

## In vivo investigations of *E. coli* chromosomal replication using single-molecule imaging

Tiruvadi Krishnan, S

**DOI**

[10.4233/uuid:9a7eddec-5e49-4268-bbcd-5f58977b3f11](https://doi.org/10.4233/uuid:9a7eddec-5e49-4268-bbcd-5f58977b3f11)

**Publication date**

2016

**Document Version**

Final published version

**Citation (APA)**

Tiruvadi Krishnan, S. (2016). *In vivo investigations of E. coli chromosomal replication using single-molecule imaging*. [Dissertation (TU Delft), Delft University of Technology]. <https://doi.org/10.4233/uuid:9a7eddec-5e49-4268-bbcd-5f58977b3f11>

**Important note**

To cite this publication, please use the final published version (if applicable).  
Please check the document version above.

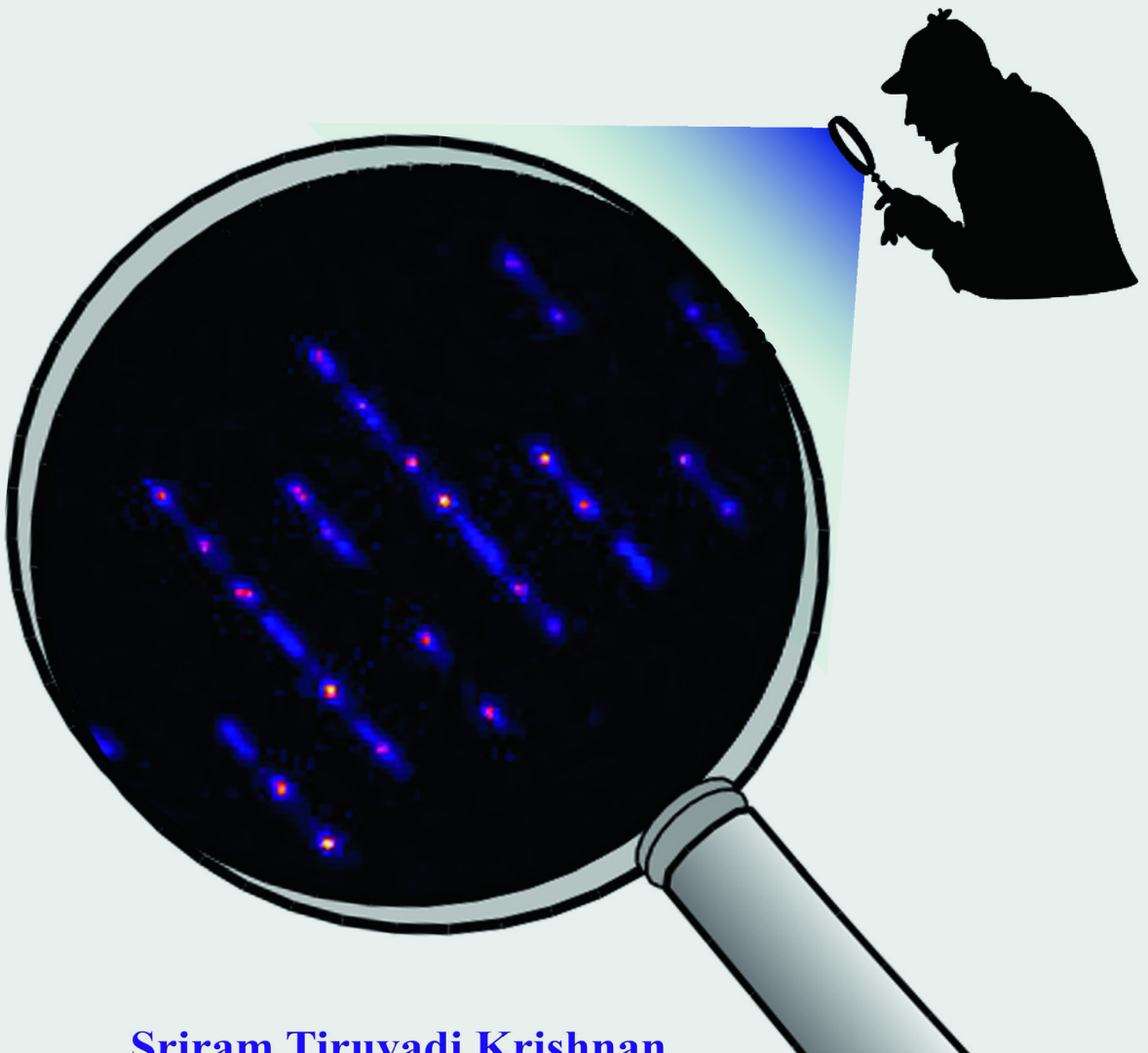
**Copyright**

Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

**Takedown policy**

Please contact us and provide details if you believe this document breaches copyrights.  
We will remove access to the work immediately and investigate your claim.

*In vivo* investigations of  
*E. coli* chromosomal replication  
using single-molecule imaging



**Sriram Tiruvadi Krishnan**



## **Propositions**

accompanying the dissertation

### ***In vivo* investigations of *E. coli* chromosomal replication using single-molecule imaging**

by

Sriram TIRUVADI KRISHNAN

1. Chromosomally engineered *E. coli* strains require further validation beyond simply PCR and DNA sequencing prior to their use in research [Chapter 2].
2. During DNA replication, the sliding clamps are actively loaded and unloaded from DNA, and their accumulation facilitates the recruitment of additional requisite proteins to DNA [Chapter 3].
3. An advancing replication fork is only hindered, and not stopped indefinitely, by the endogenous Tus-*Ter* roadblocks during DNA replication [Chapter 4].
4. The endogenous Tus proteins are expressed in low copy numbers of around eight to thirteen molecules per cell and are bound to the chromosomal *Ter* sites distributed throughout the cytoplasmic space in growing cells [Chapter 5].
5. The structure of its operon complicates the labeling of the native primase protein, requiring the development of advanced genome editing tools with higher efficiency [Chapter 6].
6. As both nature (genetic) and nurture (environmental) factors contribute significantly to the behavior of a living organism, it is a fallacy to question which of these factors plays the dominant role.
7. The responsibility of improving the reproducibility of scientific research lies not only on the researcher but also with each of the stakeholders involved [C. Glenn Begley et al., *Circ Res*, 2015].
8. At all levels, the educational system should be based on principles of cooperation rather than competition.
9. Very short, infrequent mental breaks definitely improve one's performance of a task; however, given their addictive nature, the number of such breaks must be handled with caution to avoid excessive procrastination.
10. The highest activity a human being can attain is learning for understanding, because to understand is to be free [Baruch Spinoza, *Ethica*, 1677].

These propositions are regarded as opposable and defensible, and have been approved as such by the supervisor prof.dr.N.H.Dekker.

## Stellingen

behorende bij het proefschrift

### ***In vivo* investigations of *E. coli* chromosomal replication using single-molecule imaging**

door

Sriram TIRUVADI KRISHNAN

1. Genetisch gemanipuleerde *E. coli* stammen moeten grondiger gevalideerd worden dan alleen door middel van PCR en DNA sequentie bepaling, vóórdat ze gebruikt kunnen worden in wetenschappelijk onderzoek [Hoofdstuk 2].
2. Tijdens DNA replicatie worden de ‘sliding clamp’ eiwitten actief geladen en gelost van het DNA, en deze accumulatie bevordert de rekrutering van additionele vereiste eiwitten op het DNA [Hoofdstuk 3].
3. Tijdens DNA replicatie wordt een voortbewegende replicatie vork slechts vertraagd - en niet voor onbepaalde tijd gestopt - door de endogene Tus-*Ter* blokkades [Hoofdstuk 4].
4. De endogene Tus eiwitten komen tot expressie in een laag aantal van ongeveer acht tot dertien moleculen per cel en zijn gebonden aan de chromosomale *Ter* locaties die verspreid zijn over de cytoplasmatische ruimte in groeiende cellen [Hoofdstuk 5].
5. De structuur van het bijbehorende operon bemoeilijkt de *labeling* van het endogene primase eiwit, waardoor de ontwikkeling van geavanceerde technieken met hogere efficiëntie nodig is [Hoofdstuk 6].
6. Aangezien zowel de genetische achtergrond als de omgevingsfactoren in belangrijke mate bijdragen aan het gedrag van een levend organisme, is het een misvatting om te vragen welk van deze factoren de dominante rol speelt.
7. De verantwoordelijkheid voor het verbeteren van de reproduceerbaarheid van wetenschappelijk onderzoek ligt niet alleen bij de onderzoeker, maar ook bij elk van de betrokken *stakeholders* [C. Glenn Begley et al., *Circ Res*, 2015].
8. Het onderwijssysteem moet op alle niveaus gebaseerd moeten zijn op de beginselen van samenwerking in plaats van concurrentie.
9. Zeer korte, infrequente mentale pauzes verbeteren ongetwijfeld de uitvoering van een taak; echter, gezien hun verslavende aard, moet het aantal van dergelijke pauzes met zorg omringd worden om uitstelgedrag te voorkomen.
10. De hoogst bereikbare activiteit voor een mens is het leren voor begrip, want wie begrijpt is vrij [Baruch Spinoza, *Ethica*, 1677].

Deze stellingen worden opponeerbaar en verdedigbaar geacht en zijn als zodanig goedgekeurd door de promotor prof.dr.N.H.Dekker.

***In vivo* investigations of *E. coli* chromosomal replication  
using single-molecule imaging**

**Proefschrift**

ter verkrijging van de graad van doctor  
aan de Technische Universiteit Delft,  
op gezag van de Rector Magnificus prof.ir. K.C.A.M. Luyben;  
voorzitter van het College voor Promoties,  
in het openbaar te verdedigen op  
maandag 19 september 2016 om 15:00 uur

door

Sriram TIRUVADI KRISHNAN  
biotechnologie ingenieur,  
geboren te Madurai, Indië

This dissertation has been approved by the promotor:

Prof.dr. Nynke H.Dekker

Composition of the doctoral committee:

Prof.ir. K.C.A.M. Luyben                      chairman, Rector Magnificus, TUDelft.

Prof.dr. N.H. Dekker                         promotor, TUDelft

Independent members

Prof.dr. A. H. Engel                         TUDelft

Prof.dr. T.S. Shimizu                        AMOLF/Vrije Universiteit Amsterdam

Prof.dr. H.J. Tanke                         Leiden Universiteit / TUDelft

Dr. G.E. Bokinsky                         TUDelft

Other Member:

Prof.dr. V. Lorent                         University of Paris 13, France.



Delft  
University of  
Technology



Bionanoscience Department  
Think big about life at the smallest scale



**Keywords:** DNA replication; *E. coli*; chromosomal engineering; single-molecule; epifluorescence microscopy; microfluidics; photo-activable fluorescence microscopy; *in vivo* stoichiometry;

Copyright © 2016 by S. Tiruvadi Krishnan

Casimir Ph.D. Series, Delft-Leiden 2016-23

ISBN 978-90-8593-267-3

An electronic version of this dissertation is available at

<http://repository.tudelft.nl>

# CONTENTS

1	Introduction.....	1
1.1.	References.....	4
2	Essential validation methods for <i>Escherichia coli</i> strains created by chromosome engineering.....	7
2.1.	Introduction.....	8
2.2.	Results and Discussion.....	10
2.3.	Conclusions.....	19
2.4.	Methods.....	20
I.	Strain creation methods.....	20
II.	Strain verification methods.....	22
2.5.	Contributions.....	26
2.6.	Acknowledgments.....	27
2.7.	References.....	27
3	Slow unloading leads to DNA-bound $\beta_2$ sliding clamp accumulation in live <i>Escherichia coli</i> cells.....	31
3.1.	Introduction.....	32
3.2.	Results.....	34
3.3.	Discussion.....	41
3.4.	Materials and Methods.....	45
3.5.	Contributions.....	50
3.6.	Acknowledgments.....	50
3.7.	References.....	51
3.8.	Supplementary Information.....	56
3.9.	Supplementary Figures and Captions.....	62
3.10.	Supplementary Tables.....	70
3.11.	Supplementary References.....	72
4	The progression of replication forks at natural replication barriers in live bacteria...	75
4.1.	Introduction.....	76
4.2.	Results.....	79
4.3.	Discussion.....	89



4.4.	Materials and Methods .....	91
4.5.	Acknowledgments .....	91
4.6.	Contributions .....	91
4.7.	References .....	92
4.9.	Supplementary Information .....	98
4.10.	Supplementary Figures and Captions .....	104
4.11.	Supplementary Tables .....	110
4.12.	Supplementary References .....	112
5	Dynamics and stoichiometry of Tus during the cell cycle of <i>Escherichia coli</i> .....	115
5.1.	Introduction .....	116
5.2.	Results and discussion .....	118
5.3.	Conclusions .....	124
5.4.	Outlook .....	125
5.5.	Materials and Methods .....	125
5.6.	Contributions .....	130
5.7.	References .....	130
5.8.	Supplementary Figures and Captions .....	133
6	Chromosomal engineering of <i>dnaG</i> gene for the labeling of <i>Escherichia coli</i> primase protein <i>in vivo</i> .....	137
6.1.	Introduction .....	138
6.2.	Results and Discussion .....	140
6.3.	Conclusions .....	146
6.4.	Outlook .....	146
6.5.	Materials and Methods .....	146
6.6.	Contributions .....	151
6.7.	References .....	151
	Summary .....	155
	Samenvatting .....	159
	Acknowledgements .....	163
	Curriculum vitae .....	167
	List of publications .....	169

# 1 INTRODUCTION

*It is essential for genetic material to be able to make exact copies of itself; otherwise growth would produce disorder, life could not originate, and favorable forms would not be perpetuated by natural selection.* - Maurice Wilkins on DNA replication, Nobel Lecture 1962

The bacterium *Escherichia coli* (*E. coli*) was first isolated and discovered in 1886 by a German-Austrian pediatrician named Theodor Escherich (1). This versatile microbe can grow both in the presence and absence of oxygen (facultative anaerobe) making it ideal for microbiology and genetic engineering research (2, 3) as a model organism (4). It is usually found in the colon (lower intestine) of warm-blooded organisms, which includes humans (5). Depending on the growth conditions, the width of an *E. coli* cell varies from 0.25 to 1  $\mu\text{m}$  while its length is usually around 2  $\mu\text{m}$  (6), and each cell reproduces itself in a binary fission manner (7). *E. coli* has a circular chromosome as the genetic material containing around 4.6 million bases that include more than 4200 known genes (8). The chromosome has a contour length of around 1.5 mm, and it is compactly organized within a volume of approximately 0.5  $\mu\text{m}^3$  as a nucleoid occupying most of the cell space (9). During most of the cell growth, the chromosome is replicated in a semi-conservative fashion using a complex of enzymes called 'replisome' along with the other cell components (10, 11). The predominant enzyme of the replisome complex is the DNA polymerase III (13 subunits) which adds complementary nucleotides (NTPs) to the two single strands of the parent DNA (12). The polymerase III (PolIII) is aided by the DNA unwinding enzyme helicase (six subunits), the replication initiator protein (single subunit named DnaA), the processivity increasing clamp proteins (two subunits named  $\beta$ ) and the replication initiating RNA primers which are synthesized by primase (single subunit) (12).

DNA replication process in *E. coli* proceeds in three stages (13). The first stage is initiation during when the replisome complex is formed and bound to the origin region (*ori*) of the chromosome (13). In the second stage of elongation, the two sets of replisomes independently move along the chromosome with the replication fork proceeding in two opposite directions (14). Finally, for the termination the two independent replisomes meet at the termination region (*ter*) and after completing the replication unbind from the chromosome (14). Simultaneously during this process, messenger RNA (mRNA) molecules are transcribed from the genes by RNA polymerases, and proteins are translated from the mRNA sequence by ribosomes and transfer RNA (tRNA) (15). Due to the helicase unwinding activity during the replication process topological stress is created which is relieved by the activity of topoisomerases IV (16). The single strands of DNA formed in the replication fork are protected from degradation by the single-strand DNA binding proteins (*ssb*) (17). The two DNA strands diverging from the replication fork are termed as leading and lagging strands because the leading strand is synthesized continuously, while the lagging strand is synthesized semi-discontinuously as 'Okazaki fragments' with the help of primers made by primases (18). This is due to the unidirectional addition of NTPs (i.e. only at the 3'-OH of the previously added NTP) to the two anti-parallel strands by PolIII. The

DNA polymerase I corrects the RNA-DNA primed regions, and the nicks are ligated by the DNA ligase (18).

Recent advances in single-molecule fluorescence microscopy have made it possible to understand the intricate details of the *E. coli* DNA replication mechanism and the protein-protein interactions involved that were previously unknown (19). On assembling the PolIII complex *in vitro* on DNA curtains, it was found that the components of the PolIII complex are actively recycled after every Okazaki fragment synthesis (20). With advances in genetic engineering techniques, it became possible to label native replisome components with fluorescent proteins directly inside the cell, and it opened new avenues of research by imaging the replisome proteins in action as and when the cell grows (19). In one such research, the two independent replisomes on replication forks were tracked in real time with respect to the positions of *ori* and *ter* regions of the chromosome and the phases of *E. coli* cell cycle were accurately determined (21). Fluorescence imaging of such live cells with single molecule sensitivity also gave further insights on the stoichiometry of labeled replisome components. From such data, it became evident that there are always three PolIII complexes at a replication fork which might be actively switched during the leading and lagging strand synthesis giving a new architecture for replisomes (22). Such interesting details were revealed only through the labeling of native proteins using chromosome engineering techniques.

In this thesis, I detail the methodology and research performed with the goal to expand our knowledge on chromosome replication in growing *E. coli* cells. The projects presented here are part of a collaborative research executed by me along with other researchers. At the end of each chapter, I have mentioned the contributions of each personnel related to that particular project. We employed in our research a multi-disciplinary approach using the following techniques: quantitative fluorescence microscopy, chromosome engineering, microfluidics and molecular cloning. This made us possible to investigate the DNA replication process in single cells throughout their cell cycles. The outline of this dissertation is as follows:

In Chapter 2, we discuss the strain creation techniques that we use in our research and, in detail the essential validation methods to be used on the created strain with examples. The most common and successful chromosome engineering tools for *E. coli* are the  $\lambda$ -Red recombination and P1 phage transduction. However, these tools are not completely error proof, and it necessitates validations to be performed beyond the standard methods like PCR and DNA sequence tests. The occurrence of extraneous insertions is the most common defect that can arise during  $\lambda$ -Red recombination, which can be detected easily using Southern blotting. Another way to rectify this error is to use P1 phages to transduce the sequence from engineered strain into the wildtype strain. However, errors can arise when P1 phage stock contains temperate phages that can exist inside the created strain as a plasmid, and we present the methods to detect the temperate phages. Finally, we explain the easy-to-use methodologies to identify the defects in cell growth and shape characteristics in the engineered strain and the guidelines for the strain validation as a process flow.

In Chapter 3, we detail our investigation of the dynamics of the *E. coli* sliding clamp proteins ( $\beta_2$ -clamps) *in vivo*. This protein not only increases the processivity of PolIII complex but also participates in the protein complex involved in the error correction in

DNA. Using quantitative fluorescence microscopy with microfluidics, we accurately enumerate the loading, binding and recycling times of the labeled  $\beta_2$ -clamps on DNA across many cell cycles. We show that  $\beta_2$ -clamps are recruited at the start of DNA replication and they are bound to DNA for a short period (~4 minutes) causing an accumulation at the replication fork until a steady state is attained. The accumulated  $\beta_2$ -clamps may help in the binding of other DNA modifying enzymes at the fork. The results of this study have expanded our understanding of the mechanism of DNA replication and repair processes.

In Chapter 4, we focus on the termination of DNA replication. We study the outcome of an encounter between a replication fork and a natural replication barrier at the *ter* region of the *E. coli* chromosome. We tracked the replisome (using labeled  $\beta_2$ -clamps and a PolIII subunit) in relation to specific chromosomal loci in a strain with a displaced origin region (*oriZ*) such that one of the two replisomes faces the roadblock first. From our observations, we show that even after facing the roadblock, the replisome is stably bound and the replication is slowed down but not stalled forever. We also found that the chromosome organization in daughter cells varies depending on the presence and absence of the replication barrier. Our findings provide further insights into the versatility of the cell to complete the chromosome replication.

In Chapter 5, we turn our attention to the examination of the DNA binding protein Tus which forms the natural replication barriers by binding to the *Ter* sites on the *E. coli* chromosome in live cells. We studied the labeled Tus proteins across the cell cycle in relation to the *ter* region of the chromosome. Our initial results reveal that during the cell growth the Tus proteins are expressed in very low copy numbers (8 to 13 molecules) of which the bound Tus proteins (~5 molecules) are mostly static and are distributed throughout the cell similar to the *Ter* loci. The outcomes of our research improve the understanding of the Tus-*Ter* complexes in growing cells and inspire better designs for further studies on protein-protein interactions involved in the termination of DNA replication in live cells.

In Chapter 6, we present about our various attempts to reveal the features of primase proteins in action within the live cells. Since the endogenous primase gene of *E. coli* is present in a complex operon system, it poses hurdles in labeling the native primase proteins. Through our control experiments, it is revealed that even after successful chromosome engineering the expressed fluorescent protein linked primase fusion from its native operon is non-fluorescent and may be truncated. While the same fusion gene, when expressed under an inducible promoter in a plasmid, the expressed protein is found to be fluorescent. Hence we attempted a different approach using the  $\lambda$ -Red recombination-based strategies to knock out the native primase gene while ectopically expressing the fluorescent primase gene from the genome, and they yielded only false positive results. Our trials in this study emphasize on the necessity of a more efficient genome-editing tool for further research on studying native primase proteins in live *E. coli* cells.

## 1.1. References

1. T. Escherich, *Die darmbakterien des säuglings und ihre beziehungen zur physiologie der Verdauung*. (F. Enke, 1886).
2. S. Y. Lee, High cell-density culture of *Escherichia coli*. *Trends in Biotechnology* **14**, 98-105 (1996).
3. E. Russo, Special Report: The birth of biotechnology. *Nature* **421**, 456-457 (2003).
4. F. X. Gomis-Rüth, M. Coll, Cut and move: protein machinery for DNA processing in bacterial conjugation. *Current Opinion in Structural Biology* **16**, 744-752 (2006).
5. P. Singleton, in *Bacteria in Biology, Biotechnology and Medicine*. (Wiley, 1999), chap. 16, pp. 444-454.
6. A. C. S. Yu, J. F. C. Loo, S. Yu, S. K. Kong, T.-F. Chan, Monitoring bacterial growth using tunable resistive pulse sensing with a pore-based technique. *Applied Microbiology and Biotechnology* **98**, 855-862 (2013).
7. S. Bayne-Jones, E. F. Adolph, Growth in size of micro-organisms measured from motion pictures. III. *Bacterium coli*. *Journal of Cellular and Comparative Physiology* **2**, 329-348 (1932).
8. F. R. Blattner *et al.*, The Complete Genome Sequence of *Escherichia coli* K-12. *Science* **277**, 1453-1462 (1997).
9. M. Macvanin, S. Adhya, Architectural organization in *E. coli* nucleoid. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms* **1819**, 830-835 (2012).
10. M. Meselson, F. W. Stahl, THE REPLICATION OF DNA IN *ESCHERICHIA COLI*. *Proceedings of the National Academy of Sciences of the United States of America* **44**, 671-682 (1958).
11. A. M. Breier, H.-U. G. Weier, N. R. Cozzarelli, Independence of replisomes in *Escherichia coli* chromosomal replication. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 3942-3947 (2005).
12. A. Johnson, M. O'Donnell, Cellular DNA replicases: components and dynamics at the replication fork. *Annu. Rev. Biochem.* **74**, 283-315 (2005).
13. S. Cooper, C. E. Helmstetter, Chromosome replication and the division cycle of *Escherichia coli* B/r. *Journal of molecular biology* **31**, 519-540 (1968).
14. R. Reyes-Lamothe, E. Nicolas, D. J. Sherratt, Chromosome replication and segregation in bacteria. *Annual review of genetics* **46**, 121-143 (2012).
15. T. Kogoma, Stable DNA replication: interplay between DNA replication, homologous recombination, and transcription. *Microbiology and Molecular Biology Reviews* **61**, 212-238 (1997).

16. X. Wang, R. Reyes-Lamothe, D. J. Sherratt, Modulation of Escherichia coli sister chromosome cohesion by topoisomerase IV. *Genes & development* **22**, 2426-2433 (2008).
17. R. D. Shereda, A. G. Kozlov, T. M. Lohman, M. M. Cox, J. L. Keck, SSB as an organizer/mobilizer of genome maintenance complexes. *Critical Reviews in Biochemistry and Molecular Biology* **43**, 289-318 (2008).
18. S. D. Bell, Molecular biology: Prime-time progress. *Nature* **439**, 542-543 (2006).
19. G.-W. Li, X. S. Xie, Central dogma at the single-molecule level in living cells. *Nature* **475**, 308-315 (2011).
20. N. Y. Yao, R. E. Georgescu, J. Finkelstein, M. E. O'Donnell, Single-molecule analysis reveals that the lagging strand increases replisome processivity but slows replication fork progression. *Proceedings of the National Academy of Sciences* **106**, 13236-13241 (2009).
21. R. Reyes-Lamothe, C. Possoz, O. Danilova, D. J. Sherratt, Independent positioning and action of Escherichia coli replisomes in live cells. *Cell* **133**, 90-102 (2008).
22. R. Reyes-Lamothe, D. J. Sherratt, M. C. Leake, Stoichiometry and architecture of active DNA replication machinery in Escherichia coli. *Science* **328**, 498-501 (2010).



## 2 ESSENTIAL VALIDATION METHODS FOR *ESCHERICHIA COLI* STRAINS CREATED BY CHROMOSOME ENGINEERING

*Chromosome engineering encompasses a collection of homologous recombination-based techniques that are employed to modify the genome of a model organism in a controlled fashion. Such techniques are widely used in both fundamental and industrial research to introduce multiple insertions in the same Escherichia coli strain. To date,  $\lambda$ -Red recombination (also known as recombineering) and P1 phage transduction are the most successfully implemented chromosome engineering techniques in E. coli. However, due to errors that can occur during the strain creation process, reliable validation methods are essential upon alteration of a strain's chromosome. Polymerase chain reaction (PCR)-based methods and DNA sequence analysis are rapid and powerful methods to verify successful integration of DNA sequences into a chromosome. Even though these verification methods are necessary, they may not be sufficient in detecting all errors, imposing the requirement of additional validation methods. For example, as extraneous insertions may occur during recombineering, we highlight the use of Southern blotting to detect their presence. These unwanted mutations can be removed via transducing the region of interest into the wild-type chromosome using P1 phages. However, in doing so one must verify that both the P1 lysate and the strains utilized are free from contamination with temperate phages, as these can lysogenize inside a cell as a large plasmid. Thus, we illustrate various methods to probe for temperate phage contamination, including cross-streak agar and Evans Blue-Uranine (EBU) plate assays, whereby the latter is a newly reported technique for this purpose in E. coli. Lastly, we discuss methodologies for detecting defects in cell growth and shape characteristics, which should be employed as an additional check. The simple, yet crucial validation techniques discussed here can be used to reliably verify any chromosomally engineered E. coli strains for errors such as non-specific insertions in the chromosome, temperate phage contamination, and defects in growth and cell shape. While techniques such as PCR and DNA sequence verification should standardly be performed, we illustrate the necessity of performing these additional assays. The discussed techniques are highly generic and can be easily applied to any type of chromosome engineering.*



## 2.1. Introduction

Modification of an organism's phenotypes by altering its chromosomal DNA sequence in a controlled manner provides the fundamental motivation for chromosome engineering (1). This engineering comprises a collection of techniques that can be applied to insert foreign DNA sequences at a specific locus, delete the native sequence, or alter the bases in the chromosomal DNA of a model organism. One such key model organism is the bacterium *Escherichia coli* (*E. coli*), used for many fundamental studies in molecular and cell biology, and it is also utilized for expressing novel proteins. The single, circular chromosome of *E. coli* has been fully sequenced and amply annotated, paving the way for researchers to engineer precisely its chromosome using a variety of methods (2). Most chromosome engineering techniques in *E. coli* harness the properties of the recombinase family of enzymes, expressed by plasmids or bacteriophages, which recombine homologous linear DNA fragments into the host chromosome (3).

In the past decade, numerous studies have successfully employed chromosome engineering tools in *E. coli* research. The ability to fuse a fluorescent protein gene with native genes (4) is an example that has opened up the possibility of live cell imaging to visualize the dynamics and stoichiometry of native proteins involved in key biochemical processes such as DNA replication (5, 6), transcription (7, 8), translation (7), chromosome segregation (9), cell signaling (10), and flagellar motor dynamics (11). Similarly, a high throughput study in which single genes were systematically deleted from the *E. coli* chromosome has shed light on the (non)-essential character of individual genes, and hence on the possibilities of creating a minimal cell containing only essential genes (12, 13). In addition to such studies into fundamental aspects of molecular and cellular biology, chromosome engineering is also used in industrial research to produce essential biochemicals, biofuels, and precursors for pharmaceuticals on a large scale by engineering all necessary genes into a single *E. coli* strain (14-16). In industrial research, chromosome engineering is typically preferred to conventional cloning in plasmids, as it obviates the need for antibiotics to maintain gene presence (15).

For any research in which the chromosome is engineered, it is essential to reliably verify that the process has not inadvertently introduced anomalies into the genome. For example, the use of  $\lambda$ -Red recombination or recombineering (17-20) (**Table 1**) may result in one of the following: insertions at undesired locations in the chromosome, a result of sequence heterogeneities introduced during synthesis of the requisite long primers (17, 21), the presence of an unstable genomic region (22), or the occurrence of partial gene duplication in the chromosome during the strain creation process (23). The latter is illustrated by the gene duplication errors that occurred in 0.6% of the 3864 single-gene deletion mutants of *E. coli* K12 strains in the Keio collection (23). Together with the intended insertion which usually occurs at a probability of  $\sim 10^{-4}$  to  $10^{-5}$  (24) non-specific mutations may occur and, they are not detected easily using standard PCR techniques (22). When such errors occur, P1 phage transduction (**Table 1**) can be performed to recover the strain of interest, as the region of interest can be specifically transduced into a clean wild-type strain following recombineering (25-31). However, P1 phage transduction comes with its own challenges, such as the potential contamination of temperate phages in the phage stock that can lysogenize as a large plasmid in the created strain. Hence, appropriate

validation of the chromosomally engineered strain remains a critical step in the strain creation process.

### **$\lambda$ -Red recombination**

The  $\lambda$ -Red recombination (or recombineering) approach has been successfully implemented in many studies engineering specific sites in the *E. coli* chromosome (17). In this approach, chromosomal sequences are replaced by a linear DNA fragment (flanked by sequences homologous to the region of interest) using a temperature-sensitive plasmid that expresses either the Red recombinase genes (*bet*, *gam*, and *exo*) from  $\lambda$ -phage (17) or the RecET proteins from  $\lambda$  phage (18) upon induction. The linear DNA fragment of interest is usually synthesized via PCR, in which case the homologous sequences (~50 bases) are introduced through the employed primers. The recombined strain is selected using a constitutively expressed antibiotic marker that is integrated into the chromosome along with the insert of interest. This technique may also be combined with FLP/*FRT*-based recombination, in which the antibiotic marker is flanked by *FRT* sites that allow it to be recombined out using the flippase (FLP) enzyme (19). In this way, the created strain may be employed in multiple rounds of chromosome engineering using the same antibiotic marker (17, 20).

### **P1 phage transduction**

Generalized P1 phage transduction is widely used to transfer mutations from one *E. coli* strain to another with the same genetic background (26, 27). This approach is based on the fact that virulent P1 phages commit errors while packaging their DNA into coat proteins: instead of packaging their own genome, they package lysed host chromosomal DNA fragments (28-30). Such mispackaged phages constitute approximately ~5% of the total phage population in a lysate. When they are transduced into a different host, the chromosome fragment may be inserted precisely at a homologous site using the RecA-dependent system (26). Using this approach, multiple insertions can be made into the ~4.6 Mb chromosome of same *E. coli* strain, provided that they are separated by ~100 kb (31). A combination of  $\lambda$ -Red recombination, FLP/*FRT* recombination, and P1 phage transduction methods can also be used to introduce multiple insertions into a single *E. coli* strain (6, 15).

Here, we describe in detail a number of general methods for the validation of strains with altered chromosomes and accompany this description with experimental results. We note that several of the techniques described here are individually well known to the scientific community; however, frequently only the more standard verification procedures for chromosome engineering, PCR, and DNA sequence analysis, are typically reported. Grouping together the description of these techniques, we bring to the increased attention of researchers the most common defects that can arise during strain creation, together with the appropriate methods to verify them. In doing so, we hope to make these techniques more readily accessible to a wider community, facilitating access to them by new researchers and/or those engaged in cross-disciplinary study. We have organized our description of these essential validation methods along the lines of the irregularities that may occur: (i) non-specific insertions in the chromosome; (ii) the contamination of

temperate P1 phage in the engineered *E. coli* strains; and (iii) defects in phenotypes such as cell growth and morphology. The latter physiological aspect must be taken into account if the results obtained from an engineered *E. coli* strain are to be generalized to wild-type *E. coli*. Within each of these categories, we illustrate the defects that can arise from errors in chromosome engineering and describe various methods to detect them, using as examples the creation of two *E. coli* AB1157 strains: non-motile  $\Delta$ *motAB* for use in live-cell fluorescence microscopy (32) and *pBad-DnaG* in which an inducible primase gene is inserted into the non-essential *galK* chromosomal locus.

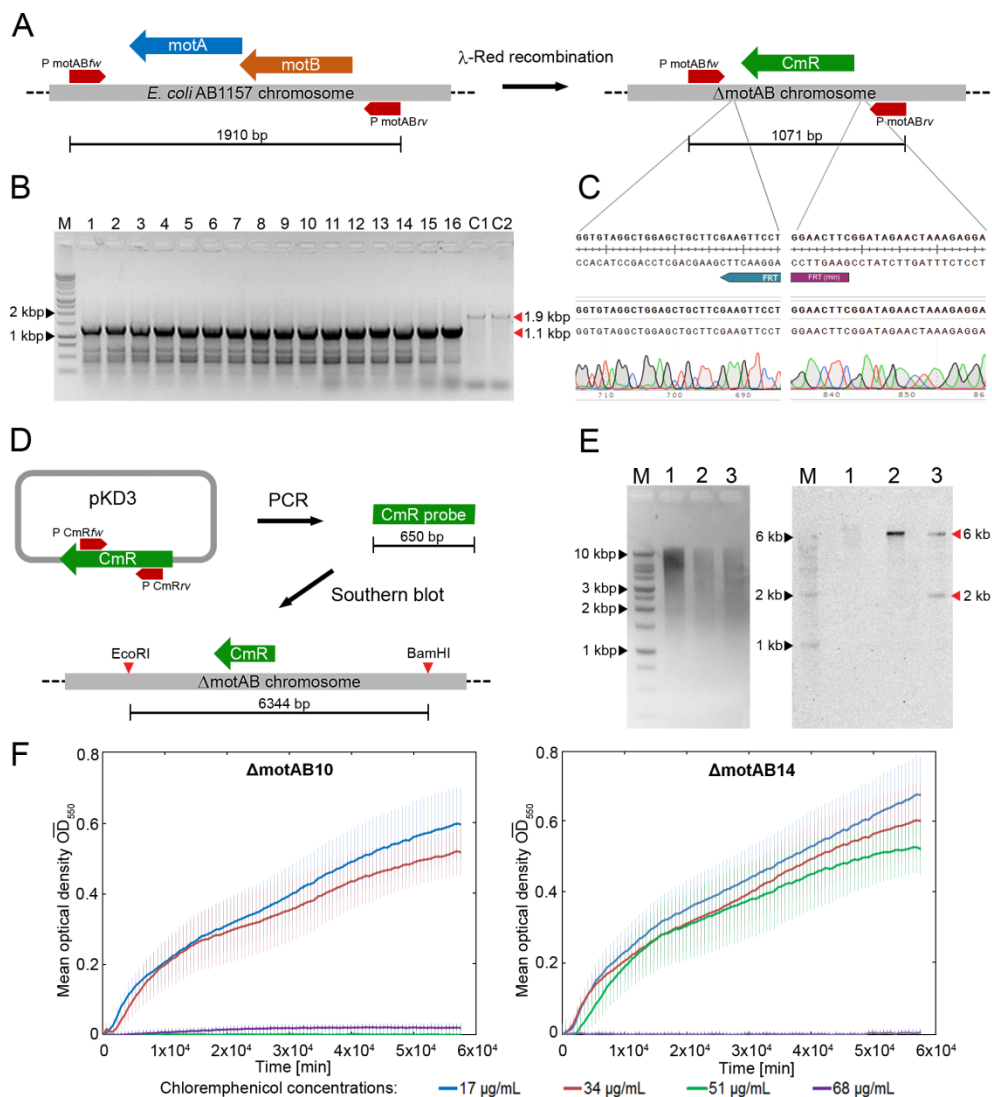
## 2.2. Results and Discussion

### 2.2.1. Verification of an engineered sequence in the chromosome

PCR and DNA sequence analysis are the techniques that are widely performed and reported to verify whether a chromosome engineering technique has successfully modified the chromosomal DNA sequence. We performed these well-known techniques as a first pass in the validation process of strains created via  $\lambda$ -Red recombination. Specifically, we performed  $\lambda$ -Red recombination to knock out the genes expressing the flagellar motor proteins (*AmotAB*) in the *E. coli* AB1157 strain (**Methods section IA**). The targeted gene was replaced with the chloramphenicol resistance gene (*CmR*), which was used as a selection marker to isolate the successfully engineered colonies. To verify whether the insert (*FRT-CmR-FRT*) was located at the intended site, we designed primers that flank the region of interest (**Figure 1A, Table 1**). The positive results of a PCR reaction performed on the 16 selected colonies using flanking primers indicated that the *motAB* genes in the chromosome were replaced with the *CmR* gene (**Figure 1B**). DNA sequence analysis was performed to verify the recombinered region of the chromosome (**Methods section II A**). The quality of sequencing results also provides insights on the integrated DNA sequence in the chromosome. For example, if the DNA sequence results show a double signal (i.e. signals for two bases at the same position), it indicates sequence heterogeneity of the integrated DNA amongst the cells of a colony (33). In our experiments, the DNA sequence and its alignment with the template sequence revealed that the *motAB* genes were successfully knocked out without sequence errors in 9 out of 16 colonies (**Table 2**). A representative DNA sequencing result of the  $\Delta$ *motAB*10 strain at the sites of integration and the corresponding alignment with the expected template DNA sequence are shown (**Figure 1C**).

Multiple copies of the insert sequence (*CmR*) could have recombined elsewhere in addition to the intended site on the chromosome, and such extraneous insertions can be detected using Southern blotting (**Methods section IIB**)(22, 24, 34). We performed this technique on nine  $\Delta$ *motAB* strains and AB1157 strain as a negative control, all of which were initially verified by PCR and DNA sequence analysis. For this specific experiment, the chromosomal DNA samples isolated from the strains were first digested with two high fidelity restriction enzymes: EcoRI-HF and BamHI-HF selected based on the criteria detailed in **Methods section IIB**. The digested DNA fragments were separated via agarose gel electrophoresis and were then transferred to a blotting membrane to probe for the *CmR* gene. The DNA sequence that serves as a probe was obtained by PCR-based amplification of a 650 bp fragment obtained from the *CmR* gene of template plasmid pKD3 (**Table 1**,

**Figure 1D).** This single-stranded probe DNA was labeled directly with a thermostable alkaline phosphatase enzyme. The following hybridization on the blotting membrane, this enzyme catalyzes a chemiluminescence reaction, thereby allowing target DNA fragments complementary to the probe to be detected on a CCD detector. Using this approach, we observed that in two of the nine strains ( $\Delta motAB11$  and  $\Delta motAB14$ ) the *CmR* gene had recombined not only at the intended region but also at another non-essential unknown region of the chromosome.



**Figure 1: Verification of an engineered sequence in the chromosome.** A) The scheme depicts the changes at the recombining site to create the *motAB* gene knockout strain ( $\Delta motAB$ ) using a chloramphenicol resistance gene (*CmR*). The positions of the flanking primers for the *motAB* region

(PmotAB<sub>fw</sub> and PmotAB<sub>rv</sub>) are marked, and the corresponding product lengths from PCR are indicated at the bottom. **B**) PCR results of the colonies obtained through recombineering. In lane M, 10  $\mu$ L of DNA ladder was loaded. In lanes 1-16, 10  $\mu$ L of PCR products from selected, individual  *$\Delta$ motAB* colonies were loaded ( *$\Delta$ motAB1* to  *$\Delta$ motAB16*). The PCR products of the control AB1157 strain were added to wells C1 and C2. PCR products of the intended sizes are visible for all 16 selected colonies (~1.1 kbp marked with red arrow) as well as the AB1157 colonies (~1.9 kbp). **C**) A representative DNA sequencing result of the  *$\Delta$ motAB10* strain at the sites of integration and the corresponding alignment with the expected template DNA sequence are shown. **D**) The important steps of making the Southern blot probes are illustrated. A 650 bp PCR product is amplified from the template plasmid pKD3 and is then labeled with alkaline phosphatase to probe the *CmR* region (expected size: 6 kbp). **E**) The ethidium bromide stained gel containing the DNA ladder (lane M), the restriction-digested AB1157 genome (lane 1), restriction-digested genomes of two  *$\Delta$ motAB* colonies (lane 2:  *$\Delta$ motAB10* and lane 3:  *$\Delta$ motAB14* which was verified by PCR and DNA sequencing). The Southern blot results show that the AB1157 sample in lane 1 has no insert, as expected; lane 2 with  *$\Delta$ motAB10* has one band (6 kbp) at the right fragment size showing that the integration was successful at the predicted site; lane 3 with  *$\Delta$ motAB14* has two bands (6 kbp and 2 kbp). **F**) The growth of  *$\Delta$ motAB10* and  *$\Delta$ motAB14* strains in 96 well plate OD reader containing LB medium with various concentrations of chloramphenicol (17  $\mu$ g/mL to 68  $\mu$ g/mL). The results show that  *$\Delta$ motAB14* strain containing the extraneous insertion grew at a higher concentration of chloramphenicol (51  $\mu$ g/mL) than the normal concentration (34  $\mu$ g/mL), while the  *$\Delta$ motAB10* did not grow at 51  $\mu$ g/mL of chloramphenicol.

**Table 1: Primer names and sequences used**

Primer name	Primer sequence
PmotAB <sub>fw</sub>	5'- GCT GAA GCC AAA AGT TCC TG -3'
PmotAB <sub>rv</sub>	5'- TGC CTG CAG CTT ATG TCA AC -3'
PcmR <sub>fw</sub>	5'- ATC ACA AAC GGC ATG ATG AA -3'
PcmR <sub>rv</sub>	5'- TCA CTA CCG GGC GTA TTT TT -3'
PgalK <sub>fw</sub>	5'- TCC ATC AGC GTG ACT ACC ATC -3'
PgalK <sub>rv</sub>	5'- CAG AAC AGG CAG CAG AGC GT -3'

A comparison of the blotting results for  *$\Delta$ motAB10* strain with single intended *CmR* insertion and  *$\Delta$ motAB14* strain with extraneous insertion is shown (**Figure 1E**). We also verified that the  *$\Delta$ motAB14* strain showed increased chloramphenicol tolerance compared to the  *$\Delta$ motAB10* strain as a result of this additional insertion. To do so, both strains were grown under constant shaking at 37 °C and 350 rpm in 96 well plates containing LB media with different concentrations of chloramphenicol, and the optical density was measured at regular intervals. While no growth was observed for the  *$\Delta$ motAB10* strain in LB medium including an increased concentration of chloramphenicol (51  $\mu$ g/mL compared to a normal dosage of 34  $\mu$ g/mL), the growth of  *$\Delta$ motAB14* strain remained unaffected (**Figure 1F**). These results demonstrate that extraneous mutations may

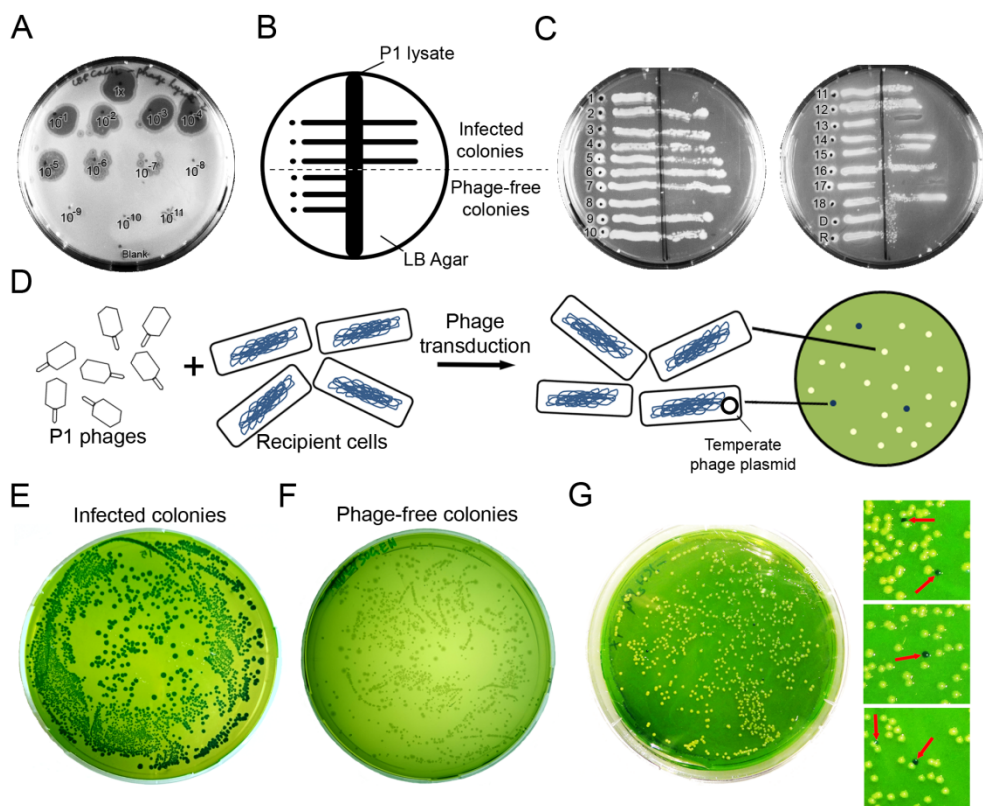
occur when performing recombinering techniques (24), supporting the need for strain verification steps like Southern blot analysis in addition to the standard methods of PCR and DNA sequence analysis. The properly verified  $\Delta motAB10$  strain generated in these experiments will be referred as  $\Delta motAB$  in what follows.

**Table 2: Summary of DNA sequence analysis results for various  $\Delta motAB$  colonies.**

Colony id	Summary of DNA sequence analysis results
$\Delta motAB1$	Positive (Good signal at both the integration sites and insert)
$\Delta motAB2$	Positive (Good signal at both the integration sites and insert)
$\Delta motAB3$	Negative (Double signal at the end)
$\Delta motAB4$	Positive (Good signal at both the integration sites and insert)
$\Delta motAB5$	Positive (Good signal at both the integration sites and insert)
$\Delta motAB6$	Negative (Low signal)
$\Delta motAB7$	Negative (Double signal at the end)
$\Delta motAB8$	Negative (Double signal at the beginning)
$\Delta motAB9$	Negative (Low signal)
$\Delta motAB10$	Positive (Good signal at both the integration sites and insert)
$\Delta motAB11$	Positive (Good signal at both the integration sites and insert)
$\Delta motAB12$	Negative (Low signal with broad peaks)
$\Delta motAB13$	Negative (Low signal)
$\Delta motAB14$	Positive (Good signal at both the integration sites and insert)
$\Delta motAB15$	Positive (Good signal at both the integration sites and insert)
$\Delta motAB16$	Positive (Good signal at both the integration sites and insert)

## 2.2.2. Detection of temperate phage contamination in a phage lysate or a transduced strain

In a phage-transduced strain, temperate P1 bacteriophage can lysogenize as a large plasmid and can replicate for generations along with chromosome of the strain. To demonstrate detection techniques for this phenomenon, we first performed a sample P1 phage transduction experiment. In this experiment, we employed a donor strain in which the  $\beta$ -clamp gene (*dnaN*) is fused with the gene for a yellow fluorescent protein adjacent to a kanamycin marker (*kanR*), as reported in a recent study on DNA replication (*Ypet-DnaN*) (5). Transduction of such a DNA sequence into the non-motile  $\Delta$ *motAB* strain can provide a general approach for live cell imaging studies, whose focus on the visualization of internal cellular dynamics benefits from the use of immobilized cells (32). As a proof-of-principle for our validation techniques, however, we employed *E. coli* AB1157 strain as the recipient strain.



**Figure 2: Detection of temperate phage contamination in a phage lysate or transduced strain.**

**A)** The spot agar assay was performed using a serially diluted P1 phage lysate starting from the undiluted lysate (labeled as 1x) to the dilution factor of  $10^{11}$  (labeled as  $10^{-11}$ ). The results reveal that plaques are observed at every concentration down to the  $10^7$ <sup>th</sup> dilution. The P1 lysate stock was thereby determined to have an infection titer value of  $7 \times 10^9$  pfu mL<sup>-1</sup>. **B)** A schematic diagram of a typical LB agar plate used for the cross-streak agar assay: The vertical dark region at the center

represents the zone of P1 lysate. The dot represents the location where the tested cells are inoculated at a safe distance from lysate zone, and the horizontal solid lines represent either the temperate phage-infected cells that are growing across the lysate zone or the phage-free cells that are not growing beyond the lysate zone. **C)** The two representative plates of the cross-streak agar assay performed with the colonies obtained from phage transduction experiment demonstrate that 14 of the 18 colonies tested are infected with temperate phages. The plate on the right side indicates that the donor (D) and recipient (R) strains used are devoid of phages. **D)** A scheme of the steps involved in Evans Blue-Uranine (EBU) plate assay, explaining the principle of this technique to screen for temperate phage contamination. **E)** Temperate phage-containing cells verified using cross-streak agar assay grew as dark green colonies on EBU plate. **F)** Cells verified to be free of phages grew as pale green colonies on EBU plate. **G)** An EBU plate assay was performed with the diluted cultures of colonies obtained from a P1 phage transduction experiment. A representative result plate and an enlarged view of the colonies obtained from various EBU plates are shown. A mix of uninfected colonies (pale yellowish green color) containing no temperate phages and colonies containing temperate phages (dark green) were observed on the plates.

Prior to phage transduction in *E. coli*, P1 lysate stock should be tested to determine both the infectivity of the P1 phages in the stock as well as the sensitivity of the *E. coli* strain used for transduction. To visualize plaque formation and determine the infection titer value of the P1 stock, we performed a spot agar assay (31, 25). In this assay, different dilutions of P1 lysate stock are spotted onto a lawn of cells grown on a soft LB agar (0.75 %). The titer values are determined in terms of plaque forming units per mL (pfu mL<sup>-1</sup>). We performed a spot agar assay on the *E. coli* AB1157 strain (**Methods section IIC**) and, by counting the plaques formed during the assay, determined the titer of the tested P1 lysate to be  $\sim 7 \times 10^9$  pfu mL<sup>-1</sup> (**Figure 2A**). The observed titer value was found to be in the optimal range for successful transduction ( $10^9$  to  $10^{10}$  pfu mL<sup>-1</sup>) (25). Phage transduction was performed using the characterized lysate on *Ypet-DnaN* as the donor strain and AB1157 as the recipient strain (**Methods section IB**).

One must carefully ensure that phage-transduced cells do not harbor temperate phages, which can result for instance from the use of a P1 lysate contaminated with temperate phages. This phenomenon can yield undesirable results, such as slow growth or abnormal physiology, in the created strains (25). Cells carrying temperate P1 phage DNA as a large plasmid are also prevented from further P1 phage infection. This principle is used in cross-streak agar assays to detect the presence of temperate phages in the sample. In this technique, the colonies to be tested are streaked across a 'P1 lysate layer' on an LB agar plate, and the plate is incubated. If colony growth is not observed on the streak beyond the lysate layer, then it confirms the absence of temperate P1 phages from the sample. However, if growth is observed beyond the lysate layer, this indicates either the presence of temperate P1 phages or immunity of the strain to P1 phage infection (**Figure 2B**) (25). We used the cross-streak agar assay to test the colonies obtained from the phage transduction experiment as well as control strains (**Methods section IID**). We observed that 14 out of the 18 tested colonies from the phage transduction experiment grew across the P1 lysate streak, thereby demonstrating the presence of temperate bacteriophages in these transduced colonies (**Figure 2C**). The remaining four colonies that were verified to be devoid of temperate phages can be used for further experiments.

The best practice to avoid contamination by temperate bacteriophages is to employ a verified virulent P1 lysate in P1 phage transduction experiments. To facilitate this verification of the P1 lysate, we have developed a rapid, easily applicable assay to detect



the presence of temperate phages in the P1 lysate or in the employed strains. This assay is derived from Evans Blue-Uranine (EBU) plate assays, which are commonly used to verify pseudo-lysogeny in P22 phage transduction experiments of *Salmonella* strains (35). We demonstrate here its first use in P1 phage transduction experiments using *E. coli* cells. When temperate phages are present in cells, a colony formed from these cells will have a pH that differs from that of uninfected cells because of pH lowering through the lysis of pseudo-lysogenic cells (36). This property has been exploited in the EBU plate assay to directly visualize colonies containing Evans blue stained pseudo-lysogenic cells (**Figure 2D**).

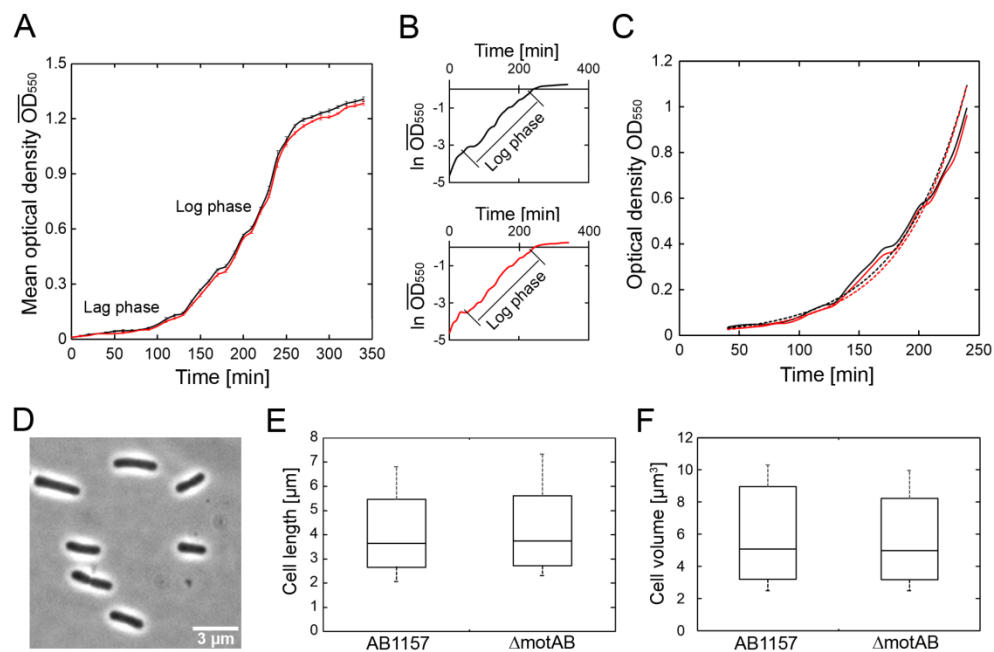
The temperate phage-infected cells verified by the cross-streak agar experiment were tested using the EBU plate assay (**Methods section IIE**): they exhibited exclusively dark green colonies (**Figure 2E**), which we associate with infection by temperate phages that result in a change in cellular pH (36). Conversely, the cells verified to be free of phages displayed exclusively pale green colonies (**Figure 2F**). To demonstrate that even minute contamination of temperate phages in P1 lysate could be detected, we mixed the contaminated lysate with a verified virulent P1 stock of the *Ypet-DnaN* strain at a ratio of 1:100. We performed a P1 phage transduction experiment as described above, and inspected 20 random colonies using EBU plate assay. The results showed that 3 of the inspected 20 EBU plates contain few colonies which are dark green in color, whereas neighboring pale yellowish green-colored colonies are free of phages (**Figure 2G**), confirming the detectability of low-level temperate phage contaminants in the P1 lysate. We find that the EBU plate assay is more convenient and reliable than the cross-streak agar assay for the detection of temperate phage contamination in transduced colonies and P1 lysates, and the phage-free colonies can be used for further research.

### 2.2.3. Evaluation of *E. coli* strains based on cellular growth or morphology characteristics

Bacterial growth curve analysis provides an overview of the growth behavior of the chromosomally engineered *E. coli* strains. A typical bacterial growth curve starts with a lag phase as the bacteria adapt to the fresh growth medium, followed by a log phase in which growth is exponential. The final phase of the growth curve displays stationary growth as a result of nutrient scarcity, after which cells eventually die (**Figure 3A**) (37). Two important parameters that can be determined using the technique of growth curve analysis are the log-phase growth rate ( $\mu$ ) and the duration of lag phase ( $\tau_l$ ) (38). The log phase doubling time (generation time,  $\tau_d$ ) is calculated from  $\mu$ . If growth defects are introduced during the strain creation process, they can be detected by comparing the generation times of the parental strain with those of the created strain. The literature suggests numerous models and tools with which to perform this analysis (37, 38). As an example, we have performed growth curve validation for the AB1157 and  $\Delta motAB$  strains (**Figure 3A, Methods section IIF**). The critical step is to determine which time points of the growth curve fall in the log phase; fortunately, this is easily achieved by determining the linear region of the semi-log plot of the same curve (**Figure 3B**). By fitting the log phase portion of the curve with an exponential function, we calculated the growth rates for each sample (**Figure 3C**). From the growth rates, the mean generation times with standard deviation (SD) for the AB1157 and  $\Delta motAB$  were found to be  $39.2 \pm 2.1$  min and  $38.7 \pm 1.9$  min, respectively. To determine

the statistical significance of this difference, we employed *t*-test statistics for two independent sample means (39). From the observed *p*-value of 0.68 (Table 3), we conclude with 95% confidence intervals that no significant difference in generation times can be attributed to the *motAB* deletion genotype.

Cell morphology can be examined using numerous methods including flow cytometry (40), atomic force microscopy (41), among others, and this essential phenotype can reveal the overall fitness of the chromosomally modified strain. Here we describe an approach that employs phase contrast microscopy and automated image analysis software MicrobeTracker (Methods section IIG) (42). Using this open-source software, numerous indicators of cellular physiology such as cell volume and cell length can be determined simultaneously in an automated fashion from phase contrast images of cells (Figure 3D).



**Figure 3: Evaluation of *E. coli* strains based on cellular growth or morphology characteristics.**

**A)** Growth curves of the AB1157 (black) and recombinered  $\Delta$ *motAB* cells (red) in shake flasks containing LB medium at 37 °C and 250 rpm. **B)** A simple method using the semi-log plot to find the log phase of the growth curve for AB1157 (black) and  $\Delta$ *motAB* (red) strains. The linear region of the semilog plot is the log phase of the growth curve. **C)** The exponential fitting of the selected log phase of the growth curve for AB1157 (black) and  $\Delta$ *motAB* (red) strains. From the fit (dotted lines), the growth rates ( $\mu$ ) are determined as  $1.11 \text{ h}^{-1}$  and  $1.13 \text{ h}^{-1}$  for one sample of AB1157 and  $\Delta$ *motAB* strains respectively. **D)** A sample phase contrast image of AB1157 cells which were grown in LB medium at 37 °C and 250 rpm is shown. Such images were analyzed by MicrobeTracker software to calculate precisely the cell length and volume for each cell. **E)** The data of cell volume and cell length of ~350 cells for each strain are plotted using a Box and Whiskers plot. The line within the box corresponds to the median value, the borders show the upper and lower quartiles (75% and 25%), and the whiskers represent the maximum and minimum values.

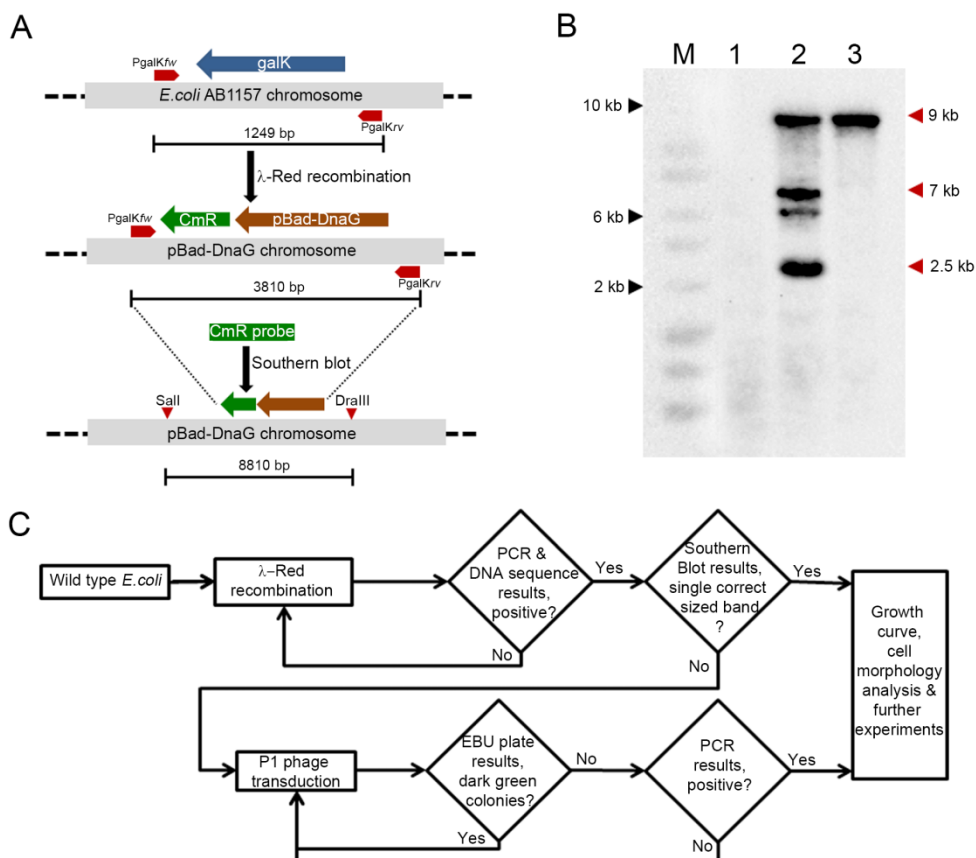
To illustrate this approach, we acquired images of AB1157 and  $\Delta motAB$  cells grown in LB medium at 37 °C, and for each strain, we analyzed approximately 350 cells. The mean cell volume (with SD) of AB1157 was found to be  $5.5 \pm 1.8 \mu\text{m}^3$  while that of  $\Delta motAB$  was  $5.3 \pm 1.6 \mu\text{m}^3$ . In the same analysis, the mean cell length for AB1157 strain was found to be  $3.3 \pm 0.9 \mu\text{m}$  (**Figure 3E**), and that of  $\Delta motAB$  was  $3.4 \pm 0.7 \mu\text{m}$  (**Figure 3F**). *t*-test statistics were used to determine any significant cell shape defects in the strains (**Table 3**) and revealed that the mean cell volume and cell length are not significantly different between the AB1157 and  $\Delta motAB$  strains using 95% confidence intervals from the observed *p*-values ( $\sim 0.1$ ).

**Table 3: “*t*-test” statistics for two independent samples of AB1157 and  $\Delta motAB$  strains’ generation time, cell volume and cell length**

Parameter	Mean		Standard Deviation ( $\sigma_{\text{Mean}}$ )		<i>t</i> -value	<i>p</i> -value
	AB1157	$\Delta motAB$	AB1157	$\Delta motAB$		
Sample	AB1157	$\Delta motAB$	AB1157	$\Delta motAB$		
Generation time (min)	39.2	38.7	2.1	1.9	0.43	0.68
Cell volume ( $\mu\text{m}^3$ )	5.5	5.3	1.8	1.6	1.55	0.12
Cell length ( $\mu\text{m}$ )	3.3	3.4	0.9	0.7	-1.64	0.1

#### 2.2.4. Application of the validation methods in a strain engineered at a different chromosome locus

To demonstrate the effectiveness of the validation methods described here, we applied them to a strain that is chromosomally engineered at a different locus using  $\lambda$ -Red recombination. In this strain, we replaced the endogenous non-essential *galk* gene in the *E. coli* AB1157 chromosome with an arabinose inducible primase gene (*pBad-DnaG*) along with *CmR* gene (**Figure 4A**). We performed Southern blotting and we observed from its results that extraneous insertions occurred in the created strain along with the intended insertion. We then applied P1 phage transduction using verified virulent P1 phages to transduce the intended insertion into the wild type *E. coli* AB1157 strain. Again using Southern blotting, we found that the P1 phage transduction step effectively produced the desired strain, fully devoid of extraneous insertions (**Figure 4B**).



**Figure 4: Application of the validation methods in a strain engineered at a different chromosome locus.** **A)** The scheme depicts the replacement of the endogenous non-essential *galk* gene with an arabinose inducible primase gene (*pBad-DnaG*) along with a *CmR* gene through  $\lambda$ -Red recombination. The positions of the flanking primers for the *galk* gene (*PgalKfv* and *PgalKrv*) are marked, and the corresponding product lengths from PCR are indicated, along with the probed region in the chromosome and the expected band size from Southern blot analysis. **B)** The Southern blot results for the different experiments. Lane 1: the AB1157 sample in lane 1 has no insert, as expected. Lane 2: the *pBad-DnaG* strain obtained through recombinering has multiple bands (mainly at  $\sim 9$  kbp, 7 kbp, 6.5 kbp, and 2.5 kbp). Lane 3: the *pBad-DnaG* strain after P1 phage transduction of the intended locus into the wild-type AB1157 strain displays one band ( $\sim 9$  kbp) at the right fragment size, showing that the extraneous insertions can be removed in the final strain using this approach. **C)** A flow diagram summarizing the sequence of the various validation techniques that should be performed prior to subsequent usage of the chromosomally engineered *E. coli* strain.

### 2.3. Conclusions

We have consolidated and explained in detail simple, yet reliable, validation techniques that may be applied to verify chromosomally engineered *E. coli* strains for non-specific insertions in the chromosome, temperate phage contamination and general phenotype defects in growth and cell shape. We have described the aspects of strain verification that

common approaches such as PCR and DNA sequence analysis do not report on, such as the presence of extraneous insertions verified by Southern blot analysis after recombineering. Additionally, we have also adapted and described in detail the EBU plate assay for the validation of *E. coli* strains created by the commonly used approach for chromosomal insertion, P1 phage transduction. In comparison to the cross-streak agar assay, we find the EBU plate assay to be more convenient and reliable. The validation methods discussed here are of widespread utility and can be applied to any chromosome engineering technique. A summary of the suggested workflow for the various validation steps that should be performed prior to using an *E. coli* strain for further experiments is shown in **Figure 4C**.

## 2.4. Methods

All the chemicals and biological reagents used in this study were ordered from Life Technologies (Europe). The *E. coli* strains used in this research are the AB1157 strain (43) and its derivatives. Optical density measurements were performed using the cell density meter Ultraspec 10™ from GE Healthcare Europe GmbH (The Netherlands). For each of the techniques described in the main text, brief, and specific descriptions are provided below, together with a detailed, stepwise protocol. Our adaptations or improvements to standardized protocols are indicated in italics.

### I. Strain creation methods

In this study, we performed  $\lambda$ -Red recombination (17) and P1 phage transduction (31) for the strain creation using the following protocols.

#### A. $\lambda$ -Red recombination (or recombineering)

This protocol is adapted from the literature (17, 24).

1. Perform PCR using a high-fidelity polymerase (e.g. Phusion polymerase from New England Biolabs® Inc) from a *linear or linearized DNA template* containing an antibiotic resistance gene (e.g. chloramphenicol resistance (*CmR*)) by specially designed primers (~70 bp) for generating homologous regions at the ends.
2. *Gel purify the PCR product.*
3. Transform a competent strain of interest with a temperature sensitive plasmid coding for  $\lambda$  phage's Red recombinase proteins (e.g. pKD46 (17)) as described (44) and grow the cells overnight on an agar plate at 30 °C in the presence of the corresponding antibiotic.
4. Pick a colony from this plate and grow the cells overnight in 5 mL LB medium containing respective antibiotics at 30 °C, *while shaking at 250 rpm.*
5. Prepare three samples of 1.4 mL LB medium with antibiotics, *50  $\mu$ L of the overnight culture of cells containing the pKD46 plasmid, 0.2% arabinose (or appropriate inducer used) along with one sample without arabinose in separate microfuge tubes.*
6. Incubate the cultures for *1.5 to 2 h at 37 °C*, shaking at 750 rpm, with the lids punctured.
7. Centrifuge the samples for 30 s at 12000 $\times$  g.

8. Discard the supernatant, and place the samples on ice.
9. Re-suspend the pellet in 1 mL chilled and sterile double distilled water.
10. Repeat the centrifugation and resuspend the pellet again in double distilled water *3 times* more.
11. Centrifuge the samples for 30 s at 12000× g, and *remove the supernatant until ~50 μL is left in the tubes.*
12. Re-suspend the pellet in remaining volume, and keep it on the ice.
13. *Add 1-3 μg of purified PCR product with a volume not more than 5 μL from step 2 to the samples on ice.*
14. *Include the following controls: a plasmid with same antibiotic selection marker as a positive control and sterile water as negative control.*
15. Pipette the entire volume from a sample to an electroporation cuvette, and electroporate the cells at 1250 Volts using an electroporator (Eppendorf®).
16. Re-suspend the cells in the cuvette with 1 mL of SOC medium *by pipetting up and down, and transfer them to a new sterile microfuge tube.*
17. *Incubate the samples in lid-punctured microfuge tubes for 70 min at 37 °C while shaking at 750 rpm, in order for the recombination to occur.*
18. Pour LB agar plates containing an appropriate concentration of the selection antibiotic. Do not add the antibiotics required for the temperature sensitive plasmid since the plasmid will be lost during the culture.
19. *Centrifuge the samples at 12000× g for 30 s, and remove 900 μL of the supernatant.*
20. Re-suspend the cells in remaining volume, and plate them onto LB agar plates.
21. Incubate the plates overnight at 37 °C.
22. *Re-streak 10 colonies on a plate with the selection antibiotic, and incubate the plate overnight at 37 °C.*

## B. P1 phage transduction

The protocol of P1 phage transduction that we adapted from the literature is explained here (31). The protocol consists of two steps: (i) P1 lysate preparation from donor strain and (ii) Phage transduction to the recipient strain. *However, the infectivity of the source P1 lysate stock should be determined first by using spot agar assay (Methods section IIC).*

### i. P1 Lysate preparation

1. Inoculate the recipient strain and the donor strain from the LB agar plates (e.g. from **Methods section IA** step 22) in 5 mL of LB medium with respective antibiotics, and grow the cells overnight at 37 °C with shaking at 250 rpm.
2. *Dilute in duplicate 0.5 mL of the overnight donor culture into 4.5 mL of LB medium containing 60 μL of 1M CaCl<sub>2</sub> and 120 μL of 1M MgSO<sub>4</sub>.*
3. Incubate the cultures at 37 °C for 45 min.
4. *Add 100 μL of the P1 phage lysate stock that has been prepared by infecting the wild type E. coli strain. The volume of P1 lysate used may vary depending on its infectivity (For example, if the infectivity value of P1 lysate is 10<sup>9</sup> pfu mL<sup>-1</sup>, then use 100 μL lysate).*

5. Continue the incubation until the culture is lysed or cell clumps are visible (usually around 3 to 4 h). The control culture without phages should show normal growth.
6. *Add 4 to 5 drops of chloroform to the lysed culture, and stir the mixture well using a vortex mixer.*
7. *After leaving the mixture to clarify for 5 min, transfer the upper liquid layer to a new sterile tube, and centrifuge the liquid for 20 min at  $4200\times g$ ,  $4\text{ }^{\circ}\text{C}$ .*
8. Pass the upper lysate layer further through a  $0.45\text{ }\mu\text{m}$  filter to remove any viable donor cells.
9. Store the P1 lysate devoid of chloroform at  $4\text{ }^{\circ}\text{C}$  for future use.
- ii. Transduction
  1. Dilute  $100\text{ }\mu\text{L}$  of the overnight recipient culture in  $900\text{ }\mu\text{L}$  LB medium containing  $75\text{ mM CaCl}_2$  and  $150\text{ mM MgSO}_4$  in 5 microfuge tubes.
  2. *Add  $5\text{ }\mu\text{L}$ ,  $50\text{ }\mu\text{L}$ ,  $100\text{ }\mu\text{L}$  or  $200\text{ }\mu\text{L}$  of P1 lysate both to the 4 tubes and, as a control, to the one tube without lysate. The volumes of P1 lysate used may vary depending on its infectivity. (For example, if the infectivity value of P1 lysate is  $10^9\text{ pfu mL}^{-1}$ , then use the volumes mentioned here).*
  3. Incubate the cultures at  $37\text{ }^{\circ}\text{C}$  while shaking at  $250\text{ rpm}$  for 30 min. Infection of recipient cells occurs in this step.
  4. Centrifuge the cells at  $12000\times g$  for 3 min, and discard the supernatant.
  5. Re-suspend the pellet in 1 mL of LB medium containing 20 mM sodium citrate ( $\text{pH } 5.5$ ) to reduce the infectivity of the adsorbed P1 phages by chelating the divalent ions. The transduction occurs during this step.
  6. Incubate the cells for 1.5 to 2 h at  $37\text{ }^{\circ}\text{C}$  with sufficient aeration and shaking at  $250\text{ rpm}$ .
  7. *Centrifuge the cells and discard the supernatant.*
  8. *Repeat step 6 and 7 twice to remove the phages as much as possible.*
  9. Re-suspend the pellet in  $100\text{ }\mu\text{L}$  LB medium containing 20 mM sodium citrate ( $\text{pH } 5.5$ ).
  10. Plate the cultures onto LB agar plates containing 20 mM sodium citrate with respective antibiotics for selection.
  11. *Re-streak  $\sim 16$  colonies onto LB agar plates containing 20 mM sodium citrate and selection antibiotics.*

## II. Strain verification methods

Here we describe the step-by-step methodology of the general validation techniques for chromosomal engineered *E. coli* strains.

### A. PCR and DNA sequencing

For the high throughput verification of the individual colonies that are obtained from the strain creation methods described above, a simple analytical PCR and DNA sequencing are the widely used verification steps. The accompanying steps are described below.

1. *Design primers for PCR in such a way that the sequence flanks the region of interest in the chromosome and the primer binding sites are not farther than 100 bases from the recombination site.*
2. *Re-suspend each colony to be verified by PCR in 50  $\mu$ L of sterile water, and streak 10  $\mu$ L on LB agar plates containing antibiotics.*
3. *Extract the DNA from the remaining cells into the water by boiling the samples for 5 min and centrifuging them at  $12000\times g$  for 1 min.*
4. *Perform a PCR reaction with this DNA using the primers designed in step 1.*
5. *Verify the length of the PCR products using agarose gel electrophoresis, and select colonies with proper insert length for DNA sequence analysis of the corresponding PCR product using the same flanking primers. *If necessary, use specific internal sequencing primers to verify the correct insertion in the chromosome.**
6. *Compare the DNA sequences by aligning it with the corresponding theoretical sequences to check for any point mutation or deletion introduced during the strain creation process.*

## **B. Southern blotting**

The copy number of the recombined DNA in the chromosome can be easily verified by using Southern blotting (34). In this study, we used the AlkPhos<sup>®</sup> Direct labeling and detection system manufactured by Amersham<sup>™</sup> (GE healthcare Europe GmbH, The Netherlands) because it is specially developed and well optimized for blotting experiments. DNA extraction was performed using Qiagen<sup>™</sup> DNA isolation kit. We used high fidelity restriction enzymes manufactured by New England Biolabs<sup>®</sup> Inc.

1. *Extract the genomic DNA from the strains to be verified. *Include the genomic DNA wild-type strain as a control.**
2. *Select two or three restriction enzymes using following criteria. *The restriction sites must flank the region of interest and should not be contained within the region itself. Since it is difficult to resolve DNA fragments larger than 10 kb through gel electrophoresis and to achieve the best resolution, it is a good practice to make sure the size difference between the restricted fragment and the region of interest (usually < 2 kb) does not exceed 7 kb.**
3. *Perform the restriction digestion of  $\sim 10\mu\text{g}$  genomic DNA samples overnight preferably using the high fidelity restriction enzymes to avoid star activity (45).*
4. *Separate the digested DNA samples in a 0.8% agarose gel by running electrophoresis overnight at a constant current of 15 mA.*
5. *Depurinate the DNA fragments by incubating the gel for 15 min in 0.1M HCl solution on a plate shaker, and wash subsequently four times with double distilled water.*
6. *Denature the fragments in denaturing buffer (1.5M NaCl and 0.5M NaOH) on a plate shaker for 15 min, and wash four times with double distilled water.*
7. *Incubate the gel for 15 min in neutralizing buffer (1.5M NaCl and 0.5M Tris base, pH 7.5) while shaking, and wash four times with double distilled water.*



8. Transfer the DNA fragments by capillary action to a pre-soaked Hybond-N+ membrane (GE Healthcare) using 20 x SSC buffer (3M NaCl and 0.3M trisodium citrate).
9. Pre-hybridize the blot with hybridization mix (35 mL AlkPhos® Direct hybridization buffer, 1g NaCl and 1.4g blocking agent-GE Healthcare) for 30 min at 55 °C in a rotary mixer.
10. Amplify the region of interest using specific PCR primers from the source DNA (usually a template plasmid or wild-type chromosome). The optimal size of the PCR product is ~200 to 1000 bp.
11. Denature the PCR product by boiling for 5 min, chill it on ice, and label it using appropriate reporters (e.g. *thermostable alkaline phosphatase*) that can catalyze non-luminescent substrates and yield luminescent products. The labeled PCR product can then be used as a DNA probe.
12. Add the DNA probe to the membrane in the hybridization buffer, and hybridize the probe in a rotary mixer at 55 °C overnight.
13. Wash the membrane in rotary shaker at 55 °C with 100 mL wash buffer 1 for 10 min (pH: 7; 2 M Urea, 0.1% SDS, 0.15 M NaCl, 0.05 M NaH<sub>2</sub>PO<sub>4</sub>, 1 mM MgCl<sub>2</sub>, and 1g Blocking reagent)
14. Wash the membrane twice with 100 mL wash buffer 2 (pH: 10; 3g Tris base, 2.8g NaCl, and 0.2M MgCl<sub>2</sub>) in plate shaker at room temperature for 10 min.
15. Incubate the membrane, with 3 mL non-luminescent substrate for 5 min and dry the membrane.
16. Wrap the membrane using Saran™ wrap, and detect chemiluminescence on the blot using appropriate detectors. *The resulting number of bands obtained is indicative of the copy number of the recombined DNA in the chromosome.*

### C. Spot agar assay

The infectivity of the phages in a P1 lysate can be determined by spot agar assay. The method is adapted from Ref. (31).

1. Add CaCl<sub>2</sub> to the overnight culture of the recipient strain from the phage transduction step such that the final concentration is 5 mM.
2. Serially dilute the P1 lysate to the order of 10<sup>-10</sup> using LB medium containing 75 mM CaCl<sub>2</sub> and 150 mM MgSO<sub>4</sub>. *Make sure to change the pipette tips during the dilution step.*
3. Mix 0.25 mL of cell culture with 2.5 mL of molten LB top agar (0.75%) containing 2.5 mM CaCl<sub>2</sub>.
4. Pour the mixture onto LB agar (1.5%) plates containing 2.5 mM CaCl<sub>2</sub> and let it solidify.
5. Spot 10 µL of each phage stock dilution onto the lawn of cells.
6. Keep the plates upright, and after the spots are dry, incubate them at 37 °C overnight.
7. Calculate the titer value in pfu mL<sup>-1</sup> by counting the number of plaques in the lowest concentration spot, then multiplying it with the order of dilution (e.g. 10<sup>7</sup>) and finally by the factor 100 (to account for 10 µL volume used).

## D. Cross-streak agar assay

The presence of temperate phages in phage-transduced colonies can be tested by cross-streak agar assay. The method is adapted from Ref.(25).

1. Prepare an LB agar plate containing  $2.5 \text{ mM CaCl}_2$ .
2. Draw a straight line on the back of plate across the middle and well-spaced dots on one side (**Figure 2B**).
3. Holding the plate in a slanting position, pour  $50 \mu\text{L}$  of phage lysate on the LB agar from one end of the straight line.
4. Tilt the plate back to uniformly distribute the lysate around the straight line.
5. Take a colony or liquid overnight culture to be tested for temperate phages with the broad side of the inoculation needle.
6. Place a dot of the colony or culture at the marked location and leaving a few millimeters space streak the colony or culture perpendicularly across the phage lysate.
7. Transfer the plate carefully to the incubator, and incubate overnight at  $37^\circ\text{C}$ .

## E. Evans Blue-Uranine (EBU) plate assay

The phage-transduced colonies can be tested for temperate phage contamination and can be cured of phages using EBU plate assay described below. The method is adapted from Ref. (46) for *E. coli* strains.

1. Make agar plates from 1L of molten LB agar containing 40 mL of 12.5%  $\text{K}_2\text{HPO}_4$ , 1.25 mL of 1% Evans Blue stain solution, 250  $\mu\text{L}$  of 10% uranine solution and if necessary antibiotics.
2. *Store the plates in dark at  $4^\circ\text{C}$ . Since Evans Blue stain is carcinogenic the plates must be handled carefully with gloves, and it is usually a good practice to autoclave and dispose of the carcinogenic waste.*
3. *Inoculate the colonies obtained from a phage transduction experiment in LB medium, and grow them for 3 h at  $37^\circ\text{C}$  with shaking at 250 rpm.*
4. *Dilute the exponentially growing LB cultures 100 times, and spread  $50 \mu\text{L}$  of the culture uniformly across the EBU plates.*
5. *Incubate overnight at  $37^\circ\text{C}$  in the dark.*
6. *Analyze the colonies. The colonies without temperate phages appear pale yellow in color while those with temperate phages are colored dark green.*

## F. Growth curve analysis

Growth curve analysis of the *E. coli* strains can be performed using shake flasks as described below. In this study, we used MS Excel and MATLAB to analyze the data.

1. *Dilute in triplicate overnight cultures of the wild type and the strain to be verified to an  $OD_{550\text{nm}}$  of 0.01 in 50 mL fresh medium.*
2. *Grow the cells at  $37^\circ\text{C}$  in a shaking incubator (250 rpm).*
3. *Use sterile disposable cuvettes and aseptic conditions to note the  $OD_{550\text{nm}}$  of a 1mL aliquot from each culture sample using a cell density meter.*

4. *Add back the culture aliquot to the sample after the measurement in order to keep the culture volume constant.*
5. *Steps 3 and 4 are repeated at an interval of 15 min until the OD<sub>550nm</sub> is a constant value.*
6. *Plot the OD<sub>550nm</sub> values against time in minutes to get a sigmoidal curve in this plot.*
7. *Plot the OD<sub>550nm</sub> values in logarithmic scale, the exponential phase of the growth curve can be identified as the distinct linear part (Figure 3B,C), and perform an exponential fit only on the OD<sub>550nm</sub> values of this phase in the growth curve for each sample to find the growth rate ( $\mu$ ) (38).*
8. *Calculate the generation time ( $\tau_d$ ) from  $\mu$  using the formula:  $\tau_d = \ln(2)/\mu$ .*
9. *To test the reproducibility of results, repeat the experiment, and average the generation times for 6 samples of the strain to be verified. The result can then be compared to the average generation time of the wild type strain using a t-test for two independent sample means (39).*

## G. Detection of cell shape defects by microscopy

Phase-contrast microscopy is a powerful technique to determine the cell shape characteristics of a bacterium. In this study, it was performed using a Nikon Ti™ microscope with a 100x objective and an Andor iXon™ Ultra 897 EMCCD camera.

- 1 *Grow the LB culture of the strain to be verified until it reaches the exponential phase OD<sub>550nm</sub> is ~ 0.3 to 0.4.*
- 2 *Centrifuge 1 mL of the culture at 12000× g for 1 min, and remove the supernatant.*
- 3 *Re-suspend in 250 μL of fresh medium, and pipette 5 μL of concentrated culture on a pad of agarose (1.5%) flattened on a microscopic slide. Let the fluid evaporate.*
- 4 *Place a sterile and clean cover glass on top of the cells, and press gently using tweezers.*
- 5 *Analyze the cells on a phase contrast microscope using an objective with a magnification >60x and a digital camera.*
- 6 *For a robust analysis, acquire the images with at least ~100 separate cells within a field of view.*
- 7 *Store the images in greyscale tiff format, and analyze them using MicrobeTracker software to accurately determine the cell shape defects (42).*
- 8 *Export the analyzed data from the software in CSV format for further analysis e.g. t-test statistics and representation of data, e.g. a Box and Whiskers plot.*

## 2.5. Contributions

STK carried out the experiments detailed in this research and drafted the manuscript; MCM aided in the conception of the study, design of microscopy experiments, and manuscript writing; TL aided in the troubleshooting of experiments; ASM provided suggestions in the conception of the study, design of P1 phage-related experiments, and manuscript writing;

and NHD contributed in the conception of the study, general design of experiments, and manuscript writing.

## 2.6. Acknowledgments

We thank Prof. David Sherratt for providing the AB1157 strain, beta-clamp labeled strain and P1 lysates for the experiments, Dr. Richard Janissen for helping in the data analysis and presentation, and Nienke Blokker for contributing to the growth curve analysis. This work was supported by the Netherlands Organisation for Scientific Research (NWO) through a Vici grant to NHD, Delft University of Technology, and the European Community's Seventh Framework Program FP7/2007–2013 under grant agreements n° 241548 (MitoSys).

## 2.7. References

1. E. C. Lee *et al.*, A Highly Efficient *Escherichia coli*-Based Chromosome Engineering System Adapted for Recombinogenic Targeting and Subcloning of BAC DNA. *Genomics* **73**, 56-65 (2001).
2. F. R. Blattner *et al.*, The Complete Genome Sequence of *Escherichia coli* K-12. *Science* **277**, 1453-1462 (1997).
3. H. M. Ellis, D. Yu, T. DiTizio, D. L. Court, High efficiency mutagenesis, repair, and engineering of chromosomal DNA using single-stranded oligonucleotides. *Proceedings of the National Academy of Sciences* **98**, 6742-6746 (2001).
4. D. Boyd, D. S. Weiss, J. C. Chen, J. Beckwith, Towards Single-Copy Gene Expression Systems Making Gene Cloning Physiologically Relevant: Lambda InCh, a Simple *Escherichia coli* Plasmid-Chromosome Shuttle System. *Journal of Bacteriology* **182**, 842-847 (2000).
5. R. Reyes-Lamothe, D. J. Sherratt, M. C. Leake, Stoichiometry and Architecture of Active DNA Replication Machinery in *Escherichia coli*. *Science* **328**, 498-501 (2010).
6. R. Reyes-Lamothe, C. Possoz, O. Danilova, D. J. Sherratt, Independent Positioning and Action of *Escherichia coli* Replisomes in Live Cells. *Cell* **133**, 90-102 (2008).
7. S. Bakshi, A. Siryaporn, M. Goulian, J. C. Weisshaar, Superresolution imaging of ribosomes and RNA polymerase in live *Escherichia coli* cells. *Molecular Microbiology* **85**, 21-38 (2012).
8. M. Kitagawa *et al.*, Complete set of ORF clones of *Escherichia coli* ASKA library (A Complete Set of *E. coli* K-12 ORF Archive): Unique Resources for Biological Research. *DNA Research* **12**, 291-299 (2006).

9. I. F. Lau *et al.*, Spatial and temporal organization of replicating *Escherichia coli* chromosomes. *Molecular Microbiology* **49**, 731-743 (2003).
10. D. Greenfield *et al.*, Self-Organization of the *Escherichia coli* Chemotaxis Network Imaged with Super-Resolution Light Microscopy. *PLoS Biol* **7**, e1000137 (2009).
11. N. J. Delalez, R. M. Berry, J. P. Armitage, Stoichiometry and Turnover of the Bacterial Flagellar Switch Protein FliN. *mBio* **5**, (2014).
12. T. Baba *et al.*, Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular Systems Biology* **2**, (2006).
13. M. Juhas, D. Reuß, B. Zhu, F. M. Commichau, *Bacillus subtilis* and *Escherichia coli* essential genes and minimal cell factories after one decade of genome engineering. *Microbiology*, (2014).
14. L. R. Jarboe *et al.*, Metabolic Engineering for Production of Biorenewable Fuels and Chemicals: Contributions of Synthetic Biology. *Journal of Biomedicine and Biotechnology* **2010**, 18 (2010).
15. D. Koma, H. Yamanaka, K. Moriyoshi, T. Ohmoto, K. Sakai, A convenient method for multiple insertions of desired genes into target loci on the *Escherichia coli* chromosome. *Appl Microbiol Biotechnol* **93**, 815-829 (2012).
16. T. Maeda, V. Sanchez-Torres, T. K. Wood, Hydrogen production by recombinant *Escherichia coli* strains. *Microbial Biotechnology* **5**, 214-225 (2012).
17. K. A. Datsenko, B. L. Wanner, One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proceedings of the National Academy of Sciences* **97**, 6640-6645 (2000).
18. Y. Zhang, F. Buchholz, J. P. P. Muyrers, A. F. Stewart, A new logic for DNA engineering using recombination in *Escherichia coli*. *Nat Genet* **20**, 123-128 (1998).
19. F. Martinez-Morales, A. C. Borges, A. Martinez, K. T. Shanmugam, L. O. Ingram, Chromosomal Integration of Heterologous DNA in *Escherichia coli* with Precise Removal of Markers and Replicons Used during Construction. *Journal of Bacteriology* **181**, 7143-7148 (1999).
20. T. E. Kuhlman, E. C. Cox, Site-specific chromosomal integration of large synthetic constructs. *Nucleic Acids Research* **38**, e92 (2010).
21. D. Chen, Z. Yan, D. L. Cole, G. S. Srivatsa, Analysis of internal (n-1)mer deletion sequences in synthetic oligodeoxyribonucleotides by hybridization to an immobilized probe array. *Nucleic Acids Research* **27**, 389-395 (1999).

22. J. L. Hobman *et al.*, Comparative Genomic Hybridization Detects Secondary Chromosomal Deletions in Escherichia coli K-12 MG1655 Mutants and Highlights Instability in the flhDC Region. *Journal of Bacteriology* **189**, 8786-8792 (2007).
23. N. Yamamoto *et al.*, Update on the Keio collection of Escherichia coli single-gene deletion mutants. *Mol Sys Biol* **5**:335 (2009).
24. S. K. Sharan, L. C. Thomason, S. G. Kuznetsov, D. L. Court, Recombineering: a homologous recombination-based method of genetic engineering. *Nat. Protocols* **4**, 206-223 (2009).
25. S. D. Moore, in *Strain Engineering*, J. A. Williams, Ed. (Humana Press, New York (USA), 2011), vol. 765, chap. 10, pp. 155-169.
26. E. S. Lennox, Transduction of linked genetic characters of the host by bacteriophage P1. *Virology* **1**, 190-206 (1955).
27. H. Ikeda, J. I. Tomizawa, Transducing fragments in generalized transduction by phage P1. 3. Studies with small phage particles. *J Mol Biol* **14**, 120-129 (1965).
28. B. M. Tyler, R. B. Goldberg, Transduction of chromosomal genes between enteric bacteria by bacteriophage P1. *J Bacteriol* **125**, 1105-1111 (1976).
29. N. Sternberg, R. Hoess, The molecular genetics of bacteriophage P1. *Annu Rev Genet* **17**, 123-154 (1983).
30. G. Satta, E. Debbia, C. Pruzzo, L. Calegari, The peculiar behaviour of coliphage P1vir mutants on restricting hosts. *Microbios* **22**, 93-102 (1978).
31. L. C. Thomason, N. Costantino, D. L. Court, in *Current Protocols in Molecular Biology*. (John Wiley & Sons, Inc., New York (USA), 2001).
32. M. D. Manson, Dynamic motors for bacterial flagella. *Proceedings of the National Academy of Sciences* **107**, 11151-11152 (2010).
33. F. Sanger, S. Nicklen, A. R. Coulson, DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences* **74**, 5463-5467 (1977).
34. E. M. Southern, Detection of specific sequences among DNA fragments separated by gel electrophoresis. *Journal of Molecular Biology* **98**, 503-517 (1975).
35. N. D. Zinder, J. Lederberg, Genetic exchange in *Salmonella*. *J Bacteriol* **64**, 679-699 (1952).
36. B. R. Bochner, Curing bacterial cells of lysogenic viruses by using UCB indicator plates. *Biotechniques*, 234-240 (1984).

37. M. H. Zwietering, I. Jongenburger, F. M. Rombouts, K. van 't Riet, Modeling of the Bacterial Growth Curve. *Applied and Environmental Microbiology* **56**, 1875-1881 (1990).
38. B. G. Hall, H. Acar, A. Nandipati, M. Barlow, Growth Rates Made Easy. *Molecular Biology and Evolution* **31**, 232-238 (2014).
39. J. H. McDonald, *Handbook of Biological statistics*. (Sparky House Publishing, Baltimore, 2008).
40. A. I. Konokhova, A. A. Gelash, M. A. Yurkin, A. V. Chernyshev, V. P. Maltsev, High-precision characterization of individual *E. coli* cell morphology by scanning flow cytometry. *Cytometry Part A* **83A**, 568-575 (2013).
41. D. Osiro, R. Bernardes Filho, O. B. G. Assis, L. A. d. C. Jorge, L. A. Colnago, Measuring bacterial cells size with AFM. *Brazilian Journal of Microbiology* **43**, 341-347 (2012).
42. O. Sliusarenko, J. Heinritz, T. Emonet, C. Jacobs-Wagner, High-throughput, subpixel precision analysis of bacterial morphogenesis and intracellular spatio-temporal dynamics. *Molecular Microbiology* **80**, 612-627 (2011).
43. S. K. DeWitt, E. A. Adelberg, The occurrence of a genetic transposition in a strain of *Escherichia coli*. *Genetics* **47**, 577-585 (1962).
44. C. A. Woodall, in *Methods in Molecular Biology*. (2003), vol. 235, pp. 55-59.
45. N. Kamps-Hughes, A. Quimby, Z. Zhu, E. A. Johnson, Massively parallel characterization of restriction endonucleases. *Nucleic Acids Research* **41**, e119-e119 (2013).
46. S. R. Maloy, V. J. Stewart, R. K. Taylor, *Genetic analysis of pathogenic bacteria.*, (Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press., 1996).

# 3 SLOW UNLOADING LEADS TO DNA-BOUND $\beta_2$ SLIDING CLAMP ACCUMULATION IN LIVE *ESCHERICHIA COLI* CELLS

*The ubiquitous sliding clamp facilitates processivity of the replicative polymerase and acts as a platform to recruit proteins involved in replication, recombination, and repair. While the dynamics of the E. coli  $\beta_2$  sliding clamp proteins have been characterized in vitro, their stoichiometry and dynamics remain unclear in vivo. To probe both  $\beta_2$  clamp dynamics and stoichiometry in live E. coli cells, we use custom-built microfluidics in combination with single-molecule fluorescence microscopy and photoactivated fluorescence microscopy. We quantify the recruitment, binding, and turnover of  $\beta_2$  sliding clamps on DNA during replication. These quantitative in vivo results demonstrate that numerous  $\beta_2$  clamps in E. coli remain on the DNA behind the replication fork for a protracted period of time, allowing them to form a docking platform for other enzymes involved in DNA metabolism.*

---

This chapter has been published as: M. Charl Moolman, Sriram Tiruvadi Krishnan, Jacob W.J. Kerssemakers, Aafke van den Berg, Pawel Tulinski, S. Martin Depken, Rodrigo Reyes-Lamothe, David J. Sherratt, and Nynke H. Dekker, Slow unloading leads to DNA-bound  $\beta_2$  sliding clamp accumulation in live *Escherichia coli* cells. *Nature Communications*, **5**, 5820 (2014)



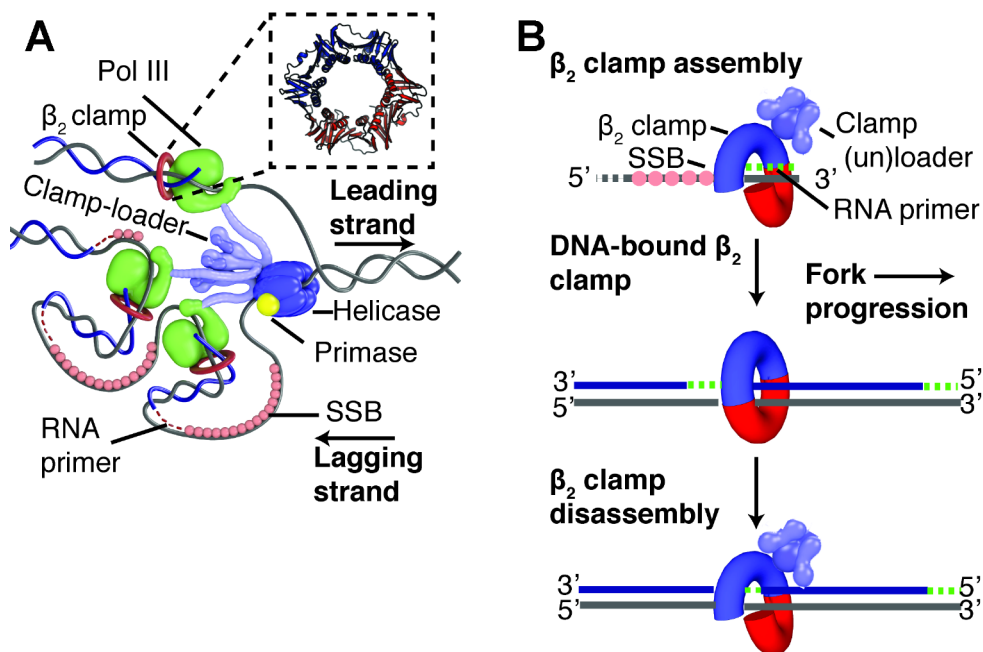
### 3.1. Introduction

The multi-protein replisome complex (replisome, **Figure 1A**) is responsible for the accurate and timely duplication of the genome prior to cell division. The sliding clamp protein complex is a key subunit of the replisome and is vital for protein-DNA interactions related to DNA metabolism in all three domains of life (1-3). Through their interaction with polymerases, DNA ligase, replication initiation protein DnaA, the dynamin-like protein CrfC, as well as different mismatch-repair proteins, sliding clamps play important roles in replication, and repair (4-15). In *E. coli*, the  $\beta_2$  sliding clamp ( $\beta_2$  clamp) is a homodimer (16) (**Figure 1A (inset)**) that encircles double-stranded DNA (dsDNA) and tethers DNA Polymerase III (DNA Pol III) to the template, thereby ensuring sufficiently high processivity during synthesis (17, 18).

The  $\beta_2$  sliding clamp is actively assembled and disassembled onto DNA during synthesis of the two complementary DNA strands (**Figure 1B**). The loading reaction of a  $\beta_2$  clamp onto each new primer-template junction (19) is catalyzed by an ATP-dependent heteropentameric clamp-loader complex (clamp-loader), also known as the  $\gamma$ -complex (20). The clamp-loader pries open the  $\beta_2$  clamp, recognizes the primer-template junction (21), and closes the  $\beta_2$  clamp around the dsDNA prior to release (22). The clamp-loader is also thought to chaperone DNA Pol III onto a newly loaded  $\beta_2$  clamp (23) and to unload inactive DNA-bound  $\beta_2$  clamps via the  $\delta$ -subunit (24). During all of these reactions, the loader complex, and the various clamp binding proteins compete for the C-terminal face of the clamp. In accordance with the proposed model in which the replisome includes three core DNA polymerase III's (25, 26), three  $\beta_2$  clamps can be active at the replication fork, one for each of the three polymerases (**Figure 1A**). While leading-strand replication is thought to be continuous, utilizing only a single  $\beta_2$  clamp, the lagging-strand template is copied in discrete 1 - 2 kb Okazaki fragments (27), each utilizing a separate  $\beta_2$  clamp. These fragments are initiated by the continuous formation of 10 - 12 nt RNA primers by the primase (DnaG) which, together with the helicase (DnaB), sets the replication fork clock (28). Since the number of Okazaki fragments (2000-4000) for the 4.6 Mbp genome is roughly an order of magnitude greater than the average number of  $\beta_2$  clamps per cell in a nutrient-rich culture (24, 29), continuous recycling of  $\beta_2$  clamps is necessary for total genome replication to occur.

Despite numerous *in vitro* and *in vivo* studies, it still remains unclear whether recycling of the *E. coli*  $\beta_2$  clamps takes place immediately following the completion of an Okazaki fragment, or at a later time. A slow recycling could permit a  $\beta_2$  clamp to fulfill additional functions while remaining bound to the newly synthesized DNA. Quantitative *in vitro* unloading assays (24, 30) indicate that in the absence of the clamp-loader, a loaded  $\beta_2$  clamp has a long half-life of  $t_{\text{unload}} > 1$  hr on the DNA (More information on the definition of unloading time is given in the **Supplementary section S9**). While this is decreased by more than an order of magnitude to  $t_{\text{unload}} \sim 127$  s/ $\beta_2$  in the presence of clamp-loader, this unloading time still remains long compared with the typical time required to complete an Okazaki fragment (on the order of seconds). Such a slow unloading time suggests that many  $\beta_2$  clamps are left behind in the wake of the replication fork (31). However, a recent *in vitro* single-molecule study indicates that lagging-strand synthesis can persist *in vitro* in the absence of excess  $\beta_2$  clamps in solution, implying that a  $\beta_2$  clamp can be directly reused at a successive primer-template junction (32). Two *in vivo* studies, one in *Bacillus subtilis*

(*B. subtilis*) (33) and the other in *E. coli* (26), provided contrasting results. Hence, to understand the regulatory mechanism that underlies the recycling of  $\beta_2$  clamps in *E. coli*, further insights into their *in vivo* dynamics are required.



**Figure 1: The *E. coli* replisome and  $\beta_2$  clamp assembly during replication.** (A) The position of the  $\beta_2$  sliding clamp within the *E. coli* replisome complex. The helicase (DnaB) unwinds dsDNA ahead of the replicative polymerase (DNA Pol III) which subsequently duplicates the template strands. Different configurations of Pol III are potentially possible. Primase (DnaG) synthesizes short RNA primers on the lagging strand for Okazaki fragment initiation. Single-stranded binding proteins (SSB) remove the secondary structure of ssDNA and protect it from digestion. To ensure sufficient processivity during replication, Pol III is tethered to the DNA by the  $\beta_2$  sliding clamp.  $\beta_2$  is assembled onto primer-template junctions by the multi-protein  $((\tau/\gamma)_3\delta\delta'\psi\gamma)$  clamp-loader complex. (Inset) A ribbon representation of the DNA-bound  $\beta_2$  sliding clamp (generated using the Protein Data Bank (PDB) file, 2POL (16)). The  $\beta_2$  sliding clamp is a homodimer that consists of six globular domains (16). The monomers are arranged in a ring that encircles the DNA (57) and can slide freely along it. Different proteins can bind to the two hydrophobic pockets of the  $\beta_2$  clamp via a conserved sequence motif (10). (B) The life cycle of the  $\beta_2$  clamp during replication. (top) The  $\beta_2$  clamp is actively loaded by the clamp-loader, which opens the closed clamp and places it onto dsDNA prior to release (22). (middle) The  $\beta_2$  clamp remains DNA-bound as long as an Okazaki fragment is being synthesized. (bottom) After the  $\beta_2$  clamp has reached the end of an Okazaki fragment, DNA Pol III is signaled to release (50). The  $\beta_2$  clamp is believed to be disassembled by the clamp-loader.

To gain detailed insight into the *in vivo* recruitment and turnover of the  $\beta_2$  clamp, we investigate its dynamics in individual live cells with single-molecule sensitivity. We employ both conventional fluorescence microscopy and Photoactivated Localization Microscopy (PALM) (34, 35), in combination with custom-built microfluidics. Single-molecule techniques have provided us with insights into the dynamics of processes – such

as replication, transcription, and translation – that are not readily accessible with conventional ensemble-averaging techniques (36). *In vivo* single-molecule fluorescence imaging, in particular, has provided detailed insights into the behavior of individual molecules in live cells (37-39). Combining single-molecule fluorescence microscopy with microfluidics allows us to image individual molecules in live cells over multiple cell cycles, without chemical fixation that could potentially perturb the dynamic behavior of the protein under investigation (40).

By employing this experimental approach, we have measured the number of DNA-bound  $\beta_2$  clamps during chromosomal replication over the entire course of the cell cycle. Additionally, we have determined the time required to unload an individual DNA-bound  $\beta_2$  clamp during replication, as well as the effective time required to load a new  $\beta_2$  clamp during replication. Our data reveal that the number of DNA-bound  $\beta_2$  clamps accumulates on the DNA after initiation and then levels off to a constant steady state number of DNA-bound  $\beta_2$  clamps on the order of minutes. This steady state is maintained throughout the rest of the replication process until termination occurs and a concomitant decrease of DNA-bound  $\beta_2$  clamps is observed. The number of DNA-bound  $\beta_2$  clamps in the steady state exceeds the estimates of a previously published *in vivo* study (26) by an order of magnitude. The measured values for the effective loading time and unloading time during replication, in the context of the live cell, are in good agreement with previous biochemical *in vitro* experiments (24). Taken together, our data indicate a  $\beta_2$  clamp remains on the DNA for a protracted period of time following the completion of an Okazaki fragment. DNA-bound  $\beta_2$  clamps that are left behind during fork progression may facilitate the recruitment of additional proteins active during the cell cycle for different processes such as DNA repair.

## 3.2. Results

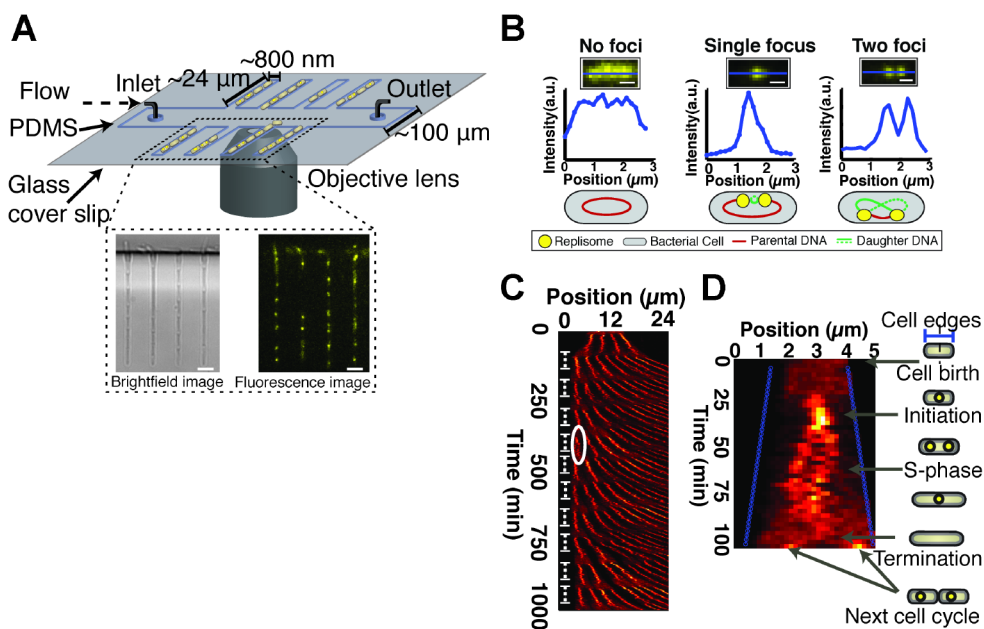
### 3.2.1. The *in vivo* dynamics of $\beta_2$ clamps measured in single cells

To study the dynamics of  $\beta_2$  clamps by wide-field fluorescence microscopy, we perform long time-lapse imaging of labeled  $\beta_2$  clamps over multiple replication cycles. During such experiments, we ensure healthy cell physiology by implementing a custom-built microfluidic device (**Figure 2A**; See **Materials and Methods**) (41, 42) in which cells growing in steady state are immobilized in micron-sized growth channels.

Through a neighboring central trench, the growth medium is continuously supplied throughout an experiment (See **Materials and Methods**). In such a microfluidic device, cells experience minimal perturbation over the course of the time-lapse experiment, as stable growth conditions remain continuously present. This contrasts with long time-lapse experiments performed on agarose pads in which nutrients and water may become depleted, leading to non-steady state cell populations as a result. Additional benefits of such a device are that daughter cells ultimately grow out of the growth channels, preventing accumulation of cells and that the cells are always aligned, which facilitates data analysis.

Labeling of the  $\beta_2$  clamp was accomplished by using a functional N-terminal YPet (43) fusion (26) (**Supplementary Figure 6**; **Supplementary Table 1**; See **Materials and Methods**) expressed from (and replacing) the endogenous *E. coli dnaN* gene locus.

Fluorescence images are acquired under shuttered 515 nm laser excitation (See **Materials and Methods**; **Figure 2A (inset)**). Fluorescence images of YPet- $\beta_2$  within individual cells either yielded no (**Figure 2B, left**), a single (**Figure 2B, middle**), or two cellular foci (**Figure 2B, right**), depending on the stage of replication, in agreement with previous reports of fluorescently labeled replisome components (44). Prior to each fluorescence image, a brightfield image is acquired to provide details of the cell periphery (**Figure 2A (inset)**). This alternating imaging sequence has a sufficiently long period to avoid giving rise to any notable deleterious growth effects, as assessed by comparing the doubling time of cell growth in a shake flask with cells grown in the microfluidic device (**Supplementary Figure 6, 7**).



**Figure 2: Long time-lapse fluorescence microscopy of the  $\beta_2$  sliding clamp at the single-cell level utilizing microfluidics.** (A) The microfluidic device used for performing long-time lapse fluorescence microscopy. *E. coli* cells are immobilized in growth channels perpendicular to the main trench through which growth medium is actively pumped. (inset) A brightfield image and corresponding YPet- $\beta_2$  fluorescence image (80 ms laser light exposure) are acquired every 2.5 min for the duration of the time-lapse experiment. Scale bars, 3  $\mu\text{m}$ . (B) YPet- $\beta_2$  molecules that are either DNA-bound or freely diffusing are studied using wide-field fluorescence microscopy. (left) Freely diffusing YPet- $\beta_2$  molecules in the cytoplasm of a cell. This signal is representative to YPet- $\beta_2$  dynamics prior to and after replication. (middle) A clear focus is observed due to DNA-bound YPet- $\beta_2$  molecules. The observation of a single focus, instead of two distinct foci, shortly after initiation results from the overlap of diffraction limited spots. (right) Two distinct foci are visible, indicative of two individual replisomes. Scale bars, 800 nm. (C) Kymograph of a single growth channel during an overnight time-lapse experiment. The cells first grow the growth channel full and maintain a steady state growth rate as can be observed from the curved shape of the fluorescence signal. The shape of the fluorescence signal is due to the individual cells growing and pushing each other in the direction of the main trench. Clear observable diffuse patterns occur at regular intervals, indicative of no DNA-bound  $\beta_2$  clamps prior to initiation or after termination. This repeating pattern is due to the multiple

cycles of replication (indicated with repeating white dashed lines). **(D)** A kymograph of an individual replication cycle indicated in panel (C). The blue lines are the cell boundaries detected from the brightfield images. The illustrations on the right-hand side indicate the different stages of replication that can be observed during the cell cycle.

Using this approach, we are able to observe numerous consecutive replication (and corresponding division) cycles of cells in the different growth channels. We examine the global replication dynamics of multiple cells within a growth channel by converting the time-lapse images into a kymograph (**Figure 2C**; See **Materials and Methods**). A distinct reoccurring pattern indicative of multiple replication cycles in generations of cells is clearly noticeable (indicated in repeating dashed lines in **Figure 2C**). Under these experimental conditions (See **Materials and Methods**), the analysis of individual cells ( $n = 137$ ) in our microfluidic system yields an average replication time (C-period) of  $t_{\text{rep}} = 68 \pm 10$  min, and a doubling time of  $t_{\text{double}} = 84 \pm 17$  min (**Supplementary Figure 7**). In both cases, the error is  $\pm$  S.D. We further analyze these kymographs to investigate the subcellular dynamics of the YPet- $\beta_2$  molecules within individual cells from cell birth till cell division (**Figure 2D**). One can clearly observe the dynamic of the two  $\beta_2$  clamp foci associated with the two independent replisomes.

### 3.2.2. The assembly and accumulation of $\beta_2$ clamp proteins on DNA

We use the fluorescence intensity from the YPet- $\beta_2$  fusion to determine the number of  $\beta_2$  clamps that are DNA-bound as well as in the total number in the cell during the life cycle of a cell. A sample montage of the YPet- $\beta_2$  fluorescence signal from cell birth till after cell division (**Figure 3A**) illustrates that there is a distinct increase in the foci following the B-period (45) of the cell cycle (represented by a diffuse signal after birth) (**Figure 3A (inset)**) and a similar decrease prior to cell division. The fraction of fluorescence that originates from DNA-bound YPet- $\beta_2$  (foci) provides clear evidence that  $> 50\%$  of  $\beta_2$  clamps are DNA-bound shortly after initiation of replication (**Figure 3B**). The steady decline in the fraction of DNA-bound  $\beta_2$  clamps that commences roughly 10 min after initiation, results from the increase in a total number of  $\beta_2$  clamps in the cell as the cell grows. In assessing this intensity fraction, we verified that very little out-of-focus fluorescence escapes detection (**Supplementary Section S1**; **Supplementary Figure 8**).

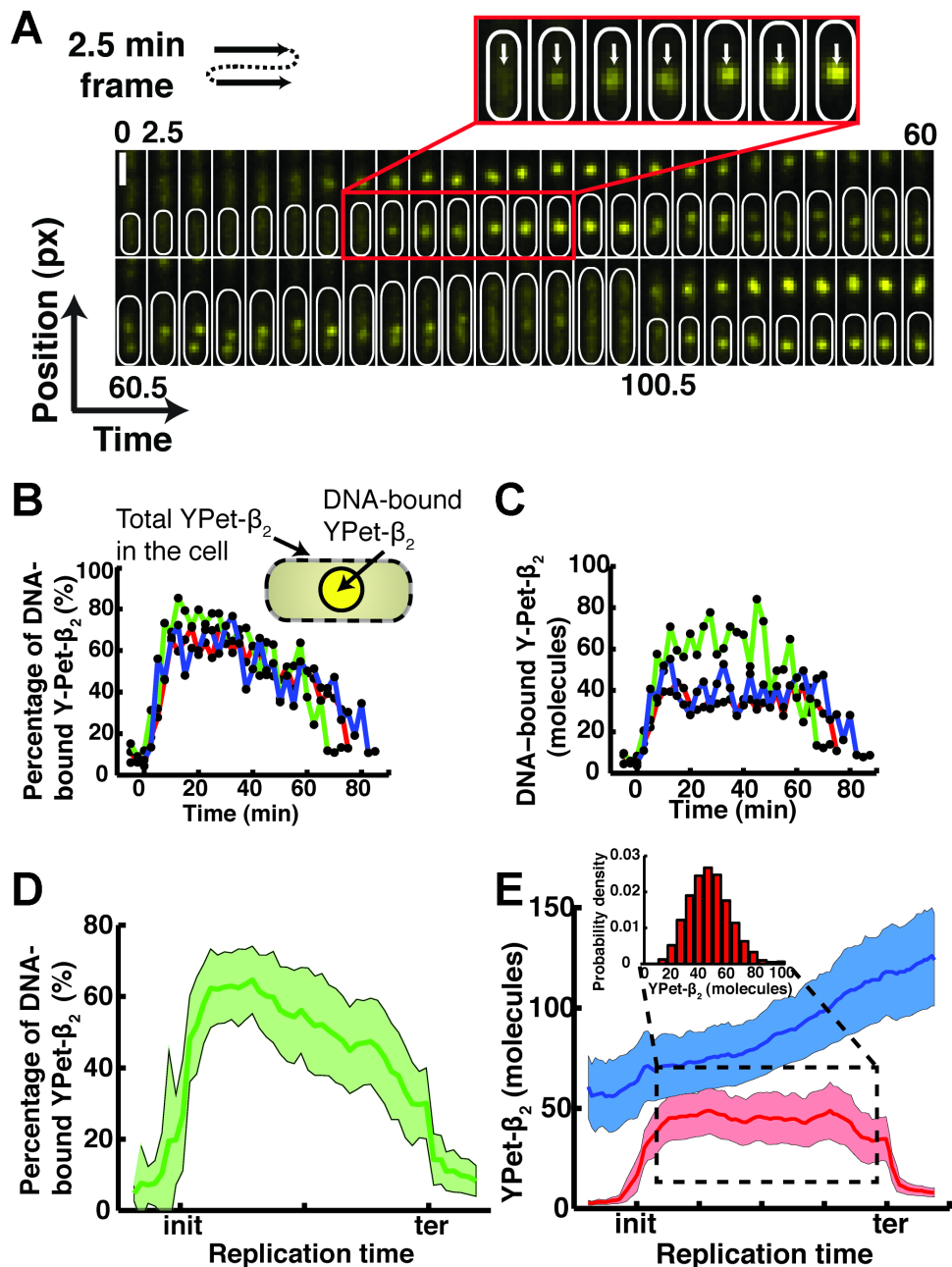
### 3.2.3. A constant number of DNA-bound $\beta_2$ clamps is maintained.

To precisely quantify the number of DNA-bound  $\beta_2$  clamps as a function of the replication cycle, we exploited a single-molecule *in vitro* calibration method (26) that allows us to reliably convert the detected YPet- $\beta_2$  signal into an absolute number of molecules (**Supplementary Section S2**). We immobilize single purified YPet molecules on a cover glass and determine the average intensity of a single YPet fluorescent protein under these conditions (**Supplementary Figure 9**). Using this calibration, we perform control stoichiometry experiments of previously studied DNA-bound replisome components (26), specifically the  $\epsilon$ -subunit of Pol III and the  $\tau$ -subunit of the clamp-loader to verify that our *in vitro* single-molecule calibration remains reliable *in vivo* (46) (**Supplementary Figure 10**). For both the proteins, we reproduced the stoichiometry for the pair of sister replisomes as previously published (26), namely  $5.74 \pm 0.04$  molecules in total for the  $\epsilon$ -subunit ( $n = 64$ ) and  $6.12 \pm 0.03$  molecules in total for the  $\tau$ -subunit ( $n = 66$ ). Here the error is  $\pm$  s.e.m. We

also verified that a YPet- $\beta$  intensity standard provides the same mean intensity value under our experimental conditions (**Supplementary Figure 9g**). Therefore, we subsequently use this average intensity value to estimate the number of  $\beta_2$  clamps in our experiments. In calculating the number of YPet- $\beta_2$  molecules, we correct for photobleaching (**Supplementary Section S3; Supplementary Figure 8**) and verify that the fraction of immature, dark YPet proteins is negligible (**Supplementary Section S4, Supplementary Figure 8**). In our conversion from intensity to molecules, we also take into account that  $\beta_2$  clamps are dimers by dividing the measured YPet signal by two. This is a realistic assumption since it is believed that  $\beta_2$  clamps are in closed conformation even when they are not DNA-bound (22).

Using this calibration standard, we quantify the absolute number of DNA-bound YPet- $\beta_2$  molecules for individual traces of DNA replication. Representative individual time-traces of single cells clearly demonstrate that following the initiation of replication, a gradual increase of the number of DNA-bound  $\beta_2$  occurs until a steady state plateau is reached (**Figure 3C**). This plateau is maintained until decrease is observed shortly after or during termination (**Figure 3C**). From the individual traces, one can observe that there is significant cell-to-cell variability in the absolute number of DNA-bound  $\beta_2$  clamps, but that the overall trend in which the number of DNA-bound  $\beta_2$  clamps is constant for a significant fraction of the cell cycle is the same for all cells. We compared this temporal behavior to that of a different replisome component, the  $\tau$ -subunit of the clamp-loader in a strain in which both the  $\beta_2$  clamp and the  $\tau$ -subunit are fluorescently labeled. (**Supplementary Section S5**). The  $\tau$ -YPet fluorescence signal fluctuates strongly and does not yield a stable plateau, in contrast to the mCherry- $\beta_2$  fluorescent signal (**Supplementary Figure 11**).

To obtain statistically significant values for both the total number of  $\beta_2$  clamps in the cell and the number of DNA-bound  $\beta_2$  clamps, we extracted the average behavior from analysis over numerous cells ( $n = 137$ ) (**Figure 3D,E**). **Figure 3D** clearly depicts that for an average cell the fraction of DNA-bound  $\beta_2$  clamps is more than half of the total content in the cell, which decreases to roughly zero after termination. During the cell cycle, an average cell doubles its YPet- $\beta_2$  content from approximately 60 to 120 molecules. This number of  $\beta_2$  clamps in the cell is in good agreement with ensemble Western estimates we performed under the same growth conditions (**Supplementary Section S6; Supplementary Figure 12**). Remarkably, the number of DNA-bound YPet- $\beta_2$  is held at a stable value of  $N_{\beta_2} = 46$  (s.d. = 12, s.e.m. = 1) (**Figure 3E (inset)**). We also observe that the number of DNA-bound  $\beta_2$  clamps is close to zero prior to initiation and after termination. We ruled out the presence of an ectopic *dnaN* gene, by verifying that only a single copy of the *dnaN* gene is present in the strain that we used for these experiments (**Supplementary Section S7; Supplementary Figure 12**). The experiment was successfully reproduced with a different fluorescent protein fusion (mCherry- $\beta_2$ ), which strengthens the argument that accumulation is unlikely the result of fluorophore aggregation (47, 48), but rather due to physiological build-up of DNA-bound clamps (**Supplementary Section S8; Supplementary Figure 13**). The slightly lower mean number of DNA-bound clamps ( $N_{\beta_2} = 34$ , s.d. = 12, s.e.m. = 1.5) as measured using the mCherry- $\beta_2$  protein fusion, in combination with the mCherry intensity calibration, possibly results from the less ideal photophysical properties of mCherry, which make it less suitable than YPet for rigorous quantitative fluorescence microscopy.



**Figure 3: Quantification of the *in vivo*  $\beta_2$  sliding clamp stoichiometry during replication.** (A) A representative temporal montage of the YPet- $\beta_2$  fluorescence signal from prior to initiation until after cell division. A clear intensity increase is observed at the focus formation following initiation (indicated with white arrows in inset) Scale bar, 1.6  $\mu\text{m}$ . (B-C) Traces of focus formation in

individual cells. In **(B)** we plot the fraction of YPet- $\beta_2$  molecules that are DNA-bound compared to the total number in the cell. More than 50% of the total YPet- $\beta_2$  molecules are DNA-bound. The gradual decline in this fraction results from the increase of  $\beta_2$  during the cell cycle. The inset indicates how the DNA-bound YPet- $\beta_2$  molecules and the total YPet- $\beta_2$  in the cell are defined. In **(C)** we plot the absolute number of DNA-bound YPet- $\beta_2$  molecules. Here the gradual increase, steady state and gradual decrease of the DNA-bound YPet- $\beta_2$  molecules can clearly be seen. In both **(B)** and **(C)** the traces have been aligned with respect to initiation. **(D-E)** The average behavior of individual YPet- $\beta_2$  molecules measured in individual cells. **(D)** The fraction of DNA-bound YPet- $\beta_2$  molecules is on average > 50% half way through the replication cycle. **(E)** The YPet- $\beta_2$  molecules in the whole cell (blue curve) approximately doubles during the cell cycle, from 60 to 120 YPet- $\beta_2$  molecules. The DNA-bound YPet- $\beta_2$  molecules (red curve) remarkably increases to a mean steady state value of 46 YPet- $\beta_2$  molecules (s.d. = 12, s.e.m = 1) following initiation. This value is maintained throughout the replication process until a concomitant decrease is observed after or during termination. Individual traces have been normalized with respect to initiation and termination to make averaging possible. (inset) A histogram of the distribution of a number of DNA-bound YPet- $\beta_2$  molecules during steady state. (n = 137).

### 3.2.4. Single $\beta_2$ clamps are not rapidly unloaded *in vivo*.

To study the *in vivo* unloading time of an individual  $\beta_2$  clamp, we utilized single-molecule PALM (**Figure 4A**). The endogenous *dnaN* gene was replaced with a functional N-terminal PAmCherry (49) fusion (See **Materials and Methods**). Fluorescence images are acquired under shuttered 561 nm excitation (See **Materials and Methods**), while activation is performed once with low 405 nm laser illumination, such that on average less than one DNA-bound PAmCherry- $\beta$  per cell is activated. Prior to fluorescence activation, a PH image is acquired to determine the cell's position and its periphery. Sample pre- and post-activation images (**Figure 4B**), together with the corresponding line-profile intensity plots (**Figure 4C**), demonstrate successful activation of individual DNA-bound PAmCherry- $\beta$  molecules in our strain. The advantage of PALM over more conventional techniques like Fluorescence Recovery After Photobleaching (FRAP) and Fluorescence Loss In Photobleaching (FLIP) for measuring protein turnover is that it allows one to directly image a single unloading event, as shown in the sample temporal montage and the corresponding integrated intensity trace (**Figure 4D**). We image a different field of view of cells for each complete PALM measurement sequence (**Figure 4A**) to ensure that the cell physiology and  $\beta_2$ -clamp behavior are not influenced by excessive 405 nm light exposure. Using the individually analyzed traces from different cells, we are able to build up a distribution for the on-time events of single PAmCherry- $\beta$  molecules (**Figure 4D**). In order to visualize a single unloading event, we only image one out of the two PAmCherry- $\beta_2$  dimer subunits. After correcting for photo-bleaching (**Supplementary Section S9; Supplementary Figure 14**), we estimate the *in vivo* unloading time to be  $t_{\text{unload}} = 195 \pm 58 \text{ s}/\beta_2$  (**Figure 4E**). This result is in good agreement with previous *in vitro* experiments ( $127 \text{ s}/\beta_2$ ) (24).

### 3.2.5. The effective *in vivo* loading rate of $\beta_2$ clamps.

The *in vivo* loading time of a  $\beta_2$  clamp during chromosomal replication provides us with insight into how frequently a new clamp  $\beta_2$  is needed for processive genome duplication. We utilize both the long-time lapse and the single-molecule PALM data to compute the effective loading time *in vivo* of a new  $\beta_2$  clamp ( $t_{\text{load}}^{\text{eff}}$ ). The designation 'effective' is



added since we do not directly measure the loading of an individual  $\beta_2$  clamp, but rather the total loading rate of  $\beta_2$  clamps onto DNA.

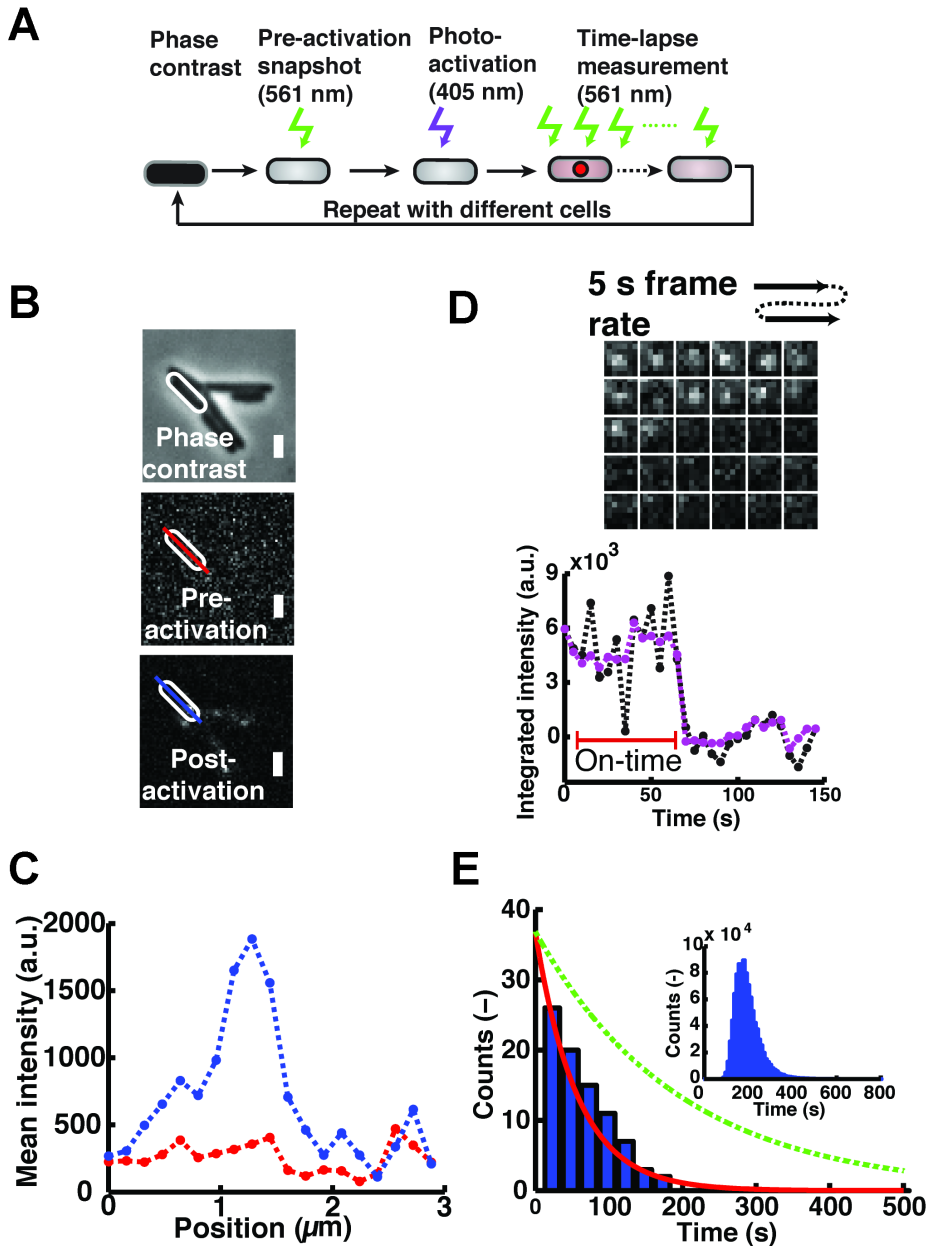


Figure 4: Direct measurement of the *in vivo* unloading time of the  $\beta_2$  sliding clamp during replication. (A) Illustration of the measurement sequence to image a single  $\beta_2$  clamp unloading event.

First, a phase contrast (PH) and pre-activation snapshot are taken, after which molecules are activated only once, and subsequently imaged until foci are no longer visible **(B-C)** Single PAmCherry- $\beta$  molecules are visualized by PALM. The sample PH image together with the respective pre-activation and post-activation fluorescence images illustrate that a single PAmCherry- $\beta$  molecule can successfully be photo-activated. The corresponding pre-activation (red) and post-activation (blue) line profile plots of the single DNA-bound PAmCherry- $\beta$  molecule. Scale bars, 1.6  $\mu\text{m}$ . **(D)** A representative example of a montage showing the fluorescence intensity of a PAmCherry- $\beta$  molecule over time and the corresponding intensity trace of the signal. The single-step disappearance is indicative of a single molecule. **(E)** On-time distribution for individual PAmCherry- $\beta$  molecules ( $n = 84$ ) fitted with an exponential (red line), and the distribution for the unloading times corrected for photobleaching (dashed green line). The inset shows the distribution of the fitted unloading time constants over the  $10^6$  bootstrapped data sets from which the confidence interval for the unloading time is determined.

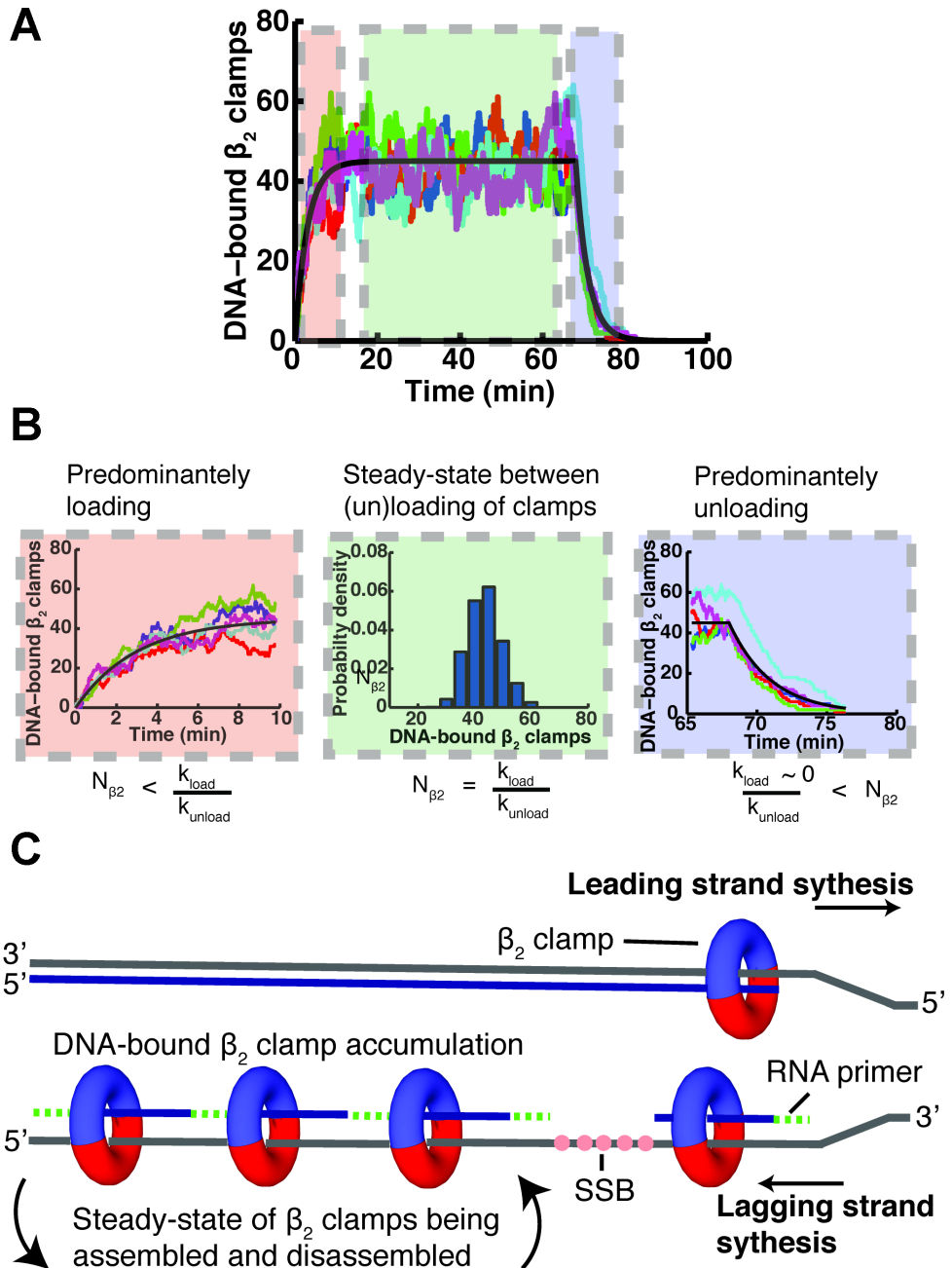
We have shown in the preceding section that the number of DNA-bound  $\beta_2$  clamps remains essentially constant ( $N_{\beta_2} = 46$ ) during  $\sim 2/3$  of the replication process. We independently determined the *in vivo* unloading time via PALM to be  $t_{\text{unload}} = 195 \text{ s}/\beta_2$  during replication. In the steady-state regime, the total unloading rate of  $\beta_2$  clamps is balanced by the effective loading rate of  $\beta_2$  clamps ( $t_{\text{load}}^{\text{eff}}$ ) onto newly formed primer for Okazaki fragment synthesis:

$$\frac{1}{t_{\text{load}}^{\text{eff}}} = N_{\beta_2} \cdot \frac{1}{t_{\text{unload}}}. \quad (1)$$

Using Equation (1), we compute the *in vivo* effective loading time for a  $\beta_2$  clamp during replication to be  $t_{\text{load}}^{\text{eff}} = 4 \pm 1 \text{ s}/\beta_2$ . The reader is referred to **Supplementary Section S10** and **Supplementary Figure 15** for a more detailed discussion of **Equation (1)**.

### 3.3. Discussion

DNA replication, orchestrated by the multi-protein replisome-complex, is a process essential to cell viability. By using *in vivo* single-molecule fluorescence microscopy in combination with microfluidics, we were able to investigate the detailed dynamics of an essential component of the replisome, the  $\beta_2$  clamp, during DNA replication in live *E. coli* cells. Lagging strand synthesis is a complex and highly dynamic process, and the sliding clamp is one of the key proteins involved. Each new primer-template junction requires a loaded  $\beta_2$  clamp to ensure processive replication by DNA Pol III, which is signaled to cycle from one Okazaki fragment to the next (50) as the replication fork progresses at approximately 600 bp/s. Given the average replication fork speed and the typical size of an Okazaki fragments (1-2 kb), one would expect a  $\beta_2$  clamp to be present on the chromosome for  $\sim 1.5 - 3 \text{ s}$ . Leading strand synthesis might be less processive than commonly believed, which would imply that a new  $\beta_2$  clamp would also need to be loaded on the leading-strand. In what follows, however, we have assumed that during normal replication the exchange of  $\beta_2$  clamps on the leading-strand is a much less frequent occurrence than  $\beta_2$  clamp exchange for the lagging-strand. It was until now not demonstrated *in vivo* whether these loaded  $\beta_2$  clamps are predominantly recycled (i.e., immediately unloaded and reloaded) between successive Okazaki fragments, or whether  $\beta_2$  clamps remain bound to a completed fragment for a prolonged period of time.



**Figure 5: Describing the  $\beta_2$  sliding clamp recycling process during replication.** (A) A Monte Carlo simulation of the  $\beta_2$  clamp assembly and disassembly reaction. For illustrative purposes we perform a Monte Carlo simulation of the proposed model, utilizing the experimentally determined

primer formation and unloading rate, as well as the replication time, and under the assumption that primer formation is rate limiting. We show the simulation results for 5 individual traces (colored lines). The black curve is the analytical solution for the average number of loaded  $\beta_2$  clamps. Here we divide the total trace into three-time regions, namely initiation (red), steady state (green) and termination (blue). **(B)** A zoom of the different sections from (A). (left) A build-up of loaded  $\beta_2$  clamps on the DNA proceeds for  $\sim 10$  min. (middle) After the gradual accumulation of loaded  $\beta_2$  clamps, a steady state plateau of 46 DNA-bound  $\beta_2$  clamps is maintained for approximately  $2/3$  of the replication process. (right) After termination, all DNA-bound  $\beta_2$  clamps are unloaded in  $\sim 5$  min. **(C)** A cartoon illustrating the DNA-bound  $\beta_2$  clamp build up during replication. As the rate at which  $\beta_2$  clamps are loaded (one every 4s) is much faster than the unloading rate of individual  $\beta_2$  clamps (once every 195 s) during replication, there will be a dynamic reservoir of  $\beta_2$  clamps that have not yet been unloaded left on the lagging strand.

Our results indicate that the number of DNA-bound  $\beta_2$  sliding clamps increases during the course of the cell cycle, peaking at more than twenty behind an individual fork. Following initiation of replication, we observe that the number of DNA-bound  $\beta_2$  clamps gradually increases until a steady state plateau is reached. This plateau, whose magnitude is such that about 50% of the total  $\beta_2$  clamps in the cell are DNA-bound, is maintained throughout the remainder of the cell cycle. We determined the number of  $\beta_2$  clamps in the cell (60 - 120) during replication, as well as the total number of DNA-bound  $\beta_2$  clamps ( $N_{\beta_2} = 46$ ). After termination,  $\beta_2$  clamps are presumably no longer being loaded, and the fraction of DNA-bound  $\beta_2$  clamps decays accordingly.

Notably, our measurements for the number of DNA-bound  $\beta_2$  clamps differ from the value measured previously by some of us in a comprehensive *in vivo* study of the whole *E. coli* replisome complex (26). In this study, the number of  $\beta_2$  clamps was estimated to be 3 for each of the two independent replisomes (44), for a total of 6 DNA-bound  $\beta_2$  clamps present during replication. While we cannot fully explain the difference in the number of DNA-bound  $\beta_2$  clamps, we can nonetheless delineate some possible contributions. The difference may result from the cell physiology due to the immobilization method, lower statistics due to the challenging nature of the 'slim-field' experiments at the time, or due to inadvertent changes of the imaging system since measurements spanned across months in the earlier study. It is thus crucial maintaining the healthy cell physiology and cell cycle synchronization during experiments, which highlights the utility of microfluidics in live cell single-molecule fluorescence measurements.

The substantial number of DNA-bound  $\beta_2$  clamps behind each replication fork suggests that  $\beta_2$  clamps are not rapidly recycled during replication. To corroborate this view, we have utilized PALM to directly measure the *in vivo* unloading rate of a single  $\beta_2$  clamp ( $t_{\text{unload}} = 195\text{s}/\beta_2$ ). Together with the number of DNA-bound  $\beta_2$  clamps in steady state this allows us to calculate the effective time of loading (**Equation (1)**) a  $\beta_2$  clamp during replication as  $t_{\text{load}}^{\text{eff}} = 4\text{s}/\beta_2$ . This result is in good agreement with our previous calculated average estimate of the primer formation time using the Okazaki fragment size range and the typical size of the *E. coli* genome. Also, this effective loading rate is in accordance with the model that DnaG sets the fork speed (28). DnaG is thought to synthesizes RNA primers at a rate of approximately one primer every one to two seconds (51), which is in good agreement with our calculated *in vivo* effective loading time. We suggest that an individual  $\beta_2$  clamp remains on the DNA for a protracted period of time during chromosomal replication, as has been proposed on this basis of *in vitro* experiments (24, 52) and plasmid

replication (53). Our results are in agreement with the behavior of the sliding clamp for the Gram-positive bacterium *B. subtilis* (33). In this bacterium, the number of DNA-bound  $\beta_2$  clamps was estimated at  $\sim 200$  during replication, indicative of clamps being left behind during fork progression. There is a slight possibility that the loading and unloading reaction could be sterically hampered by the fusion protein. However, we have no reason to believe that this is the case since our results are in very good agreement with previous *in vitro* (24) and *in vivo* work (33). Our study shows that rapid recycling of  $\beta_2$  clamps for subsequent lagging-strand synthesis (32), though observed in *in vitro* experiments in the absence of excess  $\beta_2$  clamps in solution, is not the predominant mode *in vivo*. While our data does not exclude that  $\beta_2$  clamps are rapidly recycled at the replication fork, the fact that the loading rate from solution matches the estimated primer formation rate strongly suggests that direct recycling is not the dominant mode of clamp loading.

To illustrate the overall  $\beta_2$  clamp dynamics during replication, we perform a Monte Carlo simulation (See **Materials and Methods**) that takes our experimentally determined values for  $N_{\beta_2}$ ,  $t_{\text{unload}}$ ,  $t_{\text{load}}^{\text{eff}}$  and  $t_{\text{rep}}$  as input parameters (**Figure 5A**). Since the approximate rate of clamp removal during termination (**Figure 3E**) agrees with the value measured by PALM during steady-state replication ( $\sim 195 \text{ s}/\beta_2$ ), we simply input the latter (likely more accurate) value into the simulations. The simulation starts at  $t = 0$  with no DNA-bound  $\beta_2$  clamps, after which it takes  $< 10$  min to reach a stable steady state number of  $\beta_2$  bound to DNA (**Figure 5B, left**). This value is maintained for  $\sim 60$  min (**Figure 5B, middle**), after which termination occurs and clamps are unloaded in  $< 5$  min (**Figure 5B, right**). The number of  $\beta_2$  clamps in steady state as well as the rise and fall times underline our measured results, and are depicted schematically in **Figure 5C**.

The steady state build-up of DNA-bound  $\beta_2$  sliding clamps forms a  $\beta_2$ -landing pad (33) for different proteins to dock themselves to DNA during the life cycle of the cell. Numerous different proteins utilize the  $\beta_2$  clamp via the same binding pocket (54) to perform their respective biological function. Such other  $\beta_2$  clamp-binding proteins range from DNA ligase for Okazaki fragment maturation, inactivation of DnaA through the  $\beta_2$ -Hda<sub>2</sub> interaction(11-13, 55, 56), potential screening of DNA damage due to sliding capability of the  $\beta_2$  clamps (54) the tethering of the necessary polymerases for repair (4, 5, 7, 8), overcoming replication barriers (57), as well as coupling mismatch detection and replication by positioning MutS at newly replicated DNA (58). It is still unclear which of the above mentioned(or other) proteins are the main users of the DNA-bound clamps that are not directly situated at the replication fork. Since Okazaki fragment maturation seems to be relatively fast as assessed via Ligase and Pol I dynamics (59), it is likely not these proteins that predominantly occupy the DNA-bound clamps. The extent to which DNA-bound  $\beta_2$  clamps are utilized while being docked to the Okazaki fragment will most likely be dependent on the physiological state of the cell at a particular time in its cell cycle. In the case of stress conditions, for example, MutS and the different repair polymerases might predominantly make use of DNA-bound  $\beta_2$  clamps, while under minimal stress conditions ligase, CrfC and Hda<sub>2</sub> are the likely candidates. A thorough *in vivo* investigation of the stoichiometry and dynamic of different  $\beta_2$  associated proteins over the course of the cell cycle would provide the quantitative underpinning required to provide further insight into these biological processes.

### 3.4. Materials and Methods

#### 3.4.1. Strains and strain construction.

All strains are derivatives of *E. coli* K12 AB157. Strains were constructed either by P1-transduction (60) or by  $\lambda$ -red recombination (61).

The *YPet-dnaN:ter-mCerulean* was constructed using P1 transduction by transducing the *YPet-dnaN* fusion (26) together with the adjacent *kanR* gene into a strain that contains a *tetO* array (50 kb clockwise from the *dif* site) as well as the chromosomal integrated chimeric gene *tetR-mCerulean* (44). The presence of the *YPet-dnaN* gene fusion was verified using the oligonucleotides: 5' – CGT TGG CAC CTA CCA GAA AG – 3' and 5' – ATG CCT GCC GTA AGA TCG AG – 3'. The sequence of the *YPet-dnaN* fusion was confirmed by DNA sequencing.

A chromosomal fusion of the gene encoding for the photoactivatable fluorescent protein (PAmCherry1) (49) to the N-terminus of the *dnaN* gene was created using  $\lambda$ -red recombination (61). The gene encoding for PAmCherry1 was amplified by PCR. The forward primer used contains a *XmaI* restriction site (5' – GCG GGC CCC GGG ATG GTG AGC AAG GGC GAG GAG – 3'). The reverse primer used contains a sequence coding for an 11 amino acid linker and a *SacI* site (5' – CGA TCG GAG CTC CGC GCT GCC AGA ACC AGC GGC GGA GCC TGC CGA CTT GTA CAG CTC GTC CAT GCC – 3'). The PCR product was cloned into the backbone of pROD44 (26) containing a kanamycin resistance cassette, flanked by *frt* sites, resulting in the template plasmid PAmCherry1.

This plasmid was then used as a template plasmid for generating the insert sequence used during  $\lambda$ -red recombination to create the *PAmCherry-dnaN* strain. The primer sequences used were: Forward 5' – ACG ATA TCA AAG AAG ATT TTT CAA ATT TAA TCA GAA CAT TGT CAT CGT AAC TGT AGG CTG GAG CTG CTT C – 3'; Reverse 5' – ACC TGT TGT AGC GGT TTT AAT AAA TGC TCA CGT TCT ACG GTA AAT TTC ATC GCG CTG CCA GAA CCA GCG G – 3'. The DNA fragment was gel purified and ~700 ng of the linear DNA was used for electroporation of AB1157 cells overexpressing  $\lambda$ -Red proteins from pKD46 (61). The correct insertion of the fragment into the chromosome of the resulting strain was assayed by PCR. The oligonucleotides used were 5' – CGT TGG CAC CTA CCA GAA AG – 3' and 5' – ATG CCT GCC GTA AGA TCG AG – 3'. The sequence of the fusion gene in this strain was confirmed by DNA sequencing.

Construction of the *mCherry-dnaN* strain. The *mCherry* gene was amplified by PCR. The forward primer used for this contains a *XmaI* restriction site (5' – TAG GCT CCC GGG ATG AGC AAG GGC GAG GAG GAT AAC – 3'). The reverse primer used contains a *SacI* site and sequence coding for an 11 amino acid linker (5' – AAG GAG CTC GCG CTG CCA GAA CCA GCG GCG GAG CCT GCC GAC TTG TAC AGC TCG TCC ATG CC – 3'). The *frt*-flanked kanamycin resistance gene was amplified using the following primers: Forward 5' – TTA CCC GGG CAT ATG AAT ATC CTC CTT AG-3'; Reverse 5'-TTA GGA TCC TGT AGG CTG GAG CTG CTT CG – 3'. The resulting fragment was digested with *XmaI* and *BamHI*. The *mCherry* fragment and the kanamycin fragment were cloned into pUC18 between *SacI* and *BamHI* sites.

The  $\lambda$ -red recombination was performed as mentioned in the previous section using the primers: Forward 5' – TAT CAA AGA AGA TTT TTC AAA TTT AAT CAG AAC ATT GTC ATC GTA AAC CTG TAG GCT GGA GCT GCT TCG – 3'; Reverse 5' – ACC TGT TGT AGC GGT TTT AAT AAA TGC TCA CGT TCT ACG GTA AAT TTC ATC GCG CTG CCA GAA CCA GC – 3'. The presence of the gene fusion was verified using oligonucleotides 5' – CGT TGG CAC CTA CCA GAA AG – 3' and 5' – ATG CCT GCC GTA AGA TCG AG – 3'. The sequence of the fusion gene in this strain was confirmed by DNA sequencing.

The *dnaX*( $\tau$ )-*YPet:mCherry-dnaN* strain was constructed using P1 transduction by transducing the *dnaX-YPet* fusion (26) together with the adjacent *kanR* gene into a strain that contains the *mCherry-dnaN* gene fusion. The presence of the *dnaX-YPet* gene fusion after transduction was verified using the oligonucleotides: 5' – GAG CCT GCC AAT GAG TTA TC – 3' and 5' – GGC TTG CTT CAT CAG GTT AC – 3' and similarly the *mCherry-dnaN* fusion using, 5' – CGT TGG CAC CTA CCA GAA AG – 3' and 5' – ATG CCT GCC GTA AGA TCG AG – 3'. The sequences of the fusions in this strain were confirmed by DNA sequencing.

**Supplementary Tables 2 and 3** provide an overview of the plasmids used, as well as a summary of the different strains. The cell morphology and the doubling times of the fusion strains in LB and M9-glycerol growth medium were compared to AB1157 wild type (WT). No significant differences were observed (**Supplementary Table 1** and **Supplementary Figure 6a**). The doubling times of the cells in the microfluidic device were similar (slightly faster) compared to cells grown in a shake flask (**Supplementary Figure 7**). We also confirmed that in the absence of IPTG (the experimental condition used during long-time lapse microscopy) no DNA-bound foci were detected for the *YPet-dnaN:ter-mCerulean* strain. (**Supplementary Figure 6b**).

### 3.4.2. M9 growth medium used in experiments.

The M9 growth medium used in experiments is as follows. 1 L of M9 growth medium contains 10.5 g/L of autoclaved M9 broth (Sigma-Aldrich); 0.1 mM of autoclaved CaCl<sub>2</sub> (Sigma-Aldrich); 0.1 mM of autoclaved MgSO<sub>4</sub> (J.T.Baker); 0.3% of filter-sterilized glycerol (Sigma-Aldrich) as carbon source; 0.1 g/L of filter-sterilized 5 amino acids, namely L-threonine, L-leucine, L-proline, L-histidine and L-arginine (all from Sigma-Aldrich) and 10  $\mu$ L of 0.5% filter-sterilized Thiamine (Sigma-Aldrich).

### 3.4.3. Microfluidics for extended time-lapse microscopy.

We use our own design (41) of the previously published microfluidic device known as the *mother machine* (42) for cell immobilization during long time-lapse experiments. Here we only briefly outline the main steps involved. First, we use electron-beam lithography in combination with dry etching techniques to create the structure in silicon. Next, we make a negative mold of this structure in polydimethylsiloxane (PDMS). The PDMS mold is then used to fabricate the positive structure in PDMS which is subsequently used for experiments.

### 3.4.4. Preparation of cells for microscopy.

Cells were streaked on Luria-Bertani (LB)-plates containing the appropriate antibiotic. Single colonies from these plates were inoculated overnight at 37 °C with shaking in M9 medium supplemented with 0.3% glycerol (Gly), essential nutrients together with the appropriate antibiotics. The subsequent day the overnight culture was subcultured into the same medium and grown at 37 °C with shaking until an  $OD_{600} \sim 0.2$  was reached. Cells were concentrated by centrifugation for 2 min at 16100 g. The subsequent steps are dependent on the type of microscopy experiment performed as outlined next.

For agarose pad experiments, the supernatant was decanted and the pellet was resuspended in 100  $\mu$ L M9-Gly supplemented with essential nutrients. The resuspended cells were subsequently vortexed for 2 s and immobilized on an M9-Gly 1.5% agarose pad between two coverslips. (The coverslips were ultrasonically cleaned in Acetone and Isopropyl alcohol and burned by a flame to minimize the fluorescent background prior to use).

For microfluidic device experiments, the supernatant was decanted and the pellet was resuspended in 50  $\mu$ L M9-Gly with essential nutrients and injected into the microfluidic device. After injection into the device, the device was centrifuged for 10 min at 2500 g (Eppendorf 5810R) as to load the cells into the growth channels. Following centrifugation the device was mounted on the microscope with tubing attached and incubated for  $\sim 45$  min at 37 °C. After incubation, fresh M9-Gly with essential nutrients and the appropriate antibiotics are flushed through the device. The syringe containing the medium is then attached to an automated syringe pump to continuously infuse fresh M9-Gly, essential nutrients and 0.2 mg/mL bovine serum albumin (BSA) through the device at a rate of 0.5 mL/hr.

### 3.4.5. Microscope setup.

All images were acquired on a commercial Nikon Ti microscope equipped with a Nikon CFI Apo TIRF 100x, 1.49 NA oil immersion objective and an Andor iXon 897 Electron Multiplying Charge Coupled Device (EMCCD) camera operated by a personal computer (PC) running Nikon NIS elements software. Cell outlines were imaged using the standard Nikon brightfield halogen lamp and condenser components. The fluorescence excitation was performed using custom-built laser illumination. A Cobolt Fandango 515 nm continuous wave (CW) diode-pumped solid-state (DPSS) laser was used to excite YPet, Cobolt Jive 561 nm CW DPSS laser was used to excite mCherry and PAmCherry respectively. PAmCherry was activated by a Votran Stradus 405 nm. All three laser beams were combined using dichroic mirrors (Chroma ZT405sp-xxr, 575dcspxr) and subsequently coupled into a single-mode optical fiber (KineFLEX). The output of the fiber was expanded and focused onto the back focal plane of the objective mounted on the microscope. Notch filters (Semrock NF03-405E, NF03-514E, NF03-561E) were used to eliminate any laser light leaking onto the camera. The emission of the different fluorescent proteins was projected onto the central part of the EMCCD camera using custom filter sets: Chroma z561, ET605/52m, zt561rdc (mCherry), Chroma z514, ET540/30m, zt514rd (YPet), Chroma zet405, ET480/40m, zt405rdc (CFP). A custom design commercial temperature control housing (Okolabs) enclosing the microscope body maintained the temperature at 37



°C. Sample position was controlled with a Nikon stage (TI-S-ER Motorized Stage Encoded, MEC56100) together with the Nikon Perfect Focus System (PFS) to eliminate Z-drift during image acquisition.

### 3.4.6. Cell lysate preparation for intensity calibration.

The cell lysate used for single-molecule intensity calibration was prepared as follows. Cells were grown overnight at 37 °C with shaking in M9 medium supplemented with 0.3% glycerol (Gly), essential nutrients together with the appropriate antibiotics. The subsequent day the overnight culture was subcultured into the same medium and grown at 37 °C with shaking until an  $OD_{600} \sim 0.5$  was reached. The cells were harvested by centrifugation at 6000x g (Beckman JLA 9.1000 rotor) for 15 min. Cells were subsequently resuspended in 5 mL M9-Gly and essential nutrients. The cell suspension was French pressed (Constant Systems) twice at 20000 psi. The cell lysate was then spun down at 30000x g (Beckman JA-17 rotor) for 35 min. The supernatant was shock frozen using liquid nitrogen and kept at -80 °C until needed.

### 3.4.7. Data acquisition.

All data acquisition was performed on the same microscope setup. Image acquisition was performed with Nikon NIS-elements software. The acquisition protocol was dependent on the type of experiment performed as outlined next.

Long time-lapse experiments were conducted as follows. The cell outlines were imaged using standard brightfield illumination. Subsequently, the sample was excited by laser excitation (515 nm) with an intensity of approximately  $5 \text{ W}\cdot\text{cm}^{-2}$  as calculated according to (62). The exposure time was set to 80 ms. The camera gain was set to 100. Brightfield and fluorescence images were acquired every 2.5 min. Data spanning ~10 hrs of measurement were acquired overnight.

We conducted two types of PALM experiments. First, we determined the bleaching characteristic of PAMCherry under our experimental conditions, and secondly, we measured the unloading time of a single  $\beta_2$  clamp. PALM images were acquired as follows. First, the cell outlines were imaged by taking a single phase-contrast (PH) image using a commercial Nikon external phase ring configuration. The sample was then excited for a single frame (400 ms exposure time) by a 561 nm laser with an intensity of approximately  $5 \text{ W}\cdot\text{cm}^{-2}$ , calculated according to Grünwald *et al.* (62). This image was used to determine the auto-fluorescence level due to the sample prior to activation. Photo-activation of PAMCherry was done with a single pulse (5 s) of 405 nm with an intensity of approximately  $2.5 \text{ W}\cdot\text{cm}^{-2}$ , calculated according to Grünwald *et al.* (62). Subsequently, a post-activation time-lapse of images were acquired using the 561 nm laser at the same intensity at a frame rate of either ~700 ms (bleaching experiments) or 5s (unloading experiments) with an exposure time of 400 ms per frame. Camera gain was set to 100.

### 3.4.8. Image analysis of long time-lapse experiments.

Images were analyzed with custom-written MATLAB software (MathWorks). Prior to any analysis, we subtract the uneven background using a rolling ball filter (63) and subsequently corrected for illumination heterogeneity by using the previously measured

laser beam profile (64). We also align the brightfield and fluorescence signals with respect to each other with 1-pixel accuracy. X-Y drift is corrected in both the fluorescence and brightfield images by tracking a fiducial marker in the PDMS to within 1 pixel.

Each drift-corrected region of interest, consisting out of a single growth channel, is analyzed individually. The brightfield images are used to determine the cell poles of all the cells in a given frame. For the fluorescence signal, a kymograph of the fluorescence signal is constructed by summation of the pixel intensities per image perpendicular to the channel direction for each frame. This results in summed intensity information as a function of time per growth channel (**Figure 2C**). We make use of the generated kymographs to determine individual replication and division cycles per cell (**Figure 2D**). A post-processing step is subsequently performed to eliminate cells that did not match the following selection criteria: correct cell length, sufficient growth characteristics, observation of a complete cell cycle, clear fluorescence signal that both starts and ends in a diffuse state (**Figure 2D**).

The fluorescence images of the detected individual cells that pass the above selection criteria are analyzed further. We base our fluorescence analysis on an image of an individual bacterium with its long axis aligned with the horizontal direction of the image. The width of the image is equal to the length of the bacterium. We fix the height of the image such that a sufficient area above and below the bacterium is included that is indicative of the auto-fluorescence of the sample. We analyze the fluorescent intensity counts of a single bacterium using the individual fluorescence kymographs of each cell (summed line-profiles) by calculating three types of image content for a specific bacterium, namely 'background', 'foci' and 'cytoplasm' (**Supplementary Figure 3a**). In brief, we first estimate the background fluorescence from the sample using the signal outside of the bacterium. We did not have to take into account auto-fluorescence from the bacterium itself since we conducted our experiments using a minimal medium which results in negligible levels of cellular auto-fluorescence (**Supplementary Figure 3b**). The intensity outside the bacterium is used for a threshold with the remaining pixels intensities being representative of the total bacterium fluorescent counts. We subsequently separate 'cytoplasm' and 'foci' signals by determining the median of the summed line-profiles. The signal significantly above this value is attributed to foci, while the remainder (lower values) are treated as the fluorescence signal from the cytoplasm (**Supplementary Figure 3a**). This results in an integrated intensity value for the foci and also for the cytoplasm.

### 3.4.9. Image analysis of PALM experiments.

PALM data was analyzed using custom-written MATLAB software (MathWorks) in combination with the freely available MicrobeTracker software (65). Prior to any spot analysis, the fluorescence images are subjected to illumination correction and to alignment with respect to the phase-contrast (PH) images. The resulting corrected and aligned fluorescence images are then used during further analysis.

Using the PH image, the different cells are detected in the field of view and their respective outlines are determined using MicrobeTracker. Subsequently, using the spot detection algorithm as described in Olivo-Marin *et al.* (66), the spots in each individual image of the fluorescence time-lapse series are detected, and the integrated intensity determined by summing the pixel values of each spot (67). The integrated intensities of the

spots as followed as a function of time. This results in individual time-lapse integrated intensity traces of single molecules (**Figure 3C**). The cell outlines as determined previously are overlaid with the fluorescence images. Any foci that are not situated in a bacterial cell (false positives) are rejected from further analysis. Only cells that had a clear fluorescence intensity focus were analyzed. This focus is indicative of DNA-bound clamps and thus DNA replication. Foci that exhibit multiple steps in fluorescence intensity are also rejected. For the remainder of the foci, the time it takes from the start of the data acquisition until spot disappearance is recorded (**Figure 3C**). These calculated time differences are indicative of molecule unloading (or bleaching, depending on the time of acquisition) and analyzed further as described in the following section.

### 3.4.10. Monte Carlo simulation of $\beta_2$ loading and unloading dynamics.

For illustrative purposes, we perform Monte-Carlo simulations (**Figure 5**) starting with no clamps loaded and no primers formed ( $N_{\beta_2} = 0$ ), and assuming the loading rate to be much faster than the rate of primer formation ( $N_p \approx 0$ ). In each small time-step  $dt$  we let  $N_{\beta_2} \rightarrow N_{\beta_2}+1$  with probability  $dt t_p^{-1}$  and  $N_{\beta_2} \rightarrow N_{\beta_2}-1$  with probability  $N_{\beta_2} dt t_{unload}^{-1}$ . This is repeated until the replication time is reached, upon which the primer formation rate  $t_p^{-1}$  is set to zero.

## 3.5. Contributions

M.C.M. and N.H.D. designed the research and the experiments. M.C.M. and S.T.K. undertook the experiments. S.T.K. constructed the strains. P.T. performed the western blot. S.T.K. performed the Southern blot. M.C.M. and J.W.J.K. wrote the software to analyze all the microscopy data. M.C.M., A.v.d.B., and M.D. analyzed the data. A.v.d.B. performed the Monte Carlo simulation. R.R.L. and D.J.S. provided strains and contributed to the planning and discussion of the work.

## 3.6. Acknowledgments

We thank Ilja Westerlaken for her assistance with the French Press for cell lysate preparation, Hugo Snippert for initial help with the Southern blotting, and Bronwen Cross for initial help with the Western blot, as well as Jasper van Veen, Kasper de Leeuw, and Liliane Jimenez van Hoorn for initial help with constructing the *PAmCherry-DnaN* strain. We thank Anjana Badrinarayanan for providing the *mCherry-DnaN* strain. We thank Jan Lipfert, Anne Meyer, Tim Blosser, Yaron Caspi, Antoine van Oijen, and Theo van Laar for stimulating discussions. We thank Francesco Perdaci, David Dulin, David Gruenwald, Chirlmin Joo, and Stanley Dinesh Chandradoss for initial help with microscopy. We thank Bernd Rieger for initial help with image analysis. We acknowledge contributions from Jelle van der Does, Dimitri de Roos, and Jaap Beekman for instrumentation and infrastructure. This work was supported by the European Community's Seventh Framework Programme FP7/2007-2013 under grant agreement n° 241548 (MitoSys) and a Vici grant from Netherlands Organization for Scientific Research (NWO), both to NHD, a Wellcome Trust Program Grant WT083469 to D.J.S., a TU Delft startup grant to M.D., and the Natural Sciences & Engineering Research Council of Canada (#435521-13) for R.R.L.

### 3.7. References

1. C. Indiani, P. McInerney, R. Georgescu, M. F. Goodman, M. O'Donnell, A sliding-clamp toolbelt binds high- and low-fidelity DNA polymerases simultaneously. *Molecular Cell* **19**, 805-815 (2005).
2. J. B. Vivona, Z. Kelman, The diverse spectrum of sliding clamp interacting proteins. *FEBS letters* **546**, 167-172 (2003).
3. F. J. López de Saro, Regulation of Interactions with Sliding Clamps During DNA Replication and Repair. *Current Genomics* **10**, 206-215 (2009).
4. R. Napolitano, R. Janel-Bintz, J. Wagner, R. P. Fuchs, All three SOS-inducible DNA polymerases (Pol II, Pol IV and Pol V) are involved in induced mutagenesis. *The EMBO Journal* **19**, 6259-6265 (2000).
5. J. Wagner, S. Fujii, P. Gruz, T. Nohmi, R. P. Fuchs, The beta clamp targets DNA polymerase IV to DNA and strongly increases its processivity. *EMBO reports* **1**, 484-488 (2000).
6. F. J. Lopez de Saro, M. O'Donnell, Interaction of the  $\beta_2$  sliding clamp with MutS, ligase, and DNA polymerase I. *Proceedings of the National Academy of Sciences* **98**, 8376-8380 (2001).
7. S. Kobayashi, M. R. Valentine, P. Pham, M. O'Donnell, M. F. Goodman, Fidelity of Escherichia coli DNA Polymerase IV: PREFERENTIAL GENERATION OF SMALL DELETION MUTATIONS BY dNTP-STABILIZED MISALIGNMENT. *Journal of Biological Chemistry* **277**, 34198-34207 (2002).
8. M. Tang *et al.*, Roles of E. coli DNA polymerases IV and V in lesion-targeted and untargeted SOS mutagenesis. *Nature* **404**, 1014-1018 (2000).
9. S. Jergic *et al.*, A direct proofreader-clamp interaction stabilizes the Pol III replicase in the polymerization mode. *The EMBO Journal* **32**, 1322-1333 (2013).
10. B. P. Dalrymple, K. Kongsuwan, G. Wijffels, N. E. Dixon, P. A. Jennings, A universal protein-protein interaction motif in the eubacterial DNA replication and repair systems. *Proceedings Of The National Academy Of Sciences Of The United States Of America* **98**, 11627-11632 (2001).
11. T. Katayama, T. Kubota, K. Kurokawa, E. Crooke, K. Sekimizu, The initiator function of DnaA protein is negatively regulated by the sliding clamp of the E. coli chromosomal replicase. *Cell* **94**, 61-71 (1998).
12. J. Kato, T. Katayama, Hda, a novel DnaA-related protein, regulates the replication cycle in Escherichia coli. *The EMBO Journal* **20**, 4253-4262 (2001).

13. M. Kurz, B. Dalrymple, G. Wijffels, K. Kongsuwan, Interaction of the sliding clamp beta-subunit and Hda, a DnaA-related protein. *Journal of Bacteriology* **186**, 3508-3515 (2004).
14. S. Ozaki *et al.*, A Replicase Clamp-Binding Dynammin-like Protein Promotes Colocalization of Nascent DNA Strands and Equipartitioning of Chromosomes in *E. coli*. *Cell Reports* **4**, 985-995.
15. L. A. Simmons, B. W. Davies, A. D. Grossman, G. C. Walker, Beta clamp directs localization of mismatch repair in *Bacillus subtilis*. *Molecular Cell* **29**, 291-301 (2008).
16. X.-P. Kong, R. Onrust, M. O'Donnell, J. Kuriyan, Three-dimensional structure of the  $\beta_2$  subunit of *E. coli* DNA polymerase III holoenzyme: A sliding DNA clamp. *Cell* **69**, 425-437 (1992).
17. A. Johnson, M. O'Donnell, CELLULAR DNA REPLICASES: Components and Dynamics at the Replication Fork. *Annual Review of Biochemistry* **74**, 283-315 (2005).
18. C. S. McHenry, DNA replicases from a bacterial perspective. *Annual review of biochemistry* **80**, 403-436 (2011).
19. J. Turner, M. M. Hingorani, Z. Kelman, M. O'Donnell, The internal workings of a DNA polymerase clamp-loading machine. *The EMBO Journal* **18**, 771-783 (1999).
20. M. Hedglin, R. Kumar, S. J. Benkovic, Replication Clamps and Clamp Loaders. *Cold Spring Harbor Perspectives in Biology* **5**, (2013).
21. K. R. Simonetta *et al.*, The mechanism of ATP-dependent primer-template recognition by a clamp loader complex. *Cell* **137**, 659-671 (2009).
22. J. N. Hayner, L. B. Bloom, The  $\beta$  Sliding Clamp Closes around DNA prior to Release by the *Escherichia coli* Clamp Loader  $\gamma$  Complex. *The Journal of biological chemistry* **288**, 1162-1170 (2013).
23. C. D. Downey, C. S. McHenry, Chaperoning of a Replicative Polymerase onto a Newly-Assembled DNA-Bound Sliding Clamp by the Clamp Loader. *Molecular cell* **37**, 481-491 (2010).
24. F. P. Leu, M. M. Hingorani, J. Turner, M. O'Donnell, The  $\delta$  Subunit of DNA Polymerase III Holoenzyme Serves as a Sliding Clamp Unloader in *Escherichia coli*. *Journal of Biological Chemistry* **275**, 34609-34618 (2000).
25. P. McNerney, A. Johnson, F. Katz, M. O'Donnell, Characterization of a Triple DNA Polymerase Replisome. *Molecular Cell* **27**, 527-538.

26. R. Reyes-Lamothe, D. J. Sherratt, M. C. Leake, Stoichiometry and architecture of active DNA replication machinery in *Escherichia coli*. *Science (New York, N.Y.)* **328**, 498-501 (2010).
27. T. Ogawa, T. Okazaki, Discontinuous DNA replication. *Annual review of biochemistry* **49**, 421-457 (1980).
28. K. Tougu, K. J. Mariani, The Interaction between Helicase and Primase Sets the Replication Fork Clock. *Journal of Biological Chemistry* **271**, 21398-21405 (1996).
29. P. M. Burgers, A. Kornberg, Y. Sakakibara, The *dnaN* gene codes for the beta subunit of DNA polymerase III holoenzyme of *Escherichia coli*. *Proceedings Of The National Academy Of Sciences Of The United States Of America* **78**, 5391-5395 (1981).
30. N. Yao *et al.*, Clamp loading, unloading and intrinsic stability of the PCNA,  $\beta$  and gp45 sliding clamps of human, *E. coli* and T4 replicases. *Genes Cells* **1**, 101-113 (1996).
31. A. Yuzhakov, J. Turner, M. O'Donnell, Replisome Assembly Reveals the Basis for Asymmetric Function in Leading and Lagging Strand Replication *Cell* **86**, 877 - 886 (1996).
32. N. A. Tanner *et al.*, *E. coli* DNA replication in the absence of free  $\beta$  clamps. *The EMBO Journal* **30**, 1830-1840 (2011).
33. M. Su'etsugu, J. Errington, The Replicase Sliding Clamp Dynamically Accumulates behind Progressing Replication Forks in *Bacillus subtilis* Cells. *Molecular Cell* **41**, 720-732 (2011).
34. S. T. Hess, T. P. K. Girirajan, M. D. Mason, Ultra-high resolution imaging by fluorescence photoactivation localization microscopy. *Biophysj* **91**, 4258-4272 (2006).
35. E. Betzig *et al.*, Imaging intracellular fluorescent proteins at nanometer resolution. *Science* **313**, 1642-1645 (2006).
36. D. Dulin, J. Lipfert, M. C. Moolman, N. H. Dekker, Studying genomic processes at the single-molecule level: introducing the tools and applications. *Nature Reviews Genetics* **14**, 19-22 (2013).
37. T. Xia, N. Li, X. Fang, Single-Molecule Fluorescence Imaging in Living Cells. *Annual review of physical chemistry* **64**, 459-480 (2013).
38. F. Persson, I. Barkefors, J. Elf, Single molecule methods with applications in living cells. *Current opinion in biotechnology* **24**, 737-744 (2013).

39. A. Gahlmann, W. E. Moerner, Exploring bacterial cell biology with single-molecule tracking and super-resolution imaging. *Nature Reviews Microbiology* **12**, 9-22 (2014).
40. X. Wang, R. Reyes-Lamothe, D. J. Sherratt, Visualizing genetic loci and molecular machines in living bacteria. *Biochemical Society Transactions* **36**, 749-753 (2008).
41. M. C. Moolman, Z. Huang, S. T. Krishnan, J. W. J. Kerssemakers, N. H. Dekker, Electron beam fabrication of a microfluidic device for studying submicron-scale bacteria. *Journal of Nanobiotechnology* **11**, 12-12 (2013).
42. P. Wang *et al.*, Robust growth of Escherichia coli. *Current biology : CB* **20**, 1099-1103 (2010).
43. A. W. Nguyen, P. S. Daugherty, Evolutionary optimization of fluorescent proteins for intracellular FRET. *Nature biotechnology* **23**, 355-360 (2005).
44. R. Reyes-Lamothe, C. Possoz, O. Danilova, D. J. Sherratt, Independent Positioning and Action of Escherichia coli Replisomes in Live Cells. *Cell* **133**, 90-102 (2008).
45. O. Michelsen, M. J. Teixeira de Mattos, P. R. Jensen, F. G. Hansen, Precise determinations of C and D periods by flow cytometry in Escherichia coli K-12 and B/r. *Microbiology (Reading, England)* **149**, 1001-1010 (2003).
46. V. C. Coffman, J. Wu, Counting protein molecules using quantitative fluorescence microscopy. *Trends in biochemical sciences* **37**, 499-506 (2012).
47. M. T. Swulius, G. J. Jensen, The helical MreB cytoskeleton in Escherichia coli MC1000/pLE7 is an artifact of the N-terminal yellow fluorescent protein tag. *Journal of Bacteriology* **194**, 6382-6386 (2012).
48. D. Landgraf, B. Okumus, P. Chien, T. A. Baker, J. Paulsson, Segregation of molecules at cell division reveals native protein localization. *Nature Methods* **9**, 480-482 (2012).
49. F. V. Subach *et al.*, Photoactivatable mCherry for high-resolution two-color fluorescence microscopy. *Nature Methods* **6**, 153-159 (2009).
50. Q. Yuan, C. S. McHenry, Cycling of the E. coli lagging strand polymerase is triggered exclusively by the availability of a new primer at the replication fork. *Nucleic Acids Research* **42**, 1747-1756 (2014).
51. K. J. Marians, Prokaryotic DNA replication. *Annual review of biochemistry* **61**, 673-719 (1992).

52. P. T. Stukenberg, J. Turner, M. O'Donnell, An explanation for lagging strand replication: Polymerase hopping among DNA sliding clamps. *Cell* **78**, 877-887 (1994).
53. R. Reyes-Lamothe *et al.*, High-copy bacterial plasmids diffuse in the nucleoid-free space, replicate stochastically and are randomly partitioned at cell division. *Nucleic Acids Research* **42**, 1042-1051 (2014).
54. G. Wijffels, B. Dalrymple, K. Kongsuwan, N. E. Dixon, Conservation of eubacterial replicases. *IUBMB life* **57**, (2005).
55. M. Su'etsugu, T.-R. Shimuta, T. Ishida, H. Kawakami, T. Katayama, Protein associations in DnaA-ATP hydrolysis mediated by the Hda-replicase clamp complex. *The Journal of biological chemistry* **280**, 6528-6536 (2005).
56. M. L. Mott, J. M. Berger, DNA replication initiation: mechanisms and regulation in bacteria. *Nature Reviews Microbiology* **5**, 343-354 (2007).
57. R. E. Georgescu, Yao, N. Y, M. O'Donnell, Single-molecule analysis of the Escherichia coli replisome and use of clamps to bypass replication barriers. *FEBS letters* **584**, 2596-2605 (2010).
58. J. S. Lenhart, A. Sharma, M. M. Hingorani, L. A. Simmons, DnaN clamp zones provide a platform for spatiotemporal coupling of mismatch detection to DNA replication. *Molecular Microbiology* **87**, 553-568 (2013).
59. S. Uphoff, R. Reyes-Lamothe, F. Garza de Leon, D. J. Sherratt, A. N. Kapanidis, Single-molecule DNA repair in live bacteria. *Proceedings Of The National Academy Of Sciences Of The United States Of America* **110**, 8063-8068 (2013).
60. L. C. Thomason, N. Costantino, D. L. Court, E. coli genome manipulation by P1 transduction. *Current protocols in molecular biology* **Chapter 1**, Unit 1.17 (2007).
61. K. A. Datsenko, B. L. Wanner, One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. *Proceedings Of The National Academy Of Sciences Of The United States Of America* **97**, 6640-6645 (2000).
62. D. Grünwald, S. M. Shenoy, S. Burke, R. H. Singer, Calibrating excitation light fluxes for quantitative light microscopy in cell biology. *Nature protocols* **3**, 1809-1814 (2008).
63. Sternberg, Biomedical Image Processing. *IEEE Computer* **16**, 22-34 (1983).
64. Y. Taniguchi *et al.*, Quantifying E. coli Proteome and Transcriptome with Single-Molecule Sensitivity in Single Cells. *Science* **329**, 533-538 (2010).



65. O. Sliusarenko, J. Heinritz, T. Emonet, High-throughput, subpixel precision analysis of bacterial morphogenesis and intracellular spatio-temporal dynamics *Molecular Microbiology* **80**, 612-627 (2011).
66. J. C. Olivo-Marin, Extraction of spots in biological images using multiscale products. *Pattern Recognition* **35**, (2002).
67. M. H. Ulbrich, E. Y. Isacoff, Subunit counting in membrane-bound proteins. *Nature Methods* **4**, 319-321 (2007).

### 3.8. Supplementary Information

#### 3.8.1. S1 Determining the error of estimating the total YPet- $\beta_2$ content in the cell from the detected out of focus fluorescence

Using the detected fluorescence to estimate the total amount of clamps in the whole cell, we had to verify that we do not miss a significant amount of out of focus fluorescence. We verified this by determining the total intensity measured in a growth channel at different focal positions. A z-stack ( $\pm 1 \mu\text{m}$  out of focus) was measured on the same growth channel. We determined the total intensity of these cells at the different Z-positions (**Supplementary Figure 8c**). As can be seen, the difference is  $< 10\%$ .

#### 3.8.2. S2 Single-molecule fluorescence calibration

We use a similar approach as in (2, 4) to determine the average intensity of a single YPet molecule under our experimental conditions. In brief, purified single YPet molecules were immobilized on a clean cover glass through conjugation via the anti-YPet antibody (Clonotech) (5) either in M9-Gly or cell lysate. Using these two different buffer conditions, we were able to estimate whether cytoplasm of the cell could potential influence the YPet brightness. These YPet proteins were subsequently imaged until they irreversibly photobleached using the same imaging conditions used for the long-timelapse experiments apart from a higher frame rate ( $\sim 300$  ms).

The acquired images were analyzed with custom-written MATLAB software (MathWorks). Prior to spot analyses, we subtracted the uneven background using a rolling ball filter (6, 7), and subsequently corrected for illumination heterogeneity by using a known measured laser beam profile (8). Spot detection is performed in each individual frame for the whole time-lapse series taken using the algorithms as defined in (9). For each of the detected spots, we calculate the integrated intensity by pixel summation (10) in each individual frame. The integrated intensity of each respective spot in each frame is subsequently linked, which results in individual time-lapse integrated intensity traces (**Supplementary Figure 9a**) of single molecule bleaching.

Each individual trace is subsequently processed as follows. First, we perform an edge-preserving Chung-Kennedy (C-K) filter (11) on the trace, using code adapted from (12). We then calculated the pair-wise difference distribution function (PDDF) of the C-K filtered trace (**Supplementary Figure 9b**). The single-sided power spectrum is subsequently calculated for the PDDF, which provided us with the intensity of that

respective single YPet molecule (**Supplementary Figure 9b (inset)**). Performing this analysis for subsequent YPet molecules resulted in a distribution of the intensity of a single YPet molecule in M9-Gly or cell lysate (**Supplementary Figure 9c,d**) for our imaging conditions. A similar approach was utilized to calibrate the intensity of a single mCherry molecule in M9-Gly (**Supplementary Figure 9e**) or cell lysate (**Supplementary Figure 9f**). In both the YPet and mCherry cases, the intensity value used to convert counts to molecules is the average between the means of the M9-Gly and cell lysate measurements. We experimentally confirmed that a YPet- $\beta$  fluorescence standard and a YPet standard provide the same mean fluorescent intensity in our experimental conditions (**Supplementary Figure 9g**). In other words, the YPet intensity is thus unaffected by the fusion to the clamp.

As a control to verify our calibration, we determined the stoichiometry of both the  $\epsilon$ -subunit (DnaQ) of DNA Polymerase III as well as the  $\tau$ -subunit (DnaX) of the clamp loader complex using YPet fusions to the respective proteins. In both the cases we determined the combined stoichiometry of the proteins in both the replisomes. We reproduced the numbers as previously published (2) for these replisome components (**Supplementary Figure 10a,b**). This established that our *in vitro* single-molecule calibration is applicable *in vivo*.

### 3.8.3. S3 Estimation of $\beta_2$ clamp stoichiometry from fluorescence time-lapse data

We determine the number of sliding clamps in the cell from the integrated intensity detected using the single-molecule calibration while correcting for photo-bleaching and converting from monomers to dimers.

To correct for photobleaching, we use the following approach. We assume that cell growth and protein copy number in the cell increases linearly from cell birth until cell division. The ratio of these two numbers, as a function of time, should thus remain constant throughout the cell cycle of a cell if there were no photo-bleaching. However, due to photobleaching, this ratio will decrease as a function of time (**Supplementary Figure 8d**). We fit this curve with a single exponential and multiply the detected fluorescent signal with the appropriate factor as a function of time in order to correct for this decline in fluorescence due to photobleaching.

This intensity corrected for photobleaching is subsequently converted into clamp dimers by dividing the fluorescence signal by the single-molecule calibration integrated intensity value as subsequently dividing by a factor of two to convert from monomers to dimers.

### 3.8.4. S4 Estimation of the immature fraction of YPet molecules

We estimated the immature fraction of YPet (dark fraction) using a similar fluorescence recovery after photobleaching (FRAP) protocol as was previously published by *Badrinarayanan et al.* (13), with slight variation due to our experimental approach. An experiment was prepared as described in **Materials and Methods**, and the cells were allowed to incubate for ~2 hrs without any measurement. A reference fluorescence and

brightfield image were acquired after which, M9 growth medium supplemented with 50  $\mu\text{g}/\text{mL}$  of chloramphenicol (CHL) was flushed through the microfluidic system. The fluorescence was subsequently bleached until roughly background level. This growth medium supplemented with CHL was continuously injected at a rate of 500  $\mu\text{L}/\text{hr}$  while fluorescence and brightfield images were acquired at a frame rate of 15 min with an exposure time of 80 ms for  $\sim 2$  hrs. The CHL-treated cells do not express further proteins, and thus the fluorescence recovery is indicative of the maturation of previously (at the start of the experiment) immature YPet (**Supplementary Figure 8e**). This measured value was estimated to be less than 15% of the starting value. We estimate the maturation time of YPet in *E. coli* under our experimental conditions using the method as describe in Wang *et al* (14) to be  $t_{\text{mature}} < 12.5$  min.

### 3.8.5. S5 Dual-color measurement of DnaX-YPet and mCherry-DnaN

In order to compare the dynamics of the clamp (mCherry- $\beta_2$ ) and the clamp loader complex ( $\tau$ -YPet), we performed a dual-color experiment. Images were acquired in the same way as for the YPet- $\beta_2$  experiment, with the only difference being measuring both the mCherry and YPet signals sequentially. As is evident from **Supplementary Figure 11**, the temporal behavior of the  $\tau$ -subunit is significantly different than that of  $\beta_2$  clamp. In contrast to reaching a steady-state plateau, the  $\tau$ -subunit seems to be highly dynamic and might even be exchanged during replication.

### 3.8.6. S6 Determination of cellular YPet-DnaN molecules by Western blotting

#### S6.1 His-YPet-DnaN reference standard

The cloning was conducted as follows. The *YPet-dnaN* gene from the AB1157 cells was amplified using primers 5' – CTC CAG GGA TCC GAT GTC TAA AGG TGA AGA A – 3' and 5' – GAT CAA CAA GCT TGT GAG GGA CAT TAC AGT CTC ATT GGC ATG ACA ACA TAA GCC – 3'. The PCR amplicon was digested with *HindIII* and *BamHI*, and ligated into expression vector pRSETb (Invitrogen), resulting in construct pRSETb-dnaN. The cloning was confirmed by plasmid sequencing using the T7 promoter and terminator flanking regions. BL21pLysS cells (Promega) were transformed with the pRSETb-dnaN construct to allow inducible protein expression of YPet-DnaN-His. Bacterial cells were grown in LB broth supplemented with ampicillin (100  $\mu\text{g}/\text{mL}$ ) and chloramphenicol (34  $\mu\text{g}/\text{mL}$ ) at 37 °C with shaking until an  $\text{OD}_{600} \sim 0.6$ . The protein induction was performed using 1 mM Isopropyl- $\beta$ -D-thio-galactoside (IPTG). After 4 hr of induction, bacterial cells were collected by centrifugation at 4000 g for 5 min. A 10% SDS-PAGE system was used to analyze the resulting proteins.

The protein purification was performed as follows. The bacterial pellets were thawed on ice, resuspended in buffer (50 mM  $\text{Na}_2\text{HPO}_4$  pH 8.0, 0.3 M NaCl, 8 M urea, 10 mM imidazole) and lysed by sonication. The supernatant was cleared by ultracentrifugation at 20000 g (Beckman JA-17 rotor) for 30 min and subsequently incubated with Ni-NTA beads (Thermo) for 1 h at room temperature (RT). The beads were washed twice with 10 volumes of washing buffer (8 M Urea, 50 mM  $\text{Na}_2\text{HPO}_4$  pH 8.0, 0.3 M NaCl 20 mM imidazole). The beads were subsequently washed with refolding buffer (50 mM  $\text{Na}_2\text{HPO}_4$

pH 8.0, 0.3 M NaCl 20 mM imidazole) to refold the His-Tagged protein. Finally, the N-terminally His6-tagged protein was eluted from the beads using elution buffer 50 mM Na<sub>2</sub>HPO<sub>4</sub> pH 8.0, 0.3 M NaCl and 250 mM imidazole). Proteins were subsequently stored in 50% glycerol at -80 °C until used.

### S6.2 Culture preparation for Western blot

The *YPet-dnaN* cells were grown overnight at 37 °C with shaking in M9 medium supplemented with 0.3% glycerol (Gly), essential nutrients together with the appropriate antibiotics. The subsequent day the overnight culture was sub-cultured (~2%) into 50 mL of the same medium and grown at 37 °C with shaking until an OD<sub>600</sub> ~0.35 was reached. The bacterial culture was serially diluted in Phosphate Buffered Saline (PBS) and subsequently plated on LB Agar medium. The plates were incubated overnight at 37 °C and colony forming units (CFU) were enumerated after 24 hrs. Bacteria were harvested by centrifugation at 4000 g (Eppendorf 5810R) for 5 min. The bacterial pellet was resuspended in 2 mL of B-PER Protein Extraction Reagents (Thermo Scientific) to isolate the total amount of protein using the instructions provided by the manufacturer. Isolated proteins were subsequently stored at -80 °C until used.

### S6.3 Conducting the Western blot

The protein standard (25 ng, 16.6 ng, 12.5 ng of His-YPet-DnaN) and total cell lysed were separated on 10% SDS-PAGE gel under 100V for 1 hr (Biorad). A Western blot assay was conducted using the above-mentioned purified recombinant His-YPet-DnaN protein and total cell lysed. The gel was blotted onto a nitrocellulose membrane using a transfer buffer containing 25 mM Tris (pH = 8.3), 192 mM glycine and 10% methanol at 200 mA for 1 hr at 4 °C. The blotted membrane was blocked with 5% (w/v) skim milk in PBS with 0.05% Tween-20 (T-PBS) for 1 hr at room temperature. The blot was incubated for 1 hr at room temperature (RT) with the primary antibody (Clonetech). After diluting 1:10000, the blots were washed three times with T-PBS and incubated with the secondary antibodies peroxidase conjugated goat anti-mouse IgG at a 1:7500 dilution in T-PBS. The blots were then washed three times with T-PBS and reactions were developed by SuperSignal Western Blot Enhancer (ThermoScientific) according to the manufacturer's instructions.

### S6.4 Quantification of bands

The bands corresponding to the protein standards, as well as bands of appropriate DnaN protein in the crude extract (**Supplementary Figure 12e**) were quantified using Image Lab (Biorad). The intensities of the bands were plotted versus protein concentration (ng) and fitted to a straight line, which yielded a function to related intensity and protein concentration. This linear equation was subsequently used to calculate the amount of that protein present in the crude cell extract. Finally, the number of protein molecules per cell (*C*) was calculated using the following equation:

$$C = \frac{MA}{LN}. \quad (2)$$

Here *M* is the amount of protein in the crude cell extract (grams), *A* is Avogadro's number (molecules/mol), *L* is the molecular mass of the fusion protein (~69 kDa) and *N* is the number of cells (in the volume of crude cell extract used for Western analysis). The two

independent experiments (**Supplementary Figure 12a**) resulted in a combined average of  $198 \pm 12$  YPet-DnaN monomer molecules per cell.

### 3.8.7. S7 Verification of the *dnaN* gene copy number by Southern blotting

The genomic copy number of *dnaN* in WT and YPet-DnaN:*ter*-mCerulean strains was verified by Southern blotting as follows. Genomic DNA was extracted from the strains. The respective DNA was digested by restriction enzymes *Bam*HI and *Sac*I overnight. The digested DNA samples were separated by gel electrophoresis on a 0.8% agarose gel at 30 V for 3 hr. The digested DNA in the gel of the strains was depurinated, denaturated, neutralized and subsequently transferred to a Hybond-N+ nylon membrane (GE Healthcare) using capillary blotting. The probe (870 bp) for southern blotting was obtained from the wild-type AB1157 chromosome by PCR amplifying the region in the *dnaN* gene using the primers 5' – GCA CCG CTA TAG GTA ACG TC – 3' and 5' – TCG TCC TAC GCT ACC GAT TC – 3'. Subsequently, the PCR product was labeled with thermostable alkaline phosphatase using the Amersham AlkPhos Direct kit (GE Healthcare) according to the instructions of the manufacturer. The labeled DNA probe was then added to the membrane and incubated at 55 °C overnight. After adding the luminescent substrate, imaging was performed using Gel-Doc. The resulting number of bands obtained is indicative of the copy number of the probed region in the chromosome (**Supplementary Figure 12b**). As is evident from the blotting experiments, there is only a single band, and thus only a single copy of the *dnaN* gene, in the chromosome of the WT (lane 1) and the strain (lane 3) used in the experiments.

### 3.8.8. S8 Reproducibility of the $\beta_2$ clamp dynamics with an *mCherry-dnaN* fusion

Fluorescent proteins have been shown to potentially aggregate in cells. This aggregation could lead to imaging artifacts (15, 16). We conducted the same long-time lapse experiment with an mCherry-DnaN fusion in order to strengthen the argument that the stoichiometry and accumulation of clamps are reproducible with a different fluorescent protein and less likely due to any aggregation artifact. As it is clearly visible, the result agrees within experimental error with what we have measured for the YPet-DnaN fusion strain (**Supplementary Figure 13**). In determining the mCherry-DnaN stoichiometry we used the single-molecule intensity calibration done for mCherry (**Supplementary Information Section 5.2**).

### 3.8.9. S9 Determining the unloading time of the $\beta_2$ clamp by PALM experiments

The distribution of dwell times of bound clamps is approximated with a single-exponential and has the following form:

$$p(t) = e^{-t/t_m} t_m^{-1}, t_m^{-1} = t_{\text{unload}}^{-1} + \tilde{t}_{\text{bl}}^{-1} \quad (3)$$

Here  $t_m^{-1}$  is the total decay rate of fluorescence, expressed as the sum of bleaching rate ( $\tilde{t}_{\text{bl}}^{-1}$ ) and the unloading rate ( $t_{\text{unload}}^{-1}$ ). As we are ultimately interested in extracting the

unloading time from Equation (3), we first determined the total decay rate and the bleaching rate.

To determine the decay rate of fluorescence for our experiments, the distribution in Equation (3) was fitted to the unloading data through maximum likelihood (ML) estimation, and errors were estimated over  $10^5$  bootstrapped data sets (**Figure 4e (inset)**). This resulted in  $t_m=57\pm 4$  s. Here the error is  $\pm$  s.d. from the bootstrapped data.

To determine the bleaching rate, we first note that in the unloading experiments the proteins are only exposed to light during a fraction of the time they stay bound. Thus, the bleaching time ( $\tilde{t}_{bl}$ ) in Equation (3) is related to the constant-exposure bleaching time ( $t_{bl}$ ) through:

$$\tilde{t}_{bl} = \frac{T}{t_{ex}} t_{bl}. \quad (4)$$

Here  $T$  is the experimental time between frame start times, and  $t_{ex}$  is the experimental exposure time. The constant-exposure (no shuttering, thus  $t_{ex}$  is the camera exposure in this case) bleaching time was independently determined from the unloading experiments (**Supplementary Figure 14**). Using ML estimation of the bootstrapped data sets, we arrive at  $t_{bl} = 7\pm 0.6$  s, giving  $\tilde{t}_{bl} = 80.5\pm 6.9$  s by Equation (4). Using Equation (3), we now extract  $t_{unload} = 195\pm 58$  s, where the confidence intervals were estimated through bootstrapping  $10^6$  datasets (under the assumption that we can ignore the error in our estimate of the bleaching time).

### 3.8.10. S10 Estimating the effective loading rate

During chromosomal replication, the number of DNA-bound  $\beta_2$  clamps ( $N_{\beta_2}$ ) fluctuates around an average in the steady state. We assume here that primers are formed at the rate  $t_p^{-1}$ , and that individual  $\beta_2$  clamps are loaded at a rate of  $t_{load}^{-1}$  (**Supplementary Figure 15**). The number of primers that have not been loaded with a  $\beta_2$  clamp is denoted as  $N_{pre-\beta_2}$  and the number of DNA-bound clamps is denoted as  $N_{\beta_2}$ . For each cell in the steady-state regime of replication, the rate of primer formation ( $t_p^{-1}$ ) and the total rate that  $\beta_2$  clamps are loaded ( $t_{load}^{eff-1}$ ) must balance,

$$t_p^{-1} = t_{load}^{eff-1} = N_{pre-\beta_2} t_{load}^{-1}. \quad (5)$$

This shows that the rate of primer formation sets the effective loading rate within our model in the steady state regime. According to Equation (5), a change in primer formation rate will always be compensated by a change in the steady state value of primers that have not had a  $\beta_2$  clamp loaded. Our data does not allow us to directly extract either  $N_{pre-\beta_2}$  or  $t_{load}$ . Similarly, if a  $\beta_2$  clamp unloads at a rate, ( $t_{unload}^{-1}$ ), the total effective rate at which  $\beta_2$  clamps load ( $N_{pre-\beta_2} t_{load}^{-1}$ ) must equal the total rate they unload ( $N_{\beta_2} t_{unload}^{-1}$ ), giving:

$$N_{\beta_2} t_{unload}^{-1} = N_{pre-\beta_2} t_{load}^{-1}. \quad (6)$$

Together, Equations (5) and (6) give:

$$t_p^{-1} = t_{load}^{eff-1} = N_{\beta_2} t_{unload}^{-1}. \quad (7)$$

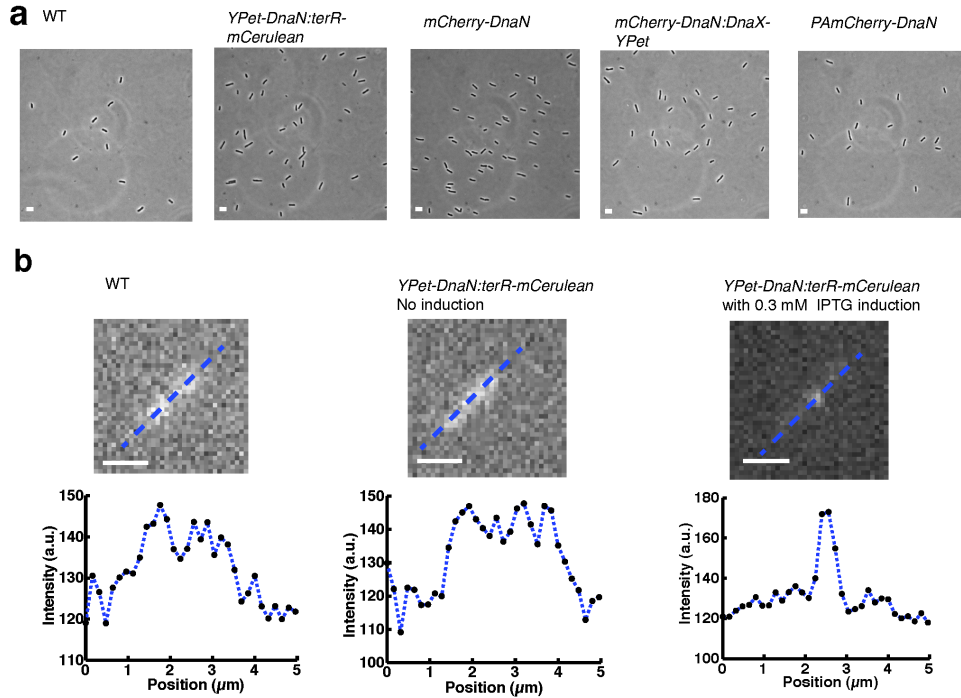
Using Equation (7) the effective loading time is  $t_{\text{load}}^{\text{eff}} = 4 \pm 1$  s. Here the error was derived via propagation from the errors in  $N_{\beta 2}$  and  $t_{\text{load}}$  and calculated from Equation (7):

$$\frac{\langle \Delta t_p^2 \rangle}{t_p^2} = \frac{\langle \Delta t_{\text{unload}}^2 \rangle}{t_{\text{unload}}^2} + \frac{\langle \Delta N_{\beta 2}^2 \rangle}{N_{\beta 2}^2},$$

This is Equation (8) where we have assumed the errors in  $N_{\beta 2}$  and  $t_{\text{unload}}$  to be uncorrelated as they originate from different experiments.

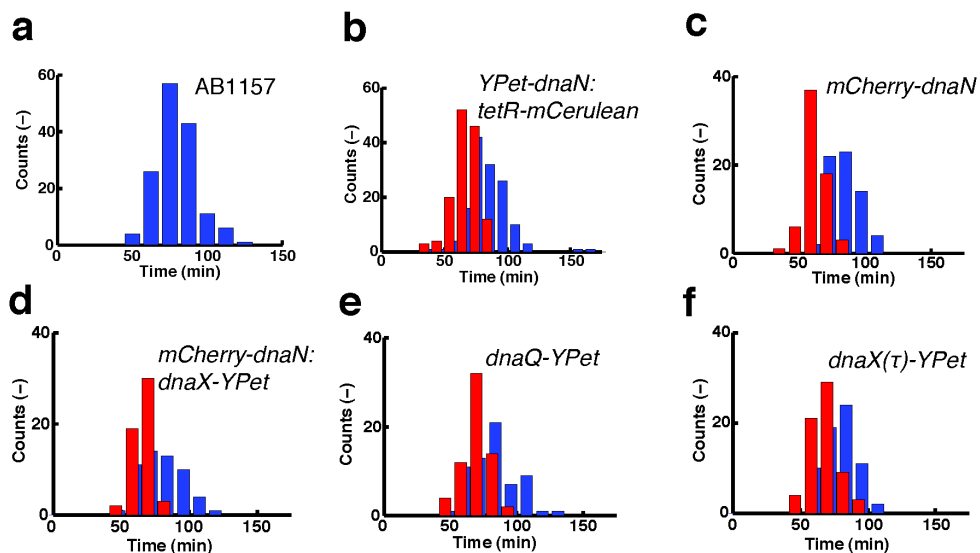
### 3.9. Supplementary Figures and Captions

#### 3.9.1. Supplementary Figure 6: Verification of the different *E. coli* variant used in this study.



(a) Sample phase contrast images of the cells used in this study, which indicate that the physical appearance of these strains is essentially indistinguishable from that of WT. Scale bars: 3  $\mu\text{m}$ . (b) Verification that in the absence of IPTG, *terR-mCerulean* is hardly expressed. Here we show the sample fluorescence images and the corresponding line profile plots. There is essentially no difference between WT and *YPet-dnaN:terR-mCerulean* in the absence of IPTG. Conversely, in the presence of IPTG, a clear focus can be observed in *YPet-dnaN:terR-mCerulean*. Scale bars: 1.6  $\mu\text{m}$

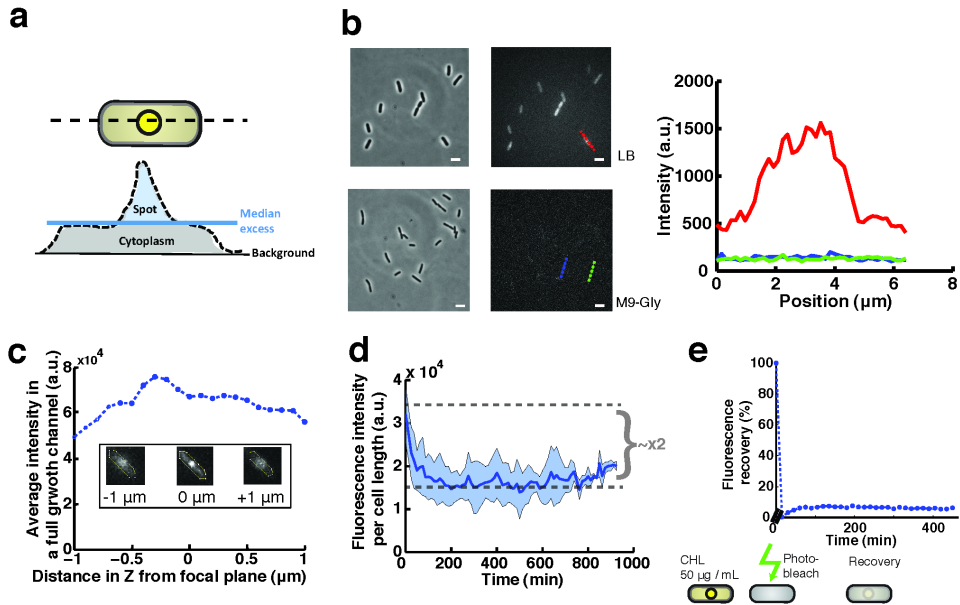
### 3.9.2. Supplementary Figure 7: Cell doubling and replication times of the different *E. coli* variants measured in the microfluidic device.



(a) The doubling distribution of the AB1157 parental strain in the microfluidic device during a long time-lapse experiment. The doubling time is  $t_{\text{double}} = 80 \pm 13$  min,  $n = 148$  (mean and  $\pm$  s.d.). (b-f) The doubling (blue) and replication (red) distributions of the different *E. coli* strains in the microfluidic device during a long time-lapse experiment. The average replication time ( $t_{\text{rep}}$ ) and doubling time ( $t_{\text{double}}$ ) are presented with the error being  $\pm$  s.d. (b) *YPet-dnaN:ter-mCerulean*.  $t_{\text{rep}} = 68 \pm 10$  min and  $t_{\text{double}} = 84 \pm 17$  min ( $n = 137$ ). (c) *mCherry-dnaN*.  $t_{\text{rep}} = 63 \pm 9$  min and  $t_{\text{double}} = 84 \pm 12$  min ( $n = 65$ ). (d) *dnaX( $\tau$ )-YPet:mCherry-dnaN*.  $t_{\text{rep}} = 67 \pm 8$  min and  $t_{\text{double}} = 79 \pm 15$  min ( $n = 54$ ). (e) *dnaQ-YPet*.  $t_{\text{rep}} = 70 \pm 10$  min and  $t_{\text{double}} = 82 \pm 17$  min ( $n = 64$ ). (f) *dnaX( $\tau$ )-YPet*.  $t_{\text{rep}} = 67 \pm 12$  min and  $t_{\text{double}} = 79 \pm 12$  min ( $n = 66$ ).

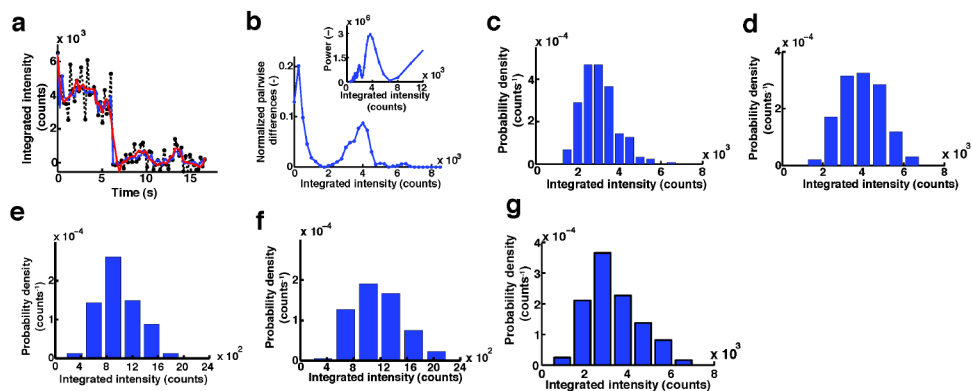


### 3.9.3. Supplementary Figure 8: The fluorescence signal definition and the different microscopy control experiments performed.



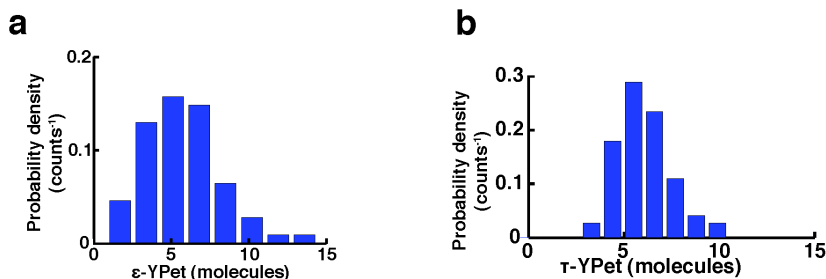
(a) Illustration of how the integrated line-profile is used to define sample background intensity, foci intensity, and total intensity in the cell. (b) Comparison of the degree of auto-fluorescence of *E. coli* AB1157 in LB and M9-Gly growth medium. (left) Example phase contrast and corresponding fluorescence images of cells grown in LB or M9-Gly. Scale bars: 3  $\mu\text{m}$ . (right) Example line profile plots of a cell in LB (red curve), M9-Gly (blue curve) and an arbitrary position in the M9-Gly fluorescence image (green curve). (c) A sample fluorescence signal of a z-stack of cells in a full growth channel (range: -1  $\mu\text{m}$  to +1  $\mu\text{m}$  with respect to the focus). (d) The bleaching behavior of the YPet- $\beta_2$  fluorescence signal during a long-time lapse experiment. The relative change in the intensity as a function of time is used to correct for YPet- $\beta_2$  bleaching. (e) Fluorescence recovery after photobleaching of cells in the presence of chloramphenicol.

### 3.9.4. Supplementary Figure 9: Single YPet fluorescence intensity calibration.



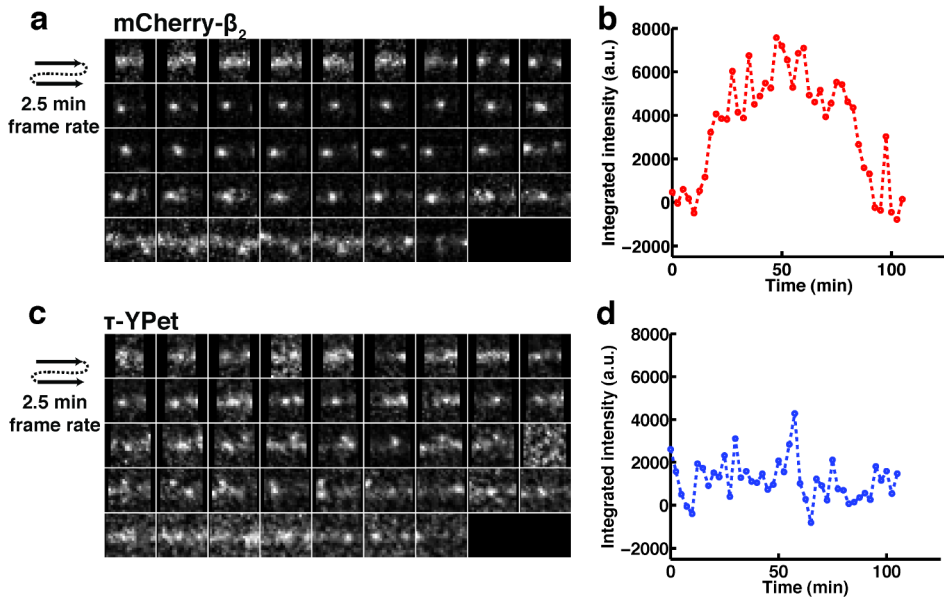
(a) A sample integrated intensity trace of a single immobilized YPet molecule as a function of time. Here we show the data points (black line), the Chung-Kennedy (CK) filtered result of the data (blue line) and the Savitsky-Golay (SK) filtered trace (red). In the further analysis steps, the CK-filtered data result was used. (b) The pairwise difference distribution function (PDDF) of the CK-filtered trace in (a) and its corresponding power spectrum (inset). (c) The integrated intensity distribution of single YPet molecules imaged in M9 growth medium supplemented with 0.3% glycerol and essential nutrients ( $n = 235$ ,  $\mu = 3070$ ,  $\sigma = 856$ ). (d) The integrated intensity distribution of single YPet molecules imaged in french-pressed cell lysate ( $n = 245$ ,  $\mu = 3884$ ,  $\sigma = 1108$ ). (e) The integrated intensity distribution of single mCherry molecules imaged in M9 growth medium supplemented with 0.3% glycerol and essential nutrients ( $n = 107$ ,  $\mu = 990$ ,  $\sigma = 318$ ). (f) The integrated intensity distribution of single mCherry molecules imaged in french-pressed cell lysate ( $n = 102$ ,  $\mu = 1158$ ,  $\sigma = 380$ ). Images were acquired, for both (e) and (f), at an exposure time of 400 ms. The resulting intensity values were subsequently converted to values equivalent to an exposure time of 80 ms by dividing them by 5. (g) The integrated intensity distribution of single YPet- $\beta$  molecules imaged in M9 growth medium supplemented with 0.3% glycerol and essential nutrients ( $n = 131$ ,  $\mu = 3305$ ,  $\sigma = 856$ ).

### 3.9.5. Supplementary Figure 10: Stoichiometry of DnaQ and DnaX.



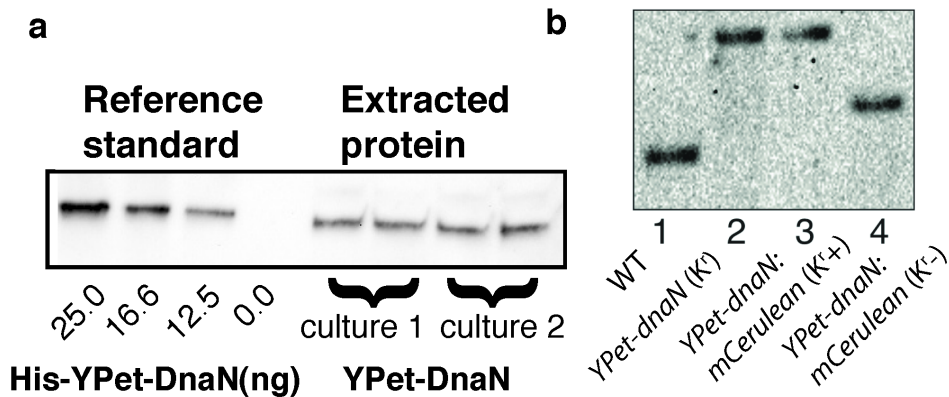
(a-b) The distributions of the stoichiometry of the DnaQ ( $\epsilon$ -YPet) ( $n = 64$ ) and DnaX ( $\tau$ -YPet) ( $n = 66$ ) molecules within the replisome.

### 3.9.6. Supplementary Figure 11: The dual-color experiment that simultaneously monitors the temporal dynamics the $\beta_2$ clamp and the clamp loader ( $\tau$ ).



(a) A sample temporal montage of the mCherry- $\beta_2$  clamp behavior for one replication cycle. (b) The integrated intensity of the foci from the mCherry- $\beta_2$  signal shown in (a). (c) A sample temporal montage of the  $\tau$ -YPet behavior for one replication cycle. (d) The integrated intensity of the foci from the  $\tau$ -YPet signal shown in (c).

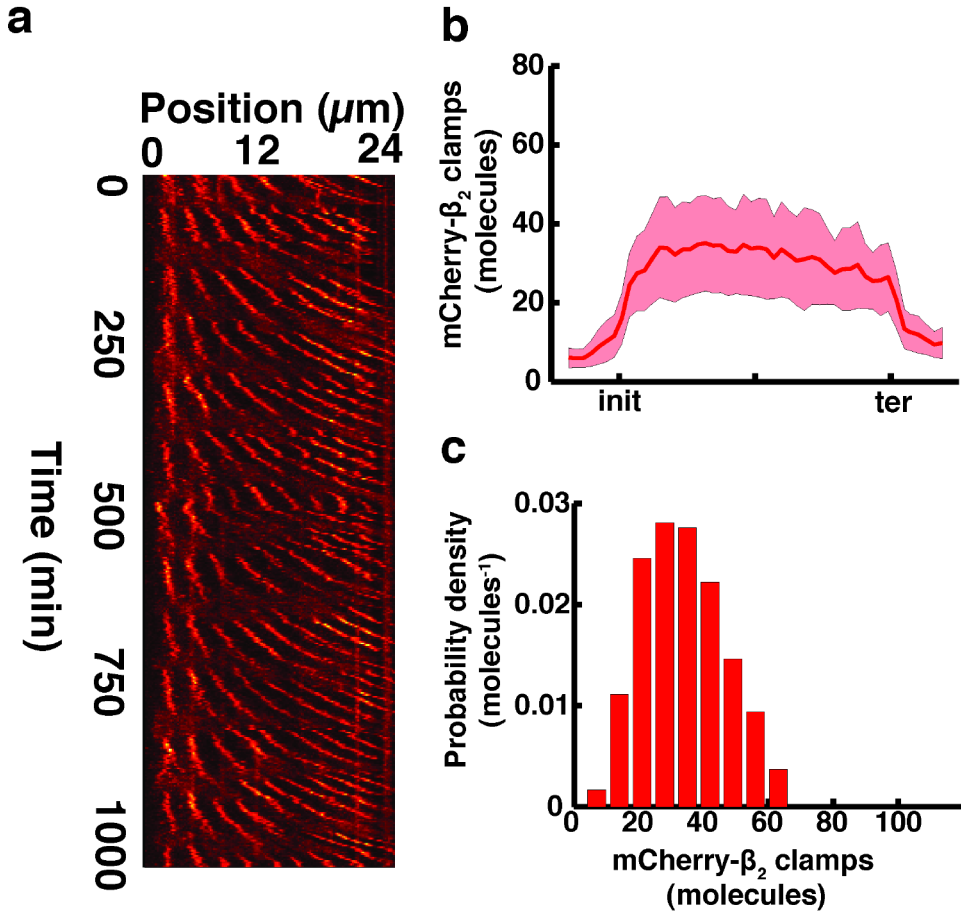
### 3.9.7. Supplementary Figure 12: Western blot and Southern blot control experiments.



**(a)** The Western blot result, indicating the average number of clamps in the cell. Using the purified His-YPet-DnaN (Lanes 1-4 from left to right) as the quantitative standard we determined the average amount of two independent cultures (Lanes 5-8 from left to right)

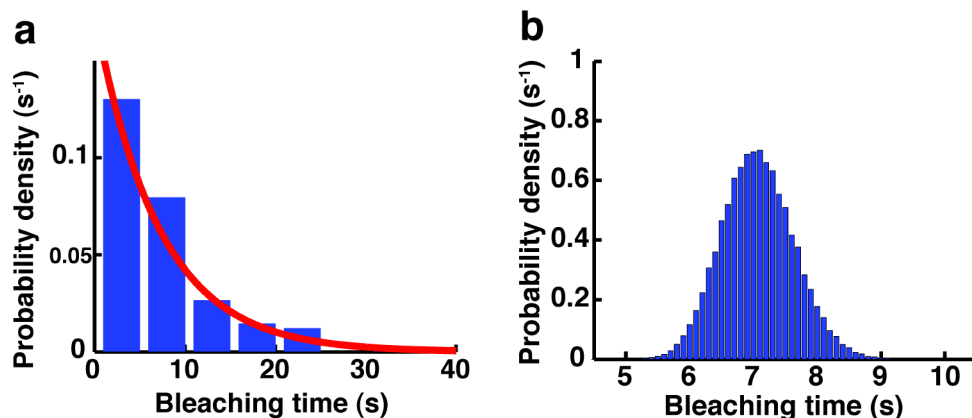
**(b)** The Southern blot result, indicating that there is only a single gene copy of *dnaN*. Here we compare WT cells (lane 1) to three different *E. coli* variant (lane 2-4) we have created. We used the strain in lane 3 (described in **Materials and Methods**) in this study.

**3.9.8. Supplementary Figure 13: The temporal dynamics of the YPet- $\beta_2$  are reproduced with an mCherry- $\beta_2$  fusion.**



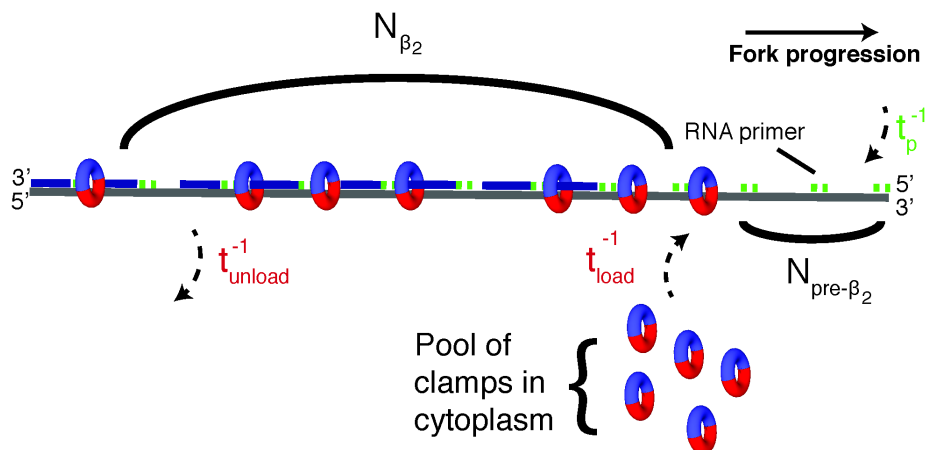
**(a)** A sample kymograph of a single growth channel during an overnight experiment. **(b)** The behavior of DNA-bound mCherry- $\beta_2$  clamps in both the replisomes during replication. Here we have normalized the time to make averaging of different cells, possible ( $n = 65$ ). The error bars represent  $\pm$  s.d., from the mean. **(c)** The distribution of DNA-bound mCherry- $\beta_2$  clamps in both replisomes during steady-state replication. The mean number of DNA-bound clamps is  $N_{\beta_2} = 34$  with s.d. = 12, s.e.m. = 1.5.

### 3.9.9. Supplementary Figure 14: The PAMCherry bleaching curve together with the bootstrapped values for uncertainty estimation.



(a) The distribution of the bleaching time measured under our experimental conditions. The red curve is a single exponential MLE fit to the data. (b) The distribution of the fitted bleaching time constants over  $10^6$  bootstrapped data sets, from which the standard deviation for the bleaching time was determined.

### 3.9.10. Supplementary Figure 15: A schematic illustration of the recycling model used.



In our model, we have five different variables. They are the number of primers having a loaded  $\beta_2$  clamp ( $N_{\beta_2}$ ), the number of primers that do not have a loaded  $\beta_2$ -clamp ( $N_{\text{p,unocc}}$ ), the rate of loading a  $\beta_2$  clamp ( $k_{\text{load}}$ ), the rate of unloading a clamp ( $k_{\text{unload}}$ ), and the rate at which primers are being formed ( $k_p$ ).

### 3.10. Supplementary Tables

#### 3.10.1. Supplementary Table 1

<i>E. coli</i> strain	Generation time (min)	
	M9-Gly	LB
AB1157	103±4	37±2
<i>YPet-dnaN:ter-mCerulean</i>	103±1	37±2
<i>PAmCherry-dnaN</i>	102±3	38±2
<i>mCherry-dnaN</i>	105±3	38±1
<i>mCherry-dnaN:dnaX-YPet</i>	103±1	39±1

Bulk generation times of the different strains in LB and M9-Gly growth medium. The numbers specified indicate the mean  $\pm$  s.d. The generation times were determined from three independent cultures for all the strains.

#### 3.10.2. Supplementary Table 2

Plasmids	Relevant genotype	Construction
pKD46	Plasmid with $\lambda$ -Red recombinase genes expressed under arabinose promoter	Created by standard cloning (1)
pROD44	<i>YPet</i> template plasmid for N-terminus fusion	Cloning <i>YPet</i> and <i>frit-kanR-frit</i> in pUC18 backbone (2)
mCherry-N	<i>mCherry</i> template plasmid for N-terminus fusion	Cloning <i>mCherry</i> and <i>frit-kanR-frit</i> in pUC18 backbone
PAmCherry1	<i>PAmCherry1</i> template plasmid	Cloning <i>PAmCherry1</i> in place of <i>YPet</i> in pROD44 backbone

Summary of different plasmids used in this study.

### 3.10.3. Supplementary Table 3

Strains	Relevant genotype	Construction
BN1110	AB1157 strain containing pKD46 plasmid	<i>E. coli</i> K-12 derivative (3)
BN1108	<i>tetO</i> array (50 kb clockwise from the <i>dif</i> -site) and <i>PlacI-tetR-mCerulean</i> in place of <i>galK</i>	Phage transduction (3)
BN1109	<i>YPet</i> fused to <i>dnaN</i>	$\lambda$ -red recombination: <i>Ypet-kanR</i> from pROD44 $\rightarrow$ BN1107
BN1219	<i>YPet-dnaN</i> and <i>ter-mCerulean</i>	Phage transduction: BN1109 $\rightarrow$ BN1108
BN1682	<i>mCherry</i> fused to <i>dnaN</i> with <i>kanR</i>	$\lambda$ -red recombination: <i>mCherry-kanR</i> from mCherry-N $\rightarrow$ BN1107
BN1683	<i>mCherry</i> fused to <i>dnaN</i>	<i>Flp-frt</i> recombination: <i>kanR</i> flipped out from BN1682
BN1684	<i>dnaX</i> ( $\tau$ ) fused to <i>YPet</i> with <i>kanR</i>	$\lambda$ -red recombination (3)
BN1864	<i>dnaX</i> ( $\tau$ )- <i>YPet</i> and <i>mCherry-dnaN</i>	Phage transduction: BN1684 $\rightarrow$ BN1683
BN1867	<i>PAmCherry1</i> fused to <i>dnaN</i> with <i>kanR</i>	$\lambda$ -red recombination: <i>PAmCherry1-kanR</i> from <i>PAmCherry1</i> $\rightarrow$ BN1107

Summary of different strains used in this study.



### 3.11. Supplementary References

1. K. A. Datsenko, B. L. Wanner, One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proceedings Of The National Academy Of Sciences Of The United States Of America* **97**, 6640-6645 (2000).
2. R. Reyes-Lamothe, D. J. Sherratt, M. C. Leake, Stoichiometry and architecture of active DNA replication machinery in *Escherichia coli*. *Science (New York, N.Y.)* **328**, 498-501 (2010).
3. R. Reyes-Lamothe, C. Possoz, O. Danilova, D. J. Sherratt, Independent Positioning and Action of *Escherichia coli* Replisomes in Live Cells. *Cell* **133**, 90-102 (2008).
4. M. C. Leake *et al.*, Stoichiometry and turnover in single, functioning membrane protein complexes. *Nature* **443**, 355-358 (2006).
5. G. I. Mashanov, D. Tacon, A. E. Knight, M. Peckham, J. E. Molloy, Visualizing single molecules inside living cells using total internal reflection fluorescence microscopy. *Methods (San Diego, Calif)* **29**, 142-152 (2003).
6. Sternberg, Biomedical Image Processing. *IEEE Computer* **16**, 22-34 (1983).
7. A. B. Loveland, S. Habuchi, J. C. Walter, A. M. van Oijen, A general approach to break the concentration barrier in single-molecule imaging. *Nature Methods* **9**, 987-992 (2012).
8. Y. Taniguchi *et al.*, Quantifying *E. coli* Proteome and Transcriptome with Single-Molecule Sensitivity in Single Cells. *Science* **329**, 533-538 (2010).
9. J. C. Olivo-Marin, Extraction of spots in biological images using multiscale products. *Pattern Recognition* **35**, (2002).
10. M. H. Ulbrich, E. Y. Isacoff, Subunit counting in membrane-bound proteins. *Nature Methods* **4**, 319-321 (2007).
11. S. H. Chung, R. A. Kennedy, Forward-backward non-linear filtering technique for extracting small biological signals from noise. *Journal of Neuroscience Methods* **40**, 71-86 (1991).
12. N. F. Reuel *et al.*, NoRSE: noise reduction and state evaluator for high-frequency single event traces. *Bioinformatics* **28**, 296-297 (2012).
13. A. Badrinarayanan, R. Reyes-Lamothe, S. Uphoff, M. C. Leake, D. J. Sherratt, *In vivo* Architecture and Action of Bacterial Structural Maintenance of Chromosome Proteins. *Science* **338**, 528-531 (2012).

14. S. Wang, J. R. Moffitt, G. T. Dempsey, X. S. Xie, X. Zhuang, Characterization and development of photoactivatable fluorescent proteins for single-molecule-based superresolution imaging. *Proceedings Of The National Academy Of Sciences Of The United States Of America* **111**, 8452-8457 (2014).
15. M. T. Swulius, G. J. Jensen, The helical MreB cytoskeleton in Escherichia coli MC1000/pLE7 is an artifact of the N-terminal yellow fluorescent protein tag. *Journal of Bacteriology* **194**, 6382-6386 (2012).
16. D. Landgraf, B. Okumus, P. Chien, T. A. Baker, J. Paulsson, Segregation of molecules at cell division reveals native protein localization. *Nature Methods* **9**, 480-482 (2012).



# 4 THE PROGRESSION OF REPLICATION FORKS AT NATURAL REPLICATION BARRIERS IN LIVE BACTERIA

*Protein-DNA complexes are one of the principal barriers the replisome encounters during replication. One such barrier is the Tus-Ter complex, which is a direction dependent barrier for replication fork progression. The details concerning the dynamics of the replisome when encountering these Tus-Ter barriers in the cell are poorly understood. By performing quantitative fluorescence microscopy with microfluidics, we investigate the effect on the replisome when encountering these barriers in live Escherichia coli (E. coli) cells. We make use of an E. coli variant that includes only an ectopic origin of replication that is positioned such that one of the two replisomes encounters a Tus-Ter barrier before the other replisome. This enables us to single out the effect of encountering a Tus-Ter roadblock on an individual replisome. We demonstrate that the replisome remains stably bound after encountering a Tus-Ter complex from the non-permissive direction. Furthermore, the replisome is only transiently blocked and continues replication beyond the barrier. Additionally, we demonstrate that these barriers affect sister chromosome segregation by visualizing specific chromosomal loci in the presence and absence of the Tus protein. These observations demonstrate the resilience of the replication fork to natural barriers and the sensitivity of chromosome alignment to fork progression.*

## 4.1. Introduction

All dividing cells must ensure the accurate and timely replication of their genome. The replication of DNA is an intricate process undertaken by the ubiquitous multi-protein complex known as the replisome (1-3). In *Escherichia coli* (*E. coli*), the two independent replisomes (4) assemble at a single origin of replication (*oriC*) and subsequently advance in opposite directions at equal rates to synthesize the 4.6 Mbp genome, while the newly replicated DNA is sequentially segregated (5-9) prior to cell division. During fork progression, the replisomes encounter a substantial number of impediments preceding their fusion in the terminus region (10). Stalled replication forks are a potential source of genome instability, and thus a risk of cell viability in general (11). The major source of a hindrance for a replisome during replication is believed to be protein-DNA complexes (12). A specific natural protein-DNA complex that is particularly involved in the termination of replication is the 36 kDa monomeric Tus protein bound to one of the ~10 specific 23 bp DNA sequences (*Ter*-sites) (13-16). The majority of these *Ter* sequences are spread across the terminus region of the chromosome, spanning ~2 Mbp across the two replichores (chromosome halves) (17) (**Figure 1A**).

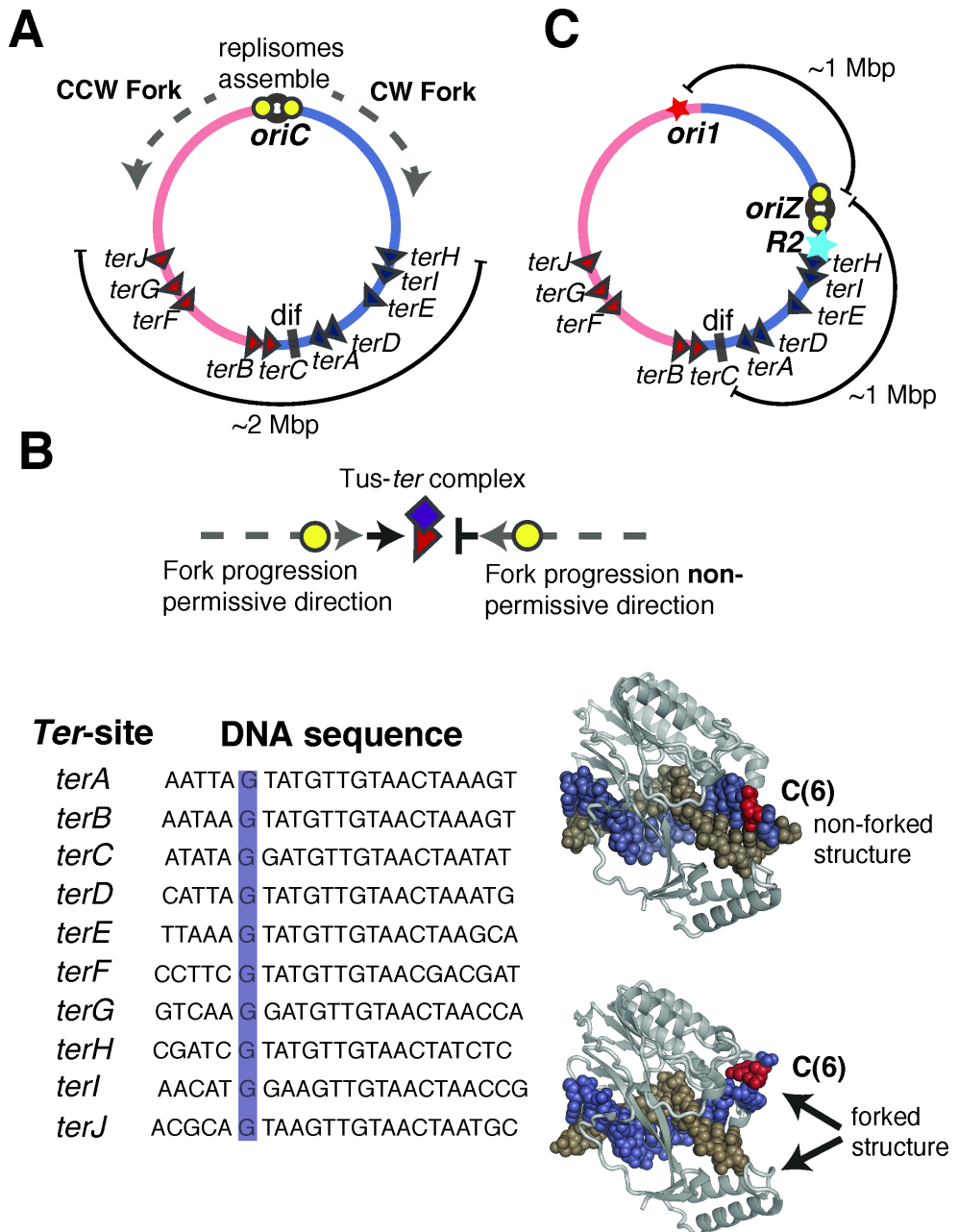
Replication forks are believed to be stalled when approaching such a Tus-*Ter* complex from one direction (non-permissive side) but not from the other direction (permissive side) (18-22) (**Figure 1B, top**), creating a replication fork barrier (RFB). The dominant mechanism of forming such an RFB is thought to be the binding of a conserved G-C(6) base pair residue (**Figure 1B, left**) to Tus. Lock formation is formed through the separation of the parental DNA strand in the non-permissive direction by the replicative helicase (DnaB) until the conserved residue is free and subsequently binds to the binding pocket of Tus (**Figure 1B, right**) (23). Other lock formation mechanisms, for example, protein-Tus interactions, have also been proposed (24) but seem unlikely to be the dominant mechanism (25). The binding of Tus to a *Ter* site is strongly dependent on ionic strength (26). For example, Tus has a dissociation constant of  $K_d \sim \text{pM}$  and a half-life of  $t_{1/2} \sim 500$  min in 0.15 M potassium glutamate, while in 0.25 M potassium chloride (KCl), it exhibits a  $K_d \sim \text{nM}$  and a  $t_{1/2} \sim 2$  min (27). In physiological ionic strength (150 mM KCl), the half-life for a Tus-*Ter* locked complex was found to be  $t_{1/2} \sim 500$  min (23). The RFB system is conserved in *B. subtilis* (28-30) and has also been demonstrated to be capable of artificially pausing a replication fork in yeast and mammalian cells (31, 32).

The biological significance of the Tus-*Ter* RFB and the fate of a replisome when encountering such a barrier in live cells remains poorly understood (21). The 'replication fork trap' (RFT) model suggests that the Tus-*Ter* complexes act as a safeguard, ensuring that forks can enter but not exit the inner termination zone, thereby restraining fork fusion to the zone roughly opposite the origin of replication (33). However, it is not clear why this would be necessary. It might assist fork fusion by ensuring that one of the two replisomes 'waits' for the other oppositely progressing replisome to complete replication, or reduce the incidence of head-on collisions between the replication and transcription machinery (10). It might also be related to the presence of the specific locus important for segregation (*dif*), situated between the innermost *Ter*-sites (**Figure 1A**). The multiplicity of *Ter*-sites on each of the two chromosome arms suggests that a replication fork may, despite the strong DNA binding of Tus, frequently overcome a non-permissive Tus-*Ter* roadblock. A bulk *in vivo* study of plasmid fork arrests (34) suggests that the different *Ter*-sites have distinct fork

arrest efficiencies, with the innermost sites (*TerA-E*) being the most effective (20 - 35%). This is in line with an *in vitro* study that demonstrated that Tus binds strongly to *TerA-E* and less strongly to other *Ter*-sites (35). It remains unclear what happens to a replisome when it encounters such a block. Given the relatively low efficiencies of the RFBs and the *in vitro* robustness of the replisome in head-on collisions (36, 37), it could be the case that the replisome remains intact after encountering a Tus-*Ter* block, or rapidly restart and proceed further with replication.

To gain insight into the response of a replisome as it encounters a Tus-*Ter* roadblock under *in vivo* conditions, we study single live *E. coli* cells that contain only an ectopic origin (38) in either the presence or absence of Tus (**Figure 1C**). The *oriZ* position is such that the CW replisome encounters a non-permissive Tus-*Ter* complex earlier on in replication than the CCW replisome. The strain also includes a genomic rearrangement of a highly transcribed region (*rrnCABE* operon cluster) that alleviates head on replication-transcription conflicts of the CCW replisome (39). Utilizing custom microfluidics (40) to ensure healthy cell physiology in combination with time-lapse wide-field fluorescence microscopy, we quantitatively investigate chromosomal replication and segregation during successive cell cycles. We track the replisome by its progression through its components DnaN (the  $\beta_2$  sliding clamp) and DnaQ (the  $\epsilon$ -subunit of DNA polymerase III), and specific chromosomal loci on the left (L) and right (R) replicore arms of the chromosome (**Figure 1C**). This enables us to investigate the effect of the Tus roadblock on both replication and segregation.

Our data indicate that while the Tus-*Ter* roadblock impedes replisome progression, the replisome is sufficiently resilient (as assessed from the DNA-bound DnaQ and DnaN molecules) to avoid complete disassembly at the barrier, proceeding to overcome it and finish the replication process in an overall time that only exceeds that of cells with the native origin of replication (*oriC-dnaN*) by 15%. The replication time in the absence of Tus (*oriZ-dnaN: $\Delta$ Tus* cells) is in excellent agreement with that of *oriC-dnaN* cells, confirming that the Tus-*Ter* complex is indeed the barrier responsible for the increased replication time in *oriZ-dnaN* cells. We found that the generation time of all  $\Delta$ Tus mutants with an ectopic or native origin of replication exceeded that of *oriC-dnaN* cells. We investigated the effect of the Tus-*Ter* block on the chromosomal segregation organization since this pattern is sensitive to fork progression (41, 42). The sister chromosome alignment (SCA) patterns after replication for the *oriZ-dnaN* and *oriZ-dnaN: $\Delta$ Tus* cells were found to differ: SCA patterns changed from a predominant Left-Right-Left-Right (LRLR) (*oriZ-dnaN* cells) configuration to an RLLR configuration (*oriZ-dnaN: $\Delta$ Tus* cells). This suggests that the RFB caused by the Tus-*Ter* complex can influence the manner in which the sister chromosomes are segregated and positioned following replication. Collectively, our results show that the replisome is stable, or rapidly restarts after having encountered a natural impediment in the cell. It does not collapse completely but rather remains DNA-bound and overcomes this barrier, possibly with assistance from other proteins.



**Figure 1. Investigating the Tus-Ter replication block efficacy in an *E. coli* strain with an ectopic origin.** (A-B) The different *E. coli* chromosome arrangements. The left (red) and right replicore (blue), together with their respective *Ter*-sites (triangles) are shown. The dashed gray lines indicate the directionality of the two independent replisomes (yellow circles) movement. (A) The parental WT

strain (*oriC-dnaN*-strain) which has roughly the equal genomic distance from *oriC* till the innermost *Ter*-sites for both replisomes. **(B)** Polar-fork arrest by the Tus-*Ter* complex. (top) A schematic representation illustrating two replisomes approaching the permissive and non-permissive faces of a Tus-*Ter* complex. The purple square represents the Tus protein. The replisome approaching from the non-permissive side is believed to be blocked by the Tus-*Ter* complex indicated here with a flat arrowhead. (left, right) The dominant mechanism of polar fork arrest by a Tus-*Ter* complex. (left) The DNA sequences of the different *Ter*-sites. Here the conserved G is highlighted by a light blue rectangle. (right) The crystal structures of a non-forked Tus-bound *Ter*-site and a forked Tus-bound *Ter*-site approached from the non-permissive direction. Here the Tus protein is depicted in light gray. Images were generated using the Protein Data Bank (PDB) files, 1ECR and 2EWJ (23, 66). DNA molecules are represented in purple and brown, while the C(6) residue is highlighted in red. **(C)** The strain with an ectopic origin (*oriZ-dnaN*-strain) (38). Here the origin of replication (*oriZ*) is positioned at 344 kb on the *E. coli* genetic map. The insertion position of a *lacO* operator array at 3908 kb (*oriI*) and a *tetO* operator array at 366 kb (*R2*) are indicated with a red and blue star respectively.

## 4.2. Results

### 4.2.1. Delineating the time when the CW replisome encounters a Tus-*Ter* complex from the non-permissive side

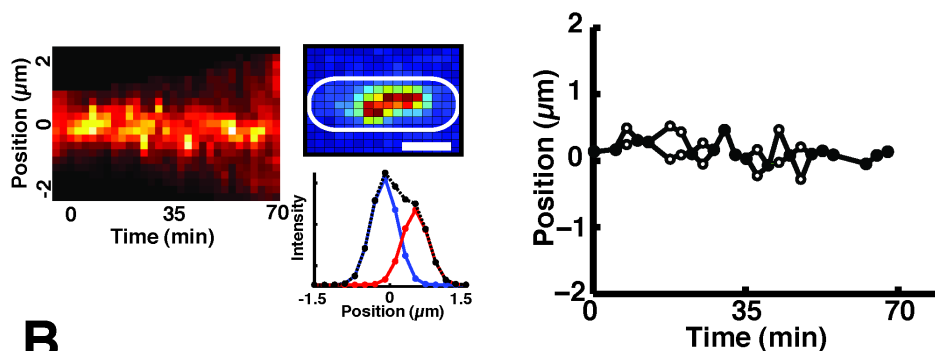
We utilize time-lapse fluorescence microscopy in combination with a custom-built microfluidic device (40) to track the independent replisomes and chromosomal loci during chromosomal replication. The individual replisomes are visualized via a functional YPet-DnaN chromosomal fusion (38), and the left- and right replicore arms are imaged using a fluorescent-repressor-operator system (FROS) (43). FROS enables one to image chromosomal loci by engineering binding sites at a defined location on the chromosome to which specific proteins bind. If these proteins are fluorescently tagged, one then visualizes that locus of the chromosome in the form of a focus. The *lacO* array (*oriI*) is positioned on the left arm, 15 kb CCW from the former position of *oriC*, and the *tetO* array (*R2*) is positioned on the right arm, 21 kb CW of *oriZ* (38) (**Figure 1C**). Fusions of LacI-mCherry and TetR-mCerulean are expressed at low levels to keep any perturbation to the generation and replication times (44) to a minimum (38, 45, 46). Kymographs of sample individual cells illustrate the YPet-DnaN and *oriI*-mCherry movements during replication (**Figure 2A-B, left**). The respective temporal montages of the individual images are shown in the **Supplementary Figure S1A-B**. We determine the positions of the respective foci for each individual image of a time-lapse measurement by fitting a 1D-Gaussian function across the long axis to the profile summed across the short axis of the bacterium (**Supplementary Section 1.8.1**). **Figure 2A-B (center)** depict sample fluorescence images and their corresponding Gaussian fits. The calculated positions per individual image are used to track the YPet-DnaN and *oriI*-mCherry foci over a complete replication cycle in each individual cell. We perform an equivalent analysis to detect and track the *R2*-mCerulean loci (**Supplementary Figure S2**).

The temporal information of *oriI*-mCherry enables us to establish the moment during replication at which the CW replisome encounters the Tus-*TerC* roadblock. This is because the distances of *oriI* from *oriZ* (~1050 kbp) and of *TerC* from *oriZ* (~1100 kbp) can be seen as equivalent (~1 Mbp) (**Figure 1C**). We determine the time-point at which the *oriI*-mCherry focus spatially doubles (**Figure 2C**). Since the two replisomes are believed to

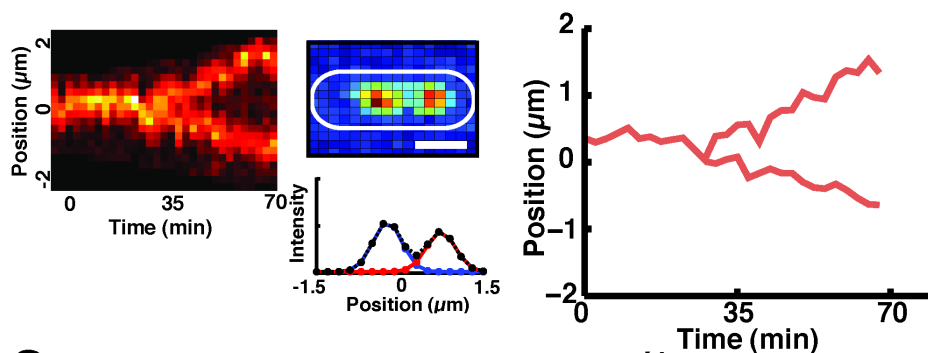


progress at the same rate on average ( $v_{\text{rep}} \sim 550$  bp/s) (**Supplementary Section 1.9**), this time-point is indicative of when the CW replisome encounters the Tus-*TerC* roadblock from the non-permissive direction (**Figure 2C, inset**).

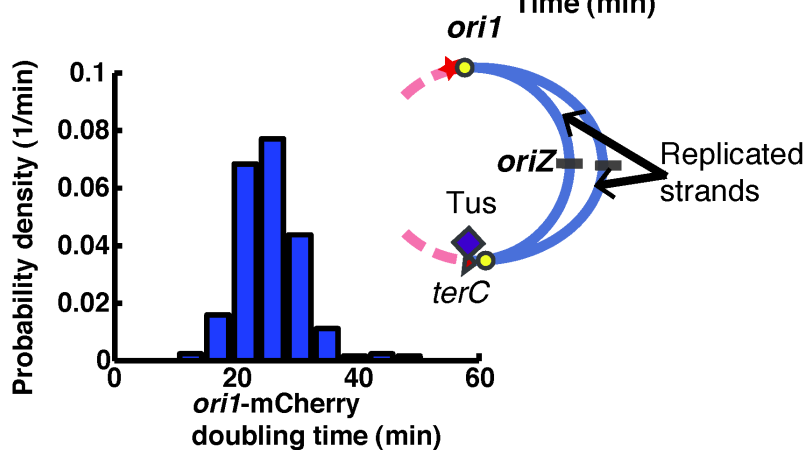
**A**



**B**



**C**



**Figure 2. An example individual replication cycle of *oriZ-dnaN* cells. (A-B)** Representative DnaN-YPet and *oriI*-mCherry fluorescent signals. (A-B, left) Individual kymographs of the replisome (YPet-DnaN) and *oriI*-mCherry locus for one complete replication cycle constructed by summation of the pixel intensities along the short axis of the cell. The replisomes remain relatively close together, while the *oriI*-mCherry loci segregate and move to opposite cell poles. (A-B, centre) Sample fluorescent foci of YPet-DnaN and *oriI*-mCherry imaged (represented here in pseudocolour) together with their respective Gaussian fits. Scale bars, 1.3  $\mu\text{m}$ . (A-B, right) The complete YPet-DnaN (A, right) and *oriI*-mCherry (B, right) traces shown in (A, left) and (B, left) determined from the Gaussian fitted positions from individual images for each time point. Here a single YPet-DnaN focus is indicated by a filled circle and two individual foci are indicated by open circles. The lines between data points are included to aid the reader. *oriI*-mCherry positions over time are indicated by solid red lines. (C) The distribution of the time point during replication when the *oriI*-mCherry focus doubles. The mean spatial doubling time of the *oriI*-mCherry focus is  $t_{\text{oriI}}=26\pm 6$  min (Error is  $\pm$  SD,  $n_{\text{cells}}=282$ ). Here  $t = 0$  is taken as the initiation of replication. (Inset) We schematically show that when the *oriI*-mCherry focus is reached it is indicative of the time point when the CW replisome reaches the Tus-*TerC* complex.

The average time following initiation required for the CCW replisome to replicate *oriI* is  $t_{\text{oriI}} = 26\pm 6$  min (Error is  $\pm$  SD;  $n = 282$ ; **Figure 2C**). This measured time is in agreement with previous measurements (38), as well as with the expected value ( $t_{\text{oriI,expected}} = 32$  min) calculated using the distance of *oriI* from *oriZ* and the average replication speed ( $v_{\text{rep}}$ ).

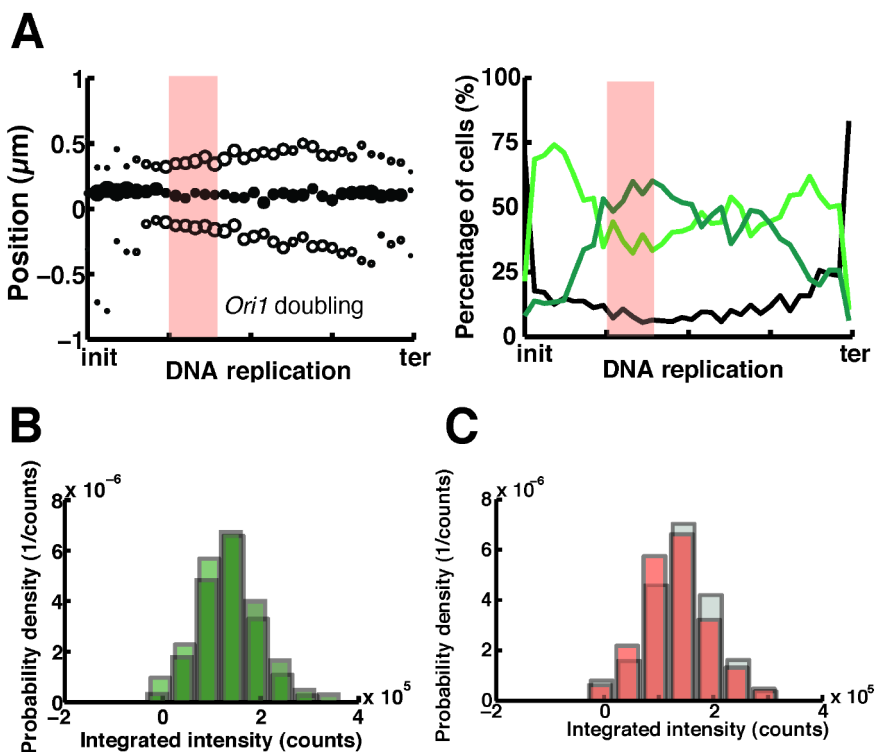
#### 4.2.2. A replisome remains DNA-bound after encountering a Tus-*Ter* roadblock

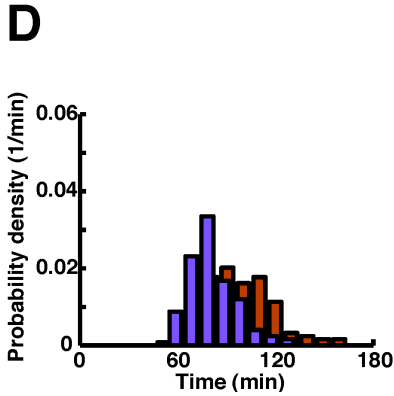
Since it is conceivable that the replisome disassembles when encountering a Tus-*Ter* complex, we examine the dynamics of the replisome in single cells from initiation till termination, with specific attention given to the time point when the replisome reaches the Tus-*TerC* roadblock. Initiation and termination are taken as focus appearance and disappearance. The DNA-bound YPet-DnaN foci are detected and tracked during the whole replication process (**Figure 3A, left**). The size of the markers in the average time-resolved trace (**Supplementary Section 1.8.2**) presented is indicative of the percentage of cells at that time point having either a single focus (black dots) or two foci (black circles). The majority of the cells exhibits a single focus at initiation and just prior to termination, while during elongation cells display single and double foci at equal frequencies (**Figure 3A, right**).

To rule out that the single focus observed in roughly 50% of the cells during elongation results from the CW replisome disassembly at the Tus-*TerC* site, we quantify and compare the DNA-bound YPet-DnaN intensity under two conditions. First, we compare the intensity of a single focus (light green) to the sum of the intensity of two foci (dark green) during replication (**Figure 3B**). We observe only a small difference between the means ( $<10\%$ ), which suggests that there is no change in the number of DNA-bound YPet-DnaN molecules regardless of whether one or two foci are detected. Thus, the differing number of foci detected does not reflect replisome disassembly, but likely results from occasional DNA conformational changes that result in sufficiently close proximity (below the resolution limit imposed by diffraction) of the two replisomes to preclude their separate identification.

Secondly, we establish whether the CW replisome is disassembled after encountering the Tus-*Ter* roadblock. This is accomplished by comparing the intensity of DNA-bound YPet-DnaN prior (grey) and post (red) doubling of the *ori1*-mCherry focus (as assessed by the appearance of a second focus, **Figure 3C**). This comparison is performed by summing the intensities of all the frames prior or post *ori1*-mCherry focus doubling, respectively (similar results are obtained when taking smaller numbers of frames into account, see **Supplementary Table S1**). We only detect slight differences ( $< 10\%$ ). This implies that there is no significant decrease in the number of DNA-bound YPet-DnaN before and after the replisome has encountered the Tus-*TerC* region. This suggests that the replisome, as assessed from the YPet-DnaN signal, remains DNA-bound after encountering a Tus-*Ter* roadblock and is not directly disassembled.

Since the sliding clamp has been shown to accumulate on the DNA (47), and is thus not necessarily always bound in the immediate vicinity of the replication fork, we also studied the fate of a different replisome component that is active at the fork: the  $\epsilon$ -subunit of DNA polymerase III (DnaQ-YPet). We observed no obvious difference in the number of DnaQ-YPet foci prior and post reaching the Tus-*TerC* roadblock, and only a minor difference ( $< 6\%$ ) in the number of DNA-bound DnaQ-YPet, as assessed by comparing the foci intensities before and after the replisome has encountered the Tus-*TerC* region (**Supplementary Figure S3**). These observations, together with the YPet-DnaN measurements, are in-line with the replisome remaining bound (or quickly restarting) after encountering the Tus-*Ter* block.





± SD. (right) The percentage of cells that have a single focus (light green line), double foci (dark green line) and no foci (black line) as a function of replication time. It is evident that the percentage of cells having a single focus or double foci is roughly equally distributed. **(B)** The integrated intensity when only a single YPet-DnaN focus (light green) is visible ( $I_{\text{single}}=1.28 \cdot 10^5 \pm 6.11 \cdot 10^4$  counts), ( $n_{\text{foci}} = 2123$ ), is essential the same as the sum of two individual YPet-DnaN foci (dark green) ( $I_{\text{double}}=1.41 \cdot 10^5 \pm 6.72 \cdot 10^4$  counts), ( $n_{\text{foci}} = 1647$ ). The difference between the two means of the individual distributions is < 10%. This intensity similarity implies that the same number of YPet-DnaN molecules are DNA-bound when a single focus or two foci are detected. **(C)** The intensity distributions of DNA-bound YPet-DnaN prior (grey) and post (red) spatial doubling of the *oriI*-mCherry focus.  $I_{\text{prior,oriI-mCherry}}=1.43 \cdot 10^5 \pm 6.58 \cdot 10^4$  counts ( $n_{\text{foci}} = 1304$ ),  $I_{\text{prior,oriI-mCherry}}=1.31 \cdot 10^5 \pm 6.28 \cdot 10^4$  counts (mean ± SD,  $n_{\text{foci}} = 3157$ ). The difference between the means of the two distributions is < 9%. The number of DNA-bound YPet-DnaN is thus practically unchanged after *oriI*-mCherry has been replicated, i.e. after the CW replisome has encountered the *Tus-TerC* roadblock. This unchanged intensity value is indicative of replisomes not being disassembled after encountering a *Tus-Ter* roadblock but rather remaining DNA-bound. **(D)** The distribution of the replication times (purple) and division time (brown) of individual *oriZ-dnaN* cells. The average replication time is  $t_{\text{rep,oriZ-dnaN}}=81 \pm 15$  min and the average division time is  $t_{\text{div,oriZ-dnaN}}=97 \pm 21$  min (Error is ± SD).

#### 4.2.3. A replisome is impeded by a *Tus-Ter* complex, but not halted indefinitely

We demonstrate that the replisome does not only remain DNA bound but is capable of progressing past the *Tus-Ter* block and continues replication. While the replisome does not disassemble following the encounter with *Tus-TerC*, it could potentially stall there without further replication; in other words, it is plausible that the CCW replisome synthesizes the remainder of the chromosome while the CW replisome is halted. To test whether the CW replisome can continue replication after having reached the *Tus-TerC* block, we compare the replication and division times of *oriZ-dnaN* cells to that of *oriC-dnaN* cells (**Table 1**). The *oriC-dnaN* cells have a division time of  $t_{\text{div,oriC-dnaN}} = 85 \pm 15$  min, and *oriZ-dnaN* cells have a  $t_{\text{div,oriZ-dnaN}} = 97 \pm 21$  min (Errors are ± SD; **Figure 3D**). These relatively long division times results from to the minimal growth medium (**Supplementary Section 1.3**) used in the experiments. Thus, the division time for *oriZ-dnaN* cells compared to that of *oriC-dnaN* cells increases by ~13%. The *oriC-dnaN* cells have a replication time of  $t_{\text{rep,oriC-dnaN}} = 70 \pm 7$  min (The errors are ± SD; **Supplementary Figure S4A-B**). The measured replication time

**Figure 3. The replisome is slowed down at a *Tus-Ter* roadblock, but not halted indefinitely.** **(A)** Average replisome behaviour during replication in the *oriZ-dnaN*-strain. (left) An average time-resolved trace of the two replisomes (YPet-DnaN) from individual cells during complete replication cycles ( $n_{\text{cells}} = 128$ ). Here we plot the single DnaN-YPet focus (filled circles) and double DnaN-YPet foci (open circles). The size of an individual element at each time point is representative of the percentage of cells having that particular distribution of foci. The traces have been aligned with respect to initiation and termination and binned. The transparent red rectangle indicates the time when *oriI*-mCherry focus spatially doubles, as determined from the distribution in **Figure 2C**. The width of the rectangle is

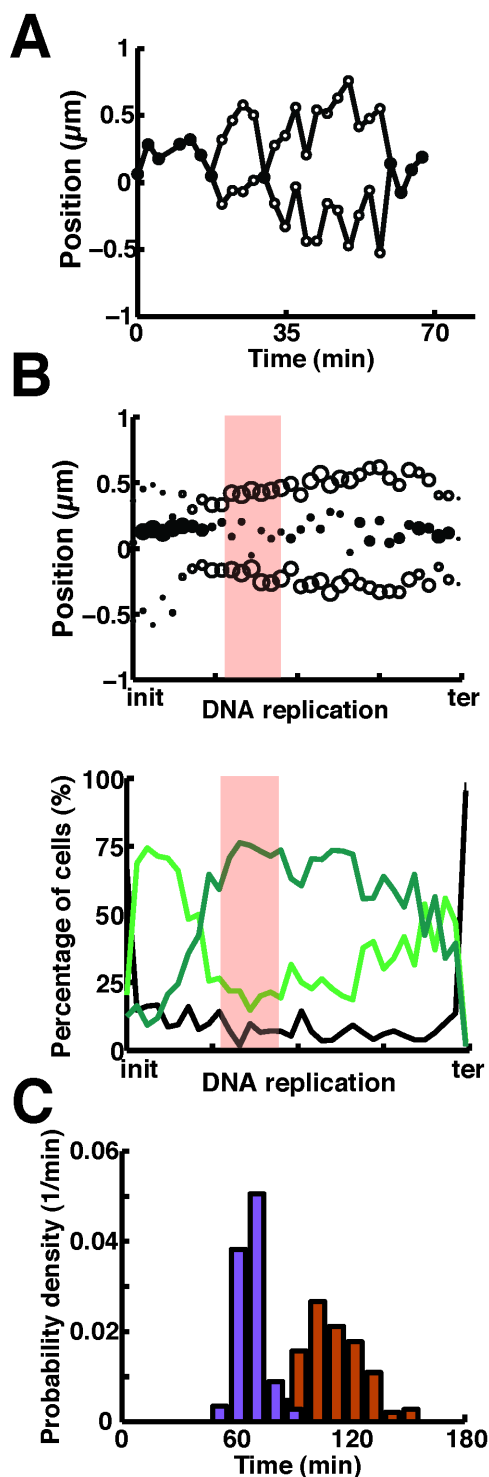
in *oriZ-dnaN* cells is  $t_{\text{rep,oriZ-dnaN}} = 81 \pm 15$  min (mean  $\pm$  SD; Figure 3D). If the CW replisome was halted indefinitely, one would expect the CCW replisome to copy  $\sim 1/4$  more of the chromosome. This should cause the replication time in *oriZ-dnaN* cells to increase by a factor of 1.5, thus to  $\sim 105$  min, a value that notably exceeds the observed increase of  $\sim 1.15\times$ . Even though we cannot fully exclude from our data that the CCW replisome might speed up during replication due to an increased availability of deoxyribonucleotide triphosphates (dNTPs) (48) by the stalling of the CW fork, we deem this to not be the dominant factor, since it has recently been shown that the CW replisome can pass a Tus-*Ter* barrier (39). This conclusion is supported by the observation that the number of DNA-bound DnaQ-YPet molecules appears unchanged when the replisome encounters the Tus-*Ter* barrier (**Supplementary Figure S3**). Thus, the marginal increase in the replication time suggests that the CW fork is, at least in a large fraction of the cases, capable of passing the Tus-*TerC* block. Given the wide distribution of replication times in *oriZ-dnaN* cells (**Figure 3D**), it is possible that there was variation in the pausing times of individual replisomes upon their encounter with the Tus-*Ter* barriers. Bridging of such a barrier could well be a stochastic process as opposed to a uniformly regulated one.

#### 4.2.4. Replisome progression is influenced by the absence of Tus in *oriZ-dnaN* cells

To verify that the longer replication time observed in *oriZ-dnaN* cells can be attributed to hampering of the CW replisome by the Tus-*Ter* complex, we investigate the replisome dynamics and replication time in *oriZ-dnaN* cells that do not express Tus (*oriZ-dnaN: $\Delta$ Tus*) (49) (**Supplementary Section 1.1.5**). This strain is incapable of forming the Tus-*Ter* complex, and hence the replisome should be able to progress unimpeded beyond the innermost terminus region.

Double foci are more frequently observed during elongation in *oriZ-dnaN: $\Delta$ Tus* cells compared to *oriZ-dnaN* cells. A sample trace (Figure 4A) and its corresponding temporal montage (**Supplementary Figure S1C-D**) illustrate this effect that foci appear to overlap less frequently in *oriZ-dnaN: $\Delta$ Tus* cells compared to *oriZ-dnaN* cells. The average time-resolved trace (**Figure 4B, top**) clearly shows that the fraction of cells displaying a single focus during elongation has decreased by a factor of  $\sim 2$  (**Figure 4B, bottom**). This is most likely due to the influence of the absence of the Tus-*Ter* roadblock on the replication-dependent chromosome dynamics and organization (41, 42).

*OriZ-dnaN: $\Delta$ Tus* cells exhibit very similar replication times to *oriC-dnaN* cells. The replication time for *oriZ-dnaN: $\Delta$ Tus* cells is  $t_{\text{rep,oriZ-dnaN: $\Delta$ Tus}} = 68 \pm 7$  min (**Figure 4C**), which is consistent with the above argument that the longer replication time in *oriZ-dnaN* cells compared to *oriC-dnaN* cells is due to the Tus-*Ter* block. The replication time of *oriZ-dnaN: $\Delta$ Tus* cells is in close similarity with the replication time for *oriC-dnaN* cells ( $t_{\text{rep,oriC-dnaN}} = 70 \pm 7$  min) (**Supplementary Figure S4B**). While the replication time for *oriZ-dnaN: $\Delta$ Tus* is decreased compared to that of *oriZ-dnaN* cells where we found  $t_{\text{rep,oriZ-dnaN: $\Delta$ Tus}} = 81$  min, the division time is slightly longer  $t_{\text{div,oriZ-dnaN: $\Delta$ Tus}} = 109 \pm 18$  min (**Figure 4C**) compared to  $t_{\text{div,oriZ-dnaN}} = 97$  min.



**Figure 4. Removing the Tus-Ter roadblock directly affects replication in *oriZ-dnaN:ΔTus* cells.** (A) A representative complete single replication cycle in a single *oriZ-dnaN:ΔTus* cell. A single YPet-DnaN focus is indicated with a solid black dot and two individual foci are indicated as circles. The lines between data points are included to aid the reader. (B) (top) An average time-resolved trace of the two replisomes (YPet-DnaN) from individual cells during complete replication cycles ( $n_{\text{cells}} = 152$ ). The single DnaN-YPet focus is represented as filled circles and double DnaN-YPet foci with open circles. The size of an individual element at each time point is again representative of the percentage of cells having that particular foci distribution. The traces have been aligned with respect to initiation and termination and binned. (bottom) The percentage of cells that have a single focus (light green line), double foci (dark green line) and no foci (black line) as a function of replication time. It is evident that the majority of the cells have two foci for the largest part of the replication process. The transparent red rectangle indicates the time when *oriI*-mCherry replicates, as determined from the distribution in **Figure 2C**. The width of the rectangle is  $\pm$  SD. (C) The distribution of the replication time (purple) and division time (brown) of individual *oriZ-dnaN:ΔTus* cells. The average replication time is  $t_{\text{rep},\text{oriZ-dnaN:}\Delta\text{Tus}} = 68 \pm 7$  min and the average division time is  $t_{\text{div},\text{oriZ-dnaN:}\Delta\text{Tus}} = 109 \pm 18$  min (Error is  $\pm$  SD).

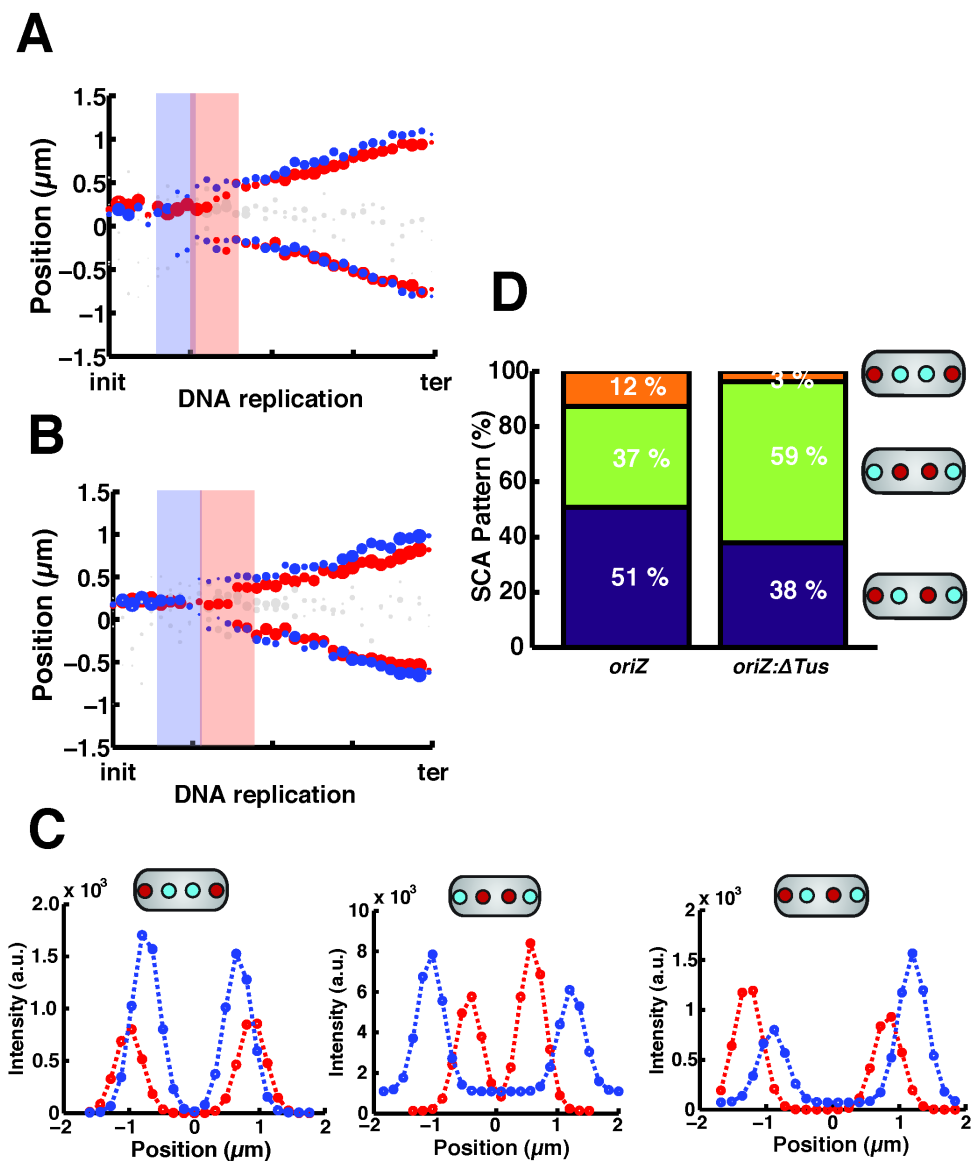
To gain further insight into the effect on the replication and division times in the absence of Tus, we investigated these times in *oriC-dnaN* cells that did not express Tus (*oriC-dnaN:ΔTus*) (**Supplementary Section 1.1.3; Figure S4**). We observed a slight increase in both the division time ( $t_{\text{div},\text{oriC-dnaN:}\Delta\text{Tus}} = 96 \pm 19$  min) and replication time ( $t_{\text{rep},\text{oriC-dnaN:}\Delta\text{Tus}} = 80 \pm 11$  min) compared to *oriC-dnaN* cells where we found  $t_{\text{div},\text{oriC-dnaN}} = 85$  min and  $t_{\text{rep},\text{oriC-dnaN}} = 70$  min

(**Supplementary Figure S4C**). This increase might be due to the more collision-prone fork fusion in the absence of Tus. No visible difference was observed in the replisome dynamics in terms of single or double foci as function of replication time (**Supplementary Figure S4D-E**). The division- and replication times of the four different strains have been summarized in **Table 1**.

#### 4.2.5. The presence and absence of Tus influences the sister chromosome alignment pattern in *oriZ-dnaN* cells

To gain more insight into the discernibility of the individual replisomes during replication we investigate the chromosome organization in *oriZ-dnaN* and *oriZ-dnaN:ΔTus* cells. Utilizing the labeled chromosomal loci on the left (*ori1*-mCherry) and right (*R2*-mCerulean) replichores, we visualize their movement during replication (**Figure 5A-B**). It is evident from the time-resolved traces that the *ori1*-mCherry and *R2*-mCerulean loci segregate and move to opposite cell halves in both cases. The times when the *ori1*-mCherry focus and *R2*-mCerulean focus spatially double are indicated. Since we did not observe a significant difference in the spatial doubling times of the respective loci in the two strains, we grouped the *oriZ-dnaN* and *oriZ-dnaN:ΔTus* data for better statistics. As stated previously, the time measured for the spatial doubling of *ori1* foci ( $t_{\text{ori1}}=26\pm 6$  min) is in excellent agreement with the expected value. In the case of *R2*, one would expect this region to be replicated  $\sim 1$  min after initiation given its position and the average speed of a replisome, but we measure  $t_{R2} = 17\pm 5$  min (Error is  $\pm$  SD,  $n=167$ ) (**Supplementary Figure S2B**), in agreement with previous measurements (38). However, this discrepancy between the measured and expected time when *R2* separates, can be attributed to the region of the chromosome that remains juxtaposed for  $\sim 10$  min and not yet segregated fully (4, 50, 51). The pattern at the end of replication between the two strains appears to be different (**Figure 5A-B**).

The sister chromosome alignment (SCA) patterns formed illustrate a change in chromosome organization in the absence of Tus. Since the movement or stagnation of the replication fork influences chromosome segregation organization (42, 52), we investigated the resulting arrangement of the left (L) and right (R) replichore loci after termination in both the *oriZ-dnaN* and *oriZ-dnaN:ΔTus* cells (**Figure 5C**). The three segregation patterns observed are LRRL (**Figure 5C, left**), RLLR (**Figure 5C, centre**), and LRLR (**Figure 5C, right**). In a minority of the cells ( $< 9\%$ ), the loci of the LR replichore arms spatially overlap (**Supplementary Figure S5A**), preventing their proper classification. Such cells are discarded from further analysis. The small number of cells displaying a RLRL pattern (**Supplementary Figure S5B**) are grouped with the LRLR (conform convention, since it is a symmetric configuration (53)). The majority of the *oriZ-dnaN* cells displayed a LRLR pattern, consistent with what has been measured previously (38), while the predominant pattern in *oriZ-dnaN:ΔTus* cells is RLLR. There is thus a switch in the predominant SCA pattern in the absence of Tus. This predominate RLLR pattern is already observed half way through replication (**Supplementary Figure S5C**). The fraction of *oriZ-dnaN:ΔTus* cells displaying RLLR is larger when determined half way through replication compared to at termination, while for *oriZ-dnaN* cells it is essentially unaffected. Through some mechanism the cell is able to revert back, but not fully, to the ‘normal’ LRLR pattern.



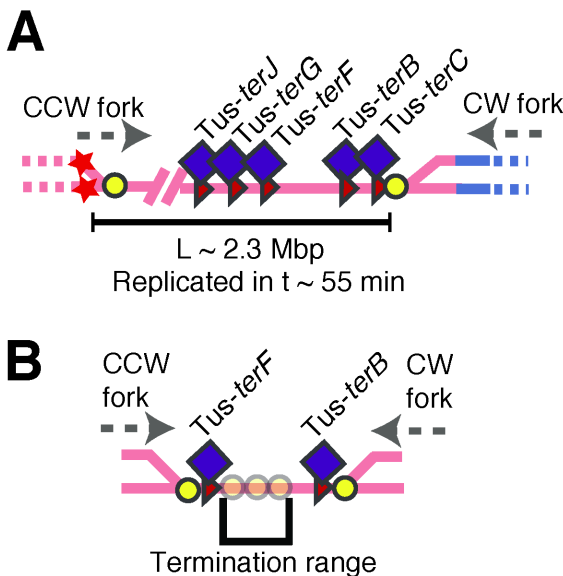
**Figure 5.** Sister chromosome alignment patterns are different for *oriZ-dnaN* and *oriZ-dnaN:ΔTus* cells. **(A-B)** Average time-resolved traces of the *ori1*-mCherry (Red) and *R2*-mCerulean (blue) loci. The blue and red rectangles indicate the time when the respective loci replicate as determined from the distributions in **Figure 2C** and **Supplementary Figure S2**. The shaded white dots are when only a single focus was detected, which was in the minority of cases. As with the previous figures, the size of the dots is representative of the percentage of cells having that particular foci distribution. The width of the shaded region is the SD. **(A)** *oriZ-dnaN*-strain ( $n_{\text{cells}} = 71$ ). **(B)** *oriZ-dnaN:ΔTus*-strain ( $n_{\text{cells}} = 79$ ). **(C)** Example sister chromosome alignment (SCA) patterns of the left (L) and right (R) replichores after replication. Here we plot the fitted spatial intensity



distributions. (left) LRRL, (centre) RLLR and (right) RLRL. (Insets) Schematic depiction of the different chromosomal loci patterns. **(D)** The distributions of the different SCA patterns for the *oriZ-dnaN*-, and *oriZ-dnaN:ΔTus*-cells after replication. The predominant pattern for *oriZ-dnaN* (51%) is the LRLR organization, while for *oriZ-dnaN:ΔTus* cells this is the RLLR organization (59%).

#### 4.2.6. Fork fusion in *oriZ-dnaN* cells occurs most likely close to *TerF*

The time point at which the CW replisome reaches *TerC* and the total replication time provide us with insight into the position where termination might occur in *oriZ-dnaN* cells. Since we demonstrated that the CW replisome is not disassembled at *Tus-TerC*, and that it most likely continues synthesis (39, 54), albeit being slowed down, it is possible to determine the most likely termination position. Calculating from the time when the CW replisome encounters *TerC* (~26 min) there remain ~55 min for the two replisomes to copy the remainder (~2300 kbp) of the genome (**Figure 6A**). The CCW replisome is not hampered as it synthesizes in the natural direction and would thus progress ~1800 kbp, assuming a constant velocity of  $v \sim 550$  bp/s. This implies that the CW replisome is effectively delayed by the *Tus-Ter* barriers for ~30 min, and replicates ~500 kbp on average. From this it follows that replication termination, i.e. fork fusion, occurs close to *TerF*. However, the distribution of replication times (**Figure 3D**) also indicates that there are events where termination might occur between *TerB-TerF* (**Figure 6B**). To illustrate this more quantitatively, we plot the distribution of the time differences between the averaged temporal occurrence of *oriI* spatial doubling and individual replication termination events (**Figure 6C**). The proportion of cells that terminates faster than ~55 min is more likely indicative of *TerB-TerF* termination events, while the proportion of cells that terminates after ~55 min is more likely associated with *TerB/C* termination events. These proportions are consistent with *TerB* and *TerC* being strong roadblocks that reduce the rate of CW replisome progression.



**Figure 6. The CW replisome crosses *TerB* and *TerC*.** **(A)** Schematic illustration of the remaining chromosome that needs to be replicated after the CW replisome has encountered the *Tus-TerC* roadblock. **(B)** A graphic representation of the position where termination occurs in *oriZ-dnaN* cells. Here the transparent yellow circles are representative of the uncertainty on the exact position of fork fusion, i.e. termination, somewhere in between *TerF* and *TerB*.

Since the CW replisome moves more slowly due to the *Tus-Ter* barriers, one might have expected the sliding clamps to unload, resulting in a reduction of their number from the steady-state

equilibrium (47). This is not observed, implying the possibility that the dynamics of the sliding clamp are altered in the terminus region in this strain, possibly due to the chromosome compaction of the highly prevalent MatP/matS site-specific system that organizes the terminus region (55).

### 4.3. Discussion

To ensure the accurate inheritance of genetic material one needs faithful completion of replication and chromosome segregation. Our work provides insight into the replisome's robustness *in vivo* and into its ability to circumvent the natural Tus-*Ter* replication barriers. We characterize the effect of the Tus-*Ter* complex on the stability and progression of the replisome when it encounters the barrier from the non-permissive side. Our data suggest that the CW replisome (as assessed by studying YPet-DnaN and DnaQ-Ypet foci) does not disassemble after having encountered the Tus-*TerC* complex from the non-permissive direction, but rather remains stably bound to the DNA (or is rapidly reloaded). This result is consistent with what has been observed *in vitro* during a head-on collision of the *E. coli* replisome and an RNA polymerase (RNAP) (36). We observed that the progression of the CW fork is impeded, but that it is capable of overcoming this barrier, as deduced from the unaffected numbers of DNA-bound YPet-DnaN and DnaQ-YPet before and after encountering the roadblock (**Figure 3C**, **Supplementary Figure S3**), together with the fact that the overall replication time only increased by 15% (**Figure 3D**). Our results are in agreement with a recent study where it was shown, in their strain with only an ectopic origin of replication ( $\Delta oriC$  *oriZ*), that termination frequently occurs at *TerB*, thus the CW replisome crosses the Tus-*TerC* barrier (39).

We demonstrate that the presence and absence of Tus affect the replication- and cell generation times in both *oriC-dnaN* and *oriZ-dnaN* cells. For both *oriZ-dnaN*: $\Delta$ Tus and *oriC-dnaN*: $\Delta$ Tus, an increase in the mean cell doubling time was observed when compared to *oriC-dnaN* and *oriZ-dnaN* cells respectively (~11% and ~12%, respectively). In both cases, this increase could result from processing events required to complete replication when the converging forks fuse at sites other than Tus-bound *Ter* sites. This hypothesis is supported by our observation that there is an increase in the D-period for these strains (**Table 1**).

**Table 1**

<i>E. coli</i> strain	Division time (min)	Replication time (min)	D-period (min)
<i>oriC-dnaN</i>	85±15	70±7	23±9
<i>oriC-dnaN</i> : $\Delta$ Tus	96±19	80±11	30±13
<i>oriZ-dnaN</i>	97±21	81±15	19±13
<i>oriZ-dnaN</i> : $\Delta$ Tus	109±18	68±7	31±10

The division- and replication times together with the respective D-periods for the different *E. coli* strains investigated. The numbers specified indicate the mean  $\pm$  SD. Here replication time is defined as the time interval between YPet-DnaN focus appearance and disappearance. D-period is defined as the time interval between YPet-DnaN focus disappearance and cell division. The true replication time and D-period are likely slightly shorter than the reported values since termination occurs prior to a focus completely disappearing.

We note that the replication time for *oriZ-dnaN: $\Delta$ Tus* was in very good agreement with *oriC-dnaN* cells, strengthening the view that the Tus-*Ter* complexes caused the measured replication fork delay in *oriZ-dnaN* cells. We observed a slightly longer replication time for *oriC-dnaN: $\Delta$ Tus* cells compared to *oriC-dnaN* cells (~14% increase). This observation of longer replication times in *oriC-dnaN* is in accordance with the RFT model (21, 34) that fork fusion and completion of replication are most effective at Tus-bound *Ter* sites (56), rather than at the *dif* site as has been proposed by bioinformatics analysis (57). The *dif* locus, located between *TerA* and *TerC*, is where the XerCD recombinase decatenates newly replicated sister chromosomes and resolves chromosome dimers of nascent daughter chromosomes to allow successful chromosome segregation (58). If fork fusion occurred at the *dif* site most frequently, the absence of Tus should have had a negligible effect on the replication and cell division. We propose that Tus-*Ter* sites around the *dif* site facilitate the localization of the proteins necessary for proper dimer resolution by ensuring that only one of the two replisomes replicates the *dif* sequence, as postulated by Duggin *et al* (21). Possibly, fork fusion and attendant collision processes in the direct vicinity of the *dif* site might influence for example decatenation or dimer resolution, and hence increase the overall duration of replication and cell division.

We observed a switch in the predominant SCA patterns after replication when the *tus* gene was deleted from *oriZ-dnaN* cells. The predominant SCA pattern in *oriC-dnaN* and *oriZ-dnaN* cells has been shown to be LRLR (38). We indeed observed this pattern for *oriZ-dnaN* cells, while interestingly the predominant pattern switched to RLLR in the absence of Tus (*oriZ-dnaN: $\Delta$ Tus* cells), an organization pattern that is associated with fork stalling (42). One would rather have anticipated that the proportion of cells displaying LRLR pattern should increase when Tus is absent from the cells; since no replication fork stalling is occurring due to Tus-*Ter* complexes. Our different observation may imply that the collision of the CW replisome with different proteins after passing the *dif* site plays a role in the final SCA in *oriZ-dnaN* cells.

A future topic of study would be to investigate the mechanistic details by which the replication fork can progress beyond an RFB caused by the Tus-*Ter* complex. The  $\beta_2$  sliding clamps have been shown to facilitate in bypassing of replication barriers (59). This might be a strategy by which the replisome bridges such an RFB when encountering the Tus-*Ter* complex from the permissive side. SSB, for example, might facilitate repair at these collisions (60) and in that way assist the replisome. It might be other non-replicative helicases that dispose of these roadblocks (61). For example, the *E. coli* helicase II (UvrD) is a 3' to 5' helicase that has been shown to be able to remove Tus *in vitro* (62). It is plausible that it could remove Tus from a *Ter* site *in vivo* as well. The FtsK translocase involved in chromosome segregation (63, 64) is another potential candidate for Tus removal, though a recent single-molecule *in vitro* study has shown that FtsK changes direction when encountering Tus and does not seem to dislodge it from the DNA (65).

Independently of which proteins assist the replisome in encountering a roadblock, our work demonstrates that the replisome is robust in continuing replication despite this hurdle.

## 4.4. Materials and Methods

The details regarding the experimental procedures can be found in the Supplementary Information. Here we briefly highlight the different strains and techniques used in this study.

### 4.4.1. Strains, strain construction and growth for microscopy

All endogenous chromosomal integration of the *lacO* and *tetO* arrays, and fusions of either YPet-DnaN, LacI-mCherry, and TetR-mCerulean were constructed in previous work and described in detail (38). All chromosomal deletions of the *tus* gene in the different strains were realized by  $\lambda$ -Red recombination and P1 phage transduction.

Cells used for microscopy were grown in M9-Gly supplemented with the necessary nutrients until an OD  $\sim$  0.2 was reached. The cells were immobilized for imaging using a custom microfluidic system made from polydimethylsiloxane (PDMS) (40). Fresh growth medium was continuously injected automatically into the device during an experiment at 0.5 mL/hr via a syringe pump.

### 4.4.2. Microscopy and image analysis

All microscopy experiments were conducted on a commercial Nikon Ti microscope with custom laser excitation. YPet was imaged using a 515 nm (Cobolt Fandango) laser, mCherry was imaged using a 561 nm (Cobolt Jive) laser and mCerulean were imaged using a 457 nm (Cobolt Twist) laser. The type of extended time-lapse microscopy that we conduct using microfluidics allows us to quantitatively compare intensities since the rate of photobleaching is constant, and hence a steady-state between bleaching and protein products is reached rapidly into the measurement (47). Image acquisition was performed with an EMCCD (Andor) using commercial Nikon NIS elements software. The whole microscope body, including sample, was kept at  $\sim$ 37 °C using a temperature controller (Okolabs). Image analysis was performed using ImageJ and custom written Matlab code.

## 4.5. Acknowledgments

We thank Christian Lesterlin for support with the *oriZ*-strain, Theo van Laar for stimulating discussions. We thank Jelle van der Does, Dimitri de Roos, and Jaap Beekman for their contribution towards instrumentation and infrastructure. This work was supported by a Wellcome Trust Program Grant WT083469 to D.J.S., and the European Community's Seventh Framework Programme FP7/2007-2013 under grant agreement n° 241548 (MitoSys) and a Vici grant by Netherlands Organisation for Scientific Research (NWO) both to N.H.D.

## 4.6. Contributions

M.C.M. and N.H.D. designed the research. M.C.M., S.T.K., and N.H.D. designed the experiments. S.T.K. undertook the experiments. R.d.L and V.L assisted during experiments

and contributed to discussions. S.T.K. constructed strains. M.C.M. and J.W.J.K. wrote software to analyze the microscopy data. M.C.M. and J.W.J.K. analyzed the data. D.J.S. provided strains and contributed to the discussion of the work. M.C.M., S.T.K., R.d.L and N.H.D wrote the paper.

#### 4.7. References

1. A. Johnson, M. O'Donnell, CELLULAR DNA REPLICASES: Components and Dynamics at the Replication Fork. *Annual Review of Biochemistry* **74**, 283-315 (2005).
2. C. S. McHenry, DNA replicases from a bacterial perspective. *Annual review of biochemistry* **80**, 403-436 (2011).
3. K. E. Duderstadt, R. Reyes-Lamothe, A. M. van Oijen, D. J. Sherratt, Replication-Fork Dynamics. *Cold Spring Harbor perspectives in biology* **6**, (2013).
4. R. Reyes-Lamothe, C. Possoz, O. Danilova, D. J. Sherratt, Independent Positioning and Action of *Escherichia coli* Replisomes in Live Cells. *Cell* **133**, 90-102 (2008).
5. S. Jun, A. Wright, Entropy as the driver of chromosome segregation. *Nature Reviews Microbiology* **8**, 600-607 (2010).
6. N. J. Kuwada, K. C. Cheveralls, B. Traxler, P. A. Wiggins, Mapping the driving forces of chromosome structure and segregation in *Escherichia coli*. *Nucleic Acids Research* **41**, 7370-7377 (2013).
7. B. Di Ventura *et al.*, Chromosome segregation by the *Escherichia coli* Min system. *Molecular systems biology* **9**, 686 (2013).
8. J. K. Fisher *et al.*, Four-dimensional imaging of *E. coli* nucleoid organization and dynamics in living cells. *Cell* **153**, 882-895 (2013).
9. B. Youngren, H. J. Nielsen, S. Jun, S. Austin, The multifork *Escherichia coli* chromosome is a self-duplicating and self-segregating thermodynamic ring polymer. *Genes & Development* **28**, 71-84 (2014).
10. H. Merrikkh, Y. Zhang, A. D. Grossman, J. D. Wang, Replication-transcription conflicts in bacteria. *Nature Reviews Microbiology* **10**, 449-458 (2012).
11. B. Michel, G. Grompone, M.-J. Florès, V. Bidnenko, Multiple pathways process stalled replication forks. *Proceedings Of The National Academy Of Sciences Of The United States Of America* **101**, 12783-12788 (2004).
12. M. K. Gupta *et al.*, Protein-DNA complexes are the primary sources of replication fork pausing in *Escherichia coli*. *Proceedings Of The National Academy Of Sciences Of The United States Of America* **110**, 7252-7257 (2013).

13. T. M. Hill, M. L. Tecklenburg, A. J. Pelletier, P. L. Kuempel, tus, the trans-acting gene required for termination of DNA replication in *Escherichia coli*, encodes a DNA-binding protein. *Proceedings Of The National Academy Of Sciences Of The United States Of America* **86**, 1593-1597 (1989).
14. M. Hidaka, T. Kobayashi, S. Takenaka, H. Takeya, T. Horiuchi, Purification of a DNA replication terminus (ter) site-binding protein in *Escherichia coli* and identification of the structural gene. *The Journal of biological chemistry* **264**, 21031-21037 (1989).
15. T. M. Hill, Arrest of Bacterial DNA Replication. *Annual review of microbiology* **46**, 603-633 (1992).
16. F. F. Coskun-Ari, A. Skokotas, G. R. Moe, T. M. Hill, Biophysical characteristics of Tus, the replication arrest protein of *Escherichia coli*. *The Journal of biological chemistry* **269**, 4027-4034 (1994).
17. C. J. Rudolph, A. L. Upton, G. S. Briggs, R. G. Lloyd, Is RecG a general guardian of the bacterial genome? *DNA repair* **9**, 210-223 (2010).
18. T. M. Hill, K. J. Marians, *Escherichia coli* Tus protein acts to arrest the progression of DNA replication forks *in vitro*. *Proceedings Of The National Academy Of Sciences Of The United States Of America* **87**, 2481-2485 (1990).
19. V. Bidnenko, S. D. Ehrlich, B. Michel, Replication fork collapse at replication terminator sequences. *The EMBO Journal* **21**, 3898-3907 (2002).
20. C. Neylon, A. V. Kralicek, T. M. HILL, N. E. Dixon, Replication termination in *Escherichia coli*: structure and antihelicase activity of the Tus-Ter complex. *Microbiology And Molecular Biology Reviews* **69**, 501-526 (2005).
21. I. G. Duggin, R. G. Wake, S. D. Bell, T. M. Hill, The replication fork trap and termination of chromosome replication. *Molecular Microbiology* **70**, 1323-1333 (2008).
22. D. L. Kaplan, D. Bastia, Mechanisms of polar arrest of a replication fork. *Molecular Microbiology* **72**, 279-285 (2009).
23. M. D. Mulcair *et al.*, A Molecular Mousetrap Determines Polarity of Termination of DNA Replication in *E. coli*. *Cell* **125**, 1309-1319 (2006).
24. D. Bastia, S. Zaman, Mechanism and physiological significance of programmed replication termination. *Seminars in cell & developmental biology* **30**, 165-173 (2014).
25. B. A. Berghuis *et al.*, Strand separation suffices to establish a long-lived, foolproof DNA-protein lock at the Tus-Ter replication fork barrier. *Nature Chemical Biology* **11**, 579-585 (2015).

26. P. A. Gottlieb *et al.*, Equilibrium, kinetic, and footprinting studies of the Tus-Ter protein-DNA interaction. *The Journal of biological chemistry* **267**, 7434-7443 (1992).
27. C. Neylon *et al.*, Interaction of the *Escherichia coli* replication terminator protein (Tus) with DNA: a model derived from DNA-binding studies of mutant proteins by surface plasmon resonance. *Biochemistry* **39**, 11989-11999 (2000).
28. P. J. Lewis, G. B. Ralston, R. I. Christopherson, R. G. Wake, Identification of the replication terminator protein binding sites in the terminus region of the *Bacillus subtilis* chromosome and stoichiometry of the binding. *Journal of Molecular Biology* **214**, 73-84 (1990).
29. D. E. Bussiere, D. Bastia, S. W. White, Crystal structure of the replication terminator protein from *B. subtilis* at 2.6 Å. *Cell* **80**, 651-660 (1995).
30. K. P. Lemon, I. Kurtser, A. D. Grossman, Effects of replication termination mutants on chromosome partitioning in *Bacillus subtilis*. *Proceedings Of The National Academy Of Sciences Of The United States Of America* **98**, 212-217 (2001).
31. N. A. Willis *et al.*, BRCA1 controls homologous recombination at Tus/Ter-stalled mammalian replication forks. *Nature* **510**, 556-559 (2014).
32. N. B. Larsen, E. Sass, C. Suski, H. W. Mankouri, I. D. Hickson, The *Escherichia coli* Tus-Ter replication fork barrier causes site-specific DNA replication perturbation in yeast. *Nature communications* **5**, 3574 (2014).
33. T. M. Hill, J. M. Henson, P. L. Kuempel, The terminus region of the *Escherichia coli* chromosome contains two separate loci that exhibit polar inhibition of replication. *PNAS*, 1-5 (1987).
34. I. G. Duggin, S. D. Bell, Termination structures in the *Escherichia coli* chromosome replication fork trap. *Journal of Molecular Biology* **387**, 532-539 (2009).
35. M. J. J. Moreau, P. M. Schaeffer, Differential Tus-Ter binding and lock formation: implications for DNA replication termination in *Escherichia coli*. *Molecular Biosystems* **8**, 2783-2791 (2012).
36. R. T. Pomerantz, M. O'Donnell, Direct restart of a replication fork stalled by a head-on RNA polymerase. *Science (New York, NY)* **327**, 590-592 (2010).
37. J. T. P. Yeeles, K. J. Mariani, The *Escherichia coli* Replisome Is Inherently DNA Damage Tolerant. *Science (New York, NY)* **334**, 235-238 (2011).
38. X. Wang, C. Lesterlin, R. Reyes-Lamothe, G. Ball, D. J. Sherratt, Replication and segregation of an *Escherichia coli* chromosome with two replication origins.

*Proceedings Of The National Academy Of Sciences Of The United States Of America* **108**, E243-250 (2011).

39. D. Ivanova *et al.*, Shaping the landscape of the *Escherichia coli* chromosome: replication-transcription encounters in cells with an ectopic replication origin. *Nucleic Acids Research* **43**, 7865-7877 (2015).
40. M. C. Moolman, Z. Huang, S. T. Krishnan, J. W. J. Kerssemakers, N. H. Dekker, Electron beam fabrication of a microfluidic device for studying submicron-scale bacteria. *Journal of Nanobiotechnology* **11**, 12-12 (2013).
41. E. Rocha, The replication-related organization of bacterial genomes. *Microbiology-Sgm* **150**, 1609-1627 (2004).
42. X. Liu, X. Wang, R. Reyes-Lamothe, D. J. Sherratt, Replication-directed sister chromosome alignment in *Escherichia coli*. *Molecular Microbiology* **75**, 1090-1097 (2010).
43. X. Wang, R. Reyes-Lamothe, D. J. Sherratt, Visualizing genetic loci and molecular machines in living bacteria. *Biochemical Society Transactions* **36**, 749-753 (2008).
44. C. Possoz, S. R. Filipe, I. Grainge, D. J. Sherratt, Tracking of controlled *Escherichia coli* replication fork stalling and restart at repressor-bound DNA *in vivo*. *The EMBO Journal* **25**, 2596-2604 (2006).
45. X. Wang, C. Possoz, D. J. Sherratt, Dancing around the divisome: asymmetric chromosome segregation in *Escherichia coli*. *Genes & Development* **19**, 2367-2377 (2005).
46. X. Wang, X. Liu, C. Possoz, D. J. Sherratt, The two *Escherichia coli* chromosome arms locate to separate cell halves. *Genes & Development* **20**, 1727-1731 (2006).
47. M. C. Moolman *et al.*, Slow unloading leads to DNA-bound  $\beta_2$ -sliding clamp accumulation in live *Escherichia coli* cells. *Nature communications* **5**, 5820 (2014).
48. Morigen, I. Odsbu, K. Skarstad, Growth rate dependent numbers of SeqA structures organize the multiple replication forks in rapidly growing *Escherichia coli*. *Genes to cells : devoted to molecular & cellular mechanisms* **14**, 643-657 (2009).
49. B. Roecklein, A. Pelletier, P. Kuempel, The *tus* gene of *Escherichia coli*: autoregulation, analysis of flanking sequences and identification of a complementary system in *Salmonella typhimurium*. *Research in microbiology* **142**, 169-175 (1991).



50. X. Wang, R. Reyes-Lamothe, D. J. Sherratt, Modulation of Escherichia coli sister chromosome cohesion by topoisomerase IV. *Genes & Development* **22**, 2426-2433 (2008).
51. M. C. Joshi *et al.*, Escherichia coli sister chromosome separation includes an abrupt global transition with concomitant release of late-splitting intersister snaps. *Proceedings Of The National Academy Of Sciences Of The United States Of America* **108**, 2765-2770 (2011).
52. H. J. Nielsen, B. Youngren, F. G. Hansen, S. Austin, Dynamics of Escherichia coli chromosome segregation during multifork replication. *Journal of Bacteriology* **189**, 8660-8666 (2007).
53. X. Wang, P. Montero Llopis, D. Z. Rudner, Organization and segregation of bacterial chromosomes. *Nature Reviews Genetics* **14**, 191-203 (2013).
54. K. A. Mettrick, I. Grainge, Stability of blocked replication forks *in vivo*. *Nucleic Acids Research*, (2015).
55. R. Mercier, M. A. Petit, S. Schbath, S. Robin, M. El Karoui, The MatP/matS site-specific system organizes the terminus region of the E. coli chromosome into a macrodomain. *Cell* **135**, (2008).
56. C. J. Rudolph, A. L. Upton, A. Stockum, C. A. Nieduszynski, R. G. Lloyd, Avoiding chromosome pathology when replication forks collide. *Nature* **500**, 608-611 (2013).
57. H. Hendrickson, J. G. Lawrence, Mutational bias suggests that replication termination occurs near the dif site, not at Ter sites. *Molecular Microbiology* **64**, 42-56 (2007).
58. C. Lesterlin, F. X. Barre, F. Cornet, Genetic recombination and the cell cycle: what we have learned from chromosome dimers. *Molecular Microbiology* **54**, 1151-1160 (2004).
59. R. E. Georgescu, Yao, N. Y, M. O'Donnell, Single-molecule analysis of the Escherichia coli replisome and use of clamps to bypass replication barriers. *FEBS letters* **584**, 2596-2605 (2010).
60. R. D. Shereda, A. G. Kozlov, T. M. Lohman, SSB as an organizer/mobilizer of genome maintenance complexes. *Critical Reviews in Biochemistry and Molecular Biology*, (2008).
61. H. Boubakri, A. L. de Septenville, E. Viguera, B. Michel, The helicases DinG, Rep and UvrD cooperate to promote replication across transcription units *in vivo*. *The EMBO Journal* **29**, 145-157 (2010).

62. H. Hiasa, K. J. Marians, Differential inhibition of the DNA translocation and DNA unwinding activities of DNA helicases by the Escherichia coli Tus protein. *The Journal of biological chemistry* **267**, 11379-11385 (1992).
63. D. J. Sherratt, L. K. Arciszewska, E. Crozat, The Escherichia coli DNA translocase FtsK. *Biochemical Society \ldots*, (2010).
64. I. Grainge, Simple topology: FtsK-directed recombination at the dif site. *Biochemical Society Transactions* **41**, 595-600 (2013).
65. J. Y. Lee, I. J. Finkelstein, L. K. Arciszewska, D. J. Sherratt, E. C. Greene, Single-molecule imaging of FtsK translocation reveals mechanistic features of protein-protein collisions on DNA. *Molecular Cell* **54**, 832-843 (2014).
66. K. Kamada, T. Horiuchi, K. Ohsumi, N. Shimamoto, K. Morikawa, Structure of a replication-terminator protein complexed with DNA. *Nature* **383**, 598-603 (1996).

## 4.9. Supplementary Information

### 4.9.1. S1.1 Construction and characterization of fluorescent fusion strains

All strains are derived from the *E. coli* K12 AB1157 (1). The strains used for this study were constructed by  $\lambda$ -red recombination (2) or P1-transduction (3). A summary of the plasmids and strains used in this study are provided in **Supplementary Tables S2 and S3**.

#### S1.1.1 $\Delta$ Tus in WTAB1157

Utilizing  $\lambda$ -red recombination we created a  $\Delta$ Tus mutant in WTAB1157 by knocking out the open reading frame of the *tus* gene. The plasmid pKD13 (GenBank: AY048744.1) was used as a template plasmid for amplifying the FRT-flanked kanamycin resistance gene (*kanR*) sequence used during  $\lambda$ -red recombination. The primer sequences used were: Forward 5' – CCA CGA CTG TGC TAT AAA ATA AGT ATG TTG TAA CTA AAG TGG TTA ATA TTT GTA GGC TGG AGC TGC TTC G – 3'; Reverse 5' – GAC AGC TGG GTA CGG CCA GAA CAG ATG GTC GGC AGT ATG AAA GCC GGG CGA TTC CGG GGA TCC GTC GAC C – 3'. The DNA fragment was gel purified, and ~700 ng of the linear DNA was used during electroporation of AB1157 cells overexpressing the  $\lambda$ -red proteins from pKD46 (2). The correct insertion of the fragment into the chromosome of the resulting strain was assayed by PCR. The oligonucleotides used were 5' – GCG CAC GAT GGT CAA GTC AC – 3' and 5' – TAC GGC CAG AAC AGA TGG TC – 3'. The sequence of the deleted region in this strain was verified and confirmed by DNA sequencing.

#### S1.1.2 YPet-dnaN:ter-mCerulean (oriC-dnaN)

The YPet-dnaN:ter-mCerulean (*oriC*) strain was constructed previously as described in Moolman *et al* (4).

#### S1.1.3 YPet-dnaN:ter-mCerulean: $\Delta$ Tus (oriC-dnaN: $\Delta$ Tus)

The *oriC*: $\Delta$ Tus was constructed using P1 transduction. Firstly, the FRT-flanked *kanR* gene in the *oriC* strain, used during previous recombineering, was recombined out using the temperature sensitive Flippase (FLP) enzyme expressed from the pCP20 plasmid as described elsewhere (2). Subsequently, the FRT-flanked *kanR* gene from the  $\Delta$ Tus strain was transduced into the *oriC* strain. The presence of the  $\Delta$ Tus knock-out was verified using the oligonucleotides: 5' – GCG CAC GAT GGT CAA GTC AC – 3' and 5' – TAC GGC CAG AAC AGA TGG TC – 3'. The sequence of the deleted region was confirmed by DNA sequencing.

#### S1.1.4 YPet-DnaN:oriZ (oriZ-dnaN)

The *oriZ*-dnaN strain was constructed previously as described in (5).

#### S1.1.5 YPet-DnaN:oriZ: $\Delta$ Tus(oriZ-dnaN: $\Delta$ Tus)

The *oriZ*-dnaN: $\Delta$ Tus was constructed using P1 transduction. Firstly, the FRT-flanked *kanR* gene in the *oriZ* strain, used during previous recombineering, was recombined out using the

temperature the sensitive Flippase (FLP) enzyme expressed from the pCP20 plasmid as described elsewhere (2). Subsequently, the FRT-flanked *kanR* gene from the  $\Delta$ *Tus* strain was transduced by P1 transduction into the out-recombined *oriZ-dnaN* strain. The presence of the *Tus* knock-out was verified using the oligonucleotides: 5'– GCG CAC GAT GGT CAA GTC AC – 3' and 5' – TAC GGC CAG AAC AGA TGG TC – 3'. The correct deletion of the *tus* gene at the desired region was confirmed by DNA sequencing.

### **S1.1.6 *dnaQ*-YPet:*oriZ* (*oriZ*-*dnaQ*)**

The *oriZ-dnaQ* was constructed using P1 transduction. The *dnaQ-YPet* gene from the previously constructed *dnaQ-YPet* strain (6) was transduced by P1 transduction into the *oriZ* strain (5). For clarity, this is the strain without any replisome component labeled. The presence of the *dnaQ-YPet* knock-out was verified using the oligonucleotides: 5' – AAT GAC CGC TAT GAG CAC TG – 3' and 5' – TTG CCT CGA CCT TCG TCA AC – 3'.

## **4.9.2. S1.2. Construction of *OriZ-dnaN* strain with *ter* locus labeled**

Previously in this chapter, we used the doubling of *oriI* foci to identify the replication speed and indirectly determine the time when the CW replication fork encounters the non-permissive face of *Tus-TerC* roadblock earlier than CCW fork. In order to verify directly when it happens with more precision, we aimed to create a new *OriZ-dnaN* strain with only the *ter* locus (chromosomal region between *TerBC* sites, 50 kb clockwise from the *dif* site) labeled with mCerulean. We successfully created the *OriZ-dnaN-ter* strains both in the presence and absence of *tus* gene (*Tus* and  $\Delta$ *Tus*) as described in the sub-section below. When we analyzed the created *OriZ* strains, we found that the *ter* locus did not segregate until before the end of cell division similar to the AB1157 strain with the normal origin (*oriC-dnaN*). Hence, the precise time when the CW replication fork faces the *Tus-TerC* roadblock was not determined using this strain.

### **S1.2.1 *LacI*-mCherry ::*LeuB***

The strain containing only the *placZ-lacI-mcherry* gene together with the adjacent *frt*-flanked *CmR* gene present in the place of the *leuB* gene of the AB1157 chromosome was constructed previously as described in (7).

### **S1.2.2 YPet-*dnaN*:*ter*-mCerulean:*LacI*-mCherry (*oriC*-*dnaN*:*LacI*-mCherry)**

The *oriC-dnaN:LacI-mCherry* was constructed using P1 phage transduction by transducing the *placZ-lacI-mcherry* gene together with the adjacent *frt*-flanked *CmR* gene present in the place of the *leuB* gene of the AB1157 chromosome into the *oriC-dnaN* strain. The presence of the *lacI-mcherry* gene fusion was verified using fluorescence microscopy. This step was performed to label any other chromosomal loci for further studies by inserting the *lacO* array sequences at the respective region.

### **S1.2.3 *OriC*-*OriZ***

The strain with only the *oriZ* insert with FRT-flanked *kanR* gene (*OriC-OriZ*) was constructed previously as described in (5).

#### **S1.2.4 YPet-dnaN:ter-mCerulean:LacI-mCherry:oriZ (oriC-oriZ-dnaN-ter)**

The *oriC-oriZ-dnaN-ter* was constructed using P1 phage transduction by transducing the *oriZ* insert sequence from *OriC-OriZ* strain into the *oriC-dnaN:LacI-mCherry* strain. The presence of the *oriZ* insert was confirmed by PCR and DNA sequencing using the oligonucleotides: 5'-GTG CTC AAC AAA CTC TCC TT- 3' and 5'-TTA CAG CAG AGA AGG GAC GC- 3' along with the external primer for *kanR* gene 5'- ATG ATT GAA CAA GAT GGA TTG CAC GC - 3'.

#### **S1.2.5 frt-oriC-CmR-frt**

The strain with only the FRT-flanked *oriC* insert with *CmR* gene (*frt-oriC*) in the place of *oriC* was constructed previously as described in (5).

#### **S1.2.6 YPet-dnaN:ter-mCerulean:LacI-mCherry:oriZ:frt-oriC-CmR-frt (frt-oriC-oriZ-dnaN-ter)**

The *frt-oriC-oriZ-dnaN-ter* was constructed using P1 phage transduction. Firstly, the FRT-flanked *kanR* and *CmR* genes in the *oriC-oriZ-dnaN-ter* strain were recombined out using the pCP20 plasmid as described elsewhere (2). Subsequently, the FRT-flanked *oriC-CmR* insert from the *frt-oriC-CmR-frt* strain was transduced by P1 transduction into the out-recombined *oriC-oriZ-dnaN-ter* strain. The presence of the *oriC-CmR* was verified by PCR and DNA sequencing using the oligonucleotides: 5'-GTG CTC AAC AAA CTC TCC TT- 3' and 5'-TTA CAG CAG AGA AGG GAC GC- 3' along with the external primer for *CmR* gene 5'-TCA TCG CAG TAC TGT TGT ATT CAT TAA GCA T- 3'.

#### **S1.2.7 YPet-dnaN:ter-mCerulean:LacI-mCherry:oriZ:ΔoriC (oriZ-dnaN-ter)**

The FRT-flanked *oriC-CmR* was recombined out using the pCP20 plasmid as described elsewhere (2). The deletion of *oriC* was verified by PCR and DNA sequencing using the oligonucleotides: 5'-TGT TGT TAA CAG TCT AAC CGG TCA ATT TTT TAT GA-3' and 5'-GCG CTT TGA TAC CGG GAT GC-3'.

#### **S1.2.8 YPet-dnaN:ter-mCerulean:LacI-mCherry:oriZ:ΔoriC:ΔTus (oriZ-dnaN-ter:ΔTus)**

The *oriZ-dnaN-ter:ΔTus* was constructed using P1 transduction. The FRT-flanked *kanR* gene from the *ΔTus* strain was transduced into the *oriC* strain. The presence of the *ΔTus* knock-out was verified using the oligonucleotides: 5' – GCG CAC GAT GGT CAA GTC AC – 3' and 5' – TAC GGC CAG AAC AGA TGG TC – 3'. The sequence of the deleted region was confirmed by DNA sequencing.

### **4.9.3. S1.3 M9 growth medium used in experiments**

1 L of M9 growth medium used in the experiments contains 10.5 g/L of autoclaved M9 broth (Sigma-Aldrich); 0.1 mM of autoclaved CaCl<sub>2</sub> (Sigma-Aldrich); 0.1 mM of autoclaved MgSO<sub>4</sub> (J.T.Baker); 0.3% of filter-sterilized glycerol (Sigma-Aldrich) as carbon source; 0.1 g/L of filter-sterilized 5 amino acids, namely L-threonine, L-leucine, L-proline, L-histidine and L-arginine (all from Sigma-Aldrich) and 10 μL of 0.5% filter-sterilized Thiamine (Sigma-Aldrich).

#### 4.9.4. S1.4 Microfluidic device fabrication

Cells are immobilized for imaging utilizing our version (8) of a previously reported microfluidic device (9). A detailed description of the fabrication procedure can be found in (8). In brief, electron-beam lithography in combination with dry etching techniques is used to fabricate the structures into a silicon wafer. This wafer is subsequently used to realize a negative mold of the structures with polydimethylsiloxane (PDMS). The resulting PDMS mold is then employed to successfully fabricate the positive structures with PDMS. Subsequently, a cover glass is attached, and the device is used for time-lapse experiments.

#### 4.9.5. S1.5 Preparation of cells for microscopy

Cells were streaked onto Luria-Bertani (LB)-plates containing the appropriate antibiotics. Single colonies from these plates were grown in M9 supplemented with 0.3% glycerol (Gly), essential nutrients (**Supplementary Section 1.2**), and with the appropriate antibiotics overnight at 37 °C with shaking. The following day, the cells were sub-cultured into the same medium and grown at 37 °C with shaking until an  $OD_{600} \sim 0.2$  was reached. Cells were concentrated by centrifugation for 2 min at 16100 g. The supernatant of the concentrated cells was decanted, and the pellet was resuspended in 50  $\mu$ L M9-Gly with essential nutrients and injected into the microfluidic device. After injection into the device, the device was centrifuged for 10 min at 2500 g (Eppendorf 5810R). This centrifugation step assists the loading of the cells into the growth channels. Following centrifugation the device was mounted on the microscope with tubing attached to the inlet and outlet and incubated for ~45 min at 37 °C. After incubation, fresh M9-Gly with essential nutrients (**Supplementary Section 1.2**) and the appropriate antibiotics are flushed through the device. The syringe containing the medium is subsequently attached to an automated syringe pump that continuously injects fresh M9-Gly, essential nutrients and 0.2 mg/mL bovine serum albumin (BSA) into the device at a rate of 0.5 mL/hr.

#### 4.9.6. S1.6 Microscope setup

Microscopy data were acquired on a commercial Nikon Ti microscope equipped with a custom-built laser excitation scheme similar to that reported previously (4). In brief, a Nikon CFI Apo TIRF 100x, 1.49 NA oil immersion objective is used for excitation and detection. A standard Nikon brightfield halogen lamp and condenser components are used for imaging cell outlines. Excitation is performed using a Cobolt Fandango 515 nm continuous wave (CW) diode-pumped solid-state (DPSS) laser for YPet, Cobolt Jive 561 nm CW DPSS laser for mCherry, and a Cobolt Twist 457 nm CW DPSS laser for mCerulean. The laser beams are combined using dichroic mirrors (Chroma 575dcspxr, zt457dcrb) and subsequently coupled into a single-mode optical fiber (KineFLEX). The output of the fiber is expanded and focused onto the back focal plane of the objective mounted on the microscope. Notch filters (Semrock NF03-514E, NF03-561E) were used to eliminate any laser light leaking onto the camera. The emission of the different fluorescent proteins are projected onto the central part of an Andor iXon 897 Electron Multiplying Charge Coupled Device (EMCCD) camera using custom filter sets: Chroma z561, ET605/52m, zt561rdc (mCherry), Chroma z514, ET540/30m, zt514rd (YPet), Chroma z457/10x, ET490/40m, zt457rdc (mCerulean). A custom design commercial temperature control housing (Okolabs) is enclosing the microscope body to regulate the temperature at

37 °C. Sample position was controlled with a Nikon stage (TI-S-ER Motorized Stage Encoded, MEC56100) together with the Nikon Perfect Focus System (PFS) to eliminate Z-drift during image acquisition. A personal computer (PC) running Nikon NIS elements software is used for controlling the acquisition.

#### 4.9.7. S1.7 Time-lapse data acquisition

All data acquisition was performed on the previously described microscope in combination with a standard PC running Nikon NIS-elements (**Supplementary Information Section S1.5**). The order and type of fluorescence excitation were dependent on the strain being imaged.

For *oriC-dnaN* and *oriC-dnaN:Atus* cells, the cell outlines were imaged using standard brightfield illumination, and the YPet proteins were subsequently excited with the 515 nm laser line (80 ms exposure time). Time-lapse images were acquired every 2.5 min with the EMCCD gain set to 100.

For *oriZ-dnaN* and *oriZ-dnaN:Atus* cells, the cell outlines were again imaged using standard brightfield illumination, but different laser lines were used successively to excite the different fluorescent proteins. The sample was first excited with 515 nm (YPet), then 561 nm (mCherry) and lastly with 457 nm (mCerulean) with an exposure time of 80 ms in all the cases. The intensity for all the different measurements was kept constant. The intensity of the 515 nm and 561 nm lasers was  $\sim 5 \text{ W}\cdot\text{cm}^{-2}$ , and the 457 nm laser was set to  $\sim 2.5 \text{ W}\cdot\text{cm}^{-2}$ . Intensity calibration was performed according to (10). Images were acquired every 2.5 min. Time-lapse acquisitions typically ran overnight, and spanning  $\sim 10$  hrs of measurement.

*oriZ-dnaQ* cells were imaged using the same procedure as for *oriZ-dnaN* cells expect that images were acquired every 5 min.

#### 4.9.8. S1.8 Data analysis

##### S1.8.1 Localization of foci during replication

The acquired images were analyzed with custom-written MATLAB software (MathWorks) in combination with ImageJ (11) as reported previously (4). Briefly, we correct for uneven background and illumination heterogeneity per image. Subsequently, we detect foci in each bacterium that have an intensity above the cytoplasmic fluorescence intensity as defined by the median of the total cytoplasmic signal. The detected foci are localized in each individual image by performing a maximum-likelihood estimation (MLE) of a two-foci Gaussian fit (12-14). The resulting fits are evaluated by rejecting the secondary fit if it is off-range, too weak compared to the brightest focus or to the total fluorescence intensity.

##### S1.8.2 Time-resolved representation of foci positions

To study the temporal behavior of the replisome and chromosomal loci, it is crucial to be able to evaluate the process of focus separation during replication. Obtaining average values of focus separation is somewhat impeded by i) the widespread for individual cells in the appearance of ‘one focus’ and ‘two foci’ observations and ii) the limited optical resolution to be able to discriminate one focus, from two closely adjacent ones. To

minimize the influence of these factors, we proceed as follows when constructing averaged time-resolved position traces as in **Figure 3A, 4B** of the main text. First, for each individual cell, we normalize the time axis such that  $t = 0$  corresponds to the moment of initiation, as determined from the first emergence of a focus, and that  $t = 1$  amounts to the last time point that one or two foci are observed. Secondly, we plot the average position for the cases of one or two foci respectively per normalized time. The size of the marker represents the percentage of cells found in this state.

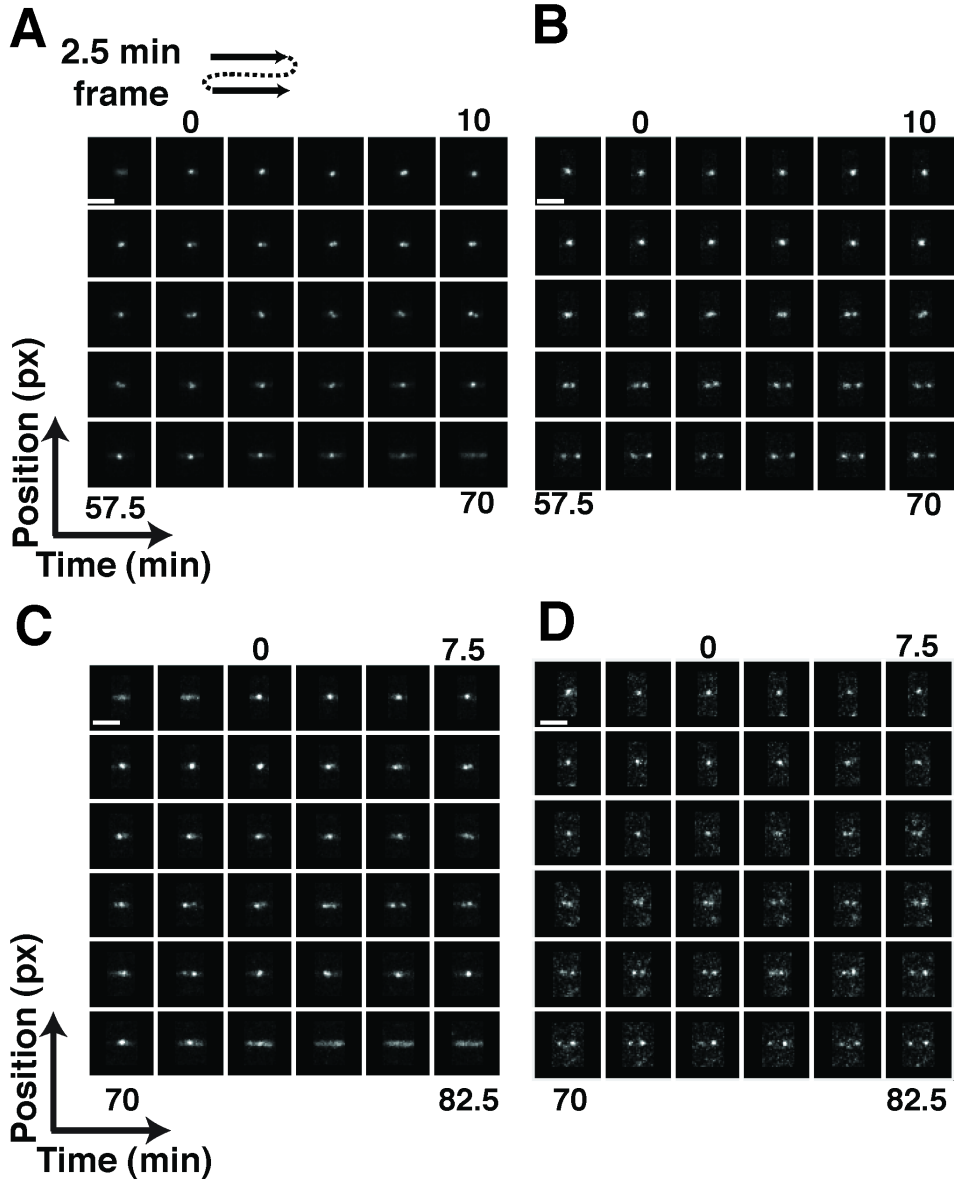
#### **4.9.9. S1.9      Calculating the average replisome velocity during replication**

We calculate the average replisome velocity in our experimental conditions by making use of the average replication time  $t_{\text{rep,oriC}} \sim 70$  min (**Supplementary Information Figure S4B**) in combination with the size of the *E. coli* genome (4.6 Mbp). This results in a velocity of  $v_{\text{rep}} \sim 550$  bp/s for an individual replisome under our experimental conditions.



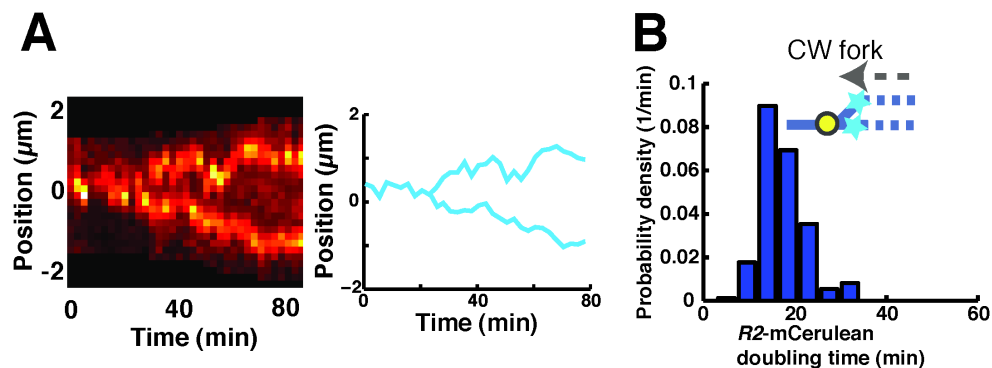
## 4.10. Supplementary Figures and Captions

4.10.1. Supplementary Figure S1. Representative temporal montages of the YPet-DnaN and *oriI*-mCherry signals for a complete replication cycle.



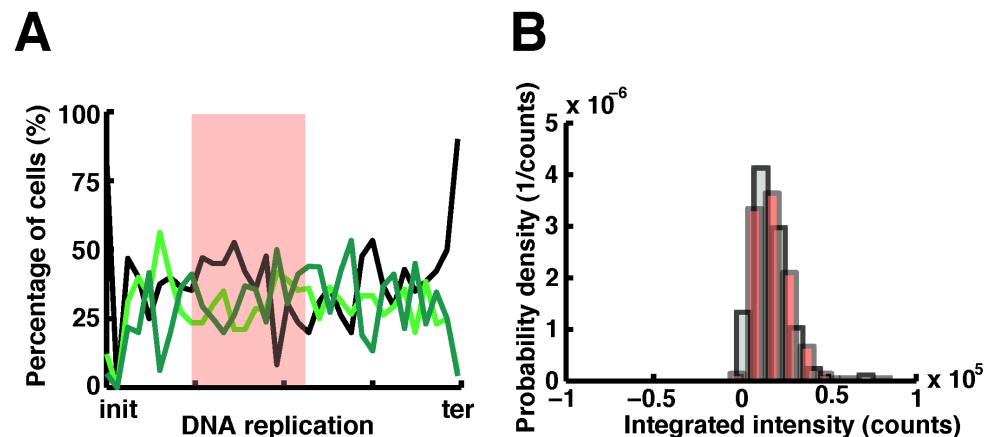
(A-B) Montages for *oriZ-dnaN* cells. (A) YPet-DnaN montage and (B) *ori1*-mCherry montage. (C-D) Montages for *oriZ-dnaN:ΔTus* cells. (C) YPet-DnaN montage and (D) *ori1*-mCherry montage. Scale bars, 3  $\mu\text{m}$ .

#### 4.10.2. Supplementary Figure S2. Focus doubling time of the *R2*-mCerulean locus.



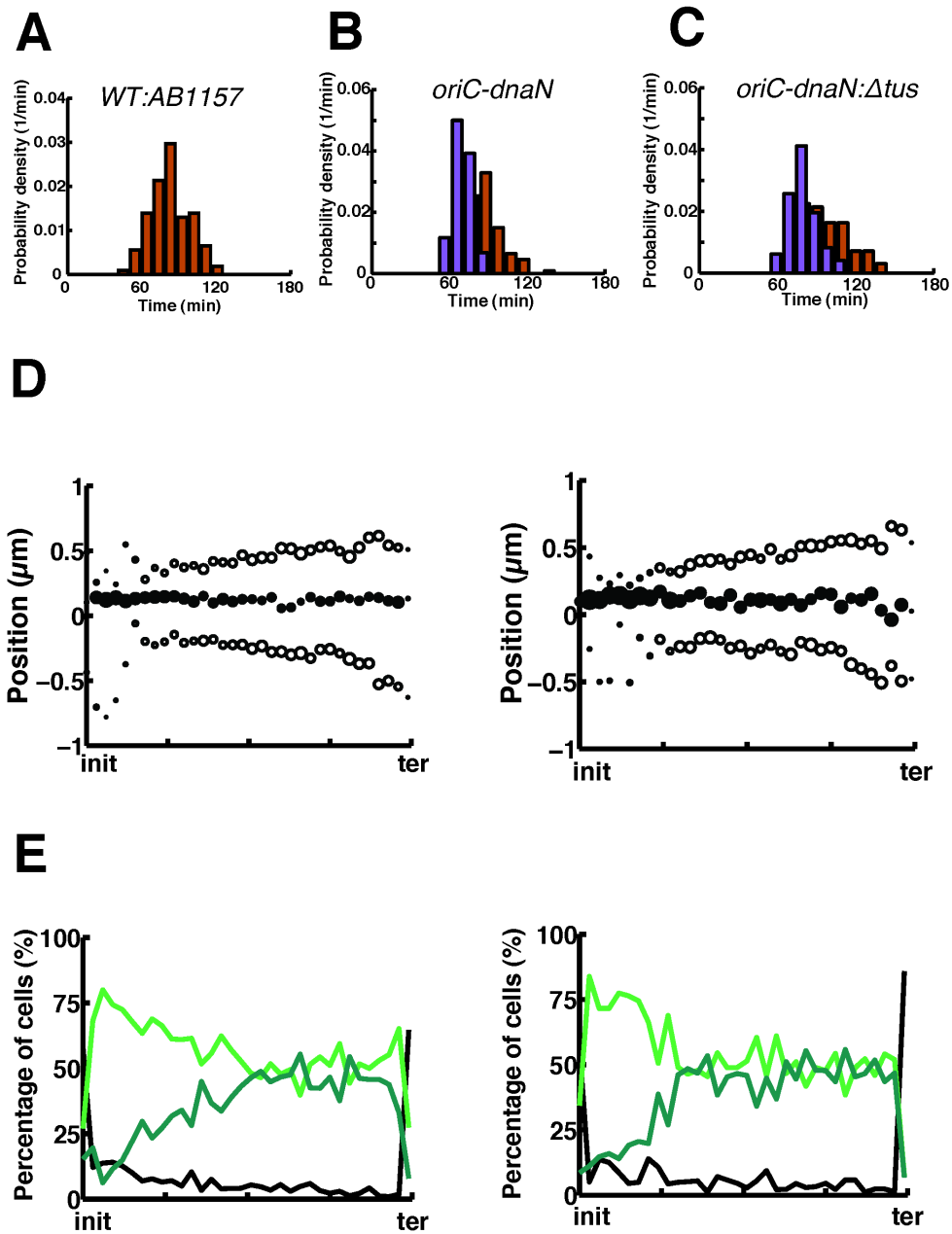
(A) A representative *R2*-mCerulean fluorescent signal. (left) The kymograph of the *R2*-mCerulean locus for one complete replication cycle constructed by summation of the pixel intensities perpendicular to the long axis of the cell. (right) The *R2*-mCherry trace shown (left) determined from the Gaussian fitted positions from individual images for each time point. Images were acquired every 2.5 min during a time-lapse experiment. (B) The distribution of the time point during replication when the *R2*-mCerulean focus doubles. The mean doubling time of *R2*-mCerulean is  $t_{\text{doubling}} = 17 \pm 5$  min. (Error is  $\pm$  SD,  $n_{\text{cells}} = 167$ ). (Inset) Schematic illustration indicating that the *R2*-mCerulean region has been replicated.

#### 4.10.3. Supplementary Figure S3. Replisome dynamics as measured in the *oriZ-dnaQ* strain.



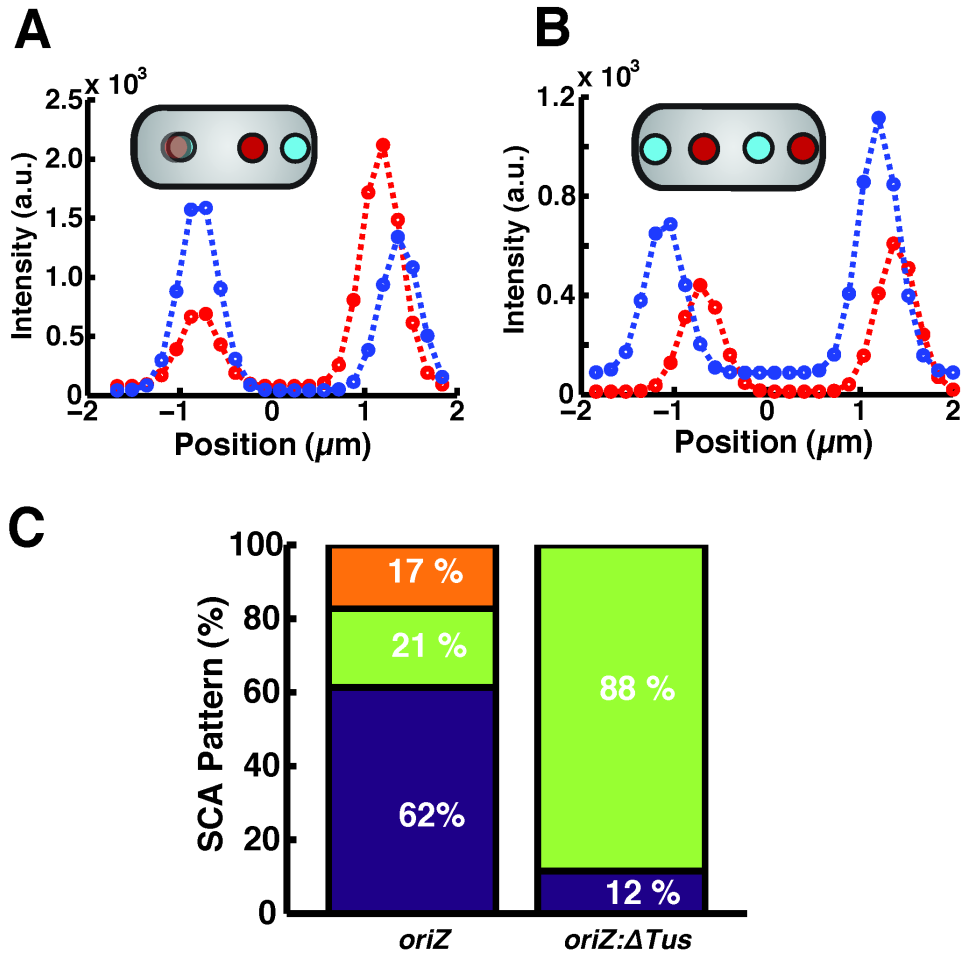
**(A)** The percentage of cells that have a single focus (light green line), double foci (dark green line) and no foci (black line) as a function of replication time. It is evident that the percentage of cells having a single focus or double foci is roughly equally distributed throughout the replication process ( $n_{\text{cells}} = 43$ ). The transparent red rectangle indicates the time when *oriI*-mCherry focus doubles spatially. The width of the rectangle is  $\pm$  SD. The SD is larger for *oriZ-dnaQ* compared to *oriZ-dnaN* due the frame rate being 5 min instead of every 2.5 min per image. **(B)** The intensity distributions of DNA-bound DnaQ-YPet prior (gray) and post (red) spatial doubling of the *oriI*-mCherry focus.  $I_{\text{prior,oriI-mCherry}} = 1.88 \cdot 10^4 \pm 2.25 \cdot 10^4$  counts ( $n_{\text{foci}} = 278$ ),  $I_{\text{post,oriI-mCherry}} = 1.79 \cdot 10^4 \pm 1.01 \cdot 10^4$  counts (mean  $\pm$  SD,  $n_{\text{foci}} = 608$ ). The difference between the means of the two distributions is  $< 6\%$ . While the error bars are more substantial for this DnaQ-YPet experiment compared to the YPet-DnaN experiment, a result of lower signal and reduced statistics. The mean number of DNA-bound DnaQ-YPet thus appears unchanged after *oriI*-mCherry has been replicated, i.e. after the CW replisome has encountered the Tus-*TerC* roadblock.

#### 4.10.4. Supplementary Figure S4. Replication in *oriC-dnaN* and *oriC-dnaN:Δtus* characteristics.



**(A)** Distribution of the division time for WT AB1157 cells. The average division time is  $t_{\text{div,AB1157}}=83\pm 16$  min (Error is  $\pm$  SD,  $n_{\text{cells}} = 125$ ). **(B)** Distribution of the division time (brown) and replication time (purple) for *oriC-dnaN* cells. The average replication time  $t_{\text{rep,oriC-dnaN}} = 70\pm 7$  min, and the average division time  $t_{\text{div,oriC-dnaN}} = 85\pm 15$  min (error is  $\pm$  SD,  $n_{\text{cells}} = 110$ ). **(C)** Distribution of the division time (brown) and replication time (purple) for *oriC-dnaN:ΔTus* cells. The average replication time  $t_{\text{rep,oriC-dnaN:ΔTus}} = 80\pm 11$  min, and the average division time  $t_{\text{div,oriC-dnaN:ΔTus}} = 96\pm 19$  min (error is  $\pm$  SD). **(D)** Replisome dynamics for *oriC-dnaN* and *oriC-dnaN:ΔTus* cells. Time-resolved traces of (left) *oriC-dnaN*, and (right) *oriC-dnaN:ΔTus* cells. Here we plot a single DnaN-YPet focus (filled circles) and double DnaN-YPet foci (empty circles). The size of an individual circle at each time point is representative of the percentage of cells having that particular distribution of foci. The traces have been aligned with respect to initiation and termination of replication and binned. **(E)** The percentage of cells having a single focus (light green line), double foci (dark green line), and no foci (black line) as a function of replication time. It is evident that the percentage of cells having a single focus or double foci is roughly equal for (left) *oriC-dnaN* and (right) *oriC-dnaN:ΔTus* cells.

## 4.10.5. Supplementary Figure S5. SCA patterns.



(A-B) The fitted spatial intensity distributions in a minority of the cells. (A) *ori1*-mCherry and *R2*-mCerulean foci *spatially* overlapped. (B) RLRL configuration, which is symmetric and taken to be equal to LRLR. The insets in (A) and (B) are schematic depictions of the different chromosomal loci patterns. (C) The observed SCA pattern determined at the time point when replication was half way.

## 4.11. Supplementary Tables

### 4.11.1. Supplementary Table S1

Frames prior and post <i>oriI</i> doubling	$I_{\text{prior,oriI-mCherry}}$	$I_{\text{post,oriI-mCherry}}$
2	$1.44 \cdot 10^5 \pm 6.16 \cdot 10^4$	$1.42 \cdot 10^5 \pm 6.52 \cdot 10^4$
5	$1.41 \cdot 10^5 \pm 6.18 \cdot 10^4$	$1.41 \cdot 10^5 \pm 6.80 \cdot 10^4$
10	$1.42 \cdot 10^5 \pm 6.53 \cdot 10^4$	$1.39 \cdot 10^5 \pm 6.83 \cdot 10^4$

Average intensities of DNA-bound YPet-DnaN prior and post spatial doubling of the *oriI*-mCherry focus. The numbers specified indicate the mean  $\pm$  SD.

### 4.11.2. Supplementary Table S2

Plasmids	Relevant genotype	Construction
pKD46	Plasmid with $\lambda$ -Red recombinase genes expressed under arabinose promoter	Created by standard cloning (2)
pCP20	Temperature sensitive plasmid with constitutively expressing flippase (FLP) enzyme.	Created by standard cloning (2)
pKD13	Template plasmid containing the FRT-flanked kanamycin resistance gene ( <i>kanR</i> ) sequence.	Created by standard cloning (2)

Summary of different plasmids used in this study.

### 4.11.3. Supplementary Table S3

Strains	Relevant genotype	Construction
BN1110	AB1157 strain containing pKD46 plasmid	<i>E. coli</i> K-12 derivative (7)
BN1219	<i>YPet-dnaN</i> and <i>termCerulean</i>	P1-phage transduction (4)
BN1516	<i>kanR</i> gene recombined out	FLP/ <i>FRT</i> recombination.
BN1598	$\Delta$ <i>Tus</i>	$\lambda$ -red recombination: <i>FRT-kanR-FRT</i> from pKD13 $\rightarrow$ BN1110
BN1868	<i>oriC-dnaN:ATus</i>	P1-phage transduction: BN1598 $\rightarrow$ BN1516
BN1861	<i>Ypet-DnaN</i> , $\Delta$ <i>oriC</i> , and <i>oriZ</i>	P1-phage transduction (5)
BN1869	<i>oriZ-dnaN</i> with the <i>kanR</i> gene out recombined	FLP/ <i>FRT</i> recombination
BN2111	<i>oriZ-dnaN:ATus</i>	P1-phage transduction: BN1598 $\rightarrow$ BN1869
BN2381	<i>oriZ-dnaQ</i>	P1-phage transduction <i>dnaQ-YPet</i> $\rightarrow$ <i>oriZ</i> without a replisome marker (7)
BN2392	<i>LacI-mCherry ::LeuB</i>	$\lambda$ -red recombination (7)
BN2393	<i>oriC-dnaN:LacI-mCherry</i>	P1-phage transduction: BN2393 $\rightarrow$ BN1516
BN2389	<i>OriC-OriZ</i>	$\lambda$ -red recombination (5)
BN2394	<i>oriC-oriZ-dnaN-ter</i>	P1-phage transduction: BN2389 $\rightarrow$ BN2393
BN2395	<i>oriC-oriZ-dnaN-ter</i> with the <i>kanR</i> gene out recombined	FLP/ <i>FRT</i> recombination
BN2390	<i>frt-oriC-CmR-frt</i>	$\lambda$ -red recombination (5)
BN2396	<i>frt-oriC-oriZ-dnaN-ter</i>	P1-phage transduction: BN2390 $\rightarrow$ BN2395
BN2397	<i>oriZ-dnaN-ter</i> with the <i>oriC-CmR</i> insert out recombined	FLP/ <i>FRT</i> recombination
BN2398	<i>oriZ-dnaN-ter:ATus</i>	P1-phage transduction: BN1598 $\rightarrow$ BN2397

Summary of different strains used in this study.



## 4.12. Supplementary References

1. B. J. Bachmann, Pedigrees of some mutant strains of *Escherichia coli* K-12. *Bacteriological Reviews* **36**, 525-557 (1972).
2. K. A. Datsenko, B. L. Wanner, One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proceedings Of The National Academy Of Sciences Of The United States Of America* **97**, 6640-6645 (2000).
3. L. C. Thomason, N. Costantino, D. L. Court, *E. coli* genome manipulation by P1 transduction. *Current protocols in molecular biology* **Chapter 1**, Unit 1.17 (2007).
4. M. C. Moolman *et al.*, Slow unloading leads to DNA-bound  $\beta_2$ -sliding clamp accumulation in live *Escherichia coli* cells. *Nature communications* **5**, 5820 (2014).
5. X. Wang, C. Lesterlin, R. Reyes-Lamothe, G. Ball, D. J. Sherratt, Replication and segregation of an *Escherichia coli* chromosome with two replication origins. *Proceedings Of The National Academy Of Sciences Of The United States Of America* **108**, E243-250 (2011).
6. R. Reyes-Lamothe, D. J. Sherratt, M. C. Leake, Stoichiometry and architecture of active DNA replication machinery in *Escherichia coli*. *Science* **328**, 498-501 (2010).
7. R. Reyes-Lamothe, C. Possoz, O. Danilova, D. J. Sherratt, Independent Positioning and Action of *Escherichia coli* Replisomes in Live Cells. *Cell* **133**, 90-102 (2008).
8. M. C. Moolman, Z. Huang, S. T. Krishnan, J. W. J. Kerssemakers, N. H. Dekker, Electron beam fabrication of a microfluidic device for studying submicron-scale bacteria. *Journal of Nanobiotechnology* **11**, 12-12 (2013).
9. P. Wang *et al.*, Robust growth of *Escherichia coli*. *Current biology : CB* **20**, 1099-1103 (2010).
10. D. Grünwald, S. M. Shenoy, S. Burke, R. H. Singer, Calibrating excitation light fluxes for quantitative light microscopy in cell biology. *Nature protocols* **3**, 1809-1814 (2008).
11. C. A. Schneider, W. S. Rasband, K. W. Eliceiri, NIH Image to ImageJ: 25 years of image analysis. *Nature Methods* **9**, 671-675 (2012).
12. R. E. Thompson, D. R. Larson, W. W. Webb, Precise nanometer localization analysis for individual fluorescent probes. *Biophysical Journal* **82**, 2775-2783 (2002).

13. C. S. Smith, N. Joseph, B. Rieger, K. A. Lidke, Fast, single-molecule localization that achieves theoretically minimum uncertainty. *Nature Methods* **7**, 373-375 (2010).
14. K. I. Mortensen, L. S. Churchman, J. A. Spudich, H. Flyvbjerg, Optimized localization analysis for single-molecule tracking and super-resolution microscopy. *Nature Methods* **7**, 377-381 (2010).



## 5 DYNAMICS AND STOICHIOMETRY OF TUS DURING THE CELL CYCLE OF *ESCHERICHIA COLI*

*The bi-directional movement of E. coli DNA replication forks on a circular chromosome raises the questions of where and how the termination of DNA replication occurs. The polar arrest of replication forks by Tus protein binding to Ter sites is a widely accepted mechanism using which some of the above questions can be answered. However, the in vivo localization, stoichiometry and protein-protein interactions of Tus-Ter complexes are unknown. To study the localization and stoichiometry of Tus during the cell cycle, we have labeled the native Tus proteins at its C-terminus with mYpet fluorescent protein and used long time-lapse single-molecule fluorescence microscopy together with microfluidics. From our initial observations of over 20 cell cycles, we conclude that Tus-mYpet proteins form an average of four foci (signal above the background) in the cells possibly due to DNA binding and that they are broadly distributed from the cell poles to mid-cell. From the fluorescence data, we observed that the Tus is expressed at low copy number, varying on average from around eight to thirteen molecules per cell during the cell cycle. These numbers confirm the results from recently published FASP-based proteomics studies, and both contradicted the higher copy number values (one hundred molecules per cell) obtained from previous bulk experiments. We found from our experiments that the Tus-mYpet foci are mostly static during the cell cycle since they are not recovered following photo-bleaching. Approximately two molecules of Tus-mYpet form the highest intensity focus, and it rarely co-localizes near the TerBC sites. No substantial dynamics in the Tus-mYpet foci intensity were observed during the cell cycle to evidently support the fork-trap model of replication termination, and further studies are necessary to infer the role played by Tus-Ter in vivo.*

## 5.1. Introduction

DNA replication of the 4.6 Mbp circular chromosome which is compactly organized as a ‘nucleoid’ in an *E. coli* cell starts at the origin region (*oriC*) by forming a ‘replication bubble’ (1). A protein complex comprising of 13 different proteins called the ‘replisome’, aids in the semi-conservative replication process during which the two single strands of DNA (ssDNA) in the bubble serve as templates for making two double strands of DNA (dsDNA) (2). Two sets of replisomes form two independent replication forks proceeding in opposite directions from the *oriC* region towards the terminus region (3, 4). The bi-directional movement of the replication forks along a circular chromosome raises two main questions: (i) how the forks meet during the termination of DNA replication and (ii) wherein the termination region of chromosome the replication ends (5).

Three important models namely the fork collision model, the fork-trap model, and the *dif*-stop model have been developed and experiments have been performed to address the above-mentioned questions (6). We now describe these in turn. Of the three models, ‘the molecular fork-trap model’ has been most widely accepted as describing the mechanism of replication termination in *Escherichia coli* (*E. coli*) and *Bacillus subtilis*. In this model, a ‘protein-DNA’ lock traps one of the replication forks while the other fork is let to pass through to avoid fork collision (7, 8). The protein candidate for the fork-trap model in *E. coli* is the ‘Termination utilization substance’ (Tus) protein which recognizes and strongly binds to the ten *Ter* sites (size: 23 bp; denoted from *TerA* to *TerJ*) that are distributed over the termination region of the chromosome (9, 10). The Tus proteins bind to *Ter* sites asymmetrically in a one-to-one manner, and depending on the direction of the replication fork, it presents one of its two sides (permissive and nonpermissive) to the approaching fork. The permissive side binds only transiently to the *Ter* site, because of which the replication fork ejects the Tus protein from the *Ter* site. The other nonpermissive side binds strongly to the cytosine residue (G-C6 base pair) of the *Ter* site, thus temporarily pauses the replication (5, 11, 12). This phenomenon is called the ‘polar arrest of the replication fork’ (more information is detailed in Chapter 4) (12, 13).

In the *dif*-stop model, the replication forks terminate precisely at the *dif* site (deletion-induced filamentation). This model is supported by the observation that the deletion of Tus causes no observable change in the phenotype of the bacterial cell (14). More supportive evidence for this model was provided by using the bioinformatics techniques to find the variations on the mutational bias patterns in bacterial chromosome sequences. The results indicated that the DNA replication termination might occur near or at *dif* site but not at the *Ter* sites. However, recent *in vitro* and *in vivo* studies in *E. coli* using normal and ectopic origin strains have shown that the replication forks do not stop at the *dif* site (15, 16).

According to the fork collision model, the replisome components complete DNA replication, and simply unbind after recognizing fork collision through an unknown mechanism. This model gained more importance following a study that included *E. coli* strains containing two origins of replication. In such strains, two types of termination were observed. One of the termination events happened near the *Ter* site similar to normal origin strains, while another termination event occurred possibly due to fork collision (6, 15).

However, further investigations are necessary to understand the exact mechanism to validate the fork collision model.

Among the three models of replication termination described here, many evidences support the role of Tus in arresting the replication forks during termination. Hence, this protein will be focused on the rest of the chapter. Many *in vitro* and *in vivo* studies have been performed to elucidate the mechanism of polar arrest in which Tus-*Ter* binding is the first step (7, 12). Tus is found to have varying binding affinities *in vitro* for all the *Ter* sites, and among them, *TerF* has the weakest binding affinity which is almost 880-fold weaker than the strongest binding *TerA* (5, 10). Since the amino acid residues of Tus that are involved in binding to *Ter* sites and those forming the Tus-*Ter* lock are different, a locking mechanism similar to that of a ‘mousetrap’ was postulated. When the replicative helicase (DnaB) unwinds the DNA, the molecular mousetrap is set and the G-C6 base-pair is flipped out of the helical DNA axis. The exposed C6 base is bound to the amino acids present in the locking domain of Tus (detailed information is presented in Chapter 4) (7, 8). Recently, Tus-*Ter* mousetrap was found to operate at a single-molecule level on a hairpin DNA with *TerB* site, using only the mechanical forces generated by the magnetic tweezer technique that mimics the strand separation by helicases (17). The direct implication from this study is that no specific interaction with helicases is necessary for successful Tus-*Ter* lock formation (17). The binding efficiency observed from these helicase-mimicking experiments is higher than that concluded from the previous *in vivo* experiments (18) performed with plasmids bearing *Ter* sites. Hence it was concluded that aspecific DnaB-Tus interactions, such as those resulting from steric effects, might actually hamper the arrest of replication forks (17). This illustrates that the exact interactions of Tus with the replication fork merits further study.

In the *in vivo* context, neither the exact number nor the distribution of the Tus proteins bound to the native chromosome is known. Different studies have yielded disparate estimates of the copy number of Tus proteins expressed in cells. Bulk biochemical analysis of *E. coli* cell extracts has shown the copy number of Tus proteins to be around one hundred per cell (19) while recent high-throughput proteomics studies based on Filter Aided Sample Preparation (FASP) technique revealed an average concentration of fourteen Tus proteins per cell (20). Similarly, the localization of Tus-*Ter* complexes within the cells is not well studied. Their localization information may be indirectly interpreted from the chromosomal macrodomains, which are the highly ordered structures of the nucleoid. The four macrodomains of an *E. coli* chromosome are named as Ori, Left, Right and Ter and various nucleoid associated proteins (NAP) like SeqA, MatP, SlmA, etc. (21) segregate them. The position of chromosomal macrodomains *in vivo* across the cell cycle is already known through labeling the macrodomain specific NAP. It is found that the Left and Right macrodomains are mostly localized near the poles while the Ori and Ter macrodomains are confined at either side of the mid-cell during the cell cycle (21). From the sequence information of *E. coli* genome, it is well-known that the ten *Ter* sites are distributed within and beyond the Ter macrodomain i.e. into the Left and Right macrodomains (5). Hence it is reasonable to assume that the Tus bound *Ter* sites, similar to their macrodomains, might be localized from cell poles to a side of the mid-cell. However, the localization of Tus-*Ter* complexes has never been precisely observed in live cells (22, 23).

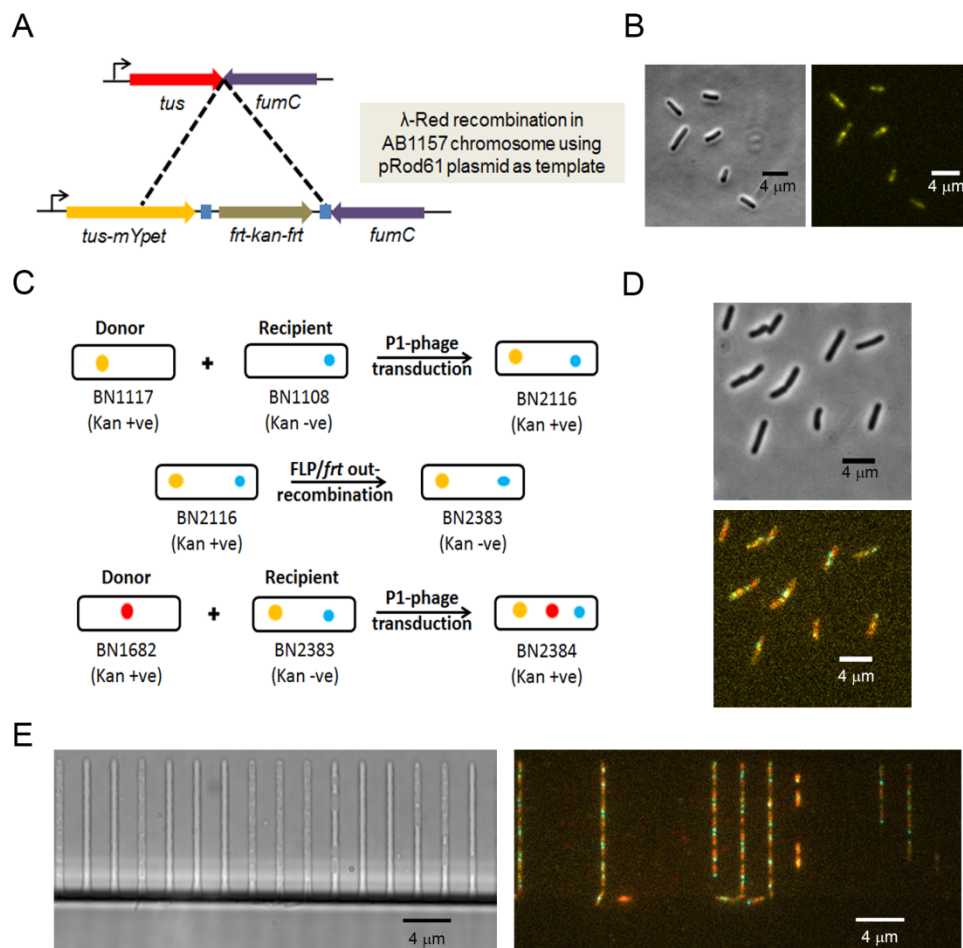
In this study, we aim to elucidate the localization of Tus bound to *Ter* sites within the chromosomal macrodomains, the stoichiometry and the dynamics of Tus-*Ter* complex especially the binding and unbinding events in live cells. Using fluorescence microscopy, we can detect the binding of fluorescently labeled Tus proteins to DNA as foci (signal above the background) inside the cell. The unbound proteins produce a diffused fluorescence signal throughout the cell. The use of microfluidic devices enables us to perform long time-lapse measurements and image the cells throughout their cell cycles. Using multi-color fluorescence imaging (yellow, cyan and red) on this strain mainly yields two types of information: (i) co-localization of Tus foci (yellow) with respect to '*ter*' region (chromosomal region between *TerBC* sites, 50 kb clockwise from the *dif* site labeled with cyan) across the replication cycle (red); and (ii) variation of the copy number of labeled Tus proteins (both as foci and diffused) during the cell growth. Our initial results on 20 cell cycles showed that only eight to thirteen molecules of Tus are expressed during a cell cycle. Around five molecules of Tus form foci and they are found to be distributed throughout the cell space during most of the cell cycle.

## 5.2. Results and discussion

### 5.2.1. Visualization of Tus-mYpet proteins throughout the cell cycle

In this study, we labeled the native Tus proteins with mYpet (monomeric yellow fluorescent protein for energy transfer) using  $\lambda$ -Red recombination as detailed in **Materials and Methods (Figure 1A)**. The chromosomally engineered strain (Tus-mYpet, BN1117) were validated (24) as discussed in Chapter 2 of this thesis (**Figures S1A, S1B, and S1C**). To visualize the intracellularly expressed Tus-mYpet proteins, we acquired fluorescence microscopy images of the cells distributed on an agarose pad (**Figure 1B**). From initial observations, we found that the fluorescence signal of Tus-mYpet is low and is distributed homogeneously throughout the cell with the occasional appearance of fluorescent foci, mostly near the poles. To further study these foci, we used P1 phage transduction to fuse the Tus-mYpet with the chromosomal '*ter*' region labeled with mCerulean fluorescent protein (cyan) (4) and native  $\beta$ -clamp proteins tagged with the mCherry fluorescent protein (cherry red) (25) in a single *E. coli* strain. The creation of this fusion strain (BN2384) is detailed in the **Materials and Methods** section (**Figure 1C**).

In the BN2384 strain, the labeled *ter* region appears as a diffraction-limited fluorescent spot in growing cells giving the real time positions of the nearby *TerBC* sites (4), and the  $\beta$ -mCherry proteins precisely determine the start and the end of DNA replication in a cell cycle as detailed in Chapter 3 (25). This strain was also validated via growth curve analysis in both LB and M9 media, and we observed no difference in the growth characteristics between BN2384, BN1117 (parent strain) and wildtype AB1157 strains (**Figure S1C**). Furthermore, we found that 0.1 mM IPTG is optimal for inducing the TetR-mCerulean protein expression, which binds to the *tetO* array sequences without affecting the DNA replication and segregation. From the fluorescence microscopy images of the BN2384 strain (**Figure 1D**) distributed on the agarose pad, we were not able to resolve the relative position of Tus-mYpet foci to the *ter* region in cells because the imaged cells were at different stages of cell cycle, and the visible foci formation might be cell cycle-dependent.



**Figure 1: Visualization of Tus-mYpet proteins during the complete cell cycle.** **A.** Scheme depicts the AB1157 chromosome region where the *tus* gene is genetically modified using  $\lambda$ -Red recombination to label the native Tus proteins with mYpet fluorescent protein. **B.** Phase contrast image (left side) and fluorescence image (right side) of the BN1117 (Tus-mYpet) cells fixed on agarose pad. **C.** Scheme depicts the strain creation steps to fuse *tus-mYpet* (BN1117) and *mCherry-dnaN* (BN1682) sequences into the *ter-mCerulean* (BN1108) strain to create the final strain BN2384 that was used extensively in this study. **D.** Phase contrast image (top) and fluorescence image (bottom) of the BN2384 cells fixed on agarose pad. Fluorescence image is a composite of Tus-mYpet (yellow), mCherry-DnaN (red) and *ter-mCerulean* (cyan) images. **E.** Brightfield image (left side) and composite fluorescence image (right side) snapshots of the BN2384 cells steadily growing in the channels of mother-machine microfluidic device.

Thus, to study the cell cycle dependence of Tus-mYpet foci, we opted to employ microfluidics for long time-lapse measurements as detailed in **Materials and Methods**. To avoid phototoxicity, image acquisition parameters had to be optimized and the final settings applied are presented in **Table 1**. A snapshot of a representative measurement is shown in



**Figure 1E.** During the imaging experiments, we faced three major challenges. First, we detected only a low recovery of the photo-bleached Tus-mYpet proteins after the long time-lapse experiments (time  $\sim 10$  hr). Second, the cells with photo-bleached Tus-mYpet were still growing normally (i.e. replication and division times were the same before and after photobleaching). These observations imply that the photo-bleached Tus-mYpet proteins compete with fluorescent Tus-mYpet proteins during foci formation. This may result in an incomplete view of Tus dynamics. Third, the step during the image analysis in which the edges of a growing cell are detected clearly across the cell cycle, (detailed in **Materials and Methods**) yielded a small sample size ( $N = 20$ , obtained from the previous step where  $N \approx 86$ ). In this chapter, we will discuss the initial key observations that result from the data analyzed of those 20 cell cycles.

**Table 1: Image acquisition settings used long time-lapse microscopy experiments**

Labeled protein or DNA site	Laser wavelength (nm)	Laser intensity* ( $\text{W}\cdot\text{cm}^{-2}$ )	Exposure time (ms)	Image acquisition time (min/frame)	EM gain of camera
Tus-mYpet	515	$\sim 5$	80	5	100
$\beta$ -mCherry	565	$\sim 1$	80	2.5	100
<i>ter</i> -mCerulean	457	$\sim 0.94$	30	2.5	100

\*The values were calculated using the method as described in literature (39).

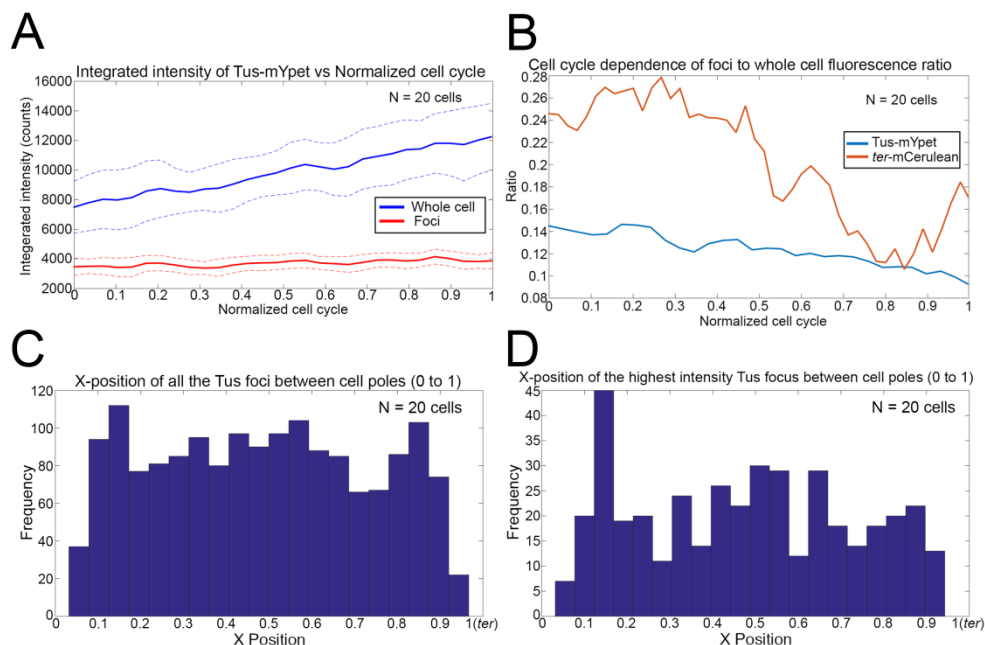
### 5.2.2. Tus-mYpet foci are mostly static and are distributed throughout the cell space

The combination of fluorescence microscopy with microfluidics enables us to visualize the labeled native proteins in action during the cell cycle. However, the analysis of the low signal from Tus-mYpet foci in the cells required a MATLAB based spot detection method, which is detailed in **Materials and Methods**. Within the constraints presented by this method, on an average four different Tus-mYpet foci were detected in the cell at a single time point. The X, Y positions of the detected foci in the cell and their mean integrated fluorescence intensities were then used for further analysis. We calculated the integrated intensity of Tus-mYpet fluorescence from the foci and whole cell and plotted them against the normalized cell cycle (**Figure 2A**). From these results, we found the cytoplasmic content of Tus per unit cell length nearly doubled by the end of the cell cycle when the Tus proteins are equally segregated to the two daughter cells, as would be expected. However, no significant change was noticed in the Tus-mYpet foci intensities during the cell cycle (**Figure 2A**). As a control when we studied the *ter*-mCerulean foci, which are mainly formed by the binding of TetR-mCerulean proteins to the *tetO* arrays near the *ter* region, we observed no significant difference in their foci intensities similar to that observed in Tus-mYpet (**Figure S2D**).

For understanding the dynamics of foci intensities better, we plotted the ratio of the fluorescence intensities in foci to the whole cell of both Tus-mYpet and *ter*-mCerulean against the normalized cell cycle (**Figures 2B**). The plots revealed that for Tus-mYpet the average ratio of foci to whole cell intensities decreased linearly during the cell cycle, while that of *ter*-mCerulean it gradually decreased only near the end of the cell cycle (at the normalized time of 0.85) after which it increased marginally until the cell has divided (**Figure 2B**). This could be because when the *ter* region is being replicated, TetR-mCerulean proteins are released from *tetO* arrays in the chromosome, which in turn could be detected as weaker foci formation until proper segregation (26). It suggests that this point of the cell cycle where the lowest foci to whole cell intensity ratio for *ter*-mCerulean was observed could possibly be the termination phase of the DNA replication. Correspondingly, during this stage of the cell cycle, only a gradual decrease of the foci to whole cell intensity ratio was observed for the Tus-mYpet fluorescence (**Figure 2B**). The decrease in the ratio may be attributed to the steady increase in the whole cell intensity of the Tus-mYpet inside the growing cell or it may be due to the Tus proteins unbinding from the *Ter* sites. Further control experiments are necessary in order to resolve the exact reason for the decrease in foci to whole cell intensity ratio.

Now that we know the static nature of the Tus-mYpet foci intensities, we attempted to find their localization information during the cell cycle. As mentioned previously the *ter*-mCerulean foci serve as an indicator for the positions of *TerBC* sites in growing cells. Since these sites showed the highest binding affinities both *in vivo* and *in vitro* (5, 23), we were interested in the co-localization of Tus-mYpet foci and the *ter* region. For this, we first changed the orientation of the cell-cycle images (referred in the **Materials and Methods** section as ‘bacpics’) in such a way that the cell pole nearest to the *ter*-mCerulean foci is oriented to the right-hand side. The detected four Tus-mYpet foci seemed to be localized over a wide range, from cell poles to mid-cell during the whole cell cycle time (**Figure 2C**). This localization is similar to that of the *Ter* sites. It should be noted that there could be a specific pattern in the position of all the Tus-mYpet foci at a specific point of the cell cycle, which we were not able to conclude because of the small sample number of 20 cell cycles.

The intensities of detected Tus-mYpet foci were not constant throughout the cell cycle and it was difficult to differentiate between the foci in the cell. Hence, to start with, we chose the focus with the highest mean integrated intensity value ( $1201.84 \pm 323.8$  counts) which appeared occasionally during the cell cycle in all the 20 cells studied. We were mainly interested in the highest intensity focus because the increased fluorescence intensity is usually correlated to the grouping of two or more Tus-mYpet molecules. Such a phenomenon can happen only when Tus molecules bind to *Ter* sites that are clustering or transiently co-localizing to each other. The relative positions of the highest intensity Tus-mYpet focus with respect to the position of *ter* region were plotted in a histogram (**Figure 2D**). From this observation, we concluded that the highest intensity Tus-mYpet focus rarely co-localized with the *ter* region where the strongest binding *Ter* sites (*TerBC*) are located. In other words, the *TerBC* sites hardly form a cluster or co-localize with each other *in vivo*. Further investigations are necessary to understand the exact reason behind such a high-intensity focus formation.

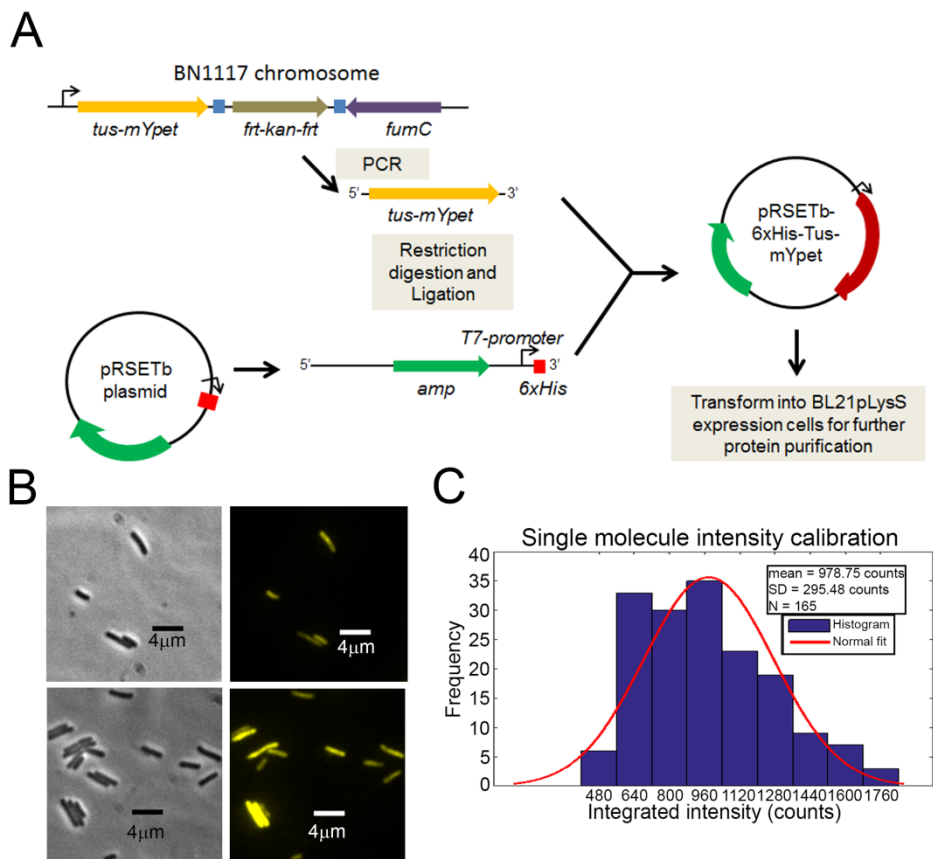


**Figure 2: Results of long time-lapse live cell imaging experiments.** **A.** Integrated intensities of Tus-mYpet fluorescence are plotted against normalized cell cycle indicated as 0 and 1 for the start and end of bacterial division cycle. Whole cell fluorescence intensities are indicated as blue lines and the foci intensities are shown as red lines. Dotted lines show the corresponding upper and lower limits of the standard error of mean values. **B.** The mean ratio of foci to whole cell fluorescence intensities are plotted against normalized cell cycle indicated as 0 and 1 for the start and end of bacterial division cycle. Data for Tus-mYpet fluorescence is indicated in blue line and that for *ter*-mCerulean fluorescence is shown as a brown line. **C.** Histogram plot shows the distribution of all the Tus-mYpet foci positions across the poles of a cell (1 refers to a pole near the *ter* region). **D.** Histogram plot shows the distribution of the highest intensity Tus-mYpet focus position across the poles of a cell (1 refers to a pole near the *ter* region).

### 5.2.3. Stoichiometry of Tus proteins during the cell cycle

Quantitative fluorescence microscopy opens up the possibility to count the number of labeled Tus-mYpet molecules inside the growing cells. *In vitro* based intensity calibration of single fluorescently fused protein molecules is a critical step for this technique. Hence, we purified the Tus-mYpet proteins by cloning the Tus-mYpet ORF into an expression vector (pRSETb) and inducing the protein expression with IPTG in BL21P1Lys cells (Figure 3A and 3B) as described in **Materials and Methods**. We then employed the experimental design as described in Chapter 3 of this thesis (25). Using the data analysis procedure for single-molecule calibration as detailed in **Materials and Methods**, the mean integrated intensity of a Tus-mYpet molecule was found to be  $978.7 \pm 295.5$  counts (SD) in a sample of 165 molecules (Figure 3C).

The number of Tus-mYpet molecules observed during the cell cycle was calculated by dividing the calibrated intensity counts of single Tus-mYpet molecules from the intensities observed from the cell cycle images mentioned previously (N = 20 cell cycles). The observed mean stoichiometry of the Tus-mYpet molecules during the cell cycle, in whole cell and foci, are indicated in **Table 2**. The average number of Tus molecules possibly bound to *Ter* sites (i.e. forming detectable foci) in the *E. coli* cell was found to be around five molecules, among which two molecules form the highest intensity focus during cell growth. The observed whole cell stoichiometry data correspond well to the range of the FASP based proteomics data that states the average number of Tus proteins to be fourteen per cell (20) and thus are in contrast with the estimate of ~100 molecules per cell as suggested in bulk biochemical experiments (19). It is of interest to note that the average number of Tus proteins during the cell cycle is on par with the number of strongest binding *Ter* sites on the replicating chromosome. However, further control experiments are necessary to conclude that the Tus proteins are minimally expressed to form only the Tus-*Ter* complexes and not bound to anywhere in the cell.



**Figure 3: Single molecule calibration of the purified Tus-mYpet proteins. A.** The scheme here

shows the steps in creating his-tagged Tus-mYpet expression plasmid. **B. Top:** Phase contrast image (left) and Fluorescence image (Top right) of the non-induced cells in which the leaky expression of his-tagged Tus-mYpet can be seen. **Bottom:** Phase contrast image (left) and Fluorescence image (Top right) of the IPTG-induced cells in which his-tagged Tus-mYpet proteins are over-expressed. **C.** The histogram shows the distribution of the integrated intensities of a Tus-mYpet protein bound to the anti-mYpet antibody in the M9 growth medium.

**Table 2: Stoichiometry of Tus proteins using single-molecule intensity values (N = 20 cell cycles)**

Number of Tus-mYpet molecules	Mean with standard error
Forming the foci (i.e. possibly bound to <i>Ter</i> site)	5±1
In the whole cell at the start of cell cycle (including those in foci)	8±2
In the whole cell at the end of cell cycle (including those in foci)	13±2
Forming the highest intensity focus (See <b>Figure S2E</b> )	2±1

From the initial observations presented in this study, we were not able to detect discernable foci intensity dynamics of the Tus-mYpet proteins that can reveal the release of Tus-*Ter* roadblock *in vivo*. Owing to the low copy-number of the Tus proteins and the static behavior of the Tus foci intensities observed in our experiments, it necessitates a better and direct evidence to conclude which of the models of DNA replication termination predominantly occurs in the *E. coli* cells.

### 5.3. Conclusions

In this study, we have successfully labeled the native Tus protein with mYpet fluorescent protein (Tus-mYpet) in *E. coli* AB1157 strain to understand its dynamics and stoichiometry. We described here in detail a methodology to study a protein like Tus, which is expressed at low levels in an *E. coli* cell. From our study, we concluded that the Tus-mYpet forms on average four different foci possibly due to Tus-*Ter* binding, and they do not change much during the cell cycle. The average number of Tus molecules forming the foci *in vivo* was found to be around five molecules and, as expected, they are distributed all around the cell similar to the *Ter* sites. Of those five molecules, we found that two molecules of Tus form the highest intensity focus during cell growth, and they rarely co-localize with the *ter* region where the *TerBC* sites are located. Average whole cell concentration of Tus molecules was found to be around eight molecules at the start of the cell cycle and around thirteen molecules towards the end of the cell cycle, which is much lower than hundred molecules per cell reported in bulk experiments (19). Noticeable Tus-mYpet foci intensity dynamics were not observed to reveal the occurrence of fork-trap mechanism based replication termination in live cells.

## 5.4. Outlook

To increase the sample number in this study an improved method for cell-pole detection and analysis has to be used in the future experiments. Brightfield Z-axis imaging is a good option to detect cell-poles clearly in microfluidic channels but it has to be first tested for compatibility with the microscope's perfect focus system (PFS) during image acquisition. The scenario of Tus-mYpet foci formation due to Tus-*Ter* binding was the only consideration in this chapter. However, being a low-intensity signal, it is possible that these foci are Tus-mYpet molecules aggregating with itself or with other proteins in the cytoplasm. To resolve this, control experiments using Chromatin Immuno-Precipitation (ChIP) sequencing technique (23) and the helicase mimicking experiments performed on purified Tus proteins (17) can be used to test the specificity of the Tus-mYpet on the *Ter*-sites. The growth media used during experiments definitely have an influence on the protein expression and DNA replication rates. For example, EZ rich defined medium has demonstrated growth characteristics similar to that of the nutrient rich LB medium and has no auto-fluorescence making it ideal for fluorescence imaging (27). Such a medium could be used for further validation of our results. Newly developed fluorescent proteins with higher quantum yield and photostability than mYpet have been reported from recent studies like mNeonGFP (28) which can be fused with Tus (similar to BN1117 strain) and tested for reproducibility of our current results.

## 5.5. Materials and Methods

### 5.5.1. Strains and strain construction

All strains created in this study are derived from *E. coli* AB1157 (29). The chromosomally engineered strains used for this study were constructed by  $\lambda$ -Red recombination (30) or P1-phage transduction (31) as described in (24).

#### BN1117 (Tus-mYpet)

In order to tag native Tus proteins at the C-terminus using mYpet fluorescent protein through chromosome engineering, the stop codon of *tus* gene was replaced with *linker-mYpet* sequence along with *frt* sequences flanked kanamycin resistance cassette (*linker-mYpet-frt-kan-frt*). For this purpose, we used the template plasmid pRod61 (a gift from Prof. David Sherratt, University of Oxford, UK). The PCR product for  $\lambda$ -Red recombination was obtained using the following primers: forward primer 5'- *CGC AGC CGC TTC GTT TGA TCA TTC CAC GGC TGC ACC TGT ATG TTG CAG ATT CGG CTG GCT CCG CTG CTG G* and reverse primer 5'- *GAC AGC TGG GTA CGG CCA GAA CAG ATG GTC GGC AGT ATG AAA GCC GGG CGT TAG CAT ATG AAT ATC CTC CTT AG*.

The homologous ends are indicated in italics. In the *E. coli* genome, the stop codon is shared between the open reading frames (ORFs) of *tus* gene and *fumA* gene. The ORF of *fumA* gene was left undisturbed by introducing a stop codon sequence into the reverse primer (indicated in bold). The PCR product was gel-purified and ~700 ng of the linear DNA was used for electroporation of AB1157 cells overexpressing  $\lambda$ -Red proteins from pKD46 (30). The correct insertion of the fragment into the chromosome of the

resulting strain was assayed by PCR. The oligonucleotides used were 5'-TCG GCT GGC TCC GCT GCT GG and 5'-CAT ATG AAT ATC CTC CTT AG (**Figure S1A**). DNA sequencing (**Figure S1B**) confirmed the sequence of the *tus-mYpet* fusion gene in this strain.

#### **BN2116 (*Tus-mYpet:ter-mCerulean*)**

This strain was constructed by P1-phage transduction of the *tus-mYpet* fusion gene together with the adjacent *kan* gene from BN1117 strain into the recipient *ter*-Cerulean strain. The *ter*-mCerulean strain was already engineered with 240 *tetO* arrays inserted at 50 kb clockwise from the *dif* site in between the *TerBC* sites, as well as the chromosomally integrated chimeric gene *tetR-Cerulean* expressing the TetR-mCerulean proteins upon IPTG induction (4). The presence of the *tus-mYpet* gene fusion was verified using the oligonucleotides: 5'-TCG GCT GGC TCC GCT GCT GG and 5'-CAT ATG AAT ATC CTC CTT AG. The fusion was confirmed by DNA sequencing.

#### **BN2384 (*mCherry-DnaN:Tus-mYpet:ter-mCerulean*)**

In order to create this strain, the kanamycin resistance cassette from the BN2116 strain was removed via the FLP-*frt* based out-recombination using flippase enzyme expressed from the temperature sensitive pCP20 plasmid (30). The strain was used as a recipient in the subsequent P1-phage transduction step. The *mCherry-dnaN* fusion gene together with the adjacent *kan* gene from the donor BN1682 strain (25) was then transduced into the recipient strain. The presence of the *mCherry-dnaN* gene fusion was verified by PCR using the oligonucleotides: 5'- CGT TGG CAC CTA CCA GAA AG and 5'-ATG CCT GCC GTA AGA TCG AG. The fusion was confirmed by DNA sequencing.

#### **BN2386 (*6xHis-Tus-mYpet*)**

The *Tus-mYpet* gene from the BN1117 cells was amplified using primers 5'- TAG GAT CCG ATG GCG CGT TAC GAT CTC GT and 5'- TAA AGC TTT TAT TTG TAC AAT TCA TTC ATA CCC TCG. The PCR amplicon was digested with *HindIII* and *BamHI*, and ligated into the pRSETb expression vector (Thermo Fisher Scientific, Waltham, MA), resulting in a construct pRSETb-6xHis-*Tus-mYpet*. The cloning was confirmed by plasmid sequencing using the T7 promoter and terminator flanking regions. The pRSETb-6xHis-*Tus-mYpet* construct was successfully transformed into BL21pLysS cells (Promega Corporation, Madison, WI) to allow inducible protein expression of 6xHis-*Tus-mYpet*.

### **5.5.2. M9 growth medium used in experiments**

We prepared 1 L of M9 growth medium which contains 10.5 g/L of autoclaved M9 broth (Sigma-Aldrich, USA); 0.1mM of autoclaved CaCl<sub>2</sub> (Sigma-Aldrich, USA); 0.1mM of autoclaved MgSO<sub>4</sub> (J.T.Baker, Europe); 0.3% of filter-sterilized glycerol (Sigma-Aldrich, USA) as carbon source; 0.1 g/L of filter-sterilized 5 amino acids, namely L-threonine, L-leucine, L-proline, L-histidine and L-arginine (all from Sigma-Aldrich, Merck KGaA, Darmstadt, Germany) and 10 μL of 0.5% filter-sterilized Thiamine (Sigma-Aldrich, Merck KGaA, Darmstadt, Germany).

### 5.5.3. Microfluidics for long time-lapse microscopy

In order to perform long time-lapse experiments, we employed our own previously published design of the microfluidic device (called the *mother machine*) (32). The standardized procedure of device production has been described previously in Chapter 3 (25). In summary: cells were grown to log-phase (up to Optical Density ( $OD_{600}$ )  $\sim 0.3$  to  $0.4$ ) in M9 growth medium with appropriate antibiotics at  $37^\circ\text{C}$ . The induction was performed with  $0.1\text{ mM}$  IPTG when the OD reaches  $\sim 0.2$ . The culture was centrifuged for 1 min at  $16,100\text{ g}$ . The supernatant was decanted and the pellet was re-suspended in  $50\ \mu\text{L}$  M9 growth medium and injected into the microfluidic device. Then the device was centrifuged for 10 min at  $2,500\text{ g}$  (Eppendorf 5810R) in order to let the cells move into the channels. The device was then mounted on the microscope after attaching the tubing and incubated for  $\sim 45\text{ min}$  at  $37^\circ\text{C}$ . After incubation, the main trench of the device was cleared of cells using a syringe filled with fresh M9 growth medium containing the appropriate antibiotics,  $0.1\text{ mM}$  IPTG and  $0.2\text{ mg}\cdot\text{ml}^{-1}$  of bovine serum albumin (BSA). Then the syringe was attached to an automated syringe pump set to a rate  $\sim 0.5\text{ ml}\cdot\text{min}^{-1}$  in order to continuously infuse fresh M9 growth medium through the device.

### 5.5.4. Microscope setup

Images were acquired on a commercial Nikon Ti microscope equipped with a Nikon CFI Apo TIRF  $\times 100$ ,  $1.49\text{NA}$  oil immersion objective and an Andor iXon 897 Electron Multiplying Charge Coupled Device (EMCCD, Andor, Belfast UK) camera operated by a personal computer running Nikon NIS elements software. Brightfield illumination is done using the standard Nikon brightfield halogen lamp. Condenser components and the brightfield images were acquired using a separate EMCCD camera. The fluorescence excitation was performed using custom-built laser illumination. A Cobolt Fandango  $515\text{ nm}$  continuous wave diode-pumped solid-state laser was used to excite mYpet; Cobolt Jive  $561\text{ nm}$  continuous wave diode-pumped solid-state laser was used to excite mCherry. Cobolt Twist  $457\text{ nm}$  continuous wave diode-pumped solid-state laser was used to excite mCerulean. Laser beams were combined by dichroic mirrors (Chroma zt457dcrb, 575dcspxr) and were subsequently coupled into a single-mode optical fiber (KineFLEX). The output of the fiber was expanded and focused onto the back focal plane of the objective mounted on the microscope. Notch filters (Semrock, UK: NF03-514E and NF03-561E) were used to eliminate any laser light leaking onto the camera. The emission of the different fluorescent proteins was projected onto the central part of the EMCCD camera using custom filter sets (from Chroma, UK): Chroma z561, ET605/52m, zt561rdc (RFP), Chroma z514, ET540/30m, zt514rdc (YFP), Chroma z457/10x, ET490/40m, zt457rdc (CFP). A custom design commercial temperature control housing (Okolabs, Pozuolli NA Italy) enclosing the microscope body maintained the temperature at  $37^\circ\text{C}$ . Sample position was controlled with a Nikon stage (TI-S-ER Motorized Stage Encoded, MEC56100) together with the Nikon Perfect Focus System (PFS) to eliminate Z-drift during image acquisition.

### 5.5.5. Data acquisition for long time-lapse experiments

Image acquisition was performed with Nikon NIS-elements software. Cell outlines were imaged using standard brightfield illumination, and the brightfield images were acquired



every 2.5 min. Subsequently for *in vivo* imaging of Tus proteins linked with mYpet,  $\beta$  clamps tagged with mCherry, and *ter* sites labeled with mCerulean the sample was excited by 515 nm, 561 nm, and 457 nm lasers respectively with the settings detailed in **Table 1**. Data spanning ~10 h of measurement were acquired overnight.

### 5.5.6. Image analysis for long time-lapse experiments

Analysis of the acquired images from long time-lapse experiments was performed in two stages with custom-written MATLAB software (MathWorks Inc., Natick, MA). The first stage consists of steps, which are detailed, in our previous publication (25). In summary: First, the uneven background was subtracted from the fluorescence images using a rolling ball filter (33) and subsequently, heterogeneity in laser illumination was corrected by a previously determined laser beam profile (34). Next, brightfield and fluorescence images were aligned with respect to each other and the X-Y drift was corrected with ~1-pixel accuracy. In an automated fashion, a growth channel was analyzed and the brightfield images were used to determine the cell poles of all the cells in a given frame. For the  $\beta$ -mCherry fluorescence signal, a kymograph was constructed by summation of the pixel intensities per image perpendicular to the channel direction for each frame. We made use of the generated kymographs to determine individual replication and division cycles per cell. A post-processing step was then performed to eliminate cells that did not match the following selection criteria: correct cell length, sufficient growth characteristics, observation of a complete cell cycle, and a clear fluorescence signal of  $\beta$ -clamps that both starts and ends in a diffuse state.

The second stage of the image analysis involved the following steps: At first, the information on cell pole position (XY values) was collected from the brightfield images of the cell cycles selected from the previous stage. Then using the cell pole position, the cells were tracked by their edges and, the time series snap-shots of the corresponding cells' fluorescence images were saved automatically. We call these snapshots as 'bacpics' hereafter. Sample montages of these bacpics are shown in **Figures S2A, S2B, and S2C**. Bacpics that showed a clear signal of *ter*-mCerulean foci were then chosen for further analysis. After determining the position of *ter* region within the cell, the cell orientation was changed accordingly. Then, we used MATLAB to detect all the Tus-mYpet foci within the cells during the complete cell cycle in an automated manner.

The foci detection process was performed as follows: We first find the highest intensity parts (which we term as 'outliers' hereafter) of each cell images in 'bacpics' using a filtering algorithm as follows: The intensity histogram of the image was first fitted with a Gaussian distribution, and the regions with the upper limit of intensities which are  $\geq 2 \cdot SD$  were removed to be stored as an outlier. This process was done iteratively until the change in SD between the iterations is  $\leq 1\%$ . The image containing only the outliers was again fitted with the 2D-Gaussian distribution in order to detect all the 'foci' within a cell at a time. After detecting a focus, the image was smoothened using a Gaussian mask to calculate the integrated intensity, and then the focus was removed by superimposing the inverted Gaussian of the same focus. This process was repeated until all the foci above an intensity threshold value of 300 counts were detected. The whole cell integrated intensity was also calculated by summing the intensity values per pixel within the cell boundary of the image. After the foci detection was performed on all the images, the cell cycle time was

first normalized. The tracking and linking of the detected foci across the cell cycle was performed based on the literature (35). The intensity values collected from ‘baccpics’ were corrected for photobleaching similar to the method detailed in Chapter 3. All the collected data were then stored in a Microsoft Access database file and plots were made using MATLAB.

### 5.5.7. Single molecule calibration

The BN2386 strain was used for purification of His-tagged Tus-mYpet proteins. We employed the commercially available Capture™ His-Tagged Purification Miniprep Kit (Clontech, CA, USA). 5 mL of LB culture containing BN2386 cells was induced with 1 mM IPTG when the OD reached ~0.3. After ~1 hour of induction when the cells reached a log phase (OD ~0.6 to 0.8), the culture was centrifuged for 1 min at 16,100 g. The cell pellet was resuspended gently in 400  $\mu$ L of xTractor buffer provided with the kit for lysis. In the meantime Capture™ miniprep nickel column was equilibrated in ~400  $\mu$ L of xTractor buffer and centrifuged for 1 min at 11,000 g to collect the buffer. The lysate was added to the column and centrifuged for 1 min at 11,000 g. The column was washed using 300  $\mu$ L of wash buffer by centrifuging for 1 min at 11,000 g. Finally, his-tagged Tus-mYpet proteins were eluted using 300  $\mu$ L of elution buffer by centrifuging for 1 min at 11,000 g. All the steps here were performed at room temperature.

In order to determine the average intensity of a Tus-mYpet molecule under our experimental conditions, we used an approach similar to studies previously described (25, 36, 37). To summarize, purified single Tus-mYpet molecules were immobilized on a clean cover glass through conjugation to the anti-YPet antibody (Clontech, CA, USA) in M9 growth medium. Next, the single Tus-mYpet proteins were imaged until they were irreversibly photobleached using the same imaging conditions as used for the long time-lapse experiments (**Table 1**) except for a faster frame rate of ~3.3 frames/s which is equivalent to the camera shutter speed.

Fluorescence images containing the Tus-mYpet spots were analyzed manually with ImageJ software version 1.50e (38) using the following methodology. Firstly, we subtracted the uneven background using a rolling ball filter (33) and subsequently corrected for illumination heterogeneity by using the 515 nm laser beam profile (34). Using the region of interest (ROI) selection tool of the ImageJ, pixels containing a Tus-mYpet spot was selected in the image series. Here we selected standard 5x5 pixels containing the fluorescent spot (**Figure S2F**). We then used the ‘Plot Z-axis profile’ tool in the selected ROI. The output is the intensity profile of the fluorescent spot across the time series which is usually a bleaching step, and these values are saved as a tab-limited text file using the ‘save’ option. Using MATLAB, the data was fitted by a Heaviside step function through which the fluorescence intensity per pixel ( $I_p$ ) was calculated. The integrated intensity ( $I_s$ ) is 25 times  $I_p$  since here the number of pixels selected in the ROI is a constant (= 25). The integrated intensity values of the various Tus-mYpet spots calculated were plotted as a histogram and a normal distribution fitting was performed using MATLAB to calculate the mean integrated intensity value of a molecule of Tus-mYpet (**Figure 3C**).

## 5.6. Contributions

M. C. Moolman, S. Tiruvadi Krishnan, and N.H. Dekker designed the research and the experiments. S. Tiruvadi Krishnan undertook the experiments. R. de Leeuw and V. Lorent assisted during experiments and contributed to discussions. S. Tiruvadi Krishnan constructed strains. M. C. Moolman and J.W.J. Kerssemakers wrote software to analyze the microscopy data. R. de Leeuw, S. Tiruvadi Krishnan, and V. Lorent analyzed the data. S. Tiruvadi Krishnan and N.H. Dekker wrote the chapter.

## 5.7. References

1. T. A. Baker, S. H. Wickner, Genetics and Enzymology of DNA Replication in Escherichia Coli. *Annual Review of Genetics* **26**, 447-477 (1992).
2. L. Fang, M. J. Davey, M. O'Donnell, Replisome Assembly at oriC, the Replication Origin of E. coli, Reveals an Explanation for Initiation Sites outside an Origin. *Molecular Cell* **4**, 541-553 (1999).
3. D. M. Prescott, P. L. Kuempel, Bidirectional Replication of the Chromosome in Escherichia coli. *Proceedings of the National Academy of Sciences of the United States of America* **69**, 2842-2845 (1972).
4. R. Reyes-Lamothe, C. Possoz, O. Danilova, D. J. Sherratt, Independent Positioning and Action of Escherichia coli Replisomes in Live Cells. *Cell* **133**, 90-102 (2008).
5. C. Neylon, A. V. Kralicek, T. M. Hill, N. E. Dixon, Replication Termination in Escherichia coli: Structure and Antihelicase Activity of the Tus-Ter Complex. *Microbiology and Molecular Biology Reviews* **69**, 501-526 (2005).
6. N. Kono, K. Arakawa, M. Tomita, Validation of Bacterial Replication Termination Models Using Simulation of Genomic Mutations. *PLoS ONE* **7**, e34526 (2012).
7. M. D. Mulcair *et al.*, A Molecular Mousetrap Determines Polarity of Termination of DNA Replication in E. coli. *Cell* **125**, 1309-1319 (2006).
8. S. Mulugu *et al.*, Mechanism of termination of DNA replication of Escherichia coli involves helicase–conrahelicase interaction. *Proceedings of the National Academy of Sciences* **98**, 9569-9574 (2001).
9. T. M. Hill, Arrest of Bacterial DNA Replication. *Annual Review of Microbiology* **46**, 603-633 (1992).
10. M. J. J. Moreau, P. M. Schaeffer, Differential Tus-Ter binding and lock formation: implications for DNA replication termination in Escherichia coli. *Molecular BioSystems* **8**, 2783-2791 (2012).
11. K. Kamada, T. Horiuchi, K. Ohsumi, N. Shimamoto, K. Morikawa, Structure of a replication-terminator protein complexed with DNA. *Nature* **383**, 598-603 (1996).

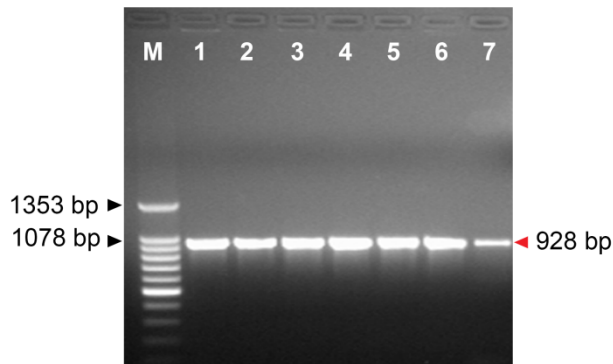
12. D. L. Kaplan, D. Bastia, Mechanisms of polar arrest of a replication fork. *Molecular Microbiology* **72**, 279-285 (2009).
13. M. Pandey *et al.*, Two mechanisms coordinate replication termination by the Escherichia coli Tus–Ter complex. *Nucleic Acids Research* **43**, 5924-5935 (2015).
14. H. Hendrickson, J. G. Lawrence, Mutational bias suggests that replication termination occurs near the dif site, not at Ter sites. *Molecular Microbiology* **64**, 42-56 (2007).
15. X. Wang, C. Lesterlin, R. Reyes-Lamothe, G. Ball, D. J. Sherratt, Replication and segregation of an Escherichia coli chromosome with two replication origins. *Proceedings of the National Academy of Sciences of the United States of America* **108**, E243-E250 (2011).
16. D. Ivanova *et al.*, Shaping the landscape of the Escherichia coli chromosome: replication-transcription encounters in cells with an ectopic replication origin. *Nucleic Acids Research*, (2015).
17. B. A. Berghuis *et al.*, Strand separation establishes a sustained lock at the Tus–Ter replication fork barrier. *Nat Chem Biol* **11**, 579-585 (2015).
18. P. A. Andersen, A. A. Griffiths, I. G. Duggin, R. G. Wake, Functional specificity of the replication fork-arrest complexes of Bacillus subtilis and Escherichia coli: significant specificity for Tus–Ter functioning in E. coli. *Molecular Microbiology* **36**, 1327-1335 (2000).
19. S. Natarajan, S. Kaul, A. Miron, D. Bastia, A 27 kd protein of E. coli promotes antitermination of replication *in vitro* at a sequence-specific replication terminus. *Cell* **72**, 113-120.
20. J. R. Wiśniewski, D. Rakus, Quantitative analysis of the Escherichia coli proteome. *Data in Brief* **1**, 7-11 (2014).
21. R. T. Dame, O. J. Kalmykova, D. C. Grainger, Chromosomal Macrod domains and Associated Proteins: Implications for DNA Organization and Replication in Gram Negative Bacteria. *PLoS Genetics* **7**, e1002123 (2011).
22. I. G. Duggin, S. D. Bell, Termination Structures in the Escherichia coli Chromosome Replication Fork Trap. *Journal of Molecular Biology* **387**, 532-539 (2009).
23. M. J. J. Moreau, Ph.D. thesis, James Cook University, Australia (2013).
24. S. Tiruvadi Krishnan, M. C. Moolman, T. van Laar, A. S. Meyer, N. H. Dekker, Essential validation methods for E. coli strains created by chromosome engineering. *Journal of Biological Engineering* **9**, 11 (2015).
25. M. C. Moolman *et al.*, Slow unloading leads to DNA-bound  $\beta$ 2-sliding clamp accumulation in live Escherichia coli cells. *Nat Commun* **5**, (2014).

26. X. Wang, C. Possoz, D. J. Sherratt, Dancing around the divisome: asymmetric chromosome segregation in *Escherichia coli*. *Genes & Development* **19**, 2367-2377 (2005).
27. F. C. Neidhardt, P. L. Bloch, D. F. Smith, Culture Medium for Enterobacteria. *Journal of Bacteriology* **119**, 736-747 (1974).
28. N. C. Shaner *et al.*, A bright monomeric green fluorescent protein derived from *Branchiostoma lanceolatum*. *Nat Meth* **10**, 407-409 (2013).
29. S. K. DeWitt, E. A. Adelberg, The Occurrence of a Genetic Transposition in a Strain of *Escherichia Coli*. *Genetics* **47**, 577-585 (1962).
30. K. A. Datsenko, B. L. Wanner, One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proceedings of the National Academy of Sciences of the United States of America* **97**, 6640-6645 (2000).
31. L. C. Thomason, N. Costantino, D. L. Court, in *Current Protocols in Molecular Biology*. (John Wiley & Sons, Inc., 2001).
32. M. C. Moolman, Z. Huang, S. T. Krishnan, J. W. J. Kerssemakers, N. H. Dekker, Electron beam fabrication of a microfluidic device for studying submicron-scale bacteria. *Journal of Nanobiotechnology* **11**, 12-12 (2013).
33. S. R. Sternberg, Biomedical Image Processing. *Computer* **16**, 22-34 (1983).
34. Y. Taniguchi *et al.*, Quantifying *E. coli* Proteome and Transcriptome with Single-Molecule Sensitivity in Single Cells. *Science* **329**, 533-538 (2010).
35. M. C. Moolman, Jacob W. J. Kerssemakers, Nynke H. Dekker, Quantitative Analysis of Intracellular Fluorescent Foci in Live Bacteria. *Biophysical Journal* **109**, 883-891 (2015).
36. M. C. Leake *et al.*, Stoichiometry and turnover in single, functioning membrane protein complexes. *Nature* **443**, 355-358 (2006).
37. R. Reyes-Lamothe, D. J. Sherratt, M. C. Leake, Stoichiometry and Architecture of Active DNA Replication Machinery in *Escherichia coli*. *Science* **328**, 498-501 (2010).
38. C. A. Schneider, W. S. Rasband, K. W. Eliceiri, NIH Image to ImageJ: 25 years of image analysis. *Nat Meth* **9**, 671-675 (2012).
39. D. Grünwald, S. M. Shenoy, S. Burke, R. H. Singer, Calibrating excitation light fluxes for quantitative light microscopy in cell biology. *Nature protocols* **3**, 1809-1814 (2008).

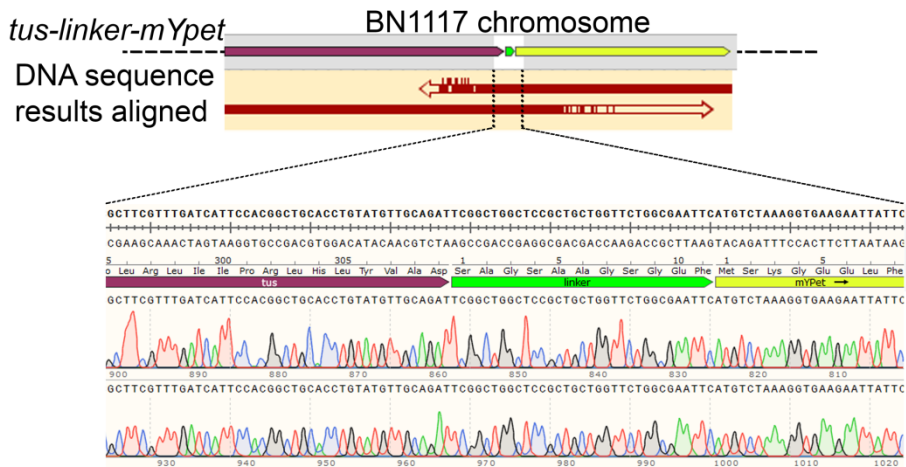
## 5.8. Supplementary Figures and Captions

### 5.8.1. Supplementary Figure S1: Strain validation results

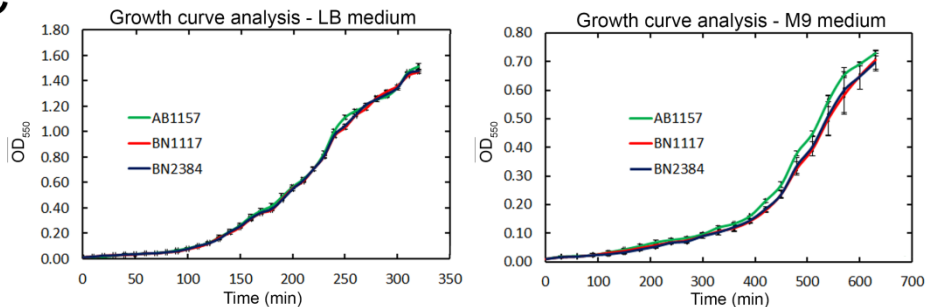
A



B

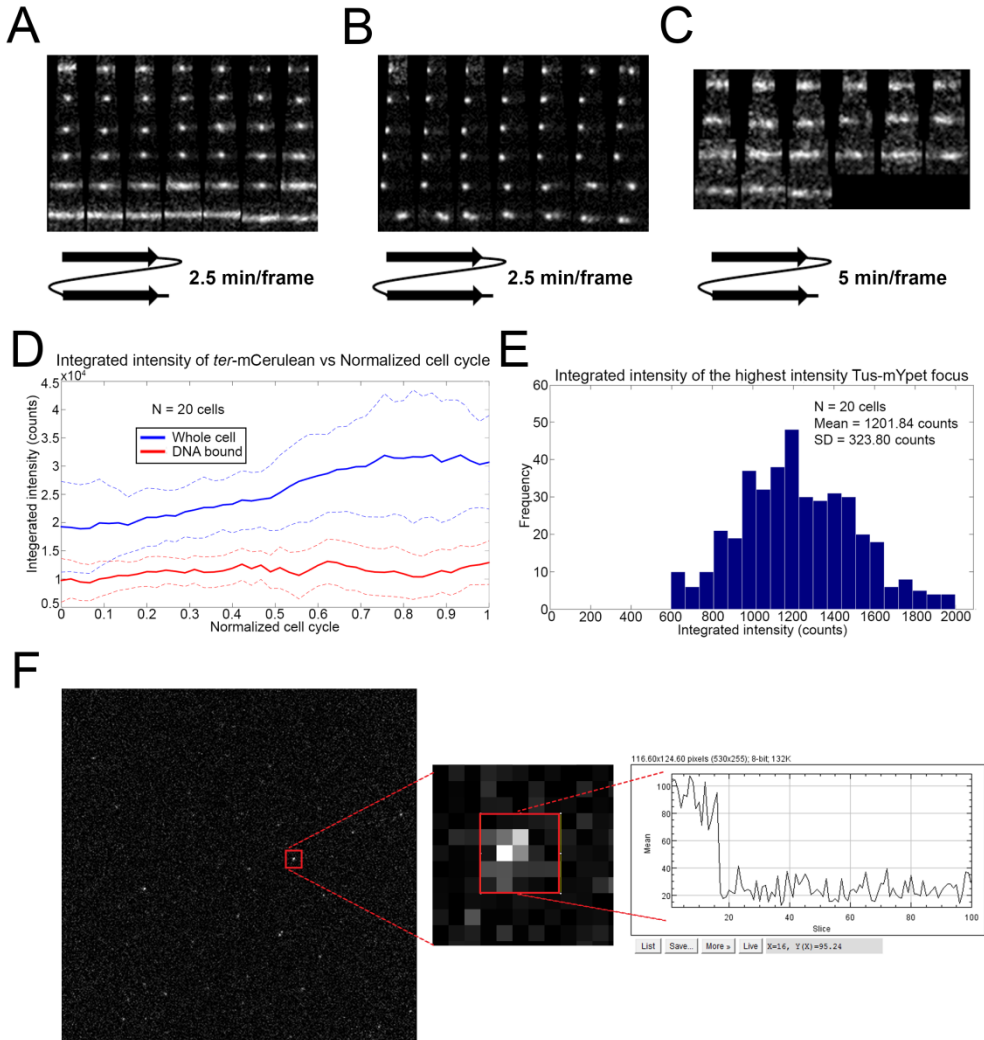


C



**A.** Colony PCR results of the  $\lambda$ -Red recombination technique while creating the BN1117 strain is shown. The primers used, specifically amplify from the middle of *tus* gene and the middle of the mYpet sequence. In lane M, 10  $\mu$ L of DNA ladder was loaded (relevant band sizes are marked using black arrowheads). Lanes 1 to 7 contain the colony PCR products which show the positive band size of 928 bp marked using red arrow head. **B.** DNA sequencing and alignment results of the BN1117 (Tus-mYpet) strain. The red arrows show the aligned sequences and the directionality of the corresponding DNA sequencing reaction. The integration site of *linker-mYpet* sequence and chromatogram result of the sequenced DNA is shown. **C.** Mean growth curves of the wild-type (AB1157), BN1117 and BN2384 (mCherry-DnaN:Tus-mYpet:*ter*-mCerulean) strains in the LB medium (left) and M9 medium (right) are shown.

### 5.8.2. Supplementary Figure S2: Analysis of the results from long time-lapse live cell imaging experiments



**A.** An example montage sequence of the mCherry-DnaN fluorescence signal during one complete cell cycle of BN2384 (mCherry-DnaN:Tus-mYpet:*ter*-mCerulean) strain. Image acquisition time (2.5 min/frame) and the directionality of the montage sequence are shown below. **B.** The montage sequence of the *ter*-mCerulean fluorescence signal during the complete cell cycle (same as panel A) of BN2384 strain. Image acquisition time (2.5 min/frame) and the directionality of the montage sequence are shown below. **C.** The montage sequence of the Tus-mYpet fluorescence signal during the complete cell cycle (same as panel A) of BN2384 strain. Image acquisition time (5 min/frame) and the directionality of the montage sequence are shown below. **D.** Integrated intensities of *ter*-



mCerulean fluorescence are plotted against normalized cell cycle indicated as 0 and 1 for the start and end of bacterial division cycle. Whole cell fluorescence intensities are indicated as blue lines and the foci intensities are shown as red lines. Dotted lines show the corresponding upper and lower limits of the standard error of mean values. **E.** The distribution of the integrated intensity of the highest intensity Tus-mYpet focus is shown in this histogram plot. **F.** Using ImageJ software, background subtracted and illumination corrected tiff stacked time-series images (left) are zoomed into the extent of seeing a pixelated spot. A standard 5x5-pixelated area that shows signals much higher than the background are selected as a region of interest (center). Plot Z-axis profile tool gives an output of a bleaching step (right) the values of which are used to calculate the integrated intensity of a Tus-mYpet protein.

## **6 CHROMOSOMAL ENGINEERING OF *dnaG* GENE FOR THE LABELING OF *ESCHERICHIA COLI* PRIMASE PROTEIN *IN VIVO***

*Semi-conservative replication of the E. coli chromosome is coordinated by the replisome, a complex consisting of 13 different enzymes. The leading strand is synthesized continuously, whereas the lagging strand is synthesized discontinuously throughout the cell cycle. DnaG, one of the subunits of the replisome, is the primase enzyme that synthesizes both the RNA primer required for initiation of the leading strand replication as well as the numerous RNA primers that support lagging strand replication. Using in vitro single-molecule approaches, it has been shown that primases are highly dynamic during replication and interact with numerous components of the replisome. However, their corresponding in vivo dynamics have not been studied in detail. Using current advances in genetic engineering, we have employed different approaches to label primase proteins in vivo with a bright fluorescent protein like YPet with the goal of elucidating primase dynamics in E. coli. Attempts to directly label the native primase proteins with the YPet fluorophore in E. coli were unsuccessful, possibly a result of post-transcriptional modifications involved in the primase expression from a native macromolecular synthesis operon. Hence, we undertook an alternative approach that involved deletion of the endogenous primase gene together with the introduction of an ectopic gene under the control of an inducible promoter. In this study, we describe three different strategies designed to knock out the native primase gene: direct ORF knockout, non-invasive site-directed mutagenesis, and indirect FLP/frt based knockout. Unfortunately, all of these strategies yielded only false positive results. From this, we conclude that improved chromosomal engineering techniques with nearly 100% efficiency and specificity in recombination might be necessary to knock out the *dnaG* gene. Ectopic insertion of the labeled primase sequence into the chromosome and knockout of the native primase will be attempted with the newly improved CRISPR-Cas9 based genome editing tools in future.*

## 6.1. Introduction

The semi-conservative replication of the *E. coli* chromosome proceeds in a continuous manner on the leading strand, and in a semi-discontinuous fashion on the lagging strand (1, 2). This asymmetry in leading versus lagging strand synthesis results from the directionality of the DNA polymerase III holoenzyme (PolIII), which can add complementary nucleotides only to the 3'-OH group of the deoxyribose sugar in a sticky end of the DNA (2). The semi-discontinuous replication of the lagging strand is aided by a specific enzyme primase (transcribed from the gene *dnaG* in *E. coli*) that synthesizes short RNA primers on the lagging strand, allowing the PolIII complex to continue the DNA replication at the 3'-end (3). These discontinuously synthesized DNA fragments on the lagging strand are called Okazaki fragments. The RNA-DNA primed regions are converted to double-stranded DNA by a different DNA polymerase (PolI), and the nicks between Okazaki fragments are joined by DNA ligase (2) (**Figure 1A**).

*E. coli* DnaG consists of three functional regions: an N-terminal zinc-finger domain, which binds to DNA, an oligoribonucleotide synthesis domain similar to those present in RNA polymerase, and a C-terminal helicase-interacting domain (4). The helicase-interacting region has been shown to interact with single-strand binding proteins (SSB) at a replication fork, thereby competing with  $\chi$ -subunit of replisome (5). DnaG is expressed, along with the 30S ribosome subunit S21 (RpsU) and  $\sigma$ 70 subunit of RNA polymerase (RpoD), from a macromolecular synthesis (MMS) operon (*rpsU-dnaG-rpoD*) (6) (**Figure 1B**). Similar to many operons involved in the regulation of cell growth, the MMS operon is under intricate regulatory control and is affected by many metabolic pathways (7). The polycistronic mRNA that is expressed by the MMS operon is processed and modified in order to express different levels of RpsU, DnaG, and RpoD proteins. One such modification, which is well studied using biochemical tools, is the cleavage performed by RNaseE on the *rpsU-dnaG-rpoD* mRNA at the intergenic site between the *dnaG* and *rpoD* ORFs. Subsequently, the *dnaG* part of mRNA is degraded in order to reduce the number of DnaG proteins (8). Additional studies should reveal further details of the post-transcriptional modifications occurring on the *dnaG* transcript expressed by the MMS operon.

The cellular stoichiometry of primases may play a crucial role in cell viability since the primases are postulated to interact with many replisome components in a concentration-dependent manner (4). One such interaction, between helicase and primases, has been well studied *in vitro*. For example, it has been shown that for each dimer of the helicase protein, a single primase is required for optimal primer synthesis (9). Therefore, since DnaB helicase forms a hexamer at the replication fork, minimally three primases are assumed to be required at the fork (9). Recent studies have also shown that SSB proteins bind to the C-terminal helicase-binding domain of primase and displace primase from the lagging strand (5, 10). *In vitro*-single-molecule studies of DNA replication fork dynamics have shown that leading strand synthesis rate is three-fold slower in the presence of primase than in its absence (11). When two or three primases cooperatively bind to the helicase hexamer, this binding weakens helicase's interaction with the  $\tau$ -subunit of replisome (11). This co-operative binding of primases leads to the replisome disassembly from leading strand and the continuation of lagging strand synthesis thereby the primases possibly improve the coordination between the leading and lagging strand synthesis (11, 12). It has

the additional effect of reducing the number of nucleotides added by PolIII and thus reducing the overall processivity. It remains unknown whether these phenomena also occur inside the cells.

Another protein that is affected by the stoichiometry of primases at the replication fork is the  $\beta$  sliding clamp protein. This protein is known to coordinate the interactions of various enzymes during the lagging strand synthesis (more information on the  $\beta$  sliding clamp proteins is presented in Chapter 3). Okazaki fragment synthesis on the lagging strand has been shown *in vitro* to occur even in the absence of free  $\beta$ -clamp proteins, which implies that free  $\beta$  proteins are not required during the replication process and that the  $\beta$ -clamp proteins can be actively recycled within the replisome itself (13). In the same study, it was shown that changing the primase concentration has a reciprocal effect on the length of Okazaki fragments, i.e. reducing the primase concentration in the system leads to longer Okazaki fragments, possibly due to fewer RNA primers bound on the lagging strand (13). As discussed in Chapter 3, we observed in *E. coli* cells that  $\beta$ -clamp proteins accumulate during DNA replication and are actively unloaded and recycled after approximately 4 s (14). It is hypothesized in *E. coli* and observed in *B. subtilis* strains that the accumulation and loading of  $\beta$ -clamp proteins are dependent on the length of Okazaki fragments, which is tightly coupled to the number of primase proteins expressed in the cells (14, 15).

To date, relatively little information is available regarding the *E. coli* primase proteins *in vivo*; for example, no data is yet available on the number of primases that are required to actively synthesize primers which might answer further questions like whether the stoichiometry of primases at the replication fork changes over the course of the cell cycle. Similarly, it is still unknown how the *in vivo* protein-protein interactions of primases occur: whether they function independently or in consensus with the other replisome proteins. In an attempt to study the dynamics and stoichiometry of DnaG *in vivo*, researchers have previously attempted to label fluorescently the native primase at its C-terminus in order to study its localization and dynamics in cells. However, the resulting strains were not viable (16, 17).

In this study, we attempted several labeling approaches in order to study the dynamics and stoichiometry of primases on replication forks *in vivo*. We first attempted to label the native DnaG protein with YPet at the N-terminus using the counter-selection based scarless two-step  $\lambda$ -Red recombination method. This approach was not successful, and we present a possible explanation for this failure. We then designed an alternative strategy in which we tried to introduce ectopically the YPet labeled DnaG under an arabinose-inducible promoter (pBAD-YPet-DnaG) in the *E. coli* genome. Simultaneously, we used three different approaches to knock out the endogenous primase gene ( $\Delta$ dnaG) in *E. coli*. By combining the deletion of endogenous *dnaG* gene and the inducible pBAD-YPet-DnaG, in a single strain, one should be able to vary the concentration of the labeled primase proteins in order to change the lengths of Okazaki fragments *in vivo* and observe its effects on the dynamics of labeled replisome components (e.g.  $\beta$ -clamp protein). In the end of the chapter we provide additional suggestions for manipulation and ultimately, successfully labeling, the *dnaG* gene in the *E. coli* chromosome.

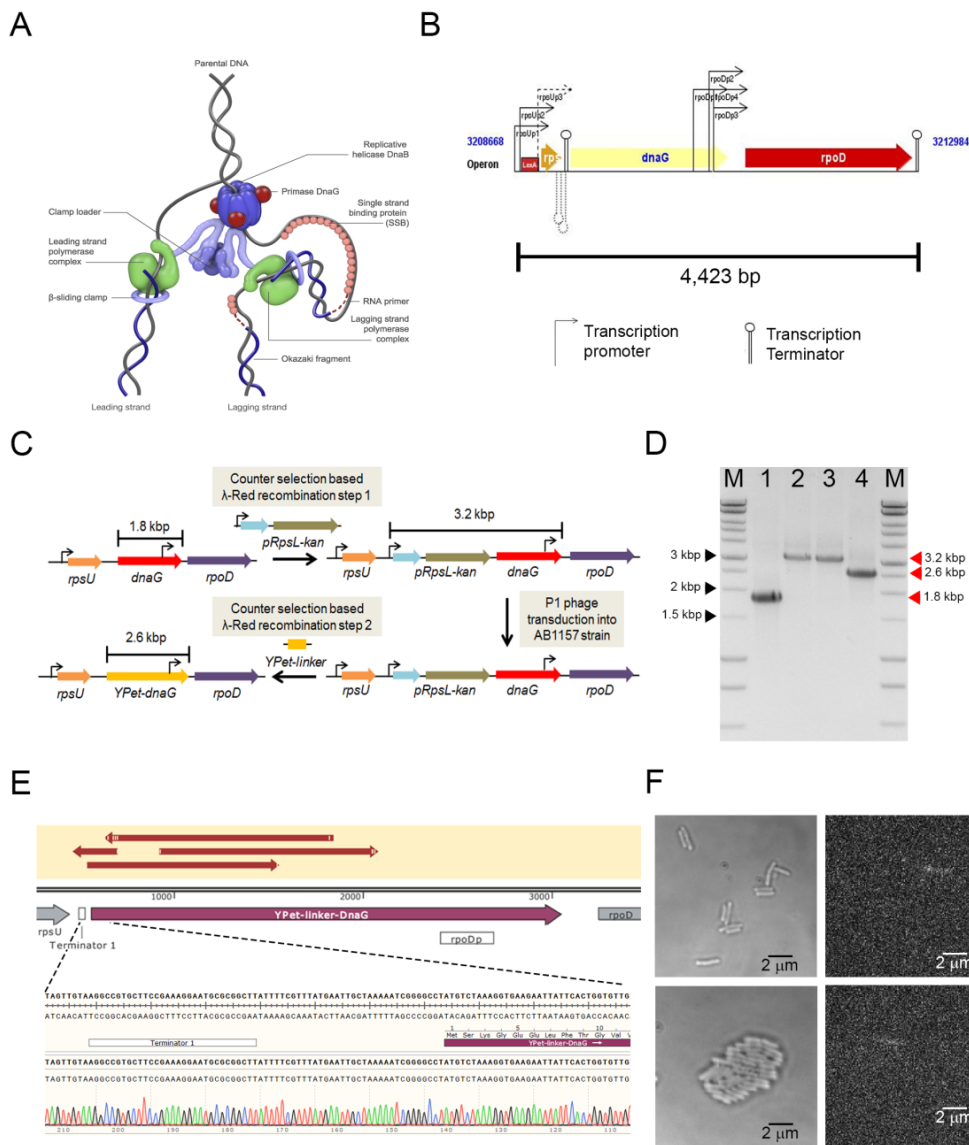
## 6.2. Results and Discussion

### 6.2.1. N-terminus fusion of native DnaG with YPet expresses a fluorescently dark protein

Whenever the endogenous gene needs to be modified for labeling the proteins *in vivo*, the details of its operons and the presence of any overlapping promoters or operator sites for the expression of neighboring genes have to be carefully considered. In the *dnaG* gene, the sequences upstream to its 5'-end contain the *rpoDp* promoter, which is responsible for the expression of  $\sigma 70$  subunits from the *rpoD* gene in response to heat shock (18). The presence of this promoter may be the reason for previously failed attempts to fuse YPet fluorescent protein to the C-terminus of DnaG (16, 18). Hence, in this study, we performed an N-terminal fusion of DnaG with YPet in AB1157 cells using a scarless counter-selection based chromosome modification process as detailed in the **Materials and Methods** (Figure 1C). After verifying the created strain BN2117 (YPet-DnaG) using PCR and DNA sequencing (Figure 1D and E), we tested the strain using widefield fluorescence microscopy. The microscopy results were negative: no fluorescence was detected from the cells when excited at the YPet excitation wavelength of 515 nm (Figure 1F).

To understand the reason for the observation of no fluorescence, we checked the transcription of the *YPet-dnaG* gene in the BN2117 strain using q-PCR. We found no obvious difference in the mRNA expression levels for the genes *dnaG*, *rpoD* and *rpsU* between the BN2117 and the parental strains, BN2112 and AB1157 (Figure 2A). From this observation, we concluded that the YPet insertion did not affect the transcription of the *dnaG* gene. Unfortunately, antibodies specific for *E. coli* DnaG are not commercially available, so we could not check the protein expression level for DnaG in these cells. However, from the growth curve analysis of the YPet-DnaG strain in LB medium, we found using *t*-statistics with 95% confidence intervals no significant difference in the growth characteristics of the YPet-DnaG strain (generation time:  $38.4 \pm 1.5$  min) and the parental AB1157 strain (generation time:  $38.9 \pm 1.4$  min). This result suggests that DnaG protein (which is an essential protein for the survival of *E. coli* (19)) is also expressed at functional levels. We performed the growth curve analysis as described in Chapter 2 (20).

To determine whether the open reading frame (ORF) of the YPet-DnaG fusion protein was intact, we cloned the ORF sequence into a vector containing the arabinose-inducible promoter (*pBAD*) as described in **Materials and Methods**. When the TOPO-pBad-YPet-DnaG vector was introduced into strain AB1157, the cloned plasmid was found to express YPet-DnaG fusion protein in the cell upon arabinose induction (Figure 2B). In other words, YPet-DnaG was found to be fluorescent when expressed from plasmids, but not when expressed from the native MMS operon. A possible explanation for these observations is that the mRNA transcript (*rpsU-YPetDnaG-rpoD*) from the MMS operon of BN2117 strain may fold in such a way that the ribosome binding site is altered, as a result of which, a truncated and fluorescently dark YPet-DnaG fusion protein containing a functional primase region appears to be expressed. The result also shows that the linker sequence and its length used for fusing the YPet and DnaG are not the reasons for the expression of fluorescently dark YPet-DnaG fusion protein.



**Figure 1: Scarless N-terminal fusion of native DnaG with YPet is non-fluorescent in cells. A.** The diagram shows various components of *E. coli* replisome replicating the parental DNA into two sister strands: leading and lagging strands. Helicase bound primases are depicted as red spheres. **B.** MMS operon in an *E. coli* genome is shown along with its transcriptional promoters and terminators (Image courtesy: [www.ecocyc.org](http://www.ecocyc.org)). **C.** Scheme depicts the steps involved in labeling the native primase with YPet-linker fusion by chromosome engineering. The 3-step chromosome engineering involves scarless and marker-less insertion of *YPet-linker* coding sequence in the place of the start codon (5'-ATG-3') of the *dnaG* gene. **D.** PCR results of the strains successfully created during each step as shown in panel C (expected product sizes are marked using red arrowheads). The primers used anneals at the region flanking *dnaG* gene as depicted in panel A. In lane M, 10  $\mu$ L of DNA ladder was

loaded (relevant band sizes are marked using black arrowheads). Lane 1 contains PCR product obtained from wild type AB1157 strain (~1.8 kbp). Lane 2 and 3 contains the PCR products obtained from *pRpsL-kan* sequence inserted at 20bp upstream to *dnaG* gene through  $\lambda$ -Red recombination and P1 phage transduction steps respectively (~3.2 kbp). Lane 4 contains the PCR product obtained from the BN2117 (YPet-DnaG) strain (2.6 kbp). **E.** DNA sequencing and alignment results of YPet-DnaG strain. The red arrows show the aligned sequences and the directionality of the DNA sequencing reaction. The integration site of YPet-linker sequence and chromatogram result of the sequenced DNA is shown. **F.** The top-left image shows the brightfield image of BN2117 (YPet-DnaG) strain grown in M9-glycerol medium and top-right image shows the corresponding fluorescence image (laser excitation at 515 nm). The bottom-left and right images show the brightfield and fluorescence images of wild-type AB1157 grown in M9-glycerol medium respectively.

We note for certain experiments, the use of labeled primase expressed from a plasmid may be perfectly fine. However, for our purpose the use of TOPO-pBad-YPet-DnaG plasmid is insufficient. Even with a low copy number (~10 – 20/cell), the plasmids hinder the fluorescence imaging of primases *in vivo* because the labeled primases bind to the plasmids along with the *E. coli* chromosome for DNA replication. This makes it difficult to distinguish the fluorescence signals from both the binding events (**Figure 2B**). Hence, we continued our efforts to obtain a chromosomal fusion.

### 6.2.2. Ectopic chromosomal insertion to control the *in vivo* expression of DnaG

We attempted to insert the *pBad-YPet-DnaG* sequence from the TOPO-pBad-YPet-DnaG plasmid into the AB1157 chromosome at the non-essential *galK* locus as described in the **Materials and Methods**. However, when we tried to confirm proper integration of the exogenous DNA, sequence results showed that the inserted fragment was not integrated at the *galK* locus. A possible explanation for the mispositioning of the integration is the strong homology between the *pBad* region of the insert sequence and the native arabinose promoter in the AB1157 strain. However, after many attempts and careful screening we did succeed in cloning the *pBad-DnaG* gene without YPet into the *galK* locus of AB1157 cells, as described in the **Materials and Methods** (**Figure 2C**). Similar to the *pBad-DnaG* insertion more trials and screening of the false positive colonies may be required for the successful *pBad-YPet-DnaG* insertion. In the meantime, we worked on various strategies with the aim of knocking out the endogenous *dnaG* gene in *E. coli*.

### 6.2.3. Various strategies attempted to knock out the endogenous *dnaG* gene in *E. coli* ( $\Delta$ *dnaG*)

Previous studies have attempted to knock out the endogenous *dnaG* gene. The Keio collection of ORF knocked-out *E. coli* strains to determine the essentiality of genes in the chromosome for cells' survival at first reported that *dnaG* as a non-essential gene by deleting the complete ORF (21). However, three years later, an update on the Keio collection stated that the previously reported *dnaG* knockout was a false positive strain due to gene duplication errors during  $\lambda$ -Red recombination, and further attempts of deletion were unsuccessful (19). These results substantiate the essentiality of the *dnaG* ORF. Hence, in this study in order to knock out the endogenous *dnaG* gene we used the BN1871 (pBad-DnaG) strain cultured in the presence of 0.2% arabinose for ectopic expression of DnaG.

We attempted three different strategies to perform the *dnaG* knockout. They are detailed below in ascending order of the number of steps involved.

### Strategy 1 - Direct knockout of native *dnaG* ORF

Owing to the complexity of the MMS operon, a direct ORF knockout of the full *dnaG* sequence might have implications for RpoD gene expression since a stress-regulated promoter *rpoDp* is present within the *dnaG* gene (18). Hence, we decided to delete only the sequence of *dnaG* upstream of the *rpoDp* promoter along with the full *dnaG* ORF knockout. We performed the  $\lambda$ -Red recombination for both full and partial *dnaG* knockouts as detailed in **Materials and Methods** and summarized in **Figure 3A**. Upon testing of the resulting colonies using PCR, we observed that neither the full nor the partial knockout methods were successful, and the native *dnaG* ORF remained intact. The reason why we obtained false positive colonies could be due to extraneous insertions that could arise during  $\lambda$ -Red recombination, as detailed in Chapter 2 (20).

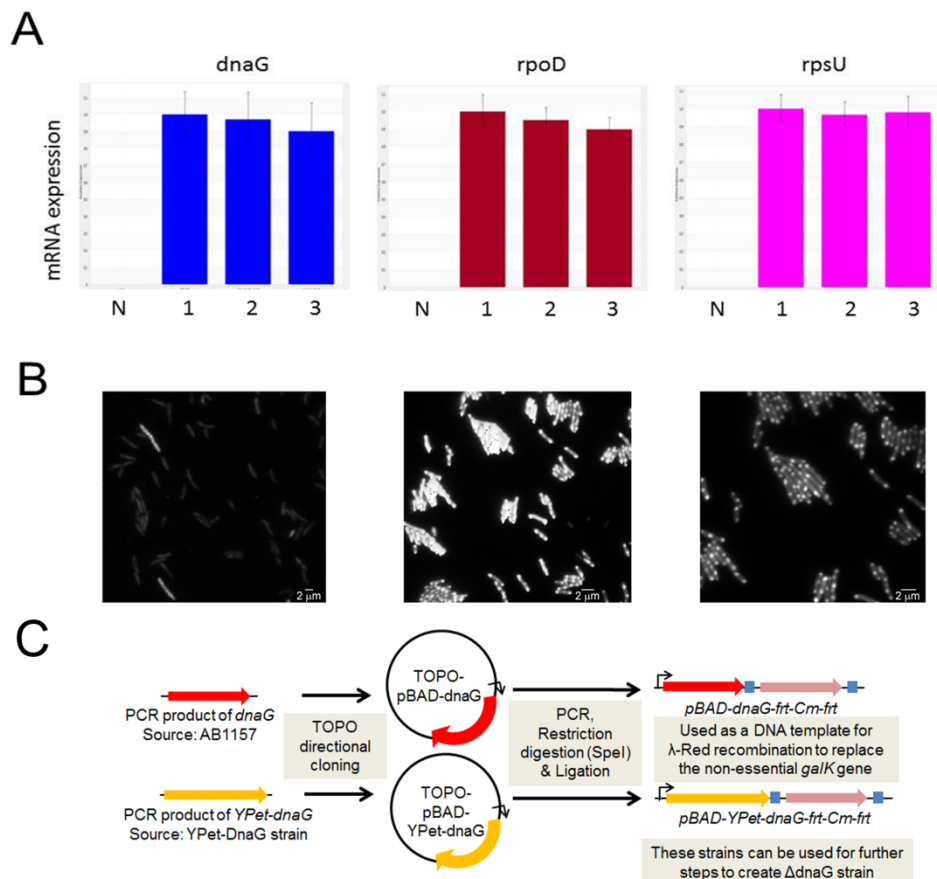
### Strategy 2 - Site-directed mutagenesis for incomplete translation of native *dnaG*

Deletion of a part of the *dnaG* gene might also affect mRNA stability and as such, expression of the *rpoD* and *rpsU* genes. Therefore, we attempted to perform the *dnaG* knockout in a non-invasive manner for the MMS operon. Using the counter selection based two-step  $\lambda$ -Red recombination technique we tried to introduce a point mutation G<sup>893</sup>→T that leads to a premature stop codon in the *dnaG* gene. Due to this mutation, a truncated catalytically inactive primase will be expressed while the expression of the other genes in MMS operon will not be affected. If the technique is successful, the cells will depend on the ectopically expressed intact DnaG (BN1871 strain) under arabinose induction. We performed this mutagenesis in the AB1157 strain as detailed in **Materials and Methods** (**Figure 3B**). However, the resulting eight colonies were found to be false positive (i.e., the *pRpsL-neo-dnaG* region was intact) upon PCR validation. One possible reason for the false positive results might be spontaneous point mutations occurring within the *pRpsL-neo* cassette, as described in literature (22).

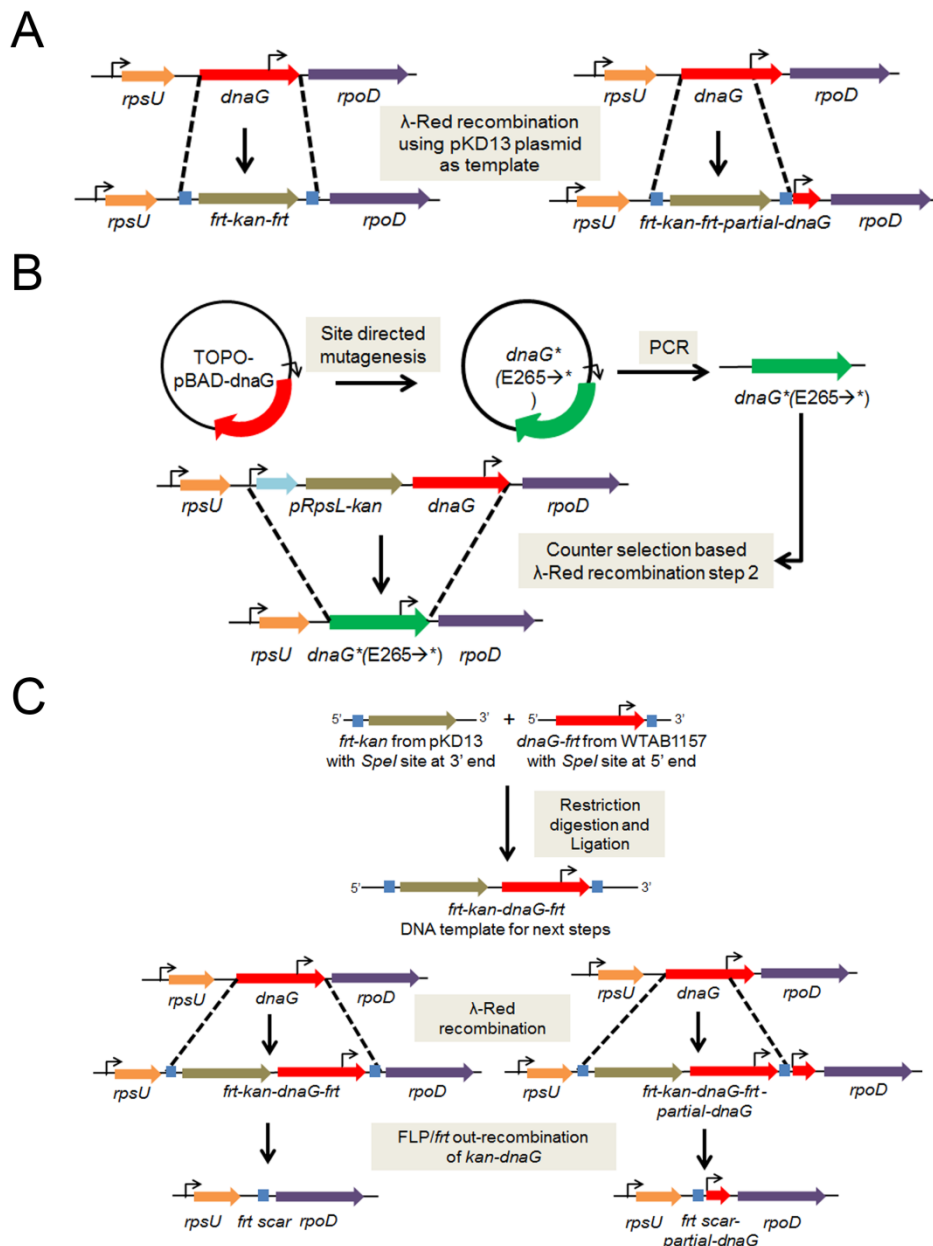
### Strategy 3 - FLP/*frt* sites based out-recombination of native *dnaG*

Due to the false positive results of the above-mentioned strategies, we attempted the out-recombination based knockout method (23). Using  $\lambda$ -Red recombination technique the native (full-length) *dnaG* ORF and a partial *dnaG* ORF (only the sequence upstream to the *rpoDp* promoter) sequences are flanked by the *frt* sites along with kanamycin cassette in the upstream region (**Figure 3C**). If these strains are successfully created, the insert can be transferred into the BN1871 strain using P1 phage transduction and then the *frt* flanked *dnaG* region can be recombined out using FLP/*frt* based knockout system. After performing the  $\lambda$ -Red recombination step as detailed in **Materials and Methods**, we found by sequencing that none of the colonies that were successfully validated using PCR contained an integrated downstream *frt* sequence. One of the possible reasons for the non-insertion of the downstream *frt* sequence in the native full *dnaG* ORF is that it might be deleterious for the expression of RpoD by changing the ribosome-binding site (18). However, the reason for the non-insertion of the downstream *frt* sequence in the partial *dnaG* ORF is unknown and needs further investigation.





**Figure 2: YPet-DnaG fusion protein is fluorescently bright when the ORF expressed from a plasmid. A.** q-PCR results to verify the mRNA expression levels of the 3 genes expressed under *rpsU-dnaG-rpoD* operon from following strains: (1) wild-type AB1157, (2) BN2112 strain (RpsL-Kan-DnaG) and (3) BN2117 (YPet-DnaG). q-PCR results of nuclease free water are referred as N. **B.** The fluorescence images of BN2385 strain (strain containing TOPO-pBAD-YPet-DnaG plasmid) grown in M9-glycerol medium for non-induced cells (left) or induced with 1% Arabinose exposed to 1 mW (middle) and 0.25 mW (right) of 515 nm laser power. **C.** The scheme depicts the steps involved in the creation of strains in which the expression of primase and labeled primase (YPet-DnaG) inside the cells can be controlled using arabinose induction.



**Figure 3: Various strategies attempted to knock out the native primase gene in *E. coli* ( $\Delta$ *dnaG*).** **A.** Steps involved in the direct knockout of native *dnaG* ORF in the pBad-DnaG strain are shown. **B.** A non-invasive knockout strategy based on site-directed mutagenesis for incomplete translation of native *dnaG* and the steps involved are described. **C.** An indirect knockout method based on FLP/*firt* out-recombination of native *dnaG* and, the steps involved are depicted.

### 6.3. Conclusions

In this study, we have summarized the challenges involved in chromosomally engineering the native primase gene using the currently available genome editing techniques in *E. coli*. We used a scarless two-step  $\lambda$ -Red recombination method to label the native primases with YPet. We did succeed in integrating the YPet sequence fused to the *dnaG* gene. From our validation experiments, we observed that the fusion expressed fluorescently dark YPet-DnaG which may be due to the complex post-transcriptional modifications of the polycistronic mRNA. We also attempted an indirect approach to express in an inducible manner the labeled or unlabeled primases from an ectopic region of the chromosome and succeeded in creating a strain expressing unlabeled primases ectopically. Then, we tried knocking out the native *dnaG* gene in this strain, using three different approaches: simple direct knockout, site-directed mutagenesis, and out-recombination of the *dnaG* gene. All the three approaches yielded many false positive colonies and further investigations are necessary to attempt the partial out-recombination of the *dnaG* gene.

### 6.4. Outlook

Firstly, since the *pBad-DnaG* insertion at the *galK* locus has worked, it is reasonable to perform more attempts and careful screening for the insertion of a *pBad-YPet-DnaG* sequence in the chromosome. Secondly, of the three strategies detailed above for knocking out the native primase gene, partial knockout using FLP/*prt* based out-recombination seems promising for re-trials since the recombination was successful but the sequencing results showed that the downstream *prt* site is absent. This error could have arisen during the cloning steps for preparing the template DNA. One method to avoid this error is to create the template DNA as a plasmid instead of the linear DNA. Even though it is a time-consuming process, using a plasmid is easier to precisely verify the template DNA sequence and circumvent any cross-contamination of partially digested or ligated DNA products. A newly developed genome-editing tool for *E. coli* using the type II CRISPR-Cas9 from *Streptococcus pyogenes* has been reported to have 100% efficiency (24, 25). This tool can be used as an alternative method or in combination with the  $\lambda$ -Red recombinase proteins in order to insert the *pBad-YPet-DnaG* sequence into the chromosome and knock out the native primase with any of the three strategies detailed here. A different fluorescent protein instead of YPet may be used to tag native primases using the methodology described in this chapter since the mRNA of such a fusion may fold in such a way to express a fluorescent primase. Newly developed fluorescent proteins like mNeonGFP (26) were described to have better quantum yield and photostability than YPet, hence they may be used as alternative fluorophores.

### 6.5. Materials and Methods

#### 6.5.1. Strains and strain construction

All strains created in this study are derived from *E. coli* AB1157 (27). The chromosomally engineered strains used for this study were constructed by  $\lambda$ -Red recombination (23) or counter selection based  $\lambda$ -Red recombination (22) or P1-phage transduction (28) as described in (20).

**BN2112 (RpsL-KanR-DnaG)**

This strain was created during the first step of counter selection based  $\lambda$ -Red recombination method involving the *pRpsL-kan* selection cassette (22). The *pRpsL-kan* cassette was generated by PCR using primers with homologous ends, which are indicated in italics: 5'-*TTG TAA GGC CGT GCT TCC GAA AGG AAT GCG CGG CTT ATT TTC GTT TAT GAG GCC TGG TGA TGA TGG CGG G* and 5'-*AAT GAA TAC GCG TGG GAT TCG TCC AGC CAT AGG CCC CGA TTT TTA GCA ATT CAG AAG AAC TCG TCA AGA A*. This cassette was integrated into the genome of AB1157 cells at the 20 bp upstream region of the start codon of the *dnaG* gene. The successful strain was selected for kanamycin resistance and tested for streptomycin sensitivity. PCR and DNA sequencing further verified the strain. Next, the *pRpsL-kan-dnaG* sequence was transferred to a clean wildtype AB1157 strain via P1 phage transduction to remove the extraneous insertions occurred during the recombinering process which was verified by Southern blotting as described previously in Chapter 2 (20).

**BN2117 (YPet-DnaG)**

This strain was constructed from the BN2112 strain from the second step of counter selection based  $\lambda$ -Red recombination method. The *pRpsL-kan* cassette and the start codon of the *dnaG* gene were replaced by the *YPet-linker* coding sequence from pROD44 (29) through  $\lambda$ -Red recombination by selecting for colonies that are resistant to streptomycin and sensitive to kanamycin. This *YPet-linker* sequence was generated by PCR with plasmid pROD44 as template DNA (24) and primers which contain 70 bases of the homologous end before the start codon of *dnaG* and 50 bases of the homologous end after the start codon (indicated in italics): 5'-*TTG TAA GGC CGT GCT TCC GAA AGG AAT GCG CGG CTT ATT TTC GTT TAT GAA TTG CTA AAA ATC GGG GCC TAT GTC TAA AGG TGA AGA ATT ATT CAC TGG TGT TG* and 5'-*TCA GTG CGT GCC AGC AGA TCA TTA ATG AAT ACG CGT GGG ATT CGT CCA GCC GCG CTG CCA GAA CCA G*. The YPet-DnaG strain was verified by PCR and DNA sequencing.

**BN2385 (TOPO-pBad-YPet-DnaG)**

The ORF sequence of YPet-DnaG fusion protein was isolated from the chromosomal DNA of BN2117 (YPet-DnaG) strain by PCR using the primers: 5'-*CAC CAT GTC TAA AGG TGA AGA ATT ATT CAC TGG T* and 5'-*TCA CTT TTT CGC CAG CTC CTG GTT TAA TGT* and cloned into the TOPO-pBAD202 directional vector (Invitrogen, USA). Obtained colonies (TOP10 strain background) were verified using PCR and DNA sequencing. The verified plasmid was then transformed into electrocompetent *E. coli* AB1157 cells to obtain the final TOPO-pBad-YPet-DnaG strain.

**BN2387 (pBad-YPet-DnaG)**

The arabinose inducible *YPet-DnaG* sequence (*pBad-YPet-DnaG*) from the TOPO-pBad-YPet-DnaG plasmid (see the above section BN2385) was amplified by PCR using the primers 5'-*GGG CGG CGC TGG AAT TCA AAC CAA TTG TCC ATA TTG C* and 5'-*GGG CGG GAG GGA CTA GTG TTT AAA CTC AAT GGT GAT G*, resulting in the presence of a *SpeI* restriction site at 3'end. Similarly, chloramphenicol resistance cassette was cloned from the pKD3 plasmid using the primers 5'-*GGA GCA CAC TAG TAG TGT*

AGG CTG GAG CTG CTT CG and 5'-GTA TCG CCG GGA TCC ATG GGA ATT AGC CAT GGT CC, containing a *SpeI* restriction site at the 5' end. The PCR products were then digested using *SpeI* enzyme and subsequently ligated together. The linear template DNA *pBad-YPet-DnaG-CmR* sequence was then purified from agarose gel using a Wizzard kit (Promega Corporation, Madison, WI). PCR was performed on the purified *pBad-YPet-DnaG-CmR* fragment using the primers containing 50 bases of homologous ends for the *galK* locus (5'-*CGG CTG ACC ATC GGG TGC CAG TGC GGG AGT TTC GTT CAG CAC TGT CCT GCA AAC CAA TTG TCC ATA TTG C* and 5'-*GTT TGC GCG CAG TCA GCG ATA TCC ATT TTC GCG AAT CCG GAG TGT AAG AAT CCG CCA AAA CAG CCA AGC T*). This fragment was integrated into AB1157 genomic DNA by  $\lambda$ -Red recombination and, the pBad-YPet-DnaG strain was selected using chloramphenicol resistance. The strain was then validated using PCR.

### **BN1840 (TOPO-pBad-DnaG)**

The ORF sequence of the *dnaG* gene was amplified from AB1157 genomic DNA by PCR using primers: 5'-*CAC CAT GGC TGG ACG AAT CCC ACG CGT ATT CAT T* and 5'-*TCA CTT TTT CGC CAG CTC CTG GTT TAA TGT* and cloned into a TOPO-pBAD202 directional vector (from Invitrogen, USA). The plasmid was then transformed into chemically competent TOP10 cells (Invitrogen, USA). Kanamycin-resistant colonies were then characterized using PCR and DNA sequencing. The verified plasmid was transformed into electrocompetent *E. coli* AB1157 cells to obtain the final TOPO-pBad-DnaG strain.

### **BN1871 (pBad-DnaG)**

The arabinose inducible *dnaG* sequence (*pBad-DnaG*) from the TOPO-pBad-DnaG plasmid (see the above section BN1840) was amplified by PCR using the primers 5'-GGG CGG CGC TGG AAT TCA AAC CAA TTG TCC ATA TTG C and 5'-GGG CGG GAG GGA *CTA GTG* TTT AAA CTC AAT GGT GAT G, resulting in a *SpeI* restriction site (in italics) at 3'end. Similarly, chloramphenicol resistance cassette is cloned from the pKD3 plasmid using the primers 5'-GGA GCA *CAC TAG TAG* TGT AGG CTG GAG CTG CTT CG and 5'-GTA TCG CCG GGA TCC ATG GGA ATT AGC CAT GGT CC, containing a *SpeI* restriction site (in italics) at 5' end. The PCR products were then digested with *SpeI* ligated together. The linear template DNA *pBad-DnaG-CmR* was then gel-purified.

Using the pBad-DnaG-CmR template DNA, PCR amplification was performed using primers containing 50 bases of homologous ends (in italics) for the *galK* locus (5'-*CGG CTG ACC ATC GGG TGC CAG TGC GGG AGT TTC GTT CAG CAC TGT CCT GCA AAC CAA TTG TCC ATA TTG C* and 5'-*GTT TGC GCG CAG TCA GCG ATA TCC ATT TTC GCG AAT CCG GAG TGT AAG AAT CCG CCA AAA CAG CCA AGC T*). The obtained fragment was integrated into the genome of AB1157 cells by  $\lambda$ -Red recombination and the pBad-DnaG strain was selected for chloramphenicol resistance. Positive clones were validated by PCR, DNA sequencing and Southern blotting (20). All strains showed extraneous insertions in Southern blotting and the final pBad-DnaG strain was obtained after removing the extraneous insertions using P1 phage transduction to AB1157 strain as described previously in Chapter 2 (20).

### $\Delta$ DnaG (Strategy 1)

ORF knockout of full *dnaG* and partial *dnaG* (deleting only the sequence upstream to the *rpoDp* promoter) were attempted using the  $\lambda$ -Red recombination technique in the pBad-DnaG strain grown in the presence of 0.2% arabinose. A kanamycin cassette was amplified from pKD13 plasmid using the following PCR primers with 50 bp of homologous ends (in italics) corresponding to the full or partial knockout of *dnaG*: Forward primer: 5'- *AAG GAA TGC GCG GCT TAT TTT CGT TTA TGA ATT GCT AAA AAT CGG GGC CTT GTA GGC TGG AGC TGC TTC G*, reverse primer for full *dnaG* knockout: 5'- *TGC GGC TGT CGG GGG CTT CCC GAT CGC TCT TCG GCA CTT AAG CCG TTA AAA TTC CGG GGA TCC GTC GAC C* and, reverse primer for partial *dnaG* knockout: 5'- *TAC GCA TGG TCG TGC GTT TTA GCT GCG GAA CAG GGC GAG AAA CGC CGC TCA TTC CGG GGA TCC GTC GAC C* (**Figure 3A**). The  $\lambda$ -Red recombination technique was performed and, the colonies obtained from kanamycin with 0.2% arabinose-containing plates were tested for proper integration by PCR.

### $\Delta$ DnaG (Strategy 2)

In this strategy, we planned to introduce a point mutation ( $G^{893} \rightarrow T$ ) that inserts a stop codon before the catalytic core of the enzyme and thereby translating a truncated catalytically inactive primase. We used the TOPO-pBad-DnaG plasmid as a template and introduced the intended point mutation ( $G^{893} \rightarrow T$ ) in the *dnaG* ORF by site-directed mutagenesis and primers: forward 5'- *CCAA TCG TCT GCT TGT GGT CTA AGG CTA TAT GGA CGT GGT GGC* and reverse 5'- *GCC ACC ACG TCC ATA TAG CCT TAG ACC ACA AGC AGA CGA TTG G*. The plasmid was verified by PCR and DNA sequencing. In parallel using P1 phage transduction, the *pRpsL-kan-dnaG* mutation from the donor strain (pRpsL-kan-dnaG strain described in the YPet-DnaG section) was transduced into a clean recipient strain pBad-DnaG. The successfully transduced strain was verified for streptomycin sensitivity and resistance for kanamycin and chloramphenicol.

The DnaG\* plasmid bearing the point mutation ( $G^{893} \rightarrow T$ ) was then used as the DNA template for the second step of the counter-selection based  $\lambda$ -Red recombination in the strain containing both the *pRpsL-kan-dnaG* and *pBad-DnaG::galK* mutations (**Figure 3B**). The primers used that contain 50 bases of homologous ends (in italics) were forward 5'- *TGC GCG GCT TAT TTT CGT TTA TGA ATT GCT AAA AAT CGG GGC CTA TGG CTG GAC GAA TCC CAC GCG TAT T* and reverse 5'- *TGC GGC TGT CGG GGG CTT CCC GAT CGC TCT TCG GCA CTT AAG CCG TTA AAT CAC TTT TTC GCC AGC TCC T*. The colonies were cultured in the presence of 0.2% arabinose and were selected for streptomycin resistance and kanamycin.

### $\Delta$ DnaG (Strategy 3)

In order to delete *dnaG* ORF completely and partially, we cloned the *frt-kan* sequence with a *SpeI* site (indicated in italics) at the 3' end from the pKD13 plasmid using the PCR primers: forward 5'- *ATTCGGGGATCCGTCGACC* and reverse 5'- *GGC TCC CAC TAG TAA GAT CCC CTT ATT AGA AGA*. In parallel, the *dnaG* ORF with a *SpeI* site (indicated in italics) at the 5' end and *frt* site (indicated in bold) at the 3' end was cloned from AB1157 genome using the PCR primers: forward 5'- *GAG GTC GAC TAG TAT GGC TGG ACG AAT CCC ACG* and reverse 5'- *TGT GTA GGC TGG AGC TGC TTC*

**GAA GTT CCT ATA CTT TCT AGA GAA TAG GAA CTT CGG AAT AGG AAC TTC TCA CTT TTT CGC CAG CTC CTG.** The PCR products were purified, digested with SpeI and ligated. The desired ligated product of length ~3 kbp was gel purified. This linear DNA was then used as a template for  $\lambda$ -Red recombination (**Figure 3C**). The primers used for full *dnaG* knockout were (50 bases of homologous ends were indicated in italics): forward 5'- *AAG GAA TGC GCG GCT TAT TTT CGT TTA TGA ATT GCT AAA AAT CGG GGC CTA TTC CGG GGA TCC GTC GAC C* and reverse 5'- *TGC GGC TGT CGG GGG CTT CCC GAT CGC TCT TCG GCA CTT AAG CCG TTA AAT GTG TAG GCT GGA GCT GCT TC*.

The primers used for partial *dnaG* knockout were (50 bases of homologous ends were indicated in italics): forward 5'- *AAG GAA TGC GCG GCT TAT TTT CGT TTA TGA ATT GCT AAA AAT CGG GGC CTA TTC CGG GGA TCC GTC GAC -3'* and reverse 5'- *TAC GCA TGG TCG TGC GTT TTA GCT GCG GAA CAG GGC GAG AAA CGC CGC TCT GTG TAG GCT GGA GCT GCT TC -3'*. The colonies obtained were verified by PCR.

### 6.5.2. qPCR verification for BN2117 strain (YPet-DnaG)

Total RNA was isolated from exponentially growing *E. coli* strain AB1157 and its derivatives BN2112 (pRpsL-kan-dnaG) and BN2117 (YPet-dnaG) using the SV total RNA isolation system (Promega Corporation, Madison, WI) as directed by the manufacturer.

cDNA synthesis and qPCR was performed in triplicate on the Illumina Eco (Illumina, Inc., San Diego, CA) using the Power SYBR Green RNA to CT 1-step kit (Applied Biosystems/Life Technologies). For each reaction, 100 ng of total RNA was used. The primer sequences used for the cDNA and qPCR reaction are presented in **Table 1**. Expression of the target genes in BN2112 and BN2117 cells were measured relative to the expression in AB1157 wild-type cells. As reference gene, *cysG* was measured. Data analysis was performed using EcoStudy v5.0.4890 software (Illumina, Inc., San Diego, CA).

**Table 1: Sequences of the oligonucleotides that contribute to the structure of the hairpin**

Oligo name	Sequence
cysG, forward	5'- CGC TTA ACG GTG AAT GCA TTA G
cysG, reversed	5'- CCT TCG ACG AGG GTT AAC ATG
dnaG, forward	5'- ATA GGG TTG CTG GTG CAA AAT C
dnaG, reversed	5'- CCA GGG AGC TTA TTT TCA TCC A
rpoD, forward	5'- AAA GTG CTG AAG ATC GCC AAA
rpoD, reversed	5'- AAT CCC CCA GAT GCG AAT C

rpsU, forward	5'- TCG TCG TCG TGA GTT CTA TGA AAA
rpsU, reversed	5'- TCT TCG CGT GAC GTT TCA CT

### 6.5.3. M9 growth medium used in experiments

1 L of M9 growth medium used in the experiments contains 10.5 g/L of autoclaved M9 broth (Sigma-Aldrich, USA); 0.1mM of autoclaved CaCl<sub>2</sub> (Sigma-Aldrich, USA); 0.1mM of autoclaved MgSO<sub>4</sub> (J.T.Baker, Europe); 0.3% of filter-sterilized glycerol (Sigma-Aldrich, USA) as carbon source; 0.1 g/L of filter-sterilized 5 amino acids, namely L-threonine, L-leucine, L-proline, L-histidine and L-arginine (all from Sigma-Aldrich, USA) and 10 µL of 0.5% filter-sterilized Thiamine (Sigma-Aldrich, USA).

### 6.5.4. Microscopy

Cells used for microscopy were grown in M9-glycerol supplemented with the necessary nutrients until an OD~ 0.2 was reached. The cells were immobilized for imaging on agarose pad as described previously (20). All microscopy experiments were conducted on a Nikon-Ti microscope with custom laser excitation. YPet was imaged using a 515nm (Cobolt Fandango) laser. Image acquisition was performed with EMCCD (Andor, Belfast UK) using commercial Nikon NIS elements software. The microscope body, including the sample, was kept at ~37°C using a temperature controller (Okolabs, Pozuolli NA Italy). Image analysis was performed using ImageJ.

## 6.6. Contributions

S. Tiruvadi Krishnan and N.H. Dekker designed the research and the experiments. S. Tiruvadi Krishnan undertook the experiments. T. van Laar assisted during experiments. T. van Laar and R. Janissen contributed to discussions. S. Tiruvadi Krishnan and T. van Laar constructed strains. S. Tiruvadi Krishnan and N.H. Dekker wrote the chapter.

## 6.7. References

1. M. Meselson, F. W. Stahl, The replication of DNA in Escherichia coli. *Proceedings of the National Academy of Sciences of the United States of America* **44**, 671-682 (1958).
2. T. Ogawa, T. Okazaki, Discontinuous DNA Replication. *Annual Review of Biochemistry* **49**, 421-457 (1980).
3. J. P. Bouché, K. Zechel, A. Kornberg, dnaG gene product, a rifampicin-resistant RNA polymerase, initiates the conversion of a single-stranded coliphage DNA to its duplex replicative form. *Journal of Biological Chemistry* **250**, 5995-6001 (1975).
4. D. N. Frick, C. C. Richardson, DNA Primases. *Annual Review of Biochemistry* **70**, 39-80 (2001).



5. N. Naue, M. Beerbaum, A. Bogutzki, P. Schmieder, U. Curth, The helicase-binding domain of Escherichia coli DnaG primase interacts with the highly conserved C-terminal region of single-stranded DNA-binding protein. *Nucleic Acids Research* **41**, 4507-4517 (2013).
6. J. R. Lupski, G. Nigel Godson, The rpsU-dnaG-rpoD macromolecular synthesis operon of E. coli. *Cell* **39**, 251-252 (1984).
7. J. Lupski, A. Ruiz, G. N. Godson, Promotion, termination, and anti-termination in the rpsU-dnaG-rpoD macromolecular synthesis operon of E. coli K-12. *Mol Gen Genet* **195**, 391-401 (1984).
8. V. Yajnik, G. N. Godson, Selective decay of Escherichia coli dnaG messenger RNA is initiated by RNase E. *Journal of Biological Chemistry* **268**, 13253-13260 (1993).
9. A. V. Mitkova, S. M. Khopde, S. B. Biswas, Mechanism and Stoichiometry of Interaction of DnaG Primase with DnaB Helicase of Escherichia coli in RNA Primer Synthesis. *Journal of Biological Chemistry* **278**, 52253-52261 (2003).
10. R. E. Georgescu, N. Yao, C. Indiani, O. Yurieva, M. E. O'Donnell, Replisome mechanics: lagging strand events that influence speed and processivity. *Nucleic Acids Research* **42**, 6497-6510 (2014).
11. N. A. Tanner *et al.*, Single-molecule studies of fork dynamics in Escherichia coli DNA replication. *Nat Struct Mol Biol* **15**, 170-176 (2008).
12. R. E. Georgescu, N. Y. Yao, M. O'Donnell, Single-molecule analysis of the Escherichia coli replisome and use of clamps to bypass replication barriers. *FEBS Letters* **584**, 2596-2605 (2010).
13. N. A. Tanner *et al.*, E. coli DNA replication in the absence of free  $\beta$  clamps. *The EMBO Journal* **30**, 1830-1840 (2011).
14. M. C. Moolman *et al.*, Slow unloading leads to DNA-bound  $\beta$ 2-sliding clamp accumulation in live Escherichia coli cells. *Nat Commun* **5**, (2014).
15. M. Su'tsugu, J. Errington, The Replicase Sliding Clamp Dynamically Accumulates behind Progressing Replication Forks in Bacillus subtilis Cells. *Molecular Cell* **41**, 720-732 (2011).
16. R. Reyes-Lamothe, C. Possoz, O. Danilova, D. J. Sherratt, Independent Positioning and Action of Escherichia coli Replisomes in Live Cells. *Cell* **133**, 90-102 (2008).
17. X. S. Xie, P. J. Choi, G.-W. Li, N. K. Lee, G. Lia, Single-Molecule Approach to Molecular Biology in Living Bacterial Cells. *Annual Review of Biophysics* **37**, 417-444 (2008).
18. M. Nesin, J. R. Lupski, G. N. Godson, Role of the 5' upstream sequence and tandem promoters in regulation of the rpsU-dnaG-rpoD macromolecular synthesis operon. *Journal of Bacteriology* **170**, 5759-5764 (1988).

19. N. Yamamoto *et al.*, Update on the Keio collection of Escherichia coli single-gene deletion mutants. *Molecular Systems Biology* **5**, n/a-n/a (2009).
20. S. Tiruvadi Krishnan, M. C. Moolman, T. van Laar, A. S. Meyer, N. H. Dekker, Essential validation methods for E. coli strains created by chromosome engineering. *Journal of Biological Engineering* **9**, 11 (2015).
21. T. Baba *et al.*, Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular Systems Biology* **2**, 2006.0008-2006.0008 (2006).
22. R. Heermann, T. Zeppenfeld, K. Jung, Simple generation of site-directed point mutations in the Escherichia coli chromosome using Red(®)/ET(®) Recombination. *Microbial Cell Factories* **7**, 14-14 (2008).
23. K. A. Datsenko, B. L. Wanner, One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. *Proceedings of the National Academy of Sciences of the United States of America* **97**, 6640-6645 (2000).
24. W. Jiang, D. Bikard, D. Cox, F. Zhang, L. A. Marraffini, CRISPR-assisted editing of bacterial genomes. *Nature biotechnology* **31**, 233-239 (2013).
25. Y. Jiang *et al.*, Multigene Editing in the Escherichia coli Genome via the CRISPR-Cas9 System. *Applied and Environmental Microbiology* **81**, 2506-2514 (2015).
26. N. C. Shaner *et al.*, A bright monomeric green fluorescent protein derived from Branchiostoma lanceolatum. *Nat Meth* **10**, 407-409 (2013).
27. S. K. DeWitt, E. A. Adelberg, The Occurrence of a Genetic Transposition in a Strain of Escherichia Coli. *Genetics* **47**, 577-585 (1962).
28. L. C. Thomason, N. Costantino, D. L. Court, in *Current Protocols in Molecular Biology*. (John Wiley & Sons, Inc., 2001).
29. R. Reyes-Lamothe, D. J. Sherratt, M. C. Leake, Stoichiometry and architecture of active DNA replication machinery in Escherichia coli. *Science (New York, N.Y.)* **328**, 498-501 (2010).



## SUMMARY

All living organisms pass on their genetic information to their offspring in the form of DNA or RNA molecules by duplicating them across generations. In the bacteria, their genes are packed in long chains of DNA molecules or chromosomes. One of the widely studied model organisms, *Escherichia coli*, replicates its circular chromosome in two directions starting from an origin region of chromosome with independent replication complexes or replisomes simultaneously synthesizing the daughter chromosomes. DNA replication is an important process of the *E. coli* life cycle because of occurrence of small errors in its mechanism will affect the cell's normal state larger. Much of our current knowledge about the dynamics of replisome complex has been obtained from *in vitro* experiments. However, the natural environment of the cell is considerably different from that of *in vitro* solutions.

Current advances in genetic engineering and fluorescence imaging enables us to image proteins *in vivo*. Single-molecule techniques provide one with added insights into the events that take place in the cell during replication, and they usually are not apparent while ensemble averaging used in bulk experiments. It is thus essential that quantitative single-molecule research is conducted before we can fully understand the details of DNA replication as it occurs in a living cell. Single-molecule imaging *in vivo* is aided by employing microfluidics in order to have spatial control over growing cells and provide precisely regulated growth conditions like temperature, media flow, etc. Similarly, the development of the accurate and automated image analysis of detected fluorescence signal has become close to reality now in spite of the heterogeneity and crowding effects in dividing cells. At the same time, the correct labeling of endogenous proteins with fluorescent proteins *in vivo* can be achieved through latest genome editing tools. In this thesis, we discuss the combination of these techniques to achieve single molecule imaging of endogenous proteins to gain insights on the DNA replication mechanism.

Chromosome engineering methods currently use homologous recombination-based techniques to modify the chromosomal DNA in *E. coli* that have application in both industrial and fundamental research. These methods are employed in this study to currently label the endogenous proteins by fusing the fluorescent protein coding sequence with the native gene of interest. However, proper validation of the created strain before using it for the research is necessary and the essential methods for such validation are discussed with example experiments in **Chapter 2**. Modification of chromosomal DNA is currently performed using the recombineering technique and the fusion of modified sequences in single *E. coli* strain is done by employing the P1-phage transduction. Beyond the standard PCR and DNA sequence analysis methods for validating the created strain, we discuss the essential verification by Southern blotting for detecting the extraneous insertions that may occur in recombineering. These unwanted mutations can be removed by transducing the gene of interest into wildtype strain using virulent P1 phages. However, phage stock used can become contaminated with temperate phages in which case they exist as a large plasmid in growing cells resulting in contaminated strains. We then describe assays to detect the presence of temperate phages in the created strain using the cross-streak agar and Evans blue-uranine agar plating. We also provide ways to detect cell growth and cell morphology defects and a stepwise flow of validating the created strain.

In order to utilize the techniques we have developed, we began by studying the  $\beta_2$  sliding clamp proteins in growing cells for an in-depth understanding of the DNA replication and repair since  $\beta_2$ -clamps play an important role in both the processes. Being one of the main components of replisome complex  $\beta_2$ -clamps, ensure high accuracy of the DNA replication and proper regulation of the processivity to coordinate the timely replication before cell division depending on the growth conditions. In **Chapter 3**, a quantitative *in vivo* investigation of the  $\beta_2$ -clamp dynamics is presented. Using the widefield and photoactivable fluorescence microscopy, we identified the number of  $\beta_2$ -clamps bound to DNA and their effective loading and unloading rates. We found from our observations that  $\beta_2$ -clamps accumulate on DNA after initiation of replication until the DNA-bound  $\beta_2$ -clamps attains a steady state that is maintained during most of the replication. This accumulation of  $\beta_2$ -clamps is due to the relatively slow unloading rate of  $\beta_2$ -clamps from DNA in comparison to the other replisome components. Through our single molecule imaging experiments we observed that each  $\beta_2$ -clamp is bound to DNA in the orders of minutes *in vivo*. Our results show the versatility of the  $\beta_2$ -clamp proteins in recruiting different proteins to the DNA to complete the replication process with high accuracy as it is involved in various processes of DNA replication and repair.

The replisome components at the fork of DNA replication face various protein barriers that are bound to DNA and surpass them to complete the DNA replication. In **Chapter 4**, we discuss the outcome of a replisome encountering the endogenous Tus-*Ter* roadblock of *E. coli* strain with ectopic origin *in vivo*. The replisomes initiated from the ectopic origin progress in such a way that the clockwise (CW) fork approaches the non-permissive side of the Tus-*Ter* roadblock considerably earlier than the counter-clockwise fork. From the results, we found that the replisomes that are detected by the labeled  $\beta_2$ -clamps continue binding to DNA even after the CW fork faces the Tus-*TerC* barrier. By observing the replication times of the ectopic origin strains we determine that the speed of replication decreases when the fork meets the Tus-*Ter* roadblocks while it is not stalled forever. We found the presence of Tus-*Ter* barriers influence the sister chromosome alignment (SCA) patterns. From these results, we demonstrate the flexibility of the replisomes at the replication fork to the Tus-*Ter* roadblocks and the sensitivity of the SCA patterns to the progression of the forks.

Polar arrest mechanism of replication forks by the Tus-*Ter* roadblocks might answer the important question of how and where the termination of chromosome replication occurs in *E. coli*. In **Chapter 5**, we attempt to ascertain the localization and stoichiometry of native Tus proteins during the cell cycle using the quantitative widefield fluorescence microscopy and microfluidics. From our initial observations, we found that Tus is expressed in low copy numbers in the order of eight to thirteen molecules per growing cell, of which around five molecules form foci (signal above the background) through possibly binding to *Ter* site. We also observe that the intensity of the foci do not change much during the cycle implicating that the Tus is constantly bound to *Ter* sites in the chromosome. Due to the low fluorescence signal of labeled Tus, no observable binding or unbinding events was detected in the cells. The labeled Tus foci were found to be distributed from cell poles to mid-cell similar to the expected localization of *Ter* sites. Our results depict an overview of the intricate details of the Tus-*Ter* dynamics *in vivo* during the cell cycle and aid in the design of future experiments for further studies on the mechanism of replication termination.

The replisome components synthesize the chromosome in a semi-conservative fashion with the forks of replication formed during the initiation phase using the RNA primers synthesized on the DNA strands by the primase. During the elongation phase, the leading strand of the replication fork is synthesized continuously while the lagging strand is synthesized discontinuously with the help of primase. In order to understand the *in vivo* dynamics of primases during the cell cycle, we attempted to label fluorescently the *E. coli* primase proteins that are detailed in **Chapter 6**. Using the scar-less recombineering technique we successfully engineered the native primase gene fused with the fluorescent protein coding sequence. However, the fusion protein is non-fluorescent due to unknown post-translational modifications while being expressed from the macromolecular synthesis operon, whereas the fusion protein is fluorescent when expressed from an inducible promoter. Hence, we attempted to express ectopically the fluorescent-primase from the chromosome in which the native primase gene is then knocked-out. The knockout strategies resulted only in false positive strains necessitating a highly efficient genome-editing tool to engineer primase gene in *E. coli*.

The research work discussed here demonstrates the utility of various methods and experiments performed to expand our knowledge on the replication of chromosome in live growing cells. An inter-disciplined approach was employed to understand the intricate details of the DNA replication in action. The applicability of *in vivo* single molecule imaging techniques ranges widely from determining the dynamics of endogenous proteins to their stoichiometry during chromosome replication. In this research work, we have validated the effectiveness of the above-said techniques in gaining valuable insights on the replication process at the single-molecule level. We have also presented the challenges in strain creation techniques and the importance of them in the implications of *in vivo* single molecule research. As with any scientific research, the valuable results obtained from our study raise more innovative questions that the researchers will be able to answer with further technical developments in future.



## SAMENVATTING

Erfelijke informatie wordt naar de volgende generatie overgebracht door middel van DNA of RNA. Genen bevinden zich op chromosomen. In *Escherichia coli*, een van de meest bestuurde modelorganismen, bevinden de genen zich op een circelvormig chromosoom die bi-directioneel gerepliceerd worden. De verdubbeling begint voor beide replicatie vorken gelijktijdig en op het zelfde punt, de ori. Kleine fouten gedurende het proces van DNA verdubbeling kunnen grote gevolgen hebben in latere stadia van de bacterie. Veel van onze huidige kennis over DNA replicatie is afkomstig van *in vitro* experimenten. Deze condities wijken echter sterk van de *in vivo* condities van de cel.

Vooruitgang in technieken voor genetische manipulatie en in technieken voor fluorescente imaging maken het mogelijk om het proces van de replicatie *in vivo* te bestuderen. Single molecule experimenten kunnen een beter inzicht geven in processen tijdens de DNA replicatie die verborgen blijven in bulk-experimenten. Kwantitatieve single-molecule experimenten zijn essentieel om het proces van replicatie volledig te doorgronden. Single molecule imaging experimenten *in vivo* worden mede mogelijk gemaakt door de ontwikkeling van microfluid kamers. Door deze toepassing kunnen de temporale en spatiale groeiomstandigheden van de delende cel optimaal gecontroleerd worden. Ook de ontwikkeling van geautomatiseerde image analyse van fluorescente signalen is sterk in opmars ondanks heterogeniteit en crowding effecten in delende cellen. En tenslotte is de correcte labeling van endogene eiwitten met fluorescerende eiwitten mogelijk gemaakt door recente ontwikkelingen in genoom modificerende technieken. In dit proefschrift worden een combinatie van genoemde technieken besproken waarbij single molecule imaging gebruikt wordt als gereedschap om inzicht te krijgen in het mechanisme van DNA replicatie *in vivo*.

Chromosoom modificatie in *E. coli* gebeurt met technieken die gebaseerd zijn op homologe recombinatie en deze vinden hun weg naar industriële toepassingen en fundamenteel onderzoek. In dit proefschrift worden deze technieken gebruikt om endogene eiwitten te labelen door de coding sequence voor de fluorescente eiwitten te fuseren aan de coding sequence van het gen of interest. Grondige validatie van de verkregen stammen is echter nodig voor dat deze gebruikt kan worden in het onderzoek en de benodigde methodes hiervoor worden beschreven in **hoofdstuk 2**.

Modificatie van chromosomaal DNA vindt plaats door middel van recombinatie technieken en de fusie van gemodificeerde sequenties in een *E. coli* stam gebeurt door middel van P1-faag transductie. Naast validatie door middel van PCR en DNA sequensen, bediscussiëren we ook de Southern blotting als techniek voor de validatie van de verkregen stammen. Eventuele extra inserties kunnen verwijderd worden door het gen of interest te transduceren naar wildtype cellen met virulente P1 fagen. Maar ook deze virulente P1 stammen kunnen besmet raken met temperate fagen, welke als plasmide kunnen voorkomen in groeiende cellen. We beschrijven hier ook methodes om besmetting met temperate fagen aan te tonen met behulp van cross streak agar en Evans-blue uranine agar platen. Daarnaast laten we zien hoe cel groei en cel morfologie defecten gebruikt kunnen worden in de validatie van de nieuwe stam.



Als toepassing van de genoemde technieken, hebben we de  $\beta_2$  sliding clamp eiwitten in groeiende cellen in detail bestudeerd.  $\beta_2$  sliding clamp eiwitten spelen een rol in DNA replicatie en DNA herstel. Ze vormen een van de belangrijkste componenten van het replisoom en als zodanig verantwoordelijk voor nauwkeurigheid en processiviteit van de replicatie. In **hoofdstuk 3** worden de resultaten van een kwantitatieve *in vivo* studie van de  $\beta_2$  sliding clamp dynamiek gepresenteerd. Met widefield en photo-activable fluorescentie microscopie hebben we het aantal  $\beta_2$  sliding clamps bepaald dat aan DNA gebonden is en de bijbehorende snelheid van DNA loading en unloading. We hebben gemeten dat de  $\beta_2$  sliding clamp ophoopt op het DNA na het begin van de replicatie en dat deze moleculen gedurende vrijwel het gehele replicatie proces aanwezig blijven op het DNA. Deze ophoping van  $\beta_2$  sliding clamp wordt veroorzaakt door een hoge loading rate maar een lage unloading rate.

In de voortschrijdende replicatie vork komen de vork eiwitten verschillende barrières tegen in de vorm van DNA gebonden eiwitten welke ze dienen te passeren om het replicatie proces te kunnen voltooien. In **hoofdstuk 4** beschrijven de we wat er gebeurd in een *E. coli* cel met artificiële ori als het replisoom een endogeen Tus-*Ter* blokkade tegen komt. We laten zien dat de replisomen, gedetecteerd door middel van een gelabeld  $\beta_2$  sliding clamp, voortdurend gebonden zijn aan het DNA, zelf wanneer er een Tus-*Ter* blokkade aanwezig is. Met behulp van metingen aan het kinetiek konden we concluderen dat de replicatiesnelheid weliswaar vertraagd wordt door de Tus-*Ter* blokkade, maar dat de replicatie niet permanent geblokkeerd wordt.

In **hoofdstuk 5** bepalen we met widefield fluorescentie microscopie de lokalisatie en stoichiometrie van het Tus eiwit bepaald gedurende de cel cyclus. Uit voorlopige resultaten blijkt dat een cel ongeveer 13 kopieën van het Tus eiwit bevat, waarvan er 5 mogelijk aan *Ter* sites gebonden zijn. De intensiteit van de foci verandert niet gedurende de celcyclus, suggererend dat Tus eiwit constant gebonden blijft aan de *Ter* site.

Replicatie in *E. coli* vind plaats in een semi-conservatief mechanisme, waarbij de replicatie vork gedurende de initiatie fase gevormd wordt met RNA primers die gesynthetiseerd zijn op de DNA template door primase. Tijdens de elongatie fase wordt de leading strand van de replicatie vork continue gesynthetiseerd terwijl de lagging strand discontinuous gemaakt wordt met de hulp van primase. Om de dynamica van het primase eiwit te kunnen bestuderen, hebben we geprobeerd om het primase eiwit in *E. coli* fluorescent te labelen. In **hoofdstuk 6** beschrijven we de scar-less techniek die we gebruikt hebben en laten we zien dat het getagde eiwit in de gemodificeerde cel geen fluorescerend signaal meer gaf. De mogelijke verklaringen hiervoor komen in dit hoofdstuk aan de orde.

Het onderzoek beschreven in dit proefschrift laat diverse technieken zien die gebruikt zijn om onze kennis over het mechanisme van DNA replicatie in levende en groeiende *E. coli* cellen uit te breiden. Een inter-disciplinaire benadering is gebruikt om details van het actieve DNA replicatie systeem in beeld te brengen. De toepassing van *in vivo* single molecule imaging technieken varieert van het vaststellen van de dynamiek van endogene eiwitten tot het vaststellen van de stoichiometrie gedurende chromosoom replicatie. In dit proefschrift hebben we de effectiviteit van de bovengenoemde technieken gevalideerd en waardevol inzicht gekregen in het replicatie proces op moleculair niveau. We hebben ook nieuwe uitdagingen laten zien in het maken van nieuwe stammen en de consequenties en het belang hiervan in *in vivo* single molecule onderzoek. En zoals bij ieder

onderzoek leiden ook hier de belangrijke vindingen tot meer uitdagende vragen die de onderzoeker kan beantwoorden met de nieuwe technieken van de toekomst.



## ACKNOWLEDGEMENTS

Ph.D. research takes away a major portion of our hard-working phase of life, and the journey towards achieving the degree requires enormous support and guidance from colleagues, family members and friends facilitating you to complete the travel successfully. I would like to express my gratitude by acknowledging the several people who have helped me in attaining my degree.

First, I would like to thank my promoter and supervisor Nynke Dekker. Nynke, I am grateful to you for trusting and helping me during this transition of performing interdisciplinary research between biology and physics. I am extremely indebted to you for supporting me during the difficult times, and your positive criticisms have always made me a better researcher. You are liberal when it comes to exploring different techniques required for the study, which has motivated me to stay in research. You have always encouraged me to widen the scope of my research by providing opportunities and funds to learn via various courses and conferences conducted both in local and abroad. You have also taught me various skills that you excel at, like academic writing, time management, and goal-oriented working. Overall, it has been a very steep and progressive learning curve for me and without your help and patience, I would not have come this far. When it comes to social activities apart from work, you made sure to build the team spirit by organizing annual group trips to exotic locations, and they were some of the most cherishing memories for me. I express here my wholehearted gratitude to you for nurturing the researcher in me.

I am very privileged to have a group of outstanding scientists on my committee: Andreas Engel, Tom Shimizu, Hans Tanke, Greg Bokinsky, and Vincent Lorent. I am very thankful to each of you for your time and interest in evaluating my thesis.

I had the opportunity to meet and work with many eminent people in Bionanoscience. First among them is Charl Moolman, who is my first research partner in lab. Charl, you were my senior by just nine months, and we both built up the live-cell imaging part of our lab, of course, along with others (mentioned below). Indeed, as you always say, it feels Awesome! I am very thankful to you for making such a fruitful and cordial collaboration, and you have been a very good inspiring friend. I wish you all the best for your career ahead. Next, I must thank Vincent Lorent, who I first met during his sabbatical in our lab. Vincent, you being a physicist with curiosity in biology and myself the other way, we both shared a good rapport in learning mutually the concepts that were alien to us. It is an honor to have worked closely with you. Then, I must thank Anne Meyer, who guided me to optimize the phage transduction technique in our department and you were always available and helpful whenever I get doubts in the strain creation techniques. Thank you very much, Anne!

I am very grateful to Richard Janissen who has been a very good friend even beyond work. Richard, I must thank you for all the friendly suggestions and encouragements you gave me whenever I was facing problems in both research and personal life. Then, I have to thank my longest office roommate, Orkide Ordu, who has been a friend and well-wisher. Orkide, I always admire your enthusiasm to finish your work rapidly and I wish you all the best for wrapping up the thesis soon. Next, I am thankful to Theo van Laar with whom I have closely collaborated on the biochemistry and strain creation work. He is one of the most trust-worthy people, and he never says no to any help asked. He has graciously aided me in

translating the Dutch part of the summary in this thesis. Thank you, Theo. Then, I must thank Jacob Kerssemakers. He took many pains to teach me the data analysis work and one of the best data analyst I have seen. I really admire his way of balancing your personal life, painting, and work. Keep up the good job and best of luck for your adventures, Jacob. I am also thankful to Margreet Docter who helped me in building the newer fluorescence microscope setup for our lab. I am also pleased and grateful to have worked with other technicians: Serge Donkers, Susanne Hage, Ilja Westerlaken, Bronwen Cross and Andrea Martorana. Thank you all.

I should thank Peter Brezda, Sumit deb Roy and Roy de Leeuw who have been helping me in the later part of my research and they are taking forward the live cell imaging projects further in our lab. Good luck folks! Then, I am thankful to Nienke Blokker who did her bachelor thesis under my guidance on growth-based strain validation and Liliane Jimenez-van der Hoorn who helped me in standardizing the biochemical techniques. Then I am also thankful to Jasper van der steen and Kasper de Leeuw, the students of the KNAW-funded project, who helped with the PALM experiments mentioned in the thesis.

I am very happy to have worked with many ‘great minds’ from our lab. Particularly, I am grateful to Zhongbo Yu who has been a good friend and advisor with similar research interests. I truly miss the opportunity of working in your lab, and I am looking forward to working or collaborating with you maybe in future. I should also thank Mariana Kober who has been a very good and helping friend. I am also thankful to my peers Bojk Berhuis, Maarten van Oene, Seungkyu Ha, Artur Kaczmarczyk, Ilyong Jung and Eugen Ostrofet, with all of whom I have had close interaction. I am also proud and happy to have interacted with senior researchers from our lab: Francesco Pedaci, Gary Skinner, Jan Lipfert, Xander Janssen, Laura Dickinson, Mina Lee, Zhuangxiong Huang, Aartjan te Velthuis, and David Dulin.

The Bionanoscience department was formed in the same month I started my Ph.D. research and in the last six years, it has grown rapidly with many young and vibrant researchers. Particularly, I must thank my peers: Stanley Chandradoss, Mahipal Ganji, Aditya Ananth, Siddharth Deshpande, Fabai Wu, Natalia Vyturina, Malwina Szczepaniak, Mathia Arens, Luuk Loeff, Mohamed Fareh, Erwin van Rijn, Anna Haagsma, Regis Flohr, Carsten Blom, Minyoung Heo, Yaron Caspi, Felix Hol, Sung Hyun Kim, Allard Katan, Jaco van der Torre, Stephanie Heerema, Jorine Eeftens, Laura Restrepo Perez, Daniel Verschueren, Jakub Wiktor, Calin Plesa, Tessa Jager, Tim Blosser, Gautam Soni, Magnus Jonsson, Rifka Vlijm, Zoreh Nourian, Pauline van Nies, Andrew Scott, Alicia Soler Canton, Helena Shomar, Fabrizio Anella, Esengul Yildirim, Aafke van den Berg, Ruben van Drongelen, Misha Klein, Dominik Schmieden, Michela de Martino, Rick Vink, Marek Noga, and Roland Kieffer.

I am grateful to the following non-scientific staff for their help in need: Dijana Boric, Amanda van der Vlist, Emmylou van Hartrop, Jolijn Leeuwenburgh, Belen Solano Hermosilla, Esther Rijnders, Jaap Beekman, Sacha Khaiboulov, Jelle van der Does and Dimitri de Roos. I am very thankful to all the faculty members for their interactions with me during forum meetings and other informal occasions, especially: Cees Dekker, Marileen Dogterom, Martin Depken, Chirlmin Joo, Bertus Beaumont, Christophe Danelon, Timon Idema, Elio Abbondanzieri, Maria-Eve Aubin Tam, Timon Idema, Andreas Engel, Greg Bokinsky, Leidewij Laan, and Hyun Youk.

I am very thankful to David Sherratt (University of Oxford) and his lab members for the inspiring collaboration and their bacterial strains formed the foundation of live-cell imaging projects we performed in our lab.

I truly believe that the strength of our professional work stems from our personal life. Now, I am going to acknowledge the people who decorate my personal life. First, I must mention about my very beautiful and understanding wife Sheela Krishnamachari. She might have come into my life only during later part of my Ph.D., but, without her support and care, I cannot imagine myself successfully completing my thesis work now. Being a biotechnologist herself, she apprenticed me in the office work too. Words are inexpressive to say how much I thank you dear. All I can say is 'I love you'. Next, I must thank my parents T.S.Krishnan and Geetha Krishnan, and my in-laws V.T.Krishnamachari and (late) Jaya Krishnamachari, for their kindness, encouragement, and blessings. They say a friend in need is a friend indeed. One such friend is Dharjath Ahamed, who has been my close friend from my bachelor studies. You and your wife Aadhil Dharjath have helped me beyond your capabilities and simple thanks will not be enough to show my gratitude to both of you. I wish you the best of luck for finishing your doctorate studies soon. I am sure we will continue our cherishing and ever-lasting friendship wherever we are. Special thanks to my good friend Vidyaramanan Ganesan and his family and only with your help I have pursued higher studies in abroad. Your occasional messages are always encouraging and motivating to my work. Thank you. I also would like to show my gratitude to my friends I made in Delft: Vasudevan Lakshminarayanan, Elly Boon, Annelies Amade, Sivananth Murugesan, Rajat Bharadwaj, Yash Joshi, Bhaskar Patil, Raj Tilak Rajan, Valia Gkeredaki, Dinakaran Sivasiddharthan, Krishna Kowlgi, Manimaran Pari, Amrit Padda, Sumit Sachdeva, Purvil Khakharia, Lalit Gangarude, and Chockalingam Veerappan.



## **CURRICULUM VITAE**

### **Sriram TIRUVADI KRISHNAN**

- 29/10/1984                      Born in Madurai, Tamil Nadu, India
- 08/2002-06/2006              B.Tech.  
Department of Biotechnology,  
Anna University, Coimbatore, India
- 07/2006-08/2008              Traineeship and successively as Software Engineer, India
- 09/2008-08/2010              M.Sc.  
Department of Biotechnology,  
Delft University of Technology, The Netherlands
- 09/2010-09/2016              Ph.D.  
Department of Bionanoscience,  
Delft University of Technology, The Netherlands  
Supervisor: Prof. Dr. Nynke H. Dekker





## LIST OF PUBLICATIONS

Moolman MC\*, Tiruvadi Krishnan S\*, Kerssemakers JWJ, de Leeuw R, Lorent V, Sherratt DJ, Dekker NH (2016) The progression of replication forks at natural replication barriers in live bacteria. *Nucleic Acids Research* **44**: 6262-6273

Tiruvadi Krishnan S, Moolman MC, van Laar T, Meyer AS, Dekker NH (2015) Essential validation methods for *E. coli* strains created by chromosome engineering. *Journal of Biological Engineering* **9**: 1-14

Moolman MC, Krishnan ST, Kerssemakers JWJ, van den Berg A, Tulinski P, Depken M, Reyes-Lamothe R, Sherratt DJ, Dekker NH (2014) Slow unloading leads to DNA-bound  $\beta_2$ -sliding clamp accumulation in live *Escherichia coli* cells. *Nature Communications* **5**: 5820

Moolman MC\*, Huang Z\*, Krishnan ST, Kerssemakers JW, Dekker NH (2013) Electron beam fabrication of a microfluidic device for studying submicron-scale bacteria. *Journal of Nanobiotechnology* **11**: 1-10

---

\* Authors have contributed equally to this work.

**Casimir PhD series 2016-23**  
**ISBN 978-90-8593-267-3**