected to result in different plasmid dynamics.

Therefore, to gain a deeper insight into the nature of the ParAB*S* system, we took advantage of a natural phenomena. It has previously been observed that ParA has an oscillatory behaviour within cells, wherein the majority of ParA is alternately localised between cell halves [142–145]. This results in a heterogeneous population of cells at birth in terms of their ParA concentration, whist the plasmids

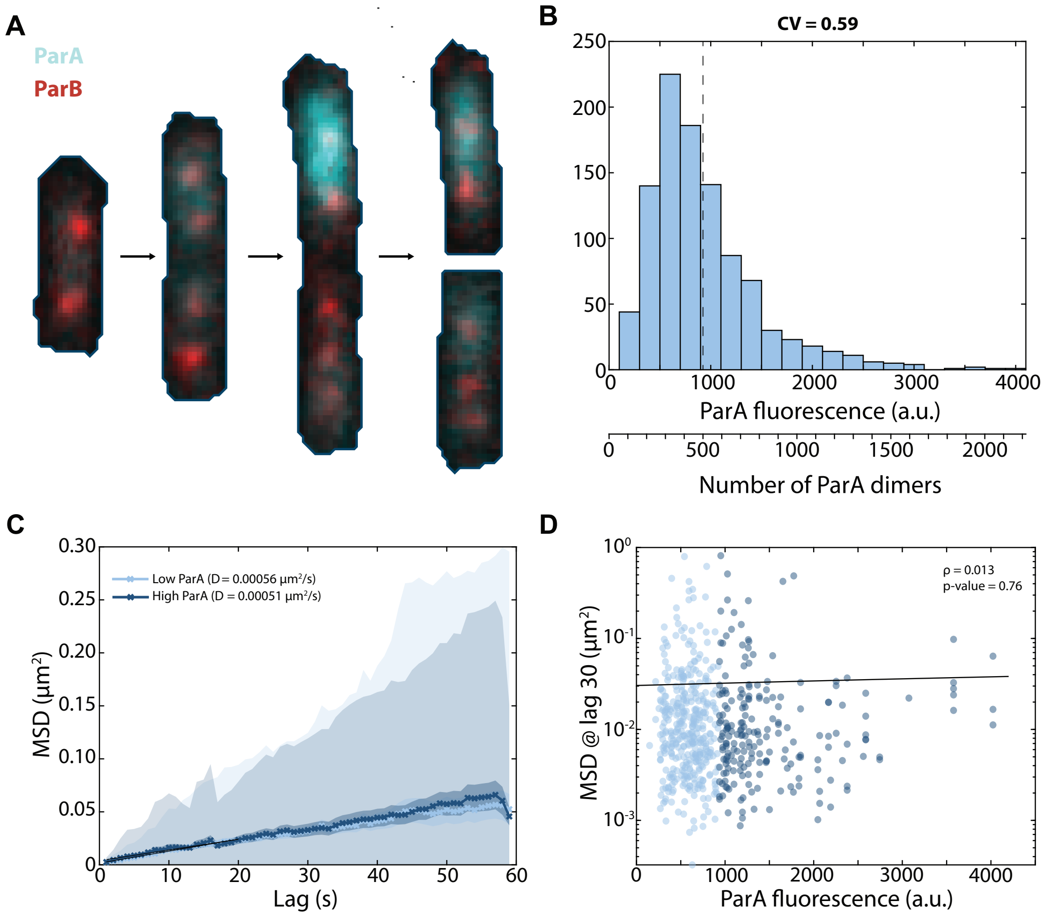


Figure 10.1: Plasmid mobility is unaffected by ParA levels. (a) Example of a cell undergoing ParA (blue) asymmetry resulting in the birth of two daughter cells with highly heterogeneous levels of ParA. The positioning of the partition complex (red) appears to be unaffected by ParA levels. (b) Histogram showing the ParA concentration across all cells. To establish a correlation with the number of ParA dimers, we compared the mean ParA concentration with the estimated mean number of dimers per cell, as determined from previous studies (mid-range estimate from [93, 99, 141]). (c) Mean square displacement for ParB-mTurquoise in cells with a high ParA concentration (> 923 fluorescence/pixel) and a low ParA concentration (< 923 fluorescence/pixel). (d) Mean square displacement at a lag of 30s as a function of the ParA concentration. Black line shows the best linear fit, with a gradient of 2×10^{-6}.

remain correctly positioned (Figure 10.1a, b). We can utilise this effect to study the behaviour of the partition complex in highly varying ParA environments.

ParA asymmetry. To quantify ParA asymmetry we measure the fractional difference in the ParA signal between each half of the cell,

$$A = \frac{S_1 - S_2}{S_1 + S_2}, \tag{10.1}$$

where S_1 is the signal in one cell half and S_2 the signal in the other cell half. We will define cells as symmetric when $|A| < 0.33$ and asymmetric when $|A| \geq 0.33$.

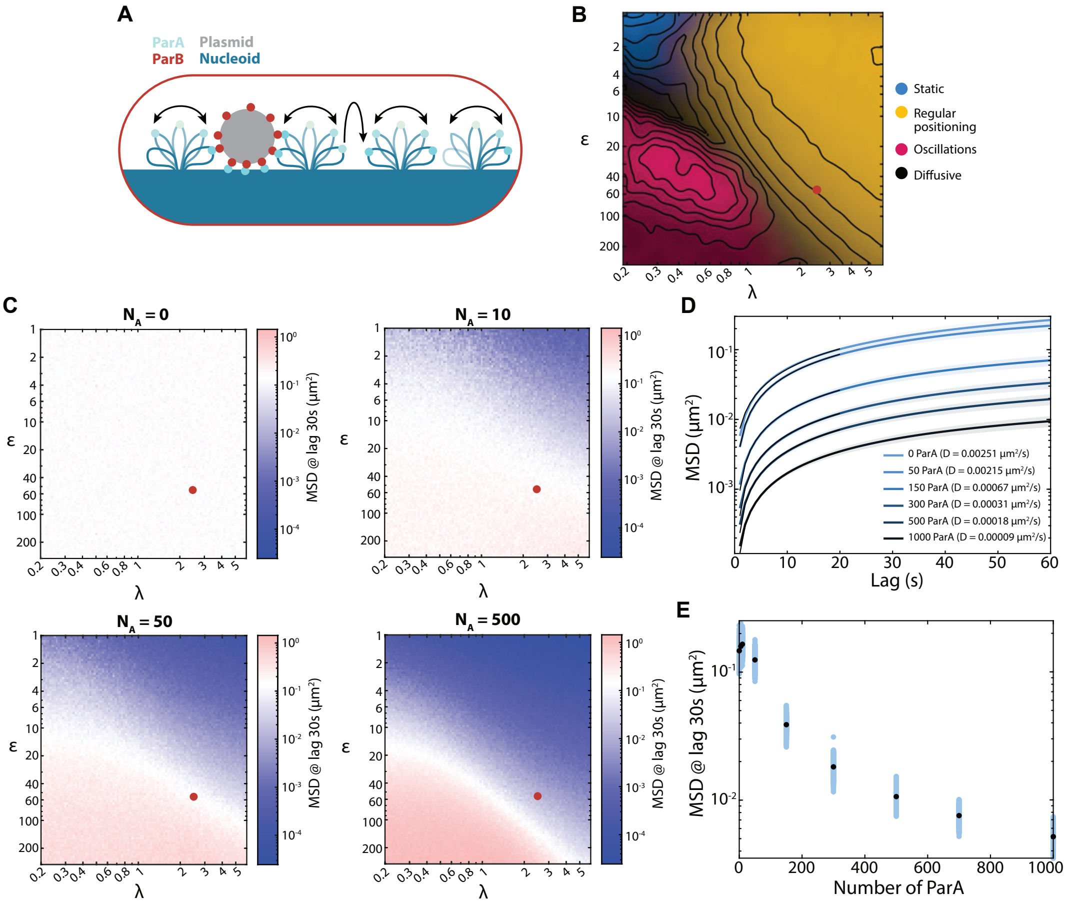


Figure 10.2: Hopping and relay model cannot reproduce experimentally observed behaviour. (a) Schematic of the hopping and relay model. Blue shading: nucleoid, blue strokes: DNA strand, light blue: nucleoid bound ParA dimers, red: partition complex bound ParB dimers. Arrows indicate dynamics of the system, chromosome fluctuations and ParA dimer hopping. (b) Phase diagram shown in terms of the dimensionless parameters λ and ϵ. The colour is based on an analysis of simulated trajectories as follows. Yellow: regular positioning, blue: static, pink: oscillations, black: diffusion. Phase diagram is for 500 ParA dimers. Red dot represents the position found to best represent F plasmid. Reproduced from [62]. (c) Phase diagram for the hopping and relay model showing the MSD at a lag of 30 s. The number of ParA dimers is varied from $0, 10, 50, 500$. The colour bar is such that red signifies an increase in MSD when compared to the mean MSD at 0 ParA dimers, and blue a decrease in MSD. Red dots marks the position best found to represent F plasmid. (d) MSD curves for different numbers of ParA dimers for the hopping and relay model at the location on the phase diagram found to best represent F-plasmid. This corresponds to $\lambda = 2.66$ and $\epsilon = 56.42$. (e) The MSD at a lag of 30 s for a varying number of ParA at the same location as in (d).

To understand the effect of ParA on the partition complex we analyse cells with a symmetric distribution of ParA and divide the cells into two populations: those with a high level of ParA (>923 fluorescence/pixel) and those with a low level of ParA (<923 fluorescence/pixel), split based on the mean ParA concentration of the population. We analyse the mean square displacement (MSD) of plasmids over a short-time interval in these different conditions. We chose this time scale to examine the plasmids diffusive behaviour, avoiding the potential influence from nucleoid tethers which becomes more pronounced at longer time scales. Both MSD curves obtained are linear, indicating that plasmid movement is diffusive, and we saw no significant difference between the MSD found in either conditions ($5.6 \times 10^{-4}\,\mu m^2/s$ in low ParA and $5.1 \times 10^{-4}\,\mu m^2/s$ in high ParA) (Figure 10.1c). Furthermore, when plotting the MSD at a lag of 30 s against the ParA fluorescence for individual cells, we find there is no evidence of correlation between the MSD and ParA fluorescence (p-value = 0.76, $\rho = 0.013$) (Figure 10.1d). It is notable that the ParA concentration across the population of cells analysed is highly heterogeneous with values spanning a ten-fold range. We cannot directly equate the fluorescence measurements to a precise count of ParA dimers, but using the mean ParA fluorescence and considering the average number of ParA dimers per cell as 500 (a mid-range estimate from [93, 99, 141]), we can estimate the number of ParA dimers in the analysed cells to range between approximately 50 and 2000 dimers (Figure 10.1b).

We compare these experimental results with the previously proposed hopping and relay model (Figure 10.2a) [62]. This model was able to capture a range of different behaviours of partition complex movement, which can clearly be seen in the phase diagram (Figure 10.2b). This phase diagram is defined by sweeping over the variables λ and ϵ. As previously defined, $\lambda = \sqrt{\frac{2D_h/k_d}{L/2}}$ where D_h is the diffusion coefficient of nucleoid bound ParA-ATP dimers, k_d is their basal unbinding rate, and L is the length of a cell. This can be understood as the average distance relative to the nucleoid length that a ParA dimer diffuses before unbinding due to basal ATP hydrolysis. $\epsilon = \frac{k_h}{k_d}$ is the ratio of the timescale of ParB-induced hydrolysis (k_h) to the timescale of basal hydrolysis [62]. Analysing the hopping and relay model in a similar manner to the experiments we can look at the MSD at a lag of 30 s across the phase space for increasing numbers of ParA dimers (Figure 10.2c). The response of the model to increasing ParA concentration is dependent on the specific regime. Note, with zero ParA dimers the partition complex is simply undergoing random diffusion as no ParA-ParB tethers can be formed. In the oscillatory regime, located in the bottom left-hand corner, the MSD of the plasmid at 30 s increases as the number of ParA dimers increases. In contrast, in the regular positioning regime or static regimes the MSD decreases. In both cases, this increase or decrease becomes more pronounced as the number of ParA dimers is increased. The different response to changing ParA levels is due to the action of tethers in these two regions. In the oscillatory regime the tethers act to move the plasmid from one side of the cell to the other. More tethers make this effect stronger and therefore cause the plasmid to move more quickly. Whereas, in the static/regular positioning regime, the tethers act to keep the plasmid at a given location. In this case, an increase in the number of tethers will lead to the plasmid being held more tightly at this location, thus resulting in a lower MSD. Notably, there is a thin white strip between these two regions where the plasmid mobility is unaffected by the number of ParA dimers. Although, it should be highlighted that the exact location of this strip changes.

The phase diagram shows that this model cannot replicate the experimental finding that the MSD of the plasmids is unchanged by varying ParA concentrations. To further illustrate this in the context of F plasmid, we examine the position within the phase diagram where the mobility of F plasmid was identified as being most similar to (this is represented by a red dot) [62]. As expected for the regular positioning regime, as the number of ParA is increased we observe a twenty-fold reduction in the MSD at a lag of 30 s (Figure 10.2e) and in the diffusion coefficient obtained from fitting to the MSD curves (Figure 10.2d). This directly contradicts our experimental observations.

10.2 ABSENCE OF PARA FOCI IN CELLS

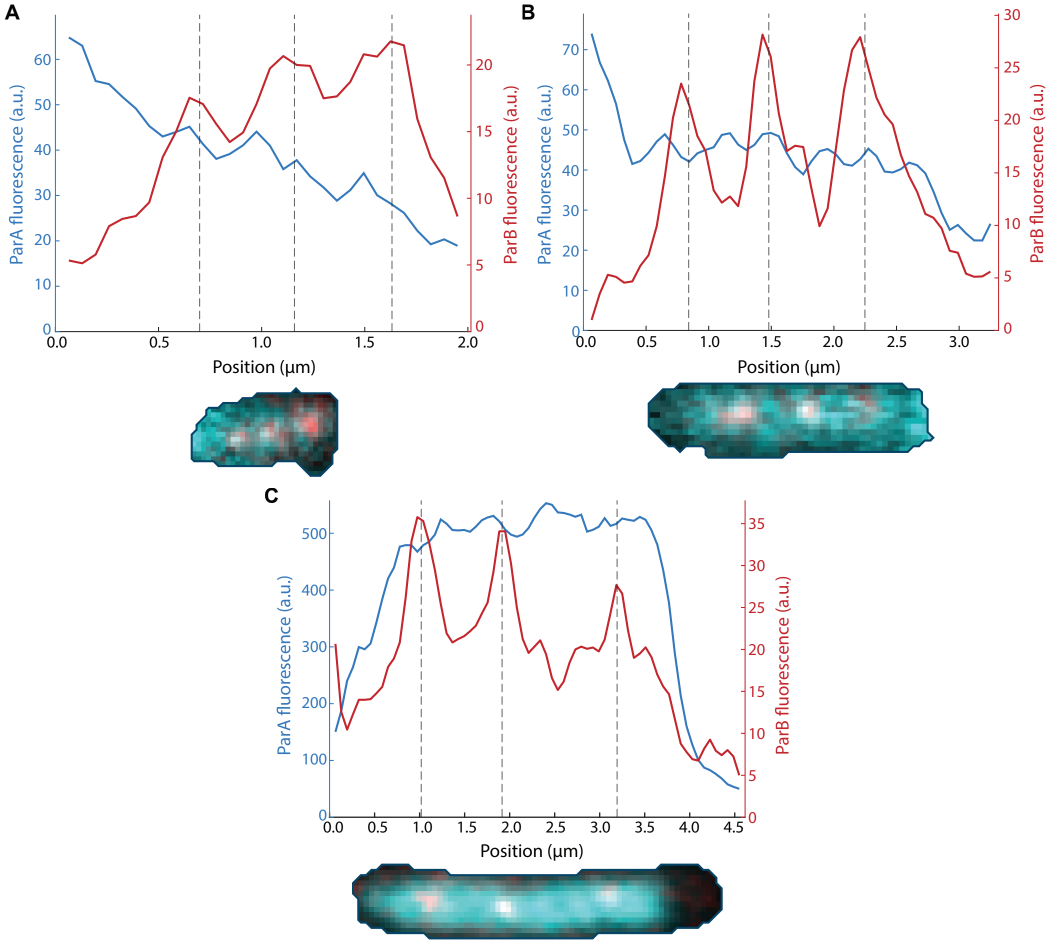


Figure 10.3: Absence of ParA foci in cells. (a) (Top) Example line profile of ParA (blue) and ParB (red) within a cell. (Bottom) Image of cell from which the line profile is taken. (b), (c) Same as in (a) for different representative cells.

In the presence of ParA-ParB tethers, one would expect to observe ParA foci colocalising with ParB since a large number of ParA dimers will be bound to the partition complex. This is the result observed in the hopping and relay model [62].

However, we see no evidence of this *in vivo*. ParB foci appear as sharp, distinct peaks but we do not observe any ParA foci or strong ParA colocalisation with ParB. ParA appears as a cloud across the nucleoid both when there is a low level of ParA fluorescence (Figure 10.3a, b) and when there is a high level of ParA fluorescence (Figure 10.3c). Looking at the average ParA surrounding a plasmid, oriented by the direction of plasmid movement, we see a similar result. The profile is almost homogeneous with only a slight peak at the location of the plasmid (Figure 10.4a).

10.3 PARA ORIENTS PLASMID MOVEMENT

Since ParA is critical for the positioning of plasmids we look to better understand its effect on plasmid movement. *In vitro* the partition complex has been observed following a ParA gradient [100]. *In vivo* this result has not been clear, although some studies have observed plasmids migrating behind a retracting gradient [97], plasmids often appear unperturbed by a ParA gradient across the cell, remaining at their regular positions rather than moving up the gradient (an example of this can be seen in Figure 10.3a).

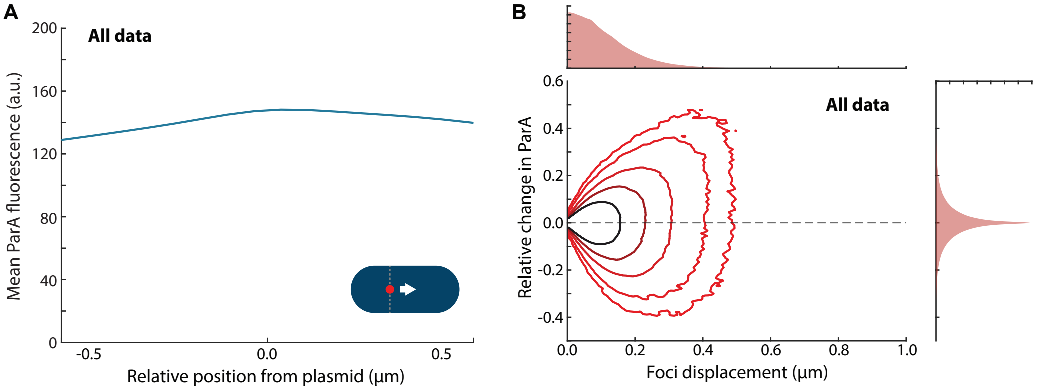


Figure 10.4: Average plasmid shows no preference for ParA concentration. (a) Mean ParA fluorescence around a plasmid for all data, the ParA fluorescence is oriented by the direction the plasmid moves in the next time step such that the plasmid always moves to the right side of the diagram. (b) Contour plot for the relative change in ParA as a function of the foci displacement for all data. The plots on the top and right hand side show the kernel density plots for each axis. Time step of 1 frame (1 min).

To investigate the action of ParA in more detail, we look at the relative change in ParA against foci displacement (Figure 10.4b). This is clearly symmetric about the change in ParA, as can be seen both from the contour plot and the kernel density plot. This suggests that on average, plasmids are just as likely to move down the ParA gradient (a negative change in relative ParA concentration) as to move up (a positive change in relative ParA concentration). We propose that this symmetry arises since the majority of plasmids being studied are at their steady state positions. The system is dynamic and plasmids exhibit slight fluctuations around these positions due to inherent noise. Given that the ParA gradient is not perfectly uniform, these fluctuations result in the plasmids moving up and down the ParA gradient in equal measure as they temporarily shift away from their steady state positions before returning. A symmetrical distribution is then the expected result for any object fluctuating around a home position in this manner.

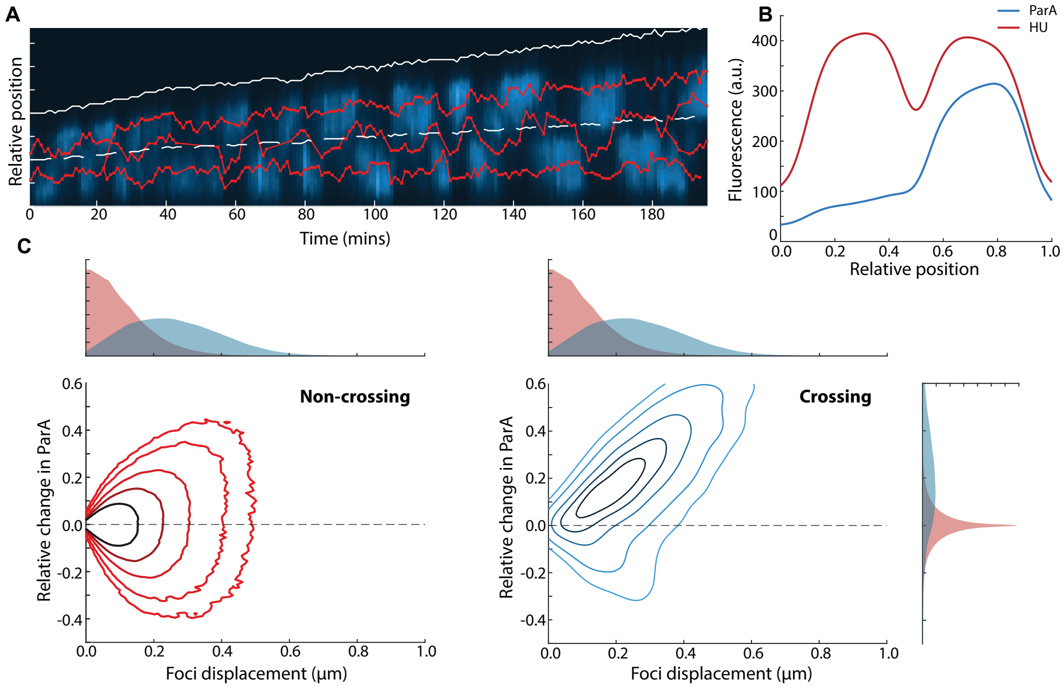


Figure 10.5: ParA directs the orientation of plasmid movement. (a) Example of a cell cycle with the ParB foci shown by the red crosses, and the ParA concentration across the cell represented by the blue contour plot. (b) Plot showing the average ParA and HU (representing the nucleoid) fluorescence across the cell, for cells with an asymmetry greater than 0.33. Cells are orientated such that the cell half with the highest ParA concentration is on the right. (c) Contour plots for the relative change in ParA as a function of the foci displacement after a time step of 1 min. Plot on the right is for plasmids crossing the centre line in cells with asymmetry > 0.33, plot on the left is for all other plasmid movement. The plots on the top and the far right-hand side show kernel density plots for the same data (blue represents the crossing plasmids, and red all other plasmids).

Due to the majority of plasmids being found at their steady state positions, experimentally assessing the impact of ParA on partition complex movement presents a challenge. Therefore, to gain an insight into the functionality of ParA, we again leverage the inherent asymmetry found in the distribution of ParA. Specifically, we take advantage of cells that naturally develop an asymmetric ParA distribution. This often occurs during the later stages of the cell cycle (Figure 10.5a). This asymmetry generates a sharp ParA gradient across the centre of the cell (Figure 10.5b) which can lead to plasmids transitioning from one cell half to another (Figure 10.5a). These transitioning plasmids present an ideal opportunity for investigating the influence of ParA on the movement of the partition complex. We separate the data for plasmid displacement against relative ParA change into two groups: plasmids crossing the centreline in asymmetric cells (crossing) and all other plasmid movements (non-crossing). The non-crossing group is nearly identical to that for all data as this forms the majority of events (Figure 10.5c (left panel) and Figure 10.4b). However, analysing the movement of crossing plasmids shows a substantial bias for foci movement up the ParA gradient (a positive change in relative ParA concentration) (Figure 10.5c (right panel). Furthermore, the movement up the gradient correlates with the relative change in ParA, a greater positive change in ParA results in a greater foci displacement.

Examining the local ParA gradient in the vicinity of plasmids in relation to the direction of plasmid movement further confirms this result (Figure 10.6). As expected, for non-crossing plasmids the gradient is very similar to that for all data, roughly symmetric about the plasmid position with the plasmid located at a minor peak. In contrast, crossing plasmids are surrounded by a distinct gradient of ParA. On average, plasmids move in the direction of increasing ParA gradient, note that Figure 10.6 is oriented such that all plasmids move to the right. This demonstrates how ParA plays an essential role in partition complex positioning, acting to, in some way, bias tether formation such that the partition complex will move up a ParA gradient.

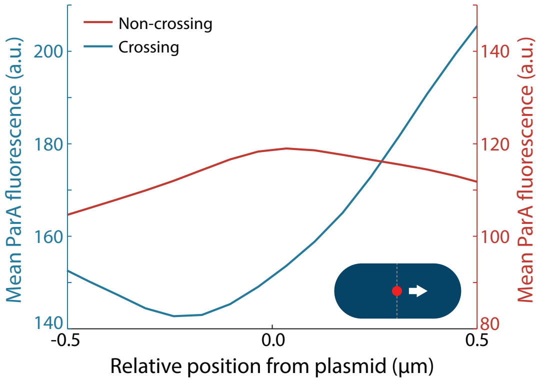


Figure 10.6: ParA directs the orientation of plasmid movement. Plot showing the mean ParA fluorescence around a plasmid for crossing (blue) and non-crossing (red) events. The ParA fluorescence is oriented such that in the next time step (1 min) the plasmid always moves to the right of the diagram.

10.4 PARB TETHERING AS AN ALTERNATIVE MODEL

There is currently no experimental data confirming that ParA can form stable tethers capable of positioning the plasmid. Furthermore, we have seen how a ParA-ParB tether model cannot explain the experimental result that the plasmid MSD remains constant, independent of the levels of ParA, and we have seen no evidence of the ParA foci we would expect if it were to form stable tethers with the plasmid. In contrast, ParB has been shown to be able to form bridges which can compact the DNA [88, 89], these ParB-ParB interactions are therefore able to remain stable under mechanical stress. Prior research has also demonstrated the ability of DNA-bound ParB dimers to recruit cytosolic ParB dimers onto spatially neighbouring non-specific DNA [114]. We propose that this mechanism could also be harnessed to recruit cytosolic ParB dimers to the nucleoid. Whilst we see no evidence of ParA forming part of the tether, we know that it is essential for the

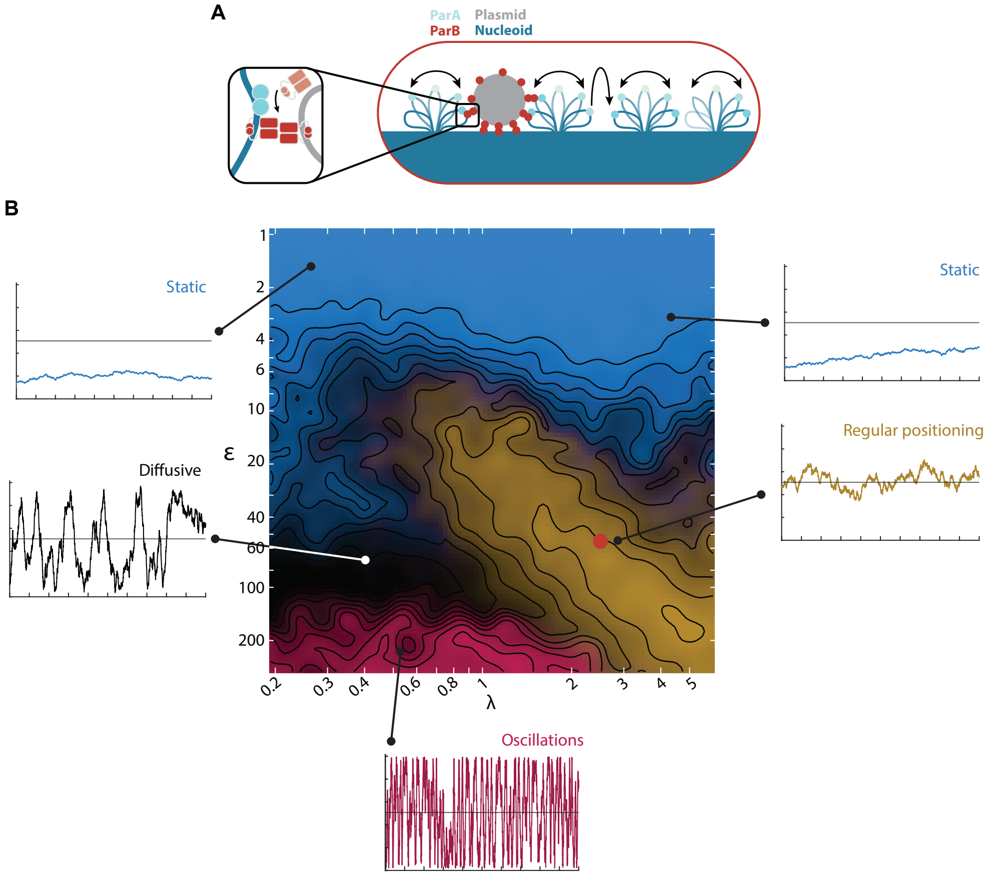


Figure 10.7: ParB-ParB tethers can produce different plasmid behaviours. (a) Schematic of the hopping and relay model with ParB-ParB tethers. Blue shading: nucleoid, blue strokes: DNA strand, light blue: nucleoid bound ParA dimers, red: ParB dimers. Arrows indicate dynamics of the system, chromosome fluctuations and ParA dimer hopping. (b) Phase diagram obtained by varying D_h and k_d, shown in terms of the dimensionless parameters λ and ϵ. As in Figure 10.2 the colour is based on an analysis of simulated trajectories as follows. Yellow: regular positioning, blue: static, pink: oscillations, black: diffusion. Phase diagram is for 500 ParA dimers. Inserts show representative trajectories at the locations indicated on the phase diagram. Red dot marks the location within the regular positioning regime taken to be representative of F plasmid.

positioning of the partition complex. We have shown above that the plasmid moves in the direction of the greatest ParA concentration. Therefore, we suggest that ParA instead acts to modulate tethers which are formed via ParB-ParB interactions. Two possible ways in which this could occur are that ParA could either modulate the lifetime of existing tethers, or ParA could stimulate the formation of tethers.

Here we will test a model whereby ParA acts to promote tether formation. We suggest that this could occur through a transient interaction whereby ParA promotes ParB recruitment onto the nucleoid (Figure 10.7a). This model adapts the scheme of the original hopping and relay model. Tethers between the partition complex and nucleoid are formed at a basal rate, and tether formation can be stimulated when the partition complex comes into contact with a ParA dimer. This results in the unbinding of the ParA dimer from the nucleoid and the instantaneous formation of a tether. Note that this is largely a reinterpretation of the tether forming mechanism of the hopping and relay model, modified such that here the ParA dimer is returned to the cytosolic pool upon tether formation rather than tether breaking. Furthermore, we introduce a limit on the number of ParB tethers, due to the number of ParB dimers within a cell.

ParB-ParB tether model. The model presented here using ParB-ParB tethers is based on the hopping and relay model first developed by Köhler and Murray [62], which is itself an extension of the DNA-relay model [109]. This model is a a 2d off-lattice stochastic model that updates positions in discrete time steps dt. It consists of the following components. ParA associates to the DNA non-specifically in its ATP-dependent dimer state with a rate k_a. Once associated, ParA-ATP dimers can move in two distinct ways. They can either undergo diffusive motion on the nucleoid with diffusion coefficient D_h, this is an effective description of the movement of dimers due to transient unbinding events that allow them to 'hop' between DNA strands, or by elastic fluctuations of the DNA strand that the dimer is bound to. These are implemented as elastic (spring-like) fluctuations around the location where the dimer initially binds.

Dimers can dissociate from the nucleoid due to either basal ATP hydrolysis at a rate k_d, or due to hydrolysis stimulated by ParB on the plasmid. In the latter case, hydrolysis and unbinding are assumed to occur instantaneously upon interaction with the plasmid. The plasmid is modelled as a ParB-coated disc. ParB-ParB tethers are formed at a basal rate k_t, or, instantaneously when the plasmid comes into contact with a ParA dimer, triggering ParA-ATP hydrolysis. ParA dimers are assumed to facilitate the formation of ParB-ParB tethers. Tethers are formed up until a maximum limit of 50 tethers is reached and dissociate with a rate k_h. The plasmid experiences the elastic force of every tether and moves according to its intrinsic diffusion coefficient D_p and the resultant force of all the tethers.

As in the previous models, simplifications are made. First, we only model two states of ParA, nucleoid associated and cytosolic. Second, cytosolic ParA dimers are assumed to be well mixed, this is justified by the slow conformational changes needed to return ParA to a DNA-binding competent state [61]. Third, no ParB dimers are individually modelled, rather the plasmid is treated as a disc coated with enough ParB for 50 tethers to form, and ParB dimers are implicitly assumed to be loaded onto the nucleoid via ParB trans loading as described in Tišma et al. [114], thus allowing ParB-ParB tethers to form between the nucleoid and the partition complex.

The nucleoid is modelled as a rectangle with the dimensions $L \times W$. To model the spring-like nature of the nucleoid we use the distribution $p(x, \delta t | x_0)$

which describes the probability that a Brownian particle, initially at position x_0, experiencing a spring-like force towards 0, is found at a position x at a time δt later [146],

$$p(x, \delta t | x_0) = \sqrt{\frac{f/k_B T}{2\pi S}} \exp\left(-\frac{f/k_B T}{2S}\left(x - x_0 e^{-\delta t/\tau}\right)^2\right), \tag{10.2}$$

where $S = 1 - e^{-2\delta t/\tau}, \tau = \frac{k_B T}{fD}, k_B$ is Boltzmann's constant and T is the absolute temperature. The stiffness of the spring is given by $f/k_B T$ and D is the intrinsic diffusion coefficient. The positions of ParA and the plasmid are updated at every time step dt as follows. Between hopping events, each nucleoid associated ParA dimer fluctuates about its home position x_h. The new position $x(t + dt)$ of each dimer is given by $x(t + dt) = x_h + \delta x$ where δx is drawn with probability $p(\delta x, dt | x(t) - x_h)$ where $x(t)$ is the dimers position at a time t, the normalised spring constant along each dimension is $1/\sigma^2_{x,y}$, and the diffusion coefficient is given by D_a. During hopping events both $x(t)$ and x_h are offset by a value drawn from a Gaussian distribution with $\mu = 0$ and $\sigma = \sqrt{2D_h dt}$ for both dimensions. The displacement of the plasmid is determined according to the resultant force acting upon it. This resultant force vector has a spring constant equal to the spring constant of a single tether multiplied by the number of tethers and acts towards an equilibrium position $x_p(t) + \sum_{i=1}^{n}(x_h - x(t))/i$ where $x_p(t)$ is the plasmid position and n is the number of tethers. The effects of torque are ignored. The intrinsic diffusion coefficient of the plasmid is given by D_p. If the plasmid has no tethers attached then it moves by normal diffusion, with displacements drawn from a Gaussian distribution with $\mu = 0$ and $\sigma = \sqrt{2D_p dt}$. The x and y components of all positions are updated independently and all simulations are run until the system reaches equilibrium before analysis is conducted.

Simulating this model and quantifying the behaviour of plasmid movement we are able to produce a similar phase diagram to that found for the original hopping and relay model (Figure 10.7b). Importantly, all the different behaviours identified before can be seen here: regular positioning, static, oscillatory, and diffusive. These different dynamics are due to the promotion of ParB tethers by ParA which acts to orient the direction of tether formation.

Analysing the MSD at a lag of 30 s across the phase diagram, for different numbers of ParA dimers, it can be seen that the MSD is almost entirely unchanged across the whole phase space (Figure 10.8a). There is only an increase in MSD in the bottom left-hand side of the diagram, corresponding to the oscillatory regime. We suggest that this is because the number of tethers in this region is not saturated. Therefore, there is an increase in the number of ParB-ParB tethers with increasing ParA concentration. ParA causes directed motion in the form of oscillations in this region, so an increasing number of tethers leads to a stronger effect of directed motion on the plasmid. At the 30 s time scale analysed this causes a visible increase in the MSD of the plasmid. In contrast, across the rest of the phase space studied,

Parameter	**Brief description**	**Value**	**Source**
k_a	Association rate to the nucleoid of cytosolic ParA	$0.19\,\mathrm{s}^{-1}$	As in [109]. Based on *in vitro* measurement from [61]. Results in 95% ParA association in the absence of a plasmid.
k_d	Dissociation due to basal hydrolysis rate of ParA	$0.01\,\mathrm{s}^{-1}$	Based on *in vitro* measurement from [105, 147].
k_t	Basal rate of tether formation	$1 \times 10^{-5}\,\mathrm{s}^{-1}$	This study.
k_h	Tether dissociation rate	$0.01\text{-}3\,\mathrm{s}^{-1}$	Swept over in this study, as in [62].
D_p	Diffusion coefficient of the plasmid	$0.003\,\mu\mathrm{m}^2\mathrm{s}^{-1}$	From [109], based on the MSD of a Δpar plasmid.
D_h	Diffusion coefficient of ParA home position on the nucleoid	$3.22 \times 10^{-4}\text{-}0.29\,\mu\mathrm{m}^2\mathrm{s}^{-1}$	Swept over in this study, as in [62].
D_a	Diffusion coefficient of DNA-bound ParA due to chromosomal fluctuations	$0.01\,\mu\mathrm{m}^2\mathrm{s}^{-1}$	As in [109]. Based on [148, 149].
W	Width of the cell	$0.95\,\mu\mathrm{m}$	Average width of a cell from [62].
L	Length of the cell	$2.5384\,\mu\mathrm{m}$	Average length of a cell from [62].
dt	Simulation time step	$0.001\,\mathrm{s}$	As in [62].
R_p	Plasmid radius	$0.05\,\mu\mathrm{m}$	As in [109], estimate from [69].
R_a	ParA radius	$0.002\,\mu\mathrm{m}$	As in [109], based on ParA crystal structure from [91].
σ_x	Width of elastic fluctuations of the chromosome along long cell axis	$0.1\,\mu\mathrm{m}$	As in [109].
σ_y	Width of elastic fluctuations of the chromosome along short cell axis	$0.05\,\mu\mathrm{m}$	As in [109].
n_A	Number of ParA dimers	500	Midrange estimate from [93, 99, 141].

Table 10.1: Parameters for ParB-ParB tether model, based on the parameters for the original hopping and relay model [62].

the number of ParB-ParB tethers is saturated even at the lowest number of ParA dimers. This is because whilst ParA promotes tether formation it is not in the tethers. Therefore, increasing the number of ParA has no effect on the MSD of the plasmid. It should be noted that this does not mean that the number of ParA dimers has no effect on the system. Since ParA promotes tether formation, and due to the limit on the number of tethers which can form, increasing the number of ParA dimers in the system increases the proportion of tethers which are formed via interaction with ParA, compared to those formed via a basal rate. Only ParA formed tethers are able to pass information of the system onto the plasmid and thus result in dynamics such as regular positioning or oscillations. Thus, increasing the number of ParA dimers, increases the input to the plasmid and so the dynamics will become stronger.

To look in more detail at the effect of a ParB-ParB tether model for the case of F plasmid we choose a representative point within the regular positioning regime,

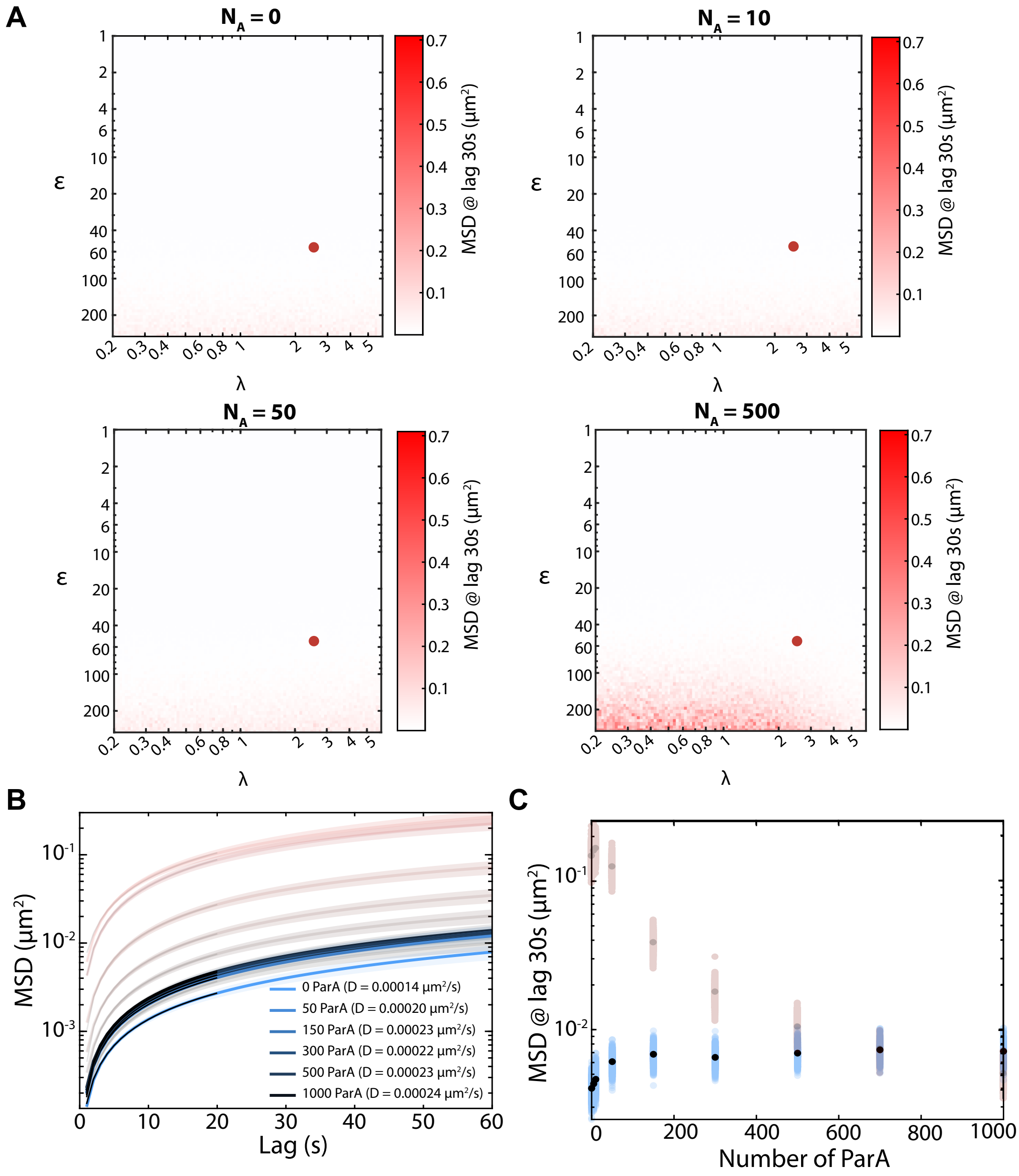


Figure 10.8: ParB tethering as an alternative model. (a) Phase diagrams for the model with ParB-ParB tethers showing the MSD at a lag of 30 s. The number of ParA dimers is varied as 0, 10, 50, 500. The colour bar is such that red signifies an increase in MSD when compared to the mean MSD at 0 ParA dimers. Red dot signifies the location taken for (b) and (c). (c) MSD curves for different numbers of ParA dimers for the ParB-ParB tether model at a representative location with the regular positioning regime for F plasmid. This corresponds to $\lambda = 2.66$ and $\epsilon = 56.42$. Light red shading shows the same result for the hopping and relay model with ParA-ParB tethers. (d) The MSD at a lag of 30 s for a varying number of ParA at the same location as in (c). Light red shading shows the same result for the hopping and relay model with ParA-ParB tethers.

which F plasmid has been found to occupy [62]. As expected, we find that the MSD at a lag of 30 s is unaffected by the number of ParA dimers (Figure 10.8c), and that there is minimal variation in the MSD curves (Figure 10.8b). This suggests that a model of ParB-ParB tether formation may better explain our experimental observations than a model based on ParA-ParB tethers. This model is consistent with both the MSD of the partition complex remaining constant across changing ParA levels and with the lack of observable ParA foci colocalised with the partition complex.

11

DISCUSSION: THE PARABS SYSTEM FROM PARTITION COMPLEX FORMATION TO DNA SEGREGATION

The sliding and bridging models presented in Chapters 7 and 9 use recent discoveries to probe the formation and structure of the partition complex. Recent *in vitro* based studies have shown that, dependent on CTP, ParB can load onto DNA at *parS* sites before sliding randomly along the DNA [71, 83–87]. It was also shown that ParB can efficiently condense DNA, again in the presence of CTP, through the formation of bridges between genomically distant DNA regions [68, 73, 88, 89]. While we have not explicitly modelled the CTP dependent nature of these processes, our model of sliding is consistent with CTP hydrolysis triggering the unbinding of ParB dimers and therefore setting the length scale [71, 84]. Our model predicts that the dynamic sliding and bridging of ParB results in two different conformational regimes, one globular, one structured for long and short bridge lifetimes respectively. The latter regime is dependent on the stiffness of the DNA. If that is ignored, short range bridges between next to neighbouring monomers dominate and DNA structures do not develop. This is consistent with the results of previous studies of chromatin organisation [150, 151] that do not incorporate stiffness and with bridging by much larger molecules [138]. We also showed how the genomic distribution of ParB could define its spatial distribution through the formation of ParB bridges. We then explicitly modelled both the sliding of ParB along the DNA and the formation of ParB-ParB DNA bridges. Importantly, we found that sufficiently short-lived bridges do not hinder sliding of ParB along the DNA and our model could reproduce both the measured genomic and spatial distribution of *C. crescentus* ParB.

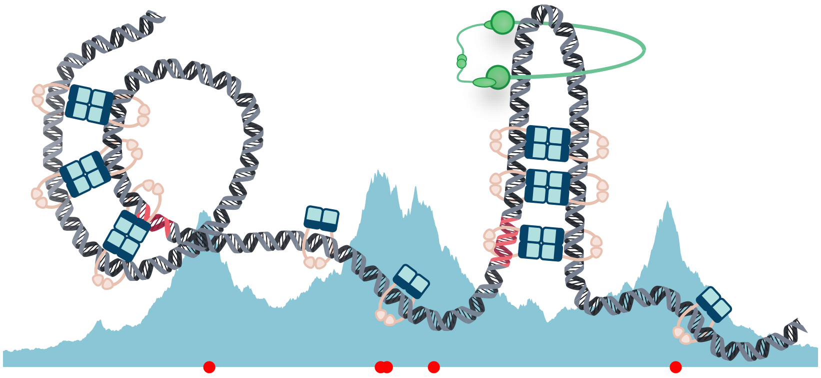


Figure 11.1: A sliding and bridging model can reproduce the genomic and spatial distribution of ParB, forming hairpins and helical structures which organise the DNA. Cartoon of the structure of the partition complex. ParB dimers load at the *parS* sites and then slide along the DNA where they can interact with genomically distant, spatially proximal dimers to form ParB-ParB bridges. These bridges can organise the DNA into hairpin and helical structures. Hairpin structures could potentially be involved in the recruitment of SMC onto the DNA.

Chapter 10 then builds on these results to question the mechanistic details of segregation by the ParAB*S* system, suggesting an alternative model by which ParB-ParB tethers form to segregate the partition complex across the nucleoid. While the prevailing paradigm suggests that ParA-ParB tethers are responsible for anchoring and driving the partition complex movement along the nucleoid, there is currently no experimental evidence for the existence of such tethers. Whilst the individual components have been found to interact, the formation of a continuous bridge, that is DNA-ParA-ParB-DNA, that is stable enough to tether the plasmid has not been shown. In contrast, the presence of ParB-ParB tethers has been experimentally demonstrated [88, 89], and they have been found to compact DNA. This indicates that ParB-ParB tethers are robust under mechanical stress. This along with the recent result that DNA bound ParB dimers can recruit cytosolic ParB dimers to neighbouring stretches of DNA [114] makes the presence of ParB-ParB tethers a compelling proposal.

Looking at our model of partition complex formation, we speculate that the two different regimes could have relevance in different biological contexts. The hairpins and helices of the structured regime may facilitate the loading of SMC (structural maintenance of chromosomes) complexes onto the DNA [152]. While this is known to be due to ParB at the *parS* sites [153], the precise mechanism is a topic of ongoing study [154, 155]. However, ParB mutations that eliminate SMC recruitment are also known to reduce the ability of ParB to form a higher-order nucleoprotein complex [43, 154, 156]. This leads us to postulate that the ParB-induced DNA structures we observe are relevant for the loading of SMC complexes (Figure 11.1). Furthermore, chromosomal ParAB*S* systems often have multiple separated *parS* sites [64] that produce a multi-peaked binding profile [67, 68, 70, 71, 79, 120–122, 157], whereas a single cluster of *parS* sites appears to be more common for plasmid-based systems [52]. Separated *parS* sites would allow the formation of multiple hairpins and could thereby be beneficial for SMC loading. In contrast, ParAB*S* -carrying plasmids, especially those of *E. coli* and other bacteria that do not carry SMC [158], would likely not require these structures. Instead it may be advantageous to form a more compact partition complex to better facilitate the partitioning function of ParAB*S* . Indeed, while F plasmid ParB spreads over a four times larger region than ParB of *C. crescentus* [69], the resultant partition complex is significantly smaller (a radius (2σ) of 35 nm) [81]. Thus, we speculate that plasmid-based ParABS systems may operate in the more compact globular region.

The bacterial chromosome is on average negatively supercoiled [159]. This results in the formation of supercoiled loops (plectonemes) that partition the chromosome into topologically isolated ~10 kb domains [160]. However, *in vitro* DNA experiments and simulations have also detected simple plectonemes in the low kilobase range [161–163]. Thus, these structures, which are topological similar to the hairpins, may be relevant at the 1-2 kb scale of the peaks of the ParB distribution. We expect they would only promote bridge formation due to bringing DNA strands into contact (helices are likely less relevant as they are thermodynamically disfavored compared to plectonemes [164]). Indeed, a previous model of F plasmid partition complex formation argued that supercoiling is required to explain the observed compactness of the partition complex of that system [82].

The conformations we observe in our simulations are similar to those recently seen using atomic force microscopy for *B. subtilis* ParB [89] but more detailed study is required to test our prediction of hairpin structures. Our model could also be better characterised by knowledge of the ParB-ParB bridge lifetime, which could be achieved *in vitro* by using magnetic tweezers to probe the relaxation time of ParB-condensed DNA upon removal of ParB from the buffer. *In vivo* characterisation of the partition complex is more challenging. While our simulated contact maps are in principle comparable to the experimental contact maps produced by chromatin conformation capture (HiC), the resolution of this technique is not yet sufficient to probe DNA structure at the short length scale of the *C. crescentus* centromeric region. This may change as the technique improves [165, 166].

Our modelling of DNA segregation shows that a ParB-ParB tether approach is theoretically feasible. Furthermore, it is consistent with experimental results showing that the mean squared displacement of the plasmid remains constant across varying levels of ParA and that ParA foci are not observed *in vivo*. We provide *in vivo* experimental evidence that plasmids move up a ParA gradient, showing that ParA is necessary to orient tether formation. Crucially, our model is able to reproduce the different behaviours observed from the previous ParA-ParB tether based hopping and relay model.

We primarily support this alternative model with the evidence that the rate of diffusion of the partition complex remains unaltered when varying the concentration of ParA. This is in contrast to what one would expect from the previous ParA-ParB tether model. Due to the technique used to study this phenomena, taking advantage of the natural asymmetry of ParA that occurs within cells, we were able to test this result down to relatively low concentrations of ParA, although not in the complete absence of ParA. This prompts consideration of another plausible scenario: the existence of ParA-ParB tethers, but with a limit on the number of tethers that falls below the numbers of ParA tested experimentally. In such a situation one may expect to also see no change in the plasmid diffusion rate once the number of tethers has saturated. However, we believe we are able to mostly eliminate this scenario. Estimates suggest that there are approximately 360 ParB dimers and 500 ParA dimers in cells [62, 99]. Given that the majority of ParB is localised within the partition complex, there is no reason to anticipate a limitation on the number of tethers below roughly 300. Below this tether threshold we would again expect to see an increase in the MSD of F plasmid. When we compare the fluorescence distribution in cells to the average number of ParA dimers, we expect to have tested numbers of ParA far lower than this and we saw no significant change in the plasmid MSD.

Further experiments could be conducted to confirm the presence or absence of ParA within tethers. For example, employing higher resolution microscopy techniques to look for ParA foci colocalised with the partition complex, which we would expect to observe in the presence of ParA tethers. Furthermore, it would be well advised to revisit previous experiments aimed at breaking down the mechanisms underlying the ParAB*S* system. Many of these experiments were conducted prior to the realisation that CTP is an important cofactor of ParB. Revisiting these experiments in the light of this new knowledge could provide valuable insights. To test our proposed model of ParB-ParB tethers, it could be insightful to revisit the use of DNA carpets. These could be used to test our hypothesis by looking at the movement of the partition complex under conditions with varying components of the ParAB*S* system; measuring the MSD of the partition complex with the full Par system, in the absence of ParA, and in the absence of both ParA and ParB. In the case of ParA-ParB tethers we would expect the MSD under conditions with the full Par system to be different to that of both no ParA and no ParA and ParB. Removing ParA and/or ParB from the system would result in no tether formation. In contrast, if ParB-ParB tethers were present, we would expect the MSD with the full Par system and with no ParA to be comparable. The MSD would only change with both no ParA and no ParB since this is the only condition where tethers would no longer be formed.

In this chapter, we have presented physical models both for the formation of the partition complex and for the mechanism of segregation of the ParAB*S* system. Our dynamical sliding and bridging model reconciles the recent result that ParB spreads along the DNA by sliding like a DNA clamp with the ability of ParB to condense the centromeric region into a nucleoprotein complex. Our model for partition complex segregation then takes this one step further proposing that ParB-ParB interactions are not only needed to structure the partition complex but also for its positioning. Future experimental investigation will be important to evaluate the validity of these models and to test their predictions.

SUMMARY

SUMMARY

In this thesis we have presented mathematical models for the mechanisms underlying two different processes involved in the internal spatial organisation of bacterial cells. Namely, the redistribution of an immobile outer membrane protein as part of the Tol-Pal system, and the segregation of genetic material and low-copy number plasmids by the Par*ABS* system. We take two broadly different approaches for these different systems. In the case of the Tol-Pal system we take a deterministic approach, where we implicitly make the assumption that the stochasticity of the system is averaged out due to the numbers of molecules involved. Whereas when investigating the mechanism of the Par*ABS* system we take a stochastic approach.

In our model of the Tol-Pal system, we have used differential equations to describe the differences observed in the localisation of Pal between dividing and non-dividing cells. Initially we formed a simplified model which we could solve analytically, this illustrated our fundamental result that localised transport is less efficient than homogeneous transport. Building on this principle, we expanded this simple model to provide a more comprehensive representation of the biological system, which we solved numerically. Here, these mathematical models allowed us to show that a 'mobilisation and capture' mechanism could explain experimental measurements of Pal mobility in both dividing and non-dividing cells. Our model was able to reproduce the effects of mutants and deletions using only wild type data for parameter estimation. Furthermore, we were able to make predictions about the mobility of TolB in the inner and outer periplasm and the effects of modulating TolB levels, which we then confirmed experimentally.

Turning to the Par*ABS* system, we formed models both for the formation of the partition complex and for the segregation of DNA. Our model for partition complex formation is inspired by the recent result that CTP is an important cofactor of ParB. This results fundamentally changed our understanding of partition complex formation and led us to test a sliding and bridging model. In order to test this model, we first looked at the effects of ParB sliding independent of the DNA conformation and ParB bridges. To do this we developed a stochastic model for ParB binding, diffusion and unbinding on a 1d lattice based on the Gillespie algorithm. Using this model we were able to reproduce the ParB distribution on the DNA for both *M. xanthus* and *C. crescentus*. Then moving our focus to the action of ParB bridges we developed a kinetic polymer model for the centromeric region of the DNA. This predicted that dependent on the bridge lifetime ParB bridging results in two different conformational regimes, globular for long lived bridges and structured, with the formation of hairpin and helices, for short lived bridges. We showed that the compaction caused by these bridges corresponded to that observed experimentally. We then combined these two models of sliding and bridging to explicitly model the diffusion of ParB dimers along the DNA and the

formation of ParB bridges. This model showed that these two processes are not mutually exclusive, sufficiently short-lived bridges do not hinder ParB sliding. This model could reproduce both the experimentally measured genomic and spatial distributions of ParB.

Building on this result, we propose a new mechanism for DNA segregation by the ParAB*S* system based on the formation of ParB-ParB tethers. Current models of segregation have been largely based on the concept of ParA-ParB tethers. We find a lack of evidence for the presence of these tethers. They have not previously been experimentally shown and ParA foci are not seen *in vivo* colocalised with ParB, as would be expected if many ParA dimers were bound to the partition complex. Furthermore, when we analyse the mean square displacement of the partition complex of F plasmid, we find that this remains constant across a wide range of ParA concentrations. This result is in direct contradiction with what we expect from the ParA-ParB tether based hopping and relay model. In contrast, our model of ParB-ParB tethers is able to reproduce these experimental results whilst still displaying the different dynamics seen in the original hopping and relay model.

Overall, this thesis attempts to delve into the mechanisms behind two fundamental processes of spatial organisation within bacterial cells. We aim to develop models based on biological experiments which describe the molecular mechanisms behind the observable spatial organisation. These models can help in understanding complex systems, allowing an encoding and decoding of reality where natural phenomena can be reduced to numerical expressions. This allows us to better understand how cells function and how organisation arises.

BIBLIOGRAPHY

[1] MR Alley, Janine R Maddock, and Lucille Shapiro. "Polar localization of a bacterial chemoreceptor." In: *Genes & Development* 6.5 (1992), pp. 825–836.

[2] Janine R Maddock and Lucille Shapiro. "Polar location of the chemoreceptor complex in the Escherichia coli cell." In: *Science* 259.5102 (1993), pp. 1717–1723.

[3] Erfei Bi and Joe Lutkenhaus. "FtsZ ring structure associated with division in Escherichia coli." In: *Nature* 354.6349 (1991), pp. 161–164.

[4] David L Williamson. "Unusual fibrils from the spirochete-like sex ratio organism." In: *Journal of bacteriology* 117.2 (1974), pp. 904–906.

[5] Zonglin Hu and Joe Lutkenhaus. "Topological regulation of cell division in Escherichia coli involves rapid pole to pole oscillation of the division inhibitor MinC under the control of MinD and MinE." In: *Molecular microbiology* 34.1 (1999), pp. 82–90.

[6] David M Raskin and Piet AJ De Boer. "Rapid pole-to-pole oscillation of a protein required for directing division to the middle of Escherichia coli." In: *Proceedings of the National Academy of Sciences* 96.9 (1999), pp. 4971–4976.

[7] Philippe Glaser, Michaela E Sharpe, Brian Raether, Marta Perego, Kari Ohlsen, and Jeff Errington. "Dynamic, mitotic-like behavior of a bacterial protein required for accurate chromosome partitioning." In: *Genes & development* 11.9 (1997), pp. 1160–1168.

[8] G Scott Gordon, Dmitry Sitnikov, Chris D Webb, Aurelio Teleman, Aaron Straight, Richard Losick, Andrew W Murray, and Andrew Wright. "Chromosome and low copy plasmid segregation in E. coli: visual evidence for distinct mechanisms." In: *Cell* 90.6 (1997), pp. 1113–1121.

[9] Chris D Webb, Aurelio Teleman, Scott Gordon, Aaron Straight, Andrew Belmont, Daniel Chi-Hong Lin, Alan D Grossman, Andrew Wright, and Richard Losick. "Bipolar localization of the replication origin regions of chromosomes in vegetative and sporulating cells of B. subtilis." In: *Cell* 88.5 (1997), pp. 667–674.

[10] Daniel T Gillespie. "A general method for numerically simulating the stochastic time evolution of coupled chemical reactions." In: *Journal of computational physics* 22.4 (1976), pp. 403–434.

[11] Zvi G Loewy, Ruth A Bryan, Sandra H Reuter, and Lucille Shapiro. "Control of synthesis and positioning of a Caulobacter crescentus flagellar protein." In: *Genes & development* 1.6 (1987), pp. 626–635.

[12] Dennis Bray, Matthew D Levin, and Carl J Morton-Firth. "Receptor clustering as a cellular mechanism to control sensitivity." In: *Nature* 393.6680 (1998), pp. 85–88.

[13] Peter Ames and John S Parkinson. "Conformational suppression of inter-receptor signaling defects." In: *Proceedings of the National Academy of Sciences* 103.24 (2006), pp. 9292–9297.

[14] Victor Sourjik. "Receptor clustering and signal processing in E. coli chemotaxis." In: *Trends in microbiology* 12.12 (2004), pp. 569–576.

[15] Fabrizio Arigoni, Kit Pogliano, Chris D Webb, Patrick Stragier, and Richard Losick. "Localization of protein implicated in establishment of cell type to sites of asymmetric division." In: *Science* 270.5236 (1995), pp. 637–640.

[16] Josefa Steinhauer, Rabia Agha, Thao Pham, Andrew W Varga, and Marcia B Goldberg. "The unipolar Shigella surface protein IcsA is targeted directly to the bacterial old pole: IcsP cleavage of IcsA occurs over the entire bacterial surface." In: *Molecular microbiology* 32.2 (1999), pp. 367–377.

[17] Ivan V Surovtsev and Christine Jacobs-Wagner. "Subcellular organization: a critical feature of bacterial cell replication." In: *Cell* 172.6 (2018), pp. 1271–1293.

[18] Ned S Wingreen and Kerwyn Casey Huang. "Physics of intracellular organization in bacteria." In: *Annual Review of Microbiology* 69 (2015), pp. 361–379.

[19] Paul E Schavemaker, Arnold J Boersma, and Bert Poolman. "How important is protein diffusion in prokaryotes?" In: *Frontiers in molecular biosciences* 5 (2018), p. 93.

[20] Thomas J Silhavy, Daniel Kahne, and Suzanne Walker. "The bacterial cell envelope." In: *Cold Spring Harbor perspectives in biology* 2.5 (2010), a000414.

[21] Daniel A Bonsor, Oliver Hecht, Mireille Vankemmelbeke, Amit Sharma, Anne Marie Krachler, Nicholas G Housden, Katie J Lilly, Richard James, Geoffrey R Moore, and Colin Kleanthous. "Allosteric β-propeller signalling in TolB and its manipulation by translocating colicins." In: *The EMBO journal* 28.18 (2009), pp. 2846–2857.

[22] Alexander JF Egan. "Bacterial outer membrane constriction." In: *Molecular microbiology* 107.6 (2018), pp. 676–687.

[23] Matthew A Gerding, Yasuyuki Ogata, Nicole D Pecora, Hironori Niki, and Piet AJ De Boer. "The trans-envelope Tol–Pal complex is part of the cell division machinery and required for proper outer-membrane invagination during cell constriction in E. coli." In: *Molecular microbiology* 63.4 (2007), pp. 1008–1025.

[24] Mélissa Petiti, Bastien Serrano, Laura Faure, Roland Lloubes, Tâm Mignot, and Denis Duché. "Tol energy-driven localization of Pal and anchoring to the peptidoglycan promote outer-membrane constriction." In: *Journal of molecular biology* 431.17 (2019), pp. 3275–3288.

[25] Joanna Szczepaniak, Cara Press, and Colin Kleanthous. "The multifarious roles of Tol-Pal in Gram-negative bacteria." In: *FEMS Microbiology Reviews* 44.4 (2020), pp. 490–506.

[26] Joanna Szczepaniak, Peter Holmes, Karthik Rajasekar, Renata Kaminska, Firdaus Samsudin, Patrick George Inns, Patrice Rassam, Syma Khalid, Seán M Murray, Christina Redfield, et al. "The lipoprotein Pal stabilises the bacterial outer membrane during constriction by a mobilisation-and-capture mechanism." In: *Nature communications* 11.1 (2020), p. 1305.

[27] A Kolmogoroff, I Petrovsky, and N Piscounoff. "Study of the diffusion equation with growth of the quantity of matter and its application to a biology problem." In: *Dynamics of curved fronts*. Elsevier, 1988, pp. 105–130.

[28] Michael Witty, Carolina Sanz, Amish Shah, J Günter Grossmann, Kenji Mizuguchi, Richard N Perham, and Ben Luisi. "Structure of the periplasmic domain of Pseudomonas aeruginosa TolA: evidence for an evolutionary relationship with the TonB transporter protein." In: *The EMBO journal* 21.16 (2002), pp. 4207–4218.

[29] Nicholas Noinaj, Maude Guillier, Travis J Barnard, and Susan K Buchanan. "TonB-dependent transporters: regulation, structure, and function." In: *Annual review of microbiology* 64 (2010), pp. 43–60.

[30] Gustavo Carrero, Ellen Crawford, Michael J Hendzel, and Gerda de Vries. "Characterizing fluorescence recovery curves for nuclear proteins undergoing binding events." In: *Bulletin of mathematical biology* 66 (2004), pp. 1515–1545.

[31] Brian L Sprague, Robert L Pego, Diana A Stavreva, and James G McNally. "Analysis of binding reactions by fluorescence recovery after photobleaching." In: *Biophysical journal* 86.6 (2004), pp. 3473–3495.

[32] Robert D Turner, Alexander F Hurd, Ashley Cadby, Jamie K Hobbs, and Simon J Foster. "Cell wall elongation mode in Gram-negative bacteria is determined by peptidoglycan architecture." In: *Nature communications* 4.1 (2013), p. 1496.

[33] Gene-Wei Li, David Burkhardt, Carol Gross, and Jonathan S Weissman. "Quantifying absolute protein synthesis rates reveals principles underlying allocation of cellular resources." In: *Cell* 157.3 (2014), pp. 624–635.

[34] Cynthia A Hale, Logan Persons, and Piet AJ de Boer. "Recruitment of the TolA protein to cell constriction sites in Escherichia coli via three separate mechanisms, and a critical role for FtsWI activity in recruitment of both TolA and TolQ." In: *Journal of Bacteriology* 204.1 (2022), e00464–21.

[35] Mohit Kumar, Mario S Mommer, and Victor Sourjik. "Mobility of cytoplasmic, membrane, and DNA-binding proteins in Escherichia coli." In: *Biophysical journal* 98.4 (2010), pp. 552–559.

[36] Grigorios Papadakos, Nicholas G Housden, Katie J Lilly, Renata Kaminska, and Colin Kleanthous. "Kinetic basis for the competitive recruitment of TolB by the intrinsically disordered translocation domain of colicin E9." In: *Journal of molecular biology* 418.5 (2012), pp. 269–280.

[37] Anastasiya A Yakhnina and Thomas G Bernhardt. "The Tol-Pal system is required for peptidoglycan-cleaving enzymes to complete bacterial cell division." In: *Proceedings of the National Academy of Sciences* 117.12 (2020), pp. 6777–6783.

[38] Patrice Rassam, Kathleen R Long, Renata Kaminska, David J Williams, Grigorios Papadakos, Christoph G Baumann, and Colin Kleanthous. "Intermembrane crosstalk drives inner-membrane protein organization in Escherichia coli." In: *Nature communications* 9.1 (2018), p. 1082.

[39] Maya I Anjur-Dietrich, Colm P Kelleher, and Daniel J Needleman. "Mechanical mechanisms of chromosome segregation." In: *Cells* 10.2 (2021), p. 465.

[40] Suckjoon Jun and Bela Mulder. "Entropy-driven spatial organization of highly confined polymers: lessons for the bacterial chromosome." In: *Proceedings of the National Academy of Sciences* 103.33 (2006), pp. 12388–12393.

[41] Hironori Niki, Aline Jaffe, Ryu Imamura, Teru Ogura, and Sota Hiraga. "The new gene mukB codes for a 177 kd protein with coiled-coil domains involved in chromosome partitioning of E. coli." In: *The EMBO journal* 10.1 (1991), pp. 183–193.

[42] Peter L Graumann. "Bacillus subtilis SMC is required for proper arrangement of the chromosome and for efficient segregation of replication termini but not for bipolar movement of newly duplicated origin regions." In: *Journal of bacteriology* 182.22 (2000), pp. 6463–6471.

[43] Stephan Gruber and Jeff Errington. "Recruitment of condensin to replication origin regions by ParB/SpoOJ promotes chromosome segregation in B. subtilis." In: *Cell* 137.4 (2009), pp. 685–696.

[44] Xindan Wang, Hugo B Brandão, Tung BK Le, Michael T Laub, and David Z Rudner. "Bacillus subtilis SMC complexes juxtapose chromosome arms as they travel from origin to terminus." In: *Science* 355.6324 (2017), pp. 524–527.

[45] Daniel Chi-Hong Lin and Alan D Grossman. "Identification and characterization of a bacterial chromosome partitioning site." In: *Cell* 92.5 (1998), pp. 675–685.

[46] Dane A Mohl, Jesse Easter Jr, and James W Gober. "The chromosome partitioning protein, ParB, is required for cytokinesis in Caulobacter crescentus." In: *Molecular microbiology* 42.3 (2001), pp. 741–755.

[47] Esteban Toro, Sun-Hae Hong, Harley H McAdams, and Lucy Shapiro. "Caulobacter requires a dedicated mechanism to initiate chromosome segregation." In: *Proceedings of the National Academy of Sciences* 105.40 (2008), pp. 15435–15440.

[48] Christopher S Campbell and R Dyche Mullins. "In vivo visualization of type II plasmid segregation: bacterial actin filaments pushing plasmids." In: *The Journal of cell biology* 179.5 (2007), pp. 1059–1066.

[49] Sue Lin-Chao, Wen-Tsuan Chen, and Ten-Tsao Wong. "High copy number of the pUC plasmid results from a Rom/Rop-suppressible point mutation in RNA II." In: *Molecular microbiology* 6.22 (1992), pp. 3385–3393.

[50] Samuel Million-Weaver and Manel Camps. "Mechanisms of plasmid segregation: have multicopy plasmids been overlooked?" In: *Plasmid* 75 (2014), pp. 27–36.

[51] Jamie C Baxter and Barbara E Funnell. "Plasmid partition mechanisms." In: *Plasmids: biology and impact in biotechnology and discovery* (2015), pp. 133–155.

[52] Jean-Yves Bouet and Barbara E Funnell. "Plasmid localization and partition in Enterobacteriaceae." In: *EcoSal Plus* 8.2 (2019), pp. 10–1128.

[53] James M Polson and Deanna R-M Kerry. "Segregation of polymers under cylindrical confinement: effects of polymer topology and crowding." In: *Soft Matter* 14.30 (2018), pp. 6360–6373.

[54] James M Polson and Qinxin Zhu. "Free energy and segregation dynamics of two channel-confined polymers of different lengths." In: *Physical Review E* 103.1 (2021), p. 012501.

[55] Stanislau Yatskevich, James Rhodes, and Kim Nasmyth. "Organization of chromosomal DNA by SMC complexes." In: *Annual review of genetics* 53 (2019), pp. 445–482.

[56] Christos Gogou, Aleksandre Japaridze, and Cees Dekker. "Mechanisms for chromosome segregation in bacteria." In: *Frontiers in Microbiology* (2021), p. 1533.

[57] Anthony G Vecchiarelli, Kiyoshi Mizuuchi, and Barbara E Funnell. "Surfing biological surfaces: exploiting the nucleoid for partition and transport in bacteria." In: *Molecular microbiology* 86.3 (2012), pp. 513–523.

[58] Joe Lutkenhaus. "The ParA/MinD family puts things in their place." In: *Trends in microbiology* 20.9 (2012), pp. 411–418.

[59] Barbara E Funnell. "ParB partition proteins: complex formation and spreading at bacterial and plasmid centromeres." In: *Frontiers in molecular biosciences* 3 (2016), p. 44.

[60] Grace E Lim, Alan I Derman, and Joe Pogliano. "Bacterial DNA segregation by dynamic SopA polymers." In: *Proceedings of the National Academy of Sciences* 102.49 (2005), pp. 17658–17663.

[61] Anthony G Vecchiarelli, Yong-Woon Han, Xin Tan, Michiyo Mizuuchi, Rodolfo Ghirlando, Christian Biertümpfel, Barbara E Funnell, and Kiyoshi Mizuuchi. "ATP control of dynamic P1 ParA–DNA interactions: a key role for the nucleoid in plasmid partition." In: *Molecular microbiology* 78.1 (2010), pp. 78–91.

[62] Robin Köhler, Eugen Kaganovitch, and Seán M Murray. "High-throughput imaging and quantitative analysis uncovers the nature of plasmid positioning by ParABS." In: *Elife* 11 (2022), e78743.

[63] STUART Austin, FRANCES Hart, A Abeles, and N Sternberg. "Genetic and physical map of a P1 miniplasmid." In: *Journal of Bacteriology* 152.1 (1982), pp. 63–71.

[64] Jonathan Livny, Yoshiharu Yamaichi, and Matthew K Waldor. "Distribution of centromere-like parS sites in bacteria: insights from comparative genomics." In: *Journal of bacteriology* 189.23 (2007), pp. 8693–8703.

[65] Oleg Rodionov, Małgorzata Łobocka, and Michael Yarmolinsky. "Silencing of genes flanking the P1 plasmid centromere." In: *Science* 283.5401 (1999), pp. 546–549.

[66] Heath Murray, Henrique Ferreira, and Jeff Errington. "The bacterial chromosome segregation protein Spo0J spreads along DNA from parS nucleation sites." In: *Molecular microbiology* 61.5 (2006), pp. 1352–1361.

[67] Adam M Breier and Alan D Grossman. "Whole-genome analysis of the chromosome partitioning and sporulation protein Spo0J (ParB) reveals spreading and origin-distal sites on the Bacillus subtilis chromosome." In: *Molecular microbiology* 64.3 (2007), pp. 703–718.

[68] Thomas GW Graham, Xindan Wang, Dan Song, Candice M Etson, Antoine M van Oijen, David Z Rudner, and Joseph J Loparo. "ParB spreading requires DNA bridging." In: *Genes & development* 28.11 (2014), pp. 1228–1238.

[69] Aurore Sanchez, Diego I Cattoni, Jean-Charles Walter, Jérôme Rech, Andrea Parmeggiani, Marcelo Nollmann, and Jean-Yves Bouet. "Stochastic self-assembly of ParB proteins builds the bacterial DNA segregation apparatus." In: *Cell systems* 1.2 (2015), pp. 163–173.

[70] Ngat T Tran, Clare E Stevenson, Nicolle F Som, Anyarat Thanapipatsiri, Adam S B Jalal, and Tung B K Le. "Permissive zones for the centromere-binding protein ParB on the Caulobacter crescentus chromosome." In: *Nucleic acids research* 46.3 (2018), pp. 1196–1209.

[71] Manuel Osorio-Valeriano, Florian Altegoer, Chandan K Das, Wieland Steinchen, Gaël Panis, Lara Connolley, Giacomo Giacomelli, Helge Feddersen, Laura Corrales-Guerrero, Pietro I Giammarinaro, et al. "The CTPase activity of ParB determines the size and dynamics of prokaryotic DNA partition complexes." In: *Molecular cell* 81.19 (2021), pp. 3992–4007.

[72] Thomas A Leonard, P Jonathan G Butler, and Jan Löwe. "Structural analysis of the chromosome segregation protein Spo0J from Thermus thermophilus." In: *Molecular microbiology* 53.2 (2004), pp. 419–432.

[73] James A Taylor, Cesar L Pastrana, Annika Butterer, Christian Pernstich, Emma J Gwynn, Frank Sobott, Fernando Moreno-Herrero, and Mark S Dillingham. "Specific and non-specific interactions of ParB with DNA: implications for chromosome segregation." In: *Nucleic acids research* 43.2 (2015), pp. 719–731.

[74] Bo-Wei Chen, Ming-Hsing Lin, Chen-Hsi Chu, Chia-En Hsu, and Yuh-Ju Sun. "Insights into ParB spreading from the complex structure of Spo0J and parS." In: *Proceedings of the National Academy of Sciences* 112.21 (2015), pp. 6613–6618.

[75] Maria A Schumacher and Barbara E Funnell. "Structures of ParB bound to DNA reveal mechanism of partition complex formation." In: *Nature* 438.7067 (2005), pp. 516–519.

[76] Gemma LM Fisher, Cesar L Pastrana, Victoria A Higman, Alan Koh, James A Taylor, Annika Butterer, Timothy Craggs, Frank Sobott, Heath Murray, Matthew P Crump, et al. "The structural basis for dynamic DNA binding and bridging interactions which condense the bacterial centromere." In: *Elife* 6 (2017), e28086.

[77] Julene Madariaga-Marcos, Cesar L Pastrana, Gemma LM Fisher, Mark Simon Dillingham, and Fernando Moreno-Herrero. “ParB dynamics and the critical role of the CTD in DNA condensation unveiled by combined force-fluorescence measurements.” In: *Elife* 8 (2019), e43812.

[78] Chase P Broedersz, Xindan Wang, Yigal Meir, Joseph J Loparo, David Z Rudner, and Ned S Wingreen. “Condensation and localization of the partitioning protein ParB on the bacterial chromosome.” In: *Proceedings of the National Academy of Sciences* 111.24 (2014), pp. 8809–8814.

[79] Roxanne E Debaugny, Aurore Sanchez, Jérôme Rech, Delphine Labourdette, Jérôme Dorignac, Frédéric Geniet, John Palmeri, Andrea Parmeggiani, François Boudsocq, Véronique Anton Leberre, et al. “A conserved mechanism drives partition complex assembly on bacterial chromosomes and plasmids.” In: *Molecular systems biology* 14.11 (2018), e8516.

[80] Gabriel David, J-C Walter, Chase P Broedersz, Jerome Dorignac, Frederic Geniet, Andrea Parmeggiani, N-O Walliser, and John Palmeri. “Phase separation of polymer-bound particles induced by loop-mediated one dimensional effective long-range interactions.” In: *Physical Review Research* 2.3 (2020), p. 033377.

[81] Baptiste Guilhas, Jean-Charles Walter, Jerome Rech, Gabriel David, Nils Ole Walliser, John Palmeri, Celine Mathieu-Demaziere, Andrea Parmeggiani, Jean-Yves Bouet, Antoine Le Gall, et al. “ATP-driven separation of liquid phase condensates in bacteria.” In: *Molecular Cell* 79.2 (2020), pp. 293–303.

[82] Jean-Charles Walter, Thibaut Lepage, Jérôme Dorignac, Frédéric Geniet, Andrea Parmeggiani, John Palmeri, Jean-Yves Bouet, and Ivan Junier. “Supercoiled DNA and non-equilibrium formation of protein complexes: A quantitative model of the nucleoprotein ParB S partition complex.” In: *PLoS Computational Biology* 17.4 (2021), e1008869.

[83] Manuel Osorio-Valeriano, Florian Altegoer, Wieland Steinchen, Svenja Urban, Ying Liu, Gert Bange, and Martin Thanbichler. “ParB-type DNA segregation proteins are CTP-dependent molecular switches.” In: *Cell* 179.7 (2019), pp. 1512–1524.

[84] Young-Min Soh, Iain Finley Davidson, Stefano Zamuner, Jérôme Basquin, Florian Patrick Bock, Michael Taschner, Jan-Willem Veening, Paolo De Los Rios, Jan-Michael Peters, and Stephan Gruber. “Self-organization of parS centromeres by the ParB CTP hydrolase.” In: *Science* 366.6469 (2019), pp. 1129–1133.

[85] Hammam Antar, Young-Min Soh, Stefano Zamuner, Florian P Bock, Anna Anchimiuk, Paolo De Los Rios, and Stephan Gruber. “Relief of ParB autoinhibition by parS DNA catalysis and recycling of ParB by CTP hydrolysis promote bacterial centromere assembly.” In: *Science advances* 7.41 (2021), eabj2854.

[86] Adam SB Jalal, Ngat T Tran, and Tung BK Le. “ParB spreading on DNA requires cytidine triphosphate in vitro.” In: *Elife* 9 (2020), e53515.

[87] Adam SB Jalal, Ngat T Tran, Clare EM Stevenson, Afroze Chimthanawala, Anjana Badrinarayanan, David M Lawson, and Tung BK Le. “A CTP-dependent gating mechanism enables ParB spreading on DNA.” In: *Elife* 10 (2021), e69676.

[88] James A Taylor, Yeonee Seol, Jagat Budhathoki, Keir C Neuman, and Kiyoshi Mizuuchi. “CTP and parS coordinate ParB partition complex dynamics and ParA-ATPase activation for ParABS-mediated DNA partitioning.” In: *Elife* 10 (2021), e65651.

[89] Francisco de Asis Balaguer, Clara Aicart-Ramos, Gemma LM Fisher, Sara de Bragança, Eva M Martin-Cuevas, Cesar L Pastrana, Mark Simon Dillingham, and Fernando Moreno-Herrero. "CTP promotes efficient ParB-dependent DNA condensation by facilitating one-dimensional diffusion from parS." In: *Elife* 10 (2021), e67554.

[90] Eugene V Koonin. "A superfamily of ATPases with diverse functions containing either classical or deviant ATP-binding motif." In: *Journal of molecular biology* 229.4 (1993), pp. 1165–1174.

[91] Thomas A Leonard, P Jonathan Butler, and Jan Löwe. "Bacterial chromosome segregation: structure and DNA binding of the Soj dimer—a conserved biological switch." In: *The EMBO journal* 24.2 (2005), pp. 270–282.

[92] Gitte Ebersbach, Simon Ringgaard, Jakob Møller-Jensen, Qing Wang, David J Sherratt, and Kenn Gerdes. "Regular cellular distribution of plasmids by oscillating and filament-forming ParA ATPase of plasmid pB171." In: *Molecular microbiology* 61.6 (2006), pp. 1428–1442.

[93] Shun Adachi, Kotaro Hori, and Sota Hiraga. "Subcellular positioning of F plasmid mediated by dynamic localization of SopA and SopB." In: *Journal of molecular biology* 356.4 (2006), pp. 850–863.

[94] Michael A Fogel and Matthew K Waldor. "A dynamic, mitotic-like mechanism for bacterial chromosome segregation." In: *Genes & development* 20.23 (2006), pp. 3269–3282.

[95] Toshiyuki Hatano, Yoshiharu Yamaichi, and Hironori Niki. "Oscillating focus of SopA associated with filamentous structure guides partitioning of F plasmid." In: *Molecular microbiology* 64.5 (2007), pp. 1198–1213.

[96] Florencia Pratto, Aslan Cicek, Wilhelm A Weihofen, Rudi Lurz, Wolfram Saenger, and Juan C Alonso. "Streptococcus pyogenes pSM19035 requires dynamic assembly of ATP-bound ParA and ParB on parS DNA during plasmid segregation." In: *Nucleic acids research* 36.11 (2008), pp. 3676–3689.

[97] Simon Ringgaard, Jeroen van Zon, Martin Howard, and Kenn Gerdes. "Movement and equipositioning of plasmids by ParA filament disassembly." In: *Proceedings of the National Academy of Sciences* 106.46 (2009), pp. 19369–19374.

[98] Jerod L Ptacin, Steven F Lee, Ethan C Garner, Esteban Toro, Michael Eckart, Luis R Comolli, WE Moerner, and Lucy Shapiro. "A spindle-like apparatus guides bacterial chromosome segregation." In: *Nature cell biology* 12.8 (2010), pp. 791–798.

[99] Hoong Chuin Lim, Ivan Vladimirovich Surovtsev, Bruno Gabriel Beltran, Fang Huang, Jörg Bewersdorf, and Christine Jacobs-Wagner. "Evidence for a DNA-relay mechanism in ParABS-mediated chromosome segregation." In: *elife* 3 (2014), e02758.

[100] Anthony G Vecchiarelli, Keir C Neuman, and Kiyoshi Mizuuchi. "A propagating ATPase gradient drives transport of surface-confined cellular cargo." In: *Proceedings of the National Academy of Sciences* 111.13 (2014), pp. 4880–4885.

[101] Antoine Le Gall, Diego I Cattoni, Baptiste Guilhas, Céline Mathieu-Demazière, Laura Oudjedi, Jean-Bernard Fiche, Jérôme Rech, Sara Abrahamsson, Heath Murray, Jean-Yves Bouet, et al. "Bacterial partition complexes segregate within the volume of the nucleoid." In: *Nature communications* 7.1 (2016), p. 12107.

[102] Hengshan Zhang and Maria A Schumacher. "Structures of partition protein ParA with nonspecific DNA and ParB effector reveal molecular insights into principles governing Walker-box DNA segregation." In: *Genes & development* 31.5 (2017), pp. 481–492.

[103] Laura Corrales-Guerrero, Binbin He, Yacine Refes, Gaël Panis, Gert Bange, Patrick H Viollier, Wieland Steinchen, and Martin Thanbichler. "Molecular architecture of the DNA-binding sites of the P-loop ATPases MipZ and ParA from Caulobacter crescentus." In: *Nucleic acids research* 48.9 (2020), pp. 4769–4779.

[104] Chen-Hsi Chu, Cheng-Yi Yen, Bo-Wei Chen, Min-Guan Lin, Lyu-Han Wang, Kai-Zhi Tang, Chwan-Deng Hsiao, and Yuh-Ju Sun. "Crystal structures of Hp Soj–DNA complexes and the nucleoid-adaptor complex formation in chromosome segregation." In: *Nucleic Acids Research* 47.4 (2019), pp. 2113–2129.

[105] Anthony G Vecchiarelli, Ling Chin Hwang, and Kiyoshi Mizuuchi. "Cell-free study of F plasmid partition provides evidence for cargo transport by a diffusion-ratchet mechanism." In: *Proceedings of the National Academy of Sciences* 110.15 (2013), E1390–E1397.

[106] Daniela Barillà, Emma Carmelo, and Finbarr Hayes. "The tail of the ParG DNA segregation protein remodels ParF polymers and enhances ATP hydrolysis via an arginine finger-like motif." In: *Proceedings of the National Academy of Sciences* 104.6 (2007), pp. 1811–1816.

[107] Takeshi Sugawara and Kunihiko Kaneko. "Chemophoresis as a driving force for intracellular organization: Theory and application to plasmid partitioning." In: *Biophysics* 7 (2011), pp. 77–88.

[108] Robert Ietswaart, Florian Szardenings, Kenn Gerdes, and Martin Howard. "Competing ParA structures space bacterial plasmids equally over the nucleoid." In: *PLoS computational biology* 10.12 (2014), e1004009.

[109] Ivan V Surovtsev, Manuel Campos, and Christine Jacobs-Wagner. "DNA-relay mechanism is sufficient to explain ParA-dependent intracellular transport and patterning of single and multiple cargos." In: *Proceedings of the National Academy of Sciences* 113.46 (2016), E7268–E7276.

[110] Longhua Hu, Anthony G Vecchiarelli, Kiyoshi Mizuuchi, Keir C Neuman, and Jian Liu. "Directed and persistent movement arises from mechanochemistry of the ParA/ParB system." In: *Proceedings of the National Academy of Sciences* 112.51 (2015), E7055–E7064.

[111] Longhua Hu, Anthony G Vecchiarelli, Kiyoshi Mizuuchi, Keir C Neuman, and Jian Liu. "Brownian ratchet mechanism for faithful segregation of low-copy-number plasmids." In: *Biophysical journal* 112.7 (2017), pp. 1489–1502.

[112] Longhua Hu, Anthony G Vecchiarelli, Kiyoshi Mizuuchi, Keir C Neuman, and Jian Liu. "Brownian ratchet mechanisms of ParA-mediated partitioning." In: *Plasmid* 92 (2017), p. 12.

[113] Andrea Volante, Juan Carlos Alonso, and Kiyoshi Mizuuchi. "Distinct architectural requirements for the parS centromeric sequence of the pSM19035 plasmid partition machinery." In: *Elife* 11 (2022), e79480.

[114] Miloš Tišma, Maria Panoukidou, Hammam Antar, Young-Min Soh, Roman Barth, Biswajit Pradhan, Anders Barth, Jaco van der Torre, Davide Michieletto, Stephan Gruber, et al. "ParB proteins can bypass DNA-bound roadblocks via dimer-dimer recruitment." In: *Science Advances* 8.26 (2022), eabn3299.

[115] Linda JS Allen. *An introduction to stochastic processes with applications to biology*. Pearson Education Inc., 2003.

[116] Mark EJ Newman and Gerard T Barkema. *Monte Carlo methods in statistical physics*. Clarendon Press, 1999.

[117] Daniel T Gillespie. “Stochastic simulation of chemical kinetics.” In: *Annu. Rev. Phys. Chem.* 58 (2007), pp. 35–55.

[118] Daniel T Gillespie, Andreas Hellander, and Linda R Petzold. “Perspective: Stochastic algorithms for chemical kinetics.” In: *The Journal of chemical physics* 138.17 (2013).

[119] Jean-Yves Bouet and David Lane. “Molecular basis of the supercoil deficit induced by the mini-F plasmid partition complex.” In: *Journal of Biological Chemistry* 284.1 (2009), pp. 165–173.

[120] Valentine Lagage, Frédéric Boccard, and Isabelle Vallet-Gely. “Regional control of chromosome segregation in Pseudomonas aeruginosa.” In: *PLoS genetics* 12.11 (2016), e1006428.

[121] Magdalena Donczew, Paweł Mackiewicz, Agnieszka Wróbel, Klas Flärdh, Jolanta Zakrzewska-Czerwińska, and Dagmara Jakimowicz. “ParA and ParB coordinate chromosome segregation with cell elongation and division during Streptomyces sporulation.” In: *Open biology* 6.4 (2016), p. 150263.

[122] Kati Böhm, Giacomo Giacomelli, Andreas Schmidt, Axel Imhof, Romain Koszul, Martial Marbouty, and Marc Bramkamp. “Chromosome organization by a conserved condensin-ParB system in the actinobacterium Corynebacterium glutamicum.” In: *Nature communications* 11.1 (2020), pp. 1–17.

[123] Johan Elf, Gene-Wei Li, and X Sunney Xie. “Probing transcription factor dynamics at the single-molecule level in a living cell.” In: *Science* 316.5828 (2007), pp. 1191–1194.

[124] Izrail Solomonovich Gradshteyn and Iosif Moiseevich Ryzhik. *Table of integrals, series, and products*. Academic press, 2014.

[125] Lakshmi Pillai-Kastoori, Amy R Schutz-Geschwender, and Jeff A Harford. “A systematic approach to quantitative Western blot analysis.” In: *Analytical biochemistry* 593 (2020), p. 113608.

[126] Jean-Charles Walter, Jerome Rech, Nils-Ole Walliser, Jerome Dorignac, Frederic Geniet, John Palmeri, Andrea Parmeggiani, and Jean-Yves Bouet. “Physical modeling of a sliding clamp mechanism for the spreading of ParB at short genomic distance from bacterial centromere sites.” In: *Iscience* 23.12 (2020), p. 101861.

[127] Pierre-Gilles De Gennes. *Scaling concepts in polymer physics*. Cornell university press, 1979.

[128] M Rubinstein and RH Colby. *Polymer Physics*. Oxford University Press, 2003.

[129] Iwao Teraoka. *Polymer Solutions. An Introduction to Physical Properties*. Wiley, 2002.

[130] I Carmesin and Kurt Kremer. “The bond fluctuation method: a new effective algorithm for the dynamics of polymers in all spatial dimensions.” In: *Macromolecules* 21.9 (1988), pp. 2819–2823.

[131] Jing-Zi Zhang, Xiang-Yao Peng, Shan Liu, Bang-Ping Jiang, Shi-Chen Ji, and Xing-Can Shen. “The persistence length of semiflexible polymers in lattice Monte Carlo simulations.” In: *Polymers* 11.2 (2019), p. 295.

[132] Manfred Bohn and Dieter W Heermann. “Diffusion-driven looping provides a consistent framework for chromatin organization.” In: *PloS one* 5.8 (2010), e12218.

[133] Annaël Brunet, Catherine Tardin, Laurence Salomé, Philippe Rousseau, Nicolas Destainville, and Manoel Manghi. “Dependence of DNA persistence length on ionic strength of solutions with monovalent and divalent salts: a joint theory–experiment study.” In: *Biophysical Journal* 110.3 (2016), 403a.

[134] Richard E Dickerson, Horace R Drew, Benjamin N Conner, Richard M Wing, Albert V Fratini, and Mary L Kopka. “The anatomy of A-, B-, and Z-DNA.” In: *Science* 216.4545 (1982), pp. 475–485.

[135] Robert S Hoy and Glenn H Fredrickson. “Thermoreversible associating polymer networks. I. Interplay of thermodynamics, chemical kinetics, and polymer physics.” In: *The Journal of chemical physics* 131.22 (2009).

[136] Michael Rubinstein and Alexander N Semenov. “Thermoreversible gelation in solutions of associating polymers. 2. Linear dynamics.” In: *Macromolecules* 31.4 (1998), pp. 1386–1397.

[137] Sharon M Loverde, Alexander V Ermoshkin, and Monica Olvera De La Cruz. “Thermodynamics of reversibly associating ideal chains.” In: *Journal of Polymer Science Part B: Polymer Physics* 43.7 (2005), pp. 796–804.

[138] Chris A Brackley, Stephen Taylor, Argyris Papantonis, Peter R Cook, and Davide Marenduzzo. “Nonspecific bridging-induced attraction drives clustering of DNA-binding proteins and genome organization.” In: *Proceedings of the National Academy of Sciences* 110.38 (2013), E3605–E3611.

[139] Julene Madariaga-Marcos, Cesar L Pastrana, Gemma LM Fisher, Mark Simon Dillingham, and Fernando Moreno-Herrero. “ParB dynamics and the critical role of the CTD in DNA condensation unveiled by combined force-fluorescence measurements.” In: *Elife* 8 (2019), e43812.

[140] Dan Song, Kristen Rodrigues, Thomas GW Graham, and Joseph J Loparo. “A network of cis and trans interactions is required for ParB spreading.” In: *Nucleic Acids Research* 45.12 (2017), pp. 7106–7117.

[141] Jean-Yves Bouet, Jérôme Rech, Sylvain Egloff, Donald P Biek, and David Lane. “Probing plasmid partition with centromere-based incompatibility.” In: *Molecular microbiology* 55.2 (2005), pp. 511–525.

[142] Yoan Ah-Seng, Jérôme Rech, David Lane, and Jean-Yves Bouet. “Defining the role of ATP hydrolysis in mitotic segregation of bacterial plasmids.” In: *PLoS genetics* 9.12 (2013), e1003956.

[143] Gitte Ebersbach and Kenn Gerdes. “The double par locus of virulence factor pB171: DNA segregation is correlated with oscillation of ParA.” In: *Proceedings of the National Academy of Sciences* 98.26 (2001), pp. 15078–15083.

[144] Gitte Ebersbach and Kenn Gerdes. “Bacterial mitosis: partitioning protein ParA oscillates in spiral-shaped structures and positions plasmids at mid-cell.” In: *Molecular microbiology* 52.2 (2004), pp. 385–398.

[145] Robin Köhler and Seán M. Murray. “Putting the Par back into ParABS: Plasmid Partitioning Driven by ParA Oscillations.” In: *bioRxiv* (2023).

[146] Masao Doi and Samuel Frederick Edwards. *The theory of polymer dynamics*. Vol. 73. oxford university press, 1988.

[147] Ling Chin Hwang, Anthony G Vecchiarelli, Yong-Woon Han, Michiyo Mizuuchi, Yoshie Harada, Barbara E Funnell, and Kiyoshi Mizuuchi. “ParA-mediated plasmid partition driven by protein pattern self-organization.” In: *The EMBO journal* 32.9 (2013), pp. 1238–1249.

[148] Stephanie C Weber, Andrew J Spakowitz, and Julie A Theriot. “Bacterial chromosomal loci move subdiffusively through a viscoelastic cytoplasm.” In: *Physical review letters* 104.23 (2010), p. 238102.

[149] Avelino Javer, Nathan J Kuwada, Zhicheng Long, Vincenzo G Benza, Kevin D Dorfman, Paul A Wiggins, Pietro Cicuta, and Marco Cosentino Lagomarsino. “Persistent super-diffusive motion of Escherichia coli chromosomal loci.” In: *Nature communications* 5.1 (2014), p. 3854.

[150] Mario Nicodemi and Ana Pombo. "Models of chromosome structure." In: *Current opinion in cell biology* 28 (2014), pp. 90–95.

[151] Mariano Barbieri, Mita Chotalia, James Fraser, Liron-Mark Lavitas, Josée Dostie, Ana Pombo, and Mario Nicodemi. "Complexity of chromatin folding is captured by the strings and binders switch model." In: *Proceedings of the National Academy of Sciences* 109.40 (2012), pp. 16173–16178.

[152] Stephan Gruber. "SMC complexes sweeping through the chromosome: going with the flow and against the tide." In: *Current opinion in microbiology* 42 (2018), pp. 96–103.

[153] Anita Minnen, Laetitia Attaiech, Maria Thon, Stephan Gruber, and Jan-Willem Veening. "SMC is recruited to oriC by ParB and promotes chromosome segregation in Streptococcus pneumoniae." In: *Molecular microbiology* 81.3 (2011), pp. 676–688.

[154] Ngat T Tran, Michael T Laub, and Tung BK Le. "SMC progressively aligns chromosomal arms in Caulobacter crescentus but is antagonized by convergent transcription." In: *Cell reports* 20.9 (2017), pp. 2057–2071.

[155] Florian P Bock, Hon Wing Liu, Anna Anchimiuk, Marie-Laure Diebold-Durand, and Stephan Gruber. "A joint-ParB interface promotes SMC DNA recruitment." In: *Cell reports* 40.9 (2022), p. 111273.

[156] Adam SB Jalal and Tung BK Le. "Bacterial chromosome segregation by the ParABS system." In: *Open biology* 10.6 (2020), p. 200097.

[157] Jong Hwan Baek, Seesandra V Rajagopala, and Dhruba K Chattoraj. "Chromosome segregation proteins of Vibrio cholerae as transcription regulators." In: *MBio* 5.3 (2014), e01061–14.

[158] Sota Hiraga, Chiyome Ichinose, Toshinari Onogi, Hironori Niki, and Mitsuyoshi Yamazoe. "Bidirectional migration of SeqA-bound hemimethylated DNA clusters and pairing of oriC copies in Escherichia coli." In: *Genes to Cells* 5.5 (2000), pp. 327–341.

[159] Andrew Travers and Georgi Muskhelishvili. "Bacterial chromatin." In: *Current opinion in genetics & development* 15.5 (2005), pp. 507–514.

[160] Lisa Postow, Christine D. Hardy, Javier Arsuaga, and Nicholas R. Cozzarelli. "Topological domain structure of the Escherichia coli chromosome." In: *Genes & Development* 18.14 (2004), pp. 1766–1779.

[161] Aleksandre Japaridze, Georgi Muskhelishvili, Fabrizio Benedetti, Agni FM Gavriilidou, Renato Zenobi, Paolo De Los Rios, Giovanni Longo, and Giovanni Dietler. "Hyperplectonemes: a higher order compact and dynamic DNA self-organization." In: *Nano letters* 17.3 (2017), pp. 1938–1948.

[162] M. T. J. van Loenhout, M. V. de Grunt, and C. Dekker. "Dynamics of DNA Supercoils." In: *Science* 338 (2012), pp. 94–97.

[163] Thibaut Lepage and Ivan Junier. "A polymer model of bacterial supercoiled DNA including structural transitions of the double helix." In: *Physica A: Statistical Mechanics and its Applications* 527 (2019), p. 121196.

[164] John F. Marko and Eric D. Siggia. "Fluctuations and Supercoiling of DNA." In: *Science* 265.5171 (1994). Publisher: American Association for the Advancement of Science, pp. 506–508.

[165] Yan Zhang, Lin An, Jie Xu, Bo Zhang, W Jim Zheng, Ming Hu, Jijun Tang, and Feng Yue. "Enhancing Hi-C data resolution with deep convolutional neural network HiCPlus." In: *Nature communications* 9.1 (2018), pp. 1–9.