University of Wollongong
Research Online

University of Wollongong Thesis Collection 2017+

University of Wollongong Thesis Collections

2023

Resolving stalled replication forks in Escherichia coli

Kelsey Sarah Whinn

Follow this and additional works at: https://ro.uow.edu.au/theses1

University of Wollongong

Copyright Warning

You may print or download ONE copy of this document for the purpose of your own research or study. The University does not authorise you to copy, communicate or otherwise make available electronically to any other person any copyright material contained on this site.

You are reminded of the following: This work is copyright. Apart from any use permitted under the Copyright Act 1968, no part of this work may be reproduced by any process, nor may any other exclusive right be exercised, without the permission of the author. Copyright owners are entitled to take legal action against persons who infringe their copyright. A reproduction of material that is protected by copyright may be a copyright infringement. A court may impose penalties and award damages in relation to offences and infringements relating to copyright material. Higher penalties may apply, and higher damages may be awarded, for offences and infringements involving the conversion of material into digital or electronic form.

Unless otherwise indicated, the views expressed in this thesis are those of the author and do not necessarily represent the views of the University of Wollongong.

Research Online is the open access institutional repository for the University of Wollongong. For further information contact the UOW Library: research-pubs@uow.edu.au

Resolving stalled replication forks in *Escherichia coli*

Kelsey Sarah Whinn

Bachelor of Medical Biotechnology (Honours)

This thesis is presented as part of the requirement for the conferral of the degree:

Doctor of Philosophy

Supervisor: Prof. Antoine van Oijen

Co-supervisor: Dr Lisanne Spenkelink

Associate supervisor: Dr Harshad Ghodke

School of Chemistry and Molecular Bioscience
University of Wollongong

This research has been conducted with the support of the Australian Government Research Training Program Scholarship

February 2023

To my Nanna. I hope to continue your passion for education, in learning, teaching and growing.

Abstract

DNA replication is essential to successful cell proliferation. Inheritance of traits during cell propagation relies on the accurate duplication of the parental double-stranded DNA (dsDNA) to form two identical daughter copies. This process is carried out by a multi-protein complex referred to as the replisome. Decades of investigations using the model *Escherichia coli* (*E. coli*) replisome have provided an overall picture of the process of DNA replication initiation, elongation and termination. However, DNA replication in cells occurs on template DNA coated in DNA-binding proteins that can act as roadblocks and stall the replisome, often resulting in drastic effects on the chromosome. However, the fate of the replisome at these sites remains poorly understood. Stalled DNA replication has been linked to the emergence of antimicrobial resistance in prokaryotes, and the development of severe physical disorders and diseases in eukaryotes. Therefore, understanding the underlying mechanisms of stalled DNA replication can inform future investigations into the maintenance of genome integrity.

This thesis focuses on the development and use of single-molecule tools to investigate stalled replication and the resolution of protein roadblocks. Single-molecule tools provide the ability to watch one molecule at a time. Extensive use of these techniques has revealed the heterogeneity that exists within complex biological pathways. Specifically, this thesis highlights the myriad of previously unknown behaviors of proteins on DNA as revealed by single-molecule tools.

This thesis describes novel single-molecule techniques to investigate stalled DNA replication *in vitro*. First, automated analysis methods of detecting replicating DNA products over time and interacting fluorescently-labeled proteins are presented. This automated method reduces the introduction of bias and ensures accurate characterization of protein-replisome interactions. Further, this thesis describes the development of single-molecule fluorescence assays to stall the *E. coli* replisome using a roadblock based on the dCas9 protein. Additionally, these assays can be performed on 2-kbp and novel 18-kbp rolling-circle DNA templates providing increased spatial resolution. Finally, the single-molecule methods established throughout this thesis are used to investigate the mechanisms of action of the accessory helicase Rep at the *E. coli* replisome. Rep is an essential helicase and translocase that functions to remove protein roadblocks and rescue stalled DNA replication, however, the mechanisms dictating this function remain to be characterized. This work reveals that Rep stochastically associates with the replisome, independently of whether the replisome is elongating or is in a stalled state. Further, using roadblocks of decreasing stability, investigations of stalled DNA replication rescue show that the rate-limiting step is a process involved with the continuation of synthesis and that Rep acts quickly to remove protein roadblocks. This thesis brings together well-established methods of investigating DNA replication with the innovation of novel single-molecule techniques to investigate the resolution of stalled DNA replication *in vitro*. Together, this thesis provides the foundation for future investigations into stalled DNA replication in prokaryotic and eukaryotic systems.

Acknowledgments

Thank you to my thesis reviewers, Dr Emma Sierecki and Prof. David Rueda for taking the time to provide insightful comments.

I want to begin by acknowledging the number of women in STEM who inspired my interest in pursuing a career in science. I would not have thought I could be here too if it had not been for your presence and encouragement in those early undergraduate years.

I first wish to sincerely thank Prof. Antoine van Oijen for sparking my interest in DNA replication and single-molecule techniques. I still remember seeing the video that showed DNA replication in real-time and thinking it was amazing. From that moment you have supported, encouraged and challenged me to be the best scientist I can be. I am truly grateful to have had you as a supervisor for the past 5 years.

Secondly, to Dr Harshad Ghodke. This work would not be possible without your guidance from my Honors year to my PhD. Your enthusiasm for science and extraordinary ability to think outside of the box is truly inspiring. I appreciate all of your support, encouragement and patience over the years, and look forward to seeing what your new science endeavors bring.

To Dr Lisanne Spenkelink, thank you for your continued support and motivation throughout the years. Your encouragement to try those ‘fun Friday’ experiments have led to some fascinating results and proves that there is never any harm in trying. Thank you for your enthusiasm, optimism and encouragement, especially over the last 6 months.

To Prof. Nick Dixon, thank you for your support and encouragement over the years. Your knowledge of DNA replication is extraordinary. I value all of the feedback and input you have provided in the development and finalization of this project.

To Boban, Dr Slobodan Jergic, thank you for teaching me patience in science. Your ability to think through each intricate step of reactions and experiments is truly amazing. I appreciate all of your support and feedback over the years.

To Dr Zhi-Qiang Xu, thank you for taking me under your wing and teaching me about protein expression and purification. Your years of expertise and work made this project possible.

A huge thank you to the members of the van Oijen, Dixon, Oakley and Tolun labs. It has been a pleasure getting to know each of you and I’m excited to see what you all do in the coming years.

A special mention to fellow van Oijen PhD candidates (some now officially/almost complete) over the past 5 years - Sarah, Megan, Caitlin, Nick M, Richard, Stefan, Gurleen, Tom, Remy, Elizabeth, Nischal, Elise. You have all provided words of advice, motivation or methods of procrastination that have made the years enjoyable. Gurleen – thank you for your guidance, always happily answering my ‘dumb’ questions and providing the sass when it was needed. Tom – your chilled nature may be polar opposite to my own, however, your ability to face everything with a ‘she’ll be right’ was always comforting in my times of need. Liz – your support over the final few years has

meant an incredible amount to me. Thank you for always being ready for a snack run, passionate conversations about the patriarchy and general life chat. I have made some incredibly intelligent friends along the way and I cannot wait to see where your passions for science take each of you.

Being a first-generation university student, and PhD candidate, has made this journey tough, but I hope I have inspired others along the way. To my friends, thank you for encouraging me and providing some fun times to help recharge. To my parents, Carmel and Allyn, and sister, Georgia, thank you for supporting me, believing in me and encouraging me to do whatever it was I set my mind to. Thank you for always being an ear to listen to my frustrations and offering solutions to not-so-big problems. To my brother, Rory, the challenges you have faced throughout your life have been an inspiration to me. My hope is that one day you find your inspiration too. To my four-legged coworkers during the lockdowns, Sugar and Nyah (*Vale*), thank you for always having a wagging tail on those tough days.

And finally, to my husband, Ryan – you are my rock. Thank you for being my biggest supporter and cheerleader. Thank you for agreeing to marry me during a PhD, mixed with a dash of a pandemic. I cannot put into words how your support and encouragement have gotten me to where I am. Here’s to continuing to chase our dreams.

Certification

I, Kelsey Sarah Whinn, declare that this thesis submitted in fulfilment of the requirements for the conferral of the degree Doctor of Philosophy, from the University of Wollongong, is wholly my own work unless otherwise referenced or acknowledged. This document has not been submitted for qualifications at any other academic institution.

Kelsey Sarah Whinn
22nd February 2023

List of Publications

1. **Kelsey S. Whinn**, Zhi-Qiang Xu, Slobodan Jergic, Nischal Sharma, Lisanne M. Spenkelink, Nicholas E. Dixon, Antoine M. van Oijen, Harshad Ghodke, ‘Single-molecule visualization of stalled replication rescue by the *Escherichia coli* Rep helicase’, *Nucleic Acids Research,* 51 (7), 2023

2. **Kelsey S. Whinn**, Nischal Sharma, Antoine M. van Oijen, Harshad Ghodke, ‘Single-molecule fluorescence imaging of DNA replication stalling at sites of nucleoprotein complexes’, *Manuscript accepted at Methods in Molecular Biology: Single Molecule Analysis*

3. **Kelsey S. Whinn**, Antoine M. van Oijen, Harshad Ghodke, ‘Single-molecule studies of helicases and translocases in prokaryotic genome-maintenance pathways’, *DNA Repair (Amst)*, 108, 2021

4. **Kelsey S. Whinn**, Antione M. van Oijen, Harshad Ghodke, ‘Spying on Cas9: Single-molecule tools reveal the enzymology of Cas9’, *Current opinions on Biomedical Engineering*, 12, 2019

5. **Kelsey S. Whinn***, Gurleen Kaur*, Jacob S. Lewis, Grant D. Schauer, Stefan H. Mueller, Slobodan Jergic, Hamish Maynard, Zhong Y. Gan, Matharishwan Naganbabu, Marcel P. Bruchez, Michael E. O’Donnell, Nicholas E. Dixon, Antoine M. van Oijen, Harshad Ghodke, ‘Nuclease dead Cas9 is a programmable roadblock for DNA replication’*, Scientific Reports*, 9 (1), 2019

* These authors contributed equally to this work

Table of Contents

List of Figures

Chapter 1

Chapter 2

Chapter 3

Chapter 4

Chapter 5

Chapter 6

List of Tables

Chapter 2

Chapter 4

Chapter 5

Chapter 6

Chapter 1

General Introduction

DNA contains all of the genetic information essential to sustaining cellular life. Characterizations of the structure of DNA, and the specific sequences it encodes, implied the existence of a sophisticated mechanism where the DNA sequence could be replicated and repaired [1-4]. High-fidelity DNA replication is required to support successful cell proliferation in all organisms. Early investigations of DNA replication revealed a semiconservative process that occurs on a distinct two-armed structure called the replication fork [3, 4]. Decades of extensive biochemical investigations have revealed a complex of proteins, termed the replisome, is responsible for DNA replication.

The replisome is a multi-protein machine present in all organisms. DNA replication begins with the assembly of the replisome at specific genomic locations, called origins [4, 5]. The replisome then translocates and simultaneously unwinds the double-stranded DNA (dsDNA) to form two template strands that are subsequently replicated in a semi-discontinuous manner to form two daughter copies [3-9]. Despite decades of investigations, and what presents as a fairly straightforward process, many of the precise molecular mechanisms of DNA replication are not fully understood. Specifically, the mechanisms underlying the failure of the replisome to replicate DNA accurately is an area still incompletely understood.

Inaccurate DNA replication can have deleterious effects on cells. This can be a result of, for example, incorrect copying of the genetic sequence, stalling of the replisome as it progresses along the template DNA, or failure to repair DNA damage [10, 11]. Each of these processes can result in mutations in the genetic sequence, some of which can be fatal or beneficial to the cell. In prokaryotes, mutations in DNA can give rise to antimicrobial resistance [12, 13]. To combat resistance to existing or future antimicrobial compounds, it is important to first understand the underlying mechanisms within prokaryotes that can promote resistance. Furthermore, the effects of inaccurate copying of genetic material in eukaryotes can be serious. In humans, these inaccuracies have been linked to the development of severe developmental disorders and diseases including premature aging, neuromuscular degeneration and cancer predisposition [14-18]. Understanding the intricate mechanisms that determine the failure or success of DNA replication can provide insight into disease development and treatment.

For decades, scientists have used prokaryotic replisome systems to characterize the mechanisms of DNA replication. These systems provide a simple alternative to complex eukaryotic systems, where homologs of vital replisome components are present. Investigations of these model replisomes in ensemble-averaging biochemical techniques involve reconstitution of the replisomes from prokaryotes. The work described in this thesis uses the *E. coli* model replisome.

1.1. The *E. coli* replisome as a model system

The *E. coli* model replisome has been extensively used to investigate the molecular processes underpinning DNA replication (reviewed in [19]). The *E. coli* replisome is comprised of more than 12 proteins that, once assembled, are capable of unwinding and duplicating DNA at rates of up to 1000 base pairs per second (bp s^{-1}) [6, 20-23]. DNA replication is initiated by the recruitment and assembly of the initiator proteins, DnaA, DnaB and DnaC, at a unique locus called the replication origin, or *OriC* [24-27]. Assembly of the initiator proteins at *OriC* results in bidirectional duplication of the circular *E. coli* genome, where two sister replisomes replicate half of the genome each [28-30]. This process is initiated by the binding of the DnaA protein that subsequently melts the duplex DNA to form a single-stranded DNA (ssDNA) bubble [24-26] (Figure 2.3). The ring-shaped DnaB helicase is then opened and loaded onto the ssDNA bubble in complex with the DnaC helicase loader [27, 31]. Binding of the DnaG primase to DnaB triggers the disassociation of DnaC and translocation of DnaB towards the fork junction, where a functional replisome is formed and elongation begins [32].

The DnaB helicase drives strand separation at the leading edge of the replication fork. Interacting with DnaB is the DNA polymerase III (Pol III) holoenzyme, itself comprising several protein complexes [19, 33] (Figure 1.1). These include the Pol III core, consisting of α, ε, θ subunits that are responsible for DNA synthesis; the β_2 sliding clamp that provides a stable association between the Pol III core and the template DNA; and the clamp loader complex, $\tau_n\gamma_{(3-n)}\delta\delta'\psi\chi$ (where n=2 or 3), which loads the β_2 sliding clamp onto the template DNA [6, 20]. The clamp loader complex interacts with two or three Pol III cores, forming Pol III* (i.e. the holoenzyme lacking β_2) [7, 34, 35].

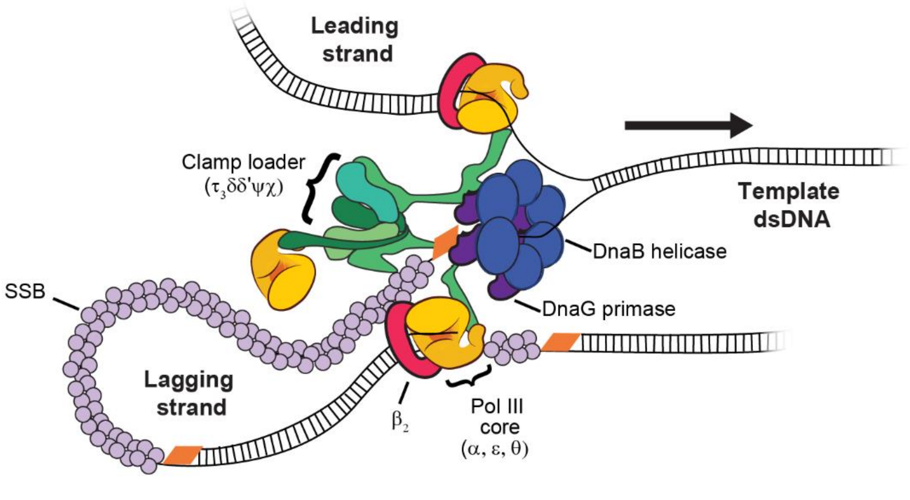


Figure 1.1. The *E. coli* replisome. Located at the front of the replisome on the lagging strand, the DnaB helicase unwinds the double-stranded template DNA into two single-strands. The DnaG primase synthesizes short RNA primers on the lagging-strand ssDNA template. The ssDNA template is protected by the binding of single-stranded DNA binding proteins (SSBs). The clamp-loader complex, comprised of up to three τ proteins and the δ,δ', ψ and χ subunits, loads the β_2 sliding clamp and $\alpha\varepsilon\theta$ polymerase III core onto

the newly primed sites. DNA replication elongation occurs in the 5' to 3' direction discontinuously on the lagging strand in the form of Okazaki fragments. Leading-strand DNA replication occurs continuously. Figure adapted from [34] with permission.

The assembly and orientation of the DNA Pol III HE allows simultaneous synthesis of both the leading and lagging strands. The antiparallel nature of DNA dictates that this coordinated synthesis occurs in a semi-discontinuous manner [8, 36]. As such, the leading strand is synthesized continuously in the direction of travel of the fork, by the incorporation of nucleotides at the 3' end of the growing chain [7]. Meanwhile, the lagging strand is synthesized discontinuously, through cycles of RNA primer synthesis, by DnaG, and DNA extension of the Pol III core to form Okazaki fragments [7, 8, 37, 38]. To overcome the directionality problem faced by coupled leading- and lagging-strand synthesis, lagging-strand loops at the replication fork reorient the lagging-strand polymerase so it synthesizes DNA in parallel with the polymerase on the leading strand [9, 39] (Figure 1.1).

In cells, replication occurs on template DNA coated in DNA-binding proteins. These protein roadblocks can pose physical barriers to the replisome, ultimately leading to replication fork collapse. These roadblocks can include proteins involved in chromosome organization and compaction, homologous recombination filaments, and transcription factors [40-43]. While each is vital to cell survival, their presence on template DNA undergoing DNA replication can prevent unwinding and inhibit DNA synthesis, leading to genome instability [11].

Transcribing RNA polymerases (RNAPs) pose a significant and interesting challenge to the progression of the replication fork. Approximately 1500-2000 molecules of RNAP are bound to DNA during the bacterial exponential growth phase [44]. Collisions between the prokaryotic transcription and replication machineries can occur as a result of differing elongation rates *in vivo* between RNAP (~50 nt s^{-1}) and the replisome (~1000 bp s^{-1}) [10, 34, 45-53]. These collisions can also result from paused, stalled or backtracked RNAPs on either damaged or undamaged template DNA [54-57]. Interestingly, these collisions occur regardless of their relative direction of travel, termed; head-on collisions, where both the leading edges of the replisome and RNAP collide; and co-directional collisions, caused by the meeting of the leading edge of the replisome with the rear edge of the RNAP [10]. These long-lived barriers significantly inhibit the progress of the replication fork and are one of the major causes of genome instability in prokaryotes [11, 58].

Fortunately, organisms harbor the ability to remove these protein roadblocks from the path of the replication fork. Accessory helicases can remove proteins bound to DNA and rescue stalled replication. In *E. coli*, accessory helicases, Rep and UvrD, are known to remove a variety of roadblocks including RNAPs [59-61]. Extensive investigations of the Rep helicase have revealed an ability to interact with the DnaB helicase and remove roadblocks from DNA through its ATPase activity [59, 62]. Despite these investigations, the intricate details underlying how Rep-replisome interactions coordinate the removal of roadblocks with reducing the occurrence of replication fork stalling remain to be fully understood.

1.2. Investigating stalled DNA replication

The fate of the replisome at protein roadblocks is not well-understood. A variety of methods have been used *in vitro* and *in vivo* to investigate stalled replication. These have included a variety of model roadblocks, including tandem repressor-operator complexes, streptavidin-biotin complexes, prokaryotic replication termination sites (Tus-*ter*), and stalled RNAP complexes [63-73]. While each method has provided insight into the mechanisms of stalled replication, their implementation suffers significant shortfalls including tedious procedures to incorporate binding sequences and poorly defined roadblock positions.

Nonetheless, investigations of replication fork arrest reveal two major open questions; does the replisome pause, remain stable and continue elongation upon removal of the roadblock? Alternatively, does the replisome disassemble upon colliding with roadblocks? Early studies suggest that the latter occurs, where replisomes were shown to dissociate after encountering tandem arrays of nucleoprotein complexes at arrest sites [66, 70, 71]. In contrast, other studies have indicated that the replisome remains stable for up to 60 min at sites of protein roadblocks and can resume elongation following the removal of the roadblock [65, 68, 72, 73].

To begin to answer these questions, and those related to the resolution of stalled replication forks, we developed a simple and well-defined model roadblock using the nuclease dead Cas9 protein (dCas9) of the CRISPR/Cas9 system [74]. This tool involves the dCas9 protein in complex with a complementary guide RNA (cgRNA) whose sequence targets the dCas9 to bind stably to a specific site. The dCas9-cgRNA roadblock can efficiently block the viral, bacterial and yeast replisomes, independent of which strand is targeted, in both bulk ensemble and single-molecule assays.

1.3. A single-molecule approach to investigate DNA replication

Our current understanding of DNA replication is largely shaped by investigations using classical biochemical techniques. While these techniques have provided details of replisome component functions, the resulting model of DNA replication is oversimplified. These textbook models describe DNA replication as being defined by steps involving proteins that are stable throughout. However, the recent use of single-molecule techniques has challenged this model of DNA replication, instead highlighting the plasticity of the replisome [34, 75-77].

Watching one molecule at a time offers a wealth of knowledge in understanding how biochemical processes are carried out. Single-molecule techniques provide increased temporal and spatial resolution, revealing the underlying complexity of biological pathways. Specifically, single-molecule techniques can reveal individual protein dynamics, pausing events and changes in behaviors that are otherwise hidden by ensemble-averaging techniques.

The single-molecule technique used in this thesis relies on fluorescence imaging of stained or labeled molecules of interest (Figure 1.2). Detection of the intensity and location of these molecules can provide real-time insights into the dynamic behaviors of individual proteins involved in DNA replication. This method has been used extensively to investigate DNA replication, revealing a highly-dynamic complex capable of pausing, exchanging components and changing rates of replication [34, 74, 75, 78].

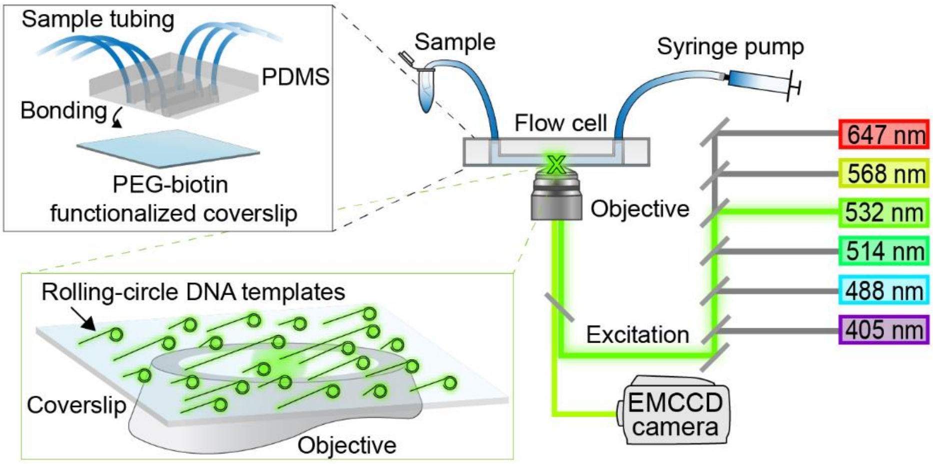


Figure 1.2. Single-molecule fluorescence experimental setup. Laser light of specific wavelengths is used to illuminate reaction components through the objective by total internal reflection fluorescence (TIRF). Reactions are carried out inside microfluidic devices (inset–top). Dye-stained rolling-circle DNA replication templates, attached to functionalized coverslips, are excited with a 532-nm laser (inset–bottom (not to scale)) and captured by the EMCCD camera. Proteins of interest can be labeled with other dyes and excited by a second wavelength to be visualized simultaneously.

1.4. Thesis outline

This thesis focuses on the development and application of techniques to investigate stalled *E. coli* DNA replication and stalled replication rescue by the accessory helicase, Rep. It aims to explore numerous hypotheses introduced above, using single-molecule fluorescence techniques, and sets out to better understand the significant importance of accessory helicases to cells.

***Chapter* 2** details the current understanding of numerous prokaryotic helicases and translocases in replication, repair and recombination pathways as highlighted by single-molecule techniques. We describe how recent advancements in single-molecule techniques have revealed interesting models of helicase and translocase behaviors that apply to proteins found in eukaryotic systems.

Chapter 3 describes the enzymology of the Cas9 protein as revealed by the use of single-molecule techniques. This literature review serves as a basis for understanding how the dCas9 roadblock, used throughout this thesis, behaves in targeting and binding dsDNA. Characterization of the dCas9 replication roadblock is provided in **Appendix I.**

Chapter 4 details the optimization of analysis protocols for *in vitro* single-molecule rolling-circle DNA replication assays. Previous studies using this assay have used manual analysis methods that introduce bias and variation between users. Here, I set out to optimize automated analysis methods to reduce the introduction of bias and accurately capture the details of single-molecule replication assays.

Chapter 5 describes in detail the development of methods to visualize stalled DNA replication at the single-molecule level. Using the rolling-circle DNA substrates, reconstituted *E. coli* replisomes and the dCas9 roadblock, we describe two assays that can be used to capture and investigate stalled DNA replication and resolution of stalled replisomes.

Chapter 6 combines the current understanding of helicases and the CRISPR/Cas9 system presented in Chapters 1 and 2, with the methods described in Chapters 4 and 5 to experimentally examine the activity of the Rep helicase at stalled replisomes. Here, a fluorescently-labeled Rep helicase is produced and used in *in vitro* single-molecule rolling-circle replication assays. These assays reveal the stochastic association of Rep to the replisome that occurs independently of whether the replisome is stalled or not. The use of the dCas9 roadblock, and roadblocks of decreasing stability, show that the rate-limiting step of stalled replication rescue is a process involved with the continuation of synthesis after roadblock removal. Together, the results presented in this chapter show that frequent Rep binding to the replisome serves as continual monitoring of the state of the replisome during elongation and enables it to act quickly to resolve roadblocks.

Chapter 7 brings together the new and existing knowledge described in this thesis to consider how our understanding of the resolution of stalled replication has progressed. Future directions of this biological question are provided.

Chapter 2

Single-Molecule Studies of Helicases and Translocases in Prokaryotic Genome-Maintenance Pathways

Kelsey S. Whinn, Antoine M. van Oijen, Harshad Ghodke

Published in *DNA Repair 2021,* 10.1016/j.dnarep.2021.103229

Helicases involved in genomic maintenance are a class of nucleic-acid-dependent ATPases that convert the energy of ATP hydrolysis into physical work to execute irreversible steps in DNA replication, repair, and recombination. Prokaryotic helicases provide simple models to understand broadly conserved molecular mechanisms involved in manipulating nucleic acids during genome maintenance. Our understanding of the catalytic properties, mechanisms of regulation, and roles of prokaryotic helicases in DNA metabolism has been assembled through a combination of genetic, biochemical, and structural methods, further refined by single-molecule approaches. Together, these investigations have constructed a framework for understanding the mechanisms that maintain genomic integrity in cells. This review discusses recent single-molecule insights into molecular mechanisms of prokaryotic helicases and translocases.

Contributions

This literature review addresses our current understanding of prokaryotic helicases and translocases in genome maintenance pathways and the significant contribution single-molecule studies have made to this knowledge. This review is divided into four sections; the first discusses the classification of helicases and translocases. The second section discusses DNA replication and the helicases and translocases involved (DnaB, Rep, RecG and PriA). The Third section discusses examples of helicases and translocases involved in DNA repair pathways (UvrD, UvrB and Mfd). Finally, the fourth section discusses examples of helicases and translocases involved in DNA recombination (RecBCD and UvrD).

I made the majority contribution to this review by exploring the literature on single-molecule studies of these proteins, drafting the manuscript, and designing and producing figures.

2.1. Introduction

DNA and RNA helicases are a diverse class of ATPases present in both prokaryotes and eukaryotes [79-82]. These enzymes are defined by their ability to use the chemical energy produced by nucleotide triphosphate hydrolysis to catalyze the remodeling of nucleic acids through directional unwinding and motion. Translocases, as the name suggests, are a related class of enzymes that transduce this chemical energy into directed motion on single-stranded or double-stranded nucleic-acid substrates.

DNA metabolism includes cellular activities such as DNA replication, repair, recombination, and transcription. In the paradigm of DNA metabolism reactions as a highly coordinated series of single reaction steps, ATPase activity serves to punctuate a sequence of reversible steps with irreversible steps that drive the reaction forward [79, 83]. These highly complex processes involve extensive manipulation of the nucleic acids through ATP-dependent functions of helicases such as separating the two strands during replication, detecting lesions in patches of single-stranded DNA (ssDNA) during repair, and processively resecting large swaths of DNA during recombination [19, 84, 85]. Additionally, helicases (and translocases) can also act on and cooperate with protein factors involved in these pathways to coordinate complex formation and disassembly [86-89]. Thus, through important interactions with proteins and nucleic acids, these enzymes play critical roles in maintaining genomic stability.

Many helicases, in both eukaryotic and prokaryotic systems, play critical roles in multiple genome-maintenance pathways. Defects in the human functional homologs of these factors can lead to severe disruptions in biochemical pathways that ultimately result in disorders. Understanding the precise roles played by helicases in individual pathways requires the ability to study these molecules in isolation, free from averaging effects of 10^7 – 10^{10} copies of molecules studied in typical bulk reactions. A detailed description of the activities of prokaryotic helicases in genome maintenance, and their regulation, is emerging through the use of single-molecule techniques. Single-molecule techniques afford significant gains in temporal and spatial resolution and consequently reveal the heterogeneity that exists in complex biological processes [90-92]. Furthermore, the recent expansion of the single-molecule toolkit to include robust approaches to quantify molecular behaviors in living cells has revealed how individual molecules conduct reactions in the cellular milieu [76, 93-99]. In this review, we showcase advances in the understanding of molecular mechanisms made using contemporary structural biology and single-molecule approaches through case studies of the best-characterized prokaryotic helicases and translocases in DNA replication, repair, and recombination pathways (Figure 2.1).

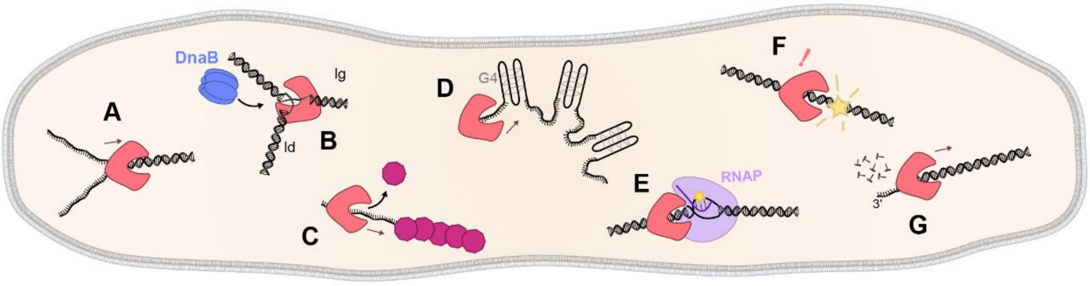


Figure 2.1: Roles of helicases and translocases in genome maintenance pathways. Helicases, and/or translocases, (pink) act in replication, repair, and recombination pathways with specific roles including **A)** unwinding dsDNA [79-82], B**)** mediating replication restart (resulting in the reloading of the DnaB helicase in *E. coli* (blue)) [100], **C)** displace nucleoprotein complexes (for example RecA recombination filaments (pink nonagons) [82, 101], **D)** resolve secondary DNA structures (for example four-stranded guanine-quadruplex (G4) structures) [102], **E)** mediate repair mechanisms (for example transcription-coupled repair) [103]**, F)** detect DNA lesions (yellow star) in DNA [104], and **G)** unwind and resect DNA (for example homologous recombination initiation) [85].

2.2. Types and classification of helicases

Helicases and translocases can be classified based on the nucleic-acid substrate, or whether they are ‘active’ or ‘passive’. Further, helicases can be classified into six distinct families based on their structural biology. Here, we briefly introduce these topics below:

2.2.1. Directionality and nucleic-acid substrate

The polarity of DNA, defined by the 5' phosphate and 3' hydroxyl groups is exploited by helicases to selectively and precisely bind and manipulate DNA. This preference results in translocase or helicase activity occurring in a 5'-to-3' or 3'-to-5' direction. Whilst directionality is a hallmark of many helicases and translocases, some RNA helicases do not move in a directional manner. Instead, these helicases act as conformational switches that catalyze unwinding by binding to and prying the duplex region apart (reviewed in [105, 106]).

2.2.2. Active vs passive helicase

Two distinct mechanisms of helicase dsDNA unwinding are defined as ‘active’ or ‘passive’ and are differentiated by the ability of the helicase to exploit thermal fluctuations that open the dsDNA [107-109]. Passive helicases can exploit these fluctuations, by waiting at the single-stranded/double-stranded DNA (ss/dsDNA) junction for the double-strand DNA (dsDNA) region to transiently fray. Whereas, active helicases directly unwind DNA by interacting with the ss-dsDNA junction resulting in the destabilization of the duplex [107-109].

2.2.3. Structural

Helicases can also be classified based on characteristic amino-acid motifs and oligomeric states, resulting in six superfamilies (SF1-6) [110]. SF1 and SF2 helicases are typically monomeric, although most require oligomerization to actively unwind dsDNA (Figure 2.2) [111, 112]. SF1 forms one of the largest classes of helicases all consisting of a characteristic structural organization of four subdomains [111] (Figure 2.2A). These domains include two highly conserved RecA-like subdomains (1A and 2A) (H1 and H2 in Figure 2.2A), and two accessory domains (1B and 2B). SF1 can be further classified into three subfamilies; UvrD/Rep-, Pif1/RecD- and Upf1-like, distinguished based on their direction of translocation on ssDNA; SF1A translocate 3'-5' and SF1B translocate 5'-3'. SF2 includes both DNA and RNA helicases that also contain RecA-like domains, however, contain a diverse range of subdomains and no characteristic organization [113] (Figure 2.2B). SF3, SF4, SF5, and SF6 are ring-shaped enzymes that usually exist as hexamers (Figure 2.2C). In monomeric form, these helicases contain either an AAA+ fold (SF3, SF6) or a RecA-like domain (SF4, SF5) [114]. In hexameric form, these ring-shaped complexes translocate DNA in either a 3'-5' (AAA+ fold-containing) or 5'-3' direction (RecA-like domain-containing) at the replication fork to drive separation of the duplex DNA.

Below we discuss the best-described prokaryotic helicases and translocases and how recent single-molecule advancements have highlighted the activity of these diverse enzymes in genome maintenance pathways.

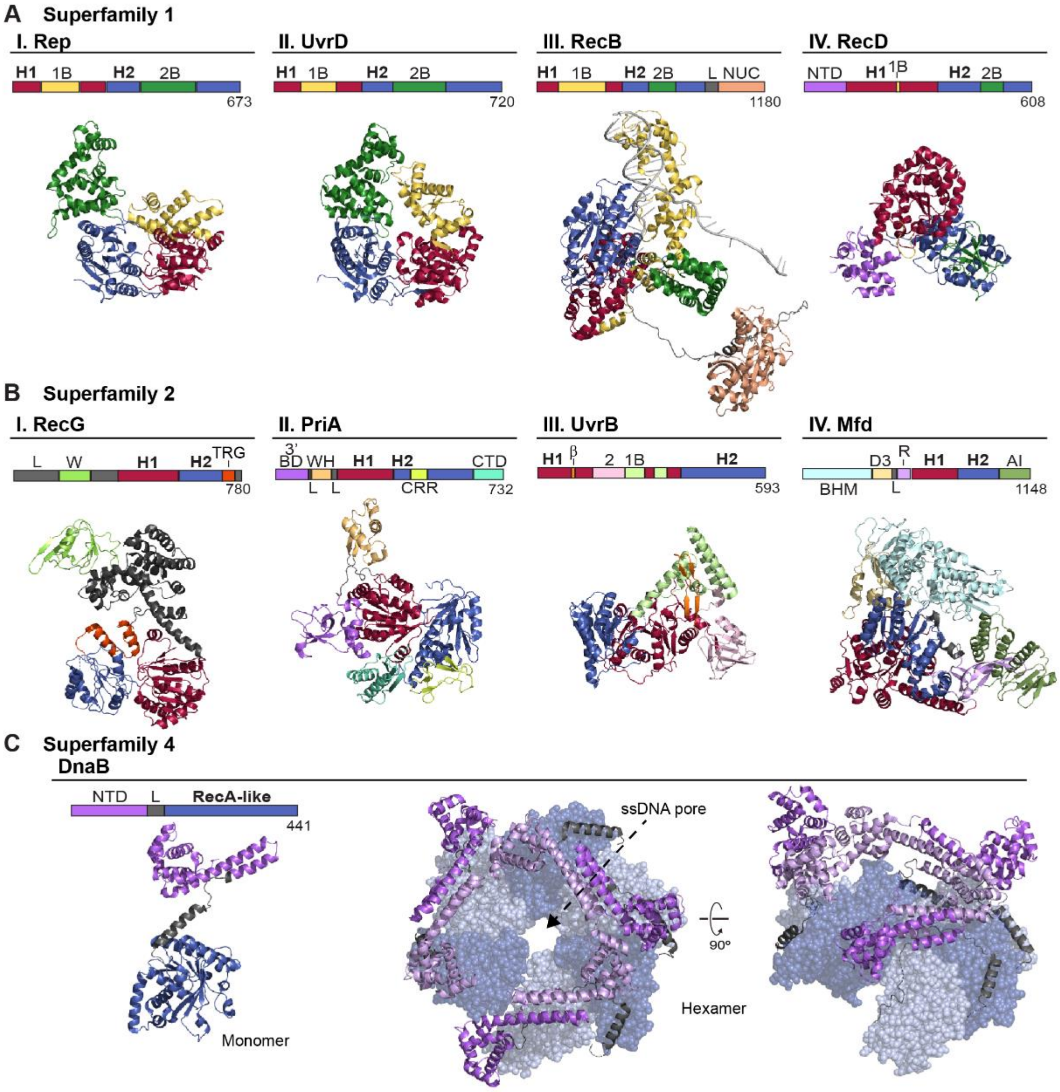


Figure 2.2: Structural classification of helicases discussed in this review. The schematics show the domain organization of each structure. In all structures, helicase domain 1 (H1) is red, and helicase domain 2 (H2) is dark blue. Accessory domains are colored according to their respective folds. **A)** Structures of superfamily 1 helicases (I) Rep (PDB 1UAA) [115], (II) UvrD (PDB 2IS2) [116], (III) RecB (PDB 5LD2) [117], and (IV) RecD (PDB 5LD2) [117]. **B)** Structures of superfamily 2 helicases (I) RecG (PDB 1GM5) [118], (II) PriA (PDB 4NL4) [119], (III) UvrB (PDB 2FDC) [120], and (IV) Mfd (PDB 6XEO) [121]. **C)** Structures of superfamily 4 helicase DnaB (PDB 4ESV) [122]. Hexameric DnaB NTD depicted as cartoon and RecA-like domain depicted as transparent surface structures from the top and 90° rotation side view. Alternating monomers depicted as light or dark color variations. (Abbreviations: L – Linker domain, NUC – Nuclease domain, NTD – N-terminal domain, W – wedge domain, TRG – translocation in RecG motif, 3' BD - 3' DNA binding domain, WH – Winged helix domain, CRR – Cysteine-rich region, CTD – C-terminal domain, β – β hairpin loop, BHM – UvrB homology module (binds UvrA), D3 – domain 3, R – RNAP binding domain, AI – Auto inhibition domain).

2.3. Replication

Successful cell proliferation in all organisms requires high-fidelity DNA replication of all genetic material. Duplication of the genome is performed by the replisome, a multi-protein machine capable of initiating DNA replication and translocating dsDNA [4-7]. During DNA replication, the dsDNA duplex is unwound to form two template strands that are subsequently replicated in a semi-discontinuous manner by the replisome [3, 8, 9]. The molecular mechanisms underlying DNA replication have been well characterized by extensive investigations of the *Escherichia coli* (*E. coli*) model replisome (reviewed by [123]).

Initiation of *E. coli* DNA replication requires the recruitment and assembly of the initiator proteins, DnaA, DnaB, and DnaC, at a unique locus called the replication origin, or *oriC* [24-27] (Figure 2.3). Initiation of this process results from the binding of the DnaA proteins that melt the duplex DNA to form an ssDNA bubble [24-26] (Figure 2.3A). The SF4 ring-shaped replicative helicase, DnaB (Figure 2.2C), is then opened and loaded onto the ssDNA bubble in complex with the helicase loader, DnaC [27, 31] (Figure 2.3B). Electron cryogenic microscopy (cryo-EM) studies of the DnaBC helicase loader complex showed that before binding DNA, six DnaC protomers bind to and crack open the DnaB hexamer [31]. Upon binding ssDNA, ATP hydrolysis by DnaC promotes closure of the DnaB ring around the DNA strand, forming a pre-translocation state capable of binding the DnaG primase [31]. Primase-DnaB interactions trigger the dissociation of DnaC and subsequent translocation of DnaB towards the fork junction [32] (Figure 2.3C). This translocation signifies the transition from the initiation to the elongation stage, during which the remaining components associate to form a functional replisome (Figure 2.3D).

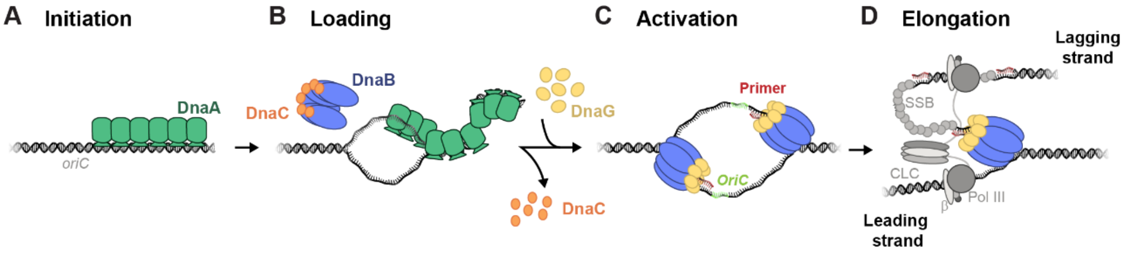


Figure 2.3: Initiation and elongation of replication. A) *Initiation,* DnaA (green) recognizes and binds *OriC.* **B)** *Loading*, melting of the dsDNA by DnaA creates an ssDNA bubble. Meanwhile, the DnaB hexamer (blue) is opened and loaded onto the ssDNA while in complex with the helicase loader DnaC (orange). **C)** *Activation,* binding of the DnaG primase (yellow) to DnaB triggers dissociation of DnaC and primer (red) extension in the direction of helicase translocation toward the fork junction. **D)** *Elongation, the* association of the replisome components (Pol III polymerase, β clamp, clamp loader complex (CLC), and SSB (grays)) forms a functional replisome capable of replicating the template DNA. Figure adapted from [124] and [125] with permission.

The antiparallel nature of DNA dictates that replication occurs in a semi-discontinuous fashion [8, 36]. The leading strand is synthesized continuously in the direction of travel of the fork, by incorporation of nucleotides at the 3' end [7]. Meanwhile, the lagging strand is synthesized discontinuously, through cycles of RNA primer synthesis and DNA extension by the DNA polymerase to form Okazaki fragments [7, 37, 38]. This textbook model of DNA replication assumes that the replisome is a stable entity that retains the same factors during the entire replication cycle. This view of an unchanging composition of the *E. coli* replisome has been overturned by a growing collection of single-molecule studies that provide access to sub-second temporal resolution and spatial resolution comparable to the size of the replication machinery [34, 75-77]. In this new paradigm, the consensus is that the replisome dynamically exchanges individual components of this machinery as it completes the replication cycle. Whereas the bulk studies suggest a highly stable complex, when viewed with powerful single-molecule tools, the picture emerges that single-stranded DNA binding protein (SSB) [75], the individual replicative polymerases, and the clamp-loader complex (PolIII*)[34, 76] rapidly exchange in a concentration-dependent manner. Additionally, individual trajectories of complexes executing leading or lagging-strand DNA synthesis revealed that each polymerase shows comparable average rates of synthesis, however, undergo stochastic pausing events resulting in switchable rates of synthesis [77]. Collectively, these observations present a fully discontinuous model, where both leading- and lagging-strand synthesis is often interrupted by the dynamic exchange of replisome components. Such a model explains how this macromolecule machine can rapidly respond to changes in cellular conditions influencing the availability of various factors or the quality of the template DNA.

How then does the replisome maintain overall composition during replication? Within a functional replisome, the DnaB helicase drives strand separation at the leading edge of the replication fork. Interestingly, *in vivo* observations of the DnaB helicase in live *E. coli* cells revealed that the helicase remains stable at the replication fork with an exchange time of approximately 900 s, even as the polymerases readily exchange [76]. Additionally, DnaB is known to bind multiple components of the replisome [33, 126-128]. Thus, DnaB may play a role in maintaining the integrity of the replisome as it acts as an interaction hub for the exchanging components of the replisome. How DnaB coordinates the exchange of components of the replisome remains unknown and awaits further investigation.

2.3.1. Rep

DNA replication occurs on crowded and often, damaged DNA templates. DnaB is faced with many potential roadblocks, including damage within the DNA template or nucleoprotein complexes, all of which can inhibit DNA synthesis resulting in genome instability [10, 11, 40-43]. The fate of the *E. coli* replisome at these sites has been investigated both *in vivo* and *in vitro* using a variety of model protein roadblocks, including tandem repressor-operator complexes, prokaryotic replication termination sites (Tus-*ter*), and stalled RNAP complexes [65, 66, 68, 70-73]. Additionally, the recruitment of accessory helicases has been shown to provide cells with a mechanism for overcoming nucleoprotein complexes [59-61].

The *E. coli* Rep helicase (herein referred to as Rep) is an SF1A helicase involved in the restart of stalled replication forks [59, 129, 130]. Similar to its structural homologs *E. coli* UvrD (Gram-negative bacteria) and *Bacillus stearothermophilus* (*B. stearothermophilus*) PcrA (Gram-positive bacteria), Rep consists of four subdomains (1A, 2A, 1B, 2B) (Figure 2.2A) [115, 116, 131]. The motor cores consist of two highly conserved RecA-like subdomains, 1A and 2A (H1 and H2 in Figure 2.2A I) [115]. ATP hydrolysis by these domains results in translocation along ssDNA in a 3' to 5' direction, with a conversion of one ATP hydrolyzed per nucleotide (nt) translocated [115, 116, 131-135]. Further, a 4-nt kinetic step size of Rep and UvrD has been resolved, with the non-uniform stepping caused by a slow step after four rapid 1-nt translocation steps [132-134].

Rep (and its homologs) has long served as a useful model for understanding how conformational changes within the protein regulate helicase activity. The 2B accessory domain can freely rotate in a swivel motion of 130° about a hinge region connected to the 2A subdomain [115]. The resulting structure adopts one of two extreme conformations termed ‘open’ and ‘closed’, with respect to the 1B subdomain [115, 136-140] (Figure 2.4A). Rotations of the UvrD 2B subdomain have recently been investigated by single-molecule Förster Resonance Energy Transfer (smFRET) [139]. In the absence of DNA, it was shown that the 2B subdomain of a FRET pair-labeled UvrD monomer can populate two intermediate states in addition to the extreme open and closed conformations that differ by a rotation of 160° [139]. Upon binding DNA, the more open state is observed, while dimerization shifts the state into the more closed conformation resulting in an active helicase [139]. Therefore, SF1A helicase and translocase activity are tightly coupled to conformational changes and oligomerization.

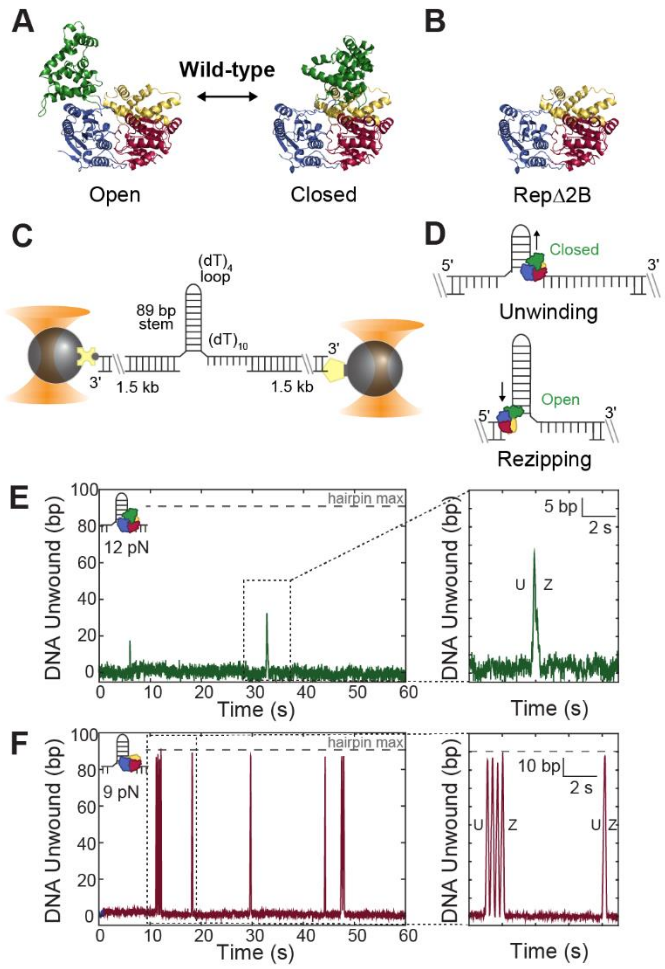


Figure 2.4: Removal of the Rep 2B subdomain activates monomer helicase activity A) The Rep wild-type (Rep WT) 2B subdomain (green, cartoon) can rotate about a hinge region between the 2B and 2A subdomains, forming either the open or closed conformations. (PDB 1UAA [115]) **B)** Structure of the RepΔ2B mutant. **C)** Schematic of dual-optical trap experimental setup. Two beads (gray spheres) held in traps (orange cones) stretch a DNA hairpin (black) at a constant tension by a biotin-streptavidin (dark gray circle - yellow cross) and digoxigenin-anti-digoxigenin (yellow pentagon – dark gray rectangle) linkages. **D)** Schematic of Rep WT (pink) monomer activity on optical trap DNA hairpin substrate. Rep WT 3' - 5' translocation with 2B subdomain in the closed results in DNA unwinding of the duplex. Rep WT translocation on the 5' strand, with 2B subdomain in the open conformation, results in rezipping behind the helicase (bottom). **E)** Representative optical trap data-trace of DNA hairpin unwinding by Rep WT dimer. Close-up inset of one round of unwinding (U) and rezipping (Z). **F)** Representative optical trap data-trace of processive unwinding by RepΔ2B monomer. Close-up inset of two rounds of full hairpin unwinding (U) (gray horizontal dashed line) and rezipping (Z). Figure adapted from [141] with permission.

Rep, UvrD, and PcrA monomers are capable of binding and translocating ssDNA, but they display no processive dsDNA unwinding activity [134, 139, 140, 142-149]. However, conformational control of the Rep 2B subdomain activates monomeric helicase activity. smFRET and optical tweezers assays demonstrated that intramolecular crosslinking of the 2B subdomain of both Rep and PcrA into the closed state activates highly processive monomer helicase capable of unwinding approximately 4 kb at rates of 138 bp/s [138]. Further, deletion of the 2B subdomain (RepΔ2B) also activates monomer helicase activity *in vitro* [141, 142, 146] (Figure 2.4B). Cells expressing RepΔ2B exhibit slower growth rates but are viable [142]. Recent optical-trap studies revealed that RepΔ2B monomers unwind dsDNA with greater processivity, reaching the limit of the hairpin substrate of approximately 90 bp, and at a higher speed (~160 bp/s) than activated Rep WT dimers (~ 30 bp at 70 bp/s) [141] (Figure 2.4C-F). Further investigations of RepΔ2B monomers unwinding 1.5 kb linear dsDNA substrates revealed a processivity of several hundred base pairs [141].

These assays also revealed that the unwinding processivity of RepΔ2B and Rep WT is limited by strand-switching activity, where the helicase alternates between the strands of the duplex DNA [141] (Figure 2.4A). The rate of strand switching is decreased in RepΔ2B mutants, which together with the faster unwinding rate, results in its increased processivity [141]. Strand switching by UvrD has previously been observed in assays combining optical tweezers and smFRET to simultaneously monitor dsDNA unwinding and the conformational states of FRET pair-labeled monomers. FRET traces of the labeled monomer and optical trap measurements revealed a correlation between closed conformation and unwinding of approximately 20 bp and transitions to a more open conformation and re-zipping of the duplex [140]. Free energy simulations attributed the re-zipping activity to a tilted conformation of the 2B subdomain, to form an opened metastable state that allows the ssDNA to escape and the motor domains to strand-switch [150]. Additionally, optical-trap assays revealed greater unwinding speeds with increasing force for both Rep WT and to a greater extent RepΔ2B, indicating a potential role of the 2B subdomain in regulating unwinding speed [141]. These mechanisms appear to be prevalent across organisms, as revealed by magnetic-trap assays investigating the unwinding activity of *Saccharomyces cerevisiae (S. cerevisiae)* SF1B helicase Pif1 [151]. Here, both the unwinding rate and processivity of Pif1 monomers were significantly enhanced when a force of 9 pN was applied [151]. Additionally, single-molecule total internal reflectance fluorescence (smTIRF) microscopy investigations of SF1A human F-box-containing DNA helicase (FBH1) revealed monomeric helicase activity on short dsDNA oligos with 3' overhangs [152]. Thus, the closure of the 2B subdomain of SF1 helicases relative to the motor core switches on the helicase activity of the protein.

Interestingly, *in vitro* ensemble studies have shown that the Rep 2B subdomain is essential for the displacement of nucleoprotein complexes [68, 130]. Positioned at the leading edge of these helicases, the 2B subdomain may play an important role in separating the primary functions of SF1A helicases. Specifically for Rep, these functions include nucleoprotein complex displacement and unwinding dsDNA. It is likely that in monomeric form, Rep acts as an ssDNA translocase using the 2B subdomain to displace proteins bound to DNA ahead of the replication fork as it makes contact with potential roadblocks [141]. The binding of another Rep monomer then activates the helicase through conformational rotation of the 2B subdomain, allowing the protein to unwind long stretches of the duplex [134, 139, 142-144, 146, 148, 149].

How and when does Rep rescue stalled replication forks? Rep removes roadblocks in the path of the replisome. Rep can physically tether itself to the replisome, by interactions between its hydrophobic C-terminus and DnaB [59, 62]. Whereas the hexameric structure of DnaB suggests a potential of six binding sites for Rep monomers [153, 154], in the absence of a detailed structural understanding of the Rep-DnaB interaction, it is unclear whether all six binding sites are simultaneously occupied. Single-molecule live-cell imaging has shed light on this issue. Early live-cell fluorescence microscopy studies failed to detect the presence of Rep monomers at progressing replication forks [153]. However, the authors proposed that either Rep interacts with the progressing replication fork in less than three copies, or Rep is recruited to the replication fork by DnaB upon fork arrest [153]. Recent *in vivo* single-molecule imaging revealed that Rep indeed co-localizes with approximately 70% of replication forks and does so with a hexameric stoichiometry via its C-terminus [155]. This interaction was also observed to be transient, with Rep associating in an ATP-dependent manner and then translocating away from the fork [155].

Dynamic association of Rep with the progressing replication fork suggests the protein is continually recruited to DnaB, as sites become available after Rep molecules bind to ssDNA and translocate ahead of the replisome [155]. Reduced co-localization of Rep with the replisome was observed in cells lacking the replication restart protein, PriC, suggesting that Rep assists in replisome reloading in the event forks stall and dissociate [155]. These authors proposed that it is likely that Rep plays two important roles at the replisome *in vivo*; 1) monitoring the replisome for potential obstacles by interacting with the DnaB helicase, and, 2) recruiting additional protein factors to initiate replisome reloading in the event of replisome collapse.

Disruption of nucleoprotein complexes by accessory helicases likely requires physical cooperation with the replisome [59, 155]. The eukaryotic Rep homolog, Pif1, is important in DNA-metabolism pathways that require the removal of nucleoprotein complexes [156]. Single-molecule TIRF microscopy investigations, using reconstituted *S. cerevisiae* replisomes and a catalytically "dead" Cas9 protein (dCas9) as a replication roadblock, revealed that Pif1 removes model nucleoprotein complexes in a replisome-dependent manner [74, 156]. Thus, cooperation of SF1 accessory helicases with the replisome is likely a requirement to ensure the successful displacement of potential roadblocks.

Replication forks can also be rescued by UvrD [101, 157-159]. UvrD has been shown to displace a variety of nucleoprotein complexes from the path of the replication fork, including RecA recombinase filaments, replication terminator proteins Tus from *ter* sites, and RNAP [68, 101, 157, 160, 161]. Rep and UvrD perform redundant roles in the removal of roadblocks although the precise mechanisms differ [59, 60, 162]. Although *rep uvrD* double mutants are lethal to the cell, single *rep* and *uvrD* mutations are viable, indicating that both proteins can function independently of each other in stalled replication rescue [59, 163, 164].

In addition to the threat of nucleoprotein complexes acting as roadblocks, potentially deleterious secondary structures of DNA also pose a barrier to replication. Recently, smFRET assays revealed a novel function of Rep, but not UvrD, in resolving four-stranded guanine-quadruplex

(G4) structures [102]. These studies suggest that Rep, and functional homologs, may act as a shield for the replisome, resolving potential roadblocks before the replisome encounters them to ensure successful replication.

2.3.2. RecG

The collapse of the replication fork at sites of roadblocks can result in deleterious DNA double-strand breaks (DSBs) [100, 165-171].To prevent these potentially fatal occurrences, cells have evolved mechanisms that detect and process abandoned replication forks for downstream replication restart mechanisms [171-173]. One replication rescue pathway in *E. coli* involves the SF2 RecG DNA helicase (Figure 2.2B I) [118, 171]. RecG catalyzes the backward movement (regression) of a stalled fork to form a Holliday Junction (HJ) that is then further processed by RuvABC to resemble DNA substrates ideal for replisome reassembly (reviewed by [171]).

RecG catalyzes both the unwinding and rewinding of a stalled fork to regress away from sites of DNA damage [174]. The duplex rewinding activity was probed in early single-molecule magnetic tweezer assays where RecG, and its functional analog, bacteriophage T4 UvsW, displayed rewinding activity of a hairpin DNA structure whilst opposing forces of up to 35 pN [174, 175]. Further, the construction of a stalled fork substrate, containing a hairpin with duplex tracks, demonstrated efficient unwinding and rewinding of the substrate to form a fully regressed fork [174].

RecG is a known interacting partner of SSB. Biochemical investigations of this interaction have revealed that the activity of RecG is enhanced and influenced by SSB [87, 176, 177]. Recently, atomic-force microscopy (AFM) investigations have revealed that this interaction triggers a remodeling of the RecG structure, thus altering the binding mode of the helicase [178-180]. Interestingly, this remodeling triggers an SSB-dependent translocation of RecG on the parental duplex DNA [178]. Further, high-speed AFM revealed ATP-dependent fork regression by RecG helicase activity is coupled to the displacement of SSB from the ssDNA [180].

Together, these investigations suggest two mechanisms by which RecG drives stalled fork regression [179]. (1) RecG loading can be ssDNA-mediated, where the helicase is loaded onto the fork and regression is initiated. (2) Alternatively, RecG loading can be SSB-mediated, where RecG is recruited to the fork by SSB and subsequently remodeled. The conformational change in the helicase alters the binding mode, allowing RecG to translocate on the parental duplex DNA, an activity that does not occur in the absence of SSB. In the presence of ATP, RecG then drives fork regression by coupling DNA unwinding with duplex rewinding whilst displacing any bound obstacles to duplex [179, 180].

2.3.3. PriA

Processing of stalled forks by initial replication restart mechanisms results in the production of DNA substrates ideal for replisome reloading. One protein involved in replisome reloading pathways in *E. coli* is the 3'-5' SF2 PriA DNA helicase (Figure 2.2B II) [100, 172]. PriA can initiate replication restart by reloading DnaB at forked DNA structures by a mechanism independent of the *oriC* [100, 181].

PriA recognizes and interacts with the SSB-coated ssDNA at abandoned replication forks [86, 100, 119, 182, 183]. Like RecG, biochemical characterization of the PriA-SSB interaction has also shown that SSB can enhance the DNA unwinding activity of PriA at abandoned replication forks [86, 184]. In atomic-force microscopy (AFM) investigations, Wang and co-workers [185] probed the influence of SSB on the ability of PriA to recognize various forked DNA substrates. In the absence of SSB and ATP, 13% of PriA proteins preferentially bound to the forked region of DNA constructs containing a gap in the leading strand and a 3' flank, compared to 8% binding to constructs containing a lagging strand gap and 5' flank [185]. However, in the presence of SSB, the specificity of binding shifted to the duplex region of the DNA constructs in 60% of complexes, with no apparent preference for the flank type [185]. Based on similar findings for RecG, the authors proposed that the SSB-mediated enhancement of PriA activity is triggered by a conformational change that shifts the selectivity of the helicase for a more favorable substrate for replication restart [185].

2.4. Repair

Uncorrected DNA damage can interfere with DNA metabolic pathways resulting in cell death or propagation of mutations in daughter cells [186, 187]. To avoid such deleterious occurrences, organisms have evolved mechanisms that repair the damage and maintain genetic stability. Numerous repair mechanisms have now been characterized and all involve four key steps: (1) damage recognition, (2) damage verification, (3) damage excision, and (4) repair synthesis and ligation. Recently, single-molecule investigations of the key steps in repair pathways have revealed the fundamental mechanisms of protein-protein and protein-DNA interactions in the repair of damaged DNA. Here, we focus on the activation and roles of the UvrD helicase in various repair pathways, DNA lesion detection by UvrB, and the extensive single-molecule characterization of the Mfd translocase in transcription-coupled repair.

2.4.1. UvrD

The SF1A *E. coli* UvrD helicase is primarily involved in global genomic nucleotide-excision repair (GG-NER) [188, 189], methyl-directed mismatch repair (MMR) [190] and transcription-coupled nucleotide excision repair (TC-NER) [158] (discussed below), in addition to roles in replication (discussed above) and recombination (discussed below) [101, 157, 191]. Like its structural homologs, Rep and PcrA, the activity of UvrD is regulated by its 2B domain, and its stoichiometry (Figure 2.2A II). UvrD monomers translocate on ssDNA at speeds of approximately 190 nt/s in an ATP-dependent manner, with no apparent helicase activity on dsDNA [132, 133, 144, 145, 148, 149, 192, 193]. The formation of a homodimer results in a functional helicase capable of unwinding dsDNA *in vitro* at a rate of ~70 bp/s [144, 148, 149, 193-196].

UvrD is a well-developed model for understanding how helicase activity is regulated by protein-binding partners. Specifically, UvrD helicase activation has been observed in *in vitro* biochemical assays by the *E. coli* MutL dimer, a regulatory protein involved in methyl-directed mismatch repair [88, 197, 198]. Recent smFRET assays have captured the activation of UvrD monomers by MutL [199, 200]. Monomeric UvrD immobilized to the coverslip surface exhibited no helicase activity when bound to a dsDNA substrate containing a 3'-ssDNA tail. However, in the presence of MutL, helicase activity was stimulated resulting in an unwinding rate of ~40 bp/s, a rate comparable to that seen in single-molecule smFRET assays of UvrD homodimers at similar ATP concentrations [139, 199]. Further investigations revealed that MutL binding stimulated a shift in the conformational state of the 2B domain to an intermediate state, between the open and closed states, capable of activating helicase unwinding [200]. Interestingly, this helicase activation mechanism is specific for UvrD, as MutL was unable to stimulate Rep monomer activity, suggesting a role for specific protein–protein interactions [199, 200]. Recent *in vitro* smTIRF investigations of MMR revealed that long-lived MutL dimers on DNA containing a mismatch can capture UvrD from solution and stimulate its helicase activity to result in strand-displacement of the mismatch-containing strand [201].

Interestingly, early AFM and smTIRF investigations revealed that monomeric PcrA can processively unwind plasmid DNA when activated by the plasmid replication initiation protein, RepD [202, 203]. RepD binds and nicks the plasmid DNA at specific sequences to initiate replication, and is proposed to complex with PcrA to activate its helicase activity at the nick site [202-205]. Thus, the emerging picture is that stimulation of SF1A helicase activity can occur by the formation of homodimers, or heterodimers comprising of accessory proteins specific to the individual pathways these helicases act in.

2.4.2. UvrB

Some helicases participating in nucleotide-excision repair have the additional ability to specifically recognize DNA damage. Here, the limited helicase activity of UvrB and its functional eukaryotic homolog (XPD) is sensitive to the presence of DNA damage. The *E. coli* UvrB SF2 helicase-like protein verifies damaged nucleotides (primarily bulky, helix-distorting lesions) in chromosomal DNA with single-nucleotide resolution in the nucleotide-excision repair pathway (Figure 2.2B III). UvrB exhibits very weak, cryptic helicase activity that is used to sense DNA damage [104, 206, 207]. Removal of bulky DNA lesions by the NER process begins with lesion recognition by UvrA and UvrB ($UvrA_2B_2$) [208-211] (Figure 2.5A). UvrB is loaded on the DNA for verification of the lesion, $UvrA_2$ is ejected and UvrC is recruited to perform incision of the DNA lesion [212-214].

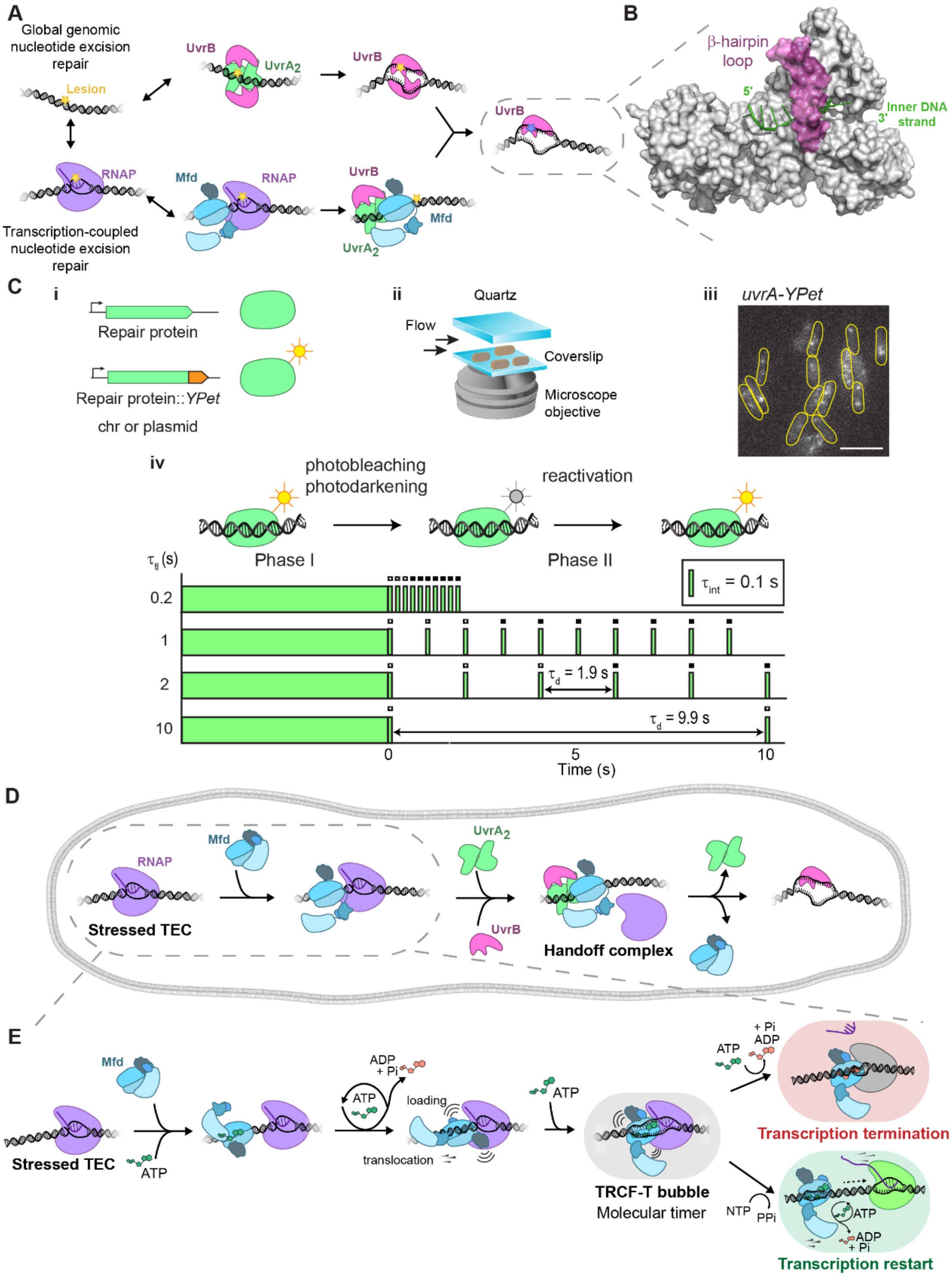


Figure 2.5: UvrB and Mfd activity in nucleotide-excision repair. A) Nucleotide-excision repair in *E. coli* occurs via global surveillance for DNA damage by UvrA$_2$ (green) and UvrB (pink), and by TEC (purple) as it transcribes DNA. Figure adapted from [215] with permission. **B)** Crystal structure of UvrB in complex with ssDNA (PDB 2FDC, [120]) highlighting the β-hairpin loop (pink) clamping down on the inner DNA strand (green). **C)** *In vivo* single-molecule imaging strategy for prolonging the observation of fluorescently-labeled repair protein (i) in live *E. coli* cells. (iii) Cells expressing fluorescent UvrA-Ypet. Figure adapted from [215] with permission. **D)** *In* vivo single-molecule live-cell model of transcription-coupled repair. Figure adapted from [216] with permission. **E)** Proposed *in vitro* model for Mfd-mediated transcription modulation. Figure

adapted from [217] with permission.

The initial lesion-recognition step of NER has been extensively characterized *in vitro* using purified proteins [208-211, 218, 219]. Resulting from these studies was the acceptance that UvrA and UvrB work together to scan the genome for UV-induced DNA lesions. However, recent *in vitro* single-molecule imaging studies have revealed that quantum dot-labeled (Qdot-labeled) UvrBC complexes can bind dsDNA tightropes independently of UvrA [220]. Here, Qdot-UvrB exhibited robust binding to dsDNA only in the presence of differentially labeled UvrC [220]. Further dsDNA tightrope investigations showed that this binding was preferential to DNA damage sites, similar to that of UvrAB complexes [221]. These studies suggested that UvrBC complexes are capable of scanning long stretches of DNA and recognizing damage. This observation was further supported by single-molecule *in vivo* imaging of UvrBC complexes in *E. coli* cells, where stable UvrBC-eGFP foci were observed in the absence of UvrA after exposure to UV [221]. These single-molecule studies have revealed additional mechanisms for DNA damage recognition during NER and provide the best evidence for a helicase (UvrB) scanning the genome and recognizing DNA damage.

How do helicases detect and verify the lesion? Insight into these mechanisms has been revealed through structural investigations of UvrB [120, 222, 223]. A potential site of lesion verification within UvrB involves a narrow passage, formed by the β– hairpin loop and 1b domain, that ssDNA must pass through for successful translocation [120, 222]. The β– hairpin loop is inserted between the two DNA strands forming a partially unwound “bubble” where the loop clamps down onto the strand termed the ‘inner’ strand (Figure 2.5B) [120, 224, 225]. ATP hydrolysis by UvrB triggers a conformational change within the helicase domains, resulting in the movement of the ssDNA through the constriction point [120, 222, 223]. It has been suggested that undamaged nucleobases can readily pass through this channel, however bulky or rigid, adjacent bases cannot and result in arrested UvrB translocation [223]. This arrested state forms the pre-incision complex required for the recruitment of incision enzymes in NER.

The eukaryotic homolog, XPD, also arrests translocation at sites of bulky lesions in a strand- and site-specific manner [226]. Structural investigations attributed this arrest to a narrow pore formed by the Arch, FeS and helicase 1 domains [227-229]. However, additional biochemical investigations of XPD lesion verification, have shown that mutations in key residues within the leading edge of the protein prevent site-specific arrest of XPD at bulky lesions [230, 231]. Thus, the authors proposed that a sensor pocket, outside of the central helicase pore, can scan the translocating DNA strand for lesions prior to entering the pore [230]. Additionally, smTIRF investigations revealed that the pore-forming domain of XPD, the Arch domain, is intrinsically dynamic producing open- and closed-pore states [232]. DNA binding does not change these dynamics, however, upon DNA damage recognition, the lifetime of the closed state is increased. These studies illustrate how ATP-driven translocation on ssDNA might play a role in the recognition of bulky lesions in DNA.

2.4.3. Mfd

Transcription is a complex process whereby dsDNA is transcribed by RNA polymerase (RNAP) to form RNA (reviewed by [233]). Prone to pausing and irreversible stalling, RNAP activity is regulated by several transcription-associated factors, one of which is a highly conserved SF2 translocase. The bacterial Mfd translocase, with functional homologs in most eukaryotes (Rad26 in *S. cerevisiae [234]* and Cockayne syndrome B protein CSB in humans [235]) modulates transcription activity through its translocase function, involving processive hydrolysis of ATP on double-stranded DNA [89, 121, 236-241] (Figure 2.2B IV). The best-characterized transcription-repair coupling factor (TRCF), Mfd exemplifies roles in recruiting the DNA repair helicase (UvrB) to the site of DNA damage and modulating catalytic activity of the transcription machinery (Figure 2.5A, d, e). Here we discuss three roles of Mfd in mediating: (1) transcription-coupled DNA repair, (2) transcription restart, and (3) mutagenesis, and propose a model to unify these three roles in the context of its known catalytic activities.

The encounter of a ternary transcription elongation complex (EC) with a bulky lesion on template DNA leads to irreversible transcriptional stalling [242-244]. TRCFs terminate transcription at sites of distressed transcription and recruit DNA repair machinery in a reaction called transcription-coupled DNA repair (Figure 2.5A) [245, 246]. In bacteria, Mfd recognizes stalled RNAPs and recruits the nucleotide excision repair (NER) factors to these sites[244]. In humans, defects in the TRCF – the so-called Cockayne syndrome B protein – lead to Cockayne syndrome [235, 247].

Mfd is an SFII translocase bearing homology to the RecG translocase[248, 249], and chromatin remodellers [121]. Existing in an autoinhibited conformation in its apo state [241, 249], engagement of Mfd with the EC leads to activation in which the DNA binding motor domains bind upstream DNA (Figure 2.5D, E) [237]. Processive ATP hydrolysis by Mfd allows it to translocate to a stalled elongation complex on the upstream edge of the EC [237, 250]. A series of ATP– and ADP– bound structures of RNAP-tethered Mfd approaching the EC have revealed the precise mechanism of double-strand DNA translocation of Mfd [237]. Much like chromatin remodelers[251], translocase domain 1 (domain 5 in Mfd) (H1 in Figure 2.2B IV) of the SFII translocase tracks the guide strand, in this case, the non-template strand, whereas translocase domain 2 (domain 6 in Mfd) (H2 in Figure 2.2B IV) swings between two conformations distinguished by the nucleotide ATP or ADP status on the tracking strand, in this case, the template strand in the 3’–5’ direction [237]. This series of structures led the authors to propose that Mfd translocates one nucleotide per ATP hydrolysis cycle. Further, the authors demonstrated that the encounter of the translocating Mfd with the nucleotide–deprived elongation complex leads to global remodeling of the RNAP [237]. Based on available structures, biochemical investigations, and *in vitro* single-molecule assays this translocase activity has been proposed to be essential for pushing RNAP off the DNA leading to transcription termination [89, 238, 252, 253], although a structural description of these transcription termination intermediates is awaited.

Meanwhile, exposure of the N-terminus of Mfd allows the recruitment of the NER proteins UvrAB

(Figure 2.5A, D) [254]. Intensive investigations of Mfd based on biochemical and biophysical studies have generated *in vitro* and *in vivo* models for how Mfd delivers the UvrB helicase to the site of a distressed elongation complex. Based on the ability of Mfd to pull down UvrA and UvrB on an Mfd-affinity column, and the known roles of UvrA and UvrB in damage recognition and verification, Selby and Sancar [244] proposed a model for handoff of the stall site to UvrB during transcription-coupled DNA repair in which Mfd releases RNAP and recruits UvrA$_2$B to the stall site, followed by simultaneous dissociation of Mfd, and UvrA$_2$ leaving behind the UvrB-DNA pre-incision complex (Figure 2.5A).

The ability to synchronize reactions and monitor events with high temporal and spatial resolution was essential to subsequently test these ideas. Single-molecule nanomanipulation assays provided remarkable insight into this reaction, specifically that Mfd in the presence of ATP forms a long-lived complex with the EC that can processively translocate on several thousands of base pairs on DNA [238, 252]. Following recognition of the EC by Mfd, the RNA transcript is lost from the complex [238]. Subsequent studies elucidated that UvrA(B) can intercept this translocating Mfd-RNAP complex on DNA and lead to the release of the RNAP[239]. The authors proposed that (1) a transient UvrB-UvrA-UvrA-Mfd-RNAP complex is formed following initial recognition of the translocating Mfd-RNAP complex by UvrA$_2$B, which converts to a UvrA$_2$B complex following the dissociation of Mfd and RNAP observed in this study, and (2) that this UvrA$_2$B complex does not recruit a second UvrB molecule[239]. In this model, the question of precisely how a single UvrB is loaded in a strand- and site-specific manner to form the pre-incision complex remains unanswered.

Observations of fluorescently-tagged Mfd in living cells revealed that Mfd is present to the extent of 22 fluorescent copies per cell in rich medium [95]. Further, these observations made in the living cellular environment, confirmed *in vitro* biochemical observations that the activity of Mfd is highly auto-regulated [95, 241, 255-257]. Disruptions in the auto-regulation of Mfd by way of introduction of a mutation in the ATPase motif of Mfd (Mfd (E730Q)) led to promiscuous DNA binding in cells [95]. The Mfd (L499R) mutant possessing a defective RNAP interaction domain was found not to interact stably with ECs [95]. Importantly, this study highlighted for the first time that Mfd interacts with RNAP in the elongation phase even in the absence of exogenous DNA damage, suggesting that this translocase serves as a house-keeping enzyme recognizing distressed elongation complexes, and not DNA damage *per se*. In agreement with this suggestion, Mfd was found to associate with RNAP on genomic loci encoding highly structured RNAs [258].

To understand the handoff between Mfd-RNAP and UvrAB leading to the pre-incision complex, measurements of the DNA-binding lifetimes of fluorescently-tagged Mfd and UvrA in single-molecule live-cell studies [215, 216] were performed using a stroboscopic illumination approach that permits reliable observation of long-lived DNA binding events in cells [259] (Figure 2.5C). Briefly, in this method, the background from the entire population of fluorescently-tagged repair factors is first reduced by exposure to laser light (Figure 2.5C IV). This exposure allows the fluorescent protein to either reversibly darken, or photobleach. Stochastic reactivation of individual fluorescent proteins allows reliable observation of individual binding events in the absence of the cellular background. Measuring the lifetime of individual interactions enables the construction of a cumulative residence time distribution which enables measurement of the lifetime of the interaction. Measurements of the DNA binding lifetimes of Mfd and UvrA in cells encoding structural or catalytic mutants of UvrA, UvrB, or Mfd resulted in the elucidation of a network of interactions yielding the conclusion that an Mfd-UvrA-UvrA-UvrB constitutes a hand-off complex *in vivo* (Figure 2.5D). Here, unlike the *in vitro* observations where UvrA alone could displace Mfd from the Mfd-RNAP complex [239], in cells lacking UvrB, Mfd was arrested on the DNA with a lifetime of ~5 minutes [216]. In contrast, the lifetime of EC-bound Mfd in the presence of the full complement of repair proteins was ~18 s [95]. Further, inside cells, the ability of UvrB to be loaded onto the DNA to form a pre-incision complex was essential for the release of Mfd from the DNA-RNAP-Mfd-UvrA$_2$-UvrB complex. Complementary measurements with fluorescently-tagged UvrA revealed a similar dependence on the residence time of UvrA on UvrB [215]. Together, the data allowed the authors to conclude that *in vivo* the DNA-RNAP-Mfd-UvrA$_2$-UvrB complex is converted to the UvrB-DNA pre-incision complex following ATP hydrolysis by UvrA$_2$ and successful loading of UvrB on the DNA (Figure 2.5D) [95, 215, 216]. These studies provided evidence for a ‘UvrB-facilitated dissociation’ model for the hand-off of the stall site to the repair machinery [216]. This model is similar to the model predicted by the Sancar lab decades earlier [244] and overcomes the limitations of the *in vitro* model that fails to accommodate strand- and site-specific loading of UvrB. Precisely how Mfd directs the loading of UvrB in the Mfd-UvrA$_2$-UvrB complex remains to be elucidated from a structural perspective.

Translocation by Mfd is further required for re-starting transcription. Biochemical foot-printing studies initially established that Mfd binds the upstream edge of the transcription bubble and causes translocation of the RNAP forward [89]. These were followed by single-molecule optical tweezers experiments that further elaborated this interaction [236]. The authors discovered that following initial engagement, Mfd rescued a paused EC, allowing it to resume transcription. Having released RNAP from a paused state, Mfd was interpreted to follow in the wake of this EC, and catch up with it at the next pause site [236]. This ‘release and catch up’ model proposed that Mfd can dynamically modulate transcription activity. Since Mfd translocates on dsDNA at a speed of 7 bp/s and elongating RNAP translocates at ~15 bp/s, the authors proposed a “speed-gating” mechanism that allows Mfd to kinetically discriminate between productively elongating and stalled ECs [236].

Puzzlingly, in addition to its role in promoting error-free repair, Mfd also promotes mutagenesis in cells [13, 260-264]. The molecular mechanism of this mutagenesis is poorly understood, but recent work has revealed that the association of Mfd with the elongating RNAP leads to the formation of a negatively supercoiled domain between the footprint of Mfd and the EC [260]. This negative supercoiling enables the 5’ end of the RNA transcript to invade the dsDNA to form an R-loop. Since both wild-type Mfd and the transcription-termination deficient R953A mutant support the formation of R-loops both *in vitro* and *in vivo*, the authors proposed that the activity leading to R-loop formation is distinct from Mfd's translocase activity [260].

A unified understanding of the three roles of Mfd – in transcription-coupled DNA repair, promoting transcription restart and R-loop formation remains unpublished. However, a recent study using a combination of single-molecule FRET, and cryo-EM provides insight into how the three activities are reconciled [217]. Here, the authors performed smFRET measurements on ECs assembled on synthetic transcription bubbles carrying a FRET dye pair on the upstream template DNA and the RNA transcript. By monitoring the smFRET signal as a proxy for the distance between the two dyes, the authors investigated the response of the EC to Mfd in the presence of wildtype or translocase deficient Mfd(R953A) in the presence of ATP or the slowly-hydrolyzable analog (ATPγS). The results revealed that the translocase-deficient R953A mutant, as well as Mfd-ATPγS, exhibited a limited ability to forward translocate the EC leading to the formation of a catalytically poised state, termed a 'molecular timer' intermediate (Figure 2.5E). ATP hydrolysis and the presence of an intact R953 enabled exit from this state accompanied by loss of the RNA transcript. Further, surface-plasmon resonance and cryo-EM studies of Mfd binding to mimics of the upstream DNA revealed that Mfd specifically binds and melts the upstream DNA and key residues in its translocation motif (R953 and R929) make specific contacts with the template overhang. Here, the Mfd-ADP-AlF$_x$ bound structure in complex with a DNA substrate containing a template strand overhang suggested that Mfd pulls the ssDNA consistent with the smFRET experiments presented here.

A consensus model of the initial stages of recognition of stalled ECs can be proposed as follows (Figure 2.5E): first, active elongation complexes (ECs) encounter DNA damage that inhibits RNA synthesis leading to an arrested complex. Mfd is recruited to the site of the RNAP through the interaction of the RNAP interacting domain (domain 4) with the β subunit of RNAP [89, 237]. Mfd is normally auto-inhibited in cells, and this interaction with RNAP releases the auto-inhibition of Mfd enabling subsequent engagement of the motor domains with the upstream DNA in the EC [95, 241, 254-257]. RNAP tethered Mfd can translocate up to the polymerase through processive ATP hydrolysis, where hydrolysis of one ATP enables forward translocation by one base pair, resulting in initial global remodeling of the EC [237]. Upon gaining access to the upstream edge of the transcription bubble, ATP-bound Mfd further melts the DNA and enlarges the transcription bubble, switching from a purely dsDNA binding mode as documented previously[237] to a partial dsDNA-ssDNA binding mode[217] (Figure 2.5E). In so doing, Mfd can bring a backtracked EC into a catalytic register, and successful incorporation of NTPs can enable transcription to restart [89, 236]. In this state, since the RNAP translocates on template DNA faster than Mfd, the EC escapes being a substrate for transcription termination. However, in the case of irreversibly stalled ECs, the translocase motif in the Mfd-ATP bound complex makes additional contacts with the template strand DNA and scrunches it. Hydrolysis of this bound ATP enables displacement of the transcript, initiating rewinding of the transcription bubble, and consequently orchestrating transcription termination[217] (Figure 2.5E). The observation that Mfd can melt DNA at the upstream edge of the fork also explains how RNAP-tethered Mfd can maintain a negatively supercoiled domain, which has subsequently been shown to enable R-loop formation [260].

2.5. Recombination

Unrepaired DNA DSBs are deleterious to the cell. Repair of DSBs via homologous recombination (HR) is a conserved repair pathway that resolves DSBs in an error-free manner. In bacteria, the RecA recombinase mediates homologous recombination through the formation of extensive nucleoprotein filaments. Formation of these filaments is predicated on two helicase activities: (1) the processive action of a helicase-exonuclease complex called RecBCD leading to the resection of DSB ends to produce the single-strand overhang substrates on which RecA can nucleate and subsequently polymerize, and (2) the action of UvrD which regulates the growth of the RecA nucleoprotein to suppress promiscuous recombination, through the disassembly of the nucleoprotein. Here we discuss the roles of these two helicases in mediating recombinational repair.

2.5.1. RecBCD

In *E. coli*, the helicase/nuclease RecBCD complex is responsible for the initiation of the HR process [85]. This molecular machine recognizes and binds to the ends of DSBs leading to the resection of the dsDNA in preparation for strand invasion. RecBCD contains three subunits, including two SF1 helicase domains: RecB, a 3'-5' SF1A helicase and nuclease (Figure 2.2A III), and RecD, a 5'-3' SF1B helicase (Figure 2.2A IV). The RecC subunit is responsible for complex stability and recognition of the regulatory *cis*-acting cross-over hotspot instigator (Chi) sequences. Recognition of a Chi sequence triggers the RecB subunit of the enzyme complex to switch from nicking the 3' end to nicking the 5' end whilst unwinding the dsDNA [265-269].

RecBCD initiates HR by binding to the blunt end of a DSB and simultaneously unwinding the duplex whilst nicking the 3' tail. The opposing polarities of the RecB and RecD subunits allow the complex to translocate along the dsDNA on opposing strands [270, 271]. Early optical trap experiments revealed that before Chi recognition, the RecBCD complex translocated dsDNA at rates of approximately 1000 bp/s [266, 272, 273]. Fast translocation speeds were confirmed in an *in vivo* fluorescence model of RecBCD end resection activity in live *E. coli* cells [274]. Additionally, processivities of approximately 100 kb were detected, and are significantly decreased upon deletion of the RecD domain [274]. Further, recognition of Chi by RecC causes the complex to pause, for approximately 5 s, and resume translocation at a significantly reduced rate (~300 bp/s) [266, 272, 273, 275]. Optical trap studies attributed this decreased translocation rate to a switch in the leading helicase motor subunit, where the slower RecB subunit leads after Chi [275].

Recently, the activity of the individual subunits of RecBCD was observed using a modified dual trap optical tweezers assay [276]. Here, a symmetric DNA substrate containing a blunt end and two dsDNA tracks is attached to beads held in separate optical traps (Figure 2.6A). This setup could detect the opposite unwinding polarities of the RecB and RecD subunits as a corresponding decrease in the dsDNA extension. Modification of the dsDNA tracks to produce asymmetric ends, where either the RecB or RecD track is significantly shorter than the other, resulted in the ability to detect the individual activity of the subunits in the context of the entire complex (Figure 2.6B). These assays revealed that each of the subunits unwind dsDNA at a rate of ~400 bp/s, a 2-3 fold decrease when compared to the entire complex (~1100 bp/s) unwinding the symmetric DNA substrate [276] (Figure 2.6C-E). Further, purified RecD subunits were unable to unwind a dsDNA fork. These results indicate that RecBCD is a synergistic complex, where it requires each functional subunit for efficient and processive DNA unwinding [276].

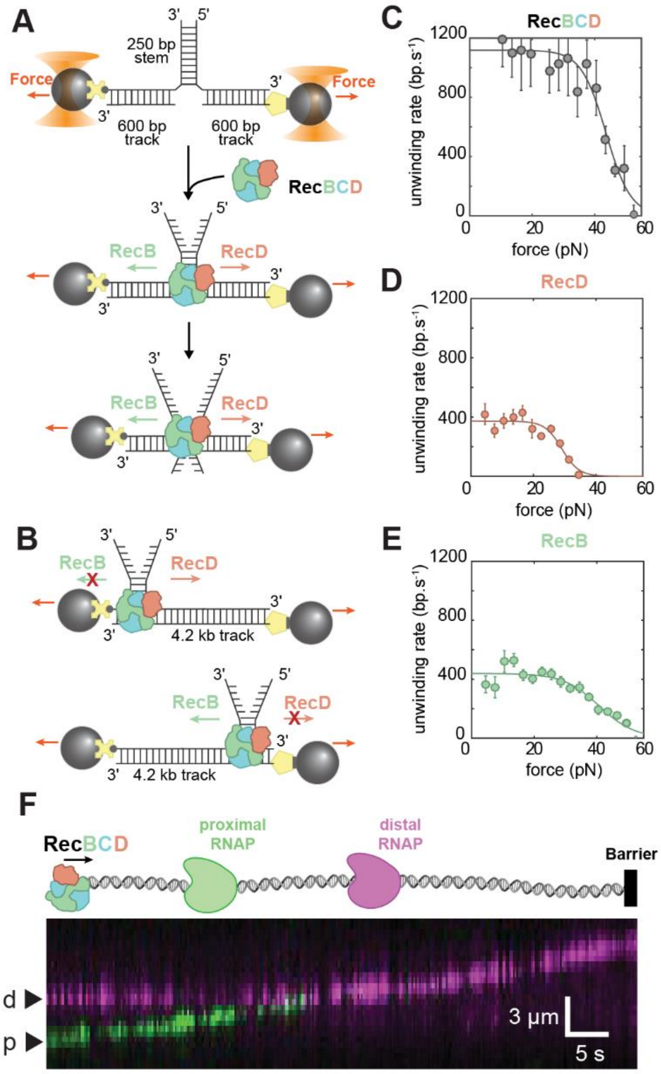


Figure 2.6: RecBCD is a synergistic complex. A) Schematic of dual-optical trap experimental setup designed to monitor RecBCD activity. (top) a DNA stem with dsDNA tracks on either side is tethered to two beads (gray spheres) held in traps (orange cones) at a constant tension by biotin-streptavidin (dark grey sphere – yellow cross) and digoxigenin-anti-digoxigenin (yellow pentagon – dark gray rectangle) linkages. (middle) addition of RecBCD (green, blue, and peach, respectively) initiates unwinding of the DNA stem region. (bottom) As RecBCD approaches the fork, the RecB and RecD subunits translocate in opposite directions. **B)** Schematics of asymmetric dual-optical trap constructs to monitor the activity of individual subunits of RecBCD. RecD construct (top) and RecB construct (bottom). **C-E)** Force velocity curves for RecBCD on symmetric DNA construct **(C)**, RecBCD on RecD asymmetric DNA construct **(D),** and RecBCD

on RecB asymmetric DNA construct, indicate maximal force is generated by the activity of the complete complex. Figure adapted from [276] with permission. **(F)** Schematic of DNA curtain assay monitoring the displacement of a proximal RNAP (green) and distal RNAP (pink) by RecBCD from long DNA substrates (top). Representative kymograph depicting the pushing of the proximal (p) RNAP by RecBCD and subsequent displacement upon encountering the distal (d) RNAP, which is then pushed by the helicase complex (bottom). Figure adapted from [277] with permission.

In the crowded environment of the cell, RecBCD must translocate on DNA coated in a variety of essential nucleoprotein complexes. As such, RecBCD has become a model protein complex for investigating how translocating enzymes deal with inevitable DNA-bound roadblocks [278, 279]. Insight into these mechanisms was first provided by Finkelstein *et. al.* [278] who used a single-tethered 'DNA curtain' assay to visualize the displacement of a variety of Qdot-labeled nucleoprotein complexes by RecBCD as it acts on dsDNA. Displacement of nucleoprotein complexes during RecBCD end resection was visualized via the disappearance of DNA stain and monitoring DNA-bound Qdot-labeled RNAP, inactive restriction enzymes (EcoRI^{E111Q}), Lac repressors, or nucleosomes [278]. Displacement of these nucleoprotein complexes occurred by either immediate ejection upon collision or by pushing the complexes for several kilobases of DNA prior to ejection [278].

DNA curtain assays determined how RecBCD travels along DNA in a more physiologically relevant crowded environment, where tandem arrays of nucleoprotein complexes are found [277]. Qdot-labeled promoter-bound RNAP holoenzymes were shown to immediately dissociate from the DNA when pushed by RecBCD into a tandem array of EcoRI^{E111Q}. Additional experiments of two-color Q-dot labeled RNAP holoenzymes revealed that the proximal RNAP is pushed by RecBCD until it encounters the distal RNAP (Figure 2.6F). Here the proximal RNAP dissociates immediately and the distal RNAP is then pushed by RecBCD [277]. Thus, the molecular crowding within cells influences the mechanism of nucleoprotein complex displacement by RecBCD, where the proximal nucleoprotein complex undergoes rapid dissociation only when it is pushed into a more distal nucleoprotein complex [277]. In situations where an isolated nucleoprotein complex is encountered by RecBCD, the roadblock is pushed along DNA until they are forced to dissociate as a result of stepping between successive nonspecific binding sites.

These stably-bound nucleoprotein complexes have been shown to require up to 30 pN of force to evict [280-282]. How does RecBCD push these nucleoprotein complexes along DNA? Recent optical-tweezer assays revealed that RecBCD complexes generate forces of up to 40 pN [276] (Figure 2.6C). These forces are a specific result of the RecB subunit but not the RecD subunit, further suggesting that the RecD subunit is a weaker helicase and synergy between the subunits is required for optimal complex activity [276]. Together these assays show that successful recombination in the crowded cellular environment relies on a physical mechanism, in that the high forces generated by the translocating RecBCD complex are strong enough to remove any potential protein roadblocks.

RecBCD is absent in *Bacillus* and *Mycobacterium*, where HR initiation is performed by functionally similar, heterodimers, AddAB and AdnAB, respectively [283, 284]. The AdnA and AdnB subunits of the mycobacterial complex each contain UvrD-like SF1 motor domains and RecB-like nuclease modules [284]. Despite being functionally homologous, the RecBCD and AdnAB complexes are not structurally homologous, as demonstrated in recent Cryo-EM structures of *M. smegmatis* AdnAB [285]. A series of structures in the absence of DNA, and in the presence of a forked DNA substrate, both before and after cleavage, revealed a distinct mechanism of strand separation and simultaneous cleavage. Specifically, the AdnB motor domain translocates on the 3' ssDNA strand which drives the displaced 5' ssDNA strand into the AdnA nuclease domain, resulting in cleavage [285]. Although the AdnB nuclease domain was not visible in structures and cleavage of the 3' ssDNA strand was not shown, successful end resection by AdnAB was successfully observed in single-molecule DNA curtain assays [285]. Observations of end resection by AdnAB revealed frequent spontaneous pausing and resumption of end resection by the complex [285]. Further investigations revealed that these pauses occurred independently of the DNA sequence, and triggered a slowing in resection rate following a pause. This observation is in stark contrast to previous work that demonstrated that RecBCD complexes pause at Chi sites, which results in variable rates post-pause [272, 273]. As no specific sequences were detected at pause sites of AdnAB, the authors hypothesized that the complex undergoes stochastic conformational changes that reduce the translocation velocity of the motor domains [285]. The exact mechanism of these pauses by AdnAB requires further investigation, however, these studies demonstrate how structurally distinct prokaryotic complexes converge on similar strategies to carry out successful end resection.

2.5.2. UvrD

In addition to its roles in DNA replication and repair, the displacement of nucleoprotein complexes from DNA by UvrD has also been revealed as an important role in modulating DNA recombination [101, 157, 161, 286]. UvrD and PcrA, but not Rep, are powerful anti-recombinases whereby their activity on ssDNA results in the displacement of RecA recombinase filaments [101, 157, 192, 287, 288]. UvrD and PcrA both show a strong affinity for 5'- ss/dsDNA junctions [192, 287]. The relative affinity of SF1 helicases (UvrD >> PcrA > Rep) for these junctions correlates with their ability to dismantle RecA filaments [192]. Further, this specificity allows the UvrD and PcrA helicases to sit at the junction and reel the ssDNA towards itself resulting in the displacement of RecA monomers, as revealed in smFRET investigations (Figure 2.7A) [192, 287].

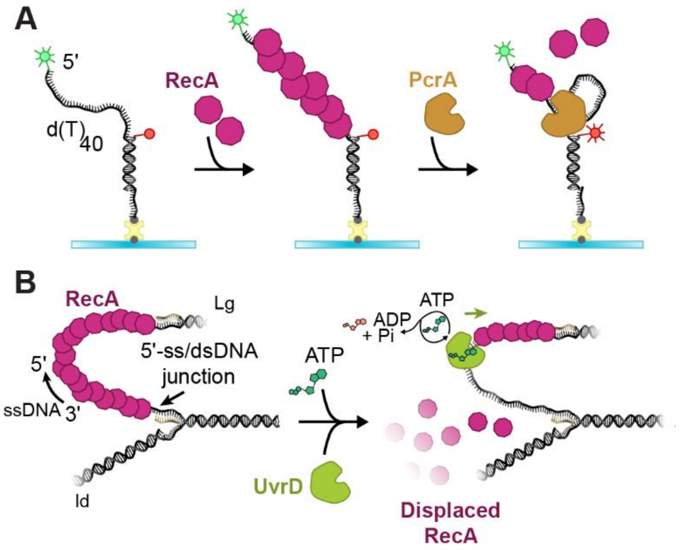


Figure 2.7: PcrA and UvrD displace RecA filaments. A) *In vitro* single-molecule FRET schematic of PcrA simultaneously reeling in the 5'-ssDNA while sitting at the 5'-ss/dsDNA junction and displacing RecA filaments. dsDNA with a 5'$(dT)_{40}$ ssDNA tail is immobilized on a PEG surface via a biotin-streptavidin (dark grey sphere – yellow cross) linkage to the duplex end. Figure adapted from [287] with permission. **B)** Proposed biological role of UvrD specificity for the 5'-ss/dsDNA junctions present on the lagging strand at abandoned replication forks. Figure adapted from [192] with permission.

Additionally, the RecA ATP hydrolysis cycle aids in the displacement of RecA monomers by PcrA [288]. As RecA hydrolyses ATP, it transitions from a high- to low-affinity ssDNA binding state, forming a filament that is readily dismantled [289]. This passive mechanism of dismantling recombinase filaments has also been reported for the eukaryotic UvrD/PcrA homolog Srs2, where physical interactions with Rad51 filaments cause ATP hydrolysis and subsequent dissociation [290]. Interestingly, UvrD-mediated disruption of RecA filaments does not depend on ATPase hydrolysis by the RecA filament. Instead, ATP hydrolysis by UvrD is required to actively remove stable RecA filaments (Figure 2.7B) [157]. Together, the relative affinity for 5'- ss/dsDNA junctions and the ability to exploit ATP hydrolysis cycles results in a mechanism by which UvrD-like helicases can strip off recombinase filaments (Figure 2.7B).

Real-time displacement of recombinase filaments by the eukaryotic Srs2 helicase was recently observed in ‘DNA curtain’ assays [291, 292]. De Tullio and colleagues [291] visualized the simultaneous translocation of mCherry-labeled Srs2 and displacement of GFP-labeled single-stranded binding protein replication protein A (RPA), or recombination protein factors Rad51 and Rad52, from ssDNA. Directional stretching of the ssDNA substrates from a tether point in a microfluidic flow cell, combined with observation using smTIRF microscopy, revealed simultaneous translocation of mCherry-Srs2 and displacement of GFP-RPA at a rate of 170 nt/s. Further observations showed the mCherry-Srs2 could also push and eventually displace GFP-labeled Rad52, at no detriment to its translocation rate [291]. Additional investigations using the ssDNA curtain assay observed GFP-labeled Srs2 displacing Rad51 filaments at a rate of 50 monomers/s in an ATP-dependent manner [292]. The displacement of these nucleoprotein filaments occurred immediately following recruitment to RPA clusters embedded between Rad51 filaments [292]. Interestingly, the SF1A human F-box-containing DNA helicase (FBH1) was shown to bind and disrupt Rad51 nucleoprotein filaments in an ATP-dependent manner in smTIRF and smFRET investigations [152, 293]. The similarities between Srs2, FBH1 and UvrD in stripping recombination filaments suggest that this is a conserved activity for UvrD-like helicases. Based on its participation in several DNA metabolism pathways, UvrD exhibits a general nucleoprotein complex clearing role across replication, repair, and recombination.

2.6. Conclusions and future perspectives

Single-molecule techniques have proven essential to understanding the fine mechanistic details of DNA helicases in replication, repair, and recombination pathways. *In vitro* single-molecule studies using purified proteins to reconstitute various genome maintenance pathway reactions, have revealed the behaviors of helicases, otherwise hidden using ensemble averaging techniques. Such behaviors include helicase activation by conformational change, force generation, and changes in velocity or processivity. The relative simplicity of reconstituting prokaryotic reactions has also provided insight into how more complex eukaryotic systems utilize and regulate powerful helicases to ensure the integrity of the genome is upheld. Additionally, bacterial systems have provided a simple model organism for investigating how helicases behave in the crowded environment of the cell.

Further single-molecule live-cell imaging of these diverse motor proteins will extend our current understanding of both prokaryotic and eukaryotic helicases as these techniques are adapted for more complex organisms. Together, *in vitro* and *in vivo* single-molecule imaging of helicases in genome maintenance pathways will contribute to a better understanding of how cells maintain and prevent the processes leading to genetic instability and disease.

Helicase activity is fundamental to successful cell proliferation and recovery from environmental DNA damage. Multiple prokaryotic helicases have been associated with the promotion of antibiotic resistance [13, 260, 294]. Dysregulated helicase activity can trigger mutagenesis [295]. The resulting mutations within the chromosome can give rise to antimicrobial resistance in prokaryotes. Further, dysfunction in the human homologs of these helicases has been linked to a variety of severe developmental disorders and diseases that often clinically present as neuromuscular degeneration, premature aging, and cancer predisposition (Table 2.1) [15-18, 296-298]. Ultimately, understanding the fundamental molecular mechanisms will enable the development of therapeutics that can translate into improved health outcomes in some of the greatest challenges to human health.

Table 2.1: The human homologs of prokaryotic DNA repair helicases and translocases cause disorders linked to genomic maintenance. For the prokaryotic helicases, Rep, UvrD and PcrA, Mfd, UvrB, RecD and RecQ, their structural classification, genome maintenance pathway they are active in, eukaryotic homolog, human homolog, and known diseases resulting from inactive or defective human homologs are provided.

Prokaryotic Helicase	Classification	Roles	Yeast homolog	Human homolog	Disease
Rep, UvrD, PcrA	SF1	Nucleoprotein complex removal	Pif1, Srs2, Fbh1	hFbh1 [296]	Cancer predisposition [15]
Mfd	SF2	Transcription-coupled repair	Rad26	CSB [16]	Cockayne syndrome B
UvrB	SF2	Nucleotide excision repair	Rad25	XPD [17]	Xeroderma pigmentosum
RecD	SF1B	DNA recombination	Rrm3	Pif1 [297, 298]	Unknown
RecQ	SF2	DNA recombination	Sgs1 (*S. cerevisiae*)	WRN [18]	Werner Syndrome
				BLM	Bloom syndrome
				RECQ4	Rothmund-Thomson Syndrome, RAPADILINO syndrome, Baller-Gerold syndrome

Chapter 3

Spy-ing on Cas9: Single-molecule tools reveal enzymology of Cas9

Kelsey S. Whinn, Antoine M. van Oijen, Harshad Ghodke

Published in *Current Opinion in Biomedical Engineering,* 10.1016/j.cobme.2019.08.013

CRISPR/Cas (clustered regularly interspaced short palindromic repeats/CRISPR-associated protein) systems are an adaptive immune response mechanism in prokaryotes that can target and cleave invading DNA or RNA. The rapid understanding of the type II CRISPR/Cas9 system through biochemical, genetic and structural investigations have contributed to the development of Cas9 for various DNA- and RNA-targeting applications. Recent single-molecule investigations of CRISPR/Cas9 mechanisms have further extended our understanding of target search, binding and cleavage. These investigations are fundamental to the further development of CRISPR/Cas9 tools. This review discusses how single-molecule techniques have illuminated the enzymology of Cas9 endonucleases.

Contributions

This literature review addresses our current understanding of the enzymology of the Cas9 protein, as highlighted by the significant contributions of single-molecule studies. This review discusses the structure of the Cas9 protein, the mechanisms of guide RNA binding, and the association and dissociation of the complex to target dsDNA regions.

I made the majority contribution to this review by exploring the literature on single-molecule studies of Cas9 proteins drafting the manuscript, and designing and producing figures.

3.1. Introduction

In recent years, the ability to target specific DNA sequences has been revolutionized by the discovery of the prokaryotic CRISPR (clustered regularly interspaced short palindromic repeats) systems [299, 300]. CRISPR systems possess a rich and complex biology. Identified as an adaptive immune response mechanism in bacteria and archaea, the CRISPR system allows the host to site-specifically target and destroy invading DNA or RNA [300-303]. Single-molecule techniques have been successfully used to identify the elemental processes that underlie biological complexity [90-92, 304, 305]. These techniques directly measure the diverse and dynamic behaviors of single molecules without the need for synchronization, thus removing the ensemble averaging of biological states. Further, by permitting observation of transient and rare intermediate states along the reaction pathway, these techniques overcome the detection limits of ensemble-based methods. More recently, an expanding toolkit of single molecule techniques has enabled *in situ* observation of reactions occurring in living organisms, demystifying how individual protein actors work inside the cellular milieu [76, 93-99, 306]. In this report, we discuss the biology of the type II CRISPR/Cas system and the power of single-molecule methods in revealing the fine mechanistic details that are central to it.

3.2. CRISPR Systems

Currently, two main classes of CRISPR systems have been described [307]: Class 1 and Class 2, with the key difference being the number of protein effectors (CRISPR-associated proteins or ‘Cas’) required to perform the reaction. Briefly, class 1 systems, containing types I, III and IV, require multi-subunit complexes for RNA-guided surveillance in bacteria and archaea [302, 303, 307, 308]. In contrast, class 2 systems, comprising types II, V and VI, use a single RNA-guided nuclease to target and cleave specific invader DNA or RNA sequences [307, 309-311].

3.3. *Spy*Cas9

The best-characterized type II CRISPR system of *Streptococcus pyogenes* is comprised of the Cas9 (*Spy*Cas9) endonuclease (Figure 3.1) [300, 301, 309, 312]. In this case, CRISPR-mediated adaptive immunity arises from a genetic locus composed of an operon encoding Cas9 and a repeat-spacer array [301]. This array of identical repeat sequences contains unique DNA sequences derived from an infecting virus or plasmid DNA. The acquisition of invader sequences forms the first step of the CRISPR-mediated immunity mechanism (Figure 3.1) [301]. The repeat-spacer array is then transcribed, producing precursor CRISPR-RNA (pre-crRNA) that is subsequently base paired with a non-coding RNA molecule, *trans*-activating RNA (tracrRNA), and processed to form a mature duplexed crRNA:trcrRNA (gRNA) [313]. Construction of a synthetic single-gRNA (sgRNA) chimera has also been used to direct *Spy*Cas9 to specific sequences [300]. The mature gRNA is then recruited by *Spy*Cas9 proteins to form mature RNA-guided ribonucleoprotein effectors. These programmed *Spy*Cas9 proteins then survey the cellular environment for the presence of a short NGG motif, termed the protospacer adjacent motif (PAM), and invader sequences that are complementary to the guide RNA [300, 302]. Successful recognition of the complementary sequence is then followed by cleavage of the double-stranded DNA (dsDNA).

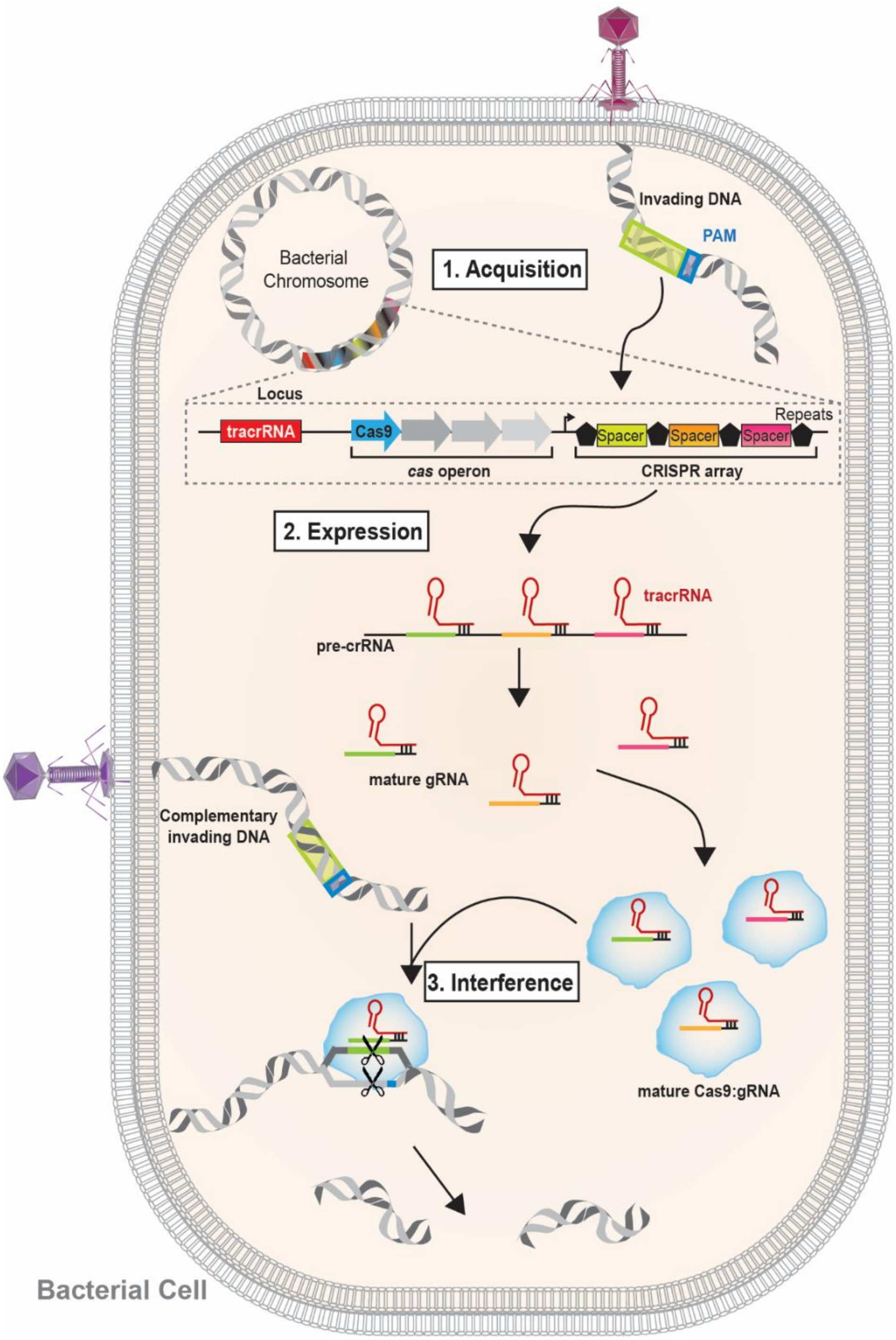


Figure 3.1: CRISPR adaptive immunity mechanism. (1) Acquisition; foreign DNA sequence (green) is incorporated as a new spacer within the CRISPR repeat-spacer array. (2) Expression; the repeat-spacer array is transcribed, expressing pre-crRNA, which anneals to tracrRNA (red). Further processing yields mature gRNA duplexes that are recruited by effector proteins (blue). (3) Interference; mature gRNA guides the effector proteins to cleave the complementary sequences of subsequent invading DNA that contain PAM sequences.

3.3.1. Structural biology of SpyCas9

*Spy*Cas9 is a 1368 amino acid, globular protein whose domain architecture is composed of two major lobes; a nuclease (NUC) lobe and a recognition (REC) lobe (Figure 3.2) [314-317]. The REC lobe contains the REC1 and REC2 domains and a bridge helix. Of these, the REC1 lobe is critical for interactions with the gRNA, whereas the bridge helix is important for interactions with both the target DNA, as well as the gRNA [317]. The NUC lobe consists of the PAM-interacting (PI) domain and two nuclease domains, HNH and RuvC, that cleave DNA targets using one metal and two metal catalytic mechanisms, respectively [317]. Within the NUC lobe, the RuvC domain interfaces with the PI domain, forming a positively-charged surface that interacts with the 3′ tail of the gRNA [317]. We direct readers to reference [318] for a detailed review on the structural biology of Cas9.

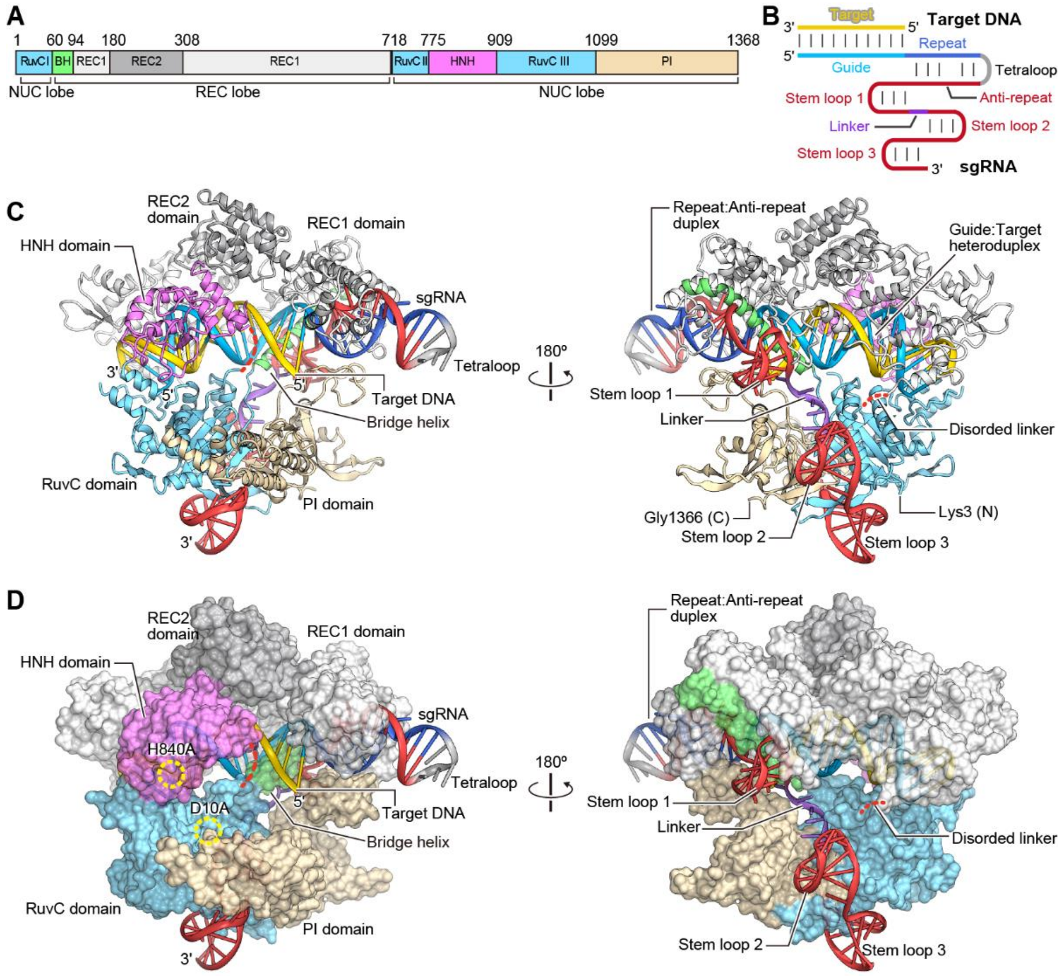


Figure 3.2: Structural biology of *Spy*Cas9 in complex with sgRNA and target DNA. A) Domain organization of *Spy*Cas9. BH, bridge helix; REC, recognition; NUC, nuclease; PI, PAM-interacting. **B)** Schematic representation of the sgRNA hybridized to the complementary target DNA sequence. **C)** Ribbon representation of the *Spy*Cas9:sgRNA:DNA complex. Red dotted line represents a disordered linker. **D)** Surface representation of *Spy*Cas9:sgRNA:DNA complex. Dashed yellow circles represent HNH and RuvC active sites, H840A and D10A, respectively. Figure adapted with permission from [317].

3.3.2. Enzymology of SpyCas9

The first step in the formation of the mature effector requires the loading of the crRNA:tracrRNA into the *Spy*Cas9 protein. Cryo-EM studies of apo-*Spy*Cas9 reveal that in the absence of RNA, *Spy*Cas9 is a highly flexible molecule, and large conformational rearrangements are triggered by the binding of RNA (Figure 3.3A, B) [316]. High-speed fluid atomic force microscopy studies further reveal that these structural transitions occur on the time scale of seconds on *Spy*Cas9 molecules immobilized on mica [319]. In agreement, enhanced MD simulations of apo-*Spy*Cas9 found that conformational rearrangement upon gRNA binding mainly involved the opening of the REC lobe with respect to the NUC lobe [320]. Opening of the bi-lobed structure results in the formation of a positively charged cavity capable of accommodating RNA. Single-molecule FRET (smFRET) studies further demonstrate that apo-*Spy*Cas9 undergoes conformational rearrangement upon gRNA binding [321-323]. Binding of sgRNA to apo-*Spy*Cas9 containing FRET-pair labeled REC1 and RuvC domains (approximately 21 Å apart), caused a 56 Å increase in distance between the REC1 and RuvC domains resulting in a decrease in FRET efficiency [323]. This movement suggested that sgRNA binding triggers a drastic opening of the REC lobe relative to the NUC lobe [323]. Bulk measurements of the kinetics of the conformational change associated with sgRNA binding were described by a single time constant, independent of the sgRNA concentration used [321]. This led the authors to suggest that an initial collision complex (apo-*Spy*Cas9:sgRNA) is formed rapidly, followed by a slower, induced conformational change leading to productive association (*Spy*Cas9:sgRNA). Further, using truncated sgRNA variants, it was found that specific motifs at both ends of the 20-nucleotide sgRNA sequence are required to stabilize the mature *Spy*Cas9:gRNA complex [322]. Collectively, these studies reveal that a complete crRNA:trcrRNA complex (or sgRNA) is required to trigger the flexible rearrangement of the REC lobe, to form a stable and mature *Spy*Cas9:gRNA complex capable of executing target search.

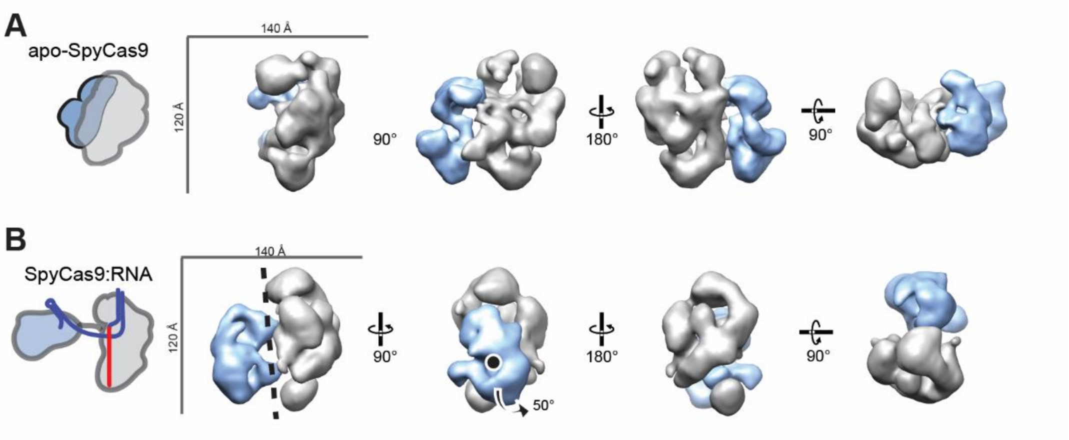


Figure 3.3: Domain flexibility of *Spy*Cas9. Cryo-EM structures of **A)** apo-*Spy*Cas9 (19-Å resolution) and **B)** *Spy*Cas9:gRNA (21-Å resolution) reveal flexibility of the bilobed structure. Cartoon representations (left) are shown alongside single-particle EM reconstructions of structures. The smaller NUC lobe (blue) rotates away from the larger REC lobe (gray) upon binding of gRNA. Figure adapted with permission from [316].

Undoubtedly, the CRISPR/Cas revolution is attributable to the remarkable ability to ‘program’ and ‘target’ *Spy*Cas9 to desired genomic loci. How does *Spy*Cas9 discriminate target sites from non-target sites? Single-molecule techniques have been extensively used to describe mechanisms involved in target search [93, 324-330]. Insight into this issue was first provided by Sternberg and co-workers who employed a double-tethered ‘DNA curtain’ assay to visualize quantum dot labeled nuclease dead *Spy*Cas9 (*Spy*dCas9):gRNA searching and binding to target sequences in λ DNA molecules, that are directionally stretched and tethered in a microfluidic flow cell using TIRF microscopy (Figure 3.4A-E) [331]. Both wild-type as well as *Spy*dCas9:gRNA constructs employed 3-dimensional diffusion to locate and specifically bind to the target site (Figure 3.4C). In addition to long-lived binding observed at target sites, *Spy*Cas9:gRNA was also found to bind transiently along the length of the DNA (Figure 3.4D). Strikingly, the probability of transient binding correlated strongly with the density of PAM sequences along the DNA substrate leading the authors to suggest that specific recognition of PAMs may be a crucial feature of target search by *Spy*Cas9:gRNA (Figure 3.4E). smFRET studies of dwell times of *Spy*Cas9:gRNA revealed that *Spy*Cas9:gRNA is longer lived on DNA substrates containing multiple PAM sequences [332]. By labeling the tracrRNA and DNA substrate with a FRET pair, this study determined that in conjunction with 3D diffusion, *Spy*Cas9:gRNA employs 1D diffusion to rapidly sample neighboring PAM sites. Indeed, measurements of association rates of *Spy*dCas9-YPet:sgRNA in live-cells containing a contiguous array of target sites revealed that it takes *Spy*dCas9:sgRNA approximately 6 hours to locate the target site in *E coli* (Figure 3.4F, G) [306, 333]. Binding of *Spy*dCas9-Ypet:sgRNA programmed to the *lacO1* operator sequence was triggered by the addition of isopropyl-β-D-thiogalactopyranoside (IPTG), which induced the dissociation of LacI from *lacO1* and allowed *Spy*dCas9-YPet:sgRNA to bind [306] (Figure 3.4G). Measurements of DNA-bound residence times of *Spy*dCas9-YPet:sgRNA revealed an average lifetime of 30 ms, that is sufficiently fast so that *Spy*dCas9:sgRNA can sample the ~10^6 PAMs present in the *E coli* chromosome [306]. Together, these studies unveil a search mechanism that is exquisitely predicated on the stability of *Spy*Cas9:PAM interactions. The intrinsically weak nature of these interactions is critical for efficient sampling of the genome.

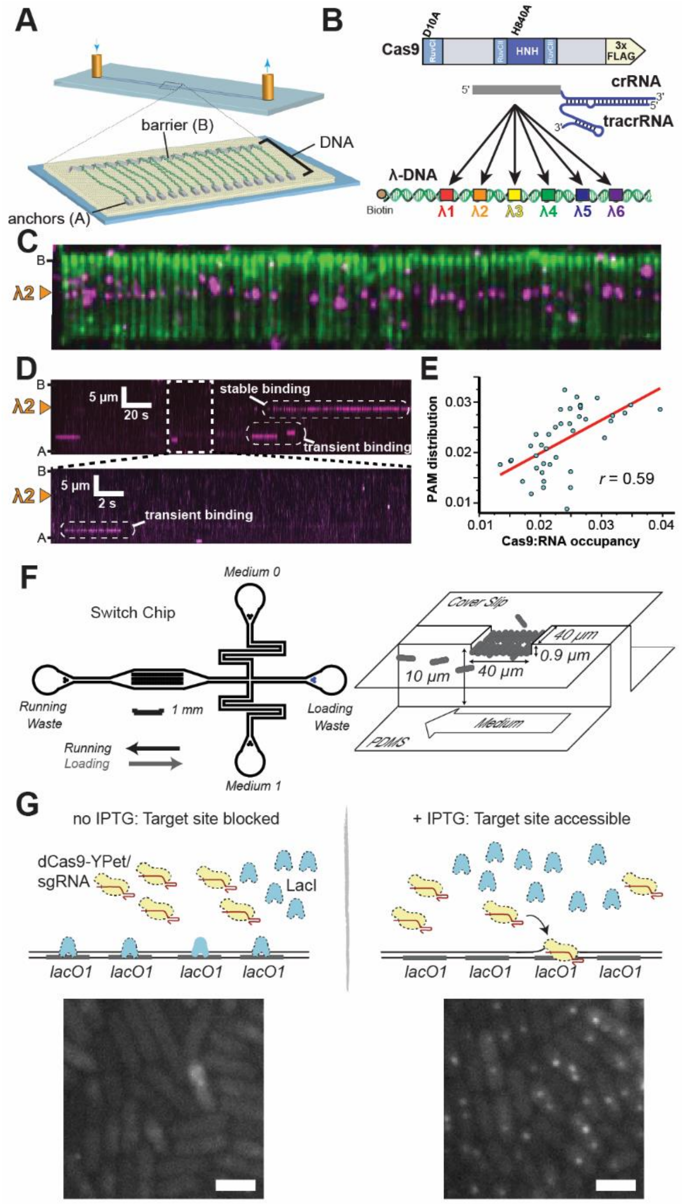


Figure 3.4: Target search and binding of *Spy*Cas9. A) Schematic of double-tethered DNA curtain assay. λ-DNA substrates (48, 502 bp) are anchored by one end to a lipid bilayer through a biotin-streptavidin linkage. Hydrodynamic force is then used to stretch the DNA molecules until they encounter antibody-coated pentagons, with which they bind to through DIG-labelled ends. **B)** WT *Spy*Cas9 or *Spy*dCas9 (top) was programmed with crRNA:tracrRNA targeting one of six specified sites (λ1-6). **C)** YOYO1-stained DNA (green) bound by QD-tagged *Spy*Cas9 (magenta) programmed with λ2 gRNA. Binding site of programmed *Spy*Cas9 corresponds to expected target site of λ2 gRNA. *Spy*Cas9 remains bound after cleavage of dsDNA. **D)** Kymographs of a single DNA molecule illustrating distinct stable and transient binding events (top). Zoomed in image of transient binding events (bottom). **E)** Pearson correlation analysis of PAM distribution and non-target *Spy*Cas9:gRNA binding distribution for λ2 gRNA (blue) and a non-complementary gRNA (green) ($r = 0.59$, $P < 0.05$). Figures adapted with permission from [331]. **F)** Schematic of live-cell imaging

experimental setup used for kinetic analysis of *Spy*dCas9-Ypet:sgRNA. The microfluidic device (left) contains three ports assigned for medium, running waste and loading waste. The chamber contains three rows, each containing 17 traps. Cells are introduced from the running waste and are caught in the traps. Each trap is a 40 X 40 X 0.9 µm compartment, confined by two rigid walls and two openings. Cells that reach the openings are freed from the traps into the 10 µm deep surrounding (right). Figure adapted with permission from [333] **G)** Schematic of single-molecule assay (top panel) where in the absence of IPTG, *lacO1* sites are occupied by LacI, preventing binding of *Spy*dCas9-Ypet:sgRNA (bottom panel). Addition of IPTG dissociates LacI, allowing *Spy*dCas9-Ypet:sgRNA to bind to the *lacO1* target, enabling detection of specific fluorescent spots using exposure times of 5 s. Bottom panel shows fluorescence images obtained before (left) and 10 min after (right) IPTG addition. Scale bar, 2 µm. Figure adapted with permission from [306].

Stable binding is only achieved on genomic loci containing PAM sequences at the correct target site. Following initial recognition, the gRNA invades the dsDNA to form an R-loop [300]. In this state, *Spy*Cas9 wraps around the R-loop while excluding the non-target strand [314-317, 334] (Figure 3.2C, d). Numerous bulk studies have revealed the requirement of essential sequence complementarity between target DNA and gRNA, termed the seed region, for stable R-loop formation [300, 335, 336]. *In vitro* plasmid cleavage assays revealed that a seed region of at least 13 base pairs between crRNA:DNA proximal to the PAM is required for efficient cleavage, whereas up to six adjacent mismatches at the PAM-distal end are tolerated [300]. In agreement, deep sequencing of *Spy*Cas9:sgRNA targeted loci in human cells showed that the specificity within the seed region ranges from 8-14 bp immediately upstream from the PAM [335]. Stopped-flow measurements of 2-aminopurine fluorescence in the seed region suggest that R-loop formation is the rate-limiting step of *Spy*Cas9 target cleavage [337]. The stepwise formation of *Spy*Cas9:gRNA R-loops has also been observed at the single-molecule level. The use of magnetic tweezers to investigate *Spy*Cas9:gRNA binding revealed that R-loop formation occurs by a directional hybridization between the crRNA and target DNA, beginning at the PAM and extending toward the distal protospacer end [338]. Further, it was found that the PAM sequence primarily affects the association rate of the R-loop, while the base pairs distal to the PAM mainly affect R-loop stability [338]. Magnetic tweezer experiments interrogating the stability of the R-loop showed that PAM distal mismatches cause R-loop propagation to stall and any further propagation competes with the collapse of the hybridized structure [339]. smFRET investigations of the *Spy*Cas9:gRNA complex on substrates containing varying numbers of mismatches, showed that the stability of the R-loop structure decreases linearly with the number of mismatches proximal to the PAM [340]. Notably, PAM-distal mismatches up to 11 bp in length still allowed the formation of a stable, but inactive complex [340].

Further, smFRET studies have identified two distinct *Spy*Cas9 domain conformations during R-loop propagation [341, 342]. The first of the two states termed the ‘open’ conformation, represents a PAM-proximal bound complex that is capable of initiating R-loop formation. The second state, termed ‘zipped’ conformation, occurs after complete R-loop formation and represents a cleavage-competent state [341]. Recently, analysis of the dwell time of the zipped conformation revealed that the lifetime increases when the base-pairing length between the crRNA:DNA reaches a maximum of 18 bp [342]. The recent combination of optical tweezers with confocal fluorescence microscopy revealed that DNA stretching induced off-target *Spy*Cas9:gRNA binding [343]. smFRET was then employed to interrogate the R-loop formation at these sites using DNA-containing bubbles, showing that off-target binding and cleavage occurred in the presence of 10 bp mismatches within the R-loop [343]. Together, these studies indicate that crRNA:DNA complementarity is essential to *Spy*Cas9 stability and cleavage competency. The directional hybridization of the R-loop ensures that when enough mismatches are formed, further propagation is stalled allowing *Spy*Cas9:crRNA to reject the site. Crucially, the cleavage-competent is only accessible when a stable R-loop with the maximum number of base pairs is formed.

*Spy*Cas9 and orthologs induce a blunt double-strand break (DSB), upstream and proximal to the guanine-rich PAM. Cleavage is catalyzed by the two nuclease domains within the NUC lobe, HNH and RuvC, that cleave the target and non-target strands, respectively [300, 315, 317] (Figure 3.2 A, C, D). Numerous *Spy*Cas9 structures have provided insight into the flexibility and structural rearrangement of the HNH catalytic domain into its active state [314, 316, 317]. smFRET studies showed that the conformational flexibility of the HNH domain directly controls cleavage [322]. Further smFRET investigations of labeled *Spy*Cas9:sgRNA complexes revealed that in the absence of DNA, *Spy*Cas9:sgRNA primarily resides (approximately 50%) in a stable state termed the R state (referring to the RNA bound *Spy*Cas9 complex) [344]. Introduction of target-containing DNA led to the formation of a new state termed the D state (referring to the correctly base-paired, target DNA-bound complex). The *Spy*Cas9:sgRNA complex was found to transiently access a third intermediate state (I state). Strikingly, transitions from the R state to the D state necessarily required visits to the I state. Considering that (1) rate of transitions to the D state are ten-fold slower, and (2) the observation that the stable residence in the D state is infrequently encountered on target DNA containing 1-3 bp mismatches, these investigators suggested that the intermediate state serves as a conformational checkpoint gating initial binding and DNA cleavage [344]. Consistent with this hypothesis, investigations of domain flexibility in the presence of mismatched target DNA substrates showed that the HNH domain has a decreased propensity for its cleavage-competent state in the presence of mismatched R-loops [345]. Indeed, greater than 4-bp mismatches in the R-loop at the PAM-distal end prevent this domain rearrangement [345]. Thus, the conformational rearrangement of the HNH domain into its cleavage-competent state is only triggered upon confirmation that the correct target site has been bound by the *Spy*Cas9 nuclease. This long-range allosteric communication between the flexible HNH domain and the PAM-distal end of the R-loop act as a final cleavage checkpoint to prevent the degradation of bound off-target sites.

*Spy*Cas9:gRNA complexes remain tightly bound to cleaved DNA products. Observations of quantum-dot labeled *Spy*Cas9 proteins in DNA curtain assays showed that *Spy*Cas9:gRNA remained bound to DNA after cleavage [331] (Figure 3.4C). Further, plasmid DNA cleavage assays at varying molar ratios of *Spy*Cas9:gRNA and target DNA revealed that *Spy*Cas9:gRNA complexes do not follow Michaelis-Menten kinetics [331]. In agreement, measurements of binding kinetics of *Spy*Cas9:gRNA interactions with substrate DNA performed using bio-layer interferometry revealed that the complex remains bound to cleaved products and displays an identical lifetime to *Spy*dCas9:sgRNA complexes [346]. Additionally, steady-state kinetic analysis of the dissociation of radiolabelled cleavage products exhibited slow multiple-turnover rates for both the HNH (k_{cat} 4.45 $\times 10^{-6}$ s^{-1}) and RuvC domains (k_{cat} 2.1 $\times 10^{-6}$ s^{-1}) [321]. Thus, *Spy*Cas9:gRNA acts as a single-turnover enzyme. Recently, bulk biochemical experiments have shown the ability of RNA polymerase to dislodge *Spy*Cas9 bound to cleavage products [347]. Displacement of the *Spy*Cas9:sgRNA complex resulted in a multiple-turnover enzyme [347]. Interestingly, displacement of the *Spy*Cas9:sgRNA complex only occurred if the sgRNA had hybridized with the RNAP template strand [347]. The *Spy*Cas9 homolog, *Staphylococcus aureus* (*Sau*Cas9) was recently identified as a multiple-turnover Cas9 enzyme [348]. *In vitro* cleavage reactions analyzed by capillary electrophoresis showed that *Sau*Cas9 cleavage resulted in 5-fold more cleavage product formation over 24 hours in comparison to *Spy*Cas9, suggesting that one *Sau*Cas9 complex can cleave multiple DNA substrates [348]. This enhanced rate of turnover is potentially attributable to an enhanced rate of product release [348]. Collectively, these studies indicate that *Spy*Cas9:gRNA complexes remain stably bound to cleavage products and are essentially single-turnover enzymes, unless acted upon by cellular machinery undertaking DNA transcription.

3.4. Conclusions and future perspectives

Single-molecule techniques have been crucial for assembling detailed models for the various kinetic mechanisms involved in effector maturation, target search and discrimination, conformational changes involved in catalysis and product release. Extensive characterization of these mechanisms has contributed to the continuous adaptation of *Spy*Cas9 for a range of molecular biology tools. Whereas *Spy*Cas9 has been extensively studied and remains at the forefront of the revolution in genetic engineering, other class I and 2 CRISPR/Cas systems remain relatively under-characterized. The single-molecule approaches used to characterize the enzymology of this exciting new family of proteins represent a broad and powerful toolkit for future investigation of other CRISPR/Cas systems. Ultimately, these approaches will enable the engineering of refined CRISPR/Cas systems for applications in genetic engineering, biotechnology, diagnostics and even fundamental discovery.

3.5. Recent investigations of the enzymology of *Spy*Cas9

Since the publication of the above review [349], significant progress has been made in unraveling the enzymology of *Spy*Cas9. Recent studies utilizing single-molecule FRET and Cryo-EM techniques have shed new light on the dynamics of sgRNA, target search, recognition and interrogation, and the conformational changes that occur throughout the catalytic cycle.

The above review highlights the extensive investigations of how *Spy*Cas9 interacts with DNA, unwinds it, and ultimately cleaves the target sites. However, these steps first require sgRNA folding and *Spy*Cas9 binding to the sgRNA. A recent study using smFRET showed that an acceptor- and donor-labeled sgRNA exhibited two main FRET states in the absence of *Spy*Cas9, with the higher state representing a more compact conformation [350]. Interestingly, the addition of *Spy*Cas9 shifted the distribution towards the higher FRET state, indicating that *Spy*Cas9 assists in sgRNA folding by promoting a compact, favorable active state while suppressing dynamic sgRNAs [350].

Single-molecule fluorescence studies have demonstrated that *Spy*Cas9 uses both 3D and 1D diffusion mechanisms to search for target sites [331, 332]. A recent combination of single-molecule fluorescence assays and Cryo-EM revealed that *Spy*Cas9 exhibits a noticeable asymmetrical search pattern that extends from approximately 10 bp upstream of the PAM to approximately 30 bp downstream of the PAM [351]. Further, Cryo-EM structures of the *Spy*Cas9:sgRNA:dsDNA complex revealed non-specific interactions between lysine residues of the PI domain and DNA 8 bp downstream of the PAM, supporting the observations of the asymmetric search pattern. According to the proposed model, *Spy*Cas9:sgRNA complexes initially interact with dsDNA through random 3D collision, and following non-specific interactions between lysines of the PI domain and dsDNA, *Spy*Cas9 then transiently binds and laterally diffuses along dsDNA to search for PAM sites [351].

Target sequence interrogation and R-loop formation have also been extensively characterized using single-molecule techniques. Cryo-EM structures of *Spy*Cas9:sgRNA:DNA complexes at different states during the interrogation pathway have revealed how *Spy*Cas9 unwinds dsDNA and samples for target complementarity [352]. Specifically, *Spy*Cas9:sgRNA binds with the PAM of a dsDNA target in an open conformation, where the REC lobe is rotated away from the NUC lobe, like that of the *Spy*Cas9:sgRNA structure (Figure 3.3B) [316, 352]. Transition to the closed state causes the REC and NUC lobes to pinch the DNA into a bent, underwound hinge where the target strand nucleotides are flipped toward the sgRNA [352]. This bend and twist mechanism allows *Spy*Cas9 to sample nucleotides for base-pair complementarity, before initiating complete strand invasion and R-loop formation [338, 339, 352].

Further, a recent study used single-molecule rotor bead tracking to simultaneously monitor DNA unwinding during *Spy*dCas9:sgRNA:DNA R-loop formation and dissociation dynamics. Measurements of the changes in torque during dsDNA unwinding, revealed that R-loop formation and collapse occur through a transient intermediate, consistent with DNA:RNA hybridization through the seed region [300, 335, 336, 353]. Additionally, PAM-proximal mismatches can destabilize both the intermediate state and the cleavage-competent open state, whereas PAM-distal mismatches destabilize the open state. However, negative torque can influence the energy landscape, allowing *Spy*Cas9 to achieve a fully unwound state and target cleavage even in the presence of mismatches [353].

Cryo-EM studies have provided additional structural insight into the *Spy*Cas9:sgRNA:DNA complex throughout its catalytic cycle [344, 354]. The structures revealed three distinct conformational states: pre-catalytic, post-catalytic and product states. In the pre-catalytic state, *Spy*Cas9 adopts a ‘checkpoint’ conformation, where the HNH nuclease domain is oriented away from the DNA [354]. Transition to the post-catalytic state shows a significant swing (~34 Å) of the HNH domain and disorder of the REC2 recognition domain. However, the product state revealed a disordered HNH domain while the REC2 domain returned to the precatalytic conformation [354]. Interestingly, recent smFRET studies revealed high flexibility of the HNH domain following target-strand cleavage, explaining why it was not visible in the Cryo-EM structures of the product state [354, 355]. Further, the high flexibility of the HNH domain suggests that the post-catalytic conformations are coupled with target and non-target strand cleavage [355].

Together, these studies highlight the fast-paced field of CRISPR/Cas9 technology. Over the past four years, these investigations have provided valuable insight into multiple mechanisms of *Spy*Cas9 enzymology. These findings not only enhance our understanding of *Spy*Cas9 mechanisms but also provide important information that can guide the optimization and application of future CRISPR/Cas9 technologies.

Chapter 4

Automated detection of fluorescent proteins in imaging of rolling-circle DNA replication

Chapter Summary

Analysis of single-molecule fluorescence imaging experiments relies on accurate tracking of fluorescently-labeled proteins through space over time. Combining rolling-circle DNA replication with fluorescently-labeled proteins that interact with the DNA is an established approach to studying DNA replication at the single-molecule level, but the analysis of these images involves manual detection methods. While useful, these analysis methods can introduce bias and variation between users. Here, we develop and optimize an automated method of tracking rolling-circle DNA replication over time. Using ImageJ/FIJI and MATLAB, this method is further developed to track dynamic interactions of fluorescently-labeled Rep proteins with the replication fork. To validate this automated method, we compare replication rates and stoichiometries of labeled proteins present at the fork detected by both the optimized automated method and the previously used manual method. We show that the automated method is capable of detecting replication rates consistent with previous characterizations and can capture the association and dissociation events of labeled proteins. The automated analysis method described in this chapter provides an approach that can be used to further characterize the proteins involved in DNA replication and stalled replication rescue.

Contributions

The manual analysis method and original automated analysis macros, described in sections 4.2.3 and 4.2.4.1, were designed and written by Dr Lisanne Spenkelink (University of Wollongong, Australia). I further adapted the automated analysis method described in sections 4.2.4.1 and 4.2.4.2 for sequential imaging of two-channel fluorescence acquisitions. I carried out all experiments described, data analysis and quantification, writing and figure design presented in this chapter.

4.1 Introduction and rationale

Rolling-circle DNA amplification enables the observation of highly processive DNA synthesis. This method involves using a nicked, covalently-closed circular double-stranded DNA (dsDNA) template containing a 5' single-stranded DNA (ssDNA) overhang that forms a forked structure able to load a fully active replisome (Figure 4.1A) [356]. The internal strand of the circle serves as the template for the leading strand, with the leading-strand product subsequently displaced by unwinding activity and serving as a template for lagging-strand synthesis. These DNA templates can support both leading-strand and simultaneous leading- and lagging-strand DNA replication, and can be made visible by both ensemble and single-molecule fluorescence assays [22, 34, 48, 74, 75, 77, 78, 356-358].

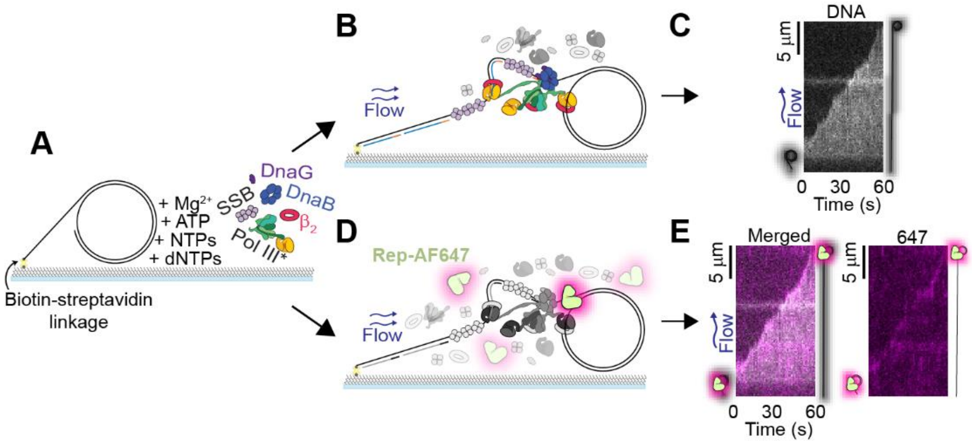


Figure 4.1: Schematics and example kymographs of rolling-circle DNA replication. A) Rolling-circle DNA templates are immobilized to the coverslip surface through a biotin-streptavidin linkage. **B)** Addition of the *E. coli* replisome components and nucleotides initiates DNA replication where the growing DNA product is stretched out in the direction of buffer flow. **C)** Replicating DNA molecules are visualized by staining with a fluorescent intercalating stain (Sytox Orange) and represented as kymographs. **D)** Fluorescently labeled proteins (here, *E. coli* Rep-AF647), can be imaged throughout the replication reaction, resulting in **E)** two-color kymographs (left) where both the Sytox orange and fluorescence of labeled proteins (alone – right) are captured.

Rolling-circle replication has been widely used to monitor the real-time synthesis of DNA templates by viral and bacterial replisomes [22, 34, 48, 74, 75, 77, 78, 356-358]. Real-time single-molecule fluorescence imaging of rolling-circle DNA replication is initiated by applying a laminar flow of buffer, containing the components required to support the DNA replication reaction and fluorescent DNA stains (Figure 4.1). Continuous buffer flow stretches the growing dsDNA product at a rate that is determined by the replication rate (Figure 4.1B). The addition of fluorescently-labeled replisome proteins provides extra information, where the individual proteins can be imaged simultaneously with DNA replication products (Figure 4.1C). Visualization of these two- or three-color experiments has allowed for the characterization of numerous protein interactions with the replication fork [34, 75, 78, 357].

Previous single-molecule rolling-circle DNA replication studies, and interacting fluorescently-labeled proteins, have relied upon manual analysis of DNA molecules [34, 75, 78]. This involves selecting replicating DNA molecules and tracking the replication progression by manually detecting individual replication rate segments by hand (described in detail below). However, this method can introduce bias and variability among users. Further, the detection of interacting fluorescently-labeled proteins with the replication fork can be missed if the tracking of the DNA molecule is not done accurately. Thus, an automated detection method of the progress of the replication fork, and potentially interacting fluorescently-labeled proteins, is required.

Automated analysis methods of single-molecule data have continued to be developed for numerous assays. Single-molecule Förster resonance energy transfer assays have been the subject of automated analysis methods, most recently involving deep learning [359, 360]. Whilst promising, these assays often involve imaging stationary fluorescent spots that do not move under constant flow. The rolling-circle DNA replication assay involves continuous movement of the DNA molecules that cannot be easily detected by simple automated analysis methods. An analysis method that accurately detects the leading edge of the replicating molecule and short-lived associations of fluorescently-labeled proteins with the replication fork is required.

In this chapter, I describe and compare the previously applied manual analysis method to that of an optimized automated method to detect replicating rolling-circle DNA molecules and interactions of fluorescently-labeled proteins (Figure 4.2). The automated analysis method described involves detecting the leading edge of the rolling-circle DNA replication product over time, using custom-written ImageJ/Fiji and MATLAB macros. The coordinates of the leading edge of the replicating molecule are then used to detect interacting fluorescently-labeled Rep proteins with the replication fork. By detecting the position of the replication fork and any interacting proteins in an automated manner, this method minimizes bias introduced through manual analysis of the experiments.

and Fiji.

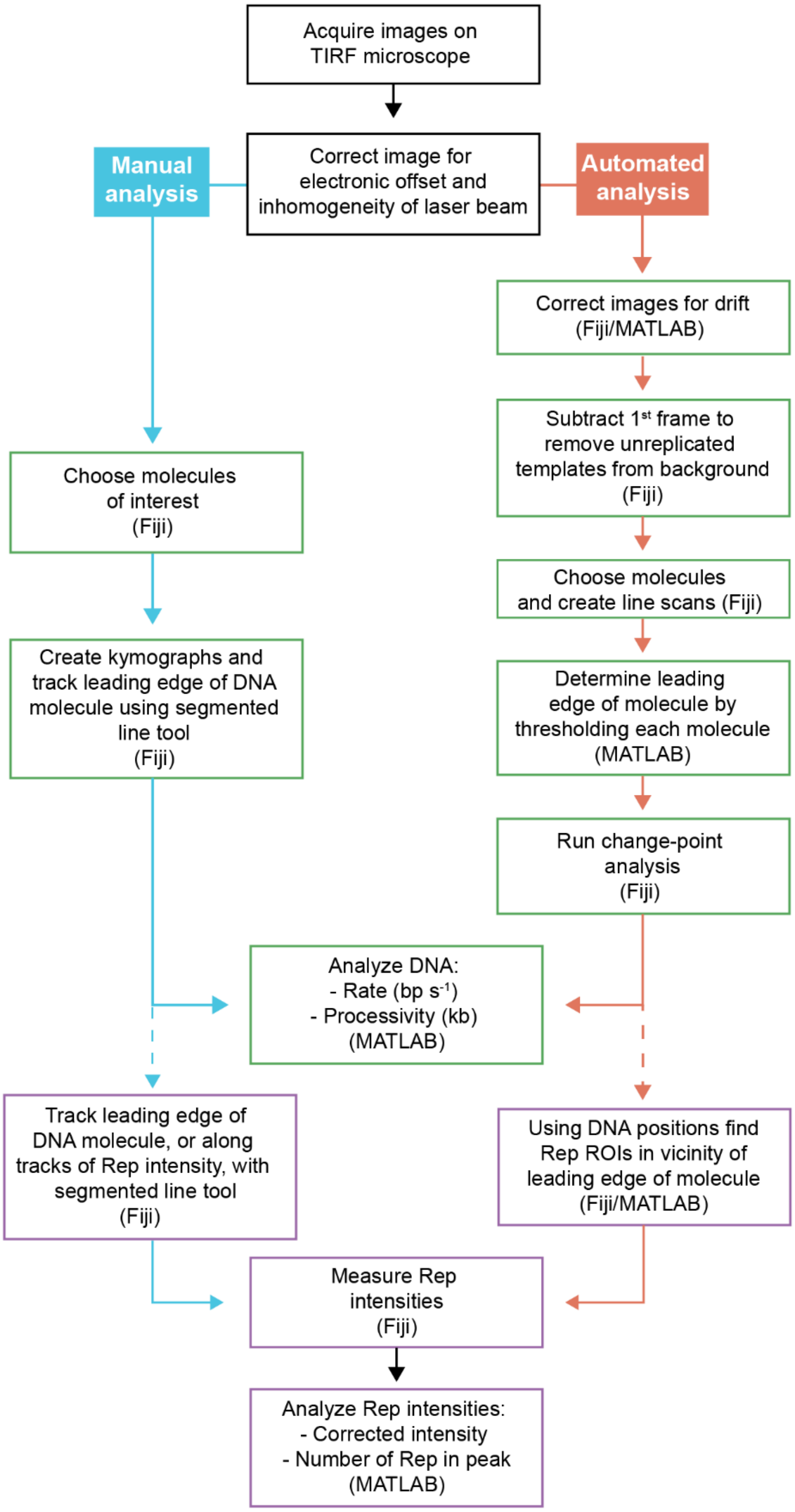


Figure 4.2: Overview of manual and automated image analysis methods of rolling-circle DNA replication. DNA replication (green) and fluorescent protein (purple) acquisitions were analyzed using manual (blue) or automated (orange) analysis methods involving various plugins and macros in MATLAB and Fiji.

4.2 Methods

4.2.1 Experimental setup

The reagents and data used for the optimization of automated analysis are described in detail in Chapter 6 (Sections 6.2.1, 6.2.3, 6.2.4 and 6.2.7). Briefly, *in vitro* single-molecule microscopy was carried out on an Eclipse Ti-E inverted microscope (Nikon, Japan) with a CFI Apo 100× oil immersion TIRF objective, as previously described [34, 74, 75, 78]. Replication reactions were carried out in microfluidic flow cells constructed from a PDMS flow chamber placed on top of a PEG-biotin-functionalized microscope coverslip.

The experiments involved visualizing fluorescently-labeled Rep proteins (Rep-AF647) in *E. coli* replication reactions where nuclease dead Cas9 (dCas9)-complementary gRNA 1 (cgRNA1) replication roadblocks are also in solution. Briefly, the dCas9-cgRNA1 complex was formed by pre-incubating the dCas9 protein with cgRNA1 for 5 min at 37°C in degassed single-molecule replication buffer (SM; 25 mM Tris-HCl pH 7.9, 10 mM $Mg(OAc)_2$, 50 mM potassium glutamate, 0.1 mM EDTA, 0.0025% Tween20, 0.5 mg mL^{-1} BSA, and 10 mM dithiothreitol). Next, the DNA-DnaBC complex was formed, by incubating 8 pM 2-kb rolling-circle DNA template with 7 nM $DnaB_6(DnaC)_6$ for 3 min at 37°C in degassed SM buffer containing 1 mM ATP. The solution, containing 150 nM Sytox orange, was adsorbed to the flow-cell surface at 10 µL min^{-1} until an appropriate surface density was achieved. Following, the replication solution was mixed as 30 nM Pol III $\alpha\varepsilon\theta$ core, 10 nM $\tau_3\delta\delta'\chi\psi$, 46 nM β_2, 75 nM DnaG and 20 nM SSB_4 in SM buffer containing 250 µM of each NTP, 50 µM of each dNTP, 20 nM Rep-AF647 and 0.25 nM pre-incubated dCas9-cgRNA1 complex. Reactions were initiated with the addition of the replication solution to the flow cell at 70 µL min^{-1} for 1 min and then slowed to 10 µL min^{-1} for 10 min. DNA and Rep-AF647 were imaged in one field of view per experiment, by sequentially exciting Sytox orange with a 532-nm laser (90 mW cm^{-2}) and Rep-AF647 with a 647-nm laser (200 mW cm^{-2}) for 200 ms once every second for 4 min.

4.2.2 Image Analysis

All analyses were carried out using ImageJ/Fiji (1.51w) and MATLAB 2016b, and plugins written in-house (the plugins are freely available at https://doi.org/10.5281/zenodo.7379064). The two platforms were used in combination as outlined in Figure 4.2. Before analyzing the data as described below, raw videos (.nd2 format) were converted to TIF files and flattened with the excitation beam profile, as described previously [361].

4.2.3 Manual analysis of rolling-circle DNA replication

4.2.3.1 Tracking of DNA trajectories

In a typical field of view (FOV), multiple rolling-circle DNA templates begin replicating during the acquisition (Figure 4.3A). To analyze the replicating DNA molecule for parameters such as rate and product length, individual molecules are manually selected and kymographs are produced in Fiji. Kymographs are 2D representations of movies of individual replication products, where all frames of the movie are placed along the time axis. Using these kymographs, individual rate segments are determined by eye, and a segmented line is drawn over the tip of the molecule. Using in-house built macros, the coordinates of each segment of the line are extracted. These coordinates can then be used in MATLAB to analyze rates, product lengths, segment durations and pausing events of individual DNA molecules.

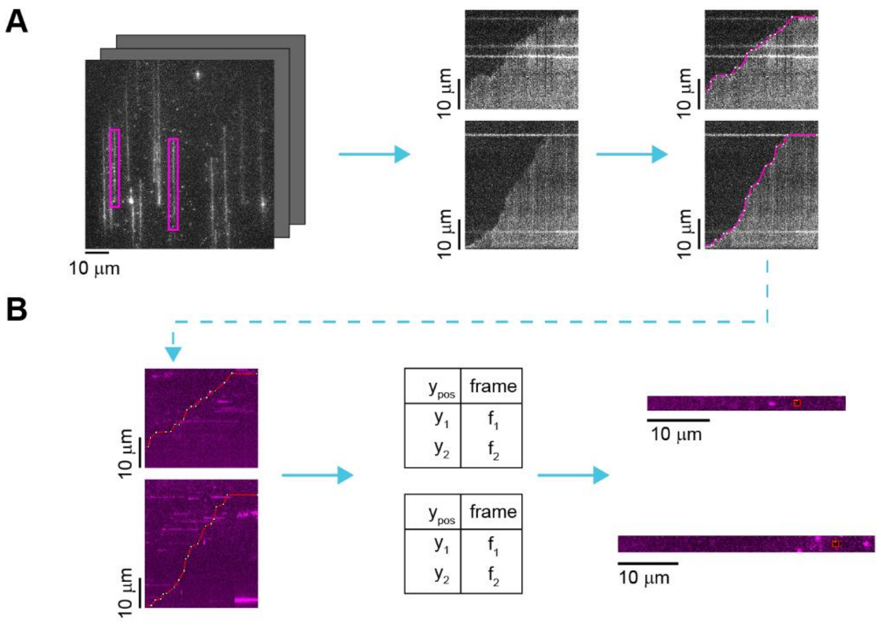


Figure 4.3: Manual analysis of DNA trajectories and fluorescent proteins. A) Individual replicating molecules are identified (magenta box) (left) from Sytox orange-stained fluorescence image stacks. Kymographs of each molecule is made (middle - top and bottom) and individual rate segments are detected by drawing segmented lines over the leading edge of the molecule (magenta line) (right). **B)** Segmented lines from the DNA channel are transposed onto the 647-nm channel (red line) (left), the y coordinates of each line segment are extracted (middle) and the 647-nm intensity is measured in an ROI for each coordinate saved (right).

4.2.3.2 Tracking of fluorescent proteins

Similarly to manual tracking of DNA trajectories, the tracking of fluorescent proteins during the same acquisition relies on drawing a segmented line over the tip of the replicating molecule in Fiji (Figure 4.3B). Either the same segmented line can be transposed into the kymograph of the fluorescent protein channel, or a new segmented line can be drawn over the top of the fluorescent protein track. Additionally, the image stack of the fluorescent-protein channel is used as described below.

First, using in-house written Fiji macros, the coordinates of the segmented line are extracted to determine the position of the tip of the replicating molecule for each frame (Figure 4.3B). Each coordinate becomes a region of interest (ROI) that is used to measure the integrated pixel intensity for each frame in the image stack of the fluorescent-protein channel. The size of the peak ROI must be consistent with the photobleaching analysis of the same fluorescent protein (described below). Here, a peak ROI of 5 × 5 pixels (width × height) is used to measure the pixel intensity inside each ROI (I_{peak}).

Next, the intensity of the local background (I_{bg}) of each peak ROI (*peak*) is measured in an ROI of at least twice the size. Here, a background ROI (*bg*) of 11 × 11 pixels is used. Using the measured pixel intensities of the peak and the background, the corrected intensity (I_{corr}) of the peak ROI is calculated as (equations 1 and 2):

$$I_{corr} = I_{peak} - \left(\left(I_{bg} - I_{peak}\right) \times constant\right) \tag{1}$$

$$where,\ constant = \frac{peak_{width}\, peak_{height}}{bg_{width}\, bg_{height} - peak_{width}\, peak_{height}} \tag{2}$$

The corrected intensities of the peak are then divided by the average intensity of a single fluorophore to determine the number of fluorophores, and thus the number of proteins, in a peak. The integrated intensity of a single fluorophore is determined by detecting photobleaching steps of non-specifically bound labeled proteins [34, 78]. For the experiment described in this chapter, the number of fluorophores in each peak provides the number of Rep monomers, as each monomer can have a maximum of 1 fluorophore as a result of site-specific single-cysteine labeling (described in detail in Chapter 6, Sections 6.2.2, 6.2.8 and 6.3.1).

Finally, to identify discrete binding and dissociation steps of fluorescent proteins from noise, a plot of the number of fluorophores over time is fitted using change-point analysis in Fiji [362-364]. Both the raw and fitted intensities are used in downstream analysis, for example, determining binding lifetimes, binding frequencies and stoichiometries.

4.2.4 Automated analysis of rolling-circle DNA replication

4.2.4.1 Tracking of DNA trajectories

To minimize the introduction of bias in detecting individual rate segments of replicating DNA molecules, an automated analysis method was required. Here, an analysis method is described that relies on both Fiji and MATLAB. The only element of bias in this method is introduced in the selection of molecules of interest (MOI), a step that is unavoidable at this stage without machine-learning algorithms.

Firstly, as is often the case with the type of single-molecule imaging setup used, there is detectable drift observed between each frame of an image stack. This drift will affect any downstream tracking of DNA molecules, where the macro possibly detects that the molecule has stopped replicating, while it has only drifted out of the ROI. To overcome this, all unreplicated

DNA templates in the first frame of the acquisition are tracked throughout the acquisition, using the in-house written plugin “finder, fitter, tracker” in Fiji (Table 4.1). The mean drift in both the x and y directions of the acquisition is calculated using MATLAB. Returning to Fiji, each frame is translated by the calculated mean drift resulting in a drift-corrected acquisition.

Table 4.1: Inputs used for acquisition drift correction. Using “finder, fitter, tracker” Fiji plugin.

Peak finder settings:	
Inner radius	1.0
Outer radius	3.0
Threshold	3.0
Selection Radius	4.0
Minimum distance between pixels	8.0
Rejection criteria:	
Max error baseline and height	5000
Max error: x, y, σx and σy	5.0
Max difference x and y	5.0
Minimum trajectory length	Number of frames - 1

In addition to actively replicating molecules, there are also several unreplicated DNA templates bound to the coverslip (Figure 4.4A). If these unreplicated templates are bound in the path of the replicating molecule, they will first be picked up by the automated tracking and determined to be either the start or end point of the replicating molecule. To overcome this, in Fiji the first frame of the acquisition is subtracted from each consecutive frame (Figure 4.4A). If acquisitions contain more than one fluorescence channel, both the drift correction and frame subtraction are applied to all channels simultaneously.

To track replicating DNA molecules, each MOI needs to be selected as an ROI in Fiji (Figure 4.4A). In the in-house built Fiji macros used here, the molecules need to replicate from the top of the FOV to the bottom and are required to do so as perfectly vertical as possible. Next, for each MOI, a vertical intensity profile for each frame is plotted and saved for downstream analysis to detect the leading edge of the replicating molecule (Y position). Next, the ROIs for each MOI are expanded by 20 pixels (height and width) and saved for downstream analysis. From this expanded ROI, a horizontal intensity profile is plotted for each position down the length of a projected average of the last 10 frames. This is later used to determine the position of the MOI on the X-axis of the acquisition.

Switching to MATLAB, both the X and Y positions of the replicating DNA molecule are determined. To detect where the leading edge of the replicating molecule is for each frame of the acquisition (Y position) a threshold for each MOI is calculated (Figure 4.4A). First, a simple moving average is calculated for the intensity profile of each frame to obtain “smoothed” data. Next, the threshold for the background of the intensity profiles is calculated. The mean background intensity is calculated (μ_{bg}) from the outer edges of the ROI, and 2.5 times the standard deviation from the center of the ROI (σ_{Int}). Ideally, the mean background intensity would not include any intensities from the DNA molecule. Using these parameters the threshold (T) is calculated, by determining the value that encompasses 99.9% of intensity values that are background values, as (equation 3):

$$T = \mu_{bg} + 0.5(\mu_{bg} + \sigma_{Int}) \tag{3}$$

The calculated threshold is used to find the leading edge of the molecule in each frame and a 3-level discrete wavelet transform is applied to reduce noise in the data.

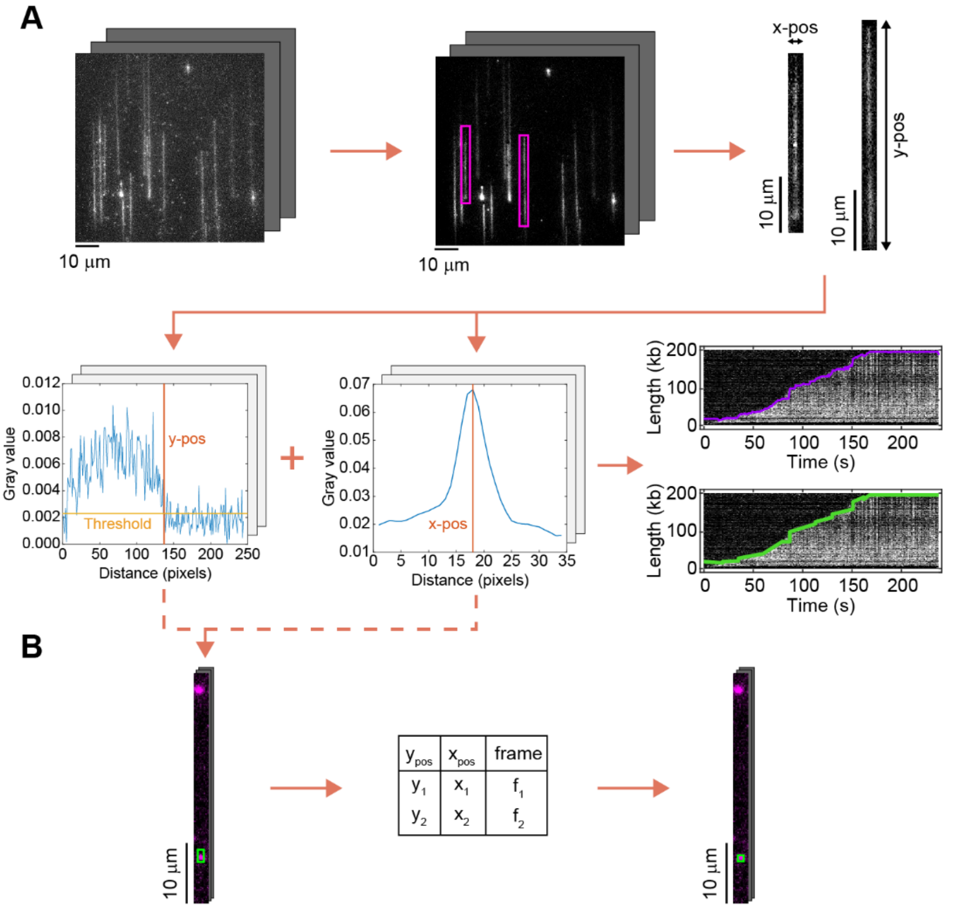


Figure 4.4: Automated analysis of DNA molecules and fluorescent proteins. A) Image stacks are corrected for drift between frames and unreplicated DNA templates are subtracted from consecutive frames (top - left and middle). Molecules of interest are selected (magenta box) (top – middle) and both vertical and horizontal line scans are produced for each frame to determine the y position (y-pos) and x position (x-pos) of the DNA product, respectively (top - right). The y-pos is then calculated by determining a threshold for each MOI (bottom – left). The x-pos is calculated by determining the max gray value (bottom – middle). The automated tracing of the DNA molecule can then be plotted and superimposed over the kymograph (bottom–right; raw trace (purple) and change-point analysis (green)). **B)** Using both y-pos and x-pos coordinates from the DNA tracing, ROIs can be created in the Rep-647 channel to find peaks within the region of the leading edge of the molecule (left). Any Rep-647 peaks found in the ROI are determined to be replication fork-associated and the coordinates saved (middle). The coordinates are used to measure the intensity of the associated Rep-647 peak (right).

To track the leading edge of the replicating molecule and detect fluorescent proteins that may be bound, the position of the replicating molecule on the X-axis of the saved acquisition needs to be determined accurately. Using the horizontal intensity profiles determined in Fiji, a moving average is calculated for each profile in MATLAB. The mean intensity value of the smoothed data is then calculated for each X position. The maximum value of the mean intensity data and corresponding X position denotes the center of the replicating molecule (Figure 4.4A). Both the calculated Y and X positions of the molecules are used in downstream analysis to track the leading edge of the replicating molecule and detect any bound fluorescent proteins.

Finally, a change-point analysis is applied to the tracked DNA data (Y positions) in Fiji (Figure 4.4A), to detect individual rate segments of the replicating molecule. The resulting data are saved and used in downstream analysis to determine rates, product lengths, and pausing events.

4.2.4.2 Tracking of fluorescent proteins

Tracking of fluorescently-labeled proteins during rolling-circle replication has been used to determine the exchange rates and stability of individual replisome components [34, 75, 78]. The method of imaging determines the analysis method required for the accurate detection of fluorescent proteins at the replication fork. When imaging Sytox Orange-stained DNA and fluorescently-labeled proteins (for example, AF647) simultaneously using a dual-view device, the position of the replication fork at any given time is the same in both color channels. In this case, the coordinates of the tip of the replicating molecules can be transposed into the 647-nm channel and intensities of fluorescent proteins can be calculated. However, when imaging different fluorescence channels sequentially (i.e. switching between laser lines), as done in the data described here, the position of the replication fork can move in the hydrodynamic flow between switching.

To account for the movement of the tip of the replicating molecule, in Fiji the coordinates determined by DNA tracking (DNA ROIs – 5 × 5 pixels) are expanded to find potentially bound fluorescent protein peaks in the 647-nm channel (Figure 4.4B). Specifically, the DNA ROIs are expanded to double the width (10 pixels) and five times the height (25 pixels). This ensures any peaks that may be associated with the DNA are captured and account for any potential movement of the DNA that might occur. The in-house built plugin, “peak finder”, is used to find peaks within the expanded ROI in the 647-nm channel for each frame. The coordinates of the detected peaks are saved and used for downstream analysis.

Next, in MATLAB, the 647-nm peaks found for each frame undergo testing to determine if they are replication fork-associated. This involves subtracting the x and y coordinates of the Rep ROIs from the corresponding DNA ROI coordinates. If the results are less than or equal to a threshold of 10 (for Y) and 5 (for X), the Rep ROI coordinates are saved for further analysis (Figure 4.4B). The thresholds used are arbitrarily defined, by looking at the average difference in pixels that the molecule moves between consecutive DNA frames.

Finally, in Fiji, the integrated intensities of each of the replication fork-associated Rep ROIs are measured in an ROI of 5 × 5 pixels. For each ROI, the local background intensity is also measured by integrating the pixel intensity of an ROI twice the size on the x-axis (11 × 5 pixels). The ROI is only expanded on the x-axis to avoid potentially picking up other Rep ROIs that might be bound behind or ahead of the replicating molecules. These measurements are saved and used to calculate the corrected intensity, the number of fluorophores in each peak, and discrete association and dissociation events, as described in section 4.2.3.2.

4.3 Validation

Rolling-circle DNA replication was visualized in the presence of both dCas9-cgRNA1 complexes and Rep-AF647 proteins in solution. This experimental setup, described in detail in chapters 5 and 6, results in the dCas9-cgRNA1 roadblock periodically stalling the replisome and subsequent rescue by Rep-AF647. Briefly, rolling-circle DNA templates were bound onto the coverslip surface and replication was initiated by the introduction of a laminar flow of buffer containing the replisome components, dCas9-cgRNA1 and Rep-AF647. Kymographs of the resulting replicating DNA molecules show frequent interspersed pausing and rescuing events, visible by imaging the Sytox-stained DNA (Figure 4.5A). Additionally, visualization of the Rep-AF647 molecules shows frequent and dynamic association to the tip of the replicating DNA molecule and associated replisome (Figure 4.5B).

In total, 21 DNA molecules were analyzed both manually and with the optimized automated analysis method. Automated tracking of the tip of the replicating molecule over time results in a trace of the kymograph (Figure 4.5A). These traces show the intricate changes in replication rates and processivity of the replisome. Further, analysis of individual rate segments, detected by change-point fitting, result in a median rate of 445 ± 37 bp s^{-1} (median ± S.E.M) (Figure 4.5C). This value is slightly lower than manual analysis of the same DNA molecules (603 ± 27 bp s^{-1}), however, both are consistent with previous characterizations of *E. coli* DNA replication rates [34, 48, 49, 75]. Differences between automated and manual analysis of the same DNA molecules are likely attributed to the over-fitting of rate segments in manual analysis, where more segments were detected by eye. This discrepancy highlights the implicit bias that can be introduced through manual analysis, which can also differ between users.

Analysis of the number of Rep-AF647 molecules associating over time with the same 21 molecules used above, resulted in similar traces when analyzed manually or using the automated algorithm (Figure 4.5B). Differences in the exact number of molecules detected and the absence of large negative numbers in the automatically analyzed traces are likely attributed to the more accurate positioning of the ROIs in the automated tracking; the automated tracking of fluorescent molecules uses the raw trace of the DNA molecules (orange lines in Figure 4.5A), whereas the manual analysis uses the manually drawn segmented line. Further, automated change-point analysis of the number of Rep-AF647 molecules over time reveals a mostly monomeric stoichiometry (Figure 4.5D), consistent with manually analyzed data.

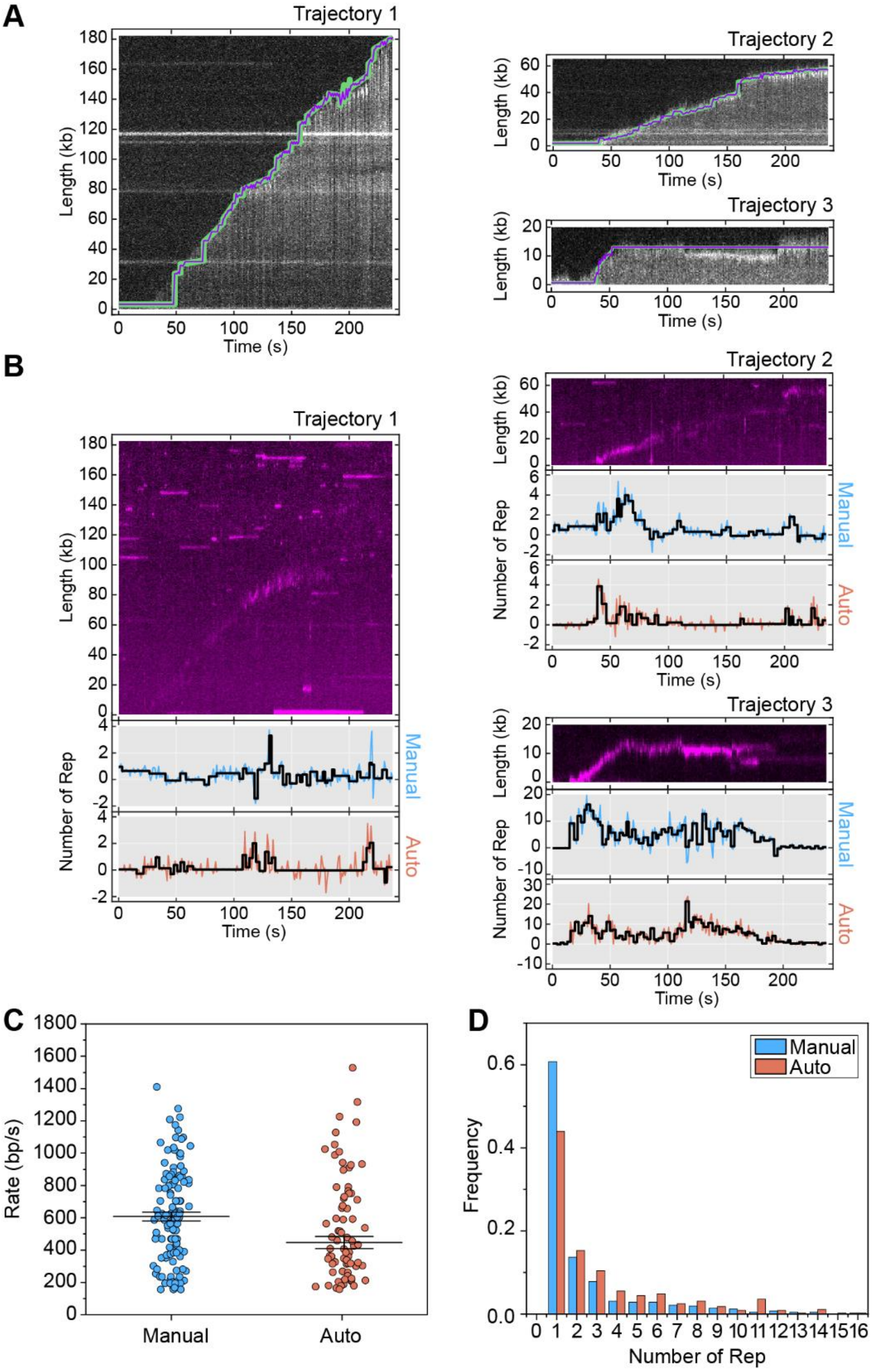


Figure 4.5: Validation of automated analysis of rolling-circle DNA replication. A) Three example kymographs of replicating DNA molecules with automated detection of the leading edge of the replicating molecule (purple) and change-point fitting (green). **B)** Example kymographs of Rep-AF647 fluorescence from DNA molecules in A) (top). Detected stoichiometries for Rep over time by manual (middle) and automated (bottom) analysis, where colored lines are intensity traces and black lines are change-point fitted. **C)** Rate of replication determined from manual (blue - 600 ± 30 bp s^{-1} (median ± S.E.M), n segments = 122) or automated (orange - 450 ± 40 bp s^{-1}, n segments = 74) analysis of 21 example DNA molecules. **D)** Histogram distribution of Rep-AF647 stoichiometry at the replication fork determined by manual (blue; n = 423 binding events) and automated (orange; n = 450 binding events) analysis methods.

We present here an automated tracking method of replicating rolling-circle DNA molecules and associated fluorescently-labeled proteins. This method of analyzing replication events results in similar replication rates and stoichiometries of associating fluorescent proteins to manually analyzed data and previously reported data. Importantly, this method eliminates the introduction of bias and over-fitting of DNA molecules that can affect the analysis of fluorescent protein association. This analysis method will prove important in the labor-intensive analysis of complex single-molecule experiments where parameters such as stoichiometries and association frequencies are reported for vital replisome components.

Chapter 5

Single-molecule fluorescence imaging of DNA replication stalling at sites of nucleoprotein complexes

Kelsey S. Whinn, Nischal Sharma, Antoine M. van Oijen, Harshad Ghodke

Version of manuscript accepted at Single Molecule Analysis. Methods in Molecular Biology

DNA replication in cells occurs on crowded and often damaged template DNA, forming potentially deleterious roadblocks to the progressing replication fork. Numerous tools have been developed to investigate the mechanisms of DNA replication and the fate of stalled replication forks. Here, we describe single-molecule fluorescence imaging methods to visualize processive DNA replication and replication fork stalling at site-specific nucleoprotein complexes. Using dCas9 as a protein barrier and rolling-circle DNA templates, we visualize effective, stable, and site-specific blocking of the *Escherichia coli* replisome. Additionally, we present a protocol to produce an 18-kb rolling-circle DNA template that provides increased spatial resolution in imaging the interplay between replisomes and roadblocks. These methods can be used to investigate encounters of the replisome with nucleoprotein complexes at the single-molecule level, providing important mechanistic details of replisome stalling and downstream rescue or restart pathways.

Contributions

This chapter describes the key methods used to visualize stalled rolling-circle DNA replication that is used throughout Chapter 6. It details the theory, experimental methods and analysis using dCas9 as a protein barrier.

I designed the experimental protocols for each of the stalled replication assays, constructed the 2-kb rolling-circle DNA template following previously established protocols, carried out all example experiments and analyses, drafted the manuscript, and designed and produced figures.

5.1. Introduction

Cell proliferation in all organisms requires high-fidelity replication of all genetic material. DNA replication occurs on crowded and often damaged DNA templates, presenting a wide range of potential roadblocks for progressing replication forks. If left unresolved, these roadblocks can lead to replication fork collapse and ultimately, genome instability.

Numerous techniques have been developed to investigate the fate of progressing replication forks when faced with these roadblocks. Specifically, multiple *in vitro* tools have been established to mimic physiologically relevant roadblocks in cells. These tools include DNA templates containing site-specific DNA lesions [365, 366] and model protein roadblocks, including tandem repressor-operator complexes, prokaryotic replication termination sites (Tus-*ter*), and stalled RNA complexes [64, 68, 367]. However, each tool possesses significant shortcomings in experimental throughput and ease of use. Recently, we developed a site-specific, stable and simple replication-blocking tool using the catalytically inactive Cas9 protein (dCas9) [74]. We show that effective and long-lived replication-blocking in bacterial, viral and yeast systems occurs independently of the strand targeted using rolling-circle DNA templates. This dCas9 tool has since been used to investigate the activity of yeast accessory helicase, Pif1, in removing nucleoprotein complexes from linear DNA templates [156].

Rolling-circle DNA amplification has been widely used as a method to investigate mechanistic questions relating to DNA replication [22, 34, 48, 74, 75, 78, 368]. Based on a natural mechanism for replication of bacteriophage DNA, rolling-circle amplification of DNA involves nicked, covalently-closed circular double-stranded DNA where the internal strand serves as the template for the leading strand [356, 369]. Loading of a bacterial replicative helicase (5'-3') is facilitated by a 5'-non-complementary single-stranded overhang. Rolling-circle DNA templates have recently been constructed using nicked double-stranded DNA (dsDNA) plasmids allowing for controllable gap size and lengths of ssDNA overhangs [356]. These DNA templates can be replicated continuously, allowing for the detection of processive replication whilst being easy to develop.

Single-molecule techniques continue to highlight the heterogeneity that exists in complex biological processes [90-92]. The rich temporal and spatial details provided by these techniques provide insight into the diverse behaviors of single molecules that would otherwise be missed in ensemble averaging methods. Numerous single-molecule studies have captured the intermediate states of DNA replication, repair, and recombination. Additionally, fluorescence-based techniques allow visualization and characterization of the activity of individual components during the reaction.

Here, we describe the experimental protocols underpinning two single-molecule fluorescence techniques that allow the visualization and quantification of replication stalled at sites of nucleoprotein complexes. Specifically, we describe the use of single-molecule rolling-circle replication assays, using either a 2-kb or 18-kb DNA template, to detect blocked replication events using the dCas9 replication-blocking tool. The two different lengths of DNA templates are used to highlight trade-offs between temporal and spatial resolution on the one hand and experimental throughput on the other.

In both assays described, the DNA templates are tethered to the functionalized surface of a microfluidic flow cell and stained with fluorescent intercalating dye to allow real-time visualization of flow-stretched DNA replication products. The two assays herein termed ‘roadblock pre-incubation assay’ and ‘roadblock in-solution assay’, use either of the two DNA templates. These two types of single-molecule replication-blocking assays, while similar in technique, each provide different outputs. The roadblock pre-incubation assay reveals blocked templates simply upon observing the difference in DNA lengths after the introduction of replisome components; unblocked templates will replicate processively, while blocked templates will undergo replication up until the roadblock is encountered. The roadblock in-solution assay introduces the roadblock at the same time as the replisome components, resulting in processive replication until the roadblock binds and is encountered by the replisome. Whereas both assays provide the DNA product length as an output, the dynamic behaviors of the replisome under the two conditions together provide insight into the affinity and stability of the roadblock. Additionally, these assays represent an experimental platform that could be adapted to visualize other genomic-maintenance pathways, for example, replication-fork collapse and stalled-replication rescue, where the rolling-circle template can provide multiple events per replicating molecule.

5.2. Materials

5.2.1. *Surface functionalization*

1. 24 × 24 mm glass microscope coverslips.
2. Microscope-coverslip staining jar.
3. 100 % Ethanol, absolute anhydrous.
4. 1 M potassium hydroxide (KOH).
5. Acetone.
6. 3- aminopropyltriethoxysilane.
7. Functionalized PEGs, here mPEG-Succinimidyl valerate (mPEG-SVA-5000) and Biotin-PEG-SVA 5000 (Laysan Bio USA). Store under N_2 gas at – 20 °C.
8. PEG coupling buffer: 100 mM $NaHCO_3$, pH 8.2.
9. Bath sonicator.
10. Compressed nitrogen gas.
11. Vacuum desiccator.

5.2.2. Rolling-circle template preparation

5.2.2.1. *2-kb rolling-circle DNA template*

1. Thermocycler
2. pSCW01 plasmid DNA (Aldevron, USA).
3. Capture oligos 1-3 (Integrated DNA Technologies, USA) (Table 5.1).
4. NEB 3.1 buffer (New England Biolabs, USA).
5. Nt.bstNBI (New England Biolabs, USA).

6. Polyethylene Glycol 8000 (PEG 8000).
7. 1 M Magnesium chloride ($MgCl_2$).
8. 70% (v/v) Ethanol.
9. Cold benchtop centrifuge.
10. CutSmart buffer (New England Biolabs, USA).
11. Oligo 4 (Integrated DNA Technologies, USA) (Table 5.1).
12. 1 M Dithiothreitol (DTT).
13. 100 mM Adenosine Triphosphate (ATP).
14. T4 ligase (New England Biolabs, USA).
15. Nuclease-free water.

5.2.2.2. 18-kb rolling-circle DNA template

1. Thermocycler.
2. pUBER plasmid DNA (Aldevron, USA).
3. Dig-competitor oligonucleotide (Sigma-Aldrich, USA) (Table 5.1).
4. CutSmart buffer (New England Biolabs, USA).
5. Nt.*Bbv*CI nickase (New England Biolabs, USA).
6. Fork oligonucleotide (Sigma-Aldrich, USA) (Table 5.1).
7. 1 M Dithiothreitol (DTT).
8. 100 mM Adenosine Triphosphate (ATP).
9. T4 ligase (New England Biolabs, USA).
10. TE buffer; 10 mM Tris.HCl pH 7.6, 1 mM EDTA, 300 mM NaCl.
11. Sepharose 4B (1× 25 cm; Sigma-Aldrich) column.

Table 5.1: Custom oligonucleotides

2-kb rolling-circle DNA template		
Capture oligo 1	5′-ATT TGA CTC C	
Capture oligo 2	5′-CAT GGA CTC GCT GCA G	
Capture oligo 3	5′-GAA TGA CTC GG	
Oligo 4	5′-/bio/AAA AAA AAA AAA AAA AGA GTA CTG TAC GAT CTA GCA TCA ATC ACA GGG TCA GGT TCG TTT GGG AGT CAA AT	
18-kb rolling-circle DNA template		
Dig-competitor oligo	5’ - GGA CTT AAG TGC TGA TTT TTT TTT T/Dig/	
Fork oligonucleotide	5′-/bio/AAA AAA AAA AAA AAA AGA GTA CTG TAC GAT CTA GCA TCA ATC ACA GGG TCA GGT TCG TTG CAC TTA AGT CC	
crRNAs		*Notes*
crRNA1	5’ - ACAATTAATAGACTGGATGG	Targets 1.2 kb from the fork in 2-kb template; 17 kb in 18-kb template.
crRNA2	5’ - GGTGTGAAAGAACACCAACA	Targets 3 kb away from fork in 18-kb template.
crRNA3	5’ - CTGGTGAACTTCCGATAGTG	Targets 13 kb away from fork in 18-kb template.

5.2.3. Experimental setup

1. Functionalized coverslip (from section 5.2.1).
2. PDMS mold: stainless-steel mold that is laser-engraved with a ridge of measurements 0.1 × 0.5 × 19 mm.
3. Polydimethylsiloxane.
4. 184 Silicone elastomer.
5. PE-60 tubing (Bioseb, 0.76 inner diameter, 1.22 mm outer diameter).
6. Microscope-compatible flow cell holder.
7. Syringe pump (SyringeONE) with 5-mL syringe.
8. Syringe pump software: SyringePumpPro.
9. Blocking buffer, 1×: 50 mM Tris-HCl (pH 7.9), 50 mM KCl, 2% (*v/v*) Tween-20.
10. Wash buffer, 1×: 25 mM Tris-HCl (pH 7.9), 10 mM magnesium acetate, 50 mM potassium glutamate, 0.1 EDTA, 0.0025 % (*v/v*) Tween-20, 0.5 mg/mL BSA.
11. 1 mg/mL Neutravidin in PBS, pH 7.3.
12. Vacuum desiccator.
13. Nikon optical microscope body, 100× TIRF objective (N.A. = 1.49, oil).
14. Compressed air or nitrogen gas.

5.2.4. Visualization and analysis of DNA replication stalling

1. Sytox Orange (S.O.) nucleic-acid stain.
2. 100 µM dNTP bundle. Mix all 4 dNTPs equally to make a 25-µM dNTP solution.

3. 100 µM NTP bundle. Mix all 4 NTPs equally to make a 25-µM NTP solution.
4. Biotinylated rolling-circle DNA template (from section 5.2.2.1 or 5.2.2.2).
5. Heating block capable of going to temperatures up to 150°C.
6. *E. coli* replication buffer: 25 mM Tris-HCl (pH 7.9), 10 mM magnesium acetate, 50 mM potassium glutamate, 0.1 EDTA, 0.0025 % (*v/v*) Tween-20, 0.5 mg/mL BSA.
7. Purified components of the *E. coli* replisome: typically 70 nM DnaB$_6$C$_6$, 75 nM DnaG, 30 nM αεθ, 46 nM β$_2$, 10 nM τ$_3$δδ'ψχ, 20 nM SSB.
8. Fluorescently-labeled gRNAs. These can be constructed by annealing fluorescently-labeled trcrRNAs to crRNAs, containing the complementary sequence to the target DNA site, purchased from Integrated DNA Technologies (IDT), USA. See section 5.3.4.1 (Table 5.1).
9. Purified dCas9 proteins. Here, proteins as described in [74] are used. dCas9 proteins can also be purchased from New England Biolabs (USA).
10. Heated microscope stage insert.
11. Lasers that match the excitation wavelength of the fluorophores used (we typically applied 532-nm for Sytox Orange, and 647-nm for Atto647).
12. Excitation and emission filters appropriate for each laser line.
13. EMCCD camera (Andor iXON 897, UK).
14. Imaging software: Nikon Elements Advanced Research with JOBS module.
15. Acquisition analysis software: ImageJ/FIJI (version 1.51w) with custom plugins used during the analysis process. These plug-ins are freely available on GitHub, https://github.com/LMSpenkelink/SingleMoleculeReplication and https://github.com/SingleMolecule.
16. Data-analysis software: MATLAB 2016b and OriginPro 2021b.

5.3. Methods

5.3.1. Glass coverslip surface functionalization

The coverslip functionalization protocol described here is adapted from previously described methods [358]. Rolling-circle DNA templates are tethered to glass microscope coverslips via a biotin-streptavidin-biotin linkage. To functionalize, glass coverslips are first reacted with the alkoxy groups of aminosilane, resulting in a surface with reactive amine groups. Following, a mixture of biotinylated and non-biotinylated succinimidyl valerate-PEG is coupled to the amine-functionalized glass, forming a layer of PEG displaying a mixture of biotin and nonreactive groups. This mixture forms an inert layer to reduce nonspecific interactions between the glass surface and the biological complexes under study.

1. To extensively clean commercially available glass microscope coverslips before functionalization, coverslips are placed in glass staining jars, or equivalent, and sonicated for 20 min in ethanol. Rinse the coverslips in the jar with Millipore water and sonicate again for 20 min in 1 M KOH. Rinse and repeat both EtOH and KOH sonications. After the second KOH sonication, refill the jar with Millipore water and sonicate for 20 min.

2. Before the silanization step, remove all traces of water from the coverslips and jar using compressed N_2 and by rinsing the jar in acetone. Refill the jar with approximately 250 mL of acetone. Add 5 mL of 3-aminopropyltriethoxysilane solution to an approximate 2% (v/v) final concentration. Return coverslips to the jar and incubate covered at room temperature for 3 min. Extensively rinse coverslips in Millipore water and dry with compressed N_2. Store coverslips briefly in a vacuum desiccator.
3. Remove PEG stocks from the freezer and allow them to warm to room temperature. This will prevent condensation, as the PEG succinimidyl esters are labile and will hydrolyze. Prepare a solution of the PEGs in 0.1 M $NaHCO_3$, pH 8.3 (50 µL per coverslip). Mix the methylated (mPEG-SVA-5000) and biotinylated PEG (biotin-PEG-SVA-5000) at a ratio of 0.6%:15% (w/v) biotin-PEG:mPEG (Note 1). Vortex mix the solution and centrifuge at max speed for 1 min to remove undissolved PEG. To avoid hydrolysis, immediately pipette 50 µL onto a silanized coverslip and place another silanized coverslip on top of the liquid bubble. This forms a coverslip-PEG-coverslip sandwich and ensures good spreading of PEG solution across the entire surface of each coverslip and prevents drying during incubation (Note 2). Place coverslips in a dark cupboard and incubate for 3-4 h.
4. Repeat step 3 PEG solution mixture. Separate each sandwich, noting which side has been functionalized, and rinse thoroughly with Millipore water. Dry using compressed air or N_2. Place the coverslip functionalized side up, and pipette 50 µL of fresh PEG solution in the center. Immediately place another clean and dry coverslip, functionalized side down to form another coverslip-PEG-coverslip sandwich. Incubate coverslips overnight in a dark cupboard. Repeating the functionalization step increases the coverage of PEG on the surface of the coverslip, increasing the number of available biotin groups for DNA immobilization.
5. Separate each sandwich and thoroughly rinse with Millipore water. Dry using compressed air or N_2. Store the coverslips at room temperature in a vacuum desiccator for at least 1 week without loss of quality (Note 3).

5.3.2. Rolling-circle DNA template construction

Rolling-circle DNA templates, using nicked plasmids as a basic structure, allow for processive replication. In the protocol described here, the 5' ssDNA overhang provides a replication fork-like structure where the helicase can load and begin unwinding the dsDNA plasmid. The internal strand then serves as a template for the leading strand, which is displaced from the circle by helicase unwinding activity and serves as the template for lagging-strand synthesis. This template is tethered to the surface of a microfluidic flow cell, with the application of a laminar flow of buffer – containing the components required to sustain the replication reaction – stretching the growing dsDNA product at a rate that is determined by the replication rate.

5.3.2.1. *2-kb rolling-circle DNA template*

The 2-kb rolling-circle DNA template is constructed from the plasmid pSCW01 (2030 bp) as described previously [356]. Briefly, this plasmid is nicked in four sites on the same strand by the nickase, Nt.*Bst*NBI. The nicked oligonucleotides are displaced to yield a 37-nt-long ssDNA region, allowing a partially complementary fork oligonucleotide to be annealed and ligated to generate a gap and overhang presenting replication-fork topology. Typically the construction of the 2-kb rolling-circle template is carried out in 5 steps.

1. *Nicking*: Treat 400 µg of pSCW01 plasmid DNA with 1.5 units/µg of Nt.*Bst*NBI, in the presence of 50 × molar excess of displacer oligonucleotides complementary to the fragments to be removed to create the gap (oligo 1-3 (Table 5.1)), in 1 × NEB buffer 3.1 at 55°C for 4 h in a thermal cycler. Immediately increase the temperature to 85°C for 10 min to inactivate the nickase. Following, anneal the displacer oligos by decreasing the temperature slowly at a rate of 1°C/min until the reaction reaches 14°C.
2. *Purification of the nicked plasmid*: Purify the nicked plasmid from excess displacer oligos by PEG purification by adding an equal volume of freshly made 2 × solution containing 26% (*w/v*) PEG-8000 and 20 mM $MgCl_2$ in Millipore water to the cooled reaction mixtures containing the DNA. Centrifuge the PEG-DNA mixture at 6°C for 1 h at 21, 000 × g. Remove the supernatant and gently resuspend and wash the pellet with 1.5 mL cold 70% (*v/*v) ethanol. Centrifuge the resuspended pellet at 6°C for 15 min at 21,000 ×g. Remove as much ethanol supernatant as possible and place the tube upside down on tissue to evaporate excess ethanol, ensuring the pellet does not completely dry. Resuspend the pellet, containing the gapped plasmid, in 100 µL Millipore water previously warmed at 65 °C (Note 4).
3. *Annealing*: Anneal the fork oligonucleotide (oligo 4 (Table 5.1)) in the presence of 3 × molar excess over the gapped DNA substrate in 1 × CutSmart buffer at 50 °C for 10 min, followed by slow cooling to 16°C at a rate of 1°C/min. Here, we use a fork oligonucleotide that is a 71-mer ssDNA molecule containing 12-nt 3'-sequence complementary to the gapped region of pSCW01. Hybridization of the oligonucleotide to pSCW01 results in a 25-nt gap (Note 5).
4. *Ligation*: Ligate the fork oligonucleotide to the gapped plasmid by adding 62.5 units per µg of DNA substrate in the reaction mixture, supplemented with 8 mM ATP and 10 mM DTT. Incubate the reaction at 16°C for 18 h in a thermal cycler, followed by inactivation of the ligase by heating at 65°C for 10 min.
5. *Purification*: Purify the rolling-circle DNA template by PEG purification as outlined in step 2. Resuspend in Millipore water and store at -20°C. For long-term storage at -80°C, resuspend in TE buffer. (Note 6)

5.3.2.2. 18-kb rolling-circle DNA template

Construction of the 18-kb rolling-circle DNA template is adapted from previously described methods using the plasmid pUBER (18.4 kb) [365, 366]. This plasmid is nicked at two sites on the same strand using the nickase Nt. *Bbv*CI. The nicked oligonucleotide is displaced to yield a 15-nt gap where a partially complementary fork oligonucleotide is annealed and ligated to resemble replication-fork topology. The 18-kb rolling-circle template increases the spatial resolution by ninefold compared to that of the 2-kb template. Typically, the construction of the 18-kb rolling-circle template was carried out by the following six steps.

1. *Nicking*: Treat 50 µg of pUBER plasmid with 1 unit/µg Nt.*Bbv*CI in 1× CutSmart buffer at 37°C for 4 h in a thermal cycler to create the ssDNA gap (Note 7).
2. *Hybridization of competitor oligo*: Add 10-fold molar excess of Dig-competitor oligonucleotides (Table 5.1) to the reaction. Raise the temperature to 65°C for 20 min and anneal the competitor oligonucleotides by decreasing the temperature slowly at a rate of 1°C/min until the temperature reaches 14°C.
3. *Displacement of hybridized oligonucleotides*: Remove the competitor oligonucleotides (either non-hybridized or hybridized with displaced 15 nucleotide oligonucleotides) from the gapped plasmid by magnetic separation using 1 µg/nmol tosyl-activated paramagnetic beads functionalized with anti-digoxigenin Fab fragments. Specifically, equilibrate the beads by washing 3 times with 1X CutSmart buffer. Add the gapped plasmid reaction to the beads and incubate for 30 min at room temperature on a hula mixer. Collect the supernatant.
4. *Annealing*: Anneal the fork oligonucleotide in the presence of 100× molar excess over the gapped DNA substrate in 1× CutSmart buffer at 50°C for 10 min, followed by slow cool to 16°C at a rate of 1°C/min.
5. *Ligation*: Ligate the fork oligonucleotide by adding 62.5 units/µg of T4 DNA ligase, and supplementing with 12 mM ATP and 10 mM DTT. Incubate the reaction at 16°C for 18 h in a thermal cycler, followed by inactivating the ligase by increasing the temperature to 65°C for 10 min.
6. *Purification*: Purify the 18-kb rolling-circle DNA template from excess fork oligonucleotides using the gel-filtration chromatography previously described [366] (Note 8). Briefly, load the ligation reaction on a Sepharose 4B (1× 25 cm; Sigma-Aldrich) column in the presence of 12 mM EDTA and 300 mM NaCl. Elute the purified 18-kb DNA fractions with TE buffer (10 mM Tris.HCl pH 7.6, 1 mM EDTA, 300 mM NaCl) (30% recovery). Store the DNA at -80°C.

5.3.3. Experimental setup

After functionalizing the glass coverslips and preparing the rolling-circle DNA templates, a flow cell can be constructed to perform single-molecule rolling-circle DNA replication experiments. DNA templates are then introduced to the flow cell and coupled to the coverslips through a DNA-biotin-streptavidin-biotin linkage, resulting in the starting template for single-molecule DNA replication stalling events. The experimental setup is adapted from previously described methods [48, 358, 370]

1. Polydimethylsiloxane (PDMS) is a silicon-based polymer that is generally inert and easy to use to create custom flow cell lids that contain imprints of a channel. To make our flow cell lids, vigorously mix PDMS with a silicone-curing agent (1:10 ratio). Remove the air bubbles formed from mixing by placing the PDMS mix in a vacuum desiccator for 30 - 60 min. Pour the PDMS mix into a metal mold that is laser-engraved with a ridge measuring 0.1 × 0.5 ×19 mm (Note 9). This ridge imprints a channel to the bottom side of the PDMS block (Note 10). Leave the PDMS to set in the mold at room temperature for at least 2 days, or bake in the oven at 70°C for 2 h. Once set, cut the block from the mold and pierce 0.6-mm diameter holes through the PDMS at each end of the channel. The PDMS channel can then be used to form a flow cell and reused afterward following a wash procedure (Note 11).
2. Heat 1 mL of 1 × blocking buffer and 1 mL of 1 × wash buffer at 42.5°C for 15 min. Degas the warmed buffers for at least 15 min in a vacuum desiccator.
3. Add 100 µL of the blocking buffer to 25 µL of neutravidin solution and immediately spread this mix over the surface of a functionalized coverslip. Incubate at room temperature for 15 min in a humid box to prevent drying.
4. Rinse the coverslip thoroughly with Millipore water to remove unbound streptavidin and dry using compressed air or $N_{2.}$ Immediately place the coverslip, functionalized surface up, in a flow-cell holder, and quickly place the PDMS flow cell on top of the functionalized side of the coverslip. This process needs to be done quickly to minimize PEG exposure to air and avoid degradation. Place the flow-cell holder lid on top of the PDMS-coverslip combination and tighten screws with slight pressure to air seal the flow cell. For single-molecule TIRF experiments, we use the flow-cell setup shown in Figure 5.1A, with PDMS containing 3 flow channels. Insert pre-cut tubing (of at least 15 cm in length) into the holes of the PDMS flow cell to form inlet and outlet tubes.
5. Place the flow cell on the microscope and connect the outlet tube to the syringe pump. Place the inlet tube in the degassed blocking buffer. Pull through approximately 300 µL of blocking buffer (manually, or by programmed syringe pump at a rate of 100 µL/min) and incubate in the flow channel for at least 10 min to reduce nonspecific interactions. Following incubation, wash the blocking buffer from the flow channel

with at least 200 µL of wash buffer at a rate of 100 µL/min. The flow cell is now ready for use in replication stalling experiments (Figure 5.1B).

6. Turn on the lasers, camera, and stage heater. The cooled EMCCD camera needs to reach the operating temperature before use. Additionally, *E. coli* replication experiments are typically performed at 31.2°C, maintained by an electrically heated chamber.

5.3.4. Visualization and Analysis of DNA replication stalling

5.3.4.1. gRNA-647 hybridization

gRNA oligonucleotides can be designed and purchased commercially. Here, crRNA and fluorescently modified tracrRNA oligonucleotides are purchased from IDT and undergo hybridization reactions to form the gRNA molecule required for dCas9 binding to DNA.

1. Centrifuge the tubes before opening to ensure dried oligonucleotides are at the bottom of the tube. Resuspend both crRNA and tracrRNA oligonucleotides to a final concentration of 100 µM in Nuclease-free duplex buffer (IDT), while working in either a laminar flow hood or biological safety cabinet to prevent degradation of the RNA oligos.
2. Combine the tracrRNA and crRNA in an equimolar amount to a final concentration of 15 µM in the nuclease-free duplex buffer.
3. Heat the mixed RNA oligonucleotides at 95°C for 5 min in a benchtop heat block and then remove to slowly cool to room temperature overnight.
4. Store at -20°C until required or long term at -80°C.

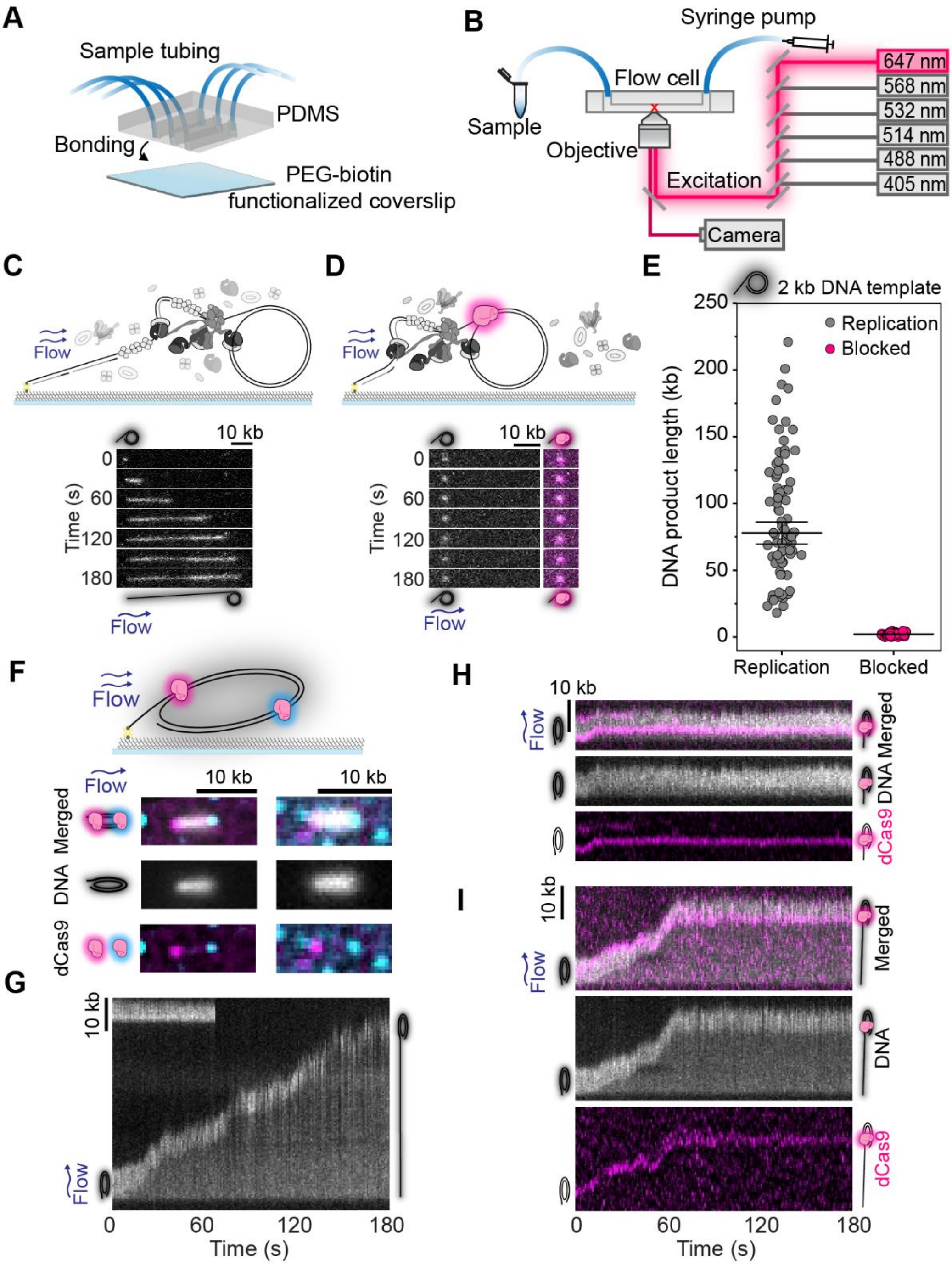


Figure 5.1: Experimental setup and imaging of rolling-circle DNA replication reactions pre-incubated with dCas9 roadblocks. A) Schematic construction of microfluidic device used in rolling-circle DNA replication experiments. **B)** The microfluidic device is placed on the microscope stage with the inlet tubing placed in the tube containing reaction components, and the outlet connected to the syringe pump. Reactions are loaded into the microfluidic device by hydrodynamic flow applied by the syringe pump. Laser light of a specific wavelength is used to illuminate the reaction components through the objective via total internal reflection fluorescence (TIRF). DNA stains and labeled reaction components fluoresce at specific emission wavelengths that are captured by the EMCCD camera and used for downstream analysis. **C)** (top) In the absence of protein roadblocks, the DNA template is immobilized onto the functionalized coverslip by biotin (gray circle) – streptavidin (yellow cross) linkage. Replication is initiated upon the addition of the replisome proteins, ATP, NTP, and dNTPs. In the presence of constant hydrodynamic flow, replicating molecules are

stretched out and visualized by exciting the Sytox Orange DNA stain. (bottom) Example montage of 2-kb rolling-circle DNA replication. **D)** (top) Rolling-circle DNA templates are immobilized onto the functionalized coverslip after incubating with dCas9-gRNA647 complexes. Initiation of the replication reaction results in replication occurring until the roadblock is encountered. (bottom) Example montage of 2-kb blocked rolling-circle DNA templates. The 2-kb DNA templates blocked by dCas9-gRNA647 (crRNA1) complexes result in short products, colocalized with dCas9-gRNA647 complexes (visualized by exciting the labeled gRNA with the 647 nm laser). **E)** Scatter plot of the distribution of DNA product length after 3 min reactions. 2-kb rolling-circle DNA template results in a mean length of 76 ± 8 kb (mean ± standard error of the fit), (*n* = 81 molecules; replication efficiency 4.0 ± 0.3 % (S.E.M.)). In the presence of dCas9-gRNA647 roadblocks, 2-kb rolling-circle DNA template results in a mean length of 2 ± 0.2 kb (*n* = 80 molecules). **F)** (Top) Schematic of 18-kb rolling-circle DNA template stretched out by hydrodynamic flow with bound dCas9-gRNA complexes, targeted to sites 3-kb (red) and 13-kb from the fork (blue). (Bottom, left and right) Averaged 30 s acquisition of 18-kb rolling-circle DNA template with bound dCas9-gRNA complexes shows the distance between pre-incubated dCas9-gRNA complexes can be resolved using the larger DNA template. **G)** Example montage of processive 18-kb rolling-circle DNA replication (*n* = 71 molecules; replication efficiency 6.5 ± 0.6 %). The 18-kb circle is stretched out under constant flow, resulting in a higher intensity observed at the tip of the replicating molecule. **H)** Example 18-kb DNA template blocked by dCas9-gRNA647 complex targeted to a site 3 kb from the replisome assembly site (crRNA2). The processivity of the replisome equals approximately 3 kb until the roadblock was encountered. **I)** Example 18-kb DNA template blocked by dCas9-gRNA647 complex targeted to a site 17 kb from the replisome assembly site (crRNA1). The processivity of the replisome equals approximately 17 kb until the roadblock was encountered. (10-kb scale bar is equivalent to approximately 3 μm under the conditions described).

5.3.4.2. dCas9-gRNA647 pre-incubation

dCas9 and the chosen gRNA oligonucleotide are pre-incubated to form a working dCas9-gRNA stock. Pre-incubation of these components increases the chance of every dCas9 protein containing a gRNA oligonucleotide, thus increasing the chance of dCas9-gRNA complexes binding DNA and blocking replication.

1. Replication buffer should be degassed before preparing the reaction mixture. To form dCas9-gRNA647 complexes, the reaction is prepared as: 10 mM DTT, 112 nM dCas9, and 450 nM gRNA-647.
2. Incubate the reactions at 37°C for 5 min in a benchtop heat block. Remove and immediately place on ice.

5.3.4.3. Reaction scheme 1: Roadblock pre-incubation assay

Pre-incubation of dCas9-gRNA complexes with DNA forms a stable roadblock before the initiation of replication reactions. This method provides a defined site where the roadblock will be encountered by the replisome, dependent on the distance between the replisome assembly site and the gRNA target site. Rolling-circle DNA templates are first incubated with dCas9-gRNA647 complexes, and then further incubated with DnaBC. Following, dCas9-gRNA647-DNA-DnaBC complexes are adsorbed to the flow cell surface until an appropriate surface density of DNA is achieved. Finally, replisome proteins, less DnaBC, are added to the flow cell to initiate replication, during which imaging of both the Sytox Orange (S.O.) stained DNA and 647-labeled dCas9-gRNA complexes allows for visualization of replication stalling (Figure 5.1C-I). The spatial resolution of the 2-kb rolling-circle DNA template allows for the detection of blocked templates by observing the absence or presence of replication products (Figure 5.1C-E). The increased spatial resolution of the 18-kb rolling-circle template supports the detection of the bound fluorescent roadblock and simultaneous movement of the DNA rolling circle as replication proceeds (Figure 5.1F-I).

1. Heat 300 µL of 1 × replication buffer and 200 µL of 1 × replication buffer at 42.5°C for 15 min. Degas the warmed buffers for at least 15 min in a vacuum desiccator.
2. In 300 µL of 1 × replication buffer, prepare solution 1 as 10 mM DTT, and 5 mM ATP.
3. In 20 µL of solution 1, incubate 5 nM of dCas9-gRNA647 (concentration determined from dCas9 concentration as pre-incubated in section 5.3.4.2.) and 40 pM of rolling-circle DNA template at 37°C for 5 min in a benchtop heat block.
4. To the dCas9-gRNA647-DNA mix, incubate 70 nM DnaBC at 37°C for 3 min in a benchtop heat block.
5. Add 180 µL of solution 1 to the dCas9-gRNA647-DNA-DnaBC mix and 50 µM S.O. dye.
6. Flow the DNA reaction mix into the flow cell chamber at a constant rate of 10 µL/min until an appropriate DNA density is achieved. Typically, 20-50 µL total volume should provide 200-500 bound 2-kb rolling-circle templates in one 100 × (80 µm x 80 µm) field of view (Note 12).
7. The *E. coli* replication reaction is performed as: 30 nM αεθ, 46 nM β (dimer), 10 nM $\tau_3\delta\delta'\chi\psi$, 75 nM DnaG, 20 nM SSB in 200 µL 1 × replication buffer containing 125 µM each of dATP, dCTP, dGTP, dTTP, 250 µM each of ATP, CTP, GTP, UTP, 10 mM DTT and 150 nM S.O. dye.
8. Flow the replication reaction mixture through the inlet tubing to the edge of the flow cell at a constant rate of 70 µL/min for 1 min. Immediately decrease the flow rate to 10 µL/min, sufficient enough to stretch dsDNA, and begin imaging the DNA, with a 532-nm laser and the dCas9-gRNA647 complexes with a 647-nm laser (Note 13).
9. Typical acquisition parameters are 1-5 frames/s for 1-5 minutes, an exposure time of 200 ms, and a laser intensity of 80-800 mW/cm^2 (Notes 14 - 16).

5.3.4.4. *Reaction scheme 2: roadblock in-solution assay*

Binding of dCas9-gRNA complexes during the replication reaction results in real-time blocking of the replisome. The output of this strategy provides various lengths of products, dependent on the time point that the dCas9-gRNA complex has formed a stable roadblock (Figure 5.2). Here, the rolling-circle DNA template is pre-incubated with DnaBC and then adsorbed to the coverslip of the flow cell. The replication reaction is then initiated upon the introduction of buffer containing the replisome components (less DnaBC) and dCas9-gRNA complexes. Reactions are imaged similarly to reaction scheme 1, where unblocked DNA products are monitored by exciting the S.O. DNA stain, and binding of dCas9-gRNA complexes is detected by imaging the 647-labeled gRNA. In addition to detecting the dCas9-gRNA complex that stalls the replisome, binding of the complex behind the fork occurs as the binding site is duplicated with each round of replication (Figure 5.2B, D, E).

1. Prepare the buffers and DNA-DnaBC mix as described in section 5.3.4.3 steps 1, 2, 4 and 5, less dCas9-gRNA647.
2. Flow DNA-DnaBC mix into the flow cell chamber at a constant rate of 10 µL/min until an appropriate DNA density is achieved. Typically, 20-50 µL total volume should provide 200-500 bound 2-kb rolling-circle templates in one 100 × (80 µm x 80 µm) field of view (Note 12)
3. The *E. coli* replication reaction is performed as described for reaction scheme 1 with the addition of dCas9-gRNA647 complexes. The concentration of dCas9-gRNA647 complexes determines the number of replicating products visible and the average product length at the end of the acquisition (Figure 5.2C). Typically concentrations below 5 nM (dCas9) will yield visibly replicating products for the 2-kb rolling-circle template.
4. Initiate the replication reaction by flowing the reaction mixture as described in section 5.3.4.3 step 8.
5. Typical acquisition parameters are: 1-5 frames/s for 1-5 minutes an exposure time of 200 ms and a laser intensity of 80-800 mW/cm^2 (Notes 14 and 16).

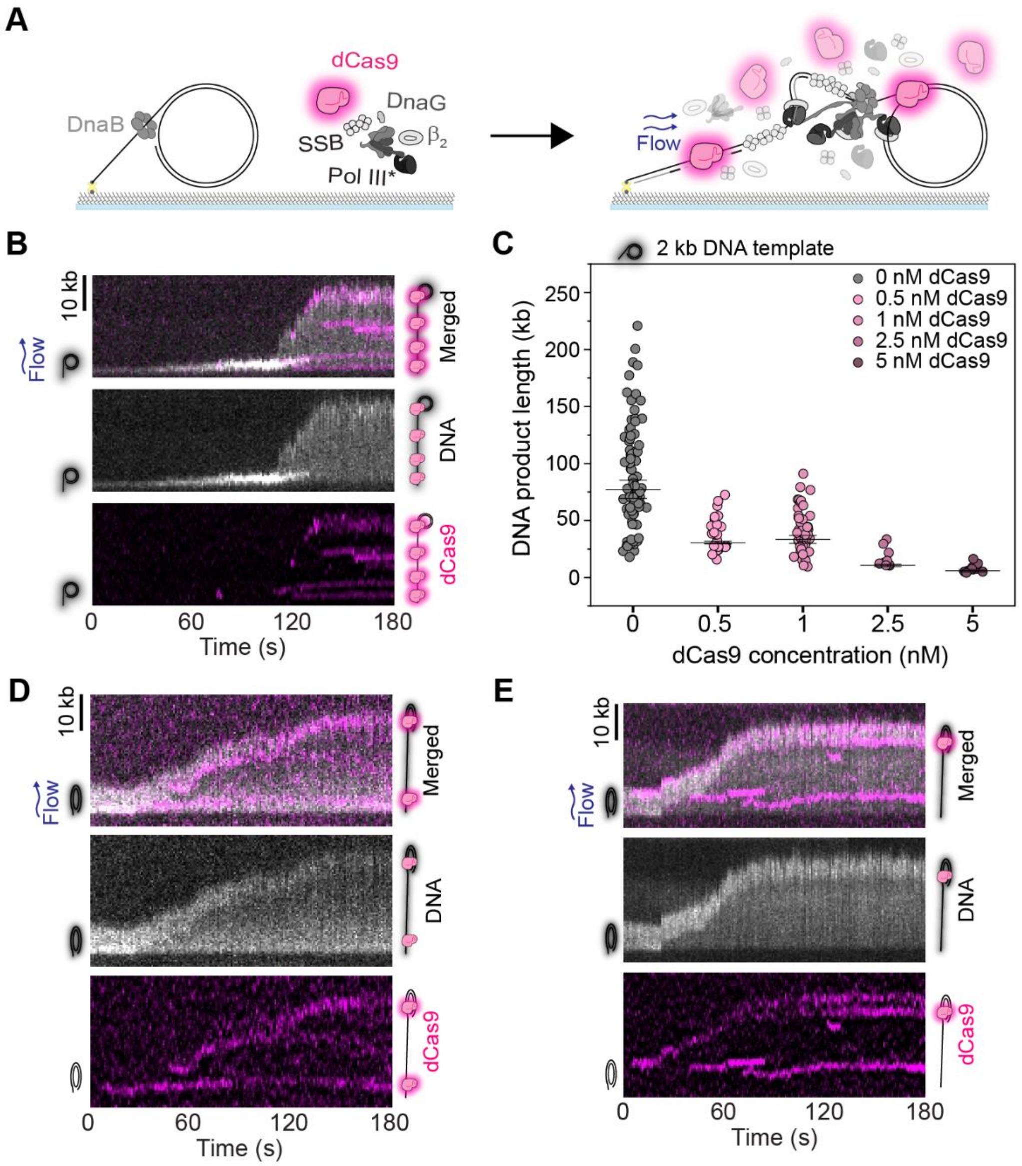


Figure 5.2: Imaging rolling-circle DNA replication in the presence of roadblocks in solution. A) Schematic representation of rolling-circle DNA template immobilized on the functionalized coverslip. Replication is initiated by the addition of the replisome proteins in the presence of dCas9-gRNA647 complexes. This reaction results in processive replication up until the dCas9-gRNA647 complex binds to the replicating template and is encountered by the replisome. Additionally, dCas9-gRNA647 complexes will also bind to target sites behind the replication fork as the number of sites increases with each round of replication. **B)** Example kymograph of a 2-kb rolling-circle template undergoing replication in the presence of dCas9-gRNA647 (crRNA1) complexes in solution. Sytox-stained DNA merged with 647-excited dCas9-gRNA647 complexes (top). Sytox-stained DNA showing processive replication begins at approximately 120 s time point and is blocked shortly after (middle). dCas9-gRNA647 complexes bind behind the replication fork and colocalize with the time point at which replication is blocked (bottom). **C)** Scatter plots of the product length distribution after 3 min reactions at titrated dCas9-gRNA647 concentrations in solution. 0 nM (76 ± 8 kb (mean ± S.E.M.)) condition is reproduced from Figure 5.1E. Mean product lengths in the presence of dCas9-gRNA647 complexes: 0.5 nM; 30 ± 1 kb (*n* = 22 molecules; replication efficiency 3.1 ± 0.4 %), 1 nM; 34 ± 4 kb (*n* = 49 molecules; replication efficiency 3.2 ± 0.3 %), 2.5 nM; 11 ± 1 kb (*n* = 9 molecules; replication

efficiency 1.0 ± 0.1 %), 5 nM; 7 ± 1 kb (*n* =11 molecules; replication efficiency 0.5 ± 0.1 %). **D)** Example kymograph of an 18-kb rolling-circle template undergoing replication in the presence of dCas9-gRNA647 complexes in solution, targeted to a site 3 kb from the replisome assembly site (crRNA1). Merged (top), Sytox stained DNA (middle) and dCas9-gRNA647 complexes (bottom). **E)** Example kymograph of an 18-kb rolling-circle template undergoing replication in the presence of dCas9-gRNA647 complexes in solution, targeted to a site 17 kb from the replisome assembly site (crRNA2). Merged (top), Sytox stained DNA (middle) and dCas9-gRNA647 complexes (bottom). (10-kb scale bar is equivalent to approximately 3 μm under the conditions described).

5.3.4.5. Data quantification

The acquisitions obtained from both reaction schemes described were analyzed for the DNA product length at the end of the acquisition time (3 min) as previously described [74] (Figure 5.1E and Figure 5.2C). Briefly, image analysis was carried out in FIJI, using the Single Molecule Biophysics plug-ins (see section 5.2.4). Here, raw acquisitions (nd2 format) were converted into TIF files and flattened with the excitation beam profile. Quantification of DNA product length was determined by deconvolving the length of the rolling-circle substrate using the calibrated pixel size in bp (here, 1 pixel = 470 bp). Product length distributions were fit with a single Gaussian function. All distributions were made and fitted using MATLAB and Origin Pro.

5.4 Notes

1. The ratio of the PEG mixture can be optimized to achieve the desired density of DNA molecules bound in the flow cell.
2. Drying can also be prevented by incubating coverslip sandwiches in a humid box, created simply by using an empty microtube rack with water on the bottom with a tight-fitting lid.
3. Long-term storage sees degradation in surface passivation. We recommend no more than 2 weeks, especially in humid climates.
4. Typically 60% of nicked 2-kb plasmid is recovered using this method of PEG purification.
5. The lengths of the fork oligonucleotides, and the gap they form once hybridized to the circle, are customizable and controllable. The size of the gap will affect the loading of the helicase and subsequent recruitment of polymerases.
6. We recommend long-term storage of purified rolling-circle DNA templates in small volumes of working-concentration stocks. This will prevent the degradation of the DNA template through repetitive freeze-thaw cycles.
7. The large size of the pUBER plasmid must be taken into consideration when pipetting. To avoid damage to the plasmid throughout construction and use, pipetting to mix solutions must be avoided where possible, and flicking the tube used as an alternative.
8. The 18-kb rolling-circle DNA template can also be purified by the following methods: Promega Wizard DNA Clean-up system (14% recovery), and PEG purification (13% recovery).

9. The PDMS and curing agent need to be mixed vigorously (stir for >3 min) to ensure uniform curing of the PDMS in the mold.
10. Our PDMS block mold contains three channels imprinted side by side (3 mm apart). This allows for multiple single-molecule experiments to be carried out in quick succession, preventing the need to dismantle and rebuild new flow cells.
11. Once the single-molecule experiment is complete, the PDMS blocks can be washed for future reuse. This is done by sonicating for 20 min in 0.5% Triton and then extensively rinsing in Millipore water. The PDMS blocks are then sonicated for a further 20 min in 1 M $NaHCO_3$, then extensively washed with Millipore. Finally, the PDMS blocks are sonicated for 20 min in 70% ethanol, rinsed with Millipore, and dried with compressed air. Clean PDMS blocks can be stored at room temp in a dust-proof container.
12. For the 18-kb rolling-circle template, 50 µL total volume at 10 µL/min typically provides approximately 200 surface-bound rolling-circle templates in one field of view (80 µm x 80 µm at 100 × magnification).
13. After opening the camera and before image capturing, move the stage to a nearby FOV. This will ensure the molecules and fluorophores being imaged have not already undergone photocleavage or begun to photobleach.
14. Sytox Orange and similar DNA stains can cause photocleavage of dsDNA, therefore we suggest performing experiments at low laser powers, typically 80-800 mW/cm^2. This can be adjusted based on DNA stain concentration and acquisition rate.
15. Experimental examples of 2-kb rolling-circle templates provided (Figure 5.1C & D) were obtained by imaging 5 fields of view consecutively after another to increase the experimental throughput. The disadvantage of this method is decreased temporal resolution of each replicating product. If the temporal resolution is of concern, we recommend capturing one field of view continuously until the end of the experiment (as done for reaction scheme 2 examples).
16. Experimental examples provided were conducted by rapidly switching between laser lines and filter sets. This method of imaging results in two-color image sets that are not perfectly correlated in time, however, provide a larger detectable field of view. Alternatively, if the study design requires temporal synchronization between the two channels, a dual-view device can be used to direct light from each emission channel onto separate halves of the camera chip. The disadvantage of this method is that the detectable field of view is effectively halved.

Chapter 6

Single-molecule visualization of stalled replication-fork rescue by the *Escherichia coli* Rep helicase

Kelsey S. Whinn, Zhi-Qiang Xu, Slobodan Jergic, Nischal Sharma, Lisanne M. Spenkelink, Nicholas E. Dixon, Antoine M. van Oijen and Harshad Ghodke

Manuscript version currently under revision at Nucleic Acids Research

Genome duplication occurs while the template DNA is bound by numerous DNA-binding proteins. Each of these proteins act as potential roadblocks to the replication fork and can have deleterious effects on cells. In *Escherichia coli*, these roadblocks are displaced by the accessory helicase Rep, a DNA translocase and helicase that interacts with the replisome. The mechanistic details underlying the coordination with replication and roadblock removal by Rep remain poorly understood. Through real-time fluorescence imaging of the DNA produced by individual *E. coli* replisomes and the simultaneous visualization of fluorescently-labeled Rep, we show that Rep continually surveils elongating replisomes. We found that this association of Rep with the replisome is stochastic and occurs independently of whether the fork is stalled or not. Further, we visualize the efficient rescue of stalled replication forks by directly imaging individual Rep molecules as they remove a model protein roadblock, dCas9, from the template DNA. Using roadblocks of varying DNA-binding stabilities, we conclude that the continuation of synthesis is the rate-limiting step of stalled replication rescue.

Contributions

This chapter brings together the single-molecule methods developed in Chapters 4 and 5 to visualize and characterize the activity of the Rep helicase at the *E. coli* replisome. The results described in this chapter have implications for how Rep interacts with stalled and elongating replisomes, the fate of the replisome at protein roadblocks and the continuation of DNA synthesis after removal of roadblocks by Rep.

I carried out the expression, purification and labeling of the Rep-AF647 protein, all ensemble replication assays, and all single-molecule assays using the Rep proteins and dCas9-gRNA complexes. I also analyzed all the data, compiled results, created and designed figures and drafted the manuscript.

6.1. Introduction

Cell proliferation requires high-fidelity duplication of the genome. In all cells, this process is achieved by a group of proteins collectively known as the replisome. In *Escherichia coli*, more than 12 proteins coordinate the unwinding and duplication of DNA at rates of up to 1000 base pairs (bp) per second [6, 20-23] and at processivities of 100s of kilobase pairs (kbp) [49, 371, 372]. DNA replication occurs on template DNA bound by numerous other nucleoprotein complexes involved in other important cellular functions such as transcription, DNA repair and recombination. However, the presence of these proteins on the template DNA creates potential roadblocks to the replisome, which can impede replication and result in fork collapse and genome instability [58, 191, 373]. Replication across roadblocks requires the action of accessory proteins or activation of DNA repair pathways and replication restart mechanisms [374].

Both prokaryotes and eukaryotes express accessory replicative helicases that remove protein roadblocks from the path of the replication fork. These enzymes often act on the strand opposite to that encircled by the ring-shaped replicative helicase to rescue stalled replication forks and prevent the collapse of the progressing replisome [374, 375]. In eukaryotes, the Pif1 helicase is important for replisome bypass of R-loops and protein roadblocks [156, 376, 377]. In *E. coli*, the Rep and UvrD helicases promote the removal of a range of roadblocks, including RNA polymerase (RNAP) [59-61, 101, 157, 158, 160]. This underlying activity is believed to be dependent on coordination with either a stalled or progressing replisome [59, 155, 156], but the precise mechanisms are not completely understood.

The *E. coli* Rep helicase is a superfamily 1A (SF1A) helicase that translocates on single-stranded DNA (ssDNA) in a 3'–5' direction [115]. Rep like its structural homologs *E. coli* UvrD and *Bacillus stearothermophilus* PcrA (reviewed in [374]) comprises four domains (1A, 2A, 1B and 2B). While the motor cores, termed 1A and 2A, consist of two highly conserved RecA-like subdomains, it is the 2B subdomain that exhibits extreme conformational changes between a closed and open state that are tightly linked to the activation of helicase activity [115, 374]. Although Rep, UvrD, and PcrA monomers can bind and translocate on ssDNA, they do not exhibit processive double-stranded DNA (dsDNA) unwinding activity [134, 144, 147]. However, intramolecular coupling of the 2B subdomain in the extreme closed state, or deletion of the subdomain altogether, has revealed activation of monomeric helicase activity [138, 142, 374]. Additionally, this subdomain has been proposed to play an important role in separating the primary functions of Rep; protein-roadblock displacement and unwinding of dsDNA [374]. The displacement of protein roadblocks has proven essential to cell viability. Notably, in *E. coli* UvrD can also rescue stalled replication forks, where single *rep* and *uvrD* mutations are viable but double *rep*, *uvrD* mutations are lethal [59].

Rep can physically interact with the replication fork through both protein-DNA and protein-protein interactions. The opposite translocation direction to that of the lagging-strand DnaB replicative helicase (5' to 3') places Rep on the leading-strand DNA template [19]. Physical interactions between the C-terminus of Rep and DnaB are important in promoting the efficient removal of protein roadblocks. Recent *in vivo* studies showed a maximal occupancy of the DnaB hexamer, revealing six Rep monomers associated with the replisome [155]. Rep proteins with a mutated C-terminus showed a decrease in this occupancy. The authors proposed that the C-terminus plays an important role in recruiting Rep monomers to the replisome by the interaction with DnaB [155]. Further, this interaction allows the Rep monomers to be loaded onto the leading strand to translocate ahead of the replication fork [155]. This was further supported by the study of an ATPase-deficient mutant of Rep that could colocalize with the replisome but showed no evidence of translocation away from the replication fork. Additionally, the 2B subdomain of Rep is crucial for the displacement of protein roadblocks [68, 130]. Positioned at the leading edge of the helicase, conformational changes of this subdomain likely cause the translocation activity of Rep to switch to protein displacement upon contacting potential roadblocks [130, 141]. Despite the extensive structural and functional knowledge of Rep and homologs, the kinetic mechanisms underlying the association with the replisome leading to the displacement of roadblocks and stalled replication rescue remain poorly understood.

Investigations of replication fork stalling have resulted in the development of many tools to mimic protein roadblocks. We have previously developed a highly stable, site-specific roadblock using the nuclease dead Cas9 (dCas9) protein [74]. This protein roadblock containing an RNA:DNA hybrid can efficiently block bacterial, viral, and yeast replisomes for long periods (> 20 min). Using this tool, in combination with *in vitro* single-molecule fluorescence imaging, we investigated replisome roadblock bypass with a high degree of spatial and temporal resolution. These single-molecule techniques reveal the heterogeneity of complex biological processes and provide insight into the individual behaviors of single molecules that are not detected by ensemble averaging methods.

To investigate the *E. coli* Rep helicase in the context of elongating and stalled replication forks, we use single-molecule assays to directly visualize the individual Rep proteins at the replisome. Using an *in vitro* reconstituted system and the dCas9 roadblock, we monitor Rep behaviors in real time by imaging fluorescently labeled Rep proteins. We set out to test the hypothesis that Rep is recruited to stalled replication forks, which has been raised by several studies [59, 153, 155], or whether it might associate with the replisome continually regardless of it being stalled or not. We find that during replication elongation, Rep frequently associates with the replisome in a predominantly monomeric state. Further, in the presence of dCas9 roadblocks, we see efficient removal of the roadblock and rescue of replication. By use of less stable roadblocks, we show that the time elapsed between stalling and rescue of replication is not dependent on the stability of the roadblock, but rather represents a process that occurs after the roadblock has been removed.

6.2. Material and Methods

6.2.1. Proteins

Escherichia coli DNA replication proteins were produced as described previously: the β_2 sliding clamp [378], SSB [379], the $DnaB_6(DnaC)_6$ helicase-loader complex (referred to as DnaBC) [78], DnaG primase [380], the Pol III $\tau_3\delta\delta'\psi\chi$ clamp loader complex [381], and Pol III $\alpha\varepsilon\theta$ core [34, 381]. dCas9-dL5, referred to as dCas9, was produced as described previously [74].

6.2.2. Overproduction and purification of His$_6$-Rep WT, His$_6$-Rep K28A, and His$_6$-Rep ΔC33

6.2.2.1. Construction of plasmids

To construct plasmid pZX2198 that encodes His$_6$-Rep under the control of the T7 promoter, the *rep* gene was amplified by PCR using plasmid pCL771, a derivative of pET3c harboring a *rep* gene, as a template and primers PET3 (5'-CGACTCACTATAGGGAGACCACAAC) and 740 (5'-AAAGAGCTCTTATTTCCCTCGTTTTGCCG) that carry a *Sac*I site immediately downstream of the *rep* gene. The PCR product was cleaned using a Qiagen PCR purification kit, digested with *Nde*I and *Sac*I, and ligated into plasmid pETMCSIII [382] that was pre-digested with *Nde*I and *Sac*I and gel purified. The ligation mixture was transformed into *E. coli* strain AN1459 [382]. Colony PCR was then performed to identify colonies harboring the desired plasmid. Plasmid DNAs were extracted and sequences of *rep* were verified by DNA sequencing.

To construct plasmid pZX2199 that encodes His$_6$-Rep ΔC33, a 3'-fragment of the *rep* gene missing the C-terminal 33 residues that are known to interact with the DnaB helicase was amplified by PCR using plasmid pCL771 as template and primers 738 (5'-TACTGGCGAGCTGATCG) and 739 (5'- AAAGAGCTCTTACCAAATCAGATCATCCTG). The PCR product was cleaned and digested with *Mlu*I and *Sca*I and ligated into pZX2198 pre-digested with the same enzymes. The plasmid was then selected and verified as above.

Plasmid pZX2200 encoding His$_6$-Rep K28A was constructed by site-directed mutagenesis using the QuikChange protocol with pZX2198 as template and primers 734 (5'-CGCGGGTTCCGGTGCAACTCGTGTTATCACC) and 735 (5'-GATAACACGAGTTGCACCGGAACCCGCGCC).

6.2.2.2. Protein expression and purification

Proteins were over-expressed using *E. coli* strain BL21(λDE3)/pLysS harboring the desired plasmids. Briefly, a bacterial colony was inoculated into LB medium (10 g tryptone, 5 g yeast extract and 10 g NaCl per liter) supplemented with 100 μg mL^{-1} of ampicillin and grown at 37°C overnight, then 5 mL of overnight culture was inoculated into 1 L of LB supplemented with 100 μg mL^{-1} of ampicillin (2 L in total). Bacteria were grown at 37°C until OD$_{600}$ of the culture was approximately 0.6. Protein expression was induced by adding IPTG to a final concentration of 0.4

mM. The cultures were then grown for 3 h at 25°C and the cells were collected by centrifugation. Cell pellets were weighed, snap-frozen in liquid nitrogen, and stored at –80°C until use.

Bacterial cells were resuspended in 60 mL of Lysis buffer containing 50 mM Tris-HCl, pH 7.6, 10% sucrose, 300 mM NaCl, 2 mM dithiothreitol, 1 mM EDTA. A tablet of EDTA-free protease inhibitor cocktail (Roche) was also added. The cells were lysed by two passages through a French Press operated at 12,000 psi. Cell debris was removed by centrifugation (40,000 × *g*, 20 min). Ammonium sulfate (0.31 g mL^{-1}) was then added to the cleared lysate and stirred for 1 h. Precipitated proteins were collected by centrifugation (40,000 × *g*, 50 min) and dissolved in 50 mL of Buffer A (50 mM Tris-HCl, pH 7.6, 10% glycerol, 700 mM NaCl, 0.5 mM dithiothreitol, 15 mM imidazole). The protein was loaded onto a 5 mL HisTrap HP column (Cytiva) equilibrated with Buffer A. The loaded column was washed with 100 mL of Buffer A and proteins were eluted with a gradient of 15–600 mM imidazole over 25 mL. Fractions containing His_6-Rep proteins were pooled and diluted 2.5-fold with Buffer A. The solution was loaded onto a 5 mL HiTrap Heparin column (Cytiva) equilibrated with Buffer B (50 mM Tris-HCl, pH 7.6, 30% glycerol, 1 mM EDTA, 1 mM dithiothreitol, 0.2 M NaCl). The loaded column was washed with 40 mL of Buffer B and proteins were eluted with a gradient of 0.2–1.0 M NaCl over 120 mL. The fractions containing pure His_6-Rep proteins were combined and dialyzed against 1 L of a storage buffer containing 50 mM Tris-HCl, pH 7.6, 30% (*v*/*v*) glycerol, 450 mM NaCl, 2 mM dithiothreitol and 1 mM EDTA, and stored at –80°C. The purity of samples was confirmed by SDS-PAGE (Supplementary Figure 6.S1A–C).

6.2.3. Expression, purification and labeling of His_6-Rep A97C

His_6-Rep A97C was expressed and purified as previously described [383], with modifications. His_6-Rep A97C was overproduced using *E. coli* BL21(λDE3) harboring the pRepA97C plasmid (43). Briefly, a bacterial colony was inoculated into 30 mL of LB broth supplemented with 30 μg mL^{-1} of kanamycin in a 100 mL flask and grown at 37°C overnight. 10 mL of overnight culture was inoculated into 1 L of LB supplemented with 30 μg mL^{-1} of kanamycin (2 L in total). Bacteria were grown at 30°C until OD_{600} ~ 0.8, and protein expression was induced by the addition of 300 μM IPTG. After growth for 2 h at 30°C, cells were collected by centrifugation, weighed, snap-frozen in liquid nitrogen and stored at –80°C.

Thawed cells were resuspended as described above, with the following modifications. Immediately before lysis, 0.5 mM phenylmethylsulfonyl fluoride (PMSF) was added. Following lysis by French Press, cell debris was removed by centrifugation (40,000 × *g*, 30 min). Dissolved, precipitated proteins were purified as described above, where Heparin affinity purification was carried out with Buffer B containing 10% glycerol. Fractions containing pure His_6-Rep A97C proteins were combined and dialyzed against 2 L of a storage buffer containing 50 mM Tris-HCl, pH 7.6, 30% (*v*/*v*) glycerol, 300 mM NaCl, 5 mM dithiothreitol, and 1 mM EDTA; and stored at –80°C. The purity of the sample was confirmed by SDS-PAGE (Supplementary Figure 6.S1D).

Purified His_6-Rep A97C was fluorescently labeled with Alexa Fluor 647 (Invitrogen) using methods adapted from [75, 384]. First, a total of 1 mg of His$_6$-Rep A97C was reduced with 5 mM tris(2-carboxyethyl)phosphine (pH 7.6) in storage buffer containing 50% (*v*/*v*) saturated ammonium sulfate solution, at 6°C for 1 h with gentle rotation to yield Fraction I. Fraction I was centrifuged (21,000 ×*g*, 15 min) at 6°C and the supernatant was carefully removed. The precipitate was washed with ice-cold labeling buffer (50 mM Tris-HCl, pH 7.6, 30% (*v/v*) glycerol, 300 mM NaCl, 1 mM EDTA) and 50% (*v/v*) saturated ammonium sulfate solution to yield Fraction II (both solutions had been extensively degassed by sonication and deoxygenated using Ar gas). Fraction II was incubated at 6°C for 1 h with gentle rotation, then centrifuged (21 000 × *g*, 15 min) at 6°C and the supernatant was removed to yield Fraction III. The labeling reaction was carried out on Fraction III, now devoid of reducing agent, using a 5-fold molar excess of maleimide-conjugated AF647 with 33 μM His$_6$-Rep A97C in 300 μL of deoxygenated and degassed labeling buffer. The reaction was allowed to proceed at 6°C overnight with gentle rotation (in the dark) and quenched with 30 mM dithiothreitol for 1 h at 6°C to yield Fraction IV. Fraction IV was split into three equal volumes and applied to Zeba Spin desalting columns (40K MWCO) (ThermoFisher) equilibrated with storage buffer. Free dye was separated from labeled protein by centrifugation (1500 × *g*, 2 min) at 6°C. The flow-through from each column was then applied to a second desalting column and centrifuged. The flow-through from each column was combined and stored at –80°C. The degree of labeling was determined by UV/Vis spectroscopy to be 1.0 fluorophore per Rep monomer. Alexa Fluor 647-labeled His$_6$-Rep A97C is henceforth referred to as Rep-AF647.

6.2.4. Surface plasmon resonance

SPR experiments were carried out on a BIAcore T200 instrument (Cytiva) using a streptavidin (SA) coated sensor chip to study the binding and dissociation of His$_6$-Rep WT, His$_6$-Rep K28A, and His$_6$-Rep ΔC33 to ssDNA substrates. All experiments were carried out at 20°C at a flow rate of 10 (chip preparation) or 20 μL min^{-1} (Rep protein binding). The SA chip was activated with four sequential 40 s injections of 1 M NaCl, and 50 mM NaOH, then stabilized by 1 min treatment with 1 M $MgCl_2$.

The 3'-biotinylated-dT_{35}, -dT_{15} or –dT_{10} substrates were dissolved in 1 × SPR buffer (25 mM Tris-HCl, pH 7.6, 50 mM NaCl, 5 mM $MgCl_2$, 0.25 mM dithiothreitol, 0.2 mM EDTA and 0.005% (*v/v*) surfactant P20) to a final concentration of 10 nM and introduced onto the SPR chip for immobilization, followed by three sequential wash steps with 1 M $MgCl_2$. The signal from the ssDNA substrates corresponded to 250 response units (RU) (dT_{35}), 170 RU (dT_{15}) and 125 RU (dT_{10}).

Following the immobilization of the ssDNA substrates, binding studies were done by injecting specified concentrations of Rep proteins in the SPR buffer. For measurements of His$_6$-Rep (WT, K28A, and ΔC33) binding to dT_{35} and dissociation in the presence of 200 μM of AMP-PNP, ADP, or ATP, 20 nM of protein was injected for 60 s, followed by injection of the specified nucleotide cofactor for 60 s. For measurements of His$_6$-Rep WT binding to dT_{35}, dT_{15}, and dT_{10}, 20 nM of His$_6$-Rep WT was injected for 60 s, followed by injection of SPR buffer lacking any Rep or nucleotide cofactor. For measurements of His$_6$-Rep WT binding to dT$_{15}$, an optimized concentration range of [0, 1, 2, 4, 8] nM was injected at 60 μL min^{-1} for 60 s.

Kinetics of the interaction of His$_6$-Rep WT with the DNA substrates were evaluated using BioEvaluation 2.0 software (Cytiva). Stoichiometries (n) of the interaction of Rep monomers with DNA substrates were approximated using Equation 1, where R_{Rep} and R_{DNA} are responses (in RU) due to Rep and the ssDNA substrates, respectively.

$$n = \frac{R_{Rep}}{R_{DNA}} \times \frac{M_{r(DNA)}}{M_{r(Rep)}} \tag{1}$$

6.2.5. Rolling-circle replication templates

The 2-kbp DNA rolling-circle substrates were prepared as previously described [356].

The 18-kbp DNA rolling-circle substrates were prepared from plasmid pUBER using methods and oligonucleotides described by Mueller et al. [366]. Briefly, 50 µg of supercoiled pUBER plasmid was treated with 1 unit μg^{-1} Nt.*Bbv*CI in 1 × Cutsmart buffer (New England Biolabs, USA) at 37°C for 4 h. A 10-fold molar excess of Dig-competitor oligonucleotides was added to the reaction and the temperature was raised to 65°C for 20 min followed by cooling to 14°C at a rate of 1°C min^{-1}. The displaced oligonucleotides were purified from the gapped plasmid by magnetic separation using 1 μg nmol^{-1} tosyl activated paramagnetic beads functionalized with anti-digoxigenin Fab fragments, equilibrated in 1 × Cutsmart buffer. The nicking reaction mixture was incubated with functionalized beads for 30 min at room temperature with gentle rotation. The fork oligonucleotide was annealed to the gapped plasmid in the presence of 100-fold molar excess over the DNA substrate in 1 × Cutsmart buffer at 50°C for 10 min followed by cooling to 16°C at 1°C min^{-1}. The fork oligonucleotide was ligated to the DNA substrate by adding 62.5 units μg^{-1} T4 DNA ligase, supplementing with 12 mM ATP and 10 mM dithiothreitol, and incubating at 16°C for 18 h followed by inactivation of the ligase at 65°C for 10 min. 18-kbp rolling-circle DNA templates were purified from excess fork oligonucleotides by gel filtration on a Sepharose 4B (1 × 25 cm; Sigma-Aldrich) column in the presence of 12 mM EDTA and 300 mM NaCl. Rolling-circle DNA templates were eluted with TE buffer containing 10 mM Tris-HCl, pH 7.6, 1 mM EDTA and 300 mM NaCl. Fractions containing the 18-kbp rolling-circle DNA template were stored at –80°C.

6.2.6. In vitro ensemble replication rescue assays

Standard leading-strand replication assays were set up as described previously [74], with the following modifications. Replication rescue assays were set up in replication buffer (RB; 30 mM Tris-HCl pH 7.5, 12 mM magnesium acetate, 50 mM potassium glutamate, 0.5 mM EDTA and 0.0025% (*v/v*) Tween20) or RB containing high magnesium (RBM; containing 24 mM magnesium acetate). Reactions contained 2 nM 2-kbp rolling-circle replication template, specified concentrations of dCas9 and gRNA, 60 nM DnaBC, 30 nM $\tau_3\delta\delta'\psi\chi$, 90 nM Pol III $\alpha\varepsilon\theta$ core, 200 nM β_2, 10 mM dithiothreitol, 1 mM ATP (in RB) or 10 mM ATP (in RBM), and 125 μM dNTPs (each) in a final volume of 12 μL. First, dCas9 was incubated with the specified cgRNA for 5 min at room temperature and further incubated with rolling-circle DNA templates for 5 min at room temperature. Other components were mixed and incubated at room temperature, then cooled on ice for 5 min before the addition of dCas9-cgRNA-DNA complexes. Reactions were initiated at 30°C. At specified time points, stated concentrations of His_6-Rep WT, His_6-Rep K28A, His_6-Rep ΔC33, or Rep-AF647 were added to the reactions in the absence or presence of 50 nM trap dsDNA, an 83-mer target DNA containing one complementary sequence to the fully complementary gRNA (cgRNA1) (Supplementary Table S1) [74]. The reactions were quenched at specified time points by the addition of 12 μL of LES (2 × DNA gel loading dye, 200 mM EDTA and 2% (*w*/*v*) SDS). The quenched reactions were loaded into a 0.6% (*w/v*) agarose gel in 2 × TAE. Products were separated by agarose gel electrophoresis at 60 V for 150 min, stained in SYBR-Gold (Invitrogen), and imaged under long-wave UV light.

E. coli leading- and lagging-strand DNA replication reactions were set up as previously described [34, 74], with minor modifications. Reactions were set up in RB and contained 4 nM 2-kbp rolling-circle DNA template, specified concentrations of dCas9 and cgRNA, 60 nM DnaBC, 80 nM DnaG, 30 nM $\tau_3\delta\delta'\psi\chi$, 10 nM SSB, 90 nM Pol III $\alpha\varepsilon\theta$ core, 200 nM β_2, 10 mM dithiothreitol, 1 mM ATP, 125 μM dNTPs and 250 μM NTPs to a final volume of 12 μL. At 10 min after initiation of the replication reaction, a specified concentration of His_6-Rep WT was added to each reaction in the presence of 50 nM trap dsDNA. At 20 min, reactions were quenched by the addition of 1.5 μL 0.5 M EDTA and 3 μL DNA loading dye (6 mM EDTA, 300 mM NaOH, 0.25% (*w*/*v*) bromocresol green, 0.25% (*w*/*v*) xylene cyanol FF, 30% (*v*/*v*) glycerol). DNA products were separated on a 0.6% (*w*/*v*) alkaline agarose gel at 14 V for 14 h. The gel was neutralized in TAE buffer, stained with SYBR-Gold and imaged under UV light.

6.2.7. In vitro single-molecule fluorescence microscopy

In vitro single-molecule fluorescence microscopy was carried out on an Eclipse Ti-E inverted microscope (Nikon, Japan) with a CFI Apo TIRF 100× oil-immersion TIRF objective (NA 1.49, Nikon, Japan) as described previously [34, 74, 75, 78]. The temperature was maintained at 31.2°C by an electronically heated flow-cell chamber coupled to an objective heating jacket (Okolab, USA). NIS-elements software was used to operate the microscope and the focus was locked into place through the Perfect Focus System (Nikon, Japan). Images were captured using a 512 × 512-pixel EM-CCD camera (either Photometrics Evolve 512 Delta or Andor iXon 897). DNA molecules stained with 150 nM Sytox orange were imaged with either a 568-nm laser (Coherent, Sapphire 568-200 CW) at 400 mW cm^{-2}, a 514-nm laser (Coherent, Sapphire 514-150 CW) at 200 mW cm^{-2}, or 532-nm laser (Coherent, Sapphire 532-300 CW) at 90 mW cm^{-2}. dCas9 complexed to cgRNA-Atto647 was imaged with a 647-nm laser (Coherent, Obis 647-100 CW) at 220 mW cm^{-2}. Rep-AF647 was imaged with the 647-nm laser at 200 mW cm^{-2}.

6.2.7.1. Preparation of flow cells for in vitro imaging

Replication reactions were carried out in microfluidic flow cells constructed from a PDMS flow chamber placed on top of a PEG-biotin-functionalized glass microscope coverslip as described previously [34, 48, 358, 381]. Once assembled, all surfaces of the flow cell including tubing were blocked against non-specific binding by the introduction of at least 300 μL blocking buffer (50 mM Tris-HCl pH 7.9, 50 mM potassium chloride, 2% (*v/v*) Tween20).

6.2.7.2. Single-molecule Rep binding assay

The single-molecule Rep binding assay was designed to investigate the association of Rep to forked DNA templates bound by DnaBC and SSB, in the absence and presence of ATP. First, 8 pM 2-kbp rolling-circle DNA template was incubated with 7 nM DnaBC in degassed single-molecule replication buffer (SM; 25 mM Tris-HCl pH 7.9, 10 mM magnesium acetate, 50 mM potassium glutamate, 0.1 mM EDTA, 0.0025% Tween20, 0.5 mg mL^{-1} BSA, 1 mM ATP, 10 mM dithiothreitol and 150 nM Sytox orange) for 3 min at 37°C. Following, 20 nM SSB was added to the DNA-DnaBC mixture. The DNA-DnaBC-SSB was adsorbed to the flow-cell surface at 10 μL min^{-1} until an appropriate surface density was achieved. The flow cell was subsequently washed with SM buffer containing Sytox orange at 70 μL min^{-1} for 2 min. Next, 10 nM Rep-AF647 in SM buffer supplemented with 10 nM SSB, 10 mM dithiothreitol, and 150 nM Sytox orange, in the presence or absence of 5 mM ATP, was added to the flow cell at 50 μL min^{-1} for 1 min and then at 10 μL min^{-1} for 5 min. To detect Rep binding to the DNA, the Sytox orange-stained rolling-circle DNA template was visualized with a 532-nm laser (90 mW cm^{-2}) sequentially for 200 ms once every second for 3 min. The Rep-AF647 protein was visualized with a 647-nm laser (200 mW cm^{-2}) sequentially for 200 ms once every second for 3 min. Fluorescence signals were captured with an EMCCD camera (Andor iXon 897) with appropriate filter sets.

6.2.7.3. Single-molecule rolling-circle replication assay

The single-molecule rolling-circle replication assays were carried out in microfluidic flow-cell devices with the same DnaBC pre-incubation step and DNA adsorption step as described above. The replication step was carried out as previously described [34, 74, 75, 78] with modifications, described below.

6.2.7.4. Visualization of Rep during replication

The replication solution contained specified concentrations of His$_6$-Rep WT or Rep-AF647, 30 nM Pol III $\alpha\epsilon\theta$ core, 10 nM $\tau_3\delta\delta'\psi\chi$, 46 nM β_2, 75 nM DnaG and 20 nM SSB$_4$ in SM buffer plus 250 μM of each NTP, 50 μM of each dNTP. Replication was initiated by injecting the replication solution into the flow cell containing immobilized DNA-DnaBC at 70 μL min^{-1} for 1 min and then slowed to 10 μL min^{-1} for 10 min. Sytox orange stained DNA molecules were imaged with a 514-nm laser (200 mW cm^{-2}) sequentially for 200 ms once every second for a period of 1 min. The Rep-AF647 protein was visualized with a 647-nm laser (200 mW cm^{-2}) sequentially for 200 ms once every second for 1 min.

6.2.7.5. *Replication rescue at pre-incubated roadblocks*

Visualization of replication rescue by His$_6$-Rep WT, His$_6$-Rep K28A, or His$_6$-Rep ΔC33 in conditions where the DNA template has been pre-incubated with the dCas9 roadblock was carried out as described previously [74] with the following modifications. 5 nM dCas9 was incubated with 20 nM cgRNA1-Atto647 for 5 min at 37°C in SM buffer (omit Sytox orange). The dCas9-cgRNA1-(Atto647) complex, referred to as dCas9-cgRNA1-Atto647, was then incubated with 80 pM 2-kbp rolling-circle DNA template for 5 min at 37°C. To form the DNA pre-incubation solution, the DNA-dCas9-cgRNA1-Atto647 complex was then incubated with 70 nM DnaBC for 3 min at 37°C. The DNA pre-incubation solution was then diluted 1:10 with SM buffer plus 150 nM Sytox orange and adsorbed to the surface of the flow cell at 10 μL min^{-1} until an appropriate DNA density was achieved. The flow cell was then washed with 200 μL of SM buffer. Following this, replication was initiated with the continuous presence of replication proteins as above in the absence or presence of 20 nM His$_6$-Rep WT, His$_6$-Rep K28A, or His$_6$-Rep ΔC33. DNA and dCas9-cgRNA1-Atto647 were imaged in multiple fields of view per experiment, by sequential excitation with 532-nm (90 mW cm^{-2}) and 647-nm (200 mW cm^{-2}) lasers for 200 ms once every 10 s for 3 min.

6.2.7.6. *Replication rescue with roadblocks in solution*

Visualization of replication rescue by Rep-AF647 in conditions where both Rep and dCas9-cgRNA complexes were in solution with the replisome components was carried out as previously described [34, 74, 75, 78], with the following modifications. First, the dCas9-cgRNA complex was formed with the specified cgRNA as described for the pre-incubation of roadblocks with DNA assay. Next, the DNA-DnaBC complex was formed and adsorbed to the flow-cell surface as described previously. Following, the replication solution was mixed as described for visualization of Rep during elongating replication assays, with the addition of specified concentrations of dCas9-cgRNA (determined by dCas9 concentration in the complex) and Rep-AF647. Reactions were initiated with the addition of the replication solution to the flow cell at 70 μL min^{-1} for 1 min and then slowed to 10 μL min^{-1} for 10 min. DNA and Rep were imaged in one field of view per experiment, by sequential excitation with 532-nm (90 mW cm^{-2}) and 647-nm (200 mW cm^{-2}) lasers for 200 ms once every second for 4 min.

6.2.7.7. *Single-molecule characterization of mismatch gRNA binding durations on short oligos*

To assess the duration of binding of mismatch gRNAs (MMgRNA) in complex with dCas9, an 83-bp oligo containing one sequence complementary to the fully complementary gRNA (cgRNA1) was used (Supplementary Table S1) [74]. First, the dCas9-MMgRNA complexes were formed by pre-incubating 5 nM dCas9 with 20 nM of the specified MMgRNA-Atto647 in SM buffer (with ATP omitted throughout) for 5 min at 37°C. Next, 1 pM of 83-bp oligo was adsorbed to the flow-cell surface in SM buffer at 10 μL min^{-1} for 2 min. The dCas9 pre-incubation mixture was then diluted 1:10 with SM buffer and flowed into the flow cell at 70 μL min^{-1} for 1 min. The buffer was then switched immediately to only SM buffer and flowed at 10 μL min^{-1} for 20 min. The DNA oligos and dCas9-MMgRNA-Atto647 complexes were imaged in multiple fields of view per experiment, by sequentially exciting with 532-nm (140 mW cm^{-2}) and 647-nm (200 mW cm^{-2}) lasers for 200 ms once every 30 s for 10 min.

6.2.8. Data Analysis

All analyses were carried out using ImageJ/Fiji (1.51w), MATLAB 2016b, OriginPro 2021b and in-house built plugins, found here: https://doi.org/10.5281/zenodo.7379064.

6.2.8.1. Degree of labeling

The number of fluorophores per labeled Rep-AF647 was quantified at the single-molecule level by immobilizing 40 pM Rep-AF647 onto the flow cell surface in SM buffer (less ATP). The fluorophores were imaged by exciting for 200 ms constantly for 3 min until the fluorophores were photobleached. Raw movies were corrected for the electronic offset of the camera and excitation-beam profile. Single molecules of Rep-AF647 were identified using an in-house built peak fitter tool and the photobleaching steps were fit using change-point analysis [362-364]. The histogram of steps per molecule showed a degree of labeling of mostly 1 dye per Rep monomer (Supplementary Figure 6.S2). It is likely observations of more than 2 fluorophores represent more than a monomer in the identified peak observed, as there is only one cysteine residue in each monomer.

6.2.8.2. Automatic analysis of rolling-circle DNA replication

To automatically track the replication of the rolling-circle DNA templates, raw videos (.nd2 format) were first converted to TIF files and flattened with the excitation beam profile, as previously described [361]. Detectable drift between subsequent frames was then corrected and any un-replicated molecules were removed by subtracting the first frame from each consecutive frame. This prevents un-replicated DNA templates on the surface from interfering with the detection and tracking of replicating molecules. Next, replicating molecules of interest were selected and analyzed individually. Positions of the replicating molecule were determined with custom-written ImageJ and MATLAB plug-ins that detect the leading edge of each replicating molecule, saving as coordinates for downstream analysis. These coordinates were used to detect individual rate segments of the replicating molecules by change-point analysis.

6.2.8.3. Automatic tracking of labeled Rep during rolling-circle DNA replication

Videos of labeled Rep proteins during rolling-circle DNA replication were flattened and prepared as above. Using the coordinates saved from automatic tracking of the replicating DNA molecule, Rep-AF647 molecules were detected within the proximity of the tip of the replicating molecule by using custom-written ImageJ and MATLAB plug-ins. Here, regions of interest around the replicating tip of the DNA molecules were expanded and transposed into the corresponding video of Rep-AF647. Peak finder was used to detect any Rep-AF647 molecules and colocalization was determined for those peaks that were within the distance of the tip of the replicating DNA molecule. The intensities of colocalized Rep-AF647 molecules were measured and individual binding events were detected by change-point analysis.

6.2.8.4. *Replication efficiency analysis*

The replication efficiency (%) was calculated by dividing the number of DNA templates replicating at the end of the reaction by the number of DNA templates immobilized on the coverslip surface at the beginning of the reaction. Briefly, the number of DNA templates immobilized on the coverslip surface was detected using the custom-written ImageJ plugin, peak finder. The number of DNA templates replicating at the end of the reaction was detected by hand. The reported means and S.E.M. represents the replication efficiency calculated from at least *N* (replicates) = 3.

6.2.8.5. *Determination of stoichiometry of Rep*

The number of labeled Rep molecules binding to DNA, at actively replicating or stalled replisomes, was calculated by dividing their initial intensities by the intensity of a single fluorophore, as previously described [34, 78]. Briefly, the average intensity per fluorophore was quantified by detecting photobleaching steps in labeled Rep non-specifically bound to the coverslip surface (Supplementary Figure 6.S2). Imaging was done under the same conditions as the experiment of interest. The integrated intensity for every fluorescently labeled Rep visible in the field of view was calculated after applying a local background subtraction. The histograms obtained were fit with a Gaussian distribution function to give the average intensity of a single molecule.

6.2.8.6. *Colocalization analysis*

Sytox orange stained DNA molecules and dCas9-cgRNA-Atto647 complexes, dCas9-MMgRNA-Atto647, or Rep-AF647 molecules were classed as being colocalized as previously described [78]. Briefly, foci of interest were classed as being colocalized if their centroid positions (determined using an ImageJ in-house built peak finder tool) fell within 2 pixels of each other. The chance of coincidental colocalization (C) was calculated using Equation 2, where A_R is the focus area, A_{FOV} is the field of view area, and n is the number of foci.

$$C = \frac{A_R}{A_{FOV}} \times n \tag{2}$$

6.2.8.7. *Determination of Rep association*

Rep association dynamics were extracted by tracking the fluorescence intensity of Rep-AF647 molecules over time. A threshold was applied for each trajectory equivalent to the intensity of half a Rep-AF647 molecule. The binding frequency of Rep-AF647 during replication was defined as the number of times per minute where the intensity exceeded the threshold.

6.2.8.8. *Statistical analysis*

Statistical analysis of data was carried out using OriginPro 2021b. Data were collected from images from at least two independent experiments for every condition. Significant differences between conditions were analyzed using a one-way ANOVA with subsequent post hoc testing with Tukey's test, with $p \leq 0.05$ determined to be statistically significant. Where distributions did not meet assumptions of parametric tests, the non-parametric Kruskal-Wallis test with Dunn's multiple comparisons was used, with $p \leq 0.05$ determined to be statistically significant.

6.3. Results

6.3.1. Purified and labeled Rep binds to ssDNA

Association of Rep and other similar SF1B helicases to various DNA structures has been widely characterized by previous studies (reviewed in [79, 385]). We used surface plasmon resonance (SPR) to assess DNA binding and ATP hydrolytic properties of the His$_6$-tagged Rep WT, Rep ∆C33 and Rep K28A variants. Since His$_6$-Rep variants have been previously used *in vitro* and it was shown that the His$_6$-tag did not affect the proteins' activity [155], His$_6$-Rep proteins used in this study are hereafter referred to as Rep WT, Rep ∆C33 and Rep K28A. First, we assayed the binding of Rep WT at a saturating concentration (400 nM) to 3'-biotinylated-dT$_{35}$ (3'-bio-dT$_{35}$) immobilized onto the surface of a streptavidin-coated SPR chip (Figure 6.1A). We found that the high-affinity interactions of Rep WT with dT$_{35}$ at this concentration showed a biphasic association and dissociation kinetics (Supplementary Figure 6.S3A). Rep is known to multimerize on and to bind tightly to ssDNA [386-388], so binding to long oligos at high concentrations is expected to be a complex process involving both protein-protein and protein-DNA interactions. Nevertheless, based on the signal amplitude we estimated (Equation 1) that four Rep monomers can bind to one dT$_{35}$. This is consistent with the observed footprint of a Rep monomer of approximately 8 nt [115].

To simplify the analysis, much lower concentrations of Rep were used to examine the binding of Rep WT to dT$_{35}$ (Figure 6.1B, Supplementary Figure 6.S3B). With Rep WT at 20 nM, equilibrium was not reached during the short injection of Rep, but kinetic data fit well to a 1:1 interaction model (with mass transfer) to give association (k_{on} = (1.03 ± 0.00) × 10^6 M^{-1} s^{-1}) and dissociation (k_{off} = (2.15 ± 0.03) × 10^{-4} s^{-1}) rate constants; $K_D = k_{off}/k_{on}$ = 0.2 nM. Further, we interrogated the binding of Rep WT to shorter ssDNA oligonucleotides (dT$_{10}$ and dT$_{15}$). We hypothesize that only Rep monomers are likely to bind to dT$_{10}$ since it cannot accommodate the predicted ≥16-nt footprint of a Rep dimer. We observed that Rep interacted very similarly with dT$_{10}$ and dT$_{15}$, but dissociated much more quickly from the short oligos than from dT$_{35}$ (Figure 6.1B). To estimate the K_D of Rep binding to dT$_{15}$, we titrated Rep WT at different concentrations (1, 2, 4, and 8 nM) (Supplementary Figure 6.S3C). Global fitting of association and dissociation (1:1 binding with mass transfer) yielded a K_D for Rep:dT$_{15}$ of 0.5 nM, with k_{on} = (1.22 ± 0.00) × 10^7 M^{-1} s^{-1} and k_{off} = (5.91 ± 0.02) × 10^{-3} s^{-1}. Thus, Rep binds to dT$_{35}$ to give a much larger (normalized) response during the observation time, is more stably bound to dT$_{35}$ than to dT$_{10}$ and dT$_{15}$, and binds to dT$_{15}$ with a 2.5-fold lower K_D. We interpret these differences among the various ssDNA lengths to be due to Rep binding as a monomer to dT$_{10}$ and dT$_{15}$, and to its assembly as a more stably bound dimer (or other multimer) on dT$_{35}$.

We looked at the effect of nucleotides on the affinity of Rep WT, K28A, and ∆C33 for dT$_{35}$. We show that ATP, but not ADP or AMP-PNP, stimulated the dissociation of Rep WT and Rep ΔC33 from ssDNA (Figure 6.1C and Supplementary Figure 6.S3D). Dissociation of Rep K28A cannot be stimulated by ATP (Supplementary Figure 6.S3D). This mutant does not bind or hydrolyze ATP [154], suggesting that ATP hydrolysis is required for the induced faster dissociation of Rep variants from ssDNA. Additionally, ssDNA-binding and ATP-stimulated dissociation of Rep WT was similar with dT_{35} immobilized via biotin attached to either the 3' or 5' end (3'-bio-dT_{35} and 5'-biotinylated-dT_{35}; data not shown), suggesting Rep directly dissociated from DNA upon ATP binding or hydrolysis rather than translocating off the 5'-end.

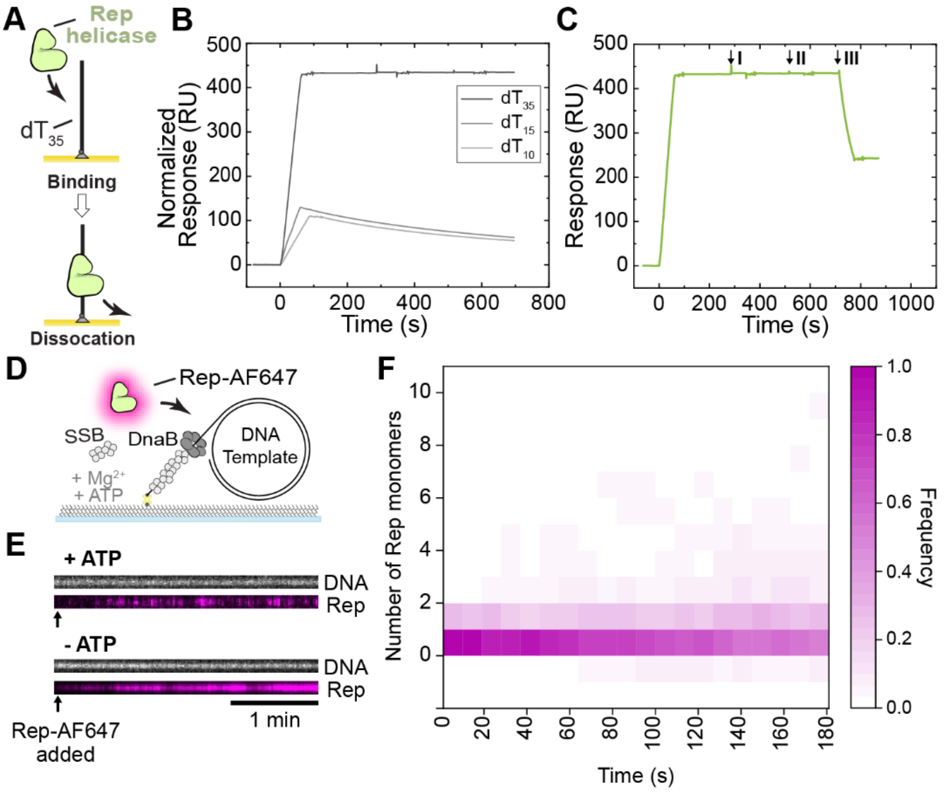


Figure 6.1. Visualization of Rep binding to ssDNA. A) Schematic representation of Rep proteins binding to dT_{35} oligos in surface plasmon resonance investigations. (B–C) SPR sensorgrams of **B)** 20 nM Rep WT association (60 s) and dissociation from dT_{35} (dark gray), dT_{15} (gray) and dT_{10} (light gray). Sensorgrams for Rep binding to dT_{10} and dT_{15} were normalized to that for dT_{35} to enable comparison of binding to the same molar surface density of immobilized oligos. Based on the responses and molecular weights of Rep and oligos, it is estimated that ~1050 RU of bound Rep corresponds to a stoichiometry n = 1.0 Rep monomers bound to each oligo (Equation 1). **C)** Response of dT_{35}-bound Rep WT to interrogation by 200 μM AMP-PNP (I), ADP (II), and ATP (III) injection at 20 μL min^{-1} for 60 s. Only ATP resulted in fast dissociation of Rep from dT_{35}. **D)** Schematic representation of the single-molecule Rep-AF647-DNA binding assay. Rolling-circle DNA templates (2030 bp) are pre-incubated with DnaBC and applied to a microfluidic flow cell. The 5'-biotinylated DNA template couples to the streptavidin functionalized coverslip. Rep-AF647, in the presence of SSB, ATP, and magnesium, is then applied to the flow cell and imaged to monitor for binding. **E)** Example kymographs of Rep-AF647 (magenta – bottom) binding to DNA templates (gray – top) in the presence and absence of 5 mM ATP. Arrows indicate the time point of Rep-AF647 addition to the flow cell. **F)** Heatmap of the number of Rep-AF647 monomers bound to the DNA template over time in the presence of ATP (n = 65).

To visualize Rep behaviors in single-molecule fluorescence assays, we used a previously-studied (44) mutant of His$_6$-Rep (A97C) site-specifically labeled with a cysteine-reactive red fluorescent dye (Alexa Fluor 647). Wild-type Rep contains five native and non-essential cysteine residues, all of which have been substituted, and a now-unique cysteine residue replaces Ala97 (44). This residue is within the 1B subdomain and predicted to be close to the 3'-end of the ssDNA. This subdomain is not responsible for ATP binding or hydrolysis, or rotation of the 2B subdomain, activities previously characterized to be important for Rep functions [68, 130, 141, 142, 155, 389]. We show in bulk biochemical assays, using the 2-kbp rolling-circle DNA template that both the unlabeled and labeled His$_6$-Rep A97C proteins are as active as Rep WT in removing the dCas9-cgRNA1 replication roadblock, an activity that is increased at increased ATP concentration under the conditions used (Supplementary Figure 6.S4A and B). Therefore, the fluorescently labeled Rep variant appears to be fully functional and is hereafter referred to as Rep-AF647.

To characterize Rep-AF647 binding to DNA in single-molecule fluorescence assays and to verify that it does not interact significantly with double-stranded DNA, we used a 5'-biotin-tailed, gapped and circular DNA template (2030 bp) that contains a replication fork structure [356]. This template is almost completely double-stranded, except for a total of 64-nt at the fork that could bind Rep. We set out to characterize the binding of Rep-AF647 in the presence and absence of ATP to DNA occupied by DnaB and SSB (Figure 6.1D). DnaBC and SSB were pre-incubated with the DNA template and then injected into a microfluidic flow cell to immobilize the complex on a streptavidin-functionalized surface. Rep-AF647, in the presence or absence of ATP, was subsequently injected into the microfluidic flow cell, and binding events were visualized by near-total internal reflection fluorescence (TIRF) imaging of Sytox orange-stained DNA and AF647-labeled Rep molecules. Colocalization and stoichiometry analysis of corresponding foci confirmed that in the absence of ATP, Rep-AF647 binds stably to the DNA template (Figure 6.1E, Supplementary Figure 6.S4C). Additionally, significant numbers of Rep-AF647 single-molecules remained bound; equivalent to approximately 8–10 molecules (Supplementary Figure 6.S4D). This observation suggests that the ssDNA in the gapped circular DNA and ssDNA overhang (total of 64 nt) is occupied by a Rep molecule(s) by the end of the acquisition, assuming that Rep has a footprint of 8 nt on ssDNA [115]. Therefore, this occupancy in the absence of ATP could suggest that Rep out-competes or interacts with SSB bound to the ssDNA regions. Similar to our observations in the SPR studies, in the presence of ATP this stable binding activity is significantly reduced (Figures 6.1E and F *cf.* Supplementary Figure 6.S4D). Rep-AF647 appears to bind only transiently to DNA in the presence of ATP, suggesting that ATP binding and hydrolysis play an important role in Rep-ssDNA affinity. In agreement with earlier studies [390], we show that ATP binding, and potentially hydrolysis, put Rep into a low-affinity state, resulting in dissociation from the ssDNA template. These data show that the labeled Rep protein is active in DNA binding, ATP binding and ATP hydrolysis. Further, our SPR and single-molecule studies confirm the stability of Rep in the absence of ATP on oligonucleotides containing more than 16 nt [385, 391]. Together, our results confirm previous hypotheses about Rep-ssDNA affinity and activity.

6.3.2. Rep associates frequently with elongating replication forks in the absence of roadblocks

Early studies of the growth characteristics of *rep* mutant strains suggested higher replication rates in the presence of Rep [392]. Additionally, previous studies have shown that Rep interacts with the DnaB helicase through its C-terminal region [59, 62, 154]. Therefore, we first set out to investigate the effect of Rep WT on the rate and processivity of replication using a single-molecule rolling-circle DNA replication assay. This assay utilizes the 2-kbp rolling-circle template used above, where the 5'-tail is anchored to the surface of the flow cell [48, 49, 356]. Replication is initiated by the introduction of a laminar flow of buffer containing the *E. coli* replication proteins required for coupled leading- and lagging-strand synthesis to form functional replisomes at the fork structure. Initiation of unwinding and synthesis at the fork results in the newly synthesized leading strand being displaced from the circle to form the template for the lagging strand. This process results in the production of a dsDNA tail that is stretched out in the buffer flow and the movement of the circle away from the anchor point, at a rate determined by the replication rate (Figure 6.2A). These replication events are visualized in real-time by TIRF imaging of Sytox orange-stained dsDNA.

Analysis of the replication rates of individual replisomes in the absence and presence of unlabeled, wild-type Rep (Rep WT) (5 nM) resulted in median rates of 580 ± 30 and 570 ± 30 bp s^{-1} (median ± SEM), respectively (Supplementary Figure 6.S5A). These rates are in agreement with previously reported *E. coli* replication rates [34, 48, 49, 75]. Further, at 10- and 100-fold higher concentrations of Rep WT, resulting replication events revealed rates of 500 ± 20 and 470 ± 30 bp s^{-1}, respectively. Taken together, no significant difference was found between median replication rates, suggesting that Rep WT has no effect on the rate of replication at the concentrations used. Further, we saw no significant effect of Rep on the processivity of the replisome, as measured by the length of individual rate segments (Supplementary Figure 6.S5B).

The absence of a clear effect of Rep on DNA replication rate and processivity raises the question of whether Rep is present at all at elongating forks in the absence of protein roadblocks. Previous studies have hypothesized that Rep is only present at replication forks in the event a roadblock is encountered, being that it is recruited to the replication fork upon stalling [59, 153, 155]. To investigate whether Rep is present at elongating replisomes, we added Rep-AF647 to the single-molecule rolling-circle DNA replication assays described above (Figure 6.2A). This assay provides additional information as it allows simultaneous imaging of fluorescently labeled replisome components and interacting partners during the replication reaction [34, 75, 368]. The addition of Rep-AF647 at a concentration of 20 nM in this assay shows that Rep-AF647 is transiently located at the tip of the growing DNA molecule consistent with its interaction with the elongating replisome (Figure 6.2B and Supplementary Figure 6.S6). Only 3% of DNA molecules analyzed showed no Rep-AF647 intensity above the threshold at a concentration of 20 nM Rep-

AF647. Further, the intensity of interacting Rep-AF647 with the replisome fluctuates throughout the elongation of the DNA product (Figure 6.2D). A single-exponential fit of binding event durations, as detected by change-point analysis of Rep fluorescence over time, revealed an approximate binding lifetime of 2.0 ± 0.2 s. If the same Rep-AF647 molecule remained bound to the active replisome, the intensity should decay at the characteristic lifetime of photobleaching (8.0 ± 0.1 s) (Supplementary Figure 6.S2E). Therefore, it is likely that Rep-AF647 associates with the replication fork and quickly dissociates to be replaced by a new Rep-AF647 molecule from solution, as observed for Rep-AF647 binding investigations above.

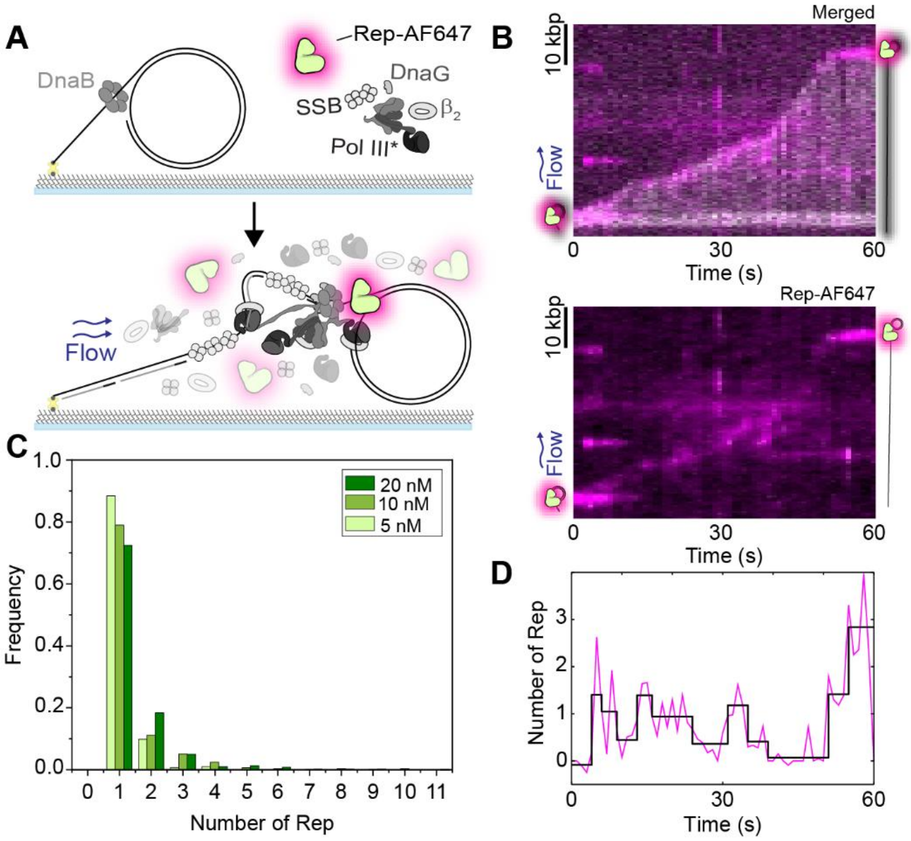


Figure 6.2. Rep interacts with processive replisomes. A) Schematic representation of the single-molecule rolling-circle DNA replication assay in the presence of Rep-AF647. Rolling-circle DNA templates are pre-incubated with DnaBC and immobilized to the flow cell surface. The addition of the replisome components, ATP, NTPs, and dNTPs, results in the initiation of DNA replication. Sytox orange-stained DNA replication products are stretched out by hydrodynamic flow and visualized simultaneously with fluorescent Rep-AF647 proteins. **B)** Example kymograph of 20 nM Rep-AF647 interacting with the tip of the Sytox orange stained DNA product (merged – top). The Rep-AF647 intensity alone (bottom) shows its frequent association with and dissociation from the replication fork. **C)** Histogram distributions of Rep-AF647 stoichiometry at the replication fork reveal monomeric stoichiometry at 20 nM (dark green; *n* = 755), 10 nM (green; *n*= 490) and 5 nM (light green; *n* = 286). **D)** Number of Rep-AF647 as a function of time for the example kymograph in (B) showing variations in Rep stoichiometry at the replication fork during processive replication. Individual steps are detected by change-point analysis (black line).

Previous studies have hypothesized the potential of six Rep monomers binding to the replication fork through the hexameric structure of the DnaB helicase [59, 155]. Our assays reveal an approximate stoichiometry of 1–2 Rep at the fork at any given time point, independent of the concentration of Rep-AF647 used (5, 10, or 20 nM) (Figure 6.2C). We determined the average binding frequency of a Rep-AF647 molecule to an elongating replisome by applying thresholding analysis to the Rep-AF647 signal to detect binding events. We found that the average binding frequency during replication elongation was 16 ± 1 min^{-1} at 20 nM Rep-AF647. This average frequency decreased with decreasing Rep-AF647 concentration, to 10 ± 1 and 7 ± 1 min^{-1} at 10 and 5 nM, respectively. Thus, the higher the local concentration of Rep is, the more frequently it associates with the fork. Given the low stoichiometry and regular binding events, these observations suggest potentially two behaviors: 1) that Rep stochastically binds to the replisome, and 2) Rep is interacting with the DnaB helicase, but not occupying all potential binding sites on the hexamer at a given time. Nonetheless, we show that Rep associates with elongating replication forks in the absence of protein roadblocks.

6.3.3. Wild-type Rep removes model roadblocks and rescues stalled replication forks

To investigate if Rep could remove nucleoprotein complexes and rescue stalled replication forks, we used dCas9 in complex with a complementary guide RNA (dCas9-cgRNA) as a protein barrier. We have previously shown we can block the reconstituted *E. coli* replisome in *in vitro* bulk and single-molecule replication assays using dCas9-cgRNA complexes targeted to a specific site in the rolling-circle template [74]. We first tested the ability of Rep WT to remove dCas9-cgRNA complexes targeted to either the leading or lagging strand, in bulk biochemical assays. Here dCas9-cgRNA complexes are incubated with the rolling-circle DNA template, to which the *E. coli* replisome proteins are added. After 10 min of the replication reaction, the indicated concentration of Rep WT is added to the reaction and allowed to proceed for a further 10 min. Finally, the reactions are quenched and products are analyzed by gel electrophoresis.

We show that in reactions containing Rep WT, replication of DNA templates occurs past the dCas9-cgRNA binding site, indicating that the dCas9-cgRNA roadblock has been removed (Supplementary Figure 6.S7). This activity occurs in reactions where either the leading or lagging strand is targeted by the dCas9-cgRNA complex (Supplementary Figure 6.S7A). The addition of a trap dsDNA acts to bind free dCas9-cgRNA1 complexes in solution, and thus prevents free dCas9-cgRNA1 complexes from rebinding to the template DNA, where in its absence a laddering in the gel is seen consistent with multiple replication blocking and rescuing events (Supplementary Figure 6.S7B, C). At all concentrations tested in bulk biochemical assays (2–300 nM Rep WT), each reaction resulted in long DNA products showing that Rep has effectively removed the dCas9-cgRNA1 complex allowing replication to proceed (Supplementary Figure 6.S7B, C). Additionally, no replication rescue was observed in reactions containing either the Rep ∆C33 or K28A mutants (Supplementary Figure 6.S7D).

To assess this activity at the single-molecule level, we repeated the above experiments by pre-incubating DNA templates with the dCas9-cgRNA1-Atto647 complex. This allowed us to simultaneously visualize both the potential loss of the dCas9-cgRNA1-Atto647 complexes and the growing replication products, by collecting the Atto647 fluorescence emission and dsDNA-bound Sytox orange fluorescence emission, respectively. Pre-incubation of the DNA template with the roadblock complex allows replication to proceed until the block is encountered (Figure 6.3A). Before adsorption to the flow-cell surface, the DNA-dCas9-cgRNA1-Atto647 complex is pre-incubated with the DnaBC complex.

Consistent with the bulk assays and previous studies, we saw effective replication blocking of the *E. coli* replisome in the presence of the dCas9-cgRNA1-Atto647 complex (Figure 6.3B, C, and G, Supplementary Figure 6.S8), indicated by the synthesis of short replication products [74]. This blocking is also reflected by the replication efficiency of the reactions, where the presence of dCas9-cgRNA1-Atto647 resulted in a replication efficiency of 0% (no replicating molecules observed). In the presence of Rep WT, long DNA products are synthesized following the removal of the dCas9-cgRNA1-Atto647 complex (Figure 6.3D and G, Supplementary Figure 6.S8). This recovery is also represented by the increase in replication efficiency (2 ± 1%) when compared to the absence of Rep WT. Observations of dCas9-cgRNA1-Atto647 complexes photobleaching when imaged every 200 ms, showed a characteristic photobleaching lifetime of 87 s (Supplementary Figure 6.S8D). However, in these experiments, the dCas9-cgRNA1-Atto647 complexes are only imaged once every 10 s, extending the lifetime of the fluorophore. Thus, we are confident that the loss of the dCas9-cgRNA1-Atto647 intensity is due to the displacement of the complex by Rep WT. Additionally, unlike the ensemble assays, displaced dCas9-cgRNA1-Atto647 complexes are carried away from replicating templates in the buffer flow, thus preventing rebinding to the template. In the presence of the Rep ΔC33 or K28A mutants, no removal of the roadblock or active replication is observed (Figures 6.3E, F, and G, Supplementary Figure 6.S8E, F). The experimental setup of these single-molecule reactions involves only pre-incubated DnaBC with the template DNA, and DnaBC is not present free in solution during the replication reaction. Therefore, these results suggest that the DnaB helicase remains bound after the replisome encounters the dCas9-cgRNA1-Atto647 roadblock, consistent with previous ensemble *in vivo* studies [71, 73, 393]. Taken together, these observations demonstrate that Rep WT effectively removes the dCas9-cgRNA1 complex and that the success of this activity is dependent on a functional ATPase domain and the presence of the C-terminus that interacts with DnaB.

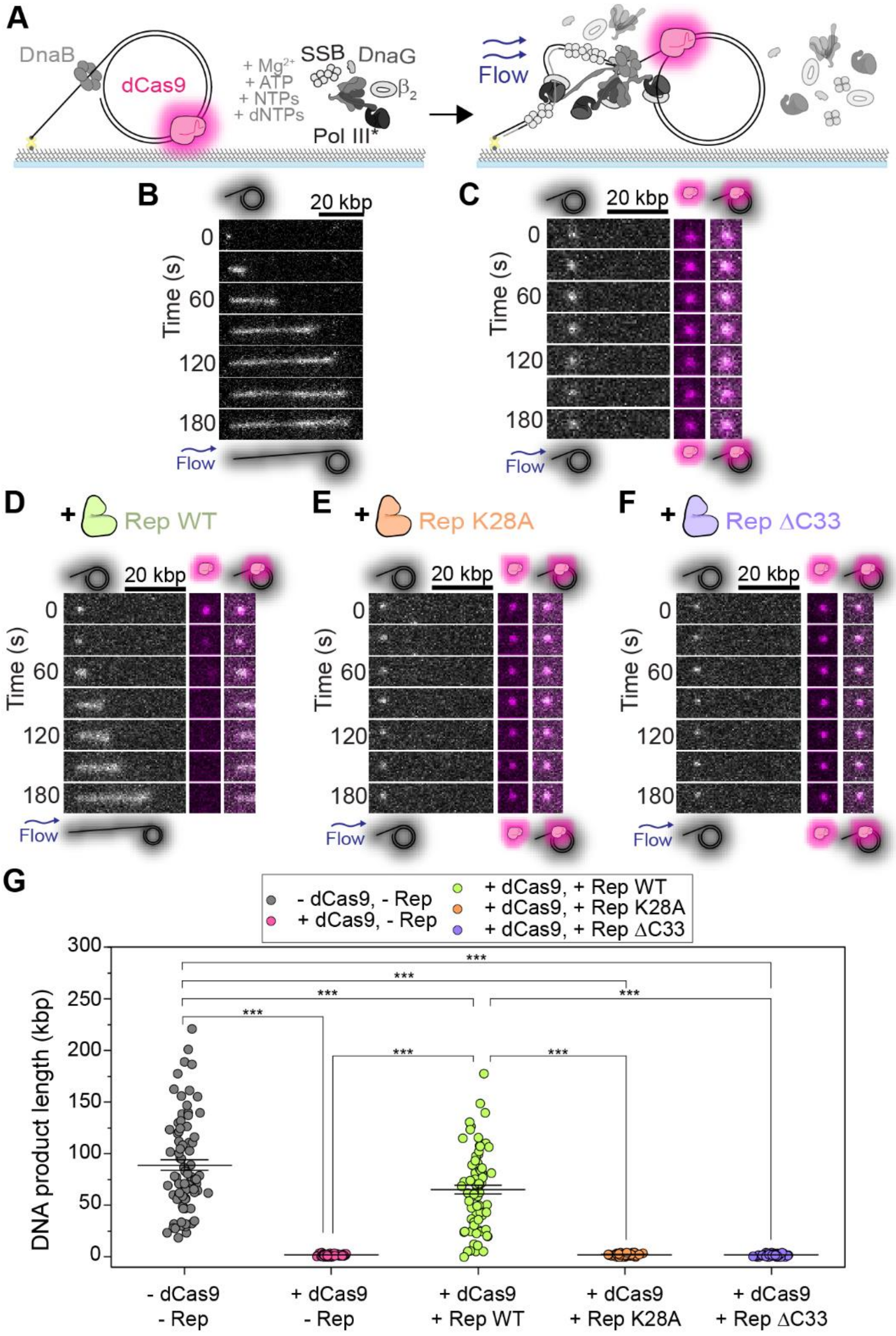


Figure 6.3. Visualization of stalled replication rescue by Rep at the single-molecule level. A) Schematic representation of single-molecule stalled rolling-circle replication assay. The dCas9-cgRNA1-Atto647 complex is pre-incubated with the rolling-circle DNA template, and further incubated with DnaBC before immobilization to the flow cell. The addition of the replisome components results in the initiation of replication until the dCas9-cgRNA1-Atto647 roadblock has been encountered. Stalled replication is imaged by visualizing the Sytox orange stained-DNA (gray) and dCas9-cgRNA1-Atto647 complex (magenta). **B–F)** Example montages and **(G)** mean DNA product length of **(B)** rolling-circle DNA replication in the absence of protein roadblocks and Rep proteins (89 ± 5 kbp; *n* = 81; replication efficiency of 4 ± 1% (S.E.M.)). **C)** Stalled DNA replication by dCas9-cgRNA1-Atto647 complex in the presence of all replisome components (2.0 ± 0.1 kbp; *n* = 80; no replicating products observed). **D)** stalled replication rescue by Rep WT following removal of

the dCas9-cgRNA1-Atto647 complex (65 ± 4 kbp; *n* = 75; 2 ± 1%). **E)** Stalled replication in the presence of Rep K28A (2 ± 1 kbp; *n* = 80; no replicating products observed). **F)** stalled replication in the presence of Rep ∆C33 (2 ± 1 kbp; *n* = 80; no replicating products observed). **G)** Total DNA product length after 3 min, where bars represent the reported mean ± S.E.M, as listed for (B–F). Comparison of distributions was conducted using one-way ANOVA with Tukey's multiple comparisons post-hoc test, where *** denotes statistical significance with $p \leq 0.001$ and the absence of markers indicates no significant difference ($p > 0.05$).

6.3.4. Rolling-circle DNA templates show periodic replication stalling and rescue events

The rolling-circle DNA template allows for the observation of replication events that proceed for extended periods, limited only by the amount of nucleotides in the solution. Therefore, we hypothesized that the introduction of both the dCas9 roadblock and Rep in solution would result in the observation of multiple cycles of dCas9 binding, fork stalling and Rep-mediated rescue on individual molecules. Here, we repeated rolling-circle DNA replication assays with two modifications: (1) adding Rep-AF647 and dCas9-cgRNA1 complexes in solution with the replisome components, and (2) using an 18-kbp rolling-circle DNA template to resolve stalling and rescue events unambiguously (Figure 6.4A). The 18-kbp rolling-circle DNA template is capable of producing long DNA products at a rate similar to that of the 2-kbp rolling-circle DNA template (Supplementary Figure 6.S9). The addition of Rep-AF647 and dCas9-cgRNA1 complexes to the single-molecule 18-kbp rolling-circle DNA replication assay resulted in long DNA products with multiple, well-defined pausing and rescue events, where trajectories resembled steps at the expected binding sites of the roadblock (Figure 6.4B and Supplementary Figure 6.S9B).

We performed change-point analysis [362, 364] of the trajectories of the position of the replicating DNA molecule in the movies using an automated tracking algorithm. Here, the individual rate segments were defined as stalled replication events or pauses, where the replication rate was below 100 bp s^{-1}. Using this definition, we could then determine the pause sites (expressed in kbp from the start of the replication reaction). Using the 2-kbp rolling-circle DNA template, we observed multiple stalling and rescue events in reactions containing dCas9-cgRNA1 and Rep in solution (Supplementary Figure 6.S10A). To confirm these observations, we repeated the experiment with the 18-kbp rolling-circle DNA template, which showed pausing at the expected target sites of the dCas9-cgRNA1 complex and subsequent rescue (Figure 6.4B and Supplementary Figure 6.S9B). Pairwise distance analysis of the pause start sites of the 2-kbp rolling-circle DNA templates shows that replisome stalling and rescue events occur at every 2 kbp, or integers of 2 kbp, evident by the clustering around these distances and consistent with our expectations (Figure 6.4C).

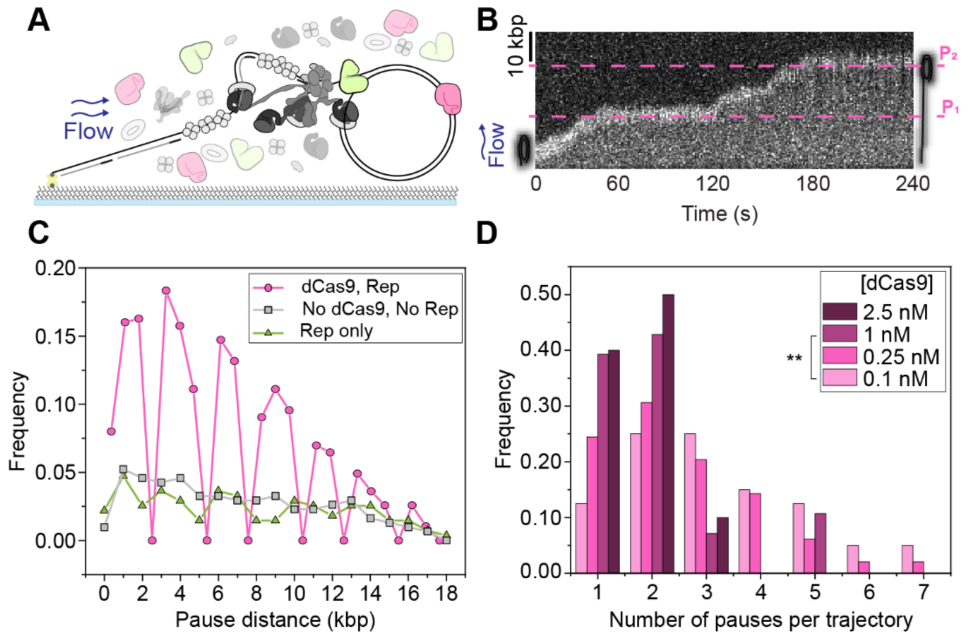


Figure 6.4. Observations of multiple stalling events. A) Schematic representation of single-molecule stalled replication rescue assays, pre-incubated with DnaBC and immobilized to the flow cell surface. Replication is initiated in the presence of Rep and dCas9-cgRNA1. **B)** Example 18-kbp rolling-circle DNA template undergoing multiple replication stalling and rescue events, at approximately 17 kbp (P_1) and 36 kbp (P_2). The target site of the dCas9-cgRNA1 complex occurs once every 18 kbp of the DNA template. **C)** Pairwise distance analysis of the pause site replication rescue events on the 2-kbp DNA template in the presence of 10 nM Rep-AF647 and 0.25 nM dCas9-cgRNA1 (magenta, 16 pauses/275 kbp), only Rep-AF647 (green, 16 pauses/275 kbp) and absence of both proteins (gray, 18 pauses/307 kbp), for the first 20 kbp of DNA products. Symbols represent the distribution of histogram bin heights, normalized to the total DNA product length. Pauses in the absence of dCas9-cgRNA1 represent spontaneous pausing of the replisome. **D)** Histograms of the number of pauses per replicating molecule at titrated dCas9-cgRNA1 complexes in the presence of 10 nM Rep-AF647 using the 2-kbp rolling-circle DNA template; 2.5 nM dCas9-cgRNA1, 17 pauses from 10 molecules (mean of 2 ± 1 (S.D.), replication efficiency of 1 ± 1% (mean ± S.E.M.); 1 nM dCas9-cgRNA1, 56 pauses from 30 molecules (2 ± 1; 1 ± 1%); 0.25 nM dCas9-cgRNA1, 128 pauses from 51 molecules (3 ± 2; 3 ± 1%); and 0.1 nM dCas9-cgRNA1, 130 pauses from 43 molecules (3 ± 2; 3 ± 1%). Comparison of the mean number of pauses per trajectory was conducted using one-way ANOVA with Tukey's comparison post-hoc test, where ** denotes statistical significance with $p \leq 0.01$ and the absence of markers indicates no significant difference ($p > 0.05$).

In the absence of dCas9-cgRNA1 complexes, the periodic pausing behavior of the replisome was absent. Spontaneous pausing of DNA replication was observed to occur at any site, regardless of the presence of Rep-AF647. Periodic replication stalling and rescue were also observed in assays containing dCas9-cgRNA4 complexes targeted to the leading strand of the 2-kbp rolling-circle DNA template (Supplementary Figure 6.S10B). Despite the lower spatial precision in pause-site identification in the 18-kbp rolling circle template experiments, clusters in the periodicity of pausing are observed only in experiments containing both dCas9-cgRNA1 and Rep-AF647

(Supplementary Figure 6.S9D). We attribute the lower precision to be a result of increasing product length, therefore the number of well-defined pause positions is lower when using the 18-kbp rolling circle template. However, this template allows for multiple pauses (2-3) to be well-defined. Having confirmed that the replication pausing occurs at the dCas9 binding sites, the remaining experiments were conducted using the 2-kbp rolling-circle DNA template due to the low experimental throughput of the 18-kbp template.

We next titrated dCas9-cgRNA1 in solution to observe the effect of the number of stalling and rescuing events per replicating molecule. We observed a significant decrease in both the number of stalling and rescue events and the replication efficiency of molecules undergoing replication at high (1–2.5 nM) dCas9-cgRNA1 concentrations (Figure 6.4D). We attribute this behavior to both the fast association of roadblocks in solution to the target site after displacement by Rep-AF647 and the low resolution of the 2-kbp DNA template making it difficult to distinguish consecutive pausing and rescue events that are only 2-kbp apart in resulting kymographs. Therefore, subsequent reactions were conducted at lower dCas9-cgRNA1 concentrations.

6.3.5. Resolution of stalled replication by Rep shows one rate-limiting step

Having established that we can observe multiple pause and rescue events using the 2-kbp rolling-circle DNA template, we next set out to investigate the activity of Rep during the pause states of the replication fork. We first identified pauses in rolling-circle DNA replication reactions containing only 0.25 nM dCas9-cgRNA1 in solution with the replisome components, or 0 nM dCas9-cgRNA1 and 0 nM Rep-AF647. Notably, in reactions containing only 0.25 nM dCas9-cgRNA1 the pause duration was recovered from a Gaussian distribution to be 140 ± 60 s, almost thirty-fold higher than spontaneous pauses identified in reactions containing 0 nM Rep-AF647 and 0 nM dCas9-cgRNA1 (5 ± 2 s), recovered from a single-exponential distribution (Figure 6.5A). This striking increase in pause duration confirms that pauses observed in reactions containing dCas9-cgRNA1 are caused by the roadblock complex. It is important to note that the duration of pauses in the dCas9-cgRNA1-only conditions reflects the period of time where the replisome was paused after initiation of replication until the end of the 4-minute acquisition. Therefore, the duration of a pause in the absence of Rep is an underestimate, given that previous estimates of the lifetime of the roadblock are on the tens of hours scale [74].

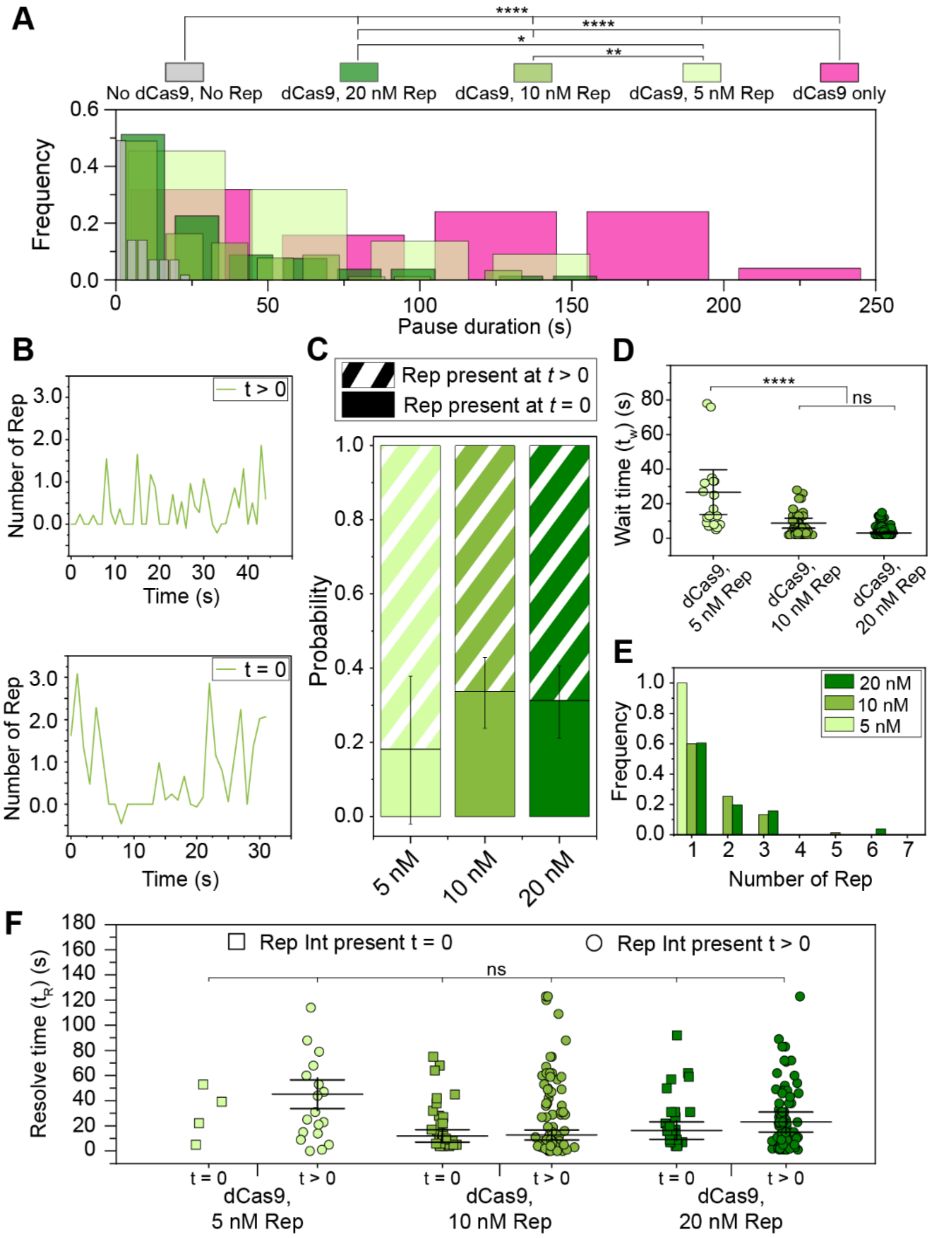


Figure 6.5. Observations of Rep at stalled replisomes. A) Duration of a pause is decreased at increasing concentrations of Rep. In the presence of 0.25 nM dCas9 only, the mean pause duration is determined by fitting a Gaussian distribution function (140 ± 60 s (S.E.M.), *n* = 25 pauses). In the presence and absence of dCas9-cgRNA1 and Rep-AF647, mean pause duration is determined by fitting single-exponential decay functions to the data (No dCas9-cgRNA1, No Rep-AF647, 5 ± 2 s, *n* = 57 pauses; dCas9-cgRNA1 and 5 nM Rep-AF647, 80 ± 40 s, *n* = 22 pauses; dCas9-cgRNA1 and 10 nM Rep-AF647, 20 ± 7 s, *n* = 92 pauses; dCas9-cgRNA1 and 20 nM Rep-AF647, 20 ± 9 s, *n* = 80 pauses). Pause durations of titrated Rep conditions represent pauses where a Rep-AF647 intensity was observed above a threshold. **B)** Example traces of the number of Rep-AF647 present during a pause as a function of time. Two distinct types are observed; Rep-AF647 intensity above the threshold is reached at *t* > 0 (top), and Rep-AF647 intensity above the threshold is reached at *t* = 0. **C)** The probability of observing the two Rep-AF647 activities during a pause for each concentration of Rep. Error bars indicate the margin of error for each concentration. **D)** Distribution of the wait time (t_w) for a Rep-AF647 molecule to associate to the replication fork in *t* > 0 events. The mean wait time for each concentration of Rep-AF647 was determined by fitting a single-exponential decay function to

the data: 5 nM Rep-AF647, 30 ± 40 s (*n* = 18 pauses); 10 nM Rep-AF647, 9 ± 3 s (*n* = 76 pauses); 20 nM Rep-AF647, 3 ± 1 s (*n* = 78 pauses). **E)** Histogram showing distributions of the number of Rep-AF647 molecules that associated first during a pause at t_W reveal predominantly monomeric stoichiometry at all concentrations used. **F)** Pause resolve time (t_R) for $t = 0$ (squares) and $t > 0$ (circles) events. The mean pause resolve times were determined by fitting a single-exponential decay function to the data of each concentration: 5 nM Rep-AF647, ($t = 0$) no fit converged (*n* = 4 pauses), and ($t > 0$) 50 ± 60 s (S.E.M.); 10 nM Rep-AF647, ($t = 0$) 12 ± 2 s (*n* = 31 pauses), and ($t > 0$) 13 ± 4 s; 20 nM Rep-AF647, ($t = 0$) 17 ± 7 s (*n* = 25 pauses), and ($t > 0$) 20 ± 8 s. Comparison of distributions was conducted by Kruskal-Wallis test for multiple comparisons with Dunn's procedure, where *, ** and **** denote statistical significance with $p \leq 0.05$, 0.01 and 0.0001, respectively. The absence of markers or ns denotes no significant difference ($p > 0.05$).

Next, we titrated Rep-AF647 in solution with dCas9-cgRNA1 complexes and visualized it at sites of stalled replication forks (Supplementary Figure 6.S10A). In these dual-channel videos, we identified the pausing of the DNA replisome in the Sytox channel and measured the lifetime of the pause where the association of Rep-AF647 to the site was detected. We observed a reduction in pause duration with increasing concentrations of Rep-AF647. Specifically, at 5 nM Rep-AF647 the pause distribution was best described by a single-exponential fit with a mean duration of 80 ± 40 s, that reduced to 20 ± 7 s and 20 ± 9 s at 10 and 20 nM Rep-AF647, respectively (Figure 6.5A). The plateauing of the pause duration at approximately 20 s for both the 10 and 20 nM Rep-AF647 conditions suggests a saturating concentration has been reached. The significant reduction in pause duration when compared to the dCas9-cgRNA1-only conditions provides evidence that Rep-AF647 is required to remove the roadblock. Further, the single-exponential distributions of the pause durations suggest there exists a single rate-limiting step governing the continuation of DNA replication.

To understand the behavior of Rep-AF647 during these pause events, we plotted the intensity profiles of Rep-AF647 over time during the pauses. This revealed two distinct types of Rep-AF647 signals: (1) events in which Rep associated with the replication fork after the pause started ($t > 0$), or (2) events in which Rep-AF647 was already present at the replication fork when the pause event began ($t = 0$) (Figure 6.5B). The likelihood of these two distinct events occurring was also dependent on the concentration of Rep-AF647 in solution: at higher Rep-AF647 concentration we detected a greater fraction of events in which Rep-AF647 was present at $t = 0$ (5 nM = 20 ± 20%; 10 nM = 30 ± 10%; 20 nM = 30 ± 10%) (Figure 6.5C).

Additionally, we quantified the average binding frequency of Rep-AF647 during the pause state of the replication fork to be slightly higher than that of an elongating replication event. At 20 nM Rep-AF647 a binding frequency of 20 ± 2 events per minute of stalled replication (mean ± S.E.M., *n* = 132 pauses) was determined. This frequency decreases at 10 and 5 nM Rep-AF647, where binding frequencies of 16 ± 1 (*n* = 128 pauses) and 2.0 ± 0.2 (*n* = 132 pauses) events per minute were determined, respectively. The reported binding frequencies during a pause are slightly higher than observed during elongation at the same concentrations (20 nM, 16 ± 1 min^{-1}; 10 nM, 10 ± 1 min^{-1}; and 5 nM, 7 ± 1 min^{-1}). These observations rule out a scenario in which Rep is specifically recruited to the stalled replisome following an encounter with a roadblock; rather the association of Rep with the replication fork is stochastic and more frequent at higher concentrations of Rep in solution.

We further analyzed the events where Rep associates at $t > 0$. These events provide the opportunity to determine the characteristics of the associating Rep molecules after the replisome has stalled, rather than the Rep molecules that were already present at the fork when stalling occurred. We quantified the wait time (t_w) for Rep-AF647 molecules to associate with the stalled replication fork using the increase in the intensity of Rep-AF647 molecules co-localizing with the fork at $t > 0$ (Supplementary Figure 6.S10E). As expected, the t_w for associating Rep-AF647 molecules decreased with increasing Rep-AF647 concentrations, where single-exponential fits to the distribution of wait times revealed a mean t_w of 30 ± 40, 9 ± 3, and 3 ± 1 s for 5, 10, and 20 nM Rep-AF647, respectively (Figure 6.5D). These waiting times reflect the association rate constant of Rep to the replisome since the wait time decreases with increasing Rep concentrations.

Further, we quantified the stoichiometry of the first associating Rep-AF647 molecules during the pause states. Our assays revealed that the predominant stoichiometric state of an associating Rep-AF647 molecule is the monomer at all concentrations of Rep-AF647 used (Figure 6.5E). Our observations of two or more Reps binding could represent either the binding of higher oligomeric states or the association of multiple monomers to the DnaB helicase. Under the experimental conditions used, the intensity quantification is robust enough to see differences between monomers and dimers, but not sufficient to distinguish between dimers and higher oligomers.

Finally, we quantified the pause resolution time (t_R) for both types of Rep-association, where t_R is defined as the duration between the time at which the Rep-AF647 intensity exceeds the background and the time at which the pause is resolved in the corresponding DNA channel (Supplementary Figure 6.S10E). This time period corresponds to the activities needed for roadblock removal and re-start of replication elongation. Interestingly for both observed categories, the pause resolution times were similar at each concentration, especially at 10 nM Rep-AF647 where t_R was determined from single-exponential fits to the resolution time distributions to be 13 ± 4 and 12 ± 2 s for $t > 0$ and $t = 0$, respectively (Figure 6.5F). This similarity, in addition to the single-exponential distributions, provides further evidence that a single rate-limiting step governs the rescue of stalled replication upon Rep association. We cannot distinguish, however, whether this rate-limiting step corresponds to roadblock removal or another process underlying the continuation of replication. Taken together our results support a scenario in which Rep interrogates the state of the replication fork through frequent and stochastic association with the replisome.

6.3.6. Pause duration does not depend on the stability of the roadblock

Having observed that Rep-AF647 can effectively remove dCas9-cgRNA1 complexes that are fully complementary to the DNA target, we set out to investigate the effect of roadblock stability on the pause duration. dCas9-cgRNA1 complexes bind their target sites tightly with 75% of complexes remaining bound after 16 h [74]. The number of mismatched gRNA-target DNA bases has a significant effect on the stability of the complex [335, 338, 340, 394]. Specifically, multiple mismatches in the PAM-distal end of the gRNA-target DNA hybrid trigger faster dissociation. Therefore, we designed a set of mismatch (MM) gRNAs containing 20–80% complementarity to the original target on the lagging strand of the 2-kbp rolling-circle template (termed gRNA80, gRNA60, gRNA40, and gRNA20) (Supplementary Figure 6.S11A).

First, we estimated the approximate binding lifetime of these MM gRNAs to an 83-mer target dsDNA sequence containing a single target site in single-molecule TIRF assays (Figure 6.6A). Here, the 83-mer target dsDNA was immobilized to the coverslip surface and videos of Sytox orange-stained dsDNA were collected as dCas9-MMgRNA-Atto647 was introduced to the flow cell. Images were acquired intermittently every 30 s for a total of 10 min. We observed efficient and stable binding of the dCas9-gRNA80-Atto647 complex until the end of the acquisition, with a mean binding lifetime of 8.6 ± 0.1 min (mean ± S.E.M., *n* = 1414) (Figure 6.6B and Supplementary Figure 6.S11B). This binding lifetime was similar in complexes containing 60% complementarity (gRNA60, 8.3 ± 0.1 min, *n* = 619), where some complexes were observed to remain bound until the end of the acquisition, where others dissociated. Interestingly, complexes containing 40 and 20% complementarity showed very weak binding and were often only observed bound to the 83-mer target DNA for one frame (gRNA40, 1.0 ± 0.1 min, *n* = 249; gRNA20, 1.0 ± 0.1 min, *n* = 109). This is consistent with observations of dCas9 stability where cgRNA-target DNA hybrids containing less than 50% complementarity, especially in the reversibility-determining region, showed significantly reduced lifetimes [394] (Supplementary Figure 6.S11A).

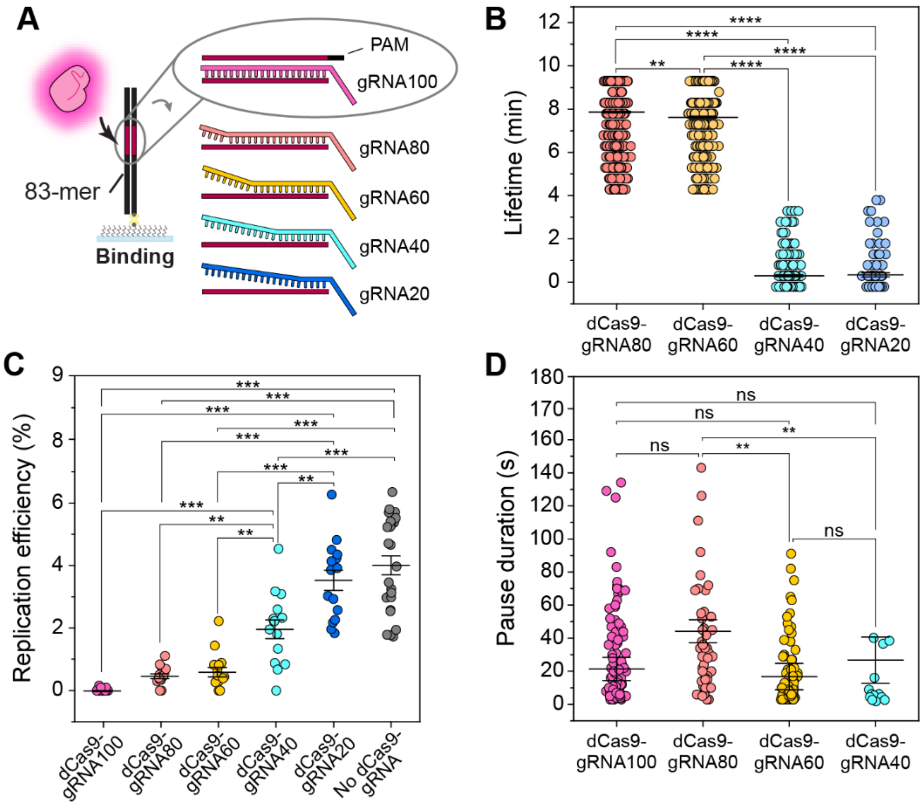


Figure 6.6. Stalled replication rescue at less stable roadblocks. A) Schematic representation of dCas9-gRNA mismatch complexes binding to 83-mer dsDNA oligonucleotides in single-molecule lifetime assays. **B)** Lifetime distributions of dCas9-MMgRNA complexes binding to 83-mer dsDNA. Mean lifetime for each complex: dCas9-gRNA80, 8.6 ± 0.1 min (S.E.M.; *n* = 1414 events); dCas9-gRNA60, 8.3 ± 0.1 min (*n* = 619 events); dCas9-gRNA40, 1.0 ± 0.1 min (*n* = 249 events); dCas9-gRNA20, 1.0 ± 0.1 min (*n* = 103 events). Comparison of distributions was conducted by the Kruskal-Wallis test for multiple comparisons with Dunn's procedure, where ** and **** denote statistical significance with p ≤ 0.01 and 0.0001, respectively. The absence of markers denotes no significant difference (p > 0.05). **C)** The efficiency of replication distributions of DNA templates following pre-incubation with dCas9-gRNA complexes. Mean efficiencies for each complex: dCas9-gRNA100, 0.02 ± 0.01%; dCas9-gRNA80, 0.5 ± 0.1%; dCas9-gRNA60, 0.6 ± 0.1%; dCas9-gRNA40, 2.0 ± 0.3%; dCas9-gRNA20, 3.5 ± 0.3%; No dCas9 complexes, 4.0 ± 0.3%. Comparison of the mean replication efficiency was done using one-way ANOVA with Tukey's comparison post-hoc test, where ** and *** denote statistical significance with p ≤ 0.01 and 0.001, respectively. The absence of markers indicates no significant difference (p > 0.05). **D)** Pause duration distributions in the presence of dCas9-MMgRNA complexes and Rep-AF647. Mean pause durations were determined by fitting a single-exponential decay function to the data of each complex: dCas9-gRNA80, (40 ± 20 s (S.E.M.), *n* = 51 pauses, replication efficiency of 3 ± 1%); dCas9-gRNA60, (20 ± 6 s, *n* = 72, 3 ± 1%); dCas9-gRNA40, (30 ± 14 s, *n* = 16, 4 ± 1%). dCas9-gRNA100 data are duplicated from Figure 6.5A (dCas9-cgRNA1 and 10 nM Rep-AF647). Pause durations represent pauses where Rep-AF647 intensity was observed above the threshold. Comparison of distributions was done by the Kruskal-Wallis test for multiple comparisons with Dunn's procedure, where ** denotes statistical significance with p ≤ 0.01 and ns denotes no significance difference (p > 0.05).

Next, we measured the efficiency of DNA replication on templates that were pre-incubated with the various dCas9-MMgRNA complexes. The 2-kbp rolling-circle DNA templates in the absence of roadblocks show a mean replication efficiency of 4% under the conditions used (Figure 6.6C). This is significantly reduced when DNA templates are pre-incubated with dCas9-cgRNA1 complexes (0.02%). Interestingly, dCas9-MMgRNA complexes containing 80 or 60% complementarity to the target exhibited a 25-fold increased efficiency of replicating products, in comparison to fully complementary complexes (gRNA80, 0.5 ± 0.1%; gRNA60, 0.6 ± 0.1 %). The replication efficiency was further increased in experiments containing dCas9-MMgRNA complexes with 40 and 20% complementarity (gRNA40, 2.0 ± 0.3%; gRNA20, 3.5 ± 0.3%). This is consistent with bulk biochemical assays where DNA replication products were only detected in reactions containing dCas9-MMgRNAs of 60–20% complementarity, and the absence of the block band at approximately 2.5 kbp, at 40 and 20% complementarity (Supplementary Figure 6.S11C).

Finally, we investigated the duration of pauses resolved by Rep-AF647 when caused by dCas9-MMgRNA complexes. Since we saw a reduced lifetime and similar efficiency to that of replication for gRNA20, we used the 80–40% MMgRNAs in these assays. As before, Rep-AF647 and the specified dCas9-MMgRNA were added to the flow cell containing DnaBC-DNA to initiate the replication reaction. Given the measured lifetimes of the complexes (Figure 6.6B), we hypothesized that a decrease in pause duration greater than two-fold would suggest that the rate-limiting step is the removal of the roadblock. However, the pauses detected and resolved by Rep-AF647 revealed similar durations despite decreasing complementarity to the target DNA (Figure 6.6D). Specifically, the pause duration corresponding to all of the MMgRNAs was similar to that of the dCas9-cgRNA1 complex (gRNA80, 40 ± 20 s; gRNA60, 20 ± 6 s; gRNA40, 30 ± 14 s,). The similar pause durations of each of the cgRNAs in the presence of RepAF647 suggest that the removal of the roadblock is not the rate-limiting process during the rescue of stalled replication. Rather, these results suggest that the activity of Rep at the stalled replication fork is quick, and that a process involved with the subsequent continuation of DNA synthesis is the rate-limiting step.

6.4. Discussion

In this study, we set out to visualize the *E. coli* Rep helicase as it interacts with elongating and stalled replisomes. Our single-molecule observations of Rep binding to DNA confirm early investigations whereby the affinity of Rep to ssDNA is significantly decreased by ATP hydrolysis. We observed frequent and stochastic association of Rep in a predominantly monomeric state to the elongating replisome as it replicates DNA. Our investigations of Rep at the stalled replisome revealed that Rep removes dCas9-cgRNA roadblocks resulting in the rescue of stalled replication. Further, we showed that the resolution of replication stalled at high-stability roadblocks occurs with kinetics that can be described with a single rate-limiting step, regardless of whether Rep was already present at the fork at the onset of the stall or whether Rep associated after the stall. Finally, we show that the duration a replisome is stalled is constant at less stable roadblocks, indicating that the rate-limiting step is a process involved in the continuation of replication and not roadblock removal. Together, these results provide insight into the activity of Rep at the replisome and allow us to propose a model describing how Rep protects the replisome and acts in the context of roadblocks.

The main aim of this study was to observe the Rep helicase at elongating and stalled replisomes. While Rep is known to interact with the replisome, the context of this interaction is not well defined. Specifically, is Rep only present at the replisome in the stalled state, or is Rep continually interacting with the replisome? We show here that Rep interacts stochastically with the elongating replisome in the absence of protein roadblocks. The addition of Rep-AF647 into single-molecule rolling-circle replication assays shows that Rep can interact with the replisome and have no effect on its rate or processivity (Figure 6.2 and Supplementary Figure 6.S5). Recent live-cell fluorescence studies proposed that Rep interacts with the replisome in low copy numbers or is only recruited to DnaB upon fork arrest [153]. Here, we show that Rep frequently interacts with the elongating replisome in a predominantly monomeric stoichiometry at all concentrations used (Figure 6.2D). However, the frequency of binding to the replisome increases with increasing concentrations of Rep, suggesting this interaction is stochastic.

The stoichiometry of Rep at the replisome has been hypothesized to be hexameric, assuming all sites are occupied on the hexameric DnaB [153, 154]. Our studies, both in the absence and presence of roadblocks, reveal a monomeric stoichiometry when Rep is associated with the replisome (Figure 6.2D and Figure 6.5E). Recent single-molecule live-cell studies observed up to six Rep monomers at the replication fork [155]. At the concentrations used in our assays, we are well below the predicted micromolar cellular concentration of Rep [155]. However, previous studies have also predicted that Rep may be present at the replisome in less than 3 copies [153]. Together, these results show the plasticity of the replisome and could suggest that while Rep could occupy all binding sites on DnaB, the likelihood of this occurring could be dependent on other DnaB interacting partners. Previous surface plasmon resonance investigations estimated the apparent K_D of the Rep-DnaB interaction to be approximately 90 nM [59]. However, these results are obtained outside the context of the functional replisome. Without structural information on the Rep-DnaB interaction, it cannot be certain that Rep does not interact in the same binding pockets as other essential components of the replisome. Therefore, we predict that the association of Rep to DnaB, and thus our reported stoichiometry, occurs not only due to the low concentrations used but also due to shared binding sites becoming available during replication.

Rep removes protein roadblocks from the path of the replication fork. In our study we observe the robust displacement of the dCas9-cgRNA roadblock in ensemble and single-molecule assays, resulting in the continuation of DNA synthesis. The dCas9-cgRNA model roadblock provides a simple alternative to using either tandem arrays of roadblocks, single RNA polymerase (RNAP) complexes or Tus-*Ter* sites to stall the replisome, which require tedious procedures to insert binding sequences [59, 65, 66, 68, 70, 71, 74, 130]. However, it is possible that the processing of stalled replication forks by Rep at tandem arrays of roadblocks varies from that of single roadblocks, for example, RNAP. Nevertheless, pre-incubation of the rolling-circle DNA template with the dCas9-cgRNA1 roadblock showed a clear continuation of replication after the disappearance of the roadblock fluorescent signal (Figure 6.3). Displacement of the dCas9-cgRNA roadblock by Rep is likely comparable to the processing of stalled forks at RNAP, where both roadblocks are stabilized by R-loop formation [68]. This activity was not observed in the presence of either the ATPase-deficient Rep K28A or DnaB interaction-deficient Rep ΔC33 mutants, thus indicating that both activities are required for the removal of protein roadblocks from the template DNA. This observation is in agreement with previous studies that showed that Rep mutants lacking these structural components could not displace RNAP or other model roadblocks [59]. Additionally, single-molecule live-cell studies showed that the C-terminus is required for the association to the replication fork, while the ATPase activity is required for translocation away from the fork [155]. Recent studies have also shown that the 2B subdomain of Rep is essential for protein roadblock displacement [68, 130]. Therefore, it is likely that the interaction with the DnaB helicase, the functional ATPase domain, and the 2B subdomain are all essential to the displacement of roadblocks by Rep.

We report here the first real-time observations of protein displacement and the rescue of stalled replication by Rep. The single-molecule rolling-circle replication assays containing both Rep-AF647 and dCas9-cgRNA1 roadblocks in solution revealed a pause duration dependent on the concentration of Rep-AF647 (Figure 6.5A). This observation suggests that the higher the local concentration of Rep, the quicker the resolution of the roadblock due to a shorter search time. Interestingly, under our conditions, we observed a plateauing of the pause duration at 10 and 20 nM Rep-AF647. Despite the predicted K_D of Rep-DnaB being much higher, this saturation suggests that the K_D of the Rep-DnaB interaction within a functional replisome might be significantly lower than previously estimated (17), potentially due to the availability of higher affinity binding sites upon replisome stalling. The predicted micromolar cellular concentration of Rep could suggest that pauses in cells are resolved quicker than in our reconstituted replisome experiments, due to the high availability of Rep to the elongating and stalled replisome. Nonetheless, our assays provide insight into the mechanisms required for Rep to displace protein roadblocks and rescue stalled replication.

We observed two distinct classes of Rep-mediated roadblock removal events: (1) where Rep is already present at the replisome upon stalling; and (2) where Rep associates after the replisome stalls (Figure 6.5B, C). While the latter activity could suggest recruitment to the stalled replisome, both activities resulted in similar pause resolution times, suggesting that once Rep is present the displacement of the roadblock occurs at the same rate through the same process. Further, we observed that the wait time for Rep to associate with stalled replisomes was less at higher concentrations of Rep (Figure 6.5D). Unlike the observed pause durations, the reported wait times did not plateau. Therefore, we hypothesize that the Rep association to the stalled replisome is independent of the affinity of the binding site and represents the concentration-dependent bimolecular association rate. Further, we predict that at more biologically relevant concentrations, Rep would interact with the replisome more frequently, further decreasing potential association wait times. The observed single-exponential distributions throughout the pause duration and pause resolution time all provide evidence that there is a single rate-limiting kinetic step governing the rescue of stalled replication once Rep associates with the replisome.

Our investigation of Rep at the sites of roadblocks with decreased stability provided further evidence that there is one rate-limiting step of stalled replication rescue. Investigations of mismatched RNA-DNA hybrids in complex with dCas9 have shown that only 8 bp of complementarity is required to establish a stable complex [394]. Interestingly, our investigations revealed comparable pause durations when the replisome was stalled by dCas9-MMgRNA complexes to fully complementary roadblocks despite the significantly lower extent of complementarity in the R-loop (Figure 6.6D). The intrinsic lifetimes of the MMgRNAs (Figure 6.6B) and the constant pause durations suggest that a slower process after the removal of the protein roadblock is the rate-liming step of the reaction. Further, these results suggest that the roadblock removal activity of Rep once associated with the stalled replisome, is relatively quick, occurring on the time scale of a few seconds.

Additionally, we detected higher replication efficiencies when DNA templates were pre-incubated with the less stable dCas9 complexes, in the absence of Rep (Figure 6.6C). While this is reflective of the observed lifetimes of these complexes, the higher efficiencies may also suggest that the replisome can bypass the less stable complexes without the need for Rep. Further investigations of the replisome, with or without Rep, at sites of stalled replication caused by unstable roadblocks will elucidate mechanisms of roadblock removal and bypass.

How does the frequent association of Rep to the replication fork result in roadblock removal? Our study allows us to propose a model of Rep activity at elongating and stalled replisomes (Figure 6.7). Our investigation provides evidence supporting a model proposed by a previous live-cell study, whereby Rep is associated with the replication fork during elongation [155]. Further, we propose that the association of Rep to the replisome is stochastic and does not occur by a recruitment mechanism. The association of Rep through the interaction with the DnaB helicase allows Rep to frequently monitor the state of the replication fork, acting as a shield to potential roadblocks (Figure 6.7A). If a roadblock is detected or encountered, then displacement activity will result (Figure 6.7B). This activity is quick, whereas the processes required to initiate the continuation of replication are relatively slow (Figure 6.7C). Given that the rescue of stalled replication in our assays did not require helicase reloading mechanisms, we hypothesize that the time required to reinitiate synthesis determines if DnaB will remain stable or if the entire replication fork will collapse. Previous ensemble *in vivo* studies of replication fork stalling have estimated that DnaB remains stable for up to 30 min after stalling [71, 73, 393].

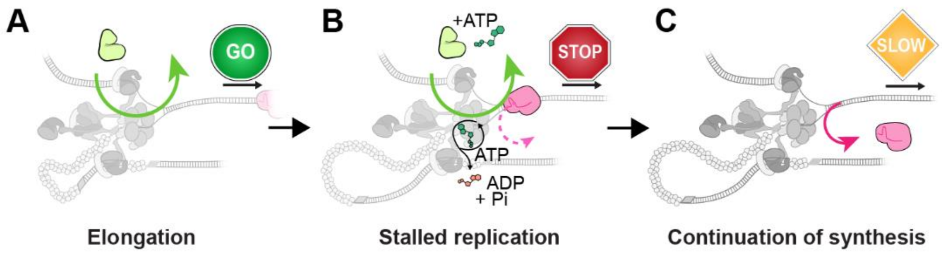


Figure 6.7: Model for Rep activity at elongating and stalled replication forks. A) Rep stochastically associates with the replication fork during elongation. **B)** Upon the replisome stalling, associated Rep molecules work to remove the roadblock from the path of the replication fork. This process is relatively quick. **C)** Continuation of replication after the removal of the roadblock is slow, representing the rate-limiting step in the resolution of the stalled state.

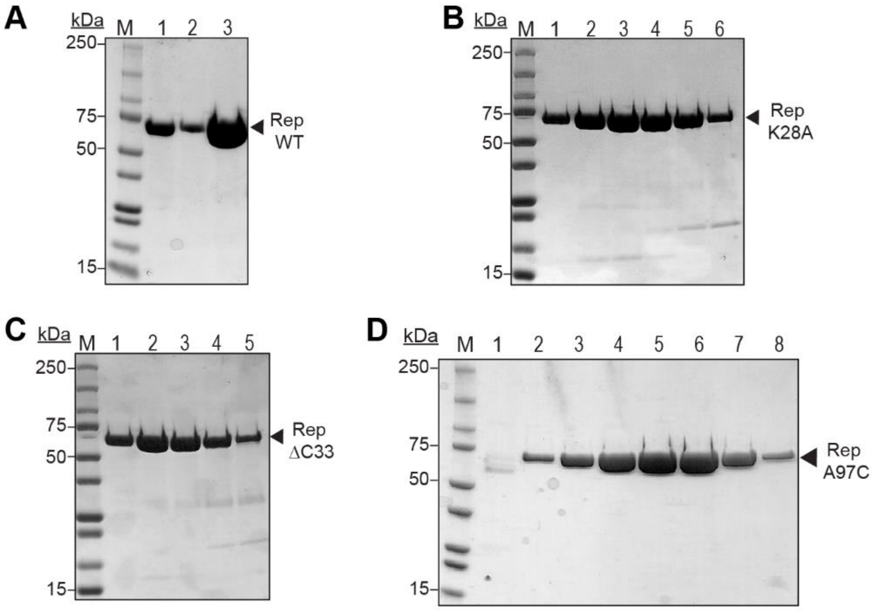
Finally, our investigations also provide insight into the stability of the DnaB helicase when the replisome encounters a protein roadblock. Each single-molecule assay described in this investigation has DnaB pre-incubated with the DNA template and omits free DnaB complexes in solution during the replication reactions. However, all other key components (Pol III holoenzyme, DnaG and SSB) are constantly present in solution throughout the duration of the experiment. The successful continuation of replication after Rep has displaced the dCas9-cgRNA roadblock, without the need for additional protein factors, suggests that the pre-incubated DnaB helicase remains bound to the template DNA after an encounter with the roadblock. The DnaB helicase has recently been visualized in similar single-molecule assays to be a stable component of a processive replisome [78]. However, other key components are able to exchange into the elongating replisome [34, 75]. While further investigations of the stability of these individual components in the stalled replisome (for example, components of the Pol III holoenzyme and DnaG) are required, these recent observations suggest that the DnaB helicase is integral for the continuation of DNA replication. A stable DnaB helicase may act as a hub for the elongating replisome, allowing for efficient reloading of replisome components if collisions result and components of the replisome dissociate.

Single-molecule observations have proven valuable in elucidating the individual behaviors of replisome components [34, 75, 78]. Our work, combined with other recent investigations, suggests a model where cooperation between Rep and the replisome is needed for efficient roadblock removal. We propose that Rep, and its homologs, fulfill their protective role for the replisome in an entirely stochastic manner that occurs independently of whether the replisome is stalled or not. Further elucidation of both the Rep-replisome and the Rep-roadblock interactions will provide insight into the significant importance of this accessory helicase to cells.

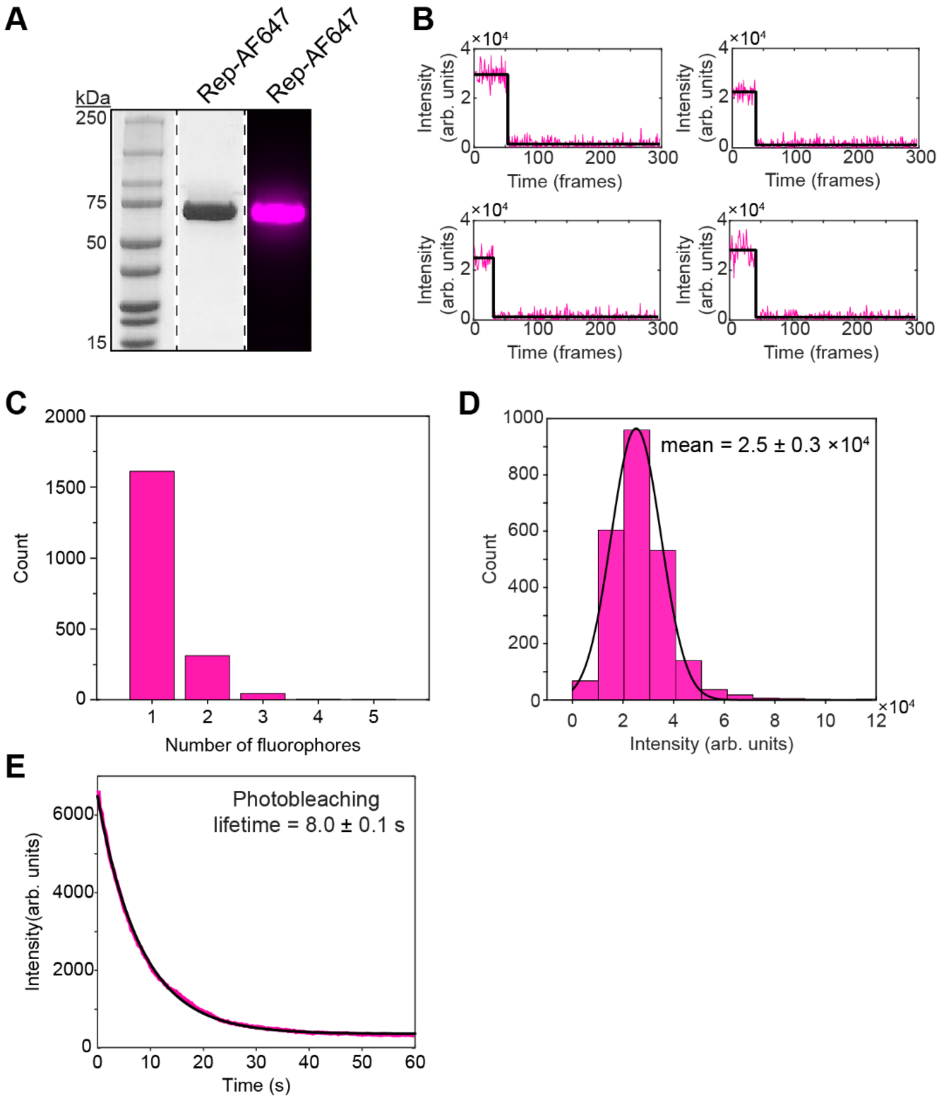
6.5. Supplementary material

Table 6.S1: Nucleic acid substrates used in this study.

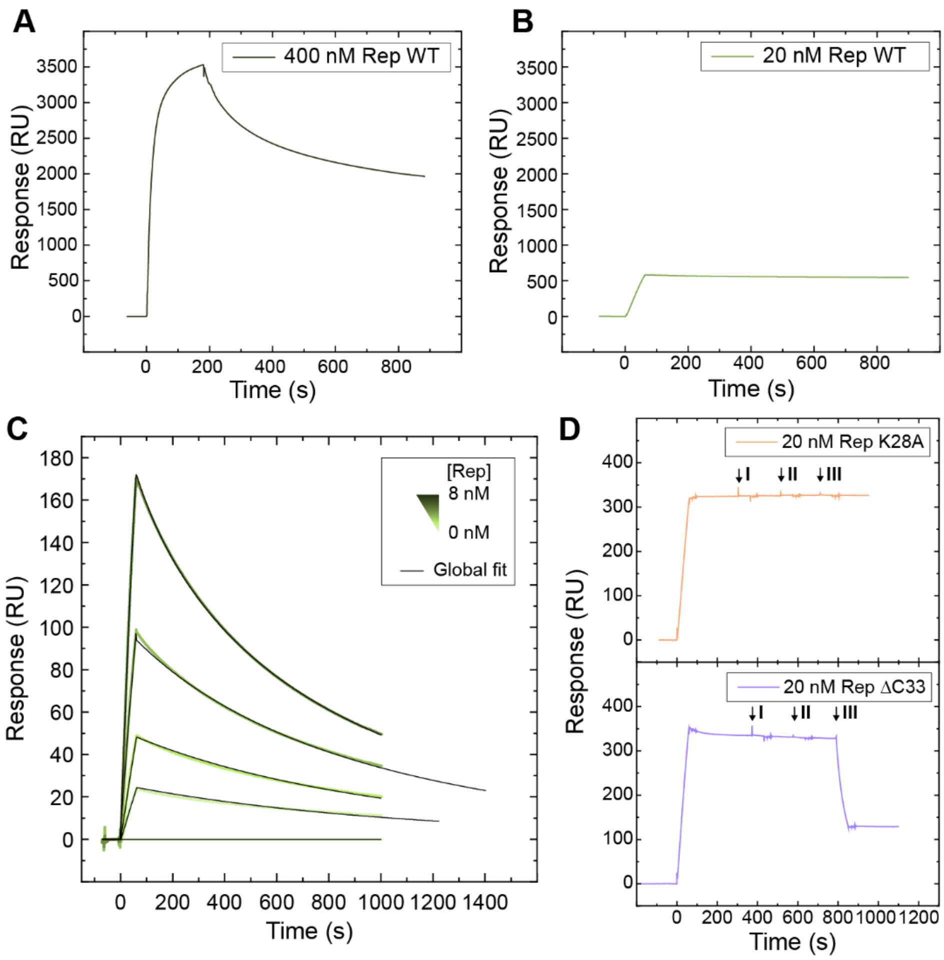
Trap dsDNA/83mer	
83_S	5’-CAC ATG CTA TGA GCT GTT GCA ATC TCT CGT ACA ATT AAT AGA CTG GAT GGT GGA TGA CAA AGC TCT ACA CTA GAT ACT CAC AC-3’
83_AS	3-/bio/-GTG TAC GAT ACT CGA CAA CGT TAG AGA GCA TGT TAA TTA TCT GAC CTA CCA CCT ACT GTT TCG AGA TGT GAT CTA TGA GTG TG-5’
gRNA	
cgRNA1	5’ - ACA AUU AAU AGA CUG GAU GG
cgRNA3	5’ - CAU UCC UGC AGC GAG UCC AU
cgRNA4	5’ - AAA CUC ACG UUA AGG GAU UU
Mismatch gRNA	
gRNA80	5’ - cac cUU AAU AGA CUG GAU GG
gRNA60	5’ - cac cgg ccU AGA CUG GAU GG
gRNA40	5’ - cac cgg ccg cuc CUG GAU GG
gRNA20	5’ - cac cgg ccg cuc agu uAU GG



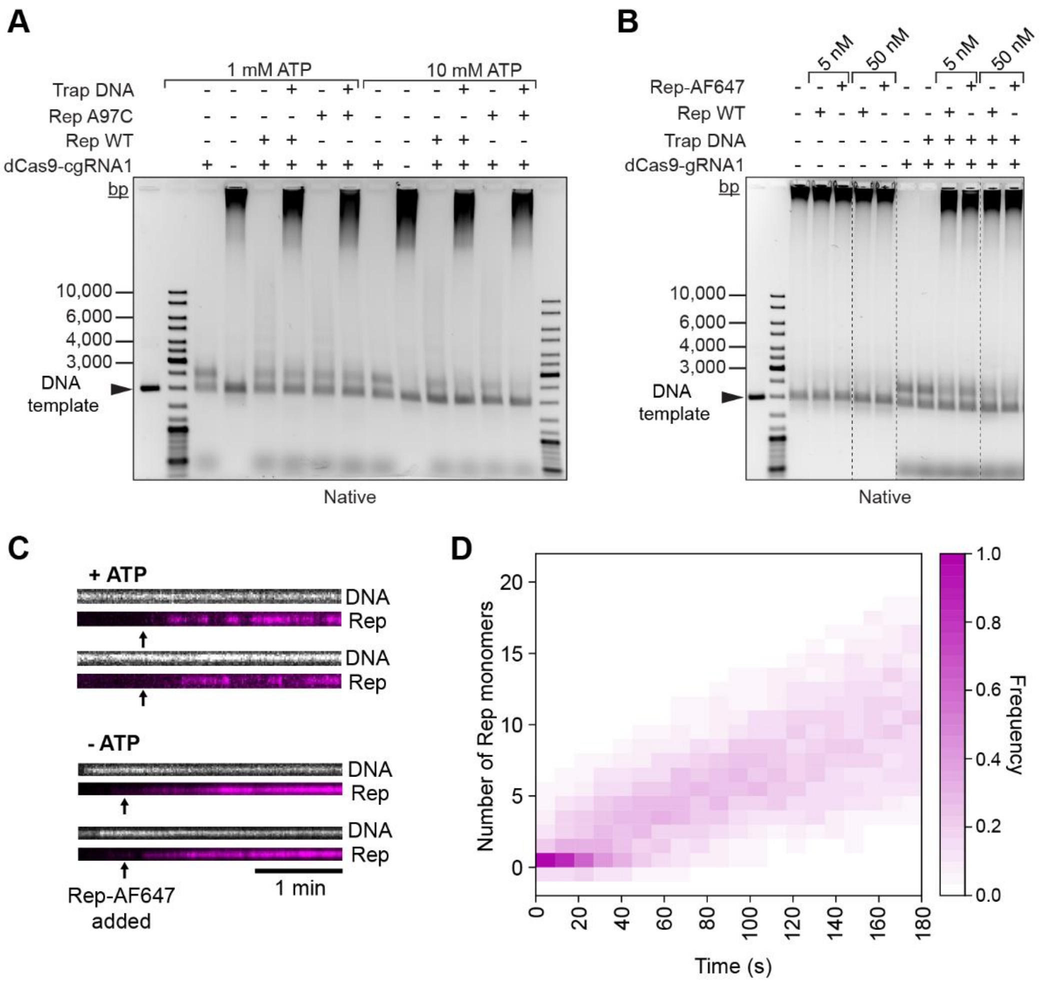
Supplementary Figure 6.S1: Purification of Rep proteins using 5 mL HiTrap Heparin columns. Samples were analyzed on 4–20% SDS-PAGE gels. **A)** Purification of His_6 Rep WT. Sample after purification on His-Trap column (lane 1), fractions of purified protein from Heparin column (lanes 2–3) used in this study. **B)** Purification of His_6 Rep K28A. Samples from successive fractions (indicated by lane numbers 1–6). **C)** Purification of His_6 Rep ΔC33. Samples from successive fractions (indicated by lane numbers 1–5). **D)** Purification of His_6 Rep A97C. Samples from column flow-through (lane 1), and successive fractions (indicated by lane numbers 2–8).



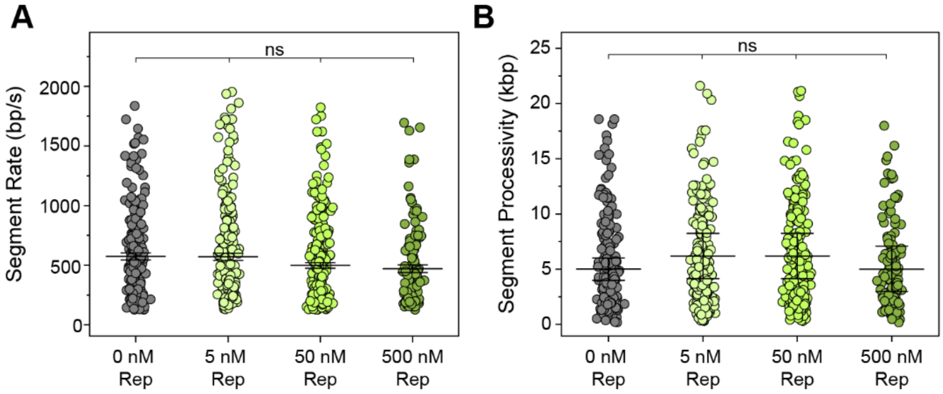
Supplementary Figure 6.S2: Quantification of fluorescent labeling of Rep-AF647. A) SDS-PAGE gel of purified Rep-AF647. Left and middle lanes are stained with Coomassie blue and imaged using a Bio-Rad Gel Doc XR. Right lane is unstained and Alexa Fluor 647 fluorescence imaged using an Amersham Imager 600. **B)** Example trajectories of Rep proteins deposited on a coverslip and subjected to photo-bleaching. The number of fluorophores per Rep monomer is detected by quantifying the single-molecule photobleaching steps using change point analysis (black line). **C)** Distribution of the number of steps and therefore the number of fluorophores per monomer (*n* = 1483). **D)** Distribution of the step size of (B) was 2.5 ± 0.3 x 10^4 (mean ± S.E.M., *n* = 1483). **E)** Average photo-bleaching trajectory for Rep-AF647 (*n* = 1483) at an excitation power density of 200 mW cm^{-2}. From a fit with a single-exponential decay function (black line), a photo-bleaching lifetime of 8.0 ± 0.1 s was obtained.



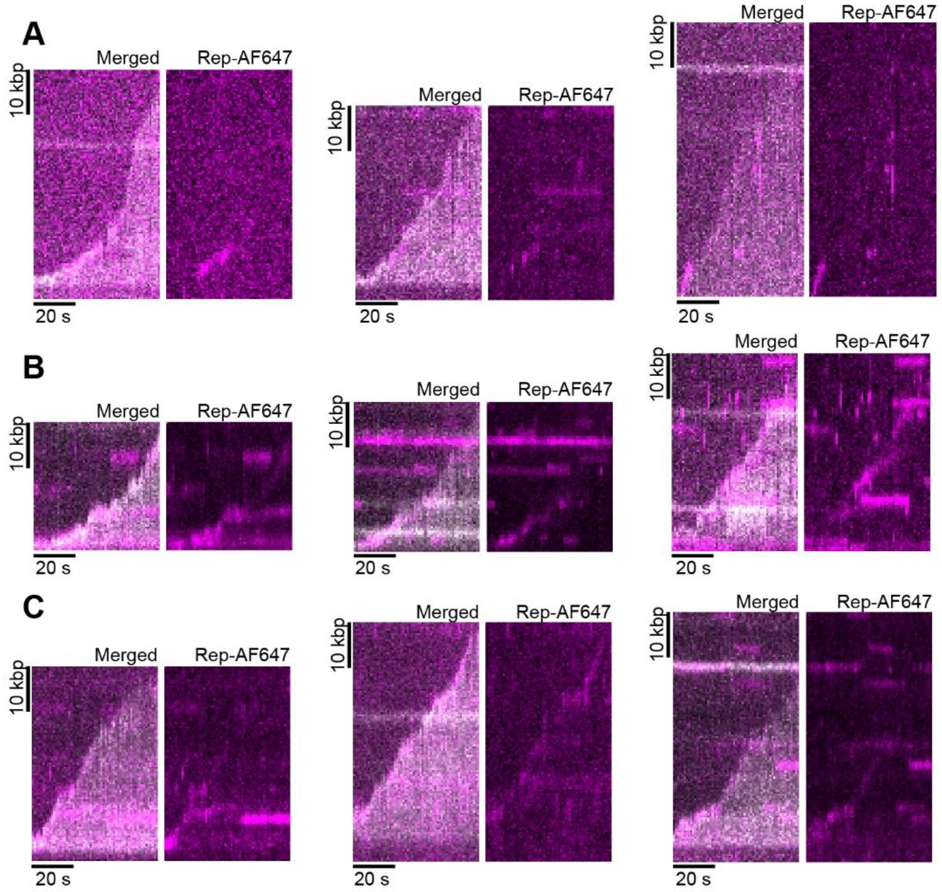
Supplementary Figure 6.S3: Observations of Rep binding to DNA. SPR sensorgrams of 400 nM **A)** and 20 nM **B)** Rep WT binding to dT_{35}. **C)** Association (60 s) and dissociation of titrated (1–8 nM) Rep WT binding to dT_{15} (global fitting of Rep binding (1:1 binding with mass transfer) shown in black yielded a K_D of approximately 500 pM). **D)** 20 nM Rep K28A (orange – top) and 20 nM Rep ΔC33 (purple – bottom) dissociation after injection of nucleotides AMP-PNP (I), ADP (II) and ATP (III).



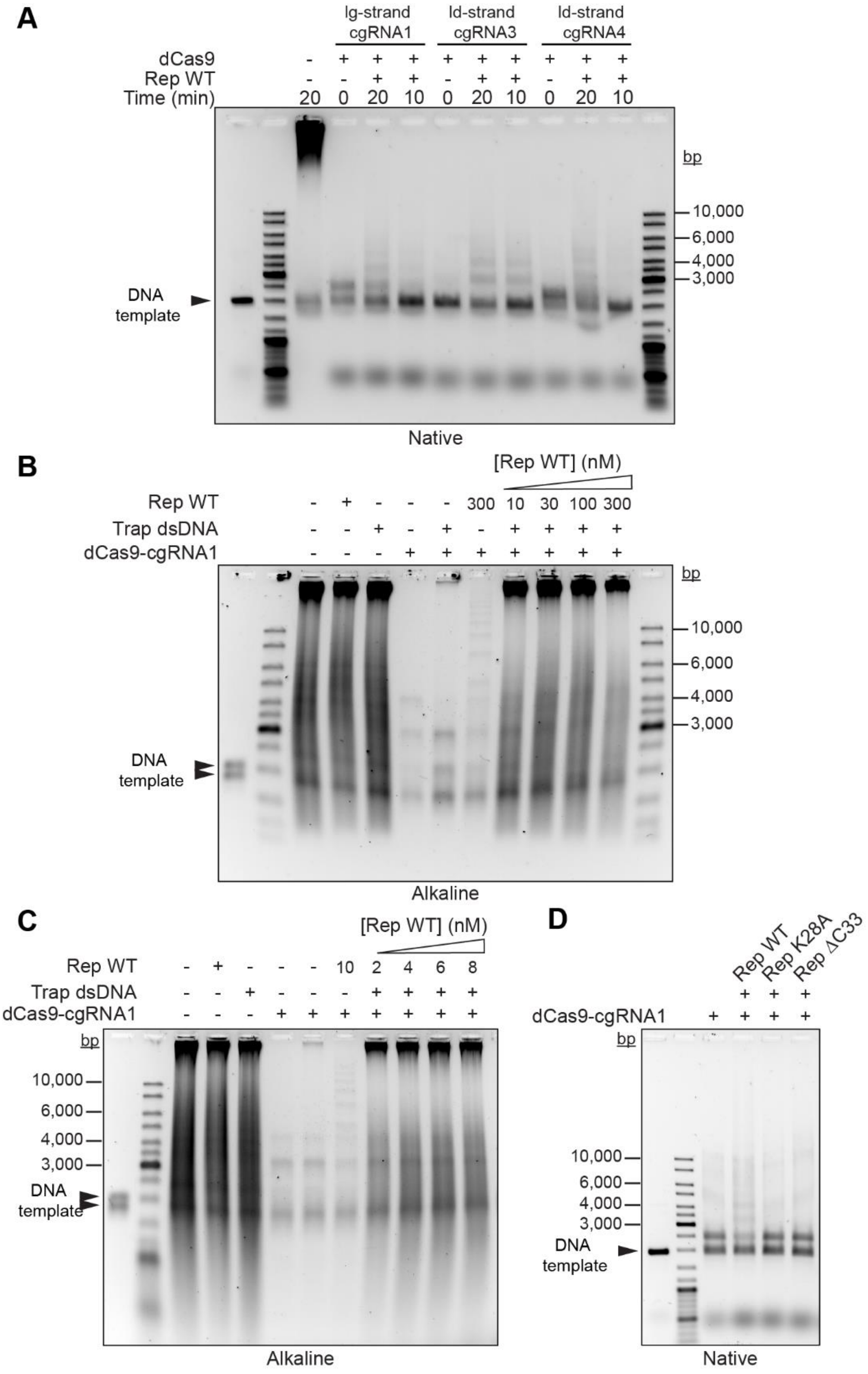
Supplementary Figure 6.S4: Ensemble replication rescue activity tests of Rep A97C and Rep-AF647. A) Rep WT and Rep A97C (50 nM) efficiently remove dCas9-cgRNA1 (50 nM) complex from DNA templates and rescue stalled DNA replication. Note that the appearance of long DNA products signifies efficient replication rescue. A 10-fold increase in ATP concentration (and 2-fold increase in $MgCl_2$ concentration) results in a greater extent of DNA replication rescue, evident by decreased intensity of DNA band at approximately 2.5 kbp. **B)** Rep-AF647 successfully removes the dCas9-cgRNA1 complex from DNA templates. Irrelevant lanes to the figure are cropped out (dashed lines). **C)** Example kymographs of Rep-AF647 binding to 2-kbp rolling-circle DNA template bound by SSB and DnaBC, in the presence (top) and absence (bottom) of ATP. Arrows indicate the time point of the addition of Rep-AF647 to the flow cell. **D)** Heatmap of the number of Rep-AF647 monomers bound to the DNA template over time in the absence of ATP (*n* = 70).



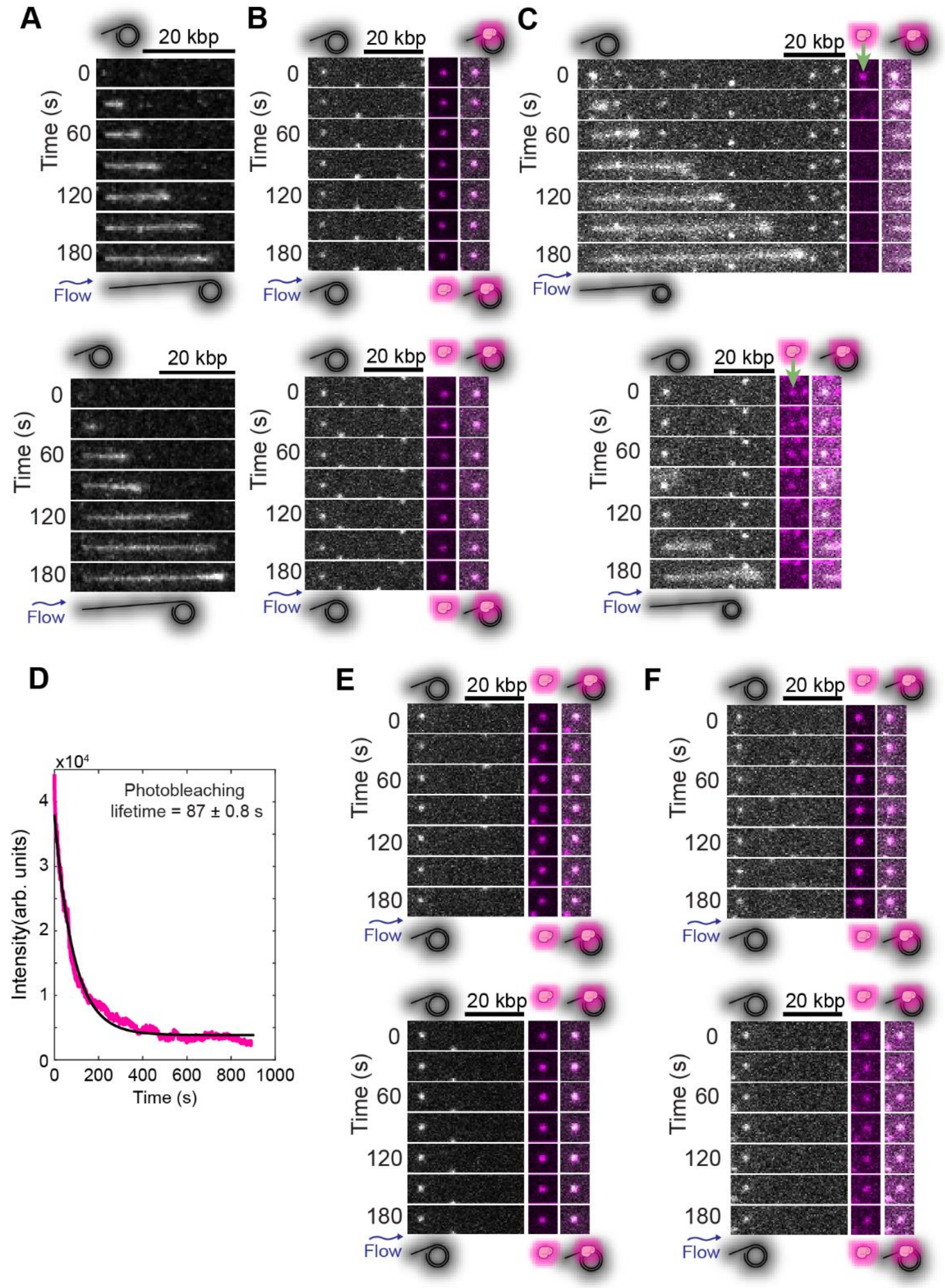
Supplementary Figure 6.S5: Quantification of the effect of Rep WT on the rate and processivity of replication. A) Median rates of replication in the absence (gray) (580 ± 30 bp s^{-1} (median ± S.E.M., *n* = 179, replication efficiency = 5 ± 1 % (S.E.M.)) and presence of titrated Rep WT (5 nM (light green) (573 ± 30 bp s^{-1} (*n* = 208, 5 ± 1%), 50 nM (olive green) (500 ± 20 bp s^{-1} (*n* = 234, 6 ± 1%), and 500 nM (dark green) (470 ± 30 bp s^{-1} (*n* = 113, 3 ± 1%)), quantified by change-point analysis of single-molecule rolling-circle DNA replication trajectories. **B)** Mean processivity of replication in the absence (gray) (5 ± 1 kbp (mean ± S.E.M.)) and presence of titrated Rep WT (5 nM (light green) (6 ± 2 kbp), 50 nM (olive green) (6 ± 2 kbp), and 500 nM (dark green) (5 ± 2 kbp)). Mean processivities were determined by fitting a single-exponential decay function to data. Comparison of distributions was conducted using the Kruskal-Wallis test for multiple comparisons with Dunn's procedure where ns indicates no significant difference ($p > 0.05$).



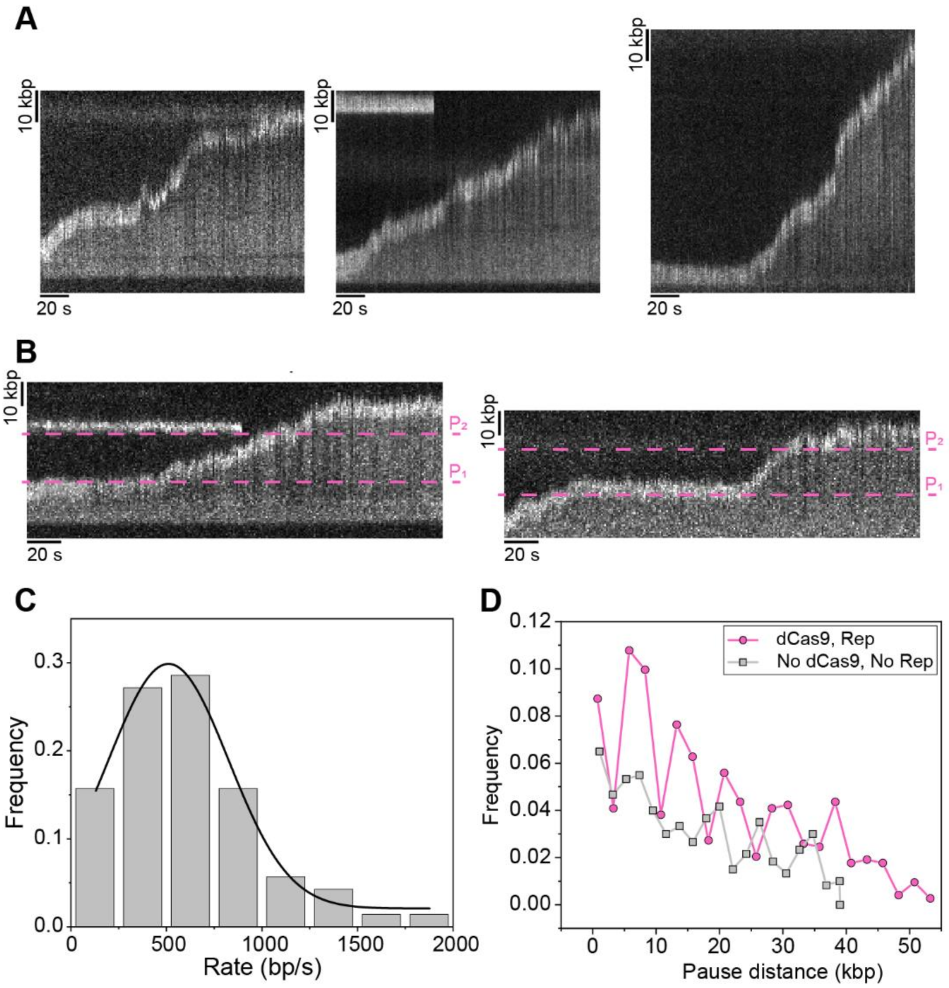
Supplementary Figure 6.S6: Example kymographs of Rep-AF647 during processive rolling-circle DNA replication. Merged kymographs of Rep-AF647 intensity (magenta) and Sytox orange-stained DNA (gray) (left) and Rep-AF647 intensity alone (right). **A)** 5 nM Rep-AF647 (replication efficiency of 7 ± 1% (S.E.M.)). **B)** 10 nM Rep-AF647 (5 ± 1%). **(C)** 20 nM Rep-AF647 (4 ± 1%).



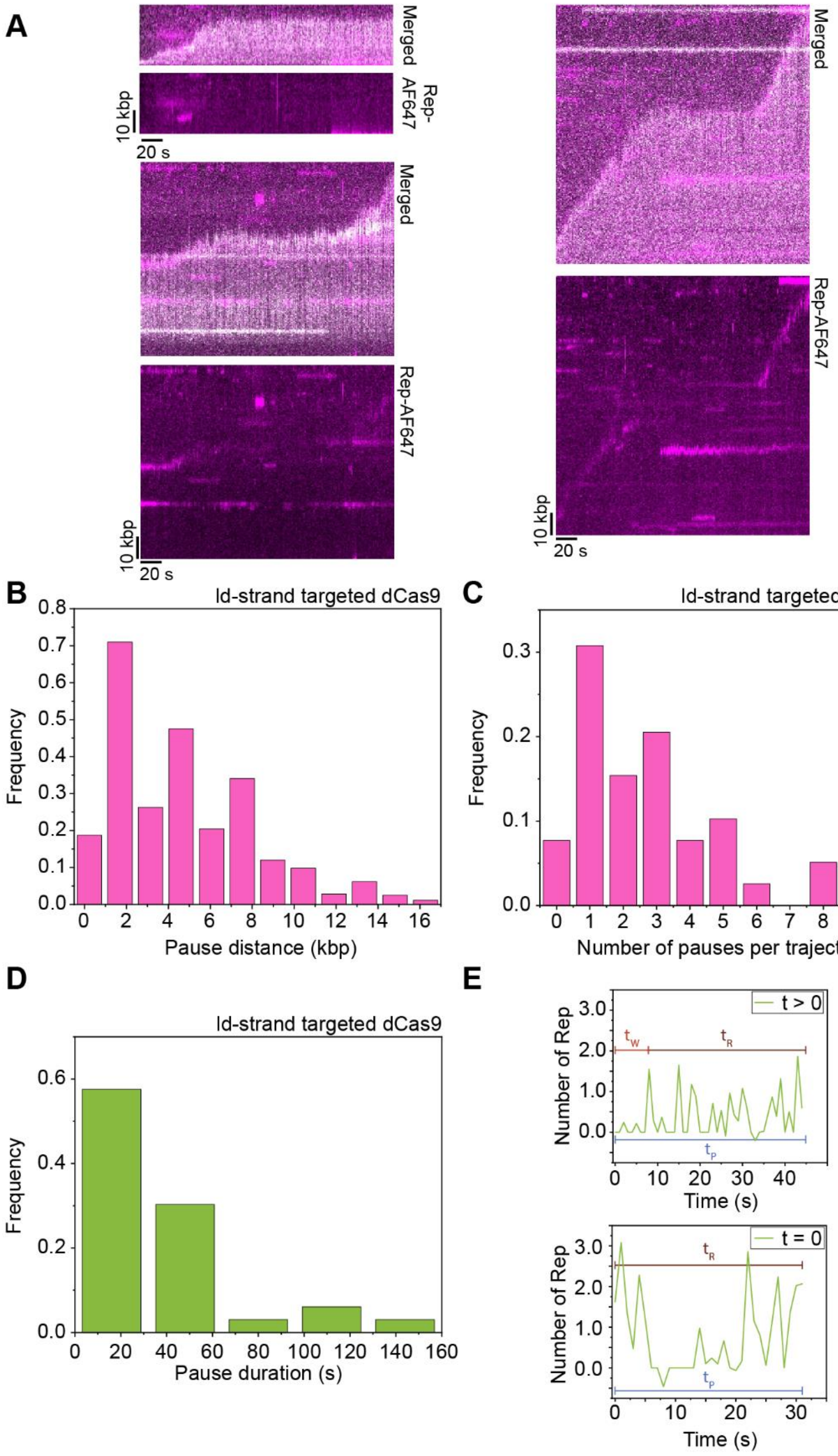
Supplementary Figure 6.S7: Rep rescues DNA replication stalled by dCas9-cgRNA complexes. Reactions contained 200 nM of specified cgRNA and 50 nM dCas9 complex. **A)** Rep WT rescues stalled DNA replication independent of the DNA strand targeted by the dCas9-cgRNA complex. Rep WT was added to reactions at the 10 min time point. *N* > 2 independent experiments. **B)** Titration of Rep WT (10–300 nM). The addition of trap dsDNA results in a higher extent of rescued DNA products. *N* > 2 independent experiments. **C)** Titration of Rep WT (2–8 nM). *N* > 2 independent experiments. **D)** Rep mutants lacking either the C-terminal domain (∆C33) or functional ATPase (K28A) cannot rescue dCas9-gRNA1 stalled DNA replication. *N* > 2 independent experiments.



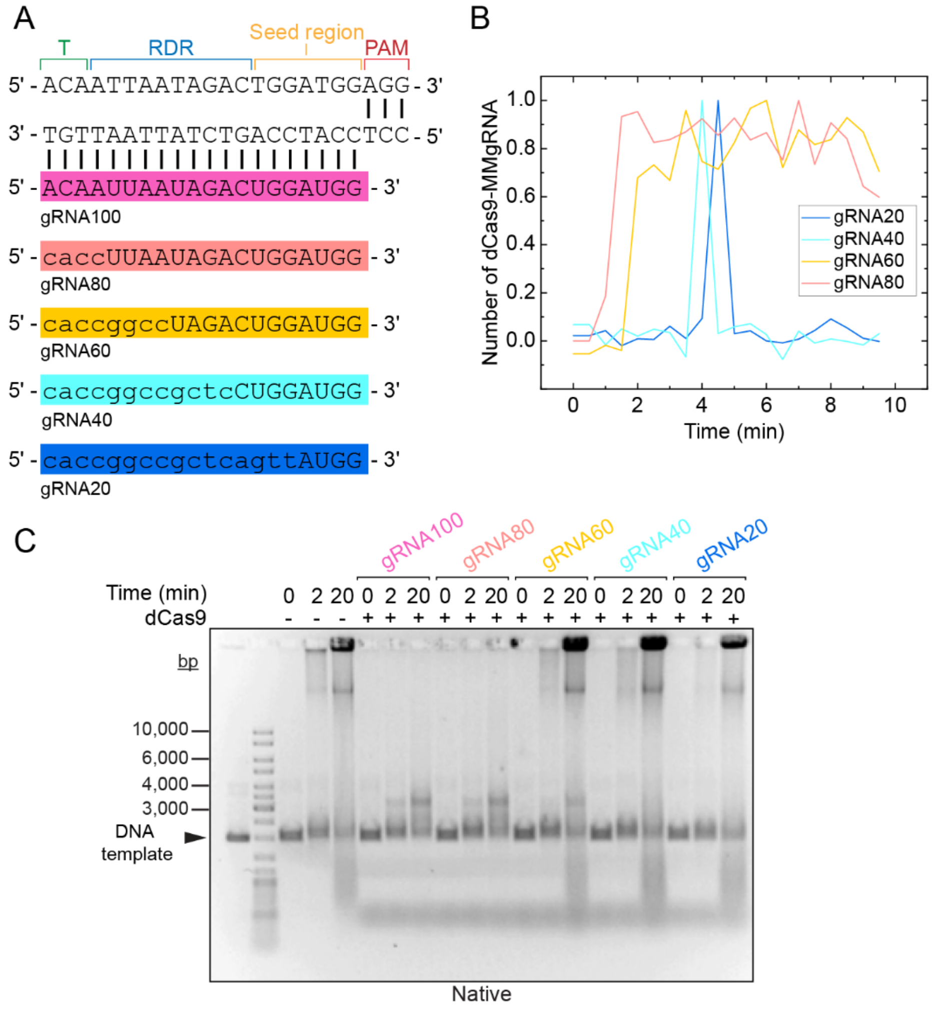
Supplementary Figure 6.S8: Additional example montages of replication rescue of DNA templates (Sytox orange stained; grayscale) pre-incubated with dCas9-cgRNA1-Atto647 (magenta) complexes. **A)** In the absence of dCas9-cgRNA1 complexes. **B)** Pre-incubation of the DNA template with dCas9-cgRNA1-Atto647. **C)** Addition of Rep WT (20 nM) results in the disappearance of dCas9-cgRNA1-Atto647 (green arrow). **D)** Average photo-bleaching trajectory for dCas9-cgRNA1-Atto647 (*n* = 139) at an excitation power density of 200 mW cm^{-2}, imaged every 200 ms. Single-exponential decay fit (black line) revealed a photo-bleaching lifetime of 87 ± 1 s. **E)** Rep K28A and **F)** Rep ΔC33 cannot remove dCas9-cgRNA1-Atto647 complexes.



Supplementary Figure 6.S9: Single-molecule rolling-circle DNA replication of an 18-kbp DNA template. A) Three example kymographs of elongating DNA replication of Sytox orange-stained 18-kbp rolling-circle DNA template (gray) (*n* = 32 molecules; replication efficiency of 7 ± 1% (S.E.M.)). The large circle of the template is resolved at the tip of the replicating molecule at a higher intensity as it is stretched out by flow. **B)** Two example kymographs of elongating DNA replication showing multiple pausing and rescue events by dCas9-cgRNA1 (0.25 nM) and Rep WT (10 nM) in solution (*n* = 26 molecules; 3 ± 1%). The dashed lines (magenta) indicate the theoretical pause start sites at approximately 17 kbp (P_1) and 36 kbp (P_2). **C)** Histograms of the rate of replication for 18-kbp rolling-circle DNA templates (520 ± 130 bp s^{-1}, *n* = 70) as in (A), fit to a Gaussian distribution. **D)** Pairwise distance analysis of the paused start sites of 18 kbp replication rescue events in the presence of 10 nM Rep-AF647 and 0.25 nM dCas9-cgRNA1 (*n* = 37 pauses/733 kbp) (magenta) and absence (*n* = 28 pauses/600 kbp) (gray) for the first 60 kbp of DNA products.



Supplementary Figure 6.S10: Single-molecule replication rescue of 2-kbp rolling-circle DNA templates. A) Three example kymographs of rolling-circle DNA replication pausing and rescue events in reactions containing dCas9-cgRNA1 (0.25 nM) and Rep-AF647 (10 nM). Merged kymographs of Sytox orange-stained DNA products (gray) and Rep-AF647 (magenta) (top), and Rep-AF647 intensity alone (bottom). (Replication efficiency of 3 ± 1%). (B–D) Assays containing leading-strand target dCas9-cgRNA4 complexes (0.25 nM) and Rep-AF647 (10 nM) show **B)** periodicity of pausing and rescue events (*n* = 43 pauses/358 kbp), **C)** the number of pausing events per replicating molecule (*n* = 40 molecules; replication efficiency of 2 ± 1%), and **D)** mean duration of pauses of 40 ± 20 s (*n* = 33 pauses). **E)** Annotated example of Rep-AF647 intensity traces over time showing the time points used for determining the pause duration (t_P – blue), association wait time (t_R – red) and pause resolve time (t_R – dark brown) for each Rep activity (*t* > 0, left; *t* = 0, right).



Supplementary Figure 6.S11: Characterization of dCas9 complexed with mismatch gRNAs. A) Designs of cgRNA1 and various mismatch gRNAs ranging from 80–20% complementarity to the target sequence of cgRNA1 (or gRNA100). Mismatches, denoted by lowercase letters, span across the DNA-RNA hybrid from the PAM distal region (or terminal region (T – green)), reversibility-determining region (RDR – blue), seed region (yellow) and to the PAM proximal region (red). **B)** Example lifetime intensity trajectories of dCas9-gRNA80 (peach), dCas9-gRNA60 (yellow), dCas9-gRNA40 (light blue) and dCas9-gRNA20 (blue). Intensity is corrected for the intensity of a single dCas9-gRNA-Atto647 molecule measured by photobleaching analysis. **C)** Time course ensemble characterization of dCas9-MMgRNA complexes blocking rolling-circle DNA replication over 20 min in the absence of Rep proteins and trap dsDNA.

Chapter 7

Discussion and Future Perspectives

DNA replication has been extensively investigated for over sixty years. From these studies, we have developed a strong knowledge of the overall composition and function of the replisome across numerous organisms during initiation, elongation and termination. Despite these impressive advances, the mechanisms that dictate the failure or success of this vital process are still unclear. Specifically, how does the elongating replisome deal with roadblocks on template DNA undergoing replication: Does the replisome stall, remain intact, and continue synthesis upon removal of the roadblock? Alternatively, does the replisome collapse upon colliding with roadblocks? Further, how do other accessory proteins function at the replisome to ensure successful replication and maintain genomic integrity?

Collisions of the replisome with protein roadblocks on DNA can have significant effects on the chromosome in all organisms. Prokaryotic systems serve as a simple alternative for investigating conserved mechanisms of the replisome and interacting proteins. Whilst the overall composition of the replisome across organisms may vary, each involves proteins with conserved mechanisms and roles. This thesis highlights the significant contributions single-molecule techniques have made in detecting intricate mechanisms of helicases and translocases involved in numerous DNA metabolism pathways (*Chapter 2*). Additionally, these techniques have been applied to characterize protein-DNA interactions of the CRISPR/Cas9 system providing essential knowledge for adaptation into various tools (*Chapter 3*).

Continued advancement of single-molecule fluorescence microscopy methods will prove vital in understanding the fine mechanistic details of biological processes. This thesis has expanded the single-molecule toolkit available to the field to investigate mechanisms of DNA replication and stalled replication rescue *in vitro*. Specifically, I further established automated data analysis protocols to detect protein interactions with replication forks in well-established single-molecule fluorescence assays (*Chapter 4*). Additionally, I established methods to stall the replisome using site-specific and stable dCas9-cgRNA roadblocks on DNA templates of increased spatial resolution (*Chapter 5*). Together, these techniques were applied to characterize the activity of the Rep helicase at the *E. coli* replisome (*Chapter 6*). For the first time, I was able to visualize stochastic interactions of the Rep helicase with the *E. coli* replisome during replication elongation and successful rescue of stalled replication *in vitro*. These studies have provided insight into the stability of the replisome at single protein roadblocks and have revealed a cooperative mechanism of an accessory helicase with the replisome to ensure successful replication in the presence of roadblocks.

7.1. The importance of studying single-molecules

Single-molecule techniques have had a significant impact on the understanding of protein-DNA interactions and their functions (*Chapters 2 and 3*) (reviewed in [395-398]). In the work described in this thesis, I expanded upon specialized methods (see Appendix I) to stall DNA replication to investigate how reconstituted *E. coli* replisomes deal with and overcome protein roadblocks. Specifically, I developed various assays to visualize stalled DNA replication in real-time using well-established 2-kb rolling-circle DNA replication templates (*Chapter 5*). Additionally, an 18-kb DNA template was designed and used in these assays to provide increased spatial resolution. Further, I established an automated analysis method to accurately and unbiasedly detect and characterize rolling-circle DNA replication and interacting fluorescently-labeled proteins (*Chapter 4*). The methods described in these chapters build upon previous methods to produce an entirely novel approach to characterizing stalled DNA replication rescue.

The analysis of single-molecule microscopy images is an exciting area of this field. With most assays, early analysis methods involved hand-picking molecules of interest and detecting sub-populations by eye. The reliance on the individual user to detect and characterize these characteristics introduces an implicit bias that can vary between users. In Chapter 4, I optimized analysis methods of rolling-circle DNA replication to result in an automated detection of DNA replication rates, processivities and the characterization of interacting fluorescent proteins. Further development of these automated methods through machine learning tools, specifically to detect replicating versus non-replicating rolling-circle DNA templates, will streamline the analysis for users while also decreasing any potential bias introduced when choosing molecules of interest.

The spatial resolution of single-molecule fluorescence techniques, specifically for investigating DNA replication, has been an aspect gaining significant interest. The development of new assays involving linear DNA substrates has revealed exciting insight into how the replisome may deal with and resolve various roadblocks that are faced within cells [156, 365, 366]. However, the limitations to these assays are significant, specifically in the construction of the templates themselves and the single-turnover aspect of visualizing only one replisome-roadblock collision per DNA template. In Chapter 5, I describe the development, construction and use of an 18-kb rolling-circle DNA template to visualize stalled DNA replication. This technique improves the established 2-kb DNA template spatial resolution by 9-fold. Further, this DNA template increases the visualization of well-defined replisome-protein roadblock collisions to multiple per DNA template, an observation not possible with established linear DNA templates which only accommodate one replisome-protein roadblock collision event per DNA molecule [156, 365, 366]. The overall structure of this DNA template would allow for easy adaptation to investigate biologically relevant protein roadblocks, for example, RNAP or Tus-*ter* sites, or other DNA secondary structures faced by the replisome in cells.

7.2. The fate of the stalled replisome

This thesis set out to expand and implement dCas9 roadblock assays to investigate how the replisome and accessory helicases resolve stalled DNA replication. As described in detail in Appendix I, the dCas9 roadblock is site-specific, stable and easy to use. In Chapter 5, I further developed single-molecule fluorescence microscopy assays to visualize stalled replication using the dCas9 tool. The two assays described rely on simultaneous imaging of fluorescently-stained DNA and fluorescently-labeled dCas9-cgRNA complexes. This imaging regime allows for the detection of blocked DNA replication (DNA imaging) and the binding site of the roadblock (dCas9-cgRNA imaging). These assays were then implemented in Chapter 6 to show the successful rescue of stalled DNA replication by Rep.

DNA replication conflicts often occur at transcription sites and result in major genome instability. The simultaneous occurrence of replication and transcription, and the faster speed of the replisome compared to RNAP, can result in conflicts of two different orientations. Co-directional and head-on collisions can have detrimental effects on the bacterial genome ultimately resulting in cell death. Despite extensive *in vivo* and *in vitro* ensemble investigations, how the bacterial replisome resolves these conflicts remains to be fully understood [10, 58, 60, 68, 72, 73, 399-407]. Interestingly, the RNA-DNA R-loop formed by the dCas9-cgRNA-DNA complex mimics the R-loop formed during transcription. The methods developed throughout this thesis place the field in a position to resolve replisome collisions with the biologically relevant roadblock, RNAP, *in vitro* at the single-molecule level. Further, this thesis offers a foundation for building a mechanistic understanding of the stability of the replisome at transcription sites and the resolution of replication-transcription conflicts.

Investigating stalled DNA replication rescue by Rep revealed interesting insights into the stability of replisome components following a collision (*Chapter 6*). All single-molecule assays described in this thesis involved the pre-incubation of the DnaBC complex with DNA templates and subsequent omission from the buffer solution during the replication reactions, where all other components were constantly available in solution. Observations of the continuation of replication following the removal of the dCas9 roadblock by Rep suggest that the DnaB helicase remains stable on the DNA after roadblock collision. Previous studies have shown that the DnaB helicase is a stable component of the replisome on DNA throughout elongation and can remain stable after colliding with a roadblock on the order of tens of minutes [65, 68, 72, 73, 78]. Further, using dCas9-gRNA roadblocks of decreasing stability revealed that the rate-limiting step of the rescue of stalled replication is likely a process involved in the continuation of synthesis (*Chapter 6*). Given that the DnaB helicase remains stable, this suggests that other vital components of the replisome, for example, the Pol III holoenzyme or DnaG, might dissociate and the re-association of these components determines when replication will continue. Further investigations into the stability of individual replisome components following replisome-roadblock collisions, will provide further insight into the composition of the stalled replisome.

7.3. The significance of accessory helicases to cells

Accessory helicases play an important role in reducing the occurrence of deleterious conflicts with roadblocks, including replication-transcription conflicts. The work described in this thesis sought to better understand the role of the Rep helicase at stalled replisomes at roadblocks that mimic transcription complexes. It is well-established that Rep has the ability to remove protein roadblocks and rescue stalled replication. However, the link between this activity and the structural components of the protein is not clear. Previous studies suggest that that the displacement of protein roadblocks requires cooperation between the ATPase domains, rotation of the 2B subdomain and association with the replisome through the C-terminus [59, 68, 130, 141, 155]. Further, it has been hypothesized that the association of Rep to the DnaB helicase is via a recruitment mechanism after the replisome collides with a roadblock [60, 153, 155]. Together these hypotheses paint a very specific and sequential mechanism of roadblock removal by Rep.

Visualizing replication rescue by Rep at the single-molecule level, we challenge the idea of recruitment by showing that Rep stochastically associates with the replisome (*Chapter 6*). This activity occurs independently of whether the replisome is in a stalled state or not. The stochastic association of Rep to the replisome suggests a simpler model of replication rescue. Frequent monitoring of the state of the replisome likely results in quick activation of the roadblock removal activity only when the replisome is in a stalled state. How Rep detects whether the replisome is elongating versus stalled remains to be characterized. Nonetheless, our visualizations of replication rescue show that the activity of Rep is quick so as to reduce the time the replisome is stalled.

It is important to consider how the activity of Rep may differ from that of prokaryotic structural homologs, UvrD and PcrA. While UvrD and PcrA can remove protein roadblocks and rescue stalled replication, no known interactions with replisome components exist [59, 408]. Further, it is well-established that the activity of Rep, UvrD and PcrA is regulated by the rotation of the 2B subdomain and stoichiometry (see section 2.3.1) [134, 139, 140, 142-149]. Despite striking similarities, one question stands out: How are the activities of Rep and UvrD in *E. coli* separated? UvrD, but not Rep, can remove RecA recombinase filaments [101, 157, 161]. Rep, but not UvrD, can unwind guanine-quadruplex structures [102]. And both helicases can remove RNAP from the path of the replication fork [68]. The clear dependence on the C-terminus of Rep determining the success of roadblock removal suggests that localization of the protein separates its activity from that of UvrD. The predicted micromolar cellular concentrations of each protein per *E. coli* cell suggest a mechanism involving localization to targets is required to prevent the over-activation of the proteins [155, 409, 410]. UvrD and PcrA have been shown to associate with RNAP and have a strong affinity to RecA recombinase filaments [158, 192, 287, 408, 411]. Therefore, it is likely that Rep is a replisome-specific accessory helicase, while UvrD is a general roadblock displacement helicase across replication, repair and recombination.

Despite decades worth of characterization, numerous questions remain about Rep and its homologs in roadblock removal activity. Specifically, how does Rep remove highly stable protein roadblocks? Eukaryotic Rep homolog, Pif1, has been shown to preferentially unwind RNA-DNA hybrids over DNA-DNA hybrids to displace telomerase holoenzymes [377, 412]. Early studies also observed this unwinding activity for UvrD, but not Rep [413]. Further investigation is required to establish if this is a conserved mechanism of SF1A helicases linking unwinding and roadblock displacement, or if another mechanism exists for Rep.

Perhaps another mechanism used by Rep is an exertion of force on the roadblock through helicase or translocase activity. Previous studies have suggested that translocation of the monomer and collision with a roadblock could result in 2B subdomain rotation and activation of helicase activity. Pushing of high-stability roadblocks along DNA has been observed for numerous helicases including bacteriophage T4 helicase Dda, and *E. coli* RecBCD [277, 278, 414]. Interestingly, single-molecule FRET investigations revealed Pif1, UvrD and Rep translocases can push *E. coli* SSB along ssDNA [415]. These observations suggest that these helicases and translocases can generate forces high enough to push and subsequently displace high-stability protein roadblocks. Interestingly, each of the helicases mentioned here are superfamily 1 helicases, suggesting this may be a conserved mechanism of protein roadblock displacement.

The importance of the Rep-DnaB interaction has been highlighted throughout this thesis. Despite numerous studies providing evidence that this is attributed to the Rep C-terminus, structural information about this interaction remains absent [59, 62]. DnaB has known interactions with multiple proteins, including elongating replisome components; the τ subunit of the clamp loader complex, and the DnaG primase [126, 127]. Given these interactions, we hypothesize that the monomeric stoichiometry of Rep at the replisome reported in this thesis, is a result of potentially shared binding pockets with other DnaB-interacting proteins. Further, these results suggest that Rep associates with DnaB as a monomer, likely by a DnaB monomer-Rep monomer interaction.

What is the stoichiometry of Rep during translocation, helicase unwinding and roadblock removal? It is well-established that the dimer is the active helicase, while the monomer displays no processive dsDNA unwinding activity (see section 2.3.1) [134, 143, 146]. However, cross-linking of the closed state or deletion of the 2B subdomain activates monomeric helicase activity [138, 141, 142, 146]. Additionally, the results presented in this thesis show that it is predominantly a monomer that associates with the replisome during elongation and stalled states (see sections 6.3.2 and 6.3.5). Further, we see no recruitment of additional monomers at the stalled replisome suggesting that a dimer, or higher order oligomer, is not required to remove the roadblock. As discussed, current models suggest that it is the monomer that translocates ahead of the replication fork and activates helicase activity through the rotation of the 2B subdomain upon colliding with the roadblock [141, 155]. Therefore, the 2B subdomain may dictate the stoichiometry of Rep. While Rep can dimerize, this state may not be required for all activities. It would be interesting to see how the different oligomeric forms and structural conformations of Rep, and homologs, determine the success of replication rescue activities.

Finally, given the vital role Rep plays in maintaining genome integrity in cells, it is tempting to suggest that Rep is not an accessory to the replisome. Bulk ensemble averaging experiments have perhaps unknowingly omitted Rep from the reconstitution of replisomes, likely due to the absence of protein roadblocks. However, *in vivo* experiments have highlighted the lethality of *rep uvrD* double mutants, while single *rep* and *uvrD* mutants are viable, suggesting a mechanism of roadblock removal is essential. Our observations of Rep at the replisome in the absence of roadblocks are the first in a reconstituted *in vitro* system and suggest that it is an exchangeable component of the replisome during elongation. That is to suggest that like other exchangeable components, Rep is always present at the replisome to remove roadblocks frequently encountered in cells.

7.4. Future perspectives

This thesis highlights both the importance of using single-molecule techniques and the complexity of stalled DNA replication rescue. Prior to this study, it was hypothesized that Rep was only recruited to the replisome upon encountering a roadblock [60, 153, 155]. However, the single-molecule methods developed and described in this thesis reveal two behaviors of the *E. coli* Rep helicase at the replisome that challenge previous hypotheses: First, that Rep is a stochastic binding partner of the replisome and acts quickly to prevent replication stalling by frequent interactions. Second, the rate-limiting step of the stalled replication rescue reaction is a process involved in the continuation of synthesis.

Activation of the roadblock-removal activity by Rep remains to be understood, however, it is possible that major differences between the composition of an elongating and stalled replisome may be the key. Whilst this potential mechanism may appear simple in the model *E. coli* replisome, how eukaryotic replisomes overcome protein roadblocks is likely to be more complex. Nonetheless, the known eukaryotic Rep functional homolog, Pif1, has been shown to have conserved mechanisms of replisome-dependent protein roadblock removal [156, 376]. Therefore, it is likely that continuous monitoring of the eukaryotic replisome is also achieved through frequent association with important eukaryotic replisome components.

While this thesis focuses on *in vitro* characterization of stalled replication rescue, these assays have revealed important behaviors of Rep at the replisome. Recent studies have used bulk ensemble assays using reconstituted proteins or live-cell single-molecule studies to characterize stalled replication rescue [68, 155]. These studies have contributed vital information to the mechanistic understanding of accessory helicases in resolving stalled replication. However, the results presented in this thesis have highlighted the importance of stripping back the assays to watch these mechanisms closer in real time. By removing roadblocks from the reaction, we observed stochastic binding of Rep to the replisome during elongation – a behavior not seen previously. Continuing to combine bulk ensemble assays with both *in vitro* and *in vivo* single-molecule assays can further elucidate the mechanisms of stalled replication rescue in prokaryotes and eukaryotes.

Overall, this thesis has highlighted the importance of characterizing the types of interactions that occur in multi-protein complexes. Using the *E. coli* replisome as an example, the extensive single-molecule fluorescence investigations over recent years have revealed a picture of the replisome that challenges the decades-old stable replisome established by ensemble techniques. Instead, the emerging picture is that replisome components are frequently exchanging during elongation while the DnaB helicase remains stable, likely acting as a binding hub throughout the reaction. These interactions are well-understood and proposed models of the replisome continue to gain evidence. Additionally, the work presented in this thesis suggests that Rep may be a constant component of the replisome, functionally required in cells to remove protein roadblocks. We are in an exciting stage of DNA replication investigations where the combination of Cryo-EM and innovative single-molecule techniques could one day provide a whole *E. coli* replisome structure. Translation of these techniques to visualize the replisome and components following collisions with protein roadblocks could ultimately change the pathway recent replication and replication-roadblock investigations have set.

This thesis focuses on the *E. coli* replisome and how it works to resolve protein roadblocks. Despite this system being relatively simple compared to that of eukaryotic systems, this work highlights the depth of knowledge that is still yet to be learned from simpler systems such as *E. coli*. In addition to addressing open questions in the prokaryotic system, continued use of these simple systems can also provide us with a fundamental understanding of how multi-protein complexes in eukaryotic systems may function. Characterization of the intricate behaviors of these complexes across species can inform future investigations of antimicrobial resistance, disease occurrence and disease treatment.

Chapter 8

References

1. Franklin, R.E. and R.G. Gosling, *Molecular Configuration in Sodium Thymonucleate.* Nature, 1953. **171**(4356): p. 740-741.
2. Watson, J.D. and F.H.C. Crick, *Molecular structure of nucleic acids: A structure for deoxyribose nucleic acid.* Nature, 1953. **171**(4356): p. 737-738.
3. Meselson, M. and F.W. Stahl, *The replication of DNA in Escherichia coli* Proc. Natl. Acad. Sci. U.S.A., 1958. **44**: p. 671-682.
4. Cairns, J., *The bacterial chromosome and its manner of replication as seen by autoradiography.* J. Mol. Biol., 1963. **6**(3): p. 208-213.
5. Bleecken, S., *"Replisome"-controlled initiation of DNA replication.* J. Theor. Biol., 1971. **32**(1): p. 81-92.
6. McHenry, C.S. and A. Kornberg, *DNA polymerase III holoenzyme of Escherichia coli. Purification and resolution into subunits.* J. Biol. Chem., 1977. **252**(18): p. 6478-6484.
7. Wu, C.A., et al., *Coordinated leading- and lagging-strand synthesis at the Escherichia coli DNA replication fork.* J. Biol. Chem., 1992. **267**(6): p. 4064-4073.
8. Okazaki, R., et al., *Mechanism of DNA chain growth, I. Possible discontinuity and unusual secondary structure of newly synthesized chains.* Proc. Natl. Acad. Sci. U.S.A., 1968. **59** p. 598-605.
9. Sinha, N.K., C.F. Morris, and B.M. Alberts, *Efficient in vitro replication of double-stranded DNA templates by a purified T4 bacteriophage replication system* J. Biol. Chem., 1980. **255**(9): p. 4290-4303.
10. Mirkin, E.V. and S.M. Mirkin, *Mechanisms of transcription-replication collisions in bacteria.* Mol. Cell Biol., 2005. **25**(3): p. 888-95.
11. Gupta, M.K., et al., *Protein-DNA complexes are the primary sources of replication fork pausing in Escherichia coli.* Proc. Natl. Acad. Sci. U.S.A., 2013. **110**(18): p. 7252-7257.
12. Blázqúez, J., J. Ródríguez-Beltrán, and I. Matic, *Antibiotic-Induced Genetic Variation: How It Arises and How It Can Be Prevented.* Annu. Rev. Microbiol., 2018. **72**: p. 209-230.
13. Ragheb, M.N., et al., *Inhibiting the evolution of antibiotic resistance.* Mol. Cell, 2019. **73**(1): p. 157-165 e5.
14. Preston, B.D., T.M. Albertson, and A.J. Herr, *DNA replication fidelity and cancer.* Semin. Cancer Biol., 2010. **20**(5): p. 281-93.
15. Jeong, Y.T., et al., *FBH1 protects melanocytes from transformation and is deregulated in melanomas.* Cell Cycle, 2013. **12**(7): p. 1128-32.
16. Xu, J., et al., *Structural basis for the initiation of eukaryotic transcription-coupled DNA repair.* Nature, 2017. **551**(7682): p. 653-657.
17. Kuper, J. and C. Kisker, *DNA helicases in NER, BER, and MMR.* Adv. Exp. Med. Biol., 2013. **767**: p. 203-24.
18. Sharma, S., K.M. Doherty, and R.M. Brosh, Jr., *Mechanisms of RecQ helicases in pathways of DNA metabolism and maintenance of genomic stability.* Biochem. J., 2006. **398**(3): p. 319-37.
19. Lewis, J.S., S. Jergic, and N.E. Dixon, *The E. coli DNA replication fork*, in *Enzymes*, L. Kaguni, S. and M. Oliveira, Editors. 2016, Elsevier Inc. : USA. p. 31-88.
20. Maki, S. and A. Kornberg, *DNA polymerase III holoenzyme of Escherichia coli. II. A novel complex including the γ subunit essential for processive synthesis.* J. Biol. Chem., 1988. **263**(14): p. 6555-6560.

21. Pham, T.M., et al., *A single-molecule approach to DNA replication in Escherichia coli cells demonstrated that DNA polymerase III is a major determinant of fork speed.* Mol. Microbiol., 2013. **90**(3): p. 584-96.
22. Jergic, S., et al., *A direct proofreader-clamp interaction stabilizes the Pol III replicase in the polymerization mode.* EMBO J., 2013. **32**(9): p. 1322-33.
23. Elshenawy, M.M., et al., *Replisome speed determines the efficiency of the Tus-Ter replication termination barrier.* Nature, 2015. **525**(7569): p. 394-8.
24. Bramhill, D. and A. Kornberg, *Duplex opening by DnaA protein at novel sequences in initiation of replication at the origin of the E. coli chromosome* Cell 1988. **52**: p. 743-755.
25. Braun, R., K. O'Day, and A. Wright, *Cloning and characterization of dnaA(Cs), a mutation which leads to overinitiation of DNA replication in Escherichia coli K-12.* J. Bacteriol., 1987. **169**(9): p. 3898-3903.
26. Zahn, G. and W. Messer, *Control of the initiation of DNA replication in Escherichia coli.* Mol. Gen. Genet., 1979. **168**(2): p. 197-209.
27. Arias-Palomo, E., et al., *The bacterial DnaC helicase loader is a DnaB ring breaker.* Cell, 2013. **153**(2): p. 438-48.
28. Bird, R.E., et al., *Origin and sequence of chromosome replication in Escherichia coli.* J. Mol. Biol., 1972. **70**(3): p. 549-566.
29. Nagata, T., *Polarity and synchrony in the replication of DNA molecules of bacteria.* Biochem. Biophys. Res. Commun. , 1962. **8**: p. 348-351.
30. Nagata, T., *The molecular synchrony and sequential replication of DNA in Escherichia coli.* Proc. Natl. Acad. Sci. U.S.A., 1963. **49**: p. 551-559.
31. Arias-Palomo, E., et al., *Physical basis for the loading of a bacterial replicative helicase onto DNA.* Mol. Cell, 2019. **74**(1): p. 173-184 e4.
32. Makowska-Grzyska, M. and J.M. Kaguni, *Primase directs the release of DnaC from DnaB.* Mol. Cell, 2010. **37**(1): p. 90-101.
33. Kim, S., et al., *Coupling of replicative polymerase and helicase: A τ-DnaB interaction mediates rapid replication fork movement* Cell 1996. **84**: p. 643-650.
34. Lewis, J.S., et al., *Single-molecule visualization of fast polymerase turnover in the bacterial replisome.* eLife 2017(6): p. e23932.
35. Onrust, R., et al., *Assembly of a chromosomal replication machine: Two DNA polymerases, a clamp loader, and sliding clamps in one holoenzyme particle.* J. Biol. Chem., 1995. **270**(22): p. 13366-13377.
36. Watson, J.D. and F.H.C. Crick, *Genetical implications of the structure of deoxyribonucleic acid* Nature, 1953. **171**: p. 964-967.
37. Bouché, J., K. Zechel, and A. Kornberg, *dnaG gene product, a rifampicin-resistant RNA polymerase, initiates the conversion of a single-stranded coliphage DNA to its duplex replicative form.* J. Biol. Chem., 1975. **250**(15): p. 5995-6001.
38. Tougu, K. and K.J. Marians, *The interaction between helicase and primase sets the replication fork clock.* J. Biol. Chem., 1996. **271**(32): p. 21398-21405.
39. Alberts, B.M., et al., *Studies on DNA replication in the bacteriophage T4 in vitro system* Cold Spring Harb. Symp. Quant. Biol., 1983. **47**(2): p. 655-668.
40. Yamazaki, K.-I., et al., *Isolation and characterization of nucleoid proteins from Escherichia coli.* Mol. Gen. Genet., 1984. **196**: p. 217-24.
41. Wang, W., et al., *Chromosome organization by a nucleoid-associated protein in live bacteria.* Science, 2011. **333**(6048): p. 1445-9.
42. Pugh, B.F. and M.M. Cox, *Stable binding of recA protein to duplex DNA* J. Biol. Chem., 1987. **262**(3): p. 1326-1336.
43. Ghosh, S. and H. Echols, *Purification and Properties of D Protein: A Transcription Factor of Escherichia coli.* Proc. Natl. Acad. Sci. U.S.A., 1972. **69**(12): p. 3660-3664.

44. Jishage, M. and A. Ishihama, *Regulation of RNA polymerase sigma subunit synthesis in Escherichia coli: Intracellular levels of* σ^{70} *and* σ^{38}*.* J. Bacteriol., 1995. **177**(23): p. 6832-6835.
45. Hirose, S., S. Hiraga, and T. Okazaki, *Initiation site of deoxyribonucleotide polymerization at the replication origin of the Escherichia coli chromosome.* Mol. Gen. Genet., 1983. **189**(3): p. 422-31.
46. Vogel, U. and K.F. Jensen, *The RNA chain elongation rate in Escherichia coli depends on the growth rate* J. Bacteriol., 1994. **176**(10): p. 2807-2813.
47. Gotta, S.L., O.L. Miller, and S. French, *rRNA transcription rate in Escherichia coli.* J. Bacteriol., 1991. **173**(20): p. 6647-6649.
48. Tanner, N.A., et al., *Real-time single-molecule observation of rolling-circle DNA replication.* Nucleic Acids Res., 2009. **37**(4): p. e27.
49. Yao, N., et al., *Single-molecule analysis reveals that the lagging strand increases replisome processivity but slows replication fork progression.* Proc. Natl. Acad. Sci. U.S.A., 2009. **106**(32): p. 13236-13241.
50. Proshkin, S., et al., *Cooperation between translating ribosomes and RNA polymerase in transcription elongation* Science, 2010. **328**(5977): p. 504-508.
51. Epshtein, V. and E. Nudler, *Cooperation between RNA polymerase molecules in transcription elongation* Science, 2003. **300**(5620): p. 801-805.
52. Tolić-Nørrelykke, S.F., et al., *Diversity in the rates of transcript elongation by single RNA polymerase molecules.* J. Biol. Chem., 2004. **279**(5): p. 3292-9.
53. Adelman, K., et al., *Single molecule analysis of RNA polymerase elongation reveals uniform kinetic behavior.* Proc. Natl. Acad. Sci. U.S.A., 2002. **99**(21): p. 13538-13543.
54. Komissarova, N. and M. Kashlev, *RNA polymerase switches between inactivated and activated states by translocating back and forth along the DNA and the RNA.* J. Biol. Chem., 1997. **272**(24): p. 15329-15338.
55. Nudler, E., et al., *The RNA-DNA hybrid maintains the register of transcription by preventing backtracking of RNA polymerase* Cell, 1997. **89**(1): p. 33-41.
56. Cheung, A.C. and P. Cramer, *Structural basis of RNA polymerase II backtracking, arrest and reactivation.* Nature, 2011. **471**(7337): p. 249-53.
57. Trautinger, B.W., et al., *RNA polymerase modulators and DNA repair activities resolve conflicts between DNA replication and transcription.* Mol. Cell, 2005. **19**(2): p. 247-58.
58. Dutta, D., et al., *Linking RNA polymerase backtracking to genome instability in E. coli.* Cell, 2011. **146**(4): p. 533-43.
59. Guy, C.P., et al., *Rep provides a second motor at the replisome to promote duplication of protein-bound DNA.* Mol. Cell, 2009. **36**(4): p. 654-66.
60. Boubakri, H., et al., *The helicases DinG, Rep and UvrD cooperate to promote replication across transcription units in vivo.* EMBO J., 2010. **29**(1): p. 145-57.
61. Byrd, A.K. and K.D. Raney, *Protein displacement by an assembly of helicase molecules aligned along single-stranded DNA.* Nat. Struct. Mol. Biol., 2004. **11**(6): p. 531-8.
62. Atkinson, J., M.K. Gupta, and P. McGlynn, *Interaction of Rep and DnaB on DNA.* Nucleic Acids Res., 2011. **39**(4): p. 1351-9.
63. Larsen, N.B., et al., *The Escherichia coli Tus-Ter replication fork barrier causes site-specific DNA replication perturbation in yeast.* Nat. Commun., 2014. **5**: p. 3574-3583.
64. Wang, X., R. Reyes-Lamothe, and D.J. Sherratt, *Visualizing genetic loci and molecular machines in living bacteria.* Biochem. Soc. Trans., 2008. **36**(Pt 4): p. 749-53.
65. Possoz, C., et al., *Tracking of controlled Escherichia coli replication fork stalling and restart at repressor-bound DNA in vivo.* EMBO J., 2006. **25**(11): p. 2596-2604.

66. McGlynn, P. and C.P. Guy, *Replication forks blocked by protein-DNA complexes have limited stability in vitro.* J. Mol. Biol., 2008. **381**(2): p. 249-55.
67. Payne, B.T., et al., *Replication fork blockage by transcription factor-DNA complexes in Escherichia coli.* Nucleic Acids Res., 2006. **34**(18): p. 5194-202.
68. Hawkins, M., et al., *Direct removal of RNA polymerase barriers to replication by accessory replicative helicases.* Nucleic Acids Res., 2019. **47**: p. 5100-5113.
69. Fu, Y.V., et al., *Selective bypass of a lagging strand roadblock by the eukaryotic replicative DNA helicase.* Cell, 2011. **146**(6): p. 931-41.
70. Bidnenko, V., S.D. Ehrlich, and B. Michel, *Replication fork collapse at replication terminator sequences* EMBO J., 2002. **21**(14): p. 3898-3907.
71. Mettrick, K.A. and I. Grainge, *Stability of blocked replication forks in vivo.* Nucleic Acids Res., 2016. **44**(2): p. 657-68.
72. Pomerantz, R.T. and M. O'Donnell, *The replisome uses mRNA as a primer after colliding with RNA polymerase.* Nature, 2008. **456**(7223): p. 762-6.
73. Pomerantz, R.T. and M. O'Donnell, *Direct restart of a replication fork stalled by a head-on RNA polymerase.* Science, 2010. **327**(5965): p. 590-2.
74. Whinn, K.S., et al., *Nuclease dead Cas9 is a programmable roadblock for DNA replication.* Sci. Rep., 2019. **9**(1): p. 13292.
75. Spenkelink, L.M., et al., *Recycling of single-stranded DNA-binding protein by the bacterial replisome.* Nucleic Acids Res., 2019. **47**(8): p. 4111-4123.
76. Beattie, T., et al., *Frequent exchange of the DNA polymerase during bacterial chromosome replication* eLife, 2017.
77. Graham, J.E., K.J. Marians, and S.C. Kowalczykowski, *Independent and stochastic action of DNA polymerases in the replisome.* Cell, 2017. **169**(7): p. 1201-1213 e17.
78. Spinks, R.R., et al., *DnaB helicase dynamics in bacterial DNA replication resolved by single-molecule studies.* Nucleic Acids Res., 2021. **49**(12): p. 6804-6816.
79. Lohman, T.M. and K.P. Bjornson, *Mechanisms of helicase-catalyzed DNA unwinding.* Annu. Rev. Biochem., 1996. **65**: p. 169-214.
80. Singleton, M.R., M.S. Dillingham, and D.B. Wigley, *Structure and mechanism of helicases and nucleic acid translocases.* Annu. Rev. Biochem., 2007. **76**: p. 23-50.
81. Lohman, T.M., E.J. Tomko, and C.G. Wu, *Non-hexameric DNA helicases and translocases: mechanisms and regulation.* Nat. Rev. Mol. Cell. Biol., 2008. **9**(5): p. 391-401.
82. Mackintosh, S.G. and K.D. Raney, *DNA unwinding and protein displacement by superfamily 1 and superfamily 2 helicases.* Nucleic Acids Res., 2006. **34**(15): p. 4154-9.
83. Lohman, T.M., *Helicase-catalyzed DNA unwinding* J. Biol. Chem., 1993. **268**(4): p. 2269-2272.
84. Van Houten, B. and A. Snowden, *Mechanism of action of the Escherichia coli UvrABC nuclease: Clues to the damage recognition problem.* BioEssays, 1993. **15**(1): p. 51-59.
85. Dillingham, M.S. and S.C. Kowalczykowski, *RecBCD enzyme and the repair of double-stranded DNA breaks.* Microbiol. Mol. Biol. Rev., 2008. **72**(4): p. 642-71, Table of Contents.
86. Cadman, C.J. and P. McGlynn, *PriA helicase and SSB interact physically and functionally.* Nucleic Acids Res., 2004. **32**(21): p. 6378-87.
87. Buss, J.A., Y. Kimura, and P.R. Bianco, *RecG interacts directly with SSB: implications for stalled replication fork regression.* Nucleic Acids Res., 2008. **36**(22): p. 7029-42.

88. Hall, M.C., J.R. Jordan, and S.W. Matson, *Evidence for a physical interaction between the Escherichia coli methyl-directed mismatch repair proteins MutL and UvrD* EMBO J., 1998. **17**(5): p. 1535-1541.
89. Park, J.S., M.T. Marr, and J.W. Roberts, *E. coli transcription repair coupling factor (Mfd protein) rescues arrested complexes by promoting forward translocation.* Cell, 2002. **109**(6): p. 757-67.
90. Stracy, M. and A.N. Kapanidis, *Single-molecule and super-resolution imaging of transcription in living bacteria.* Methods, 2017. **120**: p. 103-114.
91. Li, Y., et al., *Visualizing bacterial DNA replication and repair with molecular resolution.* Curr. Opin. Microbiol., 2018. **43**: p. 38-45.
92. Uphoff, S. and D.J. Sherratt, *Single-molecule analysis of bacterial DNA repair and mutagenesis.* Annu. Rev. Biophys., 2017. **46**: p. 411-432.
93. Elf, J., G.W. Li, and X.S. Xie, *Probing transcription factor dynamics at the single-molecule level in a living cell* Science, 2007. **316**: p. 1191.
94. Stracy, M., et al., *Single-molecule imaging of UvrA and UvrB recruitment to DNA lesions in living Escherichia coli.* Nat. Commun., 2016. **7**: p. 12568.
95. Ho, H.N., A.M. van Oijen, and H. Ghodke, *The transcription-repair coupling factor Mfd associates with RNA polymerase in the absence of exogenous damage.* Nat. Commun., 2018. **9**(1): p. 1570.
96. Stracy, M., et al., *Live-cell superresolution microscopy reveals the organization of RNA polymerase in the bacterial nucleoid.* Proc. Natl. Acad. Sci. U.S.A., 2015. **112**(32): p. E4390-9.
97. Robinson, A., et al., *Regulation of mutagenic DNA polymerase V activation in space and time.* PLoS Genet., 2015. **11**(8): p. e1005482.
98. Ghodke, H., et al., *Spatial and temporal organization of RecA in the Escherichia coli DNA-damage response.* eLife, 2019. **8**.
99. Uphoff, S., *Real-time dynamics of mutagenesis reveal the chronology of DNA repair and damage tolerance responses in single cells.* Proc. Natl. Acad. Sci. U.S.A., 2018. **115**(28): p. E6516-E6525.
100. Windgassen, T.A., et al., *Mechanisms of bacterial DNA replication restart.* Nucleic Acids Res., 2018. **46**(2): p. 504-519.
101. Veaute, X., et al., *UvrD helicase, unlike Rep helicase, dismantles RecA nucleoprotein filaments in Escherichia coli.* EMBO J., 2005. **24**: p. 180-189.
102. Paul, T., et al., *E. coli Rep helicase and RecA recombinase unwind G4 DNA and are important for resistance to G4-stabilizing ligands.* Nucleic Acids Res., 2020. **48**(12): p. 6640-6653.
103. Park, J., M.T. Marr, and J.W. Roberts, *E. coli transcription repair coupling factor (Mfd protein) rescues arrested complexes by promoting forward translocation.* Cell, 2002. **109**(6): p. 757-767.
104. Theis, K., et al., *The nucleotide excision repair protein UvrB, a helicase-like enzyme with a catch* Mutat. Res. , 2000. **460**(3-4): p. 277-300.
105. Linder, P. and E. Jankowsky, *From unwinding to clamping - the DEAD box RNA helicase family.* Nat. Rev. Mol. Cell. Biol., 2011. **12**(8): p. 505-16.
106. Gilman, B., P. Tijerina, and R. Russell, *Distinct RNA-unwinding mechanisms of DEAD-box and DEAH-box RNA helicase proteins in remodeling structured RNAs and RNPs.* Biochem. Soc. Trans., 2017. **45**(6): p. 1313-1321.
107. Manosas, M., et al., *Active and passive mechanisms of helicases.* Nucleic Acids Res., 2010. **38**(16): p. 5518-26.
108. Delagoutte, E. and P.H. von Hippel, *Helicase mechanisms and the coupling of helicases within macromolecular machines. Part I: Structures and properties of isolated helicases.* Q. Rev. Biophys., 2002. **35**(4): p. 431-78.
109. Betterton, M.D. and F. Julicher, *Opening of nucleic-acid double strands by helicases: active versus passive opening.* Phys Rev E Stat Nonlin Soft Matter Phys, 2005. **71**(1 Pt 1): p. 011904.

110. Gorbalenya, A.E. and E.V. Koonin, *Helicases: amino acid sequence comparisons and structure-function relationships.* Curr. Opin. Struct. Biol., 1993. **3**(3): p. 419-429.
111. Raney, K.D., A.K. Byrd, and S. Aarattuthodiyil, *Structure and mechanisms of SF1 DNA helicases.* Adv. Exp. Med. Biol., 2013. **767**: p. 17-46.
112. Beyer, D.C., M.K. Ghoneim, and M. Spies, *Structure and mechanisms of SF2 DNA helicases.* Adv. Exp. Med. Biol., 2013. **767**: p. 47-73.
113. Fairman-Williams, M.E., U.P. Guenther, and E. Jankowsky, *SF1 and SF2 helicases: family matters.* Curr. Opin. Struct. Biol., 2010. **20**(3): p. 313-24.
114. Lyubimov, A.Y., M. Strycharska, and J.M. Berger, *The nuts and bolts of ring-translocase structure and mechanism.* Curr. Opin. Struct. Biol., 2011. **21**(2): p. 240-8.
115. Korolev, S., et al., *Major domain swiveling revealed by the crystal structures of complexes of E. coli Rep helicase bound to single-stranded DNA and ADP.* Cell, 1997. **90**(4): p. 635-647.
116. Lee, J.Y. and W. Yang, *UvrD helicase unwinds DNA one base pair at a time by a two-part power stroke.* Cell, 2006. **127**(7): p. 1349-60.
117. Wilkinson, M., Y. Chaban, and D.B. Wigley, *Mechanism for nuclease regulation in RecBCD.* eLife, 2016. **5**.
118. Singleton, M.R., S. Scaife, and D.B. Wigley, *Structural analysis of DNA replication fork reversal by RecG.* Cell, 2001. **107**: p. 78-89.
119. Bhattacharyya, B., et al., *Structural mechanisms of PriA-mediated DNA replication restart.* Proc. Natl. Acad. Sci. U.S.A., 2014. **111**(4): p. 1373-8.
120. Truglio, J.J., et al., *Structural basis for DNA recognition and processing by UvrB.* Nat. Struct. Mol. Biol., 2006. **13**(4): p. 360-4.
121. Brugger, C., et al., *Molecular determinants for dsDNA translocation by the transcription-repair coupling and evolvability factor Mfd.* Nat. Commun., 2020. **11**(1): p. 3740.
122. Itsathitphaisarn, O., et al., *The hexameric helicase DnaB adopts a nonplanar conformation during translocation.* Cell, 2012. **151**(2): p. 267-77.
123. Kornberg, A. and T.A. Baker, *DNA Replication*. 1992: W.H. Freeman.
124. Perera, H.M., et al., *Contacts and context that regulate DNA helicase unwinding and replisome progression.* Enzymes, 2019. **45**: p. 183-223.
125. Robinson, A. and A.M. van Oijen, *Bacterial replication, transcription and translation: mechanistic insights from single-molecule biochemical studies.* Nat. Rev. Microbiol., 2013. **11**(5): p. 303-15.
126. Gao, D. and C.S. McHenry, *τ binds and organizes Escherichia coli replication proteins through distinct domains. Domain IV, located within the unique C terminus of τ, binds the replication fork, helicase, DnaB.* J. Biol. Chem., 2001. **276**(6): p. 4441-6.
127. Oakley, A.J., et al., *Crystal and solution structures of the helicase-binding domain of Escherichia coli primase.* J. Biol. Chem., 2005. **280**(12): p. 11495-504.
128. Tougu, K., H. Peng, and K.J. Marians, *Identification of a domain of Escherichia coli primase required for functional interaction with the DnaB helicase at the replication fork.* J. Biol. Chem., 1994. **269**(6): p. 4675-4682.
129. Heller, R.C. and K.J. Marians, *Unwinding of the nascent lagging strand by Rep and PriA enables the direct restart of stalled replication forks.* J. Biol. Chem., 2005. **280**(40): p. 34143-51.
130. Brüning, J.G., et al., *The 2B subdomain of Rep helicase links translocation along DNA with protein displacement.* Nucleic Acids Res., 2018. **46**: p. 8917-8925.
131. Velankar, S.S., et al., *Crystal structures of complexes of PcrA DNA helicase with a DNA substrate indicate an inchworm mechanism.* Cell, 1999. **97**(1): p. 75-84.

132. Tomko, E.J., C.J. Fischer, and T.M. Lohman, *Single-stranded DNA translocation of E. coli UvrD monomer is tightly coupled to ATP hydrolysis.* J. Mol. Biol., 2012. **418**(1-2): p. 32-46.
133. Tomko, E.J., et al., *A nonuniform stepping mechanism for E. coli UvrD monomer translocation along single-stranded DNA.* Mol. Cell, 2007. **26**(3): p. 335-47.
134. Cheng, W., et al., *E. coli Rep oligomers are required to initiate DNA unwinding in vitro.* J. Mol. Biol., 2001. **310**(2): p. 327-50.
135. Dillingham, M.S., D.B. Wigley, and M.R. Webb, *Demonstration of unidirectional single-stranded DNA translocation by PcrA helicase: Measurement of step Size and translocation speed.* Biochemistry 2000. **39**(1): p. 205-212.
136. Myong, S., et al., *Repetitive shuttling of a motor protein on DNA.* Nature, 2005. **437**(7063): p. 1321-5.
137. Jia, H., et al., *Rotations of the 2B sub-domain of E. coli UvrD helicase/translocase coupled to nucleotide and DNA binding.* J. Mol. Biol., 2011. **411**(3): p. 633-48.
138. Arslan, S., et al., *Engineering of a superhelicase through conformational control.* Science, 2015. **348**(6232): p. 344-347.
139. Nguyen, B., et al., *Large domain movements upon UvrD dimerization and helicase activation* Proc. Natl. Acad. Sci. U.S.A., 2017 **114**(46).
140. Comstock, M.J., et al., *Direct observation of structure-function relationship in a nucleic acid-processing enzyme* Science, 2015. **348**(6232): p. 352-354.
141. Makurath, M.A., et al., *Regulation of Rep helicase unwinding by an auto-inhibitory subdomain.* Nucleic Acids Res., 2019. **47**(5): p. 2523-2532.
142. Cheng, W., et al., *The 2B domain of the Escherichia coli Rep protein is not required for DNA helicase activity.* Proc. Natl. Acad. Sci. U.S.A., 2002. **99**(25): p. 16006-16011.
143. Ha, T., et al., *Inititation and re-initiation of DNA unwinding by the Escherichia coli Rep helicase.* Nature, 2002. **419**: p. 638-641.
144. Maluf, N.K., C.J. Fischer, and T.M. Lohman, *A dimer of Escherichia coli UvrD is the active form of the helicase in vitro.* J. Mol. Biol., 2003. **325**(5): p. 913-935.
145. Fischer, C.J., N.K. Maluf, and T.M. Lohman, *Mechanism of ATP-dependent translocation of E. coli UvrD monomers along single-stranded DNA.* J. Mol. Biol., 2004. **344**(5): p. 1287-309.
146. Brendza, K.M., et al., *Autoinhibition of Escherichia coli Rep monomer helicase activity by its 2B subdomain.* Proc. Natl. Acad. Sci. U.S.A., 2005. **102**(29): p. 10076-81.
147. Niedziela-Majka, A., et al., *Bacillus stearothermophilus PcrA monomer is a single-stranded DNA translocase but not a processive helicase in vitro.* J. Biol. Chem., 2007. **282**(37): p. 27076-85.
148. Lee, K.S., et al., *Direct imaging of single UvrD helicase dynamics on long single-stranded DNA.* Nat. Commun., 2013. **4**: p. 1878.
149. Yokota, H., Y.A. Chujo, and Y. Harada, *Single-molecule imaging of the oligomer formation of the nonhexameric Escherichia coli UvrD helicase.* Biophys. J., 2013. **104**(4): p. 924-33.
150. Ma, W., et al., *Free-energy simulations reveal molecular mechanism for functional switch of a DNA helicase.* eLife, 2018. **7**.
151. Li, J.H., et al., *Pif1 is a force-regulated helicase.* Nucleic Acids Res., 2016. **44**(9): p. 4330-9.
152. Masuda-Ozawa, T., et al., *Single-molecule sorting reveals how ubiquitylation affects substrate recognition and activities of FBH1 helicase.* Nucleic Acids Res., 2013. **41**(6): p. 3576-87.
153. Bentchikou, E., et al., *Are the SSB-interacting proteins RecO, RecG, PriA and the DnaB-interacting protein Rep bound to progressing replication forks in Escherichia coli?* PLoS One, 2015. **10**(8): p. e0134892.

154. Atkinson, J., et al., *Localization of an accessory helicase at the replisome is critical in sustaining efficient genome duplication.* Nucleic Acids Res., 2011. **39**(3): p. 949-57.

155. Syeda, A.H., et al., *Single-molecule live cell imaging of Rep reveals the dynamic interplay between an accessory replicative helicase and the replisome.* Nucleic Acids Res., 2019. **47**(12): p. 6287-6298.

156. Schauer, G.D., et al., *Replisome bypass of a protein-based R-loop block by Pif1.* Proc. Natl. Acad. Sci. U.S.A., 2020. **117**: p. 30354-30361.

157. Petrova, V., et al., *Active displacement of RecA filaments by UvrD translocase activity.* Nucleic Acids Res., 2015. **43**(8): p. 4133-49.

158. Epshtein, V., et al., *UvrD facilitates DNA repair by pulling RNA polymerase backwards.* Nature, 2014. **505**(7483): p. 372-7.

159. Heller, R.C. and K.J. Marians, *Non-replicative helicases at the replication fork.* DNA Repair (Amst), 2007. **6**(7): p. 945-52.

160. Bidnenko, V., R. Lestini, and B. Michel, *The Escherichia coli UvrD helicase is essential for Tus removal during recombination-dependent replication restart from Ter sites.* Mol. Microbiol., 2006. **62**(2): p. 382-96.

161. Florés, M.J., N. Sanchez, and B. Michel, *A fork-clearing role for UvrD.* Mol. Microbiol., 2005. **57**(6): p. 1664-75.

162. Lestini, R. and B. Michel, *UvrD and UvrD252 counteract RecQ, RecJ, and RecFOR in a rep mutant of Escherichia coli.* J. Bacteriol., 2008. **190**(17): p. 5995-6001.

163. Lane, H.E.D. and D.T. Denhardt, *The rep mutation III. Altered structure of the replicating Escherichia coli chromosome.* J. Bacteriol., 1974. **120**(2): p. 805-814.

164. Petit, M. and D. Ehrlich, *Essential bacterial helicases that counteract the toxicity of recombination proteins* EMBO J., 2002. **21**(12): p. 3137-3147.

165. Shima, N., et al., *A viable allele of Mcm4 causes chromosome instability and mammary adenocarcinomas in mice.* Nat. Genet., 2007. **39**(1): p. 93-8.

166. Merrikh, H., et al., *Replication-transcription conflicts in bacteria.* Nat. Rev. Microbiol., 2012. **10**(7): p. 449-58.

167. Cox, M.M., et al., *The importance of repairing stalled replication forks.* Nature, 2000. **404**(6773): p. 37-41.

168. Helmrich, A., et al., *Transcription-replication encounters, consequences and genomic instability.* Nat. Struct. Mol. Biol., 2013. **20**(4): p. 412-8.

169. Hamperl, S. and K.A. Cimprich, *Conflict resolution in the genome: How transcription and replication make it work.* Cell, 2016. **167**(6): p. 1455-1467.

170. Chang, E.Y. and P.C. Stirling, *Replication fork protection factors controlling R-loop bypass and suppression.* Genes (Basel), 2017. **8**(1).

171. Bianco, P.R. and Y. Lu, *Single-molecule insight into stalled replication fork rescue in Escherichia coli.* Nucleic Acids Res., 2021. **49**(8): p. 4220-4238.

172. Marians, K.J., *PriA-directed replication fork restart in Escherichia coli.* Trends Biochem. Sci. , 2000. **25**(4): p. 185-189.

173. Liu, J., et al., *Replication fork assembly at recombination intermediates is required for bacterial growth* Proc. Natl. Acad. Sci. U.S.A., 1999. **96**: p. 3552-3555.

174. Manosas, M., et al., *RecG and UvsW catalyse robust DNA rewinding critical for stalled DNA replication fork rescue.* Nat. Commun., 2013. **4**: p. 2368.

175. Manosas, M., et al., *Direct observation of stalled fork restart via fork regression in the T4 replication system.* Science, 2012. **338**(6111): p. 1217-20.

176. Slocum, S.L., et al., *Characterization of the ATPase activity of the Escherichia coli RecG protein reveals that the preferred cofactor is negatively supercoiled DNA.* J. Mol. Biol., 2007. **367**(3): p. 647-64.

177. Abd Wahab, S., M. Choi, and P.R. Bianco, *Characterization of the ATPase activity of RecG and RuvAB proteins on model fork structures reveals insight into stalled DNA replication fork repair.* J. Biol. Chem., 2013. **288**(37): p. 26397-409.
178. Sun, Z., et al., *Remodeling of RecG Helicase at the DNA Replication Fork by SSB Protein.* Sci. Rep., 2015. **5**: p. 9625.
179. Sun, Z., et al., *Dynamics of the interaction of RecG protein with stalled replication forks.* Biochemistry, 2018. **57**(13): p. 1967-1976.
180. Sun, Z., et al., *Nanoscale interaction of RecG with mobile fork DNA.* Nanoscale Adv., 2020. **2**(3): p. 1318-1324.
181. Windgassen, T.A., et al., *Function of a strand-separation pin element in the PriA DNA replication restart helicase.* J. Biol. Chem., 2019. **294**(8): p. 2801-2814.
182. Costes, A., et al., *The C-terminal domain of the bacterial SSB protein acts as a DNA maintenance hub at active chromosome replication forks.* PLoS Genet., 2010. **6**(12): p. e1001238.
183. Yu, C., et al., *SSB binds to the RecG and PriA helicases in vivo in the absence of DNA.* Genes Cells, 2016. **21**(2): p. 163-84.
184. Chen, H.W., S.H. North, and H. Nakai, *Properties of the PriA helicase domain and its role in binding PriA to specific DNA structures.* J. Biol. Chem., 2004. **279**(37): p. 38503-12.
185. Wang, Y., et al., *Atomic force microscopy-based characterization of the interaction of PriA helicase with stalled DNA replication forks.* J. Biol. Chem., 2020. **295**(18): p. 6043-6052.
186. Friedberg, E.C., et al., *DNA Repair and Mutagenesis*. 2005: ASM Press.
187. Roos, W.P. and B. Kaina, *DNA damage-induced cell death by apoptosis.* Trends Mol. Med., 2006. **12**(9): p. 440-50.
188. Sancar, A., *DNA excision repair* Annu. Rev. Biochem., 1996. **65**: p. 43-81.
189. Kisker, C., J. Kuper, and B. Van Houten, *Prokaryotic nucleotide excision repair.* Cold Spring Harb. Perspect. Biol., 2013. **5**(3): p. 012591.
190. Iyer, R.R., et al., *DNA mismatch repair: Functions and mechanisms* Chem. Rev. , 2006. **106**(2): p. 302-323.
191. Atkinson, J. and P. McGlynn, *Replication fork reversal and the maintenance of genome stability.* Nucleic Acids Res., 2009. **37**(11): p. 3475-92.
192. Tomko, E.J., et al., *5'-Single-stranded/duplex DNA junctions are loading sites for E. coli UvrD translocase.* EMBO J., 2010. **29**(22): p. 3826-39.
193. Ali, J.A., N.K. Maluf, and T.M. Lohman, *An oligomeric form of E. coli UvrD is required for optimal helicase activity.* J. Mol. Biol., 1999. **293**: p. 815-834.
194. Runyon, G.T., I. Wong, and T.M. Lohman, *Overexpression, purification, DNA binding, and dimerization of the Escherichia coli uvrD gene product (helicase II).* Biochemistry, 1993. **32**: p. 602-612.
195. Maluf, N.K., J.A. Ali, and T.M. Lohman, *Kinetic mechanism for formation of the active, dimeric UvrD helicase-DNA complex.* J. Biol. Chem., 2003. **278**(34): p. 31930-40.
196. Ali, J.A. and T.M. Lohman, *Kinetic measurement of the step size of DNA unwinding by Escherichia coli UvrD helicase.* Science, 1997. **275**(5298): p. 377-80.
197. Yamaguchi, M., V. Dao, and P. Modrich, *MutS and MutL activate DNAhHelicase II in a mismatch-dependent manner.* J. Biol. Chem., 1998. **273**(15): p. 9197-201.
198. Mechanic, L.E., B.A. Frankel, and S.W. Matson, *Escherichia coli MutL loads DNA helicase II onto DNA.* J. Biol. Chem., 2000. **275**(49): p. 38337-46.
199. Ordabayev, Y.A., et al., *Regulation of UvrD helicase activity by MutL.* J. Mol. Biol., 2018. **430**(21): p. 4260-4274.
200. Ordabayev, Y.A., et al., *UvrD helicase activation by MutL involves rotation of its 2B subdomain.* Proc. Natl. Acad. Sci. U.S.A., 2019. **116**(33): p. 16320-16325.

201. Liu, J., et al., *MutL sliding clamps coordinate exonuclease-independent Escherichia coli mismatch repair.* Nat. Commun., 2019. **10**(1): p. 5294.
202. Chisty, L.T., et al., *Monomeric PcrA helicase processively unwinds plasmid lengths of DNA in the presence of the initiator protein RepD.* Nucleic Acids Res., 2013. **41**(9): p. 5010-23.
203. Zhang, W., et al., *Directional loading and stimulation of PcrA helicase by the replication initiator protein RepD.* J. Mol. Biol., 2007. **371**(2): p. 336-48.
204. Noirot, P., J. Bargonetti, and R.P. Novick, *Initiation of rolling-circle replication in pT181 plasmid: Initiator protein enhances cruciform extrusion at the origin.* Proc. Natl. Acad. Sci. U.S.A., 1990. **87**: p. 8560-8564.
205. Arbore, C., L.M. Lewis, and M.R. Webb, *Kinetic mechanism of initiation by RepD as a part of asymmetric, rolling circle plasmid unwinding.* Biochemistry, 2012. **51**(17): p. 3684-93.
206. Oh, E.Y. and L. Grossman, *Helicase properties of the Escherichia coli UvrAB protein complex.* Proc. Natl. Acad. Sci. U.S.A., 1987. **84**(11): p. 3638-3642.
207. Oh, E.Y. and L. Grossman, *Characterization of the helicase activity of the Escherichia coli UvrAB protein complex.* J. Biol. Chem., 1989. **264**(2): p. 1336-1343.
208. Pakotiprapha, D., et al., *A structural model for the damage-sensing complex in bacterial nucleotide excision repair.* J. Biol. Chem., 2009. **284**(19): p. 12837-44.
209. Pakotiprapha, D., et al., *Structure and mechanism of the UvrA-UvrB DNA damage sensor.* Nat. Struct. Mol. Biol., 2012. **19**(3): p. 291-8.
210. Kad, N.M., et al., *Collaborative dynamic DNA scanning by nucleotide excision repair proteins investigated by single-molecule imaging of quantum-dot-labeled proteins.* Mol. Cell, 2010. **37**(5): p. 702-13.
211. Verhoeven, E.E.A., et al., *The presence of two UvrB subunits in the UvrAB complex ensures damage detection in both DNA strands* EMBO J., 2002. **21**(15): p. 4196-4205.
212. Orren, D.K. and A. Sancar, *The (A)BC excinuclease of Escherichia coli has only the UvrB and UvrC subunits in the incision complex* Proc. Natl. Acad. Sci. U.S.A., 1989. **86**: p. 5237-5241.
213. Yeung, A.T., et al., *Enzymatic properties of purified Escherichia coli uvrABC proteins* Proc. Natl. Acad. Sci. U.S.A., 1983. **80**: p. 6157-6161.
214. Verhoeven, E.E., et al., *Catalytic sites for 3' and 5' incision of Escherichia coli nucleotide excision repair are both located in UvrC.* J. Biol. Chem., 2000. **275**(7): p. 5120-3.
215. Ghodke, H., H.N. Ho, and A.M. van Oijen, *Single-molecule live-cell imaging visualizes parallel pathways of prokaryotic nucleotide excision repair.* Nat. Commun., 2020. **11**(1): p. 1477.
216. Ho, H.N., A.M. van Oijen, and H. Ghodke, *Single-molecule imaging reveals molecular coupling between transcription and DNA repair machinery in live cells.* Nat. Commun., 2020. **11**(1): p. 1478.
217. Paudel, B.P., et al., *Mechanism of transcription modulation by the transcription-repair coupling factor.* 2021: p. 2021.04.02.438179.
218. Webster, M.P., et al., *Crystal structure of the UvrB dimer: insights into the nature and functioning of the UvrAB damage engagement and UvrB-DNA complexes.* Nucleic Acids Res., 2012. **40**(17): p. 8743-58.
219. Koo, H., et al., *ATP-dependent partitioning of the DNA template into supercoiled domains by Escherichia coli UvrAB.* Proc. Natl. Acad. Sci. U.S.A., 1991. **88**: p. 1212-1216.
220. Hughes, C.D., et al., *Real-time single-molecule imaging reveals a direct interaction between UvrC and UvrB on DNA tightropes.* Nucleic Acids Res., 2013. **41**(9): p. 4901-12.

221. Springall, L., et al., *Recruitment of UvrBC complexes to UV-induced damage in the absence of UvrA increases cell survival.* Nucleic Acids Res., 2018. **46**(3): p. 1256-1265.
222. Theis, K., et al., *Crystal structure of UvrB, a DNA helicase adapted for nucleotide excision repair.* EMBO J., 1999. **18**(24): p. 6899-6907.
223. Lee, S.J., R.J. Sung, and G.L. Verdine, *Mechanism of DNA Lesion Homing and Recognition by the Uvr Nucleotide Excision Repair System.* Research, 2019. **2019**: p. 5641746.
224. Truglio, J.J., et al., *Prokaryotic nucelotide excision repair: The UvrABC system.* Chem. Rev., 2006. **106**(2): p. 233-252.
225. Malta, E., G.F. Moolenaar, and N. Goosen, *Base flipping in nucleotide excision repair.* J. Biol. Chem., 2006. **281**(4): p. 2184-94.
226. Mathieu, N., N. Kaczmarek, and H. Naegeli, *Strand- and site-specific DNA lesion demarcation by the xeroderma pigmentosum group D helicase.* Proc. Natl. Acad. Sci. U.S.A., 2010. **107**(41): p. 17545-50.
227. Fan, L., et al., *XPD helicase structures and activities: Insights into the cancer and aging phenotypes from XPD mutations.* Cell, 2008. **133**(5): p. 789-800.
228. Liu, H., et al., *Structure of the DNA repair helicase XPD.* Cell, 2008. **133**(5): p. 801-12.
229. Wolski, S.C., et al., *Crystal structure of the FeS cluster-containing nucleotide excision repair helicase XPD.* PLoS Biol., 2008. **6**(6): p. e149.
230. Mathieu, N., et al., *DNA quality control by a lesion sensor pocket of the xeroderma pigmentosum group D helicase subunit of TFIIH.* Curr. Biol., 2013. **23**(3): p. 204-12.
231. Pugh, R.A., C.G. Wu, and M. Spies, *Regulation of translocation polarity by helicase domain 1 in SF2B helicases.* EMBO J., 2012. **31**(2): p. 503-14.
232. Ghoneim, M. and M. Spies, *Direct correlation of DNA binding and single protein domain motion via dual illumination fluorescence microscopy.* Nano. Lett., 2014. **14**(10): p. 5920-31.
233. Belogurov, G.A. and I. Artsimovitch, *The mechanisms of substrate selection, catalysis, and translocation by the elongating RNA polymerase.* J. Mol. Biol., 2019. **431**(20): p. 3975-4006.
234. van Gool, A.J., et al., *RAD26, the functional S. cerevisiae homolog of the Cockayne syndrome B gene ERCC6.* EMBO J., 1994. **13**(22): p. 5361-9.
235. van Hoffen, A., et al., *Deficient repair of the transcribed strand of active genes in Cockayne's syndrome cells.* Nucleic Acids Res., 1993. **21**(25): p. 5890-5.
236. Le, T.T., et al., *Mfd dynamically regulates transcription via a release and catch-up mechanism.* Cell, 2017.
237. Kang, J.Y., et al., *Structural basis for transcription complex disruption by the Mfd translocase.* eLife, 2021. **10**.
238. Graves, E.T., et al., *A dynamic DNA-repair complex observed by correlative single-molecule nanomanipulation and fluorescence.* Nat. Struct. Mol. Biol., 2015. **22**(6): p. 452-7.
239. Fan, J., et al., *Reconstruction of bacterial transcription-coupled repair at single-molecule resolution.* Nature, 2016. **536**(7615): p. 234-7.
240. Selby, C.P. and A. Sancar, *Structure and function of transcription-repair coupling factor. II. catalytic properties.* J. Biol. Chem., 1995. **270**(9): p. 4890-5.
241. Smith, A.J., M.D. Szczelkun, and N.J. Savery, *Controlling the motor activity of a transcription-repair coupling factor: autoinhibition and the role of RNA polymerase.* Nucleic Acids Res., 2007. **35**(6): p. 1802-11.
242. Selby, C.P., E.M. Witkin, and A. Sancar, *Escherichia coli mfd mutant deficient in "mutation frequency decline" lacks strand-specific repair: In vitro complementation with purified coupling factor.* Proc. Natl. Acad. Sci. U.S.A., 1991. **88**(24): p. 11574-8.

243. Selby, C.P. and A. Sancar, *Transcription-repair coupling and mutation frequency decline.* J. Bacteriol., 1993. **175**(23): p. 7509-14.

244. Selby, C.P. and A. Sancar, *Molecular mechanism of transcription-repair coupling.* Science, 1993. **260**(5104): p. 53-8.

245. Mellon, I., G. Spivak, and P.C. Hanawalt, *Selective removal of transcription-blocking DNA damage from the transcribed strand of the mammalian DHFR gene.* Cell, 1987. **51**(2): p. 241-9.

246. Mellon, I. and P.C. Hanawalt, *Induction of the Escherichia coli lactose operon selectively increases repair of its transcribed DNA strand.* Nature, 1989. **342**(6245): p. 95-8.

247. Venema, J., et al., *The genetic defect in Cockayne syndrome is associated with a defect in repair of UV-induced DNA damage in transcriptionally active DNA.* Proc. Natl. Acad. Sci. U.S.A., 1990. **87**(12): p. 4707-11.

248. Mahdi, A.A., et al., *A model for dsDNA translocation revealed by a structural motif common to RecG and Mfd proteins.* EMBO J., 2003. **22**(3): p. 724-34.

249. Deaconescu, A.M., et al., *Structural basis for bacterial transcription-coupled DNA repair.* Cell, 2006. **124**(3): p. 507-20.

250. Shi, J., et al., *Structural basis of Mfd-dependent transcription termination.* Nucleic Acids Res., 2020.

251. Clapier, C.R., et al., *Mechanisms of action and regulation of ATP-dependent chromatin-remodelling complexes.* Nat. Rev. Mol. Cell. Biol., 2017. **18**(7): p. 407-422.

252. Howan, K., et al., *Initiation of transcription-coupled repair characterized at single-molecule resolution.* Nature, 2012. **490**(7420): p. 431-4.

253. Haines, N.M., et al., *Stalled transcription complexes promote DNA repair at a distance.* Proc. Natl. Acad. Sci. U.S.A., 2014. **111**(11): p. 4037-42.

254. Deaconescu, A.M., et al., *Nucleotide excision repair (NER) machinery recruitment by the transcription-repair coupling factor involves unmasking of a conserved intramolecular interface.* Proc. Natl. Acad. Sci. U.S.A., 2012. **109**(9): p. 3353-8.

255. Murphy, M.N., et al., *An N-terminal clamp restrains the motor domains of the bacterial transcription-repair coupling factor Mfd.* Nucleic Acids Res., 2009. **37**(18): p. 6042-53.

256. Manelyte, L., et al., *Regulation and rate enhancement during transcription-coupled DNA repair.* Mol. Cell, 2010. **40**(5): p. 714-24.

257. Smith, A.J., C. Pernstich, and N.J. Savery, *Multipartite control of the DNA translocase, Mfd.* Nucleic Acids Res., 2012. **40**(20): p. 10408-16.

258. Ragheb, M.N., et al., *Mfd regulates RNA polymerase association with hard-to-transcribe regions in vivo, especially those with structured RNAs.* Proc. Natl. Acad. Sci. U.S.A., 2021. **118**(1).

259. Ho, H.N., et al., *Identification of multiple kinetic populations of DNA-binding proteins in live cells.* Biophys. J., 2019. **117**(5): p. 950-961.

260. Portman, J.R., et al., *Cotranscriptional R-loop formation by Mfd involves topological partitioning of DNA.* Proc. Natl. Acad. Sci. U.S.A., 2021. **118**(15).

261. Ukkivi, K. and M. Kivisaar, *Involvement of transcription-coupled repair factor Mfd and DNA helicase UvrD in mutational processes in Pseudomonas putida.* DNA Repair (Amst), 2018. **72**: p. 18-27.

262. Martin, H.A., et al., *Transcriptional de-repression and Mfd are mutagenic in stressed Bacillus subtilis cells.* J Mol Microbiol Biotechnol, 2011. **21**(1-2): p. 45-58.

263. Gómez-Marroquin, M., et al., *Stationary-phase mutagenesis in stressed Bacillus subtilis cells operates by Mfd-dependent mutagenic pathways.* Genes (Basel), 2016. **7**(7).

264. Baharoglu, Z., A. Babosan, and D. Mazel, *Identification of genes involved in low aminoglycoside-induced SOS response in Vibrio cholerae: a role for transcription stalling and Mfd helicase.* Nucleic Acids Res., 2014. **42**(4): p. 2366-79.
265. Taylor, A.F., et al., *Control of RecBCD enzyme activity by DNA binding- and Chi hotspot-dependent conformational changes.* J. Mol. Biol., 2014. **426**(21): p. 3479-99.
266. Spies, M., et al., *A molecular throttle: The recombination hotspot χ controls DNA translocation by the RecBCD helicase* Cell, 2003. **114**: p. 647-654.
267. Wang, J., R. Chen, and D.A. Julin, *A single nuclease active site of the Escherichia coli RecBCD enzyme catalyzes single-stranded DNA degradation in both directions.* J. Biol. Chem., 2000. **275**(1): p. 507-13.
268. Anderson, D.G. and S.C. Kowalczykowski, *The recombination hotspot χ is a regulatory element that switches the polarity of DNA degradation by the RecBCD enzyme* Genes Dev., 1997. **11**: p. 571-581.
269. Dixon, D.A. and S.C. Kowalczykowski, *The recombination hotspot χ is a regulatory sequence that acts by attenuating the nucelase activity of the E. coli RecBCD enzyme* Cell, 1993. **73**: p. 87-96.
270. Taylor, A.F. and G.R. Smith, *RecBCD enzyme is a DNA helicase with fast and slow motors of opposite polarity* Nature, 2003. **423**(6942): p. 889-893.
271. Dillingham, M.S., M. Spies, and S.C. Kowalczykowski, *RecBCD enzyme is a bipolar DNA helicase* Nature, 2003. **423**(6942): p. 893-897.
272. Handa, N., et al., *Direct visualization of RecBCD movement reveals cotranslocation of the RecD motor after χ recognition.* Mol. Cell, 2005. **17**(5): p. 745-50.
273. Liu, B., R.J. Baskin, and S.C. Kowalczykowski, *DNA unwinding heterogeneity by RecBCD results from static molecules able to equilibrate.* Nature, 2013. **500**(7463): p. 482-5.
274. Wiktor, J., et al., *Direct observation of end resection by RecBCD during double-stranded DNA break repair in vivo.* Nucleic Acids Res., 2018. **46**(4): p. 1821-1833.
275. Spies, M., et al., *RecBCD enzyme switches lead motor subunits in response to χ recognition.* Cell, 2007. **131**(4): p. 694-705.
276. Zananiri, R., et al., *Synergy between RecBCD subunits is essential for efficient DNA unwinding.* eLife, 2019. **8**.
277. Terakawa, T., et al., *Sequential eviction of crowded nucleoprotein complexes by the exonuclease RecBCD molecular motor.* Proc. Natl. Acad. Sci. U.S.A., 2017. **114**(31): p. E6322-E6331.
278. Finkelstein, I.J., M.L. Visnapuu, and E.C. Greene, *Single-molecule imaging reveals mechanisms of protein disruption by a DNA translocase.* Nature, 2010. **468**(7326): p. 983-7.
279. Eggleston, A.K., et al., *Unwinding of nucleosomal DNA by a DNA helicase* J. Biol. Chem., 1995. **270**(5): p. 2024-2031.
280. Hall, M.A., et al., *High-resolution dynamic mapping of histone-DNA interactions in a nucleosome.* Nat. Struct. Mol. Biol., 2009. **16**(2): p. 124-9.
281. Jin, J., et al., *Synergistic action of RNA polymerases in overcoming the nucleosomal barrier.* Nat. Struct. Mol. Biol., 2010. **17**(6): p. 745-52.
282. Rudnizky, S., et al., *H2A.Z controls the stability and mobility of nucleosomes to regulate expression of the LH genes.* Nat. Commun., 2016. **7**: p. 12958.
283. Chédin, F. and S.C. Kowalczykowski, *A novel family of regulated helicases/nucleases from Gram-positive bacteria: insights into the initiation of DNA recombination* Mol. Microbiol., 2002. **43**(4): p. 823-834.
284. Sinha, K.M., et al., *AdnAB: a new DSB-resecting motor-nuclease from mycobacteria.* Genes Dev., 2009. **23**(12): p. 1423-37.

285. Jia, N., et al., *Structures and single-molecule analysis of bacterial motor nuclease AdnAB illuminate the mechanism of DNA double-strand break resection.* Proc. Natl. Acad. Sci. U.S.A., 2019. **116**(49): p. 24507-24516.
286. Long, J.E., N. Renzette, and S.J. Sandler, *Suppression of constitutive SOS expression by recA4162 (I298V) and recA4164 (L126V) requires UvrD and RecX in Escherichia coli K-12.* Mol. Microbiol., 2009. **73**(2): p. 226-39.
287. Park, J., et al., *PcrA helicase dismantles RecA filaments by reeling in DNA in uniform steps.* Cell, 2010. **142**(4): p. 544-55.
288. Fagerburg, M.V., et al., *PcrA-mediated disruption of RecA nucleoprotein filaments--essential role of the ATPase activity of RecA.* Nucleic Acids Res., 2012. **40**(17): p. 8416-24.
289. Menetski, J.P., D.G. Bear, and S.C. Kowalczykowski, *Stable DNA heteroduplex formation catalyzed by the Escherichia coli RecA protein in the absence of ATP hydrolysis.* Proc. Natl. Acad. Sci. U.S.A., 1990. **87**(1): p. 21-25.
290. Antony, E., et al., *Srs2 disassembles Rad51 filaments by a protein-protein interaction triggering ATP turnover and dissociation of Rad51 from DNA.* Mol. Cell, 2009. **35**(1): p. 105-15.
291. De Tullio, L., et al., *Yeast Srs2 helicase promotes redistribution of single-stranded DNA-bound RPA and Rad52 in homologous recombination regulation.* Cell Rep., 2017. **21**(3): p. 570-577.
292. Kaniecki, K., et al., *Dissociation of Rad51 presynaptic complexes and heteroduplex DNA joints by tandem assemblies of Srs2.* Cell Rep., 2017. **21**(11): p. 3166-3177.
293. Simandlova, J., et al., *FBH1 helicase disrupts RAD51 filaments in vitro and modulates homologous recombination in mammalian cells.* J. Biol. Chem., 2013. **288**(47): p. 34168-34180.
294. Amundsen, S.K., et al., *Small-molecule inhibitors of bacterial AddAB and RecBCD helicase-nuclease DNA repair enzymes.* ACS Chem. Biol., 2012. **7**(5): p. 879-91.
295. Wu, Y. and R.M. Brosh, Jr., *Helicase-inactivating mutations as a basis for dominant negative phenotypes.* Cell Cycle, 2010. **9**(20): p. 4080-90.
296. Chiolo, I., et al., *The human F-Box DNA helicase FBH1 faces Saccharomyces cerevisiae Srs2 and postreplication repair pathway roles.* Mol. Cell Biol., 2007. **27**(21): p. 7439-50.
297. Zhang, D.H., et al., *The human Pif1 helicase, a potential Escherichia coli RecD homologue, inhibits telomerase activity.* Nucleic Acids Res., 2006. **34**(5): p. 1393-404.
298. Bochman, M.L., N. Sabouri, and V.A. Zakian, *Unwinding the functions of the Pif1 family helicases.* DNA Repair (Amst), 2010. **9**(3): p. 237-49.
299. Ishino, Y., et al., *Nucleotide sequence of the iap gene, responsible for alkaline phosphatase isozyme conversion in Escherichia coli, and identification of the gene product* J. Bacteriol., 1987. **169**(12): p. 5429-5433.
300. Jinek, M., et al., *A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity* Science 2012. **337**: p. 816-821.
301. Barrangou, R., et al., *CRISPR provides acquired resistance against viruses in prokaryotes* Science, 2007. **315**: p. 1709-1712.
302. Brouns, S., et al., *Small CRISPR RNAs guide antiviral defense in prokaryotes* Science 2008. **321**: p. 960-964.
303. Marraffini, L.A. and E.J. Sontheimer, *CRISPR interference limits horizontal gene transfer in staphylococci by targeting DNA.* Science, 2008. **322**(5909): p. 1843-5.
304. Heck, C., et al., *Analytical epigenetics: single-molecule optical detection of DNA and histone modifications.* Curr. Opin. Biotechnol., 2019. **55**: p. 151-158.

305. Chen, H. and D.R. Larson, *What have single-molecule studies taught us about gene expression?* Genes Dev., 2016. **30**: p. 1796-1810.
306. Jones, D., et al., *Kinetics of dCas9 target search in Escherichia coli.* Science, 2017. **357**: p. 1420-1424.
307. Koonin, E.V., K.S. Makarova, and F. Zhang, *Diversity, classification and evolution of CRISPR-Cas systems.* Curr. Opin. Microbiol., 2017. **37**: p. 67-78.
308. Özcan, A., et al., *Type IV CRISPR RNA processing and effector complex formation in Aromatoleum aromaticum.* Nat. Microbiol., 2019. **4**(1): p. 89-96.
309. Sapranauskas, R., et al., *The Streptococcus thermophilus CRISPR/Cas system provides immunity in Escherichia coli.* Nucleic Acids Res., 2011. **39**(21): p. 9275-82.
310. Zetsche, B., et al., *Cpf1 Is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system.* Cell, 2015. **163**(3): p. 759-71.
311. Shmakov, S., et al., *Discovery and functional characterization of diverse class 2 CRISPR-Cas systems.* Mol. Cell, 2015. **60**(3): p. 385-97.
312. Garneau, J.E., et al., *The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA.* Nature, 2010. **468**(7320): p. 67-71.
313. Deltcheva, E., et al., *CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III.* Nature, 2011. **471**(7340): p. 602-7.
314. Anders, C., et al., *Structural basis of PAM-dependent target DNA recognition by the Cas9 endonuclease.* Nature, 2014. **513**(7519): p. 569-73.
315. Jiang, F., et al., *A Cas9-guide RNA complex preorganized for target DNA recognition* Science, 2015. **348**(6242): p. 1477-1481.
316. Jinek, M., et al., *Structures of Cas9 endonuclease reveal RNA-mediated conformational activation* Science 2014. **343**(6176).
317. Nishimasu, H., et al., *Crystal structure of Cas9 in complex with guide RNA and target DNA.* Cell, 2014. **156**(5): p. 935-49.
318. Jiang, F. and J.A. Doudna, *CRISPR-Cas9 Structures and Mechanisms* Annu. Rev. Biophys., 2017. **46**: p. 505-529.
319. Shibata, M., et al., *Real-space and real-time dynamics of CRISPR-Cas9 visualized by high-speed atomic force microscopy.* Nat. Commun., 2017. **8**(1): p. 1430.
320. Palermo, G., et al., *CRISPR-Cas9 conformational activation as elucidated from enhanced molecular simulations.* Proc. Natl. Acad. Sci. U.S.A., 2017. **114**(28): p. 7260-7265.
321. Raper, A.T., A.A. Stephenson, and Z. Suo, *Functional insights revealed by the kinetic mechanism of CRISPR/Cas9.* J. Am. Chem. Soc., 2018. **140**(8): p. 2971-2984.
322. Sternberg, S.H., et al., *Conformational control of DNA target cleavage by CRISPR-Cas9.* Nature, 2015. **527**(7576): p. 110-3.
323. Osuka, S., et al., *Real-time observation of flexible domain movements in CRISPR-Cas9.* EMBO J., 2018. **37**(10).
324. Gorman, J., et al., *Single-molecule imaging reveals target-search mechanisms during DNA mismatch repair.* Proc. Natl. Acad. Sci. U.S.A., 2012. **109**(45): p. E3074-83.
325. Blainey, P.C., et al., *Nonspecifically bound proteins spin while diffusing along DNA.* Nat. Struct. Mol. Biol., 2009. **16**(12): p. 1224-9.
326. Blainey, P.C., et al., *A base-excision DNA-repair protein finds intrahelical lesion bases by fast sliding in contact with DNA* Proc. Natl. Acad. Sci. U.S.A., 2006. **103**(15): p. 5752-5757.
327. Izeddin, I., et al., *Single-molecule tracking in live cells reveals distinct target-search strategies of transcription factors in the nucleus.* eLife, 2014. **3**.

328. Liao, Y., et al., *Single-molecule motions and interactions in live cells reveal target search dynamics in mismatch repair.* Proc. Natl. Acad. Sci. U.S.A., 2015. **112**(50): p. E6898-906.
329. Wang, Y.M., R.H. Austin, and E.C. Cox, *Single molecule measurements of repressor protein 1D diffusion on DNA.* Phys. Rev. Lett., 2006. **97**(4): p. 048302.
330. Ghodke, H., et al., *Single-molecule analysis reveals human UV-damaged DNA-binding protein (UV-DDB) dimerizes on DNA via multiple kinetic intermediates.* Proc. Natl. Acad. Sci. U.S.A., 2014. **111**(18): p. E1862-71.
331. Sternberg, S.H., et al., *DNA interrogation by the CRISPR RNA-guided endonuclease Cas9.* Nature, 2014. **507**(7490): p. 62-7.
332. Globyte, V., et al., *CRISPR/Cas9 searches for a protospacer adjacent motif by lateral diffusion.* EMBO J., 2019. **38**(4).
333. Hammar, P., et al., *Direct measurement of transcription factor dissociation excludes a simple operator occupancy model for gene regulation.* Nat. Genet., 2014. **46**(4): p. 405-8.
334. Jiang, F., et al., *Structures of a CRISPR-Cas9 R-loop complex primed for DNA cleavage* Science, 2016. **351**(6275): p. 867-871.
335. Hsu, P.D., et al., *DNA targeting specificity of RNA-guided Cas9 nucleases.* Nat. Biotechnol., 2013. **31**(9): p. 827-32.
336. Mali, P., et al., *CAS9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering.* Nat. Biotechnol., 2013. **31**(9): p. 833-8.
337. Gong, S., et al., *DNA unwinding is the primary determinant of CRISPR-Cas9 activity.* Cell Rep., 2018. **22**(2): p. 359-371.
338. Szczelkun, M., et al., *Direct observation of R-loop formation by single RNA-guide Cas9 and Cascade effector complexes* Proc. Natl. Acad. Sci. U.S.A., 2014. **111**(27): p. 9798-9803.
339. Rutkauskas, M., et al., *Single-molecule insight into target recognition by CRISPR-Cas complexes.* Methods Enzymol., 2017. **582**: p. 239-273.
340. Singh, D., et al., *Real-time observation of DNA recognition and rejection by the RNA-guided endonuclease Cas9.* Nat. Commun., 2016. **7**: p. 12778.
341. Lim, Y., et al., *Structural roles of guide RNAs in the nuclease activity of Cas9 endonuclease.* Nat. Commun., 2016. **7**: p. 13350.
342. Zeng, Y., et al., *The initiation, propagation and dynamics of CRISPR-SpyCas9 R-loop complex.* Nucleic Acids Res., 2018. **46**(1): p. 350-361.
343. Newton, M.D., et al., *DNA stretching induces Cas9 off-target activity.* Nat. Struct. Mol. Biol., 2019. **26**(3): p. 185-192.
344. Dagdas, Y.S., et al., *A conformational checkpoint between DNA binding and cleavage by CRISPR-Cas9.* Sci. Adv., 2017. **3**(8).
345. Yang, M., et al., *The conformational dynamics of Cas9 governing DNA cleavage are revealed by single-molecule FRET.* Cell Rep., 2018. **22**(2): p. 372-382.
346. Richardson, C.D., et al., *Enhancing homology-directed genome editing by catalytically active and inactive CRISPR-Cas9 using asymmetric donor DNA.* Nat. Biotechnol., 2016. **34**(3): p. 339-44.
347. Clarke, R., et al., *Enhanced bacterial immunity and mammalian genome editing via RNA polymerase-mediated dislodging of Cas9 from double strand DNA breaks* Mol. Cell, 2018. **71**(1): p. 42-55.
348. Yourik, P., et al., *Staphylococcus aureus Cas9 is a multiple-turnover enzyme* RNA, 2019. **25**(1): p. 35-44.
349. Whinn, K.S., A.M. van Oijen, and H. Ghodke, *Spy-ing on Cas9: Single-molecule tools reveal the enzymology of Cas9.* Curr. Opin. Biomed. Engin., 2019. **12**: p. 25-33.
350. Okafor, I.C. and T. Ha, *Single Molecule FRET Analysis of CRISPR Cas9 Single Guide RNA Folding Dynamics.* J. Phys. Chem. B., 2023. **127**(1): p. 45-51.

351. Yang, M., et al., *Nonspecific interactions between SpCas9 and dsDNA sites located downstream of the PAM mediate facilitated diffusion to accelerate target search.* Chem. Sci., 2021. **12**(38): p. 12776-12784.
352. Cofsky, J.C., et al., *CRISPR-Cas9 bends and twists DNA to read its sequence.* Nat. Struct. Mol. Biol., 2022. **29**(4): p. 395-402.
353. Ivanov, I.E., et al., *Cas9 interrogates DNA in discrete steps modulated by mismatches and supercoiling.* Proc. Natl. Acad. Sci. U.S.A., 2020.
354. Zhu, X., et al., *Cryo-EM structures reveal coordinated domain motions that govern DNA cleavage by Cas9.* Nat. Struct. Mol. Biol., 2019. **26**(8): p. 679-685.
355. Wang, Y., et al., *Real-time observation of Cas9 postcatalytic domain motions.* Proc. Natl. Acad. Sci. U.S.A., 2021. **118**(2).
356. Monachino, E., et al., *Design of DNA rolling-circle templates with controlled fork topology to study mechanisms of DNA replication.* Anal. Biochem., 2018. **557**: p. 42-45.
357. Tuan, P.M., et al., *Direct visualization of translesion DNA synthesis polymerase IV at the replisome.* Proc. Natl. Acad. Sci. U.S.A., 2022. **119**(39): p. e2208390119.
358. Geertsema, H.J., K.E. Duderstadt, and A.M. van Oijen, *Single-molecule observation of prokaryotic DNA replication*, in *Methods Mol. Biol.*, S. Vengrova and J. Dalgaard, Editors. 2015, Springer New York: New York, NY. p. 219-238.
359. Xu, J., et al., *Automated stoichiometry analysis of single-molecule fluorescence imaging traces via deep learning.* J. Am. Chem. Soc., 2019. **141**(17): p. 6976-6985.
360. Thomsen, J., et al., *DeepFRET, a software for rapid and automated single-molecule FRET data classification using deep learning.* eLife, 2020. **9**.
361. Schindelin, J., et al., *Fiji: an open-source platform for biological-image analysis.* Nat. Methods, 2012. **9**(7): p. 676-82.
362. Watkins, L.P. and H. Yang, *Detection of intensity change points in time-resolved single-molecule measurements.* J. Phys. Chem. , 2005. **109**(1): p. 617-628.
363. Duderstadt, K.E., et al., *Simultaneous real-time imaging of leading and lagging strand synthesis reveals the coordination dynamics of single replisomes.* Mol. Cell, 2016. **64**(6): p. 1035-1047.
364. Hill, F.R., A.M. van Oijen, and K.E. Duderstadt, *Detection of kinetic change points in piece-wise linear single molecule motion.* J. Chem. Phys., 2018. **148**(12): p. 123317.
365. Kaur, G., et al., *Production of long linear DNA substrates with site-specific chemical lesions for single-molecule replisome studies*, in *Methods Enzymol.* 2022, Academic Press. p. 299-315.
366. Mueller, S.H., et al., *Design of customizable long linear DNA substrates with controlled end modifications for single-molecule studies.* Anal. Biochem., 2020. **592**: p. 113541.
367. Larsen, N.B., I.D. Hickson, and H.W. Mankouri, *Tus-Ter as a tool to study site-specific DNA replication perturbation in eukaryotes.* Cell Cycle, 2014. **13**(19): p. 2994-8.
368. Geertsema, H.J., et al., *Single-molecule studies of polymerase dynamics and stoichiometry at the bacteriophage T7 replication machinery.* Proc. Natl. Acad. Sci. U.S.A., 2014. **111**(11): p. 4073-8.
369. Schröder, C.H., E. Erben, and H. Kaerner, *A rolling circle model of the in vivo replication of bacteriophage φχ174 replicative form DNA: Different fate of two types of progeny replicative form.* J. Mol. Biol., 1973. **79**(4): p. 599-613.
370. Spinks, R.R., L.M. Spenkelink, and A.M. van Oijen, *Single-molecule fluorescence methods to study protein exchange kinetics in supramolecular complexes*, in *Single Stranded DNA Binding Proteins*, M.T. Oliveira, Editor. 2021, Springer US: New York, NY. p. 49-65.

371. Fay, P.J., et al., *Size classes of products synthesized processively by two subassemblies of Escherichia coli DNA polymerase III holoenzyme.* J. Biol. Chem., 1982. **257**(10): p. 5692-5699.
372. Tanner, N.A., et al., *E. coli DNA replication in the absence of free β clamps.* EMBO J., 2011. **30**(9): p. 1830-40.
373. Mirkin, E.V. and S.M. Mirkin, *Replication fork stalling at natural impediments.* Microbiol. Mol. Biol. Rev., 2007. **71**(1): p. 13-35.
374. Whinn, K.S., A.M. van Oijen, and H. Ghodke, *Single-molecule studies of helicases and translocases in prokaryotic genome-maintenance pathways.* DNA Repair (Amst), 2021. **108**: p. 103229.
375. Gao, Y. and W. Yang, *Different mechanisms for translocation by monomeric and hexameric helicases.* Curr. Opin. Struct. Biol., 2020. **61**: p. 25-32.
376. Sparks, M.A., P.M. Burgers, and R. Galletto, *Pif1, RPA, and FEN1 modulate the ability of DNA polymerase δ to overcome protein barriers during DNA synthesis.* J. Biol. Chem., 2020. **295**(47): p. 15883-15891.
377. Chib, S., A.K. Byrd, and K.D. Raney, *Yeast helicase Pif1 unwinds RNA:DNA hybrids with higher processivity than DNA:DNA duplexes.* J. Biol. Chem., 2016. **291**(11): p. 5889-5901.
378. Oakley, A.J., et al., *Flexibility revealed by the 1.85Å crystal structure of the β sliding-clamp subunit of Escherichia coli DNA polymerase III.* Acta Crystallogr. D Biol. Crystallogr., 2003. **59**: p. 1192-1199.
379. Mason, C.E., et al., *Escherichia coli single-stranded DNA-binding protein: NanoESI-MS studies of salt-modulated subunit exchange and DNA binding transactions.* J. Am. Soc. Mass Spectrom., 2013. **24**(2): p. 274-85.
380. Stamford, N.P.J., P.E. Lilley, and N.E. Dixon, *Enriched sources of Escherichia coli replication proteins: The dnaG primase is a zinc metalloprotein.* Biochim. Biophys. Acta, 1992. **1132**(1): p. 17-25.
381. Tanner, N.A., et al., *Single-molecule studies of fork dynamics in Escherichia coli DNA replication.* Nat. Struct. Mol. Biol., 2008. **15**(2): p. 170-6.
382. Neylon, C., 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 2000. **39**(39): p. 11989-11999.
383. Rasnik, I., et al., *DNA-binding orientation and domain conformation of the E.coli Rep helicase monomer bound to a partial duplex junction: Single-molecule studies of fluorescently labeled enzymes.* J. Mol. Biol., 2004. **336**(2): p. 395-408.
384. Kim, Y., et al., *Efficient site-specific labeling of proteins via cysteines.* Bioconjug. Chem. , 2008. **19**(3): p. 786-791.
385. Lohman, T.M., *Escherichia coli DNA helicases: mechanisms of DNA unwinding* Mol. Microbiol., 1992. **6**(1): p. 5-14.
386. Chao, K. and T.M. Lohman, *DNA-induced dimerization of the Escherichia coli Rep helicase.* J. Mol. Biol., 1991. **221**(4): p. 1165-1181.
387. Wong, I., et al., *DNA-induced dimerization of the Escherichia coli Rep helicase.* J. Biol. Chem., 1992. **267**(11): p. 7596-7610.
388. Bjornson, K.P., K.J. Moore, and T.M. Lohman, *Kinetic mechanism of DNA binding and DNA-induced dimerization of the Escherichia coli Rep helicase.* Biochemistry, 1996. **35**(7): p. 2268-82.
389. Wong, I. and T.M. Lohman, *Allosteric effects of nucleotide cofactors on Escherichia coli Rep helicase-DNA binding.* Science, 1992. **256**(5055): p. 350-355.
390. Arai, N. and A. Kornberg, *Rep protein as a helicase in an active, isolatable replication fork of duplex ΦX174 DNA.* J. Biol. Chem., 1981. **256**(10): p. 5294-5298.
391. Amaratunga, M. and T.M. Lohman, *Escherichia coli Rep helicase unwinds DNA by an active mechanism.* Biochemistry 1993. **32**(27): p. 6815-20.

392. Lane, H.E. and D.T. Denhardt, *The rep mutation: IV. Slower movement of replication forks in Escherichia coli rep strains.* J. Mol. Biol., 1975. **97**(1): p. 99-112.
393. Maisnier-Patin, S., K. Nordström, and S. Dasgupta, *Replication arrests during a single round of replication of the Escherichia coli chromosome in the absence of DnaC activity.* Mol. Microbiol., 2001. **42**(5): p. 1371-82.
394. Boyle, E.A., et al., *High-throughput biochemical profiling reveals sequence determinants of dCas9 off-target binding and unbinding.* Proc. Natl. Acad. Sci. U.S.A., 2017. **114**(21): p. 5461-5466.
395. Spinks, R.R., et al., *Single-molecule insights into the dynamics of replicative helicases.* Front. Mol. Biosci., 2021. **8**: p. 741718.
396. Mohapatra, S., et al., *Single-molecule analysis and engineering of DNA motors.* Chem. Rev., 2020. **120**(1): p. 36-78.
397. Ghodke, H., H. Ho, and A.M. van Oijen, *Single-molecule live-cell imaging of bacterial DNA repair and damage tolerance.* Biochem. Soc. Trans., 2018. **46**(1): p. 23-35.
398. Monachino, E., L.M. Spenkelink, and A.M. van Oijen, *Watching cellular machinery in action, one molecule at a time.* J. Cell Biol., 2017. **216**(1): p. 41-51.
399. French, S., *Consequences of replication fork movement through transcription units in vivo.* Science, 1992. **258**(5086): p. 1362-1365.
400. Lang, K.S., et al., *Replication-Transcription Conflicts Generate R-Loops that Orchestrate Bacterial Stress Survival and Pathogenesis.* Cell, 2017. **170**(4): p. 787-799 e18.
401. Lang, K.S. and H. Merrikh, *The Clash of Macromolecular Titans: Replication-Transcription Conflicts in Bacteria.* Annu. Rev. Microbiol., 2018. **72**: p. 71-88.
402. Mangiameli, S.M., et al., *Transcription leads to pervasive replisome instability in bacteria.* eLife, 2017. **6**.
403. Million-Weaver, S., et al., *An underlying mechanism for the increased mutagenesis of lagging-strand genes in Bacillus subtilis.* Proc. Natl. Acad. Sci. U.S.A., 2015. **112**(10): p. E1096-105.
404. Million-Weaver, S., A.N. Samadpour, and H. Merrikh, *Replication restart after replication-transcription conflicts requires RecA in Bacillus subtilis.* J. Bacteriol., 2015. **197**(14): p. 2374-82.
405. Paul, S., et al., *Accelerated gene evolution through replication-transcription conflicts.* Nature, 2013. **495**(7442): p. 512-5.
406. Sankar, T.S., et al., *The nature of mutations induced by replication-transcription collisions.* Nature, 2016. **535**(7610): p. 178-81.
407. Vilette, D., S.D. Ehrlich, and B. Michel, *Transcription-induced deletions in plasmid vectors: M13 DNA replication as a source of instability.* Mol. Gen. Genet., 1996. **252**(4): p. 398-403.
408. Gwynn, E.J., et al., *The conserved C-terminus of the PcrA/UvrD helicase interacts directly with RNA polymerase.* PLoS One, 2013. **8**(10): p. e78141.
409. Yokota, H., *DNA-unwinding dynamics of Escherichia coli UvrD lacking the C-terminal 40 amino acids.* Biophys. J., 2020. **118**(7): p. 1634-1648.
410. George, J.W., R.M. Brosh, Jr., and S.W. Matson, *A dominant negative allele of the Escherichia coli uvrD gene encoding DNA helicase II: A biochemical and genetic characterization.* J. Mol. Biol., 1994. **235**(2): p. 424-435.
411. Kawale, A.A. and B.M. Burmann, *UvrD helicase-RNA polymerase interactions are governed by UvrD's carboxy-terminal Tudor domain.* Commun. Biol., 2020. **3**(1): p. 607.
412. Boulé, J.B. and V.A. Zakian, *The yeast Pif1p DNA helicase preferentially unwinds RNA DNA substrates.* Nucleic Acids Res., 2007. **35**(17): p. 5809-18.

413. Matson, S.W., *Escherichia coli DNA helicase II (uvrD gene product) catalyzes the unwinding of DNA·RNA hybrids in vitro.* Proc. Natl. Acad. Sci. U.S.A., 1989. **86**(12): p. 4430-4434.

414. Morris, P.D. and K.D. Raney, *DNA helicase displace streptavidin from biotin-labeled oligonucleotides* Biochemistry 1999. **38**(16): p. 5164-5171.

415. Sokoloski, J.E., et al., *Chemo-mechanical pushing of proteins along single-stranded DNA.* Proc. Natl. Acad. Sci. U.S.A., 2016. **113**(22): p. 6194-9.

Chapter 9: Appendix I

Nuclease dead Cas9 is a programmable roadblock for DNA replication

Kelsey S. Whinn, Gurleen Kaur, Jacob S. Lewis, Grant D. Schauer, Stefan H. Müller, Slobodan Jergic, Hamish Maynard, Zhong Yan Gan, Matharishwan Naganbabu, Marcel P. Bruchez, Michael E. O'Donnell, Nicholas E. Dixon, Antoine M. van Oijen, Harshad Ghodke

Published in *Scientific Reports,* 2019*,* 10.1038/s41598-019-49837-z

Limited experimental tools are available to study the consequences of collisions between DNA-bound molecular machines. Here, we repurpose a catalytically inactivated Cas9 (dCas9) construct as a generic, novel, targetable protein–DNA roadblock for studying mechanisms underlying enzymatic activities on DNA substrates *in vitro*. We illustrate the broad utility of this tool by demonstrating replication fork arrest by the specifically bound dCas9–guideRNA complex to arrest viral, bacterial and eukaryotic replication forks *in vitro*.

Contributions

This chapter describes the characterization of the dCas9 roadblock used in Chapters 4, 5 and 6, in ensemble and single-molecule DNA replication assays. The results of this chapter were presented in my Honors thesis, submitted in 2018 to the University of Wollongong, and as such do not form part of my PhD thesis. Reproduction of the work in this appendix serves to inform the reader on the characterization and use of the dCas9 roadblock as a critical method in this thesis.

I carried out all the *E. coli* bulk ensemble assays and Surface Plasmon Resonance investigations (under the supervision of Dr. Slobodan Jergic) presented in this study, compiled the results and analysis, and drafted the original manuscript.

9.1. Introduction

Enzymes that regulate and execute the reactions that govern life must contend with a host of other DNA-binding proteins as they perform their functions. Obtaining a detailed mechanistic understanding of how these reactions are performed in conditions approaching physiological contexts, demands an exquisite ability to precisely manipulate strand and substrate occupancy by DNA-binding proteins. Several examples of roadblocks are described in the literature that have been proven invaluable for interrogating a variety of molecular mechanisms – from understanding how site-specifically bound proteins may confine the diffusion of proteins translocating on DNA, to blocking the enzymatic activity of transcription elongation complexes, or determining whether enzymes such as ring-shaped helicases can transiently open to overcome barriers on DNA [1-5].

The impediment of the progress of DNA replication machinery on template DNA occupied by proteins is an important case in point. DNA replication occurs on chromosomal DNA while processes such as DNA repair, recombination and transcription continue. Replisomes encounter three major types of protein barriers: transcription complexes, nucleoid-associated proteins, and recombination filaments [6-8]. Successful replication across such roadblocks requires the coordinated action of several accessory factors and DNA-repair and dedicated restart proteins. Improper resolution of arrested forks can lead to replication fork collapse and eventually, genetic instability [3, 9, 10].

Several roadblocks have been developed to mimic encounters between replication forks and protein barriers. Inspired by the Tus-*ter* block that terminates replication in *Escherichia coli*, replication fork arrest has been studied at *ter* sites recombined into the *Saccharomyces cerevisiae* chromosome [11]. Other approaches have involved the introduction of repeat sequences that enable binding of transcription factors to artificially introduce repressor/operator arrays, or proteins that polymerize to form nucleoprotein filaments [4, 12-14]. Despite their tremendous utility in studying replication fork arrest, these methods suffer from several disadvantages: since the tandem binding of several roadblock proteins is required for effective stalling of the replication fork, the exact positions of the block are often poorly defined. Further, tedious recombination procedures are required to incorporate tandem arrays of terminator or repressor/operator sequences. Finally, high local concentrations of the fluorescently tagged roadblock may influence the local structure of the DNA due to a residual ability for the genetic fluorescent protein fusion to oligomerize. These limitations call for the development of a generic fluorescent protein roadblock that is monomeric, binds DNA with high affinity and specificity, and does not require extensive genetic manipulation of template DNA. Here, we describe the construction and validation of a fluorescently tagged nuclease dead Cas9 construct that serves as a monomeric roadblock for use in *in vitro* assays. Nuclease dead Cas9 blocks the progress of replication forks from viral, bacterial and eukaryotic model replisomes reconstituted *in vitro*.

9.2. Results

9.2.1. Construction of a stable roadblock that can be observed on long timescales in vitro and in vivo

We reasoned that target bound, catalytically inactivated *Streptococcus pyogenes* Cas9 (dCas9) could act as a versatile roadblock enabling easy and precise targeting, and control over site-, orientation- and strand-specific binding to template DNA. Additionally, to permit long-term visualization of nucleic acid processing enzymes at sites of dCas9 roadblocks *in vitro*, we genetically fused dCas9 to the photostable fluoromodule dL5 that becomes fluorescent upon binding the dye, malachite green [15, 16]. The fluorogen used in this work is an ester modified variant of the malachite green dye, herein referred to as malachite green-ester (MGE) (Figure 9.1A).

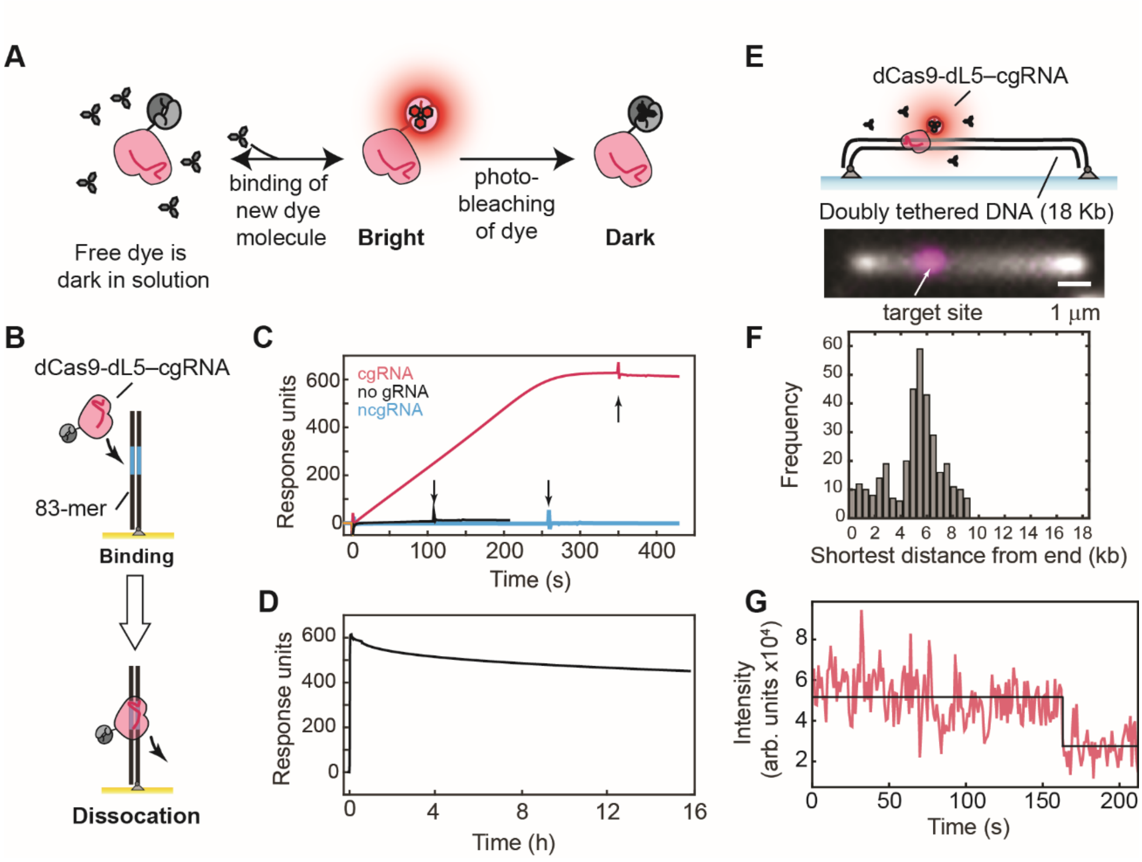


Figure 9.1. Characterization of dCas9-dL5. A) Schematic of the dCas9-dL5 probe. Free dye is dark in solution. Binding of MGE to the dL5 tag enables visualization of dCas9-dL5. **B)** Schematic of dCas9-dL5 binding to immobilized dsDNA containing the target sequence on an SPR chip. **C)** Sensorgram describing the binding of dCas9-dL5 to dsDNA substrate carrying the target sequence in the absence of gRNA or programmed with a complementary gRNA (cgRNA1) or non-complementary gRNA (ncgRNA). Arrows indicate the completion of the injection phase, and switch to running buffer. *N* = 1 independent experiment. **D)** Dissociation of dCas9-dL5–cgRNA1 bound to the dsDNA target monitored over 16 h. **E)** Schematic and examples of elongated surface bound and elongated linear dsDNA template bound to dCas9-dL5 (scale bar – 1 μm). dsDNA is stained using Sytox orange, and dCas9-dL5–cgRNA1 is stained by MGE. **F)** Histogram

of detected position of dCas9-dL5–cgRNA1 complex visualized by addition of MGE (n = 345 molecules). The shortest distance to the position of the dCas9-dL5 is plotted here. **G)** Example photo-bleaching trajectory of dCas9-dL5–cgRNA1–MGE complex (n = 345).

9.2.2. In vitro characterization of dCas9-dL5 binding to DNA

First, we purified the dCas9-dL5 fusion protein (Supplementary Table 9.S1 and Figure 9.S1) and assayed its binding to an 83-mer target DNA using surface plasmon resonance (SPR) (Figure 9.1B, Supplementary Table S2). Biotinylated target DNA was immobilized on a streptavidin-coated surface and a solution containing dCas9-dL5 pre-programmed with a complementary guide RNA (cgRNA1) was introduced (Figure 9.1B, Supplementary Table 9.S3). The dCas9-dL5–cgRNA1 complex exhibited robust and stable binding to the target DNA, whereas dCas9-dL5 alone or in presence of a non-complementary gRNA (ncgRNA) did not exhibit appreciable binding (Figure 9.1C, Supplementary Table 9.S2). We found that highly purified dCas9-dL5 alone exhibited binding to 83-mer biotinylated dsDNA in the absence of guide RNA (Figure 9.1C), consistent with previous work [17]. This minimal binding was lost when dCas9-dL5 was programmed with ncgRNA and may reflect non-specific association of dCas9-dL5 for dsDNA ends. Importantly, dCas9-dL5–cgRNA associated strongly and stably with the target DNA – only approximately 25% of the bound complexes dissociated over 16 h (Figure 9.1D; Supplementary Methods).

Next, we confirmed that dCas9-dL5 binds specifically to its target sequence. We used single-molecule total internal reflectance fluorescence (TIRF) microscopy to directly visualize dCas9-dL5 bound to its target sequence on individual DNA molecules. DNA molecules were pre-incubated with dCas9-dL5–cgRNA and doubly tethered to a streptavidin-coated glass coverslip inside a microfluidic flow cell using biotinylated oligonucleotide handles (Figure 9.1E) [18]. Addition of MGE into the flow cell enabled visualization of the dL5 tag, and positioning of the dCas9-dL5–cgRNA complex along the length of the DNA (Supplementary Methods; 349 out of 899 DNA templates had dCas9-dL5 bound). Consistent with previous work, the position of the bound dCas9-dL5–cgRNA complex was in good agreement with the expected position (Figure 9.1F) [17]. The spread in the position of the dCas9 may be attributable to incomplete extension of the doubly-tethered substrates on the surface of the flow cell. The use of the MGE allowed us to reliably visualize target-bound dCas9-dL5 for several minutes (Figure 9.1G).

9.2.3. Target-bound dCas9-dL5 blocks DNA replication in bulk measurements

These observations highlight the potential of dCas9-dL5 to be applied as a general roadblock to study details of molecular transactions on DNA *in vitro*. As a proof of principle, we ran reconstituted replisomes from model systems into this dCas9 roadblock. First, we investigated whether single dCas9-dL5–cgRNA1 molecules bound to template DNA could impede DNA replication using a rolling-circle replication assay, both at the ensemble and single-molecule levels[19-23]. This assay allows observation of robust DNA synthesis by replisomes under a variety of experimental conditions (Figure 9.2A, see Supplementary Figure 9.S3 online for raw data). Pre-incubation of template DNA with dCas9-dL5–cgRNA1 resulted in potent replication fork arrest of reconstituted *E. coli* replisomes during either leading-strand (Figure 9.2A and Supplementary Figure 9.S2) or simultaneous leading- and lagging-strand DNA synthesis (Supplementary Figure 9.S2), with an average blocking efficiency of 85 ± 2% (*N* [replicates] = 5).Importantly, neither complementary gRNA alone (Figure 9.2A and Supplementary Figure 9.S2 online) nor dCas9-dL5 alone (Figure 9.2A, see Supplementary Figure 9.S2 and Supplementary Methods online) or programmed with ncgRNA (Figure 9.2A and see Supplementary Figure 9.S2 online) could site-specifically arrest DNA replication (summarized in Figure 9.2H). Further, dCas9-dL5 targeted to the leading strand using a complementary gRNA duplex (cgRNA4 (Ld)) blocked *E. coli* leading-strand (Figure 9.2A) and leading- and lagging-strand synthesis with similar efficiencies (85 ± 2% (*N* [replicates] = 5) (see Supplementary Figure 9.S2 online). Taken together, these observations demonstrate that encounters of the replisome with either the PAM-proximal (cgRNA1 (Lg)) or PAM-distal (cgRNA4 (Ld)) side of bound dCas9-dL5–cgRNA complexes does not influence its ability to arrest replication.

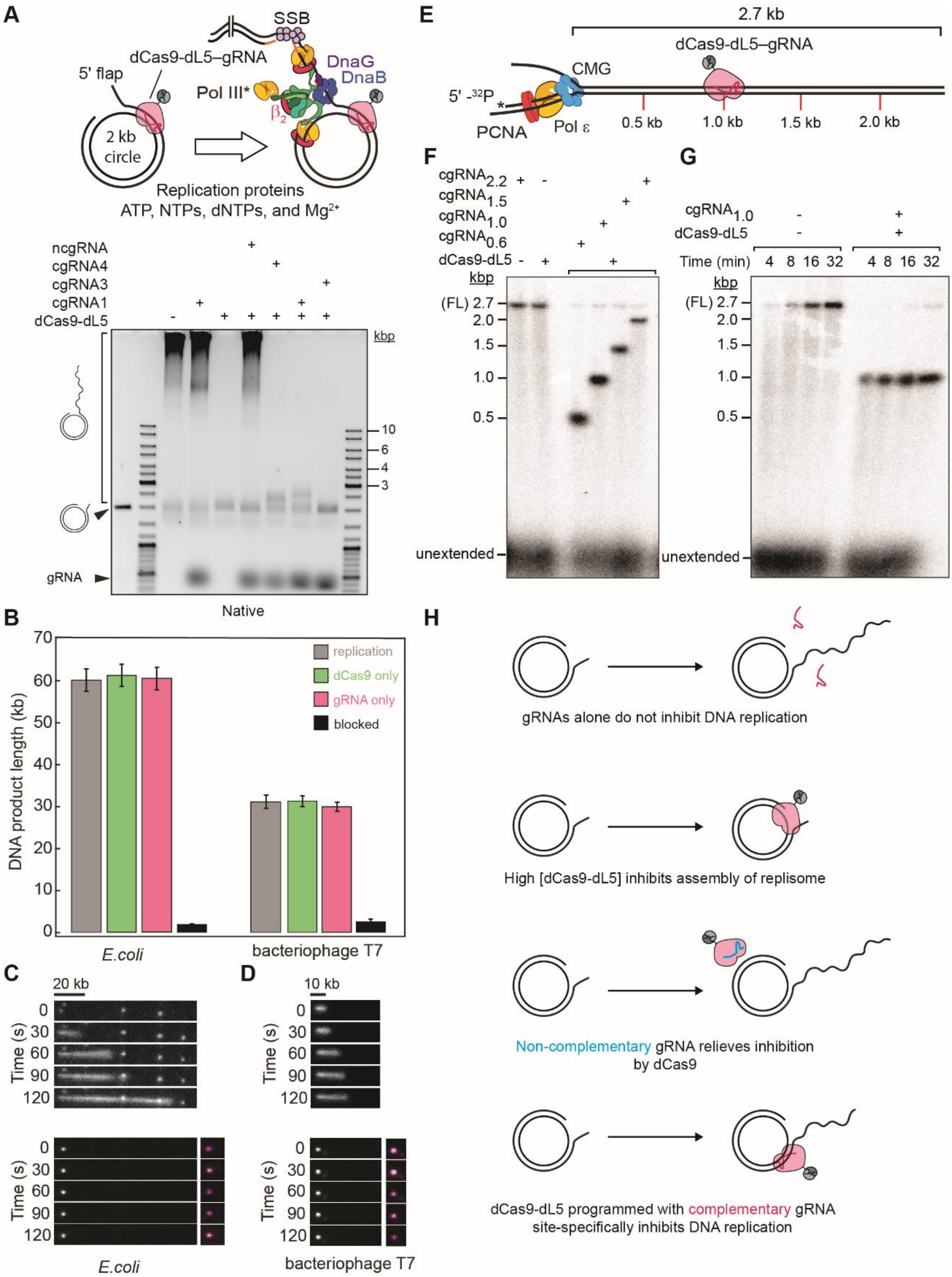


Figure 9.2. dCas9-dL5 efficiently and stably blocks bacterial, viral, and eukaryotic DNA replication regardless of the targeted strand. A) Schematic of the rolling-circle DNA replication assay. Addition of the *E. coli* or T7 replication proteins, nucleotides, and Mg^{2+} initiates DNA synthesis. The DNA products are separated by gel electrophoresis by staining with SYBR-Gold, or visualized by single-molecule fluorescence microscopy by staining with Sytox orange. dCas9-dL5 (100 nM) programmed with ncgRNA (400 nM) and cgRNAs alone (400 nM) alone do not inhibit DNA replication. At high concentrations, dCas9-dL5 (100 nM)

alone inhibits DNA synthesis. See also, Figure S2D for dCas9-dL5 titration. dCas9-dL5 programmed with cgRNAs arrest the progress of the replication fork at the target site. **B)** Bar plots of mean DNA product lengths from *E. coli* and T7 single-molecule rolling-circle DNA replication assays. Values plotted are derived from exponential fits to single-molecule DNA product length distributions (*n* > 91 molecules). Error bars indicate errors of the fit. **C)** and **D)** (Top panel) Example kymographs of an individual DNA molecule undergoing DNA replication by *E. coli* (C) (*n* = 177 molecules; replication efficiency of 26 ± 2%) and T7 replisomes D) (*n* = 136 molecules; replication efficiency of 24 ± 2%) in the absence of target bound dCas9-dL5. (Bottom panels) Example kymographs of an individual DNA molecule arrested by target bound dCas9-dL5. The grey scale indicates the fluorescence intensity of stained DNA and magenta indicates dCas9-dL5–cgRNA stained by MGE. No replication events were detected. **E)** Schematic of the eukaryotic DNA replication assay. Eukaryotic replication is blocked by dCas9-dL5 at specific positions on the replication template. **F)** dCas9-dL5 efficiently blocks eukaryotic replication. The cgRNAs used to specifically target the template are indicated. cgRNA$_{0.6}$ and cgRNA$_{2.2}$ block the leading strand; cgRNA$_{1.0}$ and cgRNA$_{1.5}$ block the lagging strand (see Supplementary Methods for details). All reactions were stopped at 16 min. **G)** Time course of eukaryotic replication in the presence or absence of dCas9-dL5 and cgRNA$_{1.0}$. Reactions were stopped at 4, 8, 16 and 32 min as indicated. **H)** Summary of interactions of dCas9-dL5 and template DNA. Only the correctly programmed dCas9-dL5-cgRNA complex site specifically inhibits DNA replication.

9.2.4. Target-bound dCas9-dL5 blocks DNA replication in single-molecule assays

Next, to demonstrate the use of this tool in single-molecule assays, we repeated these experiments in single-molecule rolling-circle assays and measured the average lengths of DNA products synthesized by individual *E. coli* replisomes in the presence of dCas9-dL5–cgRNA complexes. Consistent with the bulk experiments, target bound dCas9-dL5 was found to specifically block simultaneous leading- and lagging-strand DNA synthesis (Figure 9.2B and Figure 9.2C).

Finally, we examined the capacity of dCas9-dL5 as a universal roadblock for arresting replication forks site-specifically; the ability of dCas9-dL5 programmed with complementary gRNA to arrest replication *in vitro* was assessed using model replisomes from T7 bacteriophage (Figure 9.2B and 9.2D) and *S. cerevisiae* (Figure 9.2E–G, see Supplementary Figure 9.S3 for raw data and Supplementary Methods online). Replication reactions using both reconstituted replisomes carried out in the presence of template associated dCas9-dL5–cgRNA also exhibited replication fork arrest as observed with *E. coli*. Taken together, these results demonstrate that the dCas9-dL5–cgRNA complex binds with high specificity and stability to its target DNA sequence and can be visualized effectively in a variety of experimental conditions.

9.3. Discussion

Here, we have harnessed the specificity and programmability of the CRISPR/Cas9 system and combined it with the photo-stability of the dL5 fluoromodule to repurpose dCas9 as a tool for studying metabolic processes that occur on DNA. Our *in vitro* characterization of dCas9-dL5 binding to dsDNA indicated that this tool is a high-stability and sequence-specific roadblock. Indeed, the stability of this roadblock could be affected by PAM-distal mismatches between the gRNA and DNA, as shown previously in investigations of WT Cas9 stability [24, 25]. Interestingly, replication fork arrest is not strand-specific and occurs when dCas9-dL5 is targeted to either the leading- or lagging-strands, suggesting that the dCas9:gRNA complex may inhibit strand separation by the replicative helicase. This is in contrast to recent investigations showing that the elongating RNA polymerase is able to displace WT or nuclease dead Cas9 proteins stably bound to the template strand, but not on the non-template strand [26, 27]. Here, we demonstrate the suitability of the dCas9-dL5 tool for investigating mechanisms that underlie the protein dynamics that govern replication fork rescue at sites of protein roadblocks on template DNA undergoing replication by viral, bacterial, and eukaryotic replisomes. Indeed, this precisely tunable roadblock may prove useful in understanding fine mechanistic details of DNA helicase and translocases, repair and other sliding factors involved in DNA metabolism.

9.4. Materials and Methods

9.4.1. Replication proteins

E. coli DNA replication proteins were produced as described previously: the β_2 sliding clamp[28], SSB[29], the $DnaB_6(DnaC)_6$ helicase–loader complex[30], DnaG primase[31], the Pol III $\tau_3\delta\delta'\chi\psi$ clamp loader[32] and Pol III $\alpha\varepsilon\theta$ core[22]. *S. cerevisiae* DNA replication proteins were produced as described previously: the CMG (Cdc45/Mcm2-7/GINS) helicase [33], the Mrc1–Tof1–Csm3 (MTC) complex[34], DNA polymerase Pol ε[33], the PCNA sliding clamp[35], RPA[33] and the RFC clamp loader[36]. T7 gp2.5 was produced as described previously [37]. Highly purified T7 gp4 helicase and DNA polymerase gp5/trx were generous gifts of Charles Richardson.

9.4.2. DNA and RNA oligonucleotides.

DNA oligonucleotides and tracrRNA, unmodified crRNAs and crRNAs containing Alexa Fluor 555 were purchased from Integrated DNA Technologies (USA). Sequences of DNA oligonucleotides, crRNAs and tracrRNAs used in this study are listed in Table S2. Synthetic guide RNA (gRNA) targeting various regions of the 2.7 kb linear DNA template were produced with the EnGen sgRNA Synthesis Kit (New England Biolabs, USA) using the DNA are also described in the table S2. All DNA and RNA oligonucleotides were stored in TE buffer (10 mM Tris-HCl pH 8.0, 1 mM EDTA) at –20°C.

9.4.3. Construction of plasmid pJL001.

Plasmid pJL001 was constructed by ligation of a 1007 bp *Sac*I–*Xho*I gene block (Aldervon, USA) between the corresponding sites in pET302 (obtained from Addgene plasmid #72269), to encode dCas9-dL5 containing an N-terminal 6xHis and C-terminal 3xFLAG tags. The corresponding plasmid sequence is given in Table S1.

9.4.4. Expression and purification of dCas9-dL5.

E. coli strain Rosetta 2(DE3) containing plasmid pJL001 was grown in LB medium supplemented with thymine (25 mg/mL) and ampicillin (100 μg/mL) at 37°C. Upon growth to A_{600} = 0.8, the temperature was reduced to 16°C and protein expression induced by addition of 0.5 mM isopropyl-β-D-thiogalactoside. Cultures were further shaken for 16 h at 16°C, then chilled on ice. Cells (8 g from 2 L of culture) were harvested by centrifugation, frozen in liquid nitrogen and stored at –80°C. All subsequent steps were carried out in a cold room maintained at 6°C. After thawing, cells were resuspended in lysis buffer (20 mM Tris-HCl pH 7.6, 0.1 mM EDTA, 1 mM dithiothreitol, 150 mM NaCl, 5% (*v/v*) glycerol) and 2x Protease Inhibitor Cocktail tablets and 0.7 mM phenylmethylsulfonyl fluoride were added to inhibit proteolysis. Cells were lysed by being passed twice through a French press (12,000 psi), and cell debris were then removed by centrifugation. Crude supernatant (85 mL) was brought to 0.4% (*v/v*) in polyethylenimine (PEI) and vigorously stirred. After 40 min, the white precipitate was separated by centrifugation. The remaining pellet was homogenized by stirring in lysis buffer for 15 min. The remaining white precipitate was immediately collected by centrifugation and the supernatant discarded. The remaining pellet was further homogenized in lysis buffer + 250 mM NaCl for 15 min. After centrifugation, the high salt supernatant containing dCas9-dL5 was collected yielding Fraction I (72 mL). Proteins that were precipitated from Fraction I by addition of solid ammonium sulfate (0.32 g/mL) and stirring for 60 min, were collected by centrifugation and dissolved in 30 mL of FLAG buffer (25 mM Tris-HCl pH 7.6, 1 mM EDTA, 1 mM dithiothreitol, 200 mM NaCl and 5% (*v/v*) glycerol). The solution was dialyzed against 2 L of the same buffer overnight, to yield Fraction II. Fraction II was added to 4 mL FLAG M2 resin prepared as per manufacturer’s instructions and left to incubate with constant mixing. After 1 h, the FLAG M2 resin was poured into a PD-10 column and equilibrated in FLAG wash buffer (50 mM Tris-HCl pH 7.6, 0.5 mM dithiothreitol, 0.5 mM EDTA, 300 mM NaCl, 5% (*v/v*) glycerol). The column was washed with FLAG buffer until the A_{280} was approximately 0.05, and dCas9-dL5 was eluted using FLAG wash buffer containing 3X FLAG peptide (200 μg/mL). Fractions containing dCas9-dL5 were collected and pooled to yield Fraction III (15 mL), which was dialyzed against 2 L of HisTrap buffer (50 mM Tris-HCl pH 7.6, 0.5 mM EDTA, 2 mM dithiothreitol, 300 mM NaCl, 20 mM imidazole pH 8.0, 5% (*v/v*) glycerol). Fraction III was applied at 1 mL/min onto a 5 mL HisTrap column equilibrated in HisTrap buffer. The column was washed until A_{280} returned to baseline and dCas9-dL5 was eluted as a single peak with a step elution of 300 mM imidazole pH 8.0. Fractions under the peak were pooled and dialyzed against 2 L of storage buffer (50 mM Tris-HCl pH 7.6,1 mM EDTA, 3 mM dithiothreitol, 300 mM NaCl, 50% (*v/v*) glycerol) to give Fraction IV (4 mL, containing 6.9 mg of protein; Figure 9.S1A). Aliquots were stored at –20°C.

9.4.5. Rolling-circle replication template.

DNA rolling circle substrates were prepared as previously described [38].

9.4.6. Linear DNA substrates.

Plasmid pSupercos1 DNA[39] (7 pmol) was linearized overnight at 37°C with 100 U of *Bst*XI in 1 x buffer 3.1 (New England Biolabs, USA). The 18,345 bp fragment was purified with a Wizard SV gel and PCR clean-up kit (Promega, USA) and the concentration was measured. DNA oligonucleotides (750 pmol of arm 1, 4500 pmol arm 2, and 70 pmol capping 1, 2) were annealed by heating at 94°C for 5 min before slow cooling. The biotinylated handles were ligated to the 18,345 bp fragment in 1 X T4 ligase buffer and 2000 U of T4 ligase overnight at 16°C. Biotinylated linear DNA substrates were purified from excess DNA oligonucleotides by adjusting NaCl to 300 mM and loaded by gravity onto a Sepharose 4B (1 x 25 cm) column, equilibrated in gel filtration buffer (10 mM Tris-HCl pH 8.0, 1 mM EDTA, and 300 mM NaCl). Biotinylated linear DNA substrates eluted as a single peak in the column void volume, fractions under the peak were analyzed by agarose gel electrophoresis. Fractions containing linear DNA substrates were pooled and dialyzed overnight in 2 L of sterilized TE buffer, concentrated 2-fold in a vacuum concentrator and the concentration measured. This protocol typically yielded ~20 μg DNA. Aliquots were stored at –80°C.

9.4.7. Forked linear DNA substrates.

The eukaryotic DNA replication template, a linearized 2.7 kb plasmid ligated to a synthetic replication fork, was prepared as previously described [33, 40]. A synthetic 37-mer oligonucleotide (Fork primer) was end-labeled with ^{32}P-ATP by T4 polynucleotide kinase (New England Biolabs, USA) according to manufacturer's instructions and annealed to the forked substrate by heating to 85°C and slowly cooling.

9.4.8. Assessment of dCas9 interactions by SPR.

SPR measurements were carried out on a BIAcore T200 instrument (GE Healthcare, Sweden) at 20°C in SPR buffer (30 mM Tris-HCl pH 7.6, 0.5 mM dithiothreitol, 0.25 mM EDTA, 0.005% (*v/v*) surfactant P20) containing NaCl/$MgCl_2$ concentrations as described. A streptavidin-coated (SA) sensor chip was activated with three sequential injections of 1 M NaCl, 50 mM NaOH (40 s each at 5 µL/min). Then, a solution (2.5 nM) of the 3'-biotinylated 83-mer template dsDNA in SPR buffer containing 50 mM NaCl (SPR running buffer), assembled *in situ* by premixing 83-S and 83-AS oligonucleotides (to final concentrations of 1.2 and 1 μM, respectively) in hybridization buffer (20 mM Tris-HCl pH 7.8, 50 mM NaCl, 5 mM $MgCl_2$) at 90°C for 5 min followed by slow cooling overnight to the room temperature, was used to immobilize ~150 RU of DNA template at 5 μL/min over 456 s onto the surface of flow cell 4, whereas flow cell 3 was left unmodified and served as a control (4–3 subtraction).

To interrogate binding specificity of dCas9-dL5 for immobilized 83 dsDNA template in the presence of complementary guide cgRNA1, a solution of protein (10 nM) with cgRNA1 (50 nM) in SPR buffer supplemented with 100 mM NaCl and 10 mM $MgCl_2$ (SPR binding buffer) was made to flow at 10 μL/min for 350 s, yielding a response of ~625 RU (Figure 9.1C). Following the association phase, the slow dissociation of protein from immobilized DNA template initiated by re-introduction of the running buffer in the flow cell and monitored over >70 s indicated stable binding. Bound proteins/RNA complexes were removed and immobilized dsDNA on the chip surface regenerated by three successive 40 s injections of 3 M $MgCl_2$ at 10 µL/min. Injections of dCas9-dL5 under similar experimental conditions, either in the presence of ncgRNA (257 s injection) or in the absence of any guide gRNA (107 s), as well as the injection of cgRNA1 alone (66 s) led to barely detectable binding responses, suggesting that only the dCas9-dL5–cgRNA1 complex interacts stably and specifically with 83 template dsDNA. Furthermore, binding of dCas9-dL5–cgRNA1 is concentration dependent, since comparative injection of 30 nM dCas9-dL5 with 50 nM cgRNA led to faster association (Figure 9.S1A). Moreover, notably similar responses were measured at equilibrium when 10 nM and 30 nM dCas9-dL5 were injected with 50 nM cgRNA1 (~625 RU) implies saturation of all the template DNA molecules on the chip surface with 10 nM dCas9-dL5–cgRNA1, indicating: (a) that the K_D for the dCas9-dL5–cgRNA1–dsDNA interaction is significantly below 10 nM in buffer containing 100 mM NaCl and 10 mM $MgCl_2$, and (b) that the dCas9-dL5–cgRNA1 complex binds 83-mer template DNA in 1:1 molar ratio, *i.e.* considering that the ratio of mol. wt. between dCas9-dL5–cgRNA1 complex (~218.1 kDa; 184.5 kDa for dCas9 and ~33.6 kDa for cgRNA1) and template dsDNA (51.7 kDa; 25.5 kDa for 83-S and 26.2 kDa for 83-AS) is 4.2, and considering that ~150 RU of DNA was immobilized on the surface, ~630 RU (4.2·150 RU) of bound dCas9-dL5–cgRNA1 could be expected at saturation in case of 1:1 interaction with template DNA.

To demonstrate the strong association and long-term stability of dCas9-dL5–cgRNA1 complex with the target DNA template, the dissociation of a complex assembled on the surface during injection of 30 nM dCas9-dL5 and 50 nM cgRNA1 (as described above) from immobilized DNA in SPR running buffer, interspersed with an early 1500 s injection of SPR binding buffer to assess the complex stability in the buffer used for the association, was monitored for over 16 h (58807 s; final response was ~450 RU; Figure 9.1D). The surface (immobilized template dsDNA) was then regenerated with one 40 s injection of 3 M $MgCl_2$ at 10 µL/min. Assuming first-order dissociation and SPR responses that were measured following the injection of SPR binding buffer, at the start of measured dissociation R_0 = 575 RU and at the end R_t = 450 RU over the period of t = 57000 s, the dissociation half-life of > 44 hours (see also Supplementary Methods) was calculated using Equation 1:

$$t_{1/2} = \frac{t \cdot \ln 2}{\ln \frac{R_0}{R_t}}$$

9.4.9. Measurement of dCas9-dL5 binding specificity on long DNA substrates.

Microfluidic flow cells were prepared as described in “Preparation of flow-cells for *in vitro* imaging”. To help prevent non-specific interactions of proteins and DNA with all surfaces, they were blocked with 2% Tween20 in blocking buffer (50 mM Tris-HCl pH 7.6, 50 mM KCl). Imaging parameters are described in “*In vitro* single-molecule fluorescence microscopy”.

First, 9 nM dCas9-dL5 was incubated with 15 nM cgRNA1 at 37°C for 5 min in reaction buffer (25 mM Tris-HCl pH 7.6, 10 mM $MgCl_2$, 150 mM potassium glutamate, 0.1mM EDTA and 0.0025% (*v/v*) Tween20). The dCas9-dL5–cgRNA1 complex was further incubated with 125 pM biotinylated linear DNA substrates at 37°C for 20 min in reaction buffer supplemented with 0.5 mg/mL heparin. To reduce heterogeneity in DNA lengths upon binding to the surface, 200 µM chloroquine was added immediately prior to injection of the sample into the flow cell. The solution was injected at a constant rate of 17 µL/min until an appropriate DNA density was achieved. Next, the flow cell was washed with 2 mL of reaction buffer, supplemented with 100 mM NaCl, 15 nM gRNA and 0.5 mg/mL heparin. dCas9-dL5–cgRNA1–DNA complexes were imaged in reaction buffer containing 150 nM Sytox orange and 150 nM MGE.

9.4.10. in vitro ensemble E. coli replication assays.

Standard leading-strand replication assays were set up in replication buffer (RB; 60 mM Tris-HCl pH 7.6, 24 mM $Mg(OAc)_2$, 100 mM potassium glutamate, 1 mM EDTA and 0.005% (*v/v*) Tween20) and contained 2 nM rolling-circle replication template, specified concentrations of dCas9-dL5 and gRNA, 60 nM DnaBC, 30 nM $\tau_3\delta\delta'\chi\psi$, 90 nM Pol III αεθ core, 200 nM β_2, 10 mM dithiothreitol, 1 mM ATP, and 125 µM dNTPs in a final volume of 12 µL. First, dCas9-dL5 was incubated with gRNA for 5 min, and further incubated with rolling-circle DNA templates for 5 min at room temperature. Components (except dCas9-dL5–gRNA–DNA) were mixed and treated at room temperature, then cooled in ice for 5 min prior to addition of dCas9-dL5–gRNA–DNA complexes. Reactions were initiated at 30°C and quenched at specified time points by the addition of 200 mM EDTA and 2% (*w/v*) SDS. The quenched reactions were loaded into a 0.6% (*w/v*) agarose gel in 2x TAE. Products were separated by agarose gel electrophoresis, at 60 V for 150 min and stained in SYBR-Gold (Invitrogen) and imaged under UV light.

E. coli leading- and lagging-strand DNA replication reactions were carried out as previously described [22] with the following minor modifications. Reactions were set up in RB, and contained 4 nM rolling-circle replication template, specified concentrations of dCas9-dL5 and gRNA, 60 nM DnaBC, 80 nM DnaG, 30 nM $\tau_3\delta\delta'\chi\psi$, 10 nM SSB, 90 nM Pol III αεθ core, 200 nM β_2, 10 mM dithiothreitol, 1 mM ATP, 125 µM dNTPs, and 250 µM NTPs to a final volume of 12 µL, quenched after 30 min by addition of 1.5 µL 0.5 M EDTA and 3 µL DNA loading dye (6 mM EDTA, 300 mM NaOH, 0.25% (*v/v*) bromocresol green, 0.25% (*v/v*) xylene cyanol FF, 30% (*v/v*) glycerol). DNA products were separated on a 0.6% (*w/v*) alkaline agarose gel at 14 V for 14 h. The gel was then neutralized in TAE buffer, stained with SYBR-Gold and imaged under UV light.

9.4.11. in vitro ensemble T7 replication assays.

T7 leading-strand DNA replication assays were carried out using previously described conditions[41]. Briefly, reactions were set up in T7 replication (TR) buffer (25 mM Tris-HCl pH 7.6, 10 mM $MgCl_2$, 50 mM potassium glutamate, 0.1 mM EDTA and 0.0025% (*v/v*) Tween20) and contained 2 nM rolling-circle replication template, specified concentrations of dCas9-dL5 and cgRNA1, 180 nM gp2.5, 5 nM gp4, 40 nM gp5, 10 mM dithiothreitol, 1 mM ATP, 1 mM CTP, and 600 µM dNTPs, in a final volume of 12 µL. First, dCas9-dL5 was incubated with cgRNA1 for 5 min, and further incubated with rolling-circle DNA templates for a further 5 min at room temperature. Components (except dCas9-dL5–cgRNA1–DNA) were mixed and treated at room temperature, then cooled in ice for 5 min prior to addition of dCas9-dL5–cgRNA1–DNA complexes. Reactions were initiated at 30°C and quenched at specified time points by the addition of 200 mM EDTA and 2% (*w/v*) SDS. The quenched reactions were loaded onto the 0.6% (*w/v*) agarose gel, which was run under the same conditions as standard *E. coli* leading-strand replication assays.

9.4.12. in vitro ensemble S. cerevisiae replication assays.

Leading-strand replication assays were set up in eukaryotic replication (ER) buffer (25 mM Tris-OAc pH 7.5, 5% glycerol, 80 µg/mL BSA, 5 mM tris(2-carboxyethyl)phosphine, 10 mM $Mg(OAc)_2$, 50 mM potassium glutamate, 0.1 mM EDTA), and contained 1.5 nM DNA substrate (see section on Forked Linear DNA substrates) 30 nM CMG, 30 nM MTC, 20 nM Pol ε, 10 nM RFC, 30 nM PCNA, 600 nM RPA, 5 mM ATP and 120 µM dNTPs, and where indicated 40 nM sgRNA and 20 nM dCas9-dL5 in a final volume of 20 µL. First, DNA was incubated with CMG and MTC for 2 min at 30°C followed by an additional 2 min with dCas9-dL5 and cgRNAs. Components except ATP and RPA were added and further incubated for 5 min at 30°C. Replication was initiated by addition of ATP and RPA. The reactions proceeded for the indicated amount of time at 30ºC and were quenched with an equal volume of 2x stop solution (40 mM EDTA and 2% (*w/v*) SDS). DNA products were separated on a 1.3% (*w/v*) alkaline agarose gel at 35 V for 16 h. Gels were backed with DE81 paper (GE Healthcare), dried by compression, exposed to a phosphorimager screen, and imaged with a Typhoon FLA 9500 PhosphorImager (GE Healthcare).

9.4.13. in vitro single-molecule fluorescence microscopy.

In vitro single-molecule microscopy was performed on an Eclipse Ti-E inverted microscope (Nikon, Japan) with a CFI Apo TIRF 100x oil-immersion TIRF objective (NA 1.49, Nikon, Japan), as previously described[6]. The temperature was maintained at 31°C (unless otherwise stated) by an electronically heated flow-cell chamber coupled to an objective heating jacket (Okolab, USA). NIS-elements was used to operate the microscope and the focus was locked through Perfect

Focus System (Nikon, Japan). Images were captured using an Evolve 512 Delta EMCCD camera (Photometics, USA) with an effective pixel size of 0.16 μm. DNA molecules stained with 150 nM Sytox orange were imaged with a CW 568-nm Sapphire LP laser (200 mW max. output), and ET600/50 emission filter (Chroma, USA) at 0.76 W/cm^2. dCas9-dL5–MGE complexes were imaged with a CW 647-nm OBIS laser (100 mW max. output), and 655LP emission filter (Chroma, USA) at 57.7 W/cm^2.

9.4.14. Preparation of flow-cells for in vitro imaging.

Replication reactions were carried out in microfluidic flow-cells constructed from a PDMS flow chamber placed on top of a PEG-biotin-functionalized microscope coverslip as previously described [19, 22, 32, 41]. Once assembled, all surfaces of the flow-cell including connecting tubing were blocked against non-specific binding by introduction of 1 mL malic acid buffer (100 mM Na.maleate pH 7.5 and 250 mM NaCl) containing 1% (*w/v*) blocking reagent (Roche, Switzerland).

9.4.15. Single-molecule rolling-circle blocking replication assays.

The overall experimental scheme was to first form the dCas9-dL5–cgRNA1–DNA complex. Next, dCas9-dL5–cgRNA1–DNA complex was attached via the 5’-biotinylated flap-primed 2030-bp dsDNA circle bearing a 25-nt fork gap, to the surface via a biotin–streptavidin bond. Following a wash to remove unbound dCas9-dL5, replication was initiated by continuous flowing of reconstituted replisomes, ATP, dNTPs, and rNTPs and flow-stretching the DNA.

Specifically, 10 nM dCas9-dL5 was incubated with 200 nM cgRNA1 for ~5 min at 37°C in single-molecule imaging (SM) buffer (25 mM Tris-HCl pH 7.6, 10 mM $Mg(OAc)_2$, 50 mM potassium glutamate, 0.1 mM EDTA, 10 mM dithiothreitol and 0.0025% (*v/v*) Tween20). The dCas9-dL5–cgRNA1 complex was then incubated with 100 pM replication templates for a further 20 min at 37°C. The dCas9-dL5–cgRNA1–DNA complexes were adsorbed to the surface in SM buffer + 150 nM Sytox orange at 10 μL/min until an appropriate surface density was achieved. The flow-cell was then washed with 200 μL of SM buffer containing 50 mM NaCl. Following this replication was initiated — *E. coli* leading- and lagging-strand DNA replication reactions were carried out under the continuous presence of all proteins as previously described [22]. T7 leading- and lagging-strand DNA replication assays were carried out under the continuous presence of all proteins using previously described conditions [19]. All *in vitro* single-molecule rolling-circle blocking experiments were performed at least three times.

9.4.16. Analysis of agarose gels of replication products.

Agarose gel images were adjusted for brightness and contrast for clear visualization using FIJI [42]. Blocked replication products were quantified in FIJI using in-house built plugins, by comparing the integrated intensity of bands between control and reaction lanes; the resulting percentages were then corrected for background and for the specified control.

9.4.17. *In vitro image analysis.*

Image analysis was performed in FIJI, using the Single Molecule Biophysics plugins (available at https://github.com/SingleMolecule/smb-plugins). Raw videos (.nd2 format) were converted into TIF files and flattened with the excitation beam profile as described previously [43]. For quantification of DNA product lengths, intensity projections were generated by summing 10 frames to reduce the contribution of transverse Brownian fluctuations of the DNA. Product length was determined by deconvolving the length of the rolling-circle substrate using the calibrated pixel size in bp (here, 1 pixel = 470 bp). Product length distributions were fit with a single-exponential decay (assuming a single rate-limiting step determining the end of an event). All distributions were made and fitted using MATLAB (Mathworks, USA).

9.5. Supplementary Methods

9.5.1. *Assessment of diffusion limited binding kinetics of dCas9-dL5 binding to target containing DNA in SPR studies*

The sensorgrams obtained during the association of dCas9-dL5-cgRNA at different concentrations of dCas9-dL5 (Figure 9.S1B) exhibited a distinct biphasic profile. The linear response (RU)–time (s) relationship suggests that the association of the complex from solution could be a fast, diffusion limited process. To examine this possibility, solutions of dCas9-dL5 (10 nM) with cgRNA1 (50 nM) in SPR buffer supplemented with 150 mM NaCl and 10 mM $MgCl_2$ were injected over immobilized template DNA for 120 s at three different flow rates accessible to the BIAcore T200 instrument: 20, 30 and 80 μL/min (Figure 9.S1C). The observed increase in association rates with the increase in flow rates confirms that the interaction is indeed diffusion-limited, a situation that occurs when the diffusion of analyte from the bulk solution to the chip surface is slower than its binding to the ligand. Conversely, it suggests that the dissociation of analyte (dCas9-dL5–cgRNA) from the ligand (83-mer DNA) must also be a diffusion limited process, *i.e.* upon dissociation from the ligand, the analyte may not diffuse into bulk solution, allowing it to re-bind. This would result in an apparently slower dissociation rate; therefore, the dissociation half-life ($t_{1/2}$) of dCas9-dL5-cgRNA1 from its target dsDNA is an over-estimate of the true dissociation half-life.

9.5.2. *Measurement of position of bound dCas9-dL5-cgRNA on 18-kb template*

Measurement of the position of the bound dCas9-dL5–cgRNA complex on the 18-kb template was performed as follows:

1. Line profiles were manually drawn over all individual DNA molecules. The length of the individual molecules was defined as the distance between the maximum and minimum of the first derivative of the intensity along the drawn lines. Using these measurements, a length distribution was plotted, and values below the 25% and above the 75% percentile were classified as outliers. The resulting distribution was fit to a Gaussian distribution with a mean of 39.5 ± 0.1 pixels. This mean length was then assumed to correspond to the total length of 18,345 bp of the DNA substrate. This conversion resulted in a calibration factor of 466 ± 1 bp/pixel.
2. Next, peaks were detected along the line profile in the MGE-channel. The position of the detected peaks, relative to the ends of the DNA-molecules was then calculated. The distances to both ends of the DNA-molecule were measured. The position in base-pairs was calculated using the calibration described above. The histogram shows the smaller of the two distances from the DNA ends for each molecule.

9.5.3. Description of cgRNAs used in studies of bacterial DNA replication

Exact positions of targeting gRNA on the rolling-circle DNA replication template are as follows; cgRNA1 targeted to nucleotides 1402-1421 of the lagging-strand, cgRNA3 targeted to nucleotides 177-196 of the leading-strand, and cgRNA4 targeted to nucleotides 1046-1065 of the leading-strand.

9.5.4. Description of cgRNAs used in studies of eukaryotic DNA replication

Exact positions of targeting gRNA on the eukaryotic linear DNA replication template are as follows: cgRNA$_{0.6}$ targeted nucleotides 583–602 of the leading strand, cgRNA$_{1.0}$ targeted nucleotides 1005–1024 of the lagging strand, cgRNA$_{1.5}$ targeted nucleotides 1493–1512 of the lagging strand, and cgRNA$_{2.2}$ targeted nucleotides 2196–2215 of the leading strand.

9.6. Supplementary figures

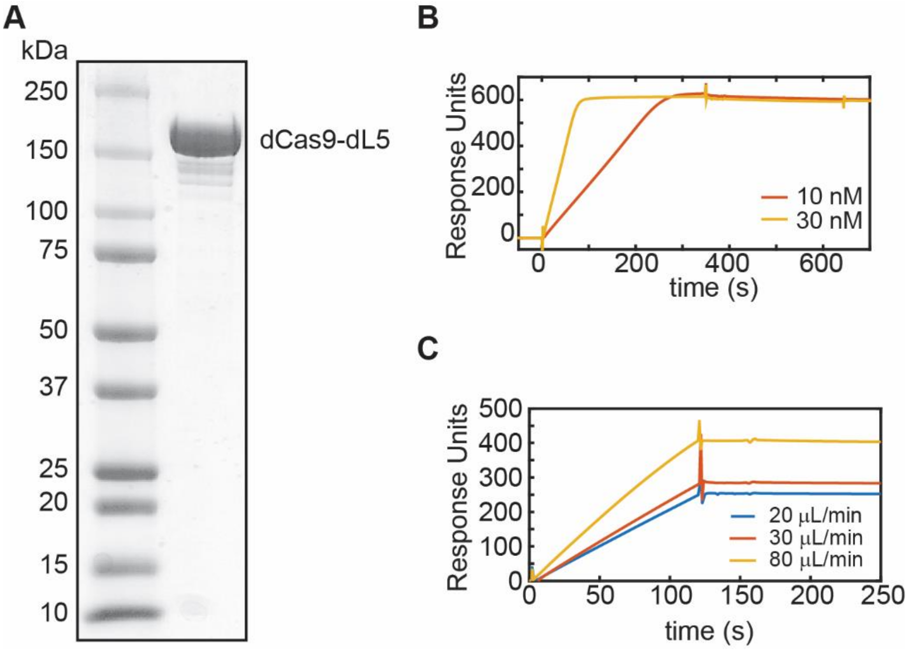


Figure 9.S1: Characterization of dCas9-dL5. A) Coomassie stained 4–20% SDS-PAGE of purified dCas9-dL5. **B)** Sensorgrams showing binding to and dissociation of 10 and 30 nM dCas9-dL5-cgRNA from 83-mer dsDNA substrate immobilized on an SPR chip. **C)** Sensorgrams monitoring the association of dCas9-dL5-cgRNA (10 nM) injected at three different flow rates (20, 30 and 80 μL/min) onto 83-mer dsDNA containing target sequence immobilized on an SPR chip. The linearity and difference in responses indicates mass transfer limitation. *N* = 1 independent experiment.

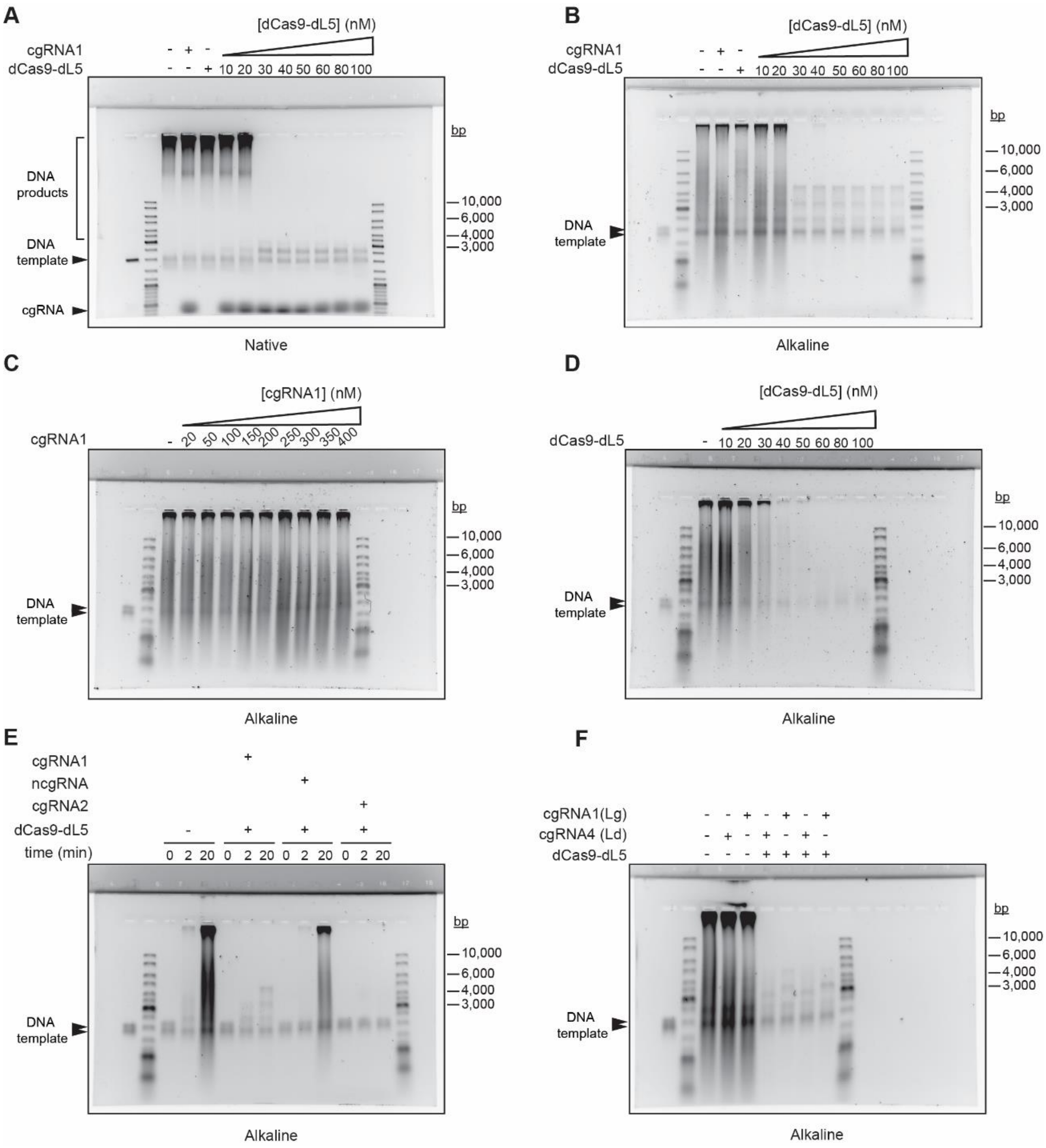


Figure 9.S2: Target-bound dCas9-dL5 site-specifically arrests *E. coli* DNA synthesis. A) Target-bound dCas9-dL5 arrests *E. coli* leading strand DNA synthesis. Unless otherwise specified reactions contained 400 nM cgRNA1 and 20 nM dCas9-dL5. At concentrations below 20 nM, dCas9-dL5 does not completely arrest leading-strand DNA synthesis. *N* > 3 independent experiments. **B)** Target-bound dCas9-dL5 arrests *E. coli* leading- and lagging-strand DNA synthesis. Unless otherwise specified, reactions contained 400 nM cgRNA1. At concentrations below 20 nM, dCas9-dL5 does not completely arrest leading- and lagging-strand DNA synthesis. *N* > 3 independent experiments. **C)** High concentrations of complementary gRNAs do not inhibit *E. coli* leading- and lagging-strand DNA synthesis. Unless otherwise specified, 50 nM dCas9-dL5 was used for all reactions. *N* > 2 independent experiments. **D)** dCas9-dL5 alone does not site-specifically inhibit *E. coli* leading- and lagging-strand DNA synthesis. Non-specific inhibition is observed at high concentrations of dCas9-dL5 alone (see Supplementary Methods, Figure 2A and summary in Figure 2H). *N* > 2 independent experiments. **E)** Only dCas9-dL5 programmed with complementary gRNAs specifically arrests *E. coli* leading- and lagging-strand DNA synthesis. Reactions

contained 50 nM dCas9-dL5 and 400 nM gRNAs. Reactions were initiated at 30°C and aliquots were removed and quenched at 0, 2, and 20 min time points. *N* > 3 independent experiments. **F)** *E. coli* leading- and lagging-strand DNA synthesis arrest by target-bound dCas9-dL5 is not strand specific. Unless otherwise specified reactions contained 400 nM cgRNAs and 50 nM dCas9-dL5. Lg denotes cgRNA targeted to the lagging strand, and Ld denotes cgRNA targeted to the leading strand. *N* > 3 independent experiments. All panels show photographic negative images of gels that had been stained with SYBR-gold nucleic acid stain.

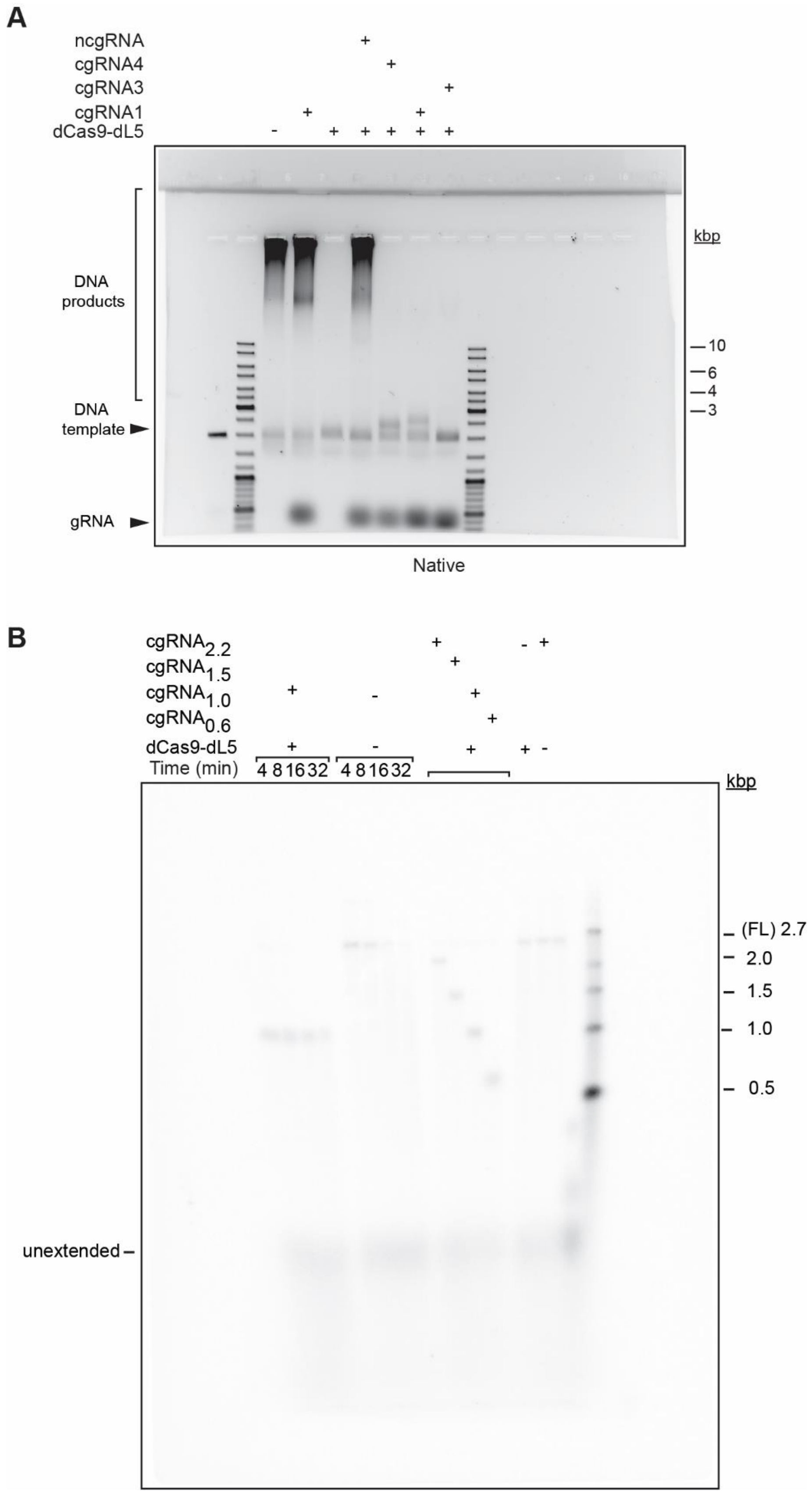


Figure 9.S3: Raw data presented in Figure 2. A) dCas9-dL5 efficiently and stably blocks bacteria DNA replication regardless of the targeted strand. dCas9-dL5 (100 nM) programmed with cgRNAs (400 nM) arrest the progress of the bacterial replication fork at target sites. **B)** dCas9-dL5 efficiently blocks eukaryotic

replication at target sites indicated by cgRNA, and time course assay of eukaryotic replication in the presence or absence of dCas9-dL5 and $cgRNA_{10}$.

9.7 References

1. Gupta, M.K., et al., *Protein-DNA complexes are the primary sources of replication fork pausing in Escherichia coli.* Proc. Natl. Acad. Sci. U.S.A., 2013. **110**(18): p. 7252-7257.
2. Mettrick, K.A. and I. Grainge, *Stability of blocked replication forks in vivo.* Nucleic Acids Res., 2016. **44**(2): p. 657-68.
3. Mirkin, E.V. and S.M. Mirkin, *Replication fork stalling at natural impediments.* Microbiol. Mol. Biol. Rev., 2007. **71**(1): p. 13-35.
4. Wang, X., R. Reyes-Lamothe, and D.J. Sherratt, *Visualizing genetic loci and molecular machines in living bacteria.* Biochem. Soc. Trans., 2008. **36**(Pt 4): p. 749-53.
5. Coster, G. and J.F.X. Diffley, *Bidirectional eukaryotic DNA replication is established by quasi-symmetrical helicase loading.* Science, 2017. **357**(6348): p. 314-318.
6. Ghosh, S. and H. Echols, *Purification and properties of D protein: a transcription factor of Escherichia coli.* Proc Natl Acad Sci U S A, 1972. **69**(12): p. 3660-4.
7. Yamazaki, K.-I., et al., *Isolation and characterization of nucleoid proteins from Escherichia coli.* Mol. Gen. Genet., 1984. **196**: p. 217-24.
8. Pugh, B.F. and M.M. Cox, *Stable binding of recA protein to duplex DNA* J. Biol. Chem., 1987. **262**(3): p. 1326-1336.
9. Dutta, D., et al., *Linking RNA polymerase backtracking to genome instability in E. coli.* Cell, 2011. **146**(4): p. 533-43.
10. Atkinson, J. and P. McGlynn, *Replication fork reversal and the maintenance of genome stability.* Nucleic Acids Res., 2009. **37**(11): p. 3475-92.
11. Larsen, N.B., et al., *The Escherichia coli Tus-Ter replication fork barrier causes site-specific DNA replication perturbation in yeast.* Nat. Commun., 2014. **5**: p. 3574-3583.
12. Possoz, C., et al., *Tracking of controlled Escherichia coli replication fork stalling and restart at repressor-bound DNA in vivo.* Embo j, 2006. **25**(11): p. 2596-604.
13. McGlynn, P. and C.P. Guy, *Replication forks blocked by protein-DNA complexes have limited stability in vitro.* J. Mol. Biol., 2008. **381**(2): p. 249-55.
14. Payne, B.T., et al., *Replication fork blockage by transcription factor-DNA complexes in Escherichia coli.* Nucleic Acids Res., 2006. **34**(18): p. 5194-202.
15. Szent-Gyorgyi, C., et al., *Malachite green mediates homodimerization of antibody VL domains to form a fluorescent ternary complex with singular symmetric interfaces.* J. Mol. Biol., 2013. **425**(22): p. 4595-613.
16. Saurabh, S., et al., *Kinetically tunable photostability of fluorogen-activating peptide-fluorogen complexes.* Chemphyschem, 2015. **16**(14): p. 2974-80.

17. Sternberg, S.H., et al., *DNA interrogation by the CRISPR RNA-guided endonuclease Cas9.* Nature, 2014. **507**(7490): p. 62-7.

18. Yardimci, H., et al., *Uncoupling of sister replisomes during eukaryotic DNA replication.* Mol. Cell, 2010. **40**(5): p. 834-40.

19. Tanner, N.A., et al., *Real-time single-molecule observation of rolling-circle DNA replication.* Nucleic Acids Res., 2009. **37**(4): p. e27.

20. Tanner, N.A., et al., *E. coli DNA replication in the absence of free β clamps.* EMBO J., 2011. **30**(9): p. 1830-40.

21. Yao, N., et al., *Single-molecule analysis reveals that the lagging strand increases replisome processivity but slows replication fork progression.* Proc. Natl. Acad. Sci. U.S.A., 2009. **106**(32): p. 13236-13241.

22. Lewis, J.S., et al., *Single-molecule visualization of fast polymerase turnover in the bacterial replisome.* eLife 2017(6): p. e23932.

23. Geertsema, H.J., et al., *Single-molecule studies of polymerase dynamics and stoichiometry at the bacteriophage T7 replication machinery.* Proc. Natl. Acad. Sci. U.S.A., 2014. **111**(11): p. 4073-8.

24. Singh, D., et al., *Real-time observation of DNA recognition and rejection by the RNA-guided endonuclease Cas9.* Nat. Commun., 2016. **7**: p. 12778.

25. Rutkauskas, M., et al., *Single-molecule insight into target recognition by CRISPR-Cas complexes.* Methods Enzymol., 2017. **582**: p. 239-273.

26. Clarke, R., et al., *Enhanced bacterial immunity and mammalian genome editing via RNA-polymerase-mediated dislodging of Cas9 from double-strand DNA breaks.* Mol. Cell, 2018. **71**(1): p. 42-55 e8.

27. Qi, L.S., et al., *Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression.* Cell, 2013. **152**(5): p. 1173-83.

28. Oakley, A.J., et al., *Flexibility revealed by the 1.85 Å crystal structure of the β sliding-clamp subunit of Escherichia coli DNA polymerase III.* Acta Crystallogr. D Biol. Crystallogr., 2003. **59**(Pt 7): p. 1192-9.

29. Mason, C.E., et al., *Escherichia coli single-stranded DNA-binding protein: NanoESI-MS studies of salt-modulated subunit exchange and DNA binding transactions.* J. Am. Soc. Mass Spectrom., 2013. **24**(2): p. 274-85.

30. Jergic, S., et al., *A direct proofreader-clamp interaction stabilizes the Pol III replicase in the polymerization mode.* EMBO J., 2013. **32**(9): p. 1322-33.

31. Stamford, N.P.J., P.E. Lilley, and N.E. Dixon, *Enriched sources of Escherichia coli replication proteins: The dnaG primase is a zinc metalloprotein.* Biochim. Biophys. Acta, 1992. **1132**(1): p. 17-25.

32. Tanner, N.A., et al., *Single-molecule studies of fork dynamics in Escherichia coli DNA replication.* Nat Struct Mol Biol, 2008. **15**(9): p. 998.

33. Georgescu, R.E., et al., *Mechanism of asymmetric polymerase assembly at the eukaryotic replication fork.* Nat. Struct. Mol. Biol., 2014. **21**(8): p. 664-70.

34. Langston, L.D., et al., *Mcm10 promotes rapid isomerization of CMG-DNA for replisome bypass of lagging strand DNA blocks.* eLife, 2017. **6**.

35. Yao, N., et al., *Replication factor C clamp loader subunit arrangement within the circular pentamer and its attachment points to proliferating cell nuclear antigen.* J. Biol. Chem., 2003. **278**(50): p. 50744-53.

36. Finkelstein, J., et al., *Overproduction and analysis of eukaryotic multiprotein complexes in Escherichia coli using a dual-vector strategy.* Anal. Biochem., 2003. **319**(1): p. 78-87.

37. Hyland, E.M., L.F. Rezende, and C.C. Richardson, *The DNA binding domain of the gene 2.5 single-stranded DNA-binding protein of bacteriophage T7.* J. Biol. Chem., 2003. **278**(9): p. 7247-56.

38. Monachino, E., et al., *Design of DNA rolling-circle templates with controlled fork topology to study mechanisms of DNA replication.* Anal. Biochem., 2018. **557**: p. 42-45.

39. van Loenhout, M.T., M.V. de Grunt, and C. Dekker, *Dynamics of DNA supercoils.* Science, 2012. **338**(6103): p. 94-7.

40. Schauer, G., J. Finkelstein, and M. O'Donnell, *In vitro assays for eukaryotic leading/lagging strand DNA replication.* Bio. Protoc., 2017. **7**(18).

41. Geertsema, H.J., K.E. Duderstadt, and A.M. van Oijen, *Single-molecule observation of prokaryotic DNA replication*, in *Methods Mol. Biol.*, S. Vengrova and J. Dalgaard, Editors. 2015, Springer New York: New York, NY. p. 219-238.

42. Schindelin, J., et al., *Fiji: an open-source platform for biological-image analysis.* Nat. Methods, 2012. **9**(7): p. 676-82.

43. Robinson, A., et al., *Regulation of mutagenic DNA polymerase V activation in space and time.* PLoS Genet., 2015. **11**(8): p. e1005482.