2018

Single-molecule studies of the replisome Visualisation of protein dynamics in multi-protein complexes

Lisanne Maria Spenkelink

UNIVERSITY
OF WOLLONGONG
AUSTRALIA

Single-molecule studies of the replisome

Visualisation of protein dynamics in multi-protein complexes

Lisanne Maria Spenkelink

This thesis is presented as part of the requirements for the conferral of the degree:
Doctor of Philosophy

Supervisor:
Antoine M. van Oijen

Co-supervisor:
Nicholas E. Dixon

The University of Wollongong
School of Chemistry

Month 2018

Declaration

I, Lisanne M. Spenkelink, declare that this thesis, submitted in fulfilment of the requirements for the conferral of the degree Doctor of Philosophy, in the School of Chemistry, from the University of Wollongong, is wholly my own work unless otherwise referenced or acknowledged. This document has been submitted for qualifications at the University of Groningen in accordance with the Joint Agreement between the University of Wollongong and the University of Groningen.

Lisanne M. Spenkelink
November 6, 2018

Abstract

DNA replication, or the duplication of parental double-stranded DNA (dsDNA) into a pair of identical copies, is essential to the transmission of hereditary information from cell to cell and thus the propagation of all life. It is a fundamental cellular process that is carried out by a multi-protein complex known as the replisome. Since the identification of the first replication proteins by Arthur Kornberg in the 1950s, ensemble-averaging biochemical techniques have been successfully used to study the roles of the various proteins within the replisome. However, the coordination of the multiple activities within the replisome involves transient intermediates and dynamic conformational changes that are difficult, if not impossible, to observe with ensemble experiments. Recently, new single-molecule techniques have been developed to study the dynamics of proteins with a high precision and without the need for population averaging. This thesis centers on the use of these approaches to study the dynamic behaviour of the replisome.

First, I provide a review of the latest advances in force- and fluorescence-based single-molecule methods both for systems based on purified protein components and for those using live cells, with a focus on their applications in studies on cytoskeletal motors and DNA replication. We describe how these technological developments allow us to study systems of increasing biological complexity. Next, a novel fluorescence imaging technique is described that enables the observation of single fluorescently labeled molecules can be observed in real time at concentrations that are too high for existing single-molecule methods. This technique allows the use of fluorescent probes up to micromolar concentrations and is compatible with a wider range of physiologically relevant affinities between biological macromolecules. Single-molecule tools are being used in increasingly broader areas of research and have now also found their use in the development of targeted drug-delivery mechanisms. I will describe a novel single-molecule fluorescence imaging approach to determine the density of proteins on functionalised liposomes. This parameter is important for the efficacy of drug delivery, and tools did not exist yet to enable its quantification. This approach will help with translating functionalised liposome techniques into the clinic as viable therapeutic approaches.

The main part of this thesis revolves around the highly dynamic behaviour of protein factors in the *E. coli* replisome. I describe single-molecule visualisation experiments that show exchange behaviour of replication proteins that is markedly more dynamic than the textbook pictures of the replisome suggest. For both the replicative DNA polymerase and the single-stranded DNA binding protein we observe rapid exchange with kinetics that depend on the concentration of competing protein in solution. This concentration dependence can be rationalised via a multi-site exchange mechanism, made possible by the presence of multiple weak protein–protein and protein–DNA interactions. This multi-site exchange mechanism reconciles a large number of observations and has a generality that suggests applicability to a large number of multi-protein complexes. The emergence of these exchange dynamics as a mechanism to balance stability with plasticity in

complex systems illustrates the importance of studying the molecular sociology of multi-protein complexes at the single-molecule level.

Furthermore, this thesis describes studies of the mechanism by which the *E. coli* RarA protein acts on DNA replication and repair. Combining *in vitro* and *in vivo* techniques, we propose that addition of RarA results in gaps in lesion-containing DNA templates. Our observations suggest a model in which the previously poorly understood RarA protein commits the cell to the translesion DNA synthesis repair pathway.

Finally, I describe how we take our single-molecule tools to the next level of biological complexity and apply them to visualise *S. cerevisiae* leading-strand synthesis at the single-molecule level for the first time. We confirm that the MTC complex modulates the speed of the replication fork and observe, surprisingly, that MTC only transiently interacts with the replisome through a weak interaction.

Further improvement of single-molecule techniques will allow us to study complex biological systems in increasingly fine detail. *In vitro* experiments give a precise handle on the experimental conditions in a controlled environment, while *in vivo* measurements provide the physiologically relevant complexity that exists in live cells. Bridging the gap between the two techniques will allow us to further elucidate the dynamics of proteins within multi-protein complexes.

Acknowledgements

My PhD has been an incredible four-year journey on two continents. During this time I have been lucky enough to work with a great group of people. Many of you made important contributions to this thesis and to my life in some way or another. Since I do not want to risk forgetting anyone, I would like to take this opportunity to thank all of you. I am very grateful for all the effort you put in. This thesis could not have been written without all of your contributions.

There are a few people who deserve special mentioning here.

First and foremost I would like to thank my supervisor, **Antoine**. Bedankt dat ik mijn PhD mocht doen in jouw groep. Ik heb enorme bewondering voor je wetenschappelijke intuïtie, je betrokkenheid met iedereen in de groep, de vrijheid die je me hebt gegeven en jouw schijnbaar oneindige hoeveelheid energie. Daarnaast heb je een ongeloofelijk talent voor het samenbrengen van een geweldige groep mensen. Het was, en is, geweldig om daar onderdeel van uit te maken. Ik wil je ook bedanken dat je je lab naar Australië hebt verhuisd. Een betere kans had ik niet kunnen krijgen!

My paranimfs **Jacob** and **Enrico**. **Jacob**, more than half of this thesis would not have existed without your help. Thank you so much for all your efforts and for teaching a physicist some biochemistry. Let's get that next paper. **Enrico**, we started this journey at the same time. We were two physicist learning to do biology, and two students moving from Groningen to Wollongong. So now it only makes sense that we finish it together. Grazie mille!

My co-supervisor in Wollongong, **Nick**. With your seemingly endless knowledge on DNA replication you have made a large contribution, not only to the work in this thesis, but to all the work done in the group. Thank you for always taking time to discuss our experiments and think about the data.

Thank you to all our collaborators who have shared their knowledge and precious proteins with us: The Laboratory of DNA Replication at Rockefeller, especially **Michael O'Donnell** and **Grant Schauer**; Our collaborators from the Molecular biology and enzymology of genetic recombination and DNA repair group at the Univeristy of Wisconsin-Madison, **Michael Cox** and **Elizabeth Wood**; From the Targeted Cancer Therapeutics Laboratory at the University of Wollongong, **Kara Perrow** and **Lisa Belfiore**.

I would like to thank the members of my assessment committee **Dagmar Klostermeier**, **Richard Fishel**, **Wouter Roos**, and **Aaron Oakley** for their critical reading of my thesis. Ook wil ik mijn co-promotor in Nederland, **Bert Poolman** bedanken voor het lezen van mijn thesis.

Hylkje, mijn SMB zus. Ik begon in de groep als jouw master student en ik herinner me nog goed jouw woorden toen je hoorde dat ik zou blijven voor een PhD.

Mede dankzij jouw goede begeleiding en enthousiasme was de keuze om in de groep te blijven snel gemaakt. En daardoor werd je ook een erg goede vriendin. Dank voor al je steun! **Jasper**, potje doen? Onze ping-pong sessies waren altijd een goede manier om even een pauze te nemen en frustraties weg te slaan. Jij ook bedankt voor al je steun!

The SMB group in Groningen: **Karl**, **Sarah**, **Jochem**, **Joris**, **Evelyn**, **Victor C**, **Giorgos**, **Jelle** (Gefeliciteerd!!), **Victor K**, **Michiel**, **Alex**, **Iullia**, **Samaneh**, **Guus**, **Marysia**, **Kostas**, **Florence**, **Atieh**, **Aartje**, **Douwe**, **Jason**, **Jan Peter**, **Margriet**, and **Jeannette**.

The Biophysics group in Wollongong (Dixon/van Oijen/Oakley/Yu/Tolun): **Harshad**, **Han**, **Sarah**, **Megan**, **Amy**, **Gurleen**, **Stefan**, **Richard**, **Slobodan**, **Zhi-Qiang**, **Bishnu**, **Varsha**, **Caitlyn**, **Nick M.**, **Tom** (Paul), **Ryan**, **Fay**, **Gökhan**, **Flynn**, **Quill**, **Parisa**, **Nick H.**, **Kate**, **David**, and **Haibo**.

Finally, I would like to thank **Stephanie** and **Andrew**.

Thank you!

List of publications

1. **Lisanne M. Spenkelink**, Jacob S. Lewis, Slobodan Jergic, Zhi-Qiang Xu, Andrew Robinson, Nicholas E. Dixon, Antoine M. van Oijen. Single-molecule visualization of SSB dynamics shows a competition between an internal-transfer mechanism and external exchange. Manuscript submitted to *N.A.R.*

2. Lisa Belfiore[†], **Lisanne M. Spenkelink**[†], Antoine M. van Oijen, Kara L. Vine. Quantification of ligand stoichiometries in liposomal drug delivery systems using single-molecule fluorescence imaging. *Journal of Controlled Release*, 28 May 2018;**2**78:80–86

3. Jacob S. Lewis[†], **Lisanne M. Spenkelink**[†], Grant D. Schauer, Flynn R. Hill, Roxanna E. Georgescu, Michael E. O'Donnell, Antoine M. van Oijen. Single-molecule visualization of leading-strand synthesis by *S. Cerevisiae* reveals dynamic interaction of MTC with the replisome. *Proc. Natl. Acad. Sci. U.S.A*, Sept 2017;**114**(40): 10630–10635.

4. Jacob S. Lewis[†], **Lisanne M. Spenkelink**[†], Slobodan Jergic, Elizabeth A. Wood, Enrico Monachino, Nicholas P. Horan, Karl E. Duderstadt, Michael M. Cox, Andrew Robinson, Nicholas E. Dixon, Antoine M. van Oijen. Single-Molecule visualisation of fast polymerase turnover in the bacterial replisome. *eLife* 2017;10.7554/eLife.23932

5. Enrico Monachino[†], **Lisanne M. Spenkelink**[†], Antoine M. van Oijen. Watching cellular machinery in action, one molecule at a time. *Journal of Cell Biology*, 02 Jan 2017; **216** (1):41–51

6. Hylkje J. Geertsema, Aartje C. Schulte, **Lisanne M. Spenkelink**, William J. McGrath, Seamus R. Morrone, Jungsan Sohn, Walter F. Mangel, Andrew Robinson, Antoine M. van Oijen. Single-Molecule imaging at high fluorophore concentrations by Local Activation of Dye. *Biophysical Journal*, 17 Feb 2015; **108** (4):949–56

[†]These authors contributed equally.

Contents

List of figures

1 Introduction

1.1 DNA replication

Nearly 150 years ago, Johann Friedrich Miescher first purified a substance from cell nuclei, which he called nuclein (1). We now know that he had discovered DNA (deoxyribonucleic acid), the molecule that carries all genetic information needed for the functioning of life (2). Aided by X-ray crystallographic images on DNA fibres obtained by Rosalind Franklin, James Watson and Francis Crick determined the 3-dimensional structure of DNA in 1953 (3, 4). They showed the molecule is structured as two right-handed helical chains each coiled around the same axis, but running in opposite directions. Each of the two chains consists of a series of nucleotides, with each nucleotide carrying one of four bases. Watson and Crick found that only specific pairs of bases will bond together: adenine with thymine and guanine with cytosine. This specific pairing immediately suggested a possible copying mechanism of the genetic material. Five years after Watson and Crick solved the structure of DNA, Arthur Kornberg's group identified the mechanism of its synthesis and the first DNA polymerase, the enzyme responsible for this process (5).

DNA replication, or accurate duplication of parental double-stranded DNA (dsDNA) into, identical daughter copies, is essential for the propagation of all terrestrial life forms as it plays a crucial role in transmitting hereditary information from cell to cell. It is a fundamental cellular process that is carried out by a multi-protein complex known as the replisome. The replisome contains enzymatic activities responsible for many more processes than just DNA synthesis. It involves the separation of the parental dsDNA into two daughter strands, both of which serve as a template for the new copies of DNA. Due to the opposing polarity of the two DNA strands and the fact that new DNA can only be synthesised in one direction, one of the strands — the lagging strand —, is synthesised in a series of short Okazaki fragments (6) in the opposite direction to the leading strand, which is synthesised continuously. Since DNA synthesis can only occur through extension of a pre-exisiting structure, the production of each individual Okazaki fragment is initiated by a priming reaction. The replication of DNA is accomplished with a remarkable speed and at high accuracy. To provide a sense of scale we can estimate the amount of DNA that our cells have to replicate during our lifetime. There are approximately $2 \cdot 10^{14}$ cells in a human body and on average each divides 50 times during our lifetime (with significant variation between tissue type) (7). If we multiply these numbers with the length of the DNA sequence contained within one nucleus (approximately $6 \cdot 10^9$ bp (8)), we find that our cells produce an astonishing total amount of 10^{16} meters of DNA. This length is roughly equal to one light year!

1.2 Building complexity

Over the past few decades, a large variety of ensemble-averaging biochemical techniques have been used to study the roles of the various proteins within the replisome. One approach to dissecting the multiple events that occur during DNA replication has been to study simple replication systems of bacteriophages such as T4, T7 and ϕ29, and bacteria such as *Bacillus subtilis* and *escherichia coli* (*E. coli*). The number of proteins required for DNA replication in these systems is relatively small, however, the basic steps in DNA replication are similar to those found in higher organisms. Building up complexity in a similar way, work described in this thesis has been done on the T7, *E. coli*, and *Saccharomyces cerevisiae* (*S. cerevisiae*) replisomes. I will, therefore, give a short overview of these systems.

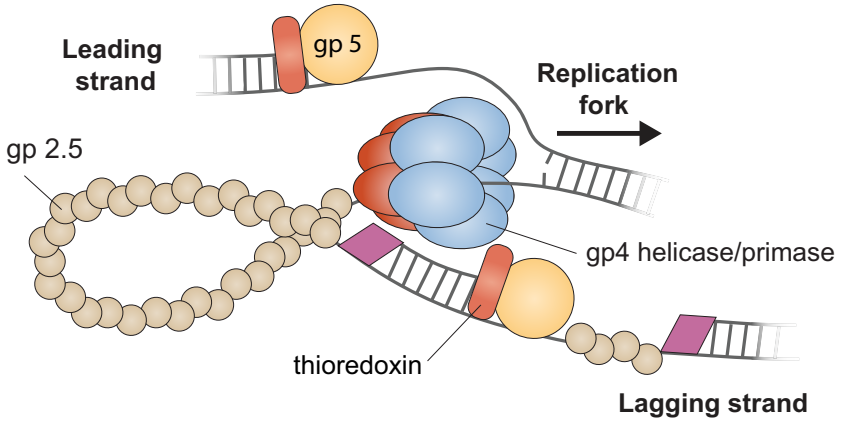
1.2.1 Bacteriophage T7

Figure 1.1a shows a schematic representation of the T7 replisome. It can be reconstituted *in vitro* from just four proteins (9). The gene product 4 (gp4) provides both helicase and primase activities. The helicase activity of gp4 is performed by the C-terminal portion of the protein and the N-terminus contains the primase. Gp4 forms a hexameric ring upon binding to ssDNA, and uses the energy derived from the hydrolysis of dTTP to translocate in a 5′ to 3′ direction (10). The DNA primase domain is comprised of two subdomains. A flexible linker connects the zinc-binding domain (ZBD) located at the N-terminus of the primase to the C-terminal RNA polymerase domain (RPD), where the RNA primers are synthesised (11). The DNA polymerase gene product 5 (gp5) synthesises new DNA on the two strands. It forms a complex with thioredoxin (trx) of the *E. coli* host. Trx functions as a processivity factor for gp5, providing a physical mechanism of stabilising the polymerase on the DNA (12). The single-stranded DNA binding protein, gene product 2.5 (gp2.5), binds exposed ssDNA and coordinates simultaneous synthesis of leading- and lagging-strands. Gp2.5 also plays a role in recombination and in the repair of double-stranded breaks in phage DNA (13). The reconstituted replisome can duplicate DNA at a rate of 80 bp/s (14).

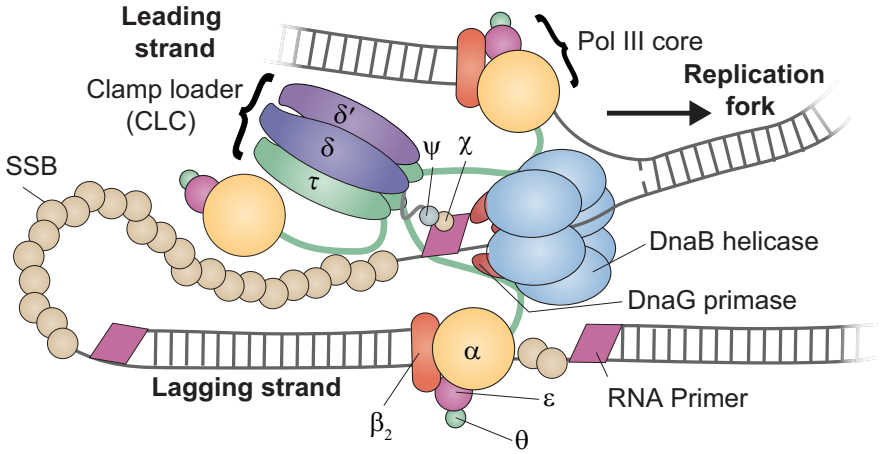
1.2.2 *escherichia coli* replisome

The *E. coli* replisome is an example of an increasingly complex replication system. The majority of work described in this PhD thesis has been done with this system. I shall, therefore, discuss it in a bit more detail. With a dozen individual subunits the *E.coli* replisome is still relatively small compared to the replication complexes of higher organisms, yet significantly more complex than the T7 system. By now, the *E. coli* replisome is perhaps the best understood across all species. Once assembled and active, the *E. coli* replisome unwinds and duplicates DNA at a very high rate, approaching 1000 bp/s with an error rate of roughly one mistake for every 10^{-6} to 10^{-7} nucleotides synthesised (17). Figure 1.1b shows a schematic representation of the *E. coli* replisome. Similar to the T7 repli-

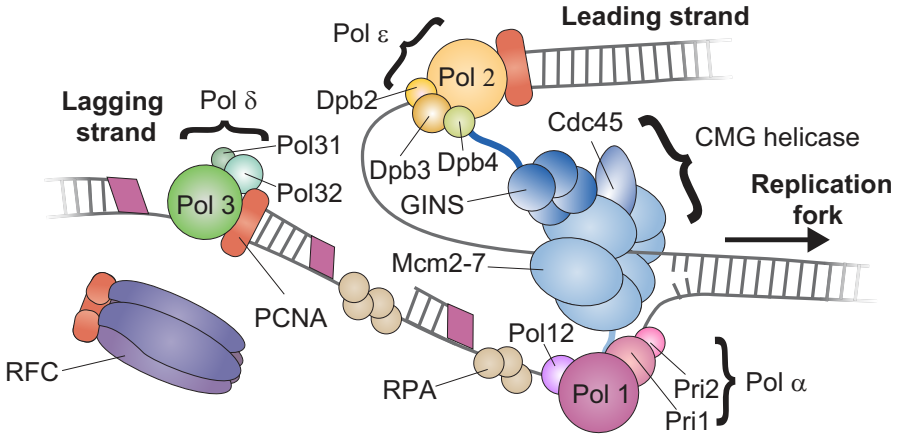
a T7 replisome



b *E. coli* replisome



c *S. cerevisiae* replisome



some, DNA is unwound by a hexameric helicase DnaB which uses the energy of ATP hydrolysis to unwind dsDNA (18). DnaB consist of six identical 52 kDa subunits oriented in the same direction, and is assembled in the presence of Mg^{2+}. It is loaded onto the DNA by the helicase loader DnaC (19). The C-terminal motor domains of DnaB, which are located at the front of the replisome, bind and hydrolyse ATP to drive the unwinding of the dsDNA. The N-terminal domain of DnaB forms a trimer of dimers, which serve as binding sites for up to three molecules of DnaG primase (20). DnaG, which synthesises short RNA primers on the lagging strand for initiation of DNA synthesis on the lagging strand, consists itself of tree domains. It has a zinc-binding domain (ZBD) at the N terminus, which is essential for primase activity and is thought to recognise priming sequences in the ssDNA (21). The central RNA polymerase domain (RPD) is responsible for NTP binding and incorporation (22). The C-terminal helicase-binding domain (HBD) binds to the N-terminal domain of DnaB to form the primosomal complex $DnaB_6$–$(DnaG)_3$ (23). The replicative polymerase in the *E. coli* replisome is the DNA Polymerase III holoenzyme (Pol III HE). The Pol III HE is able to synthesise both the leading and lagging strands simultaneously. It is arranged into three functionally distinct and stably-bound subassemblies. $\alpha\epsilon\theta$ forms the Pol III core that has DNA polymerase activity (24). By itself it will synthesise only 10–20 nt at a rate of $\sim$20 nt/s (25). β_2 is the sliding clamp, which encircles the DNA and ensures stable association of the core polymerases with the primer-template DNA, thus enabling higher processivities (26). β_2 is also know to interact with about a dozen proteins related to DNA replication, recombination, and repair (27). $\tau_n\gamma_{(3-n)}\delta\delta'\chi\psi$ (where n = 2 or 3 in the Pol III HE) is the clamp loader complex (CLC) that uses ATP hydrolysis to load β_2 onto DNA and is the central organiser of the replisome (28). Up to three Pol III cores are coupled through the τ subunits of the CLC, and the τ sub-

Figure 1.1 *(preceding page)*: Schematic representations of the T7, *E. coli*, and *S. cerevisiae* replisomes. (a) The replisome of bacteriophage T7 contains 4 proteins: the DNA polymerase (gp5) and its processivity factor, *E. coli* thioredoxin (trx), the DNA primase–helicase (gp4), and the ssDNA-binding protein (gp2.5). Gp4 unwinds the dsDNA and generates two ssDNA templates for the leading- and lagging-strand gp5/trx. Gp2.5 coats the lagging-strand ssDNA. The primase domain of gp4 catalyses the synthesis of short primers for the initiation of each Okazaki fragment. Figure adapted from Robinson et al. (15). (b) Architecture of the *E. coli* replisome at the chromosomal replication fork derived from *in vitro* studies and direct observation *in vivo*. The DnaB helicase is located at the apex of the replication fork on the lagging strand. The single-stranded lagging-strand template produced by helicase action is protected by the Single-Stranded DNA Binding protein (SSB). The DNA polymerase III holoenzyme (Pol III HE) synthesises new DNA on both strands. The β_2 sliding clamp confers high processivity on the DNA Pol III HE by tethering the Pol III $\alpha\epsilon\theta$ core complexes onto the DNA. The clamp loader complex (CLC) assembles the β_2 clamp onto RNA primer junctions on template DNA. DnaG primases synthesise RNA primers to initiate DNA synthesis on the lagging strand. Figure adapted from Lewis et al. (16). (c) Architecture of the *S. cerevisiae* replisome. The CMG helicase loads onto the leading strand and unwinds the dsDNA. CMG consists of the Mcm2–7 complex, which has the ATPase activity, and the accessory GINS complex and Cdc45. Pol α synthesises short RNA/DNA primers on both strands. Pol ϵ and Pol δ extend these primers on the leading and lagging strand, respectively. The PCNA sliding clamp confers high processivity on the DNA polymerases by tethering them onto the DNA. The clamp loader RFC assembles PCNA onto the DNA. RPA coats the transiently exposed single-stranded DNA

units also interact with DnaB, thus organising and coupling the DNA Pol III HE to DnaB (29). The minimal CLC which is proficient in clamp loading and pol III core binding is the $\tau_3\delta\delta'$ heteropentamer. The χ–ψ subunits are accessory proteins that connect the CLC with the Single-Stranded DNA Binding protein (SSB) (30). SSB binds to ssDNA in a sequence-independent manner to protect it against nucleolytic attacks and to prevent the formation of any secondary structures (31). SSB is also an important interaction partner for a large number of proteins, and therefore plays a central role in many DNA replication, recombination, and repair processes (32). It forms a homotetramer of 19 kilodalton subunits. The N-terminal domain forms an oligonucleotide binding (OB) fold responsible for ssDNA binding (33). The four ssDNA-binding domains enable it to bind tightly to ssDNA in different modes with different properties depending on salt concentrations (34). At low monovalent salt concentrations, binding in the $(SSB)_{35}$ mode is favoured. In this mode, the DNA interacts with only two of the four SSB subunits, resulting in a footprint of 35 nt per tetramer. It is suggested that the C termini of SSB may interact, at least transiently, with the ssDNA-binding sites of neighbouring SSB proteins. This interaction suggests a mechanism that enhances the ability of SSB to selectively recruit its partner proteins to sites on DNA. Also, it allows for very high cooperative binding, which results in the formation of SSB clusters along the ssDNA (35). At higher salt concentrations binding occurs mostly in the less cooperative $(SSB)_{65}$ mode (36), in which 65 nt interact with SSB (37). SSB can utilise a direct transfer mechanism through which SSB can be transferred from one ssDNA molecule to another without proceeding through a free protein intermediate. It is hypothesised that this could enable recycling of SSB tetramers between old and newly formed ssDNA regions during lagging-strand DNA replication (38).

1.2.3 *Saccharomyces cerevisiae* replisome

Very recently, the minimal *S. cerevisiae* (yeast) replisome has been reconstituted (Figure 1.1C) (39). Consisting of at least 31 individual proteins, this system is far more complex than any reconstituted replisome studied before. The dsDNA is unwound by the 11-subunit helicase CMG. The motor of CMG is the Mcm2–7 complex, a heterohexamer of AAA+ ATPase subunits. Mcm2–7 forms a ring around the leading strand and has 3′–5′ helicase activity (40), in contrast to the T7 and *E. coli* helicases, which unwind DNA in the opposite direction. The helicase is activated upon association of Mcm2–7 with Cdc45 and the four-subunit GINS (Japanese spelling of the numbers 5,1,2,3 $\underline{g}$o-$\underline{i}$chi-$\underline{n}$i-$\underline{s}$an) to form the CMG complex (41). The DNA polymerase α primase (Pol α) acts as a primase by synthesising a hybrid RNA/DNA primer of 2–30 nucleotides (42). These primers get extended by the DNA polymerase ϵ (Pol ϵ) and by DNA polymerase δ (Pol δ). Though both polymerases can function on either strand, Pol ϵ is favoured on the leading strand and Pol δ on the lagging strand (39). It has been shown that Pol ϵ directly binds to CMG, forming a stable complex (43). Pol δ requires PCNA to stabilise it on the DNA for high processivity. The clamp loader, Replication Factor C (RFC), uses ATP to load PCNA onto the DNA (44). Replication Protein A (RPA) is the eukaryotic single-stranded DNA binding protein.

Even though much is known about the structure and function of the different proteins within the replisome, the coordination of multiple components, activities, and interactions within the replisome involve transient intermediates and dynamic conformational changes that are difficult, if not impossible, to observe with ensemble experiments. Recently, new single-molecule techniques have been developed to study the dynamics of proteins with a high precision and without the need for population averaging. This thesis centres on the use of these approaches to study dynamic behaviour of the replisome.

1.3 Single-molecule techniques

1.3.1 Why single molecules?

In ensemble experiments, a measurement of a molecular property represents the measurement of the average behaviour of many individual components. Observing molecular properties at the single-molecule level allows characterisation of subpopulations, the visualisation of transient intermediates, and the acquisition of detailed kinetic information that would otherwise be hidden by ensemble averaging. This concept can be illustrated by making an analogy to the delivery trucks (Figure 1.2) (45). With ensemble-averaging methods, we can measure the average speed of the delivery trucks. We can not tell, however, if all the trucks are moving at the same speed (Figure 1.2a), or whether some trucks are speeding while others take it easy (Figure 1.2b). Furthermore, the trucks could be changing speed, or stopping for a break (Figure 1.2c). Again we would not be able to know this from the average speed.

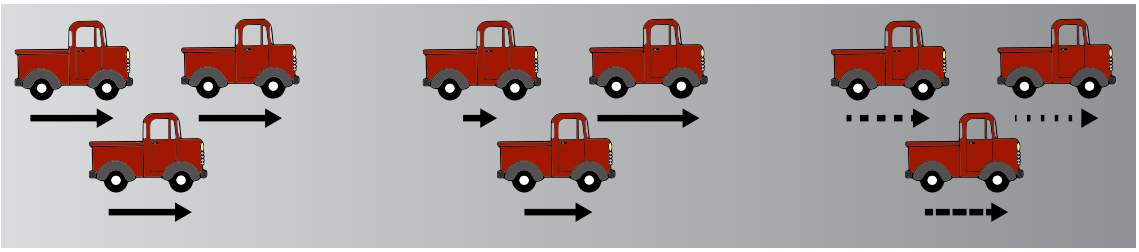


Figure 1.2: Ensemble versus single-molecule studies. Through ensemble studies we can obtain information about the average behaviour of a system, for example the average speed of all the trucks, indicated by the arrows (left). Through single-molecule studies we can see whether some molecules within the system behave differently, for example, if the trucks have different speeds (middle). We can also observe changes in behaviour of individual molecules, illustrated by the trucks changing speed, indicated by the dashed arrows (right).

The goal of single-molecule experiments is to remove this ensemble averaging and to observe the heterogeneity within the system. In the last decade, single-molecule experiments have taught us much about the dynamics within the replisome. For example, in the bacteriophage T4 system, single-molecule studies have revealed the pathway for assembly of the primosome (46) and have provided a detailed real-time visualisation of the DNA helicase unwinding activity (47).

Other studies revealed the real-time dynamics of the conformational change of the β_2 clamp (48). For the T7 system, the textbooks told us that polymerases are very stably bound to the replisome and synthesise the entire genome (49). Using single-molecule fluorescence imaging, it was shown that the polymerases exchange from solution at the rate of Okazaki fragment synthesis (50).

During my PhD, I primarily used single-molecule visualisation methods that rely on mechanically stretching individual DNA molecules and the imaging of individual fluorescent proteins acting on DNA. I will discuss these methods in more detail in the next section.

1.3.2 Single-molecule fluorescence imaging

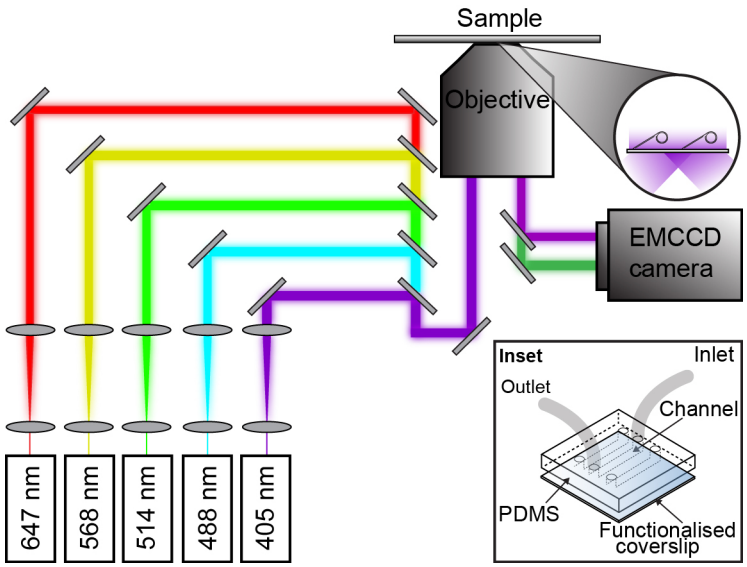


Figure 1.3: Schematic representation of the fluorescence microscope. Laser light of a specific wavelength is coupled into the microscope objective. The fluorescence signal from the sample is detected with either an sCMOS or an EMCCD camera. **Inset**, Micro-fluidic flow cell schematic. A PDMS lid containing three flow chambers is placed on top of a PEG-biotin-functionalised microscope coverslip. Tubing inserted into the PDMS provides easy access to the reaction chamber.

To enable the visualisation of the behaviour of a protein at the single-molecule level, the protein of interest is labelled with a fluorophore. This fluorophore is excited by laser light of appropriate wavelength. As a result, the molecule will act as a point source for emitted fluorescence photons which can be imaged as a focussed, diffraction-limited spot by a very sensitive EMCCD or CMOS camera (figure 1.3). By tracking the intensity and position of the fluorophore we can obtain information about the dynamic behaviour of single proteins, in real time (51). By using two fluorescent probes that emit light at different wavelengths, we can track the behaviour of two proteins simultaneously. Single-molecule fluorescence methods were originally only applicable at low nanomolar concentrations, due to the diffraction-limited nature of the optics in a fluorescence microscope and the

resultant minimal size of the excited volume. In order to resolve single molecules, fluorophores have to be spaced further apart than the diffraction limit. This limit depends on the numerical aperture (NA) of the microscope and the wavelength (λ) used and is typically a few 100 nm (equation 1.1, where n = the index of refraction and θ = the angle of the incident beam).

$$d = \frac{\lambda}{2NA}; NA = 2nsin(\theta). \tag{1.1}$$

Over the past few decades rapid developments in fluorophore stability and imaging techniques have increased the variety and complexity systems probed by single-molecule fluorescence tools tremendously.

1.3.3 Tethered-bead assay

In the tethered-bead assay, a forked linear DNA template is tethered to the surface of a microscope cover slide at one end, and to a paramagnetic bead at the other end (Figure 1.4). By applying a laminar flow across the slide, the resultant Stokes drag force on the bead will stretch out the DNA (equation 1.2, where η = the viscosity, R = the radius of the bead, and v = the velocity of the laminar flow).

$$F_d = 6\pi\eta Rv. \tag{1.2}$$

At force regimes of $\sim$1–2 pN dsDNA is much longer than ssDNA. A conversion of dsDNA to ssDNA, for example by leading-strand DNA synthesis, will therefore cause the bead to move against the direction of the flow. By visualising and tracking the movement of the bead, we can get very accurate (1–10 nm precision) information on replication kinetics at the fork. In the last years this technique has seen some major improvements both to the hardware as well as the, now automated, data analysis. Using a low-magnification wide-field microscope (52), 10,000 beads can be imaged simultaneously, making this a very high throughput single-molecule assay (53). A typical experiment will now generate $\sim$50 GB of data. Such a data volume required a stream-lined and automated analysis software to extract useful single-molecule parameters (54).

The great advances in both the fluorescence-imaging techniques and tethered-bead assay are giving rise to new insights in the behaviour of dynamic multi-protein systems. We have come to learn that their behaviour is not linear and deterministic, as previously suggested, but actually highly dynamic and subject to a great level of stochasticity (55).

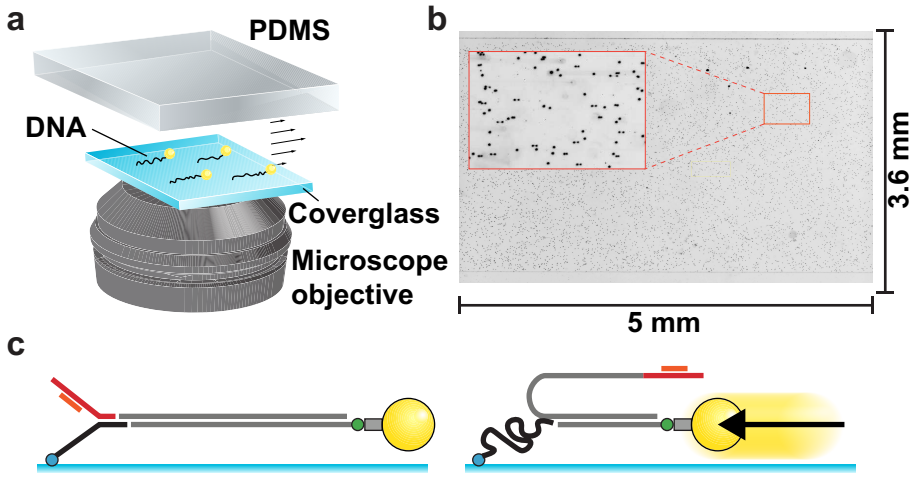


Figure 1.4: Schematic representation of the fluorescence microscope. (a) Experimental setup. Individual DNA molecules are tethered to the surface of a microfluidic flow cell. Beads are attached to the DNA ends and imaged using low-magnification wide-field microscopy. (b) A representative field of view showing 4,000 beads. (inset) Image of beads attached to DNA flow-stretched in both directions by a flow reversal. (c) Architecture of typical DNA template used in the tethered-bead assay. (left) Linear DNA with a replication fork is attached to the surface of a micro-fluidic flow cell. A bead attached to the other end of the DNA stretches the DNA in the direction of flow. (right) Length changes due to the conversion from dsDNA to ssDNA result in a movement of the bead.

1.4 Scope of this thesis

The goal of my PhD research is to develop and use single-molecule tools to understand the molecular mechanisms in DNA replication. Chapter 2 is a summary of the latest advances in force- and fluorescence-based single-molecule methods both *in vitro* and *in vivo*, with a focus on their applications in studies on cytoskeletal motors and DNA replication. We describe how these advances now allow us to study increasingly complex systems. We developed a new fluorescence imaging technique, by which single fluorescent molecules can be observed in real time at high, physiologically relevant concentrations (Chapter 3). Single-molecule tools are being used in increasingly broader areas of research and have now also found their use in the development of targeted drug delivery mechanisms. Stepping away from DNA replication for a bit, in Chapter 4 is described how we use single-molecule fluorescence imaging, to determine the density of proteins on functionalised liposomes. This density is a pharmacologically important number that had not been properly quantified before. In Chapters 5 and 6 I discuss the exchange behaviour of polymerases and SSB in the *E. coli* replisome under physiologically relevant protein concentrations. We see rapid exchange, depending on the concentration of competing protein in solution. The emergence of the concentration dependence illustrates the importance of studying the molecular sociology (56) of multi-protein complexes at the single-molecule level. In Chapter 7 we determine the effect of the *E. coli* RarA protein on DNA replication and repair. RarA is a highly conserved protein whose role in replication and repair was poorly understood. Combining *in vitro* and *in vivo* techniques, we propose that RarA activity is involved in the creation of gaps in lesion-containing DNA templates, and thereby commits the cell to the translesion DNA synthesis repair pathway. In the last step of my PhD journey of building up complexity, we describe the first single-molecule experiments done on the reconstituted *S. cerevisiae* replisome, characterising the kinetics of leading-strand synthesis (Chapter 8). We confirm a previously reported observation that the MTC complex enhances the speed of the replication fork by $\sim$2 fold. Surprisingly, however, our data suggest that MTC only transiently interacts with the replisome through a weak interaction.

2 | Watching cellular machinery in action, one molecule at a time.

Enrico Monachino[†], **Lisanne M. Spenkelink**[†], Antoine M. van Oijen.

[†]These authors contributed equally.

Published in *Journal of Cell Biology*, 02 Jan 2017;**216**(1):41–51.

Single-molecule manipulation and imaging techniques have become important elements of the biologist's toolkit to gain mechanistic insights into cellular processes. By removing ensemble averaging, single-molecule methods provide unique access to the dynamic behavior of biomolecules. Recently, the use of these approaches has expanded to the study of complex multiprotein systems and has enabled detailed characterization of the behavior of individual molecules inside living cells. In this review, we provide an overview of the various force- and fluorescence-based single-molecule methods with applications both *in vitro* and *in vivo*, highlighting these advances by describing their applications in studies on cytoskeletal motors and DNA replication. We also discuss how single-molecule approaches have increased our understanding of the dynamic behavior of complex multiprotein systems. These methods have shown that the behavior of multicomponent protein complexes is highly stochastic and less linear and deterministic than previously thought. Further development of single-molecule tools will help to elucidate the molecular dynamics of these complex systems both inside the cell and in solutions with purified components.

E.M. and I contributed equally to this review. I reviewed the single-molecule fluorescence imaging techniques, mainly focussing on their use in studies on DNA replication.

2.1 Introduction

Single-molecule approaches are transforming our understanding of cell biology. In the context of the living cell, proteins are found in various states of structural conformation and association in complexes, with the transitioning between states occurring in a seemingly chaotic fashion. Observing molecular properties at the single-molecule level allows characterization of subpopulations, the visualization of transient intermediates, and the acquisition of detailed kinetic information that would otherwise be hidden by the averaging over an ensemble of stochastically behaving constituents. Although the field is rapidly evolving, and many technical challenges still exist, methods to visualize individual proteins in purified systems, henceforth referred to as *in vitro*, contribute to a tremendous gain in mechanistic insight into many cellular processes. However, the comparatively low complexity of such *in vitro* experiments does not necessarily represent the physiology of the cell. Development of single-molecule tools has begun to enable the visualization of complex biochemical reactions with great resolution in the dynamic and crowded environment of the cell. *In vitro* single-molecule studies on reconstituted systems of high complexity are informing on how these systems may behave in a cellular environment, and live-cell single-molecule imaging is providing pictures of increasing clarity about the physiological relevance of pathways observed *in vitro*. This interplay between *in vitro* and *in vivo* assays will play a major role in future studies, with bottom-up and top-down approaches required to fill the gaps.

In this review, we provide an overview of the state of the field and discuss the main classes of single-molecule methods that have found applications in *in vitro* and *in vivo* studies. In particular, we describe the principles of both force- and fluorescence-based single-molecule methods, and we highlight how these approaches have increased our understanding of molecular machineries. Using recent work, we illustrate both the advances in methodology and new insights into the dynamic behavior of complex systems that they provide. To guide our review of the main technological developments and the biological breakthroughs they have allowed, in the context of what seems like an overwhelming amount of examples and applications, we focus on studies of the molecular motors that carry cellular cargo and the multiprotein complex involved in DNA replication, the replisome. Our focus on these studies merely represents an attempt to illustrate the methodological possibilities—the reader is advised to consult the many other excellent sources and reviews that discuss the use of single-molecule tools in other fields and systems.

2.2 Push, pull, poke and prod: Mechanical single-molecule techniques

The folding of proteins into functional structures, the manner with which they undergo conformational transitions, and their interactions between binding partners are all complex processes that are strictly ruled by the shape of the free-energy

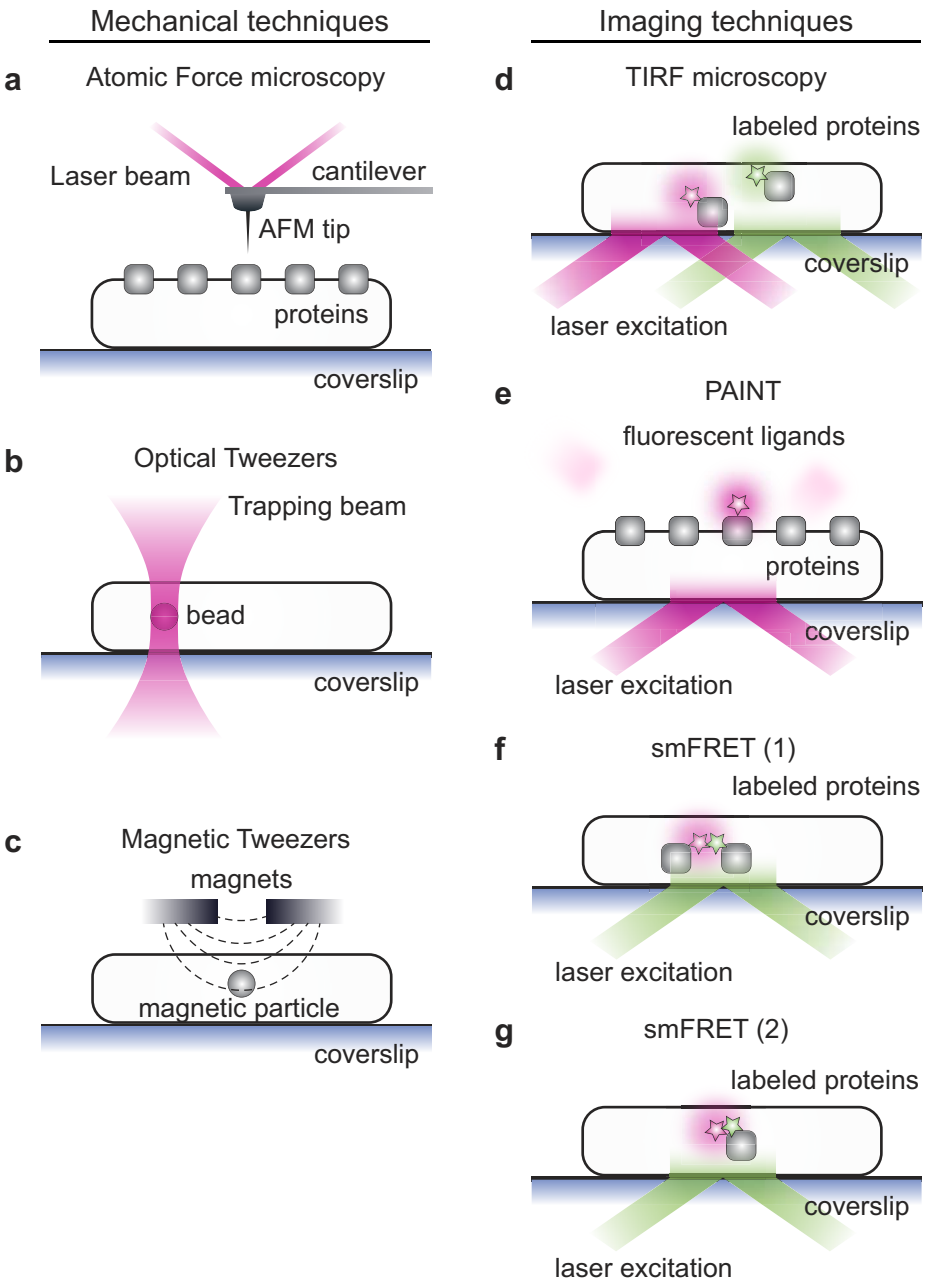

Mechanical techniques
Imaging techniques
a
Atomic Force microscopy
Laser beam
cantilever
AFM tip
proteins
coverslip
b
Optical Tweezers
Trapping beam
bead
coverslip
c
Magnetic Tweezers
magnets
magnetic particle
coverslip
d
TIRF microscopy
labeled proteins
coverslip
laser excitation
e
PAINT
fluorescent ligands
proteins
coverslip
laser excitation
f
smFRET (1)
labeled proteins
coverslip
laser excitation
g
smFRET (2)
labeled proteins
coverslip
laser excitation

Figure 2.1*(preceding page)*: Single-molecule approaches. (a) AFM. A tip is attached to a cantilever, with deflection of the tip or changes in its resonance frequency reporting on proximity to features on a cellular surface. By raster scanning the sample, an image of the 3D shape can be formed with subnanometer resolution. (b) OT. A functionalized bead is introduced into the cell. The bead is trapped and manipulated by a focused laser beam. (c) MT. Magnetic beads that specifically interact with a substrate of interest are introduced into the cell. By applying a magnetic field, the beads can be rotated or translated, thereby introducing a force to the system. (d) Fluorescence microscopy. Substrates of interest are labeled with a fluorescent tag. Their fluorescence is detected on a sensitive camera, allowing real-time visualization of spatiotemporal dynamics. (e) PAINT. This technique works by labeling a substrate that interacts transiently with a receptor. A low concentration of fluorescent ligands is introduced in the extracellular medium such that at a constant rate, receptors in the membrane are being visualized by short-lived fluorophore immobilization during the imaging sequence. (f) and (g) smFRET. (f) Two substrates of interest are labeled with two specific fluorescent tags (a donor-acceptor FRET pair). The emission of the donor tag spectrally overlaps with the absorption of the acceptor dye. The donor transfers its energy to the acceptor in a distance-dependent manner (FRET). An interaction between the two substrates will give a FRET signal, providing a dynamic observation of molecular interactions. (g) A molecule of interest is labeled with a FRET pair at known positions, one with a donor and the other with an acceptor. A change in the conformation of the substrate can be observed as a change in the FRET efficiency.

landscapes describing the thermodynamics of the system. Theoretically, there is a huge number of possible 3D conformations that a one-dimensional sequence of amino acids can assume, each characterized by a specific free energy. However, a protein assumes only those states that minimize the free energy, with preference for the absolute minimum. Thus, the number of possible protein conformations is limited to very few, if not only one (57). The application of forces to these systems introduces well-defined changes to the energetics and enables a precise interrogation of the relevant interactions and processes. Single-molecule mechanical techniques have been developed to use small forces to controllably manipulate individual biomolecules so that molecular mechanisms can be investigated at a level of detail inaccessible with conventional ensemble-averaged assays. In this paper, we focus on three main classes of these methods: atomic force microscopy (AFM), optical tweezers (OT), and magnetic tweezers (MT). Each of these techniques works in a different force regimen, with these three techniques together covering a range from femto-Newtons (fN) to nano-Newtons (nN), providing experimental access to forces that are relevant to biochemical processes and reactions. More comprehensive reviews on each technique and applications can be found elsewhere. (15, 58–67)

2.2.1 Atomic Force Microscopy

AFM is a scanning probe microscopy technique that allows visualization of the surface topography of a sample at subnanometer resolution. It uses an atomically sharp tip on the free end of a projecting arm (called cantilever) to measure the height (z axis) at a specific (x,y) position (Figure 2.1A). In biological imaging applications, AFM is typically used in the so-called tapping mode with the cantilever oscillating at a frequency close to its mechanical resonance. In this way, interactions with the surface can be detected with great sensitivity

without the tip in constant contact with the sample, thus eliminating dragging and frictional effects during the (x,y) scan and avoiding distortion of image data. Ultimately, the tapping mode helps to preserve the integrity of the soft biological sample and allows the visualization of biomolecules for periods up to hours (68).

AFM was initially limited to the imaging of static structures, but the last decade has seen the introduction of even smaller cantilevers (69) and improvements in the image acquisition rate, making it possible to scan surfaces at high speed (high-speed AFM [HS-AFM]). HS-AFM is one of the few techniques so far that allows observation of biological molecules at both subnanometer and sub-100-ms resolution. This technical breakthrough has enabled real-time observation of molecular processes, such as the movement of motor proteins along cytoskeletal filaments, and has allowed the direct study of relationships between structural and dynamic properties of biochemical reactions, at the single-molecule level, with one single technique (70). This powerful and quite unique ability of HS-AFM to relate structure to function was highlighted in a hallmark study in which the walking of myosin V on actin was imaged (Figure 2.2A) (71). Not only did the high-speed imaging visualize the hand-over-hand mechanism of myosin V translocation, but the authors of this study were also able to explain the mechanism in structural terms. They showed that the forward movement of the myosin is a purely mechanical process related to the accumulation of tension in the leading head. Recently, a further technical improvement has allowed imaging of large fields of view at high speed and visualization of biochemical reactions occurring on the outer surfaces of cells (72). *In vivo* biological imaging with AFM offers several advantages over other techniques with high spatial resolution such as scanning EM. In particular, AFM does not require dehydration steps and can provide topographic images with nanometer resolution under physiological conditions (73). These aspects position AFM as a technique with great potential to provide unique insight in various areas of cell biology such as membrane structure and dynamics, cell division, growth, and morphology. Finally, there have been attempts to bring AFM inside cells (65), opening to the use of its high spatial and temporal resolution to observe fundamental cellular processes inside the cell itself.

In addition to its topographic imaging applications, AFM is a powerful tool to perform force spectroscopy on single molecules in the 10 pN to 10 nN range. In this application, the tip of the AFM is used to capture one end of a biomolecule that is bound to a surface at its other end, apply a stretching force to it by moving the cantilever away from the surface, and thus unfolding it with a precise and controllable force (74, 75). This approach makes it possible to probe the molecular interactions that stabilize the protein in a specific conformation. The alternative conformations of proteins when subjected to mechanical forces inside the cell can now be revealed (75). Finally, by using different loading rates, researchers can model the kinetics of transitions and obtain details of the free-energy landscape controlling the various structural transitions (67). An early example of AFM-based force spectroscopy involved the unfolding of the integral

membrane protein bacteriorhodopsin out of archaeal purple membranes (76). Further, the role of ligands in stabilizing biomolecular structures can be assessed and quantified by mechanical unfolding. The interaction between a ligand and a protein affects the free-energy landscape of the system and potentially yields different unfolding profiles as a function of the ligand (77). This approach is not limited to answer fundamental questions about cellular mechanisms, but also benefits applied research. For instance, researchers have been able to study *in vivo* membrane protein–ligand interactions to facilitate drug development (78).

2.2.2 Optical Tweezers

In OT (also called optical traps), a tightly focused laser beam is diffracted by a dielectric particle, resulting in a force that traps the particle nearby the focus of the laser. At the same time, by changing the position of the focus, it is possible to move the particle, just as if the laser beam were a pair of tweezers. By tethering one end of a molecule of interest to the bead and the other end either to a surface or to a second trapped particle, a stretching force can be applied to the molecule in the 0.1–100 pN range. The applied force can be modulated by either changing the tightness of the trap or by moving the position of the particle with respect to the beam focus (Figure 2.1B). Tracking of the 3D displacement of the trapped particle allows measurements with subnanometer spatial resolution and sub-millisecond time resolution. Thanks to such precision, this technology has, for instance, enabled the visualization of the motion of motor proteins such as kinesins and dyneins along microtubules, (79, 80), myosins along actin, and nucleic-acid enzymes along DNA. (81, 82)
Anytime lasers are used, photo damage to biological samples is a reason of concern. In the case of OT, this problem is minimized because biological samples are almost transparent to the near-infrared wavelengths of the lasers that are typically used to trap particles (83). This compatibility with cellular specimens, combined with recently developed sophisticated force-calibration techniques (66, 84), allows the use of OT *in vivo* and opens the possibility of studying the same biological system both *in vitro* and *in vivo*. Such hybrid approaches will be key in filling the gap between the mechanistic understanding obtained from *in vitro* reconstituted systems and biochemical reactions that occur in a cellular environment.This strategy has been very successful already in the characterization of the motor proteins kinesins, dyneins, and myosins. (66, 85, 86). The Xie group played a pioneering role in the development and use of OT *in vivo* at the submillisecond time resolution needed to observe organelle transport (87, 88). They reported that, in living human lung cancer cells, cargoes carried by kinesins make individual steps of 8 nm, while those carried by dyneins make individual steps of 8, 12, 16, 20, and 24 nm, providing new insight into the cooperative effects of multiple dyneins carrying the same cargo (87). They also observed that kinesins and dyneins both have a stall force of around 7–8 pN (88). In a study by the Goldman group (89), it was shown that the force exerted by individual motors is the same both *in vivo* (in mouse macrophage cells) and *in vitro*. These researchers suggest, however, that the viscoelastic cell environment and the pres-

ence of cytoskeletal networks favor motor binding. By comparing *in vitro* with *in vivo* experiments, they propose that in living macrophage cells, cargo is carried by as many as twelve dyneins and up to three kinesins in a tug-of-war mechanism (Fig. 2B–F). A study by the Selvin laboratory (90) characterized the transport of lipid vesicles and phagocytosed polystyrene beads in A549 human epithelial cells and in *Dictyostelium discoideum*, allowing them to propose that a single kinesin is sufficient to carry the cargo towards the periphery of the cell, while two to three dyneins are needed to transport the cargo towards the center. During outward motion, dyneins act as a drag on the kinesin-cargo translocation by pulling the cargo in the opposite direction. During inward motion, the kinesin is still bound to the cargo but not to the microtubule, and therefore does not obstruct the action of the dynein (90).

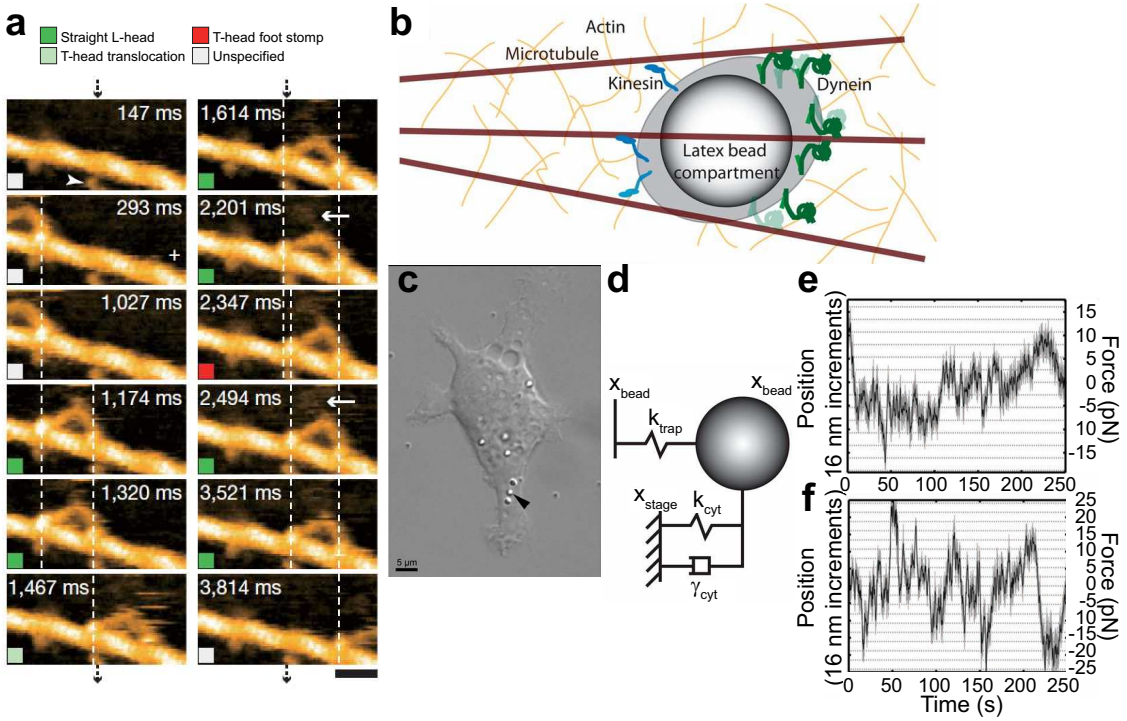


Figure 2.2: Force based measurements on motor proteins. (a) Myosin V walking on actin was directly observed using high-speed AFM. The acquisition times are indicated on each frame. Bar, 30 nm. (a is adapted with permission from (71)) (b–f) The *in vivo* transport of intracellular cargoes and the associated forces were measured with OT. (b) Cartoon describing the experiment. Multiple copies of the motor proteins dynein and kinesin carry along microtubules a bead that has been internalized by the cell. The bead was optically trapped and its movement tracked. (c) Picture of a mouse macrophage cell with internalized polystyrene beads (arrowhead is pointing at one of the beads). (d) Diagram indicating the various contributions experienced by the bead because of the trapping force and viscous drag experienced inside the cytoplasm. (e and f) Example trajectories tracking the displacement of the bead with respect to the beam focus in living cells. (b–f are adapted from (89)).

2.2.3 Magnetic Tweezers

MT are conceptually similar to OT: a magnetic field is used to trap a super-paramagnetic bead that is bound to one end of the molecule of interest (Figure 2.1C). MT can apply forces between fN up to several hundreds of pN, depending on the experimental design. Importantly, unlike OT, MT can apply torque by making use of the fact that magnetic beads act as a dipole with a preferred orientation in the external magnetic field. By applying bright-field illumination and using the interference patterns of the individual beads to provide information on their position with respect to the focal plane, the movement of the beads can be tracked with nanometer resolution. The large homogeneity of magnetic fields allows tracking of hundreds of beads simultaneously, a throughput difficult or impossible to achieve with OT. Moreover, magnetic fields are very selective for the magnetic particles and, therefore, do not interfere with the biological system under study, making MT ideal for *in vivo* investigations. The downside of this approach, compared to OT, is the difficulty of combining high forces with three-dimensional control over the magnetic bead. *In vivo* MT experiments have been reported (91), but more development is needed for the method to be employed as an alternative to OT.

Recent developments in bright, laser-based illumination sources, improvements in CMOS camera speeds, and the introduction of GPU-based calculation have made it possible to acquire bead images and track them in real time at kHz rates. These methods have made it possible for MT experiments to achieve sub-nanometer and sub-millisecond resolution and have enabled the observation of *in vitro* processes in real time at high spatiotemporal resolution (92). The combination of force and torque provided by MT has proven to be ideally suited to study DNA conformations and the activity of DNA-binding proteins. For example, it has revealed important mechanistic aspects of proteins involved in DNA replication. Studies investigating primer extension with the T7 polymerase and Escherichia coli Pol I (Pol I) resulted in a model where DNA synthesis is rate-limited by conformational changes involving multiple nucleotides on the template strand (93). Using MT to study helicase activity of the T4 bacteriophage and its coupling to partner proteins in the replisome, such as the primase and the polymerase, provided new insight into how the replisome is assembled onto DNA and how DNA replication is initiated. These experiments visualized how the synthesis of an RNA primer on the lagging strand results in the formation of loops of single-stranded DNA (ssDNA), a phenomenon that later was shown to occur in other replication systems (52, 94, 95). A study of the interplay between the T4 phage helicase and its DNA polymerase activities revealed that replication is faster than the unwinding by the helicase or synthesis by the polymerase as individual activities. Since the physical interaction between the two proved to be very weak, such synergies suggest an important role for ratchet-type mechanisms in speeding up reactions that consist of both reversible and irreversible steps (96). Recent studies on replication termination demonstrate the strength of mechanical approaches in their ability to apply external forces to rationalize mechanistic aspects of findings originally made *in vivo*. By using MT to exert different levels

of force to the *E. coli* Tus-Ter replication fork barrier *in vitro* and by observing its lifetime on DNA, a pathway describing barrier formation was proposed that reconciled previous structural, biochemical and microbiology studies (97).

Summarizing, it is clear that the various experimental platforms to apply mechanical force to individual molecules represent a powerful toolbox, each method with its own strengths and weaknesses. AFM combines high-resolution microscopy with force manipulation, with high time resolution. First, a biological sample is imaged, and then a specific part of it is directly probed. Therefore, it can provide structural, dynamic, and force information all from a single platform. OT and MT, instead, offer only force manipulation, but they can follow dynamics up to 100 times faster than AFM, thus granting access to short-lived states. Furthermore, both OT and MT can probe soft biological samples with virtually no damage at all. In the case of OT, this aspect has resulted in a mature tool for *in vivo* investigations, allowing mechanical manipulation inside the cell.

2.3 What you see is what you get: Imaging techniques

Mechanical single-molecule techniques allow the precise measurement of force and energy changes and have, therefore, been invaluable to studies on protein folding, DNA stability, and protein–DNA interactions. In this section, we describe single-molecule fluorescence imaging methods, approaches that take a more passive approach than force-based methods in that they are based on the visualization of mechanically unperturbed, fluorescently tagged molecules. Single-molecule fluorescence imaging methods are especially powerful in the visualization of molecular associations, copy numbers, conformational changes in biomolecules and enzymatic activity, often in real time. By using a fluorescence microscope equipped with a laser source to excite the fluorescent tag, and a sensitive camera to detect its fluorescence emission, a single fluorophore can be imaged with high spatiotemporal precision (10s of nanometers within 10s of milliseconds). Labeling with such fluorophores, therefore, allows direct, real-time observation of a system of interest (Figure 2.1D). The first single-molecule fluorescence experiment was performed in 1990 under cryogenic conditions (98). These low temperatures were necessary to increase the stability and lifetime of the fluorophores. Only five years later, the increase in the quality of optics and photon detectors allowed the first room temperature single-molecule experiment to be performed, showing individual ATP turnovers by myosin (99). The limited stability and lifetime of fluorophores impose significant challenges on the use of fluorescent tags to follow the dynamics of individual biomolecules, as they affect the quality of the signal and the duration of the experiment. Furthermore, the fluorophores need to be able to be specifically linked to a biomolecule of interest. Through the development of new fluorophores and photo-stabilizing compounds (100, 101), the brightness, the stability, and the lifetime of fluorescent probes have increased significantly. Current efforts are directed towards improving the compounds that

confer increased photostability to reduce their toxic effects and potential interference with the system of interest (102). Another key challenge in single-molecule fluorescence imaging experiments is the optical diffraction limit, giving rise to a lower limit of the smallest detection volume achievable. At high concentrations, this limitation results in a total number of fluorophores in the detection volume that is too large to allow single-molecule detection. As a result, single-molecule fluorescence-imaging tools were originally only useful at low nanomolar concentrations. Initial methodological advances were mainly made in the area of molecular motors, like DNA-based polymerases, myosins and kinesins (103), in part because the tight binding of these systems to their templates allows their study at very low concentrations. Over the past decade, developments in fluorophore stability and imaging techniques have increased the useful concentration range for single-molecule imaging by $\sim$10,000-fold. These developments have expanded the variety and complexity of systems probed by single-molecule fluorescence tools tremendously.

2.3.1 Total internal reflection fluorescence (TIRF)

One of the first methods introduced to increase the useful concentration range of single-molecule fluorescence imaging was TIRF (Total Internal Reflection Fluorescence) microscopy. In TIRF microscopy (104), an evanescent wave excites only those molecules in a $\sim$100-nm thin layer above a glass-water interface (105). Though TIRF can be used to study molecular and cellular phenomena at any liquid/solid interface (such as transport on membranes), it has proven to be most useful in single-molecule microscopy. The reduction of the excited volume as a result of the thin evanescent wave results in an increase of the signal-to-background ratio that allows high-contrast imaging of single molecules up to a concentration of $\sim$10s of nM. A good example of the application of TIRF microscopy in single-molecule studies is the mechanism of DNA replication. Applying TIRF imaging to purified and fluorescently labeled replication proteins acting on surface-tethered and flow-stretched DNA molecules, the dynamic behavior of bacteriophage T7 polymerases within replisomes was visualized during DNA synthesis. Though it was previously assumed that polymerases are stably bound to the replication fork, it was demonstrated that the polymerases in fact rapidly exchange with those in solution (50). TIRF microscopy has also allowed the real-time visualization of *in vitro* reconstituted eukaryotic replication-origin firing. It was shown that the helicase motor domains Mcm2–7 bind as double hexamers preferentially at a native origin sequence and that single Mcm2–7 hexamers propagate bidirectionally, monotonically, and processively as constituents of active replisomes (106). For kinesins, TIRF microscopy has been used to work out a longstanding mechanistic controversy on their walking mechanism. By labeling a single head of dimeric kinesin with a fluorophore and localizing the position of the dye, it was observed that a single kinesin head moves in alternating steps of 16.6 nm and 0 nm. This observation proves that kinesins take steps in a hand-over-hand mechanism, and not an inchworm mechanism (107).

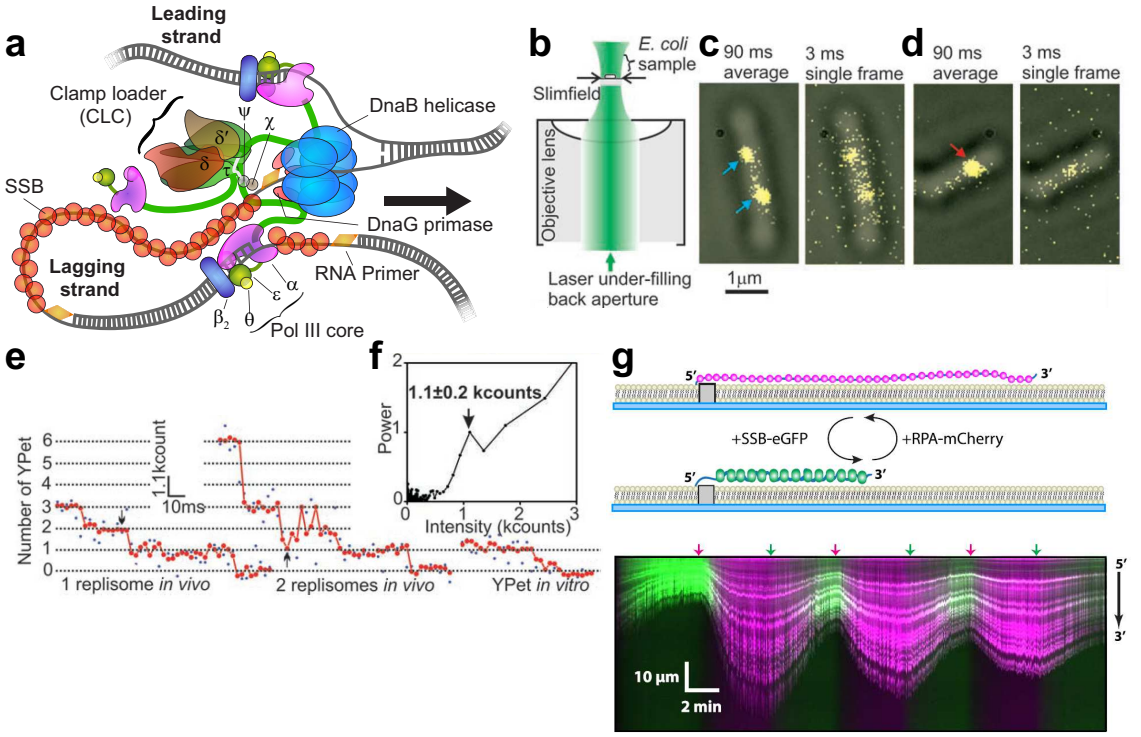


Figure 2.3: Fluorescence imaging of DNA replication. (a) Schematic representation of the *E. coli* DNA-replication machinery. Coordinated unwinding of parental double-stranded DNA and synthesis of two daughter duplexes is catalyzed by a large multiprotein complex, the replisome, built up from 12 different proteins and held together by a large number of weak and strong protein–protein and protein–DNA interactions. (b–f) Quantitative characterization of the number of polymerases per replisome in living *E. coli* using single-molecule slim-field microscopy. (b) Laser light is focused on the back aperture of the microscope objective, generating an intense Gaussian field at the sample just large enough to image a single *E. coli* cell. (c and d) Overlay of bright-field images of cells (gray) and 90-ms frame-averaged fluorescence images (yellow) of fluorescently labeled polymerases (ϵ-YPet). The blue arrows point at replisomes with three polymerases and the red arrow indicates a replisome with six polymerases. (e) Raw (blue) and filtered (red) intensity for a putative single (left panel) and double (middle panel) replisome spot were compared with the intensity of a single surface-immobilized YPet *in vitro* (right panel). Combined with the Fourier spectral analysis to find the brightness of a single YPet (f), these data show that the *in vivo* steps were integer multiples of the intensity of a single YPet molecule and replisomes contain a mean of three polymerases. (b–f adapted from (108)). (g) Two-color fluorescence imaging of the concentration-dependent exchange of ssDNA binding proteins on ssDNA. A microfluidic flow cell with ssDNA curtains was alternatingly injected with RPA-mCherry (magenta) and *E. coli* ssDNA binding protein (SSB)-EGFP (green). The exchange is evident by the change in color of the fluorescence and length of the ssDNA. Arrows placed above the kymograph indicate the time points of the injections. (g is adapted from (109)).

In vivo, near-TIRF microscopy has been used to examine the replisome stoichiometry and architecture in living cells. Using fully functional fluorescent YPet derivatives of *E. coli* replisome components expressed from their endogenous promoters, it was shown that active replisomes contain three molecules of the replicative polymerase Pol III core, rather than the historically accepted two (Figure 2.3A–F) (108). The mutagenic polymerase pol V, one of the players in the bacterial SOS response to DNA damage, was recently visualized at the single-molecule level in live *E. coli* cells. It was shown that pol V is, beyond the known

regulatory mechanisms at the transcriptional and posttranslational level, subject to a novel form of spatial regulation, in which it is transiently sequestered at the inner cell membrane (110). Movement of kinesins and dyneins has been observed inside living cells using Fluorescence Imaging with One Nanometer Accuracy (FIONA). Green fluorescence protein (GFP)-tagged peroxisomes in cultured Drosophila S2 cells were located within 1.5 nanometers in 1.1 milliseconds. Surprisingly, dyneins and kinesins do not work against each other during peroxisome transport *in vivo*. Rather, multiple kinesins or multiple dyneins work together, producing up to ten times the speed previously reported in *in vitro* measurements (111).

2.3.2 Local activation of dye (LADye), photoactivation, diffusion, and excitation (PhADE), point accumulation for imaging in nanoscale topography (PAINT)

To reduce the background fluorescence even further and to enable the visualization of individual labeled molecules at physiologically relevant concentrations, techniques have been introduced that rely on photoactivatable tags. In PhADE (PhotoActivation, Diffusion, and Excitation) (112), a protein of interest is fused to a photoactivatable protein and introduced to its surface-immobilized substrate. After photoactivation of the protein near the surface, rapid diffusion of the unbound proteins away from the detection volume reduces background fluorescence, whereupon the bound molecules are imaged. This method allowed the visualization of the micrometer-scale movement of replication forks, the spatiotemporal pattern of replication initiation along individual DNA molecules, and the dynamics of individual proteins at replication forks in undiluted cellular extracts (112). The drawback of this technique is the need for photoactivatable proteins. In an alternative method, LADye (Local Activation of Dye) (113) relies on the labeling of proteins with fluorophores that are chemically darkened (114). Only those proteins bound to their substrate are selectively activated, via a short-distance energy-transfer mechanism. Although the chemicals used to darken the fluorophores could potentially alter the behavior of the system, this approach has already allowed the observation of the sequence-independent interaction of interferon-inducible protein 16 (IFI16) with DNA and the sliding via diffusion of adenovirus protease (pVIc-AVP) on DNA in the presence of very high, μM concentrations of protein (113). PhADE and LADye have increased the useful concentration of proteins in *in vitro* single-molecule experiments to levels closer to *in vivo* conditions than ever before, thereby providing new insight into the behavior of DNA-interacting proteins at physiologically relevant concentrations.

The concentrations of most proteins inside living cells are well above the concentration limit that allows visualization using conventional single-molecule imaging methods (16). Therefore, similar techniques to reduce background fluorescence are used *in vivo*. In PAINT (Point Accumulation for Imaging in Nanoscale Topography) (115), the objects to be imaged are continuously

targeted based on many cycles of transient association by fluorescent probes present in the solution, rather than having the fluorescent probe stably bound to the objects. As a result, a fluorescent signal appears as a diffraction-limited spot on the object when a label briefly binds to it and is momentarily immobilized (Figure 2.1E). This method was employed to track endogenous AMPA glutamate receptors (AMPARs) on living neurons, revealing high receptor densities and reduced diffusion in synapses (116).

2.3.3 Single-molecule fluorescence resonance energy transfer (smFRET)

Fluorescence (Förster) Resonance Energy Transfer (FRET) is the distance-dependent non-radiative energy transfer between two fluorescent molecules that occurs when the emission spectrum of one fluorophore overlaps with the absorption spectrum of the other. Measuring the FRET efficiency allows the visualization of changes in the distance between fluorophores between $\sim$1 and 10 nm (117). By attaching two fluorophores with the appropriate spectral properties to two molecules of interest, association events and relative movements can be observed through single-molecule FRET (smFRET) (Figure 2.1F). By labeling a protein with two fluorophores at known positions within the protein, conformational changes and dynamics within a single molecule can be detected (Figure 2.1G). Since the initial development of the method (118), smFRET has rapidly evolved as an experimental platform to answer fundamental questions in all aspects of cellular biochemistry. For example, by labeling the two heads of a kinesin with a FRET pair, it was shown that the kinesin waits for ATP in a one-head-bound state and makes brief transitions to a two-head-bound intermediate as it walks along the microtubule (119).

Further, smFRET has allowed the direct observation of the conformational dynamics of single amino-acid transporters during substrate transport. (120) (121) Also, smFRET studies revealed the real-time dynamics of the conformational change of the β_2 clamp, the processivity factor in the DNA replication machinery, during loading onto DNA. The distance between the clamp and DNA was monitored by attaching a red Cy5 acceptor fluorophore to β_2 and a green Cy3 donor fluorophore to the DNA. Three successive FRET states were seen, corresponding to closure of the clamp, followed by clamp release from its loader, and diffusion on the DNA (48).

To enable *in vivo* fluorescence imaging, proteins are traditionally genetically fused to a fluorescent protein. The spectral properties and poor photostability of these fluorescent proteins, however, make their use in smFRET very challenging. Therefore, observing smFRET in living cells requires new labeling, internalization and imaging strategies. Significant progress in all these areas has been made in the last decade (122). Fluorescently labeled DNA was internalized in living *E. coli* cells using heat shock (123). By electroporating a large fragment of DNA polymerase I (Klenow fragment, KF), doubly labeled on the fingers and thumb

domains, FRET was measured between internalized, immobile KF molecules. This study shows that the distance between the two domains is preserved in live cells (124).

2.3.4 cryo-Electron Microscopy (cryo-EM)

Perhaps the most rapidly developing single-molecule imaging technique is cryo-Electron Microscopy (cryo-EM). In cryo-EM, rapid freezing techniques (vitrification) provide immobilization of biological samples embedded in amorphous ice, preserving the structure of the samples in their native state. Using electron microscopy, these biological structures can be resolved down to the atomic level. The ability to obtain near-atomic resolution structures using cryo-EM was initially shown almost three decades ago (125). By now, cryo-EM is a firmly established tool to gain structural information on both purified and cellular systems. Recent developments in both sample preparation and detection techniques have given access to resolutions as high as 2.2 Å for proteins as small as $\sim$100 kDa. (126) (127) (128) 8–9 Å resolution structures of four different states of kinesins bound to microtubules allowed precise docking of a kinesin crystal structure into the map. With this information, structural rearrangements that occur upon binding of the kinesin motor domain to the microtubules could be identified (129). The structure of the Saccharomyces cerevisiae helicase, CMG, was determined by cryo-EM at a resolution of 3.7–4.8 Å, hinting towards a new unwinding mechanism. In this mechanism, two domains of the helicase move in a pumpjack-like motion to translocate on DNA (130). 8 Å resolution structures of DNA-bound and DNA-free states of the *E. coli* polymerase complex revealed previously unknown interactions, thereby shedding light on different operational modes of the polymerase (131).

Cryo-EM and fluorescence microscopy are now being combined into Correlative Light-Electron Microscopy (CLEM) (132). This combination of techniques uses fluorescence microscopy to guide the search for specific features and to locate areas worth recording and examining by cryo-EM. Fluorescence imaging can furthermore provide valuable information about local variations in ice thickness, ice crystal contamination or other defects that could affect cryo-EM data quality. In live CLEM, proteins in a living cell are first observed using FM, followed by the observation of cellular structures, such as organelles or membranes, using cryo-EM in the same cell (133). With the combination of these two techniques, dynamic events can be observed in specific cellular structures. This potentially makes live CLEM a powerful method to provide functional and structural understanding of dynamic and complex events, such as nuclear envelope formation (134).

Summarizing, single-molecule fluorescence imaging methods and fluorescence tagging strategies have matured to the point where they can almost routinely be used to visualize biological processes, often in real time. Methods that allow the detection of individual molecules in high-concentration, crowded

environments, combined with advances in specific and selective fluorescent labeling, pave the way to a precise interrogation of molecular processes inside living cells. Combined with the advent of cryo-EM methods, in particular those that visualize cellular structures, we are now able to visualize the dynamics of a single proteins inside a living cell with access to the structural properties of its immediate environment. These methods will enable the field to study more and more complex systems in increasingly physiologically relevant environments.

2.4 Two's company, three's a crowd: multi-protein complexes in crowded environments

All molecular processes that support cellular activity arise from an intricate network of macromolecular interactions that take place in complex, crowded environments. It is therefore of fundamental importance to decipher this 'molecular sociology' (56, 135) ideally by direct visualization. The great advances that have been made in single-molecule techniques are emphasizing a view of dynamic multi-protein systems that is not linear and deterministic, but highly stochastic (55). In this review, we compared *in vitro* and *in vivo* experiments on cytoskeletal motors. It is clear that increasing the complexity of the system, for example by having multiple kinesins and dyneins acting on the same cargo, changes their dynamics. We have also described the dynamic behavior of the DNA replication machinery (replisome) during DNA synthesis. The composition of the replisome has previously been shown to be very stable and highly resistant to dilution (39, 49, 136). Single-molecule studies on the bacteriophage T7 and *E. coli* replisomes demonstrated that the composition of the replisome is in fact highly dynamic when operating in an environment with replisomal components present in solution, with proteins binding and unbinding extremely rapidly (50, 137). This suggests a mechanism in which, in a low-concentration condition, a protein remains stably bound to a complex, while being exchanged rapidly in the presence of competing protein at high concentration. Such a perhaps counter-intuitive concentration-dependent dissociative mechanism has recently also been reported for replication protein A (RPA) in *S. cerevisiae* (109). Using DNA curtains and fluorescently labeled RPA, it was shown that RPA remains bound to ssDNA for long periods of time when free protein is absent from solution. In contrast, RPA rapidly dissociates from ssDNA when free RPA or free SSB (*E. coli* Single-Stranded DNA Binding protein) is present in solution allowing rapid exchange between the free and bound states (Figure 2.3G). Further, in a study on the binding and unbinding kinetics of DNA transcription regulators in living *E. coli* cells, the kinetics of dissociation from chromosomal recognition sites was shown to be concentration-dependent (138).

The apparent paradox between stability under high dilution and plasticity at high concentrations can be rationalized through a network of many weak interactions (Fig. 4). Under dilute conditions, stochastic, transient disruptions of any one of the interactions within a protein complex will not result in dissociation of the

protein, as it is held to the complex via the other bonds, and the interaction would be rapidly reform (Figure 2.4A). Under more physiologically relevant protein concentrations, however, a protein can bind at a transiently vacated binding site and consequently compete out the original protein (Figure 2.4 B) (139). This phenomenon obeys fundamental chemical and thermodynamic principles and can be mathematically described (140, 141). This multi-site exchange mechanism would allow components of multi-protein complexes to be easily replaced. In the case of the replisome, for example, this mechanism may represent a pathway through which a defective polymerase can easily be replaced, thereby insuring replication with a high fidelity. Furthermore this concentration-dependent exchange could provide easy access to other potential binding partners, like repair polymerases (142). The upregulation of these repair polymerases will increase their copy number and stimulate the dissociation of Pol III* through the multi-site exchange mechanism, thereby guaranteeing fast DNA repair.

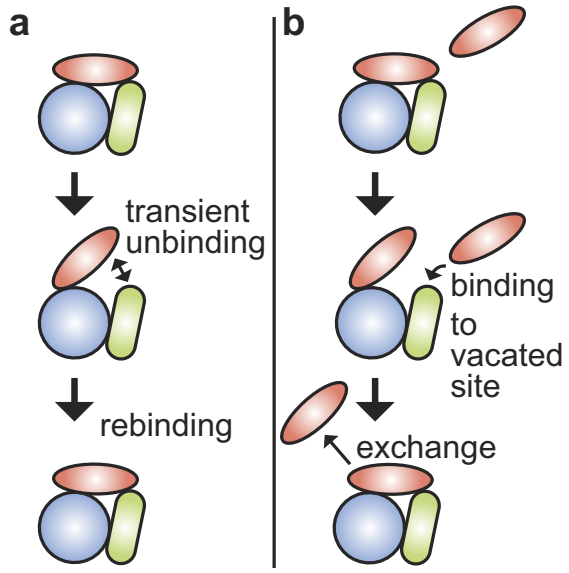


Figure 2.4: Stability versus. plasticity. (a) Under dilute conditions, transient disruption of any one of the weak interactions holding a complex together would be followed by its rapid re-formation, preventing complete dissociation of the protein from the complex. This rapid microscopic re-association would allow a protein to remain stably bound to the complex. (b) If, however, there are competing proteins in close proximity to the complex, one of these can bind at a transiently vacated binding site and consequently be at a sufficiently high local concentration to compete out the original protein.

2.5 Outlook

Single-molecule tools have enabled experimental access to the dynamic behavior of complex biomolecular systems under physiologically relevant conditions. An important next direction is to further develop the single-molecule methods to study larger, more complex systems. The *in vitro* use of force- and fluorescence-based tools described in this review has matured to a point where the complexity of the systems under study seems without limits. Single-molecule studies of complex biochemical systems have already significantly changed our view of the dynamic behavior of molecular systems. The role of stochastic processes in how biological macromolecules move and interact with each other has significant impact on how biochemical processes are controlled. Instead of deterministic pathways, multi-protein complexes seem to perform their tasks by choosing from a multitude of pathways, each made possible by the constellation of weak and strong interactions that hold such a complex together. Applications of these tools within cells

are still comparatively limited, however, in their ability to monitor structural and functional properties in real time at the single-molecule level. Further development of these tools and new labeling approaches are needed to further elucidate the molecular gymnastics of these complexes *in vivo* and bridge the gap between *in vitro* studies and observations inside living cells.

3 | Single-molecule imaging at high fluorophore concentrations by Local Activation of Dye

Hylkje J. Geertsema, Aartje C. Schulte, **Lisanne M. Spenkelink**, William J. McGrath, Seamus R. Morrone, Jungsan Sohn, Walter F. Mangel, Andrew Robinson, Antoine M. van Oijen.

Published in *Biophysical Journal*, 17 Feb 2015;**108** (4):949–56

Single-molecule fluorescence microscopy is a powerful tool for observing biomolecular interactions with high spatial and temporal resolution. Detecting fluorescent signals from individual labeled proteins above high levels of background fluorescence remains challenging, however. For this reason, the concentrations of labeled proteins in *in vitro* assays are often kept low compared to their *in vivo* concentrations. Here, we present a new fluorescence imaging technique by which single fluorescent molecules can be observed in real time at high, physiologically relevant concentrations. The technique requires a protein and its macromolecular substrate to be labeled each with a different fluorophore. Making use of short-distance energy-transfer mechanisms, only the fluorescence from those proteins that bind to their substrate is activated. This approach is demonstrated by labeling a DNA substrate with an intercalating stain, exciting the stain, and using energy transfer from the stain to activate the fluorescence of only those labeled DNA-binding proteins bound to the DNA. Such an experimental design allowed us to observe the sequence-independent interaction of Cy5-labeled interferon-inducible protein 16 with DNA and the sliding via one-dimensional diffusion of Cy5-labeled adenovirus protease on DNA in the presence of a background of hundreds of nanomolar Cy5 fluorophore.

I developed the protocol for the chemical caging of fluorophores and subsequent uncaging using UV light

3.1 Introduction

Recent developments in single-molecule fluorescence microscopy have allowed remarkable insight into the dynamic properties of biomolecular processes. The high spatial and temporal resolution of fluorescence microscopy has enabled the visualization of intermediates and time- dependent pathways in biochemical reactions that were difficult or impossible to extract from experiments at the ensemble-averaged level. However, one of the important technical challenges in single-molecule fluorescence imaging is the visualization of individual fluorescently labeled molecules at high concentrations. Using conventional diffraction-limited optics, the fluorescence of individual molecules can only be resolved if the molecules are farther apart than the diffraction limit, $\sim$ 250 nm in the lateral and $\sim$ 500 nm in the axial direction. As a consequence, the highest concentration at which single fluorescently labeled molecules can still be resolved at a sufficiently high signal/background ratio is on the order of 10–100 nM. This concentration limit reduces the applicability of single-molecule fluorescence imaging to the study of biomolecular interactions with dissociation constants in the nanomolar range or tighter (105). A common and straightforward strategy to circumvent this limitation is to use a partially labeled population of the molecules of interest and supplement with a high concentration of unlabeled molecules. However, when complicated pathways that involve many binding partners or rare molecular transitions are studied, there is a need for approaches that allow the visualization of all events instead of merely a small fraction.

Several recent experimental approaches in single-molecule fluorescence imaging have overcome this concentration limit (143) by confinement of the molecules (144), reduction of the fluorescence excitation volume (145), or temporal separation of fluorescent signals (112). Confinement of labeled molecules in a closed volume, considerably smaller than the diffraction limit, enables the detection of single molecules at concentrations much higher than the fluorescence concentration limit. As an example, trapping of proteins inside nanovesicles with a volume of $\sim 5 \times 10^{19}$ L allows single molecules to be visualized at an effective protein concentration of $\sim$3 μM (146). An alternative is to reduce the excitation volume itself to below the diffraction limit. For example, zero-mode waveguides are widely employed nanophotonic devices that enable effective excitation of a volume much smaller than the diffraction limit and support single-molecule detection at fluorescence concentrations of up to 10 μM (145). Another emerging approach for visualizing biomolecular interactions at high concentrations leverages the stability of biomolecular complex formation. In the technique known as photoactivation, diffusion, and excitation (PhADE), photoactivatable fluorescent proteins in the detection volume are activated and subsequently imaged after first allowing the activated proteins that did not bind the surface-immobilized substrate to diffuse away. This approach allowed the visualization of micrometer-scale movement of single molecules on replicating DNA templates at concentrations up to 2 μM (112). Although all of the above-mentioned techniques have dramatically pushed the limits of

single-molecule imaging at high concentrations, experimental difficulties and temporal limitations have hindered the development of a generally applicable method of observing single-molecule dynamics at high fluorescent background concentrations. The trapping of molecules in nanovesicles or nanophotonic devices allows the fluorescence of single molecules amid a high background concentration to be detected at high temporal resolution, but the spatial confinement precludes the visualization of large-scale movements of the fluorescently labeled species. Whereas the PhADE technique removes these spatial constraints, it introduces limits on time resolution as a result of the time needed for activated but noninteracting proteins to diffuse away from the observation volume.

Here, we present a fluorescence imaging technique that is based on the incorporation of an activator molecule into the binding target of a molecule, and we use energy transfer to activate the fluorescence of only those molecules that bind the target. The free molecules remain in their dark state, and only the molecules involved in complex formation are switched on and their fluorescence observed. We demonstrate our method by visualizing large-scale motions of individual DNA-binding proteins at protein concentrations exceeding the previous concentration limit by an order of magnitude.

3.2 Materials and methods

3.2.1 Dyes and proteins

Recombinant interferon-inducible protein 16 (IFI16) was synthesized in *Escherichia coli*, purified, and subsequently labeled with the Cy5 dye by maleimide chemistry, as described previously (147). Cy5-IFI16 was used at a concentration of 1 nM in combination with the M13 DNA template and at 30 nM in the experiments with λ-DNA.

The gene for arginine vasopressin (AVP) was expressed in *E. coli* and the resultant protein was purified as described previously (148, 149). AVP concentrations were determined using a calculated extinction coefficient (150) of 26,510 $M^{-1}cm^{-1}$ at 280 nm. The 11-amino acid peptide pVIc (GVQSLKRRRCF) was purchased from Invitrogen (Carlsbad, CA), and its concentration was determined by titration of the cysteine residue with Ellman's reagent (151) using an extinction coefficient of 14,150 $M^{-1}cm^{-1}$ at 412 nm for released thionitrobenzoate. Octylglucoside was obtained from Fischer Scientific (Faden, NJ) and endoproteinase Glu-C from Sigma (St. Louis, MO). Disulfide-linked AVP-pVIc complexes were prepared by overnight incubation at 4°C of 75 μM AVP and 75 μM pVIc in 20 mM Tris-HCl (pH 8.0), 250 mM NaCl, 0.1 mM EDTA, and 20 mM β-mercaptoethanol. Initially, during overnight incubation, the β-mercapto- ethanol reduces most oxidized cysteines in AVP and in pVIc. Then, the β-mercaptoethanol evaporates, thereby allowing Cys-104 of AVP and Cys-10' of pVIc to undergo oxidative condensation (152, 153).

Disulfide-linked AVP-pVIc complexes, 75 μM, were labeled in 25 mM HEPES (pH 7.0), 50 mM NaCl, and 20 mM ethanol by the addition of Cy5 maleimide (GE Healthcare, Piscataway, NJ) to 225 muM. Labeling reactions were incubated at room temperature in the dark for 2.5 h. Excess reagents were removed from the labeled sample by passage through Bio-Spin 6 chromatography columns (Bio-Rad, Hercules, CA) equilibrated in the labeling buffer. The degree of labeling was determined using molar extinction coefficients of 26,510 $M^{-1}cm^{-1}$ for AVP at 280 nm and 250,000 $M^{-1}cm^{-1}$ for Cy5 at 649 nm, with a correction factor at 280 nm of 0.05. The ratio of labeled AVP-pVIc to total AVP-pVIc was determined to be $\sim$0.8. The labeled materials were characterized by matrix-assisted laser desorption ionization time of flight mass spectrometry.

Specific enzymatic digestions followed by matrix-assisted laser desorption ionization time-of-flight mass spectrometry were used to locate cys- teinyl–Cy5 conjugates in AVP-pVIc complexes. Labeled AVP-pVIc complexes, 1.2 μg, were digested by incubation with 0.01 μg endoproteinase Glu-C or trypsin at 21°C in 25 mM Tris-HCl (pH 7.5). At 1, 2, 4, and 22 h, 0.5 μL of each reaction was removed and added to 4.5 μL of a saturated matrix solution (α-cyano-4-hydroxycinnamic acid) in 50$\%$ acetonitrile and 0.1$\%$ trifluoroacetic acid. The matrix-analyte solution was then immediately spotted onto a 100-well stainless steel sample plate. The sample plate was calibrated using Applied Biosystems peptide calibration mixtures 1 and 2. Mass spectrometric characterization was carried out on a Voyager-DE Biospectrometry Workstation (Applied Biosystems, Foster City, CA). The m/z peak list generated for each chromatogram was analyzed by the FindPept Tool (154). The Cy5 modification was entered as a posttranslational modification. AVP-pVIc complexes were found to be labeled at Cys-199 (data not shown).

Cy5 monoreactive NHS ester (GE Healthcare) was dissolved in dimethyl sulfoxide and stored at –20°C. For the high Cy5 background experiments, 270 nM of this Cy5 was caged with 1 mg/mL $NaBH_4$ and added to the single-molecule imaging buffer.

3.2.2 DNA construct

A biotinylated 60-basepair (bp) duplex oligo DNA with a 3′ amine-end modification was purchased from Integrated DNA Technology (Coralville, IA). The amine group of the oligo was covalently linked to Cy5 monoreactive N-hydroxysuccinimide (NHS) ester (GE Healthcare) by adding a 125-fold excess of Cy5 to the oligo in reaction phosphate-buffered saline (pH 7.5) at room temperature. The labeling reaction was performed for 1 h. Unreacted fluorophores were removed by immediately running the reaction products over a NAP5 column (GE Healthcare). The degree of labeling, determined by absorbance spectroscopy, was 0.8 fluorophores/oligo.

Single-stranded M13mp18 (New England Biolabs, Ipswich, MA) was bi-

otinylated by annealing a complementary biotinylated oligo to the M13 template. Subsequently, the primed M13 was filled in by adding T7 DNA polymerase (New England Biolabs), dNTPs, and a replication buffer containing $MgCl_2$. Replication proteins were removed from the filled-in DNA products by phenol/chloroform extraction and stored in 10 mM Tris-HCl and 1 mM EDTA (TE) buffer (pH 8.0) (155).

λ-DNA substrates were constructed by standard annealing reactions. The linearized DNA had 12-nucleotide single-stranded overhangs at each end. To the 5′ end of the λ-DNA, a biotinylated oligo was annealed to allow surface attachment to the functionalized glass coverslip (155).

3.2.3 Experimental setup

Single-molecule measurements were performed on an Olympus IX-71 microscope equipped with a 100x total internal reflection fluorescence (TIRF) objective (UApoN, NA 1.49 (oil), Olympus, Center Valley, PA). The sample was continuously excited with a 154 W·cm^{-2} 643 nm laser (Coherent, Santa Clara, CA) and pulsed photoactivation was applied with a 0.49 W·cm^{-2} 532 nm laser (Coherent) and a 62 W·cm^{-2} 404-nm laser (Cube) controlled by a home-built shutter program. Images were captured with an EMCCD camera (Hamamatsu, Hamamatsu City, Japan) using Meta Vue imaging software (Molecular Devices, Sunnyvale, CA) with a typical frame rate of 5 frames/s for the Cy5-oligo DNA and Cy5-IFI16 experiments and 32 frames/s for the Cy5-pVIc-AVP measurements. The gray scale is rescaled in all images to provide best contrast; thus, intensity values do not directly reflect numbers of photons.

All experiments were performed in home-built flow cells. PEG-biotin functionalized coverslips were treated with a streptavidin solution to enable DNA template binding (156). A polydimethylsiloxane flow cell with 0.5-mm-wide channels was adhered to the top of the glass coverslip and inlet and outlet tubes were inserted into the polydimethylsiloxane. After thoroughly washing the flow cell, the DNA was flushed through and the reaction buffer was subsequently flowed through at 10 μL/min. When the reaction buffers entered the flow cells, fluorescence imaging was started.

3.2.4 Buffers for single-molecule measurements

Local activation of dye (LADye) requires the predarkening of the Cy5 fluorophore, which was achieved by caging with 1 mg/mL $NaBH_4$ in 20 mM Tris at pH 7.5, 2 mM EDTA, and 50 mM NaCl before the fluorescent imaging experiments. The buffers used in the single-molecule imaging experiments were designed to suppress blinking and reduce photobleaching rates in caged Cy5 and Cy5-labeled proteins. The experiments on Cy5-oligo-DNA and Cy5-IFI16 switching were performed in a 20 mM Tris at pH 7.5, 2 mM EDTA, and 50 mM NaCl buffer, and the Cy5-pVIc-AVP experiments were peformed in 10 mM Hepes at pH 7.0, 2

mM NaCl, 5% glycerol, 20 mM EtOH, and 50 μM EDTA buffer. To increase Cy5 photostability and facilitate switching, all buffers contained 0.45 mg/mL glucose oxidase, 21 μg/mL catalase, 10% (w/v) glucose, 1 mM Trolox (157), and 143 mM β-mercaptoethanol (BME). In addition, 50 nM Sytox Orange (Molecular Probes, Eugene, OR) was added to the reaction buffers to stain the DNA to enable Cy5 switching.

3.3 Results

Our approach to visualizing bimolecular interactions at high fluorophore concentrations is based on generating fluorescence from only those molecules that have formed a bimolecular interaction; no other fluorophores will fluoresce. As a switchable label fluorophore we use Cy5, a red carbocyanine dye that has been shown to act as an efficient reversible single-molecule photoswitch supporting hundreds of cycles of switching between a dark and a bright state (158, 159). Our method is based on the selective activation of only the fluorescence of those few molecules that successfully associate to a target substrate, with the remainder of the fluorophores left in their dark state. As such, at the start of an imaging experiment, all fluorophores need to be placed in a dark, fluorescently inactive state. Darkening of Cy5 can be achieved chemically upon interaction with $NaBH_4$ (114), phosphine tris(2-carboxyethyl)phosphine (160), or primary thiols combined with red-light illumination (161). These chemicals presumably bind to the polymethine bridge of Cy5, resulting in disruption of the conjugated π-system and a drastic blue shift of the Cy5 fluoresence (160). The bright red-emitting state of Cy5 is recovered upon dissociation of the chemical darkening moieties, an event that can be triggered by irradiation at shorter wavelengths (158) or by excitation of a nearby secondary fluorophore whose emission spectrally overlaps with the Cy5 absorption (159, 162). Excitation of such a secondary fluorophore offers the opportunity to specifically recover the bright state of only those Cy5 molecules in close proximity to that secondary fluorophore. We present here the use of a DNA-intercalating stain to locally activate only those Cy5 fluorophores that are close to a DNA template while leaving the background population of Cy5-labeled proteins in their dark state (Figure 3.1a). Such a local activation of dye (LADye) allows the selective activation of fluorescence of only those labeled proteins that are bound to a DNA substrate molecule. The short activation distance of $\sim$1 nm (159) results in very small effective excitation volumes. This small excitation volume makes it possible to perform single-molecule experiments on DNA-interacting proteins in the presence of high background concentrations of labeled proteins and to still follow the large-scale motions of the protein on DNA.

As a proof of principle, we labeled a 60-bp double-stranded (ds) DNA oligonucleotide with the DNA-intercalating stain Sytox Orange and used its fluorescence emission to photoactivate a dark Cy5 fluorophore that is covalently coupled to the same DNA (Figure 3.1b). Darkening of the Cy5 fluorophore with 1 mg/mL $NaBH_4$ was found to be 69% efficient (Figure 3.4 in the Supplementary information). The

ability of the Cy5 molecule to switch on was studied both by direct activation upon irradiation with 404-nm light and by indirect, proximal activation upon excitation of the Sytox Orange with 532-nm light. Throughout the entire experiment, the sample was continuously excited with 643-nm laser light, resulting in the emission of red fluorescence whenever the Cy5 was present in the photoactivated state. This same wavelength in the presence of BME eventually resulted in darkening of Cy5. Repeats of this sequence made it possible to reactivate the Cy5 and prolong its imaging. Both 404-nm (62 W·cm^{-2}) and 532-nm (0.49 W·cm^{-2}) laser pulses, given every 8 s for a duration of 0.4 s, resulted in similar recoveries of the red Cy5 fluorescence, indicating that photoactivation of Cy5 is equally efficient for direct activation by 404-nm irradiation and for indirect activation by Sytox Orange (Figure 3.1b). Analysis of single-molecule fluorescence trajectories indicates efficient activation of Cy5 in 34 of 56 404-nm pulses and in 36 of 57 532-nm pulses, indicating a switching efficiency of 61% and 63%, respectively (see Figure 3.1c, black trace). As a control, the experiment was run in the absence of Sytox Orange (Figure 3.1c, red trace), and no Cy5 switching was observed for 532 pulses, though 404-nm pulses remained effective in Cy5 activation (45% per single 404-nm pulse). Observation of 11 different Cy5-labeled oligos for 264 532-nm pulses in total did not show any fluorescence recovery of the Cy5 fluorophores. Assays in the absence of thiols (BME) completely eliminated Cy5 switching for both 404-nm and 532-nm pulses (Figure 3.1c, blue trace).

Next, we investigated whether excitation of an intercalating stain bound to DNA could result in the activation of the fluorescence of a Cy5-labeled, DNA-bound protein. As a model system, we studied the DNA-dependent fluorescent activation of Cy5-labeled IFI16, an 82-kDa human protein that acts as a cytosolic viral DNA sensor. The physiological role of IFI16 includes nonspecific binding to cytosolic foreign dsDNA, which triggers an innate immune response that activates cell death (163). We labeled IFI16 with $NaBH_4$-darkened Cy5 and allowed it to bind nonspecifically to a circular dsDNA template (based on phage M13 DNA, with a circumference of 7.3 kb) that had been coupled to a glass surface and stained with Sytox Orange (Figure 3.2a). Applying pulsed excitation of the DNA stain with 532-nm light while continuously irradiating the sample with 643-nm light resulted in fluorescence activation and visualization of the Cy5-labeled proteins, as can be seen by the appearance of fluorescent spots. Repeated photoactivation of one or more Cy5-IFI16 proteins per single DNA template is shown in Figure 3.2b. Investigation of the fluorescence intensity of the black trace in Figure 3.2b uncovered a number of discrete intensity levels, with each level an integer multiple of a constant intensity (5900 $\pm$ 1300 counts), suggesting the binding of an integer number of individual molecules to the DNA (Figure 3.2c). Next, we imaged individual Cy5-IFI16 proteins on flow-stretched λ-phage DNA to visualize the spatial distribution and movement of the proteins on longer DNA substrates (164). Continuous illumination of the Cy5-IFI16 by 643-nm light and pulsed excitation of the DNA stain by 532-nm laser light allowed the activation and visualization of individual IFI16 proteins bound to the DNA (Figure 3.2d). Upon continuous 532-nm excitation, the total number of fluorescently labeled

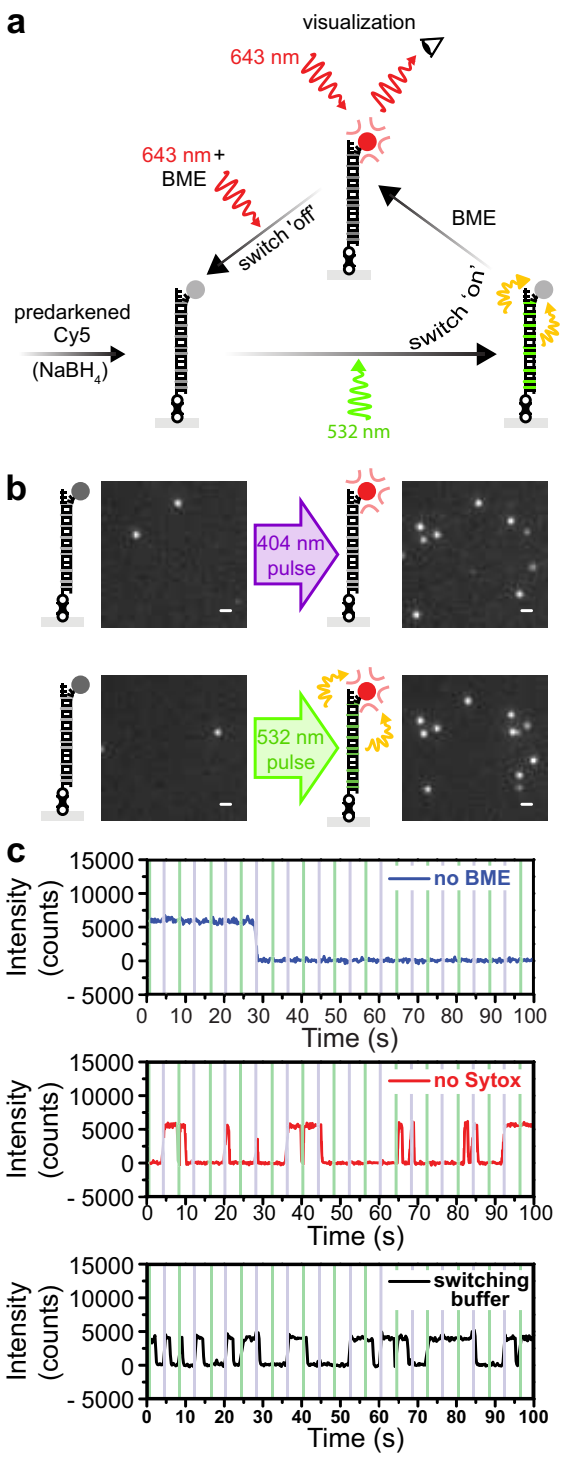
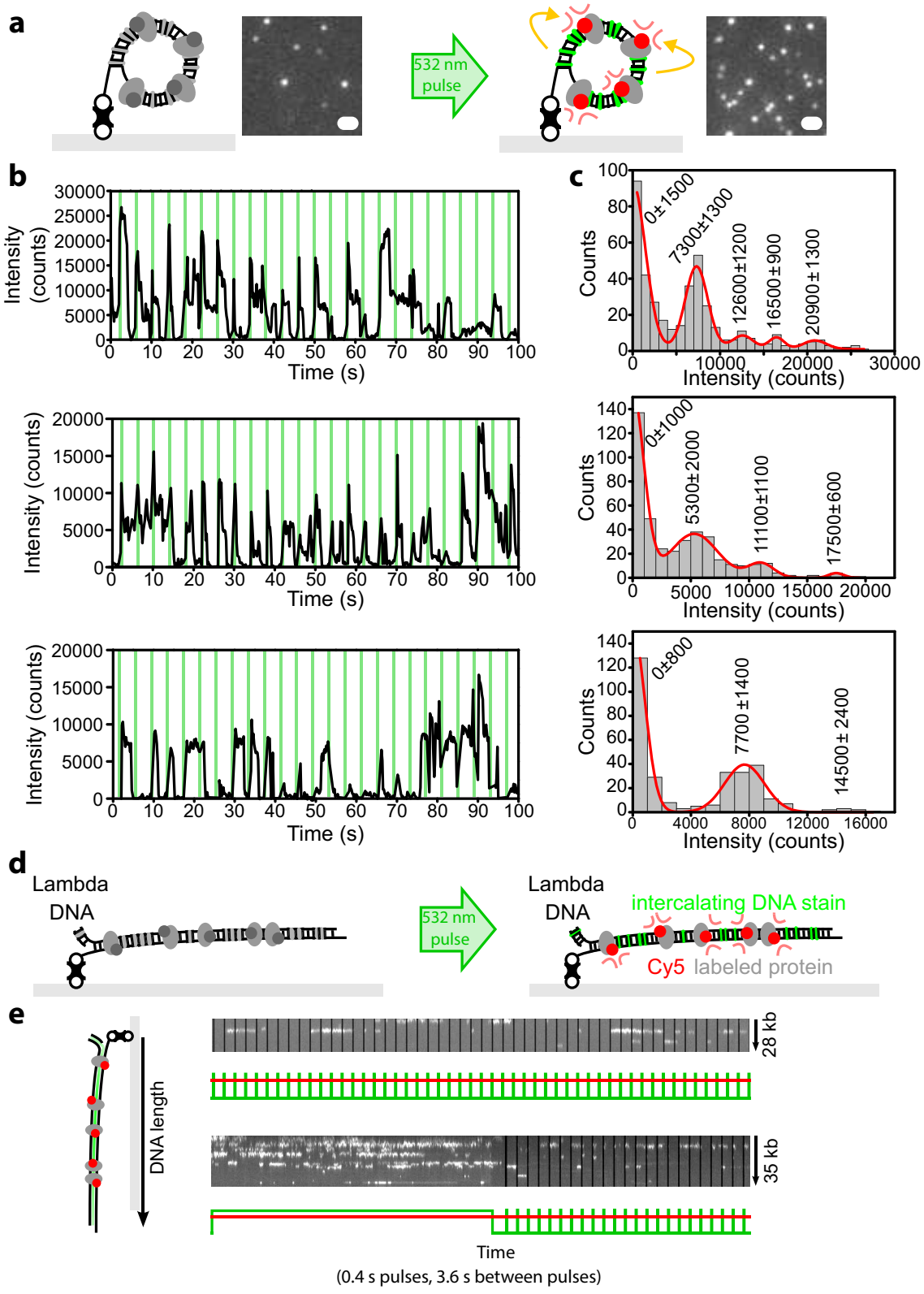


Figure 3.1: Fluorescence switching of Cy5 bound to stained DNA. (a) Schematic representation of local activation of dye (LADye) Before the experiment, the entire population of Cy5 molecules is darkened. Green excitation (532 nm) allows the DNA-bound intercalating stain to fluoresce, which in turn results in photoactivation of the DNA-proximal Cy5 molecules with BME present in solution. Subsequently, Cy5 is visualized using red laser light (643 nm), which eventually brings the Cy5 molecules back to the dark state when BME is present. (b) Single-molecule Cy5 fluorescence recovering upon a 0.4 s pulse with either 404-nm light, causing direct activation of the Cy5 fluorescence (not visualized in a), or 532-nm laser light, leading to selective activation of only those Cy5 molecules close to the DNA. Shown is the Cy5 fluorescence before (left) and right after (right) the laser pulse (Cy5 fluorescence is excited at 643 nm). Scale bar, 1 μm. (c) Reversible activation of Cy5 upon alternating 404-nm (purple bars) and 532-nm laser pulses (green bars). Black trace (lower) represents a Cy5 fluorescence trajectory in the regular reaction buffer (see paragraph 3.2: Materials and methods), whereas the red (middle) and blue (upper) traces represent Cy5 fluorescence in a reaction buffer lacking Sytox Orange and BME, respectively. The lack of photoactivation by 532-nm excitation in the absence of Sytox Orange or BME confirms the role of Sytox Orange as a local activator.

proteins observed on individual DNA molecules was increased fivefold over the number of proteins activated with pulsed 532-nm excitation (Figure 3.2e). This increase in the detected number of fluorescent proteins demonstrates that the illumination conditions (intensity and pulse duration) can be optimized to maximize the probability of photoactivation while keeping fluorescence background to a minimum.

Next, we set out to determine whether LADye can be used to visualize single labeled proteins in the presence of a high concentration of fluorophore. As a proof of principle, we imaged individual Cy5-labeled DNA-binding proteins interacting with DNA in the presence of several hundreds of nanomolar free Cy5 in solution. A key requirement to visualizing single fluorophores amid high backgrounds is the efficient initial darkening of the entire population of fluorophores in solution. As such, only those molecules that are switched to the bright state by energy transfer from the DNA intercalating stain will contribute to the observed fluorescence. Darkening of a solution of Cy5 (114) by $NaBH_4$-mediated caging resulted in a fluorescence reduction of 97% in comparison to conventional imaging of uncaged Cy5 at the same concentration. This efficient darkening allowed for the visualization of the fluorescence of single surface-bound Cy5 molecules in the presence of 270 nM Cy5 in solution (Figure 3.3a, right), conditions that are not compatible with conventional single-molecule fluorescence imaging (Figure 3.3a, left).

As a model system to visualize proteins moving along DNA amid a high background of Cy5, we chose the adenovirus protease AVP, which is a protein able to diffuse one-dimensionally along dsDNA after association with the 11-amino-acid peptide pVIc (165). The pVIc-AVP complex binds to DNA without any sequence specificity and performs a one-dimensional random walk along the duplex to locate and process DNA-bound protein substrates (165). Using darkened Cy5-labeled pVIc-AVP and pulsed 532-nm photoactivation, we specifically visualized the Cy5-pVIc-AVP bound to DNA, even in the presence of 270 nM Cy5 in solution (Figure 3.3b). The kymograph shows the spatiotemporal behavior of the protein complex as it moves along the DNA. The time-dependent position of four different Cy5-pVIc-AVP proteins along the DNA was tracked by determining the center of a two-dimensional Gaussian function fitted to the fluorescent spot at every frame. Subsequently, the diffusion constant of the Cy5-pVIc-AVPs was calculated by fitting the mean-square displacement of the Cy5-pVIc-AVPs over time, and their average was found to be (2.3 $\pm$ 0.2) $\times$ 107 bp^2/s (Figure 3.3c). This diffusion constant agrees well with a previously obtained pVIc-AVP diffusion constant of (2.1 $\pm$ 0.2) $\times$ 107 bp^2/s (165), demonstrating the applicability of LADye to investigation of binding kinetics and activity of individual molecules in the presence of high fluorescent background concentrations.


a
532 nm pulse
b
Intensity (counts)
Time (s)
c
Counts
Intensity (counts)
0±1500
7300±1300
12600±1200
16500±900
20900±1300
0±1000
5300±2000
11100±1100
17500±600
0±800
7700 ±1400
14500± 2400
d
Lambda DNA
532 nm pulse
intercalating DNA stain
Cy5 labeled protein
e
DNA length
28 kb
35 kb
Time
(0.4 s pulses, 3.6 s between pulses)

3.4 Discussion

Single-molecule fluorescence microscopy has provided remarkable insights on how biological macromolecules interact. However, the application of this technique to systems with labeled molecules at their concentrations inside the cell (105) has been made difficult due to high background fluorescence. Here, we describe a new approach to circumvent this concentration limit by selectively visualizing only the subpopulation of labeled proteins that have productively bound to a substrate and leaving the unbound population in a nonfluorescent, dark state. Chemically darkened Cy5 fluorophores coupled to proteins of interest were triggered to recover their bright state by excitation of the spectrally overlapping DNA intercalating stain Sytox Orange, resulting in the activation of only those Cy5-labeled proteins proximal to a DNA substrate. This local activation of Cy5 dyes allowed us to observe in real time the interaction of single proteins with DNA at concentrations of several hundreds of nanomolar of fluorophores in solution.

TIRF microscopy is commonly used in single-molecule fluorescence imaging to limit the penetration depth of the excitation light to $\sim$100 nm above the coverslip surface to reduce fluorescence contributions from the rest of the solution. However, long DNA molecules that are surface-anchored on one end and stretched by flow are tilted away from the surface sufficiently so that for most of their length the DNA is >100 nm away from the coverslip. To allow visualization of the full length of these molecules, instead of using a TIRF-based excitation scheme, we had to apply a near-TIRF geometry with the laser light entering the flow cell at a very steep angle of only a few degrees off with respect to the plane of the surface. Thereby, the detection volume has diffraction-limited dimensions of 250 $\times$ 250 $\times$ 500 nm^3 under the experimental conditions used here, fivefold larger than the detection volume achieved in TIRF imaging. Despite such a significantly larger detection volume, we demonstrate here the visualization of single molecules bound to DNA in the presence of 270 nM fluorophores in solution. This concentration implies the ability of our LADye technique to resolve single labeled molecules even when background concentrations are as high as 1 μM in a TIRF scheme with a 6.25 $\times$ 10^{-6} nm^3 detection volume.

Figure 3.2 *(preceding page)*: Photoactivation of Cy5-labeled IFI16 proteins bound to DNA. $(\mathbf{a})$ Cy5-coupled IFI16 proteins (1 nM) in the dark state bind the circular double-stranded M13 DNA template (left) and are switched to a bright state upon a 0.4-s 532-nm pulse (right). Scale bar, 1 μm. $(\mathbf{b})$ Fluorescence intensity over time of Cy5-labeled IFI16 on individual M13 dsDNA templates. Every 4 s, a 0.2-s green pulse is applied to recover the bright state of the Cy5 fluorophores (vertical green bars). $(\mathbf{c})$ Histograms of the Cy5 fluorescence intensity traces in b. Multiple Gaussian fits (red) revealed different fluorescence intensity levels corresponding to IFI16 proteins binding to the M13 DNA. $(\mathbf{d})$ Schematic of local activation of Cy5-labeled IFI16 proteins on a λ-phage DNA template upon application of a 532-nm laser pulse. $(\mathbf{e})$ Kymographs depicting the positions of Cy5-labeled IFI16 (20 nM) as a function of time on individual flow-stretched λ-phage DNA templates. Photoactivation was achieved by pulsed excitation of the Sytox Orange (upper) or by continuous excitation followed by pulsed 532-nm illumination (lower).

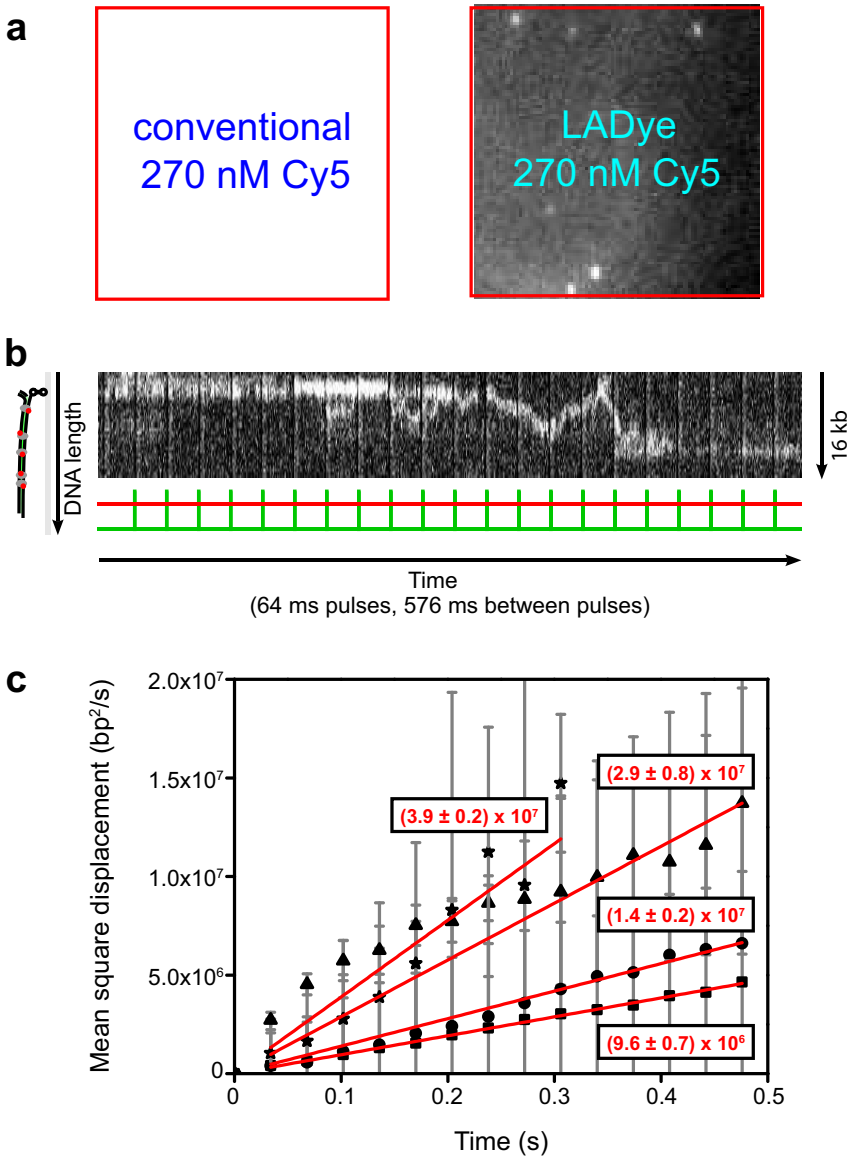


Figure 3.3: Visualization of single Cy5-labeled adenovirus protease (Cy5-pVIc-AVP) molecules sliding along DNA. (a) Caging of Cy5-pVIc-AVP and Cy5 in solution drastically reduces the fluorescence intensity. Conventional imaging of 270 nM Cy5 resulted in saturation of the camera (left), whereas caging of the Cy5 in solution allows single nonspecifically surface-bound molecules to be resolved under the same illumination and acquisition conditions (right). (b) Kymograph of the sliding motion of individual Cy5-pVIc-AVP molecules on λ-phage DNA over time. DNA-bound Cy5-pVIc-AVP was visualized by repeated photoactivation, whereas the 270-nM caged-background Cy5 remained dark. (c) Mean-square displacement of four sliding Cy5-pVIc-AVP proteins over time. The diffusion constants of the labeled pVIc-AVPs (red values) were determined by linear fitting of the data.

We demonstrate here efficient fluorescence activation of a darkened Cy5 molecule coupled to a stained dsDNA molecule. However, switching was found to be less efficient for Cy5-labeled proteins interacting with stained DNA. This observation is consistent with previous reports that the switching efficiency of Cy5 has a distance dependence much shorter and steeper than that reported for Förster resonance energy transfer (159). The physical size of the protein places the Cy5 at a slightly larger distance away from the DNA, thereby reducing the efficiency of fluorescence activation by the DNA-bound intercalating stain. However, the previously demonstrated linear dependence of switching efficiency on intensity and duration of the activation beam (159) provides a readily accessible experimental parameter to optimize the extent of switching in a population of labeled molecules (Figure 3.2e). An experimental concern in the application of fluorescence microscopy is the photoinduced degradation of biomolecules. In particular, excitation of intercalating stain has been shown to induce DNA cleavage through free-radical formation (166). However, at the low intensity of 532-nm illumination required for excitation of the intercalating stain and photo-activation of Cy5 (0.49 W·cm^{-2} at 532 nm), no DNA degradation was detected for tens of minutes.

At this point, the $NaBH_4$ caging efficiency determines the upper limit of experimentally usable fluorescent concentration. The 97% darkening efficiency implies an upper dye concentration of $\sim$1 μM for which individual fluorophores can still be resolved. However, the steep distance dependence of the Cy5 switching efficiency upon Sytox Orange excitation drastically reduces the volume in which Cy5 fluorescence can be activated to a cylinder around the DNA with a radius of $\sim$1–2 nm. Theoretically, this activation volume allows for the observation of single molecules in the presence of up to millimolar concentrations of fluorophores in solution. In comparison, FRET read-out lengths are limited to a radius of $\sim$5 nm (167) implying an upper concentration limit for single-molecule detection of 85 μM.

In principle, LADye offers a generally applicable method to study biomolecular processes including, but not limited to, DNA-based systems. The method requires only a substrate labeled with a switchable fluorophore and an immobilized binding partner coupled to a fluorophore that spectrally overlaps with the substrate's fluorophore. LADye could, for example, be employed to visualize the movement of actin- or microtubule-based motor proteins by staining the filaments with the activator dye. Alternatively, binding of ligands to transmembrane proteins could be visualized by placing activator dyes in the membrane. LADye could also be used to complement switchable FRET (168) studies not only by allowing the observation of subunit interactions within the molecular complex, but also by enabling the visualization of recruitment of subunits from solution. The ability to specifically visualize only those fluorescently labeled molecules that interact with their binding partners while the fluorescent molecules in solution remain dark enables real-time, single-molecule observations of low-affinity biomolecular

interactions under equilibrium conditions that approach the *in vivo* concentration of the reactants.

3.5 Supplementary information

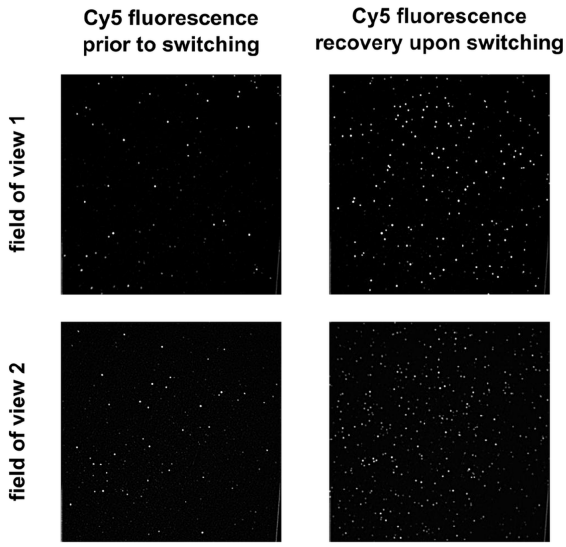


Figure 3.4: Darkening efficiency of Cy5 bound to DNA oligonucleotides. The images on the left show the fluorescence of the Cy5 fluorophores, darkened with 1 mg/mL $NaBH_4$. On the right, the recovery of the Cy5 fluorescence is shown after several activation pulses of 532 and 404 nm. Peaks were detected by fitting 2D Gaussians and resulted in 337 Cy5 fluorophore spots after activation pulses versus 71 spots prior to the experiment for the first field of view and 179 Cy5 spots after activation versus 90 prior to the experiment. Taken together, the darkening of the Cy5 fluorophores bound to the DNA oligonucleotides has been found to be successful for 355 of the 516 Cy5 fluorophores, equal to 69$\%$ darkening efficiency.

4 | Quantification of ligand stoichiometries in liposomal drug delivery systems using single-molecule fluorescence imaging

Lisa Belfiore[†], **Lisanne M. Spenkelink**[†], Marie Ranson, Antoine M. van Oijen, Kara L. Vine.

[†]These authors contributed equally.

Published in *Journal of Controlled Release*, 28 May 2018; **278**:80–86

Despite the longstanding existence of liposome technology in drug delivery applications, there have been no ligand-directed liposome formulations approved for clinical use to date. This lack of translation is due in part to the absence of molecular tools available for the robust quantification of ligand density on the surface of liposomes. Here, we report for the first time the quantification of proteins attached to the surface of small unilamellar liposomes using single-molecule fluorescence imaging. Liposomes were surface-functionalized with fluorescently-labeled human proteins previously validated to target cancer cell surface biomarkers: plasminogen activator inhibitor-2 (PAI-2) and trastuzumab (TZ, Herceptin). Protein-conjugated liposomes were visualized using a custom-built wide-field fluorescence microscope with single-molecule sensitivity. By counting the photobleaching steps of the fluorescently-labeled proteins, we calculated the number of attached proteins per liposome, which was in the range of 1–11 proteins for single-ligand liposomes. Imaging of dual-ligand liposomes revealed stoichiometries of the two attached proteins in accordance with the molar ratios of protein added during preparation. Preparation of PAI-2/TZ dual-ligand liposomes via two different methods revealed that the post-insertion method generated liposomes with a more equal representation of the two differently-sized proteins, demonstrating the ability of this method to control protein densities. We conclude that the single-molecule imaging method presented here is an accurate and reliable quantification tool for determining ligand stoichiometry on the surface of liposomes. This method has the potential to allow for comprehensive characterization of novel ligand-directed liposomes and may improve the translation of these nanotherapies through to the clinic.

L.B. and I contributed equally to this paper.I carried out all single-molecule experiments, analysed the single-molecule data, and was involved in writing the manuscript.

4.1 Introduction

Liposomes have been utilized as delivery systems for drugs and other molecules *in vivo* for several decades (169). In the context of cancer therapy, liposome-based drug formulations have demonstrated distinct advantages over free drug, including the improved solubility of encapsulated drugs, increased *in vivo* circulation time, reduction in systemic toxicity of the drug and increased delivery to the tumor site (170). The superior activity of drug-loaded liposomes relies on a multi-step process involving both passive and active targeting mechanisms. Passive targeting is primarily mediated by the enhanced permeability and retention effect (171). This phenomenon is characterized by the extravasation and retention of small particles into the tumor interstitial space due to highly porous tumor vasculature and poor lymphatic drainage from the tumor site (172). The prolonged retention of liposomes in the vicinity of the tumor increases the local drug concentration, either when drug released from the liposomes is taken up by tumor cells, or when liposomes containing the drug are internalized by tumor cells (173). Passive targeting, therefore, reduces off-target effects by preferentially accumulating drug-loaded liposomes in the vicinity of the tumor while reducing the exposure of normal cells to the cytotoxic drug.

Active targeting is achieved via conjugation of one or more ligands to the liposome surface, with that ligand binding to a target receptor(s) expressed on the tumor cell surface (174). Following liposome extravasation into the tumor interstitial space, subsequent ligand-directed surface binding and internalization (usually via receptor-mediated endocytosis) promotes liposome and drug entry into specific cell types (175). As actively targeted liposome formulations combine both passive and active drug-delivery mechanisms, actively targeted liposomes can show superior drug delivery to non-targeted liposomes (176). Liposomes with one or more targeting moieties that facilitate active uptake into cells are termed ligand-directed liposomes. In the context of cancer therapy, the development of dual-ligand-directed liposomes that can actively target more than one tumor cell subtype and/or stromal cell populations may help overcome therapeutic limitations caused by the intratumoral heterogeneity of cancer (177, 178). Despite extensive research and development of nanoparticle-based therapeutics, all clinically approved liposome formulations are non-ligand-directed, with efficacies relying solely on passive targeting and accumulation (179). A comprehensive list can be found elsewhere (180). Active targeting strategies using liposomes have been extensively explored in the preclinical setting, particularly liposomes targeting tumor-associated receptors, with many reported formulations demonstrating improved efficacy over non-ligand-directed liposomes (181, 182). Given the general movement in the field towards actively targeted nanotherapeutics, the lack of translation of ligand-directed liposome formulations into clinical practice is surprising (183).

The absence of molecular tools for the robust characterization of complex liposomes may be contributing to this deficiency. Specifically, no methodology exists

to quantify the number of ligands covalently bound to the surface of liposomes. Estimation of ligand conjugation is possible based on preparation parameters, but no methods exist to obtain direct information. Direct measurement of total protein in an actively targeted liposome formulation using biochemical assays is challenging due to phospholipid interference in the measurement of very low protein concentrations (184). Further, while such measurements could potentially quantify the total protein in a sample, they cannot provide information about the number of ligands per liposome in a formulation. Flow cytometric methods that detect the insertion of fluorescently labeled micelles into liposomes as a proxy for successful liposome functionalization have been reported, but are indirect and semi-quantitative (185). The lack of quantitative methodology poses a particular challenge for the development of liposomes with more than one surface-bound ligand, since the determination of ligand stoichiometry is important for controlling for batch-to-batch variability in the lab and for clinical production. The absence of rigorous quantification protocols hinders high-quality large-scale manufacturing of ligand-directed liposome formulations, which may introduce regulatory barriers and slow down their introduction to the clinic.

We describe here the use of single-molecule methods to enable the quantitative characterization of ligand-coupled liposomal drug delivery systems. By removing ensemble averaging, single-molecule approaches allow the direct visualization of population distributions and the precise characterization of sub-populations. These methods have already proven to be important biophysical tools to study a wide variety of biological processes (186–188). Single-molecule microscopy remains, however, an underutilized technique in therapeutics development. In this study, we report the quantification of protein attachment to the surface of single and dual ligand-directed liposomes using single-molecule fluorescence microscopy. This method allows the detection and quantification of the density of proteins attached to liposomes, facilitating the characterization and translation of ligand-directed liposomes for targeted cancer therapy, and other, applications.

4.2 Results and discussion

To visualize proteins attached to liposomes, we labeled 45 kDa human recombinant plasminogen activator inhibitor-2 (PAI-2, SerpinB2) with a small red fluorophore (CF647, 0.8 kDa). The degree of labeling (DOL) was determined by visualizing single proteins using Total Internal Reflection Fluorescence (TIRF) microscopy (Figure 4.7). Figure 4.1a shows a typical field of view of individual labeled PAI-2 proteins immobilized on a microscope coverslip. TIRF microscopy allows for the selective excitation of only the fluorescent species on the cover-slip surface and imaging of fluorescence from the surface-immobilized proteins with high contrast and low background. The intensity of the signal of every individual protein can be measured over time (Figure 4.1b, black line). These intensity trajectories show a stepwise decay towards zero, due to the photobleaching of the fluorophores on the protein. The height of a single step corresponds to the intensity of a single fluorophore. Using an unbiased change-point step-fitting

algorithm (52, 189) (Figure 4.1b, red line), we determined the intensity of a single fluorophore (Figure 4.1c). By dividing the total intensity per protein by this single-fluorophore intensity we found that there are 1.5 $\pm$ 1.2 fluorophores per protein (Figure 4.1d), with the width of the distribution in line with that expected for a Poisson distribution. These values were confirmed by electrospray ionization mass spectrometry (ESI-MS), which found an average of 3, and up to 6 total, fluorophores per protein (Figure 4.6). The same analysis was performed for PAI-2 proteins labeled with a small green fluorophore (CF488, 0.9 kDa) and we obtained an average of 4.5 $\pm$ 2.2 fluorophores per protein (Figure 4.1e). Liposomes functionalized with red labeled PAI-2 were prepared via the post-insertion method, whereby micelles containing cysteine-reactive

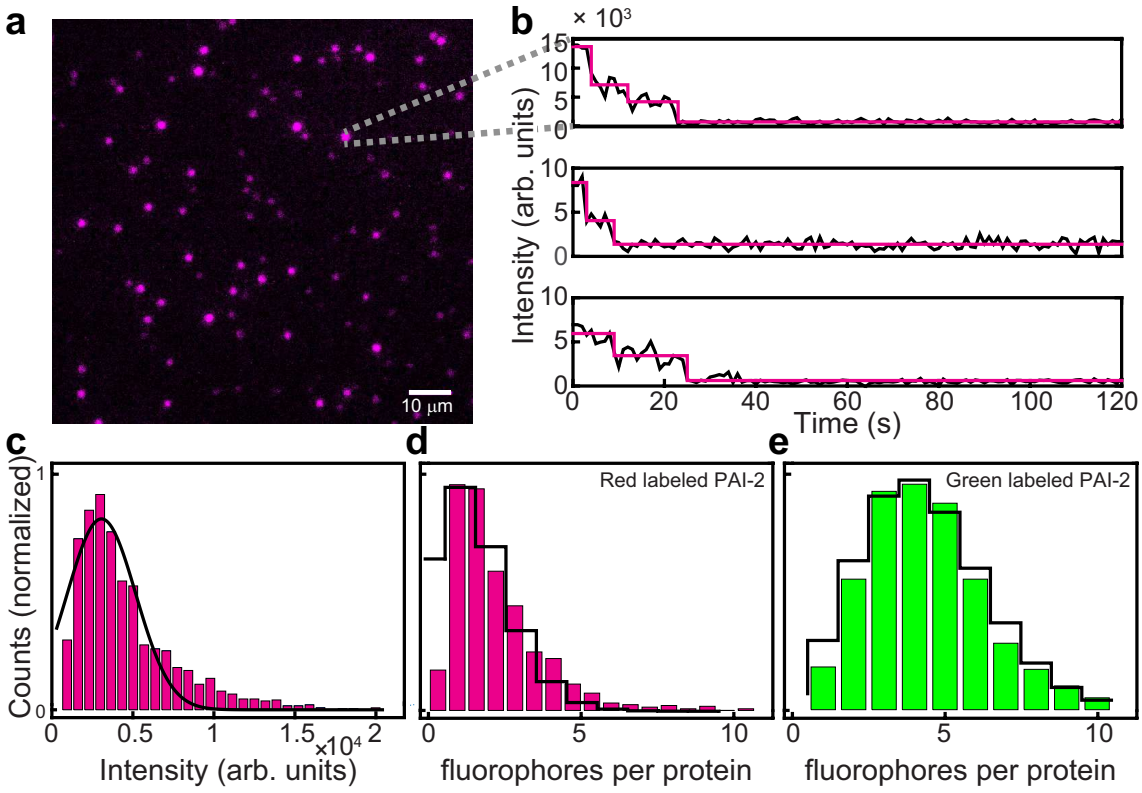


Figure 4.1: Measurement of the number of fluorophores per protein by TIRF microscopy. (a) Typical field of view — red labeled PAI-2 proteins were immobilized on cleaned coverslips. (b) Example intensity trajectories of individual labeled proteins (black line). The individual steps were identified using the change-point algorithm (magenta line) (52, 189). (c) Histogram of the intensity of a single CF647 fluorophore, fitted with a Gaussian distribution. The intensity for a single fluorophore is 3.0 $\pm$ 0.1 x 10^3 (mean $\pm$ s.e.m.). (d) Histogram of the number of CF647 fluorophores per protein, fitted with a Poisson distribution. The number of fluorophores is 1.47 $\pm$ 1.21 (mean $\pm$ s.d.). (e) Histogram of the number of CF488 fluorophores per protein, fitted with a Poisson distribution. This histogram was obtained in the same way as described for CF647 labeled proteins. The intensity for a single fluorophore is 1.2 $\pm$ 0.6 x 10^4 (mean $\pm$ s.e.m.) and the number of fluorophores per protein is 4.5 $\pm$ 2.2 (mean $\pm$ s.d.).

poly (ethylene glycol) (maleimide-PEG2000-DSPE) are reacted with protein to form functionalized micelles, before being incubated with pre-formed liposomes to promote insertion of the protein-PEG2000-DSPE conjugate into the outer leaflet

of the liposome (190). Liposomes were visualized using TIRF microscopy under the same conditions that were used to determine the number of fluorophores per protein. To confirm that the fluorescence signal observed in these experiments originates from proteins bound to single liposomes, we prepared the liposomes in the presence of the fluorophore R18 (Octadecyl Rhodamine B Chloride) so that the encapsulated R18 acts as a marker for only those liposomes that have an intact lipid bilayer (191). Using optics that split the image in a yellow and a red channel, the R18 labeled liposomes and the red labeled proteins were visualized simultaneously but each on different areas of the camera sensor.

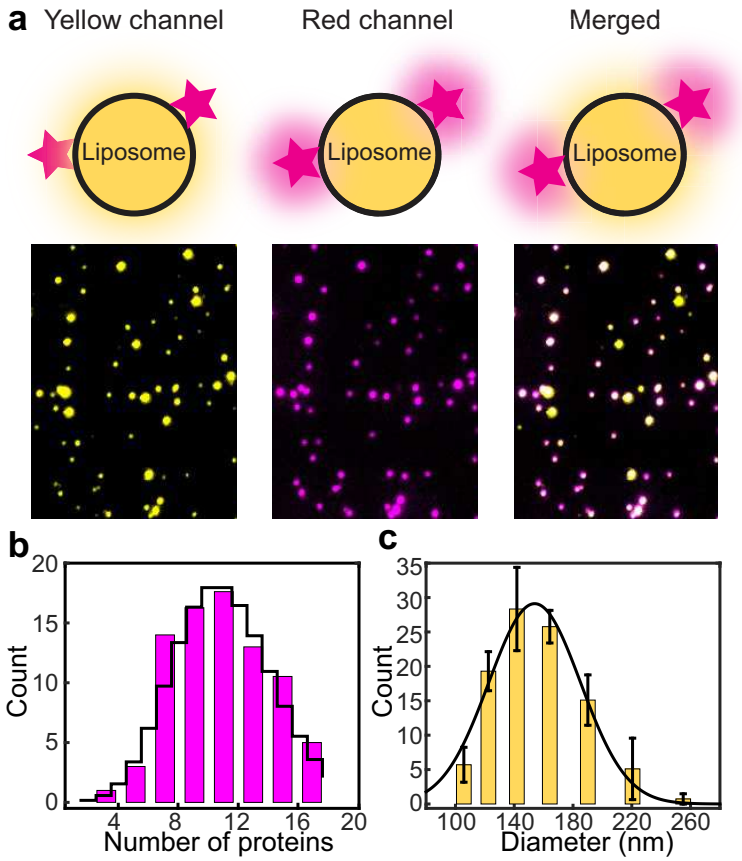


Figure 4.2: Visualization of proteins attached to liposomes. (a) Liposomes labeled with R18 (left) and proteins (middle) were imaged simultaneously. A merge of the two channels (right) showed a high degree of colocalization (white spots). (b) Histogram of the number of proteins per liposome, fitted with a Poisson distribution (black line). (c) Histogram of the diameter of the liposomes measured by dynamic light scattering, fitted with a Gaussian distribution (black line). Bars represent the mean ± s.d. (n = 3).

Figure 4.2a represents a typical field of view showing the R18 fluorescence (left), the signal from the red labeled proteins (middle), and a merge of the two signals (right), with colocalization indicated by white spots. Based on these images, we calculated that 88 % of liposomes have at least one protein attached. Liposomes prepared with non-malei-mide-functionalized micelles were used to confirm that only covalently attached proteins colocalize with liposomes in

imaging experiments (Figure 4.8). We then determined the number of proteins per liposome using the fluorescence intensity from the labeled proteins. We divided this intensity by the intensity of a single protein, obtained earlier (Figure 4.1d). We found a density of 11 $\pm$ 4 (mean $\pm$ s.d.) proteins per liposome (Figure 4.2B). Dynamic light scattering revealed a liposome diameter of 153 $\pm$ 56 nm (mean $\pm$ s.d.) (Figure 4.2c). The width of the distribution of the number of proteins per liposome correlates with the intrinsic width of the liposome size distribution. Therefore, the width of the distribution for the number of proteins per liposome is determined by the heterogeneity in liposome size.

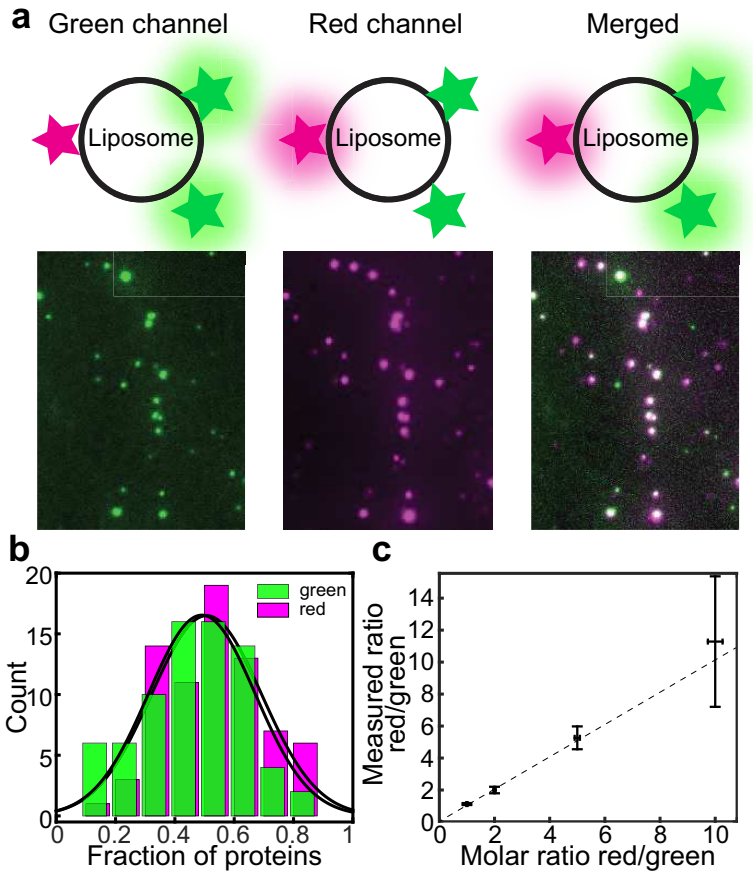


Figure 4.3: Quantification of the number of proteins per liposome. (a) Typical field of view showing dual-ligand immobilized liposomes. The green (left) and red dyes (middle) were visualized simultaneously. When the two channels are merged, colocalized spots show up as white (right). (b) Histograms of the measured fraction of green and red labeled proteins per liposome, when equal amounts of each were used during preparation. The fraction of green labeled proteins is 0.49 $\pm$ 0.02 and the fraction of red proteins is 0.51 $\pm$ 0.02. (c) Measured ratio of the fraction of red labeled proteins over the fraction of green labeled proteins as a function of the molar ratio used during preparation. The errors in the molar ratio are pipetting errors. The errors in the measured ratio are the s.e.m.

To explore the ability of single-molecule imaging to quantify differences in protein density, we varied the stoichiometry of two differently labeled proteins and quantified their ratio on the liposome surface. To negate any potential effects that would arise from using two different proteins, such as size and reactivity, we used

only PAI-2 proteins. Dual-ligand liposomes were prepared via the post-insertion method, using red and green labeled PAI-2 at molar ratios of 1:1, 2:1, 5:1 and 10:1, while keeping the total amount of protein added constant. The two proteins were visualized simultaneously using dual-color imaging (Figure 4.3a) and the protein density was determined as above. At a 1:1 molar ratio, we found 51 $\pm$ 2 $\%$ of the total number of proteins per liposome had a red label and 49 $\pm$ 2 $\%$ had a green label (Figure 4.3b). This observation indicates that the fluorophores do not affect protein attachment, and that the two proteins are incorporated in the same 1:1 ratio as their input stoichiometry in the formulation process. Further analysis revealed that changing the ratios of the two labeled proteins during preparation similarly altered the ratios of proteins incorporated into the liposome (Figure 4.3c, Figure 4.9).

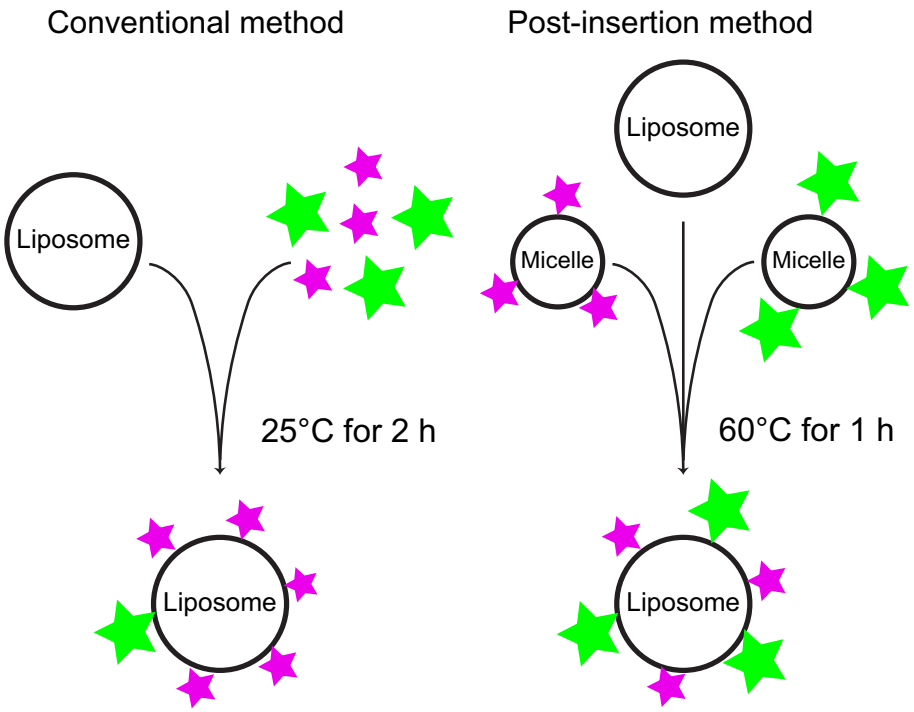


Figure 4.4: Conventional and post-insertion methods for dual-ligand liposome preparation. The conventional method involves the incorporation of polyethylene glycol (PEG) chains with a terminal maleimide functional group (maleimide-PEG-DSPE) into the lipid bilayer of the liposome during formation. Pre-formed liposomes are then incubated with two different thiolated proteins (represented by green and magenta stars) at 25°C, which attach covalently to the liposome surface via the maleimide moiety. The post-insertion method involves the creation of maleimide-PEG-DSPE micelles to which proteins are covalently attached as per the conventional method. Micelles are then incubated with pre-formed liposomes at a temperature of 60°C to facilitate the transfer of the micelle PEG-DSPE and attached ligands into the outer leaflet of the liposome bilayer. Figure not to scale.

These results highlight the accuracy of the single-molecule measurements, and illustrate the ability of this method to report on small differences in protein densities and ratios. Finally, we demonstrated the utility of single-molecule quantification in the characterization of novel clinically-relevant ligand-directed liposomes. Dual-ligand liposomes were prepared via both the conventional and the post-insertion

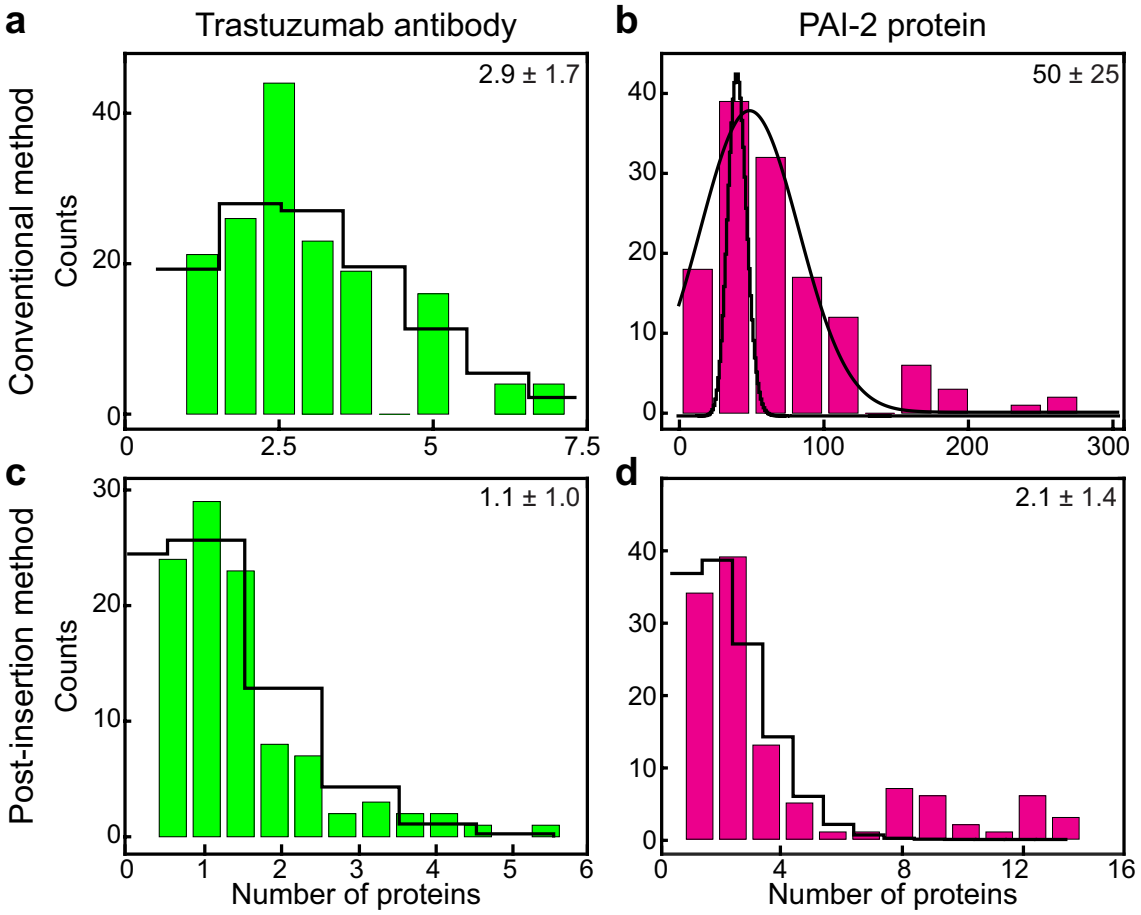


Figure 4.5: Comparison of the number of proteins per liposome prepared via the conventional and post-insertion methods. When using a 1:1 ratio of trastuzumab antibody to PAI-2 protein in the conventional preparation method, the number of PAI-2 proteins per liposome (**b**) was ∼17 times higher than the number of trastuzumab antibodies (**a**). These numbers were much more similar when the post-insertion method was used (**c**) and (**d**). The black lines represent Poisson distribution fits to the histograms. Due to the large number of proteins in panel B, heterogeneities within the sample broaden the histogram and obscure the Poisson distribution. This histogram was therefore fitted with a Gaussian distribution.

methods of liposome functionalization (Figure 4.4). PAI-2 and trastuzumab (TZ, Herceptin, 145 kDa) were labeled with red and green dyes, respectively, and added to pre-formed liposomes in a 1:1 molar ratio. Imaging and data analysis were performed as outlined above. Using our single-molecule imaging approach, we determined that the ratio of the PAI-2 and TZ incorporated into liposomes was closer to 1 for liposomes prepared via the post-insertion method (ratio = 2.1 $\pm$ 2.5) than for liposomes prepared via the conventional method (ratio = 17 $\pm$ 18) (Figure 4.5). The conventional method involves incubation of a small protein and a large antibody with pre-formed liposomes, where differences in protein size and reactivity may affect their equal incorporation into the liposomes. In contrast, the post-insertion method helps negate effects of these protein differences through the simultaneous insertion of two separate protein-functionalized micelles (192). These results provide a rationale for use of the post-insertion method in the production of dual-ligand liposomes functionalized with two very different proteins in

terms of their size and/or reactivity. The application of the single-molecule quantification informing on the preparation protocol allows for a better control of the stoichiometry of the liposomes produced.

4.3 Conclusion

In conclusion, we have demonstrated the practical utility of single-molecule fluorescence imaging in the quantification of the density of ligands attached to the surface of liposomes. This method enables the quantitative characterization of protein densities and the ability to detect changes therein, and permits future experiments to elucidate further characteristics of ligand-directed liposomes, including the quantification of inner leaflet and outer leaflet labeling of liposomes using environmentally sensitive dyes (193). The use of single-molecule imaging as a quantification technique is expected to improve the characterization of preclinical ligand-directed liposomes, assist with large-scale manufacturing processes and allow for batch-to-batch quality control in a commercial production setting. Using this technique, we showed that the post-insertion method of ligand-coupled liposome preparation is the preferred method for dual-ligand liposomes when using proteins of different sizes — an aspect relevant to the clinical setting, where liposomes used to target heterogeneous tumor cell populations would likely bear two different targeting ligands. By enabling the quantification of surface-bound ligands, and informing on optimal preparation protocols for ligand-directed liposomes, this single-molecule quantitative approach may help improve the translation of targeted liposomal drug delivery systems from the laboratory through to clinic use.

4.4 Materials and Methods

4.4.1 Labeling proteins with fluorophores.

Human recombinant plasminogen activator inhibitor-2 (PAI-2, SerpinB2), produced in-house by previously published methods (194), and trastuzumab (TZ, Herceptin; Genentech, CA, USA) were labeled with CF488 or CF647 succinimidyl ester fluorescent dyes (Sigma-Aldrich, MO, USA) as per the manufacturer's instructions. Absorbance at 280 nm (protein) and 488 nm or 647 nm (dye) was used to calculate the protein concentration and degree of labeling (DOL). DOL was further confirmed by electrospray ionization mass spectrometry (ESI-MS).

4.4.2 Electrospray ionization mass spectrometry (ESI-MS).

Positive ion mass spectra of unlabeled and labeled proteins were acquired on a quadrupole time of flight mass spectrometer (Q-TOF-MS) (Micromass Q-TOF Ultima; Waters, MA, USA) fitted with a Z-spray ionization source. Samples in phosphate-buffered saline (PBS, pH 7.4) were exchanged into deionized water containing 0.1% formic acid and made up to a final concentration of approximately

10 μM. The mass spectra were acquired with a capillary voltage of 2.6 kV, cone voltage of 50 V, source block temperature of 40°C, and a resolution power of 5000 Hz. Cesium iodide was used for external calibration. The mass spectrum data are presented as raw data, on an m/z scale. Mass was calculated using MassLynx MS V4.1 (Waters, MA, USA).

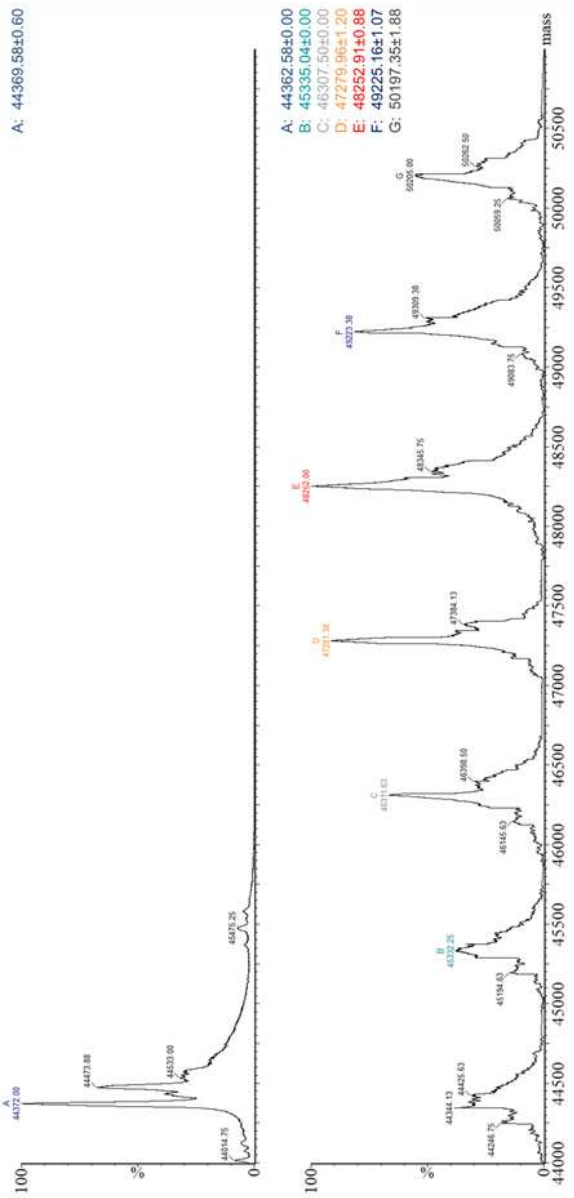


Figure 4.6: Positive ion ESI-MS spectra of unlabeled PAI-2 (top) and CF647 labeled PAI-2 (bottom). The difference in mass for each species represents the molecular weight of PAI-2 plus CF647 ($\sim$0.8 kDa). A = mass of PAI-2, B = mass of PAI-2 + 1 CF647 dye molecule, C = mass of PAI-2 + 2 CF647 dye molecules, and so on. Mass was calculated using MassLynx V4.1 (Waters, MA, USA).

4.4.3 Preparation of liposomes.

Liposomes were prepared using the thin film hydration method as described previously (195). Dipalmitoylphosphatidylcholine (DPPC), cholesterol, 1,2-distearoyl-sn-glycero-3-phosphoethanolamine-N-[(polyethylene- glycol)-2000] (mPEG2000-DSPE) and 1,2-distearoyl-sn-glycero-3-phos- pho ethanolamine-N-[maleimide(polyethylene glycol)-2000] (Avanti Polar Lipids, AL, USA) in a

20:10:0.8:0.2 molar ratio (conventional method) or DPPC, cholesterol and mPEG2000-DSPE in a 20:10:0.6 molar ratio (post-insertion method) were dissolved in chloroform/methanol (2:1 v/v). For colocalization experiments, liposomes were labeled with Octadecyl Rhodamine B Chloride (R18; Invitrogen, CA, USA) by adding R18 to the chloroform/methanol solution in a 160:1 molar ratio (liposome phospholipid:R18). Organic solvents were removed by rotary evaporation and subsequent freeze drying to form a lipid film. Phospholipids were reconstituted in degassed HEPES buffer (115 mM NaCl, 20 mM HEPES, 2.4 mM K_2PO_4, 1.2 mM $CaCl_2$, 1.2 mM $MgCl_2$; pH 7.4) at a concentration of 20 mM. Once reconstituted, liposomes were passed once through a 0.22 μm PVDF membrane (Merck Millipore, MA, USA) and then serially extruded 11 times through a 0.1 μm PVDF membrane using a syringe-driven extruding apparatus (Avanti Polar Lipids, AL, USA) at a temperature of 50°C (above the phase-transition temperature of DPPC). Liposomes were analyzed by dynamic light scattering to determine particle diameter using a Zetasizer APS (Malvern Instruments, Malvern, UK). Liposomes were surface-functionalized with CF647 labeled PAI-2 and/or CF488 labeled PAI-2 or TZ using either the conventional method or the post-insertion method.25 For the conventional method, pre-formed liposomes were incubated with thiolated labeled PAI-2 or TZ (at a molar ratio of 3333:1 liposome phospholipid:protein) for 2 hours at room temperature. For the post-insertion method, micelles composed of 0.8 mM mal-PEG2000-DSPE and 0.2 mM mPEG2000-DSPE were prepared as per previously reported methods (196), and labeled PAI-2 or TZ added to the micelles (at a molar ratio of 10:1, mal-PEG2000-DSPE:protein) to form functionalized micelles. Functionalized micelles were added to pre-formed liposomes and heated to 60°C for 1 hour to facilitate post-insertion of micelle lipids into the outer leaflet of the liposomes. Following the liposome functionalization steps, unbound protein was removed from liposomes via repeated centrifugation at 20,000 x g for 1.5 hours at 4°C. Liposomes were resuspended in HEPES buffer (pH 7.4) for single-molecule imaging.

4.4.4 Intensity measurements for labeled proteins.

Microscope coverslips were thoroughly cleaned to remove any hydrophobic and hydrophilic contaminants that could cause background fluorescence from the glass. They were first sonicated for 30 min in ethanol (Chem-Supply, SA, AUS) and then rinsed with deionized water. Subsequently they were sonicated for 30 min in 1 M potassium hydroxide (KOH; Sigma-Aldrich, MO, USA) and rinsed with deionized water again. After these sonication steps were repeated, the coverslips were dried with N2 (155). CF labeled proteins were diluted to a concentration of approximately 10 pM and immobilized on the surface of the cleaned microscope coverslip for visualization on an inverted microscope (Nikon Eclipse Ti-E) with a CFI Apo TIRF 100x oil-immersion TIRF objective (NA 1.49, Nikon) (Figure 4.7). The green and red labeled proteins were excited at 1.5 W cm^{-2} with 488 nm

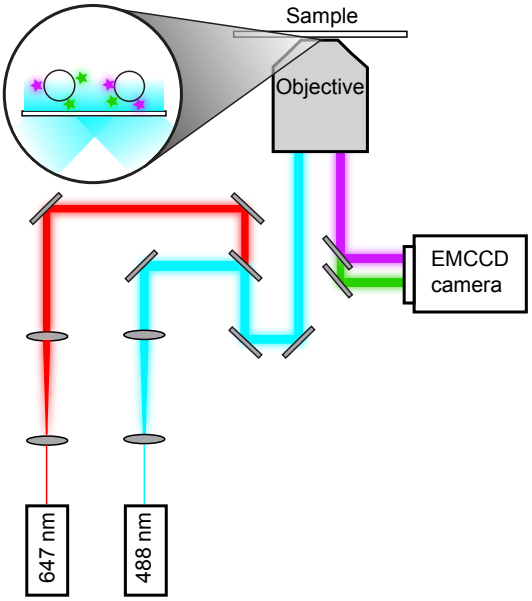


Figure 4.7: Schematic overview of the single-molecule fluorescence microscope. Laser light of a specific wavelength is coupled into the microscope objective. The fluorescence signal from the sample is detected with an EMCCD camera.

(Coherent, Sapphire 488-200 CW) and 647 nm (Coherent, Obis 647-100 CW) lasers, respectively (Figure 4.7). The signals were separated via dichroic mirrors (Photometrics, DVΛ Multichannel Imaging System) and appropriate filter sets (Chroma). The imaging was done with an EMCCD (Photometics, Evolve 512 Delta). Using ImageJ (National Institutes of Health, USA) with in-house built plugins, we calculated the integrated intensity for single CF dyes over time, after applying a local background subtraction. Using a change-point step-fitting algorithm, we calculated the intensity distributions for a single CF fluorophore (Figure 4.1b).21,22 The histograms obtained were fit with a Gaussian distribution function using MATLAB 2014b, to give a mean intensity of 2030 $\pm$ 40 for the CF647 (Figure 4.1c) and 1340 $\pm$ 50 for the CF488. To measure the number of fluorophores per protein, we divided the initial fluorescence intensity per protein by the intensity of a single fluorophore (Figure 4.1d, e).

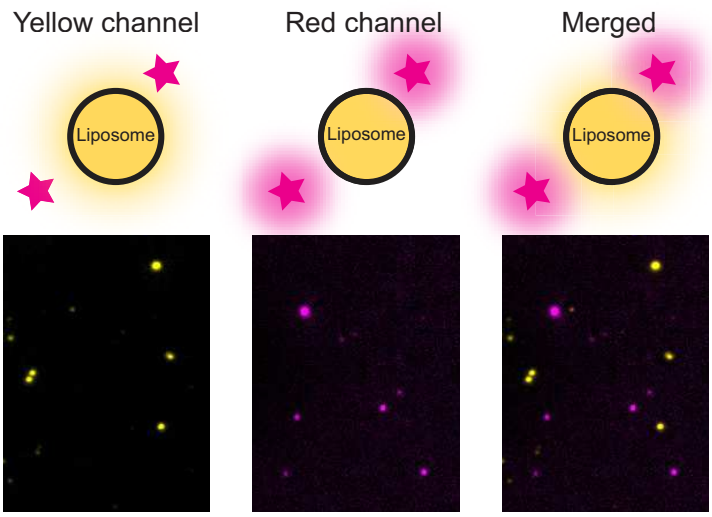


Figure 4.8: Imaging of liposomes labeled with R18 (yellow) and prepared with red labeled PAI-2 (magenta) in the absence of maleimide-PEG2000-DSPE in micelles. Liposomes (left) do not colocalize with PAI-2 proteins (middle). This is shown by a merge of the two channels (right), which does not show the white colocalization spots seen in Figure 4.2. The lack of colocalization demonstrates that nonspecific binding of PAI-2 proteins to liposomes is minimal.

4.4.5 Measurement of protein density on liposomes.

To find the number of proteins per liposome, we imaged the liposomes under the same conditions and calculated the fluorescence intensity per liposome analogously. We obtained the number of proteins per liposome by dividing these intensities by the intensity of a single protein (Figure 4.1).

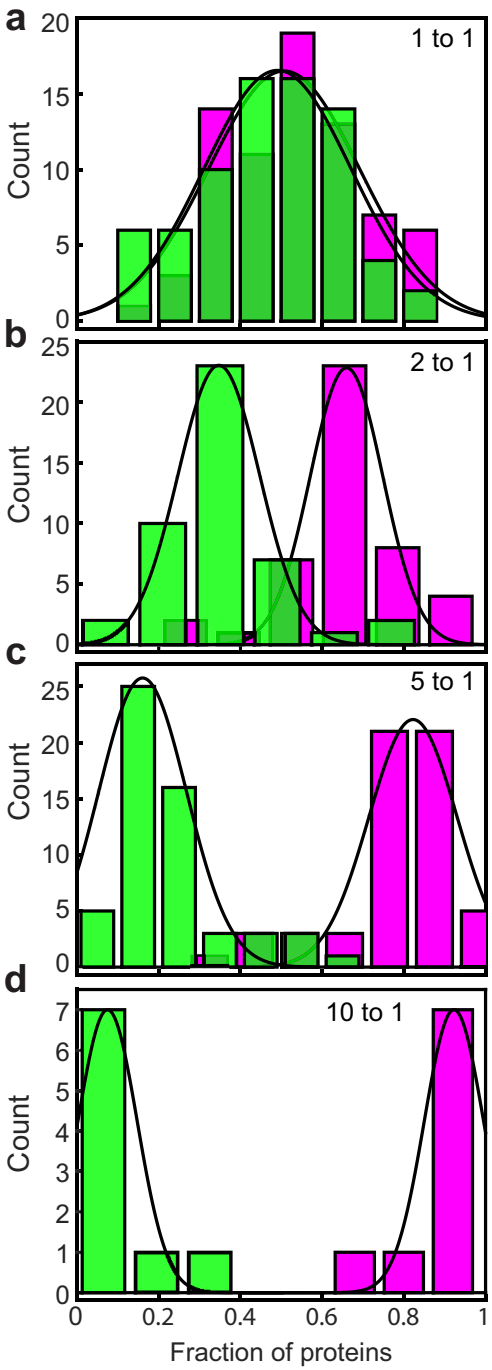


Figure 4.9: Histograms of the number of labeled PAI-2 proteins per liposome. Histograms of the number of red (magenta) and green (green) labeled proteins per liposome using a (a) 1 to 1, (b) 2 to 1, (c) 5 to 1 and (d) 10 to 1 molar ratio of red labeled protein to green labeled protein during preparation. The black lines represent Gaussian fits to the data.

5 | Single-molecule visualisation of fast polymerase turnover in the bacterial replisome

Jacob S. Lewis[†], **Lisanne M. Spenkelink**[†], Slobodan Jergic, Elizabeth A. Wood, Nicholas P. Horan, Karl E. Duderstadt, Michael M. Cox, Andrew Robinson, Nicholas E. Dixon, Antoine M. van Oijen.

[†]These authors contributed equally.

Published in *eLife* 2017;10.7554/eLife.23932

The *Escherichia coli* DNA replication machinery has been used as a road map to uncover design rules that enable DNA duplication with high efficiency and fidelity. Although the enzymatic activities of the replicative DNA Pol III are well understood, its dynamics within the replisome are not. Here we test the accepted view that the Pol III* holoenzyme remains stably associated within the replisome. We use *in vitro* single-molecule assays with fluorescently labeled polymerases to demonstrate that the Pol III* complex (holoenzyme lacking the β_2 sliding clamp), is rapidly exchanged during processive DNA replication. Nevertheless, the replisome is highly resistant to dilution in the absence of Pol III* in solution. We further show similar exchange in live cells containing labeled clamp loader and polymerase. These observations suggest a concentration-dependent dissociative mechanism providing a balance between stability and plasticity, facilitating replacement of replisomal components dependent on their availability in the environment.

J.S.L. and I contributed equally to this paper. I led the development and implementation of all in vitro and in vivo single-molecule experiments and wrote the paper.

5.1 Introduction

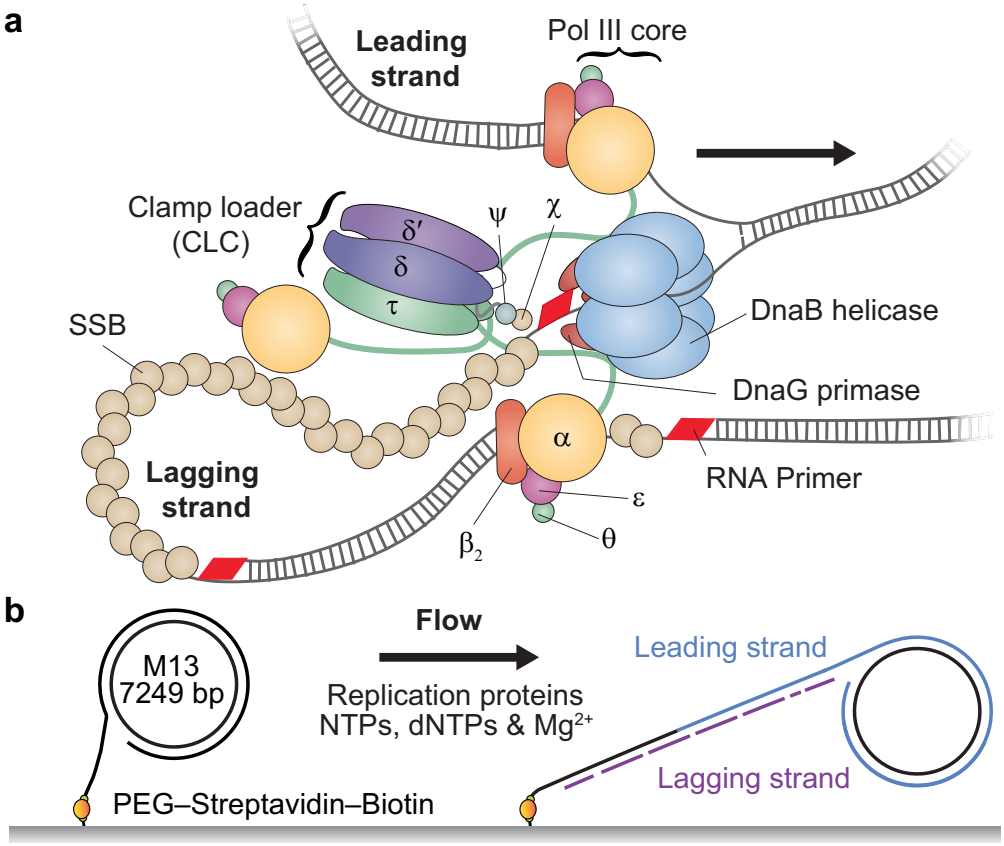


Figure 5.1: Single-molecule rolling-circle replication assay. (a) Canonical view of the organization of the *E. coli* replication fork. The DnaB helicase encircles the lagging strand, facilitates unwinding of dsDNA through ATP hydrolysis, and recruits DnaG primase for synthesis of RNA primers that initiate synthesis of 1–2 kb Okazaki fragments on the lagging strand. The extruded single-stranded (ss) DNA is protected by ssDNA-binding protein, SSB. The Pol III holoenzyme (HE) uses the ssDNA of both strands as a template for coupled, simultaneous synthesis of a pair of new DNA duplex molecules. The β_2 sliding clamp confers high processivity on the Pol III HE by tethering the $\alpha\epsilon\tau$ Pol III core onto the DNA. The clamp loader complex (CLC) assembles the β_2 clamp onto RNA primer junctions. Up to three Pol III cores interact with the CLC through its τ subunits to form the Pol III* complex, and the τ subunits also interact with DnaB, thus coupling the Pol III HE to the helicase. (b) Schematic representation of the experimental design. 5'-Biotinylated M13 DNA is coupled to the passivated surface of a microfluidic flow cell through a streptavidin linkage. Addition of the *E. coli* replication proteins and nucleotides initiates DNA synthesis. The DNA products are elongated by hydrodynamically flow, labeled with intercalating DNA stain, and visualized using fluorescence microscopy.

The *E. coli* replisome requires participation of 13 different proteins. Ten of them form the DNA polymerase III (Pol III) holoenzyme (HE), which is arranged into three functionally distinct and stably-bound subassemblies: $\alpha\epsilon\theta$ forms the Pol III core that has DNA polymerase activity; β_2 is the sliding clamp needed for stable association with the primer-template DNA; and $\tau_n\gamma_{(3n)}\delta\delta'\chi\psi$ where n = 2 or 3 in the HE) is the clamp loader complex (CLC) that loads β_2 onto DNA and is the central organizer of the replisome (Figure 5.1a) (15, 16, 197). The CLC interacts

with two or three Pol III cores via the α–τ interaction, forming stable complexes termed Pol III* (i.e., HE lacking only the sliding clamp). Pol III* ensures the organization of the cores needed for coordinated DNA synthesis on the two template strands (198, 199) and is essential for cell survival (200). Although physical coupling of leading and lagging strand cores in one HE particle requires the lagging strand polymerase to undergo cycles of release and rebinding from one Okazaki fragment to the next, the molecular mechanisms underlying its cycling are still debated (201). There is, however, consensus that Pol III is reused rather than replaced for successive Okazaki fragment synthesis (14, 136, 202–204). Thus, the replisome is believed to be a highly stable entity. The key observations that support efficient Pol III recycling derive from *in vitro* replication assays in the absence of free polymerase (136, 203, 204), and are consistent with the high stability of the α–τ interaction that binds cores to the CLC (K_D = 0.3 nM; $t_{1/2}$ = 29 min in 300 mM NaCl) (205). Nevertheless, the introduction of high concentrations of catalytically dead Pol III* (still able to bind primed DNA) inhibits ongoing replication (130). Reconciling these different observations, we here demonstrate the presence of a novel exchange mechanism that allows Pol III* to remain stably associated with the replisome under conditions of high dilution, yet facilitates rapid exchange at nanomolar concentrations.

5.2 Results

5.2.1 *In vitro* single-molecule observation of Pol III dynamics

We use a single-molecule approach to directly visualize the dynamics of Pol III complexes at the replication fork (14, 206). A rolling-circle DNA amplification scheme is used to observe highly processive DNA synthesis in real time, while imaging Pol III complexes entering and leaving the replisome. Using the minimal set of 12 proteins required to support coupled leading and lagging strand synthesis, we allow active replisomes to self assemble onto pre-formed replication forks (14, 202). A 5′-flap within a 7.2 kb double-stranded (ds) circular DNA substrate is anchored to the surface of a microfluidic flow cell and replication is initiated by introducing a laminar flow of buffer with the components required for coupled leading and lagging strand synthesis (Figure 5.1b). As replication proceeds, the newly synthesized leading strand becomes part of the circle and later acts as a template for lagging strand synthesis. With the lagging strand attached to the surface and the continuously growing DNA product stretched in the buffer flow, the dsDNA circle moves away from the anchor point. Replication is visualized by real-time near-TIRF fluorescence imaging of stained dsDNA (Figure 5.2a, Figure 5.11). This strategy allows quantification of the rates of individual replisomes and their processivities (Figure 5.2b). We fluorescently labeled the Pol III α subunit following its fusion to a SNAP tag (Figure 5.7) and covalently coupled it separately in >80$\%$ yields to red and green fluorophores (206). Fluorescently labeled Pol III cores were reconstituted from individual SNAP-α, ϵ and θ subunits and isolated chromatographically (202), then assembled into single-color Pol III*s in situ with separately-isolated τ_3-CLC (202). The labeled Pol III*s were active in

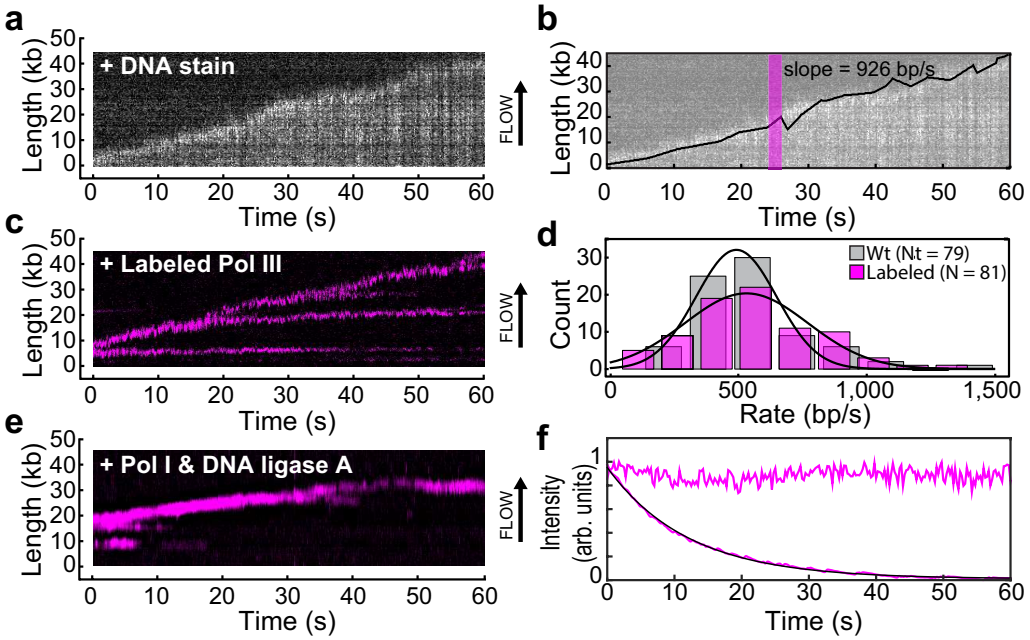


Figure 5.2: Real-time fluorescence imaging of coupled DNA replication. (a) Kymograph of an individual DNA molecule undergoing coupled leading and lagging strand replication. The grey scale indicates the fluorescence intensity of stained DNA. (b) Single-molecule trajectory obtained from the kymograph in (a), used to quantify the rates and processivities of replication events. The magenta box represents an example line segment used to determine rates. (c) Kymograph of the dynamics of red-labeled Pol IIIs on an individual DNA molecule. The Pol III moves with the replisome in the direction of flow as it elongates the DNA, visible as a bright magenta spot moving away from the surface anchor point. Additional Pol IIIs are left behind the moving replisome, seen as horizontal lines on the kymograph. (d) Histograms of the rate of replication for wild-type Pol III (492 ± 23 bp s^{-1}) and red Pol III (561 ± 27 bp s^{-1}) fit to Gaussian distributions. (e) Kymograph of the distribution of red Pol III on an individual DNA molecule in the presence of 150 nM Pol I and 100 nM DNA ligase. Prolonged Pol III spots behind the replisome are no longer observed due to the action of Pol I in Okazaki fragment processing. (f) Fluorescence intensity as a function of time of individual red Pol IIIs immobilized on the surface of a coverslip (lower trace; black line is an exponential fit with lifetime = 14.1 ± 0.4 s), and of the replisomal spot in (c) (upper trace). The fluorescence lifetime of red Pol III at the replisome is much longer than the photobleaching lifetime of the dye. The errors represent the standard errors of the mean.

coupled DNA replication, producing Okazaki fragments of similar sizes to wild-type polymerase (Figure 5.8b). A kymograph (Figure 5.2c) shows the fluorescence of the red Pol III* during rolling-circle replication; it supports replication at rates similar to the untagged wild-type enzyme (Figure 5.2d). Simultaneous imaging of the stained DNA and red Pol III* shows that the polymerase spot is located at the tip of the growing DNA, confirming that the labeled Pol III is a functional component of reconstituted replisomes (Figure 5.12). We also observe Pol III that remains bound to the DNA on the lagging strand behind the replication fork, evident as horizontal lines in Figure 5.2c. We reasoned that these correspond to polymerases bound to the 3′-termini of Okazaki fragments. We repeated the experiment in the presence of Pol I and/or DNA ligase; Pol I replaces RNA primers with DNA and ligase seals the remaining nick. In the presence of Pol I (with or without ligase), Pol III binding behind the replisome is no longer observed (Figure 5.2e), consistent with Pol I efficiently displacing Pol III during Okazaki fragment

maturation. Surprisingly, the fluorescent Pol III at the growing tip of the rolling circle is highly resistant to photobleaching. Its fluorescence in the replisome has a much longer lifetime compared to that of labeled Pol III cores immobilized on a surface and subjected to the same excitation intensity (Figure 5.2f). Since the experiments in Figure 5.2 are performed with 6.7 nM Pol III* in solution, this observation suggests that the polymerase exchanges into the replisome from solution to replace photobleached Pol III.

5.2.2 Exchange of Pol III* complexes *in vitro*

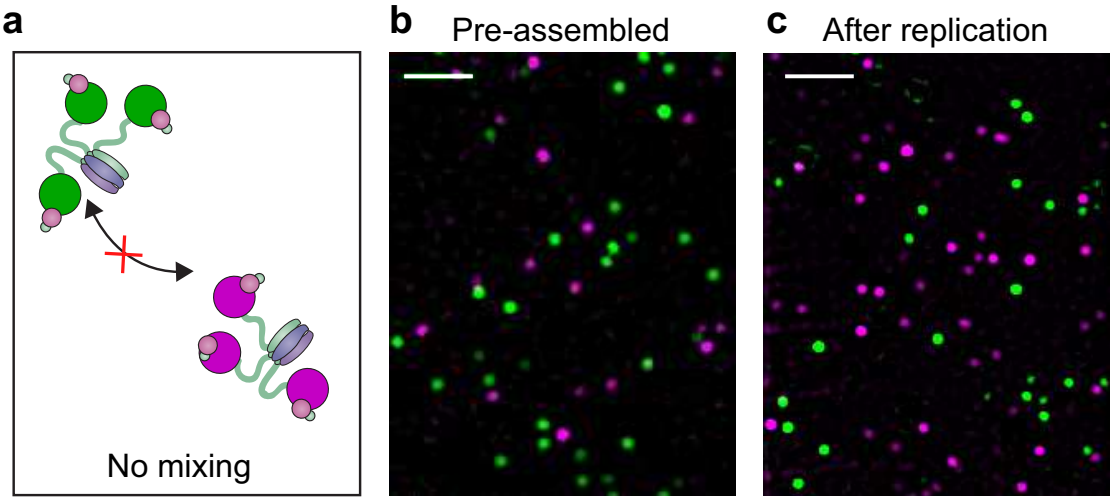


Figure 5.3: Pre-assembled Pol III* complexes do not exchange Pol III core. (a) Red and green Pol III* are separately pre-assembled by treatment at 37°C for 15 min (30 nM Pol III core and 10 nM τ_3-CLC). These are then mixed in equal ratios and kept at 37°C for 1 hour prior to dilution to 6 pM Pol III* for imaging. (b) Red Pol III* complexes and green Pol III* complexes do not co-localize to produce any white spots as seen in Figure 5.18, demonstrating the α–τ interaction within the Pol III* complex remains intact for the duration of the DNA replication assays. (c) Pre-assembled red and green Pol III* complexes that have participated in DNA replication (at 3.3 nM of each) do not co-localize, demonstrating that the Pol III cores within a Pol III* do not exchange with cores from other Pol III*s at the replication fork during active DNA synthesis. White scale bars represent 5 μm.

To characterize the dynamic behavior of Pol III at the fork and directly visualize its exchange in real time, we used mixtures of red and green Pol III*s. To demonstrate that green Pol III cores in a Pol III* complex do not exchange with the red ones from another Pol III*, we combined them in a 1:1 ratio for 1 hour at 37°C (Figure 5.3a), then imaged the mixture on the surface of a coverslip at the single-molecule level (Figure 5.3, Figure 5.19). Consistent with the stable interaction between α in the core and τ in the CLC, exchange of Pol III cores was not observed. It remained possible, however, that the nature and strength of the α–τ interaction is different at a step in lagging strand replication that involves exchange of Pol III cores within the Pol III* complex (207). To test this possibility, we mixed pre-assembled red and green Pol III* complexes in a 1:1 ratio and used them in a bulk rolling-circle replication experiment in an 8-fold molar excess

over 5′-biotinylated flap-primed dsDNA template. Under these conditions, most Pol III*s will have participated in replication at the fork, as long leading strand and shorter lagging strand products are generated. Next, the newly synthesized DNA was removed from proteins by its immobilization on streptavidin beads, and subsequent single-molecule imaging of the released protein fraction on the surface of a coverslip, showed no co-localization of red and green Pol III cores (Figure 5.3). This result confirms that the functional unit exchanging at the replication fork is the entire Pol III* complex; the interaction between the τ subunit of the CLC and α of Pol III must remain intact during DNA replication, thus challenging the previously suggested model of a τ processivity switch on the lagging strand (207). We visualized exchange of Pol III* at the replication fork by measuring the fluorescence intensity at the replisome spot as a function of time using 1:1 mixtures of red and green Pol III*s (Figure 5.4A,B, Figure 5.19). At a total Pol III* concentration of 6.7 nM, the replisomal spot exhibits fast dynamics displaying both colors, while at a lower concentration of 0.3 nM, the dynamics appear slower and distinct exchange events are visible. The longer persistence of a single color at the lower concentration demonstrates that Pol III* exchange is concentration dependent. Given that Pol III* remains intact on time scales much longer than the duration of our experiment, these observations can only be explained by wholesale exchange of Pol III* at the replication fork. Our demonstration of rapid exchange of entire Pol III*s, however, seems difficult to reconcile with observations that both leading and lagging strand Pol III cores remain stably associated during coupled DNA replication (136, 203, 204). Those studies used assays in which replisomes were assembled, replication initiated, and the reactions rapidly diluted to measure the stability of synthesising replisomes on DNA. To place our observations of dynamic exchange of Pol III* in context of the previous work, we carried out single-molecule pre-assembly replication assays (136, 203) using the red Pol III*. In this experiment, the replisome is pre-assembled onto the rolling-circle template in solution. Subsequently, the template is attached to the surface of a flow cell, which is then washed to remove all unbound proteins. Replication is initiated by introduction of a replication solution that omits Pol III* and helicase. Since the absence of free Pol III* in solution makes polymerase exchange impossible, we hypothesized that Pol III would be recycled within the replisome, enabling its sustained participation in processive DNA replication. These conditions indeed support highly processive DNA replication (Figure 5.4c), with synthesis rates and processivities identical to a situation with Pol III* in solution and consistent with values reported previously (14, 203, 206, 208) (Figure 5.4d,e). Further, we observed photobleaching without recovery, consistent with the original, pre-assembled Pol III* remaining stably associated within the replisome. As further confirmation of the robustness of the pre-assembled replisome in the absence of competing polymerases and the easy displacement of Pol III* upon challenge, we initiate replication by pre-assembly of replisomes, normally supporting highly processive synthesis, and challenge them with Pol III core. The observation of a sharp reduction in processivity is consistent with the displacement of the Pol III* from the replication fork by the Pol III cores, which are unable to support coordinated leading and lagging strand synthesis (198) (Figure 5.20).

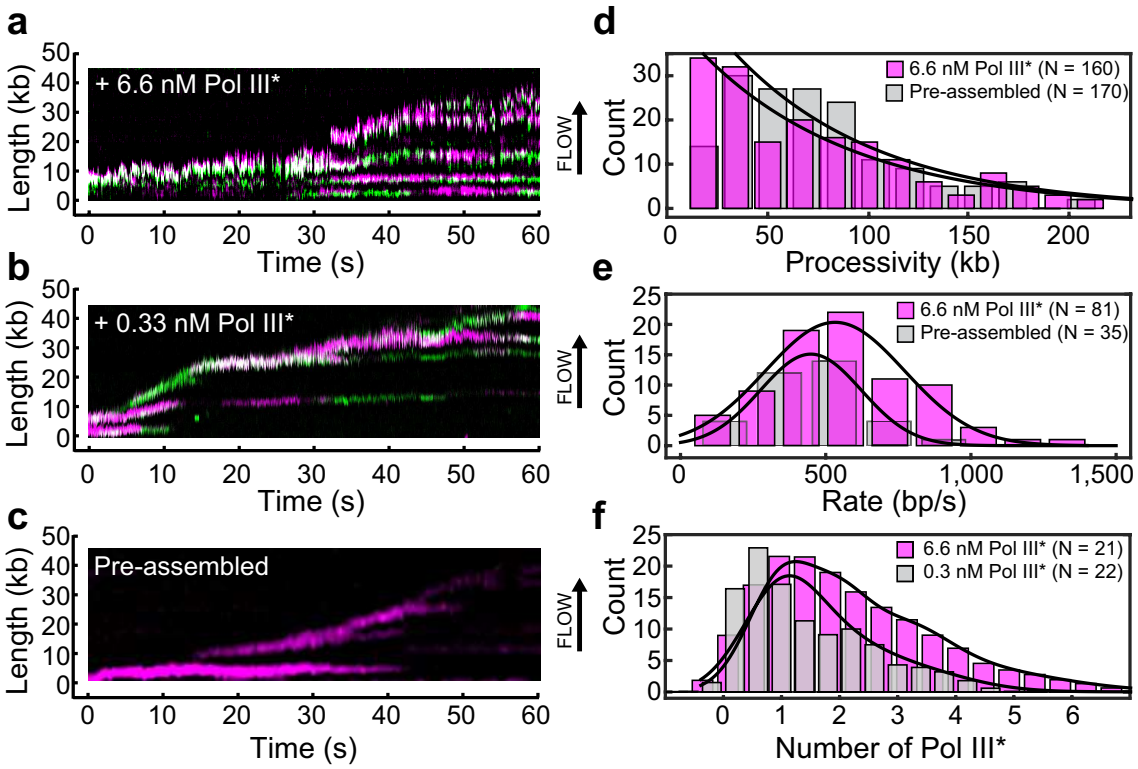


Figure 5.4: Rapid and frequent exchange of Pol III* is concentration dependent. (a) and (b) Kymographs of the distributions of red Pol III* (magenta) and green Pol III* (green) on an individual DNA molecule at a total Pol III* concentration of 6.7 (A) or 0.3 nM (B). Co-localization of the two signals is shown as a bright white fluorescent spot. (c) Kymograph of a pre-assembled replisome containing red Pol III*. The intensity of the signal from the replisomal spot decreases after a Pol III* is left behind. It subsequently bleaches and the signal does not recover. (d) Histograms of the processivity of replication with Pol III* present in solution (73 $\pm$ 25 kb) and under pre-assembly conditions (76 $\pm$ 26 kb), each fit with a single exponential decay function. (e) Histograms of the rates of replication with Pol III* present in solution (561 $\pm$ 27 bp s^1) and under pre-assembly conditions (445 $\pm$ 33 bp s^1), each fit to a Gaussian distribution. (f) Histograms of the stoichiometry of Pol III* at the replication fork, fit to four (6.7 nM) or three (0.3 nM) Gaussians centred at integral numbers of Pol III* calculated from single Pol III core intensities (Figure 5.15). The black lines represent the sums of these distributions. The errors represent the standard errors of the mean.

5.2.3 Quantification of exchange time of Pol III* *in vitro*

To quantify the concentration-dependent exchange times of Pol III* during coupled DNA replication, we performed *in vitro* single-molecule FRAP (fluorescence recovery after photobleaching) experiments. We visualized red Pol III* entering and leaving the replication fork at different concentrations using the same rolling-circle amplification scheme as described in Figure 5.1b. Instead of continuous imaging at constant laser power, we periodically bleached all Pol III* at the replication fork using a high laser power (Figure 5.5a). By bleaching the fluorescence signal of Pol III* complexes, we can monitor the recovery of the florescence signal as unbleached Pol III*s from solution exchange into the replisome (Figure 5.5b). We monitored the recovery of the fluorescence signal and calculated the average intensity after each FRAP pulse over time (Figure 5.5c).

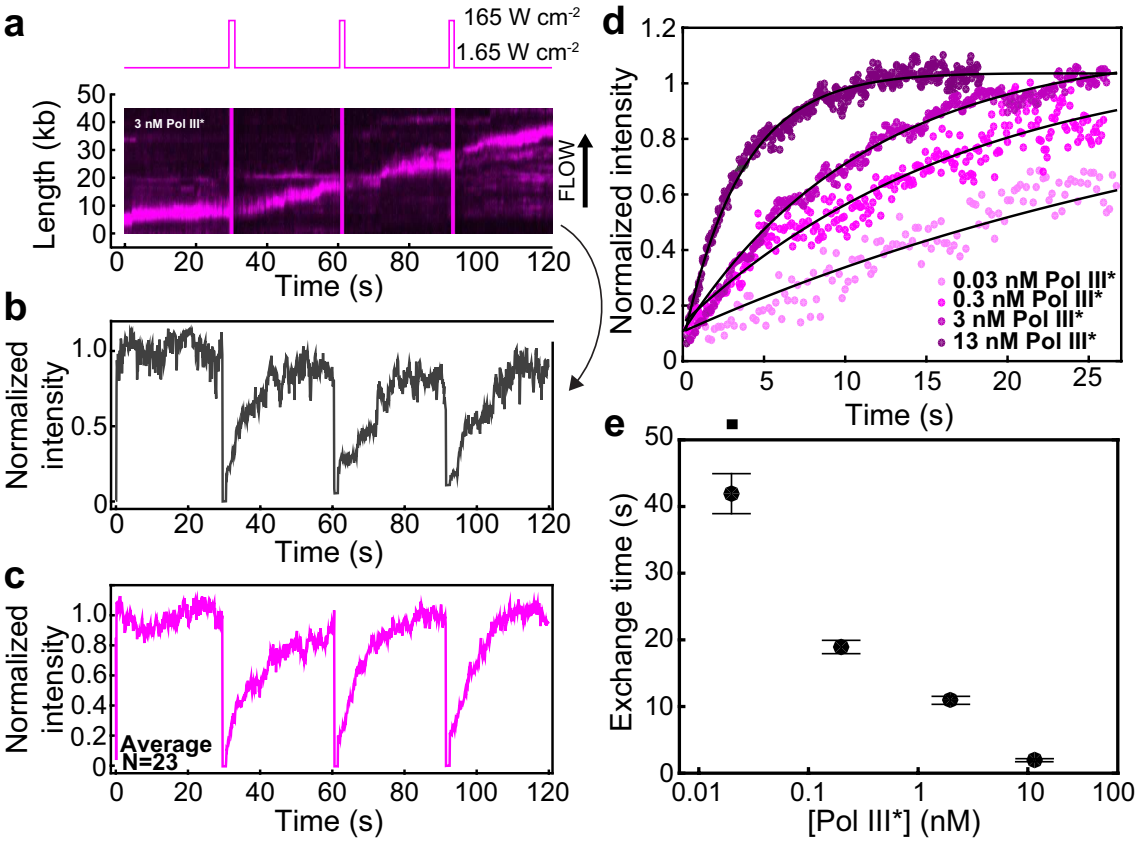


Figure 5.5: Quantification of Pol III* exchange time using single-molecule FRAP. (a) (Top Panel) Imaging sequence used during the FRAP experiments. Periodically, a FRAP pulse of high laser power was used to rapidly photobleach all the Pol III* in the field of view. (Bottom panel) A representative kymograph of red Pol III*s at the replication fork. After each FRAP pulse (indicated by the magenta line) all Pol III*s have bleached, but the fluorescence intensity recovers as unbleached Pol III*s exchange into the replisome. (b) Normalized intensity over time for an individual replisome in the presence of 3 nM Pol III* in solution. (c) The average intensity over time from 23 replisomes with 3 nM Pol III* in solution. (d) The three recovery phases in (c) were averaged again to give the final averaged normalized intensity over time after a FRAP pulse. This curve was then fit to provide a characteristic exchange time. This was done for four concentrations of Pol III* ranging from 13–0.03 nM. (e) Exchange time as a function of Pol III* concentration.

By measuring the single-molecule FRAP of Pol III* over a concentration series spanning four orders of magnitude and fitting the rate of signal recovery, we obtained the characteristic exchange time of Pol III* into active replisomes (Figure 5.5d). At a total Pol III* concentration of 13 nM, the fluorescence signal recovers rapidly (characteristic exchange time, τ = 1.85 s), while at 30 pM the fluorescence signal is 20-fold slower to recover (τ = 42 s). These observations are in agreement with our previous two-color experiments, indicating that the rate of exchange is dependent on Pol III* concentration.

5.2.4 Exchange of Pol III* complexes in live cells

Inspired by our observations of rapid exchange of Pol III* *in vitro*, we used *in vivo* single-molecule measurements to determine whether Pol III* exchange also occurs in live *E. coli* cells. We imaged cells in which the clamp loader and core

complexes were labeled at their C-termini with yellow and red fluorescent proteins, respectively (τ-YPet, ϵ-mKate2). As observed previously, these fusions remain fully functional (108) (Figure 5.16). These cells were immobilized on a (3-aminopropyl)triethoxysilane-treated coverslip and τ-YPet and ϵ-mKate2 foci were imaged simultaneously. Fluorescent proteins bound to large structures such as the nucleoid diffuse slowly and thus present in our images as diffraction-limited foci, whereas the signal from proteins freely diffusing through the cytosol blurs over the entire cell (110). To monitor exchange of polymerase molecules at replisomes, the fluorescence intensity of individual replication foci was tracked over time. We noticed that once the population of fluorescent molecules had become partially nonfluorescent by irreversible photobleaching, instances of synchronized intensity fluctuations of τ-YPet and ϵ-mKate2 within the replisome foci could be observed (Figure 5.6a). To determine whether these intensity fluctuations were truly correlated, i.e. whether they could be explained by the exchange of Pol III*, we used cross-correlation analysis, a powerful unbiased method that enables the calculation of (i) the extent of similarity between two fluctuating signals and (ii) at which timescales that similarity occurs. The average cross-correlation function calculated for 1210 foci in 480 cells showed a clear positive cross-correlation peak, consistent with synchronous exchange of τ-YPet and ϵ-mKate2 (Figure 5.5b, black line). To show that this peak arises due to protein dynamics, we fixed cells with formaldehyde to arrest all cellular processes and demonstrate the absence of a cross-correlation peak (Figure 5.6b, gray line). To eliminate the possibility of correlated intensity changes due to laser fluctuations, we calculated the cross-correlation function for random pairs of τ-YPet and ϵ-mKate2 foci within the same field of view. Also here, no cross-correlation peak was detected (Figure 5.6b, red line). Our experimental data cannot be explained by exchange of core only, the uncoupled exchange of both ϵ and τ, or even the complete absence of exchange (Figure 5.21). In support of Pol III* exchange, a positive cross-correlation peak can, however, be explained by simultaneous exchange of ϵ and τ (Figure 5.21). We then calculated the *in vivo* exchange time by fitting the cross-correlation function with an exponential decay. From this we found an exchange time of 4 $\pm$ 2 s (5.6C), consistent with measurements performed under similar experimental conditions (Beattie et al., 2017). Furthermore, the concentration of ϵ and τ was determined in the cell (Figure 5.17). Similarly to previous observations (108), we found under our experimental conditions a total concentration of 72 $\pm$ 3 nM of ϵ and 67 $\pm$ 5 nM of τ. Assuming all ϵ and τ form functional Pol III* complexes within the cell, these concentrations of ϵ and τ would correspond to $\approx$23 nM Pol III* per cell. These *in vivo* measurements are consistent with exchange times measured for the highest concentration of Pol III* *in vitro* (a few seconds at 13 nM). Given these observations, we conclude that Pol III* exchange occurs during coupled DNA replication *in vivo*.

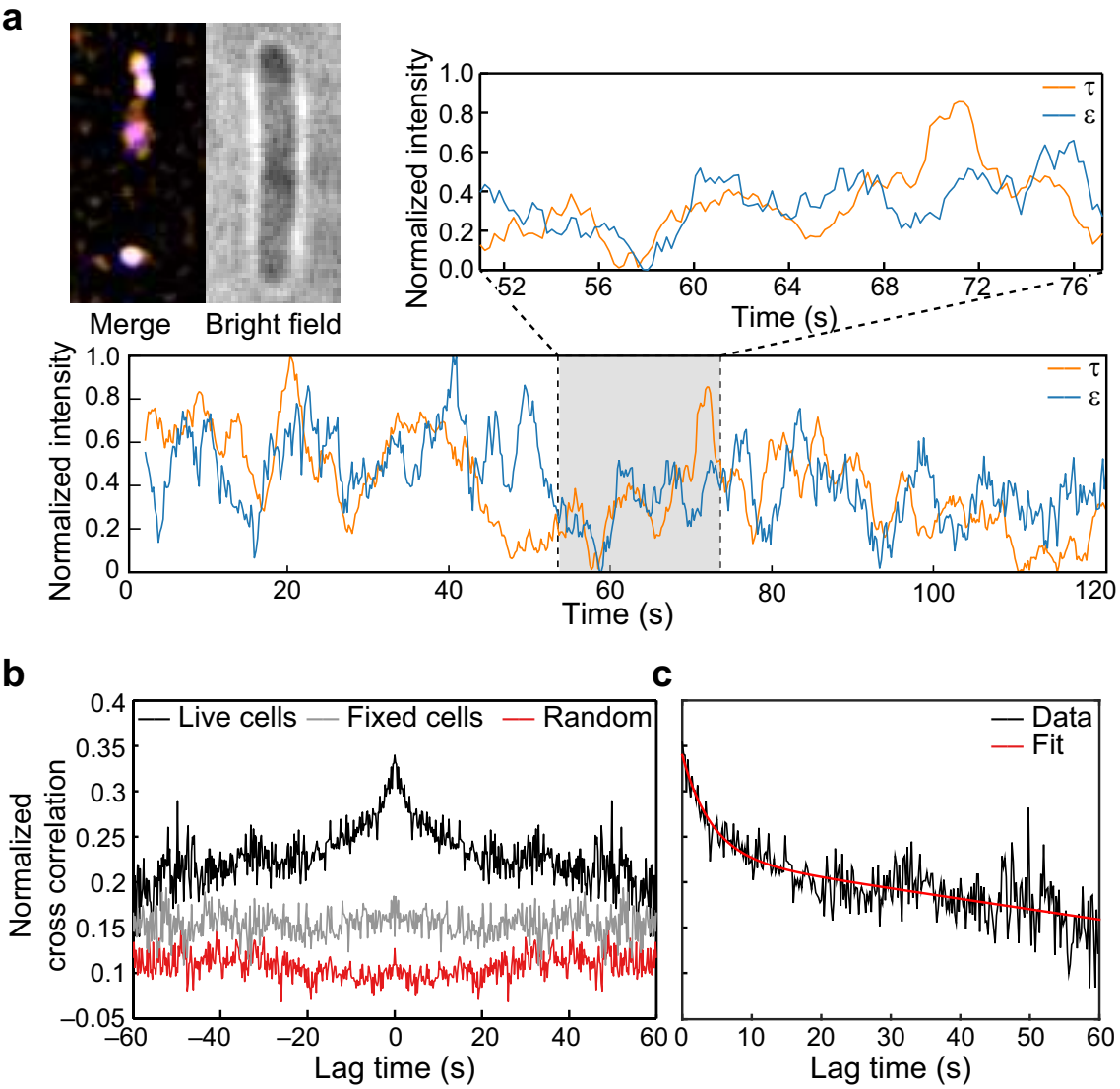


Figure 5.6: Visualization of Pol III* exchange *in vivo*. (a) Left: image of τ (orange) and ϵ (blue) foci within a single E. coli cell, averaged over 40 s. Co-localization of the two signals is shown as a white spot. Middle: bright field image of the same cell. Right and below: fluorescence intensity of τ (orange) and ϵ (blue) over time. The trajectories are averaged using a 2-s moving average filter. (b) Averaged, normalized cross-correlation functions. The cross-correlation function of 1210 pairs of foci in living cells shows a clear positive peak (black line). The cross-correlation function for 297 pairs of foci in fixed cells (grey line) and the cross-correlation function of 1210 pairs of foci, randomized within the same field of view (red line) show no positive cross correlation. Cross-correlation functions have been vertically offset for clarity. (c) Exponential fit (red) to the cross-correlation function in (b). We obtained an exchange time scale of τ = 4 $\pm$ 2 s. The error represents the error of the fit.

5.3 Discussion

We conclude that the *E. coli* replisome strikes a balance between stability and plasticity. In the absence of Pol III* in solution, it retains its original polymerase and forms a highly stable complex resistant to dilution. In its presence, Pol III*

readily exchanges into the replisome at a rate that is dependent on its concentration. Such a concentration dependent dissociative mechanism seems counterintuitive, but can be rationalized through a complex protein–protein and protein–DNA interaction network controlled and maintained by multiple dynamic interactions. Under dilute conditions, transient disruption of any one of these interactions would be followed by its rapid re-formation, preventing dissociation. If, however, there are competing Pol III*s in close proximity to the fork, one of these can bind at a transiently vacated binding site (e.g., on the β_2 sliding clamp or DnaB helicase) and consequently be at a sufficiently high local concentration to compete out the original Pol III* for binding to the other sites. Such concentration dependent exchange has recently been reported for other systems (50, 109, 137, 138, 209) and mathematically described by multisite competitive exchange mechanisms (140, 141). Further evidence for a multisite mechanism comes from comparison of the number of Pol III*s in or near the replisome at different concentrations. We quantified the number of Pol III*s at the replication fork in the *in vitro* experiments by normalizing the fluorescence intensity of the replisomal spot to the intensity of a single Pol III* (Figure 5.13). The peaks of the distributions are at one Pol III* per replisome (Figure 5.4F), consistent with *in vivo* observations (108). Nevertheless, we find that often more than one Pol III* is present in the replisome. As its concentration increases, the binding equilibria are pushed towards occupancy of all binding sites and more than one Pol III* is associated with the replisome. At lower concentrations, Pol III* still exchanges, but the average number of Pol III*s is reduced. Our observation of Pol III* exchange in living cells shows that such a multisite exchange mechanism is a physiologically relevant pathway accessible to the replisome during coupled DNA replication. Such a mechanism may have direct implications to the mechanisms used by the replisome to deal with obstacles such as DNA damage and transcription. The ability for the replisome to rapidly exchange components in the presence of competing factors in a concentration-dependent manner could allow for components to be easily replaced from solution and provide frequent but limited access to other binding partners, such as translesion synthesis polymerases (142), without violating fundamental chemical and thermodynamic principles.

5.4 Materials and Methods

5.4.1 Protein expression and purification

E. coli DNA replication proteins were produced as described previously: the β_2 sliding clamp (210), SSB (211), the $DnaB_6(DnaC)_6$ helicase–loader complex (212), DnaG primase (213), the Pol III $\tau_n\gamma_{(3n)}\delta\delta'\chi\psi$ clamp loader (202) and Pol III $\alpha\epsilon\theta$ core (202). Highly purified *E. coli* Pol I and DNA ligase A were gifts of Yao Wang (214).

5.4.2 Expression plasmids

Construction of plasmid pBOB1 (encoding w.t. α): The *tac* promoter plasmid pND517 contains the *dnaE* gene between a pair of *Bam*HI restriction sites (215). In addition, the *Bam*HI site following the gene overlaps with an *Nco*I site such that previous digestion with *Nco*I eliminates it. To incorporate an *Nde*I site at the start codon of *dnaE*, pND517 was used as template for PCR amplification of the 5′-portion of *dnaE* gene using primers 671(5′-AAAA*GGATCC*TAAGGAGGTTTG*CATATG*TCTGAACCACGTTTC; the *Bam*HI and *Nde*I sites are italicized, ribosome-binding sites are underlined) and 673 (5′-CGTTTGGCGATCTCAACGGTGT-3′). The PCR product (Fragment I; 522 bp) was isolated from an agarose gel following digestion with *Bam*HI and *Xho*I. Next, pND517 was digested with *Nco*I, and the purified linearized product digested independently with *Xho*I to generate Fragment II (3063 bp) and with *Bam*HI to yield Fragment III (5129 bp). Fragments I–III were ligated to yield pBOB1. *Construction of plasmid pJSL2197 (encoding SNAP-α):* A modified *snap 26 b* gene was amplified from pSNAP-tag(T7)-2 (New England Biolabs) by strand overlap PCR. In the first PCR, an *Nde*I site was incorporated at the start codon and an internal *Mlu*I site was removed by silent mutation using primers 728 (5′-AAAAAAAACATATGGACAAAGATTGCGAA) and 729 (5′-TGAAAATAGGCGTTCAGCGCGGTCGCC), yielding Fragment I. A second PCR used primers 730 (5′-TGGCTGAACGCCTATTTTCATCA GCCGGAAGC) and 732 (5′-AAAA*GGATCC*GATAGAGCCAGACTC*ACG CGT*TCCCAGACCCGG-3′) to generate Fragment II, removing the TGA stop codon and incorporating a sequence encoding a flexible peptide linker (sequence: TRESGSIGS (216)) flanked by *Mlu*I and *Bam*HI sites at the 3′ end of *snap26b*Δ *MluI*. Equimolar amounts of isolated Fragments I and II were then used as templates for PCR with the outside primers 728 and 732 to generate a product that was digested with *Nde*I and *Bam*HI and isolated from a gel. This fragment was ligated with the 502 bp *Bam*HI–*Xho*I fragment of pKO1479wt (217) encoding the N-terminal segment of α and the large *Nde*I–*Xho*I fragment of pBOB1, encoding the remainder of α, to generate pJSL2197, which directs overproduction of SNAP-α. Plasmid constructions were confirmed by nucleotide sequence determination.

5.4.3 Expression and purification of SNAP-α

An affinity resin for purification of full-length (unproteolysed) α was prepared by conjugation of biotinylated τ_C16 (α-binding domain V of τ) (205) to high-capacity streptavidin-agarose (Pierce Biotechnology). Biotinylated τ_C16 (15 ml; 12 mg) was added dropwise with gentle stirring into a suspension of 6 ml of resin in 11 ml of 50 mM Tris–HCl pH 7.6, 2 mM dithiothreitol, 1 mM EDTA, 50 mM NaCl at 6°C over 20 min. Unconjugated streptavidin-agarose resin (2 ml) was added to a column and allowed to settle, then the suspension of τ_C16-conjugated resin was poured over it. The column (1 $\times$ 10 cm) was then washed with 150 ml of 50 mM Tris–HCl pH 7.6, 2 mM dithiothreitol, 1 mM EDTA, 50 mM NaCl and stored at 4°C in 50 mM Tris–HCl, 5 mM dithiothreitol, 1 mM EDTA, 50 mM NaCl, 0.03% NaN_3. *E. coli* strain BL21(λDE3)*recA*/pJSL2197 was grown at 30°C in LB medium sup-

plemented with thymine (25 mg l^{-1}) and ampicillin (100 mg l^{-1}). Upon growth to A_{600} = 0.6, 1 mM isopropyl-β-D-thiogalactoside (IPTG) was added and cultures were shaken for a further 3.5 h, then chilled in ice. Cells (30 g from 6 l of culture) were harvested by centrifugation, frozen in liquid N_2 and stored at -80°C. After thawing, cells were lysed and SNAP-α was purified through Fraction IV essentially as described for wild-type α (215). Fraction IV (50 ml) was dialysed against two changes of 2 l of buffer Cα (25 mM Tris–HCl pH 7.6, 2 mM dithiothreitol, 1 mM EDTA, 10% (v/v) glycerol) and applied at 1 ml min^{-1} onto a column (2.5 $\times$ 12 cm) of heparin-Sepharose (215) that had been equilibrated with buffer Cα. The column was washed with 30 ml of buffer Cα and proteins were eluted using a linear gradient (150 ml) of 0–400 mM NaCl in buffer Cα. SNAP-α eluted as a single peak at $\sim$40 mM NaCl. Fractions were collected and pooled to yield Fraction V, which was applied directly at 1 ml min^{-1} onto the column of τC16-agarose affinity resin that had been equilibrated in buffer Dα (50 mM Tris–HCl pH 7.6, 10 mM dithiothreitol, 1 mM EDTA, 5% (v/v) glycerol) containing 20 mM $MgCl_2$. After the column had been washed with 15 ml of buffer Dα + 0.6 M MgCl2 and unbound proteins had been washed away, SNAP-α was eluted using a linear gradient (20 ml) of 0.6–4.0 M $MgCl_2$ in buffer Dα. SNAP-α eluted as a single peak at $\sim$2.8 M $MgCl_2$ (Figure 5.7). Fractions under the peak were immediately pooled and dialysed against two changes of 2 l of buffer Eα (50 mM Tris–HCl pH 7.6, 1 mM EDTA, 3 mM dithiothreitol, 100 mM NaCl, 20% (v/v) glycerol) to give Fraction VI (40 ml, containing 68 mg of protein; Figure 5.7). Aliquots were frozen in liquid N_2 and stored at –80°C.

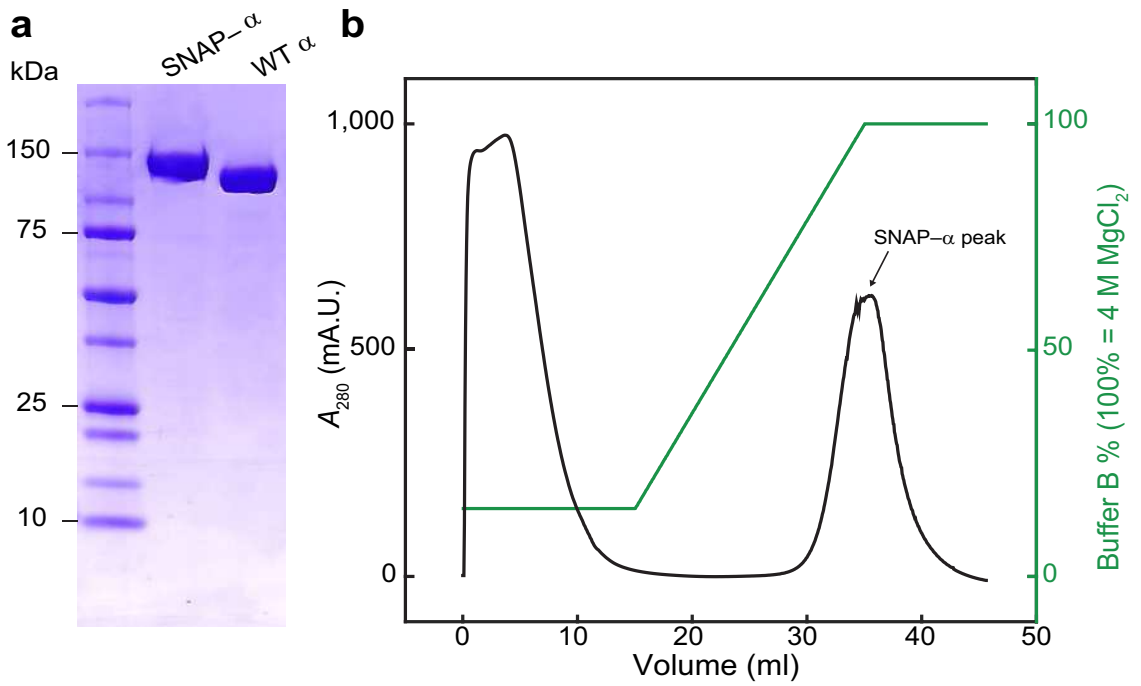


Figure 5.7: Separation of proteolytic fragments of SNAP-α from full-length SNAP-α. (a) SDS-PAGE of final fraction from the τC16 affinity chromatography, pooled from successive samples from the peak in the chromatography profile in (b).

5.4.4 Fluorescent labeling of SNAP-α

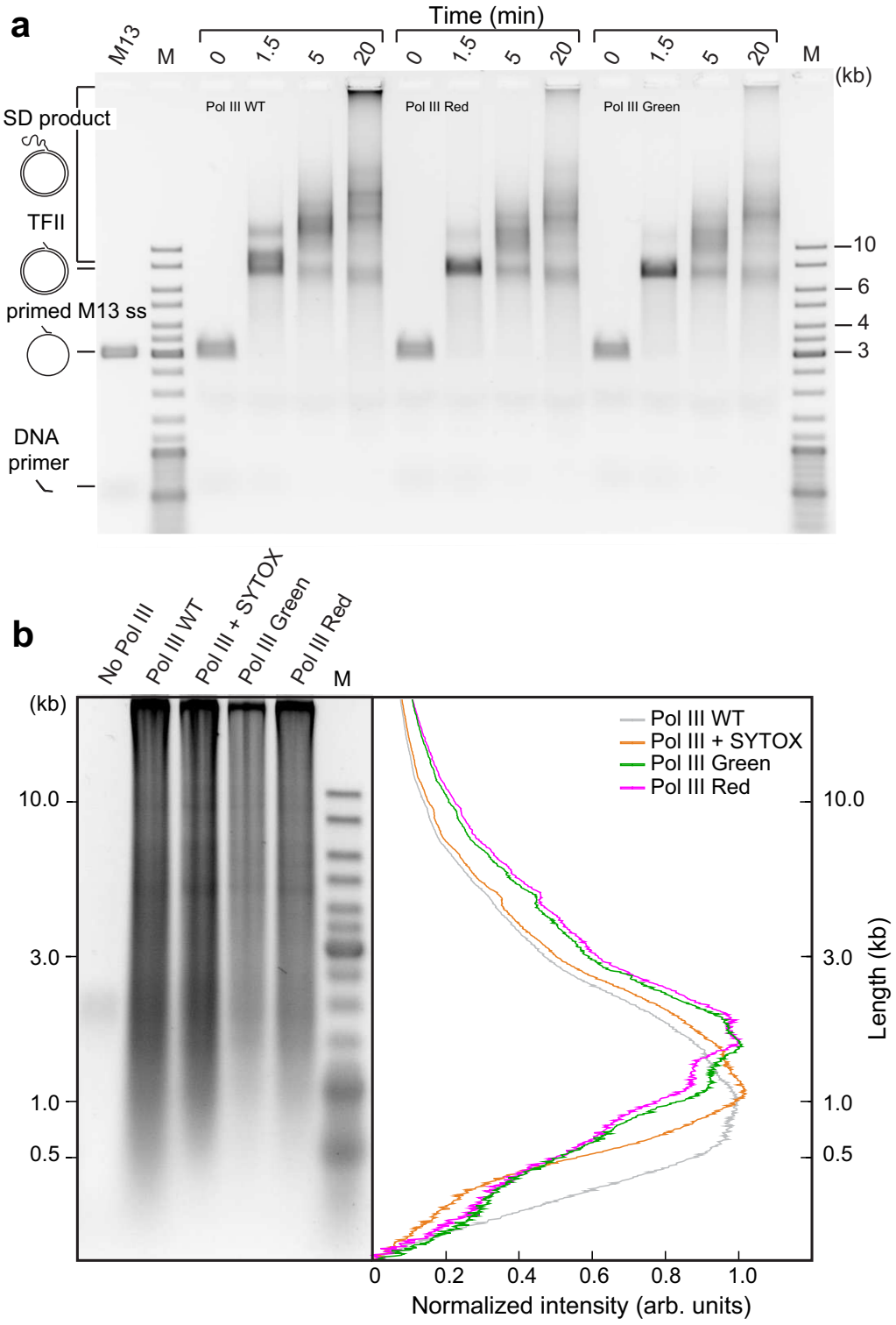
Two different fluorescent probes, SNAP-Surface 649 (red) and SNAP-Surface Alexa Fluor 488 (green; New England Biolabs), were used to label SNAP-α. All labeling reactions were carried out using a 2-fold molar excess of dye with 27 μM SNAP-α in 1 ml of 50 mM Tris–HCl pH 7.6, 2 mM dithiothreitol, 100 mM NaCl, 5$\%$ (v/v) glycerol (buffer Fα) for 2 h at 23°C, followed by 6°C overnight with gentle rotation. Following the coupling, the reaction mixture was supplemented with 1 mM EDTA and excess dye was removed by gel filtration at 1 ml min^{-1} through a column (1.5 ÃŮ 10 cm) of Sephadex G-25 (GE Healthcare) in buffer Fα + 1 mM EDTA. Fractions containing the labeled SNAP-α were pooled and dialysed against 2 l of buffer Eα, frozen in liquid N_2 and stored in aliquots at –80°C. The degree of labeling was measured to be 90$\%$ for SNAP-α649 and 83$\%$ for SNAP-α488 by UV/vis spectrophotometry.

5.4.5 Ensemble strand-displacement DNA replication assays

The flap-primed ssDNA template was made as previously described (212). Conditions for the standard coupled strand extension and Pol III strand-displacement (SD) reaction were adapted from described methods (212). Briefly, reactions contained 2.5 nM primed DNA template, 1 mM ATP, 0.5 mM of each dNTP, 30 nM $\tau_3\delta\delta'\chi\psi$, 150 nM Pol III (wild-type or SNAP-labeled), 200 nM β_2 and 800 nM SSB_4 in 25 mM Tris–HCl pH 7.6, 10 nM $MgCl_2$, 10 mM dithiothreitol and 130 mM NaCl, in a final volume of 13 mul. Components (except DNA) were mixed and treated at room temperature, then cooled in ice for 5 min before addition of DNA. Reactions were initiated at 30°C, and quenched at time points by addition of EDTA to $\sim$100 mM and SDS to $\sim$1$\%$. Products were separated by agarose gel electrophoresis and stained with SYBR-Gold (Invitrogen) (Figure 5.8a).

5.4.6 Ensemble leading and lagging strand DNA replication assays

Coupled leading and lagging strand DNA synthesis reactions were set up in replication buffer (25 mM Tris–HCl pH 7.9, 50 mM potassium glutamate, 10 mM $Mg(OAc)_2$, 40 μg/ml BSA, 0.1 mM EDTA and 5 mM dithiothreitol) and contained 1.5 nM of a 2-kb circular dsDNA template, 1 mM ATP, 250 μM CTP, GTP, and UTP, and 50 μM dCTP, dGTP, dATP, and dTTP, 6.7 nM wild-type or SNAP-labeled Pol III*, 30 nM β_2, 300 nM DnaG, 100 nM SSB_4, and 30 nM $DnaB_6(DnaC)_6$ in a final volume of 12 mul. Components (except DNA) were mixed and treated at room temperature, cooled in ice for 5 min before addition of DNA. Reactions were initiated at 30°C, and quenched after 30 min by addition of 7 μl 0.5 M EDTA and 6 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). The quenched mixtures were loaded into a 0.6$\%$ (w/v) agarose gel in alkaline running buffer (50 mM NaOH, 1 mM EDTA). Products were separated by agarose gel electrophoresis at 14 V for 14 h. The gel was then neutralized in 1 M Tris–HCl, pH 7.6, 1.5 M NaCl and stained with SYBR Gold. Okazaki fragment length distribution was calculated by


a
Time (min)
M13
M
0
1.5
5
20
Pol III WT
Pol III Red
Pol III Green
(kb)
SD product
TFII
primed M13 ss
DNA primer
10
6
4
3
b
No Pol III
Pol III WT
Pol III + SYTOX
Pol III Green
Pol III Red
(kb)
10.0
3.0
1.0
0.5
Pol III WT
Pol III + SYTOX
Pol III Green
Pol III Red
Length (kb)
0
0.2
0.4
0.6
0.8
1.0
Normalized intensity (arb. units)

normalizing the intensity as a function of DNA length (Figure 5.8b). Conditions for testing the stability of the α–τ interaction in Pol III* during replication were performed as above with modifications. First, 125 μg streptavidin-coupled magnetic beads (Invitrogen) were washed and equilibrated in replication buffer containing 200 μM AMP-PNP (replication buffer B). $DnaB_6(DnaC)_6$ was first loaded at the fork by incubation of 7.5 nM rolling-circle DNA and 75 nM $DnaB_6(DnaC)_6$ (reaction A) at 37°C for 5 min in replication buffer B (40 μl), before being immobilized on streptavidin-coupled magnetic beads for 30 min at room temperature (reaction B). Unbound DNA was removed by washing reaction B three times in 200 μl replication buffer B. Replication was initiated by resuspending reaction B in replication buffer containing 1.25 mM ATP, 250 μM CTP, GTP and UTP, 200 μM dCTP, dGTP, dATP and dTTP, 3.35 nM each of red and green labeled Pol III, 200 nM β2, 300 nM DnaG, 50 nM SSB_4, and 30 nM $DnaB_6(DnaC)_6$, and allowed to proceed for 20 min at 37°C. Reactions were quenched with 2.1 μl 2.5 M NaCl and 5 μl 0.5 M EDTA. Following quenching, the supernatant was removed, diluted 100-fold in replication buffer then imaged on the surface of a coverslip at the single-molecule level. The remaining DNA products coupled to the beads were washed three times in replication buffer, then resuspended in replication buffer (23 μl) and 7 μl DNA loading dye then heated to 70°C for 5 min. The DNA was loaded onto the alkaline agarose gel, which was run under the same conditions as before (Figure 5.9).

Figure 5.8 *(preceding page)*: Comparison of activities of wild-type and SNAP-labeled Pol III cores. (a) Agarose gel of products of Pol III strand-displacement (SD) DNA synthesis, a demanding assay for Pol III* activity (212). The time course of flap-primer extension on M13 ssDNA depicts products larger than unit length of dsDNA (TFII) products generated by SD DNA synthesis. (b) Alkaline agarose gel of coupled DNA replication. Reactions were performed on a 2-kb circular dsDNA template with wild-type (WT) Pol III*, WT Pol III* + SYTOX Orange, red SNAP-labeled Pol III*, and green SNAP-labeled Pol III*. (Left panel) The gel was stained with SYBR-Gold. (Right panel) Intensity profiles of lanes 2–5 of the left panel. The Okazaki fragment size distribution is centred at 1.3 $\pm$ 0.4 kb. Intensity profiles have been corrected for the difference in intensity of different size fragments using the ladder as a standard.

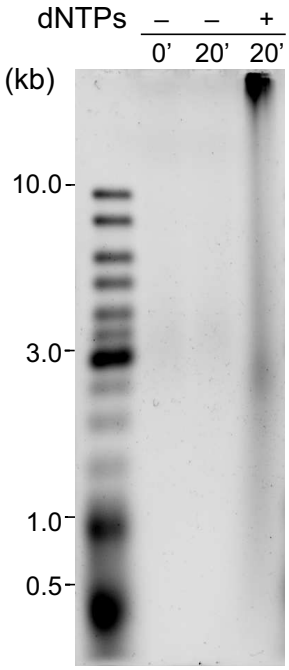


Figure 5.9: Alkaline gel showing leading- and lagging-strand products using pre-assembled red and green Pol III*s. Reactions were performed on a 2-kb circular dsDNA template without dNTPs (lanes 1 and 2) and with dNTPs (lane 3). Lane 3 shows long leading strand and shorter lagging strand products are generated after 20 min; the leading strand products remain bound to beads in the well. The gel was stained with SYBR-Gold.

5.4.7 *In vitro* single-molecule rolling-circle DNA replication assay

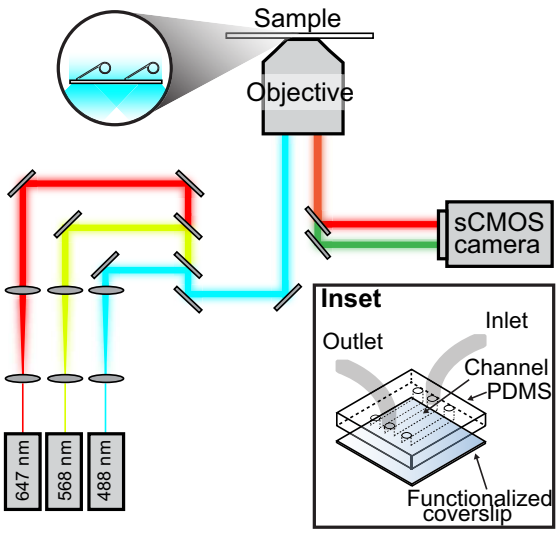


Figure 5.10: Schematic overview of the single-molecule fluorescence microscope. Laser light of a specific wavelength is coupled into the microscope objective. The fluorescence signal from the sample is detected with a EMCCD or sCMOS camera. (**Inset**) Micro-fluidic flow cell schematic. A PDMS lid containing three flow chambers is placed on top of a PEG-biotin-functionalized microscope coverslip. Tubing is inserted into 1 mm holes in the PDMS.

To construct the rolling circle template (14), a 66-mer 5′-biotin-T36AATTC GTAATCATGGTCATAGCTGTTTCCT-3′ (IDT) was annealed to M13 mp18 ssDNA (NEB) in TBS buffer (40 mM Tris–HCl pH 7.5, 10 mM MgCl2, 50 mM NaCl) at 65°C. The primed M13 was then extended by adding 64 nM T7 polymerase gp5 (New England Biolabs) in 40 mM Tris–HCl pH 7.6, 50 mM potassium glutamate, 10 mM $MgCl_2$, 100 μg ml^{-1} BSA, 5 mM dithiothreitol and 600 μM dCTP, dGTP, dATP and dTTP at 37°C for 60 min. The reaction was quenched with 100 mM

EDTA and the DNA was purified using a PCR Purification Kit (Qiagen). Microfluidic flow cells were prepared as described (51). Briefly, a PDMS flow chamber was placed on top of a PEG-biotin-functionalized microscope coverslip (Supplementary Figure 5.10 inset). To help prevent non-specific interactions of proteins and DNA with the surface, the chamber was blocked with buffer containing 20 mM Tris–HCl pH 7.5, 2 mM EDTA, 50 mM NaCl, 0.2 mg/ml BSA, and 0.005% Tween-20. The chamber was placed on an inverted microscope (Nikon Eclipse Ti-E) with a CFI Apo TIRF 100x oil-immersion TIRF objective (NA 1.49, Nikon) and connected to a syringe pump (Adelab Scientific) for flow of buffer. Conditions for coupled DNA replication under continuous presence of all proteins were adapted from previously described methods (14, 202). Briefly, 30 nM $DnaB_6(DnaC)_6$ was incubated with 1.5 nM biotinylated ds M13 substrate in replication buffer (25 mM Tris–HCl pH 7.9, 50 mM potassium glutamate, 10 mM $Mg(OAc)_2$, 40 μg ml^{-1} BSA, 0.1 mM EDTA and 5 mM dithiothreitol) with 1 mM ATP at 37°C for 30 s. This mixture was loaded into the flow cell at 100 μl min^{-1} for 40 s and then at 10 μl min^{-1}. An imaging buffer was made with 1 mM UV-aged Trolox, 0.8% (w/v) glucose, 0.12 mg ml^{-1} glucose oxidase, and 0.012 mg ml^{-1} catalase (to increase the lifetime of the fluorophores and reduce blinking), 1 mM ATP, 250 μM CTP, GTP, and UTP, and 50 μM dCTP, dGTP, dATP, and dTTP in replication buffer. Pol III* was assembled in situ by incubating $\tau_3\delta\delta'\chi\psi$ (410 nM) and SNAP-labeled Pol III cores (1.2 μM) in imaging buffer at 37°C for 90 s.

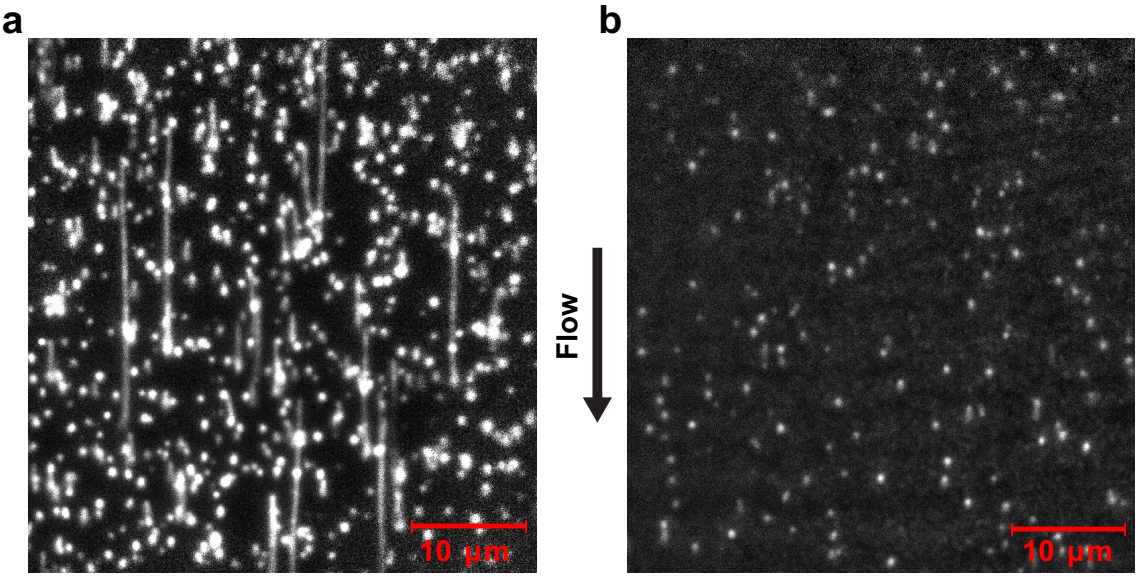


Figure 5.11: Representative field of view of SYTOX Orange-stained dsDNA from the single-molecule rolling-circle DNA replication assay. (a) Efficient DNA replication proceeds in the presence of the full complement of replication reaction mix, including the reconstituted *E. coli* replisome, NTPs and dNTPs. Note both the length and number of products. (b) No DNA products are evident in the entire flow cell in the absence of dNTPs from the replication reaction mix. Note some rolling-circle templates become linearized due to photodamage, visible as lines shorter than 7 kb (2.1 μm).

Replication was initiated by flowing in the imaging buffer containing 6.7 nM Pol III* (unless specified otherwise), 30 nM β_2, 300 nM DnaG, 250 nM SSB_4, and 30

nM $DnaB_6(DnaC)_6$ at 10 μl min^{-1}. Reactions were carried out 31°C, maintained by an electrically heated chamber (Okolab). Double-stranded DNA was visualized in real time by staining it with 150 nM SYTOX Orange (Invitrogen) excited by a 568-nm laser (Coherent, Sapphire 568-200 CW) at 150 μW cm^{-2} (Figures 5.10 and 5.11). The red and green Pol III* were excited at 700 mW cm^{-2} with 647 nm (Coherent, Obis 647-100 CW) and 488 nm (Coherent, Sapphire 488-200 CW) lasers, respectively (Figures 5.12 and 5.19). The signals were separated via dichroic mirrors and appropriate filter sets (Chroma). Imaging was done with either an EMCCD (Photometics, Evolve 512 Delta) or sCMOS camera (Andor, Zyla 4.2). The analysis was done with ImageJ using in-house built plugins. The rate of replication of a single molecule was obtained from its trajectory and calculated for each segment that has constant slope. Conditions for the pre-assembly replication reactions were adapted from published methods (136, 204). Solution 1 was prepared as 30 nM $DnaB_6(DnaC)_6$, 1.5 nM biotinylated ds M13 substrate and 1 mM ATP in replication buffer. This was incubated at 37°C for 3 min. Solution 2 contained 60 μM dCTP and dGTP, 6.7 nM red Pol III*, and 74 nM β_2 in replication buffer (without dATP and dTTP). Solution 2 was added to an equal volume of solution 1 and incubated for 6 min at 37°C. This was then loaded onto the flow cell at 100 μl min^{-1} for 1 min and then 10 μl min^{-1} for 10 min. The flow cell was washed with replication buffer containing 60 μM dCTP and dGTP. Replication was finally initiated by flowing in the imaging buffer containing 50 nM β_2, 300 nM DnaG and 250 nM SSB_4 at 10 μl min^{-1}.

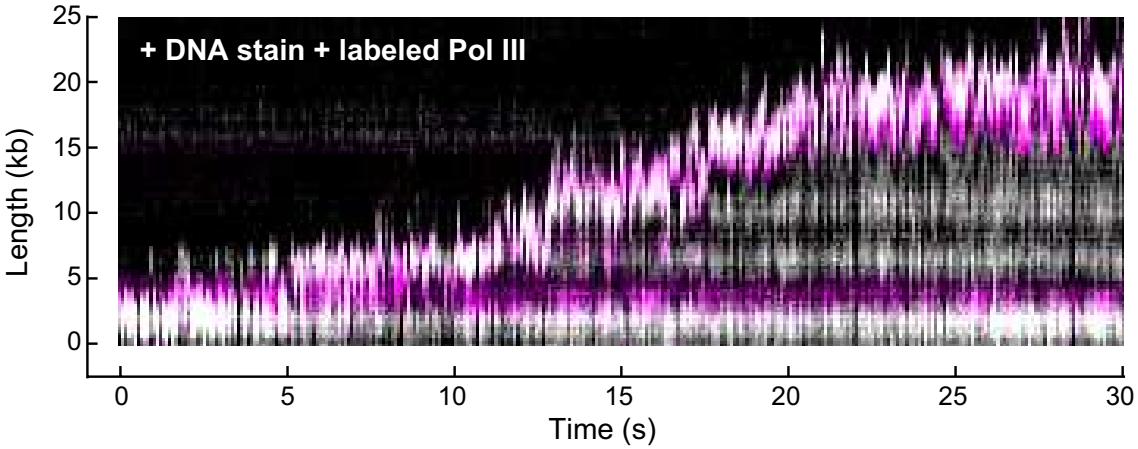


Figure 5.12: Representative kymograph of simultaneous staining of double-stranded DNA. Kymograph of SYTOX Orange visualisation (grey scale) and fluorescence imaging of Pol III labeled with a red fluorophore (magenta) in real time. The kymograph demonstrates the fluorescent spot corresponding to Pol III co-localizes with the tip of the growing DNA product (evident as a white spot) where the replication fork is located.

5.4.8 Measurement of the stoichiometry of Pol III*s at the replisome.

The average intensity of a single labeled Pol III core (6 pM) was calculated by immobilization on the surface of a cleaned microscope coverslip in imaging buffer. The imaging was under the same conditions as used during the single-molecule

rolling-circle experiments. Using ImageJ with in-house built plugins, we calculated the integrated intensity for every Pol III core in a field of view after applying a local background subtraction. The histograms obtained were fit with a Gaussian distribution function using MATLAB 2014b, to give a mean intensity of 5100 $\pm$ 2000 for the red Pol III core and 1600 $\pm$ 700 for the green Pol III core (Figure 5.13).

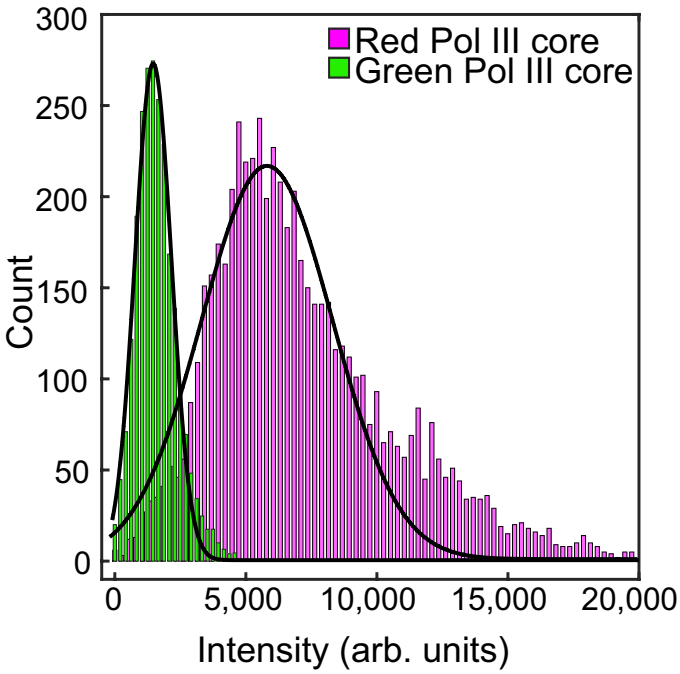


Figure 5.13: Histograms of the intensity distribution of single Pol III cores. The histograms are fit with Gaussian distribution functions to give a mean intensity of 5100 $\pm$ 2000 for the red Pol III core and 1600 $\pm$ 700 for the green Pol III core. The errors represent the standard errors of the mean.

To measure the intensity of the fluorescent spot at the replication fork, we tracked its position and integrated the intensity for both colors simultaneously over time. Given there is no decay in fluorescence intensity of labeled Pol III cores as a function of DNA length under near-TIRF imaging conditions during DNA replication (Figure 5.14), we calculated the total number of Pol III*s at every time point during coupled DNA replication by dividing these intensities by the intensity of a single Pol III*. Subsequent histograms were fit to four (6.7 nM) or three (0.3 nM) Gaussians centred at integral numbers of Pol III* (Figure 5.15) using MATLAB 2014b.

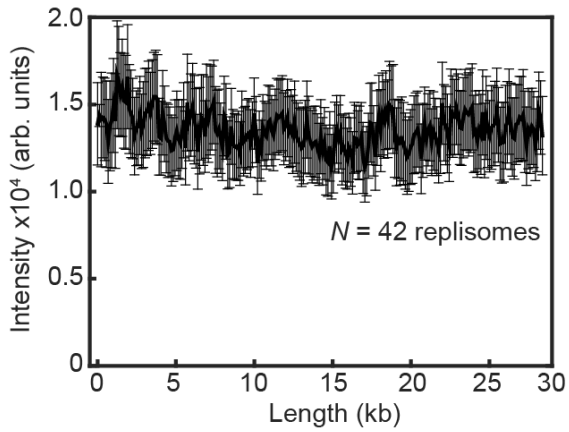


Figure 5.14: Fluorescence intensity of replicating Pol III* complexes does not change at longer DNA lengths under near-TIRF imaging conditions. The fluorescence intensity of labeled Pol III* complexes does not change as a function of DNA length during single-molecule rolling-circle DNA replication under constant flow. The errors represent the standard deviation.

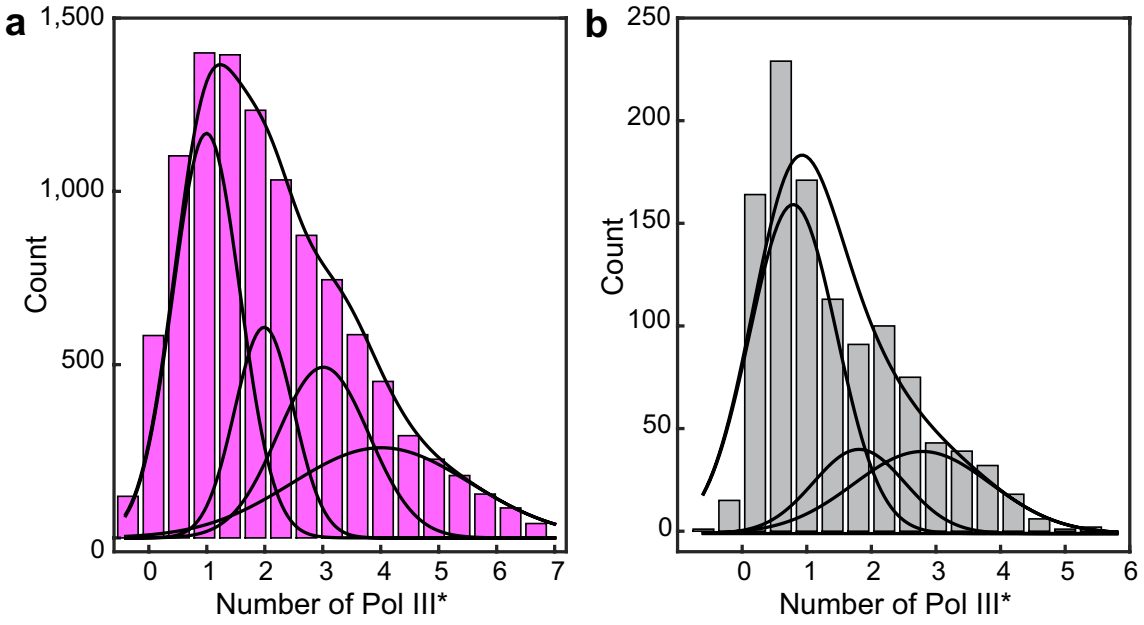


Figure 5.15: Histograms of the stoichiometry of Pol III* at the replication fork. (a) Intensity distribution at 6.7 nM Pol III* and (b), intensity distribution at 0.3 nM Pol III*. The histograms are fit with either four (6.7 nM) or three (0.3 nM) Gaussian distribution functions centred at integral numbers of Pol III*.

5.4.9 Fluorescent chromosomal fusions.

The strain EAW192 (*dnaQ-mKate2*) was constructed using a modified version of the λ RED recombination system (218), introducing a mutant FRT–Kan^R–wtFRT cassette. To select for recombinants, cells were plated on LB-agar supplemented with 40 μg ml^{-1} of kanamycin and grown over- night. Kanamycin-resistant strains were further screened for ampicillin sensitivity, to ensure that cells had been cured of the λ RED plasmid pKD46. The two-color strain EAW203 (*dnaX-YPet*, *dnaQ-mKate2*) was constructed by P1 transduction. JJC5945 cells (*dnaX-YPet*) (110) were treated with pLH29 (218) first to remove existing Kan^R markers, then infected with P1 grown on EAW192 (*dnaQ-mKate2*) cells. Positive transductants were isolated by selecting for kanamycin resistance.

5.4.10 Growth rates of fluorescent chromosomal fusions.

Single colonies of wild-type *E. coli* MG1655 and derivatives containing the C-terminal chromosomal dnaQ and dnaX fusions were used to inoculate 5 ml of LB broth (with 34 μg mL^{-1} kanamycin, if required) and grown at 37°C with shaking overnight. LB broth (100 ml) was inoculated with 1.0 $\cdot$ 10^5 cells ml^{-1} from overnight cultures. Subsequent growth of each strain was monitored at 37°C with shaking by determining OD_{600} every 30 min for 9.5 h. The doubly labeled *dnaX-YPet dnaQ-mKate2* cells grew only slightly slower than wild-type cells (division time = 33 $\pm$ 4 min cf. 33 $\pm$ 8 min) (Figure 5.16), indicating that labeling the ϵ and τ components of the replisome does not significantly disrupt DNA replication.

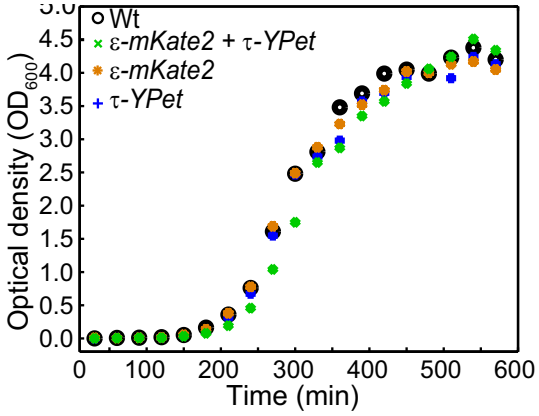


Figure 5.16: Growth curves for *E. coli* strains. Wild-type *E. coli* (black), cells expressing both C-terminal derivatives of τ (dnaX-YPet) and ϵ (*dnaQ-mKate2*) subunits under control from their endogenous promoters (green), and cells expressing only *dnaX-YPet* (blue) and *dnaQ-mKate2* (orange). Growth curves were measured for 9.5 h. The division times were obtained from a linear fit of the exponential growth phase. They are 33 $\pm$ 8 min for wild-type, 32 $\pm$ 5 min for *dnaX-YPet*, 32 $\pm$ 8 min for *dnaQ-mKate2*, and 33 $\pm$ 4 min for the double mutant. The errors represent the errors of the fit.

5.4.11 *In vivo* single-molecule visualization assays.

The cells were grown at 37°C in EZ rich defined medium (Teknova) that included 0.2% (w/v) glucose. For imaging, a PDMS well was placed on top of a coverslip that was functionalized with 3-aminopropyl triethoxy silane (BioScientific) (110). The cells were immobilized on the surface of the well, which was then placed on the heated stage (Okolab) of the microscope. Imaging was done at 37°C. The τ-YPet and ϵ-mKate2 were excited at 0.03 mW cm^{-2} with 514 nm (Coherent, Sapphire 514-150 CW) and 3 W cm^{-2} with 568 nm (Coherent, Sapphire 568-200 CW) lasers, respectively. The signals were separated via a beam splitter (Photometrics, DVΛ Multichannel Imaging System) and appropriate filter sets (Chroma). Imaging was done with an EMCCD camera (Photometrics, Evolve Delta). The image processing was done with ImageJ using in-house built plugins. The concentrations for ϵ and τ were determined as described (110) by measuring the total fluorescence within each cell and dividing by the mean intensity of a single molecule. To measure the total fluorescence, we first imaged 141 wild-type MG1655 cells to determine the cellular autofluorescence (Figure 5.17A). We found the autofluorescence to be constant for the duration of our measurements. We then imaged 273 τ-YPet, ϵ-mKate2 cells in 20 fields of view. The background fluorescence from the microscope coverslip was determined by fitting the photobleaching per field of view with a single-exponential decay (Figure 5.17B). The mean cellular intensities were corrected for the background and cellular autofluorescence. We then fitted individual cell photobleaching curves to obtain the amplitudes (Figure 5.17c). To find the intensity of a single-molecule photobleaching trajectories of single foci were determined. The ϵ and τ foci were identified by making average projections of movies. The intensity over time trajectories for each focus as it photobleached was measured. Next, the local background fluorescence around each focus was subtracted. Trajectories showed step-wise intensity transitions corresponding to photobleaching of single fluorescent molecules (Figure 5.17d inset).

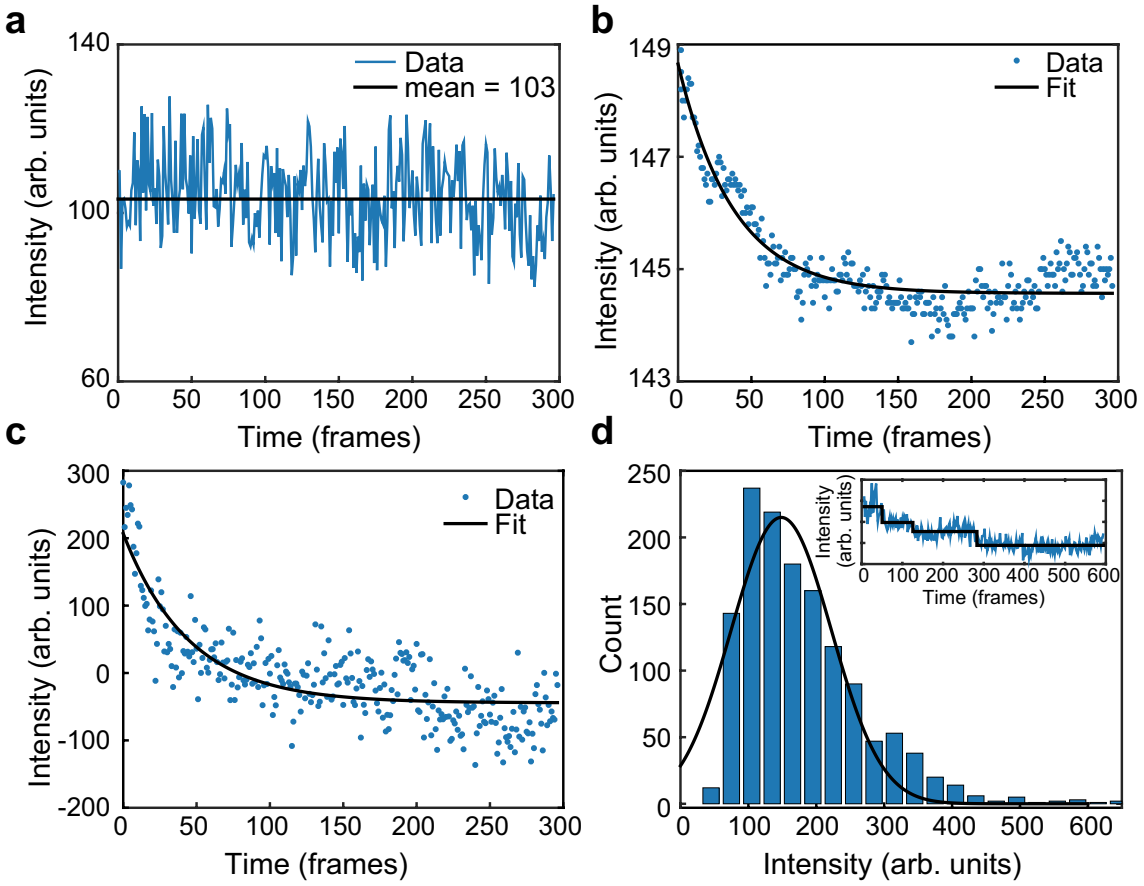


Figure 5.17: Measurement of concentrations of ϵ and τ in live cells. (**a**) Mean fluorescence signal during photobleaching of wild-type MG1655. (**b**) Bleaching of the coverslip background signal within a single field of view. (**c**) Bleaching of ϵ-mKate2 fluorescence within a single cell, corrected for the cellular autofluorescence (a) and the background fluorescence of the coverslip (b). This was fit with a single exponential decay (black line) to determine the maximum intensity. **(d)** Histogram of the single-molecule intensities obtained from the change-point step-fitting algorithm (inset). This was fit to a Gaussian distribution to find the mean intensity of a single mKate2 molecule.

These transitions were fit by change-point analysis (189), (52). A histogram of the step sizes, showed a relatively narrow distribution (Figure 5.17). We found the mean intensity of a single molecule by fitting with a Gaussian distribution. These were 158 $\pm$ 2 for ϵ and 130 $\pm$ 5 for τ (mean $\pm$ s.e.m). To find the total fluorescence intensity per cell, the mean cell intensity was multiplied by the area of the cell. This was then divided by the single-molecule intensity. It was determined that there are 104 $\pm$ 3 copies of ϵ and 96 $\pm$ 6 copies of τ per cell. This corresponds to a concentration of 72 $\pm$ 3 nM for ϵ and 67 $\pm$ 5 nM for τ. If we assume that all ϵ and τ are part of a Pol III* complex this tells us that there is 23 nM Pol III* in the cell.The intensities of the foci were measured by integrating the intensity of the peak and subtracting the mean local background intensity. The average cross-correlation functions were calculated using MATLAB 2014b (Mathworks). The cross-correlation was fit with an exponential decay and gave a characteristic time scale of 4 $\pm$ 2 s (mean $\pm$ error of the fit) 5.6.

5.5 Supplementary figures

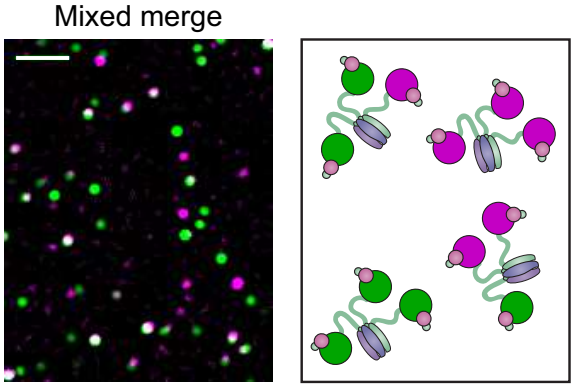


Figure 5.18: Pol III* complexes of mixed Pol III core composition show co-localization. (**Right**) Red and green Pol III cores are mixed before adding the CLC (30 nM Pol III core and 10 nM τ_3-CLC). Pol III* is formed by treatment at 37°C for 15 min. Complexes are then allowed to equilibrate for 1 hour at 37°C prior to dilution to 6 pM for imaging. (**Left**) Red and green Pol III cores co-localize (white spots). White scale bar represents 5 μm.

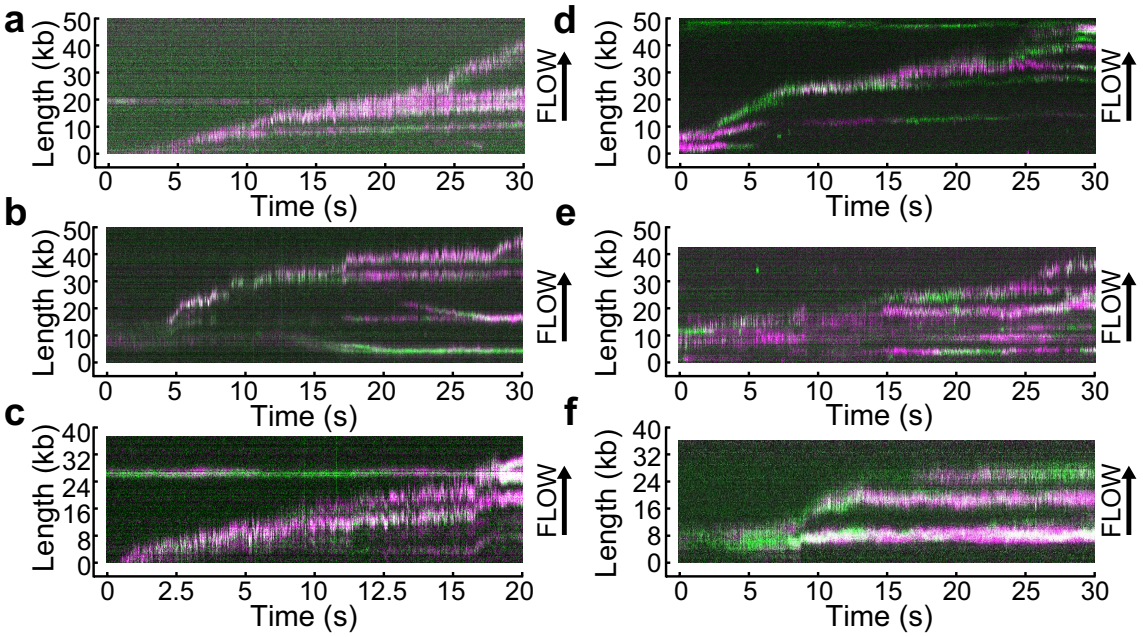


Figure 5.19: Example kymographs. Kymographs of the distributions of red Pol III* (magenta) and green Pol III* (green) on individual DNA molecules at a total Pol III* concentration of 6.7 (a–c) or 0.3 nM (d–f). Co-localization of the two signals is shown as a white fluorescent spot.

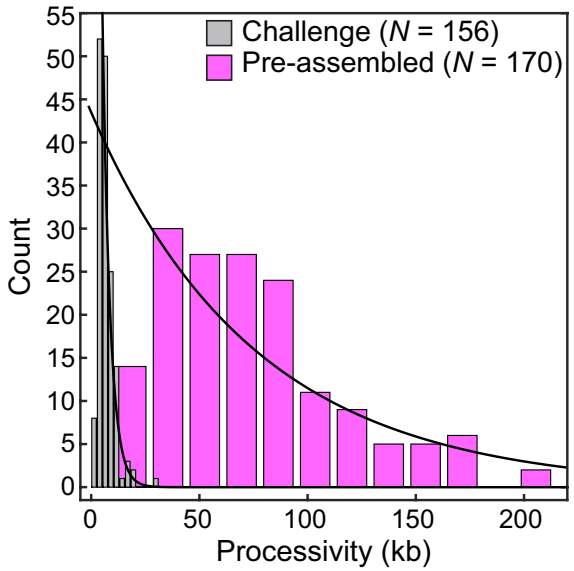


Figure 5.20: Histograms comparing pre-assembly conditions with chasing with Pol III core. Histograms for conditions with pre-assembled replisomes (no polymerases in solution) (76 $\pm$ 26 kb) and under conditions where pre-assembled replisomes are challenged with 10 nM Pol III core (3.5 $\pm$ 0.6 kb), each fit with a single exponential decay function. The data show that actively replicating Pol III* can be easily displaced when challenged with entities that bind to the replisome, but cannot support coordinated leading and lagging strand synthesis. The errors represent the errors to the fit.

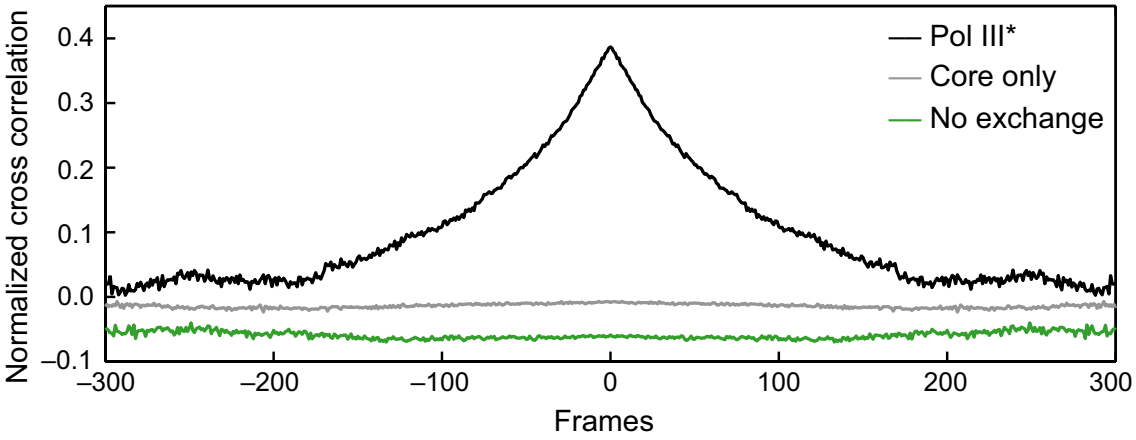


Figure 5.21: Cross-correlation analysis of simulated intensity trajectories for pairs of ϵ and τ foci. Individual intensity trajectories for 300 ϵ and τ foci were simulated in MATLAB 2014b. The simulation allows us to set k_{on} and k_{off} (in units of frames-1) for ϵ exchanging into Pol III*, and k_{on} and k_{off} for Pol III* exchanging into the replisome. By changing these rate constants we can simulate different exchange mechanisms. The black line represents the average cross-correlation function for Pol III* exchange (both ϵ and τ). Here k_{on} and k_{off} for Pol III* were set to 0.01 and the rate constants for ϵ were set to k_{on} = 1 and k_{off} «1 to simulate stable binding of core to Pol III*. A clear positive peak can be seen. The green line represents the average cross-correlation function for simulated trajectories without any exchange. Here kon = 1 and koff «1 for all rate constants, to simulate stable binding to the replisome. In this case there is no positive cross correlation. The grey line represents the average cross-correlation function for core exchange. In this case k_{on} =1 and k_{off} «1 for Pol III* and k_{on} and k_{off} for ϵ were set to 0.01. Again, we do not observe a positive cross correlation. Cross-correlation functions have been vertically offset for clarity.

6 | Single-molecule visualization of SSB dynamics shows a competition between an internal-transfer mechanism and external exchange.

Lisanne M. Spenkelink, Jacob S. Lewis, Slobodan Jergic, Zhi-Qiang Xu, Andrew Robinson, Nicholas E. Dixon, Antoine M. van Oijen.

Manuscript submitted to *Molecular Cell*.

Single-stranded DNA-binding proteins (SSBs) are integral to DNA replication by protecting single-stranded DNA (ssDNA) from nucleolytic attacks, preventing intra-strand pairing events, and playing many other regulatory roles within the replisome. Recent developments in single-molecule approaches have led to a picture of the replisome that is much more dynamic in how the complex retains or recycles protein components. Here we use *in vivo* and *in vitro* single-molecule fluorescence imaging to show that the replisome does not exclusively recruit new SSBs from solution to coat newly formed single-stranded DNA on the lagging strand, but is also able to recycle SSBs from one Okazaki fragment to the next. We show that this internal transfer mechanism is balanced with recruitment from solution in a manner that is concentration dependent. By visualizing SSB dynamics in life cells, we show that both internal transfer and external exchange mechanisms are physiologically relevant.

I carried out and analysed all in vitro and in vivo single-molecule experiments and drafted the manuscript.

6.1 Introduction

Almost all processes associated with DNA metabolism involve the generation of single-stranded DNA (ssDNA). As a transient species to be reconverted into more stable dsDNA, ssDNA acts as a substrate for a large number of pathways. A key protein in the initial steps of ssDNA processing is the Single-Stranded DNA-binding protein (SSB), a protein that coats naked ssDNA and thus protecting it from nucleolytic attacks and preventing intra-strand pairing events such as hairpin formation. Further, it plays a critical role in the organization of protein-protein and protein-DNA interactions within the replisome, the protein machinery responsible for DNA replication (31, 32, 219, 220).

In *Escherichia coli* (*E. coli*), SSB is a stable homotetramer with each 18.9-kDa subunit consisting of 177 amino acids and separated into two distinct domains (221). The N-terminal domain (112 residues) forms an oligonucleotide-binding (OB) fold responsible for ssDNA binding (33). The C-terminal residues form a highly conserved acidic tail, which serves as an interaction site to many binding partners (31, 32, 35, 222). Through its four ssDNA-binding domains, SSB can bind to ssDNA in different modes depending on the concentration of cations and the SSB/ssDNA stoichiometry (223). The prevalent binding modes as observed in *in vitro* studies are the SSB_{65}, and SSB_{35} forms, corresponding to the binding of 65 and 35 nucleotides to each individual SSB tetramer, respectively (Figure 6.1b, left) (34). In the SSB_{65} mode, favored by moderately high salt concentrations (37), all four ssDNA-binding sites are bound to ssDNA. In the SSB35 binding mode (favored in low salt concentrations (224)), however, only two ssDNA-binding sites are occupied (Figure 6.1B, right) (33).

During DNA replication, ssDNA is produced when the helicase unwinds the parental double-stranded DNA (dsDNA). On one of the ssDNA daughter strands, the leading strand, new DNA is synthesized continuously by a copy of the DNA Polymerase III (Pol III) core closely tracking and traveling in the same direction as the helicase (Figure 6.1A), thereby minimizing the amount of time ssDNA is exposed. On the other strand, the lagging strand, DNA is synthesized discontinuously. Due to the opposite polarities of the two DNA strands and the inability of polymerases to synthesize in both direction, the Pol III core on the lagging strand synthesizes DNA in a direction opposite to that of the moving fork (6, 225). As a result, stretches of ssDNA are generated on the lagging strand that are not converted into dsDNA until the next Okazaki fragment is primed and synthesized. During the period these stretches of ssDNA are exposed, SSB is bound to them to prevent formation of secondary structure and nucleolytic digestion. As new DNA is synthesized on the lagging strand, SSB has to be displaced, likely through an interaction of the C-terminal tail of SSB with the χ subunit of the Pol III complex (226, 227).

Biochemical studies performed in the last decades suggest two different models of the dynamics of SSB binding to and dissociating from ssDNA within the

replication complex. In the first model, newly exposed ssDNA is bound by SSBs from the cytosol. It has been shown that SSB binds to free ssDNA in a fast diffusion-controlled process (38, 228). With the estimated *in vivo* SSB concentrations of 300–600nM (16, 221, 229–231), such a rapid binding process would give rise to an efficient coating of newly exposed ssDNA within milliseconds. In this model, subsequent displacement of SSB during filling in of the gap by the lagging-strand Pol III core will cause the SSB to diffuse back into the cytosol.

In an alternative model, SSB is effectively recycled within the replisome through an internal-transfer mechanism. Using Surface Plasmon Resonance (SPR) and nano-electrospray ionization mass spectrometry (nano-ESI-MS), it was shown that the SSB_{35} mode supports transfer of SSB tetramers between discrete oligonucleotides. (211). The timescales observed in stopped-flow experiments demonstrate similar behavior, suggesting that transfer occurs without proceeding through a protein intermediate that is free from DNA. Instead, transfer involves a transiently paired intermediate during which SSB is 'handed' from the first to the second ssDNA while going through a state in which the tetramer is bound to two strands simultaneously (Figure 6.1B, right) (38). These observations have led to speculations on the existence of a direct-transfer or internal-transfer mechanism, in which transfer of SSBs could occur from in front of the Pol III to newly exposed ssDNA behind the helicase (Figure 6.1a). The biochemical studies that have led to the kinetic detail supporting these models have been performed on short oligonucleotides, outside the context of active replisomes. While they have provided invaluable insight into the basic molecular mechanisms underlying the interactions between SSB and ssDNA, they have been unable to directly visualize the dynamic behaviour of SSB within the replisome. As a result, it is unknown how the replication machinery recruits SSB and whether it may retain it during multiple cycles of Okazaki-fragment synthesis. It is unclear whether the approximately 1,000 copies of SSB available within the cell (16, 221, 229–231) are sufficient to support rapid coating of all ssDNA during fast growth, with up to 12 replisomes active simultaneously (232), or whether internal recycling mechanisms are operative that enable a replisome to maintain its own local pool of SSBs.

We report here the use single-molecule fluorescence imaging to visualize the dynamics of SSB during active DNA replication, both *in vitro* in a reconstituted replication reaction and inside living bacterial cells. We rely here on the strength of the single-molecule approach to visualize transient intermediates and acquire detailed kinetic information that would otherwise be hidden by the averaging nature of ensemble measurements (186, 233–235). Particularly, we show that SSB is recycled within the replisome on time scales corresponding to the synthesis of multiple Okazaki fragments, verifying the existence of a mechanism that uses internal transfer to retain SSBs at the fork. At higher SSB concentrations, however, we see that this mechanism is in competition with external exchange to and from solution. Using *in vivo* single-molecule imaging, we show that both processes occur at the replication fork. Our observations suggest that the interactions control-

ling association and dissociation of SSB within the replisome provide a balance between plasticity and stability, enabling rapid exchange of protein factors while ensuring stability in the absence of available factors in the cellular environment.

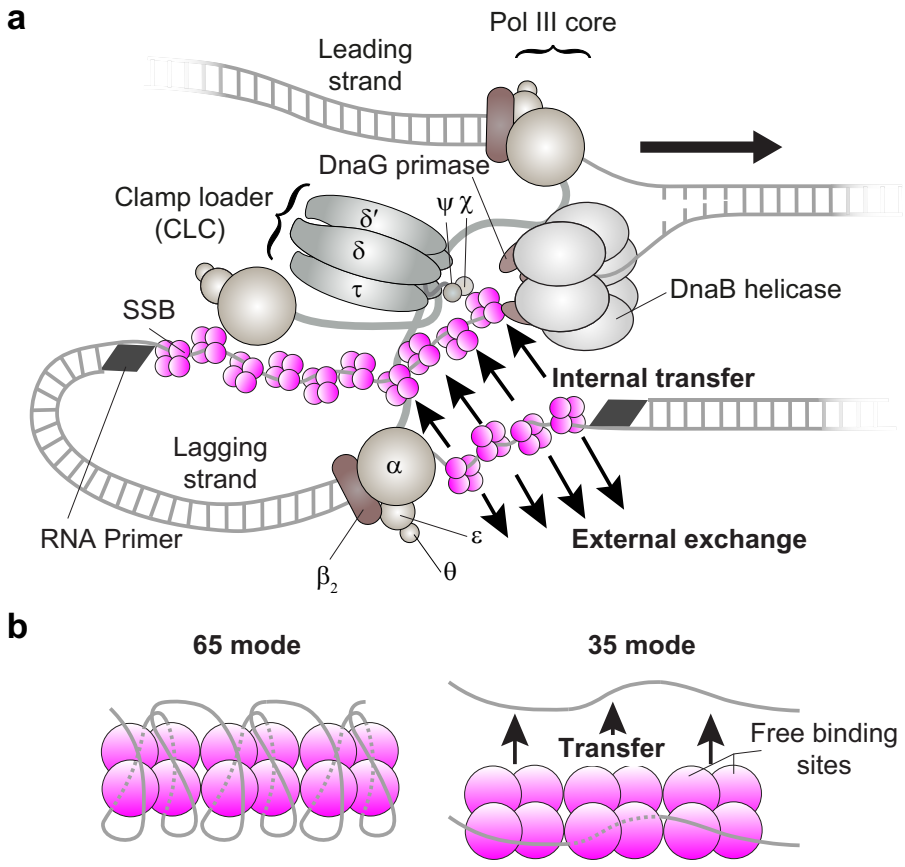


Figure 6.1: *E. coli* replisome (a) Schematic representation of the organization of the *E. coli* replication fork. The DnaB helicase encircles the lagging strand, facilitates unwinding of dsDNA through ATP hydrolysis, and recruits DnaG primase for synthesis of RNA primers that initiate synthesis of 1–2 kb Okazaki fragments on the lagging strand. The Pol III holoenzyme (HE) uses the ssDNA of both strands as a template for coupled, simultaneous synthesis of a pair of new DNA duplex molecules. The β_2 sliding clamp confers high processivity on the Pol III HE by tethering the $\alpha\epsilon\theta$ Pol III core onto the DNA. The clamp loader complex (CLC) assembles the β_2 clamp onto RNA primer junctions. Up to three Pol III cores interact with the CLC through its τ subunits to form the Pol III* complex, and the τ subunits also interact with DnaB, thus coupling the Pol III HE to the helicase. The ssDNA extruded from the DnaB helicase is protected by SSB. (15, 236) (b) Different DNA-binding modes of SSB. In the SSB_{65} mode all four OB domains are bound to DNA (left). In the SSB_{35} mode only two DNA-binding sites are occupied. The observation of transfer of SSB between discrete oligos in this mode, suggests a possible internal-transfer mechanism.

6.2 Results

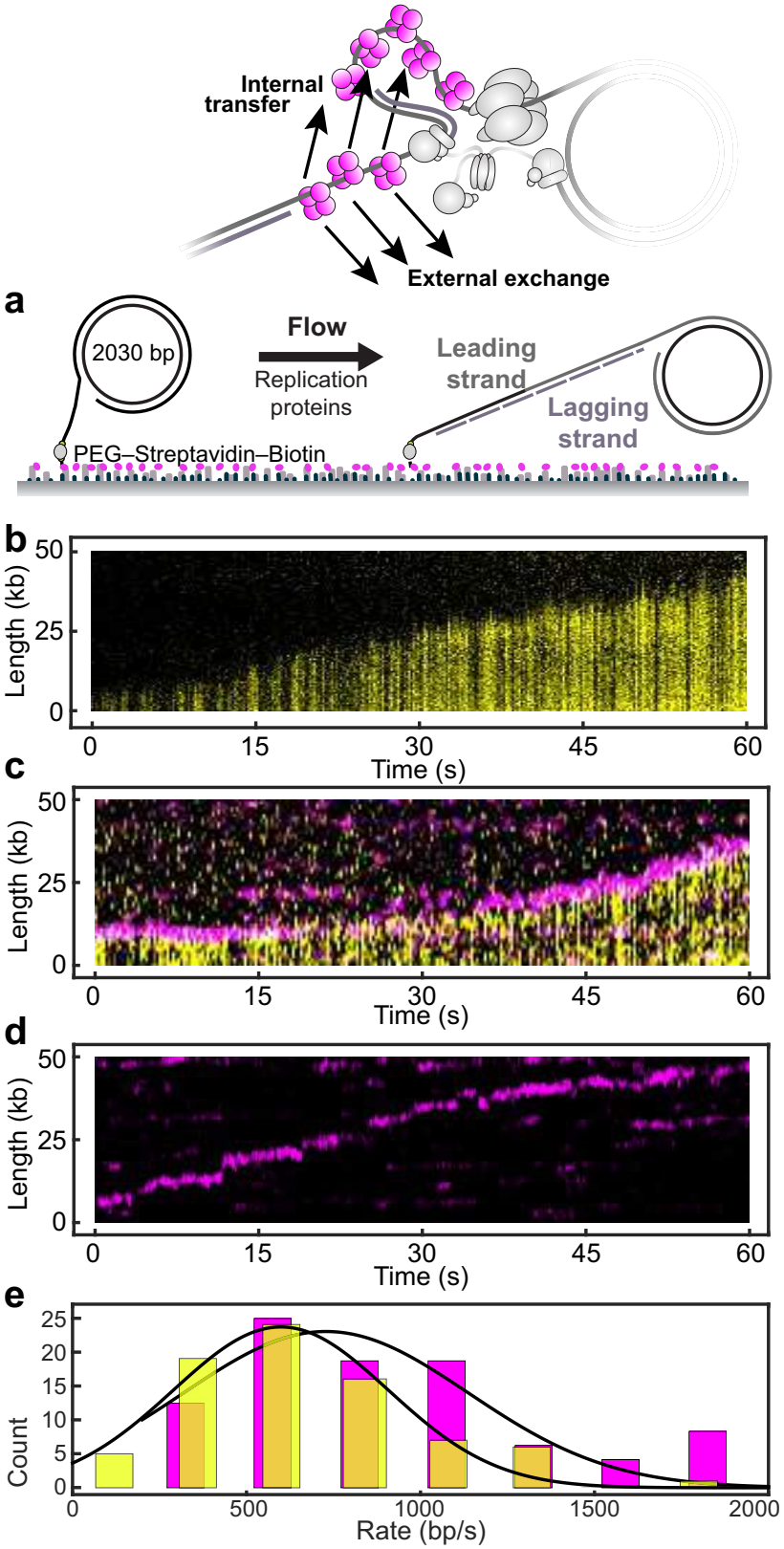
6.2.1 Vizualisation of SSB *in vitro*

We use a single-molecule fluorescence imaging approach to directly visualize DNA replication in real time and monitor the dynamics of SSB at the replication

fork. We performed single-molecule rolling-circle assays, a method that provides information on the rate of production of new DNA by individual replisomes (14) while simultaneously enabling the visualization of fluorescently labeled replisome components (50, 236). A 5′-flap within a 3.0-kb double-stranded (ds) circular DNA substrate is anchored to the surface of a microfluidic flow cell (Figure 6.2a). Replication is initiated by introducing a laminar flow of buffer, containing the minimal set of 12 replication proteins required for coupled leading- and lagging-strand synthesis (Figure 6.1a). Replisomes will assemble onto the fork structure within the circle and initiate unwinding and synthesis (14, 202). As replication proceeds, the newly synthesized leading strand becomes part of the circle to then act as a template for lagging-strand synthesis. The net result of this process is the generation of a dsDNA tail that is stretched in the buffer flow and whose growth moves the tethered dsDNA circle away from the anchor point at a rate equal to the replication rate (Figure 6.2a). Replication is visualized by real-time near-TIRF fluorescence imaging of stained dsDNA (Figure 6.2b). Quantification of the instantaneous rates of individual replisomes resulted in an average single-molecule rate of 626 $\pm$ 73 bp/s with a distribution that reflects intrinsic difference between individual replisomes (Figure 6.1e). These results are similar to those that have been obtained before in ensemble (237) and single-molecule experiments (14, 236).

To visualize the behavior of SSB during rolling-circle replication, we labeled a mutant of SSB containing a single cysteine, SSB-K43C (211), with a red fluorophore (AlexaFluor 647). The labeled SSB was active in coupled leading- and lagging-strand synthesis, producing Okazaki fragments of size distributions identical to those obtained with wild-type SSB in an ensemble-averaging solution-phase reaction (Figure 6.6c). We then used the fluorescently labeled SSB at a concentration of 20 nM (tetrameric concentration) in the rolling-circle assay. Simultaneous imaging of the stained DNA and labeled SSB shows that the SSB is located at the tip of the growing DNA, consistent with the labeled SSB integrated into active, reconstituted replisomes (Figure 6.1c). The single-molecule replication rates obtained in the presence of the labeled SSBs are similar to the rates obtained using the unlabeled wild-type SSB (Figure6.2 d,e), in agreement with our ensemble assays showing that the label does not affect the behavior of SSB in a fully reconstituted DNA-replication reaction, supporting coupled leading- and lagging-strand synthesis. We have reported previously that, under the same conditions, polymerases bind to gaps between Okazaki fragments behind the replication fork (236). Interestingly, we do not observe SSB signals on stationary positions on the product DNA, behavior that would be expected to result in horizontal lines in the kymographs as observed before for labeled Pol III* (236). Our observation of the absence of SSB in gaps left behind the replisome is still consistent with a model in which a replication loop is released before the Okazaki fragment is finished (52).

To be able to resolve SSB left in the wake of the replisome at unfinished Okazaki fragments separately from SSB at the fork, the two populations would have to be separated well beyond the diffraction limit of the microscope used, or at least


Internal transfer
External exchange
a
2030 bp
Flow
Replication proteins
Leading strand
Lagging strand
PEG–Streptavidin–Biotin
b
Length (kb)
50
25
0
0
15
30
45
60
Time (s)
c
Length (kb)
50
25
0
0
15
30
45
60
Time (s)
d
Length (kb)
50
25
0
0
15
30
45
60
Time (s)
e
Count
25
20
15
10
5
0
0
500
1000
1500
2000
Rate (bp/s)

500 nm (equal to ∼1,800 bp). During the time it takes the replisome to cover this distance, polymerases will have finished the previous Okazaki fragment. Our observation also suggests that upon completion of an Okazaki fragment, there is no ssDNA gap remaining between Okazaki fragments sufficiently large (>35 nt) for SSB to bind.

The intensity of the fluorescence signal from the SSBs at the replisome remains constant throughout the experiment (Figure 6.1d). If all SSBs were internally recycled and retained in the replisome, the fluorescence intensity should decay at the characteristic timescale of photobleaching. Given that the photobleaching lifetime of the fluorophores is 9.5 $\pm$ 0.8 s s under these conditions (Figure 6.6b), at least some SSBs at the replication fork are replaced by new proteins from solution. This exchange needs to take place at a rate that is high enough to keep the steady-state level of unbleached SSBs sufficiently high to be observable in the imaging.

6.2.2 Dynamic behaviour of SSB *in vitro*

To quantify the dynamic behavior of SSB during DNA replication, we performed *in vitro* single-molecule FRAP (Fluorescence Recovery After Photobleaching) experiments (236). We used the same rolling-circle scheme as described in Figure 6.1a. Instead of continuous imaging at constant laser power, we periodically bleached all SSBs in our field of view using 100-fold higher laser power (Figure 6.3a, left). Due to the buffer flow and high diffusional mobility, bleached SSBs that are free in solution will rapidly move away and will be replaced by unbleached, bright SSBs. After the photobleaching pulse, we monitor the recovery of the fluorescence signal at the replisome as a readout for the kinetics of introduction of new, unbleached SSB into the replisome. This measurement allows us to distinguish between internal transfer and external exchange of SSB, as they should result in a different recovery behavior. If SSBs are transferred internally and retained at the fork, the fluorescence should not recover after photobleaching

Figure 6.2 *(preceding page)*: Visualization of SSB in the sinlge-molecule rolling-circle assay. (a) Schematic representation of the experimental design. 5′-biotinylated circular DNA is coupled to the passivated surface of a microfluidic flow cell through a streptavidin linkage. Addition of the *E. coli* replication proteins and nucleotides initiates DNA synthesis. The DNA products are elongated hydrodynamically by flow, labeled with intercalating DNA stain, and visualized using fluorescence microscopy (14). (b) Kymograph of an individual DNA molecule undergoing coupled leading- and lagging-strand replication. The gray scale indicates the fluorescence intensity of stained DNA. (c) Representative kymograph of simultaneous staining of double-stranded DNA and fluorescence imaging of labeled SSB (red) in real time. The kymograph demonstrates the fluorescent spot corresponding to SSB co-localizes with the tip of the growing DNA product where the replication fork is located (See also Figure 6.6). (d) Kymograph of the red-labeled SSBs on an individual DNA molecule. The intensity of the SSB signal remains constant for the duration of the experiment, indicating at least some SSBs are exchanged. (e) Histograms of the rate of replication for wild-type SSB (626 $\pm$ 73 bp/s) and labeled SSB (720 $\pm$ 55 bp/s) fit to Gaussian distributions. The similarity between these rates show that the label does not affect the behavior of SSB during replication.

(Figure 6.3a, top right); if dark, bleached SSBs are exchanged with fluorescent SSBs from solution, however, we should observe a recovery of the fluorescence intensity at the replication fork. Figure 6.3b shows a kymograph of a FRAP experiment using 10 nM labeled SSB. The vertical lines correspond to the high-intensity FRAP pulses.

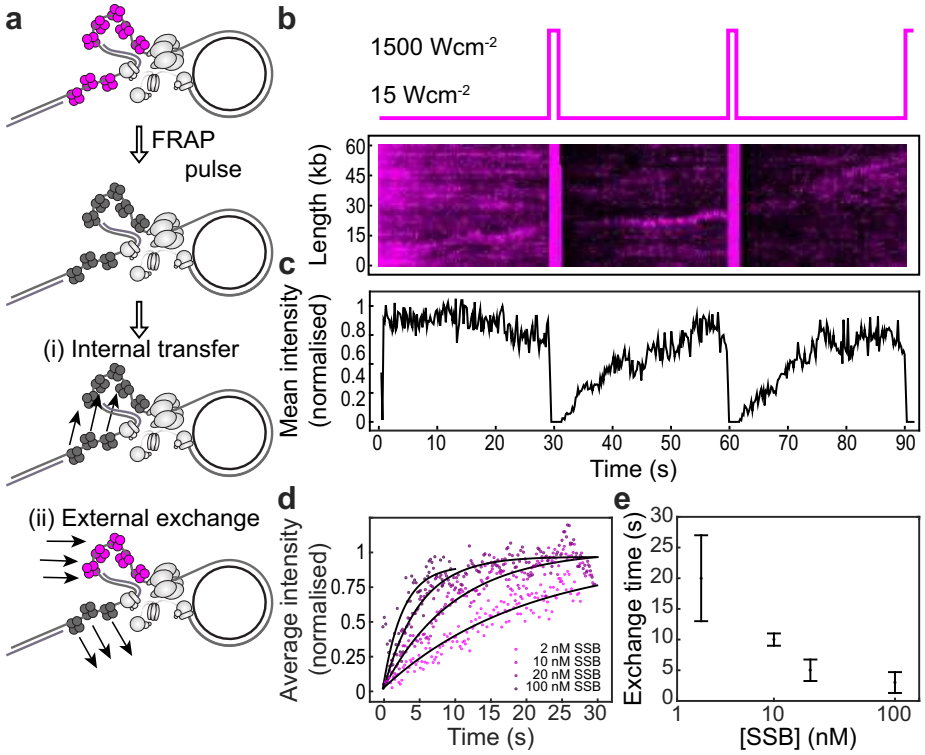


Figure 6.3: Quantification of the SSB exchange time using single-molecule FRAP. (a) Schematic representation of the FRAP experiments. SSBs are initially in a bright state (top left). After a high intensity FRAP pulse all SSBs in the field of view are photobleached (bottom left). If SSBs are internally transferred, no fluorescence recovery should be observed (top right). If SSBs are externally exchanged the fluorescence should recover rapidly (bottom right). (b) Imaging sequence used during the FRAP experiments (Top Panel). A representative kymograph of labeled SSBs at the replication fork (Bottom panel) in a FRAP experiment. After each FRAP pulse (indicated by the vertical red line) all SSBs have bleached. The fluorescence intensity recovers as unbleached SSBs exchange into the replisome. (c) The average intensity over time from 20 replisomes with 10 nM SSB in solution. (d) The three recovery phases in (c) were averaged again to give the final averaged normalized intensity over time after a FRAP pulse. This curve was then fit to provide a characteristic exchange time. This was done for four concentrations of SSB ranging from 100–2 nM. (e) Exchange time as a function of SSB concentration shows a concentration dependent exchange time.

After each FRAP pulse, the fluorescence of the SSB spot decreases to zero and the population is bleached. This bleaching is followed by a gradual increase in intensity, indicating that SSBs from solution associate with the replisome. We determined the intensity after each FRAP pulse over time by averaging over 24 replisomes (Figure 6.3c). At this SSB concentration of 10 nM, we find that the recovery time is 10 $\pm$ 1 s (Figure 6.1d). We then repeated this measurement for SSB concentrations varying from 2 nM to 100 nM (Figure 6.3e). At a total SSB concentration of 2 nM, the fluorescence signal recovers slowly (characteris-

tic exchange time, τ = 20 $\pm$ 7 s, *N* = 20), while at a concentration of 100 nM, the fluorescence signal is $\sim$10 fold faster to recover (τ = 2.9 $\pm$ 1.7 s, *N* = 18). These data show that SSB exchange is concentration dependent, with faster exchange at high concentrations and slower at low concentrations.

6.2.3 SSB is recycled for many Okazaki fragments

Having obtained information on the timescale of SSB turnover at the replisome, we then set out to characterize the number of Okazaki fragment priming and synthesis cycles that occur during that time window. We did so by determining rates of replication and the lengths of the Okazaki fragments. First, we used the single-molecule rolling-circle assay to obtain the DNA replication rates for the different SSB concentrations we used in the FRAP experiments (Figure 6.7). At all concentrations of SSB, the replication rate was 750 bp/s, with no statistically significant differences in rate for the various SSB concentrations. The observation that SSB recovery times can be as high as tens of seconds (Figure 6.3e) suggests that the protein is recycled within the replisome for a period that corresponds to the synthesis of many thousands of base pairs. With an Okazaki-fragment length of 1–2 kb (198), our observations suggest that the replisome retains the SSB for a duration that may very well be beyond the time need to synthesize an Okazaki fragment. Such a long retention time can only be explained by a mechanism that would allow internal transfer of SSBs from one Okazaki fragment to the next. To verify this interpretation, we measured the length of Okazaki fragments generated under our experimental conditions, using both an ensemble-averaging biochemical approach and direct single-molecule observation. It has previously been reported that the SSB concentration has an effect on Okazaki-fragment length (238). To recapitulate this concentration effect, we first performed ensemble rolling-circle replication experiments. Replication reactions containing all proteins required to support coupled leading- and lagging-strand synthesis, with SSB at different concentrations, were allowed to proceed for 30 min. The resulting products were run on a denaturing alkaline agarose gel and stained with an ssDNA stain for visualization (Figure 6.4a). The intensities distributions were normalized to correct for the fact that the intensity per mole of product DNA scales linearly with length. The product length distributions show that the Okazaki fragments are shorter for lower SSB concentrations (1.4 $\pm$ 0.2 knt at 2 nM versus 2.8 $\pm$ 1.0 knt at 200 nM), a factor of two difference in product length for a 100-fold change in SSB concentration.

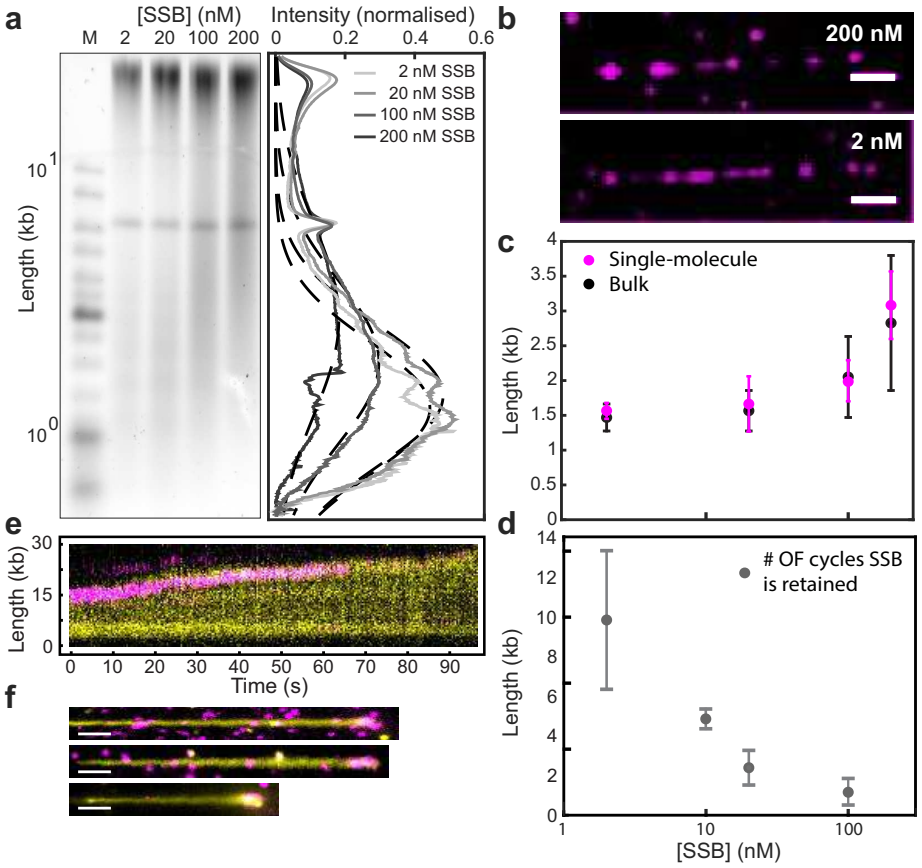


Figure 6.4: Internal transfer of SSB. (a) Alkaline agarose gel of coupled DNA replication for concentrations of SSB used in the FRAP experiments (left panel). Intensity profiles of lanes 2–5 of the left panel. The Okazaki fragment size distributions are centered at 1.4 ± 0.2 kb, 1.5 ± 0.3 kb, 2.0 ± 0.6 kb and 2.8 ± 1.0 kb, in the presence of 2 nM, 10 nM, 20 nM, and 100 nM SSB respectively (mean ± standard deviation). Intensity profiles have been corrected for the difference in intensity of different size fragments using the ladder as a standard. (b) Representative images showing SSBs bound in the gaps between Okazaki fragments. Replication was performed using a polymerase pre-assembly assay, with different concentrations of labeled SSB (top: 200 nM, bottom: 2 nM). Since there is no polymerase in solution to fill in the gaps between nascent OF fragments, SSB will bind there. Therefore the distance between two SSB spots is a measure for OF length. All unbound proteins were washed out for imaging. (c) Comparison of Okazaki-fragment lengths measured in the ensemble assay described in panel (a) (black) and at the single-molecule level (red, Figure 6.9). (d) The number of Okazaki-fragment synthesis cycles that are supported by the same pool of SSB, as a function of SSB concentration. The numbers were obtained by dividing the SSB recovery times in Figure 6.3e, by the time it takes to synthesize one Okazaki fragment using the lengths found in (c). (e) Kymograph of the simultaneous imaging of DNA and SSB with SSB pre-assembled, but not present in solution during replication. Leading- and lagging-strand synthesis continues after the SSB signal disappears, suggesting that SSB is still present at the fork. (f) Representative examples of long DNA products with labeled SSB present at the fork upon conclusion of an SSB pre-assembly experiment.

It can be argued that the effect of SSB concentration on the Okazaki-fragment length may be different when using the single-molecule assay. In the single-molecule experiments, SSB is continuously replenished through the buffer flow, whereas in an ensemble experiment SSB is sequestered from solution as more ssDNA is generated. To test whether such a difference exists, we set out to measure the Okazaki-fragment lengths in our single-molecule assays. However,

in our continuous-flow rolling-circle experiments using DNA staining, we do not have the spatial resolution to observe the gaps between Okazaki fragments in the product DNA. Furthermore, as discussed above, DNA polymerase activity in solution will rapidly fill in the gaps between the Okazaki fragments, making them too small for SSB to bind, and preventing us from using fluorescent SSBs to detect junctions between Okazaki fragments. To resolve this issue, we use conditions for our single-molecule rolling-circle assay that prevent free polymerases from filling in Okazaki-fragment gaps. We achieve this by performing pre-assembly experiments, in which polymerases are present in solution during the initial phase of establishing replisomes at the forks, but left out of the solution during the phase in which the pre-assembled replication complexes produce DNA. Such a design forces the replisome to retain the polymerase holoenzyme (236) and allows labeled SSB to bind ssDNA gaps between the Okazaki fragments without being displaced by other DNA polymerases (Figure 6.4b). The distance between the SSB spots can then be used as a measure for Okazaki-fragment length. By measuring the distances between many pairs of SSB spots (N = 239), we obtain distributions of distances for different SSB concentrations. The distributions were fitted with single-exponential decay functions to obtain the Okazaki-fragment lengths (Figure 6.9). These lengths are similar to those measured in the ensemble experiment, showing that the Okazaki-fragment distributions are identical between the ensemble and single-molecule experiments, with the same dependence on SSB concentration (Figure 6.4c).

We can now use the single-molecule observations of Okazaki-fragment length for different SSB concentrations to directly compare the time required for the replisome to synthesize a single fragment to the SSB recovery time. Converting the information on Okazaki-fragment lengths (Figure 6.4) into times by using the replication rate (Figure 6.7), and by dividing SSB recovery times (Figure 6.3e) by this Okazaki-fragment time, we visualize the number of Okazaki-fragment synthesis cycles that are supported by the same pool of SSB (Figure 6.4d). This analysis shows that at low concentrations, SSBs are retained within the replisome for the duration of multiple ($\sim$10) Okazaki fragments. This number decreases as the SSB concentration is increased, suggesting a competition between internal transfer and external exchange.

As SSBs are continuously displaced from the ssDNA by the lagging-strand polymerase, retention must mean that SSBs are transferred internally to newly exposed ssDNA behind the helicase. To see if we could push the equilibrium between internal transfer and external exchange completely towards internal transfer, we performed a pre-assembly experiment eliminating all free SSBs from solution. In this assay, replication was initiated in the presence of SSB and allowed to proceed for 1 min. We then switched to a buffer containing all replication proteins, but omitting SSB and thereby preventing any external exchange of SSB. Simultaneous imaging of the stained product DNA and the labeled SSB shows that the DNA tail keeps growing after the SSB signal disappears due to photo bleaching of the dye (Figure 6.4e). We conclude that under these conditions, the lifetime of

SSB on ssDNA is longer than the photo-bleaching lifetime of many tens of seconds. In support of this observation, we detect long DNA product molecules with SSB foci at the tip when we apply excitation light and visualize products not until after the replication reaction has finished (Figure 6.4f). To calculate the number of SSBs present on the end of these product molecules, we measured the intensity of these spots. When we divide their average intensity by the intensity of a single SSB (Figure 6.8a), we find that the average number of SSBs stably bound at the end of the DNA products corresponds to 35 $\pm$ 3 SSBs (Figure 6.8b). This number of SSB tetramers corresponds to a ssDNA footprint of slightly more than 1 kb (assuming 35 nt per SSB tetramer), the same length scale as an Okazaki fragment. Remarkably, this observation suggests that upon removal of SSB from solution, the replisome retains its original complement of SSBs for many 10s of kbs of synthesis, supporting highly efficient internal transfer.

6.2.4 Dynamic behavior of SSB *in vivo*

Inspired by our *in vitro* observations of the competition of internal transfer and external exchange processes for SSB in the replisome, we set out to study the dynamics of SSB in live *E. coli* cells. To this end, we used *in vivo* single-molecule FRAP experiments. *In vivo* FRAP has previously been used to measure the dynamics of other replisome components such as the Pol III holoenzyme and the DnaB helicase (239). We used *E. coli* cells in which the chromosomal SSB gene is replaced by a gene that generates a C-terminal fusion of the protein with a yellow fluorescent protein (YPet) (108). To verify that the fluorescent protein does not affect the function of SSB, we show that the growth rate of the SSB-YPet cells is similar to wt *E. coli* cells (Figure 6.10). Furthermore, to confirm that the labeled SSBs form part of active replisomes, we performed colocalization experiments using a dual-color strain with DnaQ-mKate2, producing red-labeled Pol III polymerases, and SSB-YPet. We find that as much as 100% and on average 67% of DnaQ foci per cell colocalize with SSB foci (N = 65 cells). Measurement of the fluorescence recovery of SSB within single cellular foci requires the ability to specifically bleach the fluorescence within a single replisome focus without bleaching the SSB in the rest of the cell. To this end, we placed a pinhole in a motorized filter wheel in the excitation path, producing a tight, diffraction-limited excitation focus (Full Width at Half Maximum = 500 nm). Using this pinhole and a high (200 W/cm^2) laser power, we can bleach a single focus with high spatial specificity (Figure 6.5a). The subsequent fluorescence recovery was visualized by lowering the laser power (2 W/cm^2) and by moving the pinhole out of the beam path. Figure 6.5b shows bleaching and recovery of an SSB-YPet focus within a single cell (green arrow). The first frame was acquired before applying the FRAP pulse. The image acquired immediately after the pulse (t = 0 s) shows that the fluorescence from the single focus has bleached, while the SSBs in the cytosol remain fluorescent. In subsequent frames, we see that the fluorescence recovers, indicating that non-bleached SSBs from the cytosol exchange in to the focus. To quantify the exchange time, we measure the intensity of the foci over time after the bleaching pulse. An average intensity trajectory (N=29 foci) shows an initial

recovery of the fluorescence after the photobleaching pulse, followed by a decay in intensity (Figure 6.5c). This later decay is due to the rapid photobleaching of the YPet probe during visualization, even at the lower imaging intensities after the high-intensity bleaching pulse. To correct for this, we measured the average photobleaching behavior of the probe by monitoring the fluorescence from other cells within the same field of view (Figure 6.5d). Since these cells were not subject to the high-power bleaching pulse, their fluorescence signals provide an ideal internal benchmark for the gradual photobleaching induced by the lower-power imaging illumination. These photobleaching data were fit with a single-exponential decay function (green line). This fit was then used to correct the FRAP intensity trajectories, with the corrected trajectory showing behavior that now is representive of the recovery of the pool

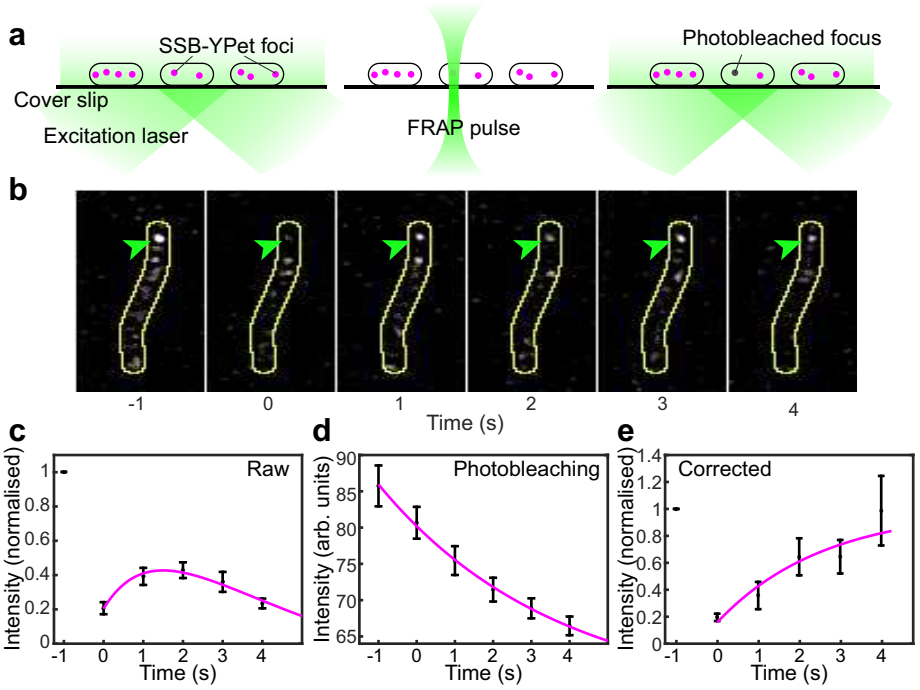


Figure 6.5: Visualization of SSB dynamics *in vivo*. (a) Schematic representation of the *in vivo* FRAP setup. Using near-TIRF illumination SSB-YPet foci (red) are visualized before FRAP (left). By placing a pinhole in the beam path a single focus will be darkened, without bleaching cytosolic SSB-YPet (middle). After the FRAP pulse the recovery of fluorescence can be monitored. (b) Representative images of *in vivo* FRAP experiment. At before t = 0 the focus (indicated by the green arrow) is bleached using a high-intensity FRAP pulse. The fluorescence recovers as fluorescent SSBs from the cytosol exchange into the replisome focus. The cell boundaries are indicated by the yellow line. (c) Averaged normalized FRAP intensity trajectory (*N = 29*). After initial recovery, the fluorescence intensity decreases due to photobleaching. (d) Average intensity over time for SSB-YPet cells outside of the FRAP volume (*N* = 40). These data were fitted with a single-exponential decay function (green line) to obtain the photobleaching lifetime. (e) Averaged normalized FRAP intensity trajectory, corrected for photobleaching. The green line represents a fit to the data, from which we obtained the characteristic *in vivo* exchange time for SSB (τ = 2.5 $\pm$ 1.7 s).

of unbleached SSB at the replisomal spot (Figure 6.5e). By fitting these recovery data (green line), we obtain a recovery time of 2.5 $\pm$ 1.7 s. This value is similar to

the time scale we obtained from the *in vitro* experiments at high SSB concentrations, a similarity that was expected since the estimated concentration of SSB *in vivo* is 300–600 nM during mid-log growth phase (16, 221, 229–231). Assuming that Okazaki fragments produced in the cell are 1,000 – 2,000 nt in length (240) and the replication rate is ∼1000 bp/s, such an exchange time would suggest that for every Okazaki fragment cycle, roughly half of the SSB is internally recycled for the next fragment and the other half exchanged with free SSB.

6.3 Discussion

Biochemical studies suggest two different models that describe how SSB binds to and dissociates from ssDNA within the replisome. In the external-exchange model, newly exposed ssDNA is bound by SSBs from the cytosol. In an alternative model, SSB is recycled within the replisome through an internal-transfer mechanism. Using an *in vitro* single-molecule visualization approach we show here that SSB can be recycled within the replisome on time scales corresponding to the synthesis of multiple Okazaki fragments, thereby verifying the existence an internal-transfer mechanism of SSBs at the fork. At higher SSB concentrations, however, we observe that this mechanism is in competition with external exchange of SSBs with those present in solution. Using single-molecule imaging of labeled SSB in live bacterial cells, we show that both processes occur at the replication fork in a cellular context and that roughly half of the SSB is internally recycled for the next Okazaki fragment.

We conclude that the *E. coli* replisome strikes a balance between internal transfer and external exchange of SSB. In the absence of SSB in solution, the original population of SSB is retained within the replisome and is efficiently recycled from one Okazaki fragment to the next (Figure 6.4e). The existence of such an internal-transfer mechanism has been hypothesized, as it has been shown that SSBs can be transferred between DNA strands through a transient paired intermediate (16, 38, 211). Internal transfer has, however, not been shown before in the context of active DNA replication. We show here that, in the absence of competing, free SSB in solution), SSBs are recycled by the replisome for many 10s of kb. Estimates of the total concentration of SSB per cell in *E. coli* have ranged from 50–600 nM (31, 38, 229). The concentrations of available SSB within the cytosol could be significantly lower with SSB bound to the various ssDNA substrates within the cell. At high growth rates, the cell could contain up to 12 active replisomes (232), leaving little free SSB. This lack of readily available SSB may make binding of SSBs from solution too slow to coat the rapidly produced ssDNA, resulting in exposure of vulnerable ssDNA that can be nucleolytically attacked, can form secondary structure, or can act as a substrate for ssDNA-binding proteins that trigger undesired pathways or responses (e.g., RecA). The internal-transfer mechanism could be a way to ensure rapid SSB coating of newly exposed ssDNA, thereby allowing replication to continue at normal rates without creating large amounts of naked ssDNA.

In the presence of competing SSBs in solution, however, this internal-transfer mechanism is in competition with external exchange of SSBs at a rate that is dependent on its concentration in solution (Figure 6.3e). Such a concentration-dependent exchange mechanism has recently been observed for other proteins that form part of multi-protein complexes (50, 109, 138, 209, 236, 241–244). Under highly diluted conditions, these proteins can remain stably bound within the complex for long periods of time. Yet, rapid (subsecond) exchange is observed at nanomolar concentrations. Such concentration-dependent dissociation can be explained (139) and mathematically described (140, 141, 245, 246) by a multi-site exchange mechanism in which a protein is associated with a complex via multiple weak binding sites, as opposed to a single strong one. At low concentrations, the transient disruption of any of these interactions would not result in dissociation, as the protein is still bound to the complex through the other binding sites. When competing proteins are present, however, a protein in close proximity could bind at the transiently vacated site. The competing protein will then eventually displace the initially bound protein, resulting in full exchange of the two. Examples of concentration-dependent exchange mechanisms can be found in the bacterial flagellar motor (241), with DNA-binding proteins such as Fis, HU (209), and RPA (109), and in transcription regulation (138). Similarly, such a multi-site exchange mechanism has been demonstrated for the association of the replicate DNA polymerase with the replisome in the phage T7 and *E. coli* systems. Instead of a conventional picture in which these proteins are stably bound to the replisome, single-molecule imaging has shown that these polymerases use a multi-site exchange mechanism to rapidly exchange in and out of the replisome at physiologically relevant concentrations (50, 137, 236, 239).

A competition between stability and plasticity that depends on concentration seems harder to comprehend for SSB. Under any circumstance, dilute or not, the SSB–ssDNA interaction has to be disrupted as new dsDNA is synthesized on the lagging strand. Therefore, stability, defined as retention within the replisome, cannot be achieved in the same way as described above, but instead needs to rely on a mechanism of internal transfer. The disruption of the SSB–DNA interaction due to lagging-strand synthesis would be followed by rapid rebinding of SSB to the next Okazaki-fragment template produced behind the helicase, thereby preventing dissociation of the SSB from the replisome. If, however, there are competing SSBs in close proximity to the fork, one of these can bind at the newly exposed ssDNA, thereby blocking that binding site for other SSBs. Consequently, disrupted SSBs from the lagging strand can no longer rebind and are effectively competed out from the replisome.

Our observations of SSB dynamics in living cells show that both internal transfer and external exchange are physiologically relevant pathways accessible to the replisome during coupled DNA replication. In our measurements, during mid-log growth phase (estimated intracellular SSB concentrations of 300–600 nM), the balance seems is towards external exchange, with relatively fast exchange times of 2.5 $\pm$ 1.7 s. This timescale is consistent with the timescales we obtained in

our *in vitro* measurements.

A multi-site exchange mechanism confers both stability and plasticity to the replication machinery, allowing the replisome to operate under different cellular conditions. Our work combined with other recently published studies on the replisome presents a much more dynamic picture of the replisome, distinctly different from the deterministic models generated over the last few decades. It is important to point out that the stochasticity and plasticity observed in recent single-molecule experiments are all consistent with fundamental chemical principles and can be readily explained by hierarchies of weak and strong interactions (55). The apparent generality of the models emerging from these studies suggests that the behaviors of other complex, multi-protein systems might also be governed by such exchange processes and might suggest that evolution of complex interaction networks has arrived at an optimal balance between stability and plasticity.

6.4 STAR Methods

6.4.1 Experimental model and subject details

Source organism for DNA replication proteins
DNA replication proteins were expressed and purified from MG1655 *E. coli* cells.

Cell lines
Wild-type (MG1655) and DnaQ-mKate2 (EAW192) *E. coli* cells were cultured in LB. SSB-YPet (JJC5380), DnaQ-mKate2 SSB-YPet (LMS001), and DnaX-YPet (JJC5945) strains were grown in LB supplemented with 25 μg/ml kanamycin.

6.4.2 Method details

Replication proteins
E. coli DNA replication proteins were produced as described previously: the β_2 sliding clamp (210); SSB (211); the $DnaB_6(DnaC)_6$ helicase–loader complex (212); DnaG primase (213); the Pol III $\tau_3\delta\delta'\chi\psi$ clamp loader (202); and Pol III $\alpha\epsilon\theta$ core (202).
Expression and purification of SSB K43C
Expression and purification procedures of a single cysteine mutant of SSB, SSBK43C were performed as previously described in (211)

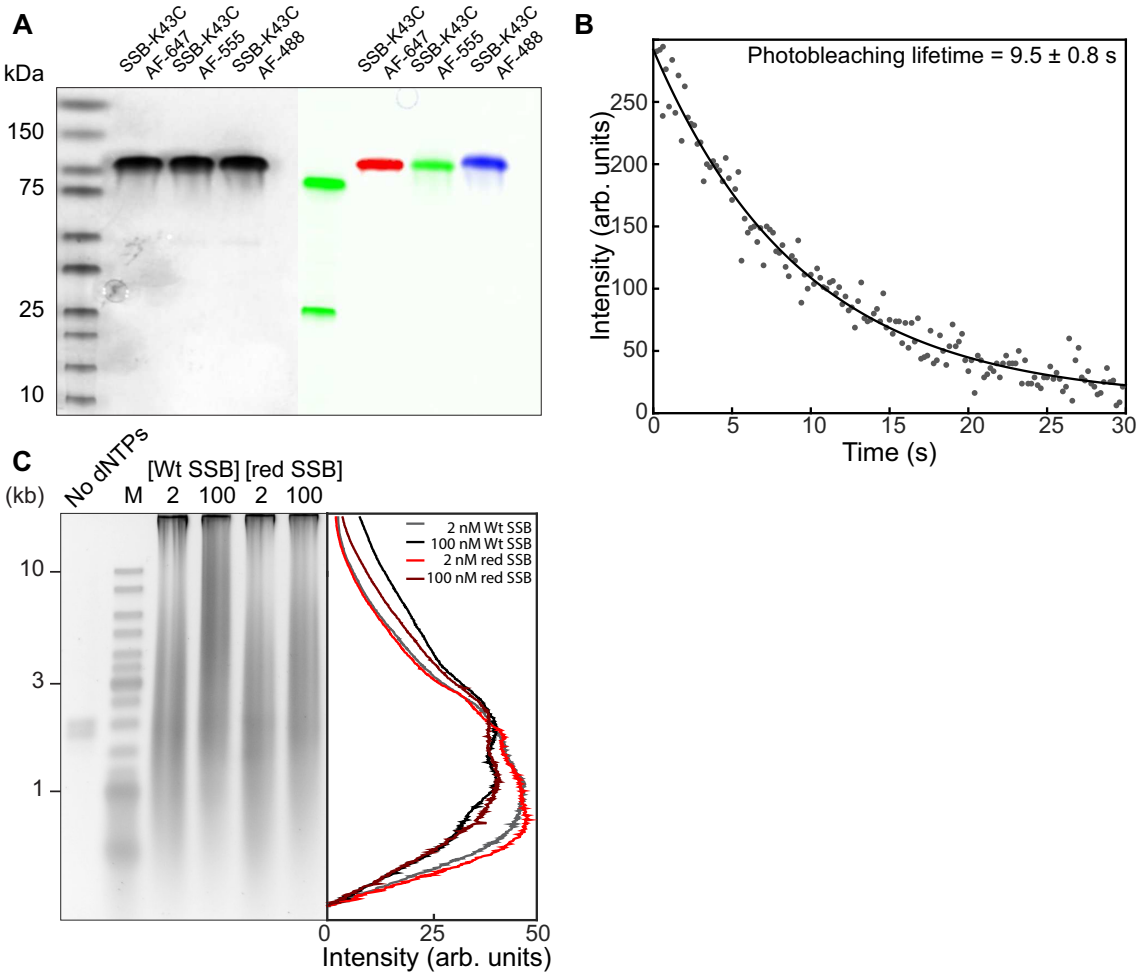


Figure 6.6: Characterization of fluorescent SSB (a) SDS page gel of labeled SSB-K43C. Fluorescence imaging shows no free dye is present after purification. (b) Average photo-bleaching trajectory for SSB-AF647 (N = 4 fields of view, 568 molecules) at excitation power density of 700 mW/cm^2. From a fit with single-exponential decay function (black line) we obtained a characteristic photobleaching lifetime of 9.5 $\pm$ 0.8 s. c Comparison of activities of wild-type and labeled SSBs. (left) Alkaline agarose gel of coupled DNA replication. Reactions were performed on a 2-kb circular dsDNA template with 2nM and 100 nM of either wild-type (Wt) SSB or red labeled SSB. The gel was stained with SYBR-Gold. (right) Intensity profiles of lanes 3–6. Intensity profiles have been corrected for the difference in intensity of different size fragments using the ladder as a standard.

SSBK43C labeling

Methods described below were adapted from (247). Three different fluorescent probes were used to label SSBK43C: Alexa Fluor 488, 555, and 647 (Invitrogen). First, a total of 6.3 mg of SSB K43C was reduced with 3 mM tris(2-carboxyethyl)phosphine (pH 7.6) in labelling buffer (100 mM sodium phosphate pH 7.3, 200 mM NaCl, 1 mM EDTA, 70% (w/v) ammonium sulphate) at 6°C for 1 h with gentle rotation to yield Fraction I. Fraction I was centrifuged (21,000 x g; 15 min) at 6 °C and the supernatant carefully removed. The precipitate was washed with ice cold labelling buffer that had been extensively degassed by sonication and deoxygenated using Ar gas, then pelleted by centrifugation (21,000 x g; 15 min) at 6 °C and supernatant removed to yield Fraction III. The labelling reaction

was carried out on Fraction III, now devoid of reducing agent, using 40 μM of maleimide conjugated dyes with 84 μM SSBK43C in 500 μL of deoxygenated and degassed buffer (100 mM sodium phosphate pH 7.3, 200 mM NaCl, 1 mM EDTA). The reaction was allowed to proceed for 3 h at 23 °C in the dark. The reaction was subsequently quenched using 30 mM dithiothreitol for 1 h at 6 °C yielding Fraction IV. Fraction IV was applied at 1 ml/min to a column (1.5 x 10 cm) of Superdex G-25 (GE-Healthcare) resin equilibrated with gel filtration buffer (50 mM Tris.HCl pH 7.6, 3 mM dithiothreitol, 1 mM EDTA, 100 mM NaCl). Fractions containing the labelled SSBK43C were pooled and dialysed into storage buffer (50 mM Tris.HCl pH 7.6, 3 mM dithiothreitol, 1 mM EDTA, 100 mM NaCl, 20 % (v/v) glycerol). The degree of labeling determined by UV/Vis spectroscopy to be between 1 and 2 fluorescent dyes per SSB tetramer.

Single-molecule rolling-circle assay
To construct the rolling circle template (14), the 66-mer 5′-biotin-T36AATTCGTAATC ATGGTCATAGCTGTTTCCT-3′ (Integrated DNA Technologies) was annealed to M13mp18 ssDNA (NEB) in TBS buffer (40 mM Tris-HCl pH 7.5, 10 mM $MgCl_2$, 50 mM NaCl) at 65°C. The primed M13 was then extended by adding 64 nM T7 gp5 polymerase (New England Biolabs) in 40 mM Tris-HCl pH 7.6, 50 mM potassium glutamate, 10 mM $MgCl_2$, 100 μg/ml BSA, 5 mM dithiothreitol and 600 μM dCTP, dGTP, dATP and dTTP at 37°C for 60 min. The reaction was quenched with 100 mM EDTA and the DNA was purified using a PCR purification kit (Qiagen). Microfluidic flow cells were prepared as described (51). Briefly, a PDMS flow chamber was placed on top of a PEG-biotin-functionalized microscope coverslip. To help prevent non-specific interactions of proteins and DNA with the surface, the chamber was blocked with buffer containing 20 mM Tris-HCl pH 7.5, 2 mM EDTA, 50 mM NaCl, 0.2 mg/ml BSA, and 0.005% Tween-20. The chamber was placed on an inverted microscope (Nikon Eclipse Ti-E) with a CFI Apo TIRF 100x oil-immersion TIRF objective (NA 1.49, Nikon) and connected to a syringe pump (Adelab Scientific) for flow of buffer.

Conditions for coupled DNA replication under continuous presence of all proteins were adapted from previously described methods (14, 202, 236). All *in vitro* single-molecule experiments were performed at least four times. Briefly, 30 nM $DnaB_6(DnaC)_6$ was incubated with 1.5 nM biotinylated ds M13 template in replication buffer (25 mM Tris-HCl pH 7.9, 50 mM potassium glutamate, 10 mM $Mg(OAc)_2$, 40 μg/ml BSA, 0.1 mM EDTA and 5 mM dithiothreitol) with 1 mM ATP at 37°C for 30 s. This mixture was loaded into the flow cell at 100 μl/min for 40 s and then at 10 μl/min. An imaging buffer was made with 1 mM UV-aged Trolox, 0.8% (w/v) glucose, 0.12 mg/ml glucose oxidase, and 0.012 mg/ml catalase (to increase the lifetime of the fluorophores and reduce blinking), 1 mM ATP, 250 μM CTP, GTP and UTP, and 50 μM dCTP, dGTP, dATP and dTTP in replication buffer. Pol III* was assembled in situ by incubating $\tau_3\delta\delta'\chi\psi$ (410 nM) and Pol III cores (1.2 μM) in imaging buffer at 37°C for 90 s. Replication was initiated by flowing in the imaging buffer containing 6.7 nM Pol III*, 30 nM β_2, 300 nM DnaG,

30 nM $DnaB_6(DnaC)_6$ and SSB4 where specified at 10 μl/min. Reactions were carried out 31°C, maintained by an electrically heated chamber (Okolab).

Double-stranded DNA was visualized in real time by staining it with 150 nM SYTOX Orange (Invitrogen) excited by a 568-nm laser (Coherent, Sapphire 568-200 CW) at 150 μW/cm^2. The red labeled SSBs were excited at 700 mW/cm^2 with a 647 nm (Coherent, Obis 647-100 CW) lasers. For simultaneous imaging of DNA and SSB, the signals were separated via dichroic mirrors and appropriate filter sets (Chroma). Imaging was done with an EMCCD (Photometics, Evolve 512 Delta) camera. The analysis was done with ImageJ using in-house built plugins. The rate of replication of a single molecule was obtained from its trajectory and calculated for each segment that has constant slope.

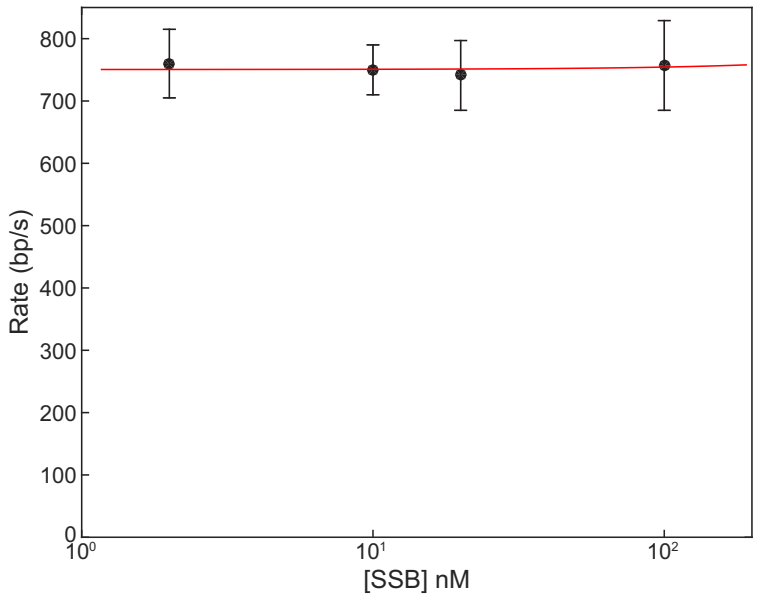


Figure 6.7: Rate of replication is independent of SSB concentration. Replication rate distributions were obtained and fitted as described in Figure 6.2. The points represent the mean of the distribution, the error bars are the s.e.m. The red line represents a linear fit to the data.

Conditions for the pre-assembly replication reactions for the Okazaki fragment length measurements were adapted from published methods (136, 204, 236). Solution 1 was prepared as 30 nM $DnaB_6(DnaC)_6$, 1.5 nM biotinylated ds M13 substrate and 1 mM ATP in replication buffer. This was incubated at 37°C for 3 min. Solution 2 contained 60 μM dCTP and dGTP, 6.7 nM Pol III*, and 74 nM β_2 in replication buffer (without dATP and dTTP). Solution 2 was added to an equal volume of solution 1 and incubated for 6 min at 37°C. This was then loaded onto the flow cell at 100 μl/min for 1 min and then 10 μl/min for 10 min. The flow cell was washed with replication buffer containing 60 μM dCTP and dGTP. Replication was finally initiated by flowing in the imaging buffer containing 50 nM β_2, 300 nM DnaG and SSB_4 where specified at 10 μl/min. Conditions for the chase replication reactions omitting SSB from solution during replication were set up as a normal continuous flow experiment. Reactions were allowed to proceed for 1 min before a replication reaction omitting only SSB was loaded at 10 μl/min.

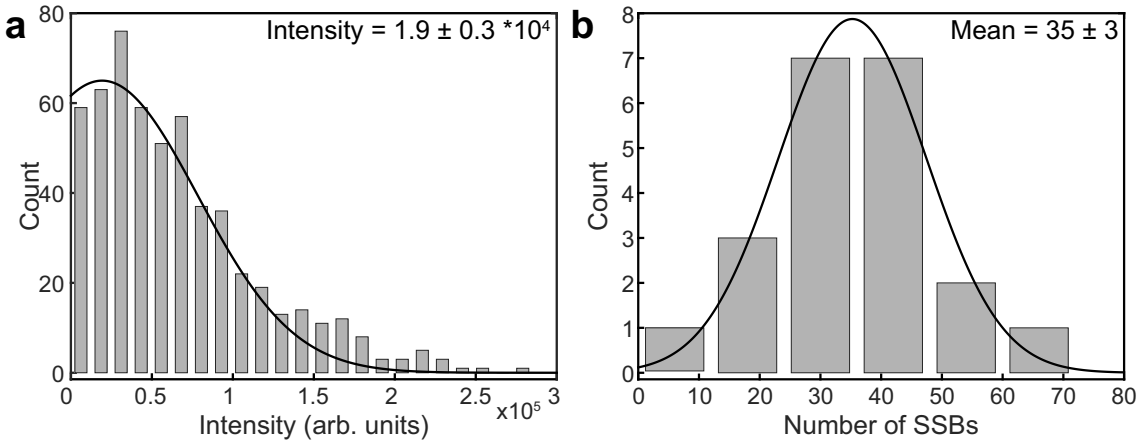


Figure 6.8: Number of SSBs at the fork(a) Histogram of the intensity distribution of single SSBs. The average intensity of a single labeled SSB was calculated by immobilization on the surface of a cleaned microscope coverslip in imaging buffer. The imaging was under the same conditions as used during the single-molecule rolling-circle experiments. Using ImageJ with in-house built plugins, we calculated the integrated intensity for every SSB in a field of view after applying a local background subtraction. The histogram was fit with a Gaussian distribution function to give a mean intensity of $(1.9 \pm 0.3) \cdot 10^4$. The error represents the standard error of the mean. (b) Histogram of the number of SSBs at the fork at the conclusion of an SSB pre-assembly experiment. The numbers were obtained by dividing the intensities at the fork, by the intensity of a single SSB found in (a). From Gaussian fit (black line) we find that there are 35 ± 3 (mean ± s.e.m.) SSBs at the fork (*N* = 31).

Ensemble Okazaki-fragment length measurements

Coupled leading- and lagging-strand DNA synthesis reactions were set up in replication buffer (25 mM Tris-HCl pH 7.9, 50 mM potassium glutamate, 10 mM $Mg(OAc)_2$, 40 μg/ml BSA, 0.1 mM EDTA and 5 mM dithiothreitol) and contained 1.0–1.5 nM of a 5′-biotinylated flap-primed 2-kb circular dsDNA template, 1 mM ATP, 250 μM CTP, GTP, and UTP, and 50 μM dCTP, dGTP, dATP, and dTTP, 6.7 nM wild-type or SNAP-labeled Pol III*, 30 nM β_2, 300 nM DnaG, 100 nM SSB4, and 30 nM $DnaB_6(DnaC)_6$ in a final volume of 12 μl. Components (except DNA) were mixed and treated at room temperature, then cooled in ice for 5 min before addition of DNA. Reactions were initiated at 30°C, and quenched after 30 min by addition of 7 μl 0.5 M EDTA and 6 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). The quenched mixtures were loaded into a 0.6% (w/v) agarose gel in alkaline running buffer (50 mM NaOH, 1 mM EDTA). Products were separated by agarose gel electrophoresis at 14 V for 14 h. The gel was then neutralized in 1 M Tris-HCl, pH 7.6, 1.5 M NaCl and stained with SYBR Gold. The Okazaki fragment length distribution was calculated by normalizing the intensity as a function of DNA length.

E. coli strains with fluorescent chromosomal fusions

The strain EAW192 (dnaQ-mKate2) is a fusion of dnaQ with mKate2 with a xx

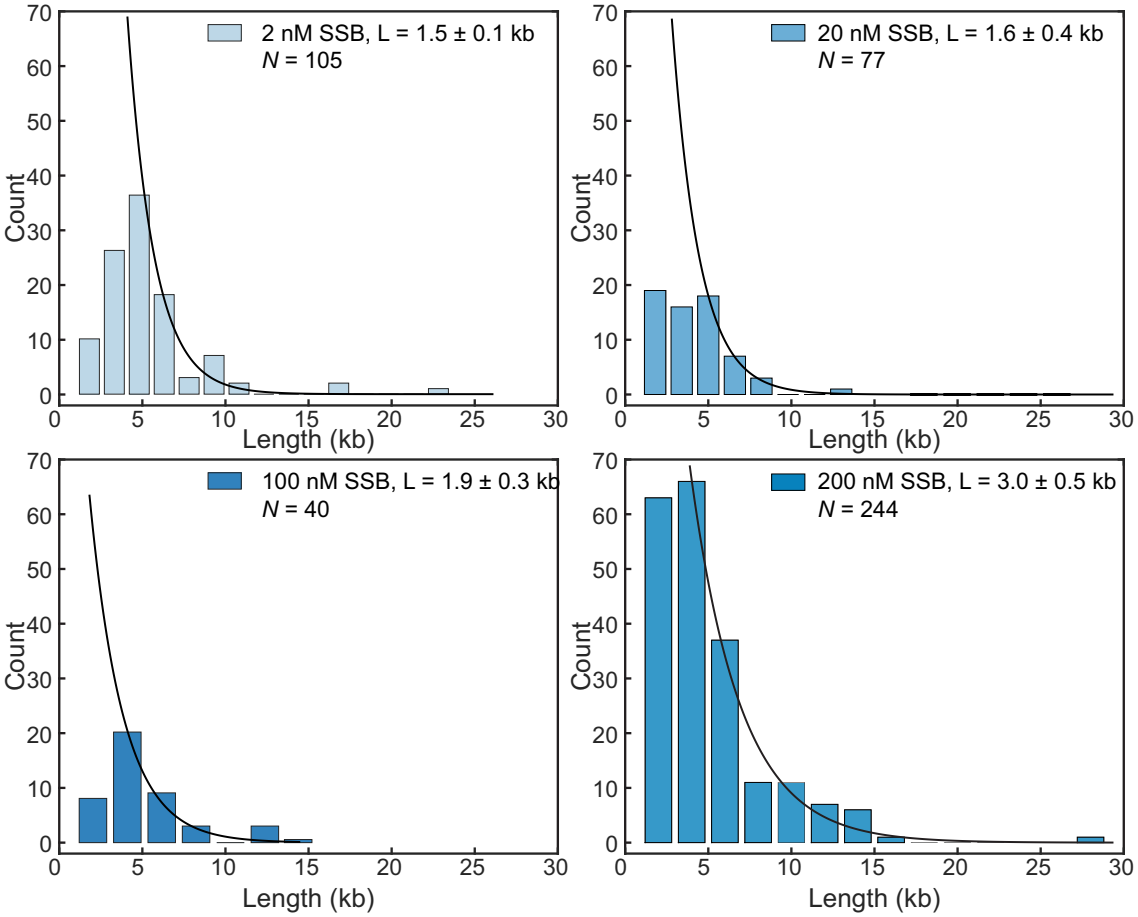


Figure 6.9: Single-molecule measurement of Okazaki-fragment length for different concentrations of SSB. The histograms represent distributions of distance measured between SSB spots. The black lines are a single-exponential fit to the data. The first bars are not included in the fit to take into account undersampling at distances shorter or comparable to the diffraction limit.

linker (236). The JJC5380 (ssb-YPet) is MG1655 ssb-YPet KanR obtained by P1 co-transduction of the ssb-YPet fusion with the adjacent KanR marker from the AB1157 ssb-YPet KanR strain (248), and was a gift from Bénédicte Michel. The two-color strain LMS001 (ssb-YPet, dnaQ-mKate2) was constructed by P1 transduction. JJC5380 cells (ssb-YPet) were infected with P1 grown on EAW192 (dnaQ-mKate2) cells. Transductants were selected for kanamycin resistance. The strain JJC5945 (dnaX-YPet) MG1655 dnaX-YPet (110)

Growth rates of strains with fluorescent chromosomal fusions

To verify that the C-terminal labeling of SSB does not affect cell growth, we compared growth rates of 5 *E. coli* strains. We compared wild-type *E. coli* cells with DnaQ-mKate2, SSB-YPet, and the doubly labelled DnaQ-mKate2 + SSB-YPet strains. We added the DnaX-YPet as a control. Single colonies of wild-type *E. coli* MG1655 and derivatives containing the C-terminal chromosomal dnaX, dnaQ and ssb fusions were used to inoculate 5 ml of LB broth (with 25 μg/ml kanamycin, if required) and grown at 37°C with shaking overnight. LB broth (100 ml) was inoculated with 1.0 · 105 cells/ml from overnight cultures. Subsequent

growth of each strain was monitored at 37°C on a plate reader (POLARstar Omega, BMG Labtech) determining OD_{700} every 20 min for 10 h. The labeled ssb-YPet and dnaQ-mKate2 cells have similar growth rates to wild-type cells (Figure 6.10), indicating that labeling the SSB and DnaQ components of the replisome does not significantly disrupt DNA replication.

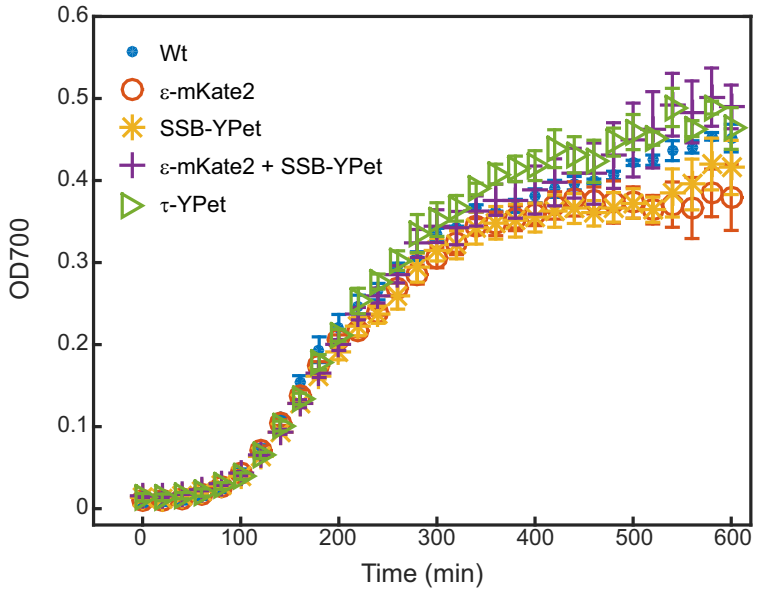


Figure 6.10: Growth curves for *E. coli* strains Wild-type *E. coli* (blue), cells expressing a C-terminal derivatives of ϵ (dnaQ-mKate2, red), ssb-YPet (yellow), dnaQ-mKate2 + ssb-YPet (purple), and dnaX-YPet (green) subunits under control from their endogenous promoters. Growth curves were measured for 10 h. Experiments were performed in triplicate. The errors represent the experimental variation.

In vivo FRAP measurements

The cells were grown at 37°C in EZ rich defined medium (Teknova) that included 0.2% (w/v) glucose. For imaging, cells were immobilized on coverslips that were functionalized with 3-aminopropyl triethoxy silane (Sigma Aldrich) (Robinson et al., 2015) and then placed on the heated stage (Pecon) of the microscope (Olympus IX81, equipped with UAPON 100XOTIRF). Imaging was done at 37°C. FRAP measurements were performed using an automated fast filterwheel (Olympus U-FFWO) with a 50 μm pinhole in the back focal plane of the microscope. A 514 nm laser (Coherent, Sapphire 514-150 CW) was used for visualization and photobleaching. FRAP pulses were 200 ms at 200 W/cm^2 with the pinhole in place. Subsequent visualization without the pinhole was done at 2 W/cm^2. Imaging was done with an EMCCD camera (Hamamatsu c9100-13). The FRAP experiments were performed in triplicate, resulting in a total of 30 photobleached foci that were used for analysis. The image processing was done with ImageJ using in-house built plugins.

7 | The RarA protein of *Escherichia coli* creates DNA gaps behind the replisome

Tyler H. Stanage, Megan E. Cherry, **Lisanne M. Spenkelink**, Jacob S. Lewis, Elizabeth A. Wood, Susan T. Lovett, Antoine M. van Oijen, Michael M. Cox, Andrew Robinson.

Manuscript in preparation for submission to Elife.

The RarA/Mgs1/WRNIP1 protein family is highly conserved from bacteria to humans, yet its cellular function remains enigmatic. We demonstrate that RarA acts directly on the replisome, interrupting replisome activity to generate daughter-strand gaps. Both the gaps and β_2 sliding clamps are left behind as fork progress continues. *In vivo*, this activity creates substrates for translesion DNA synthesis and RecFOR-mediated daughter-strand gap repair. RarA function ensures optimal rates of replisome progress, and rarA deletion results in a substantial growth defect. Loss of RarA function partially suppresses the sensitivity of $\triangle$recF, $\triangle$recO, or $\triangle$uvrA strains to DNA damage. RarA loss completely suppresses the sensitivity of strains lacking the function of any translesion DNA polymerase to DNA damaging agents. The action of RarA effectively commits the cell to the repair of DNA damage within gaps left behind the replication fork. If gap repair processes are compromised, survival is enhanced by elimination of RarA.

I conducted and analysed all in vitro single-molecule experiments and was involved in the fluorescence labelling of β_2 and writing of the manuscript.

7.1 Introduction

Replication forks can stall when encountering roadblocks, such as DNA lesions, template-strand breaks or DNA-bound proteins. The outcomes of stalling may include fork collapse and replisome dissociation (249–259), and these outcomes can have catastrophic consequences for genomic integrity and cell viability. Although estimates vary, replication forks in bacteria may stall as often as once per cell generation during normal growth conditions (249, 250, 260–267). Most of the adverse replication-fork encounters are resolved using a variety of pathways that do not introduce mutations (249–260, 268, 269).

When bacterial cells are stressed by conditions that inflict higher levels of DNA damage, the SOS response is induced. In the early stages of SOS, nonmutagenic pathways for DNA repair still predominate. If the SOS response is prolonged however, a different set of pathways for DNA damage tolerance becomes more prominent. These pathways involve specialized DNA polymerases that carry out translesion DNA synthesis (TLS) (110, 270–284). In *E. coli*, TLS is carried out by DNA polymerases II, IV, and V (282). In bacteria growing under normal conditions, TLS can become important when non-mutagenic pathways for replication-fork repair are blocked (285).

Unlike DNA polymerase V, DNA polymerases II and IV are present in significant concentrations under conditions of normal cellular growth (30–50 molecules per cell of DNA polymerase II and about 250 molecules of DNA polymerase IV) (286–289). The reason for this constitutive presence of these TLS polymerases has been enigmatic. When growing in log phase, bacterial cells lacking DNA polymerase IV ($\triangle$*dinB*) function are highly sensitive to the genotoxic agents methyl methanesulfonate (MMS), nitrofurazone (NFZ), and 4-nitroquinoline-1-oxide (4-NQO) (290–295). Loss of DNA polymerase II introduces sensitivity to oxidative DNA damage (287). DNA damage tolerance via TLS has a greater role during normal DNA replication in eukaryotes (296–302). The current study explores the function of the RarA protein. The *Escherichia coli* RarA protein is a AAA+ ATPase (447 amino acid residues; 49.6 kDa), and is part of a family with close homologs in eukaryotes (Mgs1 in yeast, WRNIP1 in humans). Sequence conservation within the family is extensive, with RarA sharing roughly 40$\%$ identity and 56–58$\%$ similarity with its *Saccharomyces cerevisiae* (Mgs1) and *Homo sapiens* (WRNIP1) homologs (303, 304). This extensive homology suggests a conserved function in DNA metabolism. RarA also shares considerable sequence homology with the τ, δ and δ' subunits of the DNA polymerase III clamp-loader complex, placing RarA in the clamp-loader AAA+ clade. The protein has also been referred to as MgsA, a reference to its homology with the yeast protein Mgs1 (305). As the RarA designation was proposed first (303), and to avoid confusion with the mgsA acronym previously assigned to the gene encoding methylglyoxal synthase (306), we use the rarA nomenclature. It is well documented that the RarA family of proteins is involved in the maintenance of genome stability in cells, but its function

and mechanism of action remain uncertain in spite of nearly two decades of work.

Several dozen studies have now been published on the RarA/Mgs1/WRNIP1 protein family. Although hese have yielded a complex, and sometimes contradictory plethora of observations, several themes are evident. First, RarA family members localize to the replisome through interactions with either the single-stranded DNA binding protein, SSB (RarA), or ubiquitinated processivity clamp PCNA (Mgs1 and WRNIP1) (303, 304, 307–312). Second, the sequence and structure of RarA (and by extension the other members of this family) place it in the clamp-loader clade of AAA+ ATPases (308). However, it appears to function as a tetramer rather than having the usual pentameric structure (308). Third, RarA has an effect on replisome stability and somehow promotes TLS (305, 313–316). Fourth, RarA, Mgs1, and WRNIP1 all exhibit a DNA-dependent ATPase activity *in vitro* that specifically targets duplex DNA ends and gap boundaries (305, 308, 317–321). Fifth, RarA function appears to complement a range of DNA damage tolerance pathways (304, 305, 317, 318, 322–327). These genetic results suggest that RarA does not belong to any currently defined repair pathway. Using a combination of *in vitro* single-molecule DNA replication assays, single-cell microscopy and cell growth assays, we provide evidence that RarA acts to transiently disengage or inhibit part of the replisome to create gaps in the lagging strand product. *In vivo*, RarA-mediated gap creation results in a situation in which DNA damage must be dealt with in gaps behind the fork, where translesion DNA synthesis plays a major role in damage tolerance. Observations made during previous studies of RarA/Mgs1/WRNIP1 function can be harmoniously reinterpreted from the perspective of catalyzed gap formation.

7.2 Results

7.2.1 Rationale and outline

The initial goal of this study was to more directly test a working hypothesis that RarA is involved in a switch between normal and TLS DNA replication. As the work progressed, the hypothesis became more focused: RarA creates gaps behind the replication fork, creating substrates for TLS polymerases and for daughter strand gap repair. The study has both *in vitro* and *in vivo* components. Using an *in vitro* single-molecule assay we directly assessed the effects of RarA action on active DNA polymerase III replisomes. To assess the effects of RarA activity *in vivo*, we have used both direct observation by single-molecule microscopy and a range of additional *in vivo* assays to document the cellular effects of a rarA deletion. These include (a) effects on growth and fitness, (b) suppression of the UV sensitivity of recF and recO mutations, (c) suppression of the sensitivity of cells lacking TLS polymerases to particular DNA damaging agents, and (d) partial suppression of the sensitivity of cells lacking UvrA function to nitrofurazone. These five sets of experiments are considered in succession below.

7.2.2 RarA *in vitro*: RarA action creates gaps during DNA polymerase III-mediated DNA synthesis

To observe the effects of RarA at replication forks, we utilized a single-molecule DNA replication assay (14, 202, 236). This assay employs a rolling-circle DNA amplification scheme, allowing observation of processive DNA synthesis by the *E. coli* replisome in real time (Figure 7.1a). A double-stranded (ds) circular DNA substrate is anchored to the surface of a microfluidic flow cell through a biotinylated 5′-flap. This flap also facilitates loading of the DnaB helicase. Replication is then initiated by introducing a laminar flow of buffer with the components required for DNA synthesis. As replication proceeds, the newly synthesized leading strand becomes part of the circle and later acts as a template for lagging-strand synthesis. With the leading strand attached to the surface and the continuously growing DNA product stretched in the buffer flow, the dsDNA circle moves away from the anchor point. Replication is visualized by real-time near-TIRF fluorescence imaging of stained dsDNA (Figure 7.1b, top). This strategem allows quantification of the instantaneous rates of individual replisomes and their processivities.

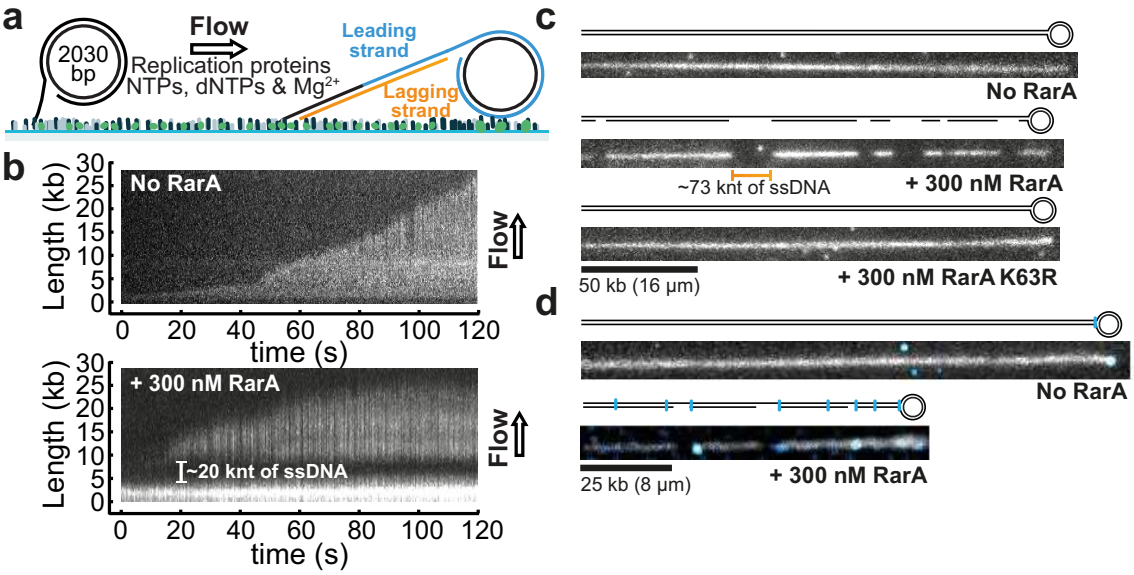


Figure 7.1: RarA induces the formation of ssDNA gaps on the lagging strand. (a) Schematic representation of the experimental design. 5′-Biotinylated DNA is coupled to the passivated surface of a microfluidic flow cell through a streptavidin linkage. Addition of the E. coli replication proteins and nucleotides initiates DNA synthesis. The DNA products are elongated hydrodynamically by flow, labeled with intercalating DNA stain, and visualized using fluorescence microscopy. (b) Kymographs of individual DNA molecules undergoing rolling circle replication in the absence or presence of 300 nM RarA. (c) Examples of individual DNA molecules produced by rolling circle replication in the absence of RarA, or in the presence of 300 nM RarA or its ATPase-dead mutant RarA K63R. The gray scale indicates the fluorescence intensity of stained DNA. (d) Examples of individual DNA molecules produced by rolling circle replication in the presence and absence of RarA in which the β_2 clamp was fluorescently labeled with Alexa Fluor 647.

In the experiments documented in Figure 7.1b, the replisome was pre-assembled onto the rolling-circle template in solution. Subsequently, the template was attached to the surface of a flow cell. The flow cell was then washed to remove

all unbound proteins. Replication is initiated by introduction of a replication solution that omits Pol III* and helicase, but includes SSB, DnaN (the β_2 clamp), DnaG primase, rNTPs, and dNTPs. The absence of free Pol III* in solution makes polymerase exchange impossible (236). Nonetheless, these conditions support highly processive DNA replication with synthesis rates and processivities identical to a situation with Pol III* in solution (Figure 7.7) (236) and are consistent with values reported previously (14, 203, 206, 328).

When RarA was included at a concentration of 300 nM (tetramer) in the replication reaction solution, numerous gaps appeared in the rolling circle products synthesized by individual replisomes (Figure 7.1b, bottom, c, middle). The stain used to visualize the duplex DNA binds poorly to single-stranded (ss) DNA. Thus, any gaps in the product can be visualized as breaks in the fluorescence signal along the growing DNA molecules. Introducing the same concentration of an ATPase defective RarA protein (RarA K63R) did not produce gaps and had no evident effect on replisome progress (Figure 7.1c, bottom). Thus, the appearance of gaps is dependent upon both the presence of RarA and its ATPase activity.

RarA and RarA K63R bind to the SSB C-terminus with very similar affinities (Figure 7.9). The absence of gaps when the ATPase mutant is added provides evidence that gap formation is not an artifact of strong binding of RarA to SSB. On average, RarA-induced gaps were 2.2 $\pm$ 0.5 μm long (mean $\pm$ s.e.m.) and appeared at an average frequency of once every 78 nm (Figure 7.1b, Figure 7.8). Under the experimental conditions, dsDNA has a length of 3 nt/nm (measured by visualization of tethered 20-kb linear dsDNA under the same conditions). In comparison, ssDNA is much more compact. Based on previous measurements of SSB-coated ssDNA (329), we estimate that the ssDNA within gaps has a length of approximately 12 nt/nm. Applying these length conversions, RarA-induced gaps had an average length of 26 $\pm$ 6 knt and appeared at a frequency of once every 37 kb (Figure 7.1b, Figure 7.8). These values imply that under these conditions leading- and lagging-strand synthesis become uncoupled for very long periods.

We next examined the dependence of these parameters on the RarA concentration. When RarA was included in the reaction mixture at 100 nM, gaps appeared less frequently (once per 100 kb). However, the lengths of the gaps were unaffected (average length of 1.8 $\pm$ 0.7 μm; (Figure 7.8). At 30 nM RarA, few gaps were observed. Thus the gap frequency is dependent on RarA concentration, whereas the gap length is independent. When RarA was included at 300 nM, but a five-fold higher concentration of β_2 was used, the length of the gaps reduced to 0.95 $\pm$ 0.19 μm (Figure 7.8). Based on these observations (together with results described below), we attribute the very long gaps to slow restart of Okazaki-fragment synthesis as a result of slow loading of β_2 clamps from solution under these experimental conditions.

From our experiments we can discern that the gaps appear exclusively in the

lagging-strand product. A consequence of the rolling-circle construct used in these assays is that any gap formed in the leading-strand product would lead to rapid termination of DNA synthesis. This termination arises because any replisomes that collide with a leading-strand gap on a subsequent trip around the circle would displace the circle from the growing product (Figure 7.14). Any gaps formed in the leading-strand product as a result of RarA action would therefore manifest as a reduction in the overall length of the products observed at the completion of the assay. No such reduction in product length was observed (Figure 7.7). Movies of actively growing DNA molecules indicate that gaps form immediately behind the replication fork (Figure 7.1c). When RarA was added late, at 20 min after most polymerization reactions had been completed, no gaps were observed (Figure 7.15). Thus we conclude that RarA acts directly at the replisome and transiently affects the function of the Pol III engaged in lagging-strand DNA synthesis. There is no nuclease function or contamination in the RarA preparation that would create gaps randomly in the DNA.

The lagging-strand bias demonstrated by RarA can potentially be explained by two different mechanisms. In the first, RarA acts selectively on the lagging-strand polymerase. Strand selectivity could be imparted through the interaction of RarA with SSB, most likely binds exclusively on the lagging-strand template. Among the many proteins that interact with the C-terminal segment of SSB (32), RarA has one of the highest affinities (308). Interestingly, Mgs1 and WRNIP1 have been reported to interact specifically with the eukaryotic lagging-strand DNA polymerase δ (316, 317, 320). In the second mechanism, RarA acts on both the leading- and lagging-strand polymerases. Here, the selectivity is explained by polymerase dissociation having different outcomes on each strand. RarA-mediated disengagement of a Pol III on the lagging strand would result in release of the DNA loop tethering lagging- and leading-strand synthesis, precluding re-assembly of a polymerase at the now distant and exposed 3′ terminus (7.16). Similar disengagement on the leading strand would likely halt the progress of the replisome, allowing rapid re-assembly of another polymerase and resumption of DNA synthesis without producing a gap. In principle the second mechanism would manifest as a reduction in the average rate of the replisome, although the effect could be subtle. There is no evidence in the kymographs of growing DNA molecules (Figure 7.1b, bottom) to suggest that RarA slows the rate of replisome progression.

RarA could act to disengage the Pol III core with its associated β_2 clamp from the DNA template, or it could act to separate that Pol III core from its β_2 clamp. If RarA disengages both the Pol III core and associated clamp, and both are re-used when lagging strand DNA synthesis re-initiates, no β_2 clamps would be left behind the fork at RarA-mediated gaps. If, on the other hand, RarA separates Pol III core from its clamp, one would expect that clamps would be left behind at gaps. To distinguish these possibilities, we repeated the rolling-circle assays using fluorescently labeled β_2 clamp. As expected, no gaps were observed in the absence of RarA (Figure 7.1d, top). Fluorescently labeled β_2 clamps were

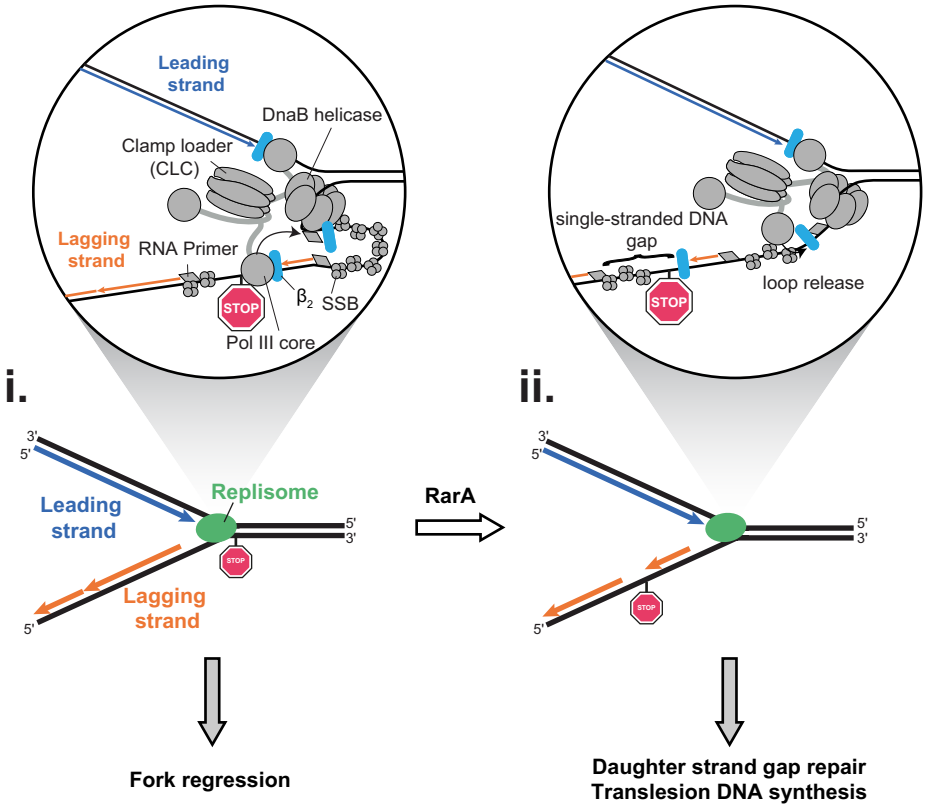


Figure 7.2: A new model of RarA activity: gap creation. The data presented in this study show that RarA creates gaps behind the replication fork. This scheme outlines a potential mechanism for this activity and highlights its implications for DNA repair in vivo. In this scheme, polymerase stalling at lesions, or perhaps other barriers, on the lagging strand triggers the action of RarA (**i**). RarA detaches the lagging strand polymerase from its β clamp, allowing replisome to skip over the lesion and leaving a short ssDNA gap it its wake (**ii**). These lesion-containing gaps are ideal substrates for translesion DNA synthesis by DNA polymerases II, IV and V, or for recFOR-dependent daughter strand gap repair. Unlike lesions at stalled replication forks, lesions within single strand gaps cannot be bypassed through fork regression mechanisms. Thus, TLS and daughter strand gap repair represent the only available options for repair.

only visible at the tips of DNA molecules, corresponding to the position of the replication fork. This indicates that under the conditions of the assay (20 nM β_2 clamp is provided in the replication solution), clamps are predominantly recycled by Pol III* during synthesis of each new Okazaki fragment (136). Introduction of 300 nM RarA to the reaction led to products containing numerous gaps, as expected. However multiple β_2 clamps were now visible on each product DNA (Figure 7.1d, bottom). This indicates that in the rolling-circle assay RarA action disengages Pol III core from the β_2 clamp which remains associated with the dsDNA upstream of each gap. Many of the abandoned clamps were seen near gaps. It is possible that all of them were thus associated, since gaps of less than 500 nt would not be observed in this experiment. We note that no DNA lesions have been purposefully introduced into these DNA substrates. RarA protein is acting to create gaps behind the replication fork. The levels of RarA employed, and perhaps the absence of other factors that may affect RarA function in the cell, may amplify an activity that normally addresses replisome stress with more targeted precision *in vivo*.

We conclude that RarA has a direct and ATPase-dependent destabilizing effect on the replisome that results in the formation of (in this system) lagging-strand gaps behind the replication fork. A working model for this activity is shown in Figure 7.2. The detailed mechanism remains to be determined, but the simplest interpretation is that gap formation involves transient disengagement of the Pol III responsible for lagging-strand synthesis such that the β_2 clamp is separated from the polymerase core. *In vivo*, such an activity would need to be balanced with Pol III function, such that Pol III would only be displaced when a significant barrier was encountered. The remainder of this study involved an examination of the effects of RarA *in vivo*, focusing on the consequences of deletion of rarA.

7.2.3 RarA *in vivo*

In considering the results below, the implications of gap creation by RarA need to be kept in mind. If RarA creates gaps behind the replication fork, these become substrates for both translesion DNA synthesis and RecA-mediated daughter strand gap repair. The balance between these pathways of DNA damage tolerance could depend upon the type of lesion, levels of DNA damage (and thus numbers of gaps) and the lengths of gaps. Lesions within gaps could not be directly repaired by nucleotide excision repair. However, both translesion DNA synthesis and daughter strand gap repair via recombination would leave the lesion in place and create a substrate for nucleotide excision repair (NER). Thus, nucleotide excision repair would also have a late role in gap repair. Gaps would rapidly become distant from the replication fork as the replisome continues on its way downstream. This would make lesion bypass through fork regression impossible. RecA-mediated daughter strand gap repair would be available, although any lesions within gaps shorter than $\sim$7 nt presumably could not be repaired in this way because RecA-mediated DNA pairing would be blocked (330–332). The apparent advantage of creating the gaps is to allow optimal progression of the replication fork, with lesion repair occurring behind it. Under normal growth conditions, such gaps may be generated when occasional lesions are encountered, or perhaps at sites where replication pauses for other reasons. Gap generation would increase with DNA damage levels. These gaps would generally go undetected by most methods currently applied to *in vivo* replication monitoring. When RarA was not present, gap generation would decline. The same barriers would be more likely to stall the replisome, forcing resolution via other pathways such as fork regression. Pathways such as nucleotide excision repair, when available, would address most genomic damage while the stalled fork was resolved. The importance of both TLS and RecFOR-mediated daughter strand gap repair would decline in the absence of RarA. With nucleotide excision repair needed both in the gaps and in the broader genomic DNA, it is hard to predict how elimination of RarA function would affect cells lacking key components of NER. If the DNA damage sensitivity of recFOR or TLS polymerase mutants occurs in whole or in part because their function is required in gaps created by RarA action, a rarA deletion should eliminate most or all of the gaps and thus suppress that DNA damage

sensitivity while repair was diverted to other slower pathways.

7.2.4 RarA *in vivo*: (a) Effects of rarA deletions on cell growth.

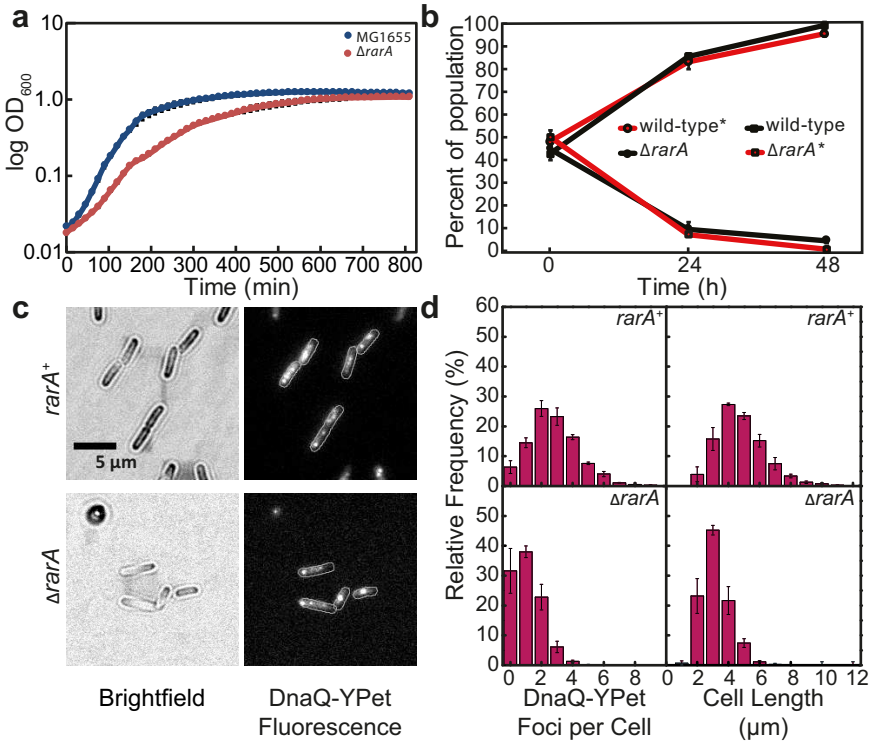


Figure 7.3: Strains lacking rarA exhibit a growth defect compared to wild type MG1655 cells, exhibit smaller cell size and contain a reduced number of replisome foci. (a) Overnight cultures were grown in LB medium, diluted 100-fold and allowed to grow for 1000 min. OD600 values were recorded over this time period. Traces represent OD600 values averaged over a minimum of three biological replicates. (b) Cells lacking the rarA gene are outcompeted by wild type MG1655 cells. Using a growth competition assay, equal amounts of wild type MG1655 and ΔrarA strains are incubated together at t = 0 h. This co-culture was allowed to grow for 48 h, with samples taken at 0, 24, and 48 h. These samples were serially diluted and plated onto tetrazolium agar plates. Deletion of the araBAD operon acts as a marker (*) and is able to be differentiated from araBAD+ cells. This marker was combined with both the wild type and ΔrarA genotypes, with no discernable effect on the growth rate of either strain. Colonies representing each genotypic population were counted and divided by the total number of colonies to determine the percentage of the total population each strain inhabited. These experiments were conducted in triplicate, with error bars representing the standard deviation from the mean. (c) Single-molecule fluorescence imaging of rarA+ (EAW170, top) and ΔrarA (THS04, bottom) strains containing DnaQ-YPet labeled replisomes. Cells were grown at 37°C in flow cells and imaged every 5 min for 180 min. (d) Histograms of DnaQ-YPet replisome foci per cell and cell length for the rarA+ (light gray) and ΔrarA (dark gray) strains. Error bars represent the standard error of the mean values for each bin across at least two replicates. DnaQ-YPet foci per cell for rarA+ cells: mean = 2.8; SEM = 0.03; n = 2738 cells. DnaQ-YPet foci per cell for ΔrarA cells: mean = 1.3; SEM = 0.03; n = 1424 cells. Cell size for rarA+ cells: mean = 5.5; SEM = 0.03 μm in length; n = 2892 cells. Cell size for ΔrarA cells: mean = 3.7; SEM = 0.03 μm in length; n = 1660 cells.

A growth defect. We hypothesize that the main function of a gap creation function is to promote optimal progression of the replication fork when lesions or other

barriers are encountered. Whereas truly stalled forks may occur at most a few times per cell cycle, there may be many more circumstances where lesions or replication pause sites are simply bypassed, leaving daughter strand gaps behind the fork. Loss of RarA function may thus result in a growth defect. No growth or viability phenotype has previously been ascribed to strains with a rarA gene deletion (303–305, 315, 333). However, no growth curves comparing the deletion mutant to wild type cells have been published. Previous work has focused on a modified rarA gene in which a chloramphenicol-resistance cassette replaced either the first 600 nucleotides of the gene (303, 304, 315, 333) or codons 113–349 (305), both in an *E. coli* AB1157 background. As most of our own constructs are based on *E. coli* strain MG1655, we constructed a complete rarA gene deletion in the MG1655 background. As detailed below, deletion of rarA resulted in a small cell phenotype and this could be complemented by moderate expression of rarA from an arabinose-inducible plasmid. We compared the growth of the ΔrarA strain to wild-type cells in rich growth medium. The ΔrarA cells grew more slowly than wild-type, exhibiting a doubling time of 42 versus 29 min for the WT cells (Figure 7.3a), a result not previously reported. To determine the relative fitness cost of rarA deletion, we carried out direct competition assays between the wild type strain and the ΔrarA strain using an approach developed by Lenski and colleagues (334) (Figure 7.1b). Wild type or mutant cells were modified to carry a neutral Ara– mutation (which confers a red color on colonies when grown on tetrazolium arabinose, TA indicator plates) to permit color-based scoring of mixed populations. Overnight cultures of the ΔrarA strain were mixed in a 50/50 ratio with isogenic wild type cells carrying the Ara– mutation. The mixed culture was then diluted and grown up again on successive days, with plating to count red and white colonies occurring once each day. Earlier work (334, 335) demonstrated that the Ara– mutation does not affect growth rates by itself. We nonetheless carried out the competitions twice with the Ara– mutation in one strain or the other to control for any anomalous effects the rarA deletion might exhibit in the Ara– background. In both experiments, the wild-type cells outgrew the ΔrarA cells and dominated the mixed cultures almost completely within 48 hours (Figure 7.1b). Based on the 24h time-point, we calculated that ΔrarA had a relative fitness w = 0.5 (336), indicating a significant loss of fitness relative to wild type cells.

Reduced cell size. To investigate why ΔrarA cells grow more slowly than wild type, we carried out single-molecule single-cell microscopy. To see if there were any replication defects, we made use of strains expressing two different replisome markers, fusions of either the ϵ or τ subunits of the replicative DNA polymerase III holoenzyme complex (dnaQ-YPet or dnaX-YPet, respectively). We compared the number and position of replisome foci in rarA+ and ΔrarA cells. In each case the replisome protein was labeled at its C-terminus with the bright yellow fluorescent protein YPet, and was expressed from its normal chromosomal locus under its native promoter, as in previous work (110). The functionality of these constructs has been assessed in multiple ways. The constructs are modeled after, and the fusion protein sequence identical to, constructs used previously by Sherratt and colleagues (248). Using an AB1157 background, these workers found little or no

effect of the fusions on cell growth rate (in minimal media), DNA content profiles, cell length distributions, or numbers of oriC foci (248). An identical dnaX-YPet derivative of MG1655 (the one also used in the current study) and a similar dnaQ-mKate2 MG1655 derivative (identical but for the nature of the fluorophore) produced no observable growth defect when grown in LB medium (236). In the current study, we again confirmed that the dnaQ-YPet and dnaX-YPet alleles have little or no effect on log-phase growth rates: dnaQ-YPet cells grew only slightly slower than wild type cells, whereas dnaX-YPet cells grew at wild type rates (Figure 7.10).

We observed that $\triangle$rarA cells were substantially smaller than rarA+ cells (3.7 [SEM = 0.03] versus 5.6 [SEM = 0.04] μm in length) and divided less frequently within the flow cell environment used for imaging (division time = 53 [SEM = 2] versus 26 [SEM = 1] min; n = 20 cells each) (Figure 7.1c). Additionally, dnaQ-YPet rarA+ cells had between 0 and 10 replication foci (mean = 2.8; SEM = 0.03; n = 2371) (Figure 7.1d), consistent with multi-fork DNA replication due to growth in rich imaging medium. Cells carrying a rarA deletion (dnaQ-YPet $\triangle$rarA) had fewer foci (mean = 1.2; SEM = 0.03; n = 1116), consistent with their slower growth rate. Approximately 7$\%$ of cells contained more than 2 replication foci, indicating that ÎŤrarA cells are capable of multi-fork replication, but grow slow enough that this mode of replication is not usually necessary. Cells carrying a rarA deletion could be complemented with leaky expression of rarA from a pBAD plasmid, returning cells to near wild-type in terms of cell size (mean = 4.0; SEM = 0.17 μm in length; n = 50) and number of DnaQ-YPet foci (mean = 3.0; SEM = 0.2; n = 50) (Figure 7.12). Similar results were obtained using the dnaX-YPet strains (Figure 7.11a,b). We also measured cell size in wild type and $\triangle$rarA cells containing unaltered replisome proteins lacking the YPet fusions (Figure 7.11c), and in each case found the distributions to be essentially identical to those seen for dnaQ-YPet and dnaX-YPet derivatives. The cell size result is not affected by the YPet fusions. The $\triangle$rarA cells did not show any outward signs of replication defects, such as filamentation or abnormal numbers of replication foci relative to cell size. Rather they resemble cells growing in a nutrient-poor medium, showing slower growth and fewer replisomes. The imaging results are consistent with the hypothesis that RarA facilitates optimal rates of replisome progression and that this helps cells to grow quickly.

We also examined cell size using flow cytometry, with results that completely substantiated the results of the microscopy experiments. Forward area light scattering and DNA dye fluorescence were measured for wild type (MG1655) and $\triangle$rarA cells grown in rich medium to exponential phase. Cells were not fixed or synchronized prior to observation by flow cytometry. Data were gated to exclude cell debris and doublets (Figure 7.13a,b). Cell size was significantly reduced in cells lacking RarA, as measured by forward scattering area (Figure 7.13c,e). DNA content was also significantly reduced and exhibited a bimodal distribution in $\triangle$rarA cells as opposed to a unimodal distribution seen in wild type cells (Figure 7.13d,f). These results further underline a model in which RarA function is

required for optimal replication fork progression *in vivo*.

7.2.5 RarA *in vivo*: (b) A rarA deletion suppresses the UV sensitivity of recF and recO mutations.

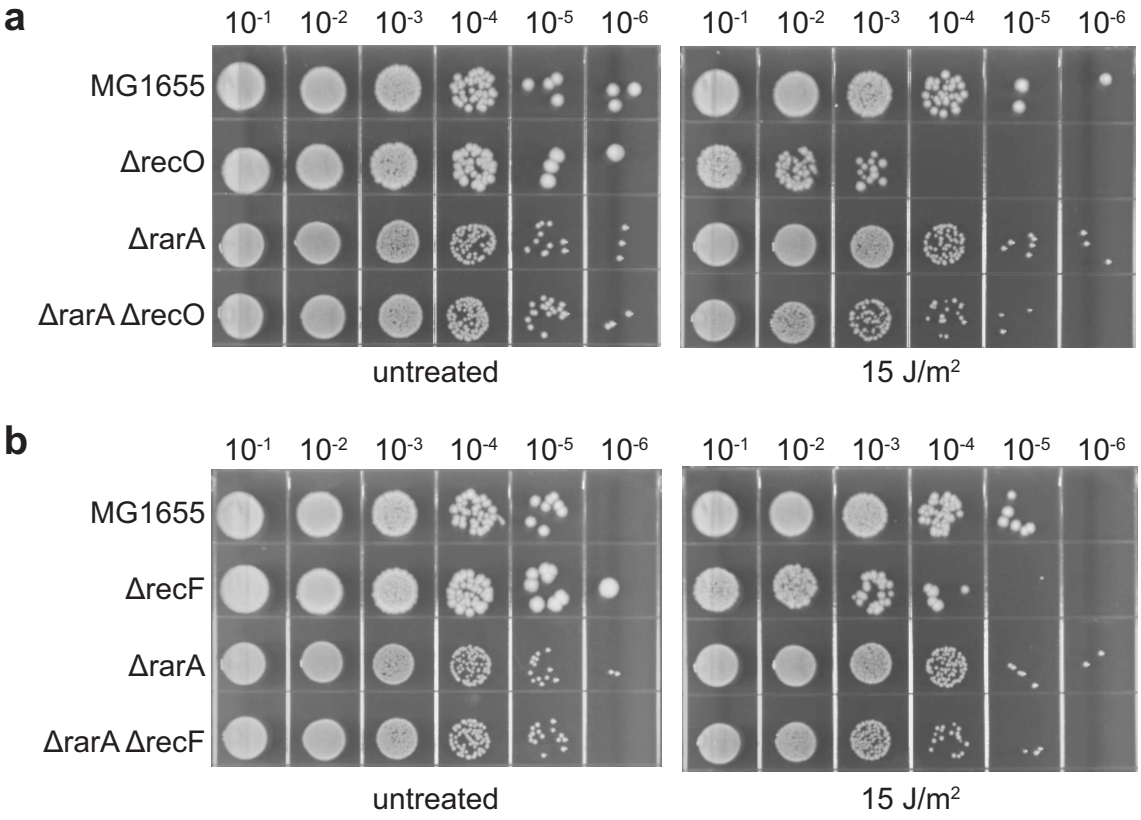


Figure 7.4: Deletion of rarA suppresses the sensitivity of ΔrecF and ΔrecO strains to low levels of ultraviolet light. Indicated strains were grown to exponential growth phase (OD_{600} = 0.2), serially diluted, spot plated onto LB agar plates, and irradiated at a dose of 15 J/m^2. (a) ΔrecO cells are sensitized to low levels of UV light by 2 orders of magnitude compared to wild type cells. Deletion of rarA results in no observable decrease in cell viability and restores cell viability to wild type levels in a ΔrecO background. (b) ΔrecF cells are sensitized to low levels of UV light by 1–2 orders of magnitude compared to wild type cells. Deletion of rarA in a ΔrecF background restores cell viability to wild type levels.

The RecF, O and R proteins have been implicated in the repair of daughter-strand gaps. These proteins all have a role in the loading of RecA protein onto SSB-coated single-stranded DNA at gaps and in some cases at DNA ends. Loss of function of any of the RecFOR proteins results in sensitivity to UV irradiation (255, 337–340). If a significant proportion of the gaps that act as RecFOR substrates in UV-irradiated cells are created by RarA action (via core polymerase disengagement to leave pyrimidine dimers in gaps), then loss of RarA function could decrease the numbers of UV-associated gaps. Fork stalling would be more likely and UV lesion repair would be channeled into pathways other than RecFOR-mediated daughter-strand gap repair. In this case, a rarA deletion could

suppress or partially suppress recFOR mutations, depending upon how much RecFOR function was focused on gap repair. As shown in Figure 7.4, this is indeed the case. When the recO gene was deleted, and cells were subjected to UV irradiation at a dose of 15 J/m2, cell survival declined by approximately 2 orders of magnitude relative to wild type cells. Cells lacking the rarA gene exhibited no decline in survival, although the cells grew more slowly as is evident by the colony size in Figure 7.4. When both recO and rarA were deleted, the sensitivity of the $\triangle$recO cells to UV at this dose was completely suppressed. This indicates that the sensitivity conferred by a lack of RecO function is not due entirely to the absence of RecO, but instead reflects an activity of RarA. This is the result that would be expected if RarA was creating the major substrate for RecO action following UV irradiation. In the absence of RarA, RecO is no longer needed as repair is shunted to alternative (slower) pathways. A similar result was obtained with cells lacking RecF function, although survival in the $\triangle$recF strain declined somewhat less at this UV dose than was seen with the $\triangle$recO cells (Figure 7.4). Each of these experiments was repeated 3 times with consistent results.

When the UV dose was increased to 30 J/m^2, a similar result was obtained (Figure 7.17). Survival by the cells lacking RecO or RecF function declined further at the higher dose, as expected. However, in this case, suppression of the UV sensitivity of the $\triangle$recO and $\triangle$recF cells by $\triangle$rarA, although readily apparent, was partial. We attribute this to the creation of substrates requiring RecFOR action by pathways that do not involve RarA at the higher UV dose. These experiments were repeated 3 times with consistent results.

7.2.6 RarA *in vivo*: (c) A rarA deletion suppresses the DNA damage sensitivity of TLS polymerase mutants. RarA-mediated commitment to translesion DNA synthesis.

Bacterial cells lacking DNA polymerase IV ($\triangle$*dinB*) function and growing in log phase are highly sensitive to the agents methyl methanesulfonate (MMS), nitrofurazone (NFZ) and 4-nitroquinoline-1-oxide (4-NQO) (290–295). DNA polymerase IV can bypass the lesions at guanine-N2 resulting from treatment with these agents (291,341,342). The sensitivity of dinB mutants to these agents has been parsimoniously interpreted as reflecting the absence of DNA polymerase IV. If RarA is creating suitable substrates for TLS action, eliminating these substrates should eliminate the need for TLS, and a rarA deletion should suppress the sensitivity of dinB mutants to these agents. This is what is observed (Figure 7.5, Figure 7.18); a rarA deletion strain was no more sensitive to NFZ than a wild type control (Figure 7.5). But as in the case of the recO and recF deletions, combining the dinB and rarA deletions led to a complete suppression of NFZ sensitivity normally found in $\triangle$dinB cells. This result implies that the sensitivity of dinB mutants to NFZ is not due simply to the absence of DNA polymerase IV. Instead, the sensitivity of dinB mutants to NFZ reflects a function of the RarA protein. The ATPase function of RarA is required to create suitable substrates for TLS; a RarA mutant that is unable to hydrolyze ATP, K63R, is as effective at suppressing the

sensitivity of dinB mutants to NFZ as complete deletion of the rarA gene (Figure 7.5). This is again consistent with a pathway where RarA creates a substrate for DNA polymerase IV action. When RarA is present but DNA polymerase IV is not, lesions are left behind the replisome in gaps but are not addressed by Pol IV and many cells die. When RarA is not present, the substrates for Pol IV action are not created and alternative pathways act to remove and/or bypass the lesions.The suppression effect of a rarA mutant is not specific to DNA polymerase IV. Mutants of DNA polymerase II (polB) are also sensitive to NFZ (343). Mutants of DNA polymerase V (umuCD) are sensitive to high doses of UV irradiation (344). In both cases, the sensitivity is suppressed by including the rarA deletion, or the rarA K63R mutant, in a Pol II or Pol V deficient background (Figure 7.5). These results indicate that the RarA activity leading to TLS function involves ATP hydrolysis. The lack of specificity for any one of the TLS polymerases, along with the similar effects of rarA deletions in recO and recF backgrounds, make it unlikely that RarA directly recruits TLS polymerases to the replisome (or to the gaps it creates). Instead, RarA is creates a substrate on which TLS polymerases can function; β_2 sliding clamps left behind at the gaps as a result of RarA action may play a role in recruiting TLS DNA polymerases.

7.2.7 RarA *in vivo*: (d) A *rarA* deletion partially suppresses the DNA damage sensitivity of a *uvrA* deletion mutant.

NER can not address a lesion present in single-stranded DNA, and thus could not function in RarA-generated gaps. However, after the gap DNA is converted to duplex DNA by translesion DNA synthesis or recombination-mediated daughter-strand gap repair, the lesion would still be present and become a substrate for NER. The importance of NER in cells lacking RarA function would then become a complex reflection of the alternatives that might exist for DNA lesion resolution or tolerance in front of the fork. We examined the effects of a rarA deletion on cells that also lack the function of the nucleotide excision protein UvrA. Cells lacking nucleotide excision repair are extremely sensitive to a range of DNA damaging agents. In the case of NFZ, extreme sensitivity led us to utilize a range of nitrofurazone concentrations from 2.5–10 μM. NFZ reduced survival in $\triangle$uvrA cells by approximately 2–5 orders of magnitude, depending on the concentration used (Figure 7.6). There was no reduction in survival in $\triangle$rarA cells, as seen in Figure 7.5, although colonies again grew more slowly. The same $\triangle$rarA deletion modestly suppressed the deleterious effects of the $\triangle$uvrA deletion, improving survival by about 1 order of magnitude at the lowest concentration of NFZ, with slightly less suppression evident at higher levels of NFZ. The cells continued to grow slowly without RarA function. The results indicate that RarA function — creating gaps — is moderately deleterious in cells lacking nucleotide excision repair.

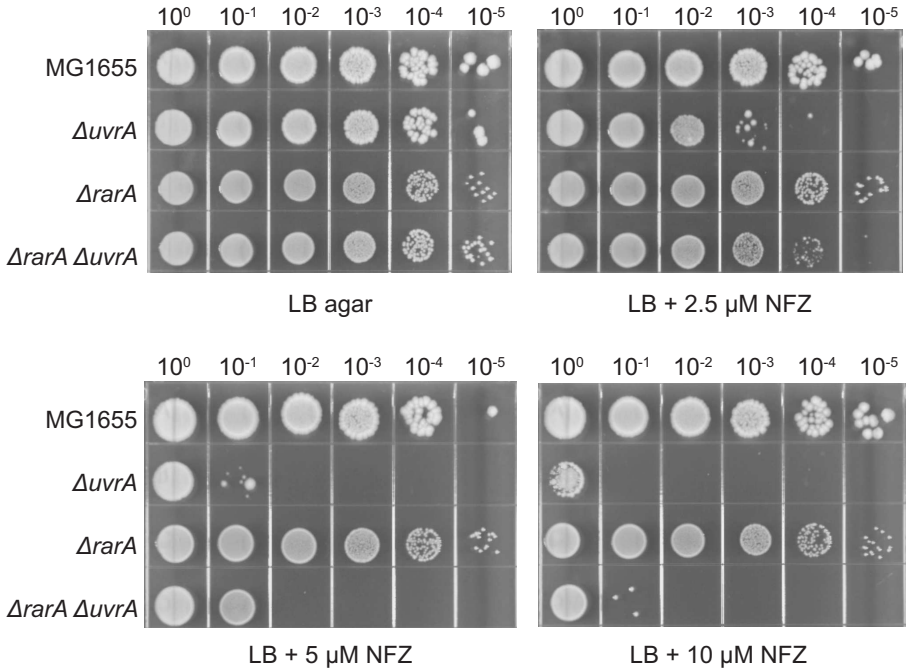


Figure 7.6: Deletion of rarA partially suppresses the sensitivity of a ΔuvrA strain to DNA damage. Cultures of indicated strains were grown to exponential growth phase (OD_{600} = 0.2), serial diluted, and spot plated onto LB agar containing indicated concentrations of nitrofurazone (NFZ). Cells lacking the nucleotide excision pathway protein uvrA are highly sensitive to nitrofurazone. As in Figure 7.5, cells lacking the rarA gene are not sensitive to nitrofurazone at concentrations up to 10 μM. At low (2.5 μM) and medium (5 μM) concentrations of nitrofurazone, deletion of the rarA gene in a ΔuvrA background partially suppresses these cells' sensitivity to NFZ by one order of magnitude. The effect of suppression is diminished at high concentrations (10 μM) of nitrofurazone. In all cases, cell viability in a double rarA/uvrA deletion strain is not restored to wild type levels.

7.3 Discussion

We conclude that the action of RarA generates gaps behind the replication fork. *In vivo*, RarA activity (a) facilitates normal rates of cell growth and (b) sets the stage for TLS function and daughter strand gap repair requiring the RecFOR proteins. Addition of RarA to a single molecule system of rolling-circle DNA replication leads to the creation of gaps and abandonment of β_2 sliding clamps behind the replication fork. Absence of RarA function confers a substantial growth defect in rich medium, albeit without a decline in cell survival. The ΔrarA mutation suppresses the DNA damage sensitivity of both recF/recO and TLS polymerase mutations by eliminating a substrate for TLS and daughter-strand gap repair and channeling DNA damage tolerance into alternative, albeit slower, repair pathways. Our working model for RarA is outlined in Figure 7.2 and provides obvious paths for further experimentation. When a lesion or perhaps another barrier is encountered, the fork may often not stall. Instead, RarA disengages part of the replisome, halting DNA synthesis at that location. DnaG-promoted restart then allows replication to continue upstream, and the lesion or other barrier is left behind in a gap. TLS polymerases can then act to fill in the gap,

or at least the portion containing the lesion. In some instances, the gap may be addressed by RecFORA-mediated daughter-strand gap repair. Nucleotide excision repair can contribute after the single-stranded and lesion-containing gap has been converted to duplex DNA. Notably, whenever a process that potentially contributes to gap repair is compromised, survival is enhanced by eliminating RarA and its gap-generating function. Alternative processes that deal with lesions in front of the fork, possibly including pathways involving fork stalling, collapse, and/or regression, increase chances for survival even while slowing cell growth.

Our model for RarA action has changed. RarA does not promote a switch between normal and TLS replication. Instead, it facilitates a switch between repair in front of the replisome, involving lesions or barriers that might otherwise lead to fork stalling or collapse, to repair in gaps left behind the fork. The repair in both cases is likely to involve both nucleotide excision repair and aspects of recombinational DNA repair. However, the contexts for repair ahead of and behind the fork have notable differences, and repair in gaps may represent a special purview of translesion DNA synthesis. Nucleotide excision repair cannot act in gaps until they have undergone processing by either TLS or recombination-mediated daughter strand gap repair. Recombination pathways could well be overwhelmed by higher DNA damage loads that lead to the generation of many gaps. Recent work has documented that DNA polymerases IV (345) and V (110) spend relatively little time at the replication fork, but instead form foci at many other locations in the genomic DNA. The need for TLS in RarA-generated gaps may explain the enigmatic presence of significant constitutive levels of DNA polymerases II and IV in *E. coli* (286–289). There is no evidence that RarA plays a direct role in TLS polymerase recruitment. Instead, the connection between RarA and TLS appears to be indirect; RarA simply generates an ideal substrate for TLS polymerases to act upon, i.e. one that results in a β_2 clamp being left behind at the edge of a ssDNA gap that may often include a template lesion (Figure 7.1d).

RarA shows strong lagging-strand bias *in vitro*. We do not know if this bias is real, or if action on the leading strand is obscured by the design of the *in vitro* system. It is possible that lesion skipping takes another form on the leading strand, either with no separate catalyst (275, 346) or with catalysis by an enzyme distinct from RarA. If RarA does maintain a lagging strand bias during its normal operations in cells, this would create specific substrates for TLS on the lagging strand and could conceivably be the origin of the lagging strand bias exhibited by TLS polymerases (347–352).

7.3.1 Why do cells maintain a gap creating activity?

The action of RarA necessitates RecF- and O-dependent daughter-strand gap repair, as well as TLS. In the absence of RarA the requirements for these pathways are greatly reduced. In fact, cells appear to survive UV damage somewhat better in the absence of RarA in some experiments (Figures 7.4, 7.5),

although cell growth is greatly slowed. If the gaps disappear and TLS is no longer required, repair would presumably also be less mutagenic. So why would cells maintain an enzyme that creates gaps in the DNA?

Almost all organisms have a gene encoding a RarA homolog, suggesting that gap creation offers a significant advantage. The results of the current study suggest that the advantage endowed by RarA is that it ensures optimal rates of chromosomal replication (allowing lesions to be taken care of behind the fork). *E. coli* mutants lacking RarA activity exhibit a significant growth defect and consequently exhibit reduced fitness relative to wild type cells. Creation of gaps behind the replication fork by RarA may represent a simple trade-off for rapid DNA replication and prolific cell growth.

7.3.2 What is the trigger for gap formation?

There are no lesions purposefully introduced into the rolling-circle DNA substrate used in the single molecule experiments of Figure 7.1. Gap creation in this system is a function of RarA concentration, and requires concentrations greater than those likely to be present *in vivo*. At present, the trigger for RarA action is unknown. *In vivo*, there is a clear connection between RarA function and DNA damage. A replisome pause at the site of a lesion is a plausible trigger *in vivo*, and pausing due to other types of barriers could also contribute. *In vitro*, replisome pausing is regularly observed in single molecule studies of replication, even on undamaged templates (52, 353–357). The gap generation of Figure 7.1 may amplify a function that is more focused on pausing at particular types of lesions *in vivo*.

7.3.3 Promotion of lagging-strand gap creation

The concept of gap creation as a means of DNA damage tolerance has a five-decade history in the literature of bacterial DNA metabolism (358–362). It was first embodied in the concept of post-replication repair, proposed by Howard-Flanders and colleagues (361, 363). These researchers found that low-level UV irradiation of cells defective in nucleotide excision repair did not entirely block DNA replication, but instead led to the appearance of shorter nascent DNA strands. The lengths roughly corresponded to the inter-lesion distance (364, 365). The observations implied that lesions had been bypassed so as to leave them behind in single-strand gaps to be repaired after replication had moved on. At longer times, the shorter DNAs were gradually incorporated into chromosome-sized DNA molecules. Many predictions of the model were subsequently borne out *in vivo* (366, 367). As the idea matured, it was later called daughter-strand gap repair (367), and it became the basis for studies of RecA-mediated gap repair and the specialized functions of proteins such as RecF, O and R (338, 339, 359, 368–371). The lesions left behind by lesion skipping would be repaired either by TLS (372) or by recombinational DNA repair (362, 373). These ideas, and the interplay between the different paths for

DNA repair in bacteria continue to be explored (264, 274, 282, 366, 374, 375).

In principle, TLS polymerases could act not only in gaps, but also at the replisome itself via polymerase exchange. DNA polymerase switching (from a DNA polymerase core to DNA polymerase II or IV) has been demonstrated by several groups (142, 275, 278, 279) and may also contribute to the bypass of lesions encountered by the replisome. This type of replisomal TLS would circumvent the need for lesion skipping and TLS in the resulting gaps. However, a role for TLS in gaps created by lesion skipping has also become prominent in the recent literature. Marians and colleagues have demonstrated that DNA polymerase III has an inherent capacity to undergo lesion-skipping, focusing primarily on the leading strand, and on a timescale of minutes (256, 275, 346, 376–378). The results of an *in vitro* study suggest that replisomal TLS and leading strand lesion skipping are competing mechanisms (275). The importance of DNA polymerases II, IV and V to carry out postreplicative TLS in the resulting gaps has been pointed out (272, 274). Recent single-molecule imaging studies (110, 345) have indicated that DNA polymerases IV and V actually spend little time at the replisome, with many non-replisomal genomic foci present under many conditions. The results are consistent with a frequent use of postreplicative TLS-mediated gap filling during cellular replication.

The idea that an enzymatic activity may exist to facilitate lesion skipping on the lagging strand has no precedent, and to our knowledge is unanticipated. Previous studies conducted *in vitro* have concluded that lagging strand lesion skipping is an inherent property of the replisome (328, 379), but the current study suggests a much different situation. The activities of RecF and RecO become partly dispensable upon deletion of rarA (Figure 7.4), as do the activities of all three TLS polymerases (Figure 7.5). The cells also grow much more slowly. This suggests that *in vivo* few, if any, ssDNA gaps are formed in the absence of RarA. Conversely, many ssDNA gaps are formed in its presence. In a previous *in vivo* study, PagÃÍs and Fuchs found that the presence of a lesion on the lagging strand triggered the formation of a lesion-containing ssDNA gap, but did not hinder the progress of the replisome (372). Since the *E. coli* strain used in the study (JM103) was rarA+, it now appears likely that the gaps were created by the action of RarA.

In the absence of RarA, lesions on the lagging strand template could stall the replication fork if the lagging strand polymerase were to become stuck and the exchange of Pol III* into the replication fork become limiting. In the absence of exogenous DNA damage, Pol III* exchanges readily *in vivo* (Beattie et al., 2017, Lewis et al., 2017). Exchange could easily become limiting in the presence of damage, however. There are only $\sim$20 molecules of Pol III* available in each cell (236, 239). If an exchange-driven lesion-skipping mechanism were at play in cells, Pol III* complexes would initially be deposited on lesions left behind the replisome. This would soon deplete the cell's supply of Pol III* and exchange would cease. By detaching Pol III* from its β_2 clamp, RarA could liberate the lagging strand

polymerase from the lesion and allow the Pol III* complex to progress downstream without exchanging.

7.3.4 What is the mechanism of polymerase detachment?

The RarA protein is homologous to clamp loader proteins (308), is present at the replication fork (304), and has an important role in the observed instability of at least two mutant proteins in the replisome (305, 315). All of these observations are consistent with the proposed gap-creation function of RarA. The *in vitro* single molecule data indicate that RarA does not disassemble the entire replisome, as replication continues apace. The appearance of gaps *in vitro* requires the RarA ATPase, and thus is not a simple function of SSB binding. Although more work is needed to explore mechanistic details, the simplest model is one where the clamp-loader-like structure of RarA is utilized to separate a Pol III core from its bound β_2 clamp. Consistent with this model, *in vitro* assays (Figure 7.1) showed that β_2 clamps are left at gaps behind the fork in the presence of RarA. It has not escaped our attention that the N-terminal portion (NH2-SNLSLDF) of *E. coli* RarA is reminiscent of a hexapeptide motif present in other proteins to interact with the β_2 clamp (e.g. QLSLPL in the *E. coli* Hda protein; bold characters indicate residues that come into intimate contact with a hydrophobic pocket on the β_2 clamp) (Dalrymple et al., 2001, Wijffels et al., 2004). Based on a demonstrated interaction of Mgs1/WRNIP1 with DNA polymerase δ, Enomoto and colleagues suggested that Mgs1 and WRNIP1 might act to detach DNA polymerase δ from PCNA (316).

Detachment of the lagging-strand polymerase from the clamp could have a destabilizing effect on the replisome, and explain why dnaE486 and holD mutants support growth under nonpermissive conditions only when RarA function is absent (305, 315). The dnaE486 variant encodes a temperature-sensitive allele in the α-catalytic subunit of DNA polymerase III (313, 314) that limits growth at 38°C. The loss of rarA function suppresses the lethality of the dnaE486 allele at restrictive temperatures (305). Loss of rarA also suppresses the effects of mutations that inactivate holD (315). These results indicate that RarA is somehow destabilizing the replisome, and that the phenotype of these replisome alterations is dependent on RarA. The holD gene encodes the ψ subunit of DNA polymerase III, which while nonessential, destabilizes the replisome when it is absent (315). When Pol III replisomes are destabilized, TLS polymerase activity can potentially increase to fill the resulting void.

All RarA-family proteins exhibit a DNA-dependent ATPase activity *in vitro* (305, 308, 317, 318, 380). The DNA interaction underlying the ATPase activity specifically targets duplex DNA ends and gap boundaries (320, 321). Recently, we demonstrated that RarA possesses an ATP-dependent DNA flap creation activity, and proposed a model for how this might facilitate DNA damage tolerance in some situations (321). This activity also provides a molecular basis for an elevated rate of damaged chromatid loss that was recently observed in recA mutants (381).

Although it could be most easily rationalized in the context of the replication fork, and represents another potential link between RarA and the replisome, it is not yet clear how this activity would mesh with the results described here. However, a flap forming activity, particularly one targeted to gaps, might be useful in facilitating aspects of lesion skipping or RecA loading.

7.3.5 Implications of gap creation for TLS

The results of the current study raise a host of new questions about the actions of TLS polymerases. To what extent does TLS polymerase activity require RarA? Is RarA action sufficient to form substrates for TLS polymerases in the absence of exogenous DNA damage? What types of DNA lesions are subject to lesion skipping, and what types represent barriers that require replisome stalling and repair by other paths? Does RarA have a lagging strand bias as observed in the *in vitro* assays, and could this explain the tendency for TLS polymerases to act on the lagging strand? Is the SOS response, which is triggered by the formation of RecA* nucleoprotein filaments on ssDNA, in effect governed by RarA activity?

The current study leads us to propose a new idea: that an enzymatic activity exists to facilitate lesion skipping by the replisome. This idea has strong implications for our understanding of TLS, but also more generally for our understanding of how cells channel damaged DNA substrates into different repair pathways. Under normal growth conditions, many more DNA lesions may be encountered by the cellular replication fork than previously appreciated, most of them skipped over in a process undetected by most approaches used to date.

7.4 Materials and methods

7.4.1 Replication proteins

E. coli DNA replication proteins were produced as described previously: the β_2 sliding clamp (210), SSB (211), the $DnaB_6(DnaC)_6$ helicase–loader complex (212), DnaG primase (213), the Pol III $\tau_n\gamma_{(3n)}\delta\delta'\chi\psi$ clamp loader (202), Pol III $\alpha\epsilon\theta$ core (202), and wild type, and K63R mutant RarA proteins (308). All proteins were carefully tested for endo- and exonuclease contamination using gel-based DNA degradation assays utilizing supercoiled and linear dsDNAs and circular and linear ssDNAs. No contaminating endo- or exonucleases were detected. Aliquots of purified proteins were thawed fresh from –80°C stocks prior to each experiment. RarA protein concentration was determined using the native extinction coefficient ϵ = 5.44 · 104 $M^{-1}cm^1$ (308).

7.4.2 Labeling of β_2 with AF647

β_2 labeling reactions were carried out at a protein concentration of 140 μM (as a dimer) at room temperature in 500 μL of labeling buffer (50 mM Tris.HCl pH 7.6, 3 mM dithiothreitol, 1 mM EDTA, 100 mM NaCl, 20% (v/v) glycerol). A 4-fold

molar excess of Alexa Fluor 647 carboxylic acid, succinimidyl ester (Invitrogen) dissolved in anhydrous DMSO was added to the protein solution and allowed to react for 1.5 h in the dark, yielding Fraction I. Fraction I was centrifuged (21,000 x g; 15 min) at 6°C and the supernatant carefully removed to yield Fraction II. Fraction II was applied at 1 ml/min to a column (1.5 x 10 cm) of Superdex G-25 resin (GE-Healthcare) equilibrated with gel filtration buffer (50 mM Tris.HCl pH 7.6, 3 mM dithiothreitol, 1 mM EDTA, 100 mM NaCl, 5% (v/v) glycerol) to remove unreacted fluorophores. Fractions containing the labeled β_2 were pooled and dialyzed into storage buffer (50 mM Tris.HCl pH 7.6, 3 mM dithiothreitol, 1 mM EDTA, 100 mM NaCl, 20% (v/v) glycerol). The degree of labeling was determined by UV/vis spectroscopy to be $\sim$1 fluorophore per β dimer.

7.4.3 *In vitro* single-molecule rolling-circle DNA replication assay

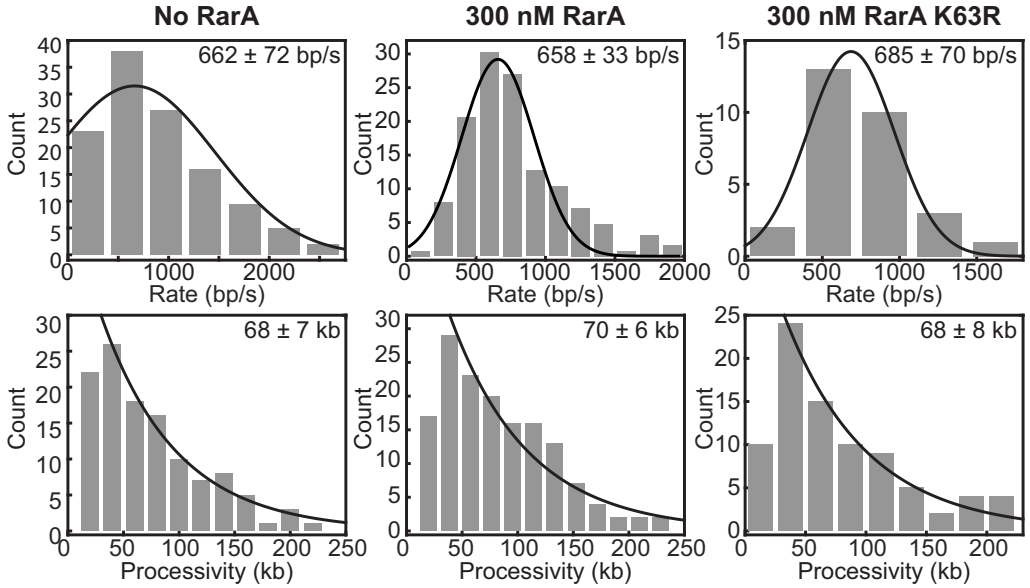


Figure 7.7: Histograms of rates and processivities for pre-assembled replisomes. The rate histograms were fit with a Gaussian distribution (black line) to obtain the mean rate. The rates are 662 $\pm$ 72 bp/s without RarA (top, left), 658 $\pm$ 33 bp/s with 300 nM RarA (top, middle), and 685 $\pm$ 70 bp/s with 300 nM RarA K63R (top right). Processivity distributions were fit with a single-exponential decay function (black line). The processivities are 68 $\pm$ 7 kb without RarA (bottom, left), 70 $\pm$ 6 kb in the presence of 300 nM RarA (bottom, middle), and 68 $\pm$ 8 kb with 300 nM RarA K63R (bottom, right). These data show that RarA does not affect the rate of replication and processivity under these conditions. The errors represent the s.e.m.

Microfluidic flow cells were prepared as described (51). Briefly, a PDMS flow chamber was placed on top of a PEG-biotin-functionalized microscope coverslip. To help prevent non-specific interactions of proteins and DNA with the surface, the chamber was blocked with blocking agent (NEB, Ipswich, MA). The chamber was placed on an inverted microscope (Nikon Eclipse Ti-E) with a CFI Apo TIRF 100x oil-immersion TIRF objective (NA 1.49, Nikon, Japan) and connected to a syringe pump (Adelab Scientific, Australia) for flow of buffer. Reactions

were carried out 31°C, maintained by an electrically heated chamber (Okolab, Burlingame, CA). Double-stranded DNA was visualized in real time by staining it with 150 nM SYTOX orange (Invitrogen) excited by a 568 nm laser (Coherent, Santa Clara, CA; Sapphire 568–200 CW) at 150 mW/cm^2. The red labeled β_2 was excited at 700 mW/cm^2 with a 647 nm (Coherent, Obis 647–100 CW) laser. For dual-color imaging the signals were separated via dichroic mirrors and appropriate filter sets (Chroma, Bellows Falls, VT). Imaging was done with an EMCCD camera(Photometics, Tucson, AZ; Evolve 512 Delta).

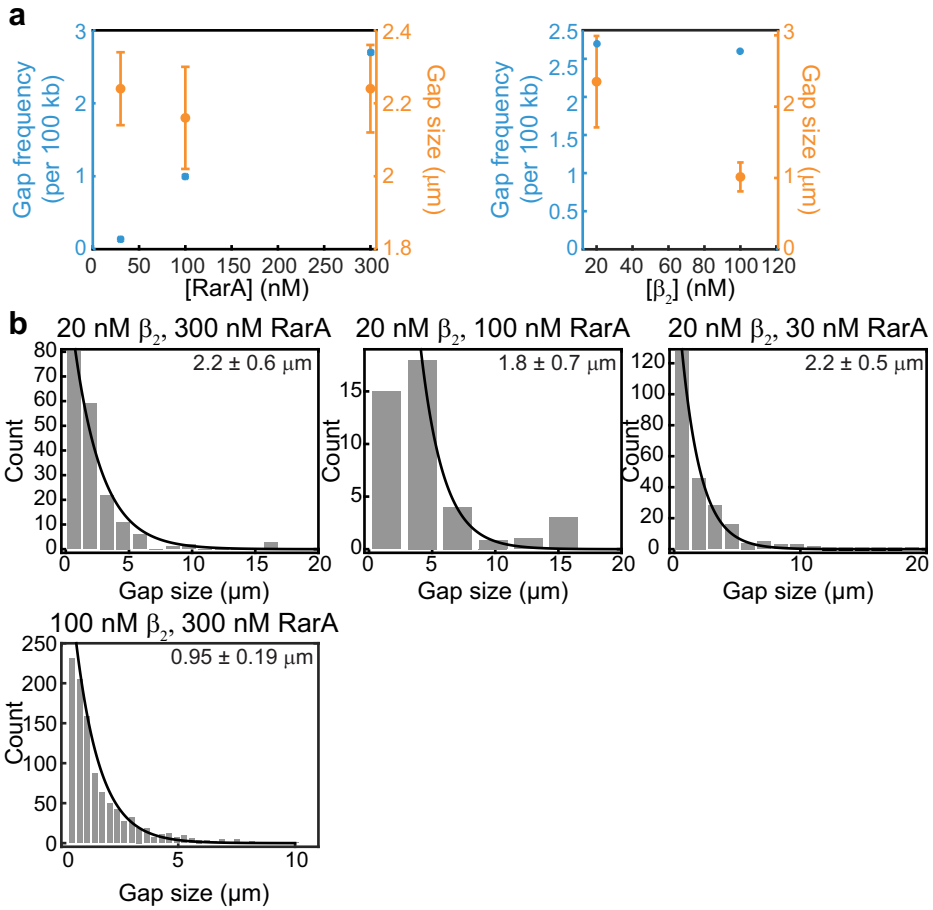


Figure 7.8: Comparison of gap sizes and gap frequencies for RarA concentrations of 30, 100 and 300 nM. (a) Gap frequency (blue) and gap size (orange) as a function of RarA concentration (left). The gap frequency increases with increasing concentrations of RarA. The gap size, obtained from histograms in (b) remains constant within this range of RarA concentrations. Gap frequency and gap size for two concentrations of β_2. The gap size changes as a function of β_2 concentration, whereas the frequency remains constant. The error bars represent the s.e.m. (b) Histograms of the gap size for RarA concentrations of 300 nM (top, left), 100 nM (top, middle), and 30 nM (top, right). The histograms were fit with single-exponential decay functions (black lines) to obtain the average gap size. The gap sizes are 2.2 $\pm$ 0.6, 1.8 $\pm$ 0.7, and 2.2 $\pm$ 0.5 μm for 300, 100, and 30 nM RarA, respectively, in the presence of 20 nM β_2. In the presence of 100 nM β_2 and 300 nM RarA (bottom, left) the gap size changed to 0.95 $\pm$ 0.19 μm. The errors represent the s.e.m.

Conditions for the pre-assembly replication reactions were adapted from published methods (136,204,236). Solution 1 was prepared as 30 nM $DnaB_6(DnaC)_6$, 1.5 nM biotinylated circular 2 kb dsDNA substrate and 1 mM ATP in replication buffer. This was incubated at 37°C for 3 min. Solution 2 contained 60 μM dCTP and dGTP, 3.3 nM Pol III* (assembled in situ by incubating $\tau_3\delta\delta'\chi\psi$ (410 nM) and Pol III cores $\alpha\epsilon\theta$ (1.2 μM) in replication buffer at 37°C for 90 s), and 74 nM β_2 in replication buffer (without dATP and dTTP). Solution 2 was added to an equal volume of solution 1 and incubated for 6 min at 37°C. This was then loaded onto the flow cell at 100 μl/min for 1 min and then 10 μl/min for 10 min. The flow cell was washed with replication buffer containing 60 μM dCTP and dGTP. An imaging buffer was made with 1 mM UV-aged Trolox, 0.8% (w/v) glucose, 0.12 mg/ml glucose oxidase, and 0.012 mg/ml catalase (to increase the lifetime of the fluorophores and reduce blinking), 1 mM ATP, 250 μM CTP, GTP and UTP, and 50 μM dCTP, dGTP, dATP and dTTP in replication buffer. Replication was finally initiated by flowing in the imaging buffer containing 20 nM β_2, 75 nM DnaG, 250 nM SSB_4, and RarA when indicated, at 10 μl/min. All *in vitro* single-molecule experiments were performed at least four times. The analysis was done with ImageJ using in-house built plugins. The rate of replication of a single molecule was obtained from its trajectory and calculated for each segment that has constant slope.

7.4.4 Fluorescence polarization assay

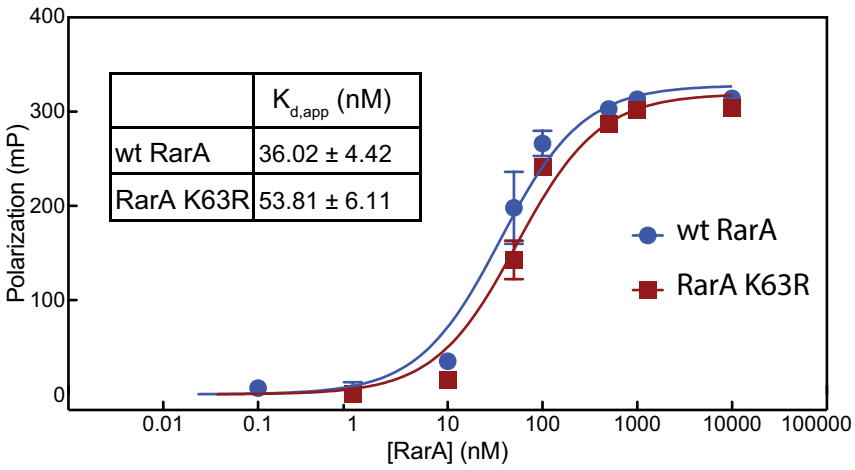


Figure 7.9: RarA K63R binds the SSB C-terminal tail peptide with an affinity similar to wild type RarA. Wild type RarA or RarA K63R protein was incubated with fluorescently labeled *E. coli* SSB peptide at room temperature. Following incubation for 30 min, fluorescence polarization values were measured. Each point represents an average polarization value for a reaction containing the indicated concentration of RarA protein, while the error bars represent one S.D. from the average polarization value. Data were fit to a simple single-site binding curve and apparent dissociation constants ($K_{d,app}$) were calculated.

Indicated concentrations of wild type RarA or RarA K63R were incubated with 5 nM fluorescein-labeled *E. coli* SSB C-terminal peptide (Fluor-Trp-Met-Asp-Phe-Asp-Asp-Ile-Pro-Phe) for 30 min at room temperature in reaction buffer (25 mM Tris acetate pH = 7.5, 3 mM potassium glutamate, 10 mM magnesium acetate,

5% (w/v) glycerol, and 1 mM dithiothreitol) supplemented with 100 ng/ml BSA. Fluorescence polarization was measured for each sample using a Beacon 2000 fluorescence polarization system (PanVera Corporation, Madison, WI). The polarization values of experimental reactions were background corrected by subtracting the value of SSB peptide alone (44 mP) from each experimental value. Binding data were fit to a simple one-site binding specific interaction model and apparent Kd values were calculated using GraphPad Prism software (GraphPad Software, San Diego, CA).

7.4.5 Reagents and growth conditions

All cell cultures were grown with shaking and aeration in Lysogeny Broth (LB) made fresh from components (Benton, Dickinson and Company, Franklin Lakes, NJ). Antibiotics were used at the following concentrations: kanamycin (40 μg/mL); chloramphenicol (25 μg/mL); nitrofurazone (10 μM); 4-nitroquinoline-1-oxide (10 μM).

7.4.6 Strain construction

All strains are *E. coli* MG1655 derivatives and are listed in Table 1. All parent strains were constructed using Lambda Red recombination as described by Datsenko and Wanner (382). All chromosomal mutations were confirmed using Sanger sequencing. When required, antibiotic resistance of a given strain was eliminated using FLP recombinase encoded by the pLH29 plasmid as described previously (218). Strains were transformed to harbor indicated plasmids using conventional methods.

7.4.7 Growth curves — plate reader

Overnight cultures of indicated strains were diluted 1:100 in LB. Three biological replicates were prepared in a clear bottom 96 well plate (Corning, Corning, NY). Cultures were grown at 37°C with continuous orbital shaking in a BioTek Synergy 2 (BioTek Instruments Inc., Winooski, VT) plate reader. OD_{600} values were taken every 10 minutes for over the course of the experiment. OD_{600} values were normalized by subtracting out the OD_{600} value of a blank sample containing only LB. Growth curves represent the average of at least three separate experiments, with error bars representing one S.D. from the average value.

7.4.8 Growth curves — spectrophotometer

Overnight cultures of indicated strains were diluted 1:100 in LB. Three biological replicates were used to measure a growth curve for a particular strain. Cultures were grown in 50 ml culture flasks at 37°C with shaking. 1 ml aliquots were taken at indicated time points and the OD_{600} value was read using a Cary 50 UV-vis spectrophotometer, diluting in fresh LB as necessary to achieve an accurate measurement. The three biological replicates were averaged to produce

the curves shown, with error bars representing the standard deviation from the averaged OD_{600} value.

7.4.9 Growth competition assays

Growth competition assays were conducted as previously described (319) using a method originally described by Lenski (334). The $\triangle$araBAD $\triangle$ParaB marker was included on either wild type (MG1655) or mutant ($\triangle$rarA) in separate experiments to control for any effect the marker may have had on cell fitness. Points represent average colony counts from at least three biological replicate experiments, with error bars representing one standard deviation from the average colony count.

7.4.10 Single-molecule time-lapse imaging and analysis

Cells were grown at 37°C in EZ rich defined medium (Teknova, Hollister, CA) that included 0.2% (w/v) glycerol. Flow cells were constructed by affixing a quartz piece, embedded with inlet and outlet tubing, to a (3-aminopropyl)triethoxysilane (APTES, Alfa Aesar, Haverhill, MA) functionalized glass coverslip according to the design specifications and procedure detailed in (110).
For time-lapse imaging, flow cells were mounted to the microscope where temperature was maintained at 37°C by a combination of stage and objective lens heating. Cells were grown in shaking culture until reaching mid-log phase, then loaded into mounted flow cells by pulling through the outlet with a syringe. The inlet was then placed into fresh, constantly aerated, EZ medium. Aerated medium was then pulled through the flow cell using a syringe pump, at a rate of 50μl/min.

We used a microscope constructed to the specifications detailed in (110). All DnaQ- and DnaX-YPet replisome labeled strains were excited at 18 W/cm^2 with 514 nm laser excitation light for 500 ms. Imaging was done with a 512 x 512 pixel EM-CCD camera (C9100-13, Hamamatsu, Japan). Image processing was performed in ImageJ using custom plugins. All cell outlining was manually done in the open source MicrobeTracker Suite in MATLAB R2013b (Mathworks).

7.4.11 Single-molecule fluorescence imaging of cells grown in shaking culture

For comparison of *rarA+* and $\triangle$*rarA* phenotypes, EAW170 (*dnaQ-YPet rarA+*) and THS04 (*dnaQ-YPet* $\triangle$*rarA*) cells were grown overnight at 37°C in EZ glycerol, diluted 1000-fold in fresh medium, then grown an additional 3 h prior to imaging. To determine if leaky expression of rarA from a pBAD plasmid could complement the $\triangle$rarA small cell phenotype, THS56 (dnaQ-YPet $\triangle$rarA pBAD-rarA) cells were prepared in EZ glycerol and EZ glycerol + 0.2% glucose. Cells were grown overnight at 37°C in EZ glycerol with 0.2% glucose, diluted 1000-fold, then grown an additional 3 h to repress expression from the pBAD plasmid.

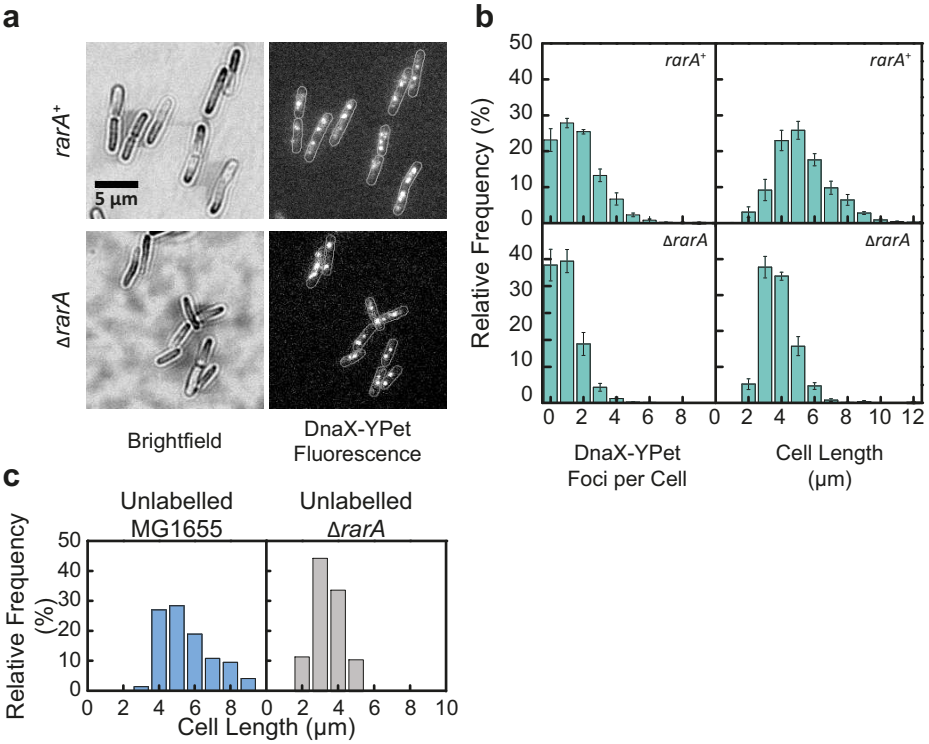


Figure 7.11: DnaX-YPet replisome labeled rarA+ 1499 and ΔrarA cells display similar trends in focus count and cell size as DnaQ-YPet labeled cells. (a) Representative images. (b) Histograms of DnaX-YPet replisome foci per cell and cell length are shown for rarA+ (JJC5945) and ΔrarA (THS05) strains. Cells were grown at 37°C in flow cells and imaged every 5 min for 180 min. Error bars represent the standard error of the mean values for each bin across at least two biological replicates. DnaX-YPet foci per cell for rarA+ 1505 cells: mean = 1.8; SEM = 0.03; *n* = 2933 cells. DnaX-YPet foci per cell for ΔrarA cells: mean = 1.2; SEM = 0.02; *n* = 1695 cells. Cell size for rarA+ cells: mean = 5.4; SEM = 0.03 μm in length; *n* = 2927 cells. Cell size for ΔrarA cells: mean = 3.9; SEM = 0.02 μm in length; *n* = 2369 cells. (c) Cells expressing wild type replisome proteins display similar rarA-dependent trends in cell size as DnaQ-YPet labeled cells. Histograms of cell length are shown for rarA+ (MG1655, light gray) and ΔrarA (EAW98, dark gray) strains. Cells in each condition were grown in EZ glycerol at 37°C for 3 h prior to imaging in coverslip sandwiches (See Methods). Cell size for rarA+ cells: mean = 5.4; SEM = 0.16 μm in length; *n* = 76 cells. Cell size for ΔrarA cells: mean = 3.1; SEM = 0.08 μm in length; *n* = 86 cells.

An aliquot (100 μl) was removed for imaging. The remaining cells were pelleted by centrifugation, resuspended in EZ glycerol, then diluted 100-fold in fresh EZ glycerol medium. This culture was grown at 37°C for 3 additional hours prior to imaging. For imaging, glass coverslips were either functionalized with APTES according to the procedure detailed in (110) or cleaned by sonicating with 5 M KOH for 25 min, rinsing with fresh MilliQ water, and drying with N2. Coverslip sandwiches were prepared by flattening 20 Âţl of cell suspension between an APTES-functionalized (bottom) and KOH-cleaned (top) cover slide. Cells were then imaged in both brightfield and fluorescence channels on the microscope described in the previous section. Image processing was again preformed in ImageJ using custom plugins after manually outlining cells in the MicrobeTracker Suite in MATLAB R2013b (Mathworks).

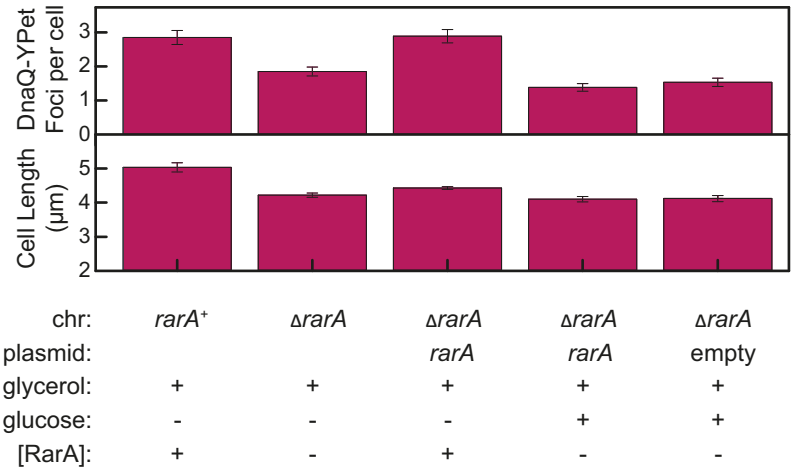


Figure 7.12: Leaky expression of rarA from a pBAD plasmid returns rarA+ phenotype to ΔrarA cells. Quantification of the mean DnaQ-YPet replisome foci per cell (top) and cell length (bottom) are shown for rarA+ (EAW170) and ΔrarA (THS04) cells grown in EZ glycerol and ΔrarA pBAD-rarA (THS56) cells grown in both EZ glycerol and EZ glycerol with 0.2% glucose. Cells in each condition were grown at 37°C for 3 h prior to imaging in coverslip sandwiches (See Methods). Error bars represent the standard error of the mean for > 50 cells. DnaQ-YPet foci per cell for rarA+ cells: mean = 2.4; SEM = 0.2; *n* = 58 cells. DnaQ-YPet foci per cell for ΔrarA cells: mean = 1.9; SEM = 0.1; *n* = 67 cells. DnaQ-YPet foci per cell for ΔrarA pTHS02 cells in EZ glycerol + glucose: mean = 1.4; SEM = 0.1; *n* = 69 cells. DnaQ-YPet foci per cell for ΔrarA cells pTHS02 in EZ glycerol: mean = 2.9; SEM = 0.2; n = 50 cells. Cell size for rarA+ cells: mean = 4.7; SEM = 0.23 μm in length; *n* = 57 cells. Cell size for ΔrarA cells: mean = 3.6; SEM = 0.10 μm in length; *n* = 67 cells. Cell size for ΔrarA pTHS02 cells in EZ glycerol + glucose: mean = 3.4; SEM = 0.13 μm in length; *n* = 69 cells. Cell size for ΔrarA cells pTHS02 in EZ glycerol: mean = 4.0; SEM = 0.17 μm in length; *n* = 50 cells.

7.4.12 Flow cytometry

Overnight cultures of indicated strains were diluted 1:100 in fresh LB medium in biological triplicate from three separate bacterial colonies. Cultures were grown at 37°C with aeration and shaking until the OD600 measured 0.2. Aliquots (500 μl) were taken from each culture and placed on ice. Cells were collected via centrifugation and washed with 1 ml of phosphate buffered saline (PBS) three times. These aliquots were diluted 1:10 in fresh PBS, then either mock stained or stained with the DNA detection dye Vybrant DyeCycle Green (Thermo Fisher Scientific, Waltham, MA) according to manufacturer's instructions (final dye concentration 10 μM).

Samples were measured using a BD Accuri Flow Cytometer (BD Biosciences, San Jose, CA) equipped with a 488 nm wavelength excitation laser and 533 nm wavelength emission filter with a 30 nm bandwidth. A sample threshold of 10,000 FSC-A was set. At least 50,000 cells were measured for each biological replicate, with three total biological replicates being measured per strain. Light scattering and fluorescence data were imported into FlowJo (FlowJo LLC, Ashland OR) and gated as indicated to exclude cell debris or doublets.

7.4.13 Spot plate drug/UV sensitivity assays

Overnight cultures of indicated strains were diluted 1:100 in fresh LB medium. Cultures were grown in biological triplicate at 37°C with aeration and shaking until the OD_{600} reached 0.2. Aliquots (1 ml) were taken from each culture, serially diluted in LB to 10–6, then 10 μl were spot plated on agar plates containing the indicated media. When UV irradiation was used, plates were exposed to the indicated amount of UV irradiation (254 nm) using a Spectrolinker XL-1500 UV crosslinker (Spectronics Corporation, Westbury, NY). Plates were incubated overnight at 37°C and imaged the next day using an ImageQuant LAS 4010 imaging system (GE Healthcare, Marlborough, MA).

7.5 Supplementary figures

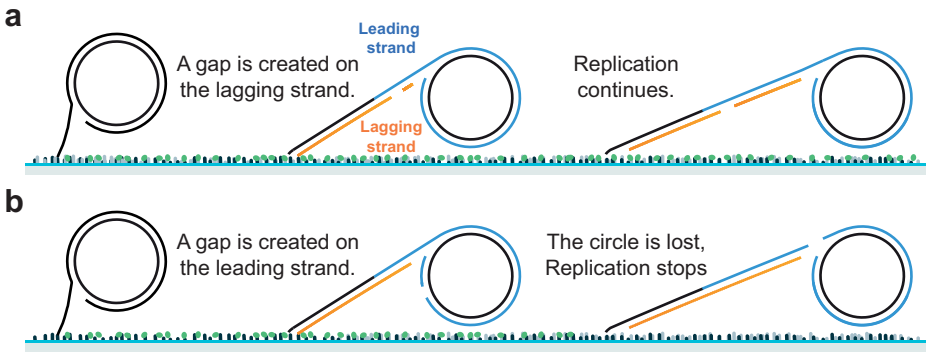


Figure 7.14: Diagram explaining termination of the rolling circle reaction by a gap in the leading-strand template. (a) When a gap is created in the lagging strand, lagging-strand synthesis can restart by re-priming farther downstream and replication will continue. (b) If a gap is created in the leading strand and replication is restarted by re-priming, the replisome will encounter this gap the next time it goes around the circle. This will result in loss of the circle and termination of replication. As a result we would observe a significant reduction in the processivity.

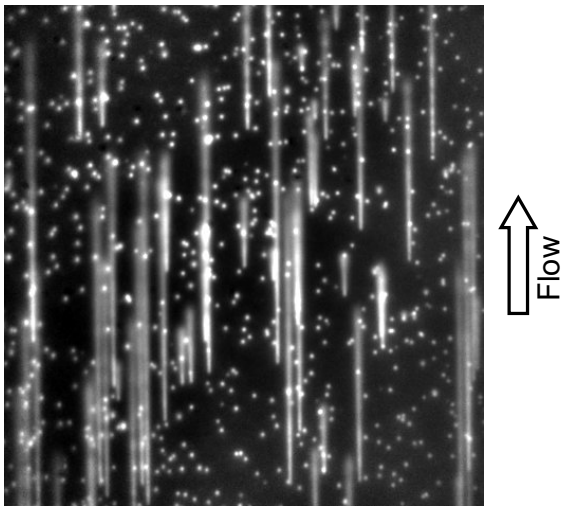


Figure 7.15: Post-replicative addition of RarA does not lead to gap formation. Representative field of view showing DNA when RarA was loaded after the replication reaction had stopped. RarA does not create gaps on the DNA outside the context of the replisome.

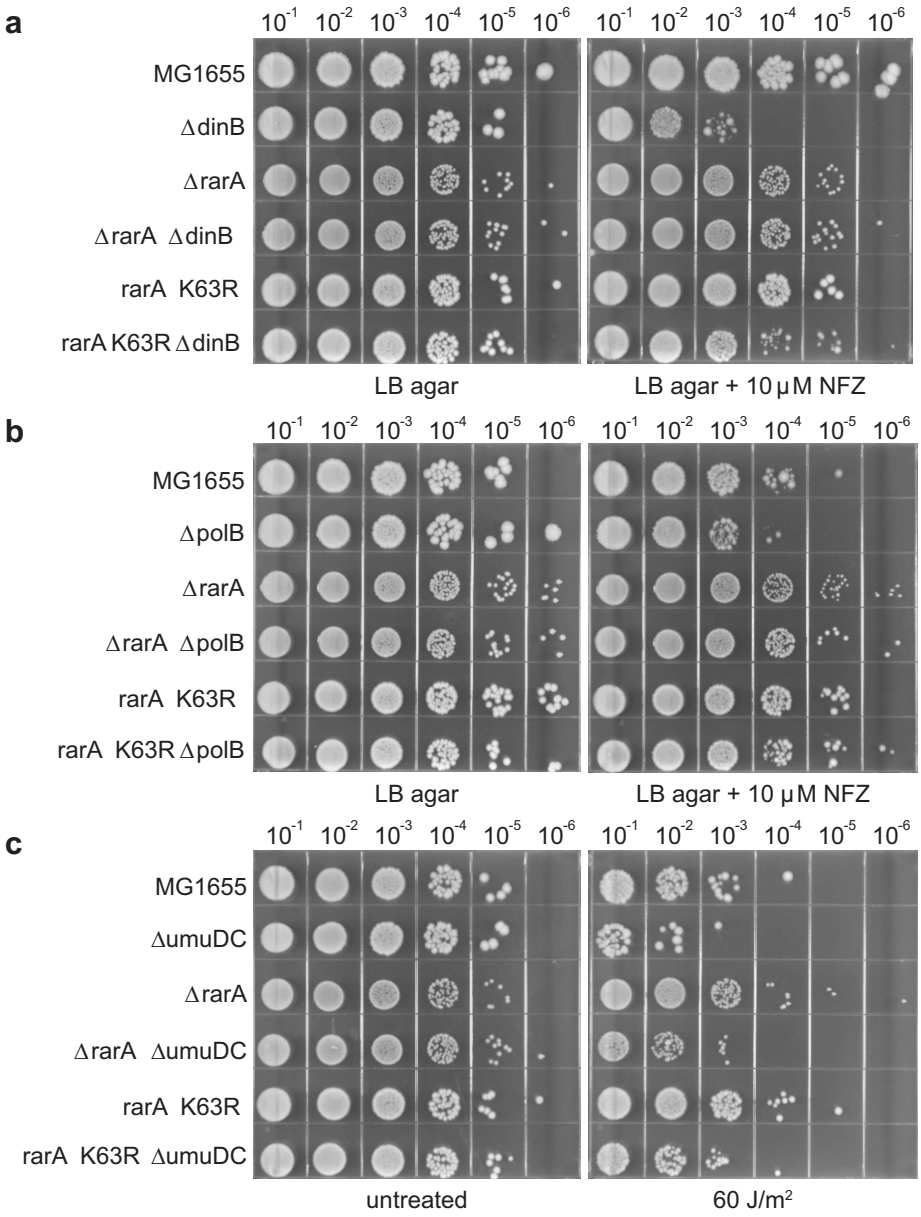


Figure 7.5 ***(preceding page)*****: Deletion of rarA suppresses the sensitivity of ΔpolB, ΔdinB, and ΔumuDC strains to DNA damage.** Indicated strains were grown to exponential growth phase (OD_{600} = 0.2), serially diluted and spot plated onto LB agar plates containing indicated media. (a) Loss of RarA ATPase activity suppreses the sensitivity of ΔdinB (Pol IV) cells to the mutagen nitrofurazone (NFZ). Cells lacking the dinB gene are sensitized to NFZ by 2–3 orders of magnitude compared to dinB+ cells. Loss of rarA or its ATPase activity (K63R) confers no sensitivity to NFZ. In a ΔdinB background, loss of rarA or its ATPase activity restores cell viability to wild type levels. (b) Loss of RarA ATPase activity suppresses the sensitivity of ΔpolB (Pol II) cells to NFZ. Cells lacking polB are mildly sensitized to NFZ by 1–2 orders of magnitude compared to wild type cells. In a ΔpolB background, loss of rarA or its ATPase activity restores cell viability to wild type levels. (c) Loss of RarA ATPase activity partially suppresses the sensitivity of ΔumuDC (Pol V) mutants to ultraviolet light. Cells lacking Pol V are sensitized to high levels of UV light (60 J/m^2) compared to wild type cells. Deletion of rarA or inactivation of its ATPase activity (K63R) results in a 1 order of magnitude increase in resistence to UV light at this incident dose. Deletion of rarA or inactivation of its ATPase activity in a ΔumuDC background restores cell viability to wild type levels at this incident dose of UV light.

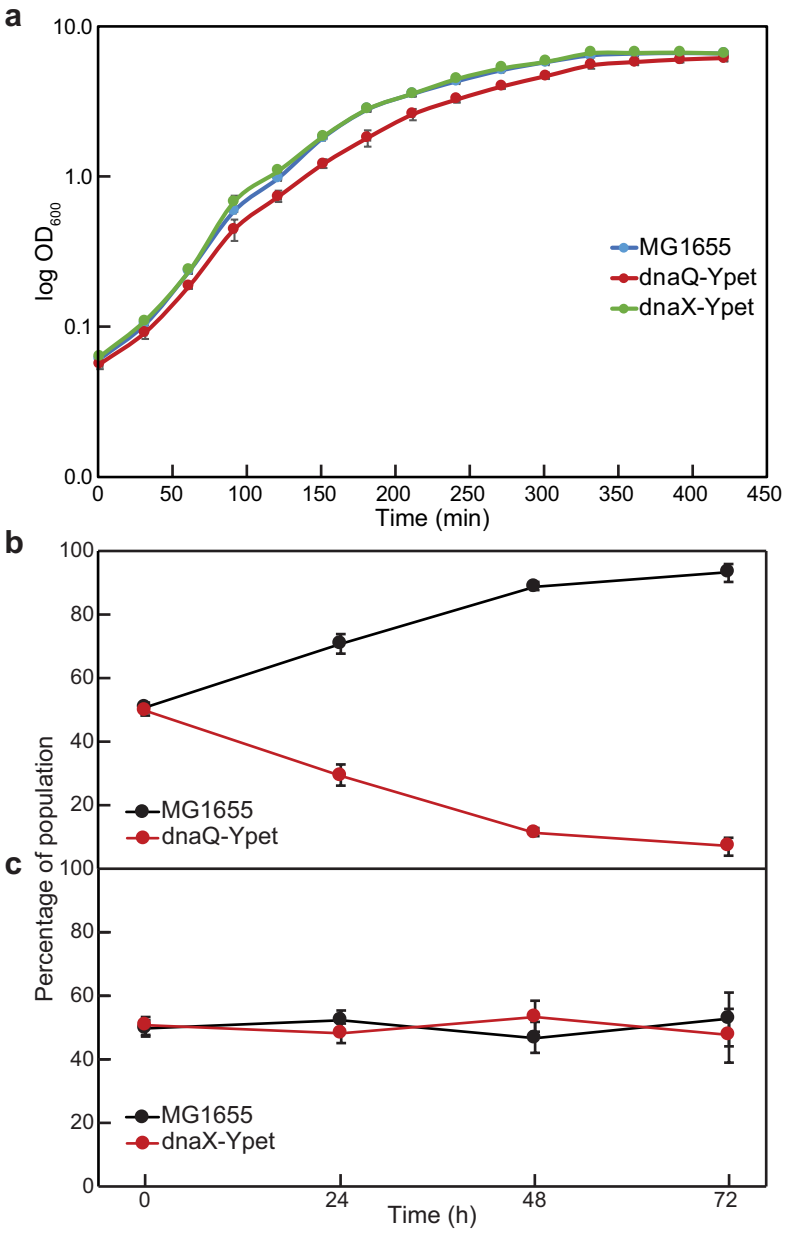


Figure 7.10 ***(preceding page)*****: Replication protein fusion strains exhibit differences in growth and fitness.** (a) Overnight cultures of indicated genotype were diluted 1:100 in fresh LB and grown at 37°C with shaking. OD_{600} values were read every 30 min for 7 h. Growth curves represent averages of at least three independent biological replicates. (b) Growth competition assays were conducted as in Figure 3B. dnaQ-YPet cells are outcompeted by wild type MG1655 cells within 72 h. Each data point represents the mean of at least three biological replicates, with error bars representing the standard deviation from the mean value. (c) Growth competition assays were conducted as in Figure 7.3b. dnaX-YPet cells appear to have no growth defect when compared to wild type MG1655 cells. Each data point represents the mean of at least three biological replicates, with error bars representing the standard deviation from the mean value.

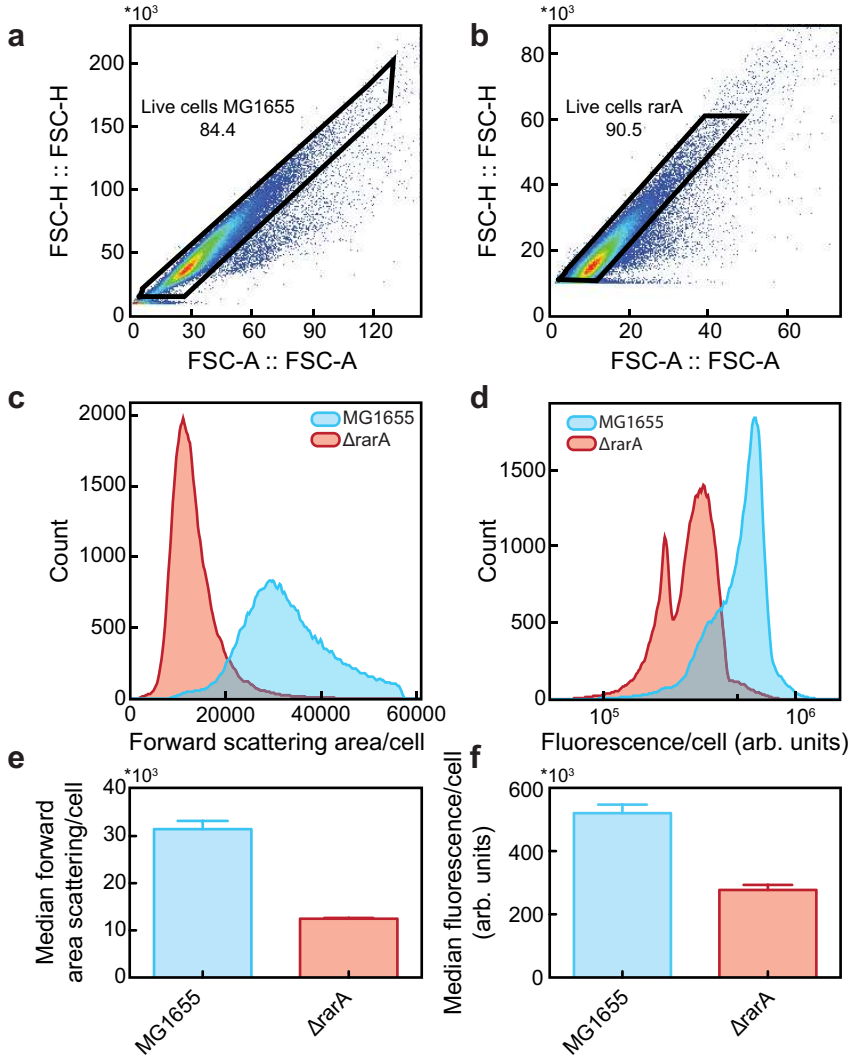


Figure 7.13 ***(preceding page)*****: ΔrarA cells in exponential growth phase are smaller and contain less DNA than wild type cells as assessed by flow cytometry.** Wild type (MG1655) or ΔrarA cells were grown to an OD_{600} of 0.2. Cells were diluted to an appropriate concentration ($\sim$1 x 10^6 cells/ml) and washed with PBS. DNA content was assessed using Vybrant DyeCycle Green as a stain. Each strain was grown up and assayed in biological triplicate, with >50,000 cells being counted for each replicate. (**a**) Wild type cells were further analysis. (**b**) ΔrarA cells were gated to exclude cell debris and doublet as indicated; 90.5% of all cells were included in further analysis. (**c**) Forward scattering area histograms of representative MG1655 (blue) and ΔrarA (red) samples are shown. (**d**) Fluorescence signal representing DNA content for MG1655 (blue) and ΔrarA (red) samples are shown. (**e**) Quantification of the forward scattering area data graphed in (c) is shown for MG1655 (blue) and ΔrarA (red) samples. Bar graphs represent the average of the median forward scattering area values across biological replicates, while error bars represent one standard deviation from the average median value. (**f**) Quantification of the fluorescence data graphed in (d) is shown for MG1655 (blue) and ΔrarA (red) samples. Bar graphs represent the average of the median fluorescence values across biological replicates, while error bars represent one standard deviation from the average median value.

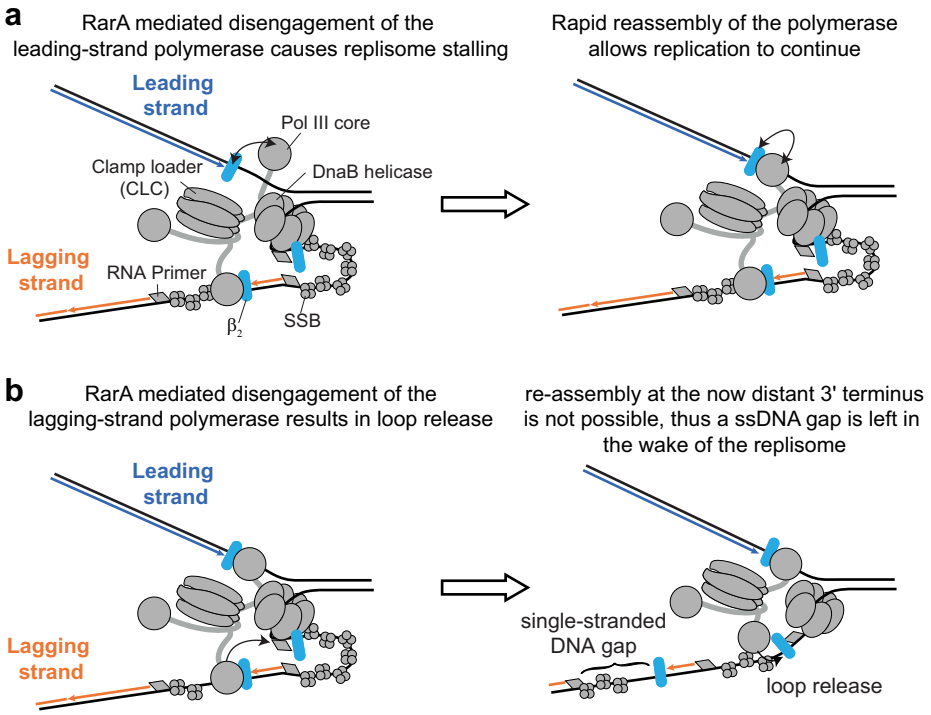


Figure 7.16: Hypothetical mechanisms of RarA-mediated disengagement of the leading- and lagging-strand polymerases. (a) RarA mediated disengagement of the leading-strand polymerase would halt the replisome. Subsequent rapid reassembly of the polymerase would allow replication to continue. A consequence of frequent replisome stalling would be a reduction in the average rate of replication. (b) Disengagement of the lagging-strand polymerase would result in the release of the replication loop. The exposed 3′ terminus is thereby moved away from the replisome, preventing the polymerase from re-engaging. Binding of the lagging-strand polymerase to a newly synthesized primer allows replication to continue, but a ssDNA gap is left in the wake of the replisome.

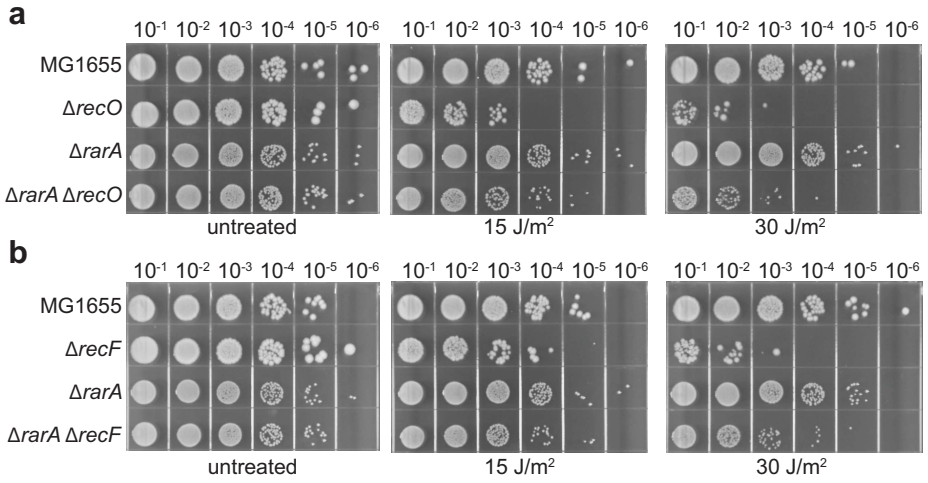


Figure 7.17: Deletion of rarA partially suppresses the sensitivity of ΔrecF and ΔrecO strains to higher levels of UV light. Cells were grown and plated as in Figure 7.4, but also irradiated at an incident dose of 30 J/m^2. (a) ΔrecO cells are sensitized to higher doses of UV light by 2–3 orders of magnitude compared to wild type cells. Deletion of rarA in a ΔrecO background results in a partial (1 order of magnitude) rescue of cell viability, but does not return viability to wild type levels. (b) ΔrecF cells are sensitized to higher doses of UV light by 2–3 orders of magnitude compared to wild type cells. Deletion of rarA in a ΔrecF background results in a partial (1 order of magnitude) rescue of cell viability, but does not return viability to wild type levels.

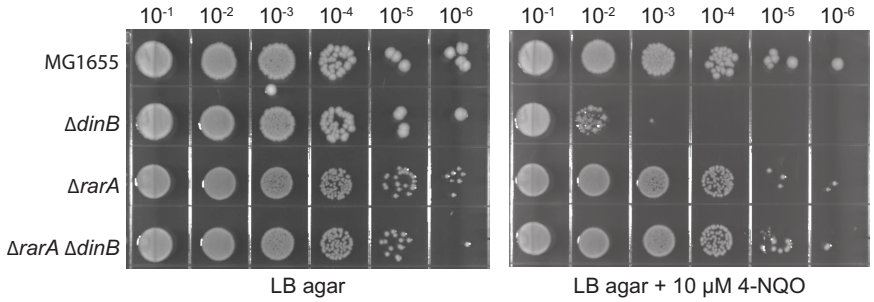


Figure 7.18: Deletion of rarA suppresses the sensitivity of a ΔdinB strain to 4-nitroquinoline-1-oxide (4-NQO). Cells were grown and plated as in Figure 7.5, but instead of NFZ, 4-nitroquinoline-1-oxide (4-NQO) was incorporated into the agar. Cells lacking dinB (Pol IV) are hypersensitized to 4-NQO by 3 orders of magnitude compared to wild type cells. Deletion of rarA results in no observable loss in cell viability on plates containing 4-NQO, and completely rescues the sensitivity of ΔdinB cells to 4-NQO.

8 | Single-molecule visualization of leading-strand synthesis by *S. Cerevisiae* reveals dynamic interaction of MTC with the replisome

Jacob S. Lewis†, **Lisanne M. Spenkelink**†, Grant D. Schauer, Flynn R. Hill, Roxanna Gerogescu, Michael E. O'Donnell, Antoine M. van Oijen.

†These authors contributed equally.

Published in *Proc. Natl. Acad. Sci. U.S.A.*, Sept 2017;114(40): 10630–10635.

The replisome, the multi-protein system responsible for genome duplication, is a highly dynamic complex displaying a large number of different enzyme activities. Recently, the *Saccharomyces cerevisiae* minimal replication reaction has been successfully reconstituted *in vitro* and provides an opportunity to uncover the enzymatic activities of many of the components in a eukaryotic system. Their dynamic behavior and interactions in the context of the replisome, however, remain unclear. We use a tethered-bead assay to provide real-time visualization of leading-strand synthesis by the *S. cerevisiae* replisome at the single-molecule level. The minimal reconstituted leading-strand replisome requires 24 proteins, forming the CMG helicase, the Pol ϵ DNA polymerase, the RFC clamp loader, the PCNA sliding clamp and the RPA single-stranded DNA binding protein. We observe rates and product lengths similar to those obtained from ensemble biochemical experiments. At the single-molecule level, we probe the behavior of two components of the replication progression complex and characterize their interaction with active leading-strand replisomes. The Minichromosome maintenance protein 10 (Mcm10), an important player in CMG activation, increases the number of productive replication events in our assay. Furthermore, we show that the fork protection complex Mrc1–Tof1–Csm3 (MTC) enhances the rate of the leading-strand replisome 3 fold. The introduction of periods of fast replication by MTC leads to an average rate enhancement of a factor of 2, similar to observations in cellular studies. We observe that the MTC complex acts in a dynamic fashion with the moving replisome, leading to alternating phases of slow and fast replication.

J.S.L. and I contributed equally to this paper. I led the development and implementation of all single-molecule experiments and wrote the paper.

8.1 Introduction

The replisome is the molecular machine that coordinates the enzymatic activities required for genome duplication. It contains the proteins responsible for unwinding the duplex, depositing primers, and synthesizing DNA, all while coordinating DNA production on both strands, removing incorrectly incorporated nucleotides, and maintaining high fidelity. The replisome in eukaryotes has evolved to be a sophisticated multi-protein and highly regulated machine; its assembly is performed by origin initiation proteins and kinases that restrict replisome formation and chromosome duplication to a single round to ensure proper ploidy across multiple chromosomes. The operation of the replisome must be finely tuned to adjust to various cellular conditions, and accurately duplicate the large genome size and diploid nature of eukaryotes, the need to maintain epigenetic inheritance and to interface with a large number of repair pathways. While the minimal operating machinery able to advance a replication fork has been established (39, 383), the reactions using minimal sets of reconstituted proteins were unable to achieve rates measured *in vivo*, suggesting additional components that facilitate replisome function. This deficiency is not surprising considering the packaging of chromosomes in eukaryotic cells acting as a physical barrier to the replisome and the abundance of other DNA-bound machineries, such as transcription complexes, that interfere with replication fork progression (384–386). The evolution of checkpoints has provided eukaryotic cells with surveillance mechanisms that orchestrate the recruitment of many other proteins to replication forks whose task it is to modulate replisome activity and facilitate efficient replication of the genome. Using simplified *in vitro* assays, our appreciation for the requirement of these additional proteins, such as those that form the replisome progression complex (RPC), has resulted in the reconstitution of efficient leading- and lagging-strand DNA replication on naked and chromatinized templates *in vitro* (383, 385, 387–390).

Once CMG helicase and the DNA polymerase (Pol) ϵ leading-strand DNA polymerase (together called the CMGE complex) is assembled at the replication fork, many additional proteins are conscripted to the complex to form the RPC. These proteins include Ctf4, Csm3, FACT, Mrc1, Pol ϵ, Tof1, and Top1. This large macromolecular assembly can be isolated from S-phase *S. cerevisiae* cells (391). It has been shown that Mrc1, a yeast homolog of Claspin and an S-phase specific mediator protein of the DNA damage response, is recruited to the fork (392–394) and increases the rate of replication *in vivo* about 2 fold (395–397). *In vitro* studies confirm that Mrc1 increases the speed of replication forks to rates similar to those measured *in vivo* (388). Moreover, inclusion of the Csm3/Tolf1 proteins acted to stimulate the functional association of Mrc1 with the replisome. It has also been demonstrated that Mrc1 binds both the N- and C-terminal halves of Pol2, the subunit that contains the polymerase/exonuclease of Pol ϵ (398). Given that we have only begun to determine the exact roles of the individual proteins at the fork, understanding basic mechanisms during DNA replication that coordinate enzymatic activity has thus far been very challenging. To date,

all *in vitro* methods used to study *S. cerevisiae* DNA replisome activity have relied on traditional biochemical techniques such as gel electrophoresis and/or incorporation of labeled nucleotides (or primers) (39, 383, 385, 387–389). Such experiments have generated a wealth of knowledge, uncovering the enzymatic activities of many of the components, and have provided the molecular mechanisms that target the replicative polymerases to their respective strands during bulk DNA synthesis (39, 383, 387, 399). However, these methods only report on total amount of DNA synthesis and provide averages, both over time and over the ensemble of molecules. The dynamic behaviors that actually govern transitions through multiple conformational states, driven by a hierarchy of strong and weak interactions, are inaccessible using traditional biochemical assays, and this knowledge is essential to understand these processes in biophysical detail. Single-molecule based approaches of DNA replication allow for the real-time observation of the kinetics from individual replisomes, revealing rare intermediates and often surprising dynamics during replication that can not be otherwise detected (52, 139, 236).

Here, we use single-molecule tethered-bead assays to study the kinetics of the leading strand replisome of a eukaryote, *S. cerevisiae*, which has surprisingly homologous replication machinery to that used in human. The minimal replisome system is reconstituted from the helicase complex Cdc45, MCM2–7, GINS (CMG), the leading-strand DNA polymerase Pol ϵ, the clamp loader Replication Factor C (RFC), the sliding clamp Proliferating Cell Nuclear Antigen (PCNA) and single-stranded DNA (ssDNA) binding protein (SSB). In the current report we observe the synthesis of the leading strand in real time, at rates near that expected from cellular observations. In the presence of Minichromosome maintenance protein 10 (Mcm10), we observe a 3-fold increase in the number of productive replication events and an increase in the basal rate of the minimal replisome, supporting the role of Mcm10 in fork rate and stability after origin firing (9). Mrc1 forms a complex with Tof1 and Csm3, often referred to as the fork protection complex, and referred to herein as the MTC complex. The MTC complex is generally thought to function when the replication fork is challenged with DNA damage or at replication fork barriers (392–394, 400). In the presence of RPC component MTC, a mediator of the DNA damage response, we observe significantly increased rates of replication. This increase is consistent with values previously published (388), and suggest that this protein complex may regulate fork speed as shown previously (398). Unexpectedly, the MTC complex causes multiple changes in rate over time during a single leading-strand replication reaction observed at the single-molecule level. In sum, the observations documented herein provide a picture of a highly dynamic interaction between the MTC complex and the leading strand replisome and have important implications about how the cell may regulate fork speed in response to DNA damage.

8.2 Results

8.2.1 Single-molecule visualization of leading-strand synthesis.

We use a single-molecule tethered-bead flow-stretching assay (156, 354) to directly visualize the replication kinetics of individual *S. cerevisiae* leading-strand replisomes. A linear and double-stranded DNA substrate containing a replication fork is attached to a microbead on one end and the top surface of a microscope cover slip on the other (Figure 8.1 c). With these DNA-bead tethers placed inside a microfluidic channel, we can apply a laminar flow to exert a controllable drag force on the beads, and thus stretch out the DNA molecules. We use ultra-wide-field, low-magnification microscopy to image thousands of beads and relate bead movement to changes in DNA length (52) (Figure 8.1 a and b). At drag forces lower than 6 pN, single-stranded DNA (ssDNA) is approximately six times shorter than double-stranded DNA (dsDNA) (156, 401). Movement of the bead against the direction of flow, therefore, reports on the conversion from dsDNA to ssDNA. With leading-strand synthesis effectively converting parental DNA into ssDNA on the lagging strand, we can now monitor leading-strand synthesis by attaching the 5′ lagging strand of the fork to the surface and visualizing the accumulation of ssDNA lagging-strand product by a gradual shortening of individual DNA molecules. Topoisomerase is not required because the DNA is free to rotate at both ends, thus preventing accumulation of supercoils. Automated fitting of the bead images and tracking of the bead positions as a function of time provides a read out for these interconversions with high precision ($\sim$50 nm, corresponding to $\sim$200 bp; (52). Simultaneously tracking thousands of beads enables high data throughput and the characterizations of sub populations within individual experiments. With access to this type of kinetic single-molecule data, we can characterize properties of individual replisomes such as rate (and changes therein) and the product length (the total number of nucleotides synthesized per replisome during the experiment). Previous work has shown that a minimal leading-strand replisome reconstituted from *S. cerevisiae* CMG, Pol ϵ, RFC, PCNA, and Replication Protein A (RPA), comprising 24 different polypeptides was competent in DNA synthesis of linear forked substrates (39). To recapitulate these results at the single-molecule level, we visualized leading-strand DNA synthesis using the single-molecule tethered-bead flow-stretching assay (Figure 8.1 d). As a template for leading-strand synthesis we use a 20-kb long DNA amplicon generated by PCR from λ phage (Figure 8.1 c). Annealing and ligating its ends with functionalized oligonucleotides allows us to construct a replication fork at one side with a biotinylated 5′ end and a digoxigenin on the other end of the DNA (Materials and Methods, Figure 8.7). The biotinylated fork is tethered to the streptavidin coated surface of the flow cell and the digoxigenin couples to a 2.8-μm, anti-digoxigenin coated bead. The leading-strand arm of the fork contains a 3′ ssDNA tail that is exposed to the solution to facilitate loading of the CMG helicase. Further, the 3′ end of the leading-strand arm includes three phosphorothioated thymine bases to stop exonuclease degradation by Pol ϵ during replisome assembly (Figure 8.1

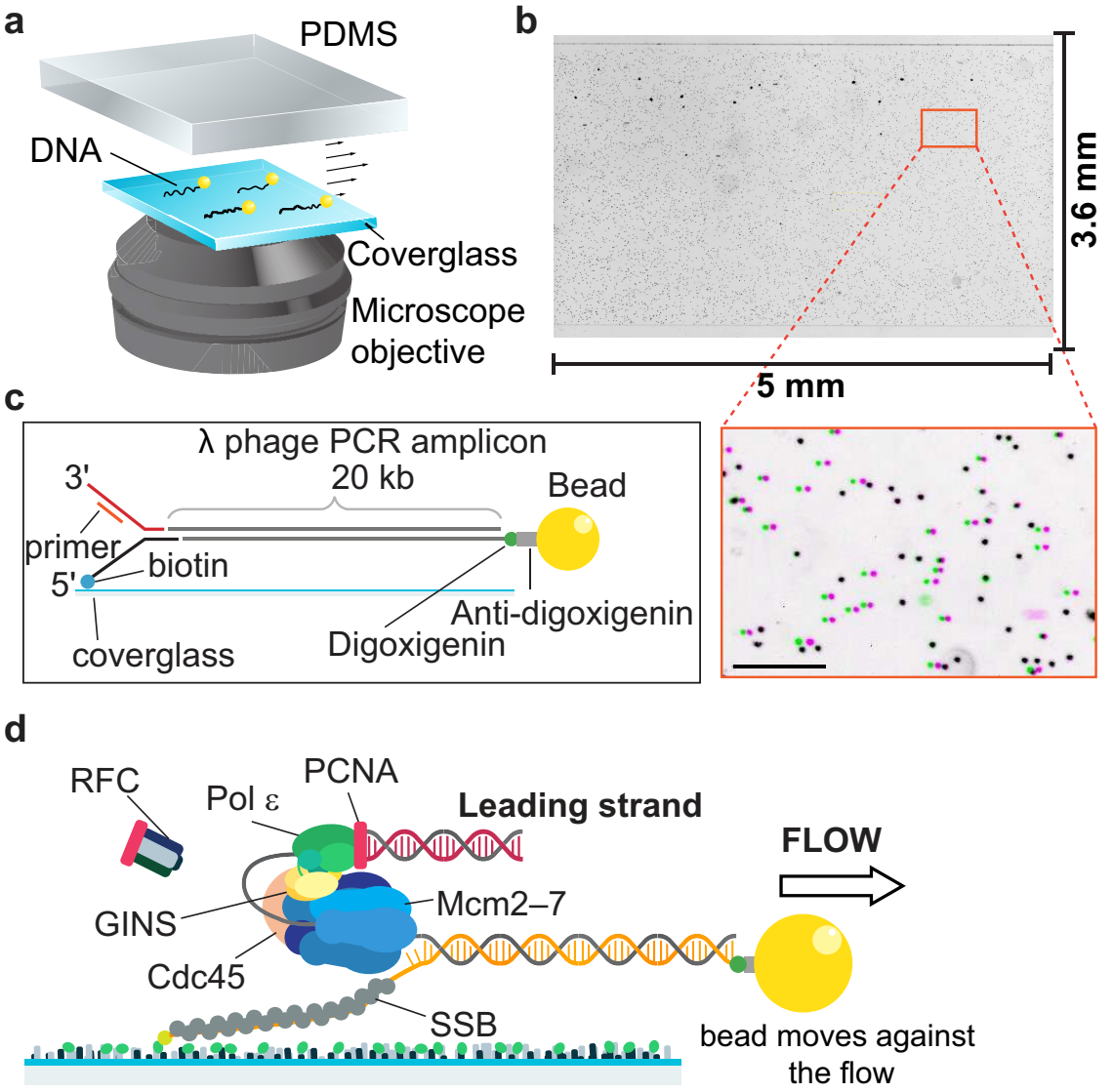


Figure 8.1: Single-molecule tethered-bead DNA-stretching assay. (a) Experimental setup. Individual DNA molecules are tethered to the surface of a microfluidic flow cell. Beads are attached to the DNA ends and imaged with wide-field optical microscopy using a low-magnification, high-numerical-aperture objective. DNA molecules are stretched close to and parallel to the surface by applying a laminar flow of buffer. (b) A representative field of view showing 4,000 beads. (Inset) Image of beads attached to DNA flow-stretched in one direction (magenta) superimposed with an image of the same bead-attached DNA molecules stretched in the opposite direction (green) shows the presence of a large number of DNA-bead tethers. The distance between the pairs of differently colored spots corresponds to twice the total DNA-tether length. The beads that are improperly tethered are shown in black. The black scale bar is 150 μm. (c) DNA template. A replication fork was introduced at one end of a 20-kb linear substrate, with a bead attachment site at the other end. The fork is attached to the surface via a biotin on the 5′ tail. (d) Schematic representation of leading-strand replication by the minimal *S. cerevisiae replisome*. As dsDNA is converted into ssDNA, the DNA shortens and the bead moves against the direction of flow.

b). Measuring the length difference between ssDNA and dsDNA provides a ratio between the number of processed nucleotides by the DNA polymerase and the amount of observed shortening. In the absence of a single-stranded DNA-binding protein, this ratio is 3670 $\pm$ 40 nt/μm. Due to the presence of single-stranded DNA binding proteins in the reaction mixtures, however, the measured
contour length of ssDNA generated during leading-strand replication will be higher than that for naked ssDNA. To correct for this difference, we measured the change in ssDNA length upon RPA binding and the correct ratio was derived between the lengths of dsDNA and RPA-coated ssDNA (Figure 8.8). This value was measured to be 106 $\pm$ 10 $\%$, making RPA-coated ssDNA almost the same length as dsDNA. RPA is therefore incompatible with the visualization of changes in DNA length during leading-strand synthesis, as conversion of binding of RPA to ssDNA would result in very little net movement of the bead. Consequently, we used *E. coli* SSB in all replication assays, as ssDNA coated with SSB has a contour length that is 24 $\pm$ 2 $\%$ of that of dsDNA, corresponding to an experimental conversion factor of 5596 $\pm$ 73 nt/μm (Figure 8.8). *E. coli SSB* and *S. cerevisiae* RPA give indistinguishable results in the leading-strand synthesis reaction (Figure 8.11).

8.2.2 Single-molecule replication rates and processivities of pol ϵ dependent leading strand synthesis.

Our overall experimental strategy for establishing the leading-strand replisome is outlined in figure 8.1 d. First, CMG is loaded onto the fork under a flow of buffer in the absence of ATP and other nucleotides. Subsequently, leading-strand synthesis is initiated by introducing a flow of buffer containing CMG, Pol ϵ, PCNA, RFC, PCNA, Mg^{2+}, all four dNTPs and ATP (See Materials and Methods for precise assembly and reaction conditions). Figure 8.2 a shows length changes of two individual DNA molecules as a function of time after replication has been initiated. A gradual shortening of the DNA is clearly visible, indicating sustained conversion of dsDNA into ssDNA. To detect rate changes and to identify different operational modes of the leading-strand replisome, we used an unbiased, multiline-fitting algorithm based on change-point theory (52, 189) (Figure 8.2 a, black lines). The rates obtained from this algorithm are weighted by the DNA segment length, reflecting the number of nucleotides that were synthesised at this rate. This places more significance to the longer rate segments, as they have a higher signal-to-noise ratio compared to shorter ones. The rate of Pol ϵ-dependent leading-strand synthesis was determined by fitting the rate histogram with a Gaussian function, resulting in a rate of 5.4 $\pm$ 0.7 nt/s (mean $\pm$ s.e.m) (Figure 8.2 b), consistent with earlier ensemble reactions using ^{32}P-dNTPs that were corrected for the length bias inherent in use of ^{32}P-dNTPs that exaggerate observed lengths due to increased amounts of ^{32}P-dNMPs in longer DNA fragments (39). These ^{32}P-dNMP incorporation corrections are not required in single molecule measurements. The rate values in these single molecule experiments are consistent with previously reported ensemble experiments (39), (383) and use of yeast extracts in single-molecule experiments (106). Instead of using processivity, a term that is am-

biguous in definition when comparing experiments with very different protein and DNA concentrations, we leverage the precise nature of our single-molecule measurements to define the product length of individual DNA products. The overall product length for an experiment is determined by measuring the total amount of dsDNA converted into ssDNA for every trajectory and fitting their distribution with a single-exponential decay (assuming a single rate-limiting step determining the end of an event). In this fitting procedure, the bins with very short product lengths are excluded to prevent their undersampling from affecting the fits. The product length of Pol ϵ-dependent leading-strand synthesis was measured to be 0.9 $\pm$ 0.2 kilonucleotides (knt) in the 20-minute observation time (Figure 8.2 c). These rate values for a minimal leading strand replisome lacking Mcm10, Mrc1, Tof1 and Csm3 are consistent with previously reported ensemble experiments (39,383). In the absence of the four dNTPs and ATP no

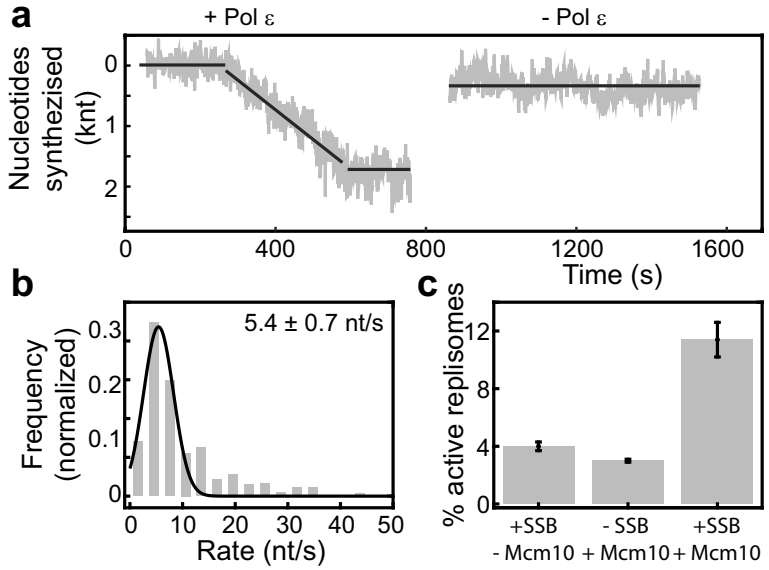


Figure 8.2: Single-molecule visualization of leading-strand synthesis by *S. cerevisiae*. (a) Representative trajectory showing Pol ϵ-dependent leading-strand synthesis (left). When Pol ϵ is omitted no replication events are observed (right). The black lines represent the rate segments identified by the change-point algorithm. (b) Histogram of the instantaneous single-molecule rates, weighted by segment length. The black line represents a Gaussian fit to the histogram with a rate of 5.4 $\pm$ 0.7 nt/s (mean $\pm$ s.e.m.) (N = 161 trajectories). (c) Efficiencies of leading-strand synthesis. The efficiency is defined as the number of beads that show replication events over the total number of correctly tethered beads. The efficiency is $\sim$3-fold higher (11.4 $\pm$ 0.2%, N = 3 experiments) when SSB and Mcm10 are present compared to experiments without Mcm10 (4.0 $\pm$ 0.3%, N = 4 experiments) or without SSB (3.0 $\pm$ 0.1%, N = 2 experiments). The errors represent the experimental error.

replication events were observed, demonstrating that the observed bead movements are enzyme dependent.

We note that previous ensemble assays of recombinant CMG show that CMG binds DNA for up to one hour, and that these longer time windows enable CMGE-PCNA to eventually complete synthesis of a 3-kb template (39, 383, 387). In our current setup, however, the typical observation time is 20 minutes, and we cannot directly observe when enzyme binding and/or unbinding occurs. The processiv-

ity/stability of these components on DNA and proficiency to exchange with components from solution will be the subject of a future study.
To exclude the possibility of Pol ϵ-independent unwinding of dsDNA by CMG, we performed the experiment in the absence of Pol ϵ under the same experimental conditions. As expected, we do not see any replication events (Figure 8.2 a). We have also performed the experiment lacking CMG, but detect no replication events, consistent with the previously described inability of Pol ϵ to perform strand displacement synthesis. (402) Combining these results, we conclude that the effective shortening of the DNA substrate arises from CMG-Pol ϵ dependent leading-strand synthesis. This observation provides us with the ability to monitor leading-strand replication of *S. cerevisiae* in real time at the single-molecule level. Additionally, it affords us the opportunity to characterize interactions between proteins within the leading-strand replisome, one replisome at a time.

8.2.3 Mcm10 increases the number of productive replication events.

Mcm10 has been identified as an important player in CMG activation (403, 404) and maintenance of the replication fork. (405, 406) Studies using reconstituted purified proteins have demonstrated that Mcm10 is not absolutely required for leading/lagging strand fork function *in vitro* (383). However, to understand the effect of Mcm10 during leading-strand replication at the single-molecule level, we added equimolar amounts of CMG and Mcm10 during initial CMG loading and in the subsequent replication reaction. The addition of Mcm10 did not result in any Pol ϵ-independent unwinding of dsDNA by CMG. However, addition of Mcm10 to the leading-strand replication reaction resulted in a $\sim$1.3-fold increase in rate (11.0 $\pm$ 0.6 nt/s, Figure 8.3 b), consistent with previous ensemble observations (9). Interestingly, we did notice a significant $\sim$3-fold increase in the number of trajectories that show productive replication events, suggesting Mcm10 enhanced the formation of productive replication complexes (Figure 8.2 d). The efficiency is defined as a percentage of the number of correctly tethered beads that show replication. The average number of correctly tethered beads is 981 $\pm$ 147 (N = 5 experiments). This increase in efficiency suggests that Mcm10 facilitates the assembly of an active leading-strand complex, or enhances its stability as observed earlier (390). Consequently, all further experiments included Mcm10. We observe a 2-fold increase in the average replication rate (19.7 $\pm$ 1.2 nt/s) in the presence of MTC consistent with *in vivo* observations of fork speed in the presence and absence of Mrc1 (Figure 8.3 b), and an increase in product length (compare Figure 8.3 E with Figure 8.3 F). The overall 1.8-fold rate increase is consistent with several *in vivo* studies of Mrc1 deficient cells (139, 236, 399). Previous work reported that Mrc1 is responsible for the increased fork speed, even though the interaction of MTC with the replisome is largely mediated by Tof1, and that Mrc1 function is largely aided by the presence of Tof1 (388, 392, 397, 398, 400).

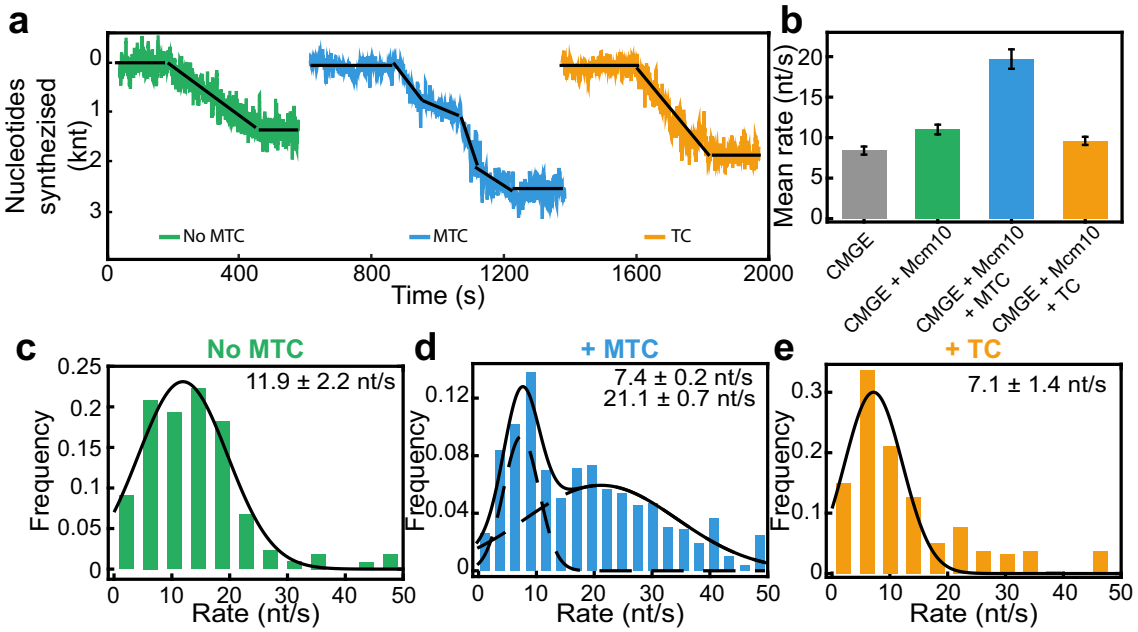


Figure 8.3: Effect of MTC on replication kinetics. (a) Representative trajectory showing Pol ϵ-dependent leading-strand synthesis without MTC (left - green), with MTC present (middle - blue), and with TC present (right - orange). The black lines represent the rate segments identified by the change-point algorithm. (b) The average single-molecule rates (mean $\pm$ s.e.m.), calculated by taking the mean of all segments determined by the change-point algorithm, using CMGE (8.4 $\pm$ 0.5 nt/s), CMGE + Mcm10 (11.0 $\pm$ 0.6 nt/s), CMGE + Mcm10 + MTC (19.7 $\pm$ 1.2 nt/s) and CMGE + Mcm10 + TC (9.6 $\pm$ 0.5 nt/s). (c) Histogram of the instantaneous single-molecule rates, weighted by segment length for replication by CMGE + Mcm10. The black line represents a Gaussian fit to the data with a mean rate of 11.9 $\pm$ 2.2 nt/s, similar to the rates obtained without Mcm10 (Figure 8.2B) (N = 96 trajectories). (d) Histogram of the instantaneous single-molecule rates, weighted by segment length for replication by CMGE + Mcm10 + MTC. The histogram shows a bimodal distribution of the rates. The data were fit with the sum of two Gaussian distributions (black line), resulting in a rate of 7.4 $\pm$ 0.2 nt/s for the slow population and 21.1 $\pm$ 0.7 nt/s for the fast population (N = 225 trajectories). (e) Histogram of the instantaneous single-molecule rates, weighted by segment length for replication by CMGE + Mcm10 + TC (omitting Mrc1). The fast population associated with MTC activity is not present.

8.2.4 Addition of MTC increases replication rates of Pol ϵ dependent leading-strand synthesis.

Previous studies demonstrated that the fork protection complex comprised of Mrc1–Tof1–Csm3 (MTC) is required for maximal fork speed *in vivo*, and enhances the rate of *in vitro* reconstituted replisomes to the rates observed *in vivo* (388). These ensemble experiments, however, did not inform on the lifetime of MTC binding to a leading-strand replisome or its effect on the instantaneous replication rates. It is interesting to note that the Mrc1, Tof1 and Csm3 proteins are present in replisome progression complexes in a substoichiometric fashion, suggesting they are not present in every replisome or only transiently associated (391). To provide access to this important kinetic information, we repeated the tethered-bead assay in the presence of 30 nM MTC (Figure 8.3 a). Since the instantaneous rates as extracted from our single-molecule experiments are weighted by the length of the rate segments, the mean of all instantaneous rates corresponds to the time-

and population-averaged rate. Consistent with this observation, the increase in leading-strand fork speed we observe requires MTC and is not observed using only the TC complex. Leading-strand replication performed in the presence of the Tof1–Csm3 complex (TC), resulted in the loss of the higher rates and loss of the increase in product length compared to the MTC complex that includes Mrc1 (Figure 8.3 b, g, h).

8.2.5 MTC induces multiple rate changes within a single leading-strand replication complex.

The single-molecule rate distribution for MTC-mediated leading-strand synthesis shows a bimodal rate distribution, comprised of a slow population with a rate of 7.4 $\pm$ 0.2 nt/s (mean $\pm$ s.e.m) and a fast population with a rate of 21.1 $\pm$ 0.7 nt/s. While these observations appear to suggest an unsaturated reaction, we have titrated MTC into bulk assays from 0–120 nM MTC and observe saturation at 15 nM, less than the 30 nM MTC used in the experiments of this report (Figure 8.14). Later in this report we compare traces using subsaturating concentrations of MTC with saturating MTC. The appearance of these two populations highlights the importance of using single-molecule techniques, as this bimodal distribution would not be visible in traditional ensemble-averaging assays. This bimodal distribution can be explained by two possible mechanisms — one in which MTC speeds up a subset of replisomes, or one in which MTC interacts with all replisomes, but only transiently. If the first mechanism were true, a subset of trajectories would exhibit faster rates consistent with these replisomes associated with MTC, whereas the rest would exhibit the slow rate observed in the absence of MTC. If the second mechanism were true, we should see leading-strand replisomes exhibiting both slow and fast rates within a single trajectory, resulting in multiple rate changes per replisome. To distinguish between these two possibilities, we first quantified the number of rate changes for each replisome. Rate changes were defined by the change-point line-fitting analysis. We compared the number of rate changes with and without MTC (Figure 8.4 a). On average we observe 4.5 times more rate changes when MTC is present. This high frequency of rate changes within individual reactions identifies the second mechanism is at work, in which MTC interacts with all replisomes, but only transiently. To further support this observation of rate changes within a single trajectory, we examined the distribution of rates associated with individual switches between rates. We did so by plotting the rate of a change-point segment within a single-molecule trace versus the rate of the previous change-point segment in the same trajectory (Figure 8.4 b). The points in this transition plot represent rate pairs from trajectories with multiple rate changes. This transition plot shows that there are more rate changes and more fast rates when MTC is present (Figure 8.4 B bottom), compared to the rates without MTC (Figure 8.4 b top). While we do observe some rate changes in the absence of MTC (Figure 8.4 a, b top), the points in the transition plot are clustered much closer to the diagonal. This clustering indicates that the rate changes are only minor, and are probably due to the small intrinsic rate variations of Pol ϵ-dependent synthesis. In contrast, when MTC is present the points in the transi-

tion plot lie much further away from the diagonal. This deviation from the diagonal shows that the change in the rate between two segments in a single trajectory is large. These large changes imply that the replisome can transition from fast, MTC-mediated rates to the slow rates, and vice versa.

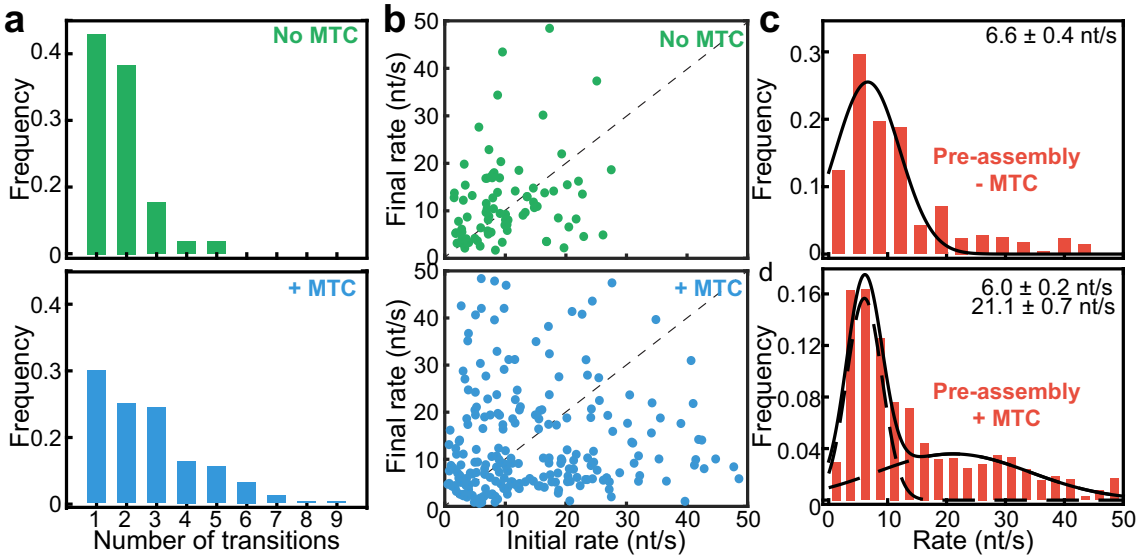


Figure 8.4: MTC interaction with the replisome is transient. (a) The number of rate changes per trajectory without MTC (top) is 4.5 times lower than with MTC present (bottom). (b) Transition plots showing the rate of a segment as a function of the rate of the previous segment for trajectories with multiple segments, with (top) and without (bottom) MTC present. The distance from the diagonal (dashed line) is ∼2.5-fold higher when MTC is present (13.6 ± 1.1 nt/s, mean ± s.e.m) than without MTC (5.9 ± 0.6 nt/s, mean ± s.e.m.) (c) Histogram of the instantaneous single-molecule rates, weighted by segment length obtained from a pre-assembly experiment in which MTC was present during loading but omitted from the replication phase. The rate is 6.6 ± 0.4 nt/s, which is similar to the rates we obtained for our continuous flow experiments without MTC. No MTC mediated fast rate population was observed (N = 101 trajectories). (d) Histogram of the instantaneous single-molecule rates, weighted by segment length obtained from a pre-assembly experiment where Pol ϵ was present during loading but omitted from the replication phase. In contrast to the previous experiment in (d), the faster population is present here. Fitting with the sum of two Gaussians gives rates of 6.0 ± 0.2 nt/s and 21.1 ± 0.7 nt/s (N = 196 trajectories).

To quantify the average change in rate between transitions within a single trajectory, we calculated the average distance from the diagonal for all the points in the transition plot. In the presence of MTC the average rate change between transitions is ∼2.5-fold higher with a rate change of 13.6 ± 1.1 nt/s (mean ± s.e.m) compared to 5.9 ± 0.6 nt/s in the absence of MTC. Moreover, the fact that the off-diagonal points are symmetrically distributed around the diagonal illustrates that it is just as likely for a slow rate segment to be followed by a fast rate segment, as it is for a fast rate segment to be followed by a slow one. This lack of bias towards a fast rate being followed by a slow rate or vice versa, reveals that MTC can bind and unbind from the replisome after replication has started. This observation further supports that MTC does not bind to the replisome through a tight and stable interaction, but instead undergoes cycles of binding to the replisome from solution and dissociation. We verified that placing the 3XFLAG tag on the C-terminus of Mrc1 and no tag on Csm3 within MTC instead of the N-terminal FLAG-Mrc1 (i.e.

compared to use of a C-terminal FLAG tag on Mrc1 and a C-terminal calmodulin tag on Csm3 in (388, 389)) did not result in appreciable differences in the ability of MTC to induce multiple rate changes within single leading-strand replication complexes (Figure 8.12).

8.2.6 MTC is transiently associated to the CMGE leading-strand replication fork complex.

We reasoned that if MTC is indeed weakly bound to the replisome, we should be able to decrease the frequency of rate transitions within single replisomes by lowering the concentration of MTC present during the replication reaction. Replication reactions performed in the presence of either 10 nM or 3 nM MTC, showed a reduction in the number of fast rates as well as the frequency of transitions within a single trajectory (Figure 8.13 A–C). To confirm and extend these observations, we performed leading-strand synthesis under conditions permitting preassembly of replisomes at the fork. If indeed MTC transiently associates with the replisome, under these conditions, we should not see the faster rates when we include MTC during the assembly phase, but omit it from the subsequent replication reaction as it would dissociate by the time the replication reaction started. As predicted, the rate distribution did not show the fast population (Figure 8.4 c). In contrast, when the CMGE complex is assembled on DNA and Pol ϵ is omitted from the subsequent replication reaction, followed by the supply of MTC (but no Pol ϵ) in the buffer flow, the faster population is evident (Figure 8.4 d). This result indicates that Pol ϵ remains stably bound to the replisome, consistent with previous reports (387). Combining these results, we hypothesize that MTC has a weak affinity for the leading-strand replisome and interacts in a dynamic fashion to increase the rate of the replication fork.

8.3 Discussion

We have used a DNA-stretching assay to visualize *in vitro* leading-strand synthesis by the *S. cerevisiae* replisome at the single-molecule level. Similar experiments have been reported for the T7 and *E. coli* replisomes (52, 202, 354, 407), but a detailed kinetic analysis of the eukaryotic replisome at the single-molecule level has been unavailable thus far. The leading-strand synthesis rates observed here are similar to those previously reported in ensemble biochemical reactions (383, 388) and within the range of replication fork movement observed inside the cell (397, 408, 409).This single-molecule assay has allowed us to probe the effect of the RPC components Mcm10 and MTC on leading-strand replisome activity, confirming reports that Mcm10 stimulates the minimal replisome in the absence of MTC (390), and that MTC stimulates the replisome by an average of 1.8-fold dependent specifically on MTC containing Mrc1 compared to TC lacking Mrc1 (388), as summarized in Figuere 8.5. The observations of Mrc1 dependent stimulation of fork rate are also consistent with cell biology studies of fork movement in Mrc1 cells (395–397). Interestingly, observation at the

single-molecule level has revealed unexpected kinetic behaviors that would have been impossible to observe with conventional biochemical assays.

We observe that Mcm10 does not substantially increase the rate or product length of leading-strand synthesis over the time frame used in our assays, but does increase the number of productive replication events, consistent with a recent study indicating that Mcm10 stabilizes the CMG on DNA (390). This Mcm10 associated increase in efficiency could be relevant to the conclusions that Mcm10 functions as an activator of the CMG complex throughout DNA replication (390, 410), assisting stabilization of CMG or helping it overcome possible obstacles. Mcm10 is also known to activate Mcm2–7 during replication initiation (403, 411, 412). It is conceivable that Mcm10 helps activate Mcm2–7 in our *in vitro* leading-strand system through a similar mechanism. It also remains possible that Mcm10 stabilizes Pol ϵ or otherwise enhances synthesis activity.

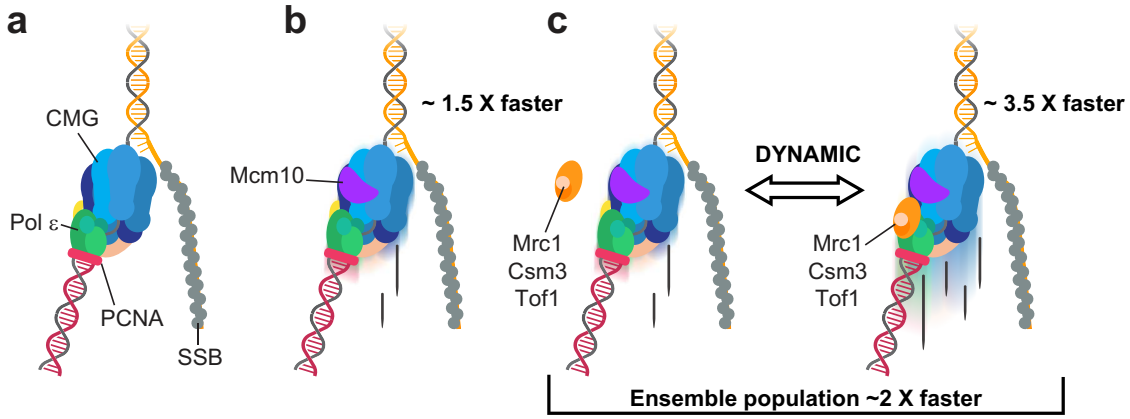


Figure 8.5: Leading-strand synthesis by the *S. cerevisiae replisome*. (a) The minimal reconstituted leading-strand replisome, containing the CMG helicase, the Pol ϵ DNA polymerase, the RFC clamp loader, the PCNA sliding clamp and the SSB single-stranded DNA-binding protein supports leading-strand synthesis at a rate of 5.4 $\pm$ 0.7 nt/s. (b) When Mcm10 is added, the rate increases by $\sim$1.5 fold (11.9 $\pm$ 2.2 nt/s). Furthermore, Mcm10 increases the number of productive replication events in our assay. (c) The fork protection complex MTC speeds up the leading-strand replisome by $\sim$3.5 fold. Our single-molecule measurements demonstrate that MTC has a weak affinity for the replisome and only transiently interacts to speed up replication.

The current study demonstrates that the MTC complex increases the rate of leading-strand synthesis in an unexpected fashion (summarized in Figure 8.5). The MTC complex appears to act in a highly dynamic fashion, only transiently active at the replisome. This observation is consistent with the substoichiometric presence of these subunits in the RPC complex (391). We observe processive leading-strand reactions at the single-molecule level with short-lived phases of higher rates that we interpret as corresponding to MTC binding to the replisome. Interestingly, the instantaneous rate during these phases increases by 3-fold, but average out to an approximately 2-fold average rate increase because they do not persist throughout the entire kinetic trace of a single molecule. Further, the 2-fold average rate increase is consistent with observations of Mrc1 in the cellular context (395–397). The fluctuating rate changes per single-molecule

trajectory suggests that MTC is distributive, and does not bind to the replisome in a stable manner. Distributive behavior of critical replisome components has precedent in bacterial DnaG primase (413, 414). When we compare the rates of successive segments within one trajectory we see that fast MTC-mediated rates can be followed by slow rates and vice versa. Furthermore, the amplitude of these rate changes is on average $\sim$2.5 fold higher than the rate changes without MTC. This observation demonstrates that MTC or one of its components (e.g Mrc1) associates with and dissociates from the replisome multiple times during leading-strand synthesis. The fact that we do not observe any fast rates when MTC is omitted during the replication phase, but present during the CMG loading phase of a preassembly experiment, further supports the conclusion that transient interactions exist between MTC and the replisome. These data contrast with previous models that suggest that Mrc1 stably binds to both CMG and Pol ϵ (392), though we note that cross-linking in ChIP assays prevents dynamics, and pull downs are not quantitative. It was proposed that this pair of interactions could be responsible for the faster rates, by tethering Pol ϵ to CMG (398, 415). The higher kinetic resolution of our experiments reveals the dynamic interaction of MTC with the replisome, with the population-averaged observables consistent with earlier biochemical assays. From the current study, however, we can not conclude whether MTC acts to stimulate the DNA polymerase, or the rate of unwinding, or both.

It is important to note that the exact phosphorylation state of the proteins is expected to play a role in MTC–replisome interactions (416). In addition, the current study focuses on the enzymes of leading-strand synthesis while additional proteins could play a role in MTC behavior. Replication proteins in *S. cerevisiae* undergo many post-translational modifications before and during replication (417, 418). For example, Dbf4-dependent kinase (DDK) and S-phase cyclin-dependent kinase (CDK) are two S phase-specific kinases that phosphorylate replication proteins (417). DDK and CDK are known to control replication initiation by phosphorylation of many proteins involved in forming the origin recognition complex (417–419). Additionally, phosphorylation of replisome components plays an important role in programmed fork arrest through phosphorylation of Mcm2–7, which promotes recruitment of Tof1–Csm3 by the replisome (416). Post-translational modifications can be expected to affect the behavior of proteins during active DNA replication. In this connection, *S. cerevisiae* recombinant expressed CMG and *S. cerevisiae* recombinant expressed Tof1–Csm3 have been demonstrated to be phosphorylated and thereby enable their interaction, while dephosphorylation of either component prevents detection of an interaction between them (416). Indeed, upon co-expression of CMG, the vast majority of expressed proteins are free Mcm2–7 complexes, Cdc45 and GINS tetramers and these factors, when mixed together, do not simply self-assemble into a CMG complex. Thus, it is possible that the small amount of recombinant CMG recovered from expression cells is in fact assembled at origins (39). We note that Mrc1 has previously been shown to be the only component of MTC that enhances replisome rate (388) and since our

observations rely on the conversion of dsDNA to ssDNA, it is possible that the dynamic interactions reported herein are of Mrc1 instead of the entire MTC complex.

It is tempting to speculate a possible biological reason for a dynamic interaction of Mrc1/MTC with the replisome. We presume that the different replication rates that correspond to the association state of Mrc1/MTC (bound to CMG or not) reflect different conformations of the replisome. Different replisome conformations may in turn facilitate active site configurations (i.e. enzymatic velocities), additional protein interactions or exchange with other partners, in a differential manner. An interesting aspect of MTC activity is its phosphorylation state. For example, it is well known that Mrc1 mediates the DNA-damage response through phosphorylation of Mrc1 by the Mec1/Rad53 kinases (420). The advantage of a dynamic interaction of Mrc1/MTC with CMG could provide an interesting type of regulation. The dynamic interaction between MTC and the replisome documented herein could ensure a complete sampling of the phosphorylation state of MTC by all replisomes, as opposed to only a subset of replisomes carrying a fully phosphorylated MTC complex. Therefore, a dynamic mechanism of MTC–replisome interaction would allow the MTC to act as a potentiometer for damage, such that the ratio of modified and unmodified MTC (i.e. in response to DNA damage) would be "sensed" by all replisomes equally, instead of a stark division in the case of a stable interaction of MTC with CMG in which some replisomes will have the phosphorylated form of MTC and other replisomes will not, resulting in different fork speeds within the same cell. Hence, a dynamic interaction with a regulatory component that is modulated by the damage response would provide a gradual titration of phosphorylated MTC being equally sampled by all replisomes, and a more uniform fork speed as a consequence.

Despite these studies revealing the dynamic nature of a replisome modulator with a moving replisome, further studies are needed to enable a greater understanding of how replication proteins interact *in vitro* and in living cells. Therfore, such approaches are intimately positioned to address questions that have an important impact on genomic integrity.

8.4 Materials and Methods

8.4.1 Protein expression and purification.

CMG, Pol ϵ, RFC, PCNA, RPA and SSB were purified as previously described (39).

Mcm10: The gene encoding Mcm10 was amplified by PCR along with an N-terminal 10 histidine tag and a C-terminal 3X FLAG tag and cloned into pET-16b. *E.coli* cells (BL21-CodonPlus (DE3)-RIL) containing the dual tagged Mcm10 expression plasmid were grown to an OD600 of 0.6 (ampicillin and kanamycin resistant), induced with isopropyl β-D-1-thiogalactopyranoside (IPTG) for 18 h

at 15°C, and collected and lysed using a pressure cell. Supernatant of lysed cells was applied to a 10 mL Chelating Sepharose Fast Flow (GE Healthcare) column charged with 50 mM nickel(II) sulfate and equilibrated with equilibration buffer (20 mM Tris-HCl pH 7.9, 5 mM imidazole, 500 mM sodium chloride, 0.01% (v/v) NP-40). After loading, the column was washed with equilibration buffer, then proteins were eluted in equilibration buffer containing 375 mM imidazole. Fractions were monitored by SDS-PAGE and peak fractions containing Mcm10 were applied to a 6 mL Anti-Flag M2 resin (Sigma) equilibrated in 20 mM Tris-HCl pH 7.5, 10% (v/v) glycerol, 500 mM sodium chloride, 1 mM dithiothreitol, 1 mM magnesium chloride, 0.01% (v/v) NP-40. After loading and subsequent washing with 50 mL equilibration buffer, the proteins were eluted using two 6 mL treatments of elution buffer containing 0.2 mg/mL 3X FLAG peptide (EZ Biolab) for 30 min. Fractions (100 μL) were collected and Mcm10 containing fractions were aliquoted, flash frozen in liquid N_2, and stored at –80°C (Figure 8.6).

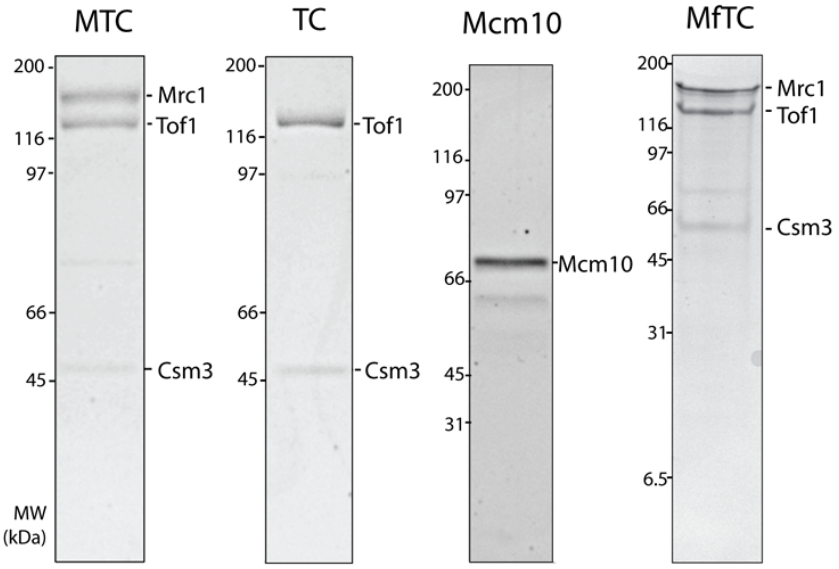


Figure 8.6: Purification of MTC, TC, Mcm10, and MfTC. Coomassie blue stained SDS-PAGE gels of MTC (left), TC (second from left), Mcm10 (third from left) and MfTC (right panel) are shown. The three left panels are the protein preparations used for this work, with the exception of Figure 8.12, which used the MfTC preparation in the rightmost panel. The left two panels (MTC and TC) are an 8% SDS-PAGE, the third (Mcm10) is a 12% SDS-PAGE, and the last (MfTC) is a 4–20% gradient SDS-PAGE. All proteins contain one or two tags, as documented in Materials and Methods.

Mrc1–Tof1–Csm3 complex: The genes encoding 3XFLAG-Mrc1 (^{Flag}Mrc1), Tof1 and 6xHis Csm3 (Csm3^{His}) were amplified from genomic DNA by PCR and inserted into yeast integration vectors similar as described for CMG in (39). Briefly, ^{Flag}Mrc1 was integrated at the *Ade2* locus, untagged Tof1 at the *His3* locus, and Csm1^{His} at the *Leu2* locus, each under control of the *Gal1/10* promoter. We also produced a yeast strain having a C-terminal Mrc1 3X FLAG tag and untagged Tof1, Csm3; the same procedure was used to purify both complexes (MfTC; Figure 8.12). Cells were initially grown at 30°C in SC-glucose under selection and then they were divided into flasks containing YP-glycerol. Cell strains

were grown at 30°C to OD600 of approximately 0.7, then induced by adding 20 g galactose/L for 6 h. After induction, cell pellets were collected by low speed centrifugation, resuspended in a minimal volume of 20 mM HEPES, pH 7.6, 1.2% (w/v) polyvinylpyrrolidone and flash frozen by dripping into liquid N_2. Induced cells were lysed using a cryogenic grinding mill (SPEX), powder was thawed in the cold room and then resuspended in 250 mM potassium glutamate, 50 mM HEPES pH 7.5, 1 mM EDTA pH 8.0 plus protease inhibitors (P8215, Roche). Cell debris was removed by centrifugation at 4ÂřC and the supernatant was mixed with 1.5 mL of anti-Flag M2 affinity resin (Sigma-Aldrich) for 1 h. The anti-FLAG resin was pelleted at 1000 x g and washed five times with 250 mM potassium glutamate, 50 mM HEPES pH 7.5, 1 mM EDTA pH 8.0 followed by centrifugation. Then the anti-FLAG affinity resin was resuspended in 2 mL of buffer containing 250 mM potassium glutamate, 50 mM HEPES pH 7.5, 1 mM EDTA pH 8.0 containing 10% (v/v) glycerol. This protein solution was then loaded by gravity onto the same column and washed with buffer containing 250 mM potassium glutamate, 50 mM HEPES pH 7.5, 1 mM EDTA pH 8.0, 750 mM sodium chloride. The MTC complex was then eluted with the same buffer containing 0.2 mg/mL 3X FLAG peptide (EZ Biolab), but without sodium chloride. Eluted proteins were concentrated and further purified using a Superose 12 gel filtration column in 2X PBS containing 10% (v/v) glycerol. Fractions were analyzed by SDS-PAGE and MTC-containing fractions were flash frozen in liquid N_2 and stored at –80°C (Figure 8.6).

8.4.2 Linear fork DNA substrate

DNA replication templates used in ensemble leading-strand experiments were prepared as previously described (39) (399). The replication substrate used for surface tethering and bead attachment in single-molecule experiments was constructed using a 19,979 bp PCR λ-phage product and the HPLC purified oligonucleotides listed in Table S1 (Integrated DNA Technologies). Briefly, bacteriophage λ DNA (New England Biolabs) was used as a template for PCR. A NheI site and digoxigenin modification were incorporated using primers 20kbF and 20 kbR. Next, the PCR product was digested with NheI (New England Biolabs), and the enzyme was heat inactivated and concentrated by ethanol precipitation. To assemble the fork duplex, a 1:1 molar ratio of both complementary-fork arm and bio-fork arm oligonucleotides were annealed by heating to 94°C for 5 min in hybridisation buffer (10 mM HEPES pH 7.5, 5 mM magnesium acetate, 50 nM sodium chloride, 0.5 mM EDTA) in the presence of 1.5-fold molar excess of the C2 primer oligonucleotide. Next, a 1.1-fold molar excess of pre-formed fork duplex was ligated with the NheI-digested PCR fragment hybridisation buffer for 48 hours at 16°C with 400 U of T4 DNA ligase (New England Biolabs). Linear fork substrates were diluted to 1.5 nM for subsequent use.

8.4.3 Single-molecule tethered-bead assay

Flow cells were prepared as described previously (52) (51). All single-molecule tethered bead assays were performed at 30°C. Briefly, a PDMS lids was placed

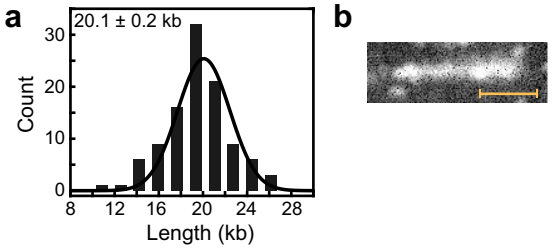


Figure 8.7: Length quantification of linear DNA substrate used in tethered bead assay using single-molecule fluorescence imaging. (a) Histogram showing the length of linear DNA templates. The black line represents a Gaussian fit to the data with a mean length of 20.1 $\pm$ 0.2 kb (N = 104 molecules). Error represents standard error of the mean. (b) Fluorescence image of a single linear DNA template labeled with SYTOX Orange. Imaging was performed as described in Lewis et al. 2017 (236). Scale bar represents 10 μm.

on top of a PEG-biotin-functionalized microscope slide (24 x 60 mm) to create a 3 mm wide and 100 μm high flow channel with Y junctions at both inlets and outlets. Polyethyleme tubes (PE-60: 0.76 mm inlet diameter and 1.22 mm outer diameter) were inserted to allow for a buffer flow. To help prevent non-specific interactions of proteins and DNA with the surface, the chamber was blocked with blocking buffer (20 mM Tris-HCl pH 7.5, 2 mM EDTA, 50 mM sodium chloride, 0.5 mg/ml BSA, and 0.0025% (v/v) Tween-20). The forked DNA substrates (20 pM) were flowed through the chamber for 12 min at 10 μL/min. After a brief wash with blocking buffer, tosylactivated paramagnetic beads (2.8 μm diameter, Life Technologies) functionalized with anti-digoxigenin (Roche) were introduced at 60 μL/min, until even surface coverage was achieved. Untethered beads were washed out of the chamber at 60 μL/min with $\sim$600 μL of buffer. The leading-strand replication reaction was performed in stages. First, CMG was loaded at the fork by pre-loading 30 nM CMG, 30 nM Mcm10, and MTC (where indicated) in replication buffer (25 mM Tris-HCl pH 7.6, 10 mM magnesium acetate, 50 mM potassium glutamate, 40 μg/ml BSA 0.1 mM EDTA, 5 mM dithiothreitol and 0.0025% (v/v) Tween-20) at 15 μL/min for 10 min. Next, a magnet was introduced above the flow cell to limit surface interactions. Finally, replication reactions were initiated by introducing 30 nM CMG, 30 nM Mcm10, 40 nM Pol ϵ, 20 nM PCNA, 6 nM RFC, 250 nM *E. coli* SSB and MTC (where indicated) in replication buffer supplemented with 5 mM ATP and 60 μM dCTP, dGTP, dATP, and dTTP. The beads were illuminated with a fiber illuminator (Thorlabs) and movies were collected at 4 fps using a 29 megapixel CCD camera (Prosilica GX6600; Allied Vision Technologies; 5.5 μm pixel size) with StreamPix imaging software (NorPix) at 4 x magnification with a lens (TL12K-70-15; Lensation) mounted directly to the camera. Leading-strand replication was monitored over a period of 20 min by tracking the movement of the beads and converting changes in position to nucleotides using custom Python software programmed in house.

8.4.4 Bead selection and processing

Typical tethered bead experiments generated movies between 1000 and 10 000 beads per frame having file sizes of $\sim$50 GB, making rapid and efficient data

analysis challenging. Trajectories were selected using a highly-staged and automated set of processing steps. First, beads with high fitting error and those that could not be tracked were rejected. The remaining beads were further filtered to remove those that are immobile due to surface interactions, those that dissociated prior to enzymes reaching the flow cell, and those with any additional nearby beads influencing movement. Remaining beads were then analyzed for activity using kinetic change point analysis, and those with greater than 1000 nucleotides of total synthesis were kept. Remaining trajectories were corrected for instabilities in the flow by taking all the trajectories present from start to finish, centering these at zero, and then taking the median x and y positions at each time point. These are then boxcar smoothed to create a drift trajectory, which is subsequently subtracted from each trajectory.

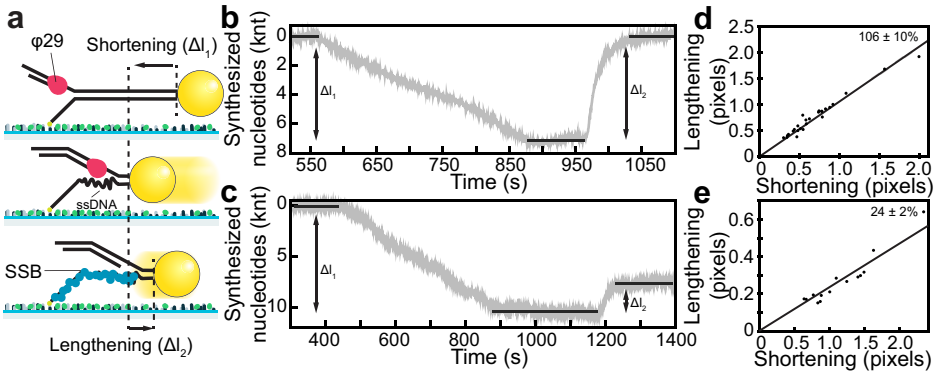


Figure 8.8: Determination of conversion factors of ssDNA coated with either RPA or SSB. (a) First, leading-strand synthesis shortens (ΔI1) the DNA by converting the lagging-strand DNA to ssDNA. Next, SSB (or RPA) coats the lagging strand resulting in lengthening (ΔI2) of the DNA. In experiments where SSB is present all the time, only an effective shortening is seen, i.e., (ΔI1 – ΔI2). To generate ssDNA, strand-displacement synthesis was performed using 60 U/mL of ϕ29 DNA polymerase (New England Biolabs) on surface-tethered forked DNAs containing replication forks in replication buffer as described in (3). After strand-displacement synthesis, the flow cell was washed excessively with replication buffer to remove any residual ϕ29 DNA polymerase. Then either *S. cerevisiae* RPA or *E. coli* SSB was flowed in at 250 nM at 15 μL/min. (b) RPA coated ssDNA has a similar length to dsDNA. (c) *E. coli* SSB coated ssDNA is shorter than dsDNA. (d) Ratio between *S. cerevisiae* RPA lengthening and shortening for 25 DNA molecules. Mean ratio 106 $\pm$ 10%. (e) Ratio between *E. coli* SSB lengthening and shortening for 14 DNA molecules. Mean ratio 24 $\pm$ 2%. Errors represent the error of the fit.

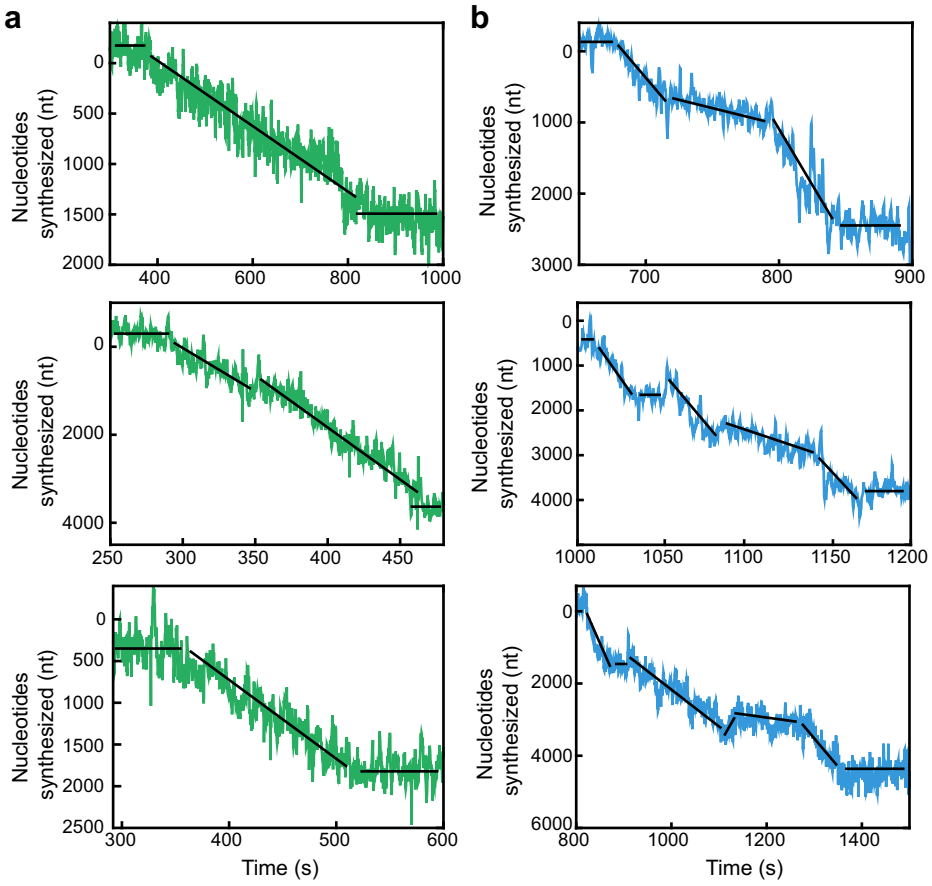


Figure 8.9: Six representative trajectories of enzymatic events observed. (a) Three example trajectories showing Pol ϵ-dependent leading-strand synthesis without MTC. (b) Three example trajectories showing Pol ϵ-dependent leading-strand synthesis in the presence of 30 nM MTC. The black lines represent the rate segments identified by the change-point algorithm.

Next, these filtered trajectories were manually curated to remove those with bead movement greater than 0.3 μm in the transverse direction, those that do not have a clear start or end point, and those that start before enzymes have reached the flow cell. Regions of enzymatic activity were processed using kinetic change-point analysis using a global noise threshold (52) (189) and replication rates determined using the conversion factor derived in Figure 8.8. Kinetic change-point segments containing less than 8 data points were also rejected. Once all kinetic change points were determined, rate distributions were constructed using slopes from single change-point segments. Each rate was weighted by its segment length to attach more significance to the longer segments, as they have a higher signal-to-noise ratio compared to shorter ones. The Gaussian distributions were binned using the square-root rule to generate the final distributions seen in Figures 8.2–8.4 and Figures 8.12–8.13. Product-length histograms were generated using the total length of DNA synthesized in a single trajectory to produce the final distributions seen in Figures 8.10, 8.12, and 8.13. These 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 2014b (Mathworks). The transition plots in Figure 8.4 and Figures 8.12–8.13 were gen-

erated by plotting the rate of a segment as a function of the rate of the previous segment, using only trajectories that have multiple rate segments.

8.4.5 Efficiency of leading-strand synthesis

The number of tethered beads was determined by counting the number of beads showing a proper flow reversal giving a length of 6.6 $\pm$ 0.1 μm (mean $\pm$ s.e.m). A typical experiment yielded 981 $\pm$ 147 correctly tethered beads (N = 5 experiments). The efficiency is defined as the number of replication events meeting the selection criteria outlined in molecule selection and processing (Materials and Methods), divided by the average number of correctly tethered beads.

8.4.6 Ensemble leading-strand replication assays

Replication reaction volumes were 25 μL. All reactions contained 25mM Tris-acetate pH 7.5, 5% (v/v) glycerol, 40 μg/ml BSA, 3 mM DTT, 2mM TCEP, 10 mM magnesium acetate, 50 mM potassium glutamate, 0.1 mM EDTA, 5 mM ATP, and 400 μM of each dNTP. The 2.8 kb linear forked templates (39) were primed with a 5′-^{32}P-labeled 37-mer oligonucleotide. Replication assays were performed by first incubating 30 nM CMG with 1.25 nM primed linear forked template in the presence or absence of 60

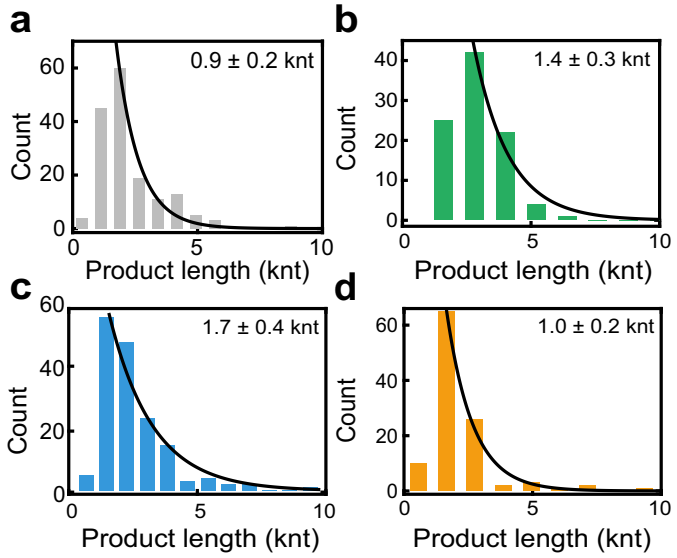


Figure 8.10: Distributions of product lengths for leading-strand replication.(a) Histogram of the total product length per trajectory (0.9 $\pm$ 0.2 knt). The fit represents a single-exponential decay function (black line; the first two bins are undersampled and not included in the fit). (b) Histogram of the total product length per trajectory for replication by CMGE + Mcm10. A single-exponential fit (black line) shows that the average product length is the same as without Mcm10 (1.4 $\pm$ 0.3 knt). (c) Histogram of the total product length per trajectory for replication by CMGE + Mcm10 + MTC, fitted to a single-exponential decay function. The total product length (1.7 $\pm$ 0.4 knt) is 1.5-fold higher than the value found in (b). (d) Histogram of the total product length per trajectory using CMGE + Mcm10 + TC, fitted to a single-exponential decay function. The total product length (1.0 $\pm$ 0.2 knt) is similar as obtained without TC (N = 111 trajectories). In all, the short values are undersampled and not included in the fits.

nM MTC for 5 min at 30°C, followed by addition of 5 nM RFC, 25 nM PCNA, and 10 nM Pol ϵ for 4 min in the presence of dATP and dCTP to promote clamp loading and polymerase binding and to prevent the 3′–5′ exonuclease activity of Pol ϵ from removing the primer. Reactions were started by addition of the withheld nucleotides (dGTP and dTTP), 5 mM ATP, and either 600 nM *S. cerevisiae* RPA or 600 nM *E.coli* SSB as indicated. Reactions were allowed to proceed for the indicated times at 30°C and then quenched by adding an equal volume of 2X stop solution (40 mM EDTA and 1% (w/v) SDS). Reactions were run on 1.3% (w/v) alkaline agarose gels at 35 V for 16 h, backed with DE81 paper, and compressed for 12 h. Gels were exposed to a phosphorimager screen and imaged with a Typhoon FLA 9500 (GE Healthcare).

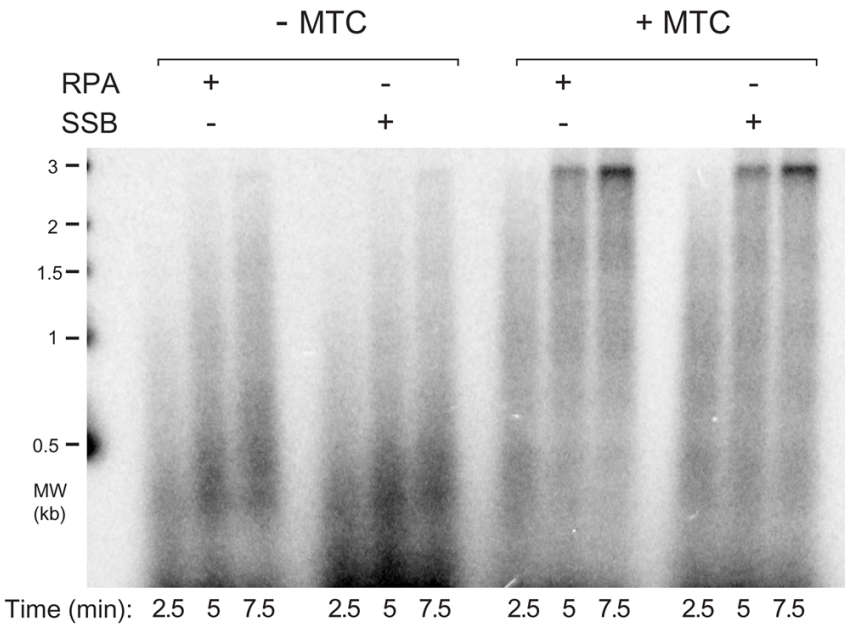


Figure 8.11: SSB and RPA are interchangeable for leading-strand replication. Alkaline agarose gel of leading-strand products by CMGE leading-strand replisomes Reactions were performed as described in Materials and Methods but included 400 μM dNTPs. Reactions in the presence of either RPA or SSB are shown, both with (lanes 7–12) and without MTC (lanes 1–6). Reactions were stopped at the indicated times.

8.4.7 Code availability

Analysis was performed using custom Python software programmed in house. Source code for most analysis tools is available at GitHub under Single-Molecule Biophysics beadpy. Additional source code used for transition density analysis is available upon request.

8.5 Supplementary figures

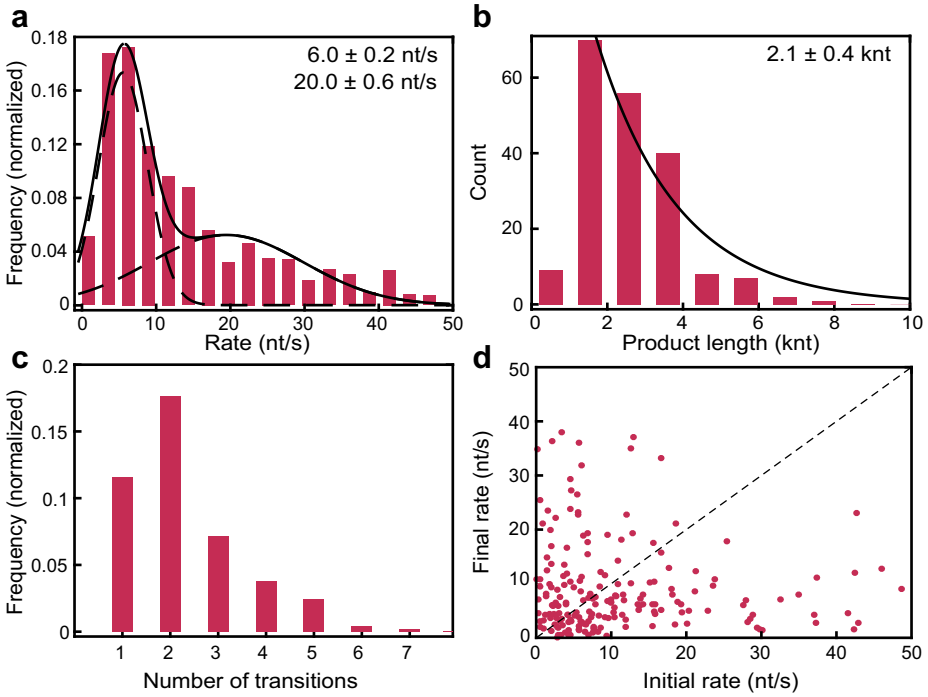


Figure 8.12: C-terminally tagged MTC (MfTC) transiently interacts with the replisome. (a) Histogram of the instantaneous single-molecule rates, weighted by segment length for replication by CMGE + Mcm10 + 30 nM MfTC. The histogram shows a bimodal distribution of the rates. The data were fit with the sum of two Gaussian distributions (black line), resulting in a rate of 6.0 ± 0.2 nt/s for the slow population and 20.0 ± 0.6 nt/s for the fast population (N = 195 trajectories). (b) Histogram of the total product length per trajectory for replication by CMGE + Mcm10 + 30 nM MfTC. A single-exponential fit (black line) shows that the total product length is similar to the value measured with MTC (1.7 ± 0.4 knt). (c) The number of rate changes per trajectory with MfTC is similar to MTC (Figure 8.4). (d) Transition plot showing the rate of a segment as a function of the rate of the previous segment for trajectories with multiple segments, with MfTC present. The distance from the diagonal (dashed line) is (12.1 ± 0.9 nt/s, mean ± s.e.m) similar to MTC (Figure 8.4).

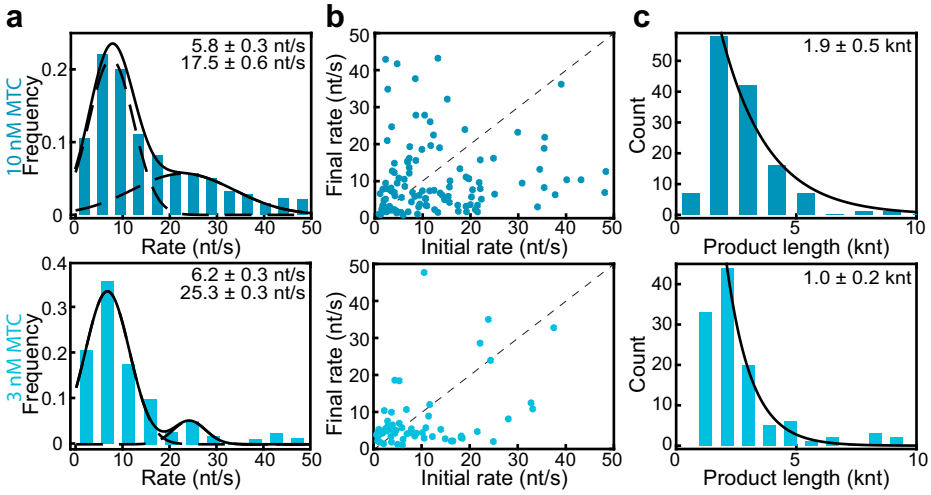


Figure 8.13: Lower MTC concentrations result in a reduction in the number of fast rates as well as the frequency of transitions within a single trajectory. (a) (Top) Histogram of the instantaneous single-molecule rates, weighted by segment length for replication by CMGE + Mcm10 + 10 nM MTC. The histogram shows a bimodal distribution of the rates. The data were fit with the sum of two Gaussian distributions (black line), resulting in a rate of 5.8 $\pm$ 0.3 nt/s for the slow population and 17.5 $\pm$ 0.6 nt/s for the fast population (N = 251 trajectories). (Bottom) Histogram of the instantaneous single-molecule rates, weighted by segment length for replication by CMGE + Mcm10 + 3 nM MTC. The data were fit with the sum of two Gaussian distributions (black line), resulting in a rate of 6.2 $\pm$ 0.3 nt/s for the slow population and 25.3 $\pm$ 0.3 nt/s for the fast population (N = 184 trajectories). (b) Transition plots showing the rate of a segment as a function of the rate of the previous segment for trajectories with multiple segments, with 10 nM (top) and 3 nM (bottom) MTC present. The perpendicular distance from the diagonal (dashed line) is $\sim$2-fold lower when 3 nM MTC is present (4.5 $\pm$ 0.7 nt/s, mean $\pm$ s.e.m) compared with 30 nM MTC (9.5 $\pm$ 0.6 nt/s). (c) (Top) Histogram of the total product length per trajectory for replication by CMGE + Mcm10 + 10 nM MTC. A single-exponential fit (black line) shows that the product length is the same as the the value measured with 30 nM MTC (1.9 $\pm$ 0.5 knt). (Bottom) Histogram of the total product length per trajectory for replication by CMGE + Mcm10 + 10 nM MTC. A single-exponential fit (black line) shows that the total product length is similar to the value measured without MTC (1.0 $\pm$ 0.2 knt).

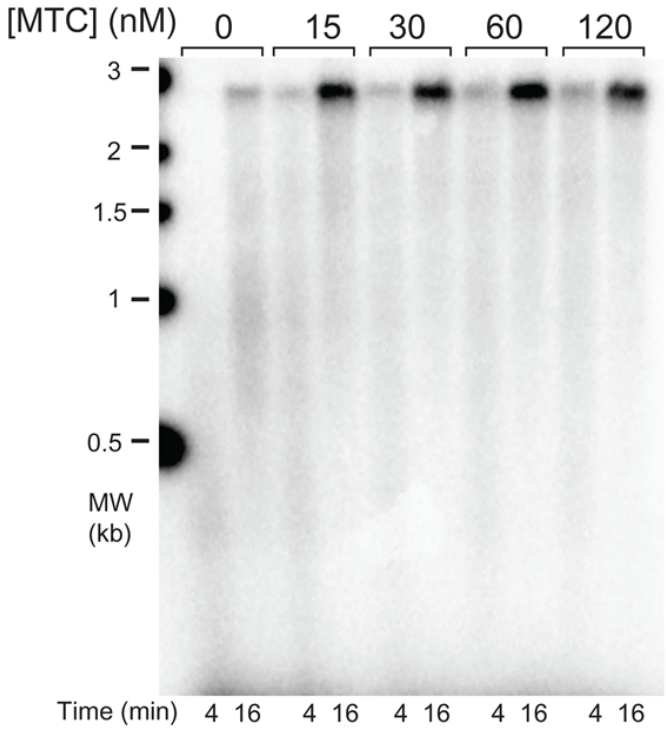


Figure 8.14: MTC titration into leading-strand replisome reactions. Alkaline agarose gel of leading-strand products at different concentrations of MTC indicated above the gel. Reactions were stopped at the indicated times below the gel. See Methods for details.

9 Discussion

Ever since Watson and Crick solved the structure of DNA, people have studied its mechanism of replication. Through decades of study, we now have a good understanding of the function of the different proteins within the replisome, the machinery responsible for DNA replication. Initially with studies of bacterial and bacteriophage replisomes and more recently of eukaryotic systems, a wealth of information has been gathered on their function and structure. The ensemble-averaging nature of classical approaches have made it difficult, though, to gain access to the dynamics of individual proteins and the interactions between them. Over the past decade, single-molecule biophysical techniques have advanced enormously and they are now regularly being used to gain critical mechanistic insights into the molecular processes underlying life (Chapter 2).

9.1 Improving single-molecule techniques

We have developed a fluorescence imaging technique by which single fluorescent molecules can be observed in real time at high, physiologically relevant concentrations, conditions that are typically incompatible with single-molecule imaging approaches (Chapter 3). The technique requires a protein and its macromolecular substrate to be labelled each with a different fluorophore. The fluorophore bound to the protein is chemically darkened and can therefore be used at a very high concentration, without contributing to background fluorescence. Making use of short-distance energy-transfer mechanisms between the two fluorophores, only the fluorescence from those proteins that bind to their substrate is activated. We have shown that this technique allows the use of fluorescent probes up to μM concentrations — concentrations well above most biologically relevant concentrations. This approach opens up the use of single-molecule fluorescence imaging to study weak interactions between macromolecules.

Single-molecule techniques are mostly used in basic research (Chapter 2), but could (or should) be used in translational research as well. As an example, in Chapter 4 we use singe-molecule fluorescence imaging to quantify the density of proteins on functionalised liposomes. Ligand-directed liposomes have been proposed as a potential drug-delivery vehicle, but despite promising work, no ligand-directed liposomes have been translated into the clinic as a viable drug-delivery strategy. This gap is partly due to the lack of a robust characterisation method for the ligand density on the liposome surface. We show here that we can accurately determine the number of labelled ligands on liposomes functionalised with two different ligands. We then used these data to inform us on the quality of different liposome preparation methods. In the future, however, this technique could be used in a pharmaceutical setting as an accurate, batch-to-batch quality control method. To make this a more generally applicable method, the ligands could be labelled using fluorescent antibodies, instead of covalently labelling the ligands with a fluorophore.

9.2 Multi-site exchange mechanisms

The *E.coli* DNA polymerase Pol III* is the protein complex within the replisome responsible for synthesising new DNA on the two daughter strands. Based on solid biochemical data, textbooks suggest that Pol III* remains stably bound to the replisome (2). We used fluorescently labelled polymerases to show that under physiologically relevant concentrations the Pol III* complex exchanges very rapidly during coupled leading- and lagging-strand DNA synthesis (Chapter 5). If there are no competing polymerases available in solution, however, the original Pol III* complex remains very stably bound to the replisome. Consistent with the original biochemical studies that relied on pre-assembly of the replisome followed by a rapid dilution, this observation recapitulates the textbook model: a Pol III* complex that is tightly bound to the replisome under 'infinite dilution'.

This concentration-dependent exchange seems irreconcilable with the concept that dissociative mechanisms are independent on concentration, but can be explained by a multi-site exchange mechanism. Such a mechanism depends on multiple weak protein–protein and protein–DNA interactions, and is best illustrated using the analogy with the monkey on a tree branch (Figure 9.1) (139). Proteins (monkeys) are bound to the larger protein complex (tree branch) via multiple binding sites. Due to the weak nature of the interactions, the proteins will transiently unbind from each site. Under physiologically relevant protein concentrations, a new protein can bind at a vacated site and then completely displace the original protein. Under dilute conditions, however, the original protein is retained within the complex through its other binding sites. It can then rapidly re-associate with the vacated binding site, and thus remain stably bound to the complex. Pol III* has several weak binding sites to facilitate this mechanism. The α and ϵ subunits of the core polymerase both have binding sites on the β_2 clamp and on the DNA 3′ end. The χ subunit of the CLC interacts with the C-terminal tail of SSB. Finally, the τ subunit has three binding sites on the DnaB helicase (unpublished work Dixon, van Oijen groups) and could therefore function as a reservoir for Pol III* to increase the local concentration, allowing exchange to happen even faster.

This balance between stability and plasticity may seem messy and inelegant at first. The ability to exchange polymerases could, however, be a pathway through which a Pol III* can easily be replaced, without the need for other proteins to remove it, and without the need for replication stalling. Exchange thereby ensures the continuation of replication with high fidelity. Exchange could also provide a way to bypass DNA lesions. It has been shown that Pol III* can not synthesise past certain types of DNA damage (421). When it encounters these types of damage, it could halt, or rather, go through futile cycles of exonuclease and polymerisation. If, however, the DnaB helicase is able to move past the leasion, Pol III* exchange would allow the binding of a new Pol III*, and replication could continue downstream, past the leasion (see also Chapter 7). Furthermore, exchange could give other binding partners access to the replisome when needed. Examples of

another binding partners that might need access to the replisome are translesions synthesis (TLS) DNA polymerases involved in lesion bypass and repair. Some of these polymerases are known to interact with the replisome and displace Pol III* at sites of DNA damage (279). Before the cellular detection of DNA damage, the concentrations of repair polymerases within the cell are low. Therefore they do not compete with Pol III*. When DNA damage occurs, however, the TLS polymerases are up regulated and their concentration increases (422). This up-regulation will not only increase the chance that a TLS polymerase will bind at the replisome, but it will actually stimulate the dissociation of Pol III* through the multi-site exchange mechanism. Though seemingly complicated at first glance, the existence of this exchange pathway eliminates the need for complex signalling pathways and obeys fundamental chemical and thermodynamic principles.

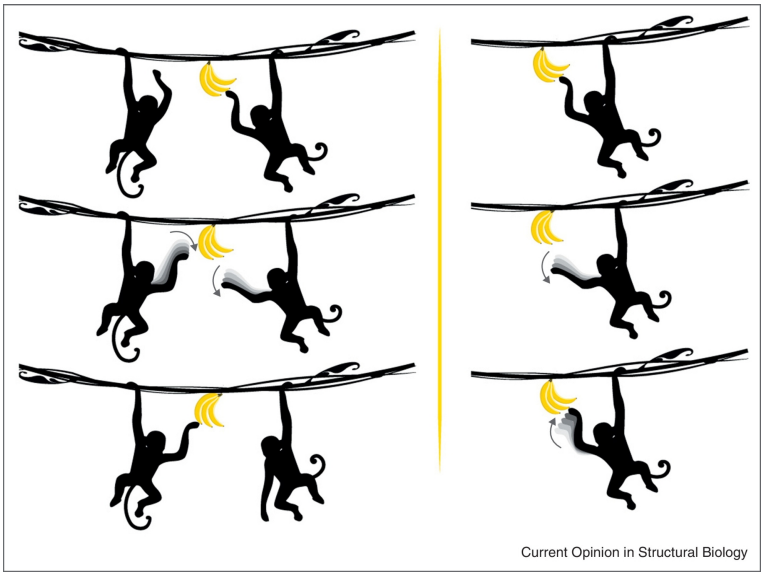


Figure 9.1: Monkey analogy for the multi-site exchange mechanism. The monkeys can hold on to the tree in two positions (sites); the branch and the bananas. (left) Competition of two monkeys (high concentration) for the same hand of bananas (binding site). Transient dissociation from the bananas by the monkey on the right allows the left monkey to compete for the same bananas. (right) With just one monkey present (dilute conditions), temporary unbinding from the bananas still allows a rapid re-association by the same monkey. Figure adapted from (139).

In Chapter 6, we demonstrate the existence of a similar multi-site exchange mechanism for single-stranded DNA binding proteins (SSBs). By binding to transiently exposed single-stranded DNA, SSBs prevent nucleolytic attacks and the formation of secondary structures (32). As new DNA is synthesised on the single-stranded DNA template, the SSBs have to be removed. We show that, at low SSB concentrations in solution, these removed SSBs can immediately rebind at newly exposed single-stranded DNA that is nearby. As a result, the SSB is effectively recycled within the replisome. This internal transfer is facilitated by the fact that SSB can utilise up to four DNA binding sites per functional SSB tetramer. Under our experimental conditions only two of these binding sites will be occupied. SSBs can transfer between DNA strands by binding of the second DNA strand via the remaining two binding sites. As in the case

of the polymerases, SSBs will exchange with SSBs from solution at high SSB concentrations.

Recycling of SSB can not be explained by the monkey model as it is described above. Since the SSB has to be removed from the DNA template when new DNA is synthesised, the model needs to include dissociation of a factor from its original binding site as part of a retention process. In other words, the monkey has to move. A revised model (Figure 9.2) illustrates how a protein can be recycled or retained within a multi-protein complex, even though it has to unbind from its initial binding site. If a protein has multiple weak interactions within the replisome, like the monkey has two hands and feet, it can use these interactions to transfer from one site to the next. Due to the high local concentration of binding sites within the replisome, this internal transfer is the most likely pathway when the concentration of proteins in solution is low. At high protein concentrations in solution, however, proteins from solution can compete for the same binding sites, like the monkeys compete for the branch. This results in a competition between internal transfer and external exchange.

The multi-site exchange mechanism is presumably not limited to DNA-replication machineries, but can likely be extended to any multi-protein complex. For example, concentration-dependent exchange has been observed for other DNA-binding proteins (209), the subunits of the flagellar motor stator (241, 423), and in transcription regulation (138). The apparent generality of the models emerging from this work suggests that the behaviours of even more complex systems are also governed by this process.

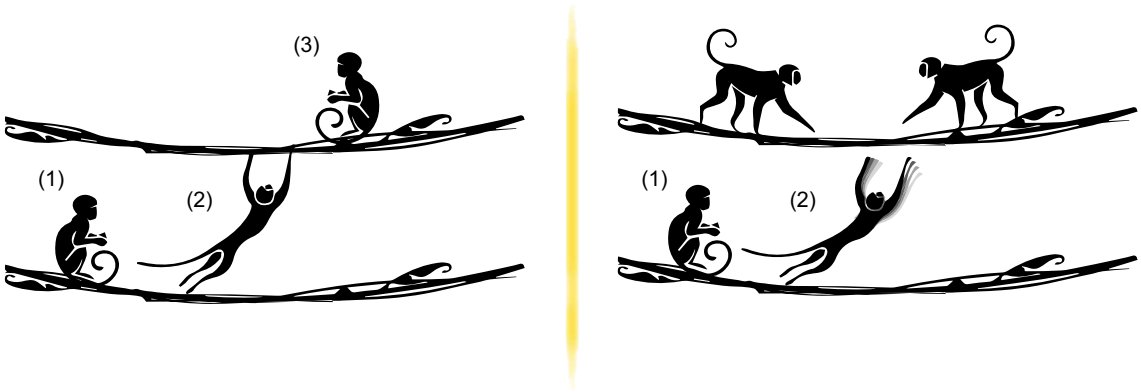


Figure 9.2: New Monkey model. (left) When there are no other monkeys around (low concentration), the monkey is free to jump between the branches of the tree (retention/recycling). (right) When there are other monkeys in the tree (high concentration) these monkeys might jump to the second branch first (competition), leaving no space for the first monkey. When the first monkey jumps, it will now fall out of the tree (exchange).

9.3 Replication and repair

DNA replication is intimately linked to other processes related to genomic metabolism, such as repair and transcription. There is a large number of

mechanisms in place to regulate all these processes and to resolve any conflicts between them (424). In Chapter 7, we look at the effect of *E.coli* RarA on DNA replication and repair. RarA is known to be one of the strongest binding partners to SSB (308). Furthermore, RarA is known to have an effect on replisome stability and promote translesion DNA synthesis (TLS) (315,316).The occurrence of homologous proteins in other domains of life suggest an important role, but its precise function and mechanism of action remain unknown. We use a combination of *in vitro* and *in vivo* assays to determine the role of RarA in DNA replication and repair. Using single-molecule fluorescence imaging, we show that the presence of high concentrations of RarA in an *in vitro* replication assay generates large ssDNA gaps on the lagging-strand product. Using fluorescently labelled β_2 clamps, we show that β_2 clamps get left behind at the 3′ end of these gaps. Furthermore, we show that the gap size depends on β_2 concentration. In live-cell imaging assays we show that deleting RarA suppresses the sensitivity to various DNA damaging agents in strains lacking TLS polymerases . We also show that deleting RarA results in a substantial decrease in growth rate. From our observations, we conclude that RarA is involved in creating a substrate for TLS polymerases. This last conclusion is consistent with our *in vitro* data, as the lagging-strand gaps with β_2 clamps, would act as a very suitable substrate for TLS polymerases. Combining all this information, we propose a model in which RarA enables the replisome to skip over lesions, leaving behind an ssDNA gap, but allowing replication to continue. RarA, thereby commits the cell to TLS or daughter-strand gap repairby providing optimal substrates for these processes.

TLS polymerases are known to be low-fidelity polymerases (17) and are responsible for a high frequency of mutations. RarA pushes the cell to use these polymerases to do DNA repair instead of using other, perhaps less mutagenic pathways. As cells are viable without RarA, one could ask the question what the benefit is of having RarA. The answer to this question could be the increased growth rate. The fact that cells with RarA grow significantly faster, probably provides them with an advantage over cells without RarA, a benefit that might offset the negative impact of TLS.

The textbook models of translesion synthesis suggests that TLS polymerases work in the context of the replisome. It is assumed that TLS polymerases exchange into the replisome, synthesize past the lesion, and then allow pol III replication to continue. Recently, it has been shown that the TLS polymerases mostly act outside of the replisome (110, 345). The action of RarA, allowing the replisome to skip the lesion, may very well facilitate this behaviour.

9.4 A more complex replisome

Single-molecule studies on relatively simple replication systems such as the bacteriophage T7 and *E. coli* have not only taught us about the important dynamic processes governing DNA replication, but they have also enabled us to develop the single-molecule techniques needed to visualise these processes.

These technological and conceptual advances are particularly relevant now, with the recent achievement of the successful *in vitro* reconstitution of a functional *S. cerevisiae* replisome. We are, therefore, now at a point where we can apply the lessons we have learned from the simpler systems to more complex, eukaryotic replication systems. In Chapter 8, we used a single-molecule tethered-bead assay to study the kinetics of the yeast leading-strand replisome. This work was the first single-molecule visualization of real-time DNA replication by a reconstituted eukaryotic system.

We revealed a highly dynamic interaction between the MTC complex and the leading-strand replisome. These dynamics were not observed in previous ensemble-averaging biochemical studies (388), highlighting the importance of single-molecule experiments. Since it is known that MTC undergoes many post-translational modifications, such as phosphorylation, we hypothesize that the dynamic interaction of MTC with the replisome could play a role in the functional coupling between MTC phosphorylation state and replisome speed. If MTC were stably bound to the replisome, it is possible that only a subset of replisomes could carry a fully phosphorylated MTC complex. The dynamic interaction, however, could ensure a complete sampling of the phosphorylation states of MTC by all replisomes.

In eukaryotes, post-translational modifications play an important role in all processes related to replication, such as replication initiation (417–419), replisome stability (405, 406), and fork stalling (416). Replication proteins may undergo several rounds of modification during the different phases of the cell cycle. Obtaining the physiologically relevant post-translational modifications, and their timing during replication, is very challenging when using reconstituted systems. Reconstituted systems could, therefore, be useful platforms to gain detailed information on post-translational modifications.

Even though it is tempting to speculate that the dynamic interaction between MTC and the replisome has a biological function, as explained above, it is also important to note that our experiments were done with the simplified, leading-strand replisome. We should consider that the presence of the other replisome components could change the interaction. It is therefore essential to develop an assay, that allows us to monitor coupled leading- and lagging-strand synthesis in the context of a fully reconstituted replisome.

Throughout this thesis, I have explained how we used a knowledge base founded on the study of relatively simple systems, to increase biological complexity and finally study highly complex systems such as the yeast replisome. It is important to emphasize that, although studying the more complex systems from higher organisms is very important, we should not stop studying the simpler systems. There are still a lot of open questions related to DNA replication in the simpler systems such as that of *E. coli* that will teach us about basic operating principles of molecular machines.

Answering these questions is not only important to later inform us on the more complicated systems — it is just as important to obtain a deep molecular understanding of the behaviour of the bacteria and viruses. A case in point is the emergence of antibiotic resistance. According to some predictions antibiotic resistance will be responsible for the number one cause of death by 2050 if no additional action is taken (425). Currently, there are no known approaches to stop the development of resistance. DNA replication and repair play an important role in both cell survival and the development of resistance (426). Bacterial replication proteins are, therefore, obvious drug targets. Having an accurate understanding of their interactions is going to be crucial in the development of new antimicrobial drugs.

The challenges in the near future lie in improving single-molecule techniques even further, allowing us to study biological systems in even more detail. In the last few years, we have started combining single-molecule *in vitro* techniques, with single-molecule imaging in live cells (233) (Chapters 5, 6, and 7). The *in vitro* experiments offer a precise handle on the experimental conditions in a controlled environment, while the *in vivo* measurements provide the physiologically relevant complexity that exists in live cells. Bridging the gap between the two approaches will allow us to further elucidate the dynamic behaviour of proteins within multi-protein complexes.

10 References

(1) J.F. Miescher. Miescher letter I; to Wilhelm His. In W. His, editor, *Die Histochemischen und Physiologischen Arbeiten von Friedrich Miescher - Aus dem wissenschaftlichen Briefwechsel von F. Miescher*, volume 1, pages 33–38. FCW Vogel, 1869.

(2) B. Alberts, A. Johnson, J. Lewis, M. Raff, K. Roberts, and P. Walter. *Molecular Biology of the Cell*. Garland Science, 6 edition, 2014.

(3) J.D. Watson and F.H.C. Crick. Molecular structure of nucleic acids: a structure for deoxyribose nucleic acid. *Nature*, 171(4356):1197–1200, 1953.

(4) R.E. Franklin and R.G. Gosling. Molecular configuration in sodium thymonucleate. *Nature*, 171(4356):740–741, Apr 1953.

(5) I.R. Lehman, M.J. Bessman, E.S. Simms, and A. Kornberg. Enzymatic synthesis of deoxyribonucleic acid. i. preparation of substrates and partial purification of an enzyme from *Escherichia coli*. *J. Biol. Chem*, 233(1):163–170, 1958.

(6) R. Okazaki, T. Okazaki, K. Sakabe, K. Sugimoto, and A. Sugino. Mechanism of DNA chain growth. I. Possible discontinuity and unusual secondary structure of newly synthesized chains. *Proc. Natl. Acad. Sci. U.S.A.*, 59(2):598–605, Feb 1968.

(7) L. Hayflick. The limited *in vitro* lifetime of human diploid cell strains. *Exp. Cell. Res.*, 6:614–636, 3 1965.

(8) International Human Genome Sequencing Consortium. Finishing the euchromatic sequence of the human genome. *Nature*, 431(7011):931–945, 2004.

(9) Richardson CC. Bacteriophage T7: minimal requirements for the replication of a duplex DNA molecule. *Cell*, 33(2):315–317, 1983.

(10) A.K. Satapathy, A.W. Kulczyk, S. Ghosh, A.M. van Oijen, and C.C. Richardson. Coupling dTTP hydrolysis with DNA unwinding by the DNA helicase of bacteriophage T7. *J. Biol Chem*, 286(39):34468–34478, 2011.

(11) J.A. Bernstein and C.C. Richardson. A 7-kDa region of the bacteriophage T7 gene 4 protein is required for primase but not for helicase activity. *Proc.Natl.Acad.Sci. U.S.A.*, 85(2):396–400, 1988.

(12) S. Tabor, H.E. Huber, and C.C. Richardson. *Escherichia coli* thioredoxin confers processivity on the DNA polymerase activity of the gene 5 protein of bacteriophage T7. *J. Biol. Chem*, 262(33):16212–16223, 1987.

(13) M. Yu and W. Masker. T7 single strand DNA binding protein but not T7 helicase is required for DNA double strand break repair. *J. Bacteriol*, 183(6):1862–1869, 2001.

(14) N.A. Tanner, J.J. Loparo, S.M. Hamdan, S. Jergic, N.E. Dixon, and A.M. van Oijen. Real-time single-molecule observation of rolling-circle DNA replication. *Nucleic Acids Research*, 37(4):e27, 2009.

(15) A. Robinson and A.M. van Oijen. Bacterial replication, transcription and translation: mechanistic insights from single-molecule biochemical studies. *Nature Reviews Microbiology*, 11(5):303–315, 2013.

(16) J.S. Lewis, S. Jergic, and N.E. Dixon. The *E. coli* replication fork. In Lauri S. Kaguni and Marcos Tulio Oliveira, editors, *The Enzymes*, chapter 2, pages 31–87. Elsevier, Oxford, 2016.

(17) I.J. Fijalkowska, R.M. Schaaper, and Jonczyk P. DNA replication fidelity in *Escherichia coli*: a multi-DNA polymerase affair. *FEMS Microbiol Rev*, 36(6):1105–1121, 2012.

(18) T.A. Baker, B.E. Funnell, and A. Kornberg. Helicase action of DnaB protein during replication from the *Escherichia coli* chromosomal origin *in vitro*. *J. Biol. Chem.*, 262:6877–6885, May 1987.

(19) F. Dong, E.P. Gogol, and P.H. von Hippel. The phage T4-coded DNA replication helicase (gp41) forms a hexamer upon activation by nucleoside triphosphate. *J. Biol. Chem.*, 270(13):7462–7473, 1995.

(20) S. Bailey, W.K. Eliason, and T.A. Steitz. Structure of hexameric DnaB helicase and its complex with a domain of DnaG primase. *Science*, 318:459–63, Oct 2007.

(21) K. Yoda and T. Okazaki. Specificity of recognition sequence for *Escherichia coli* primase. *Mol. Gen. Genet.*, 227(1):1–8, 1991.

(22) L. Aravind, D.D. Leipe, and E.V. Koonin. Toprim - a conserved catalytic domain in type IA and II topoisomerases, DnaG-type primases, old family nucleases and RecR proteins. *Nucleic Acids Research*, 26(18):4205–4213, 1998.

(23) A.J. Oakley, K.V. Loscha, P.M. Schaeffer, E. Liepinsh, G. Pintacuda, M.C. Wilce, G. Otting, and N.E. Dixon. Crystal and solution structures of the helicase-binding domain of *Escherichia coli* primase. *J. Biol. Chem.*, 280(12):11495–504, Mar 2005.

(24) C.S. McHenry and W. Crow. DNA polymerase III of *Escherichia coli*. purification and identification of subunits. *J. Biol. Chem.*, 254(5):1748–1753, 1979.

(25) H. Maki and A. Kornberg. The polymerase subunit of DNA polymerase III of *Escherichia coli*. II. Purification of the alpha subunit, devoid of nuclease activities. *J. Biol. Chem.*, 260(24):12987–12992, Oct 1985.

(26) X.P. Kong, R. Onrust, M.E. O'Donnell, and J. Kuriyan. Three-dimensional structure of the beta subunit of *E. coli* DNA polymerase III holoenzyme: a sliding DNA clamp. *Cell*, 69(3):425–237, 1992.

(27) A. Robinson, R.J. Causer, and N.E. Dixon. Architecture and conservation of the bacterial DNA replication machinery, an underexploited drug target. *Curr. Drug Targets*, 13(3):352–372, 2012.

(28) C.S. McHenry. DNA replicases from a bacterial perspective. *Annu. Rev. Biochem.*, 80:403–436, 2011.

(29) D. Gao and C.S. McHenry. tau binds and organizes *Escherichia coli* replication proteins through distinct domains. Domain IV, located within the unique C terminus of tau, binds the replication fork, helicase, DnaB. *J. Biol. Chem*, 276(6):4441–4446, 2001.

(30) J.M. Gulbis, S.L. Kazmirski, J. Finkelstein, Z. Kelman, M.E. O'Donnell, and J. Kuriyan. Crystal structure of the chi:psi sub-assembly of the *Escherichia coli* DNA polymerase clamp-loader complex. *Eur. J. Biochem.*, 271(2):439–449, 2004.

(31) R.R. Meyer and P.S. Laine. The single-stranded DNA-binding protein of *Escherichia coli*. *Microbiol. Rev.*, 54(4):342–380, 1990.

(32) R.D. Shereda, A.G. Kozlov, T.M. Lohman, M.M. Cox, and J.L. Keck. SSB as an organizer/mobilizer of genome maintenance complexes. *Crit. Rev. Biochem. Mol. Biol.*, 43(5):289–318, 2008.

(33) S. Raghunathan, A.G. Kozlov, T.M. Lohman, and G. Waksman. Structure of the DNA binding domain of *E. coli* SSB bound to ssDNA. *Nat. Struct. Biol*, 7(8):648–652, 2000.

(34) T.M. Lohman and M.E. Ferrari. *Escherichia coli* single-stranded DNA binding protein: Multiple DNA-binding modes and cooperativities. *Annual Review of Biochemistry*, 63:527–570, 1994.

(35) A.G. Kozlov, M.M. Cox, and T.M. Lohman. Regulation of single stranded DNA binding by the c-termini of *E. coli* SSB protein. *J. Biol. Chem.*, 285(22):17246–17252, 2010.

(36) R.R. Meyer, J. Glassberg, and Kornberg A. An *Escherichia coli* mutant defective in single-strand binding protein is defective in DNA replication. *Proc. Natl. Acad. Sci.*, 76(4):1702–1705, 4 1979.

(37) LB Overman, W Bujalowski, and Timothy M. Lohman. Equilibrium binding of *Escherichia coli* single-strand binding protein to single-stranded nucleic acids in the (SSB)65 binding mode. cation and anion effects and polynucleotide specificity. *Biochemistry*, 27(1):456–471, 01 1988.

(38) A.G. Kozlov and T.M. Lohman. Kinetic mechanism of direct transfer of *Escherichia coli* SSB tetramers between single-stranded DNA molecules. *Biochemistry*, 41(39):11611–11627, 2002.

(39) R.E. Georgescu, L. Langston, N.Y. Yao, O. Yurieva, D. Zhang, J. Finkelstein, T. Agarwal, and O'Donnell M.E. Mechanism of asymmetric polymerase assembly at the eukaryotic replication fork. *Nat. Struct. Mol. Biol.*, 21(8):664–670, 2014.

(40) M.L. Bochman and A. Schwacha. The Mcm2–7 complex has *in vitro* helicase activity. *Mol. Cell*, 31(2):287–293, 2008.

(41) I. Ilves, T. Petojevic, J.J. Pesavento, and M.R. Botchan. Activation of the mcm2-7 helicase by association with cdc45 and gins proteins. *Mol Cell*, 37(2):247–258, 2010.

(42) H. Singh, R.G. Brooke, M.H. Pausch, G.T. Williams, C. Trainor, and L.B. Dumas. Yeast DNA primase and DNA polymerase activities. an analysis of rna priming and its coupling to DNA synthesis. *J. Biol. Chem.*, 261(18):8564–8569, 1986.

(43) J. Sun, Y. Shi, R.E. Georgescu, Z. Yuan, B.T. Chait, H. Li, and M.E. O'Donnell. The architecture of a eukaryotic replisome. *Nat. Struct. Mol. Biol*, 22(12):976–982, 2015.

(44) D. Jeruzalmi, M.E. O'Donnell, and J. Kuriyan. Clamp loaders and sliding clamps. *Curr. Opin. Struct. Biol*, 12(2):217–224, 2002.

(45) A.E. Knight. *Single Molecule Biology*. Elsevier, 2009.

(46) Z. Zhang, M.M. Spiering, M.A. Trakselis, F.T. Ishmael, S.J. Xi, J. abd Benkovic, and G.G. Hammes. Assembly of the bacteriophage t4 primosome: single-molecule and ensemble studies. *Proc. Natl.Acad. Sci. U.S.A.*, 102(9):3254–3259, 2005.

(47) W. Lee, D. Jose, C. Phelps, A.H. Marcus, and P.H. von Hippel. A single-molecule view of the assembly pathway, subunit stoichiometry, and unwinding activity of the bacteriophage T4 primosome (helicase-primase) complex. *Biochemistry*, 52(18):157–170, 2013.

(48) W.K. Cho, S Jergic, D. Kim, N.E. Dixon, and J.B. Lee. Loading dynamics of a sliding DNA clamp. *Angew. Chem. Int. Ed. Engl.*, 53(26):6768–6771, 2014.

(49) Z. Debyser, S. Tabor, and C. C. Richardson. Coordination of leading and lagging strand DNA synthesis at the replication fork of bacteriophage T7. *Cell*, 77(1):157–166, Apr 1994.

(50) Geertsema H.J., Kulczyk. A.W., C.C. Richardson, and A.M. van Oijen. Single-molecule studies of polymerase dynamics and stoichiometry at the

bacteriophage t7 replication machinery. *Proc. Natl. Acad. Sci. U.S.A.*, 111(11):4073–4078, 2014.

(51) Geertsema H.J., K.E. Duderstadt, and A.M. van Oijen. Single-molecule observation of prokaryotic DNA replication. *Methods Mol. Biol.*, 1300:219–238, 2015.

(52) K.E. Duderstadt, H.J. Geertsema, S.A. Stratmann, C.M. Punter, A.W. Kulczyk, C.C. Richardson, and A.M. van Oijen. Simultaneous real-time imaging of leading and lagging strand synthesis reveals the coordination dynamics of single replisomes. *Molecular Cell*, 64:1–13, 2016.

(53) F.R. Hill, E. Monachino, and A.M. van Oijen. The more the merrier: high-throughput single-molecule techniques. *Biochem. Transact. Soc*, 3(45):759–769, Jun 2017.

(54) F.R. Hill, A.M. van Oijen, and K.E. Duderstadt. Detection of kinetic change points in piece-wise linear single molecule motion. *J. Chem. Phys.*, 148, 2018.

(55) A.M. van Oijen and N.E. Dixon. Probing molecular choreography through single-molecule biochemistry. *Nat. Struct. Mol. Biol.*, 22(12):948–952, 12 2015.

(56) C.V. Robinson, A. Sali, and W. Baumeister. The molecular sociology of the cell. *Nature*, 450(7172):973–982, 2007.

(57) J. N. Onuchic, Z. Luthey-Schulten, and P. G. Wolynes. Theory of protein folding: the energy landscape perspective. *Annu Rev Phys Chem*, 48:545–600, 1997.

(58) K. C. Neuman and A. Nagy. Single-molecule force spectroscopy: optical tweezers, magnetic tweezers and atomic force microscopy. *Nat. Methods*, 5(6):491–505, Jun 2008.

(59) W. J. Greenleaf, M. T. Woodside, and S. M. Block. High-resolution, single-molecule measurements of biomolecular motion. *Annu Rev Biophys Biomol Struct*, 36:171–190, 2007.

(60) D. Dulin, J. Lipfert, M. C. Moolman, and N. H. Dekker. Studying genomic processes at the single-molecule level: introducing the tools and applications. *Nat. Rev. Genet.*, 14(1):9–22, Jan 2013.

(61) N. de Souza. Pulling on single molecules. *Nat. Methods*, 9(9):873–877, Sep 2012.

(62) T. Ando. High-speed AFM imaging. *Curr. Opin. Struct. Biol.*, 28:63–68, Oct 2014.

(63) Y. L. Lyubchenko and L. S. Shlyakhtenko. Imaging of DNA and Protein-DNA Complexes with Atomic Force Microscopy. *Crit. Rev. Eukaryot. Gene Expr.*, 26(1):63–96, 2016.

(64) T. Ando, T. Uchihashi, and S. Scheuring. Filming biomolecular processes by high-speed atomic force microscopy. *Chem. Rev.*, 114(6):3120–3188, Mar 2014.

(65) D. J. Muller and Y. F. Dufrene. Force nanoscopy of living cells. *Curr. Biol.*, 21(6):R212–216, Mar 2011.

(66) B. H. Blehm and P. R. Selvin. Single-molecule fluorescence and *in vivo* optical traps: how multiple dyneins and kinesins interact. *Chem. Rev.*, 114(6):3335–3352, Mar 2014.

(67) A. M. Whited and P. S. Park. Atomic force microscopy: a multifaceted tool to study membrane proteins and their interactions with ligands. *Biochim. Biophys. Acta*, 1838(1 Pt A):56–68, Jan 2014.

(68) Sergio Santos, Victor Barcons, Hugo K. Christenson, Daniel J. Billingsley, William A. Bonass, Josep Font, and Neil H. Thomson. Stability, resolution, and ultra-low wear amplitude modulation atomic force microscopy of DNA: Small amplitude small set-point imaging. *Applied Physics Letters*, 103(6), 2013.

(69) D. A. Walters, J. P. Cleveland, N. H. Thomson, P. K. Hansma, M. A. Wendman, G. Gurley, and V. Elings. Short cantilevers for atomic force microscopy. *Review of Scientific Instruments*, 67(10):3583–3590, 1996.

(70) T Ando, T Uchihashi, and T Fukuma. High-speed atomic force microscopy for nano-visualization of dynamic biomolecular processes. *Progress in Surface Science*, 83(7-9):337 – 437, 2008.

(71) N. Kodera, D. Yamamoto, R. Ishikawa, and T. Ando. Video imaging of walking myosin V by high-speed atomic force microscopy. *Nature*, 468(7320):72–76, Nov 2010.

(72) H. Watanabe, T. Uchihashi, T. Kobashi, M. Shibata, J. Nishiyama, R. Yasuda, and T. Ando. Wide-area scanner for high-speed atomic force microscopy. *Rev. Sci. Instrum.*, 84(5):053702, May 2013.

(73) C. L. Essmann, M. Elmi, M. Shaw, G. M. Anand, V. M. Pawar, and M. A. Srinivasan. In-vivo high resolution AFM topographic imaging of Caenorhabditis elegans reveals previously unreported surface structures of cuticle mutants. *Nanomedicine*, 13(1):183–189, Oct 2016.

(74) H. Li, W. A. Linke, A. F. Oberhauser, M. Carrion-Vazquez, J. G. Kerkvliet, H. Lu, P. E. Marszalek, and J. M. Fernandez. Reverse engineering of the giant muscle protein titin. *Nature*, 418(6901):998–1002, Aug 2002.

(75) J. Alegre-Cebollada, P. Kosuri, D. Giganti, E. Eckels, J. A. Rivas-Pardo, N. Hamdani, C. M. Warren, R. J. Solaro, W. A. Linke, and J. M. Fernandez. S-glutathionylation of cryptic cysteines enhances titin elasticity by blocking protein folding. *Cell*, 156(6):1235–1246, Mar 2014.

(76) F. Oesterhelt, D. Oesterhelt, M. Pfeiffer, A. Engel, H. E. Gaub, and D. J. Muller. Unfolding pathways of individual bacteriorhodopsins. *Science*, 288(5463):143–146, Apr 2000.

(77) M. Zocher, J. J. Fung, B. K. Kobilka, and D. J. Muller. Ligand-specific interactions modulate kinetic, energetic, and mechanical properties of the human I2 adrenergic receptor. *Structure*, 20(8):1391–1402, Aug 2012.

(78) J. Zhang, G. Wu, C. Song, Y. Li, H. Qiao, P. Zhu, P. Hinterdorfer, B. Zhang, and J. Tang. Single molecular recognition force spectroscopy study of a luteinizing hormone-releasing hormone analogue as a carcinoma target drug. *J Phys Chem B*, 116(45):13331–13337, Nov 2012.

(79) K. Svoboda, C. F. Schmidt, B. J. Schnapp, and S. M. Block. Direct observation of kinesin stepping by optical trapping interferometry. *Nature*, 365(6448):721–727, Oct 1993.

(80) R. Mallik, B. C. Carter, S. A. Lex, S. J. King, and S. P. Gross. Cytoplasmic dynein functions as a gear in response to load. *Nature*, 427(6975):649–652, Feb 2004.

(81) E.A. Abbondanzieri, W.J. Greenleaf, J.W. haevitz, S R. Landick, and S.M. Block. Direct observation of base-pair stepping by rna polymerase. *Nature*, 438:460–465, 2005.

(82) D. S. Johnson, L. Bai, B. Y. Smith, S. S. Patel, and M. D. Wang. Single-molecule studies reveal dynamics of DNA unwinding by the ring-shaped T7 helicase. *Cell*, 129(7):1299–1309, Jun 2007.

(83) K. C. Neuman, E. H. Chadd, G. F. Liou, K. Bergman, and S. M. Block. Characterization of photodamage to *Escherichia coli* in optical traps. *Biophys. J.*, 77(5):2856–2863, Nov 1999.

(84) Y. Jun, S. K. Tripathy, B. R. Narayanareddy, M. K. Mattson-Hoss, and S. P. Gross. Calibration of optical tweezers for *in vivo* force measurements: how do different approaches compare? *Biophys. J.*, 107(6):1474–1484, Sep 2014.

(85) E. L. Holzbaur and Y. E. Goldman. Coordination of molecular motors: from *in vitro* assays to intracellular dynamics. *Curr. Opin. Cell Biol.*, 22(1):4–13, Feb 2010.

(86) G. Bhabha, G. T. Johnson, C. M. Schroeder, and R. D. Vale. How Dynein Moves Along Microtubules. *Trends Biochem. Sci.*, 41(1):94–105, Jan 2016.

(87) X. Nan, P. A. Sims, and X. S. Xie. Organelle tracking in a living cell with microsecond time resolution and nanometer spatial precision. *Chem. Phys. Chem*, 9(5):707–712, Apr 2008.

(88) P. A. Sims and X. S. Xie. Probing dynein and kinesin stepping with mechanical manipulation in a living cell. *Chem. Phys. Chem.*, 10(9-10):1511–1516, Jul 2009.

(89) A. G. Hendricks, E. L. Holzbaur, and Y. E. Goldman. Force measurements on cargoes in living cells reveal collective dynamics of microtubule motors. *Proc. Natl. Acad. Sci. U.S.A.*, 109(45):18447–18452, Nov 2012.

(90) B. H. Blehm, T. A. Schroer, K. M. Trybus, Y. R. Chemla, and P. R. Selvin. In vivo optical trapping indicates kinesin's stall force is reduced by dynein during intracellular transport. *Proc. Natl. Acad. Sci. U.S.A.*, 110(9):3381–3386, Feb 2013.

(91) A. H. de Vries, B. E. Krenn, R. van Driel, and J. S. Kanger. Micro magnetic tweezers for nanomanipulation inside live cells. *Biophys. J.*, 88(3):2137–2144, Mar 2005.

(92) D. Dulin, T. J. Cui, J. Cnossen, M. W. Docter, J. Lipfert, and N. H. Dekker. High Spatiotemporal-Resolution Magnetic Tweezers: Calibration and Applications for DNA Dynamics. *Biophys. J.*, 109(10):2113–2125, Nov 2015.

(93) B. Maier, D. Bensimon, and V. Croquette. Replication by a single DNA polymerase of a stretched single-stranded DNA. *Proc. Natl. Acad. Sci. U.S.A.*, 97(22):12002–12007, Oct 2000.

(94) M. Pandey, S. Syed, I. Donmez, G. Patel, T. Ha, and S. S. Patel. Coordinating DNA replication by means of priming loop and differential synthesis rate. *Nature*, 462(7275):940–943, Dec 2009.

(95) M. Manosas, M. M. Spiering, Z. Zhuang, S. J. Benkovic, and V. Croquette. Coupling DNA unwinding activity with primer synthesis in the bacteriophage T4 primosome. *Nat. Chem. Biol.*, 5(12):904–912, Dec 2009.

(96) M. Manosas, M. M. Spiering, F. Ding, V. Croquette, and S. J. Benkovic. Collaborative coupling between polymerase and helicase for leading-strand synthesis. *Nucleic Acids Res.*, 40(13):6187–6198, Jul 2012.

(97) B. A. Berghuis, D. Dulin, Z. Q. Xu, T. van Laar, B. Cross, R. Janissen, S. Jergic, N. E. Dixon, M. Depken, and N. H. Dekker. Strand separation establishes a sustained lock at the Tus-Ter replication fork barrier. *Nat. Chem. Biol.*, 11(8):579–585, Aug 2015.

(98) M. Orrit and J. Bernard. Single pentacene molecules detected by fluorescence excitation in a p-terphenyl crystal. *Phys. Rev. Lett.*, 65(21):2716–2719, Nov 1990.

(99) T. Funatsu, Y. Harada, M. Tokunaga, K. Saito, and T. Yanagida. Imaging of single fluorescent molecules and individual ATP turnovers by single myosin molecules in aqueous solution. *Nature*, 374(6522):555–559, Apr 1995.

(100) R. Dave, D. S. Terry, J. B. Munro, and S. C. Blanchard. Mitigating unwanted photophysical processes for improved single-molecule fluorescence imaging. *Biophys. J.*, 96(6):2371–2381, Mar 2009.

(101) T. Ha and P. Tinnefeld. Photophysics of fluorescent probes for single-molecule biophysics and super-resolution imaging. *Annu Rev Phys Chem*, 63:595–617, 2012.

(102) J. H. van der Velde, J. Oelerich, J. Huang, J. H. Smit, A. Aminian Jazi, S. Galiani, K. Kolmakov, G. Guoridis, C. Eggeling, A. Herrmann, G. Roelfes, and T. Cordes. A simple and versatile design concept for fluorophore derivatives with intramolecular photostabilization. *Nat Commun*, 7:10144, Jan 2016.

(103) E. J. Peterman, H. Sosa, and W. E. Moerner. Single-molecule fluorescence spectroscopy and microscopy of biomolecular motors. *Annu Rev Phys Chem*, 55:79–96, 2004.

(104) D. Axelrod, T.P. Burghardt, and N.L. Thompson. Total internal reflection fluorescence. *Annu. Rev. Biophys. Bioeng.*, 13:247–68, 1984.

(105) A.M. van Oijen. Single-molecule approaches to characterizing kinetics of biomolecular interactions. *Curr Opin Biotechnol*, 22(1):75–80, 02 2011.

(106) D. Duzdevich, M. D. Warner, S. Ticau, N. A. Ivica, S. P. Bell, and E. C. Greene. The dynamics of eukaryotic replication initiation: origin specificity, licensing, and firing at the single-molecule level. *Mol. Cell*, 58(3):483–494, May 2015.

(107) A. Yildiz, M. Tomishige, R. D. Vale, and P. R. Selvin. Kinesin walks hand-over-hand. *Science*, 303(5658):676–678, Jan 2004.

(108) R. Reyes-Lamothe, D. J. Sherratt, and M. C. Leake. Stoichiometry and architecture of active DNA replication machinery in *Escherichia coli*. *Science*, 328(5977):498–501, Apr 2010.

(109) B. Gibb, L. F. Ye, S. C. Gergoudis, Y. Kwon, H. Niu, P. Sung, and E. C. Greene. Concentration-dependent exchange of replication protein A on single-stranded DNA revealed by single-molecule imaging. *PLoS ONE*, 9(2):e87922, 2014.

(110) A. Robinson, J. P. McDonald, V. E. Caldas, M. Patel, E. A. Wood, C. M. Punter, H. Ghodke, M. M. Cox, R. Woodgate, M. F. Goodman, and A. M. van Oijen. Regulation of Mutagenic DNA Polymerase V Activation in Space and Time. *PLoS Genet.*, 11(8):e1005482, Aug 2015.

(111) C. Kural, H. Kim, S. Syed, G. Goshima, V. I. Gelfand, and P. R. Selvin. Kinesin and dynein move a peroxisome *in vivo*: a tug-of-war or coordinated movement? *Science*, 308(5727):1469–1472, Jun 2005.

(112) A.B Loveland, S. Habuchi, J.C. Walter, and A.M. van Oijen. A general approach to break the concentration barrier in single-molecule imaging. *Nature Methods*, 9(10):987–992, 2012.

(113) H. J. Geertsema, A. C. Schulte, L.M. Spenkelink, W.J. McGrath, S.R. Morrone, J. Sohn, W.F. Mangel, A. Robinson, and A.M. van Oijen. Single-molecule imaging at high fluorophore concentrations by local activation of dye. *Biophys. J.*, 108(4):949–956, Feb 2015.

(114) J.C. Vaughan, S. Jia, and X. Zhuang. Ultrabright photoactivatable fluorophores created by reductive caging. *Nature Methods*, 9(12):1181–1184, 2012.

(115) A. Sharonov and R. M. Hochstrasser. Wide-field subdiffraction imaging by accumulated binding of diffusing probes. *Proc. Natl. Acad. Sci. U.S.A.*, 103(50):18911–18916, Dec 2006.

(116) G. Giannone, E. Hosy, F. Levet, A. Constals, K. Schulze, A. I. Sobolevsky, M. P. Rosconi, E. Gouaux, R. Tampe, D. Choquet, and L. Cognet. Dynamic superresolution imaging of endogenous proteins on living cells at ultra-high density. *Biophys. J.*, 99(4):1303–1310, Aug 2010.

(117) T. Ha. Single-molecule fluorescence resonance energy transfer. *Methods*, 25(1):78–86, Sep 2001.

(118) T. Ha, T. Enderle, D. F. Ogletree, D. S. Chemla, P. R. Selvin, and S. Weiss. Probing the interaction between two single molecules: fluorescence resonance energy transfer between a single donor and a single acceptor. *Proc. Natl. Acad. Sci. U.S.A.*, 93(13):6264–6268, Jun 1996.

(119) T. Mori, R. D. Vale, and M. Tomishige. How kinesin waits between steps. *Nature*, 450(7170):750–754, Nov 2007.

(120) G. B. Erkens, I. Hanelt, J. M. Goudsmits, D. J. Slotboom, and A. M. van Oijen. Unsynchronised subunit motion in single trimeric sodium-coupled aspartate transporters. *Nature*, 502(7469):119–123, Oct 2013.

(121) N. Akyuz, E.R. Georgieva, Z. Zhou, S. Stolzenberg, M.A. Cuendet, G. Khelashvili, R.B. Altman, D.S. Terry, J.H. Freed, H. Weinstein, O. Boudker, and S.C. Blanchard. Transport domain unlocking sets the uptake rate of an aspartate transporter. *Nature*, 518:68–73, 2015.

(122) M. Sustarsic and A. N. Kapanidis. Taking the ruler to the jungle: single-molecule FRET for understanding biomolecular structure and dynamics in live cells. *Curr. Opin. Struct. Biol.*, 34:52–59, Oct 2015.

(123) T. Fessl, F. Adamec, T. Polivka, S. Foldynova-Trantirkova, F. Vacha, and L. Trantirek. Towards characterization of DNA structure under physiological conditions *in vivo* at the single-molecule level using single-pair FRET. *Nucleic Acids Res.*, 40(16):e121, Sep 2012.

(124) R. Crawford, J. P. Torella, L. Aigrain, A. Plochowietz, K. Gryte, S. Uphoff, and A. N. Kapanidis. Long-lived intracellular single-molecule fluorescence using electroporated molecules. *Biophys. J.*, 105(11):2439–2450, Dec 2013.

(125) R. Henderson, J. M. Baldwin, T. A. Ceska, F. Zemlin, E. Beckmann, and K. H. Downing. Model for the structure of bacteriorhodopsin based on high-resolution electron cryo-microscopy. *J. Mol. Biol.*, 213(4):899–929, Jun 1990.

(126) A. Merk, A. Bartesaghi, S. Banerjee, V. Falconieri, P. Rao, M. I. Davis, R. Pragani, M. B. Boxer, L. A. Earl, J. L. Milne, and S. Subramaniam. Breaking Cryo-EM Resolution Barriers to Facilitate Drug Discovery. *Cell*, 165(7):1698–1707, Jun 2016.

(127) Y. Cheng. Single-Particle Cryo-EM at Crystallographic Resolution. *Cell*, 161(3):450–457, Apr 2015.

(128) R. Fernandez-Leiro and S. H. Scheres. Unravelling biological macromolecules with cryo-electron microscopy. *Nature*, 537(7620):339–346, Sep 2016.

(129) C. V. Sindelar and K. H. Downing. An atomic-level mechanism for activation of the kinesin molecular motors. *Proc. Natl. Acad. Sci. U.S.A.*, 107(9):4111–4116, Mar 2010.

(130) Z. Yuan, L. Bai, J. Sun, R. Georgescu, J. Liu, M. E. O'Donnell, and H. Li. Structure of the eukaryotic replicative CMG helicase suggests a pumpjack motion for translocation. *Nat. Struct. Mol. Biol.*, 23(3):217–224, Mar 2016.

(131) R. Fernandez-Leiro, J. Conrad, S. H. Scheres, and M. H. Lamers. cryo-EM structures of the *E. coli* replicative DNA polymerase reveal its dynamic interactions with the DNA sliding clamp, exonuclease and ÏĎ. *eLife*, 4, Oct 2015.

(132) A. Sartori, R. Gatz, F. Beck, A. Rigort, W. Baumeister, and J. M. Plitzko. Correlative microscopy: bridging the gap between fluorescence light microscopy and cryo-electron tomography. *J. Struct. Biol.*, 160(2):135–145, Nov 2007.

(133) S. Kobayashi, M. Iwamoto, and T. Haraguchi. Live correlative light-electron microscopy to observe molecular dynamics in high resolution. *Microscopy (Oxf)*, 65(4):296–308, Aug 2016.

(134) T. Haraguchi, T. Kojidani, T. Koujin, T. Shimi, H. Osakada, C. Mori, A. Yamamoto, and Y. Hiraoka. Live cell imaging and electron microscopy reveal dynamic processes of BAF-directed nuclear envelope assembly. *J. Cell. Sci.*, 121(Pt 15):2540–2554, Aug 2008.

(135) J. Mahamid, S. Pfeffer, M. Schaffer, E. Villa, R. Danev, L. K. Cuellar, F. Forster, A. A. Hyman, J. M. Plitzko, and W. Baumeister. Visualizing the molecular sociology at the HeLa cell nuclear periphery. *Science*, 351(6276):969–972, Feb 2016.

(136) N. A. Tanner, G. Tolun, J. J. Loparo, S. Jergic, J. D. Griffith, N. E. Dixon, and A. M. van Oijen. *E. coli* DNA replication in the absence of free beta clamps. *EMBO J.*, 30(9):1830–1840, May 2011.

(137) J. J. Loparo, A. W. Kulczyk, C. C. Richardson, and A. M. van Oijen. Simultaneous single-molecule measurements of phage T7 replisome composition and function reveal the mechanism of polymerase exchange. *Proc. Natl. Acad. Sci. U.S.A.*, 108(9):3584–3589, Mar 2011.

(138) T. Y. Chen, A. G. Santiago, W. Jung, P. Krzemiski, F. Yang, D. J. Martell, J. D. Helmann, and P. Chen. Concentration- and chromosome-organization-dependent regulator unbinding from DNA for transcription regulation in living cells. *Nature Commun.*, 6:7445, Jul 2015.

(139) Geertsema H.J. and A.M. van Oijen. A single-molecule view of DNA replication: the dynamic nature of multi-protein complexes revealed. *Current Opinion in Structural Biology*, 23(5):788–793, 2013.

(140) C. E. Sing, M. Olvera de la Cruz, and J. F. Marko. Multiple-binding-site mechanism explains concentration-dependent unbinding rates of DNA-binding proteins. *Nucleic Acids Res.*, 42(6):3783–3791, Apr 2014.

(141) C. Aberg, K.E. Duderstadt, and A.M. van Oijen. Stability versus exchange: a paradox in DNA replication. *Nucleic Acids Res.*, 44:4846–4854, 2016.

(142) M. D. Sutton. Coordinating DNA polymerase traffic during high and low fidelity synthesis. *Biochim. Biophys. Acta*, 1804(5):1167–1179, May 2010.

(143) P. Holzmeister, G.P. Acuna, D. Grohmann, and P. Tinnefeld. Single-molecule approaches to characterizing kinetics of biomolecular interactions. *Chem Soc Rev.*, 43(4):1014–1028, 02 2014.

(144) E. Boukobza, A. Sonnenfeld, and G. Haran. Immobilization in surface-tethered lipid vesicles as a new tool for single biomolecule spectroscopy. *The Journal of Physical Chemistry B*, 105(48):12165–12170, 2001.

(145) M. J. Levene, J. Korlach, S. W. Turner, M. Foquet, H. G. Craighead, and W. W. Webb. Zero-mode waveguides for single-molecule analysis at high concentrations. *Science*, 299(5607):682–686, 2003.

(146) J.J. Benitez, A.M. Keller, P. Ochieng, L.A. Yatsunyk, D.L. Huffman, A.C. Rosenzweig, and P. Chen. Probing transient copper chaperone–wilson disease protein interactions at the single-molecule level with nanovesicle trapping. *Journal of the American Chemical Society*, 130(8):2446–2447, 2008.

(147) S. R. Morrone, T. Wang, L. M. Constantoulakis, R. M. Hooy, M. J. Delannoy, and J. Sohn. Cooperative assembly of IFI 16 filaments on dsDNA provides insights into host defense strategy. *Proc. Natl. Acad. Sci. U.S.A.*, 111(1):62–71, Jan 2014.

(148) Mangel WF, McGrath WJ, Toledo DL, and Anderson CW. Viral DNA and a viral peptide can act as cofactors of adenovirus virion proteinase activity. *Nature*, 111(6409):274–275, Jan 1993.

(149) W.F. Mangel, D.L. Toledo, M.T. Brown, J.H. Martin, and W.J. McGrath. Characterization of three components of human adenovirus proteinase activity *in vitro*. *Journal of Biological Chemistry*, 271(1):536–543, 1996.

(150) S.C. Gill and P.H. von Hippel. Calculation of protein extinction coefficients from amino acid sequence data. *Analytical Biochemistry*, 182(2):319–326, 1989.

(151) PW Riddles, RL Blakeley, and B. Zerner. Reassessment of ellman's reagent. *Methods Enzymol*, 91:49–60, 1983.

(152) W.J. McGrath, M.L. Baniecki, E. Peters, D.T. Green, and W.F. Mangel. Roles of two conserved cysteine residues in the activation of human adenovirus proteinase. *Biochemistry*, 40(48):14468–14474, 2001.

(153) W.J. McGrath, K.S. Aherne, and W.F. Mangel. In the virion, the 11-amino-acid peptide cofactor pvic is covalently linked to the adenovirus proteinase. *Virology*, 296(2):234–240, 2002.

(154) E. Gasteiger, A. Gattiker, C. Hoogland, I. Ivanyi, R.D. Appel, and A. Bairoch. Expasy: The proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Research*, 31(13):3784–3788, 2003.

(155) N.A. Tanner and A.M. van Oijen. Chapter Eleven - Visualizing DNA replication at the single-molecule level. In Nils G. Walter, editor, *Single Molecule Tools, Part B:Super-Resolution, Particle Tracking, Multiparameter, and Force Based Methods*, volume 475 of *Methods in Enzymology*, pages 259 – 278. Academic Press, 2010.

(156) A.M. Van Oijen, P.C. Blainey, D.J. Crampton, C.C. Richardson, T. Ellenberger, and X.S. Xie. Single-molecule kinetics of lamda exonuclease reveal base dependence and dynamic disorder. *Science*, 301(5637):1235–1238, 2003.

(157) I. Rasnik, S.A. McKinney, and T. Ha. Nonblinking and long-lasting single-molecule fluorescence imaging. *Nature Methods*, 3(11):891–893, 2006.

(158) M. Heilemann, E. Margeat, R. Kasper, M. Sauer, and P. Tinnefeld. Carbocyanine dyes as efficient reversible single-molecule optical switch. *Journal of the American Chemical Society*, 127(11):3801–3806, 2005.

(159) M. Bates, T.R. Blosser, and X. Zhuang. Short-range spectroscopic ruler based on a single-molecule optical switch. *Phys. Rev. Lett.*, 94:108101, Mar 2005.

(160) J.C. Vaughan, G.T. Dempsey, E. Sun, and X. Zhuang. Phosphine quenching of cyanine dyes as a versatile tool for fluorescence microscopy. *Journal of the American Chemical Society*, 135(4):737–738, 2013.

(161) G.T. Dempsey, M. Bates, W.E. Kowtoniuk, D.R. Liu, R.Y. Tsien, and X. Zhuang. Photoswitching mechanism of cyanine dyes. *Journal of the American Chemical Society*, 131(51):18192–18193, 2009.

(162) N.R. Conley, J.S. Biteen, and W.E. Moerner. Cy3-cy5 covalent heterodimers for single-molecule photoswitching. *Journal of Physical Chemistry B*, 112(38):11878–11880, 2008.

(163) K.M. Monroe, Z. Yang, J.R. Johnson, X. Geng, G. Doitsh, N.J. Krogan, and W.C. Greene. Ifi16 DNA sensor is required for death of lymphoid cd4 t cells abortively infected with hiv. *Science*, 343(6169):428–432, 2014.

(164) A. Tafvizi, F. Huang, A.R. Fersht, L.A. Mirny, and A.M. van Oijen. A single-molecule characterization of p53 search on DNA. *Proceedings of the National Academy of Sciences of the United States of America*, 108(2):563–568, 2011.

(165) P.C. Blainey, V. Graziano, A.J. Perez-Berna, W.J. McGrath, S.J. Flint, C.S. Martin, X.S. Xie, and W.F. Mangel. Regulation of a viral proteinase by a peptide and DNA in one-dimensional space iv: Viral proteinase slides along DNA to locate and process its substrates. *Journal of Biological Chemistry*, 288(3):2092–2102, 2013.

(166) M.A. Tycon, C.F. Dial, K. Faison, W. Melvin, and C.J. Fecko. Quantification of dye-mediated photodamage during single-molecule DNA imaging. *Analytical Biochemistry*, 426(1):13–21, 2012.

(167) I. Braslavsky, B. Hebert, E. Kartalov, and S.R. Quake. Sequence information can be obtained from single DNA molecules. *Proceedings of the National Academy of Sciences of the United States of America*, 100(7):3960–3964, 2003.

(168) S. Uphoff, S.J. Holden, L. Le Reste, J. Periz, S. Van De Linde, M. Heilemann, and A.N. Kapanidis. Monitoring multiple distances within a single molecule using switchable FRET. *Nature Methods*, 7(10):831–836, 2010.

(169) N. Grimaldi, F. Andrade, N. Segovia, L. Ferrer-Tasies, S. Sala, J. Veciana, and N. Ventosa. Lipid-based nanovesicles for nanomedicine. *Chem Soc Rev*, 45(23):6520–6545, Nov 2016.

(170) T. M. Allen and P. R. Cullis. Liposomal drug delivery systems: from concept to clinical applications. *Adv. Drug Deliv. Rev.*, 65(1):36–48, Jan 2013.

(171) J. W. Nichols and Y. H. Bae. EPR: Evidence and fallacy. *J Control Release*, 190:451–464, Sep 2014.

(172) L.E. Gerlowski and R.K. Jain. Microvascular permeability of normal and neoplastic tissues. *Microvasc. Res.*, 31(3):288–305, May 1986.

(173) Y. Barenholz. Doxil–the first FDA-approved nano-drug: lessons learned. *J Control Release*, 160(2):117–134, Jun 2012.

(174) T. Ishida, D. L. Iden, and T. M. Allen. A combinatorial approach to producing sterically stabilized (Stealth) immunoliposomal drugs. *FEBS Lett.*, 460(1):129–133, Oct 1999.

(175) B. S. Pattni, V. V. Chupin, and V. P. Torchilin. New Developments in Liposomal Drug Delivery. *Chem. Rev.*, 115(19):10938–10966, Oct 2015.

(176) S. Wilhelm, A. J. Tavares, Q. Dai, S. Ohta, J. Audet, H. F. Dvorak, and W. C. W. Chan. Analysis of nanoparticle delivery to tumours. *Nature Reviews Materials*, 1:16014, May 2016.

(177) K. Laginha, D. Mumbengegwi, and T. Allen. Liposomes targeted via two different antibodies: assay, B-cell binding and cytotoxicity. *Biochim. Biophys. Acta*, 1711(1):25–32, Jun 2005.

(178) E. Doolittle, P. M. Peiris, G. Doron, A. Goldberg, S. Tucci, S. Rao, S. Shah, M. Sylvestre, P. Govender, O. Turan, Z. Lee, W. P. Schiemann, and E. Karathanasis. Spatiotemporal Targeting of a Dual-Ligand Nanoparticle to Cancer Metastasis. *ACS Nano*, 9(8):8012–8021, Aug 2015.

(179) M. Estanqueiro, M.H. Amaral, J. Conceicao, and J.M. Sousa Lobo. Evolution of liposomal carriers intendet to anticancer drug delivery: an overview. *Int. J. Curr. Pharm. Res.*, 6:3–10, Oct 2014.

(180) J. Shi, P. W. Kantoff, R. Wooster, and O. C. Farokhzad. Cancer nanomedicine: progress, challenges and opportunities. *Nat. Rev. Cancer*, 17(1):20–37, 01 2017.

(181) A. N. Lukyanov, T. A. Elbayoumi, A. R. Chakilam, and V. P. Torchilin. Tumor-targeted liposomes: doxorubicin-loaded long-circulating liposomes modified with anti-cancer antibody. *J Control Release*, 100(1):135–144, Nov 2004.

(182) J. W. Park, D. B. Kirpotin, K. Hong, R. Shalaby, Y. Shao, U. B. Nielsen, J. D. Marks, D. Papahadjopoulos, and C. C. Benz. Tumor targeting using anti-her2 immunoliposomes. *J Control Release*, 74(1-3):95–113, Jul 2001.

(183) T. J. Anchordoquy, Y. Barenholz, D. Boraschi, M. Chorny, P. Decuzzi, M. A. Dobrovolskaia, Z. S. Farhangrazi, D. Farrell, A. Gabizon, H. Ghandehari, B. Godin, N. M. La-Beck, J. Ljubimova, S. M. Moghimi, L. Pagliaro, J. H. Park, D. Peer, E. Ruoslahti, N. J. Serkova, and D. Simberg. Mechanisms and Barriers in Cancer Nanomedicine: Addressing Challenges, Looking for Solutions. *ACS Nano*, 11(1):12–18, Jan 2017.

(184) M. E. Klegerman, A. J. Hamilton, S. L. Huang, S. D. Tiukinhoy, A. A. Khan, R. C. MacDonald, and D. D. McPherson. Quantitative immunoblot assay for assessment of liposomal antibody conjugation efficiency. *Anal. Biochem.*, 300(1):46–52, Jan 2002.

(185) K. Mack, R. RÃijger, S. Fellermeier, O. Seifert, and R.E. Kontermann. Dual targeting of tumor cells with bispecific single-chain fv-immunoliposomes. *Antibodies*, 1(2):199–214, 2012.

(186) E. Monachino, L. M. Spenkelink, and A. M. van Oijen. Watching cellular machinery in action, one molecule at a time. *J. Cell Biol.*, 216(1):41–51, Jan 2017.

(187) M. Marchetti, A. Malinowska, I. Heller, and G. J. L. Wuite. How to switch the motor on: RNA polymerase initiation steps at the single-molecule level. *Protein Sci.*, 26(7):1303–1313, Jul 2017.

(188) V. Aggarwal and T. Ha. Single-molecule fluorescence microscopy of native macromolecular complexes. *Curr. Opin. Struct. Biol.*, 41:225–232, Dec 2016.

(189) L. P. Watkins and H. Yang. Detection of intensity change points in time-resolved single-molecule measurements. *J Phys Chem B*, 109(1):617–628, Jan 2005.

(190) T. M. Allen, P. Sapra, and E. Moase. Use of the post-insertion method for the formation of ligand-coupled liposomes. *Cell. Mol. Biol. Lett.*, 7(3):889–894, 2002.

(191) A. P. Serro, A. Carapeto, G. Paiva, J. P. S. Farinha, R. ColaÃğo, and B. Saramago. Formation of an intact liposome layer adsorbed on oxidized gold confirmed by three complementary techniques: QCM-D, AFM and confocal fluorescence microscopy. *Surface and Interface Analysis*, 44(4):426–433, 2012.

(192) D. L. Iden and T. M. Allen. *in vitro* and *in vivo* comparison of immunoliposomes made by conventional coupling techniques with those made by a new post-insertion approach. *Biochim. Biophys. Acta*, 1513(2):207–216, Aug 2001.

(193) J. J. Otterstrom, B. Brandenburg, M. H. Koldijk, J. Juraszek, C. Tang, S. Mashaghi, T. Kwaks, J. Goudsmit, R. Vogels, R. H. Friesen, and A. M. van Oijen. Relating influenza virus membrane fusion kinetics to stoichiometry of neutralizing antibodies at the single-particle level. *Proc. Natl. Acad. Sci. U.S.A.*, 111(48):E5143–5148, Dec 2014.

(194) B. J. Cochran, L. P. Gunawardhana, K. L. Vine, J. A. Lee, S. Lobov, and M. Ranson. The CD-loop of PAI-2 (SERPINB2) is redundant in the targeting, inhibition and clearance of cell surface uPA activity. *BMC Biotechnol.*, 9:43, May 2009.

(195) P. S. Uster, T. M. Allen, B. E. Daniel, C. J. Mendez, M. S. Newman, and G. Z. Zhu. Insertion of poly(ethylene glycol) derivatized phospholipid into pre-formed liposomes results in prolonged *in vivo* circulation time. *FEBS Lett.*, 386(2-3):243–246, May 1996.

(196) J. N. Moreira, T. Ishida, R. Gaspar, and T. M. Allen. Use of the post-insertion technique to insert peptide ligands into pre-formed stealth liposomes with retention of binding activity and cytotoxicity. *Pharm. Res.*, 19(3):265–269, Mar 2002.

(197) K.E. Duderstadt, R. Reyes-Lamothe, A.M. van Oijen, and D.J. Sherratt. *Replication-fork dynamics.* Cold Spring Harb. Perspect. Biol, 6 edition, 2014.

(198) C. A. Wu, E. L. Zechner, A. J. Hughes, M. A. Franden, C. S. McHenry, and K. J. Marians. Coordinated leading- and lagging-strand synthesis at the *Escherichia coli* DNA replication fork. IV. Reconstitution of an asymmetric, dimeric DNA polymerase III holoenzyme. *J. Biol. Chem.*, 267(6):4064–4073, Feb 1992.

(199) R. Onrust, J. Finkelstein, J. Turner, V. Naktinis, and M. O'Donnell. Assembly of a chromosomal replication machine: two DNA polymerases, a clamp loader, and sliding clamps in one holoenzyme particle. III. Interface between two polymerases and the clamp loader. *J. Biol. Chem.*, 270(22):13366–13377, Jun 1995.

(200) A. Blinkova, C. Hervas, P. T. Stukenberg, R. Onrust, M. E. O'Donnell, and J. R. Walker. The *Escherichia coli* DNA polymerase III holoenzyme contains both products of the *dnaX* gene, tau and gamma, but only tau is essential. *J. Bacteriol.*, 175(18):6018–6027, Sep 1993.

(201) P. R. Dohrmann, C. M. Manhart, C. D. Downey, and C. S. McHenry. The rate of polymerase release upon filling the gap between Okazaki fragments is inadequate to support cycling during lagging strand synthesis. *J. Mol. Biol.*, 414(1):15–27, Nov 2011.

(202) N. A. Tanner, S. M. Hamdan, S. Jergic, K. V. Loscha, P. M. Schaeffer, N. E. Dixon, and A. M. van Oijen. Single-molecule studies of fork dynamics in

Escherichia coli DNA replication. *Nat. Struct. Mol. Biol.*, 15(2):170–176, Feb 2008.

(203) N. Y. Yao, R. E. Georgescu, J. Finkelstein, and M. E. O'Donnell. Single-molecule analysis reveals that the lagging strand increases replisome processivity but slows replication fork progression. *Proc. Natl. Acad. Sci. U.S.A.*, 106(32):13236–13241, Aug 2009.

(204) R.E. Georgescu, I. Kurth, and M.E. O'Donnell. Single-molecule studies reveal the function of a third polymerase in the replisome. *Nat. Struct. Mol. Biol.*, 19(1):113–116, Dec 2011.

(205) S. Jergic, K. Ozawa, N.K. Williams, X.C. Su, D.D. Scott, S.M. Hamdan, J.A. Crowther, G. Otting, and N.E. Dixon. The unstructured C-terminus of the tau subunit of *Escherichia coli* DNA polymerase III holoenzyme is the site of interaction with the alpha subunit. *Nucleic Acids Res.*, 35(9):2813–2824, 2007.

(206) M. Mok and K. J. Marians. The *Escherichia coli* preprimosome and DnaB helicase can form replication forks that move at the same rate. *J. Biol. Chem.*, 262(34):16644–16654, Dec 1987.

(207) F.P. Leu, R.E. Georgescu, and M.E. O'Donnell. Mechanism of the *E. coli* processivity switch during lagging-strand synthesis. *Mol. Cell*, 11:315–327, 2003.

(208) P. McInerney, A. Johnson, F. Katz, and M. O'Donnell. Characterization of a triple DNA polymerase replisome. *Mol. Cell*, 27(4):527–538, Aug 2007.

(209) J. S. Graham, R. C. Johnson, and J. F. Marko. Concentration-dependent exchange accelerates turnover of proteins bound to double-stranded DNA. *Nucleic Acids Res.*, 39(6):2249–2259, Mar 2011.

(210) A. J. Oakley, P. Prosselkov, G. Wijffels, J. L. Beck, M. C. Wilce, and N. E. Dixon. Flexibility revealed by the 1.85 Åcrystal structure of the beta sliding-clamp subunit of *Escherichia coli* DNA polymerase III. *Acta Crystallogr. D Biol. Crystallogr.*, 59(Pt 7):1192–1199, Jul 2003.

(211) C. E. Mason, S. Jergic, A. T. Lo, Y. Wang, N. E. Dixon, and J. L. Beck. *Escherichia coli* single-stranded DNA-binding protein: nanoESI-MS studies of salt-modulated subunit exchange and DNA binding transactions. *J. Am. Soc. Mass Spectrom.*, 24(2):274–285, Feb 2013.

(212) S. Jergic, N. P. Horan, M. M. Elshenawy, C. E. Mason, T. Urathamakul, K. Ozawa, A. Robinson, J. M. Goudsmits, Y. Wang, X. Pan, J. L. Beck, A. M. van Oijen, T. Huber, S. M. Hamdan, and N. E. Dixon. A direct proofreader-clamp interaction stabilizes the Pol III replicase in the polymerization mode. *EMBO J.*, 32(9):1322–1333, May 2013.

(213) N. P. Stamford, P. E. Lilley, and N. E. Dixon. Enriched sources of *Escherichia coli* replication proteins. The DnaG primase is a zinc metalloprotein. *Biochim. Biophys. Acta*, 1132(1):17–25, Aug 1992.

(214) Y. Wang. *Single-stranded DNA-binding protein and its role in Okazaki fragment maturation*. University of Wollongong, 2015.

(215) G. Wijffels, B. P. Dalrymple, P. Prosselkov, K. Kongsuwan, V. C. Epa, P. E. Lilley, S. Jergic, J. Buchardt, S. E. Brown, P. F. Alewood, P. A. Jennings, and N. E. Dixon. Inhibition of protein interactions with the beta 2 sliding clamp of *Escherichia coli* DNA polymerase III by peptides from beta 2-binding proteins. *Biochemistry*, 43(19):5661–5671, May 2004.

(216) N. K. Williams, P. Prosselkov, E. Liepinsh, I. Line, A. Sharipo, D. R. Littler, P. M. Curmi, G. Otting, and N. E. Dixon. *In vivo* protein cyclization promoted by a circularly permuted *Synechocystis* sp. PCC6803 DnaB mini-intein. *J. Biol. Chem.*, 277(10):7790–7798, Mar 2002.

(217) K. Ozawa, N. P. Horan, A. Robinson, H. Yagi, F. R. Hill, S. Jergic, Z. Q. Xu, K. V. Loscha, N. Li, M. Tehei, A. J. Oakley, G. Otting, T. Huber, and N. E. Dixon. Proofreading exonuclease on a tether: the complex between the *E. coli* DNA polymerase III subunits alpha, epsilon, theta and beta reveals a highly flexible arrangement of the proofreading domain. *Nucleic Acids Res.*, 41(10):5354–5367, May 2013.

(218) L.C. Huang, E.A. Wood, and M.M. Cox. Convenient and reversible site-specific targeting of exogenous DNA into a bacterial chromosome by use of the flp recombinase: the flirt system. *J. Bacteriol*, 179(19):6076–6083, 10 1997.

(219) J. W. Chase and K. R. Williams. Single-stranded DNA binding proteins required for DNA replication. *Annu. Rev. Biochem.*, 55:103–136, 1986.

(220) P. Markiewicz, C. Malone, J. W. Chase, and L. B. Rothman-Denes. *Escherichia coli* single-stranded DNA-binding protein is a supercoiled template-dependent transcriptional activator of N4 virion RNA polymerase. *Genes Dev.*, 6(10):2010–2019, Oct 1992.

(221) A. Sancar, K. R. Williams, J. W. Chase, and W. D. Rupp. Sequences of the *ssb* gene and protein. *Proc. Natl. Acad. Sci. U.S.A.*, 78(7):4274–4278, Jul 1981.

(222) E. Antony, E. Weiland, Q. Yuan, C. M. Manhart, B. Nguyen, A. G. Kozlov, C. S. McHenry, and T. M. Lohman. Multiple C-terminal tails within a single *E. coli* SSB homotetramer coordinate DNA replication and repair. *J. Mol. Biol.*, 425(23):4802–4819, Nov 2013.

(223) W. Bujalowski, L. B. Overman, and T. M. Lohman. Binding mode transitions of *Escherichia coli* single strand binding protein-single-stranded DNA

complexes. Cation, anion, pH, and binding density effects. *J. Biol. Chem.*, 263(10):4629–4640, Apr 1988.

(224) T. M. Lohman, W. Bujalowski, L. B. Overman, and T. F. Wei. Interactions of the *E. coli* single strand binding (SSB) protein with ss nucleic acids. Binding mode transitions and equilibrium binding studies. *Biochem. Pharmacol.*, 37(9):1781–1782, May 1988.

(225) M. J. Bessman, I. R. Lehman, J. Adler, S. B. Zimmerman, E. S. Simms, and A. Kornberg. Enzymatic synthesis of deoxyribonucleic acid. III. the incorporation of pyrimidine and purine analogues into deoxyribonucleic acid. *Proc. Natl. Acad. Sci. U.S.A.*, 44(7):633–640, Jul 1958.

(226) B. P. Glover and C. S. McHenry. The chi psi subunits of DNA polymerase III holoenzyme bind to single-stranded DNA-binding protein (SSB) and facilitate replication of an SSB-coated template. *J. Biol. Chem.*, 273(36):23476–23484, Sep 1998.

(227) A. H. Marceau, S. Bahng, S. C. Massoni, N. P. George, S. J. Sandler, K. J. Marians, and J. L. Keck. Structure of the SSB-DNA polymerase III interface and its role in DNA replication. *EMBO J.*, 30(20):4236–4247, Aug 2011.

(228) S. Kunzelmann, C. Morris, A. P. Chavda, J. F. Eccleston, and M. R. Webb. Mechanism of interaction between single-stranded DNA binding protein and DNA. *Biochemistry*, 49(5):843–852, Feb 2010.

(229) K. R. Williams, J. B. Murphy, and J. W. Chase. Characterization of the structural and functional defect in the *Escherichia coli* single-stranded DNA binding protein encoded by the *ssb-1* mutant gene. Expression of the *ssb-1* gene under lambda pL regulation. *J. Biol. Chem.*, 259(19):11804–11811, Oct 1984.

(230) E. V. Bobst, A. M. Bobst, F. W. Perrino, R. R. Meyer, and D. C. Rein. Variability in the nucleic acid binding site size and the amount of single-stranded DNA-binding protein in *Escherichia coli*. *FEBS Lett.*, 181(1):133–137, Feb 1985.

(231) A. Yuzhakov, Z. Kelman, and M. O'Donnell. Trading places on DNA–a three-point switch underlies primer handoff from primase to the replicative DNA polymerase. *Cell*, 96(1):153–163, Jan 1999.

(232) S. Fossum, E. Crooke, and K. Skarstad. Organization of sister origins and replisomes during multifork DNA replication in *Escherichia coli*. *EMBO J.*, 26(21):4514–4522, Oct 2007.

(233) H Ghodke, H.N. Ho, and van Oijen A.M. Single-molecule live-cell imaging of bacterial DNA repair and damage tolerance. *Biochem Soc Trans.*, Dec 2017.

(234) S. Shashkova and M. C. Leake. Single-molecule fluorescence microscopy review: shedding new light on old problems. *Biosci. Rep.*, 37(4), Aug 2017.

(235) M. J. Scherr, B. Safaric, and K. E. Duderstadt. Noise in the Machine: Alternative Pathway Sampling is the Rule During DNA Replication. *Bioessays*, 40(2), Feb 2018.

(236) J. S. Lewis, L. M. Spenkelink, S. Jergic, E. A. Wood, E. Monachino, N. P. Horan, K. E. Duderstadt, M. M. Cox, A. Robinson, N. E. Dixon, and A. M. van Oijen. Single-molecule visualization of fast polymerase turnover in the bacterial replisome. *Elife*, 6, Apr 2017.

(237) M. E. O'Donnell and A. Kornberg. Complete replication of templates by *Escherichia coli* DNA polymerase III holoenzyme. *J. Biol. Chem.*, 260(23):12884–12889, Oct 1985.

(238) K. Arai and A. Kornberg. A general priming system employing only DnaB protein and primase for DNA replication. *Proc. Natl. Acad. Sci. U.S.A.*, 76(9):4308–4312, Sep 1979.

(239) T. R. Beattie, N. Kapadia, E. Nicolas, S. Uphoff, A. J. Wollman, M. C. Leake, and R. Reyes-Lamothe. Frequent exchange of the DNA polymerase during bacterial chromosome replication. *Elife*, 6, Mar 2017.

(240) T. Ogawa and T. Okazaki. Discontinuous DNA replication. *Annu. Rev. Biochem.*, 49:421–457, 1980.

(241) N. J. Delalez, G. H. Wadhams, G. Rosser, Q. Xue, M. T. Brown, I. M. Dobbie, R. M. Berry, M. C. Leake, and J. P. Armitage. Signal-dependent turnover of the bacterial flagellar switch protein FliM. *Proc. Natl. Acad. Sci. U.S.A.*, 107(25):11347–11351, Jun 2010.

(242) T. Paramanathan, D. Reeves, L. J. Friedman, J. Kondev, and J. Gelles. A general mechanism for competitor-induced dissociation of molecular complexes. *Nat Commun*, 5:5207, Oct 2014.

(243) C. J. Ma, B. Gibb, Y. Kwon, P. Sung, and E. C. Greene. Protein dynamics of human RPA and RAD51 on ssDNA during assembly and disassembly of the RAD51 filament. *Nucleic Acids Res.*, 45(2):749–761, Jan 2017.

(244) T. Y. Chen, Y. S. Cheng, P. S. Huang, and P. Chen. Facilitated Unbinding via Multivalency-Enabled Ternary Complexes: New Paradigm for Protein-DNA Interactions. *Acc. Chem. Res.*, Jan 2018.

(245) M. Y. Tsai, B. Zhang, W. Zheng, and P. G. Wolynes. Molecular Mechanism of Facilitated Dissociation of Fis Protein from DNA. *J. Am. Chem. Soc.*, Oct 2016.

(246) K. Dahlke and C. E. Sing. Facilitated Dissociation Kinetics of Dimeric Nucleoid-Associated Proteins Follow a Universal Curve. *Biophys. J.*, 112(3):543–551, Feb 2017.

(247) Y. Kim, S. O. Ho, N. R. Gassman, Y. Korlann, E. V. Landorf, F. R. Collart, and S. Weiss. Efficient site-specific labeling of proteins via cysteines. *Bioconjug. Chem.*, 19(3):786–791, Mar 2008.

(248) R. Reyes-Lamothe, C. Possoz, O. Danilova, and D. J. Sherratt. Independent positioning and action of *Escherichia coli* replisomes in live cells. *Cell*, 133(1):90–102, Apr 2008.

(249) M. M. Cox, M. F. Goodman, K. N. Kreuzer, D. J. Sherratt, S. J. Sandler, and K. J. Marians. The importance of repairing stalled replication forks. *Nature*, 404(6773):37–41, Mar 2000.

(250) M. M. Cox. Historical overview: searching for replication help in all of the rec places. *Proc. Natl. Acad. Sci. U.S.A.*, 98(15):8173–8180, Jul 2001.

(251) S. C. Kowalczykowski. Initiation of genetic recombination and recombination-dependent replication. *Trends Biochem. Sci.*, 25(4):156–165, Apr 2000.

(252) A. Kuzminov. Recombinational repair of DNA damage in *Escherichia coli* and bacteriophage lambda. *Microbiol. Mol. Biol. Rev.*, 63(4):751–813, Dec 1999.

(253) A. Kuzminov. DNA replication meets genetic exchange: chromosomal damage and its repair by homologous recombination. *Proc. Natl. Acad. Sci. U.S.A.*, 98(15):8461–8468, Jul 2001.

(254) B. Michel. Replication fork arrest and DNA recombination. *Trends Biochem. Sci.*, 25(4):173–178, Apr 2000.

(255) B. Michel, H. Boubakri, Z. Baharoglu, M. LeMasson, and R. Lestini. Recombination proteins and rescue of arrested replication forks. *DNA Repair (Amst.)*, 6(7):967–980, Jul 2007.

(256) R. C. Heller and K. J. Marians. Replication fork reactivation downstream of a blocked nascent leading strand. *Nature*, 439(7076):557–562, Feb 2006.

(257) H.L. Klein and K.N. Kreuzer. Replication, recombination, and repair: Going for the gold. *Molecular Cell*, 9(3):471 – 480, 2002.

(258) M. Lopes, C. Cotta-Ramusino, A. Pellicioli, G. Liberi, P. Plevani, M. Muzi-Falconi, C. S. Newlon, and M. Foiani. The DNA replication checkpoint response stabilizes stalled replication forks. *Nature*, 412(6846):557–561, Aug 2001.

(259) H. Merrikh, Y. Zhang, A. D. Grossman, and J. D. Wang. Replication-transcription conflicts in bacteria. *Nat. Rev. Microbiol.*, 10(7):449–458, Jun 2012.

(260) M.M. Cox. The nonmutagenic repair of broken replication forks via recombination. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis*, 510(1-2):107 – 120, 2002.

(261) A. Kuzminov. Collapse and repair of replication forks in *Escherichia coli*. *Mol. Microbiol.*, 16(3):373–384, May 1995.

(262) J. D. McCool, E. Long, J. F. Petrosino, H. A. Sandler, S. M. Rosenberg, and S. J. Sandler. Measurement of SOS expression in individual *Escherichia coli* K-12 cells using fluorescence microscopy. *Mol. Microbiol.*, 53(5):1343–1357, Sep 2004.

(263) B. Michel, M. J. Flores, E. Viguera, G. Grompone, M. Seigneur, and V. Bidnenko. Rescue of arrested replication forks by homologous recombination. *Proc. Natl. Acad. Sci. U.S.A.*, 98(15):8181–8188, Jul 2001.

(264) B. Michel, G. Grompone, M. J. Flores, and V. Bidnenko. Multiple pathways process stalled replication forks. *Proc. Natl. Acad. Sci. U.S.A.*, 101(35):12783–12788, Aug 2004.

(265) A. H. Syeda, M. Hawkins, and P. McGlynn. Recombination and replication. *Cold Spring Harb Perspect Biol*, 6(11):a016550, Oct 2014.

(266) S. M. Mangiameli, C. N. Merrikh, P. A. Wiggins, and H. Merrikh. Transcription leads to pervasive replisome instability in bacteria. *Elife*, 6, Jan 2017.

(267) J. Courcelle, B. M. Wendel, D. D. Livingstone, and C. T. Courcelle. RecBCD is required to complete chromosomal replication: Implications for double-strand break frequencies and repair mechanisms. *DNA Repair (Amst.)*, 32:86–95, Aug 2015.

(268) E. V. Mirkin and S. M. Mirkin. Replication fork stalling at natural impediments. *Microbiol. Mol. Biol. Rev.*, 71(1):13–35, Mar 2007.

(269) A. Aguilera and T. Garcia-Muse. Causes of genome instability. *Annu. Rev. Genet.*, 47:1–32, 2013.

(270) M. F. Goodman. Error-prone repair DNA polymerases in prokaryotes and eukaryotes. *Annu. Rev. Biochem.*, 71:17–50, 2002.

(271) A. J. Gruber, A. L. Erdem, G. Sabat, K. Karata, M. M. Jaszczur, D. D. Vo, T. M. Olsen, R. Woodgate, M. F. Goodman, and M. M. Cox. A RecA protein surface required for activation of DNA polymerase V. *PLoS Genet.*, 11(3):e1005066, Mar 2015.

(272) C. Indiani and M. O'Donnell. A proposal: Source of single strand DNA that elicits the SOS response. *Front. Biosci. (Landmark Ed)*, 18:312–323, Jan 2013.

(273) Q. Jiang, K. Karata, R. Woodgate, M. M. Cox, and M. F. Goodman. The active form of DNA polymerase V is UmuD'(2)C-RecA-ATP. *Nature*, 460(7253):359–363, Jul 2009.

(274) R.P. Fuchs. Tolerance of lesions in *E. coli*: Chronological competition between Translesion Synthesis and Damage Avoidance. *DNA Repair (Amst.)*, 44:51–58, Aug 2016.

(275) C. B. Gabbai, J. T. Yeeles, and K. J. Marians. Replisome-mediated translesion synthesis and leading strand template lesion skipping are competing bypass mechanisms. *J. Biol. Chem.*, 289(47):32811–32823, Nov 2014.

(276) M. Ikeda, A. Furukohri, G. Philippin, E. Loechler, M. T. Akiyama, T. Katayama, R. P. Fuchs, and H. Maki. DNA polymerase IV mediates efficient and quick recovery of replication forks stalled at N2-dG adducts. *Nucleic Acids Res.*, 42(13):8461–8472, Jul 2014.

(277) H. A. Jeiranian, B. J. Schalow, C. T. Courcelle, and J. Courcelle. Fate of the replisome following arrest by UV-induced DNA damage in *Escherichia coli*. *Proc. Natl. Acad. Sci. U.S.A.*, 110(28):11421–11426, Jul 2013.

(278) J. E. Kath, S. Jergic, J. M. Heltzel, D. T. Jacob, N. E. Dixon, M. D. Sutton, G. C. Walker, and J. J. Loparo. Polymerase exchange on single DNA molecules reveals processivity clamp control of translesion synthesis. *Proc. Natl. Acad. Sci. U.S.A.*, 111(21):7647–7652, May 2014.

(279) J. E. Kath, S. Chang, M. K. Scotland, J. H. Wilbertz, S. Jergic, N. E. Dixon, M. D. Sutton, and J. J. Loparo. Exchange between *Escherichia coli* polymerases II and III on a processivity clamp. *Nucleic Acids Res.*, 44(4):1681–1690, Feb 2016.

(280) S. Mallik, E. M. Popodi, A. J. Hanson, and P. L. Foster. Interactions and Localization of *Escherichia coli* Error-Prone DNA Polymerase IV after DNA Damage. *J. Bacteriol.*, 197(17):2792–2809, Sep 2015.

(281) L. M. Margara, M. M. Fernandez, E. L. Malchiodi, C. E. Argarana, and M. R. Monti. MutS regulates access of the error-prone DNA polymerase Pol IV to replication sites: a novel mechanism for maintaining replication fidelity. *Nucleic Acids Res.*, 44(16):7700–7713, Sep 2016.

(282) M. F. Goodman and R. Woodgate. Translesion DNA polymerases. *Cold Spring Harb Perspect Biol*, 5(10):a010363, Oct 2013.

(283) M. F. Goodman, J. P. McDonald, M. M. Jaszczur, and R. Woodgate. Insights into the complex levels of regulation imposed on *Escherichia coli* DNA polymerase V. *DNA Repair (Amst.)*, 44:42–50, Aug 2016.

(284) M. F. Goodman. Better living with hyper-mutation. *Environ. Mol. Mutagen.*, 57(6):421–434, Jul 2016.

(285) K. Naiman, V. Pages, and R. P. Fuchs. A defect in homologous recombination leads to increased translesion synthesis in *E. coli*. *Nucleic Acids Res.*, 44(16):7691–7699, Sep 2016.

(286) C. A. Bonner, S. Hays, K. McEntee, and M. F. Goodman. DNA polymerase II is encoded by the DNA damage-inducible *dinA* gene of *Escherichia coli*. *Proc. Natl. Acad. Sci. U.S.A.*, 87(19):7663–7667, Oct 1990.

(287) M. Escarceller, J. Hicks, G. Gudmundsson, G. Trump, D. Touati, S. Lovett, P. L. Foster, K. McEntee, and M. F. Goodman. Involvement of *Escherichia coli* DNA polymerase II in response to oxidative damage and adaptive mutation. *J. Bacteriol.*, 176(20):6221–6228, Oct 1994.

(288) H. Iwasaki, A. Nakata, G. C. Walker, and H. Shinagawa. The *Escherichia coli polB* gene, which encodes DNA polymerase II, is regulated by the SOS system. *J. Bacteriol.*, 172(11):6268–6273, Nov 1990.

(289) Z. Qiu and M. F. Goodman. The *Escherichia coli polB* locus is identical to *dinA*, the structural gene for DNA polymerase II. Characterization of Pol II purified from a *polB* mutant. *J. Biol. Chem.*, 272(13):8611–8617, Mar 1997.

(290) I. Bjedov, C. N. Dasgupta, D. Slade, S. Le Blastier, M. Selva, and I. Matic. Involvement of *Escherichia coli* DNA polymerase IV in tolerance of cytotoxic alkylating DNA lesions*in vivo*. *Genetics*, 176(3):1431–1440, Jul 2007.

(291) D. F. Jarosz, V. G. Godoy, J. C. Delaney, J. M. Essigmann, and G. C. Walker. A single amino acid governs enhanced activity of DinB DNA polymerases on damaged templates. *Nature*, 439(7073):225–228, Jan 2006.

(292) D. F. Jarosz, P. J. Beuning, S. E. Cohen, and G. C. Walker. Y-family DNA polymerases in *Escherichia coli*. *Trends Microbiol.*, 15(2):70–77, Feb 2007.

(293) D. F. Jarosz, S. E. Cohen, J. C. Delaney, J. M. Essigmann, and G. C. Walker. A DinB variant reveals diverse physiological consequences of incomplete TLS extension by a Y-family DNA polymerase. *Proc. Natl. Acad. Sci. U.S.A.*, 106(50):21137–21142, Dec 2009.

(294) K. R. Ona, C. T. Courcelle, and J. Courcelle. Nucleotide excision repair is a predominant mechanism for processing nitrofurazone-induced DNA damage in *Escherichia coli*. *J. Bacteriol.*, 191(15):4959–4965, Aug 2009.

(295) G.C. Walker, S.E. Cohen, D.F. Jarosz, and J.J. Foti. Control and function of translesion DNA polymerases. *FASEB Journal*, 24(1), 2010.

(296) S. Cruet-Hennequart, K. Gallagher, A. M. Sokol, S. Villalan, A. M. Prendergast, and M. P. Carty. DNA polymerase eta, a key protein in translesion synthesis in human cells. *Subcell. Biochem.*, 50:189–209, 2010.

(297) J. G. Jansen, A. Tsaalbi-Shtylik, and N. de Wind. Roles of mutagenic translesion synthesis in mammalian genome stability, health and disease. *DNA Repair (Amst.)*, 29:56–64, May 2015.

(298) J. E. Sale. Translesion DNA synthesis and mutagenesis in eukaryotes. *Cold Spring Harb Perspect Biol*, 5(3):a012708, Mar 2013.

(299) E. M. Boehm, M. Spies, and M. T. Washington. PCNA tool belts and polymerase bridges form during translesion synthesis. *Nucleic Acids Res.*, 44(17):8250–8260, Sep 2016.

(300) M. Hedglin, B. Pandey, and S. J. Benkovic. Characterization of human translesion DNA synthesis across a UV-induced DNA lesion. *Elife*, 5, Oct 2016.

(301) J. McIntyre and R. Woodgate. Regulation of translesion DNA synthesis: Posttranslational modification of lysine residues in key proteins. *DNA Repair (Amst.)*, 29:166–179, May 2015.

(302) A. Quinet, D. J. Martins, A. T. Vessoni, D. Biard, A. Sarasin, A. Stary, and C. F. Menck. Translesion synthesis mechanisms depend on the nature of DNA damage in UV-irradiated human cells. *Nucleic Acids Res.*, 44(12):5717–5731, Jul 2016.

(303) F. X. Barre, B. Soballe, B. Michel, M. Aroyo, M. Robertson, and D. Sherratt. Circles: the replication-recombination-chromosome segregation connection. *Proc. Natl. Acad. Sci. U.S.A.*, 98(15):8189–8195, Jul 2001.

(304) D. J. Sherratt, B. Soballe, F. X. Barre, S. Filipe, I. Lau, T. Massey, and J. Yates. Recombination and chromosome segregation. *Philos. Trans. R. Soc. Lond., B, Biol. Sci.*, 359(1441):61–69, Jan 2004.

(305) T. Shibata, T. Hishida, Y. Kubota, Y. W. Han, H. Iwasaki, and H. Shinagawa. Functional overlap between RecA and MgsA (RarA) in the rescue of stalled replication forks in *Escherichia coli*. *Genes Cells*, 10(3):181–191, Mar 2005.

(306) S. Totemeyer, N. A. Booth, W. W. Nichols, B. Dunbar, and I. R. Booth. From famine to feast: the role of methylglyoxal production in *Escherichia coli*. *Mol. Microbiol.*, 27(3):553–562, Feb 1998.

(307) A. Costes, F. Lecointe, S. McGovern, S. Quevillon-Cheruel, and P. Polard. The C-terminal domain of the bacterial SSB protein acts as a DNA maintenance hub at active chromosome replication forks. *PLoS Genet.*, 6(12):e1001238, Dec 2010.

(308) A. N. Page, N. P. George, A. H. Marceau, M. M. Cox, and J. L. Keck. Structure and biochemical activities of *Escherichia coli* MgsA. *J. Biol. Chem.*, 286(14):12075–12085, Apr 2011.

(309) R. A. Bish and M. P. Myers. Werner helicase-interacting protein 1 binds polyubiquitin via its zinc finger domain. *J. Biol. Chem.*, 282(32):23184–23193, Aug 2007.

(310) N. Crosetto, M. Bienko, R. G. Hibbert, T. Perica, C. Ambrogio, T. Kensche, K. Hofmann, T. K. Sixma, and I. Dikic. Human Wrnip1 is localized in replication factories in a ubiquitin-binding zinc finger-dependent manner. *J. Biol. Chem.*, 283(50):35173–35185, Dec 2008.

(311) I. Saugar, J. L. Parker, S. Zhao, and H. D. Ulrich. The genome maintenance factor Mgs1 is targeted to sites of replication stress by ubiquitylated PCNA. *Nucleic Acids Res.*, 40(1):245–257, Jan 2012.

(312) I.F. Lau, S.R. Filipe, . Soballe, O. Okstad, F-X. Barre, and D.J. Sherratt. Spatial and temporal organization of replicating *Escherichia coli* chromosomes. *Molecular Microbiology*, 49(3):731–743, 2003.

(313) Z. Kelman and M. O'Donnell. DNA polymerase III holoenzyme: structure and function of a chromosomal replicating machine. *Annu. Rev. Biochem.*, 64:171–200, 1995.

(314) D. Vandewiele, A. R. Fernandez de Henestrosa, A. R. Timms, B. A. Bridges, and R. Woodgate. Sequence analysis and phenotypes of five temperature sensitive mutator alleles of *dnaE*, encoding modified alpha-catalytic subunits of *Escherichia coli* DNA polymerase III holoenzyme. *Mutat. Res.*, 499(1):85–95, Jan 2002.

(315) B. Michel and A. K. Sinha. The inactivation of *rfaP*, *rarA* or *sspA* gene improves the viability of the *Escherichia coli* DNA polymerase III holD mutant. *Mol. Microbiol.*, 104(6):1008–1026, Jun 2017.

(316) A. Yoshimura, M. Seki, and T. Enomoto. The role of WRNIP1 in genome maintenance. *Cell Cycle*, 16(6):515–521, Mar 2017.

(317) D. Branzei, M. Seki, F. Onoda, and T. Enomoto. The product of *Saccharomyces cerevisiae WHIP*/*MGS1*, a gene related to replication factor C genes, interacts functionally with DNA polymerase delta. *Mol. Genet. Genomics*, 268(3):371–386, Nov 2002.

(318) T. Hishida, H. Iwasaki, T. Ohno, T. Morishita, and H. Shinagawa. A yeast gene, *MGS1*, encoding a DNA-dependent AAA(+) ATPase is required to maintain genome stability. *Proc. Natl. Acad. Sci. U.S.A.*, 98(15):8283–8289, Jul 2001.

(319) T. Kim, S. Chitteni-Pattu, B. L. Cox, E. A. Wood, S. J. Sandler, and M. M. Cox. Directed Evolution of RecA Variants with Enhanced Capacity for Conjugational Recombination. *PLoS Genet.*, 11(6):e1005278, Jun 2015.

(320) T. Tsurimoto, A. Shinozaki, M. Yano, M. Seki, and T. Enomoto. Human Werner helicase interacting protein 1 (WRNIP1) functions as a novel modulator for DNA polymerase delta. *Genes Cells*, 10(1):13–22, Jan 2005.

(321) T. H. Stanage, A. N. Page, and M. M. Cox. DNA flap creation by the RarA/MgsA protein of *Escherichia coli*. *Nucleic Acids Res.*, 45(5):2724–2735, Mar 2017.

(322) D Branzei, M. Seki, F. Onoda, H. Yagi, Y-I. Kawabe, and T. Enomoto. Characterization of the slow-growth phenotype of *S. cerevisiae whip*/*mgs1 sgs1* double deletion mutants. *DNA Repair*, 1(8):671 – 682, 2002.

(323) T Hayashi, M Seki, E. Inoue, A. Yoshimura, Y. Kusa, S. Tada, and T. Enomoto. Vertebrate wrnip1 and blm are required for efficient maintenance of genome stability. *Genes and Genetic Systems*, 83(1):95–100, 2008.

(324) T. Hishida, T. Ohno, H. Iwasaki, and H. Shinagawa. *Saccharomyces cerevisiae* MGS1 is essential in strains deficient in the RAD6-dependent DNA damage tolerance pathway. *EMBO J.*, 21(8):2019–2029, Apr 2002.

(325) N. D. Vijeh Motlagh, M. Seki, D. Branzei, and T. Enomoto. Mgs1 and Rad18/Rad5/Mms2 are required for survival of *Saccharomyces cerevisiae* mutants with novel temperature/cold sensitive alleles of the DNA polymerase delta subunit, Pol31. *DNA Repair (Amst.)*, 5(12):1459–1474, Dec 2006.

(326) A. Yoshimura, M. Seki, T. Hayashi, Y. Kusa, S. Tada, Y. Ishii, and T. Enomoto. Functional relationships between Rad18 and WRNIP1 in vertebrate cells. *Biol. Pharm. Bull.*, 29(11):2192–2196, Nov 2006.

(327) A. Yoshimura, M. Seki, M. Kanamori, S. Tateishi, T. Tsurimoto, S. Tada, and T. Enomoto. Physical and functional interaction between WRNIP1 and RAD18. *Genes Genet. Syst.*, 84(2):171–178, Apr 2009.

(328) P. McInerney and M. O'Donnell. Functional uncoupling of twin polymerases: mechanism of polymerase dissociation from a lagging-strand block. *J. Biol. Chem.*, 279(20):21543–21551, May 2004.

(329) J.S. Lewis, L.M. Spenkelink, G.D. Schauer, F. R. Hill, R.E. Georgescu, M.E. O'Donnell, and A.M. van Oijen. Single-molecule visualization of *Saccharomyces cerevisiae* leading-strand synthesis reveals dynamic interaction between MTC and the replisome. *Proc. Natl. Acad. Sci. U.S.A*, 114(40):10630–10635, 2017.

(330) C. Danilowicz, E. Feinstein, A. Conover, V. W. Coljee, J. Vlassakis, Y. L. Chan, D. K. Bishop, and M. Prentiss. RecA homology search is promoted by mechanical stress along the scanned duplex DNA. *Nucleic Acids Res.*, 40(4):1717–1727, Feb 2012.

(331) L. Jiang and M. Prentiss. RecA-mediated sequence homology recognition as an example of how searching speed in self-assembly systems can be optimized by balancing entropic and enthalpic barriers. *Phys Rev E Stat Nonlin Soft Matter Phys*, 90(2):022704, Aug 2014.

(332) Z. Qi, S. Redding, J. Y. Lee, B. Gibb, Y. Kwon, H. Niu, W. A. Gaines, P. Sung, and E. C. Greene. DNA sequence alignment by microhomology

sampling during homologous recombination. *Cell*, 160(5):856–869, Feb 2015.

(333) R. Lestini and B. Michel. UvrD controls the access of recombination proteins to blocked replication forks. *EMBO J.*, 26(16):3804–3814, Aug 2007.

(334) R.E. Lenski, M.R. Rose, S.C. Simpson, and S.C. Tadler. Long-term experimental evolution in *Escherichia coli*. I. Adaptation and divergence during 2,000 generations. *The American Naturalist*, 138(6):1315–1341, 1991.

(335) R. T. Byrne, A. J. Klingele, E. L. Cabot, W. S. Schackwitz, J. A. Martin, J. Martin, Z. Wang, E. A. Wood, C. Pennacchio, L. A. Pennacchio, N. T. Perna, J. R. Battista, and M. M. Cox. Evolution of extreme resistance to ionizing radiation via genetic adaptation of DNA repair. *Elife*, 3:e01322, Mar 2014.

(336) M. J. Wiser and R. E. Lenski. A Comparison of Methods to Measure Fitness in *Escherichia coli*. *PLoS ONE*, 10(5):e0126210, 2015.

(337) K. C. Smith and R. C. Sharma. A model for the recA-dependent repair of excision gaps in UV-irradiated *Escherichia coli*. *Mutat. Res.*, 183(1):1–9, Jan 1987.

(338) Y. C. Tseng, J. L. Hung, and T. C. Wang. Involvement of RecF pathway recombination genes in postreplication repair in UV-irradiated *Escherichia coli* cells. *Mutat. Res.*, 315(1):1–9, Jul 1994.

(339) T. C. Wang and K. C. Smith. Mechanisms for recF-dependent and recB-dependent pathways of postreplication repair in UV-irradiated *Escherichia coli* uvrB. *J. Bacteriol.*, 156(3):1093–1098, Dec 1983.

(340) M. M. Cox. Regulation of bacterial RecA protein function. *Crit. Rev. Biochem. Mol. Biol.*, 42(1):41–63, 2007.

(341) L. H. Sanders, A. Rockel, H. Lu, D. J. Wozniak, and M. D. Sutton. Role of *Pseudomonas aeruginosa dinB*-encoded DNA polymerase IV in mutagenesis. *J. Bacteriol.*, 188(24):8573–8585, Dec 2006.

(342) B. Yuan, H. Cao, Y. Jiang, H. Hong, and Y. Wang. Efficient and accurate bypass of N2-(1-carboxyethyl)-2'-deoxyguanosine by DinB DNA polymerase *in vitro* and *in vivo*. *Proc. Natl. Acad. Sci. U.S.A.*, 105(25):8679–8684, Jun 2008.

(343) A. B. Williams, K. M. Hetrick, and P. L. Foster. Interplay of DNA repair, homologous recombination, and DNA polymerases in resistance to the DNA damaging agent 4-nitroquinoline-1-oxide in *Escherichia coli*. *DNA Repair (Amst.)*, 9(10):1090–1097, Oct 2010.

(344) C. T. Courcelle, J. J. Belle, and J. Courcelle. Nucleotide excision repair or polymerase V-mediated lesion bypass can act to restore UV-arrested

replication forks in *Escherichia coli*. *J. Bacteriol.*, 187(20):6953–6961, Oct 2005.

(345) S.S. Henrikus, E.A. Wood, J.P. McDonald, M.M. Cox, R. Woodgate, M.F. Goodman, A.M. van Oijen, and A. Robinson. DNA polymerase IV primarily operates outside of DNA replication forks in *Escherichia coli*. *PLoS Genetics*, 2017.

(346) J. T. Yeeles and K. J. Marians. Dynamics of leading-strand lesion skipping by the replisome. *Mol. Cell*, 52(6):855–865, Dec 2013.

(347) M. Banach-Orlowska, I. J. Fijalkowska, R. M. Schaaper, and P. Jonczyk. DNA polymerase II as a fidelity factor in chromosomal DNA synthesis in *Escherichia coli*. *Mol. Microbiol.*, 58(1):61–70, Oct 2005.

(348) I.J. Fijalkowska, P. Jonczyk, M.M. Tkaczyk, M. Bialoskorska, and R.M. Schaaper. Unequal fidelity of leading strand and lagging strand DNA replication on the *Escherichia coli* chromosome. *Proc. Natl. Acad. Sci. U.S.A*, 95(17):10020–10025, 1998.

(349) D. Gawel, P. Jonczyk, M. Bialoskorska, R.M. Schaaper, and I.J. Fijalkowska. Asymmetry of frameshift mutagenesis during leading and lagging-strand replication in *Escherichia coli*. *Mutation research*, 501(1-2):129–136, April 2002.

(350) W. Kuban, M. Banach-Orlowska, M. Bialoskorska, A. Lipowska, R. M. Schaaper, P. Jonczyk, and I. J. Fijalkowska. Mutator phenotype resulting from DNA polymerase IV overproduction in *Escherichia coli*: preferential mutagenesis on the lagging strand. *J. Bacteriol.*, 187(19):6862–6866, Oct 2005.

(351) M. Maliszewska-Tkaczyk, P. Jonczyk, M. Bialoskorska, R. M. Schaaper, and I. J. Fijalkowska. SOS mutator activity: unequal mutagenesis on leading and lagging strands. *Proc. Natl. Acad. Sci. U.S.A.*, 97(23):12678–12683, Nov 2000.

(352) P. V. Shcherbakova and I. J. Fijalkowska. Translesion synthesis DNA polymerases and control of genome stability. *Front. Biosci.*, 11:2496–2517, Sep 2006.

(353) J. E. Graham, K. J. Marians, and S. C. Kowalczykowski. Independent and Stochastic Action of DNA Polymerases in the Replisome. *Cell*, 169(7):1201–1213, Jun 2017.

(354) J. B. Lee, R. K. Hite, S. M. Hamdan, X. S. Xie, C. C. Richardson, and A. M. van Oijen. DNA primase acts as a molecular brake in DNA replication. *Nature*, 439(7076):621–624, Feb 2006.

(355) J. A. Morin, F. J. Cao, J. M. Lazaro, J. R. Arias-Gonzalez, J. M. Valpuesta, J. L. Carrascosa, M. Salas, and B. Ibarra. Active DNA unwinding dynamics during processive DNA replication. *Proc. Natl. Acad. Sci. U.S.A.*, 109(21):8115–8120, May 2012.

(356) J.A. Morin, B. Ibarra, and F.J. Cao. Kinetic modeling of molecular motors: pause model and parameter determination from single-molecule experiments. *Journal of Statistical Mechanics: Theory and Experiment*, 2016(5):054031, 2016.

(357) J. J. Schwartz and S. R. Quake. Single molecule measurement of the "speed limit" of DNA polymerase. *Proc. Natl. Acad. Sci. U.S.A.*, 106(48):20294–20299, Dec 2009.

(358) P. Howard-Flanders. Repair by genetic recombination in bacteria: overview. *Basic Life Sci.*, 5A:265–274, 1975.

(359) R. H. Rothman and A. J. Clark. Defective excision and postreplication repair of UV-damaged DNA in a *recL* mutant strain of *E. coli* K-12. *Mol. Gen. Genet.*, 155(3):267–277, Oct 1977.

(360) R. H. Rothman and A. J. Clark. The dependence of postreplication repair on uvrB in a recF mutant of *Escherichia coli* K-12. *Mol. Gen. Genet.*, 155(3):279–286, Oct 1977.

(361) W. D. Rupp and P. Howard-Flanders. Discontinuities in the DNA synthesized in an excision-defective strain of *Escherichia coli* following ultraviolet irradiation. *J. Mol. Biol.*, 31(2):291–304, Jan 1968.

(362) W. D. Rupp, C. E. Wilde, D. L. Reno, and P. Howard-Flanders. Exchanges between DNA strands in ultraviolet-irradiated *Escherichia coli*. *J. Mol. Biol.*, 61(1):25–44, Oct 1971.

(363) P. Howard-Flanders, W. D. Rupp, B. M. Wilkins, and R. S. Cole. DNA replication and recombination after UV irradiation. *Cold Spring Harb. Symp. Quant. Biol.*, 33:195–207, 1968.

(364) S. G. Sedgwick. Genetic and kinetic evidence for different types of postreplication repair in *Escherichia coli* B. *J. Bacteriol.*, 123(1):154–161, Jul 1975.

(365) D. A. Youngs and K. C. Smith. Genetic control of multiple pathways of postreplicational repair in uvrB strains of *Escherichia coli* K-12. *J. Bacteriol.*, 125(1):102–110, Jan 1976.

(366) J. Courcelle and P. C. Hanawalt. RecA-dependent recovery of arrested DNA replication forks. *Annu. Rev. Genet.*, 37:611–646, 2003.

(367) P. C. Hanawalt, P. K. Cooper, A. K. Ganesan, and C. A. Smith. DNA repair in bacteria and mammalian cells. *Annu. Rev. Biochem.*, 48:783–836, 1979.

(368) G. Grompone, N. Sanchez, S. Dusko Ehrlich, and B. Michel. Requirement for RecFOR-mediated recombination in priA mutant. *Mol. Microbiol.*, 52(2):551–562, Apr 2004.

(369) T. V. Wang and K. C. Smith. recF-dependent and recF recB-independent DNA gap-filling repair processes transfer dimer-containing parental strands to daughter strands in *Escherichia coli* K-12 uvrB. *J. Bacteriol.*, 158(2):727–729, May 1984.

(370) T. C. Wang and K. C. Smith. reca (srf) suppression of recf deficiency in the postreplication repair of uv-irradiated *Escherichia coli* k-12. *J. Bacteriol.*, 168:940–946, 1986.

(371) B. L. Webb, M. M. Cox, and R. B. Inman. Recombinational DNA repair: the RecF and RecR proteins limit the extension of RecA filaments beyond single-strand DNA gaps. *Cell*, 91(3):347–356, Oct 1997.

(372) V. Pages and R. P. Fuchs. Uncoupling of leading- and lagging-strand DNA replication during lesion bypass *in vivo*. *Science*, 300(5623):1300–1303, May 2003.

(373) A. R. Lehmann and R. P. Fuchs. Gaps and forks in DNA replication: Rediscovering old models. *DNA Repair (Amst.)*, 5(12):1495–1498, Dec 2006.

(374) J. Courcelle, J. J. Belle, and C. T. Courcelle. When replication travels on damaged templates: bumps and blocks in the road. *Res. Microbiol.*, 155(4):231–237, May 2004.

(375) A. Kuzminov. The precarious prokaryotic chromosome. *J. Bacteriol.*, 196(10):1793–1806, May 2014.

(376) R. C. Heller and K. J. Marians. Replisome assembly and the direct restart of stalled replication forks. *Nat. Rev. Mol. Cell Biol.*, 7(12):932–943, Dec 2006.

(377) J. T. Yeeles and K. J. Marians. The *Escherichia coli* replisome is inherently DNA damage tolerant. *Science*, 334(6053):235–238, Oct 2011.

(378) P. Nevin, C. C. Gabbai, and K. J. Marians. Replisome-mediated translesion synthesis by a cellular replicase. *J. Biol. Chem.*, 292(33):13833–13842, Aug 2017.

(379) K. Higuchi, T. Katayama, S. Iwai, M. Hidaka, T. Horiuchi, and H. Maki. Fate of DNA replication fork encountering a single DNA lesion during oriC plasmid DNA replication *in vitro*. *Genes Cells*, 8(5):437–449, May 2003.

(380) J. H. Kim, Y. H. Kang, H. J. Kang, D. H. Kim, G. H. Ryu, M. J. Kang, and Y. S. Seo. *in vivo* and *in vitro* studies of Mgs1 suggest a link between genome instability and Okazaki fragment processing. *Nucleic Acids Res.*, 33(19):6137–6150, 2005.

(381) L. Laureti, J. Demol, R. P. Fuchs, and V. Pages. Bacterial Proliferation: Keep Dividing and Don't Mind the Gap. *PLoS Genet.*, 11(12):e1005757, Dec 2015.

(382) K. A. Datsenko and B. L. Wanner. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc. Natl. Acad. Sci. U.S.A.*, 97(12):6640–6645, Jun 2000.

(383) R.E. Georgescu, G.D. Schauer, N.Y. Yao, L.D. Langston, O. Yurieva, D. Zhang, J. Finkelstein, and M.E. O'Donnell. Reconstitution of a eukaryotic replisome reveals suppression mechanisms that define leading/lagging strand operation. *Elife*, 4:e04988, Apr 2015.

(384) I. F. Azmi, S. Watanabe, M. F. Maloney, S. Kang, J. A. Belsky, D. M. MacAlpine, C. L. Peterson, and S. P. Bell. Nucleosomes influence multiple steps during replication initiation. *Elife*, 6, Mar 2017.

(385) S. Devbhandari, J. Jiang, C. Kumar, I. Whitehouse, and D. Remus. Chromatin Constrains the Initiation and Elongation of DNA Replication. *Mol. Cell*, 65(1):131–141, Jan 2017.

(386) J. M. Dewar and J. C. Walter. Chromosome biology: conflict management for replication and transcription. *Curr. Biol.*, 23(5):R200–202, Mar 2013.

(387) L. D. Langston, D. Zhang, O. Yurieva, R. E. Georgescu, J. Finkelstein, N. Y. Yao, C. Indiani, and M. E. O'Donnell. CMG helicase and DNA polymerase ÃŐÂţ form a functional 15-subunit holoenzyme for eukaryotic leading-strand DNA replication. *Proc. Natl. Acad. Sci. U.S.A.*, 111(43):15390–15395, Oct 2014.

(388) J. T. Yeeles, A. Janska, A. Early, and J. F. Diffley. How the Eukaryotic Replisome Achieves Rapid and Efficient DNA Replication. *Mol. Cell*, 65(1):105–116, Jan 2017.

(389) C. F. Kurat, J. T. Yeeles, H. Patel, A. Early, and J. F. Diffley. Chromatin Controls DNA Replication Origin Selection, Lagging-Strand Synthesis, and Replication Fork Rates. *Mol. Cell*, 65(1):117–130, Jan 2017.

(390) M. Lõoke, M. F. Maloney, and S. P. Bell. Mcm10 regulates DNA replication elongation by stimulating the CMG replicative helicase. *Genes Dev.*, 31(3):291–305, Feb 2017.

(391) A. Gambus, R.C. Jones, A. Sanchez-Diaz, M. Kanemaki, F. van Deursen, R.D. Edmondson, and K. Labib. GINS maintains association of Cdc45 with MCM in replisome progression complexes at eukaryotic DNA replication forks. *Nat. Cell Biol.*, 8(4):358–366, Apr 2006.

(392) Y. Katou, Y. Kanoh, M. Bando, H. Noguchi, H. Tanaka, T. Ashikari, K. Sugimoto, and K. Shirahige. S-phase checkpoint proteins Tof1 and Mrc1 form a stable replication-pausing complex. *Nature*, 424(6952):1078–1083, Aug 2003.

(393) A. A. Alcasabas, A. J. Osborn, J. Bachant, F. Hu, P. J. Werler, K. Bousset, K. Furuya, J. F. Diffley, A. M. Carr, and S. J. Elledge. Mrc1 transduces signals of DNA replication stress to activate Rad53. *Nat. Cell Biol.*, 3(11):958–965, Nov 2001.

(394) A. J. Osborn and S. J. Elledge. Mrc1 is a replication fork component whose phosphorylation in response to DNA replication stress activates Rad53. *Genes Dev.*, 17(14):1755–1767, Jul 2003.

(395) S. J. Szyjka, C. J. Viggiani, and O. M. Aparicio. Mrc1 is required for normal progression of replication forks throughout chromatin in *S. cerevisiae*. *Mol. Cell*, 19(5):691–697, Sep 2005.

(396) H. Tourriere, G. Versini, V. Cordon-Preciado, C. Alabert, and P. Pasero. Mrc1 and Tof1 promote replication fork progression and recovery independently of Rad53. *Mol. Cell*, 19(5):699–706, Sep 2005.

(397) B. Hodgson, A. Calzada, and K. Labib. Mrc1 and Tof1 regulate DNA replication forks in different ways during normal S phase. *Mol. Biol. Cell*, 18(10):3894–3902, Oct 2007.

(398) H. Lou, M. Komata, Y. Katou, Z. Guan, C. C. Reis, M. Budd, K. Shirahige, and J. L. Campbell. Mrc1 and DNA polymerase epsilon function together in linking DNA replication and the S phase checkpoint. *Mol. Cell*, 32(1):106–117, Oct 2008.

(399) G. D. Schauer and M. E. O'Donnell. Quality control mechanisms exclude incorrect polymerases from the eukaryotic replication fork. *Proc. Natl. Acad. Sci. U.S.A.*, 114(4):675–680, Jan 2017.

(400) M. Bando, Y. Katou, M. Komata, H. Tanaka, T. Itoh, T. Sutani, and K. Shirahige. Csm3, Tof1, and Mrc1 form a heterotrimeric mediator complex that associates with DNA replication forks. *J. Biol. Chem.*, 284(49):34355–34365, Dec 2009.

(401) C. Bustamante, J. F. Marko, E. D. Siggia, and S. Smith. Entropic elasticity of lambda-phage DNA. *Science*, 265(5178):1599–1600, Sep 1994.

(402) R.A. Ganai, X.P. Zhang, W.D. Heyer, and E. Johansson. Strand displacement synthesis by yeast DNA polymerase ÃŐÂţ. *Nucleic Acids Res.*, 44(17):8229–8240, Sep 2016.

(403) P. Perez-Arnaiz, I. Bruck, and D. L. Kaplan. Mcm10 coordinates the timely assembly and activation of the replication fork helicase. *Nucleic Acids Res.*, 44(1):315–329, Jan 2016.

(404) F. van Deursen, S. Sengupta, G. De Piccoli, A. Sanchez-Diaz, and K. Labib. Mcm10 associates with the loaded DNA helicase at replication origins and defines a novel step in its activation. *EMBO J.*, 31(9):2195–2206, May 2012.

(405) R. M. Ricke and A. K. Bielinsky. Mcm10 regulates the stability and chromatin association of DNA polymerase-alpha. *Mol. Cell*, 16(2):173–185, Oct 2004.

(406) A. K. Bielinsky. Mcm10: The glue at replication forks. *Cell Cycle*, 15(22):3024–3025, Nov 2016.

(407) S. M. Hamdan, D. E. Johnson, N. A. Tanner, J. B. Lee, U. Qimron, S. Tabor, A. M. van Oijen, and C. C. Richardson. Dynamic DNA helicase-DNA polymerase interactions assure processive replication fork movement. *Mol. Cell*, 27(4):539–549, Aug 2007.

(408) M. D. Sekedat, D. Fenyo, R. S. Rogers, A. J. Tackett, J. D. Aitchison, and B. T. Chait. GINS motion reveals replication fork progression is remarkably uniform throughout the yeast genome. *Mol. Syst. Biol.*, 6:353, 2010.

(409) C. Conti, B. Sacca, J. Herrick, C. Lalou, Y. Pommier, and A. Bensimon. Replication fork velocities at adjacent replication origins are coordinately modified during DNA replication in human cells. *Mol. Biol. Cell*, 18(8):3059–3067, Aug 2007.

(410) G. S. Chadha, A. Gambus, P. J. Gillespie, and J. J. Blow. Xenopus Mcm10 is a CDK-substrate required for replication fork stability. *Cell Cycle*, 15(16):2183–2195, Aug 2016.

(411) Y. Quan, Y. Xia, L. Liu, J. Cui, Z. Li, Q. Cao, X. S. Chen, J. L. Campbell, and H. Lou. Cell-Cycle-Regulated Interaction between Mcm10 and Double Hexameric Mcm2-7 Is Required for Helicase Splitting and Activation during S Phase. *Cell Rep*, 13(11):2576–2586, Dec 2015.

(412) R. M. Baxley and A. K. Bielinsky. Mcm10: A Dynamic Scaffold at Eukaryotic Replication Forks. *Genes (Basel)*, 8(2), Feb 2017.

(413) K. Tougu and K. J. Marians. The interaction between helicase and primase sets the replication fork clock. *J. Biol. Chem.*, 271(35):21398–21405, Aug 1996.

(414) D. Chen, H. Yue, M. M. Spiering, and S. J. Benkovic. Insights into Okazaki fragment synthesis by the T4 replisome: the fate of lagging-strand holoenzyme components and their influence on Okazaki fragment size. *J. Biol. Chem.*, 288(29):20807–20816, Jul 2013.

(415) K. Labib. Making connections at DNA replication forks: Mrc1 takes the lead. *Mol. Cell*, 32(2):166–168, Oct 2008.

(416) D. Bastia, P. Srivastava, S. Zaman, M. Choudhury, B. K. Mohanty, J. Bacal, L. D. Langston, P. Pasero, and M. E. O'Donnell. Phosphorylation of CMG helicase and Tof1 is required for programmed fork arrest. *Proc. Natl. Acad. Sci. U.S.A.*, 113(26):E3639–3648, 06 2016.

(417) S. P. Bell and K. Labib. Chromosome Duplication in *Saccharomyces cerevisiae*. *Genetics*, 203(3):1027–1067, Jul 2016.

(418) K. Siddiqui, K. F. On, and J. F. Diffley. Regulating DNA replication in eukarya. *Cold Spring Harb Perspect Biol*, 5(9), Sep 2013.

(419) I. Bruck, N. Dhingra, and D. L. Kaplan. A Positive Amplification Mechanism Involving a Kinase and Replication Initiation Factor Helps Assemble the Replication Fork Helicase. *J. Biol. Chem.*, 292(8):3062–3073, Feb 2017.

(420) B. B. Zhou and S. J. Elledge. The DNA damage response: putting checkpoints in perspective. *Nature*, 408(6811):433–439, Nov 2000.

(421) V. Pages, R. Janel-Bintz, and R. P. Fuchs. Pol III proofreading activity prevents lesion bypass as evidenced by its molecular signature within *E. coli* cells. *J. Mol. Biol.*, 352(3):501–509, Sep 2005.

(422) O. J. Becherel and R. P. Fuchs. SOS mutagenesis results from up-regulation of translesion synthesis. *J. Mol. Biol.*, 294(2):299–306, Nov 1999.

(423) M. C. Leake, J. H. Chandler, G. H. Wadhams, F. Bai, R. M. Berry, and J. P. Armitage. Stoichiometry and turnover in single, functioning membrane protein complexes. *Nature*, 443(7109):355–358, Sep 2006.

(424) S. Million-Weaver, A. N. Samadpour, and H. Merrikh. Replication Restart after Replication-Transcription Conflicts Requires RecA in Bacillus subtilis. *J. Bacteriol.*, 197(14):2374–2382, Jul 2015.

(425) Jim O'Neill. Tackling drug-resistant infections globally, May 2016.

(426) A.E. McGrath. *Structural and Functional Studies of the beta2-Sliding Clamp from Pathogenic Bacteria*. PhD thesis, University of Wollongong, March 2017.