Technische Universität München
Fakultät für Physik

The Dynamics of Eukaryotic Replication Initiation Revealed at the Single-Molecule Level

Matthias Josef Scherr

Vollständiger Abdruck der von der Fakultät für Physik der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften (Dr. rer. nat.)

genehmigten Dissertation.

Vorsitzender:	Prof. Dr. Ulrich Gerland
Prüfer der Dissertation:	1. Prof. Dr. Karl Duderstadt
	2. Prof. Dr. Michael Sattler

Die Dissertation wurde am 08.10.2021 bei der Technischen Universität München eingereicht und durch die Fakultät für Physik am 02.12.2021 angenommen.

Excerpts of this thesis have appeared in the following publications / manuscripts:

Scherr, M.J., Abd Wahab, S., Remus, D., and Duderstadt, K.E. (2021). Born to Slide: Mobile Origin Licensing Factors Confer Resistance to Transcription Conflicts. *Preprint at SSRN*. doi:10.2139/ssrn.3775178. (*in revision*)

Dequeker, B.J.H., Brandão, H.B.*, **Scherr, M.J.***, Gassler, J., Powell, S., Gaspar, I., Flyamer, I.M., Tang, W., Stocsits, R., Davidson, I.F., Peters, J.M., Duderstadt, K.E., Mirny, L.A., and Tachibana, K. (2020). MCM complexes are barriers that restrict cohesin-mediated loop extrusion. *Preprint at bioRxiv*. doi:10.1101/2020.10.15.340356. (*in revision*)

Scherr, M.J.*, Safaric, B.*, and Duderstadt, K.E. (2018). Noise in the Machine: Alternative Pathway Sampling is the Rule During DNA Replication. *Bioessays* **40**. doi:10.1002/bies.201700159.

*equal contribution

Following work shown in this thesis was not performed by myself:

- Yeast strain ySA4 expressing ybbR-tagged Cdt1-MCM was generated by Syafiq Abd Wahab (Remus Lab, MSKCC)

- Data and figures in chapter 5.1 were generously provided by Bart Dequeker (Tachibana Lab, IMBA)

- JF546-Cohesin^{STAG1} was generously provided by Iain Davidson (Peters Lab, IMP)

- Sortase A, SFP Synthase and PCD were provided by the Max Planck Institute of Biochemistry Core Facility

- LD555-CH(10)LPETGG peptide was synthesized by Stefan Pettera (Max Planck Institute of Biochemistry Core Facility)

Table of Contents

Zusammenfassung

Chromosomen erinnern an belebte Straßen, die von einer Vielzahl zellulärer Maschinen, die für die Chromosomenduplikation (Replikation) und Genexpression (Transkription), aber auch für die Erhaltung und Organisation des Genoms verantwortlich sind, befahren werden. Aufgrund sich rasch weiterentwickelnden Technologien beginnen wir unsere klassische Sichtweise dieser Mechanismen, von denen man ursprünglich annahm, dass sie sich statisch und robust verhalten, zu überdenken. Stattdessen wird immer deutlicher, dass zelluläre Mechanismen ein gewisses Maß an Plastizität zulassen, indem sie durch Ausnutzung von Proteindynamik, dynamischer Zusammensetzung und schnellen Austausch von Faktoren alternative Wege wählen, um mit dieser äußerst komplexen Umgebung zurechtzukommen, die unweigerlich zu Konflikten führt.

Grundlegend für unser Verständnis der Chromosomenduplikation ist die Vorstellung, dass Replikationsursprünge sowohl als Orte fungieren, an denen MCM-Helikasen (MCMs) während der G1-Phase geladen werden, als auch an denen Synthese später in der S-Phase beginnt. Durch den zeitlichen Abstand zwischen diesen Phasen ist die Assemblierung des Replisoms jedoch möglichen Störungen vor der Replikation ausgesetzt. Im ersten Teil dieser Arbeit untersuchen wir mit Hilfe von mehrfarbiger Einzelmolekül-Fluoreszenzmikroskopie systematisch die Folgen von Begegnungen zwischen aktiv transkribierenden RNA-Polymerasen (RNAPs) und wichtigen Zwischenstufen der Replikationsinitiierung im Kontext von Chromatin. Wir stellen fest, dass MCMs spontan in einen diffusiven DNA-Bindungsmodus wechseln können, wenn sie hoher Ionenstärke ausgesetzt werden und zeigen, dass MCMs diesen diffusiven Modus nutzen können, um Konflikte mit RNAPs zu überwinden. Bemerkenswerterweise zeigen wir, dass RNAP mehrere auf DNA geladene MCMs über weite Strecken schieben kann, wobei einzelne Nukleosomen durch Ausstoßen oder Verlagerung überwunden werden. Erstaunlicherweise stellen wir zudem fest, dass Intermediate der MCM-Assemblierung auch durch RNAP mobilisiert und neu positioniert werden können, was ein Netz alternativer Spezifizierung von Replikationsursprüngen eröffnet. Zusammenfassend zeigen unsere Beobachtungen eine erstaunliche Mobilität der für die Lizensierung von Replikationsursprüngen verantwortlichen Faktoren, um den komplexen Herausforderungen durch verschiedene Hindernisse auf den Chromosomen standhalten zu können.

In Eukaryoten organisiert Cohesin das Genom mittels Schleifenextrusion in Schleifen und topologisch assoziierte Domänen (TADs), welche wichtige Rollen bei der Entwicklung, der zeitlichen Regulierung der Replikation, der Genregulation oder der Antikörpervielfalt spielen. Obwohl sie durch CTCF-Grenzen beschränkt wird, ist wenig darüber bekannt, ob andere auf der DNA wirkende Mechanismen die Cohesin-vermittelte Schleifenextrusion beeinflussen. Im zweiten Teil dieser Arbeit zeigen wir, dass MCMs häufig vorhandene und robuste Barrieren für die Schleifenextrusion in der

G1-Phase darstellen. In Zygoten von Mäusen zeigen wir mittels Ein-Kern-Hi-C, dass Lizensierung von Replikationsursprüngen die durchschnittliche Stärke von Schleifen und TADs in einer CTCF-abhängigen Weise reduziert, was darauf hindeutet, dass MCMs die Schleifenextrusion vor den CTCF-Grenzen behindern. Mit Hilfe von Einzelmolekül-Imaging zeigen wir, dass MCMs Barrieren sind, welche die Cohesin-Translokation *in vitro* einschränken. Wir vermuten, dass MCMs einerseits physische Barrieren und andererseits aktive Grenzen darstellen, indem sie Cohesin durch spezifische Interaktionen, die durch ein YDF-enthaltendes Motiv in der humanen Mcm3-Untereinheit vermittelt werden, an sich binden. Zusammenfassend zeigen wir, dass MCMs eine von der Replikation abweichende Rolle bei der 3-dimensionalen Organisation von Genomen spielen.

Abstract

Chromosomes are reminiscent of busy roads crowded with a variety of cellular machineries responsible for chromosome duplication (replication) and gene expression (transcription) but also genome maintenance and organization. Rapidly evolving techniques start to revolutionize our classical view of these machineries, initially thought to behave in a static and robust manner. Instead, it becomes more and more clear that, to cope with this exceedingly complex environment, which inevitably causes conflicts, cellular machineries allow for a certain degree of plasticity through alternative pathway sampling by exploiting protein dynamics, dynamic composition and rapid exchange of factors.

Fundamental to our understanding of chromosome duplication is the idea that replication origins function both as sites where MCM helicases (MCMs) are loaded during the G1 phase and where synthesis begins later on in S phase. However, the temporal delay between these phases exposes the replisome assembly pathway to potential disruption prior to replication. In the first part of this thesis, we use multicolor, single-molecule fluorescence microscopy to systematically study the consequences of encounters between actively transcribing RNA polymerases (RNAPs) and key replication initiation intermediates in the context of chromatin. We find that MCMs can spontaneously adopt a diffusive DNA binding mode when challenged by high ionic strength and show that MCMs utilize this diffusive mode to overcome conflicts with RNAPs. Remarkably, we demonstrate that RNAP can push multiple licensed MCMs over long distances with individual nucleosomes overcome by ejection or displacement. Unexpectedly, we observe that MCM loading intermediates can also be mobilized and repositioned by RNAP, providing a web of alternative origin specification pathways. Taken together, our observations reveal a surprising mobility in origin licensing factors that confers resistance to the complex challenges posed by diverse obstacles encountered on chromosomes.

In eukaryotes, cohesin organizes genomes by loop extrusion into loops and topologically associating domains (TADs) which play important roles in development, replication timing, gene regulation or antibody diversity. Although restricted by CTCF boundaries, little is known whether other machineries operating on DNA affect cohesin-mediated loop extrusion. In the second part of this thesis, we demonstrate that MCMs are abundant and robust barriers to loop extrusion in G1 phase. In mouse zygotes, we show by single-nucleus Hi-C that origin licensing reduces the average loop and TAD strength in a CTCF-dependent manner, suggesting that MCMs impede loop extrusion prior to CTCF boundaries. Using single-molecule imaging, we demonstrate that MCMs are barriers that restrict cohesin translocation *in vitro*. We suggest that MCMs are physical roadblocks on the one hand and active boundaries on the other hand via tethering cohesin through specific interactions mediated by a YDF-containing motif in the human Mcm3 subunit. Together, we reveal a role of MCMs distinct from replication in shaping 3D genomes.

Chapter 1

Introduction

Life is the product of fundamental physical laws and randomness (Tsimring, 2014). This interplay shapes biology on all time and length scales, from the dynamics of populations to the most basic sub-cellular processes. On the molecular scale, where key transactions require energies not far above those of thermal fluctuations, noise is intrinsic and dominant. This raises the question: when are molecular reactions guided by stochastic, as opposed to deterministic, events? Compelling answers to this question have emerged for simple unimolecular reactions, but large multicomponent systems, like the deoxyribonucleic acid (DNA) replication machinery, which have numerous specialized factors with diverse functions, are more challenging to study and model physically. In many cases, molecular biologists have therefore relied on physical intuition to develop working models based on limited observations. However, our physical intuition is based on our experience in the macro scale world and often betrays us when considering molecular scale events (Purcell, 1977). Fortunately, single-molecule methods have emerged that provide a more direct, real-time view of individual molecular events and pathways. While still incomplete, these observations provide a detailed picture of how stochastic events shape biological reactions. These new developments call for a fundamental shift in how we approach mechanistic modeling of large multicomponent systems. The DNA replication machinery is an example of a multicomponent system for which our mechanistic understanding is rapidly evolving due to single-molecule observations.

1.1 Faithful genome duplication is essential for all life

DNA replication is indispensable for all living organisms and essential for the survival of all life, including simple viruses, bacteria and complex organisms like humans. To allow proliferation and growth, DNA, as the carrier of genetic information, has to be inherited to each daughter cell. The high demand on the replication machinery in terms of fidelity becomes apparent when recalling the amount of DNA produced in a human body during a lifetime. During human development, upon fertilization of the oocyte, a zygote represents the first diploid human cell, containing the remarkable amount of approximately two meters of DNA (Hill, 2012). Even more remarkable is the fact that a human undergoes about 10^{16} cell divisions during a lifetime (Weinberg, 2014), thus synthesizing the astronomical amount of $2x10^{16}$ meters of DNA which is equal to approximately two light years.

Given the high importance of DNA replication, a complex machinery of concerted protein complexes (often referred to as “the replisome”) has developed during evolution. In fact, cooperation of proteins within the replisome evolved to high perfection, ensuring fast, complete and accurate genome duplication. The importance of faithful genome duplication is crucial as malfunctions may lead to mutations, chromosomal aberrations and rearrangements, leading to genome instability and genetic diversity – central hallmarks of cancer (Hanahan and Weinberg, 2011). Thus, it is not surprising that defects in the replication machinery are the underlying reason for more than 80 human diseases including cancer but also other hyperproliferative diseases such as uncontrollable (bacterial and viral) infections or autoimmune diseases (Aladjem et al., 2006). Consequently, a detailed understanding of cellular mechanisms involved in faithful DNA replication is essential to develop novel therapeutic approaches for a variety of human diseases.

Our current understanding of the mechanism of DNA replication dates all the way back to the 19th century. In 1869, during his effort to determine cellular components and thus to understand cellular life, Friedrich Miescher discovered a precipitate different from known proteins in the nuclei of lymphocytes, which he initially named nuclein (Miescher, 1869, 1871). Nuclein was later renamed to nucleic acid by Miescher’s student Richard Altmann in 1889 (Altmann, 1889). Although Albrecht Kossel demonstrated that DNA consists of four different building blocks (adenine, thymine, guanine and cytosine), so called nucleobases (Kossel, 1879, 1891), DNA was not of great interest to the scientific community yet. DNA was long considered being built too simple to carry the vast amount of genetic information and thus assumed to be encoded by the more complex proteins, which were known to be composed of 20 different amino acids. Only much later in the mid 1940s, by revisiting Frederick Griffith’s experiment which led to the “transforming principle” in bacteria (Griffith, 1928), Oswalt Avery, Colin MacLeod and Maclyn McCarty could show that the genetic information is not encoded by proteins but most likely by DNA (Avery et al., 1944). These results were consolidated with studies on bacteriophage T2 by Al Hershey and Marta Chase (Hershey and Chase, 1952) and DNA became widely

accepted as the carrier of genetic information. Only one year later, a milestone was set by work from Rosalind Franklin and Maurice Wilkins as well as James Watson and Francis Crick when they deciphered the double-helical structure of DNA with its specific base pairing (Watson and Crick, 1953). Still in the same decade, Arthur Kornberg and colleagues discovered the DNA replicating enzyme (DNA polymerase) (Bessman et al., 1956) and the principle of semiconservative replication was put forward by Matthew Meselson and Franklin Stahl (Meselson and Stahl, 1958). Finally, a decade later, Har Gobind Khorana, Robert Holley, Marshall Nirenberg and colleagues decrypted the genetic code, marking the beginning of the era of modern molecular biology for which they received the Nobel Prize in 1968 (Singer, 1968).

The foundation of our understanding of replication mechanism derived from these and other seminal experiments conducted in the second half of the 20th century which suggested DNA replication is performed by a single static complex with highly defined operating principles. In contrast to this view, single-molecule observations have revealed that dynamic exchange of core components, pausing events, and several types of DNA loops may underlie coordination of daughter-strand synthesis (Beattie et al., 2017; Duderstadt et al., 2016; Geertsema et al., 2014; Georgescu et al., 2014a; Graham et al., 2017; Kapadia et al., 2020; Lewis et al., 2017a; Lewis et al., 2020). Individually, these events support different mechanistic models which cannot always be reconciled in a single reaction mechanism. However, when viewed as alternative, sometimes parallel, pathways, a more complete and coherent picture emerges. Nonetheless, this multi-pathway view of replication is often at odds with experimental observations that still dominate our textbook view of the process. Given the numerous compelling studies supporting this dynamic view (Beattie et al., 2017; Duderstadt et al., 2016; Geertsema et al., 2014; Georgescu et al., 2014a; Graham et al., 2017; Kapadia et al., 2020; Lewis et al., 2017a; Lewis et al., 2020) we must re-examine the early experiments that underlie our core theories.

The view that DNA replication is a highly coordinated process first emerged from autoradiography and pulse-chase experiments conducted by John Cairns and others (Friedberg et al., 2006; Kornberg and Baker, 1992), which revealed that DNA unwinding is coupled to the synthesis of the daughter strands. The pioneering work of Arthur Kornberg, Tsuneko and Reiji Okazaki and many others that followed, led to the discovery of the first DNA polymerases and the asymmetric mechanism of daughter-strand synthesis (Bessman et al., 1956; Nelson and Cox, 2005; Okazaki et al., 1968). This revealed that while the polymerase on the leading strand can synthesize co-directionally with unwinding, the polymerase on the lagging stand performs synthesis discontinuously, in the opposite direction, by repeatedly restarting on ribonucleic acid (RNA) primers generated by primase. Discontinuous synthesis on the lagging strand leads to the formation of a series of Okazaki fragments (OFs), which are later converted into a continuous strand.

The fundamental asymmetry in the synthesis of the daughter strands presents a

coordination challenge: How can a single protein complex accommodate the opposite directionalities of synthesis? Based on rapid dilution experiments suggesting replication is performed by a single protein complex, Bruce Alberts came up with the elegant proposal that each cycle of lagging-strand synthesis involves the formation of a “trombone loop” that allows both polymerases to reside in the same complex while moving in opposite directions (Alberts et al., 1983). These early observations created a paradigm in the field that efficient DNA replication is a consequence of highly refined operating principles. In the decades that followed, intensive research focused on clarifying the role of the key players and the precise sequence of events that underlies each replication cycle.

As our understanding of the replication machinery has advanced, many central mechanistic questions have nonetheless persisted. In particular, several divergent models have been put forward to explain how daughter-strand synthesis remains synchronized given that lagging-strand synthesis involves a series of slow enzymatic steps (priming and polymerase cycling) while leading-strand synthesis is continuous (Corn et al., 2005; Dixon, 2009; Frick and Richardson, 1999; Hamdan et al., 2009; Lee et al., 2006; Lee and Richardson, 2002; Li and Marians, 2000; Mangiameli et al., 2017; Manosas et al., 2009; Pandey et al., 2009; Tanner et al., 2008; Yuzhakov et al., 1999). Several recent observations of replication with single-molecule approaches have revealed more intrinsic stochasticity in replisome function suggesting many of these models are not mutually exclusive. Instead, they may represent different pathways whose sampling depends on environmental conditions and the current configuration of the replication machinery.

1.2 Single-molecule imaging reveals dynamic events at the replication fork

1.2.1 Mechanical manipulation

Numerous force manipulation techniques are now available that provide a direct view of biological pathways by revealing the real-time dynamics of individual molecules and molecular assemblies (Dulin et al., 2013; Neuman and Nagy, 2008). Optical trapping, magnetic tweezing, and flow stretching are the techniques of choice for studying replisome components having well-established DNA manipulation protocols (Figures 1.1A, 1.1B and 1.1C). These approaches typically rely on the attachment of a micron-sized bead to one end of an individual DNA molecule and surface attachment of the other via biotin-streptavidin linkages or antigen-antibody interactions. The attached beads allow for the application of controlled forces in the 1 to 10 pico-Newton (pN) regime and real-time tracking of DNA length changes. These methods have been widely applied to study nucleic acid manipulating enzymes and each provides unique advantages depending on the application.

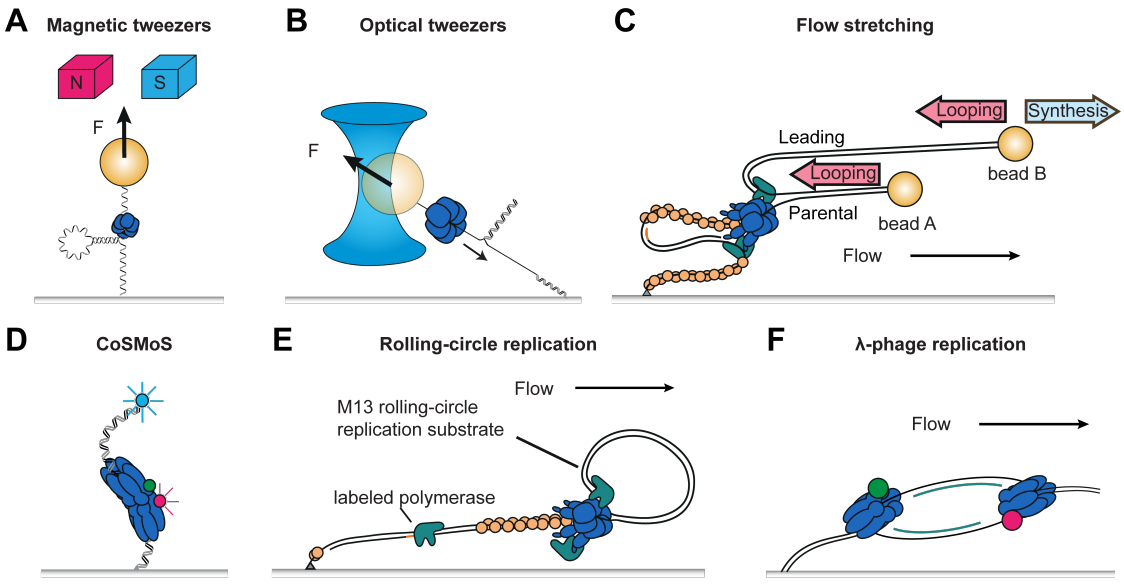

Figure 1.1. Single-molecule approaches to study replisome dynamics
(**A**) Magnetic tweezers: DNA molecules are stretched by a magnetic force (F) applied on paramagnetic beads attached to the DNA end. Length changes are monitored by tracking the bead height. (**B**) Optical tweezers: DNA molecules are stretched by optically trapping a dielectric bead on one end. Changes in length and force are monitored using feedback. (**C**) Flow stretching: A constant flow applies a drag force on two beads that are attached to the ends of surface immobilized DNA molecules. Replisome dynamics are monitored through the beads. (**D**) Co-localization single-molecule spectroscopy (CoSMoS): fluorescent labels on DNA substrates and replication factors reveal the timing of binding and dissociation events. (**E**) Rolling-circle replication: a circular substrate allows for continuous replication. A constant flow stretches out the product to follow synthesis in real-time. (**F**) Origin replication: a 48 kilobases λ-phage DNA substrate is stretched with flow and replication is initiated from an origin in the center. Fluorescent immunostaining reveals the kinetics of synthesis. In (**E**) and (**F**), fluorescent labels provide spatial and temporal information about replisome factors.

To date, optical and magnetic tweezers have offered the highest spatial and temporal resolution information about the individual enzymatic activities that underlie replisome function. These methods have been used to study the mechanism of DNA unwinding by helicases (Burnham et al., 2019; Johnson et al., 2007; Lionnet et al., 2007; Sun et al., 2011), the kinetics of nucleotide incorporation during polymerase synthesis (Maier et al., 2000; Wuite et al., 2000), and DNA looping dynamics by partially reconstituted replication complexes (Manosas et al., 2009). Optical tweezers operate by holding a dielectric bead attached to DNA in the focus of an intense laser beam (Moffitt et al., 2008; Neuman and Nagy, 2008; Svoboda and Block, 1994). The other end of the DNA molecule can be attached either to a surface (Johnson et al., 2007), a pipette tip (Wuite et al., 2000), or a second optically trapped bead (Larson et al., 2008). DNA is then manipulated by moving the position of the optical trap in either a constant force or constant extension mode. In contrast, magnetic tweezers intrinsically operate in constant force mode with no need for feedback conferring greater stability and providing access to a much broader low force regime (Dulin et al., 2013). Optical tweezers can also be combined with fluorescence imaging (see chapter 1.2.2) (correlative single-molecule fluorescence and force microscopy), a powerful approach which allows to manipulate and visualize single-molecule interactions simultaneously in real-time. For instance, this method elucidated different DNA binding modes of the eukaryotic helicase and revealed nucleosome remodeling to specify helicase loading (Li et al., 2021; Wasserman et al., 2019). While only one molecule at a time can be studied using optical tweezers, which makes studies of large multicomponent systems challenging, hundreds of molecules have been observed simultaneously with magnetic tweezers (Berghuis et al., 2016; Ribeck and Saleh, 2008). Magnetic tweezers are routinely used in studies of DNA topology manipulating enzymes where two closely spaced block magnets are rotated to apply controlled torques on individual DNA molecules (Gore et al., 2006; Strick et al., 1998). Recently, magnetic tweezers were combined with flow stretching to monitor gyrase activity on ten thousands of DNA molecules (Agarwal and Duderstadt, 2020).

Flow stretching is a complementary DNA manipulation approach in which a constant flow through a flow cell exerts a drag force on surface tethered DNA molecules with beads at their end (Figure 1.1C) (Tanner and van Oijen, 2009; van Oijen et al., 2003). Similar to magnetic tweezers, flow stretching is intrinsically a constant force method. For a given flow speed, the stretching force on the DNA is constant. Although typical flow stretching configurations provide lower temporal (seconds) and spatial (10s of nanometers) resolution, the method has the unique advantage that massive multiplexing is possible, with a study achieving simultaneous imaging of up to 10^5 molecules (Duderstadt et al., 2016). The multiplexing potential of flow stretching has allowed for the study of primer synthesis regulation and DNA looping dynamics during lagging-strand synthesis (Hamdan et al., 2009; Lee et al., 2006; Tanner et al., 2008). A two-bead flow stretching approach has been developed allowing for independent observation of leading-strand synthesis and lagging-strand looping providing a more complete view of coordination at the replication fork (Duderstadt et al., 2016).

1.2.2 Fluorescence imaging

The unparalleled design flexibility of fluorescence imaging approaches has allowed for their wide application in studies of a diverse range of biological pathways both *in vivo* and *in vitro*. In contrast to mechanical manipulation techniques that only report on functional output, fluorescence imaging techniques also can provide information about the spatial and temporal organization of individual enzymes on substrates and within large complexes (Figures 1.1D, 1.1E and 1.1F).

Many of the first single-molecule fluorescence studies of replisome components were conducted using fluorescence resonance energy transfer (FRET), which provides a real-time view of structural changes during enzymatic events. FRET has been used to study replisome assembly pathways (Smiley et al., 2006; Xi et al., 2005; Zhang et al., 2005), conformational states adopted by polymerases and helicases (Abid Ali et al., 2016; Luo et al., 2007; Ticau et al., 2017), and transient DNA looping events during priming (Pandey et al., 2009). These approaches typically rely on surface attachment of single replication substrates with total internal reflection fluorescence (TIRF) microscopy (see chapter 2.6.9 and Figure 2.5) being the most frequently used method of imaging due to its high-throughput and low background characteristics. Approaches that provide additional colors allow for simultaneous monitoring of conformational changes and assembly events by FRET. Co-localization single-molecule spectroscopy (CoSMoS) is one such method (Friedman et al., 2006), which has been used to study helicase loading and activation during assembly of the eukaryotic replisome (Figure 1.1D) (Champasa et al., 2019; De Jesus-Kim et al., 2021; Ticau et al., 2017; Ticau et al., 2015).

One of the most common fluorescence approaches used to study replisome function leverages a combination of flow stretching and rolling-circle replication to visualize the products of synthesis in real-time (Figure 1.1E). In this assay, replication proceeds continuously around a circular substrate generating a long lagging-strand tail that is surface attached and fluorescently stained for imaging (Tanner et al., 2009). This approach has been used to systematically dissect bacterial replisome function, demonstrating recycling of processivity factors (Tanner et al., 2011), the consequence of additional polymerases residing at the replication fork (Georgescu et al., 2011), and the stochastic nature of daughter-strand synthesis coordination (Graham et al., 2017; Yao et al., 2009). Combining this approach with fluorescently labeled replisome components clarifies the unique structural configurations that underlie key functional intermediates. Experiments conducted using this powerful combination are challenging longstanding views about the stability of the replisome revealing frequent polymerase exchange (Duderstadt et al., 2016; Geertsema et al., 2014; Lewis et al., 2017a; Lewis et al., 2020; Loparo et al., 2011).

Due to increased complexity, the functional output of the eukaryotic replication machinery has been challenging to study at the single-molecule level. The most promising studies have been performed using *X. laevis* and *S. cerevisiae* extracts in combination with flow stretched replication substrates generated using λ-phage DNA

(Figure 1.1F) (Duzdevich et al., 2015; Gruszka et al., 2020; Sparks et al., 2019; Yardimci et al., 2012). In these assays, replication is followed by immune-staining of products or tracking individual labeled components. Recently, the *S. cerevisiae* replisome has been fully reconstituted from purified components (Devbhandari et al., 2017; Georgescu et al., 2014a; Kurat et al., 2017; Yeeles et al., 2015; Yeeles et al., 2017) offering an exciting new direction for single-molecule studies of eukaryotic replication. Albeit full origin activation at the single-molecule level has not been achieved so far, first dynamic features of the *S. cerevisiae* replisome have been revealed by using pre-assembled CMG helicase (Lewis et al., 2017b; Lewis et al., 2020; Schauer et al., 2020; Wasserman et al., 2019).

1.3 A dynamic interplay drives initiation of bacterial DNA replication

The first step towards genome duplication consists of replisome assembly and activation during the initiation process. The existence of origins, or specific genomic loci, that serve as initiation sites for replication is a cornerstone in our understanding of chromosome duplication. However, defining universal characteristics of origins across the domains of life has proven challenging. Already in 1963, before the mechanism of the initiation process was understood, François Jacob, Sydney Brenner and François Cuzin proposed their theory on how chromosome replication in bacteria is initiated and coordinated with respect to cell cycle and division, known as the “replicon model” (Jacob et al., 1963). The replicon model, which guided our initial understanding, postulated the existence of specific DNA sequence elements, termed replicators, that serve as start sites through engagement of an initiator protein. Consistent with this model, highly-refined DNA sequence elements have been discovered in bacterial origins (Bramhill and Kornberg, 1988; Kowalski and Eddy, 1989; Mackiewicz et al., 2004).

In bacteria, AT-rich DNA sequence elements, the origin of replication (oriC) (replicator), are specifically recognized by multiple copies of DnaA, replication initiation protein. Their arrangement triggers formation of a complex that melts oriC double-stranded DNA (dsDNA) by an adenosine triphosphate (ATP)-dependent stretching mechanism (Duderstadt and Berger, 2013; Duderstadt et al., 2011). Following, DnaA and the helicase loader DnaC cooperatively recruit two copies of a ring-shaped, homohexameric helicase DnaB on either strand of the single-stranded DNA (ssDNA) in a 5′-3′ orientation, positioning it on the “future” lagging strand (Arias-Palomo et al., 2019; Bleichert et al., 2017). Next, the primase DnaG transiently binds DnaB to synthesize a short RNA primer. To complete replisome assembly, clamp loader binds activated helicase and recruits β clamp, a processivity factor for DNA polymerase Pol III, to primed DNA. Intriguingly, efficient recruitment of the clamp loader to DnaB depends on the presence of DnaC and DnaG (Monachino et al., 2020). The heteropentameric clamp loader acts as the central hub of the replication machinery as it binds DnaB as well as Pol III core via its three τ subunits, thus coupling helicase unwinding and Pol III synthesis activity. Successful initiation from oriC leads to bidirectional DNA replication with Pol III synthesizing nascent DNA in 5′-3′ direction on both strands. Due to the antiparallel nature of DNA only the leading strand can be synthesized continuously. In contrast, the lagging strand is replicated discontinuously as Pol III synthesizes DNA in the direction opposite from DnaB progression, leading to the formation of an SSB-coated trombone loop between DnaB and the lagging strand Pol III (Figure 1.2). Discontinuous replication includes repeated cycles of primer synthesis, clamp recruitment and Pol III loading, generating a series of 1-2 kilobases (kb) long OFs which are later ligated into a continuous strand.

Upon completion of genome duplication, termination occurs via encounter of two opposed forks, disassembly of the replisome, final gap filling and decatenation (Dewar and Walter, 2017). On the opposite site of oriC, *E. coli*’s circular chromosome possesses

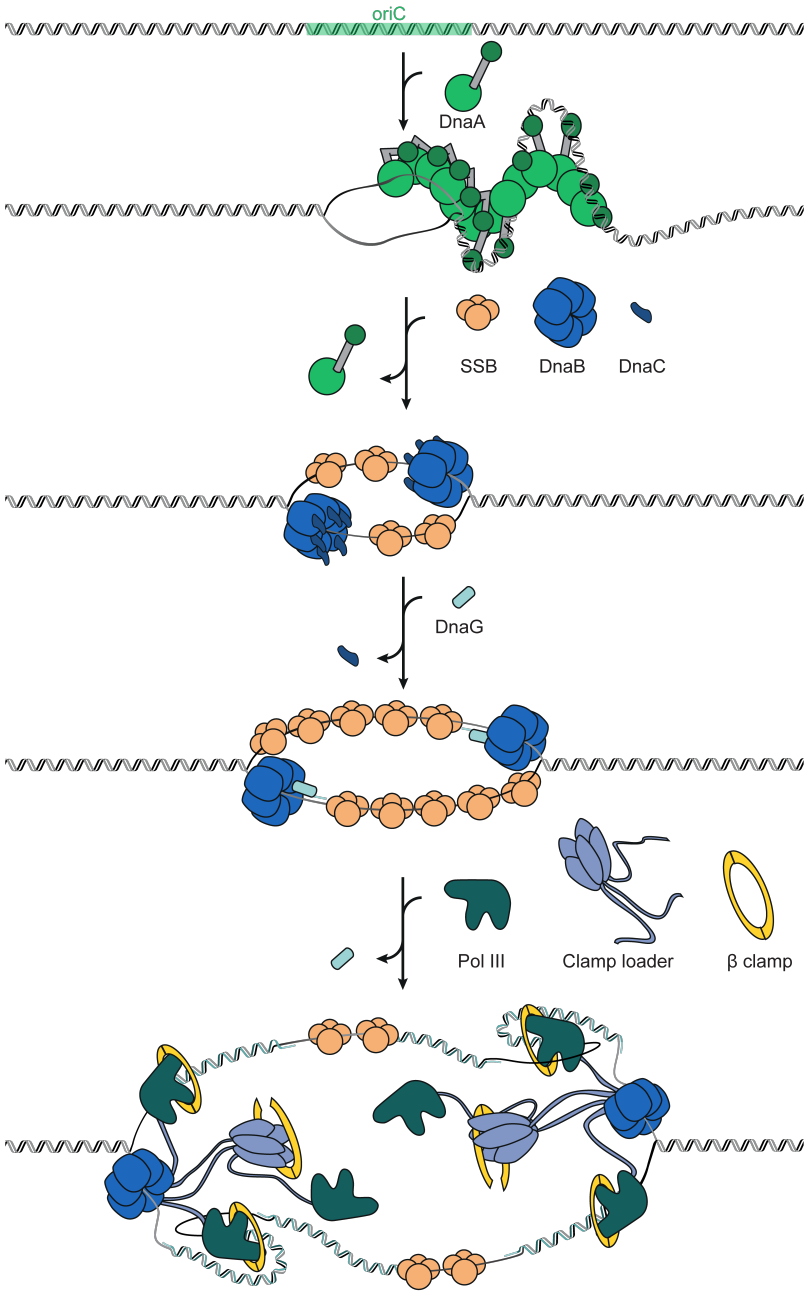

Figure 1.2. Bacterial DNA replication
Bacterial replication starts with the replication initiator DnaA recognizing and remodeling the replication origin oriC. Once the origin is melted, DnaA together with the helicase loader DnaC recruit two copies of the DnaB helicase. Next, primase DnaG binds DnaB to synthesize an RNA primer and stimulates DnaC dissociation. Following, the clamp loader complex is recruited to load the β processivity clamps onto the primers for Pol III cores. Once assembled, bidirectional replication occurs with simultaneous leading- and lagging-strand synthesis. Repeated trombone loop formation allows for lagging-strand synthesis to be conducted within a single complex.

a termination zone consisting of ten ter sites, each forming a one-way fork barrier by binding terminus site-binding protein (Tus). Each fork is able to pass the first five Tus-ter complexes which determines the final site where the two opposing replication forks converge. One model for termination in *E. coli* suggests that both DnaB pass each other

to proceed as an inactive helicase on dsDNA to generate 3′ flaps. Subsequently, DnaB dissociates, 3′ flaps are removed, gaps filled and the last OF is finished by Pol I. The final step of replication generates a catenane which is then removed. Interestingly, it has been shown that failure of 3′ flap removal causes pathological DnaA-independent loading of DnaB at termination sites, leading to re-replication and thus genome instability (Rudolph et al., 2013).

1.3.1 Stochastic events shape the bacterial replication cycle

How daughter-strand synthesis remains temporally coordinated in spite of the slow enzymatic steps that underlie lagging-strand duplication has remained a long-standing mystery in the replication field for which several different models have been put forward. One model proposes that priming pauses leading-strand synthesis to enforce synchronization (Lee et al., 2006). In another, leading-strand synthesis continues during priming, and a faster rate of lagging-strand synthesis ensures coordination (Manosas et al., 2009; Nelson et al., 2008; Pandey et al., 2009). In a third and related model, the rate of helicase unwinding is thought to limit the rate of leading-strand synthesis (Georgescu et al., 2014b). Compelling evidence exists in support of each of these models (Corn et al., 2005; Dixon, 2009; Frick and Richardson, 1999; Hamdan et al., 2009; Lee et al., 2006; Lee and Richardson, 2002; Li and Marians, 2000; Mangiameli et al., 2017; Manosas et al., 2009; Pandey et al., 2009; Tanner et al., 2008; Yuzhakov et al., 1999) and a decisive resolution of these mechanistic disputes has proven illusive. The ability to directly follow events at single replication forks has led to a fundamental shift in our understanding of these events. These observations have revealed that leading- and lagging-strand synthesis are not functionally coordinated and instead stochastic events play a major role in shaping each replication cycle (Duderstadt et al., 2016; Graham et al., 2017; McInerney and O'Donnell, 2004; Yao et al., 2009).

Observations of single replisomes in action have forced us to think again about the coordination of primer synthesis during replication. DNA flow stretching experiments revealed frequent pauses in leading-strand synthesis during priming (Lee et al., 2006) suggesting a possible mechanism to synchronize daughter-strand synthesis. However, contrasting observations emerged from single-molecule FRET (Pandey et al., 2009) and magnetic tweezers (Manosas et al., 2009) studies, which favored continuous leading-strand synthesis supported by formation of single-stranded priming loops. Differences in experimental setup offered one possible explanation for these differing models, but subsequent studies conducted with complete replisomes have demonstrated that aspects of both models support robust replication. Investigations conducted with a two-bead flow stretching assay (Figure 1.1C), using the phage T7 replisome, which serves as a simplified model system for studies of replication mechanism, demonstrated frequent priming loop formation during most replication cycles as well as pauses in leading-strand synthesis that correlated with priming (Figure 1.3A). Similarly, studies conducted using *E. coli* rolling-circle replication have revealed continuous synthesis

during most replication cycles and infrequent pauses in leading-strand synthesis (Graham et al., 2017). Interestingly, pauses in leading-strand synthesis were accompanied with continued helicase unwinding albeit with a dramatically reduced rate. The observation of helicase unwinding uncoupled from leading-strand synthesis has led to the proposal of a third DNA loop type, called a leading-strand loop, that forms between the leading-strand polymerase and helicase (Figure 1.3B).

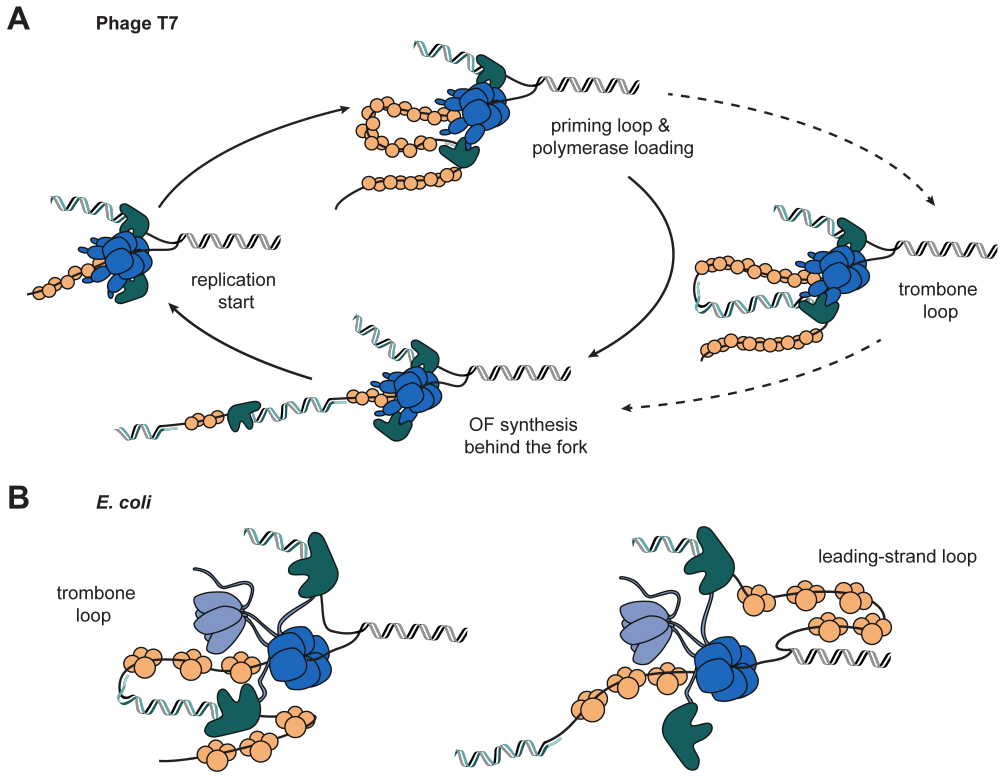

Figure 1.3. Multiple loop types and alternative pathways during bacterial replication
Multiple looping possibilities offer pathways to maintain polymerase coordination. (**A**) In phage T7, ssDNA priming loops form during primer synthesis. Once priming is completed, polymerase loading and Okazaki fragment (OF) synthesis are performed. Following, multiple parallel pathways may be sampled during each replication cycle involving either trombone loop formation or polymerase release and synthesis behind the replication fork. (**B**) Two loop types have been proposed to form during *E. coli* replication: trombone loops and leading-strand loops. Leading-strand loops mainly occur when the leading-strand polymerase stalls and the helicase continues. Priming loops, similar to those in T7, have not been detected so far.

Several strands of experimental evidence now suggest that leading- and lagging-strand synthesis are not functionally coupled, but instead the replication cycle is dominated by independent and stochastic polymerase action. *E. coli* rolling-circle replication assays conducted in the absence of primase showed no difference in leading-strand synthesis kinetics (Graham et al., 2017), consistent with earlier work showing no effect on the progression of the leading-strand polymerase upon blocking lagging-strand synthesis (McInerney and O'Donnell, 2004). Furthermore, single-molecule analysis of individual *E. coli* replisomes has revealed only small speed differences between combined leading- and lagging-strand synthesis, compared to leading-strand synthesis alone (Graham et al., 2017; Yao et al., 2009). These findings suggest that stochastic polymerase action

plays a major role in supporting a similar duplication rate on each daughter strand in spite of the asymmetry in synthesis mechanism. Pauses in leading-strand synthesis and a variable helicase unwinding rate further enforce synchronization without the need for direct communication.

1.3.2 Variation in polymerase number may enhance coordination

Rapid dilution experiments conducted by Bruce Alberts and others in the 1980s and 1990s suggested the core replisome complex is highly stable and capable of conducting processive synthesis in the absence of excess components (Alberts et al., 1983; Debyser et al., 1994; Kim et al., 1996b). This was a striking discovery at the time because it suggested that the complex acrobatics of priming, polymerase loading and OF synthesis on the lagging strand are coordinated by a single complex that continually recycles factors. This observation favored the idea that replisome function is guided by a highly refined set of rules that enforce a uniform enzymatic cycle. The concept triggered intensive efforts by many groups to clarify the molecular wiring that orders enzymatic events during replication, which ultimately led to numerous mechanistic disputes. In particular, the number of polymerases that reside at the replication fork and the trigger for polymerase cycling during OF synthesis have remained outstanding issues. Recent single-molecule observations have fundamentally altered our understanding of these events by revealing frequent polymerase exchange at active replication forks. Incorporating the possibility of polymerase exchange into these molecular scenarios opens up many additional pathways to achieve coordination.

Variation in the number of polymerases that reside at active replication forks has been observed depending on experimental conditions and the replication system studied. Clearly, at least two polymerases are required to synthesize the two daughter strands simultaneously, however, the presence of additional polymerases may enhance lagging-strand synthesis (Georgescu et al., 2011). In the case of phage T7, the polymerases are tethered to the replisome through the hexameric helicase that consists of six subunits, each containing polymerase binding sites (Gao et al., 2019). Consistent with this stoichiometry, single-molecule rolling-circle replication assays demonstrated up to six polymerases bound to the helicase (Geertsema et al., 2014). In contrast, the *E. coli* polymerase Pol III is tethered to the replisome through tight association with the τ subunits of the clamp loader (Jergic et al., 2007). Clamp loader is a heteropentameric complex composed of four possible subunits $\delta\delta'(\tau/\gamma)_3$. The τ and γ subunit are different translational frame shift variants of the same gene. Only τ, the full-length version (Flower and McHenry, 1990), can support binding of Pol III core. Therefore, only up to three polymerases can be simultaneously bound to the clamp loader complex. Consistent with this stoichiometry, single-molecule and *in vivo* measurements have revealed three polymerases both in *E. coli* (Reyes-Lamothe et al., 2010) and *B. subtilis* (Liao et al., 2016). The additional polymerase may allow for alternating action of two polymerases on the lagging strand (Georgescu et al., 2011; Lia et al., 2012; McInerney et al., 2007;

Reyes-Lamothe et al., 2010), providing alternative synchronization pathways. Nevertheless, there are still differing views on the polymerase stoichiometry and only two polymerases are used in many mechanistic studies (Alberts et al., 1983; Graham et al., 2017; Yeeles and Marians, 2011). The large uncertainty in the *in vivo* estimates allows for the possibility of a mixture of two and three polymerase replisomes (Liao et al., 2016; Reyes-Lamothe et al., 2010).

1.3.3 Frequent polymerase exchange supports alternative pathway sampling

In contrast to the long-standing view that replisome architecture is static, single-molecule fluorescence studies have demonstrated dynamic exchange of polymerases on the seconds to minutes timescale (Geertsema et al., 2014; Lewis et al., 2017a; Loparo et al., 2011; Monachino et al., 2020). Interestingly, studies conducted with the phage T7 replisome revealed that the number of polymerases at the replication fork inversely correlates with priming frequency: Up to six polymerases remain bound to the helicase under low priming conditions, whereas two to three polymerases were detected under frequent priming conditions (Geertsema et al., 2014; Loparo et al., 2011). Taken together with observations conducted using a two-bead flow stretching assay (Figure 1.1C), these findings support a revised coordination mechanism for T7 replication involving multiple dynamic pathways and DNA looping events (summarized in Figure 1.3A). In contrast to past proposals, the formation of priming loops was observed more frequently than trombone loops, with only a small fraction of lagging-strand synthesis occurring within trombone loops. Fluorescence imaging experiments revealed that, in addition to exchange (Figure 1.1E), polymerases are released onto the lagging strand consistent with completion of OF synthesis behind the replisome.

Frequent polymerase exchange (Lewis et al., 2017a; Monachino et al., 2020) and uncoupled lagging-strand synthesis (Graham et al., 2017) have also been detected during *E. coli* replication. In contrast to the exchange of individual T7 polymerases, in *E. coli,* entire Pol III* (Pol III holoenzyme lacking β clamp) complexes exchange (Beattie et al., 2017; Lewis et al., 2017a; Monachino et al., 2020). Consistent with the variation seen in T7 polymerase number, more than one Pol III* subassembly has been detected at the replication fork (Lewis et al., 2017a) with its number being controlled by the state of DnaB-DnaG interaction (Monachino et al., 2020). The observation of dynamic polymerase exchange and uncoupled lagging-strand synthesis suggest there is no single trigger that regulates the end of OF synthesis (Hacker and Alberts, 1994; Wu et al., 1992). Instead, a stochastic sampling of different OF synthesis pathways is most consistent with experimental observations.

These new findings suggest that while the replisome remains stable in the absence of excess components in solution, the addition of competing factors can drive dynamic exchange. Such a phenomenon has also been observed for the β clamp, which can be recycled in the absence of excess clamps (Tanner et al., 2011), but is left behind the

replication fork in the presence of excess clamps (Moolman et al., 2014; Stukenberg et al., 1994; Yuzhakov et al., 1996). Describing the replisome as a stable, dynamic machine might sound like an oxymoron, however, one does not exclude the other. A multisite exchange mechanism may provide an explanation for how replisomes can retain stability upon dilution, yet undergo rapid exchange in the presence of excess components (Aberg et al., 2016). Moreover, such a mechanism provides an explanation for the results of early observations, which have long promoted the idea of a static replisome assembly. These were based on rapid dilution experiments, which clearly favor recycling pathways (Alberts et al., 1983) and electron microscopy (EM) experiments (Chastain et al., 2003; Park et al., 1998) using crosslinkers to capture only a subset of possible pathways.

1.3.4 A dynamic replisome architecture ensures robustness

The dynamic nature of replisomes represents an essential feature to deal with rapidly changing demands during replication. In the cell, replication is challenged by a variety of factors, which can generally be divided into two classes: DNA bound proteins and DNA damage. DNA bound proteins, like the transcription machinery, interfere with continuous replisome progression. Indeed, an *in vivo* single-molecule study suggests that DNA replication is actually a predominantly discontinuous process reflecting replisome distortion upon replication-transcription conflicts, shown to occur multiple times per cell cycle (Mangiameli et al., 2017). In *E. coli*, DNA replication under unchallenged conditions is mainly performed by Pol III. However, when Pol III encounters damaged DNA, replisome progression pauses, allowing the replisome to reshape and bypass the obstacle (Indiani et al., 2009).

In the case of a leading-strand lesion, the replisome is able to reinitiate synthesis downstream at a new primed site (Yeeles and Marians, 2011, 2013). Upon more severe DNA damage, *E. coli* utilizes specialized polymerases Pol II, IV and V, also known as translesion synthesis (TLS) polymerases (Bonner et al., 1990; Tang et al., 1999; Wagner et al., 1999). TLS polymerases can also replace Pol III on β clamp, forming stochastic alternative replisomes (Kath et al., 2014), which are favored upon DNA damage due to higher abundance of TLS polymerases (Indiani et al., 2009). Notably, during TLS, helicase unwinding is uncoupled from DNA synthesis on either strand. In *E. coli*, the τ subunit of the clamp loader tightly coordinates helicase, leading- and lagging-strand synthesis (Figure 1.2). In contrast, TLS polymerases do not interact with the clamp loader and one model for TLS suggests that clamp loader also dissociates from the replisome along with Pol III. Hereby, again an alternative replisome forms containing a different clamp loader with all τ replaced by γ subunits (a shorter translational frame shift variant of τ, see chapter 1.3.2), which could perform clamp loading (Flower and McHenry, 1990) but does not contain Pol III or DnaB binding domains (Gao and McHenry, 2001), enabling uncoupled DNA synthesis. These scenarios emphasize the importance of broad replisome plasticity including the ability of fast exchange of components.

1.4 Staged assembly ensures robust eukaryotic DNA replication

The initiation process in eukaryotes is divided in two major steps: origin licensing and origin firing which separately happen in G1 and S phase of the cell cycle, respectively (Blow and Laskey, 1988; Diffley et al., 1994). During origin licensing, the replicative helicase is loaded on origins. In contrast to bacteria, eukaryotes have multiple origins. The genomes of pro- and eukaryotes differ a lot, with the latter ones being larger and more complex. For instance, the genome size of *E. coli* and humans is ~5x10^6 and ~6x10^9 base pairs (bps), respectively (Blattner et al., 1997; Lander et al., 2001; Venter et al., 2001). Combined with the massively reduced DNA replication rate in eukaryotes (~1000 and ~25 nucleotides (nt)/s in *E. coli* and humans, respectively (Conti et al., 2007; Kim et al., 1996a)), it becomes apparent that multiple origins per chromosome are required to duplicate the eukaryotic genome in a timely manner, decreasing the total time from 15 days (one origin per chromosome) to ~8 hours.

Eukaryotic origins seem to be more defined by chromatin organization and epigenetic marks rather than on a primary DNA level (Ganier et al., 2019; Prioleau and MacAlpine, 2016). Origins in *S. cerevisiae* represent an exception as they show well-defined, called autonomously replicating sequences (ARS) (Stinchcomb et al., 1979), serving as replicator according to the replicon model (Jacob et al., 1963). Each ARS is composed of four key elements: an 11-17 bps long ARS consensus sequence and the B1 element which are both recognized by the origin binding protein (Li et al., 2018; Yuan et al., 2017); a B2 element that facilitates helicase loading; and finally, a B3 element to position nucleosomes around the otherwise nucleosome free ARS (Marahrens and Stillman, 1992). ARS are recognized by the initiator, heterohexameric origin recognition complex (ORC), which subsequently recruits the helicase loader Cdc6 to close a hexameric ring around dsDNA (Figure 1.4A).

Contrary to bacterial DnaA, the ORC-Cdc6 complex has no DNA unwinding activity, instead its primary role is helicase loading. The eukaryotic helicase is a heterohexameric complex with a defined order of minichromosome maintenance (Mcm)2-7 subunits (Li et al., 2015). Mcm2-7, henceforth MCM, helicases, are built of an N- and C-terminal tier (N- and C-tier) of which the C-tier is responsible for ATP-driven DNA unwinding (Eickhoff et al., 2019). To load MCM on dsDNA, the ring must be in an open state, which is widely assumed to occur spontaneously between subunit 2 and 5 (Mcm2/5 gate) followed by Cdt1 stabilization (Frigola et al., 2017; Zhai et al., 2017). Hence, two MCM complexes are loaded around dsDNA in a final head-to-head double-hexamer (DH) configuration stabilized by N-tier interactions. This complex is known as the pre-replicative complex (pre-RC) (Evrin et al., 2009; Remus et al., 2009).

Origin spacing and usage varies with changing demands throughout development in higher eukaryotes (Rausch et al., 2020). Although all licensed origins in the genome represent potential origins, only a small portion is activated in a temporally controlled

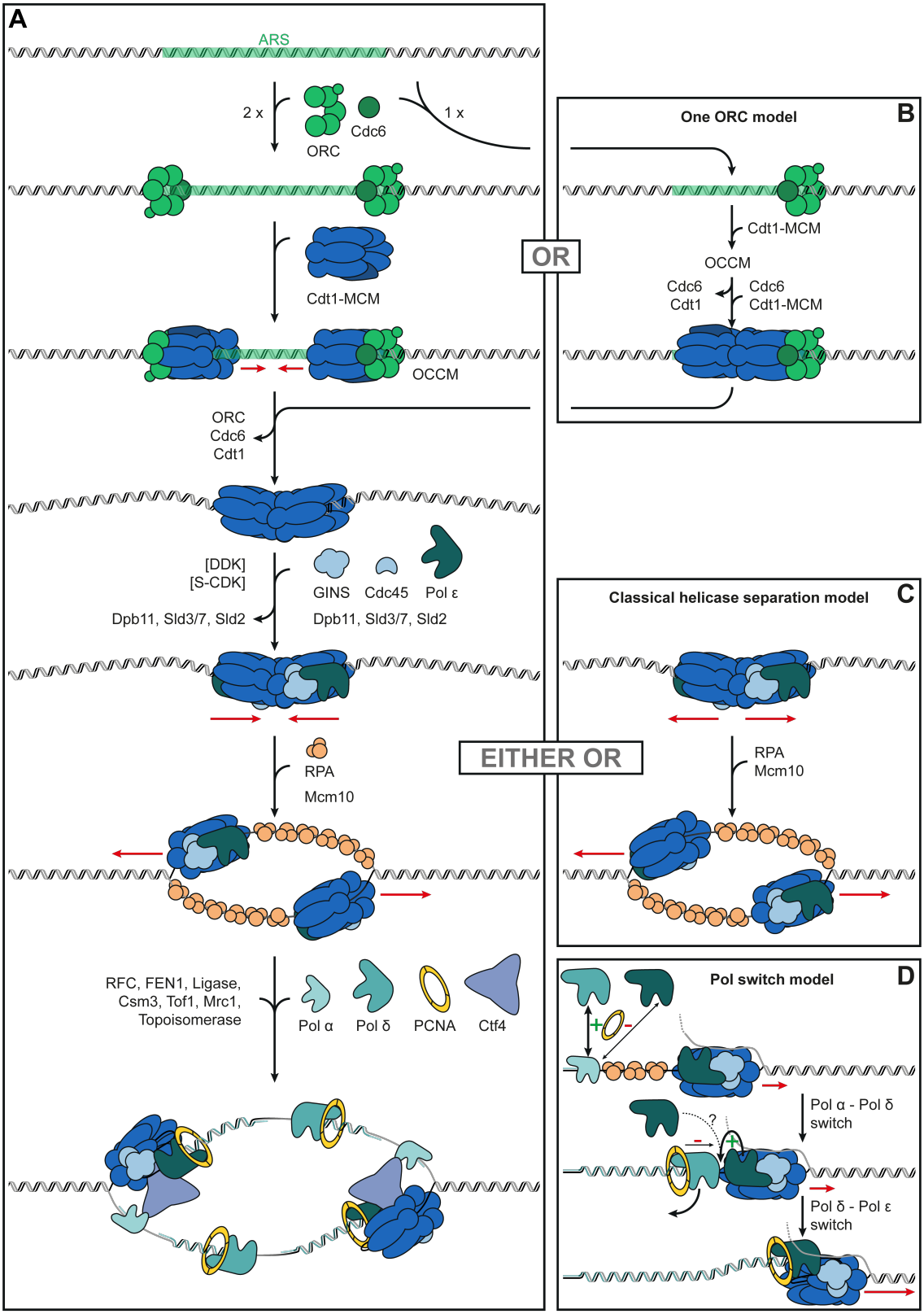

Figure 1.4. Eukaryotic DNA replication

(**A**) Eukaryotic replication begins in G1 phase with licensing of origins, recognition by two ORCs, and loading of two MCM helicases, which form a head-to-head double-hexamer. Next, during S phase, two kinases, DDK and S-CDK, stimulate CMG formation, dsDNA melting, and translocation. This leads to recruitment of additional replisome components including the ssDNA-binding protein RPA and three replicative

(legend continued on next page)

polymerases Pol ε, Pol δ, and Pol α, which perform leading-strand, lagging-strand, and primer synthesis, respectively. There is no compelling evidence strongly supporting trombone loop formation. (**B**) An alternative helicase loading mechanism in which one ORC sequentially loads two MCM hexamers. (**C**) Classical helicase separation model in which the C-tier of MCM is oriented towards dsDNA. Intrinsic structural features of MCM either allow model in (**A**) or (**C**) to be correct. (**D**) Mechanism of polymerase switching during replisome assembly. After priming by Pol α, Pol δ is switched in to continue synthesis, then switched out to allow for leading-strand synthesis by Pol ε stably associated with CMG.

manner (early and late origins). Based on the frequency of origin usage, origins are divided into active, flexible and dormant origins. The choice of origins seems to follow a stochastic-competition model in which origins compete for the limited amount of firing factors in the cell (Mantiero et al., 2011). Flexible, dormant and late origins play an important role in maintaining genome integrity as replication forks can stall or collapse, for instance due to obstacles on DNA (DNA damage, protein barriers). By subsequent firing of a not yet activated, neighboring origin, complete genome duplication is ensured. The large excess of MCMs compared to the number of origin sites further complicates classification and has become known as the “MCM paradox” (Burkhart et al., 1995; Edwards et al., 2002; Lei et al., 1996; Mahbubani et al., 1997). The observation of repetitive loading and spreading of MCMs from origin sites provides an explanation (Douglas et al., 2018; Edwards et al., 2002; Harvey and Newport, 2003; Powell et al., 2015). Consistent with the MCM paradox, further studies indicate that one ORC might be competent for multiple rounds of MCM loading (Bowers et al., 2004). However, the mechanism of MCM redistribution in the context of chromatin remains mysterious.

The MCM DH is an inactive helicase in which the N-tier shows a twisted and tilted form to the C-tier (Figure 1.4A) (Li et al., 2015). Helicase activation involves separation of the DH, dsDNA melting and exclusion of one strand to allow ssDNA translocation in 3′-5′ direction along the leading strands. Two S phase kinases play critical roles during this process: Dbf4-dependent kinase (DDK) and S phase cyclin-dependent kinase (S-CDK). First, DDK phosphorylates Mcm4/6 subunits, which exposes a binding site for Sld3 (in complex with Sld7). Cdc45 is then recruited to MCM, presumably by Sld3 (De Jesus-Kim et al., 2021). Second, S-CDK phosphorylates Sld2 and Sld3 which enables binding to Dpb11. Subsequently, the tetrameric complex GINS and Pol ε are recruited, forming the full 11-subunit replicative helicase CMG (Cdc45-Mcm2-7-GINS).

In eukaryotes, re-replication at a few but not all origins would lead to partially over-replicated chromosome regions with subsequent chromosome segregation inducing abnormal chromosomal rearrangements, imposing a threat to genomic integrity. Thus, to ensure that the whole genome is only replicated once during the cell cycle, eukaryotes tightly regulate the initiation process by restricting helicase loading and activation into G1 and S phase of the cell cycle, respectively (Blow and Laskey, 1988; Diffley et al., 1994). Thereby, S-CDK plays a major regulatory role: 1) Cdc6 phosphorylated by S-CDK is ubiquitylated and degraded (Drury et al., 2000); 2) Mcm3 phosphorylation excludes non-bound MCM from the nucleus (Nguyen et al., 2000) and 3) phosphorylated ORC is unable to load MCM to origins (Frigola et al., 2013). In higher eukaryotes, additional

mechanisms are present, for instance geminin (which is not present in *S. cerevisiae*) inhibits re-replication by directly targeting Cdt1 (McGarry and Kirschner, 1998; Wohlschlegel et al., 2000). Interestingly, studies show that tumor cells are more sensitive to re-replication than healthy cells (Vaziri et al., 2003; Zhu and Depamphilis, 2009).

Following helicase activation and CMG formation, the ssDNA-binding protein RPA and three different polymerases are recruited to the replication fork: Pol ε and Pol δ, which are thought to perform leading- and lagging-strand synthesis, respectively (Burgers and Kunkel, 2017), and Pol α which performs priming. OFs produced during lagging-strand synthesis are processed similarly as described for the bacterial system. Pol δ displaces the 5′ end of the preceding OF, generating a 5′ flap. Flap endonuclease 1 (FEN1) removes the displaced strand (synthesized DNA by error prone Pol α) which is finally ligated by DNA ligase. Proliferating cell nuclear antigen (PCNA) is a homotrimer and can interact with Pol δ, FEN1 and DNA ligase, suggesting a toolbelt model in which all three proteins could be simultaneously attached to PCNA (Kornberg and Baker, 1992).

Termination in eukaryotes occurs when two opposing forks encounter, thus most termination cites are chosen stochastically. CMG disassembly is the key step during DNA replication termination. Polyubiquitylation of Mcm7 by the E3 ubiquitin ligase SCF^{Dia2} triggers Ufd1-Npl4-dependent Cdc48 recruitment (Maric et al., 2014; Maric et al., 2017). Cdc48 is a segregase that finally drives CMG disassembly, thus CMG and associated leading-strand machinery can dissociate. SCF^{Dia2} constitutively travels with the replication fork, raising the question how CMG disassembly is restricted to termination. Single-molecule studies suggested one possible mechanism arguing that an elongation-specific DNA structure, which is lost upon fork encounter, inhibits CMG ubiquitylation (Low et al., 2020).

While in bacteria, the clamp loader acts as central hub in the replisome by tethering Pol III on both strands to the helicase, CMG plays a crucial role in organizing the eukaryotic replisome (Gambus et al., 2006; Gambus et al., 2009). Direct association of Pol ε to CMG ensures robust leading-strand synthesis (Georgescu et al., 2017; Georgescu et al., 2014a; Langston et al., 2014; Sengupta et al., 2013; Sun et al., 2015; Zhou et al., 2017) and three central complexes have been implicated in coordinating daughter-strand synthesis. First, Ctf4, a homotrimeric complex, has been shown to couple CMG and Pol α (Gambus et al., 2009; Kang et al., 2013; Simon et al., 2014; Sun et al., 2015). Second, Mcm10 has a versatile role in origin melting, CMG assembly/stimulation, and Pol α recruitment to CMG (Perez-Arnaiz et al., 2017). Finally, Csm3/Tof1/Mrc1 may link helicase to polymerase activity (Baretic et al., 2020; Cho et al., 2013; Gambus et al., 2006; Katou et al., 2003) and have been shown to simulate replication to *in vivo* rates both in bulk and single-molecule assays (Lewis et al., 2017b; Yeeles et al., 2017). Taken together, these observations highlight the numerous interactions between CMG and other replisome components that can serve to modulate replication fork progression. Single-molecule studies of bacterial replication strongly hint that dynamic exchange and a constantly evolving contact network underlie this

modulation. In fact, recent pioneering single-molecule studies of eukaryotic replication revealed frequent, availability-dependent polymerase exchange (Lewis et al., 2020) but further studies are needed to elucidate the unique importance of each core subassembly and the global architecture of the eukaryotic replisome.

1.4.1 Helicase loading and activation may involve several alternative pathways

The sequence of molecular events that underlie loading and activation of two MCM helicases at the origin of replication remains a central area of inquiry. Models involving either independent or simultaneous loading by one or two ORCs have been proposed based on various strands of experimental evidence (Abid Ali et al., 2016; Coster and Diffley, 2017; Evrin et al., 2009; Frigola et al., 2013; Remus et al., 2009; Ticau et al., 2015; Yuan et al., 2016; Yuan et al., 2017). Single-molecule fluorescence studies conducted using CoSMoS have demonstrated that a single ORC-Cdc6 can recruit one Cdt1-MCM complex to form an intermediate OCCM (ORC-Cdc6-Cdt1-MCM) (Ticau et al., 2015), which was also visualized by cryo-EM (Yuan et al., 2017). During DNA association, the Mcm2/5 gate is open and ring closure requires ATP hydrolysis, accompanied by sequential Cdc6 and Cdt1 release (Ticau et al., 2017). In these assays, recruitment of a second Cdt1-MCM complex was then observed in a Cdc6-dependent manner with additional studies revealing a conserved motif in Mcm4 being essential for stable MCM DH formation (Figure 1.4B) (Champasa et al., 2019).

Another prominent model for helicase loading proposes that two ORCs are required and each ORC loads one MCM (Figure 1.4A). Support for this model derives from the observation that mutations in the Mcm3 C-terminus which abolish the interaction with ORC, also prevent loading of a second MCM (Frigola et al., 2013). A study where the endogenous ARS was altered, lends further support, suggesting the MCMs are loaded quasi-symmetrically on distant sites (Coster and Diffley, 2017) followed by translocation to form the DH. Moreover, a model based on ORC dimerization was suggested (Amin et al., 2020). The reason for the discrepancy with single-molecule observations, which suggest a one ORC model, is still debated, but may reflect an alternative pathway that depends on experimental conditions and origin sequence. Recently, in an elegant study Thomas Miller and coworkers partially reconciled these contradicting models of MCM DH formation (Miller et al., 2019). By using time-resolved EM, they identified additional intermediates before (OC-MC complex) and after (MO and MOC-MC complex) OCCM formation. These new intermediates revealed that both MCMs are in fact recruited via the same interaction between ORC and the C-tier of MCM. Although they could identify origins containing two inverted OCCMs or intermediates thereof, as predicted by the quasi-symmetrical recruitment model (Figure 1.4A), this fraction was low. Instead, they found that upon loading of the first MCM, ORC is able to engage with the N-tier of MCM, promoting loading of the second MCM in an inverted orientation. Although these observations reconcile previous biochemical and single-molecule observations, whether the same ORC molecule recruits the second MCM and how different origins (with

different ORC affinities for the B2 element) affect those pathways remains elusive.

Intriguingly, another alternative, but still speculative pathway could exist under certain conditions. Although every ORC subunit is essential in yeast and fruit flies, two studies indicate that human cancer cell lines depleted to a non-detectable level of either Orc1 or Orc2 (Shibata et al., 2016) or Orc2 and Orc5 (Shibata and Dutta, 2020) are still viable. Orc1 and Orc2 are positioned at the ends of an open Orc1-5 ring and contain Cdc6 binding sites, suggesting that a pentameric ORC-Cdc6 could still form in those mutants. However, simultaneous Orc2/5 depletion would destabilize the other subunits, rendering an intact ORC-Cdc6 interaction unlikely. Alternatively, an unknown ORC-independent MCM recruitment pathway could exist, e.g. mediated by Cdc6 directly (Bell, 2017).

In contrast to bacterial replisomes where the dsDNA melting and unwinding functions are conducted by two different proteins, DnaA and DnaB, during eukaryotic replication these activities are both performed by the CMG (Duderstadt and Berger, 2013). However, the sequence of structural changes that occur within the CMG during these steps remains unclear. One mechanistic proposal relies on DNA kinking in the central channel of the MCM hexamer due to the twisted conformation of N- and C-tiers. This may trigger melting of dsDNA upon DH separation (Zhai et al., 2017). After initial DNA melting, strand exclusion by the CMG may involve ring opening and subsequent re-closure setting up a ssDNA translocation configuration (Ticau et al., 2017) for further DNA unwinding.

Although the head-to-head loading orientation of the MCM DH is widely accepted (Evrin et al., 2009; Remus et al., 2009), the orientation of DH separation remains less clear (Georgescu et al., 2017; Langston and O'Donnell, 2017). In the classical model of MCM separation, the C-tier faces the replication fork with the N-tier behind (Figure 1.4C). This model is supported by structural and FRET studies (Costa et al., 2014; Itsathitphaisarn et al., 2012; McGeoch et al., 2005; Rothenberg et al., 2007). In contrast, more recent work, including high-resolution structural analysis, suggests the N-tier faces the replication fork (Figure 1.4A) (Douglas et al., 2018; Georgescu et al., 2017; Lee et al., 2014). Thus, upon DH separation, the MCM helicases would pass each other translocating on opposing ssDNA strands, a suggested quality control mechanism to ensure that MCMs translocate correctly on opposing ssDNA strands (Douglas et al., 2018; Georgescu et al., 2017). Since strand engagement relies on intrinsic structural features of CMG, only one of the two orientation models can be correct. This stands in contrast to other molecular events at the replication fork, which are influenced by dynamic events and the evolving configuration of the replication machinery. The now accepted view that the MCM N-tier is facing the fork means that the N-terminus histone binding domain of Mcm2 is exposed to nucleosomes, which has been found to be implicated in nucleosome inheritance (Petryk et al., 2018). However, which dynamic events occur within the replisome to deal with nucleosomes in conjunction with chromatin remodelers and histone chaperones such as FACT is still highly investigated (Devbhandari et al., 2017; Gruszka et al., 2020; Kurat et al., 2017; Liu et al., 2020).

1.4.2 Conformational flexibility ensures robust helicase unwinding

Single-molecule approaches are playing a key role in distinguishing between several possible helicase unwinding mechanisms. Currently, three major helicase unwinding models are discussed: 1) steric exclusion (SE), 2) side channel extrusion (SCE) and 3) modified steric exclusion (MSE) model. In particular, studies have explored the consequence of placing obstacles on either the translocating or non-translocating strand using origin replication assays conducted with λ-phage DNA (Figure 1.1F) (Fu et al., 2011). Direct encounters between single helicases and steric blocks revealed that both *X. laevis* and *S. cerevisiae* CMG are only affected by blocks on the translocating strand and not the non-translocating strand (Fu et al., 2011; Kose et al., 2019; Langston and O'Donnell, 2017; Langston et al., 2017; Sparks et al., 2019), supporting the SE model. Mcm10 is thereby required to bypass blocks on the non-translocating strand (Langston et al., 2017) either by partial N-tier ring opening or a dynamic transition between different helicase unwinding modes. In addition to a classic steric exclusion mechanism, the SCE model suggests that dsDNA enters the central channel of the helicase where it is separated internally in two ssDNA strands. In this model, the second strand is extruded via a side channel. Although a small gap between Mcm2/6 subunits was proposed as potential side channel, to date this gap could not be confirmed directly (Zhai et al., 2017). However, a recent single-molecule study in fact revealed that CMG is able to switch between a ssDNA and dsDNA binding mode, suggesting a ssDNA gate within CMG different from the Mcm2/5 gate used during licensing (Wasserman et al., 2019). This to date unidentified gate might play a role in bypassing obstacles and also in CMG activation (chapter 1.4.1). Another alternative proposal (MSE model) involves dsDNA entering the N-tier and unwinding in the central channel with strand extrusion occurring out the same N-tier entry (Georgescu et al., 2017; Langston and O'Donnell, 2017). Consistent with this proposal, several studies, including single-molecule FRET experiments, revealed interactions between the emerging ssDNA strand and exterior parts of the helicase, which appear to modulate the helicase unwinding rate (Carney and Trakselis, 2016; Graham et al., 2011).

The ability to bypass blocks on DNA is another example of the highly dynamic nature of helicases which sample numerous alternative conformations. Structural studies conducted on hexameric helicases both from bacteria as well as eukaryotes are consistent with this idea revealing multiple closed and open ring conformations, with varying pore diameters as well as translocation on both ssDNA and dsDNA (Abid Ali et al., 2016; Gros et al., 2015; Itsathitphaisarn et al., 2012; Wasserman et al., 2019; Yuan et al., 2016). These observed ring dynamics provide alternative pathways that could lead to helicase stalling or direct bypass of blocks on either the translocating, non-translocating, or both strands (Yuan et al., 2016). The precise conformational changes that occur within the CMG that might underlie these pathways are not clear but they could occur in either the N-tier, C-tier, or both. In summary, further studies will be required to gain insight into the molecular details of CMG unwinding, the potential switching of CMG unwinding modes and how DNA blocks, which occur frequently in the cell, are overcome.

1.4.3 Stochastic events may shape the eukaryotic replication cycle

In contrast to bacteria where Pol III core preforms synthesis of both daughter strands, in eukaryotes three polymerases (Pol ε, Pol δ and Pol α) play distinct roles during leading- and lagging-strand synthesis. This additional functional specialization presents a coordination challenge: How is each polymerase targeted to the correct strand? Numerous regulatory pathways and exchange events appear to be critical for establishing the final polymerase configuration at the replication fork.

After initial CMG activation and unwinding, the leading- and lagging-strand polymerases, which are generally believed to be Pol ε and Pol δ, respectively, both require primers made by Pol α to start synthesis. Pol α initially synthesizes a short RNA strand which is then extended by DNA, forming an RNA-DNA primer. Two mechanisms are proposed to regulate DNA synthesis by Pol α: First, Pol α prefers A-form DNA shape (RNA-DNA hybrid) over B-form DNA (dsDNA), thus Pol α is released as DNA synthesis continues. Second, the clamp loader replication factor C (RFC) impedes Pol α-DNA association by loading PCNA. Only Pol δ and Pol ε, but not Pol α, are able to interact with PCNA. The frequency of primer synthesis by Pol α on the lagging strand as well as exact primer length is thereby not strictly defined. Subsequently, Pol ε and Pol δ, both in complex with PCNA, take over leading- and lagging-strand synthesis. Notably, PCNA is especially important for Pol δ as it increases its processivity 100-fold (Chilkova et al., 2007).

The prevailing view in the field is that eukaryotic replication proceeds with continuous synthesis of the leading strand and discontinuous synthesis of the lagging strand. However, a closer look at leading-strand synthesis also shows discontinuous synthesis, as the leading-strand polymerase is not always stably attached to DNA (Zhou et al., 2017). In the past, a variety of studies using biochemical and genetic techniques generated the widely accepted view that Pol ε performs leading-strand synthesis and Pol δ performs lagging-strand synthesis (Burgers and Kunkel, 2017). However, more recent *in vitro* studies suggest that after priming by Pol α, Pol δ establishes DNA synthesis on the leading and lagging strand (Figure 1.4D) (Aria and Yeeles, 2019; Yeeles et al., 2017). PCNA is loaded on primed ssDNA which favors Pol δ over Pol ε, thus excluding Pol ε from lagging- and initial leading-strand synthesis (Georgescu et al., 2014a). In line with these findings, Pol δ can take over leading-strand synthesis in a dysfunctional Pol ε background (Burgers et al., 2016; Devbhandari and Remus, 2020; Goswami et al., 2018; Johnson et al., 2015; Kesti et al., 1999). Additionally, structural evidence has emerged showing two conformations of Pol ε in complex with CMG, suggestive of a Pol δ-Pol ε switch on the leading strand (Georgescu et al., 2017; Sun et al., 2015; Zhou et al., 2017). In this model, initial leading-strand synthesis is uncoupled from CMG unwinding activity. As soon as Pol δ catches up with CMG, Pol δ is released and Pol ε takes over leading-strand synthesis. Hereby, the Pol δ-Pol ε switch could be triggered by a collision release mechanism. CMG only interacts with Pol ε but not with Pol δ, thus CMG excludes Pol δ from further leading-strand synthesis (Georgescu et al., 2014a).

Polymerase switching also plays a crucial role during TLS to overcome DNA damage or other obstacles on DNA. In *S. cerevisiae* there are three TLS polymerases (Pol η, Rev1 and Pol ζ) and additional ones in higher eukaryotes. TLS polymerases generally do not possess proofreading activity (except Pol II in *E. coli*), an essential feature to assure antibody diversity in higher eukaryotes. Due to their lack in proofreading activity, TLS is tightly regulated on different levels. First, the cellular concentration of TLS polymerases is regulated on a transcriptional and protein level. In *S. cerevisiae* TLS transcription is enhanced upon DNA damage (McDonald et al., 1997) but ubiquitylation seems to be the key regulator for TLS. For instance, ubiquitylated Pol η undergoes rapid proteasomal degradation, thereby restricting TLS (Skoneczna et al., 2007). Second, TLS is also controlled by recruitment of TLS polymerases to the DNA damage site in which ubiquitylation exhibits a stimulatory role. PCNA plays a central role as it recruits TLS polymerases through their PCNA-binding motifs (PIP-box). Upon DNA damage, PCNA is monoubiquitylated which tightens the TLS polymerase – PCNA interaction (Bienko et al., 2005; Hoege et al., 2002). Therefore, a TLS polymerase is able to replace Pol δ or ε in a moving replisome which can led to helicase-polymerase uncoupling (Devbhandari and Remus, 2020; Guilliam and Yeeles, 2020; Taylor and Yeeles, 2018).

The observation of frequent polymerase exchange during replication in bacteria, combined with emerging evidence that Pol δ can replace Pol ε, strongly hints that polymerase exchange dynamics may play an important role throughout the eukaryotic replication cycle. Single-molecule approaches will provide powerful avenues to directly investigate many alternative polymerase exchange pathways that support robust replication in eukaryotes (Lewis et al., 2020; Monachino et al., 2020).

1.5 Genomic DNA architecture regulates cellular functions

All living organisms are built by cells, the smallest living units, with their numbers ranging from as little as one up to trillions. Each individual cell relies on identical information encoded by their genomic DNA which, if stretched out, exceeds the cellular dimension by many folds. For instance, a human cell with a diameter of tens of micrometers contains approximately two meters of DNA (Hill, 2012). This means that an average human with around 37 trillion cells (Bianconi et al., 2013) contains a total DNA length spanning tens of times the distance from earth to sun (see chapter 1.1). To fit this remarkable amount of DNA into each individual cell, genomic DNA is highly compacted on several levels (Figure 1.5A). On a first level, DNA is wrapped around histone octamers, forming nucleosomal chromatin fibers (Kornberg, 1974). On a second level, chromatin fibers are further organized into loops or topologically associating domains (TADs), representing intramolecular interactions of distant regions within the chromosome (Dixon et al., 2012; Nora et al., 2012) as commonly identified by sequencing techniques like Hi- and Micro-C (Hsieh et al., 2015; Lieberman-Aiden et al., 2009). Besides DNA compaction, genome organization is implicated in regulatory roles in development, replication timing, gene regulation or V(D)J recombination in the process of B-cell maturation as part of the adaptive immune system (Flyamer et al., 2017; Kosak et al., 2002; Lupianez et al., 2015; Pope et al., 2014). At even larger scale, chromatin is organized into transcriptionally more active euchromatic A compartments and less active heterochromatic B compartments (Lieberman-Aiden et al., 2009).

1.5.1 SMC complexes fold genomic DNA via loop extrusion

Genome organization into loops and TADs is mediated by structural maintenance of chromosomes (SMC) protein complexes, which are thought to act via a process called loop extrusion (Davidson and Peters, 2021). In eukaryotes, loop extrusion is performed by the SMC proteins condensin and cohesin, acting in different cell cycle phases (mitosis and interphase, respectively). Direct evidence for the loop extrusion hypothesis was provided recently by *in vitro* studies using recombinant proteins, demonstrating active loop extrusion by condensin and cohesin at the single-molecule level (Davidson et al., 2019; Ganji et al., 2018; Kim et al., 2019).

Cohesin is a multisubunit protein complex, which is composed of a heterodimer of the long coiled-coil proteins SMC1-SMC3, the kleisin subunit SCC1/RAD21 and the “HEAT repeat proteins associated with kleisins” (HAWK) protein STAG1 or STAG2, together forming a ring-shaped structure (Figure 1.5B). Importantly, active forms of cohesin require additional, but mutually exclusive, regulatory HAWK subunits, called NIPBL and PDS5A/PDS5B. NIPBL is required to stimulate an ATPase activity that is associated with the SMC1-SMC3 heterodimer and is required for loop extrusion (Camdere et al., 2015; Davidson et al., 2019; Kim et al., 2019; Murayama and Uhlmann, 2014; Petela et al., 2018; Shi et al., 2020). In contrast, PDS5 proteins cannot stimulate cohesin’s ATPase

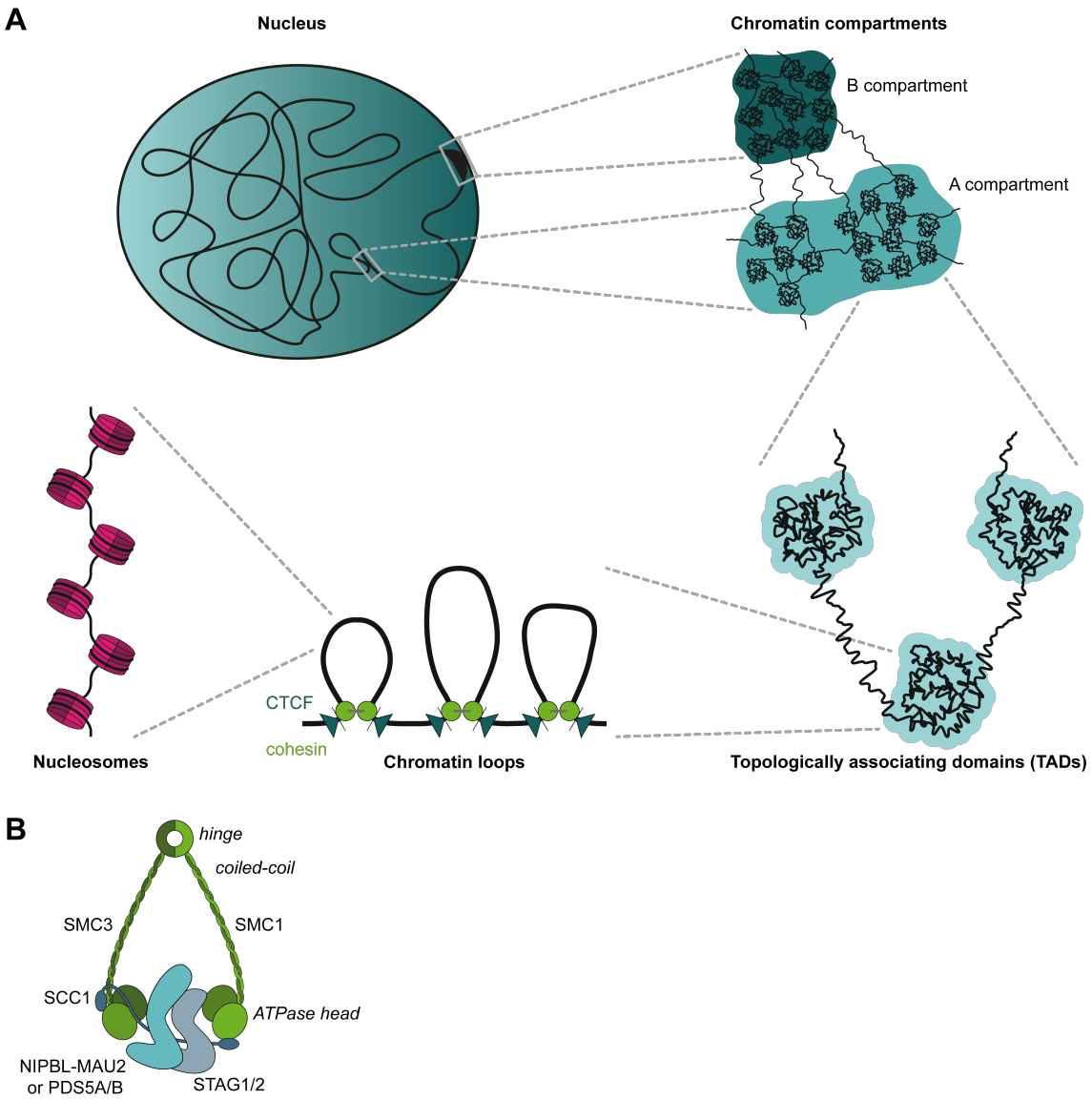

Figure 1.5. 3D genome organization
(**A**) Genomic DNA is organized hierarchically on several levels. First, all chromosomes are packaged by histones into nucleosomes and nucleosomal chromatin fibers. Chromatin fibers are further organized into chromatin loops and topologically associating domains (TADs). In eukaryotes, organization into loops and TADs is mediated by structural maintenance of chromosomes (SMC) protein complexes, in interphase by the cohesin complex. TADs are bordered by CCCTC binding factor (CTCF). At larger scale, chromatin is further organized into euchromatic A and heterochromatic B compartments, corresponding to transcriptionally more and less active regions, respectively. (**B**) Schematic of the human cohesin complex. Human cohesin is a multimeric complex, consisting of an SMC1-SMC3 heterodimer, the kleisin subunit SCC1 and HAWK protein STAG1 or STAG2, together forming a ring-shaped structure to which regulatory proteins (NIPBL and its binding partner MAU2 or PDS5A/B) can bind.

and loop extrusion activities, even though they are structurally related to NIPBL (Davidson et al., 2019; Murayama and Uhlmann, 2015; Petela et al., 2018). Instead, PDS5 proteins have roles in releasing cohesin from DNA and in allowing acetylation of cohesin by two acetyltransferases, called ESCO1 and ESCO2. Cohesin acetylation has been reported to stabilize cohesin on chromatin, both when it mediates loop formation,

but also when it connects replicated DNA molecules to generate sister chromatid cohesion (Davidson et al., 2019; Haering et al., 2008; Kim et al., 2019).

The mechanism of loop extrusion is unknown. Cohesin is known to exist in multiple conformations such as open rings and closed rods (Hons et al., 2016; Huis in 't Veld et al., 2014; Soh et al., 2015), and several speculative “walking”, “pumping” and “scrunching” models have been proposed to explain how alternations between these conformations could lead to loop extrusion (Hassler et al., 2018; Higashi et al., 2021; Yatskevich et al., 2019). Single-molecule techniques, single-molecule FRET in particular, will be key to elucidate the conformational cycle of loop-extruding cohesin in the future.

1.5.2 Loop extrusion is regulated and restricted by various mechanisms

The regulatory functions of loop extrusion, for example in gene regulation and V(D)J recombination, require that the process of genome organization into loops and TADs is well controlled. Cohesin-mediated loop extrusion is indeed regulated by various mechanisms, most importantly by controlling cohesin’s residence time on chromatin and by the presence of boundaries on DNA to loop extrusion. Loop extrusion initiates by formation of a small loop which expands dependent on cohesin’s processivity. The protein Wings apart-like (WAPL) was shown to act as an antagonist to loop extrusion by releasing cohesin complexes from chromatin, thus typically limiting the duration and length of loop formation to around 20 minutes in cells (Haarhuis et al., 2017; Wutz et al., 2017), however some long-lived loops were also observed (Vian et al., 2018). Interestingly, acetylation of the SMC3 cohesin subunit by the acetyltransferase ESCO1 protects cohesin from WAPL and thus increases the residence time on chromatin of a subpopulation of cohesin to hours in a CCCTC binding factor (CTCF)-dependent manner (Wutz et al., 2020).

CTCF is a DNA binding protein bound to specific sites throughout the genome at which cohesin accumulates and at which loop anchors and TAD boundaries are formed. CTCF is therefore thought to act as a boundary to loop extrusion mediated by cohesin (Busslinger et al., 2017; Nora et al., 2017; Parelho et al., 2008; Rubio et al., 2008; Stedman et al., 2008; Wendt et al., 2008; Wutz et al., 2017). How CTCF functions as a boundary for loop extruding complexes is incompletely understood. Several mechanisms explaining how boundary factors act in general could be envisioned. The boundary could be a passive, physical barrier, which would function as a “roadblock” to extruding cohesin (Hypothesis I). Alternatively, a boundary could contain a binding site for cohesin, thus restricting loop extrusion by tethering cohesin to boundaries and/or protecting from WAPL-mediated release (Hypothesis II). Moreover, boundaries could actively inhibit loop extrusion by shielding essential DNA binding sites in cohesin, by preventing conformational changes, which are required for loop extrusion, or by inhibiting cohesin’s ATPase activity, e.g. by replacing NIPBL with PDS5 proteins (Hypothesis III). Finally, the

underlying mechanism could be a combination of several mechanisms or different boundary factors could use distinct mechanisms.

At loop anchors, CTCF binding sites are usually oriented convergently ("convergence rule"), suggesting that CTCF exerts its boundary function asymmetrically, with the CTCF N-terminus facing cohesin. In fact, several studies have shown that the N-terminus of CTCF is required for its boundary function (Nishana et al., 2020; Nora et al., 2020; Pugacheva et al., 2020). Furthermore, structural studies revealed binding sites in the N-terminus of CTCF for SCC1 and STAG2, which also mediate binding of cohesin to WAPL, indicating that the CTCF N-terminus blocks cohesin-mediated loop extrusion by protecting cohesin's release by WAPL (Li et al., 2020) (consistent with *Hypothesis II*). However, mutation of the cohesin binding site in the N-terminus of CTCF does not fully abrogate enrichment of cohesin at CTCF sites (Li et al., 2020). This suggests that additional mechanisms to block loop extrusion must be in place, consistent with the observation that CTCF, but also other factors such as RNA polymerase (RNAP) and nucleosomes (to a different extent), can be a physical barrier to cohesin that is diffusing on DNA *in vitro* (Davidson et al., 2016; Stigler et al., 2016) (supporting *Hypothesis I*).

Furthermore, other proteins may contribute to the formation of loop anchors and TAD boundaries because PDS5 proteins and ESCO1 were shown to also be required for anchoring loops and forming TAD boundaries at CTCF sites in cells (Wutz et al., 2020; Wutz et al., 2017). Similar observations have been made in yeast, in which no CTCF orthologue has been identified to date, but where the related Pds5 and Eco1 orthologues similarly regulate cohesin-mediated loop extrusion (Bastié et al., 2021; Dauban et al., 2020). Together, these observations indicate that CTCF may be required but not sufficient for determining where loops formed by cohesin become anchored. How PDS5 proteins and ESCO1 contribute to this function remains elusive, but one interesting possibility is that PDS5 proteins enable ESCO1-dependent acetylation of cohesin at CTCF sites and by doing so prevent release of cohesin by WAPL. Another interesting possibility is that cohesin acetylation by ESCO1 promotes exchange of NIPBL by PDS5 proteins (*Hypothesis III*), for example by stabilizing interactions between cohesin and PDS5 proteins, thereby inactivating cohesin's ATPase and thus stopping loop extrusion.

In summary, all factors bound to DNA are potential barriers to cohesin-mediated loop extrusion, thus cohesin must either possess mechanisms to circumvent these barriers or cohesin-mediated loop extrusion will be restricted. Given the variety of proteins operating and residing on chromosomes, including replisome components, the presence of additional, not yet identified barriers besides CTCF seems likely and is currently actively investigated.

1.6 Aims of this thesis

Metaphorically speaking, genomic DNA within a cell is reminiscent of a busy road as a variety of cellular machineries such as replication, transcription as well as repair and maintenance machineries compete for the same DNA to perform their cellular function. In the last years it has become more and more clear that these cellular machineries are adapted to cope with this exceedingly complex environment. By exploiting protein dynamics, dynamic composition and exchange, cellular machineries allow for parallelism and redundancy which ensures robustness of complex multistage pathways. As extensively outlined in the introduction (chapters 1.3 and 1.4), the replisome evolved to a highly plastic machinery with alternative assembly and operation pathways.

In eukaryotes, the temporal gap between origin licensing and firing provides a window of time during which dynamic events on the chromosome can influence origin specificity and replication initiation. To search large exposed chromatin regions for an origin, licensing factors are thought to use global diffusion interspersed with local sliding (Chen et al., 2014; Duzdevich et al., 2015; Mirny et al., 2009; Stracy et al., 2021). This trial-and-error based approach involves more failure than success, resulting in frequent non-specific encounters on short time scales, but how these dynamic events influence the origin licensing pathway is still poorly understood. Conflicts between transcription and origin licensing are one example of how local dynamics can have potentially disastrous and far reaching consequences on replication (Gros et al., 2015; Looke et al., 2010; Tanaka et al., 1994). Beyond the dynamics of individual factors, the rapidly evolving chromatin landscape can act to positively or negatively regulate these events (Azmi et al., 2017; Foss et al., 2019). Building a complete understanding and predictive models for how dynamic encounters regulate essential pathways, such as replication, require novel approaches that reveal the dynamics on the single-complex level with high temporal and spatial resolution.

To date, while conflicts between transcription and active replisomes have been extensively investigated, it remains completely unclear how origin licensing is altered upon encounters with RNAPs. Moreover, no single-molecule studies have investigated origin licensing intermediates in the context of chromatin, known to be a key regulator of the process. Numerous indirect observations have suggested that MCM hexamers can reposition to sites away from origins under altered salt conditions (Remus et al., 2009) or when directly challenged by multiple rounds of loading or otherwise (Douglas et al., 2018; Gros et al., 2015), but the stability and mechanics of these events is not known. Thus, the first aim of this thesis was to reconstitute origin licensing at the single-molecule level (chapter 3) and develop a single-molecule transcription assay (chapter 4.1). Next, based on these assays, the second aim of this thesis was to establish a combined single-molecule origin licensing-transcription assay, which allows for temporal and spatial resolution of the dynamics of origin licensing during encounters with RNAP and to extend these observations in the context of chromatin (chapters 4.2-4.5).

Besides the replication and transcription machinery, also other machineries operating on DNA can be impeded by factors bound to DNA. While folding the genome via loop extrusion, SMC complexes most likely encounter a variety of obstacles on chromatin in addition to the known predominant barrier CTCF (introduced in chapter 1.5). Potential additional barriers include RNAPs and nucleosomes but also other macromolecular machines. Although RNAPs were shown to be moving barriers to cohesin-mediated loop extrusion in eukaryotes (Busslinger et al., 2017; Heinz et al., 2018), how cohesin deals with all other obstacles on chromatin still remains mysterious. Given the variety of obstacles on chromatin, it is conceivable that some of these might affect loop extrusion and thus the overall genome architecture with implications for regulatory functions such as in development and gene regulation.

Chromosomes display a large excess of MCM DHs over replication initiation events, marking dormant origins. This is known as the “MCM paradox” (Burkhart et al., 1995; Edwards et al., 2002; Lei et al., 1996; Mahbubani et al., 1997). Thus, the third aim of this thesis was to elucidate whether licensed origins containing MCM DHs represent a novel barrier to cohesin-mediated loop extrusion and whether this would affect the overall genome architecture (chapter 5). To test this hypothesis, multiple complementary approaches were pursued. Spearheaded by the Tachibana lab (IMBA), the first approach was to utilize the oocyte-to-zygote transition (OZT) to study the effect of MCM loss on cohesin-mediated loop extrusion, loops and TADs as determined by single-nucleus Hi-C (snHi-C) (chapter 5.1). A second, more direct approach was to test whether MCM DHs are physical barriers to cohesin. To this end, another aim was to develop a single-molecule approach to directly visualize encounters between cohesin and MCM DHs on DNA and thus to determine the barrier function of MCM DHs for cohesin translocation *in vitro* (chapters 5.2 and 5.3). This is achieved by combining a previously described cohesin translocation assay (Davidson et al., 2016; Stigler et al., 2016) with the single-molecule origin licensing assay developed for aim 1 (chapter 3).

Chapter 2

Materials and Methods

2.1 Materials

2.1.1 Antibodies

Table 2.1. Antibodies

Antibody	Source	Identifier
Anti-Cdc6, mAb, mouse	Abcam	Cat# ab20150
Anti-Mcm4, mAb, mouse	Santa Cruz Biotechnology	Cat# sc-166036
Anti-Orc6, mAb, mouse	Stephen Bell Lab (MIT)	SB49
THE His Tag Antibody, mAb, mouse	GenScript	Cat# A00186

2.1.2 Chemicals, Peptides and Recombinant Proteins

Not listed chemicals were purchased from Sigma-Aldrich or Carl Roth.

Table 2.2. Chemicals, Peptides and Recombinant Proteins

Chemical / Peptide / Recombinant Protein	Source	Identifier
1 Kb Plus DNA Ladder	Thermo Fisher Scientific	Cat# 11578636
3-Aminopropyltriethoxysilane (APTES)	Carl Roth	Cat# 2328.1
3,4-Dihydroxybenzoic acid (PCA)	Sigma-Aldrich	Cat# 03930590; CAS: 99-50-3
AF488/LD555/LD655-T7 RNAP	This study	N/A
Amersham Protran Premium 0.45 NC nitrocellulose membrane	Cytiva	Cat# 10600013
Amicon Ultra-15 Centrifugal Filter units	Merck	Cat# UFC901024 & UFC905024
Amicon Ultra-4 Centrifugal Filter units	Merck	Cat# UFC801024 & UFC805024
Anti-FLAG M2 Affinity Gel	Sigma-Aldrich	Cat# A2220
ATP-γ-S	Sigma-Aldrich	Cat# 11162306001
Biotin-PEG-Succinimidyl Carbonate (MW 5000)	Laysan Bio	Cat# Biotin-PEG-SC-5000-1g
Calmodulin Affinity Resin	Agilent Technologies	Cat# 214303
Cdc6	(Frigola et al., 2013)	N/A
Cdt1-LD655-MCM	This study	N/A
Cdt1-LD655-MCM$^{Mcm3\text{-}YDF}$	This study	N/A
Cdt1-MCM	(Frigola et al., 2013)	N/A
CHT Ceramic Hydroxyapatite	Bio-Rad	Cat# 1582200
dNTP Bundle	Jena Bioscience	Cat# NU-1005S
DYNAL Dynabeads KilobaseBINDER Kit	Thermo Fisher Scientific	Cat# 60101
Exonuclease I	New England Biolabs	Cat# M0293L
Gemini 5 µm C6-Phenyl 110 Å LC	Phenomenex	Cat# 00D-4444-E0
Gibson Assembly Master Mix	MPIB Core Facility	N/A
Herring sperm DNA	Thermo Fisher Scientific	Cat# 15634017
HiLoad 16/600 Superdex 200 pg	Cytiva	Cat# 28989335
HiLoad 26/600 Superdex 200 pg	Cytiva	Cat# 28989336
HiLoad 26/600 Superdex 75 pg	Cytiva	Cat# 28989334
HisTrap HP 5 ml	Cytiva	Cat# 17524802
HiTrap Chelating HP 1 ml	Cytiva	Cat# 17040801
HiTrap Desalting 5 ml	Cytiva	Cat# 17140801
HiTrap Heparin HP 5 ml	Cytiva	Cat# 17040703
HiTrap SP HP 1 ml	Cytiva	Cat# 17115101
Janelia Fluor 549 HaloTag	Promega	Cat# GA1110
JF546-Cohesin^{STAG1}	(Davidson et al., 2019)	N/A
Lambda DNA	New England Biolabs	Cat# N3011S

LD555 maleimide mono-reactive dye	Lumidyne Technologies	Cat# LD555-MAL
LD555-CH(10)LPETGG peptide	This study	N/A
LD555-CoA	Lumidyne Technologies	Custom synthesis
LD555-H3 histone octamers	This study	N/A
LD555-ORC	(Ticau et al., 2015) & This study	N/A
LD655-CoA	Lumidyne Technologies	Custom synthesis
mPEG-Succinimidyl Carbonate (MW 5000)	Laysan Bio	Cat# MPEG-SC-5000-5g
Ni Sepharose HP beads	Cytiva	Cat# 17526801
NTP Bundle	Jena Bioscience	Cat# NU-1014L
ORC	(Frigola et al., 2013)	N/A
PageRuler Prestained Protein Ladder	Thermo Fisher Scientific	Cat# 26616
PfuUltra High-Fidelity DNA polymerase	Agilent Technologies	Cat# 600380
Phusion High-Fidelity DNA polymerase	New England Biolabs	Cat# M0530S
Pierce ECL Western Blotting Substrate	Thermo Fisher Scientific	Cat# 32106
Platinum SuperFi DNA polymerase	Thermo Fisher Scientific	Cat# 12351010
Poly FLAG Peptide	Bimake	Cat# B23112
PreScission Protease	Cytiva	Cat# 27084301
Protein Assay Dye Reagent Concentrate	Bio-Rad	Cat# 5000006
Protino Glutathione Agarose 4B	Macherey-Nagel	Cat# 745500.10
Protocatechuate 3,4-Dioxygenase from *Pseudomonas putida* (PCD)	(Senavirathne et al., 2018)	N/A
QIAquick Gel Extraction Kit	Qiagen	Cat# 28706
QIAquick PCR Purification Kit	Qiagen	Cat# 28106
Quick CIP	New England Biolabs	Cat# M0525S
Restriction endonucleases	New England Biolabs	N/A
RNase Inhibitor, Murine	New England Biolabs	Cat# M0314L
Sephacryl S-1000 SF Tricorn 10/300 GL	GE Healthcare	N/A
SFP synthase	(Yin et al., 2006)	N/A
Slide-A-Lyzer MINI Dialysis Device, 3.5K MWCO, 0.1 mL	Thermo Fisher Scientific	Cat# 69550
SNAP-Surface Alexa Fluor 488	New England Biolabs	Cat# S9129S
Sortase A	(Freiburger et al., 2015)	N/A
Streptavidin	Sigma-Aldrich	Cat# S4762
Superdex 200 Increase 10/300 GL	Cytiva	Cat# 28990944
SYLGARD 184 Silicone Elastomer Kit (Polydimethylsiloxane, PDMS)	VWR	Cat# DOWC634165S
SYTOX Orange Nucleic Acid Stain	Thermo Fisher Scientific	Cat# S11368
T4 DNA Ligase	New England Biolabs	Cat# M0202S
Trolox	Sigma-Aldrich	Cat# 648471; CAS: 53188-07-1
α-factor	MPIB Core Facility	N/A

2.1.3 Oligonucleotides

All oligonucleotides used in this study were synthesized by Eurofins Genomics (Ebersberg, Germany).

Table 2.3. Sequences of oligonucleotides

Oligonucleotide	Sequence (5′->3′)
MS_144	GGTTTTCCATGGGTTCCAGCCATCACCACCACCATCATGGCAGTTCGATGG ACAAAGATTGCGAAATGAAACG
MS_145	GGTTTTCTCGAGTCCCAGACCCGGTTTACCCAGAC
MS_170	GGTTTTGCTAGCTGAAACTGGCGCGTGAGATG
MS_171	GGTTTTGGCCGGCCGCCAGCGCCGTCAGTG
MS_173	GGTTTTGCTAGCGGGCAAAACTCAGCTCACC
MS_178	GGTTTTGGATCCACAATCAATCAAAAAGCC
MS_179	GGTTTTGGATCCACTCTAACAAAATAGCAAATTTC
MS_180	[PHO]CTAGTCCCGCGAAATTAATACGACTCACTATAGGGAGACCACAACGG TTTGCGAT
MS_181	[PHO]CGCAAACCGTTGTGGTCTCCCTATAGTGAGTCGTATTAATTTCGCGG GA
MS_182	GGTTTTGCGATCGCCCCGCGCAGACGATC
MS_183	GGTTTTCTCGAGCCGTGGAATGAACAATGGAAGTC
MS_184	GGTTTTACTAGTCATTGCCACGGGTAAAGTTGG
MS_200	[PHO]GGCCGGTGAGTGTGTTTGAGTTGATTTTGTGTGG
MS_201	[BIOTEG]CCACACAAAATCAACTCAAACACACTCACC
MS_202	[PHO]CTAGGGTGAGTGTGTTTGAGTTGATTTTGTGTGGGCGTCGGTAAGTG AGAGG
MS_203	CCACACAAAATCAACTCAAACACACTCACC
MS_204	[BIOTEG]CCTCTCACTTACCGACGC
MS_221	[PHO]CATGGATAGCCTGGAGTTCATCGCCTCGAAGTTAGCCGGCAGTTCCC ACCATCACCACCATCATGGTGGTGGTGGCTCTC
MS_222	[PHO]TCGAGAGAGCCACCACCACCATGATGGTGGTGATGGTGGGAACTGC CGGCTAACTTCGAGGCGATGAACTCCAGGCTATC
MS_223	GGTTTTGCTAGCCTGGAGAATCCCGGTGC
MS_224	GGTTTTGCTAGCACAGGATGTATATATCTGACACG
MS_226	[BIOTEG]GCTGCTGCGTGTGGATGAG
MS_227	CCGCGTGCCTGAGTGTTC
MS_231	GTGGTGCTCTGGGAAATCGACCTGACATGTGTCAGAGCACGTCC
MS_232	AAAATAACGTTCTCCACCGACCTCTGTCGACGGTGATGACGGTGAAAACC
MS_238	CCAAACAAACTGCCCGCAAATGCACGGGTGGCAAAGCACCGC
MS_239	GCGGTGCTTTGCCACCCGTGCATTTGCGGGCAGTTTGTTTGG
MS_257	[PHO]GATCTATAATGTCACCGCTTCCTGGTTTTAGAGCTAG
MS_258	[PHO]CTAGCTCTAAAACCAGGAAGCGGTGACATTATA

2.1.4 Recombinant DNA

Table 2.4. Recombinant DNA

Recombinant DNA	Source	Identifier
12ADE-B	Scott Gradia	Addgene Plasmid #48298
p850	(Gros et al., 2015)	N/A
pARS/WTA	(Marahrens and Stillman, 1992)	N/A
pBH161	(He et al., 1997)	N/A
pET29-Sortase 4M	(Freiburger et al., 2015)	N/A
pGEX-6P-1_Cdc6	(Frigola et al., 2013)	N/A
pML104	(Laughery et al., 2015)	Addgene Plasmid #67638
pMS145_SNAP-T7RNAP	This study	N/A
pMS173_ybbR-T7RNAP	This study	N/A
pMS184_H2A_H2B	This study	N/A
pMS185_H3_H4	This study	N/A
pMS186_H3-S11C_H4	This study	N/A
pMS194_Cas9-ADE2	This study	N/A
pMS200_Cas9-ADE2_scMcm3-guide9	This study	N/A
pMSuperCos151_empty	This study	N/A
pMSuperCos159	This study	N/A
pMSuperCos162_ARS1	This study	N/A
pMSuperCos165_T7P-ARS1	This study	N/A
pMSuperCos177_T7P-ARS1-Widom601	This study	N/A
pMSuperCos182_T7P-ARS1-Widom601-5xT7T	This study	N/A
pSNAP-tag (T7)-2	New England Biolabs	Cat# N9181S
pVP91A-pcaHG	(Knoot et al., 2015)	Addgene Plasmid #113766
Sfp pet29b C-terminal His Tag	(Worthington and Burkart, 2006)	Addgene Plasmid #75015
SuperCos1 cosmid vector	Agilent Technologies	Cat# 251301

2.1.5 Strains and Organisms

Table 2.5. Strains and Organisms

Strain / Organism	Genotype	Source	Identifier
E. coli BL21(DE3)	F⁻ *dcm ompT hsdS*(r_B^- m_B^-) *gal* λ(DE3)	Agilent Technologies	Cat# 200131
E. coli BL21(DE3) codon plus RIL	F⁻ *ompT hsdS*(r_B^- m_B^-) *dcm*$^+$ Tet^r *gal* λ(DE3) *end*A Hte [*argU ileY leuW* Cam^r]	Agilent Technologies	Cat# 230245
E. coli BL21(DE3) star	F⁻ *ompT hsdS*$_B$(r_B^- m_B^-) *galdcmrne*131 (DE3)	Thermo Fisher Scientific	Cat# C601003
E. coli dam$^-$*/dcm*$^-$	*ara-14 leuB6 fhuA31 lacY1 tsx78 glnV44 galK2 galT22 mcrA dcm-6 hisG4 rfbD1 R(zgb210::Tn10) TetS endA1 rspL136 (StrR) dam13::Tn9 (*Cam^R*) xylA-5 mtl-1 thi-1 mcrB1 hsdR2*	New England Biolabs	Cat# C2925H
E. coli DH5α	F⁻ Φ80*lacZ*ΔM15 Δ(*lacZYA-argF*) U169 *rec*A1 *end*A1 *hsd*R17(r_k^-, m_k^+) *pho*A *sup*E44 *thi*-1 *gyr*A96 *rel*A1 λ⁻	Thermo Fisher Scientific	Cat# 18265017
E. coli Stellar	F⁻, *end*A1, *sup*E44, *thi*-1, *rec*A1, *rel*A1, *gyr*A96, *pho*A, Φ80d *lacZ*Δ M15, Δ(*lacZYA-argF*) U169, Δ(*mrr - hsd*RMS - *mcr*BC), Δ*mcr*A, λ⁻	Takara Bio	Cat# 636763
yJF38	*W303-1a pep4::KanMx4 bar1::Hph-NT1 his3-11::HIS3pJF2 trp1-1::TRP1pJF3 leu2-3::LEU2pJF4 ura3-1::URA3pJF6*	(Frigola et al., 2013)	N/A
yMS1	*MATa ade2-1 can1-100 pep4::kanMX bar1::hphNAT1 (hygromycinB) Mcm3-YDF ura3::Mcm2-PGal1,10-Mcm3-YDF (URA3) his3-11,15::Gal4-PGal1,10-Cdt1 (HIS3) trp1-1::Mcm5-PGal1,10-Mcm4 (TRP1) leu2-3,112::Mcm7-PGal1, 10-ybbR-Mcm6-FLAG (LEU2, clonNAT)*	This study	N/A
ySA4	*MATa ade2-1 can1-100 pep4::kanMX bar1::hphNAT1 (hygromycinB) ura3::Mcm2-PGal1,10-Mcm3 (URA3) his3-11,15::Gal4-PGal1,10-Cdt1 (HIS3) trp1-1::Mcm5-PGal1,10-Mcm4 (TRP1) leu2-3,112::Mcm7-PGal1, 10-ybbR-Mcm6-FLAG (LEU2, clonNAT)*	This study (generated by Syafiq Abd Wahab, Remus Lab)	N/A
ySD-ORC	*MATa ade2-1 ura3-1 his3-11,15 trp1-1 leu2-3,112 can1-100 bar1::Hyg pep4::KanMX TRP1 Gal1-10 ORC5,ORC6 HIS3 Gal1-10 ORC3,ORC4 URA3 Gal1-10 CBP-TEV ORC1,ORC2*	(Frigola et al., 2013)	N/A
yST163	*ade2-1 ura3-1 his3-11,15 trp1-1 leu2-3,112/LEU2::ORC1 can1-100 MATa orc1::hisG lys2::pLys2Gal1,10ORC2,5 pep4::KanMX Orc4-6XGly-His6 tag(hph) his::pST015(GAL1,10-3xFlag-GGG-ORC1, ORC6)*	(Ticau et al., 2015)	N/A

2.1.6 Essential micromirror TIRF microscope parts

Table 2.6. Essential micromirror TIRF microscope parts

Microscope part	Source	Identifier
808 nm 250 mW single mode laser module	Lumics	Cat# LU0808M250
ET520/40m emission filter	Chroma	Cat# ET520/40m
iXon Ultra 888 EMCCD camera	Andor	Cat# iXon Ultra 888
Laser Box: OBIS	Coherent	Cat# 1228877
Micromirror TIRF (mmTIRF) system	Mad City Labs	Cat# MicroMirror TIRF System
OBIS 488nm LS 120mW laser	Coherent	Cat# 1277614
OBIS 532nm LS 120mW laser	Coherent	Cat# 1280765
OBIS 637nm LX 100mW laser	Coherent	Cat# 1193838
OptoSpin25 filter wheel	Cairn Research	Cat# Optospin25
OptoSplit II dual emission image splitter	Cairn Research	Cat# OptoSplit II
T635lpxr dichroic mirror	Chroma	Cat# T635lpxr
TIRF Lock	Mad City Labs	Cat# TIRF Lock
ZET532/640m emission filter	Chroma	Cat# ZET532/640m
ZT488rdc dichroic mirror	Chroma	Cat# ZT488rdc
ZT532rdc dichroic mirror	Chroma	Cat# ZT532rdc
ZT647rdc dichroic mirror	Chroma	Cat# ZT647rdc
ZET488/10x excitation filter	Chroma	Cat# ZET488/10x
ZET532/10x excitation filter	Chroma	Cat# ZET532/10x
ZET640/10x excitation filter	Chroma	Cat# ZET640/10x
Apo N TIRF 60 x oil-immersion TIRF objective (NA 1.49)	Olympus	Cat# UPLAPO60XOHR

2.1.7 Software and Algorithms

Table 2.7. Software and Algorithms

Software / Algorithm	Source	Identifier
Example BeanShell scripts for Micromanager control	This study	https://github.com/mjosch/Dobby-control
Fiji	(Schindelin et al., 2012)	https://imagej.net/Fiji
ImageJ	(Schneider et al., 2012)	https://imagej.nih.gov/ij/
Lasergene	DNASTAR	https://dnastar.com/software/lasergene/
Micromanager v1.4	(Edelstein et al., 2010)	https://micro-manager.org/
Molecule ARchive Suite (MARS)	(Agarwal and Duderstadt, 2020)	https://github.com/duderstadt-lab
Python scripts / Jupyter notebooks for analysis	This study	https://github.com/mjosch/Born-to-slide https://github.com/mjosch/MCM-LE-barrier
Single-molecule example movies	This study	https://osf.io/z3trb/ doi:10.17605/osf.io/z3trb
SnapGene	GSL Biotech	https://snapgene.com/

2.1.8 Additional equipment

Table 2.8. Additional equipment

Equipment	Source	Identifier
Avanti J-26S XP Centrifuge	Beckman Coulter	Cat# B22987
LAS-3000 imaging system	Fujifilm	N/A
NanoPhotometer NP80	IMPLEN	Cat# NP80
NGC	BioRad	N/A
Optima XL-100K ultracentrifuge	Beckman Coulter	N/A
Polyethylene tubing (0.58 x 0.96 mm)	A. Hartenstein	Cat# SX05
Precision coverslips #1.5H, 22x22 mm	Marienfeld	Cat# 0107052
SamplePrep 6875 Freezer/Mill	SPEX	Cat# 6875 Freezer/Mill
Sonopuls HD 2200	Bandelin	Cat# 2531
Syringe Pump	World Precision Instruments	Cat# AL1000-220
Typhoon FLA 9000	GE Healthcare	N/A
Zepto plasma cleaner	Diener Electronic	Cat# Model 2 base unit type A

2.2 Molecular biology

2.2.1 Polymerase Chain Reaction (PCR)

Polymerase chain reaction (PCR) was used to amplify specific DNA fragments including target genes, or to introduce specific mutations into a gene of interest (site-directed mutagenesis).

To amplify specific DNA fragments, Phusion High-Fidelity or Platinum SuperFi DNA polymerase were used. Site-directed mutagenesis was performed using PfuUltra High-Fidelity DNA polymerase. All reactions were performed according to manufacturer's instructions with annealing temperature and elongation time adjusted based on the theoretical primer melting temperatures and DNA length, respectively.

2.2.2 PCR-directed restriction cloning and Gibson Assembly

DNA fragments were inserted into plasmids by PCR-directed restriction cloning or Gibson Assembly. All subsequent steps were performed according to manufacturer's instructions unless specified otherwise. For both strategies, the target plasmid was linearized with suitable restriction enzyme(s) and optionally dephosphorylated with Quick CIP. PCR-amplified DNA fragments were also digested (PCR-directed restriction cloning) or used directly. Linearized plasmid and PCR-amplified DNA fragments were subsequently purified from an agarose gel or directly (using QIAquick Gel Extraction or QIAquick PCR Purification Kit). Final DNA concentrations were determined on a spectrophotometer (NP80).

For PCR-directed restriction cloning, 100 ng linearized vector DNA were mixed with a three times molar excess of digested insert DNA and 400 units (U) T4 DNA Ligase in a total volume of 20 µl 1x T4-DNA Ligase buffer and incubated at room temperature (RT) for 1 hour or at 16 °C overnight.

Gibson Assembly reactions were set up with 100 ng linearized vector DNA and a three times molar excess of insert DNA in a total volume of 20 µl containing 15 µl Gibson Assembly Master Mix and incubated at 50 °C for 75 min.

2.2.3 Transformation of *E. coli*

Foreign DNA was incorporated into chemically competent *E. coli* by heat shock transformation. Chemically competent cells were either purchased or generated as previously described (Green and Rogers, 2013). 50 µl of competent cells were mixed with 100 ng plasmid DNA or 2-5 µl of a ligation or Gibson assembly reaction (chapter 2.2.2) and incubated on ice for 30 min. Subsequently, a heat shock was performed at 42 °C for 45 s and the cells were immediately put on ice for another 2 min. The cells

were supplemented with 900 µl Luria-Bertani (LB) or Super Optimal broth with Catabolite repression (SOC) medium and shaken at 37 °C for 1 hour. Finally, cells were plated on LB agar plates containing a suitable antibiotic selection. Note that DH5α and Stellar cells were used for cloning and plasmid propagation, whereas BL21 derivatives were chosen for protein expression.

2.2.4 Plasmid DNA preparation from *E. coli*

To amplify plasmid DNA, a single transformed colony was picked and grown in liquid LB medium containing a suitable antibiotic selection. Plasmid DNA was subsequently isolated based on the principle of alkaline cell lysis (Birnboim and Doly, 1979). Based on the desired scale, plasmid DNA was purified with the QIAGEN Plasmid Mini or Maxi Kit according to the manufacturer's protocol. The final DNA concentration was determined on a spectrophotometer (NP80) by measuring the absorbance at 260 nm. DNA purity was assessed by 260/230 nm and 260/280 nm ratios.

2.2.5 Sequencing

All newly generated plasmids were ultimately verified by Sanger sequencing. Sequencing reactions were performed externally by Eurofins Genomics (Ebersberg, Germany) and samples were prepared according to the respective guidelines. Sequencing reads were analyzed using the Lasergene software package.

2.3 Protein purification and labeling

2.3.1 Reagent preparation for protein labeling with SFP synthase

SFP synthase recognizes a short peptide tag (here the “ybbR-tag” was used) to which it covalently attaches a CoA-functionalized probe through a phosphopantetheinyl linker, allowing site-specific labeling of proteins. SFP synthase was purified similarly as previously described (Yin et al., 2006) with an additional purification step. *E. coli* BL21(DE3) star (carrying Sfp pet29b C-terminal His Tag, a gift from Michael Burkart (Worthington and Burkart, 2006)) were grown in Terrific Broth (TB) medium and expression was induced with 0.1 mM IPTG at 18 °C for 18 hours. SFP synthase was first purified on a HisTrap HP 5 ml equilibrated in 20 mM Tris-HCl, pH 7.9, 500 mM NaCl, 5 mM imidazole, 1 mM DTT. SFP synthase was further purified on a HiLoad 16/600 Superdex 200 pg equilibrated in 50 mM HEPES-NaOH, pH 7.5, 150 mM NaCl, 10 % (volume per volume, v/v) glycerol, 1 mM DTT. Peak fractions were pooled, concentrated with a MWCO 10000 Amicon Ultra Centrifugal Filter unit, frozen in aliquots in liquid N_2 and stored at -80 °C. The final protein concentration was determined on a spectrophotometer (NP80) by measuring the absorbance at 280 nm. Protein purity was assessed by the 260/280 nm ratio as well as sodium dodecyl sulphate–polyacrylamide gel electrophoresis (SDS-PAGE) analysis and intact mass measurement.

LD555- and LD655-CoA-functionalized dyes were purchased from Lumidyne Technologies.

2.3.2 Reagent preparation for protein labeling with Sortase

A previously developed, highly efficient version of Sortase A (Δ2-59, P94S/D160N/D165A/K196T) obtained from *Staphylococcus aureus* was used and essentially purified as described elsewhere (Chen et al., 2011; Freiburger et al., 2015). Sortase A was expressed in *E. coli* BL21(DE3) (transformed with pET29-Sortase 4M) in ZYP-5052 auto-induction medium (Studier, 2005) at 18 °C. Cleared cell lysate was applied to a HisTrap HP 5 ml equilibrated in 50 mM Tris-HCl, pH 8, 500 mM NaCl, 0.02 % (v/v) NaN_3, 10 mM imidazole. Unspecific bound protein and nucleic acids were removed by washing with 50 mM Tris-HCl, pH 8, 0.02 % (v/v) NaN_3, 25 mM imidazole + 500 mM (wash 1) or 1000 mM (wash 2) NaCl. Sortase A was then eluted with 250 mM imidazole and dialyzed overnight against 50 mM Tris-HCl, pH 8, 150 mM NaCl, 1 mM DTT in the presence of TEV protease. The reaction mixture was incubated with Ni Sepharose HP beads to recover untagged Sortase A from the supernatant. Sortase A was finally purified via gel filtration chromatography on a HiLoad 26/600 Superdex 75 pg, equilibrated in 50 mM Tris-HCl, pH 8, 150 mM NaCl, 0.02 % (v/v) NaN_3. Monomeric Sortase A was pooled, concentrated with a MWCO 10000 Amicon Ultra Centrifugal Filter unit, frozen in aliquots in liquid N_2 and stored at -80 °C. The final protein concentration was determined on a spectrophotometer (NP80) by measuring the absorbance at 280 nm. Protein purity

was assessed by the 260/280 nm ratio as well as SDS-PAGE analysis and intact mass measurement.

To site-specifically label proteins at the N-terminus using Sortase A, the peptide with the sequence H2N-CH(10)LPETGG-COOH was synthesized by solid-phase peptide synthesis and conjugated with a maleimide mono-reactive dye similar as previously described (Ticau et al., 2015). To prevent oxidation, monomethoxytrityl-protected cysteine (Fmoc-Cys(Mmt)-OH) was used for peptide synthesis. To conjugate peptide and dye, 3000 nmols crude product from peptide synthesis were deprotected, mixed with 1000 nmols LD555 maleimide mono-reactive dye in DMSO at pH 7 and incubated at RT overnight. Fluorescently labeled, full-length peptide was purified by HPLC on a Gemini 5 µm C6-Phenyl 110 Å LC column using a gradient between 0.1 % (v/v) TFA in ddH_2O and 0.08 % (v/v) TFA in acetonitrile. Clean fractions containing full-length, labeled peptide (analyzed by mass spectrometry) were pooled, lyophilized and stored at -20 °C.

2.3.3 ORC and LD555-ORC purification

S. cerevisiae ORC expression and purification was based on previously described protocols (Frigola et al., 2013; Ticau et al., 2015). Unlabeled ORC was expressed in yeast strain ySD-ORC (CBP tag at the N-terminus of Orc1). To produce fluorescently labeled ORC, UbSORT-FLAG-ORC (tags at the N-terminus of Orc1) was expressed in yST163. Note that the N-terminal ubiquitin is cleaved off in the cells, exposing three N-terminal glycines on Orc1 for Sortase A recognition (SORT-ORC). Cells were grown in 8 l Yeast extract Peptone (YP) medium supplemented with 2 % (v/v) raffinose at 30 °C. At OD_{600} = 1.2, cells were arrested at G1 phase with α-factor at a final concentration of 150 ng/ml and incubated for another 3 hours. Subsequently, protein expression was induced by addition of 2 % (v/v) galactose and incubated for another 4 hours. Cells were harvested by centrifugation (4000 x g, 10 min), washed once with 25 mM HEPES-KOH, pH 7.6, 1 M sorbitol and resuspended in 0.5 packed cell volumes of buffer A (25 mM HEPES-KOH, pH 7.6, 0.05 % (v/v) Nonidet P40 Substitute, 10 % (v/v) glycerol, 1 mM DTT) + 500 mM KCl supplemented with 1x protease inhibitor cocktail (2 µM pepstatin, 2 µM leupeptin, 1 mM PMSF, 1 mM benzamidine, 1 µg/ml aprotinin) and frozen dropwise in liquid N_2. Frozen cells were lysed in a SamplePrep 6875 Freezer/Mill and subsequently mixed with 0.5 packed cell volumes buffer A + 500 mM KCl supplemented with 1x protease inhibitor cocktail. All subsequent steps were performed at 4 °C unless stated otherwise. Thawed cell lysate was cleared by ultracentrifugation (235000 x g, 60 min).

Cleared cell lysate from ySD-ORC cells was supplemented with 2 mM $CaCl_2$ and incubated with 1.5 ml bed volumes (BV) calmodulin affinity resin equilibrated in buffer A + 500 mM KCl for 3 hours. The resin was washed with 20 BV buffer A + 200 mM KCl, 2 mM $CaCl_2$ and ORC was eluted 10 times with 1 BV of buffer A + 200 mM KCl, 2 mM EGTA, 1 mM EDTA. Fractions were pooled and applied to a HiTrap SP HP 1 ml, equilibrated in buffer B (50 mM HEPES-KOH, pH 7.6, 5 mM $Mg(OAc)_2$, 0.01 % (v/v)

Nonidet P40 Substitute, 10 % (v/v) glycerol, 1 mM DTT) + 200 mM KCl. The column was washed with 10 column volumes (CV) buffer B + 200 mM KCl and ORC was subsequently eluted with buffer B + 500 mM KCl. Peak fractions were pooled and further purified on a Superdex 200 increase 10/300 GL gel filtration column equilibrated in buffer B + 300 mM potassium glutamate (KGlu). Fractions containing stoichiometric ORC were pooled, concentrated with MWCO 50000 Amicon Ultra Centrifugal Filter unit, frozen in aliquots in liquid N_2 and stored at -80 °C.

Cleared cell lysate from yST163 cells was incubated with 2 ml BV Anti-FLAG M2 Affinity Gel equilibrated in buffer A + 500 mM KCl for 3 hours. The resin was washed with 20 BV buffer A + 200 mM KCl and SORT-ORC was eluted 5 times with 1 BV buffer A + 200 mM KCl + 0.15 mg/ml poly FLAG peptide. Elution fractions were pooled and further purified on a HiTrap SP HP 1 ml as described above. To produce LD555-ORC, SORT-ORC was diluted 2-fold with buffer B + 10 mM $CaCl_2$, incubated with Sortase A and LD555-CH(10)LPETGG peptide at a 1:1.3:25 molar ratio at 25 °C for 8 min. The reaction was terminated with 20 mM EDTA and purified on a Superdex 200 increase 10/300 GL gel filtration column equilibrated in buffer B + 300 mM KGlu, 10 mM imidazole. To remove unlabeled SORT-ORC, peak fractions were pooled and applied to a HiTrap Chelating HP 1 ml charged with Co^{2+} equilibrated in buffer B + 300 mM KGlu, 10 mM imidazole. The column was washed with 10 CV buffer B + 300 mM KGlu, 10 mM imidazole and LD555-ORC was eluted with 10 CV buffer B + 300 mM KGlu, 150 mM imidazole. Peak fractions were pooled, concentrated with MWCO 50000 Amicon Ultra Centrifugal Filter unit, frozen in aliquots in liquid N_2 and stored at -80 °C. The labeling efficiency of LD555-ORC was ~90 % based on extinction coefficients of ORC and LD555. The final protein concentration was determined with a Bradford protein assay (Bradford, 1976) using the Protein assay dye reagent concentrate and BSA as standard. Protein purity was assessed by SDS-PAGE analysis.

2.3.4 Cdc6 purification

S. cerevisiae Cdc6 was purified similar as previously described (Frigola et al., 2013). *E. coli* BL21(DE) codon plus RIL were transformed with pGEX-6P-1_Cdc6, grown in LB medium to OD_{600} = 0.6 at 37 °C and subsequently induced with 0.5 mM IPTG for 5 hours at 18 °C. All subsequent steps were performed at 4 °C. Cells were harvested by centrifugation (4000 x g, 10 min), resuspended in buffer C (50 mM K_2HPO_4/KH_2PO_4, pH 7.5, 5 mM $MgCl_2$, 0.02% (v/v) Tween20, 1 mM DTT) + 150 mM potassium acetate (KOAc), 2 mM ATP supplemented with 1x protease inhibitor cocktail and lysed by sonication. The cell lysate was cleared by centrifugation (48000 x g, 30 min) and the supernatant was incubated with 2 ml BV Protino glutathione agarose 4B for 2 hours. Beads were washed with 20 BV buffer C + 150 mM KOAc, 2 mM ATP and finally diluted to a 50 % slurry with 1 BV of the same buffer. 150 U PreScission protease (GST-tagged 3C protease) were added and the mixture was incubated overnight. Cleaved Cdc6 was recovered from the flow-through. The final concentration of KOAc was then adjusted to

75 mM by diluting with buffer C + 2 mM ATP and incubated with 2 ml BV CHT ceramic hydroxyapatite resin equilibrated in buffer C + 75 mM KOAc, 2 mM ATP for 2 hours. Subsequently, the resin was washed with 5 BV buffer C + 75 mM KOAc and 5 BV buffer C + 150 mM KOAc, 15 % (v/v) glycerol. Cdc6 was eluted 4 times with 1 BV of buffer C + 400 mM KOAc, 15 % (v/v) glycerol. Elution fractions were pooled, concentrated with MWCO 10000 Amicon Ultra Centrifugal Filter unit, frozen in aliquots in liquid N_2 and stored at -80 °C. The final protein concentration was determined with a Bradford protein assay (Bradford, 1976) using the Protein assay dye reagent concentrate and BSA as standard. Protein purity was assessed by SDS-PAGE analysis.

2.3.5 Cdt1-MCM, Cdt1-LD655-MCM and Cdt1-LD655-MCM$^{Mcm3\text{-}YDF}$purification

Unlabeled *S. cerevisiae* Cdt1-MCM was expressed using yeast strain yJF38 (Frigola et al., 2013). To generate fluorescently labeled Cdt1-MCM, yeast strain ySA4 was generated (performed by Syafiq Abd Wahab, Remus Lab). In brief, strain ySA4 was generated from yJF38, using linearized plasmids with standard genetic and cloning procedures, to overexpress Cdt1 and Mcm2-7 subunits with a ybbR and 3xFLAG tag fused to the N- and C-terminus of Mcm6, respectively. To obtain fluorescently labeled Cdt1-MCM$^{Mcm3\text{-}YDF}$, the corresponding similar motif within *S. cerevisiae* Mcm3 (identified by pairwise sequence alignment, Figure 5.7B) was replaced by an extended version of a previously described short YDF-containing motif in human Mcm3 (Li et al., 2020). The YDF motif was incorporated into the endogenous and integrated allele of Mcm3 to ensure the absence of wildtype Mcm3 in the subsequent preparation. For this, a CRISPR-Cas9-based editing approach was chosen and essentially performed as previously described (Laughery et al., 2015). Since no compatible auxotrophic selection marker was available, the URA3 marker in pML104 (a gift from John Wyrick) was replaced by an ADE2 marker (the non-URA3-containing fragment of PvuI-digested pML104 was ligated with the ADE2-containing fragment of PvuI-digested 12ADE-B (a gift from Scott Gradia)), generating pMS194_Cas9-ADE2. Next, annealed oligonucleotides MS_257 and MS_258 (containing the gRNA targeting the motif in *S. cerevisiae* Mcm3 corresponding to the YDF-containing motif in human Mcm3) were ligated into BclI-SwaI-digested pMS194_Cas9-ADE2. The resulting pMS200_Cas9-ADE2_scMcm3-guide9 was co-transformed into ySA4 with following repair template (purchased from Eurofins Genomics, Ebersberg, Germany), encoding the 19 residues YDF-containing motif in human Mcm3 (underlined): 5′-ACTCCAAGAAGGTCAACGGCATCTTCCGTTAATGCC ACGCCATCGTCAGCACGCAGAATATTACGTTTTCAAGATGACGAACAGAACGCTG GTGAAGAC<u>GATGGGGATTCATACGACCCCTATGACTTCAGTGACACAGAGGAGGA AATGCCTCAA</u>AGGCTTCAACTGGGGTTGAGAGTGTCTCCAAGACGTAGAGAACAT CTTCACGCACCTGAGGAAGGTTCGTCGGGACCTCTTACCGAGGTCGGTACTCCA-3′. Successful modification of the integrated and endogenous allele of Mcm3 was confirmed by sequencing. Finally, the new strain yMS1 was cured from pMS200_Cas9-ADE2_scMcm3-guide9 and confirmed to be Ade auxotroph. Strain yMS1 grew

comparably to ySA4, confirming that the YDF motif did not alter the MCM function.

Cells were grown in 6 l YP medium supplemented with 2 % (v/v) raffinose at 30 °C and arrested at G1 phase with α-factor at a final concentration of 150 ng/ml at OD_{600} = 1.2. After 3 hours, protein expression was induced by addition of 2 % (v/v) galactose and incubated for another 4 hours. Cells were harvested by centrifugation (4000 x g, 10 min), washed once with cold 0.3 mM PMSF in ddH_2O, once with buffer D (100 mM HEPES-KOH, pH 7.6, 800 mM Sorbitol, 10 mM $Mg(OAc)_2$, 750 mM KGlu) and finally resuspended in 1 packed cell volume of buffer D + 1 mM DTT supplemented with 1x protease inhibitor cocktail and frozen dropwise in liquid N_2. Frozen cells were lysed in a SamplePrep 6875 Freezer/Mill and subsequently mixed with 1 packed cell volume buffer E (45 mM HEPES-KOH, pH 7.6, 0.02 % (v/v) Nonidet P40 Substitute, 5 mM $Mg(OAc)_2$, 10 % (v/v) glycerol, 1 mM ATP, 1 mM DTT) + 300 mM KGlu supplemented with 1x protease inhibitor cocktail. All subsequent steps were performed at 4 °C unless stated otherwise. Thawed cell lysate was cleared by ultracentrifugation (235000 x g, 60 min) and incubated with 0.5 ml BV Anti-FLAG M2 Affinity Gel equilibrated in buffer E + 300 mM KGlu for 3 hours. To remove non-specifically bound protein, the resin was washed twice with 20 BV buffer E + 300 mM KGlu and twice with 20 BV buffer E + 100 mM KGlu. Cdt1-MCM was eluted 5 times with 1 BV buffer E + 100 mM KGlu, 0.5 mg/ml poly FLAG peptide. To produce Cdt1-LD655-MCM or Cdt1-LD655-$MCM^{Mcm3\text{-}YDF}$, following FLAG elution, ybbR-tagged Cdt1-MCM or Cdt1-$MCM^{Mcm3\text{-}YDF}$ was supplemented with 10 mM $MgCl_2$ and incubated with SFP synthase and LD655-CoA at a 1:3:6 molar ratio for 2 hours at 30 °C. Unlabeled or LD655-labeled Cdt1-MCM was further purified on a Superdex 200 increase 10/300 GL gel filtration column equilibrated in buffer E + 100 mM KOAc. Fractions containing stoichiometric Cdt1-MCM/$MCM^{Mcm3\text{-}YDF}$ complex were pooled, concentrated with MWCO 50000 Amicon Ultra Centrifugal Filter unit, frozen in aliquots in liquid N_2 and stored at -80 °C. The labeling efficiencies of Cdt1-LD655-MCM and Cdt1-LD655-$MCM^{Mcm3\text{-}YDF}$ were estimated to be ~92 % and ~90 %, respectively, based on the extinction coefficients of Cdt1-MCM and LD655. The final protein concentration was determined with a Bradford protein assay (Bradford, 1976) using the Protein assay dye reagent concentrate and BSA as standard. Protein purity was assessed by SDS-PAGE analysis.

2.3.6 AF488-, LD555- and LD655-T7 RNA polymerase purification

To site-specifically label T7 RNA polymerase (RNAP) at its N-terminus, plasmid pBH161 (He et al., 1997) was cut with NcoI and XhoI and the leader sequence was replaced with either SNAP- or ybbR-tag sequence. His_6-tagged SNAP26b sequence was amplified from pSNAP-tag (T7)-2 with oligonucleotides MS_144 and MS_145 (pMS145_SNAP-T7RNAP). The ybbR-tag followed by a His_6-tag and a GS-linker was added with annealed oligonucleotides MS_221 and MS_222 (pMS173_ybbR-T7RNAP).

E. coli BL21(DE3) star were transformed with pMS145_SNAP-T7RNAP or pMS173_ybbR-T7RNAP and grown in TB medium at 37 °C to OD_{600} = 1.5. The

temperature was lowered to 18 °C and T7 RNAP expression was induced with 0.1 mM IPTG for 16 hours. All subsequent purification steps were performed at 4 °C. Cells were harvested by centrifugation (4000 x g, 10 min), resuspended in buffer F (20 mM K_2HPO_4/KH_2PO_4, pH 8, 10 % (v/v) glycerol, 1 mM DTT) + 300 mM KCl, 10 mM imidazole supplemented with 1x protease inhibitor cocktail and lysed by sonication. The cell lysate was cleared by centrifugation (48000 x g, 30 min) and applied to a HisTrap HP 5 ml equilibrated in buffer F + 300 mM KCl, 10 mM imidazole. After sample application, the column was washed with 20 CV buffer F + 300 mM KCl, 10 mM imidazole, 5 CV buffer F + 1000 mM KCl and 10 CV buffer F + 50 mM KCl, 20 mM imidazole. T7 RNAP was eluted on a 20-300 mM imidazole gradient in buffer F + 50 mM KCl. Peak fractions were pooled and applied to two HiTrap Heparin HP 5 ml, equilibrated in buffer F + 50 mM KCl. The columns were washed with 10 CV buffer F + 50 mM KCl and protein was eluted on a 50-1000 mM KCl gradient in buffer F. Peak fractions were pooled, concentrated with a MWCO 50000 Amicon Ultra Centrifugal Filter unit and applied to a HiLoad 16/600 Superdex 200 pg equilibrated in buffer G (25 mM HEPES-KOH, pH 8, 150 mM KCl, 10 % (v/v) glycerol, 5 mM DTT). Peak fractions were pooled, spin concentrated and directly used for labeling.

To produce AF488-T7 RNAP, SNAP-T7 RNAP was labeled with a 5-fold molar excess of SNAP-Surface Alexa Fluor 488 in buffer G at 30 °C for 2 hours. To produce LD555- or LD655-T7 RNAP, ybbR-T7 RNAP was mixed with SFP synthase and LD555-CoA or LD655-CoA at a 1:2:5 molar ratio, respectively, in buffer G + 10 mM $MgCl_2$ and incubated at 30 °C for 2 hours. Labeled T7 RNAP was further purified on a Superdex 200 increase 10/300 GL gel filtration column equilibrated in buffer G. Peak fractions were pooled, dialyzed against 25 mM HEPES-KOH, pH 8, 150 mM KCl, 50 % (v/v) glycerol, 0.05 % (v/v) Tween20, 1 mM EDTA, 10 mM DTT and stored at -20 °C. Labeling efficiencies were estimated to be ~94 %, 88 % and 87 % for AF488-, LD555-, and LD655-T7 RNAP, respectively based on extinction coefficients of T7 RNAP and dyes. The final protein concentration was determined on a spectrophotometer (NP80) by measuring the absorbance at 280 nm. Protein purity was assessed by the 260/280 nm ratio as well as SDS-PAGE analysis and intact mass measurement.

2.3.7 LD555-H3 histone octamers purification

Synthetic genes coding for *S. cerevisiae* histones H2A/H2B and H3/H4 (Eurofins Genomics, Ebersberg, Germany) were cloned into pETDuet-1 (pMS184_H2A_H2B) and pCDFDuet-1 (pMS185_H3_H4) (both Novagen), respectively. Site-specific labeling at H3 was achieved by changing serine11 to cysteine by site-directed mutagenesis in pCDFDuet_H3_H4 (pMS186_H3-S11C_H4) using oligonucleotides MS_238 and MS_239. *E. coli* BL21(DE3) codon plus RIL were co-transformed with pMS184_H2A_H2B and pMS186_H3-S11C_H4 and grown in ZYP-5052 auto-induction medium at 37 °C to OD_{600} = 0.8. The temperature was lowered to 18 °C and growth continued for another 18 hours. All subsequent purification steps were performed at 4 °C.

Cells were harvested by centrifugation (4000 x g, 10 min), resuspended in buffer H (20 mM HEPES-NaOH, pH 7.6, 10 % (v/v) glycerol, 1 mM EDTA) + 800 mM NaCl, 2 mM DTT, supplemented with 1x protease inhibitor cocktail and lysed by sonication. The cell lysate was cleared by centrifugation (48000 x g, 30 min) and applied to two HiTrap Heparin HP 5 ml equilibrated in buffer H + 800 mM NaCl, 2 mM DTT. The columns were washed with 15 CV buffer H + 800 mM NaCl, 2 mM DTT and histone octamers were eluted on an 800-2000 mM NaCl gradient in buffer H + 2 mM DTT. Peak fractions were pooled, spin concentrated with a MWCO 10000 Amicon Ultra Centrifugal Filter unit and applied to a HiLoad 16/600 Superdex 200 pg equilibrated in buffer H + 2000 mM NaCl, 2 mM DTT. Peak fractions containing histone octamers were pooled and spin concentrated. Concentrated histone octamers were treated with 10 mM TCEP for 2 hours and DTT was removed with a HiTrap Desalting 5 ml equilibrated in buffer H + 2000 mM NaCl. Subsequently, histone octamers were immediately mixed with a 50-fold molar excess of LD555 maleimide mono-reactive dye. After 20 hours, the reaction was quenched with 10 mM DTT and labeled histone octamers were purified on a Superdex 200 increase 10/300 GL gel filtration column equilibrated in buffer H + 2000 mM NaCl, 1 mM DTT. Peak fractions were pooled, spin concentrated, frozen in aliquots in liquid N_2 and stored at -80 °C. The labeling efficiency was estimated to be ~1.5 dye molecules per histone octamer based on extinction coefficients. The final protein concentration was determined on a spectrophotometer (NP80) by measuring the absorbance at 280 nm. Protein purity was assessed by the 260/280 nm ratio as well as SDS-PAGE analysis and intact mass measurement.

2.3.8 JF549-Cohesin^{STAG1} purification

Recombinant human recombinant cohesin^{STAG1} with a HaloTag fused to the SCC1 subunit was purified from baculovirus infected *Sf9* cells and labeled with Janelia Fluor 549 HaloTag as previously described (Davidson et al., 2019). Human JF549-Cohesin^{STAG1} was generously provided by Iain Davison (Peters Lab, IMP).

2.3.9 Protocatechuate 3,4-Dioxygenase (PCD) purification

Single-molecule fluorescence experiments rely on continuous emission of photons from a single fluorophore. The maximum observation time is mainly defined by the total amount of emitted photons which is limited by photobleaching. Since dissolved oxygen is a main reason for photobleaching, enzymatic oxygen scavenging systems have been shown to significantly improve dye lifetime during single-molecule experiments (Aitken et al., 2008). One described enzymatic oxygen scavenging system relies on Protocatechuate 3,4-Dioxygenase (PCD) catalyzed oxidation of 3,4-Dihydroxybenzoic acid (PCA) under oxygen consumption. PCD from *Pseudomonas putida* was basically expressed and purified as previously described (Senavirathne et al., 2018). *E. coli*

BL21(DE3) carrying pVP91A-pcaHG (a gift from John Lipscomb (Knoot et al., 2015)) were grown in ZYP-5052 auto-induction medium containing 10 mg/L Fe(II)SO_4 at 18 °C. Cleared cell lysate was applied to a HisTrap HP 5 ml equilibrated in 50 mM Na_2HPO_4/NaH_2PO_4, pH 8, 500 mM NaCl, 10 % (v/v) glycerol, 10 mM imidazole, washed with 20 mM imidazole and PCD was eluted with 120 mM imidazole. PCD containing fractions were pooled and further purified via gel filtration chromatography on a HiLoad 26/600 Superdex 200 pg, equilibrated in 50 mM Tris-HCl, pH 7.5, 100 mM NaCl, 10 % (v/v) glycerol, 0.1 mM EDTA. Peak fractions were pooled, concentrated to ~20 mg/ml with a MWCO 10000 Amicon Ultra Centrifugal Filter unit, frozen in aliquots in liquid N_2 and stored at -80 °C. The final protein concentration was determined on a spectrophotometer (NP80) by measuring the absorbance at 280 nm. Protein purity was assessed by the 260/280 nm ratio as well as SDS-PAGE analysis and intact mass measurement. PCD activity (in U/ml) was determined by PCA conversion measured by decreasing absorption at 290 nm as described elsewhere (Aitken et al., 2008).

2.4 DNA substrates preparation

Single-molecule assays have been widely used to study protein-DNA interactions with temporal and spatial resolution. The latter, however, requires long DNA molecules with 10s of kb in length of high quality, often with no tolerance for a single nick in the entire DNA molecule (Agarwal and Duderstadt, 2020). Chemical modifications like biotin or digoxigenin are commonly used to specifically bind one end of a linear DNA molecule to a device (e.g. microscope slide or beads) coated with streptavidin or anti-digoxigenin antibodies, respectively. DNA from bacteriophage λ (λ DNA) fulfills these requirements, is commercially available and is hence commonly used in single-molecule assays. Although chemical modification or introduction of specific sequences to the λ DNA end is easily feasible, sequence alterations throughout the entire DNA are highly restricted by intrinsic features of the phage genome. However, more complex biochemical reaction, including assembly pathways of several macromolecular machineries, require highly customizable DNA.

To meet these requirements, we constructed a series of easily modifiable, large (pMSuperCos) plasmids and developed a highly reproducible purification and functionalization procedure, generating long and highly customizable DNA of excellent quality.

2.4.1 pMSuperCos plasmids construction

The SuperCos1 cosmid vector was used as template to construct all pMSuperCos plasmids. Initially, both cos recognition sequences were removed and replaced by a multiple cloning site of unique restriction enzymes (XbaI-XhoI-SpeI-AsiSI-NheI-FseI-NotI), generating the plasmid pMSuperCos151_empty (Figure 2.1A). To increase the total plasmid size, λ DNA sequences were amplified by PCR and cloned into pMSuperCos151_empty as follows: λ(177-3007) – amplified with MS_183 and MS_184 – between XhoI and SpeI, λ(3008-10142) – amplified with MS_182 and MS_173 – between AsiSI and NheI and λ(10143-21084) – amplified with MS_170 and MS_171 – between NheI and FseI. The resulting plasmid pMSuperCos159 was 27 kb in length containing ~21 kb λ DNA sequence (Figure 2.1B). Insertion of all λ DNA fragments was confirmed by sequencing the ligated junctions. Furthermore, a series of analytical restriction digests were performed, confirming not only correct plasmid assembly but also conservation of unique restriction sites in pMSuperCos159 (Figure 2.1C).

Plasmid pMSuperCos159 was further modified by inserting specific sites required for single-molecule assays described in this study. To construct pMSuperCos162_ARS1, yeast origin ARS1 (amplified from pARS/WTA (Marahrens and Stillman, 1992) using oligonucleotides MS_178 and MS_179 was cloned into the naturally occurring BamHI site at position 5506 in the phage sequence. This plasmid was further modified by inserting the T7 phi10 promoter (T7P) sequence with the first thymidine occurring at

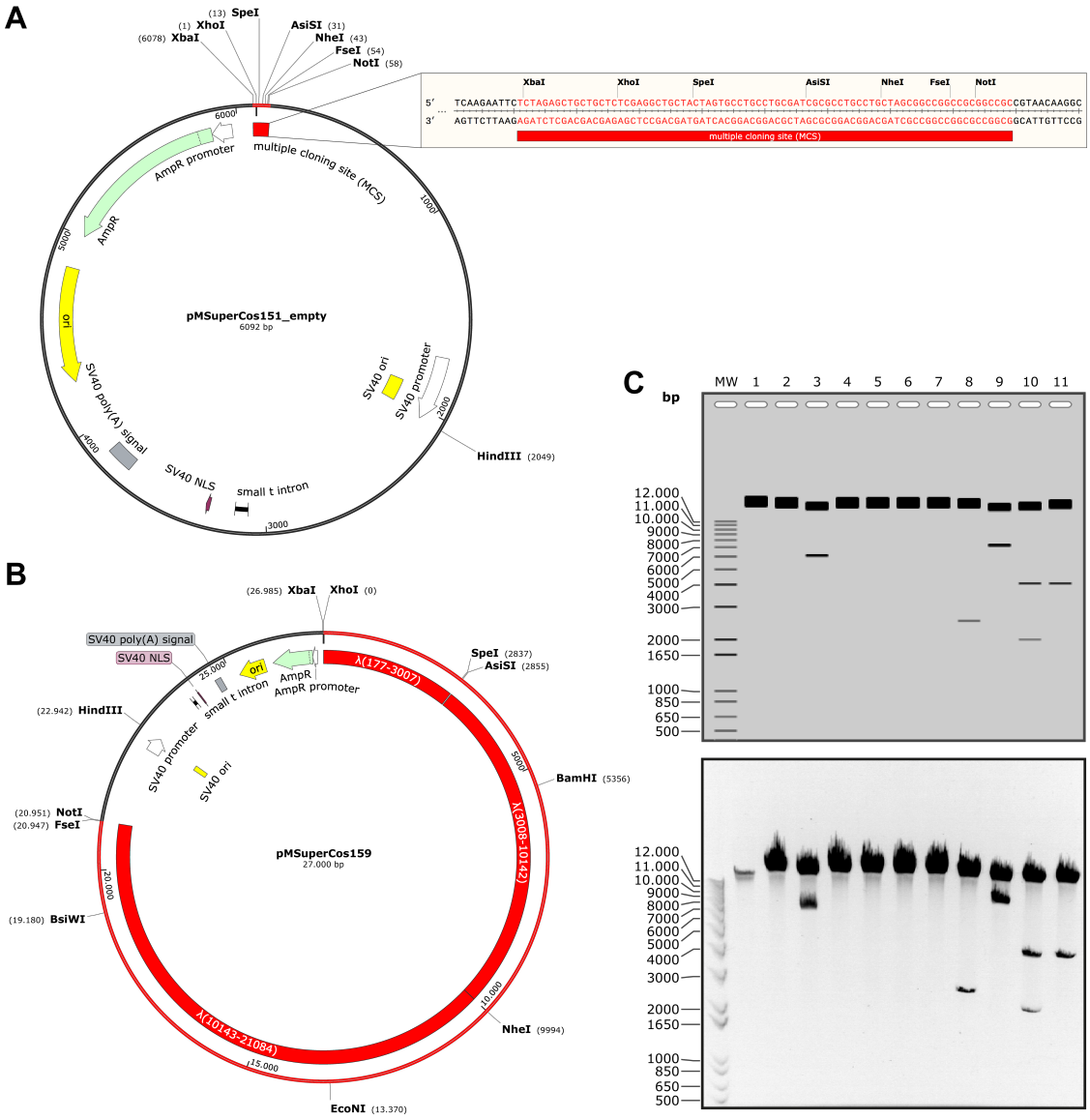

Figure 2.1. Construction of a highly customizable plasmid for single-molecule studies
(**A**) Plasmid map of pMSuperCos151_empty, which was generated from the pSuperCos1 cosmid vector by replacing both cos recognition sites with a customizable multiple cloning site (red). (**B**) Plasmid map of pMSuperCos159, which was constructed by inserting a total of ~21 kb of λ DNA sequence (red). (**C**) Analytical restriction digest confirming correct pMSuperCos159 plasmid assembly and conservation of unique restriction sites. Samples were run on a 0.5 % agarose gel in TBE (bottom) and compared to an *in silico* simulation (top). Lanes: MW, molecular weight marker (1 kb plus DNA ladder); 1, control (undigested plasmid); 2, +XhoI; 3, +XbaI/NotI; 4, +SpeI; 5, +BamHI, 6, +NheI; 7, +BsiWI; 8, +SpeI/BamHI; 9, +AsiSI/NheI; 10, +XhoI/HindIII/NotI; 11, +XhoI/HindIII.

position +16 between SpeI and AsiSI site using annealed oligonucleotides MS_180 and MS_181 (pMSuperCos165_T7P-ARS1). A single Widom601 nucleosome positioning sequence (Lowary and Widom, 1998) was amplified with oligonucleotides MS_223 and MS_224 and subsequently inserted at NheI site (pMSuperCos177_T7P-ARS1-Widom601). Finally, five tandem T7 terminator sequences (T7T) (amplified from p850 (Gros et al., 2015) using oligonucleotides MS_231 and MS_232) were introduced by Gibson Assembly at EcoNI site at position 13514 in the phage sequence. The final

plasmid pMSuperCos182_T7P-ARS1-Widom601-5xT7T and the nucleotide sequences of all specific sites are depicted in Figure 2.2.

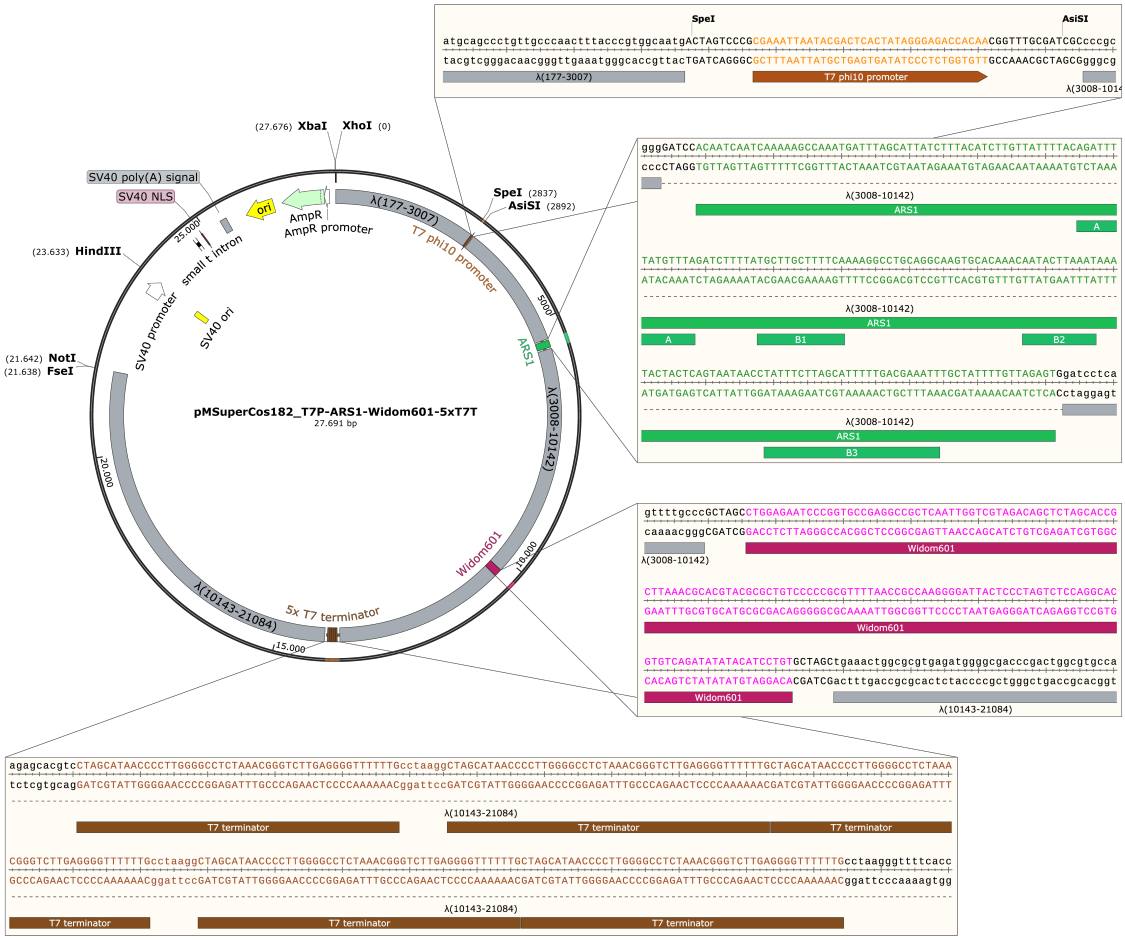

Figure 2.2. Introduction of specific sites into pMSuperCos159 for single-molecule assays
Plasmid map and sequence details of pMSuperCos182_T7P-ARS1-Widom601-5xT7T. Following specific sites were inserted: The yeast origin ARS1 (green), a T7 phi10 promoter (light brown), a single Widom601 nucleosome positioning sequence (magenta) and five tandem T7 terminator sequences (dark brown).

2.4.2 Biotinylated linear DNA preparation for ensemble assays

For ensemble helicase / RNAP loading and collision assays (chapter 2.5), biotinylated 5 kb DNA substrates were prepared by PCR using oligonucleotides MS_226 and MS_227 with pMSuperCos162_ARS1 (5 kb-ARS1 DNA) or pMSuperCos165_T7P-ARS1 (5 kb-T7P-ARS1 DNA) as template. The PCR product was purified on an agarose gel using QIAquick Gel Extraction Kit and stored in aliquots at -20 °C.

2.4.3 Biotinylated linear DNA preparation for single-molecule assays

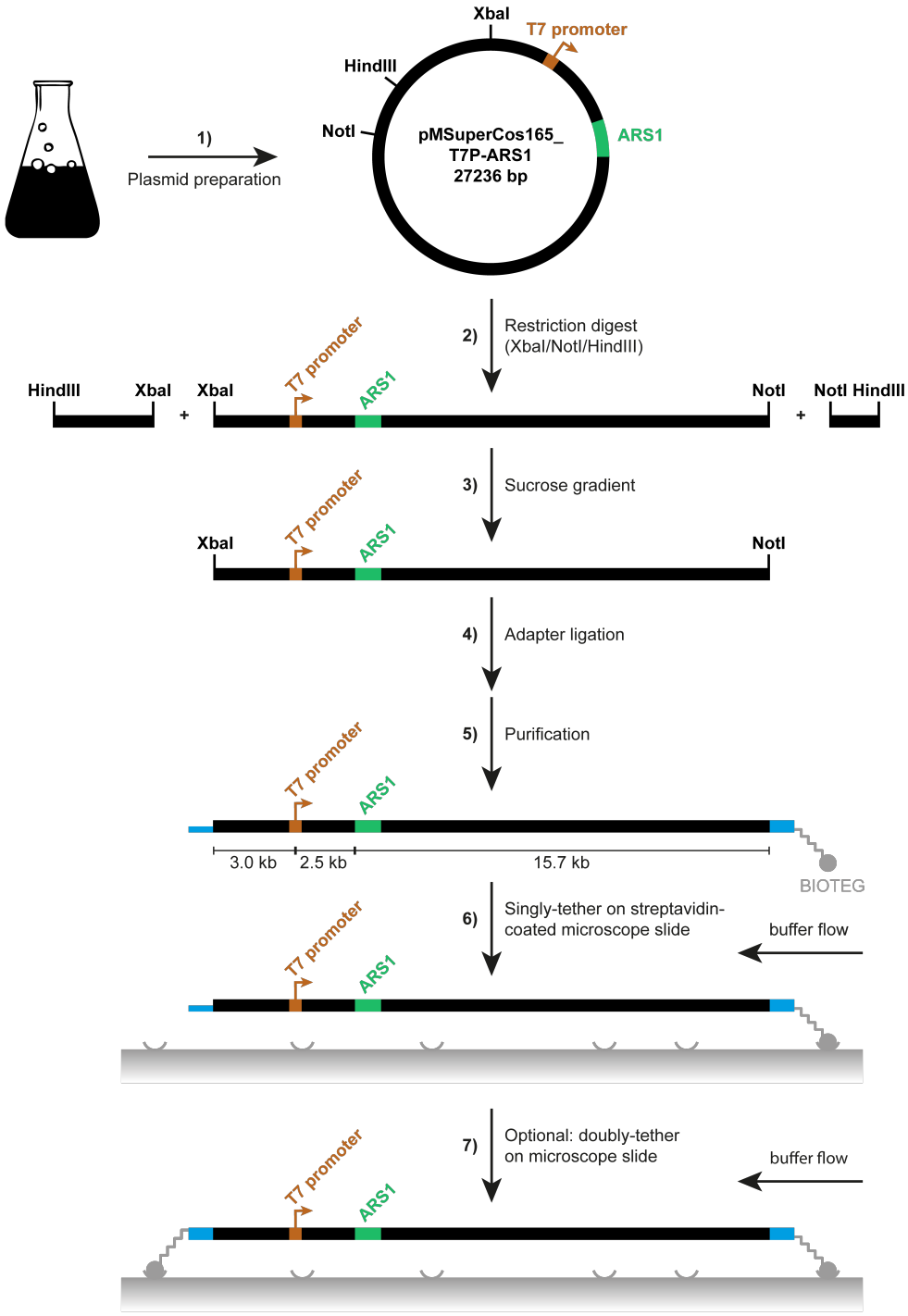

Figure 2.3. Biotinylated linear DNA preparation and tethering for single-molecule assays
Schematic example of all steps to prepare a 21 kb functionalized linear DNA substrate from pMSuperCos165_T7P-ARS1 which allows tethering to a microscope slide. The plasmid was purified from *E. coli* (step 1), digested with XbaI, NotI and HindIII (step 2) and the resulting 21 kb fragment was isolated on a sucrose gradient (step 3). Afterwards, oligonucleotide adapters were ligated to the fragment (step 4) and free adapters removed on a gel filtration column (step 5). The resulting DNA could then be singly- (step 6) or doubly-tethered (step 7) on a microscope slide.

To meet the high DNA quality criteria required for single-molecule studies, we developed a highly reproducible and gentle preparation procedure for biotinylated linear DNA (Figure 2.3, steps 1-5). For single-molecule TIRF assays all DNA substrates were restriction digest fragments (21-22 kb in length) of different pMSuperCos plasmids as indicated (see chapter 2.4.1 and Table 2.4).

First, plasmids were isolated from *E. coli* DH5α with QIAGEN Plasmid Maxi Kit according to manufacturer's instructions with one exception. All steps were performed at 4 °C which has been shown to increase plasmid quality (nick-free as indicated by a supercoiled state) (Carbone et al., 2012) (step 1). Second, 100 µg plasmid were digested with 100 U XbaI, NotI-HF and HindIII-HF in 1x CutSmart buffer at 37 °C for 7 hours (step 2). The resulting XbaI-NotI fragment was separated from the plasmid backbone on a 10-40 % (weight per volume, w/v) sucrose gradient in 20 mM Tris-HCl, pH 7.5, 1000 mM NaCl, 5 mM EDTA using a SW41 Ti rotor (Beckman Coulter) at 30000 rpm, 20 °C for 22 hours (Figure 2.4A). The gradient was separated in 500 µl fractions. Fractions containing pure XbaI-NotI fragment were precipitated with ethanol at -20 °C and DNA was reconstituted in 10 mM Tris-HCl, pH 8, 0.1 mM EDTA (step 3).

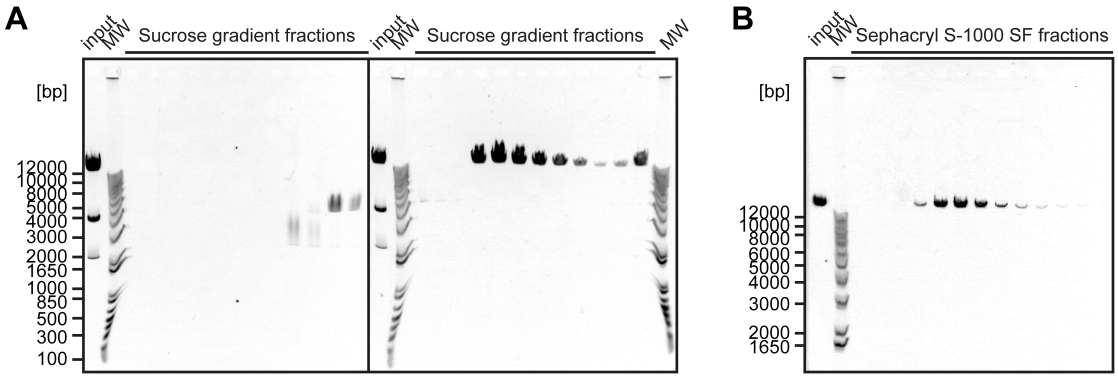

Figure 2.4. Purification strategies of the 21 kb linear DNA fragment
(**A**) Separation of the 21 kb XbaI-NotI fragment after restriction digest on a 10-40 % sucrose gradient (Figure 2.3, step 3). Fractions were analyzed on a 0.5 % agarose gel in TBE (increasing sucrose concentration from left to right). (**B**) Separation of the functionalized 21 kb fragment from excess adapters on a Sephacryl S-1000 SF gel filtration column (Figure 2.3, step 5). Fractions were analyzed on a 0.5 % agarose gel in TBE (increasing elution volume from left to right). MW, molecular weight marker (1 kb plus DNA ladder).

Subsequently, DNA handles were prepared by annealing equimolar amounts of oligonucleotides MS_200 with MS_201 and MS_202 with MS_203 in 30 mM HEPES-KOH, pH 7.5, 100 mM KOAc, incubating at 95 °C for 5 min and cooling to 4 °C at a rate of -1 °C/min. Annealed DNA handles were mixed with the purified XbaI-NotI fragment at a molar ratio of 15:1 and ligated with T4 DNA Ligase in 1x T4 DNA Ligase buffer at 16 °C overnight (step 4). Excess DNA handles were removed on a Sephacryl S-1000 SF Tricorn 10/300 gel filtration column equilibrated in 10 mM Tris-HCl, pH 8, 300 mM NaCl, 1 mM EDTA (Figure 2.4B). Finally, peak fractions were pooled, precipitated with ethanol and reconstituted in 10 mM Tris-HCl, pH 8, 1 mM EDTA. Small aliquots of the final DNA substrates were snap frozen in liquid N_2 and stored at -80 °C. All final DNA substrates

for single-molecule assays are functionalized with biotin at NotI site and an 18 bp ssDNA overhang at XbaI site which was used for orientation specific doubly-tethering where indicated.

2.4.4 Biotinylated, chromatinized linear DNA preparation for single-molecule assays

Biotinylated, linear DNA was prepared from pMSuperCos177_T7P-ARS1-Widowm601 as described in chapter 2.4.3. 0.5 µg DNA were incubated with a 50- ("low density") or 75-fold ("high density") molar excess of LD555-H3 histone octamers in a total volume of 20 µl buffer H + 2000 mM NaCl, 1 mM DTT on ice for 2 hours. Subsequently, nucleosomes were reconstituted by salt gradient dialysis as follows: The reaction mix was transferred into Slide-A-Lyzer MINI Dialysis Unit, 3.5K MWCO and put in a beaker containing 0.5 l buffer H + 2000 mM NaCl, 1 mM DTT. Using a continuous pump flow system at 3 ml/min, buffer H + 50 mM NaCl, 1 mM DTT was continuously added to the dialysis beaker, mixed and removed for 20 hours. Finally, the reaction mix was dialyzed against fresh buffer H + 50 mM NaCl, 1 mM DTT for another 2 hours. Biotinylated, chromatinized linear DNA was stored at 4 °C.

2.5 Ensemble pulldown assays

2.5.1 Ensemble helicase and RNA polymerase loading assay

For each reaction, 0.2 pmol biotinylated DNA (5 kb-ARS1 or 5 kb-T7P-ARS1 DNA) were tethered to 50 µg Dynabeads M-280 Streptavidin using DYNAL Dynabeads KilobaseBINDER Kit according to manufacturer's instructions. For helicase loading, 0.2 pmol tethered DNA were mixed with 1 pmol ORC, 1 pmol Cdc6 and 2 pmol Cdt1-MCM in 20 µl assay buffer (30 mM HEPES-KOH, pH 7.6, 8 mM $Mg(OAc)_2$, 0.05 % (v/v) Tween20, 0.1 mg/ml BSA, 5 mM DTT) + 200 mM KOAc, 3 mM ATP. For RNAP loading, 0.2 pmol tethered DNA were mixed with 10 pmol T7 RNAP in 20 µl assay buffer + 200 mM KOAc, 1 U/µl RNase inhibitor and 0.6 mM of each nucleoside triphosphate (NTP) as indicated. Samples were incubated at 30 °C, 1250 rpm in a thermoshaker for 20 min. The supernatant was collected and beads were washed once with 100 µl assay buffer containing either 200 mM KOAc (low salt wash) or 500 mM NaCl (high salt wash) and once with 100 µl assay buffer + 200 mM KOAc. Retained proteins were eluted in 1x SDS loading buffer at 95 °C, 1250 rpm for 5 min. Samples were run on a 4-12 % Bis-Tris SDS gel and visualized by Western Blot.

To this end, proteins were wet-blotted on a nitrocellulose membrane in transfer buffer (48 mM Tris, 39 mM Glycine, pH 8.3, 0.0375 % (w/v) SDS, 10 % (v/v) methanol) at 90 V, 4 °C for 90 min. Subsequently, the membrane was incubated with a primary antibody (anti-Orc6 and anti-Cdc6 1:500, anti-Mcm4 and anti-His Tag 1:1000, see Table 2.1) diluted in Superblotto buffer (2.5 % (w/v) milk powder, 0.5 % (w/v) BSA, 0.5 % (v/v) Nonidet P40 Substitute, 0.1 % (v/v) Tween20 in TBS (50 mM Tris-HCl, pH 7.5, 150 mM NaCl)) at 4 °C overnight. The membrane was washed three times with TBS-T (TBS + 0.1 % (v/v) Tween20) and incubated with an anti-mouse HRP-conjugated secondary antibody at RT for 1 hour. The membrane was washed again three times with TBS-T. Finally, the Western Blot was developed with Pierce ECL Western Blotting substrate according to manufacturer's instructions and chemiluminescence detected with a LAS-3000 imaging system.

2.5.2 Ensemble RNA polymerase - helicase collision assay

Collision assays were set up in two consecutive steps. First, helicase loading was performed on 5 kb-T7P-ARS1 DNA as described in chapter 2.5.1. Second, 0.2 pmol licensed DNA were incubated with 10 pmol T7 RNAP in 20 µl assay buffer + 200 mM KOAc, 1 U/µl RNase inhibitor, 3 mM ATP and 0.6 mM guanosine/cytidine/uridine triphosphate (GTP/CTP/UTP) as indicated and incubated at 30 °C, 1250 rpm in a thermoshaker for another 20 min. The supernatant was collected and beads were washed twice with 100 µl assay buffer + 200 mM KOAc, 1 U/µl RNase inhibitor. Retained proteins were eluted and samples analyzed by Western Blot as described in chapter 2.5.1.

2.6 Single-molecule assays

2.6.1 PEG-Biotin microscope slides preparation

Glass coverslips (22 x 22 mm) were cleaned with a Zepto plasma cleaner and incubated in acetone containing 2 % (v/v) 3-aminopropyltriethoxysilane for 5 min. Silanized coverslips were rinsed with ddH_2O, dried and baked at 110 °C for 30 min. Coverslips were then covered with a fresh solution of 0.4 % (w/v) Biotin-PEG-Succinimidyl Carbonate (MW 5000) and 15 % (w/v) mPEG-Succinimidyl Carbonate (MW 5000) in fresh 0.1 M $NaHCO_3$ and incubated overnight at RT. Coverslips were rinsed with ddH_2O, dried and incubated again with in a fresh Biotin-PEG/mPEG solution as described above. Functionalized PEG-Biotin microscope slides were again washed and dried and finally stored under vacuum.

2.6.2 Flow cell preparation

A functionalized PEG-Biotin microscope slide was covered with 0.2 mg/ml streptavidin in blocking buffer (20 mM Tris-HCl, pH 7.5, 50 mM NaCl, 2 mM EDTA, 0.2 mg/ml BSA, 0.005 % (v/v) Tween20) for 30 min. To assemble a flow cell, a polydimethylsiloxane block was placed on top of the previously washed and dried microscope slide, generating a 0.5 mm wide and 0.1 mm high flow cell. A polyethylene tube (inner diameter 0.58 mm) was inserted on either end with a buffer tube and a syringe pump connected to the inlet and outlet tube, respectively to generate directed buffer flow (Figure 2.5). The assembled flow cell was rinsed with blocking buffer for 5 min. Biotinylated linear DNA was then tethered to the slide surface at 5 pM in blocking buffer for 15 min in the absence of buffer flow and the flow cell was subsequently washed with blocking buffer. In case biotinylated, chromatinized linear DNA was used, blocking buffer was supplemented with 0.4 mg/ml herring sperm DNA to remove free histone octamers. In experiments using doubly-tethered DNA, the flow cell was flushed with 100 nM oligonucleotide MS_204 in blocking buffer at 100 µl/min for 12 min. To further reduce non-specific protein binding to the slide surface, the flow cell was washed with assay buffer + 2 mg/ml casein (+ 500 U/ml Exonuclease I for doubly-tethered DNA) and incubated for 40 min.

2.6.3 Single-molecule helicase loading assay

To achieve helicase loading, unless stated otherwise, 0.25 nM (LD555-)ORC, 4 nM Cdc6 and 10 nM Cdt1-LD655-MCM in assay buffer + 200 mM KOAc, 3 mM ATP or adenosine-5'-o-(3-thio-triphosphate) (ATPγS) were introduced to a prepared flow cell and incubated for 25 min. To favor multiple MCM loading also occurring at non-ARS1 sites, ORC concentration and incubation time were increased to 1 nM and 35 min, respectively. The flow cell was washed with 150 µl assay buffer + 0.6 mM ATP or ATPγS containing either 200 mM KOAc (low salt wash) or 500 mM NaCl (high salt wash) followed by 300 µl assay

buffer + 200 mM KOAc, 0.6 mM ATP or ATPγS supplemented with an oxygen scavenging system (OSS; consisting of 1 mM Trolox, 2.5 mM PCA, 0.21 U/ml PCD) (Aitken et al., 2008). Subsequently, imaging was started at a constant flow of 50 µl/min. DNA was post-stained with 50 nM SYTOX Orange in the same buffer as during imaging.

2.6.4 Single-molecule spontaneous helicase sliding assay

Helicases were loaded as described in chapter 2.6.3 using a prepared flow cell with doubly-tethered DNA. Upon helicase loading, the flow cell was washed with 300 µl assay buffer + 500 mM NaCl, 0.6 mM ATP supplemented with an OSS. To eliminate any external force applied by buffer flow, the flow was stopped before imaging (at higher resolution; see imaging conditions in chapter 2.6.9) was started. DNA was post-stained with 50 nM SYTOX Orange in the same buffer as during imaging.

2.6.5 Single-molecule transcription assay

To visualize RNAP on all DNA molecules in a synchronous manner, transcription assays were set up in two major steps. First, T7 RNAP was loaded on DNA to form stalled elongation complexes. To achieve this, the downstream region of the T7 promoter was designed to not contain deoxythymidine until position +16. By omitting UTP in the initially supplied NTP mix, T7 RNAP can initiate transcription, switch to its stable elongation mode but is then stalled at position +15 (a distance which should not to support stable loading of a second T7 RNAP) (Tahirov et al., 2002; Yin and Steitz, 2002). Second, after removing unbound T7 RNAP from solution, transcription was resumed by supplying all NTPs, which allowed synchronous observation of one transcription event per DNA molecule.

For this, 5 nM labeled T7 RNAP (AF488-, LD555- and LD655-T7 RNAP were used interchangeably) in assay buffer + 200 mM KOAc, 40 U/ml RNase inhibitor, 0.6 mM GTP/CTP/ATP (or ATPγS) were introduced to a prepared flow cell and incubated for 25 min. Unbound T7 RNAP was washed out with 300 µl assay buffer + 200 mM KOAc, 40 U/ml RNase inhibitor, 0.6 mM GTP/CTP/ATP (or ATPγS) supplemented with an OSS. To resume transcription, the flow cell was flushed at 50 µl/min with assay buffer + 200 mM KOAc, 40 U/ml RNase inhibitor, 0.6 mM each NTP (containing ATP or ATPγS) supplemented with an OSS and imaging was started immediately. DNA was post-stained with 50 nM SYTOX Orange in the same buffer as during imaging.

2.6.6 Single-molecule RNA polymerase and origin licensing factor collision assay

Collision assays were set up similar as described for helicase loading and transcription assay (chapters 2.6.3 and 2.6.5). Unless stated otherwise, 0.25 nM (LD555-)ORC, 4 nM Cdc6, 10 nM Cdt1-LD655-MCM and 5 nM labeled T7 RNAP in assay buffer + 200 mM KOAc, 40 U/ml RNase inhibitor, 3 mM ATP or ATPγS, 0.6 mM GTP/CTP were introduced to a prepared flow cell and incubated for 25 min. The flow cell was washed with 150 µl assay buffer + 40 U/ml RNase inhibitor, 0.6 mM GTP/CTP/ATP (or ATPγS) containing either 200 mM KOAc (low salt wash) or 500 mM NaCl (high salt wash) followed by 300 µl assay buffer + 200 mM KOAc, 40 U/ml RNase inhibitor, 0.6 mM GTP/CTP/ATP (or ATPγS) supplemented with an OSS. To continue transcription, the flow cell was flushed at 50 µl/min with assay buffer + 200 mM KOAc, 40 U/ml RNase inhibitor, 0.6 mM each NTP (containing ATP or ATPγS) supplemented with an OSS and imaging was started immediately. DNA was post-stained with 50 nM SYTOX Orange in the same buffer as during imaging.

2.6.7 Single-molecule cohesin translocation assay on licensed DNA

To visualize passive (salt-driven) cohesin translocation on licensed DNA molecules, helicase loading was performed as described in chapter 2.6.3 using a prepared flow cell with doubly-tethered DNA.

Cohesin loading and sliding was essentially performed as previously described (Davidson et al., 2016). 0.7 nM JF549-cohesin^{STAG1} were incubated with licensed DNA in cohesin binding buffer (35 mM Tris, pH 7.5, 25 mM NaCl, 25 mM KCl, 1 mM $MgCl_2$, 10 % (v/v) glycerol, 0.1 mg/ml BSA, 0.003 (v/v) Tween20, 1 mM DTT, 0.2 mM ATP) for 10 min. To remove unbound protein, remove MCM loading intermediates and induce cohesin translocation, the flow cell was washed with 150 µl assay buffer + 500 mM NaCl, 1 mM DTT, 0.6 mM ATP supplemented with an OSS. Imaging was either started directly (high salt condition) or after lowering the salt concentration to 150 mM NaCl (150 µl wash) in an otherwise identical buffer as described for high salt condition. DNA was post-stained with 50 nM SYTOX Orange in the same buffer as during imaging.

2.6.8 Assays to determine the stoichiometry of fluorescently labeled proteins

To determine the stoichiometry of fluorescently labeled proteins, photobleaching experiments were performed. All experiments were essentially performed as described above, except that the frame rate was increased ~10-fold (see imaging conditions in chapter 2.6.9) and buffers were not supplemented with OSS to allow photobleaching during the standard imaging time. T7 RNAP bleaching experiments were performed with stalled elongation complexes at the T7 promoter.

2.6.9 Micromirror TIRF microscope design and imaging conditions

Numerous biochemical processes being indispensable for life, e.g. ATP synthesis or DNA replication, require complex, macromolecular machines. Although structural approaches are suited to give insight into these machines at the atomic level, they only draw static pictures of these complex processes. However, given the high complexity and dynamic variability of macromolecular machines, structural as well as ensemble approaches fail to provide insights into unique dynamic aspects of assembly pathways, reaction kinetics and heterogeneous pathway choices. Single-molecule techniques overcome these limitations and thus have been widely used to study macromolecular machines.

Total internal reflectance fluorescence microscopy (TIRFM) is one common method to study single molecules (also see chapter 1.2.2). In multiwavelength TIRFM, a common strategy to generate TIR is based on focusing the excitation beam through a microscope objective (MO) with high numerical aperture (>1.4) (Selvin and Ha, 2008). Consequently, the MO is passed by excitation and emission light, whereas the former has to be removed from the emission path. Besides special emission filters, this is usually achieved by a dichroic mirror (DM), which reflects the excitation beam right below the back aperture of the MO while it transmits the emission fluorescence. A major disadvantage of this DM-based strategy is the requirement of multiple or complex DMs in multicolor experiments, leading to loss of emitted photons and hence reduced sensitivity. An alternative, elegant strategy is to spatially separate excitation from emission light in the absence of dichroic mirrors using micromirror TIRF (mmTIRF). In contrast to a DM-based setup, in a mmTIRF microscope the excitation beam is directed into and out of the back aperture of the MO by two small broadband mirrors placed right below the MO (Figure 2.5).

For all single-molecule assays shown in this thesis, we used co-localization single-molecule spectroscopy (CoSMoS) on a mmTIRF microscope (“Dobby”) based on a system from Mad City Labs with custom modifications similar as previously described (Larson et al., 2014) (Figure 2.5). All essential mmTIRF microscope parts are listed in Table 2.6. The mmTIRF microscope was equipped with four individual laser beams with wavelengths of 488, 532, 637 and 808 nm (OBIS 488nm LS 120mW, OBIS 532nm LS 120mW, OBIS 637 nm LX 100mW and 808 nm 250 mW single mode laser, respectively). All laser beams were first expanded and collimated with a Keplerian beam expander consisting of two lenses (L1 and L2). Attenuation of individual beams was achieved by multiple neutral-density filters (NDF) with each one reducing beam intensity by 10-fold. Next, quarter-wave plates (QWP) were incorporated to circularly polarize all beams which were subsequently passed through a respective band-pass filter (BPF; excitation filters) to cut each beam to its designated wavelength. The individual laser beams were then combined by an aligned broadband mirror (M1) and multiple long-pass dichroic mirrors (DM1-3; ZT488rdc, ZT532rdc and ZT647rdc dichroic mirror, respectively) which were selected for spectral properties to reflect and transmit wavelengths as required. The combined excitation beams were reduced to desired beam diameter with an iris (I1)

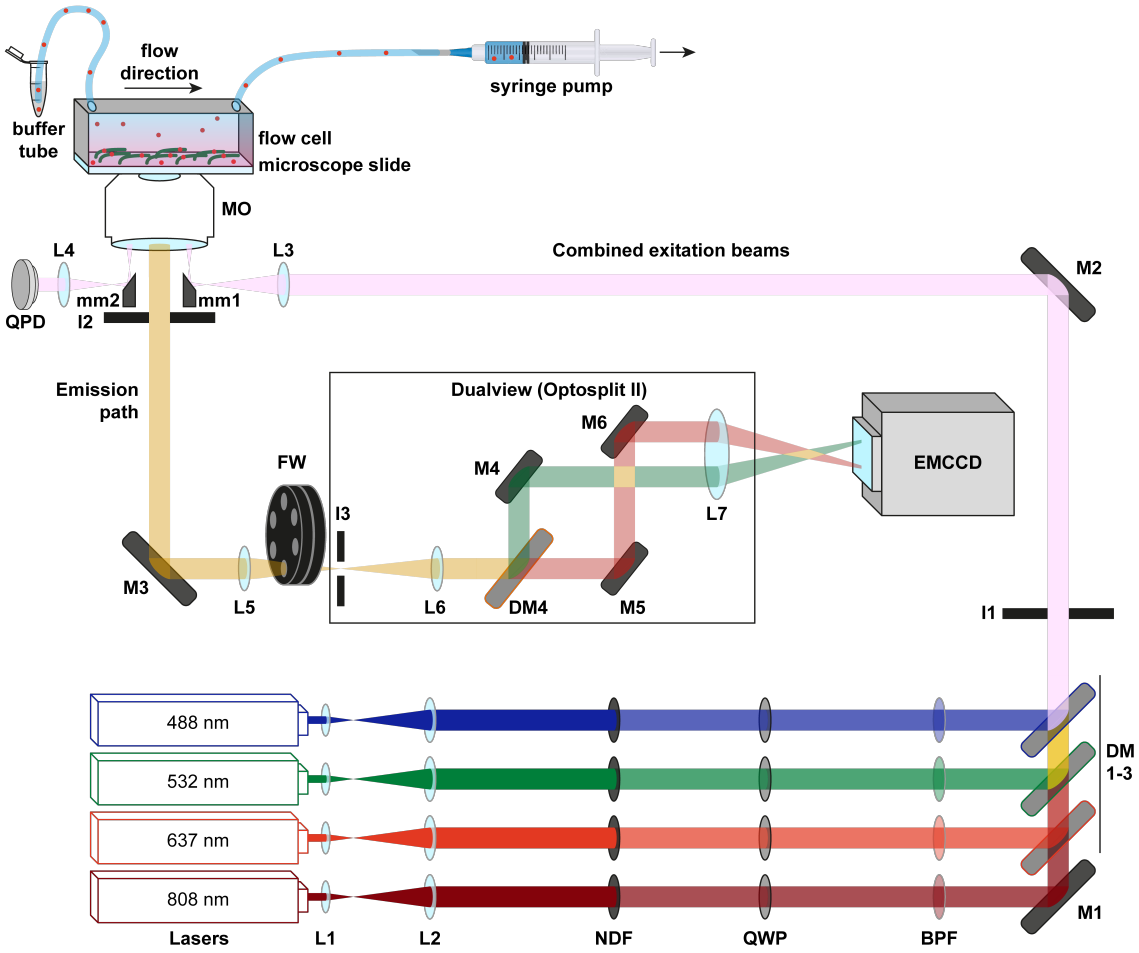

Figure 2.5. Schematic of micromirror TIRF microscope and flow cell setup
The micromirror TIRF microscope was set up as previously described (Larson et al., 2014) with custom modifications. Four individual laser beams (488, 532, 637 and 808 nm) were expanded and collimated by two lenses (L1 and L2). Neutral-density filters (NDF) attenuated individual laser beams. Quarter-wave plates (QWP) circularly polarized all beams which were passed through a band-pass filter (BPF). All laser beams were combined by a broadband mirror (M1) and long-pass dichroic mirrors (DM1-3) to reflect and transmit respective wavelengths. The beam diameter of combined excitation beams was reduced with an iris (I1) and directed to the entry micromirror (mm1) by a broadband mirror (M2). A lens (L3) focused excitation beams near the back aperture of the microscope objective (MO). Samples were mounted on a piezo-controlled microscope stage. A flow cell on a microscope slide (see chapter 2.6.2) is depicted with a buffer tube and a syringe pump connected to inlet and outlet tube, respectively. Exit beams were directed to a quadrant photo-diode (QPD) detector by the exit micromirror (mm2) which allowed for autofocusing. Scattered light from the MO was reduced with an iris (I2). A broadband mirror (M3) directed the emission light to a filter wheel (FW) with filters to remove residual scattered light. The emission light was collimated by two lenses (L5 and L6) and dualview was generated by the OptoSplit II system as follows: Iris (I3) defined the field of view. The emission light was split into a <635 nm and >635 nm channel by a long-pass DM (DM4). The two channels were directed to a lens (L7) by three broadband mirrors (M4, M5 and M6). To generate separated images for both channels, they were finally focused to different areas of the detector of an EMCCD camera.

and further directed to the entry micromirror (mm1) by another broadband mirror (M2). A final lens in the excitation pathway (L3) focused the combined excitation beams near the back aperture of the MO (Apo N TIRF 60 x oil-immersion TIRF objective, NA 1.49). Notably, mm1 position could be adjusted to change the effective TIRF angle. Samples

(here a flow cell is shown) were mounted onto a piezo-controlled microscope stage (part of Mad City Labs' mmTIRF system) allowing for positioning with low nanometer precision in three directions. The exit beams were collected by the exit micromirror (mm2) and directed to a quadrant photo-diode (QPD) detector (part of Mad City Labs' TIRF Lock system). The QPD was incorporated in the main microscope controlling software (Micromanager) and thus directly connected to the microscope stage which enabled computer-controlled autofocusing. Autofocusing was usually achieved with the 808 nm laser to avoid photobleaching of dyes with excitation at significant lower wavelengths.

To increase the signal-to-noise ratio of the emission fluorescence, scattered light from the MO in the emission path was reduced with an iris (I2) installed right below mm1 and mm2. Using another broadband mirror in a 45° mount (M3), the emission light was subsequently directed towards an Optospin25 filter wheel (FW) containing respective emission filter sets to further remove residual scattered light from the emission path (ET520/40m and ZET532/640m emission filters were used for 488 nm and 532/637 nm excitation, respectively). The emission light was collimated by a pair of lenses (L5 and L6) whereas the second one was part of the OptoSplit II dual emission image splitter system. This system contained additional optics to generate a dualview as follows: First, another iris (I3) placed right after the FW at an intermediate image plane was used to define the field of view. Second, the collimated emission light was split at 635 nm into two channels using a long-pass DM (DM4; T635lpxr dichroic mirror). The resulting <635 nm and >635 nm channels were directed to separate positions on the final lens (L7) by three additional broadband mirrors (M4, M5 and M6). Thus, the two channels were finally focused to different areas of the detector of an iXon Ultra 888 EMCCD camera which generated separated images for <635 nm and >635 nm channels.

All single-molecule experiments in this study were performed on the above described mmTIRF microscope in a temperature-controlled room at 22.5 ± 0.5 °C. To allow visualization of the entire DNA molecule, mm1 was adjusted to generate an angle to image slightly out of TIRF in the highly inclined and laminated optical sheet (HILO) mode. Alexa Fluor 488, JF549 / LD555 / SYTOX Orange and LD655 dyes were excited with a 488 nm, 532 nm and 637 nm laser respectively. In all multicolor experiments, dyes were excited sequentially, except for cohesin translocation assays in which JF549 and LD655 were imaged simultaneously. All proteins were visualized every 5-7 s for a period of 10-20 min (~150-200 frames) with following exceptions: Helicase sliding and photobleaching assays were performed at 2-4 fps with a 200 ms integration time (200 and 1500 frames, respectively). Cohesin translocation assays were performed at 6-7 fps with a 100 ms integration time for 220 seconds (1200 frames). During imaging, all microscope parts were controlled using Micromanager v1.4 for ImageJ (Edelstein et al., 2010; Schneider et al., 2012) and custom BeanShell scripts (Table 2.7).

2.7 Single-molecule data analysis

2.7.1 Raw data processing and organization in Molecule Archives

All single-molecule raw data were processed in Fiji (Schindelin et al., 2012) using Molecule ARchive Suite (MARS) commands (Agarwal and Duderstadt, 2020). First, individual channels were split and corrected for the laser beam profile to lower position-specific differences in fluorescence intensity. Second, to generate Single Molecule Archives, individual molecules were tracked with subpixel resolution with simultaneous integration of fluorescence intensity (inner radius: 2 pixels, outer radius: 4 pixels – for subtracting local background). Stuck molecules on the slide surface in the LD655 channel were used to accurately overlay post-stained DNA videos and to correct all channels for stage drift. Third, all individually separable DNA molecules were fit and checked for co-localization with individual molecule trajectories and finally organized into one combined DNA Molecule Archive. In cohesin translocation assays on licensed DNA, all doubly-tethered DNA molecules containing cohesin were chosen for further analysis. Correct tracking and co-localization with DNA were further evaluated manually.

For transcription experiments, protein position on DNA versus time was fit with a kinetic change point (KCP) algorithm (Hill et al., 2018). Individual regions were assigned to distinguish between different pushed proteins and numbers within one transcription trajectory.

All kymographs were generated using Fiji. For this, individual DNA ends were fit with subpixel localization and the kymograph was generated along the connecting line. In experiments using doubly-tethered DNA, individual DNA molecules were tethered with different extension to the slide surface and as a consequence, kymographs differ in heights. These length differences were accounted for throughout all analysis steps.

2.7.2 Collision outcomes for RNAP, ORC, MCM and nucleosome

Collision outcomes were determined manually by evaluating trajectories and raw videos. Proteins with starting positions closer than ~1 kb were excluded from collision analysis. Collisions were analyzed when proteins approached to <0.5 kb and classified as follows: Push – displacement >2 kb; bypass – displacement <2 kb with transcription >4 kb; pause – displacement >2 kb with transient pause within <2 kb upon collision; stall – displacement <2 kb with permanent pause; eject – displacement <2 kb and loss of protein signal. To analyze collisions of two MCM or nucleosome foci, survival of both foci were evaluated by the total fluorescence intensity after collision (which approximately corresponds to the sum of individual intensities prior to collision). Note that 3-color experiments with nucleosome collisions were analyzed independent of AF488 trajectories.

2.7.3 Stability at the T7 termination site

Stability at the T7 termination site was determined manually by evaluating trajectories and raw videos. Stability of MCM DHs and origin licensing intermediates was analyzed as follows: Once transcription-driven displacement stopped permanently at the T7 termination site (DNA region 12-15 kb considered), trajectories were classified into three groups. (1) Dissociate – protein stayed on DNA for less than 120 s. (2) Remain – protein stayed within less than 2 kb of T7 termination site for more than 120 s. (3) Slide back – protein stayed on DNA for more than 120 s but did not stay within less than 2 kb of the T7 termination site for more than 120 s. Note that 3-color experiments were analyzed independent of AF488 trajectories.

2.7.4 Spatial-temporal protein dynamics and kinetics

The Molecule Archives described above contained all information required for subsequent data analysis. Data were further analyzed directly from Molecule Archives with custom Python scripts and Jupyter notebooks (Table 2.7).

Protein loading sites were determined by their initial position on DNA. Transcript lengths and distances pushed were calculated by subtracting the protein loading site from the maximum detected position on DNA. Reported transcription rates correspond to burst rates which excluded region of transcription pauses (definition see below). To determine burst transcription rates, first, poorly fit segments derived from fitting with the KCP algorithm (Hill et al., 2018) were excluded (standard deviation >10 nt/s or rate <-10 nt/s). Second, all remaining segments showing a >3-fold reduced velocity compared to the mean of non-pause segments were classified as pause segments. Finally, burst transcription rates were calculated from the time-weighted average of non-pause segments. Pausing probability was derived from pauses occurring until position 19 kb on DNA (DNA substrates with termination sites were not analyzed for pausing probability) with segment lengths of >20 s. Upon pausing, if transcription failed to resume within the observation time, the pause was classified as permanent, otherwise as transient.

Pauses during cohesin translocation were also determined by fitting cohesin trajectories (position on DNA versus time) with the KCP algorithm (confidence value: 0.6; global sigma: 300 bps/s). Subsequently, resulting segments with rates <200 bps/s, standard deviations <30 bps/s and length >1 s were classified as pause segments. For both, transcription and cohesin pauses, if two adjacent segments were classified as pause and the end and start position on DNA of the first and second pause segment, respectively, were within 1 kb, these segments were merged to one pause segment. These pauses were excluded when calculating diffusion coefficients (see below) and cohesin-MCM passing probabilities (see chapter 2.7.6).

Experimentally determined diffusion coefficients (D) were calculated with

$$D = \frac{< x >^2}{2\,t} \tag{1}$$

where $<x>^2$: mean squared displacement (kbp^2 or μm^2) and t: time (s).

The theoretical upper limit of non-helical ($D_{non\text{-}hel}$) and helical (D_{hel}) diffusion was estimated using Stokes-Einstein equation (Schurr, 1979):

$$D_{non-hel} = \frac{k_B\,T}{\xi_{trans}} = \frac{k_B\,T}{6\,\pi\,\eta\,r} \tag{2}$$

$$D_{hel} = \frac{k_B\,T}{\xi_{trans} + \xi_{rot}} = \frac{k_B\,T}{6\,\pi\,\eta\,r + \left(\frac{2\,\pi}{10.5\,bp \cdot 0.34\ nm/bp}\right)^2 \cdot 8\,\pi\,\eta\,r^3} \tag{3}$$

where k_B: Boltzmann constant, T: absolute temperature, $\xi_{trans/rot}$: translational / rotational friction coefficient, η: dynamic viscosity and r: radius of protein (along diffusion axis).

The dynamic viscosity of buffer was approximated by calculating the dynamic viscosity of water using Vogel-Fulcher-Tammann-Hesse equation:

$$\ln \eta = A + \frac{B}{T - T_0} \tag{4}$$

where η: dynamic viscosity of water (mPa s); A, B and T0 are constants (-3.7188, 578.919 K and 137.546 K for water, respectively).

The population variance (σ^2) to distinguish DNA-bound from surface-stuck proteins was calculated with

$$\sigma^2 = \frac{\sum_{i=1}^{N}(x_i - \mu)^2}{N} \tag{5}$$

where x_i: value of i^{th} element, μ: population mean and N: population size.

2.7.5 Fluorescently labeled protein stoichiometry

Labeled ORC, MCM, T7 RNAP and nucleosome stoichiometry on DNA was determined by photobleaching experiments as described in chapter 2.6.8. Subsequently, photobleaching steps were fit with the KCP algorithm (Hill et al., 2018). In transcription-helicase collision assays, the number of MCM DHs in an MCM foci was estimated based on their initial fluorescence intensity and the mean fluorescence intensity of one MCM DH obtained from photobleaching analysis. A similar procedure was applied to estimate the number of nucleosomes to detect more than one nucleosome within the diffraction limit (mostly in “high density” chromatin). In cohesin translocation assays on licensed

DNA, the number of MCMs per foci was directly determined from these datasets as the imaging time was sufficient to allow for efficient photobleaching.

2.7.6 Cohesin-MCM bypassing probability

The probability of cohesin bypassing MCM was addressed as follows: Frames in which cohesin co-localized with MCM (median position) within less than $thresh1_{dynamic}$ (see equation 6) were classified as an encounter. Upon an encounter, if cohesin passed MCM in the consecutive frame by at least $thresh2_{dynamic}$ (see equation 6), the encounter was determined as successful bypassing. All remaining frames (distance > $thresh1_{dynamic}$ to MCM) were further evaluated for MCM bypassing as described above and additionally counted as encounter with successful bypassing. DNA molecules with only cohesin and no MCM were analyzed the same way using the theoretical ARS1 position on DNA. All frames within the cohesin trajectory which were part of a translocation pause (identified as described in chapter 2.7.4) were excluded from this analysis described above and instead classified as one encounter with failed bypassing.

To account for different resolution at different extensions, two dynamic thresholds ($thresh1_{dynamic}$ or $thresh2_{dynamic}$), adjusted for the individual length of the DNA molecule were used and determined with:

$$thresh1/2_{dynamic} = \frac{thresh1/2 \cdot length_{global}}{length_{molecule}} \quad (6)$$

where $length_{global}$: mean length of all DNA molecules (pixels); $length_{molecule}$: length of the DNA molecule (pixels) and thresh: set threshold at $length_{global}$ (thresh1 = 1.5 kb and thresh2 = 0.5 kb).

2.7.7 Quantification and statistical analysis

The number of observations (n) analyzed is indicated in the figure or figure legends. Errors reported in this study represent the estimated standard error of the mean (SEM) or a 95 % confidence interval (CI) determined from 10000 cycles of bootstrapping except for errors of rates which represent the standard deviation (SD) derived from a Gaussian fit as mentioned in the figure legends or text. A detailed representation of combined datasets to generate each figure panel is outlined in the supplied Jupyter notebooks (Table 2.7).

Chapter 3

Reconstitution of Origin Licensing at the Single-Molecule Level

In 1963, François Jacob, Sydney Brenner and François Cuzin proposed their theory on how chromosome replication in bacteria is initiated and coordinated with respect to cell cycle and division, since then known as the “replicon model” (Jacob et al., 1963). The core claim of the replicon model, postulating the existence of specific DNA sequence elements, termed replicators, that serve as start sites through engagement of an initiator protein, turned out to be true not only for bacteria but also for phages, archaea and even for eukaryotes to some extent (Bramhill and Kornberg, 1988; Kowalski and Eddy, 1989; Mackiewicz et al., 2004; Myllykallio et al., 2000; Stinchcomb et al., 1979; Weigel and Seitz, 2006). However, in contrast to bacteria, in which these replicator sequences (origins) are well-defined, eukaryotic origins are far more diverse and difficult to classify. In *S. cerevisiae*, origins are largely defined by ARS elements, but in other eukaryotes no specific sequences have been found; instead chromatin structure and epigenetic marks are the defining features of origins (Ganier et al., 2019; Prioleau and MacAlpine, 2016).

Origin licensing, replisome assembly and firing has been long viewed as a complex but highly-coordinated, static sequence of events. The temporal gap between licensing and firing provides a window of time during which additional events on the chromosome can influence origin specificity and the process of replication initiation. A more flexible and dynamic view of assembly pathways of macromolecular machines has emerged over the last decade, providing a framework on how replication initiation could accommodate other processes operating on the same chromosome (Chen et al., 2014; Mirny et al., 2009). It remains completely unclear how origin licensing is altered upon encounters with RNA polymerases. Studies suggested that MCMs can reposition to sites away from origins under altered salt conditions (Remus et al., 2009) or when directly challenged by multiple rounds of loading (Douglas et al., 2018; Gros et al., 2015), but the stability and mechanics of these events is unknown. Thus, we developed a TIRF-based single-molecule assay to directly visualize dynamic events during origin licensing in real-time.

3.1 MCM loading occurs at ARS1 and requires ATP hydrolysis

To reconstitute origin licensing at the single-molecule level, we first purified recombinant origin licensing factors ORC, Cdc6 and Cdt1-MCM and introduced a site-specific fluorophore at Orc1 (LD555-ORC) and Mcm6 (Cdt1-LD655-MCM) (Figure 3.1A). Next, we used an ensemble helicase loading assay with a 5 kb long ARS1-containing DNA bound to a magnetic bead (Figure 3.1B) to confirm the functionality of our recombinant origin licensing factors. In line with previous findings, in the presence of ORC, Cdc6 and ATP, MCM was efficiently retained on DNA and resistant to high salt challenge (Figures 3.1C and 3.1D), a characteristic of successful Cdt1 release and ring closure in conjunction with ATP hydrolysis (Evrin et al., 2009; Remus et al., 2009). Importantly, we did not detect any difference in loading efficiency between fluorescently labeled and

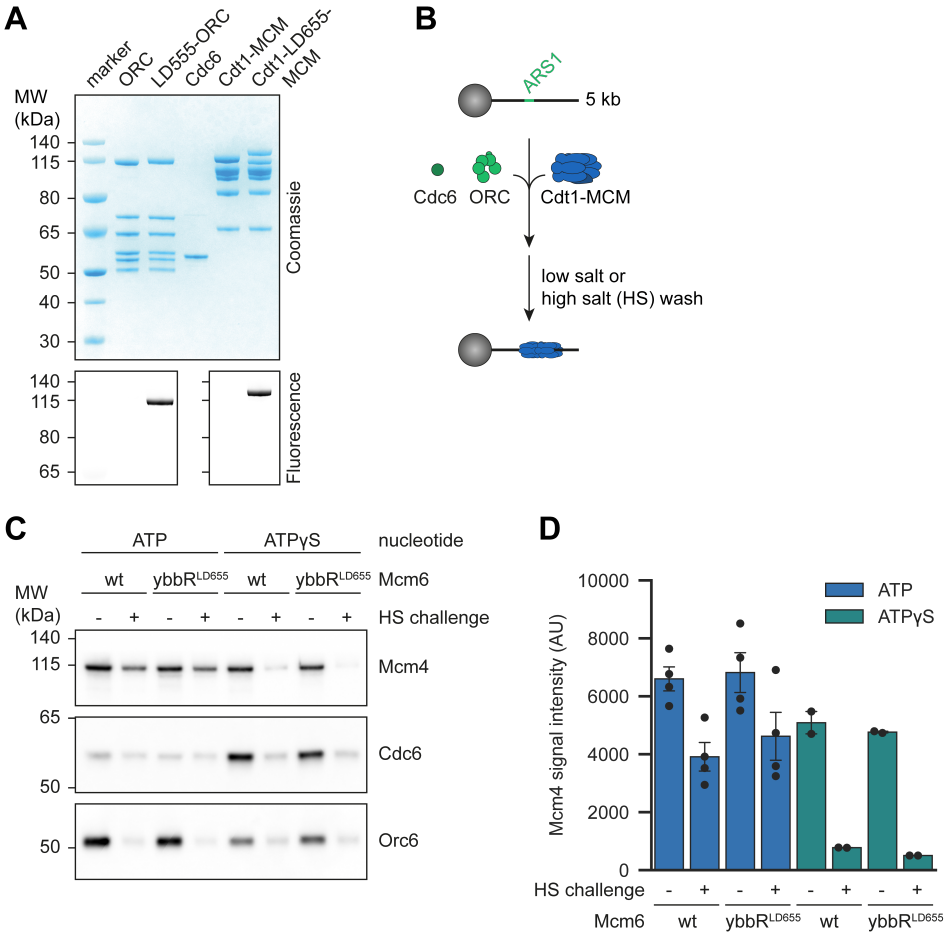

Figure 3.1. MCM DHs form on DNA in an ATP-dependent manner in ensemble assays
(**A**) SDS-PAGE analysis of purified fluorescently labeled and unlabeled licensing factors by Coomassie Blue staining and fluorescence detection. (**B**) Schematic of the ensemble helicase loading assay. ARS1-containing 5 kb DNA bound to a magnetic bead was incubated with licensing factors ORC, Cdc6 and Cdt1-MCM and subsequently washed with a low or high salt (HS) buffer. (**C** and **D**) Ensemble helicase loading assay as described in (**B**) in the presence of ATP or ATPγS. (**C**) Eluted protein was analyzed by Western Blot using protein specific primary antibodies for Orc6, Cdc6 and Mcm4. (**D**) Integration of Mcm4 signals showed no difference in loading efficiency for Cdt1-MCM containing wildtype (wt) or ybbR-LD655 Mcm6 (n = 4 and n = 2 for ATP and ATPγS condition, respectively). Error bars represent SEM.

wildtype MCMs in our ensemble assays (Figures 3.1C and 3.1D). Thus, we concluded that our recombinant origin licensing factors were functional and excluded interference of introduced fluorophores with the origin licensing process.

Having established efficient origin licensing in ensemble assays, we next developed a TIRF-based single-molecule assay to directly visualize dynamic events during origin licensing in real-time. Origin licensing at the single-molecule level was reconstituted using purified components in a stepwise manner (Figure 3.2A). First, large ARS1-containing DNA molecules, 21 kb in length, were immobilized with one end on the surface of functionalized coverslips via biotin-streptavidin-biotin interactions. Second, purified licensing factors ORC, Cdc6 and Cdt1-MCM (Figure 3.1A) were introduced to load MCM on dsDNA. Finally, the dynamics of licensing factors were temporally and spatially resolved on flow-stretched DNA molecules.

In the presence of ORC and Cdc6, fluorescently labeled MCMs readily co-localized with dsDNA in our single-molecule setup (Figure 3.2B). Most MCM complexes loaded specifically at the replication origin ARS1, coincident with fluorescently labeled ORC, but a subpopulation of complexes was observed at alternative sites throughout the DNA (Figure 3.2C), in line with previous studies showing MCM loading in the absence of origin sequences *in vitro* (Gros et al., 2014; On et al., 2014). Importantly, ORC has been shown to preferentially bind and initiate DNA replication at AT-rich regions, not only in *S. cerevisiae* with well-defined ARS sequences meeting this preference, but also in *S. pombe*, *X. leavis* and *D. melanogaster* for instance (Austin et al., 1999; Broach et al., 1983; Kong et al., 2003; Lee and Bell, 1997; Theis and Newlon, 1997). In line with these observations, the high observed origin specificity in our setup resembles the balanced GC-content of our DNA substrate which only contained a pronounced AT-rich region at the introduced ARS1 site (Figure 3.2D). Robust loading of MCMs was again verified by resistance to high salt challenge. Interestingly, differences in retained MCMs were only observed between low and high salt challenge in ensemble assays using shorter DNA substrates but not in single-molecule experiments (Figures 3.1C, 3.1D and 3.2E), indicating more robust MCM retention on longer DNAs.

Despite still ongoing debates about the exact underlying mechanism of MCM DH formation, a loading intermediate of OCCM could be confirmed by various studies (Miller et al., 2019; Ticau et al., 2015; Yuan et al., 2017). The OCCM complex occurs prior to ATP hydrolysis which triggers Cdc6 and Cdt1 release and hence stable association with dsDNA. Recently, Miller and coworkers elegantly identified additional intermediates appearing before and after OCCM, which reconciles contradicting models of MCM DH formation (Coster and Diffley, 2017; Miller et al., 2019; Ticau et al., 2015). To study OCCM dynamics, we performed MCM loading reactions in the presence of ATPγS, halting origin licensing at the OCCM state. As expected, OCCM formation also occurred preferentially at ARS1 (Figure 3.2F), however observed loading efficiencies were reduced (Figures 3.1C, 3.1D and 3.2E), indicating less stable association with DNA compared to MCM DHs. Importantly, in line with ensemble experiments and previous

studies (Evrin et al., 2009; Remus et al., 2009), only a small fraction of OCCM complexes survived high salt challenge, confirming the incomplete loading of MCM around DNA in the OCCM (Figures 3.1C, 3.1D and 3.2E). Establishing not only MCM DH but also OCCM formation at the single-molecule level, allowed us to investigate their dynamic properties.

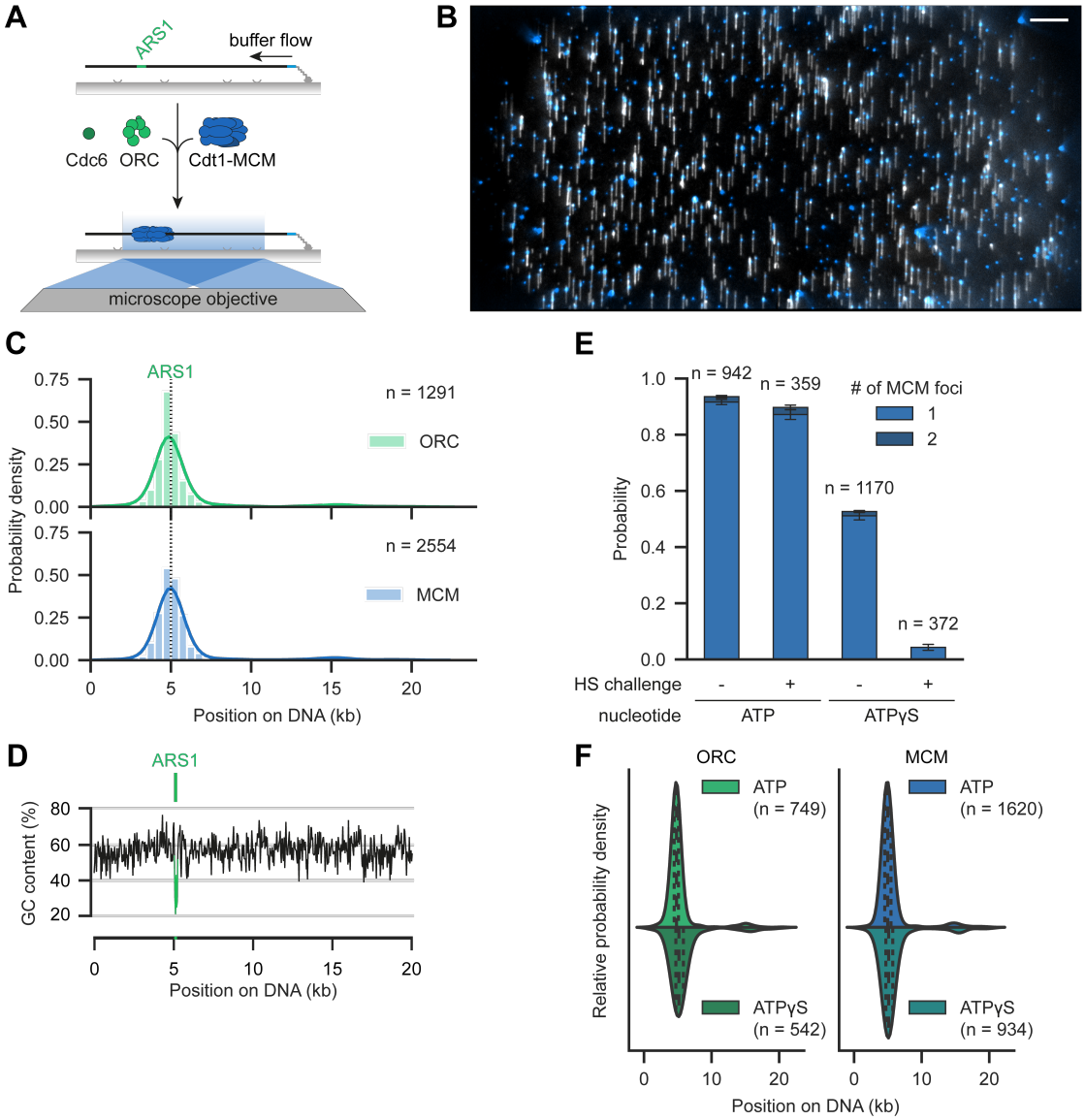

Figure 3.2. MCM loading occurs at ARS1 and requires ATP hydrolysis

(**A**) Schematic of the single-molecule helicase loading assay. ARS1-containing 21 kb DNA was incubated with licensing factors ORC, Cdc6 and Cdt1-MCM and imaged on flow-stretched DNA after removal of excess protein. (**B**) Representative field of view showing MCM (blue) co-localization with ARS1-containing dsDNA (gray) after origin licensing reaction. Scale bar represents 10 µm. (**C**) ORC (green) and MCM (blue) binding distribution on ARS1-DNA. Data from all experiments irrespective of ATP or ATPγS are shown. Lines represent the kernel density estimation (KDE). (**D**) GC content plot of ARS1-DNA used in single-molecule assays. ARS1 position is annotated and highlighted in green. (**E**) Number of MCM foci on ARS1-DNA challenged with low or high salt (HS) after helicase loading in the presence of ATP or ATPγS. Error bars represent the estimated SEM by bootstrapping. (**F**) ORC and MCM binding distribution by KDE on ARS1-DNA after helicase loading in the presence of ATP or ATPγS. Lines represent the quartiles of the distribution with the middle line corresponding to the median. Data depicted are the same as shown in (**C**).

3.2 Origins are sequestered by multiple MCMs during licensing

Initiation of DNA replication is widely known to involve loading of a single MCM DH on replication origins during G1 with subsequent activation in S phase. However, studies in *X. laevis*, *S. cerevisiae* and humans showed matching numbers for chromatin bound ORC and replication initiation events but detected a vast excess of MCMs on chromatin, a phenomenon known as the “MCM paradox” (Burkhart et al., 1995; Edwards et al., 2002; Lei et al., 1996; Mahbubani et al., 1997). In line with these observations, further studies suggested that one ORC might be competent for multiple rounds of MCM loading (Bowers et al., 2004).

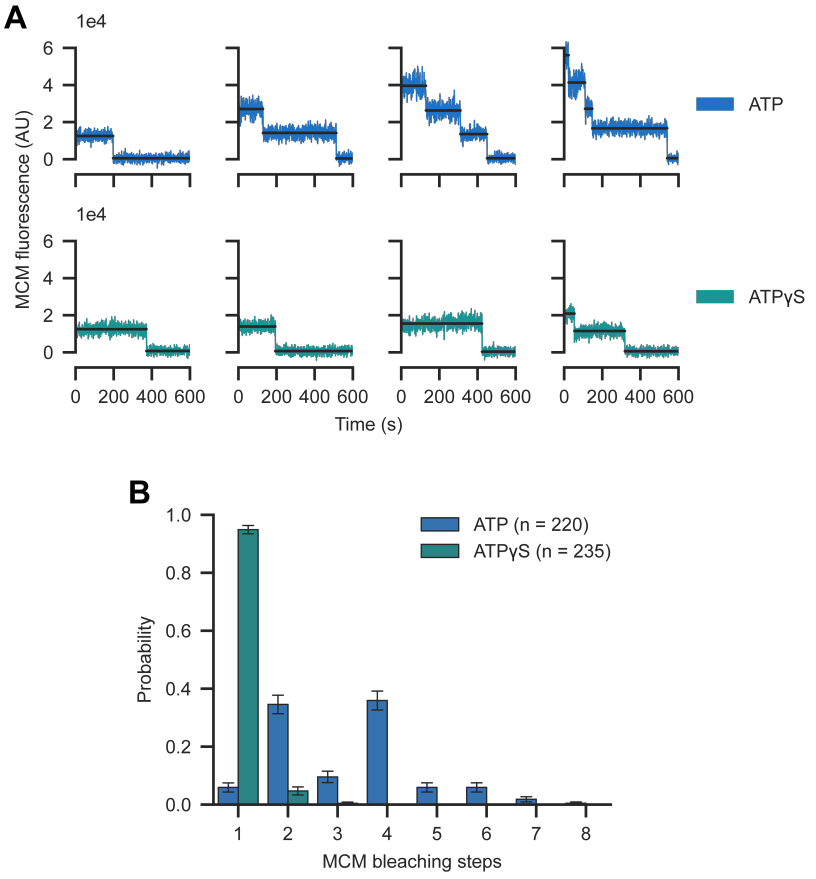

Figure 3.3. Origins are sequestered by multiple MCMs during licensing
(**A** and **B**) Example MCM bleaching trajectories (**A**) and distribution of MCM bleaching steps (**B**) after helicase loading in the presence of ATP (MCM DH formation) or ATPγS (OCCM formation). Black lines in (**A**) correspond to the fit with the kinetic change point algorithm. Error bars in (**B**) represent the estimated SEM by bootstrapping.

To investigate this possibility, we counted the number of loaded MCMs by photobleaching. Interestingly, besides the expected single MCM DH per origin, we also detected a distinct population harboring two MCM DHs at the origin or even higher order complexes in the presence of ATP (Figure 3.3). To rule out the possibility of either non-specific MCM oligomerization or parallel, independent loading events around the ARS1 within the diffraction limit in our assay, we performed the same experiment in the

presence of ATPγS. Strikingly, we almost exclusively observed single-step photobleaching, consistent with the formation of a single OCCM per origin (Figure 3.3). Only a small fraction (~5 %) showed two-step photobleaching, which may be due to ATP contamination of the ATPγS lot that allows limited MCM DH formation or indicate symmetrical MCM recruitment to ARS1 (Coster and Diffley, 2017; Miller et al., 2019). These data demonstrate that multiple MCM DHs are loaded in an origin-dependent manner in which subsequent MCMs can only be loaded once the previous MCM DH clears the origin.

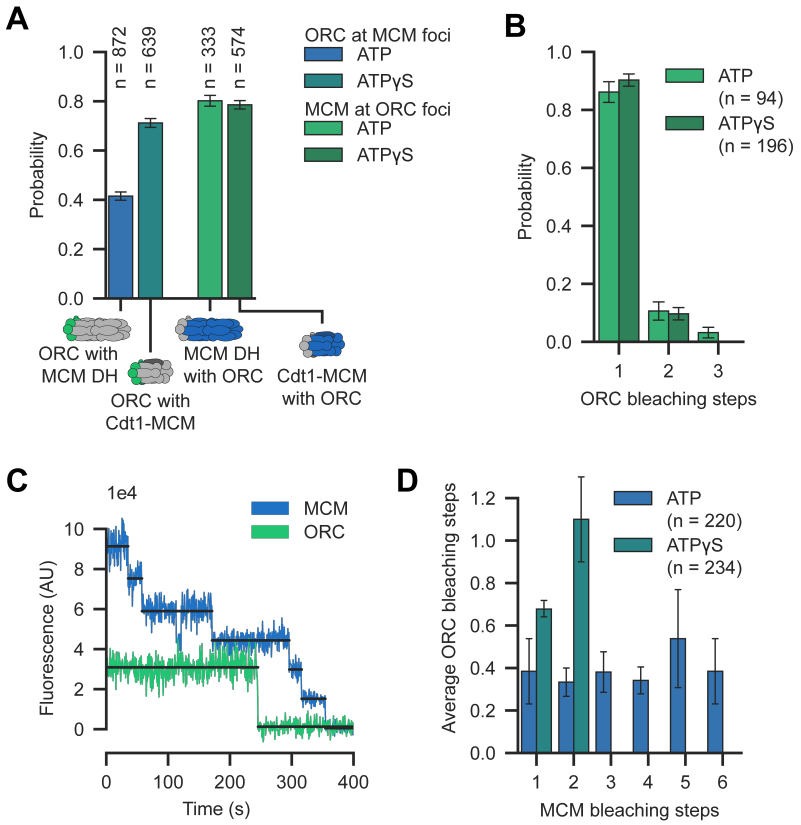

Figure 3.4. A single ORC is competent for multiple rounds of MCM DH loading at the origin
(**A**) Probability of MCM foci co-localizing with ORC and ORC foci co-localizing with MCM after helicase loading in the presence of ATP or ATPγS. (**B**) Distribution of ORC bleaching steps after helicase loading in the presence of ATP or ATPγS. (**C**) Example bleaching trajectory of ATP-loaded, co-localizing MCM and ORC showing six- and one-step bleaching, respectively. Black lines correspond to the fit with the kinetic change point algorithm. (**D**) Average number of ORC molecules co-localizing with MCM foci containing one to six (ATP) or one and two (ATPγS) MCM after helicase loading in the presence of ATP or ATPγS. Error bars in (**A**), (**B**) and (**D**) represent the estimated SEM by bootstrapping.

We further explored the mechanism of multiple MCM DH loading using labeled ORC. While only ~40 % of MCM foci co-localized with ORC in the presence of ATP, we detected ~70 % in the presence of ATPγS (Figure 3.4A), consistent with a more stable engagement of ORC and MCM within the OCCM complex. In contrast, ~80 % of all DNA-bound ORCs recruited MCM in both conditions, showing high efficiency of MCM recruitment. We also predominantly detected single-step photobleaching for ORC in both conditions (Figure 3.4B) and no correlation (Pearson r = 0.012) between MCM and ORC

bleaching (Figures 3.4C and 3.4D) for the ATP condition. Intriguingly, using ATPγS, MCMs showing two-step bleaching were more likely to be associated with ORC, further supporting a small fraction with symmetrical MCM loading. Together, these data support a model in which a single ORC is competent for MCM DH formation (Ticau et al., 2015) and potentially also for multiple rounds of MCM loading.

3.3 MCM DHs can switch to a diffusive DNA binding mode

Our data presented in chapter 3.2 suggested that upon completion of origin licensing an MCM DH can be repositioned to permit the loading of an additional MCM DH at the origin. EM studies have shown the formation of trains of MCMs on DNA under conditions that favor multiple loading events at or outside the origin (Douglas et al., 2018; Hill et al., 2020; Remus et al., 2009). Moreover, the suggested two ORC model of origin licensing involves formation of symmetrical helicase loading followed by short-range sliding to finally form activation-competent MCM DHs (Coster and Diffley, 2017). A recent EM study by Thomas Miller and coworkers identified additional licensing intermediates in the licensing process and suggested that short-range sliding in the range of tens of bps of a single MCM intermediate was required to recruit the second MCM (Miller et al., 2019). Additionally, recent studies have also detected an ability of the CMG helicase complex to slide along DNA (Douglas et al., 2018; Wasserman et al., 2019). Nevertheless, the

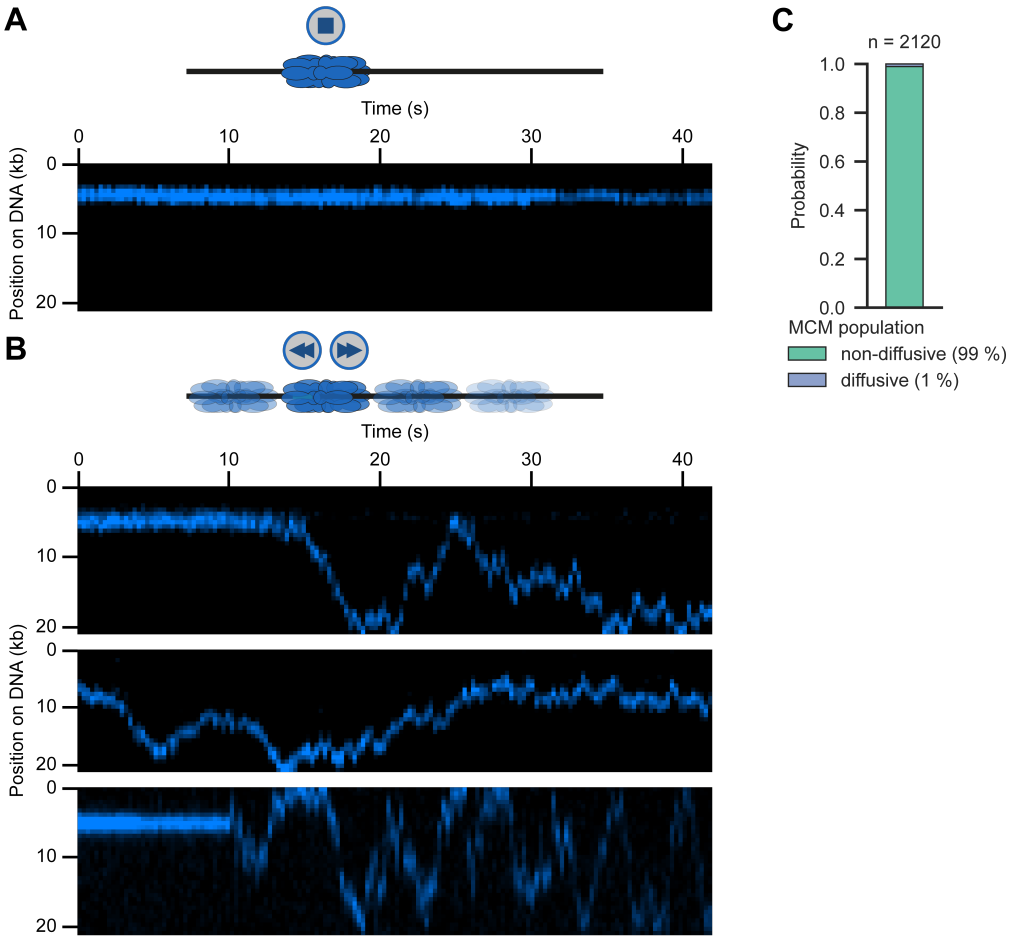

Figure 3.5. MCM DHs can switch to a diffusive DNA binding mode
(**A** and **B**) Representative kymographs showing MCM DH dynamics in the presence of 0.5 M NaCl. Most MCM DHs remained stably bound to the origin ARS1 (**A**) but a subset switched to a diffusive DNA binding mode (**B**). (**C**) Observed populations of MCM DHs in the presence of 0.5 M NaCl showed 99 % and 1 % in non-diffusive and diffusive DNA binding mode, respectively.

dynamics of MCM sliding have not been well characterized owning to lack of direct observations.

Remarkably, we observed very rare spontaneous sliding events at the completion of origin licensing under normal buffer conditions. We sought to increase the frequency of sliding (diffusive) events by continuously exposing MCM DHs to high salt (Remus et al., 2009). To eliminate the influence of external forces, we tethered DNA at both ends in an orientation-specific manner to image MCM mobility in the absence of buffer flow. Although the majority of MCMs remained stationary (non-diffusive) under these conditions (Figure 3.5A; Movie 1), the fraction of diffusive MCMs increased to 1 % (Figures 3.5B and 3.5C; Movie 2).

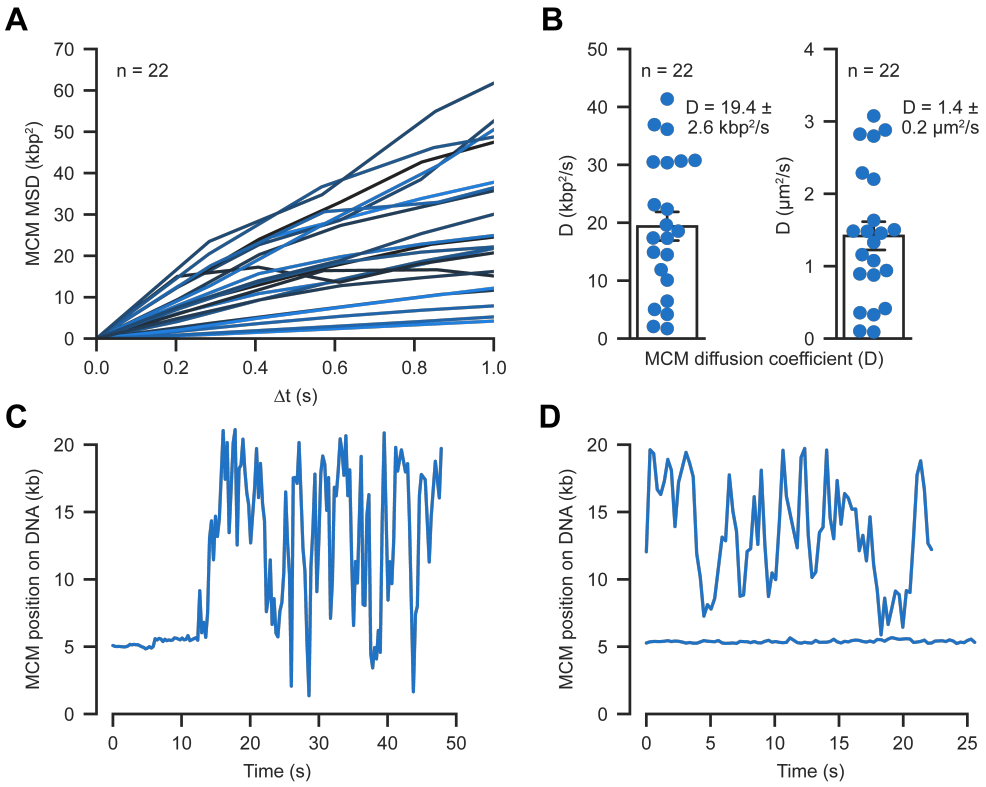

Figure 3.6. Diffusive MCM DHs approach the theoretical upper limit of diffusion
(**A**) Plot showing the mean squared displacement (MSD) vs Δt of 22 diffusive MCM DHs in the presence of 0.5 M NaCl. (**B**) MCM diffusion coefficients in the presence of 0.5 M NaCl. Analysis of 22 molecules gave D = 19.4 ± 2.6 kbp^2/s or 1.4 ± 0.2 μm^2/s (mean ± SEM). Error bars represent the estimated SEM by bootstrapping. (**C** and **D**) Example tracking results of diffusive MCM DHs in the presence of 0.5 M NaCl. (**C**) ARS1-bound MCM DHs spontaneously switched to a diffusive DNA binding mode. (**D**) DNA molecule containing two MCM DHs suggests ARS1-bound MCM DHs being a barrier to diffusive MCM DHs.

In general, diffusive translocation of a protein on dsDNA can be mediated by an actual sliding but also hopping mechanism (Berg et al., 1981). During actual sliding, the protein continuously interacts with dsDNA through electrostatic interactions and tracks the helical pitch, leading to a movement with coupled rotation and translocation (helical diffusion). In contrast, during hopping, the protein transiently loses all electrostatic interactions with dsDNA, resulting in translocation by frequent dissociation and rebinding

events without coupled rotation (non-helical diffusion). Higher salt concentration thereby increases the frequency of dissociation, extends the diffusion time before rebinding and thus increases the diffusion coefficient of proteins undergoing translocation based on the hopping but not the sliding mechanism (Givaty and Levy, 2009).

To characterize the physical properties of diffusive MCM DHs, we next calculated the mean squared displacement (MSD) of 22 diffusive MCM molecules (Figure 3.6A). The linear relationship between MSD and Δt observed for the majority (21/22) of molecules indicates that the diffusive DNA binding mode of MCM DHs is a random walk. Using linear extrapolation of calculated MSD of diffusive MCM molecules (Figure 3.6A), we determined the diffusion coefficient to 19.4 ± 2.6 kbp^2/s or 1.4 ± 0.2 $\mu m^2/s$ (mean ± SEM) at room temperature but observed a wide range of velocities ranging from to 2-40 kbp^2/s or 0.1-3 $\mu m^2/s$ (Figures 3.5B, 3.6B, 3.6C and 3.6D). Interestingly, in these conditions, diffusive MCM DHs approach the theoretical upper limit of free, non-helical diffusion of ~35 $\mu m^2/s$ and exceed the limit of helical diffusion of ~0.2 $\mu m^2/s$ by almost an order of magnitude (based on an MCM DH diameter of 13 nm along the diffusion axis (Li et al., 2015) and the viscosity of water). Thus, we speculate that diffusive MCM DHs can translocate on dsDNA through a hopping mechanism and that this might be combined with sliding as suggested by the broad distribution of observed diffusion coefficients with a slower population (Figure 3.6B). Intriguingly, a combination of hopping and sliding was also observed for diffusive PCNA (Kochaniak et al., 2009). MCM DHs topologically encircle dsDNA, a hopping mechanism, however, is based on full dissociation from dsDNA (by loss of all electrostatic interactions). This would require opening of both MCM rings which is rather unlikely given the high stability of MCM DHs on DNA. Nevertheless, the difference between the diameter of dsDNA (2.0 nm) and the diameter of the central channel in an MCM DH (3.0-4.0 nm) (Li et al., 2015) is likely to allow for disruption of all electrostatic interactions between MCM and DNA by shielding via respective counterions.

Table 3.1. Diffusion coefficients (D) of DNA-binding proteins along DNA

Protein	Salt / molarity	D ($\mu m^2/s$) (mean ± SEM)	Reference
Cohesin (+ Quantum Dot)	100 mM KCl	0.95 ± 0.20	(Stigler et al., 2016)
	500 mM KCl	3.8 ± 0.2	(Stigler et al., 2016)
Cohesin	150 mM NaCl	1.2 ± 0.1	This study (chapter 5.2)
	500 mM NaCl	3.2 ± 0.2	This study (chapter 5.2)
	750 mM NaCl	1.7 ± 0.1	(Davidson et al., 2016)
MCM double-hexamer	500 mM NaCl	1.4 ± 0.2	This study
PCNA	150 mM KGlu	1.2 ± 0.1	(Kochaniak et al., 2009)
p53	25-200 mM KCl	0.22 ± 0.18	(Tafvizi et al., 2008)
CMG	50 mM KGlu	0.18 ± 0.05	(Wasserman et al., 2019)

Notably, MCM diffusion was also an order of magnitude faster than observed for CMG helicase (and p53), albeit at lower ionic strength, but comparable to other dsDNA binding or encircling proteins sliding on DNA like cohesin or PCNA (Table 3.1) (Davidson et al., 2016; Kochaniak et al., 2009; Stigler et al., 2016; Tafvizi et al., 2008; Wasserman et al., 2019). However, further studies are required to characterize the underlying mechanism in detail, e.g. the effect of salt concentration on the diffusion coefficient, which we did not determine due to the low fraction of diffusive MCMs even at 0.5 M NaCl. Interestingly, on some DNA molecules one MCM DH remained tightly bound to ARS1 while a second MCM DH rapidly diffused along DNA (Figure 3.6D) consistent with the idea that sliding is required for loading of multiple MCMs. However, MCM sliding was observed in a very small population of molecules at non-physiologically high salt concentration, so questions remained open about the relevance of this activity under more physiological conditions.

Chapter 4

Mobile Origin Licensing Factors Confer Resistance to Transcription Conflicts

MCM sliding was observed in a very small population of molecules, which raised questions about the relevance of this activity under more physiological conditions and in the context of a more complex DNA substrate. Previous studies have demonstrated that MCM DHs can be pushed along DNA after collision with other DNA translocases, including RNA polymerase (RNAP), CMG helicase, and FtsK, suggesting that MCM DH mobility contributes to the maintenance of origin activity in the face of chromosomal traffic (Douglas et al., 2018; Gros et al., 2015). However, it has not been possible in these ensemble experiments to quantitatively assess the impact of MCM DHs on the progression of other DNA-translocating protein machineries or, vice versa, the impact of colliding DNA translocases on MCM DH mobility. Moreover, the effect of chromatin on those events has also not been addressed. We, therefore, set out to investigate controlled RNAP-MCM DH collisions on both bare and chromatinized DNA templates under single-molecule conditions.

4.1 Time-coordinated single-molecule transcription

To directly evaluate the consequences of encounters between RNAP and MCM DHs, we developed a time-coordinated, single-molecule transcription assay using T7 RNAP as a model (Figure 4.1A). To this end, we introduced a T7 promoter (T7P) sequence upstream of ARS1 into our DNA substrate. Notably, we engineered the sequence downstream of the T7 promoter to contain the first deoxythymidine at position +16, allowing transcription up to 15 nucleotides in the absence of UTP. This design was guided by structural studies, suggesting that transcription of 15 nucleotides is sufficient for a single RNAP to switch from an unstable transcription initiation to a stable transcription elongation complex which can be stalled and restarted based on nucleotide availability

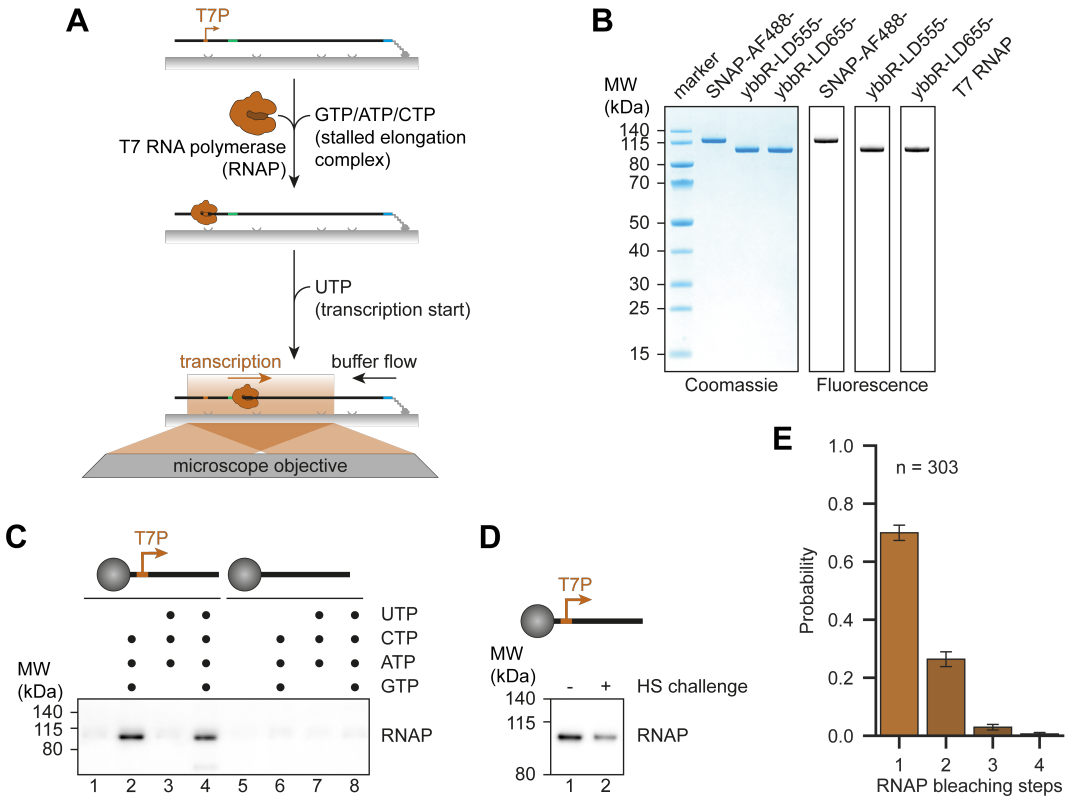

Figure 4.1. Time-coordinated single-molecule transcription

(**A**) Schematic of the time-coordinated single-molecule transcription assay. T7 promoter (T7P)-containing 21 kb DNA was incubated with T7 RNA polymerase (RNAP) in the presence of GTP, ATP and CTP, forming stalled elongation complexes. Transcription was started after UTP addition and imaged on flow-stretched DNA. (**B**) SDS-PAGE analysis of purified SNAP-AF488-, ybbR-LD555- and ybbR-LD655-T7 RNAP by Coomassie Blue staining and fluorescence detection. (**C** and **D**) Ensemble RNAP loading assay on bead-bound 5 kb linear DNA with (**C**, lanes 1-4 and **D**) or without (**C**, lanes 5-8) an integrated T7P. Eluted RNAP was analyzed by Western Blot using a primary antibody against His tag present in RNAP. (**C**) Formation of stable RNAP elongation complexes was only observed in the presence of an integrated T7P and GTP, ATP and CTP (lanes 2 and 4) but not in the absence of T7P or GTP (lanes 1, 3, 5-8). (**D**) A large fraction of formed RNAP elongation complexes showed resistance to a high salt (HS) challenge (lane 2) when compared to complexes challenged with low salt (lane 1). (**E**) Distribution of RNAP bleaching steps for those loaded at T7P as stalled elongation complexes. Error bars represent the estimated SEM by bootstrapping.

(Tahirov et al., 2002; Yin and Steitz, 2002). Thus, this strategy should provide the possibility to simultaneously load RNAP and MCM DHs on the same DNA molecule, to subsequently resume transcription, and cause collisions in a time-coordinated manner.

First, we engineered different versions of site site-specifically labeled T7 RNAP which were used interchangeably in all single-molecule assays (Figure 4.1B). In ensemble assays with 5 kb long DNA substrates, RNAP only engaged with DNAs containing the promoter but independent of the presence of UTP, confirming formation of stable elongation complexes according to our design (Figure 4.1C). Interestingly, a large fraction of RNAP elongation complexes was resistant to a high salt challenge, revealing a remarkable stability of RNAP upon switching to its elongation mode (Figure 4.1D). In our single-molecule setup (Figure 4.1A), photobleaching experiments of RNAP foci upon RNAP loading in the presence of GTP, CTP and ATP but in the absence of UTP revealed that the majority only contained one molecule (Figure 4.1E), confirming formation of single elongation complexes on our DNA substrates. Importantly, RNAP foci were almost exclusively bound to the T7 promoter (Figure 4.2B) as expected due to two specificity-

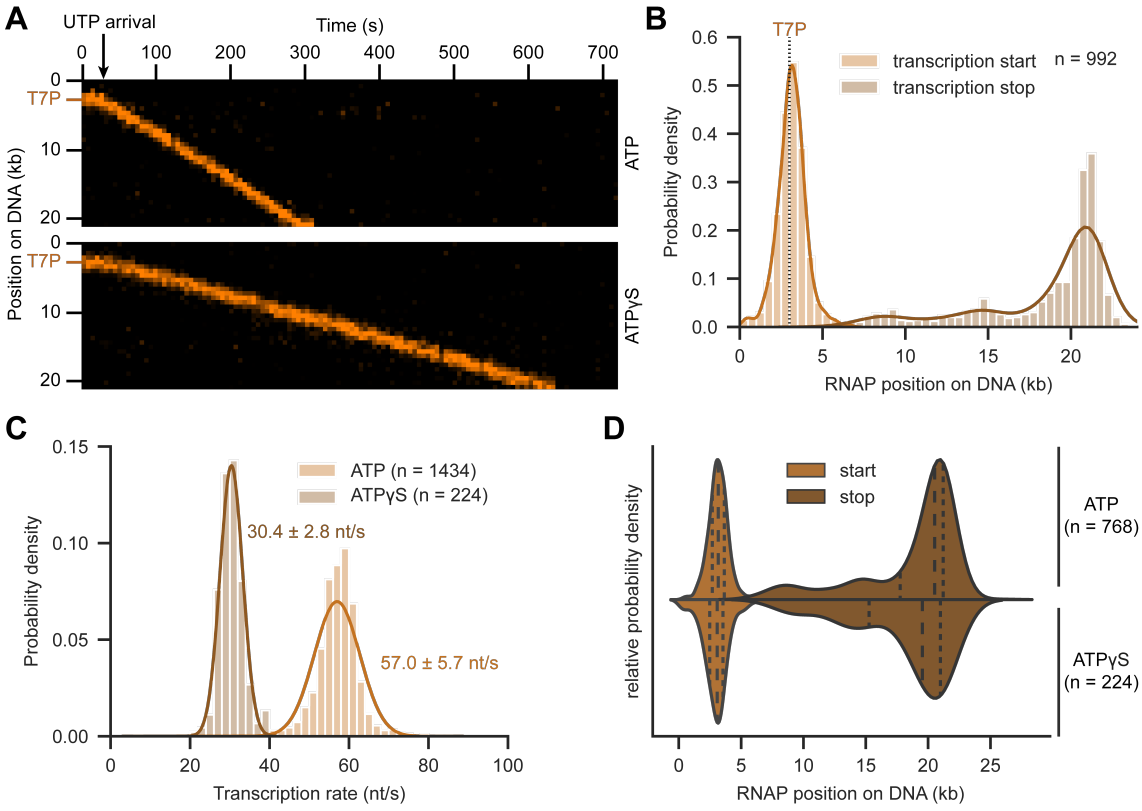

Figure 4.2. RNA polymerase transcribes highly processively at homogenous rates
(**A**) Representative kymographs showing transcribing RNAP upon UTP arrival in the presence of ATP (top) or ATPγS (bottom). (**B**) RNAP transcription start and stop sites distribution on T7P-DNA. Data from all experiments irrespective of ATP or ATPγS are shown. Lines represent the kernel density estimation (KDE). (**C**) Transcription rate distribution in the presence of ATP or ATPγS. Values indicate the mean transcription rates (± SD) derived from a Gaussian fit (lines). (**D**) RNAP transcription start and stop sites distribution by KDE on T7P-DNA in the presence of ATP or ATPγS. Lines represent the quartiles of the distribution with the middle line corresponding to the median. Data depicted are the same as shown in (**B**).

conferring structural domains (specificity and AT-rich recognition loop) in T7 RNAP (McAllister and Carter, 1980; Oakley et al., 1979; Sousa et al., 1993; Tahirov et al., 2002; Yin and Steitz, 2002). Upon addition of UTP, stalled RNAP transcription elongation complexes resumed transcription and transcribed highly processively to the opposite DNA end (Figures 4.2A and 4.2B; Movie 3). We observed a mean transcription rate of 57.0 ± 5.7 nt/s (± SD) with a distribution ranging from ~45-70 nt/s which is well in line with previous studies (Figure 4.2C) (Thomen et al., 2008).

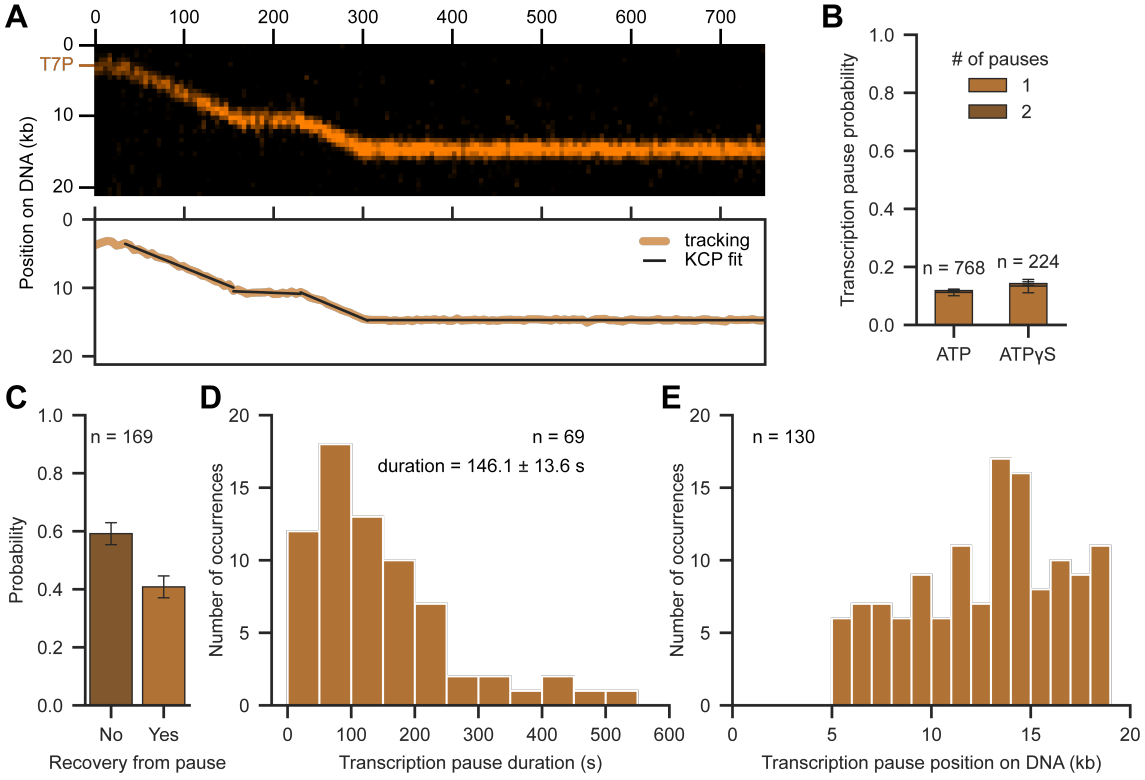

Figure 4.3. RNA polymerase transcribes long DNA continuously
(**A**) Representative kymograph (top) with corresponding tracking result (bottom) showing transcribing RNAP with one transient and one permanent pause detected by a fit with the kinetic change point (KCP) algorithm (black line). (**B**) Transcription pause probability on T7P-21 kb DNA in the presence of ATP or ATPγS. (**C**) Fraction of permanent and transient transcription pauses determined by pause recovery within the imaging time. (**D**) Distribution of transcription pause durations determined from transient pauses. Values indicate the mean pause duration (± SEM). (**E**) Distribution of transcription pause positions on DNA of permanent and transient pauses. Error bars in (**B**) and (**C**) represent the estimated SEM by bootstrapping.

Next, we evaluated the influence of ATPγS, required for OCCM formation, on transcription kinetics. Since synthesis depends only on hydrolysis of the β bond, we did not expect disruption of activity when replacing ATP with ATPγS. In fact, RNAP still readily transcribed without changing its overall processivity, but the structural difference in the substrate reduced the rate by ~50 % (30.4 ± 2.8 nt/s, mean ± SD) (Figures 4.2A, 4.2C and 4.2D; Movie 3). Although we detected molecules which paused during transcription (Figure 4.3A), almost 90 % of all molecules transcribed the entire DNA continuously in the presence of ATP and ATPγS (Figure 4.3B), further demonstrating the high processivity of RNAP. Upon pausing, approximately 40 % of molecules only paused

transiently with a mean duration of 146.1 ± 13.6 s (± SEM), whereas the remainder did not resume transcription within the imaging time (Figures 4.3A, 4.3C and 4.3D). Importantly, pausing occurred in a stochastic manner independent of the transcribed DNA sequence (Figure 4.3E). Notably, all labeled RNAPs generated showed similar enzymatic characteristics in terms of processivity, rate and pausing in our experiments (Figure 4.4). Together, this assay provides a powerful platform to study encounters between RNAP and other protein complexes.

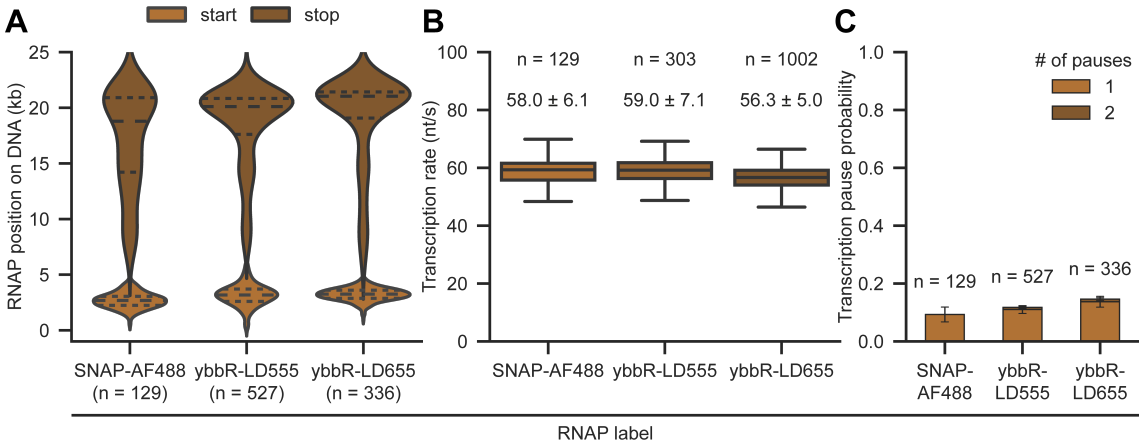

Figure 4.4. Differently labeled RNA polymerases show similar enzymatic characteristics
(**A**, **B** and **C**) Transcription start and stop sites (**A**), rates (**B**) and pause probability (**C**) did not differ between SNAP-AF488-, ybbR-LD655- and ybbR-LD655-T7 RNAP. (**A**) Lines in the violin plot represent the quartiles of the distribution with the middle line corresponding to the median. (**B**) Values above the box plots indicate the mean transcription rates (± SD) in the presence of ATP derived from a Gaussian fit. Data depicted in (**A**), (**B**) and (**C**) are the same as shown in Figure 4.2B, 4.2C and 4.3B, respectively. Error bars in (**C**) represent the estimated SEM by bootstrapping.

4.2 RNA polymerase can robustly reposition MCM DHs

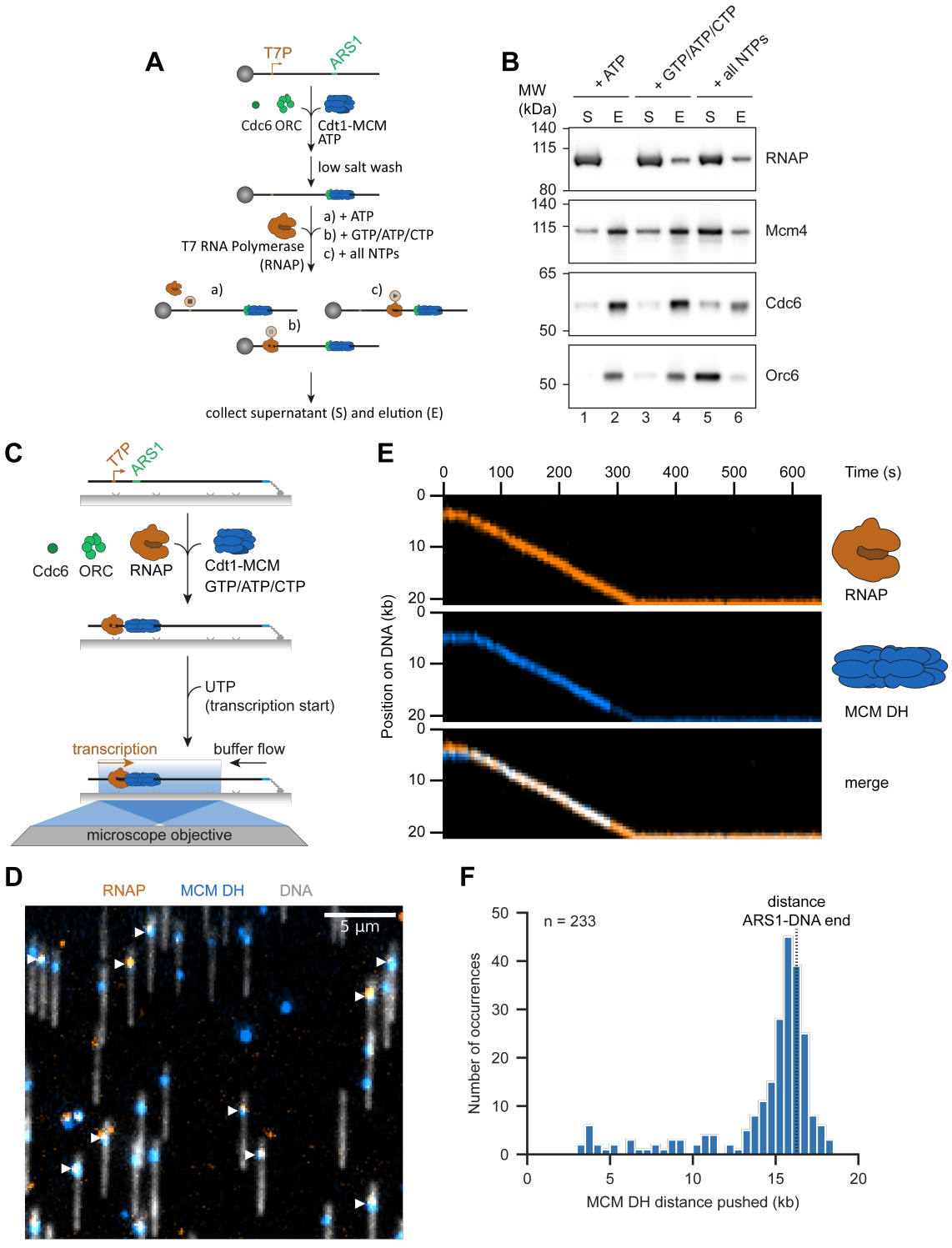

Figure 4.5. RNA polymerase can robustly reposition MCM DHs

(**A**) Schematic of the ensemble RNAP-MCM DH collision assay. MCM DHs were loaded on bead-bound 5 kb T7P-ARS1-containing DNA and unbound protein was removed. The reaction was split and incubated with RNAP in the presence of ATP, GTP/ATP/CTP or all NTPs. Finally, the supernatant (S) and elution (E) were collected, corresponding to displaced and DNA-bound protein, respectively. (**B**) Western Blot analysis of an ensemble RNAP-MCM DH collision assay as described in (**A**). In the presence of ATP alone, RNAP

(legend continued on next page)

could not engage with T7P (lanes 1+2) whereas in the presence of GTP/ATP/CTP, RNAP formed a stalled elongation complex (lanes 3+4). Both conditions did not show extensive MCM, Cdc6 or ORC displacement from DNA. In the presence of all NTPs, allowing for active transcription, MCM and ORC were efficiently displaced from linear DNA (lanes 5+6). (**C**) Schematic of the single-molecule RNAP-MCM DH collision assay. RNAP and MCM DHs were loaded on T7P-ARS1-containing 21 kb DNA. Transcription was started by addition of UTP and imaged on flow-stretched DNA. (**D**) Example of a partial field of view, showing simultaneous loading of RNAP (amber) and MCM DHs (blue) to the same DNA molecule (gray) as marked by white arrows. (**E**) Representative kymograph demonstrating that RNAP (amber) could push MCM DH (blue) upon collision. (**F**) Distribution of MCM DH distance pushed by RNAP.

The temporal delay between origin licensing and firing provides a window of opportunity for encounters between loaded MCMs and other machineries operating on chromosomes. Most notably among these are the prolific RNAPs which remain highly active during this period. In budding yeast, more than one third of replication origins sit in actively transcribed regions, raising the necessity for mechanisms that reconcile origin licensing with transcription (Looke et al., 2010). Indeed, previous reports have suggested that encounters between RNAP and MCM DHs do not derail the competence of MCM DHs for origin firing (Gros et al., 2015).

To investigate encounters between RNAP and MCM DHs, we performed transcription on licensed DNA substrates. In ensemble experiments using a linear DNA substrate, RNAP efficiently displaced ORC and MCM DHs in a transcription-dependent manner (Figures 4.5A and 4.5B). To directly visualize and determine the different outcomes of encounters between individual RNAPs and MCM DHs, which are not distinguishable in ensemble assays, we sought to performed single-molecule transcription experiments on licensed DNA with labeled RNAP and MCM (Figure 4.5C). For this, we combined our previously established single-molecule helicase loading (Figure 3.2A) and time-coordinated transcription (Figure 4.1A) assays, which simultaneously loaded stable, stalled RNAP elongation complexes and MCM DHs to the same DNA molecule in the presence of GTP, ATP and CTP but in the absence of UTP (Figure 4.5D). Remarkably, approximately four out of five collisions between RNAP and MCM DHs resulted in robust repositioning, with most MCM DHs being displaced by more than 10 kb until the DNA end (Figures 4.5E, 4.5F and 4.6A; Movie 4). Nevertheless, we observed additional collision outcomes including RNAP ejection (13.8 %) and transcription pausing (1.4 %) or stalling (3.5 %) (Figure 4.6). Strikingly, we never observed MCM ejection, showing the high robustness of the origin licensing pathway and MCM DHs. Importantly, a preceding high salt challenge did not change collision outcomes, demonstrating that MCM DHs loaded under physiological conditions are inherently prone to slide (Figure 4.6A).

Several studies, including our data shown in chapter 3.2 (Figure 3.3), suggest that multiple MCM DHs can be loaded at an origin site. We, therefore, asked if numerous MCM DHs might serve as a barrier to transcription. For this, we altered our MCM loading conditions to favor multiple MCM loading at non-ARS1 sites (see chapter 2.6.3). Surprisingly, but consistent with the high displacement efficiency of individual MCM DHs, even multiple MCM DHs could be repositioned by a single RNAP (Figure 4.7A; Movie 5).

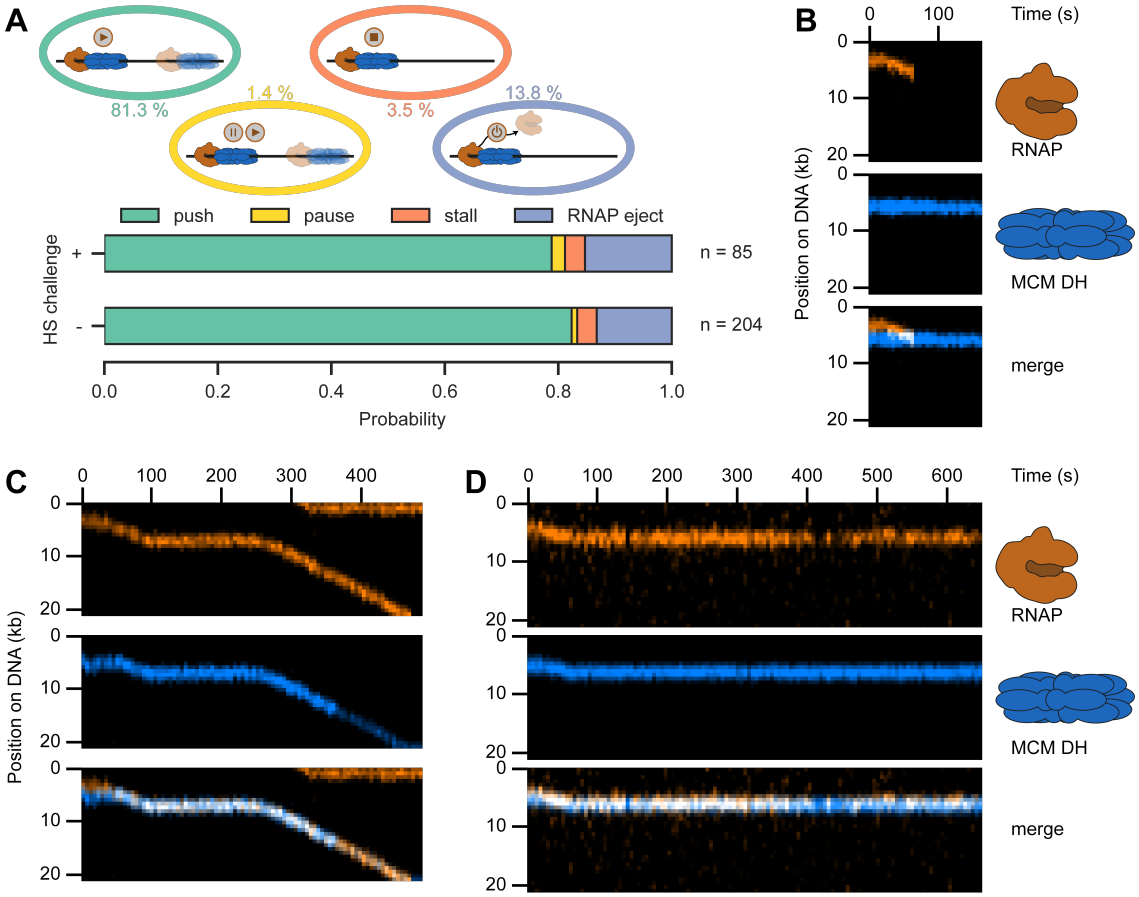

Figure 4.6. Collisions with RNA polymerase do not cause loss of MCM DHs
(**A**) Quantification of the outcomes of RNAP collisions with MCM DHs with (top) or without (bottom) a high salt (HS) wash prior to transcription start. Displayed percentages represent the combined probability of both conditions. (**B**, **C** and **D**) Representative kymographs showing RNAP ejection (**B**), transient transcription pausing (**C**) or transcription stalling (**D**) upon collision between RNAP (amber) and MCM DH (blue).

To estimate the cumulative barrier strength of MCM DHs, we approximated the number of MCM DHs in MCM foci by fluorescence intensity. Remarkably, transcription rates were not altered in the presence of single MCM DHs (57.0 ± 5.7 and 55.1 ± 5.2 nt/s, mean ± SD, in the absence or presence of one MCM DH, respectively) and only reduced by ~20 % (44.0 ± 6.7 nt/s, mean ± SD) with five or more MCM DHs (Figure 4.7B). Although transcription was more prone to pausing with increasing number of pushed MCM DHs, the probability of transient pausing and pause duration remained unchanged (Figures 4.7C, 4.7D and 4.7E) when compared to encounters with single MCM foci at ARS1. Together, our data demonstrate that even multiple MCM DHs do not pose a strong barrier to the transcription machinery, which is consistent with the observation of trains of MCM DHs visualized by EM (Douglas et al., 2018).

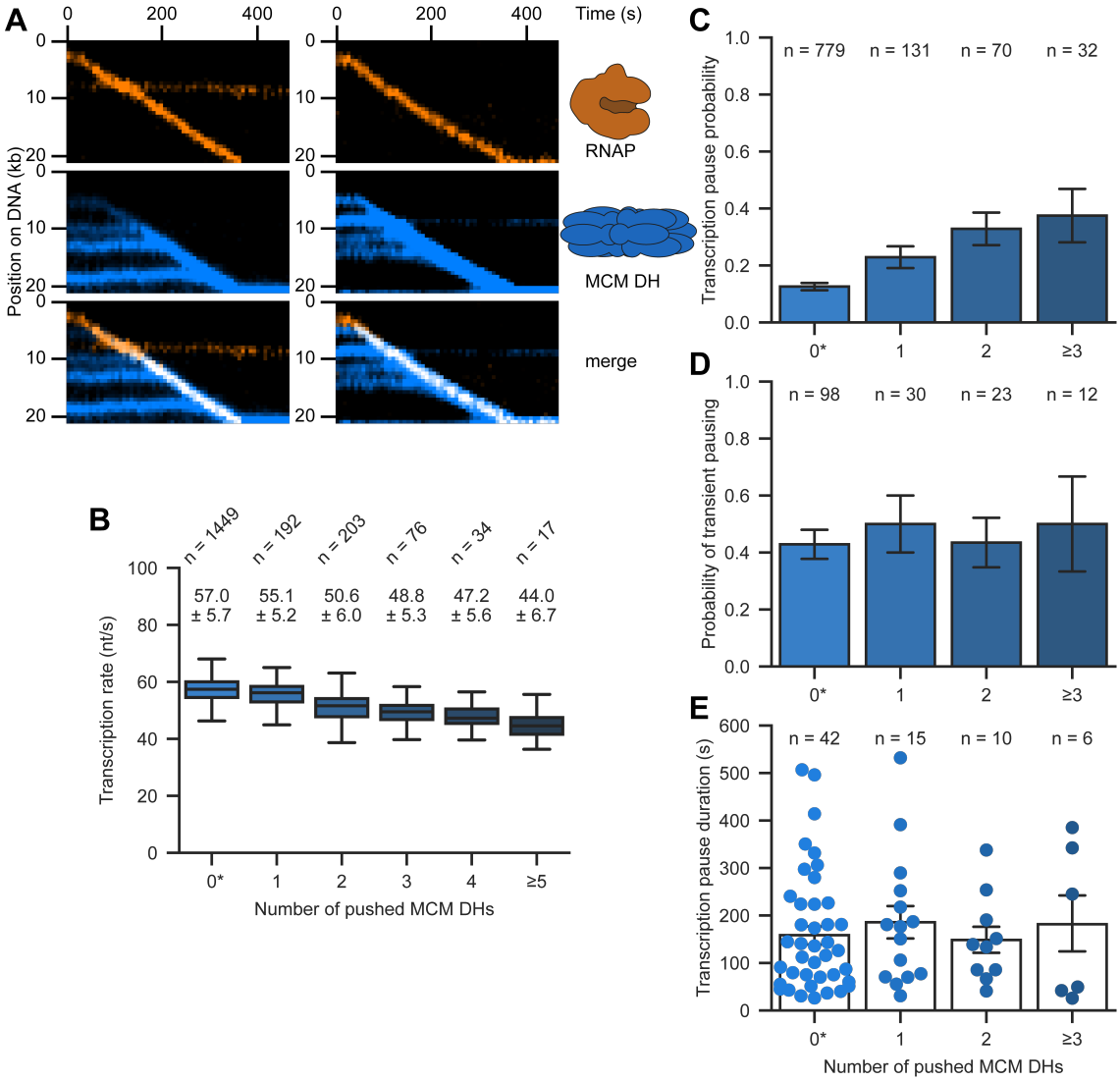

Figure 4.7. RNA polymerase can push trains of MCM DHs at slightly reduced rates
(**A**) Representative kymographs demonstrating that RNAP (amber) can push multiple MCM DHs (blue) over long distances. (**B**) Boxplot of transcription rates in the absence (0) or presence of (1 to ≥5) pushed MCM DHs. Values above the box plots indicate the mean transcription rates (± SD) derived from a Gaussian fit. (**C**, **D** and **E**) Influence of MCM DHs on transcription pausing. Transcription pause probability (**C**), probability of transient pausing (**D**) and transcription pause duration (**E**) in the absence (0) or presence of (1 to ≥5) pushed MCM DHs. Only DNA molecules with one MCM foci located at ARS1 were analyzed. Error bars represent the estimated SEM by bootstrapping. *Data displayed for 0 pushed MCM DHs in (**B**), (**C**), (**D**) and (**E**) were combined with data shown in Figure 4.2C, 4.3B, 4.3C and 4.3D (ATP condition only), respectively.

4.3 RNA polymerase can reposition MCM DHs together with multiple nucleosomes

We demonstrated that MCM DHs can switch to a diffusive mode to facilitate transcription through licensed origins. However, our observations of MCM DH displacement as well as previously reported CMG diffusion are based on bare DNA, lacking a more cellular-like context (Wasserman et al., 2019). Thus, we cannot exclude that MCM displacement is impeded in cells which contain additional obstacles on DNA. To investigate MCM displacement in a more physiological context, we sought to directly observe MCM displacement in the presence of chromatin. We introduced a widom601 nucleosome positioning sequence (Lowary and Widom, 1998) into our DNA template and reconstituted chromatin using fluorescently labeled histone octamers at two different densities (low and high) (Figures 4.8A and 4.8B). Low-density chromatin substrates with well-separable nucleosome foci revealed an enrichment of nucleosomes at the widom601 sequence, despite a broad distribution of nucleosomes on DNA (Figure 4.8C). Furthermore, individual nucleosome foci showed one- and two-step photobleaching, consistent with correctly formed nucleosomes at our labeling efficiency (~1.5 dyes/nucleosome) (Figure 4.8D).

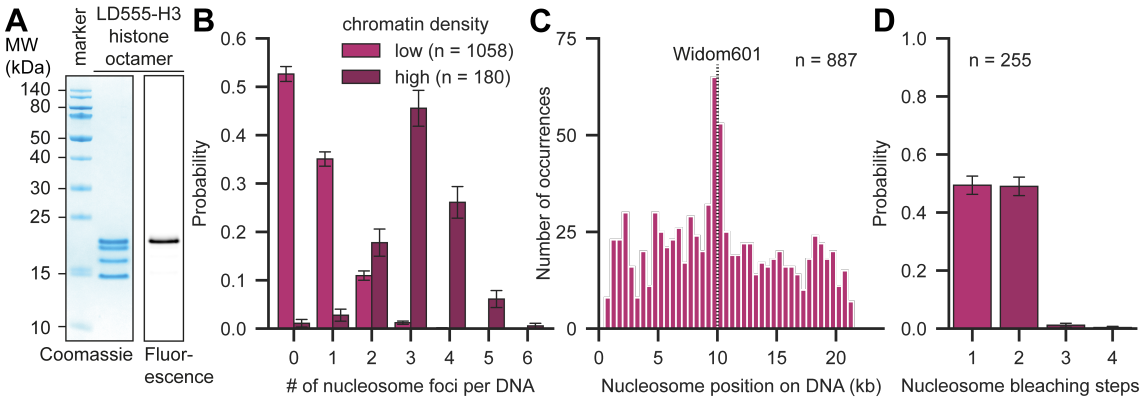

Figure 4.8. Reconstitution of sparse chromatin arrays for single-molecule assays
(**A**) SDS-PAGE analysis of purified LD555-H3 histone octamers by Coomassie Blue staining and fluorescence detection. (**B**) Number of nucleosome foci per DNA in low- and high-density chromatin. (**C**) Nucleosome binding distribution on T7P-ARS1-Widom601-containing 21 kb DNA for low-density chromatin. (**D**) Distribution of nucleosome bleaching steps in low-density chromatin. Error bars in (**B**) and (**D**) represent the estimated SEM by bootstrapping.

Next, we performed transcription on chromatin substrates in the presence or absence of pushed MCM DHs in front of RNAP (Figures 4.9A and 4.9B). Independent of MCM DH being in front of RNAP, we observed a similar global reduction in transcript lengths with increasing number of downstream nucleosomes (Figures 4.9C and 4.9D). Collisions between RNAP and individual nucleosomes revealed continuous transcription through chromatin by either pushing or ejecting nucleosomes (Figures 4.10A, 4.10B and 4.10C; Movie 6), consistent with previous studies observing short-range transcription through nucleosomes (Hodges et al., 2009; O'Neill et al., 1993; Studitsky et al., 1994; Studitsky

et al., 1997). Remarkably, individual nucleosomes were readily pushed over long distances to the DNA end (Figure 4.10D). However, roughly half of RNAP-nucleosome collisions resulted in RNAP ejection (6.9 %), transient pausing (9.4 %) or permanent stalling (30.6 %) with the latter being increased by multiple nucleosomes (Figures 4.10E and 4.10F).

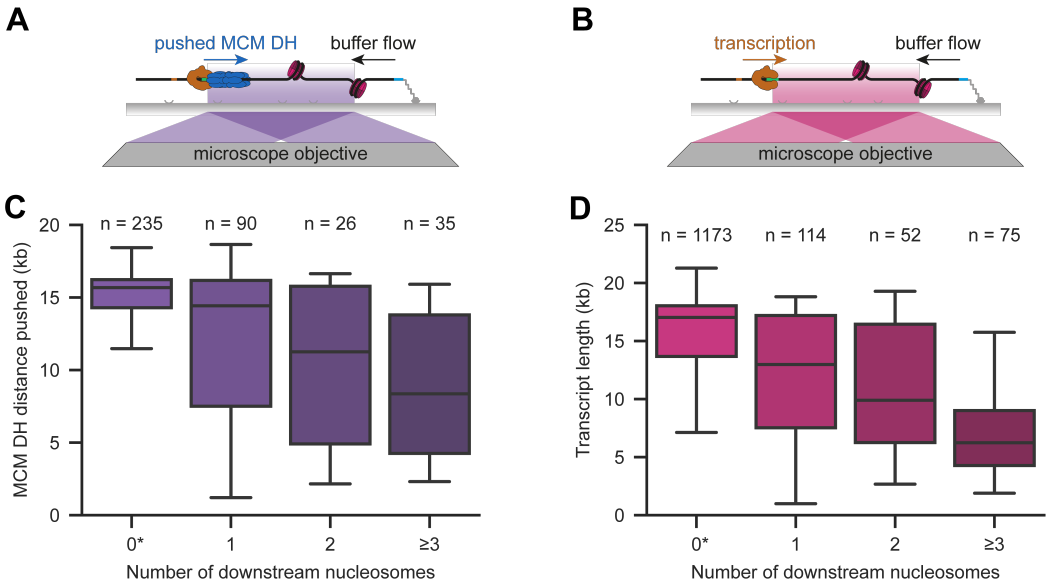

Figure 4.9. Transcript lengths reduce with increasing number of downstream nucleosomes
(**A** and **B**) Schematic of the single-molecule MCM DH displacement (**A**) and transcription (**B**) assay in the presence of nucleosomes. (**A**) MCM DHs were loaded and pushed by RNAP as described in Figure 4.5C but on chromatinized DNA. (**B**) Time-coordinated transcription was set up as described in Figure 4.1A but on chromatinized DNA. (**C** and **D**) Boxplot of the global MCM distance pushed (**C**) and transcript length (in the absence of MCM) (**D**) in the absence (0) or presence of (1 to ≥3) nucleosomes downstream of ARS1 (**C**) and T7P (**D**). *Data displayed for 0 downstream nucleosomes in (**C**) and (**D**) were combined with data shown in Figure 4.5F and 4.2B, respectively.

Surprisingly, individual nucleosomes did not present an obstacle for MCMs as RNAP could also readily push MCM DHs through nucleosomes by nucleosome ejection or pushing (Figures 4.11A, 4.11B and 4.11C; Movie 7). Pushed MCM DHs were not affected by nucleosomes in about half the cases, whereas MCM pushing also paused transiently (11.1 %) or stalled permanently (36.6 %), similar to collisions between RNAP itself and nucleosomes (Figures 4.10F, 4.11D and 4.11E). Importantly, collisions between MCM DHs and nucleosomes never triggered MCM unloading, again demonstrating the high robustness of the origin licensing pathway. However, MCM repositioning was obstructed by larger numbers of nucleosomes with increased probability of stalling but again not unloading (Figure 4.11E). Finally, we determined rates of RNAP and RNAP-MCM complexes with increasing number of pushed nucleosomes. Interestingly, RNAP-MCM complex rates remained mostly unaffected upon pushing individual nucleosomes with no slowly transcribing population (compare 0-25 % quartiles in Figures 4.11F and 4.11 G) (51.9 ± 6.5, 52.8 ± 8.4, 50.3 ± 10.0, 46.4 ± 9.7 nt/s, mean ± SD, for 0, 1, 2 and ≥3 nucleosomes, respectively) (Figure 4.11F). Conversely, we

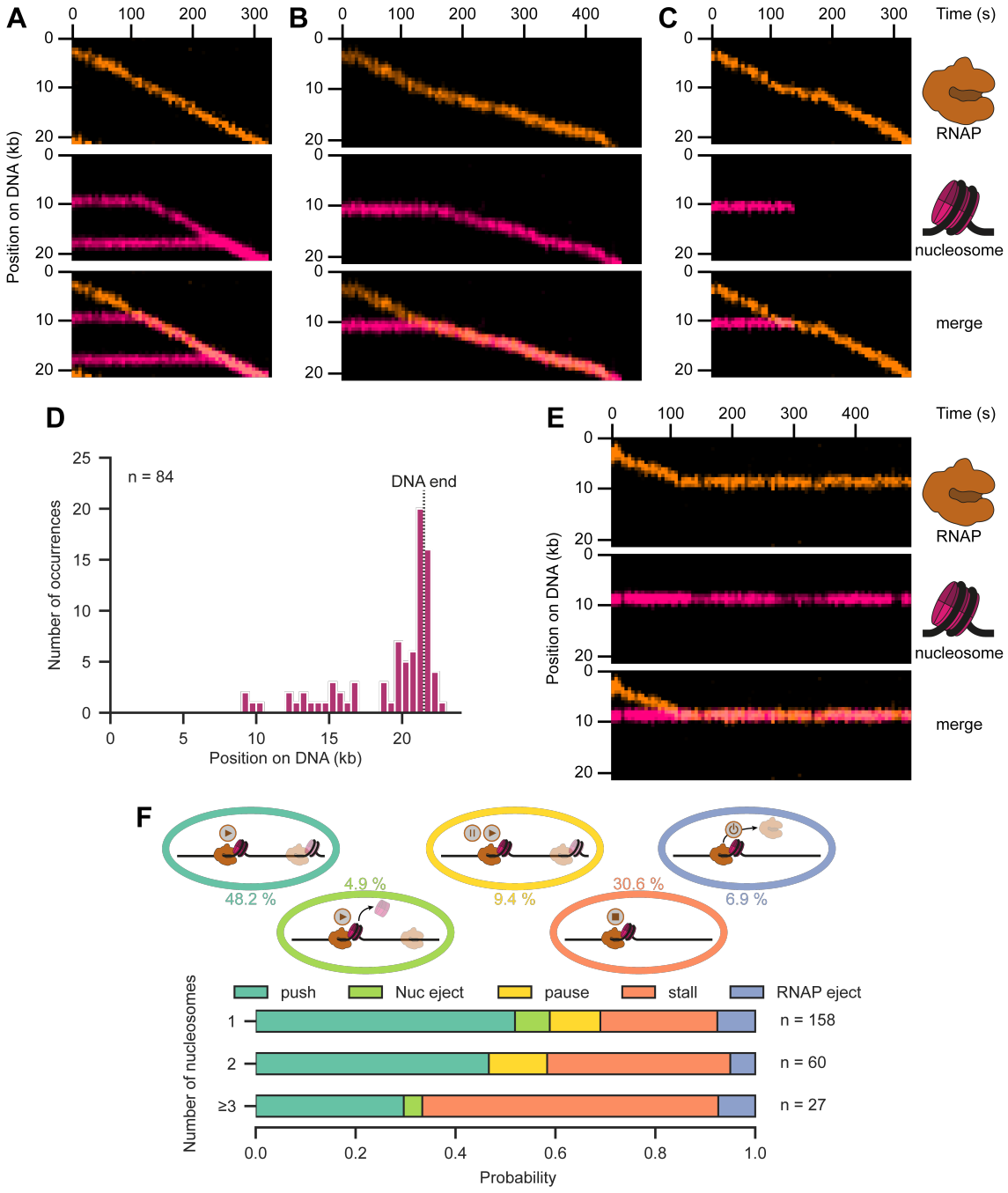

Figure 4.10. RNA polymerase can transcribe through individual nucleosomes
(**A**, **B** and **C**) Representative kymographs demonstrating that RNAP (amber) could transcribe through one or multiple nucleosomes (pink) by nucleosome pushing (**A** and **B**) or ejection (**C**). (**D**) Distribution of DNA positions of pushed nucleosomes after transcription termination. Only DNA molecules with one nucleosome foci were considered. (**E**) Representative kymograph displaying transcription stalling upon collision between RNAP (amber) and nucleosome (pink). (**F**) Quantification of the outcomes of RNAP collisions with a total of 1, 2 or ≥3 nucleosomes. Displayed percentages represent the combined probability irrespective of the number of nucleosomes.

observed a population of RNAP molecules (0-25 % quartile) with substantially decreased transcription rates upon pushing individual nucleosomes (56.9 ± 6.3, 49.7 ± 15.6, 48.1 ± 15.8, 39.1 ± 14.9 nt/s, mean ± SD, for 0, 1, 2 and ≥3 nucleosomes, respectively)

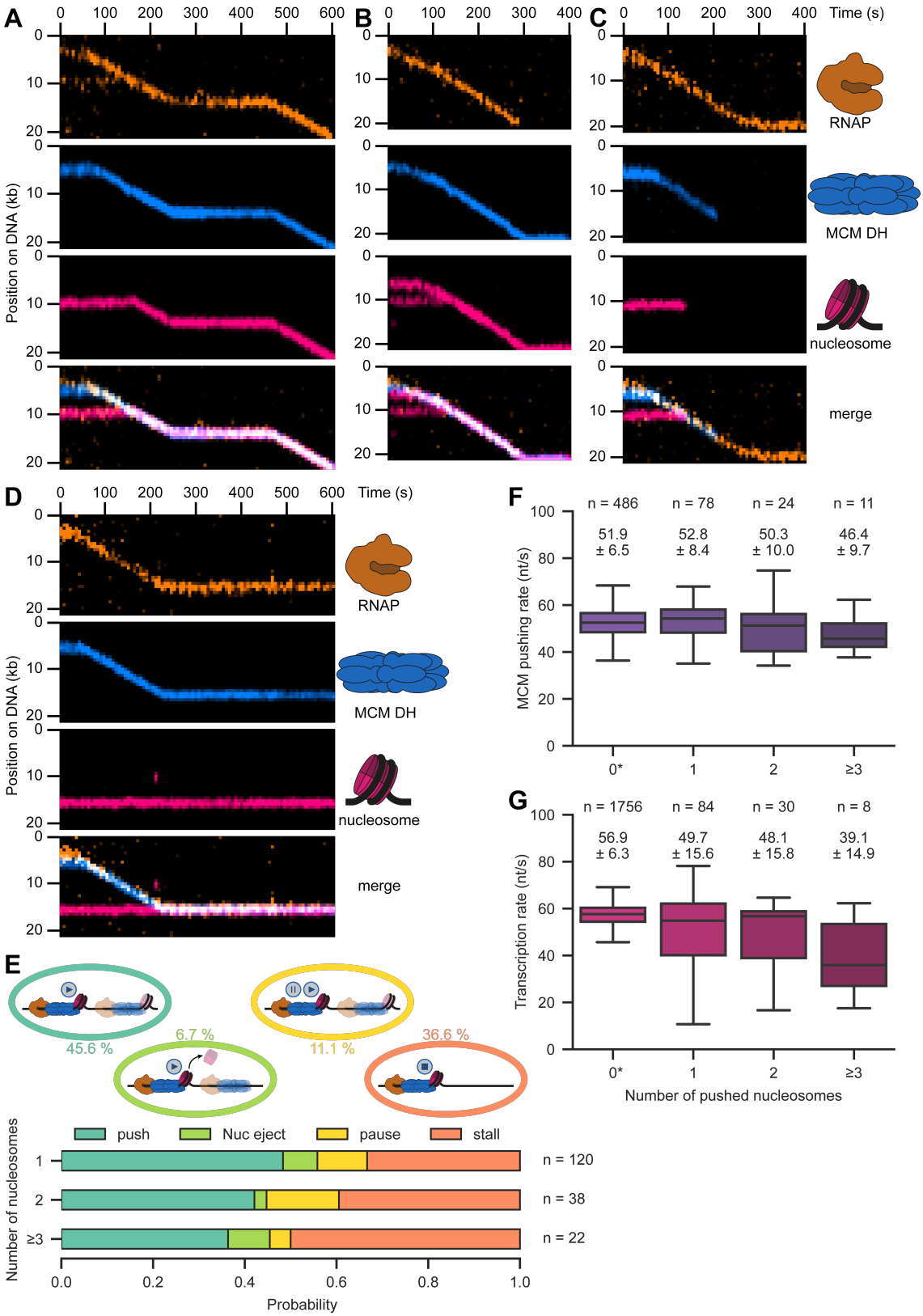

Figure 4.11. RNA polymerase can reposition MCM DHs together with multiple nucleosomes
(**A**, **B**, **C** and **D**) Representative kymographs demonstrating that RNAP (amber) could displace MCM DHs (blue) through one or multiple nucleosomes (pink) by nucleosome pushing (**A** and **B**) or ejection (**C**) besides transcription stalling (**D**) upon collision. (**E**) Quantification of the outcomes of pushed MCM DH collisions

(legend continued on next page)

with a total of 1, 2 or ≥3 nucleosomes. Displayed percentages represent the combined probability irrespective of the number of nucleosomes. (**F** and **G**) Boxplot of MCM pushing rates (**F**) and transcription rates (in the absence of MCM) (**G**) in the absence (0) or presence of (1 to ≥3) pushed nucleosomes. Values above the box plots indicate the mean rates (± SD) derived from a Gaussian fit. *Data displayed for 0 pushed nucleosomes in (**F**) and (**G**) were combined with data shown in Figure 4.7B and 4.2C (ATP condition only), respectively.

(Figures 4.10B and 4.11G). Thus, our observations suggest that the DNA extruding from MCM DHs is more suitable for RNAP. Alternatively, the absence of a slow transcription population could result from MCM DHs helping to destabilize nucleosomes, consistent with previous studies showing histone chaperone activity of the MCM2 subunit (Huang et al., 2015; Petryk et al., 2018). However, the histone binding region lies in the N-terminal tail of MCM2 and thus is not exposed towards nucleosomes in the MCM DH state.

4.4 OCCM and ORC are repositioned or bypassed by RNAP

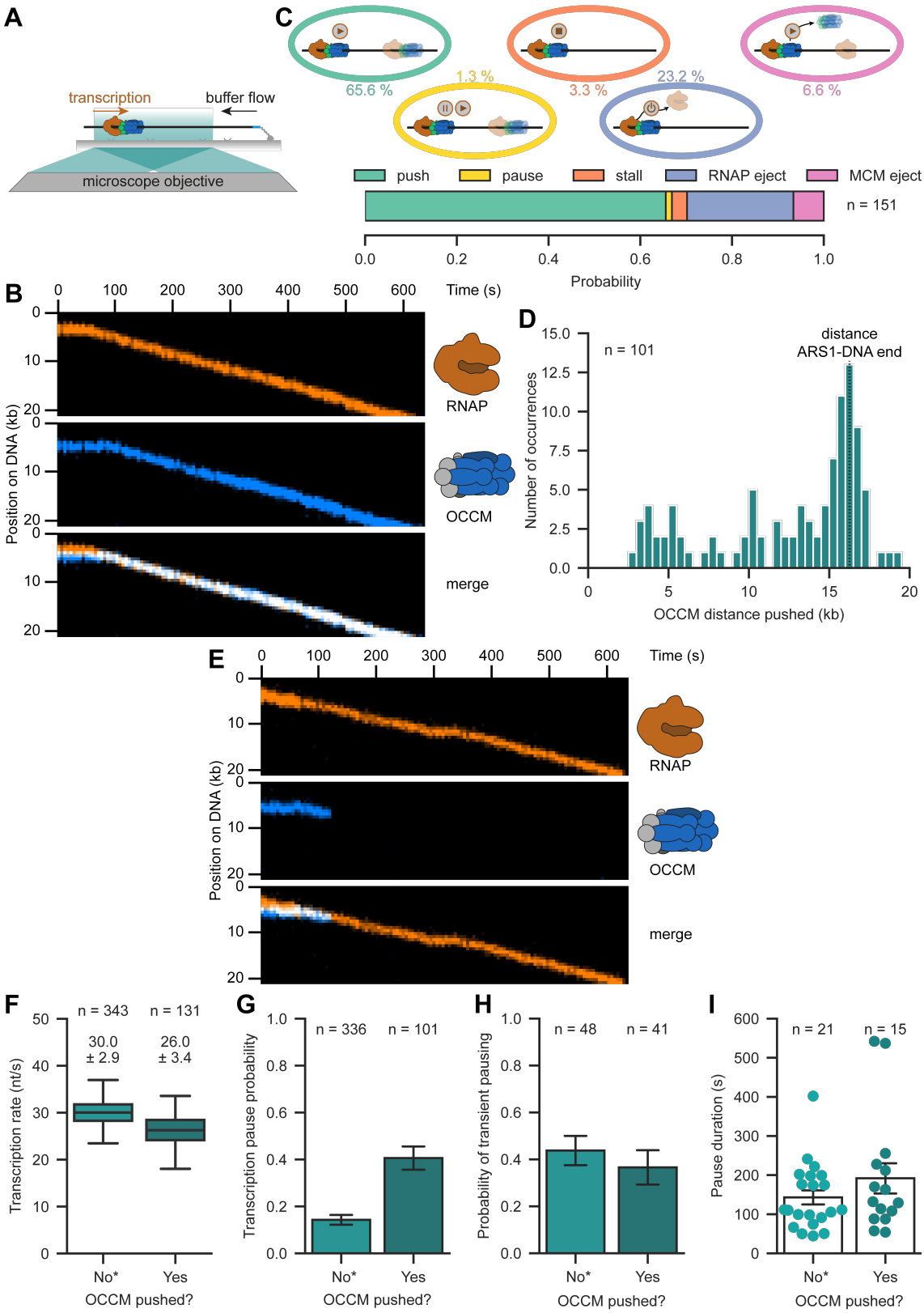

Figure 4.12. OCCM origin licensing intermediates can be repositioned by RNA polymerase
(**A**) Schematic of the single-molecule RNAP-OCCM collision assay. Assay was performed as described in Figure 4.5C but ATPγS was used instead of ATP in all steps. (**B**) Representative kymograph showing that

(legend continued on next page)

RNAP (amber) could push OCCM (blue) upon collision. (**C**) Quantification of the outcomes of RNAP collision with OCCM. (**D**) Distribution of OCCM distance pushed by RNAP. (**E**) Representative kymograph showing OCCM ejection upon collision between RNAP (amber) and OCCM (blue). (**F**) Boxplot of transcription rates in the absence or presence of pushed OCCM. Values above the box plots indicate the mean transcription rates (± SD) derived from a Gaussian fit. (**G**, **H** and **I**) Influence of OCCM on transcription pausing. Transcription pause probability (**G**), probability of transient pausing (**H**) and transcription pause duration (**I**) in the absence or presence of pushed OCCM. Error bars represent the estimated SEM by bootstrapping. *Data displayed for no pushed OCCM in (**F**), (**G**), (**H**) and (**I**) were combined with data shown in Figure 4.2C, 4.3B, 4.3C and 4.3D (ATPγS condition only), respectively.

The remarkable resilience of MCM DHs during encounters with RNAP led us to wonder if other stages of the loading pathway might exhibit the same properties. In particular, a series of complex conformational changes are required for recruitment, and final encircling of the origin by MCM. Due to those complex conformational changes MCM DH formation is a slow process with long-lived intermediates containing only one MCM prior to recruitment of the second MCM (Miller et al., 2019; Ticau et al., 2015). Thus, encounters between RNAP and loading intermediates are likely to occur during origin licensing.

To examine encounters between RNAP and the less stable OCCM intermediate, we performed MCM loading and transcription entirely in the presence of ATPγS (Figure 4.12A). Remarkably, two out of three RNAP-OCCM collisions led to robust OCCM repositioning with many complexes being pushed to the DNA end (Figures 4.12B, 4.12C and 4.12D). However, besides transcription pausing and stalling, a small fraction of OCCM was ejected (6.6 %) which was never observed for MCM DHs (Figures 4.6A, 4.12C and 4.12E). Although the OCCM complex appeared less stable than an MCM DH,

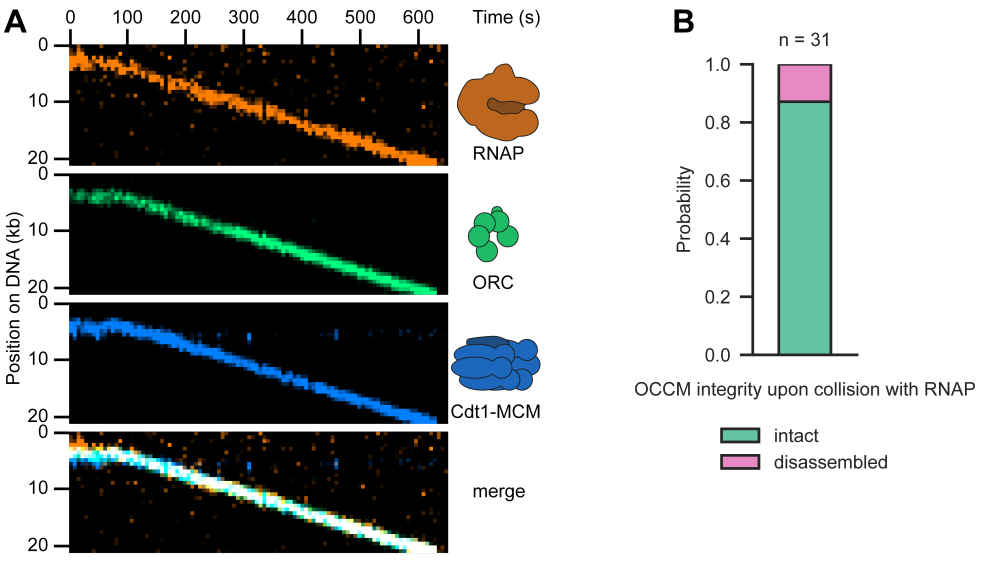

Figure 4.13. OCCM stays fully intact during encounters with RNAP and during displacement
(**A**) Representative kymograph demonstrating that OCCM stayed intact upon encounter and when being pushed by RNAP (amber), as judged by the presence of ORC (green) and Cdt1-MCM (blue) signal. (**B**) Quantification of OCCM integrity upon collision with RNAP. Probability of OCCM staying intact or being disassembled as judged by the presence of ORC and Cdt1-MCM.

we observed a higher probability of RNAP ejection upon collision with an OCCM (23.2 %). Thus, OCCM might engage more strongly with ARS1 or be in a more unfavorable conformation for sliding than MCM DHs. Alternatively, ORC, which is an integral component of the OCCM but not of MCM DHs, may present a greater obstacle to RNAP (Candelli et al., 2018). In fact, in the presence of a single OCCM, although the transcription rate only slightly decreased from 30.0 ± 2.9 to 26.0 ± 3.4 nt/s (mean ± SD) (Figure 4.12F), pausing was almost three times more frequent, comparable to three or more pushed MCM DHs (Figures 4.7C and 4.12G). Importantly, pausing properties in terms of recovery and duration remained similar (Figures 4.12H and 4.12I).

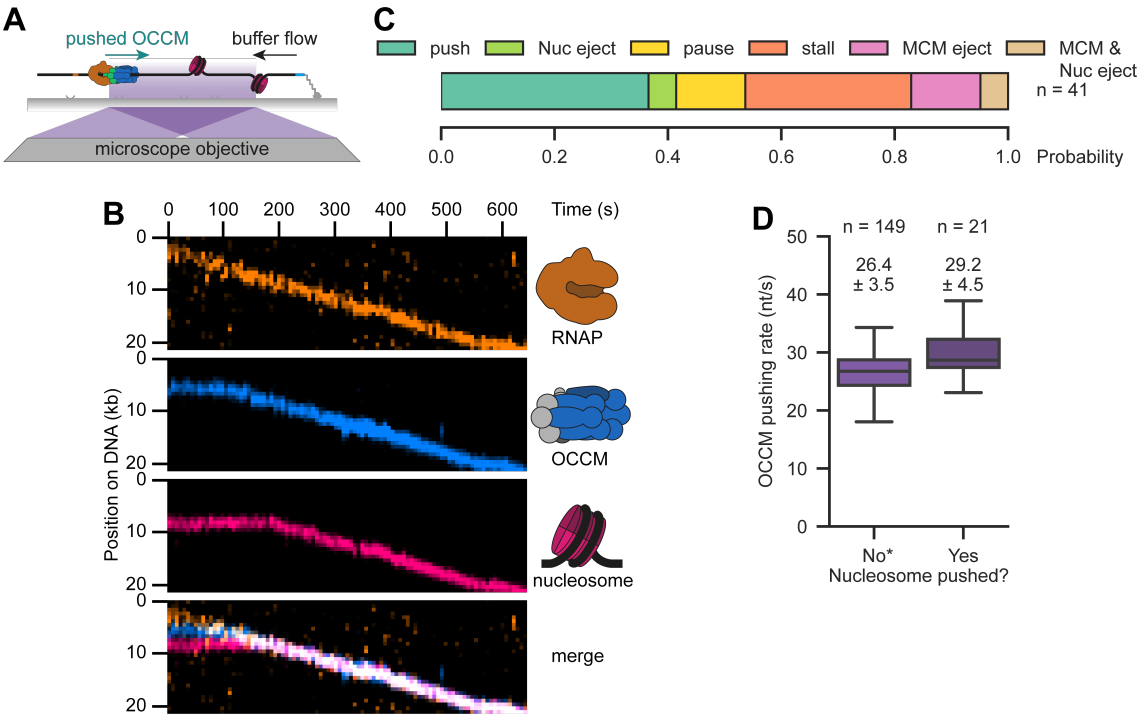

Figure 4.14. RNA polymerase can reposition OCCM in the presence of nucleosomes
(**A**) Schematic of the single-molecule OCCM displacement assay in the presence of nucleosomes. OCCM was loaded and pushed by RNAP as described in Figure 4.12A but on chromatinized DNA. (**B**) Representative kymograph demonstrating that RNAP (amber) could displace OCCM (blue) through nucleosomes (pink) by nucleosome pushing. (**C**) Quantification of the outcomes of pushed OCCM collisions with nucleosomes. (**D**) Boxplot of OCCM pushing rates in the absence or presence of nucleosomes. Values above the box plots indicate the mean OCCM pushing rates (± SD) derived from a Gaussian fit. *Data displayed for non-pushed nucleosomes were combined with data shown in Figure 4.12F.

Although we demonstrated that OCCM complexes are also robustly displaced by RNAP, OCCM integrity was only judged by the presence of MCM. To exclude the possibility of OCCM disassembly, we additionally monitored ORC in three-color experiments containing labeled RNAP, ORC and MCM. The majority of OCCM complexes stayed fully intact during encounters with RNAP and during displacement (Figure 4.13; Movie 8). We also tested collisions between pushed OCCMs and nucleosomes (Figure 4.14A). Although around 17 % of collisions led to OCCM loss, the majority stayed intact and were displaced through nucleosomes by nucleosome pushing or ejection (Figures 4.14B,

4.14C and 4.14D).

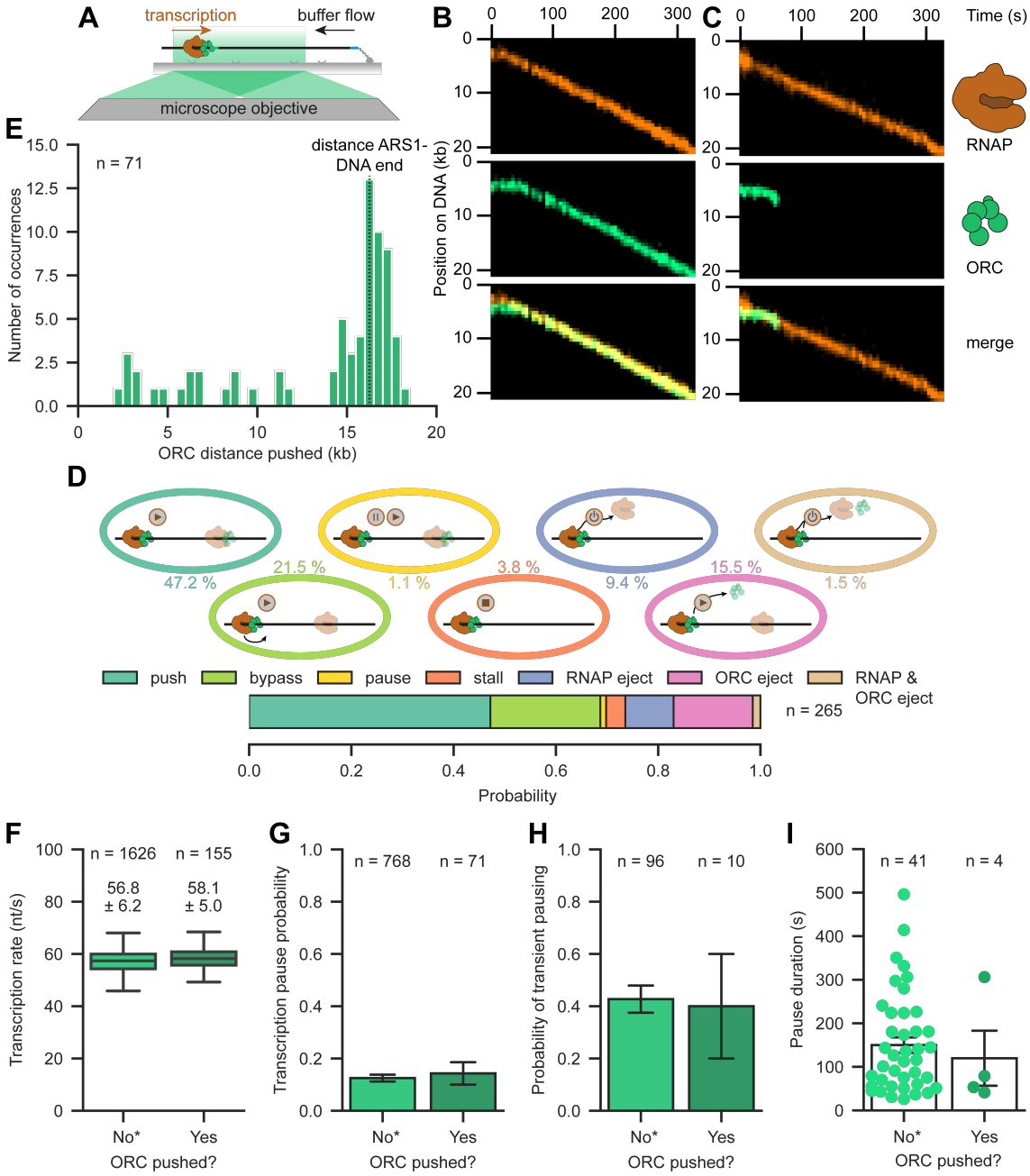

Figure 4.15. The origin licensing factor ORC can be repositioned by RNA polymerase
(**A**) Schematic of the single-molecule RNAP-ORC collision assay. Assay was performed as described in Figure 4.5C but only RNAP and ORC were loaded on DNA. (**B** and **C**) Representative kymographs showing that ORC (green) could also be displaced by RNAP (amber) (**B**) but was ejected more frequently (**C**). (**D**) Quantification of the outcomes of RNAP collision with ORC. (**E**) Distribution of ORC distance pushed by RNAP. (**F**) Boxplot of transcription rates in the absence or presence of a pushed ORC. Values above the box plots indicate the mean transcription rates (± SD) derived from a Gaussian fit. (**G**, **H** and **I**) Influence of ORC on transcription pausing. Transcription pause probability (**G**), probability of transient pausing (**H**) and transcription pause duration (**I**) in the absence or presence of pushed ORC. Error bars represent the estimated SEM by bootstrapping. *Data displayed for no pushed ORC in (**F**), (**G**), (**H**) and (**I**) were combined with data shown in Figure 4.2C, 4.3B, 4.3C and 4.3D (ATP condition only), respectively.

Having established that OCCM could be readily displaced, we next investigated the influence of transcription on ORC bound to DNA, the very first step in the loading pathway (Figure 4.15A). Surprisingly, although ORC is not topologically bound to DNA (unlike OCCM and MCM DH) and thus ejected more frequently (~17 %), RNAP could also push ORC in almost half of RNAP-ORC collisions, predominantly to the DNA end (Figures 4.15B, 4.15C, 4.15D and 4.15E; Movie 9). Consistent with ORC being less stable, especially at non-ARS1 sites, transcription in the presence of ORC was not altered in terms of rate, pausing probability and pausing properties (Figures 4.15F, 4.15G, 4.15H and 4.15I).

Unexpectedly, besides transcription pausing, stalling and RNAP ejection, we observed a major fraction (21.5 %) of RNAPs bypassing ORC (Figures 4.15D, 4.16A and 4.16B; Movie 10). Although we never observed bypassing for collisions between RNAP and MCM DHs, OCCM or nucleosomes, RNAP-ORC bypass could represent an artifact by surface stuck ORC or overlapping DNA. To exclude this possibility, we determined the variance in x direction (perpendicular to buffer flow) of stuck (not co-localizing with DNA), pushed and bypassed ORC molecules. As expected, pushed and bypassed ORC showed an order of magnitude higher mean variance (0.156 pixel^2) caused by DNA fluctuations than surface stuck ORC molecules (0.015 pixel^2), demonstrating that bypassed ORCs are in fact bound to DNA (Figure 4.16C). A small fraction of ORC failed to stably engage with ARS1 but instead co-localized with RNAP at the T7 promoter. Interestingly, RNAP pushed ORCs in this fraction to ARS1 and subsequently bypassed them, demonstrating that ORC bypass occurs on the same DNA molecule (Figure 4.16B). Together, we demonstrate that origin licensing intermediates OCCM and ORC are resilient to conflicts with RNAP by displacement or bypass mechanisms.

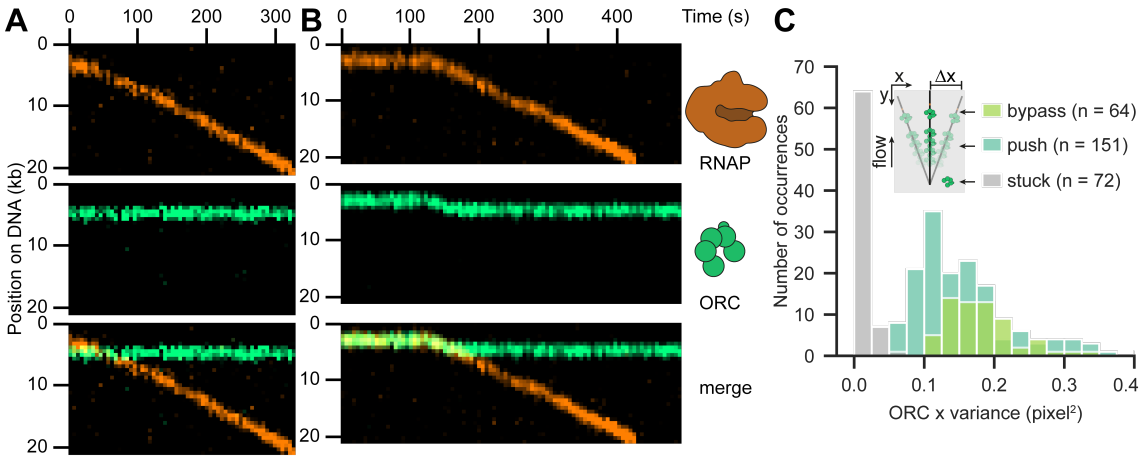

Figure 4.16. RNA polymerase can bypass ARS-bound ORC
(**A**) Representative kymograph demonstrating that RNAP (amber) was able to bypass ORC (green) bound to ARS1. (**B**) Representative kymograph displaying that ORC (green) which co-localized with RNAP (amber) could initially be pushed to ARS1 where it then was bypassed by RNAP. (**C**) Distribution of x variance perpendicular to buffer flow for bypassed, pushed and surface stuck ORC molecules.

4.5 Origin licensing factors can reside at new locations after transcription termination

The unexpected stability of origin licensing intermediates and their ability to be pushed upon encountering transcribing RNAP offered a potential novel pathway for origin specification. Consistent with our assays, although origin licensing in higher eukaryotes is not strictly defined by the primary DNA sequence, licensing in *S. cerevisiae* occurs at defined locations (Figures 3.2C and 3.2F). Thus, we tested if displaced licensing factors can reside at new locations after transcription termination. Since previous reports have suggested rather inefficient transcription termination at single termination sequences (Jeng et al., 1990; Macdonald et al., 1994; Macdonald et al., 1993; Mairhofer et al., 2015), we introduced five tandem T7 terminators (T7T) downstream of ARS1 (Figure 4.17A). As expected, transcription efficiently terminated at T7T (~95 %) but interestingly, a small fraction of RNAPs (~5 %) still escaped termination (Figures 4.17B, 4.17C and 4.17D). Based on our single-molecule data, we determined a termination efficiency of ~45 % for a single T7T, which is well in line with previous studies in bulk suggesting efficiencies of 40-70 % (Jeng et al., 1990; Macdonald et al., 1994; Macdonald et al., 1993).

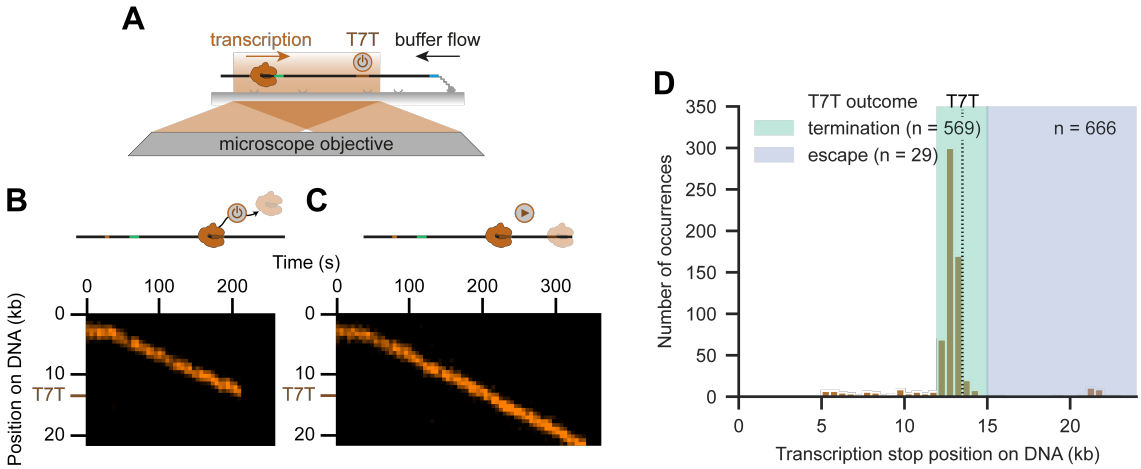

Figure 4.17. Five tandem terminators efficiently terminate transcription
(**A**) Schematic of the single-molecule transcription assay on 5xT7 terminator (T7T)-containing 21 kb DNA. Time-coordinated transcription was set up as described in Figure 4.1A. (**B** and **C**) Representative kymographs displaying efficient transcription termination (**B**) with rare cases of RNAP escaping transcription termination (**C**) at the introduced T7T site. (**D**) RNAP transcription stop site distribution on T7T-containing DNA. Bins classified as termination or escape are shaded green or violet, respectively.

Having established efficient transcription termination in our assay, we next investigated the stability of displaced ORC at T7T. In fact, ORC could remain stable at T7T (~3 %), however the majority (~87 %) dissociated upon transcription termination (Figures 4.18A, 4.18B and 4.18D (row 3); Movie 11). Strikingly, after ORC displacement, a population of molecules (~10 %) showed flow-driven diffusion of ORC in search of the origin (Figure 4.18C; Movie 12). This behavior suggests that 1D diffusion can be used to locate a new potential origin by ORC after encounters with other machineries on chromosomes, revealing a remarkable pathway for recovery at the earliest stage of origin licensing.

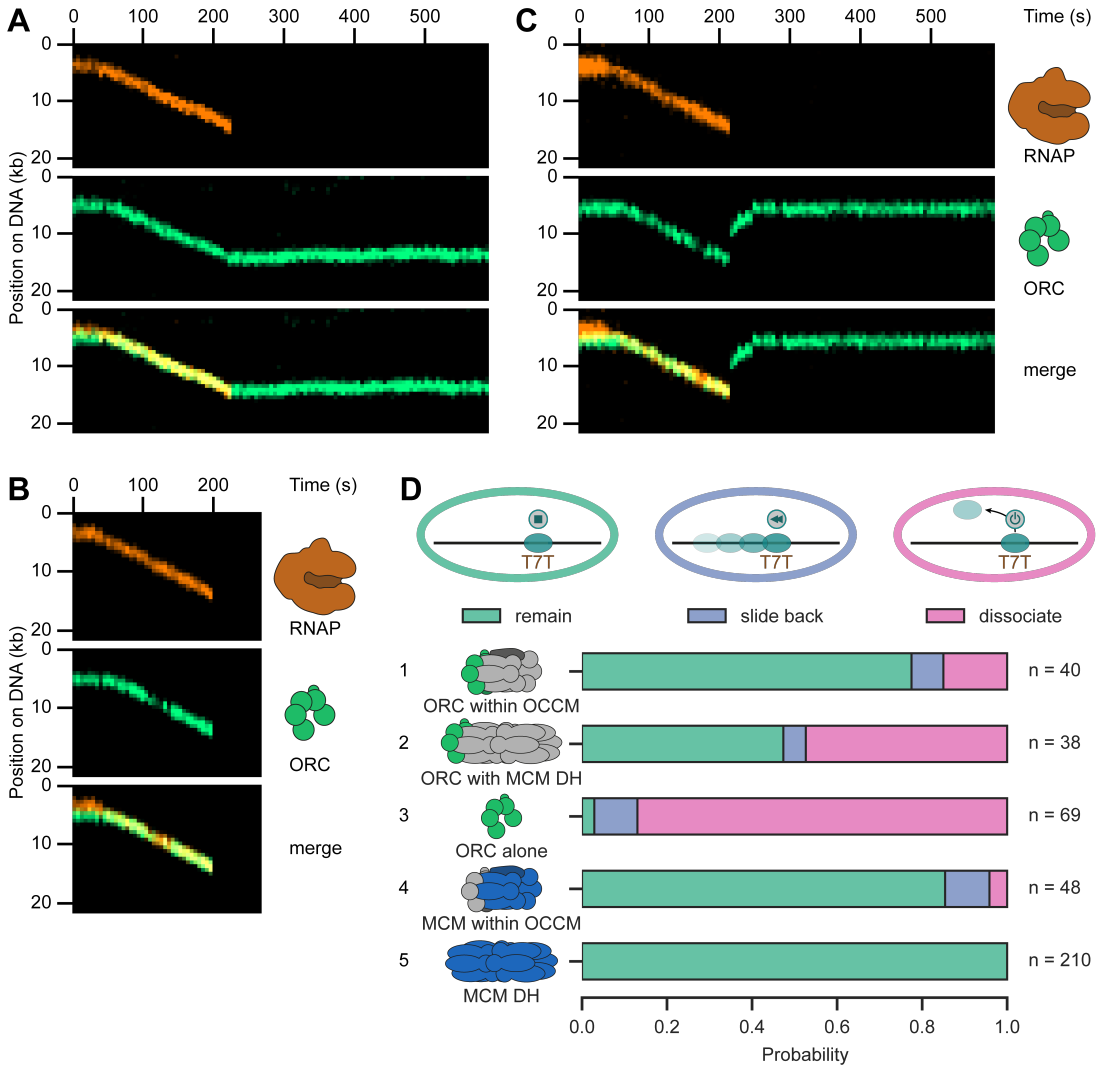

Figure 4.18. A small fraction of ORC can reside at new locations after termination
(**A**, **B** and **C**) Representative kymographs demonstrating that displaced ORC (green) could remain stable at the T7T site (**A**) but was more likely to dissociate (**B**) or slide back to ARS1 after a 1D diffusion in search of the origin (**C**) upon RNAP (amber) transcription termination. (**D**) Quantification of the stability of displaced origin licensing factors at the T7T site upon transcription termination.

In cells, we expect ORC diffusion to be limited and hence to be stabilized by other factors on chromosomes. Thus, we tested whether MCM DHs could stabilize associated ORC. MCM DHs always remained stable at T7T, in line with studies showing successful MCM DH activation at non-origin sites (Figures 4.18D (row 5), 4.19A, 4.19B and 4.19C; Movie 13) (Gros et al., 2015). Surprisingly, ORC associated with MCM DHs was far more stable with equal populations remaining and dissociating at T7T (Figure 4.18D (row2)). This indicates that ORC could hitch a ride on sliding MCM DHs to new locations. We cannot fully exclude the possibility that ORC which remained stable at T7T was part of a licensing intermediate (e.g. OCCM or MO complex). However, photobleaching was consistent with MCM DHs and we detected no difference in mean MCM fluorescence between populations with remaining or dissociating ORC (Figures 4.19C and 4.19D).

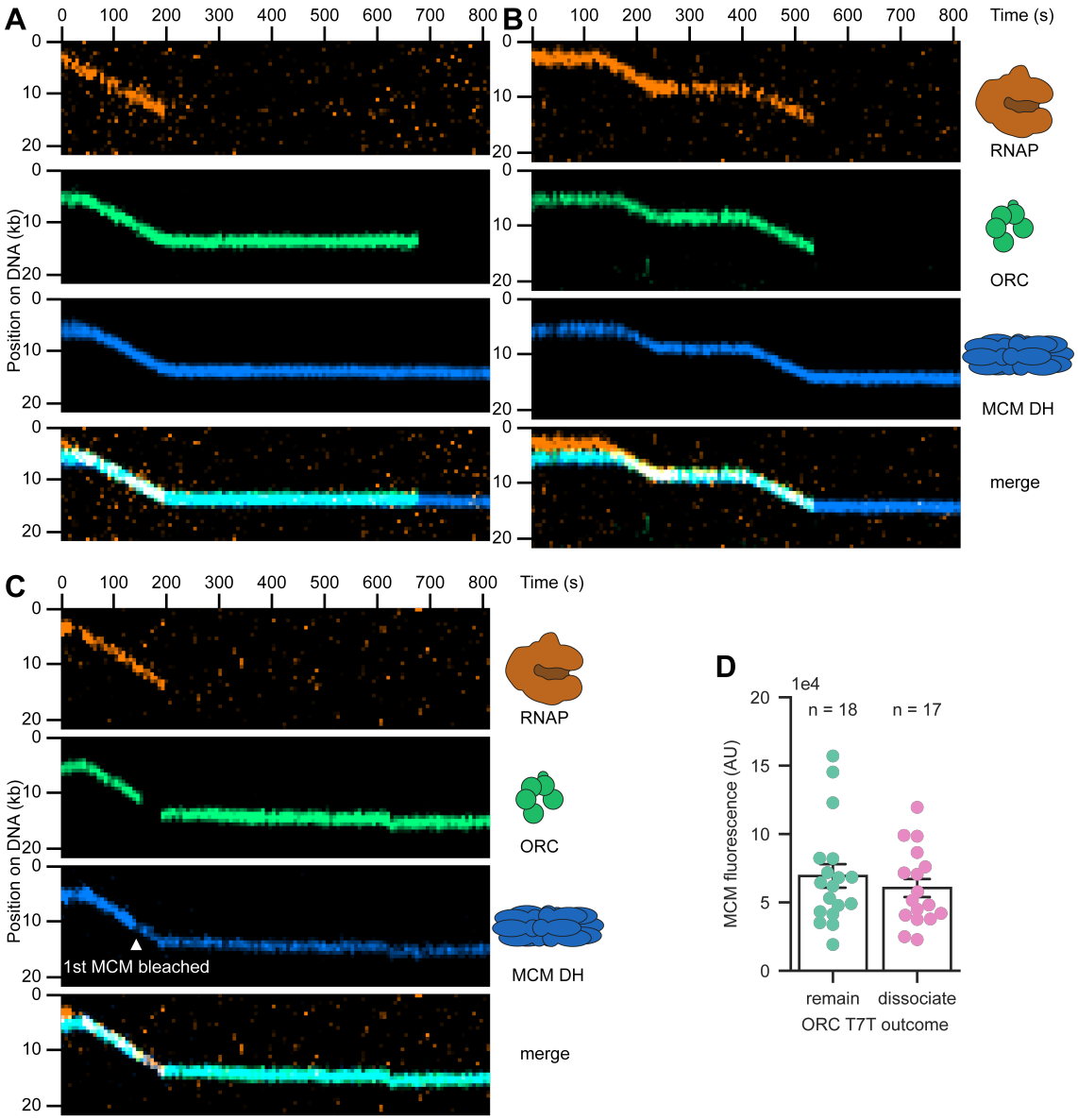

Figure 4.19. ORC can hitch a ride on sliding MCM DHs to new locations
(**A**, **B** and **C**) Representative kymographs showing that displaced MCM DHs (blue) remained stable whereas associated ORC (green) could hitch a ride on MCM DHs and be stabilized at the T7T site (**A** and **C**) but still dissociated at times (**B**) upon RNAP (amber) transcription termination. (**C**) MCM signal and photobleaching corresponded to an MCM DH and not an ORC-containing loading intermediate. (**D**) Observed MCM fluorescence for MCM DHs associated with ORC molecules which remained or dissociated at T7T upon transcription termination.

Finally, we determined the stability of the OCCM intermediate at T7T. The majority of OCCM complexes stayed intact and remained stably bound at T7T (78 % and 85 % for ORC and Cdt1-MCM, respectively) (Figures 4.18D (row 1 and 4) and 4.20A; Movie 14). Intriguingly, a small population (~10 %) of OCCM complexes diffused along DNA, however much slower than observed for ORC alone (Figure 4.20B; Movie 15). Together, based on the remarkable stability of ORC with MCM DHs and OCCM we speculate that nearly all intermediates in the MCM loading process can remain bound at non-canonical origins after displacement.

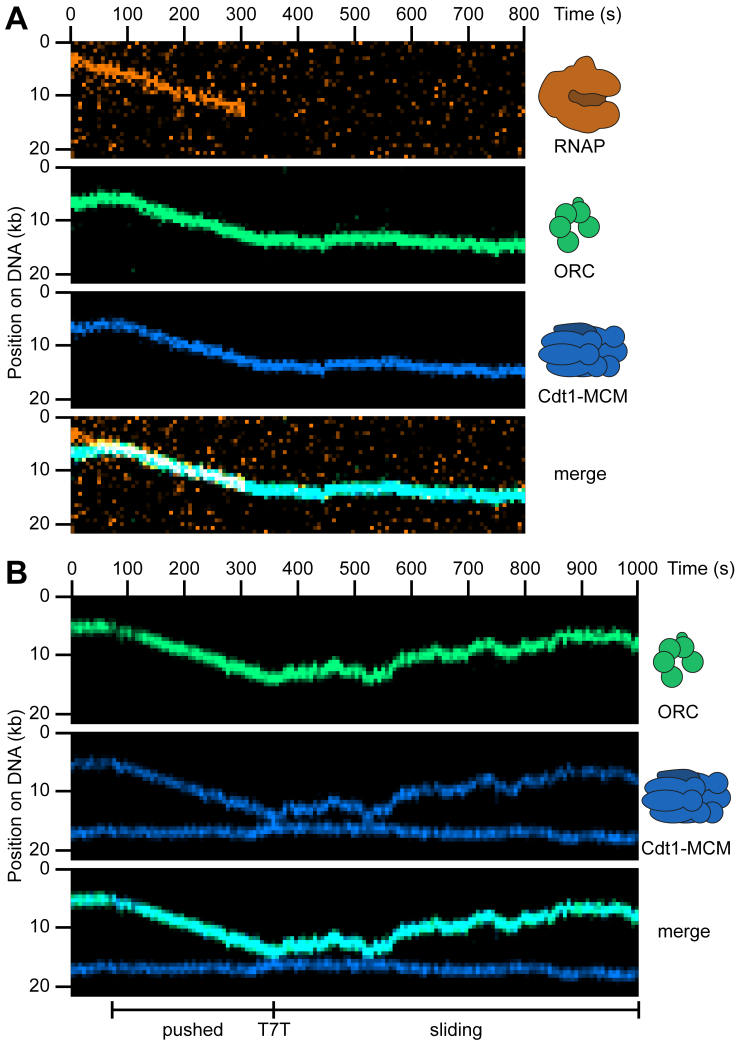

Figure 4.20. The majority of OCCM complexes reside at new locations after termination
(**A**) Representative kymograph showing that displaced OCCM remained stable at the T7T site, as judged by the presence of ORC (green) and Cdt1-MCM (blue) signal, upon RNAP (amber) transcription termination. (**B**) Representative kymograph showing that displaced OCCM could switch to a diffusive mode while staying intact, as judged by the presence of ORC (green) and Cdt1-MCM (blue) signal, at the T7T site upon transcription termination.

Chapter 5

MCM Complexes are Barriers That Restrict Cohesin-Mediated Loop Extrusion

Data and figures in chapter 5.1 were generously provided by Bart Dequeker and are part of a manuscript together with data shown in chapters 5.2 and 5.3 (Dequeker et al., 2020)

Consistent with our finding that multiple MCM DHs sequester and spread around origins in our *in vitro* single-molecule setup (chapter 3.2), MCM DHs are found in large excess over replication initiation events on chromosomes, known as the “MCM paradox” (Burkhart et al., 1995; Edwards et al., 2002; Lei et al., 1996; Mahbubani et al., 1997). We have demonstrated that MCM DHs as well as origin licensing intermediates overcome conflicts with the transcription machinery by adopting a diffusive DNA binding mode, leading to their displacement and specification of new origins (chapter 4). While MCM complexes do not pose a barrier to transcription, it remained unknown if other machineries operating on chromosomes are impeded by inactive MCM DHs or vice versa, which are loaded in G1 phase of the cell cycle. Cohesin, a member of the SMC complexes, is likely to encounter MCM DHs while folding the genome via loop extrusion. While CTCF and RNAPs are thought to act as barriers to cohesin-mediated loop extrusion (Busslinger et al., 2017; Heinz et al., 2018; Nora et al., 2017; Parelho et al., 2008; Rubio et al., 2008; Stedman et al., 2008; Wendt et al., 2008; Wutz et al., 2017), the outcome of encounters between cohesin complexes and MCM DHs was not known. Intriguingly, MCM DHs could potentially adopt a diffusive DNA binding mode similar to transcription conflicts but also act as a stable barrier in that scenario. Thus, we investigated whether licensed origins containing MCM DHs represent a novel barrier to cohesin-mediated loop extrusion and whether this would affect the overall genome architecture.

5.1 Origin licensing restricts loop and TAD formation by cohesin in G1 phase zygotes

To elucidate whether MCM DHs are barriers to cohesin-mediated loop extrusion, we first investigated the influence of chromatin-bound MCMs on loop extrusion, loops and TADs by utilizing the oocyte-to-zygote transition (OZT). Upon fertilization, oocytes undergo the meiosis II division to form zygotes (one cell embryo stage). In zygotes, parental and maternal DNA are still separated in pronuclei with their chromatin being organized by cohesin-mediated loop extrusion into loops and TADs (Flyamer et al., 2017; Gassler et al., 2017). One of the key advantages of the OZT for this study lies in the ability to inhibit origin licensing (pre-RC assembly) while completion of the cell cycle until the end of

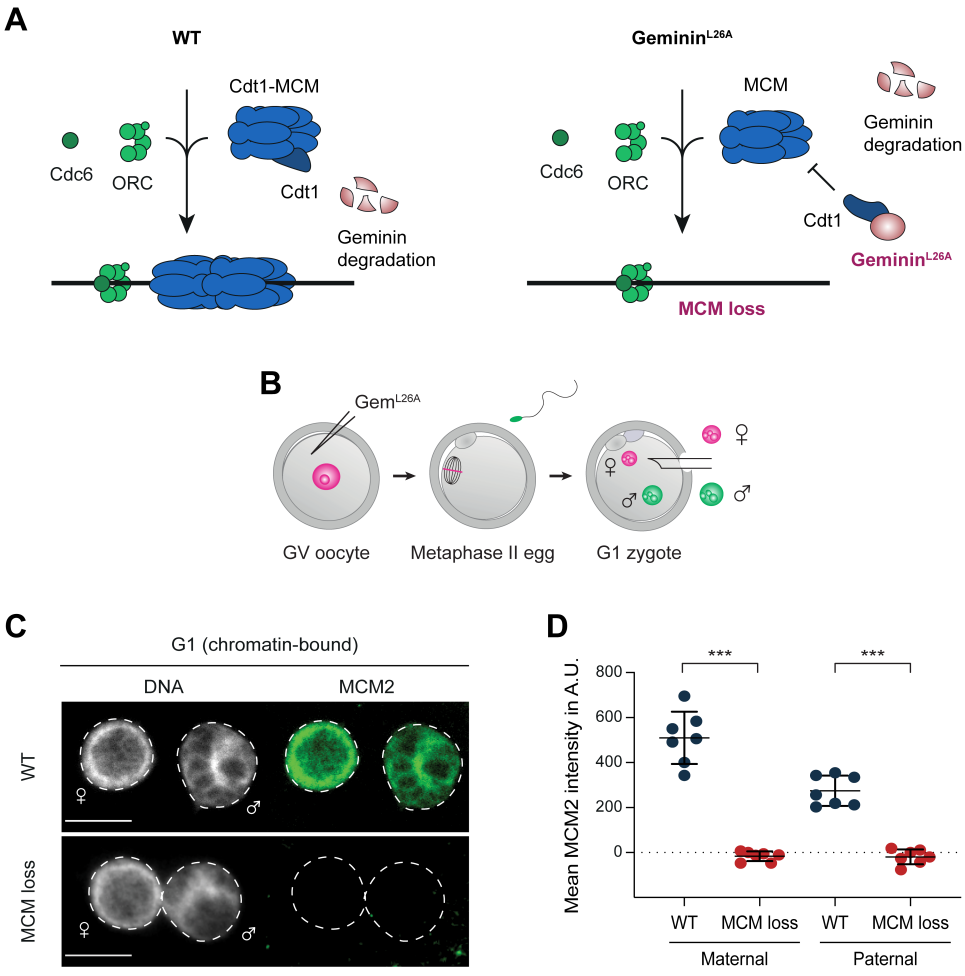

Figure 5.1. Generation of MCM-deficient zygotes
(**A**) Geminin targets Cdt1 and inhibits MCM loading outside of G1 phase. Upon degradation of geminin, Cdt1 can reengage with MCM, allowing for MCM loading on chromatin (left), whereas a non-degradable version of geminin (geminin^{L26A}) inhibits MCM loading throughout the cell cycle ("MCM loss", right). (**B**) MCM loss zygotes were generated by injecting geminin^{L26A} in germinal vesicle (GV)-stage oocytes, *in vitro* maturation and fertilization. Maternal (magenta) and paternal (green) nuclei were extracted and analyzed by snHi-C. (**C** and **D**) Immunofluorescence staining of chromatin-bound MCM2 in wildtype (WT) and MCM loss G1 phase zygotes. (**C**) Representative image with DAPI-stained DNA (scale bar 10 µm). (**D**) Quantification of mean chromatin-bound MCM2 intensity (n = 7, all conditions). *** p<0.001, calculated using unpaired t-test.

meiosis II remains unaffected (no DNA replication is occurring between meiosis I and II).

To generate zygotes, whose chromatin is deficient in MCMs, we directly inhibited the origin licensing pathway. To avoid re-replication, origin licensing and firing are naturally tightly controlled and separated into G1 and S phase, respectively, by S-CDK-dependent mechanisms (Drury et al., 2000; Frigola et al., 2013; Nguyen et al., 2000). Higher eukaryotes, including the utilized mouse OZT system, possess an additional regulator

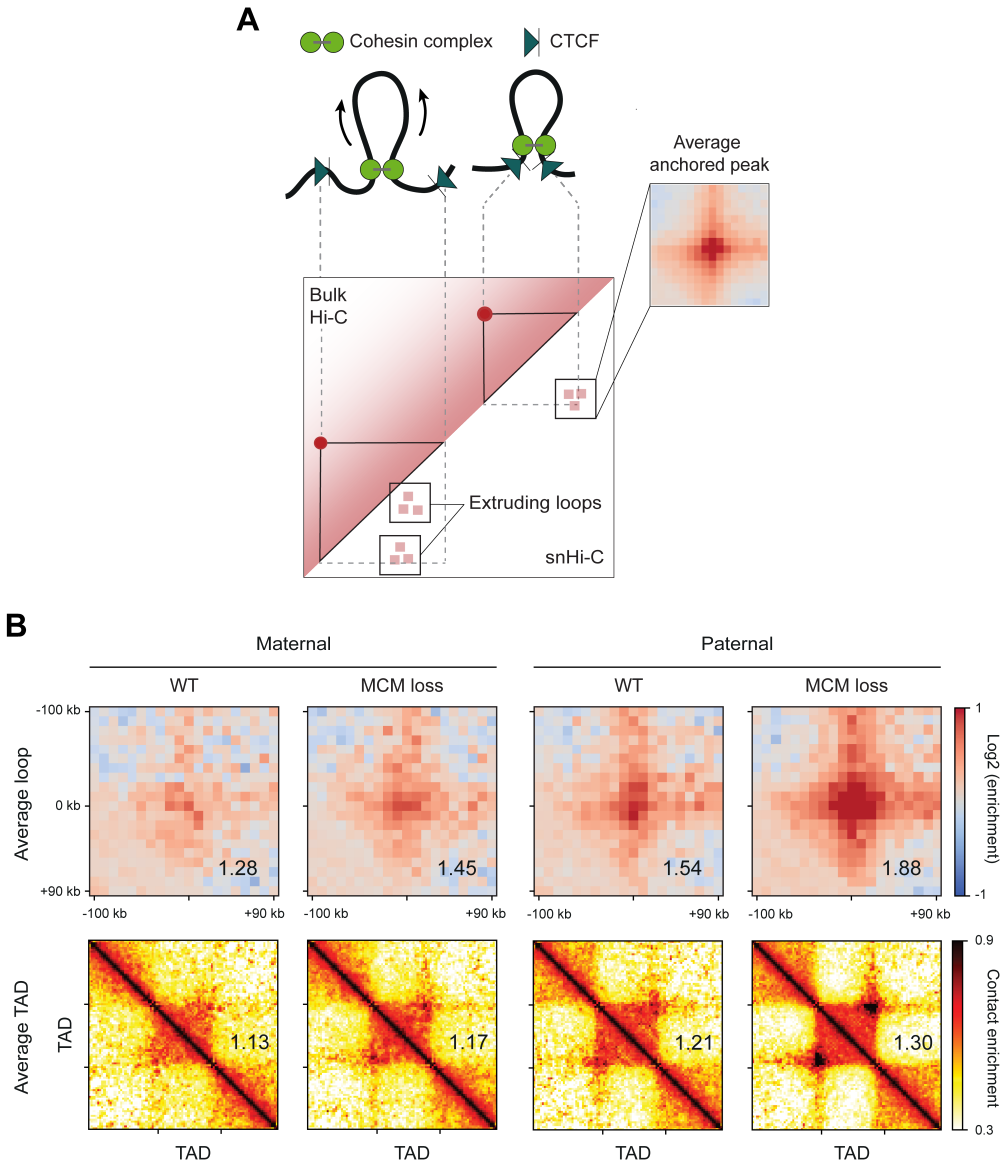

Figure 5.2. Origin licensing restricts loop and TAD formation in G1 phase zygotes
(**A**) Schematic of loop extrusion as detected via snHi-C versus bulk Hi-C. Extruding DNA loops result in stochastically distributed contacts in corresponding maps. CTCF oriented in a convergently manner can act as boundary to loop extrusion, leading to anchored loops which are discovered through snHi-C by averaging loop coordinates detected in mouse embryonic fibroblasts. (**B**) Average loop and TAD strength for maternal and paternal nuclei isolated from wildtype (WT) and MCM loss G1 phase zygotes. Data displayed are based on n = 13, 16, 16 and 15 (from left to right) from 4 independent experiments using 4-6 females. All heat maps were normalized to an equal number of reads.

called geminin, which directly targets Cdt1 and thus inhibits MCM loading (Figure 5.1A, left) (McGarry and Kirschner, 1998; Wohlschlegel et al., 2000). Upon cell cycle progression, geminin is degraded by the anaphase-promoting complex/cyclosome (APC/C), which reinitiates origin licensing. A point mutation in the destruction box generates a version of geminin (geminin^{L26A}) resistant to degradation by APC/C, causing geminin-mediated inhibition of origin licensing in G1 phase (Figure 5.1A, right) (Wohlschlegel et al., 2000). To utilize this in our OZT system, we injected mouse oocytes with mRNA coding for either GFP alone or with geminin^{L26A} (Figure 5.1B). Subsequently, these oocytes were matured and fertilized *in vitro* to form zygotes. Importantly, almost no chromatin-bound MCMs were present in geminin^{L26A}-expressing ("MCM loss") G1 zygotes as determined by immunofluorescence after extracting non-chromatin-bound protein (Figures 5.1C and 5.1D), showing efficient inhibition of origin licensing.

Next, we used this strategy to compare loops and TADs in G1 phase of maternal and paternal pronuclei in control versus MCM loss zygotes as detected by single-nucleus Hi-C (snHi-C) (Figure 5.1B). Unfortunately, *de novo* loop calling was not feasible as not enough zygotic Hi-C data were available. Thus, we utilized loop coordinates derived from previously published Hi-C data in mouse embryonic fibroblasts, revealing cohesin-dependent contacts of chromatin in zygotes (Figure 5.2A) (Silva et al., 2020). Strikingly, the strength of average loops and TADs was highly increased in MCM loss zygotes with the effect being even more pronounced in paternal than maternal chromatin (Figure 5.2B). The observed global increased strength of loops and TADs in MCM loss zygotes could derive from increased accessibility of cohesin complexes to CTCF sites in the genome, triggered by loss of MCM which is potentially acting as a chromatin-bound barrier.

To clarify whether the increased strength of loops and TADs observed by snHi-C in MCM loss zygotes was in fact mediated by cohesin complexes, we used a conditional genetic knockout approach in our OZT system. For this, alleles of the cohesin subunit *Scc1* were floxed which allowed their deletion upon expression of Cre recombinase from a Zp3 promoter in growing oocytes (Gassler et al., 2017; Ladstatter and Tachibana-Konwalski, 2016). To engineer maternal *Scc1* knockout zygotes (*Scc1*$^{\Delta(m)/+(p)}$), we isolated *Scc1*$^{\Delta\Delta}$ oocytes from *(Tg)Scc1*$^{fl/fl}$ *Zp3-Cre* female mice, injected mRNA coding for geminin^{L26A} and performed *in vitro* maturation and fertilization as described above (Figure 5.3A). Importantly, in line with a previous study (Gassler et al., 2017), no loops and TADs could be detected in *Scc1*$^{\Delta(m)/+(p)}$ zygotes which remained unaffected in the absence of chromatin-bound MCM (MCM loss condition) (Figure 5.3B). Thus, we infer that MCM DHs impede the formation of cohesin-dependent loops and TADs.

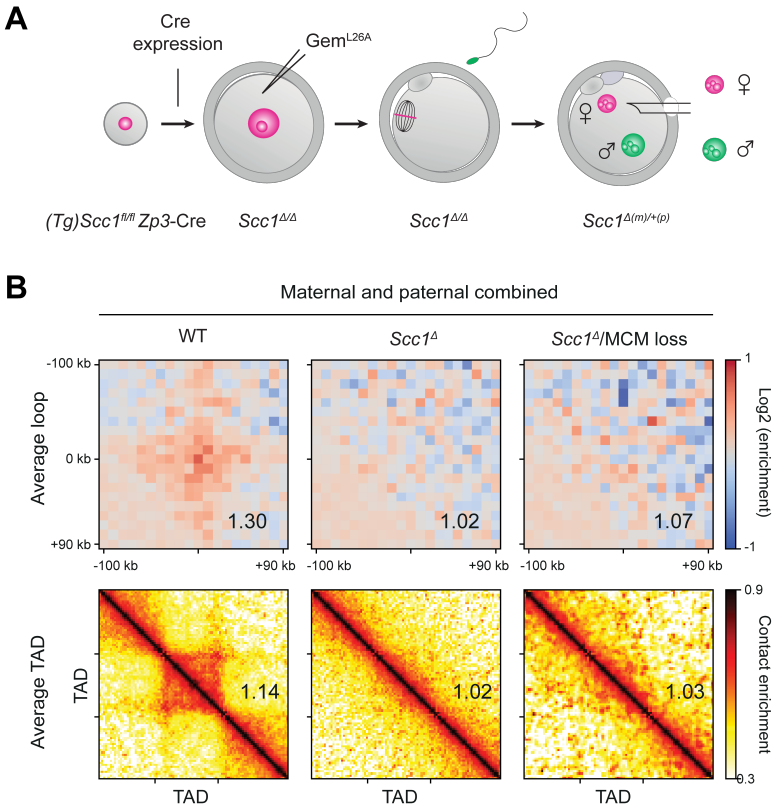

Figure 5.3. MCM impedes the formation of cohesin-dependent loops and TADs

(**A**) Generation of zygotes with a conditional genetic maternal *Scc1* knockout. During oocyte growth, Cre recombinase is expressed from a Zp3 promoter which leads to deletion of floxed *Scc1* alleles and yields maternal knockout *Scc1*$^{\Delta(m)/+(p)}$ zygotes upon further maturation and fertilization. (**B**) Average loop and TAD strength in control, *Scc1* depleted (*Scc1*$^{\Delta}$) and *Scc1* depleted, MCM loss (*Scc1*$^{\Delta}$/MCM loss) zygotes. Data displayed are pooled from paternal and maternal nuclei and based on n = 26, 44 and 10 nuclei (from left to right). All heat maps were normalized to an equal number of reads. Control and 38 *Scc1*$^{\Delta}$ samples were previously published (Gassler et al., 2017).

5.2 MCM DHs are barriers to cohesin translocation *in vitro*

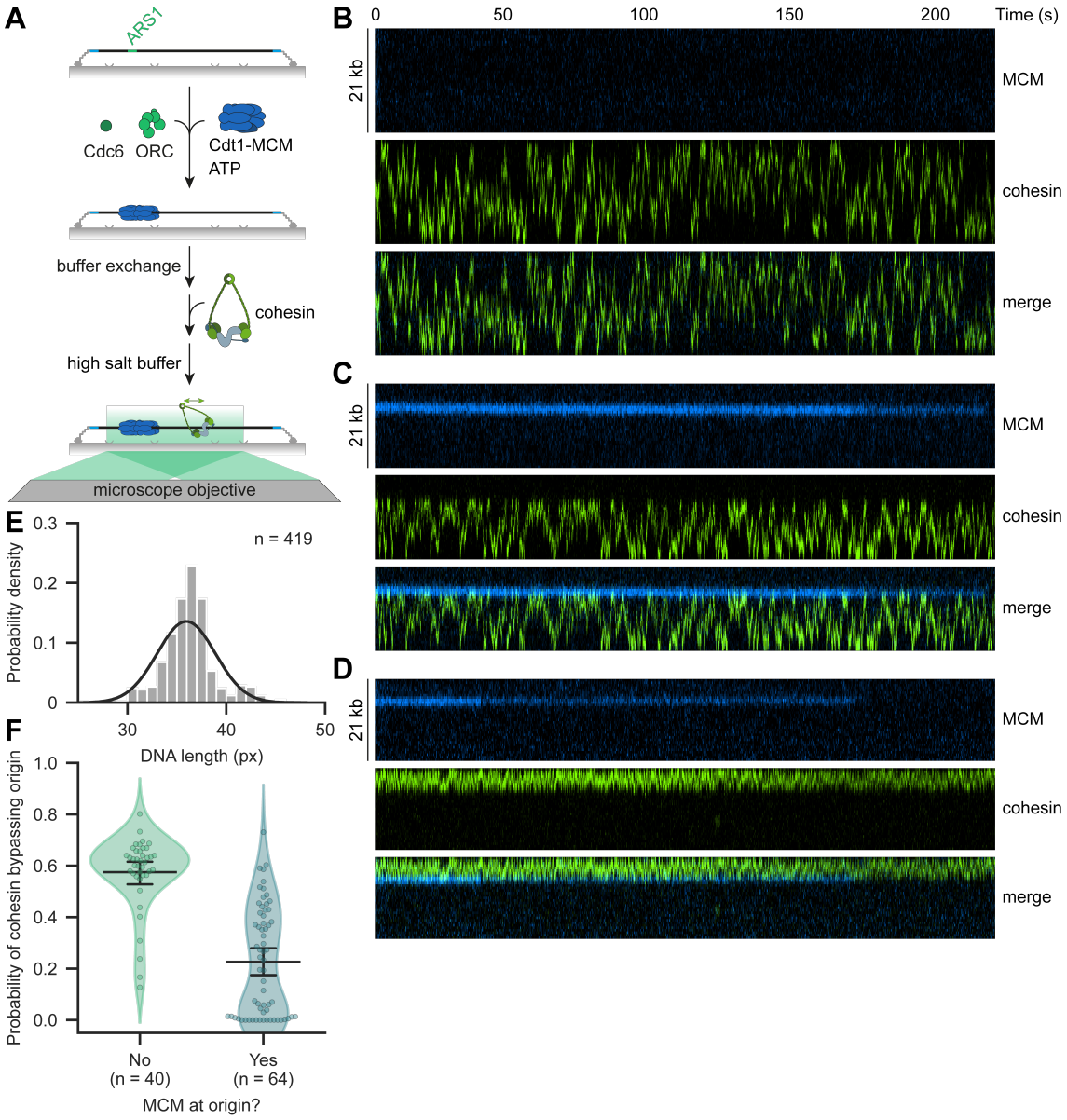

Figure 5.4. MCM DHs are barriers to cohesin translocation at high salt *in vitro*
(**A**) Schematic of the single-molecule cohesin translocation assay at high salt on licensed DNA. First, MCM DHs were loaded on doubly-tethered 21 kb DNA, containing the origin ARS1, in the presence of licensing factors ORC and Cdc6. Subsequently, cohesin was loaded and cohesin translocation was visualized at high salt (0.5 M NaCl) in the absence of free protein and buffer flow. (**B**, **C** and **D**) Representative kymographs of translocating cohesin on bare (**B**) and licensed (**C** and **D**) DNA, demonstrating that origin-bound MCM DHs were efficient barriers to translocating cohesin during a 220 s observation interval. (**E**) Length distribution of doubly-tethered DNA in pixels (px). The line represents a Gaussian fit. (**F**) Probability of translocating cohesin bypassing the origin in the absence or presence of MCM calculated from 40 or 64 molecules with 7802 or 9829 visualized encounters, respectively. Black lines display the mean within a 95 % confidence interval (CI) by bootstrapping. See chapter 2.7.6 for details. All data displayed were generated exclusively in the presence of 0.5 M NaCl, except in (**E**).

The most straightforward hypothesis to explain the effect of MCM loss on average loops and TADs in our snHi-C experiments (chapter 5.1) is that MCM interferes with cohesin-mediated loop extrusion by acting as a barrier being randomly distributed on chromosomes. According to this hypothesis, MCMs should then represent physical barriers to cohesin on DNA. To directly test this hypothesis, we developed an MCM roadblock assay for passive cohesin translocation using our TIRF microscopy setup which is able to detect cohesin-MCM interactions at the single-molecule level with high real-time spatial resolution. To this end, we utilized the established single-molecule origin licensing assay as described and characterized in detail in chapter 3 with some modifications. First, we immobilized origin (ARS1)-containing 21 kb DNA at both ends to the surface of a functionalized coverslip and subsequently added origin licensing factors ORC, Cdc6 and fluorescently labeled Cdt1-MCM, leading to robust MCM DH formation (Figure 5.4A) (Remus et al., 2009). Upon removal of unbound licensing factors, the flow cell was flushed with fluorescently labeled cohesin under mild salt conditions to allow cohesin loading on DNA. Finally, the flow cell was washed with a high salt buffer containing 0.5 M NaCl, to select for topologically engaged MCMs, remove loading intermediates (Figures 3.1C, 3.1D and 3.2E) and promote cohesin translocation on fast timescales in the absence of buffer flow (Figure 5.4B; Movie 16) (Davidson et al., 2016; Stigler et al., 2016).

Having established a cohesin translocation assay, we were able to directly visualize cohesin encounters with MCM DHs acting as potential roadblocks. Remarkably, MCM DHs highly constrained cohesin translocation in our setup (Figures 5.4C and 5.4D; Movie 17). To quantify the probability of cohesin bypassing the origin, we applied subpixel localization and tracking to automatically detect cohesin-MCM encounters and classify the outcomes into bypassing and blocking events (see chapter 2.7.6 for details). In doing so, we accounted for different resolution derived from differently stretched doubly-tethered DNA (Figure 5.4E) by determining the mean DNA extension and adjusting thresholds for the individual length of each DNA molecule. This detailed quantification revealed that the probability of cohesin bypassing the origin is about 2.5 times reduced (~57 % to ~23 %) in the presence of an MCM DH at the origin (Figure 5.4F). Notably, the observation that cohesin is only able to bypass an MCM-free origin in approximately half of the cases is in line with passive cohesin translocation being a random walk. Cohesin was also able to frequently bypass MCM DHs (Figures 5.5A and 5.5B; Movie 18). Importantly, cohesin showed intervals of efficient and inefficient MCM passage on the same DNA molecule, excluding the possibility of surface stuck DNA at origin-bound MCM DHs being the reason for observed barrier characteristics.

We found that MCM DHs are strong barriers (defined as a 5-fold reduction in origin bypassing probability compared to cohesin alone) for ~45 % of cohesin molecules under these conditions (29/64, Figure 5.4F). The observation that, despite frequent attempts, some cohesin molecules were not able to bypass the origin even once during the 220 seconds imaging window (12/64, Figures 5.4C, 5.4D and 5.4F; Movie 17), indicates that MCM DHs can also act as impermeable barriers over these timescales. Importantly,

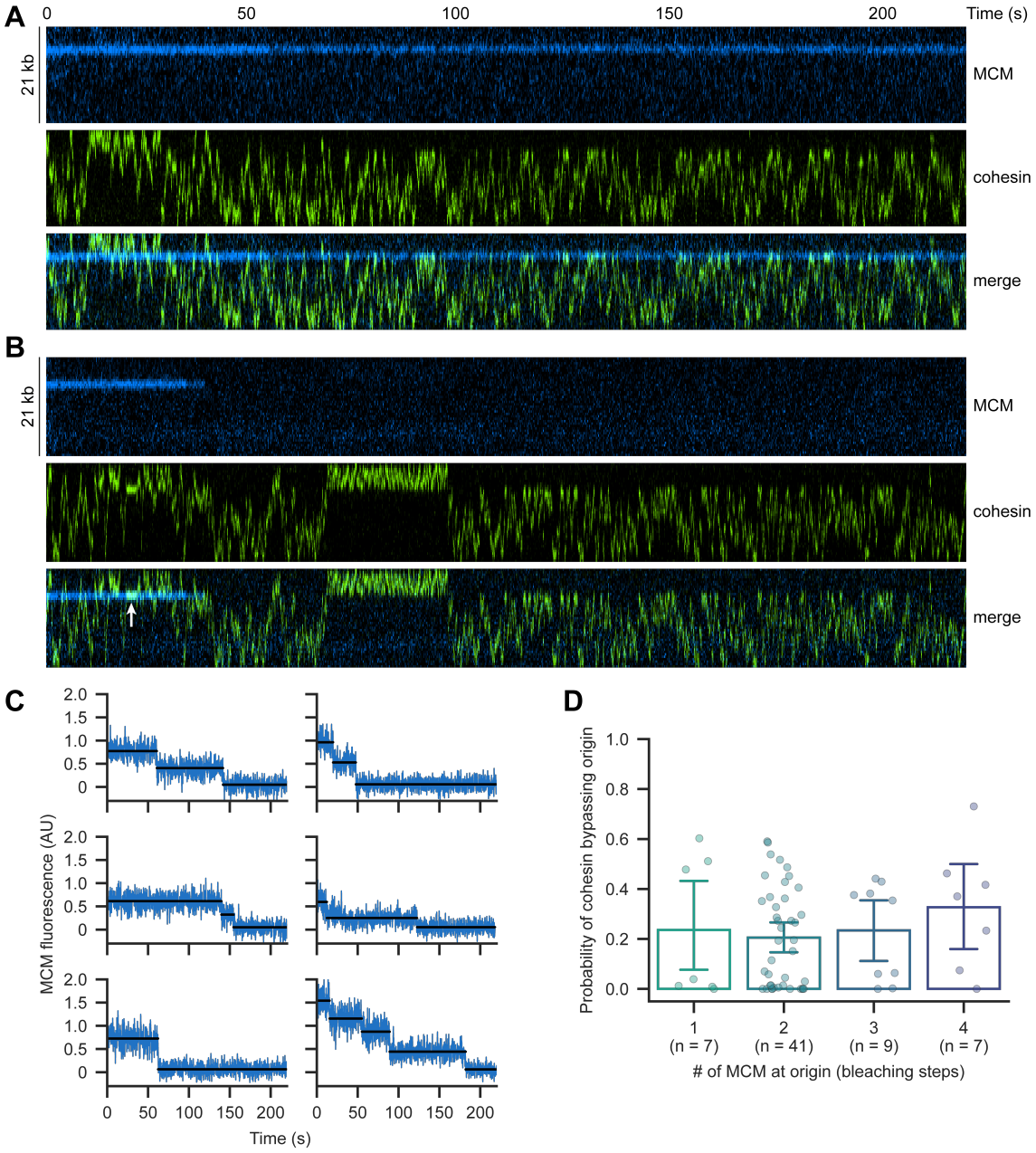

Figure 5.5. Translocating cohesin can bypass MCM DHs with reduced efficiency at high salt
(**A** and **B**) Representative kymographs of translocating cohesin on licensed DNA, demonstrating that while MCM DHs were efficient barriers, intervals of efficient and inefficient MCM passage were observed during a 220 s interval. White arrow in (**B**) indicates a short interval of cohesin translocation pausing upon MCM encounter. (**C**) Example MCM bleaching trajectories during the cohesin translocation assay confirming MCM DH formation but also multiple MCM loading as analyzed by a fit with the kinetic change point algorithm (black line) (also see chapter 3.2). (**D**) Probability of cohesin bypassing the origin versus the number of MCM present at the origin as determined by photobleaching. Error bars represent a 95 % CI by bootstrapping. All data displayed were generated exclusively in the presence of 0.5 M NaCl.

the failure to bypass the origin even once was only the case for licensed origins, since all cohesin molecules readily translocated over origins without MCM DH (40/40, Figures 5.4B and 5.4F; Movie 16). ORC and other origin licensing intermediates were expected to not or only very rarely be present at origins as these complexes dissociate from DNA

when challenged with high salt (Figures 3.1C, 3.1D and 3.2E). However, we have shown that origins are sequestered by multiple MCMs during licensing (chapter 3.2).

To study the effect of additional MCM complexes at the origin, which might form artificial obstacles, we analyzed the photobleaching behavior of MCMs during translocation assays to distinguish between single, double and multiple hexamers (Figure 5.5C). Although the majority of MCM foci contained two MCMs (MCM DH), we found additional species, as expected. Importantly, we did not observe an increased barrier strength of multiple MCMs (Figure 5.5D), however the number of observations other than for MCM DHs was low. Notably, single MCM hexamers (as determined by photobleaching) seemed sufficient to restrict cohesin translocation, suggesting that a single MCM could also be a barrier to loop extrusion, thus active CMG could displace cohesin ahead of the replication fork. However, and although single MCM hexamers are known to be resistant to high salt (Champasa et al., 2019), based on our labeling efficiency of ~90 %, a total fraction of 10-20 % could be accounted to MCM DHs with one MCM not being labeled.

Boundary factors could act through different mechanisms (see chapter 1.5.2). One possibility is that the boundary could be a passive, physical barrier, simply functioning as a “roadblock”. Alternatively, the boundary could contain a binding site for cohesin, thus restricting loop extrusion by directly tethering cohesin and/or protecting from release by WAPL, hence acting as an “active boundary”. Importantly, different mechanisms are not mutually exclusive. Our observation that multiple MCMs do not form a larger barrier are surprising as an increased barrier strength would be intuitive for both mechanisms. Nevertheless, this might be an indication for MCM rather being a physical roadblock than an active boundary but this assumption highly depends on specific characteristics of the underlying, hypothetical model of cohesin translocation. The conformation and DNA binding mode of cohesin during passive translocation are unknown, but cohesin could be a closed rod or open ring (Hons et al., 2016; Huis in 't Veld et al., 2014; Soh et al., 2015) and switch between these two states while being topologically engaged with dsDNA. In its open state, cohesin has a large central pore with a diameter of ~35 nm (Huis in 't Veld et al., 2014) which would easily accommodate an MCM with a diameter of ~13 nm (Li et al., 2015). If cohesin translocates on DNA based on a hopping mechanism, not only all electrostatic interactions with dsDNA but also with MCMs could be shielded by counterions and lost (see chapter 3.3). Thus, once cohesin accommodated one MCM, it might readily bypass additional MCMs (to a certain extent) as only the distance increases with multiple stacked MCMs but not the 1-dimensional surface along DNA. Intriguingly, the probability of bypassing MCM might then be dominated by the equilibrium between an open and closed state of cohesin, which is in line with our observation that cohesin showed intervals of efficient and inefficient MCM passage on the same DNA molecule (Figures 5.5A and 5.5B). Alternatively, interaction motifs of additional MCMs might be buried by flanking MCMs, leading to no additional boundary function.

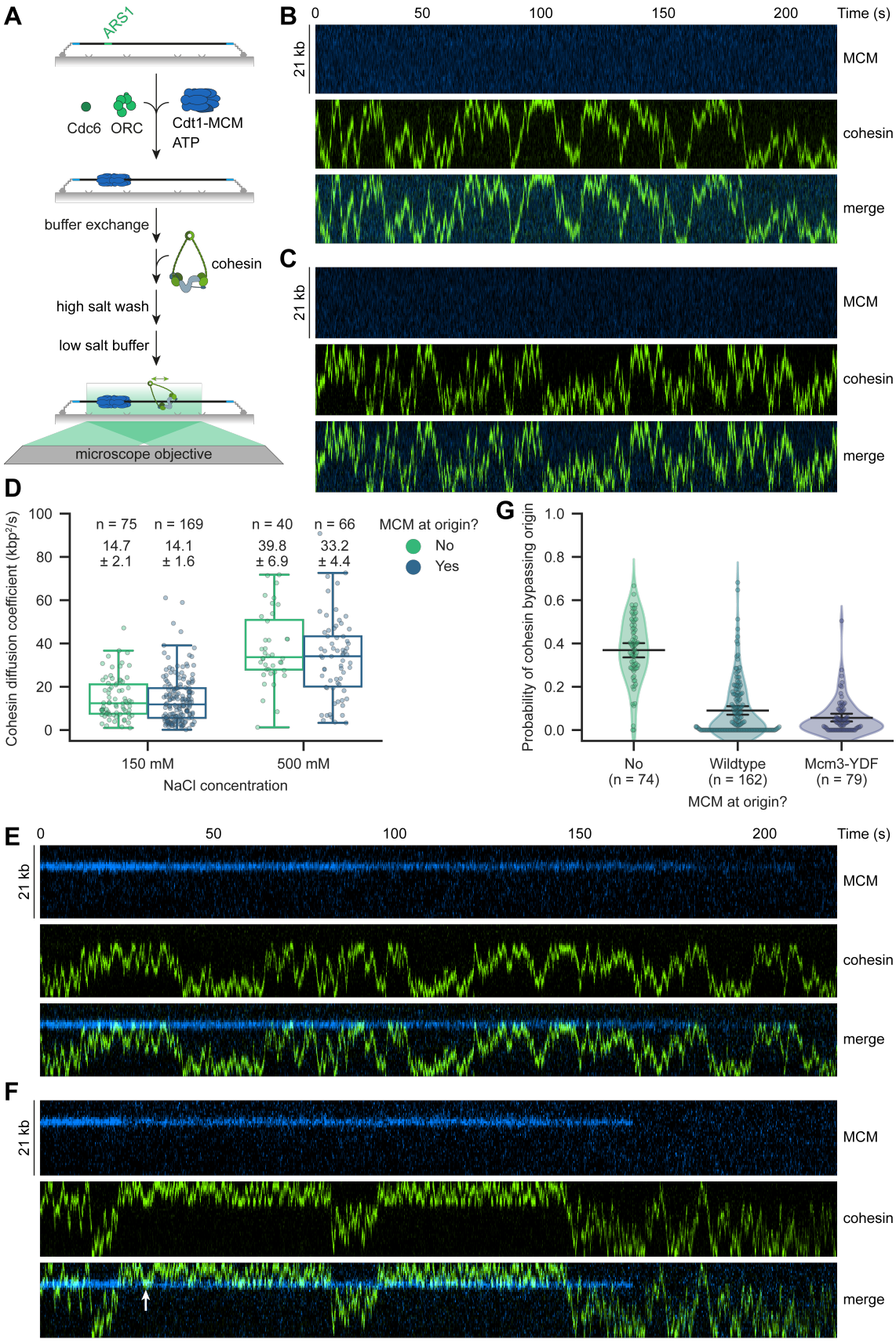

Figure 5.6. Cohesin translocation is similarly affected by MCM DHs at physiological salt

(**A**) Schematic of the single-molecule cohesin translocation assay at physiological salt concentration on

(legend continued on next page)

licensed DNA. MCM and cohesin loading was performed as described in Figure 5.4A and cohesin translocation was visualized at physiological salt concentration (0.15 M NaCl) upon performing a high salt wash. (**B** and **C**) Representative kymographs of translocating cohesin on bare DNA at physiological salt concentration. (**D**) Salt concentration but not MCM alters observed cohesin translocation rate. Values above the box plots indicate the mean diffusion coefficients in kbp^2/s and the range of a 95 % CI by bootstrapping. (**E** and **F**) Representative kymographs of translocating cohesin on licensed DNA at physiological salt concentration, showing that origin-bound MCM DHs were efficient barriers to cohesin translocation (**E**) but intervals of efficient and inefficient passage were observed (**F**) during a 220 s interval. White arrow in (**F**) indicates a short interval of cohesin translocation pausing upon MCM encounter. (**G**) Probability of translocating cohesin bypassing the origin in the absence or presence of either $MCM^{wildtype}$ or $MCM^{Mcm3\text{-}YDF}$ at physiological salt concentration calculated from 74, 162 or 79 molecules with 12175, 15348 or 9455 visualized encounters, respectively. Black lines display the mean within a 95 % CI by bootstrapping.

We also considered another alternative, caused by the design of our setup, explaining why multiple MCMs do not increase the barrier strength. All cohesin translocation experiments shown so far were performed at high, non-physiological salt concentration (0.5 M NaCl) which could not only disrupt a specific interaction between MCMs and cohesin, but also not properly resemble the barrier function under physiological conditions present in cells. To consolidate our observations at high salt, we modified our single-molecule MCM roadblock assay to allow visualization of cohesin translocation under more physiological conditions. For this, we changed to a low salt imaging buffer (0.15 M NaCl) upon performing a high salt wash (to select for topologically engaged MCMs, remove loading intermediates and promote cohesin translocation as before; Figure 5.6A). Again, we observed cohesin translocation on fast timescales in the absence of buffer flow and MCM DH (Figures 5.6B and 5.6C; Movie 19). However, translocation occurred at substantially lower rates compared to high salt (Figure 5.4B) with the diffusion coefficient decreasing from 39.8 ± 6.9 kbp^2/s or 3.2 ± 0.4 $\mu m^2/s$ at high salt to 14.7 ± 2.1 kbp^2/s or 1.2 ± 0.2 $\mu m^2/s$ at low salt (mean in range of a 95 % CI) (Figure 5.6D; Table 3.1).

The dependence of diffusion coefficients on the underlying salt concentration observed here and in a previous study (Stigler et al., 2016) suggests that passive cohesin translocation might in fact be based on a hopping mechanism. The theoretical upper limit of non-helical and helical diffusion cannot be accurately calculated for cohesin due to the unknown conformational state during translocation. However, even when assuming the simplest shape of a sphere (which is not the case), cohesin with a molecular weight of ~500 kilodaltons would have a minimal diameter of 10.4 nm (Erickson, 2009) and thus an upper limit of ~44 $\mu m^2/s$ and ~0.4 $\mu m^2/s$ for non-helical and helical diffusion, respectively. Since observed cohesin diffusion coefficients exceed the upper limit of helical diffusion and combined with the dependence on ionic strength, we conclude that cohesin translocates on dsDNA through a hopping mechanism. In contrast, the presence of MCMs at origins did not substantially alter observed cohesin diffusion coefficients (33.2 ± 4.4 and 14.1 ± 1.6 kbp^2/s (mean in range of a 95 % CI) at high and low salt, respectively, Figure 5.6D). The slight decrease in observed diffusion most likely derives from apparent confined diffusion (time-resolution of our assay) of cohesin molecules diffusing on the short DNA side of the origin.

Next, we investigated the ability of MCM DHs acting as barriers to cohesin translocation under physiological salt conditions. Similar as observed at high salt, MCM DHs highly constrained cohesin translocation (Figure 5.6E; Movie 20) but cohesin was still able to frequently bypass MCM DHs (Figure 5.6F; Movie 21). However, the effect was even more pronounced with the probability of cohesin bypassing the origin being decreased about 4-fold (~37 % to ~9.0 %) in the presence of an MCM DH at the origin (Figure 5.6G, $MCM^{wildtype}$). In line with the higher barrier strength at low salt, we also detected an increased fraction of molecules in which MCM DHs were strong (~60 %, 97/162) or impermeable (~41 %, 67/162) barriers within our imaging time (Figures 5.6E and 5.6G; Movie 20) compared to high salt (Figure 5.4F). Notably, cohesin molecules again readily translocated over origins without MCM DH (72/74, Figures 5.6B and 5.6C; Movie 19), whereas two molecules did not, which can be explained by fully unlabeled MCM DH at the origin and being entirely consistent with our labeling efficiency of ~90 %. Moreover, MCM photobleaching analysis revealed that the barrier strength remained unchanged with increasing number of MCMs, based on a reliable number of observations (Figure 5.7A). Together, these observations at physiological salt concentration support the model of MCM being a passive physical roadblock rather than an active boundary like CTCF, however the possibility of buried, specific binding sites could not be excluded.

5.3 Cohesin translocation pauses upon transient MCM binding

Our model in which MCM is a passive physical roadblock predicts that cohesin passage is simply altered by steric hindrance and thus cohesin would not show an increased residence time at MCM-bound origins mediated by specific interactions. Surprisingly, although very rarely, we detected segments in trajectories in which cohesin translocation paused at MCM DHs (Figures 5.5B and 5.6F, white arrow). Intriguingly, these pauses could represent a specific interaction, favoring a model of MCM also being an active boundary instead of exclusively being a passive barrier.

Recent structural studies by Yan Li and coworkers revealed a YDF-containing motif in the N-terminus of CTCF which binds to the cohesin subunits SCC1 and STAG2 in a competitive manner with WAPL, suggesting that CTCF's N-terminus blocks cohesin-mediated loop extrusion by protecting cohesin from release by WAPL (Li et al., 2020). However, mutations in this YDF-containing motif did not fully abrogate enrichment of cohesin at CTCF sites which indicates the presence of additional mechanisms to block loop extrusion. In line with this observation, this study identified a variety of proteins besides CTCF containing a general [Y/F]xF motif, among them WAPL, Sororin (which also protects cohesin from WAPL), but also human MCM with a YDF-containing motif in the C-terminal region of Mcm3 (Li et al., 2020). Pairwise sequence alignment with *S. cerevisiae* Mcm3 revealed that this YDF motif in human Mcm3 is not conserved in budding yeast (Figure 5.7B). Since we are using a mixed system of yeast MCM and human cohesin in our *in vitro* single-molecule assay and hence might have disrupted species-specific interactions between cohesin and MCM, we asked whether the YDF motif in Mcm3 would influence the translocation behavior of cohesin.

To investigate this, we engineered a yeast strain to express a "humanized" version of MCM in which the corresponding similar motif of yeast Mcm3 was replaced by an extended version of the described YDF motif in human Mcm3 ($MCM^{Mcm3\text{-}YDF}$) (Figure 5.7B; chapter 2.3.5). We chose a 19 amino acid stretch of the human sequence (instead of 11 as identified in the peptide screen in (Li et al., 2020)) because while the sequences located N- and C-terminally have been lost, interestingly, the 19 amino acid stretch containing the YDF motif seems conserved in the evolution from yeast to human (Figure 5.7B). The generated yeast strain with all Mcm3 alleles modified to Mcm3-YDF was viable and grew comparably to the parental strain containing wildtype Mcm3, confirming that the YDF motif did not alter the replicative function of MCM (data not shown).

Next, we examined the effect of the YDF motif by comparing $MCM^{Mcm3\text{-}YDF}$ to $MCM^{wildtype}$ in our *in vitro* single-molecule cohesin translocation assay under physiological salt conditions (Figure 5.6A). Remarkably, $MCM^{Mcm3\text{-}YDF}$ DHs constrained cohesin translocation to an even higher extent with the probability of cohesin bypassing the origin being reduced around 6.6-fold (~37 % to ~5.6 %) or 1.6-fold (~9.0 % to ~5.6 %) compared to the absence or presence of $MCM^{wildtype}$ at the origin, respectively (Figure 5.6G). In line with the increased barrier strength observed, the fraction of molecules in

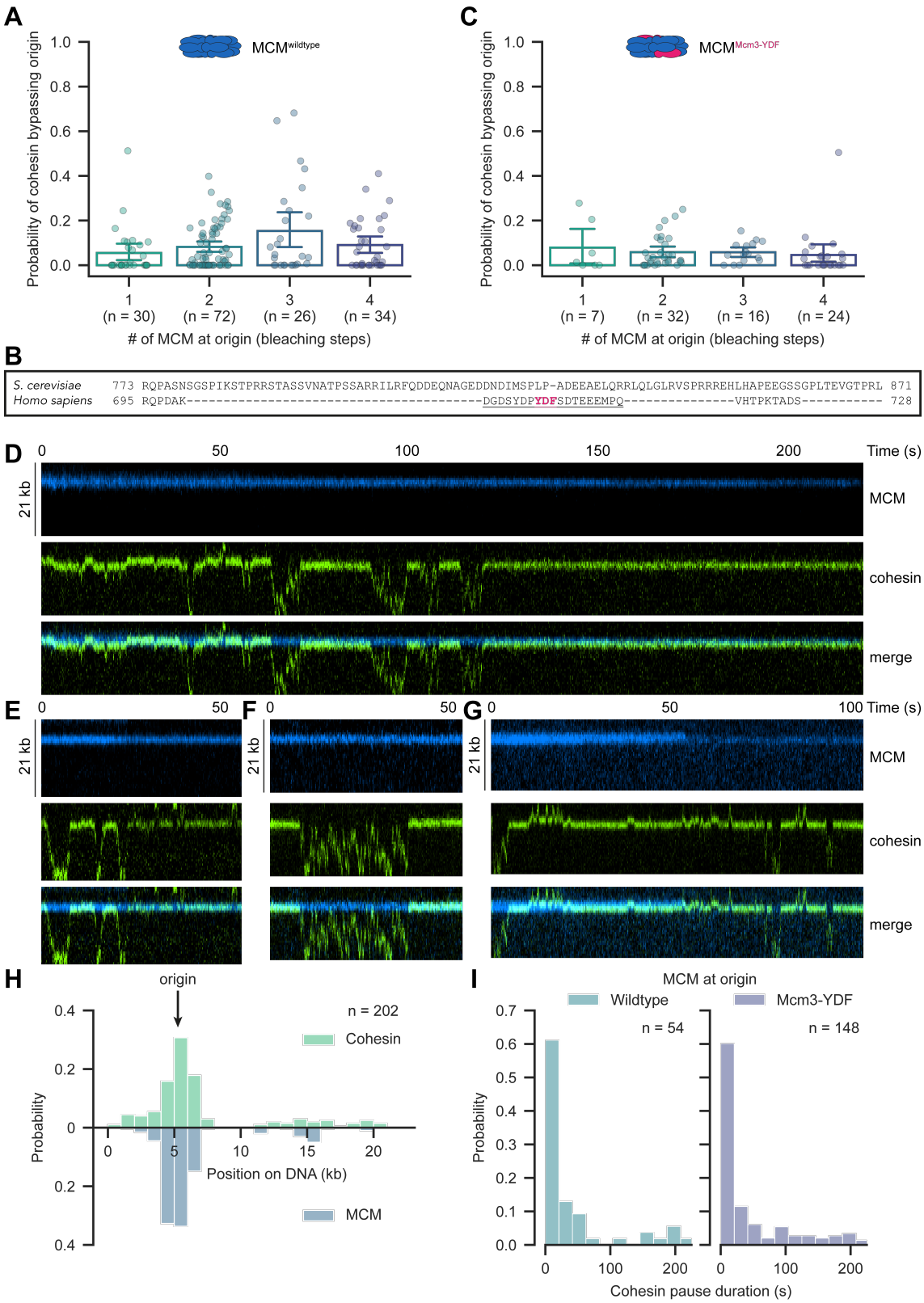

Figure 5.7. Cohesin translocation frequently pauses at MCM-bound origins

(**A** and **C**) Probability of cohesin bypassing the origin versus number of MCM^{wildtype} (**A**) or MCM$^{Mcm3\text{-}YDF}$ (**C**) present at the origin as determined by photobleaching. Error bars represent a 95 % CI by bootstrapping. (**B**) Excerpt of a pairwise sequence alignment (using the Needleman-Wunsch algorithm) of *S. cerevisiae* and *Homo sapiens* Mcm3. The YDF-containing motif in human Mcm3 replaced with the yeast sequence to

(legend continued on next page)

generate MCM$^{Mcm3\text{-}YDF}$ is underlined. (**D**, **E**, **F** and **G**) Representative kymographs demonstrating that translocating cohesin frequently pauses at origins bound by an MCM$^{Mcm3\text{-}YDF}$ DH. (**H**) Distribution of cohesin pause (top) and corresponding MCM positions (bottom). (**I**) Distribution of cohesin pause durations (in seconds) in the presence of MCM^{wildtype} (left) or MCM$^{Mcm3\text{-}YDF}$ (right) at the origin. All data displayed were generated exclusively in the presence of 0.15 M NaCl.

which MCM$^{Mcm3\text{-}YDF}$ DHs were strong (~73 %, 58/79) or impermeable (~32 %, 25/79) barriers within our imaging time slightly increased or were comparable to MCM^{wildtype} DHs (Figure 5.6G). However, cohesin was still able to occasionally bypass MCM$^{Mcm3\text{-}YDF}$. Interestingly, in contrast to MCM^{wildtype}, we observed a mild decrease in probability of cohesin bypassing the origin with increasing number of MCM$^{Mcm3\text{-}YDF}$ being present (MCM^{wildtype}: 5.4 %, 8.1 %, 15.3 % and 9 % – MCM$^{Mcm3\text{-}YDF}$: 7.9 %, 5.8 %, 5.8 % and 4.6 % for 1, 2, 3 and 4 MCMs at the origin, respectively; Figures 5.7A and 5.7C). Thus, we speculate that besides a passive barrier function, MCM could act as an active boundary by containing a specific binding site for cohesin.

To test this hypothesis, we inspected all cohesin trajectories on DNA molecules containing MCM$^{Mcm3\text{-}YDF}$ with regard to pauses occurring during translocation as also rarely seen in the presence of MCM^{wildtype} (Figures 5.5B and 5.6F, white arrow). Strikingly, we detected many cohesin trajectories in which translocation was interspersed with frequent pausing at MCM$^{Mcm3\text{-}YDF}$-bound origins (Figures 5.7D, 5.7E, 5.7F and 5.7G; Movie 22). To quantify and further characterize the pausing behavior, we applied an unbiased KCP fit to all cohesin trajectories to detect pauses. Notably, we excluded identified pause segments from all barrier analysis (instead each pause was counted as one failed bypassing event) and diffusion coefficient calculations throughout the study (see chapters 2.7.4 and 2.7.6). Intriguingly, KCP analysis revealed that cohesin translocation pausing in fact almost exclusively occurred at the origin position on DNA, coinciding with MCM binding sites, demonstrating that MCMs are responsible for cohesin translocation pausing (Figure 5.7H). Next, we determined the mean duration of pauses, however, we observed a similar distribution for pauses occurring at MCM^{wildtype} and MCM$^{Mcm3\text{-}YDF}$ with a mean duration of 42.2 and 41.8 seconds, respectively (Figure 5.7I).

Given our observation that cohesin translocation was interspersed with pauses, cohesin can switch between two modes in our assay (Figure 5.8A). If the underlying ionic strength is sufficient, cohesin translocates readily along DNA but translocation pausing or stalling can occur upon encountering an MCM DH, presumably mediated by a transient, direct interaction between cohesin and MCM. Assuming the Mcm3-YDF motif is involved in cohesin interaction, we predict that MCM$^{Mcm3\text{-}YDF}$ would shift the equilibrium between these two modes more towards cohesin pausing/stalling than MCM^{wildtype} (Figure 5.8A). Strikingly, the mean fraction of cohesin pausing highly increased from 6.4 % to 43 % in the presence of MCM^{wildtype} and MCM$^{Mcm3\text{-}YDF}$, respectively, while we did not observe any (0 %) pausing in the absence of MCM (Figure 5.8B). Furthermore, while the mean fraction of cohesin pausing was independent of the number of MCM^{wildtype} (8.9 %, 5.4 %, 10 % and 3.5 % for 1, 2, 3 and 4 MCMs, respectively; Figure 5.8C), we observed a tendency of increased pausing with increasing number of MCM$^{Mcm3\text{-}YDF}$ at the origin

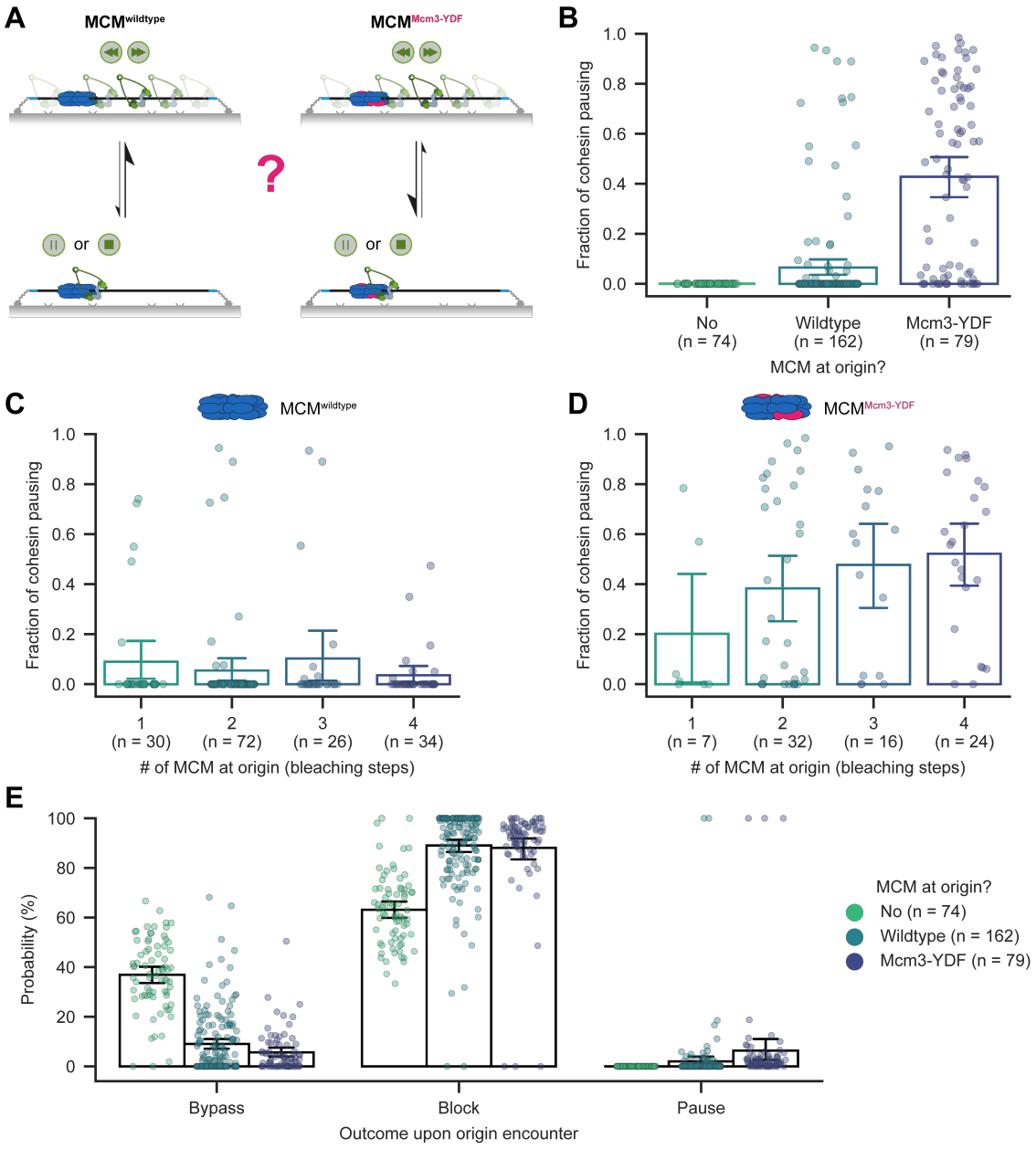

Figure 5.8. A YDF-containing motif in human Mcm3 mediates transient binding of cohesin
(**A**) Schematic of the working hypothesis in which cohesin can switch between two modes. In the presence of sufficient ionic strength, cohesin translocates rapidly along DNA (top). Upon encountering an MCM DH acting as active boundary element, interaction with cohesin can lead to translocation pausing or permanent stalling (bottom). If the Mcm3-YDF motif is involved in cohesin interaction, $MCM^{Mcm3\text{-}YDF}$ is predicted to shift the equilibrium between these two modes more towards cohesin pausing/stalling than $MCM^{wildtype}$. (**B**, **C** and **D**) Fraction of cohesin pausing of the total observation time in the absence or presence of either $MCM^{wildtype}$ or $MCM^{Mcm3\text{-}YDF}$ (**B**) or versus number of $MCM^{wildtype}$ (**C**) or $MCM^{Mcm3\text{-}YDF}$ (**D**) present at the origin as determined by photobleaching. (**E**) Probabilities of different outcomes (bypass, block or pause) of translocating cohesin upon encountering the origin in the absence or presence of either $MCM^{wildtype}$ or $MCM^{Mcm3\text{-}YDF}$ calculated from 74, 162 or 79 molecules with 12175, 15348 or 9455 visualized encounters, respectively. Bypass outcome is the same as shown in Figure 5.6G. Error bars in (**B**, **C**, **D** and **E**) represent a 95 % CI by bootstrapping. All data displayed were generated exclusively in the presence of 0.15 M NaCl.

(20 %, 38 %, 48 % and 52 % for 1, 2, 3 and 4 MCMs, respectively; Figure 5.8D). In line with these findings, the probability of translocation pausing upon encountering the origin steadily increased from 0 % in the absence of MCM to 1.9 % and 6.3 % in the presence of $MCM^{wildtype}$ and $MCM^{Mcm3\text{-}YDF}$, respectively (Figure 5.8E). Importantly, both MCM variants blocked translocating cohesin in a similar fashion, increasing the probability of a directional change of translocation at the origin around 1.4-fold from 63 % in the absence to 89 % or 88 % in the presence of either $MCM^{wildtype}$ or $MCM^{Mcm3\text{-}YDF}$, respectively (Figure 5.8E).

Together, our single-molecule data demonstrate that MCM DHs are barriers that restrict cohesin translocation *in vitro*. MCM complexes exert their barrier function by two different mechanisms. First, MCM DHs represent passive, physical barriers by functioning as roadblocks, blocking cohesin passage due to steric hindrance and/or by topologically engaging dsDNA, independently of specific binding motifs (same behavior of $MCM^{wildtype}$ and $MCM^{Mcm3\text{-}YDF}$). Second, due to specific binding sites for cohesin, MCM DHs can tether cohesin to the origin as active boundaries (different behavior of $MCM^{wildtype}$ and $MCM^{Mcm3\text{-}YDF}$). Thereby, the YDF-containing motif in human Mcm3 might mediate the interaction with cohesin, but additional, yet to be identified sites might be present in MCM complexes as pausing also occurred in the absence of the YDF motif. However, our interpretations with regard to loop extrusion rely on the assumption that a barrier to passive cohesin translocation will also impair active cohesin-mediated loop extrusion. Thus, a limitation of our single-molecule assay used here is that we cannot necessarily infer from passive translocation to active loop extrusion as we do not know if these two cohesin modes are similar. Therefore, future studies should aim to elucidate whether active loop extrusion is restricted by MCM-licensed origins. However, this will require a combined assay that has proven technically challenging with respect to identifying suitable reaction conditions for both processes *in vitro*. Taken together, we conclude that MCMs are both physical barriers and active boundaries to cohesin translocation, which may occasionally be bypassed.

Chapter 6

Discussion

Different cellular processes rely on the same DNA template, such as replication and transcription but also genome organization which is performed by SMC complexes such as cohesin. Progressing development of novel, more sensitive techniques in the last years greatly changed our mechanistic view of these cellular processes, which initially were thought to be performed by exclusively static, robust machineries. It now became more and more clear that cellular machineries cope with this exceedingly complex environment by allowing a certain degree of plasticity such as alternative pathways and dynamic exchange processes to mitigate conflicts between them. However, conflicts concomitant with the interplay between two partially static acting machineries can disrupt, delay but also modify their function which has shown to be part of essential regulatory mechanisms in cells.

In the first part of this thesis, we clarified the consequences of dynamic challenges to origin licensing by reconstituting the process at the single-molecule level with high temporal and spatial resolution (chapter 3). This allowed for direct observations of the departure of licensing factors, changes in the composition of loading intermediates, and tracking of the positions of individual factors at and around replication origins as a function of time. Our findings strongly support a model in which origin licensing intermediates overcome transcription conflicts by adopting a diffusive DNA binding mode and repositioning in front of advancing RNAPs (chapter 4). We observe that the stability of origin licensing intermediates increases as the pathway progresses with the final MCM DH being the most stable and ORC the least stable. Nevertheless, we find that RNAP can frequently bypass origin-bound ORC, compensating for its lower sliding stability. Taken together, our observations reveal numerous additional pathways for origin specification and resistance to transcription, which are summarized in Figure 6.1.

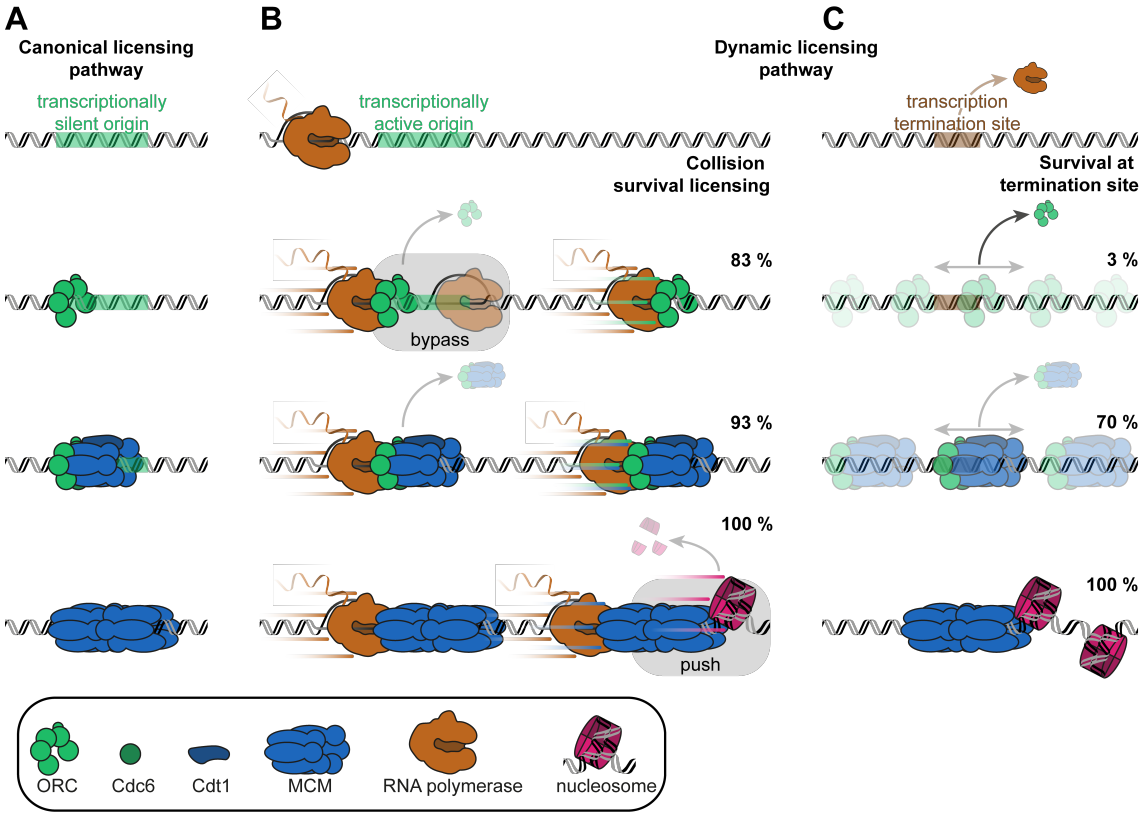

Figure 6.1. Mobile origin licensing factors confer resistance to transcription conflicts
(**A**) Canonical origin licensing pathway at transcriptionally silent origins. The MCM DH is loaded at the origin sequentially via multiple licensing intermediates. (**B**) Dynamic origin licensing pathway at transcriptionally active origins. All origin licensing intermediates can overcome transcription conflicts by adopting a diffusive DNA binding mode and repositioning in front of advancing RNA polymerases. The stability of origin licensing intermediates increases as the pathway progresses. However, ORC compensates for its lower sliding ability by an RNA polymerase bypass mechanism at the origin. Importantly, repositioning of licensing factors is possible in the presence of chromatin by forming mobile super complexes with nucleosomes adopting a diffusive DNA binding mode or being ejected. (**C**) Transcription termination sites could serve as non-canonical origins. Although ORC appears unstable at non-ARS sites, additional factors on chromosomes could restrict dissociation and 1D diffusion. Strikingly, the downstream licensing intermediate OCCM and final MCM DHs robustly remain on DNA upon transcription termination. Together, the origin licensing and firing pathway adapts to the complex challenges on chromosomes by proceeding at distant sites from canonical origins.

6.1 Mobile origin licensing factors confer resistance to transcription conflicts

Origins of replication are broadly distributed throughout eukaryotic chromosomes to ensure all regions are copied in a timely manner. This essential feature of chromosome architecture places additional demands on origin licensing factors that must conduct helicase loading in diverse and evolving local environments each with unique challenges. Encounters with polymerases and translocases, known to operate at the same cell cycle stage, pose significant risks. While some of the mechanisms that ensure genome integrity by overcoming conflicts at the replication fork are becoming clear (Branzei and Foiani, 2010; Sparks et al., 2019), the pathways of resistance at earlier stages of replication have not been well elaborated. Nevertheless, many studies, primarily on artificial substrates, have demonstrated that transcription is an acute example of an orthogonal process that can disrupt origin function (Looke et al., 2010; Tanaka et al., 1994). To reduce the frequency of conflicts, most origins have evolved in locations outside transcribed regions (Figure 6.1A), but some overlap appears unavoidable (Edwards et al., 2002; Harvey and Newport, 2003). Moreover, defects in transcription termination (Gros et al., 2015; Mischo and Proudfoot, 2013), pervasive transcription (Jensen et al., 2013) and heterogenous mRNA transcription (Pelechano et al., 2013), may lead to collisions between RNAP and origin licensing factors even outside actively transcribed genes.

Our observations reveal that origin licensing intermediates are not frequently disassembled during encounters with transcribing RNAP, but instead are mobilized and readily repositioned (Figure 6.1B). Among the loading intermediates we evaluated, MCM DHs exhibited an extreme robustness. No properly loaded MCM DHs were ejected by RNAP in hundreds of observations individually or where multiple MCMs collided with one another (chapter 4.2). This extreme stability explains the previously observed replication competence of MCM DHs laterally displaced by RNAP (Gros et al., 2015), and is entirely consistent with the observation of trains of MCMs visualized by EM (Douglas et al., 2018), evidence suggesting MCMs can be pushed by advancing replication forks (Sedlackova et al., 2020), and the observation that an additional helicase is required to offload MCMs (Hill et al., 2020; Schauer et al., 2020). MCM stability appears to be an adaptation to ensure origin licensing is not disrupted on crowded chromosomes prior to S phase (Kumar and Remus, 2016).

Structural studies have demonstrated that MCM DHs extensively engage both strands of the DNA duplex inside the axial channel through multiple subunit contacts (Abid Ali et al., 2017; Noguchi et al., 2017). Moreover, these studies revealed that the DNA is slightly bent inside the channel due to the offset and slight tilted stacking of the hexamers. Our observations of spontaneous MCM DH sliding (Figures 3.5 and 3.6) and unimpeded transcription rates even after RNAP collision with multiple MCM DHs (Figure 4.7) are therefore surprising. On the other hand, this type of DNA engagement likely explains our observation that spontaneous MCM DH sliding is a rare event at physiological salt

conditions and may underlie our observation that MCMs aid in global genome organization by functioning as roadblocks and active barriers to cohesin-mediated loop extrusion (chapter 5). Thus, collisions between different protein machineries and MCM DHs are not equivalent and are likely influenced by the force exerted on the MCM DH by a colliding DNA translocase. It is noteworthy that orthogonal pathways such as the block to cohesin translocation may provide additional selective pressure promoting the extreme stability of MCMs once loaded.

Our collision experiments unexpectedly showed that several loading intermediates formed prior to completion of MCM DH formation can be mobilized and repositioned (chapter 4.4). In particular, the stalled ATPγS-bound intermediate OCCM, comprised of ORC, Cdc6, Cdt1 and a single MCM hexamer in a cracked-ring conformation (Yuan et al., 2017), was also readily mobilized and easily repositioned despite its anticipated lower stability (Figures 4.12 and 4.13). This observation allows for the possibility that completion of loading could occur at sites distant from the origin (Figure 6.1C). Among the intermediates evaluated in this study, ORC exhibited the lowest stability, and was most frequently ejected (Figure 4.15). This is consistent with ORC exhibiting an open ring structure on DNA, while binding of Cdc6 to ORC in the OCCM results in topological enclosure of the DNA by ORC/Cdc6 (Li et al., 2018; Yuan et al., 2017). Nevertheless, to our surprise, many encounters with RNAPs did promote ORC mobilization and repositioning (Figure 4.15).

Diffusive sliding has emerged as a vital and intrinsic feature of numerous factors that operate on chromosomes (Chen et al., 2014; Duzdevich et al., 2015; Mirny et al., 2009; Stracy et al., 2021). One explanation for this common property is the benefit conferred when searching the chromosome. Modeling has demonstrated a combination of global 3D diffusion and local 1D sliding reduces the time required to locate specific sites by an order of magnitude (Mirny et al., 2009). Our findings support the notion that sliding is equally important once sites have been located and downstream pathways have been activated. Under these circumstances, sliding serves an important additional function providing an intrinsic resistance to disruption which allows pathways to proceed within the same local vicinity. Our choice of T7 RNA polymerase as an orthogonal machinery to challenge origin licensing removes the possibility of any specific contacts playing an important role while exerting forces comparable to eukaryotic RNA polymerase II (Galburt et al., 2007; Thomen et al., 2008). Therefore, our observations provide a broad framework for predicting the dynamic events and outcomes during conflicts with diverse families of polymerases and translocases beyond origin licensing.

6.2 ORC dynamics reveal alternative resistance and origin specification pathways

The dynamics of ORC are unique among origin licensing factors. In sharp contrast to complexes containing the MCMs, ORC alone exhibits lower stability at locations outside ARS1 (Figure 6.1C), consistent with its primary role in origin identification and MCM recruitment, and in line with previous studies (Duzdevich et al., 2015). Displacement of ORC by RNAP frequently results in ejection, as one might expect for a factor adapted to rapidly sample local regions on large chromosomes interspersed with 3D diffusion. In those cases where ORC remained bound to random sequences, rapid sliding and relocalization at the origin was observed, driven by buffer flow, or in rare cases pushing by RNAP (Figure 4.18). Strikingly, on frequent occasions, RNAP bypassed ARS1-bound ORC, which was never observed for intermediates containing MCMs (Figures 4.15D and 4.16). Structural characterization of DNA-bound ORC has revealed extensive sequence specific as well as large non-sequence specific ORC-DNA contact surfaces that likely facilitate ORC sliding (Hu et al., 2020; Jaremko et al., 2020; Li et al., 2018; Schmidt and Bleichert, 2020). For RNAP to pass, ORC would have to partially or fully disengage from the origin. Interestingly, several reports have demonstrated that ORC binds single-stranded DNA (Hoshina et al., 2013; Kawakami et al., 2019), which triggers large scale conformational rearrangements (Lee et al., 2000). We therefore speculate that ORC could largely dissociate from the origin while remaining firmly bound to the excluded single-strand of the transcription bubble. After passage of RNAP, ORC could rapidly rebind the duplex origin. We predict that the contacts between ORC and DNA are greatly reduced on random DNA sequences based on our observation of low stability in regions outside the origin. Further studies beyond the scope of the current work are needed to delineate the importance of distinct ORC contact surfaces for each mode of engagement and bypass. Moreover, it will be interesting to test if the ability to engage DNA in multiple modes is conserved in metazoan ORC, which exhibits a significantly reduced DNA binding surface and greatly reduced DNA binding specificity compared to yeast ORC (Jaremko et al., 2020; Lee et al., 2021; Schmidt and Bleichert, 2020).

The lower stability of ORC at regions outside the origin would seem to disfavor loading at alternative sites. However, our observations revealed that ORC can hitch a ride on sliding MCMs or within the OCCM complex (Figures 4.19 and 4.20). When in complex with MCMs, ORC remains more stably associated at non-origin sites (Figure 4.18D). Therefore, repositioning through association with the MCMs may allow ORC to perform further cycles of helicase loading at sites distant from origins. In higher eukaryotes, origins contain few sequence specific elements, and instead the local chromatin environment is a defining feature. Chromatin likely confers additional stability to ORC (De Ioannes et al., 2019; Kuo et al., 2012; Shen et al., 2010), allowing it to remain stably associated at alternative sites where further cycles of helicase loading could be conducted, as supported by recent single-molecule studies (Li et al., 2021).

6.3 Origin selection in the chromatin landscape

Our observation that nucleosomes are frequently pushed during encounters with MCMs demonstrates that origin licensing factors can reshape the chromatin landscape (Figure 6.1B). These observations suggest that the origin could be reorganized with nucleosomes displaced or ejected during MCM DH assembly. Consistent with this idea, separate loading of the two MCM helicases, followed by subsequent sliding and formation of the MCM DH has been proposed as a mechanism for MCM DH formation (Coster and Diffley, 2017). This mechanism would seem to demand reorganization of local chromatin structure and ejection of intervening nucleosomes similar to the ejection pathway visualized in this study. To our surprise, RNAP can push MCMs together with multiple nucleosomes (Figures 4.11 and 4.14), however, the likelihood of stalling increases with each nucleosome collision (Figure 4.11E). The presence of chromatin chaperones and remodelers, which were excluded in this study, may further mobilize nucleosomes encountered by MCMs, facilitating large scale reorganization of chromatin during displacement by RNAP.

MCM sliding could have dramatic consequences for many pathways by reorganizing chromatin. The local chromatin architecture at the origin could be displaced in front of MCMs, leading to complete transplantation of the origin to a distant site. In higher eukaryotes, where sequence only plays a minor role in defining origins, this transplantation could help to promote ORC recruitment and further rounds of MCM loading at new sites (Li et al., 2021). Moreover, if key chromatin hallmarks of relocated origins could be propagated to new generations, this may allow for the rapid birth of new origins. MCM sliding could also have tragic consequences, potentially leading to complete loss or reprogramming of critical gene expression patterns. The presence of MCM ahead of RNAP is likely to disrupt histone inheritance pathways by blocking engagement by the histone processing machinery that travels with the polymerase. Further studies are required to explore these potential outcomes.

Based on the structure of MCM DH, we would predict that the nucleosome encounters we have directly visualized occur on the C-terminal face of MCM. In the context of replication, MCM adopts the opposite orientation in the CMG helicase with the N-terminus facing parental DNA and nucleosomes (Li and O'Donnell, 2018). As a consequence, the exposed regions that could engage histones will be different. In particular, the H3-H4 binding site residing in the N-terminus of Mcm2 is distant from nucleosomes in the MCM DH (Huang et al., 2015). Therefore, the outcomes we observed likely do not rely on specific histone contacts, but instead are the result of non-specific physical encounters. Future studies contrasting our observations with the outcomes of CMG-nucleosome encounters could help clarify the roles of specific histone interacting regions in shaping histone processing pathways at the replication fork.

We have demonstrated that the mobilization of origin licensing factors provides resistance pathways to overcome challenges from orthogonal processes on the

chromosome. However, this mobilization also provides further avenues for specialized regulation. Interestingly, recent work has uncovered a mechanistic link between transcriptional silencing and late replication wherein histone deacetylation can trigger transcription-mediated displacement of MCMs in rDNA repeats to modulate rDNA origin efficiency (Foss et al., 2019). Pathways like these may be part of larger programs in which mobilization of origin licensing factors is leveraged to further regulate or augment downstream pathways. Therefore, we anticipate that the dynamics we see are likely to have significant implications for numerous essential pathways beyond replication.

6.4 Implications of the MCM paradox for 3D genome organization

By investigating collisions with RNAPs, we demonstrated a high degree of dynamic flexibility of origin licensing intermediates. However, collisions with other DNA translocases are not necessarily equivalent, for instance due to different forces being exerted. Consistent with our finding that multiple MCM DHs sequester and spread around origins (chapter 3.2), MCM DHs are found in large excess over replication initiation events on chromosomes, which is known as the “MCM paradox” (Burkhart et al., 1995; Edwards et al., 2002; Lei et al., 1996; Mahbubani et al., 1997). Thus, while folding the genome via loop extrusion, cohesin complexes also encounter origin licensing intermediates, most prominently MCM DHs. Whether MCM DHs could affect cohesin-mediated loop extrusion and thus the overall genome architecture by acting as barriers in addition to CTCF sites was not known.

Therefore, in the second part of this thesis, we investigated the influence of MCM DHs on loop extrusion using multiple complementary approaches. In collaboration and using the OZT, we revealed that MCM loss from chromatin in G1 phase zygotes (Figure 5.1) leads to an increased strength of average loops and TADs (Figure 5.2) and that this is entirely dependent on cohesin function (Figure 5.3). Moreover, we developed a single-molecule MCM roadblock assay for passive cohesin translocation which demonstrated that MCM DHs represent in fact direct barriers to cohesin (Figures 5.4, 5.5 and 5.6). Nevertheless, we frequently saw cohesin bypassing MCM DHs, suggesting that the barrier function of MCM underlies stochastic principles. A limitation of our single-molecule approach is that we do not know if passive translocation and loop extrusion are connected. Given the dynamic repositioning of MCM DHs upon conflicts with RNAPs, it is conceivable that active loop extruding cohesin would act in a similar fashion. However, while RNAP exerts forces of ~17 pN (Galburt et al., 2007; Thomen et al., 2008), loop extrusion by human cohesin already stalls at forces of ~1 pN (Kim et al., 2019). Thus, cohesin is less likely to reposition MCM DHs and is not expected to displace MCM DHs over long distances, especially in the presence of nucleosomes.

A recent study indicates that loop-extruding yeast condensin can bypass nucleosomes, *E. coli* RNAP and dCas9, even when tethered to 200 nanometers large nanoparticles, with high efficiency (Pradhan et al., 2021). However, the same study found that human cohesin was already blocked by 30 nanometers large nanoparticles in ~50 % of the cases. The authors suggest a non-topological model for loop extrusion which seems likely but requires external binding of DNA, both at the heads and hinge region of SMC subunits (Bisht et al., 2013; Chiu et al., 2004; Shi et al., 2020). In contrast to the barrier types used in that study, with the nanoparticles being rather artificial, they are not topologically engaged with dsDNA and thus still allow contacts between SMC complexes and DNA at barrier regions. In contrast, MCM DH represents a physiological barrier which is topologically bound to dsDNA and thus entirely shields at least 60 bps of DNA in its central channel (Li et al., 2015; Noguchi et al., 2017). Therefore, we speculate that

cohesin-mediated loop extrusion, which relies on continuous DNA contacts, will be affected by MCM DHs as predicted by our observations based on passive cohesin translocation *in vitro*.

We considered different, not mutually exclusive mechanisms how MCM DHs could perform their boundary function. Some possibilities, which we could not investigate with our assay, are that MCM DHs could shield essential DNA binding sites in cohesin, prevent conformational changes in cohesin required for loop extrusion or inhibit cohesin's ATPase activity, e.g. by mediating exchange of NIPBL with PDS5 proteins. Future studies utilizing an active single-molecule loop extrusion assay on licensed DNA molecules and FRET-based approaches will be required to investigate these possibilities. Moreover, MCM could be a passive, physical barrier, simply functioning as a "roadblock". Our observation that MCM DHs efficiently impede cohesin translocation, while multiple MCMs did not from a larger barrier (Figures 5.4F, 5.5D, 5.6G and 5.7A) might support the passive roadblock mechanism which is also consistent with previous observations with RNAP and nucleosomes (Davidson et al., 2016; Stigler et al., 2016). However, we cannot exclude that specific motifs of additional MCMs might be buried by flanking MCMs, thus not increasing the overall boundary efficiency.

Alternatively, MCM could also be an "active boundary" by containing one or multiple binding sites for cohesin, thus restricting loop extrusion by directly tethering cohesin and/or protecting from release by WAPL. For instance, the canonical barrier CTCF was shown to exert its boundary function asymmetrically where an N-terminal motif binding to cohesin is concomitant with its boundary function (Li et al., 2020; Nishana et al., 2020; Nora et al., 2020; Pugacheva et al., 2020). This N-terminal YDF-containing motif in CTCF was also identified in human but not in yeast MCM. Since we used a mixed system of human cohesin and yeast MCM and to further investigate the effect of this YDF motif in MCM, we generated a "humanized" version of yeast MCM containing this motif in the C-terminal region of Mcm3 (Figure 5.7B). Strikingly, we found that cohesin translocation frequently pauses at MCM-bound origins (Figures 5.7D, 5.7E, 5.7F, 5.7G and 5.7H) and that the likelihood of pausing is highly increased in the presence of the YDF motif in Mcm3 (Figure 5.8B). Combined with our data showing increased barrier strength and likelihood of pausing with increasing number of MCMs containing the YDF motif (Figures 5.7C and 5.8D), we suggest that MCM complexes exert their barrier function by two different mechanisms. Besides being passive, physical barriers, MCM complexes are also active boundaries to loop extrusion, by tethering cohesin through specific binding sites such as the YDF motif in a similar manner as CTCF.

Notably, we used cohesin^{STAG1} in this study, whereas both SCC1 and STAG2 subunits were shown to be required for stable interaction with CTCF and [Y/F]xF motifs in pulldown assays and peptide screens (Li et al., 2020). Due to the high temporal and spatial sensitivity of our single-molecule assay we were able to detect short-lived interactions mediated by SCC1 alone and would expect stronger, more stable interaction with cohesin^{STAG2}. Intriguingly, we did not observe a difference in pause durations in the

presence or absence of the YDF motif (Figure 5.7I). This could be explained by an increased on-rate but an unchanged off-rate or by short-lived transient binding of cohesin which leads to pausing exceeding the cohesin-MCM interaction time. In the latter case, the detected duration would then rather be dependent on spontaneous restart of cohesin translocation, which is equally likely at identical ionic strength. However, and in line with our observations, in the presence of the YDF motif a subsequent pause would occur more likely due to higher chance of transient interaction (Figure 5.8B). It remains to be elucidated whether MCMs contain additional, not yet identified interaction sites for cohesin. Additionally, it will be important to investigate the boundary function of CMG in which the YDF motif in the C-terminal region of Mcm3 does not face the replication fork (Douglas et al., 2018; Georgescu et al., 2017; Lee et al., 2014).

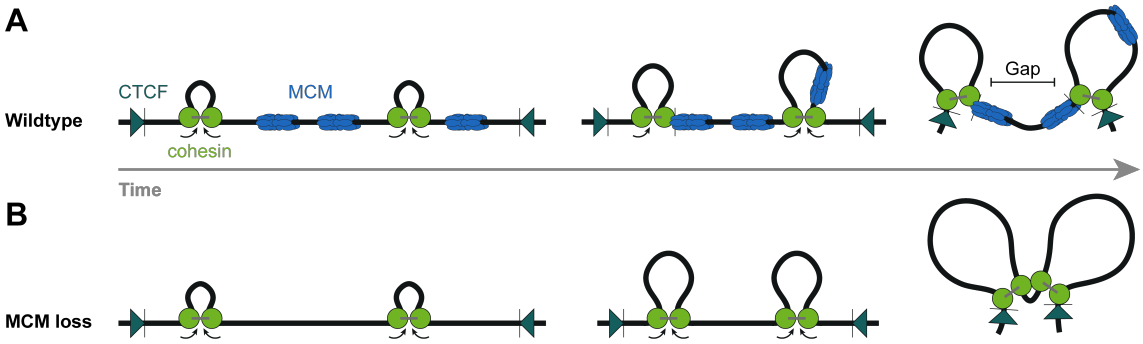

Figure 6.2. MCM complexes are barriers that restrict cohesin-mediated loop extrusion
Cohesin complexes organize the chromosome via loop extrusion. Upon encountering convergently-oriented CTCF sites, loop extrusion can be blocked. (**A**) In wildtype G1 phase cells, MCM DHs can also act as semi-permeable barriers to cohesin-mediated loop extrusion on DNA. Cohesin may bypass a subset of MCM DHs (and also CTCF sites). Whether a barrier is bypassed underlies stochastic processes and may vary in time. (**B**) MCM loss from chromatin removes one type of barrier, allowing cohesin-mediated loop extrusion to proceed unimpeded until cohesin reaches CTCF sites.

In summary, in addition to known CTCF sites, we have identified MCM DHs as novel barrier to cohesin-mediated loop extrusion by *in vivo* and *in vitro* experiments. In unperturbed, wildtype cells, MCMs can act as semi-permeable passive, physical barriers and active boundaries which still permit cohesin to bypass in a stochastic manner (Figure 6.2A). Upon MCM loss from chromatin, this type of barrier is lost, allowing cohesin-mediated loop extrusion to proceed unimpeded until reaching convergently-oriented CTCF sites (Figure 6.2B). Our data demonstrate that the MCM paradox has important implications for chromatin organization and thus for regulatory roles in development, replication timing, gene regulation or V(D)J recombination (Flyamer et al., 2017; Kosak et al., 2002; Lupianez et al., 2015; Pope et al., 2014). Therefore, the MCM paradox might also be relevant in human pathologies caused by mutations affecting the origin licensing pathway such as the Meier-Gorlin syndrome (de Munnik et al., 2015). Moreover, changes in MCM distribution on chromatin, as found in different cell types but also caused by altered transcription patterns or origin usage such as in hyperproliferative cancer cells, has far ranging consequences for the overall 3D genome architecture, essential cellular functions and in the end faithful genome maintenance and replication.

Appendix

A1 List of Publications

Scherr, M.J., Abd Wahab, S., Remus, D., and Duderstadt, K.E. (2021). Born to Slide: Mobile Origin Licensing Factors Confer Resistance to Transcription Conflicts. *Preprint at SSRN*. doi:10.2139/ssrn.3775178. (*in revision*)

Bauernfried, S., **Scherr, M.J.**, Pichlmair, A., Duderstadt, K.E., and Hornung, V. (2021). Human NLRP1 is a sensor for double-stranded RNA. *Science* **371**. doi:10.1126/science.abd0811.

Safaric, B., Chacin, E., **Scherr, M.J.**, Rajappa, L., Gebhardt, C., Kurat, C.F., Cordes, T., and Duderstadt, K.E. (2021). The fork protection complex recruits FACT to reorganize nucleosomes during replication. (*in revision*)

Dequeker, B.J.H., Brandão, H.B.*, **Scherr, M.J.***, Gassler, J., Powell, S., Gaspar, I., Flyamer, I.M., Tang, W., Stocsits, R., Davidson, I.F., Peters, J.M., Duderstadt, K.E., Mirny, L.A., and Tachibana, K. (2020). MCM complexes are barriers that restrict cohesin-mediated loop extrusion. *Preprint at bioRxiv*. doi:10.1101/2020.10.15.340356. (*in revision*)

Scherr, M.J.*, Safaric, B.*, and Duderstadt, K.E. (2018). Noise in the Machine: Alternative Pathway Sampling is the Rule During DNA Replication. *Bioessays* **40**. doi:10.1002/bies.201700159.

Eckl, J.M.*, **Scherr, M.J.***, Freiburger, L., Daake, M.A., Sattler, M., and Richter, K. (2015). Hsp90.Cdc37 Complexes with Protein Kinases Form Cooperatively with Multiple Distinct Interaction Sites. *J Biol Chem* **290**, 30843-30854. doi:10.1074/jbc.M115.693150.

*equal contribution

A2 Movie captions

All movies can be accessed at https://osf.io/z3trb/ (doi:10.17605/osf.io/z3trb)

Movie 1. Most MCM DHs are stably bound to the origin ARS1, Related to Figure 3.5A
Loaded MCM DHs (blue) were challenged and imaged in the presence of 0.5 M NaCl.

Movie 2. MCM DHs can switch to a diffusive DNA binding mode, Related to Figure 3.5B
Loaded MCM DHs (blue) were challenged and imaged in the presence of 0.5 M NaCl.

Movie 3. RNA polymerase transcribes highly processively, Related to Figure 4.2A
Upon UTP arrival, RNAP (amber) resumed transcription in the presence of ATP (left) or ATPγS (right).

Movie 4. RNAP can push a single MCM DH over long distances, Related to Figure 4.5E
Encounters with a single RNAP (amber) led to robust repositioning of a single MCM DH (blue).

Movie 5. RNAP can robustly reposition multiple MCM DHs, Related to Figure 4.7A
Multiple encounters of RNAP (amber) with individual MCM DHs (blue) did not alter transcription.

Movie 6. RNAP can transcribe through individual nucleosomes, Related to Figure 4.10A
Encounters between transcribing RNAP (amber) and individual nucleosomes (pink) did not stall transcription, instead led to repositioning of nucleosomes.

Movie 7. RNAP can reposition MCM DHs together with nucleosomes, Related to Figure 4.11A
Three-way encounters between transcribing RNAP (amber), MCM DH (blue) and nucleosome (pink) led to formation of a mobile super complex.

Movie 8. OCCM stays intact during encounters with RNAP and displacement, Related to Figure 4.13A
The OCCM intermediate stayed intact when being pushed by RNAP (amber) as judged by the presence of ORC (green) and Cdt1-MCM (blue).

Movie 9. RNAP can push ORC to distant sites, Related to Figure 4.15B
Encounters between transcribing RNAP (amber) led to robust repositioning of ORC (green).

Movie 10. ORC can be bypassed by RNAP, Related to Figure 4.16A
Transcribing RNAP (amber) frequently bypassed ARS1-bound ORC (green).

Movie 11. ORC can reside at new locations after transcription termination, Related to Figure 4.18A
Displaced ORC (green) could remain stable at the T7T site upon RNAP (amber) transcription termination.

Movie 12. Displaced ORC can undergo 1D diffusion and rebind to origins, Related to Figure 4.18C
Displaced ORC (green) could slide back and rebind to ARS1 after a 1D diffusion in search of the origin upon RNAP (amber) transcription termination.

Movie 13. ORC can hitch a ride on sliding MCM DHs to new locations, Related to Figure 4.19A
Displaced MCM DH (blue) and associated ORC (green) remained stable at the T7T site upon RNAP (amber) transcription termination.

Movie 14. OCCM remains stable after displacement, Related to Figure 4.20A
Displaced OCCM remained stable at the T7 terminator site, as judged by the presence of ORC (green) and Cdt1-MCM (blue), upon RNAP (amber) transcription termination.

Movie 15. OCCM can diffuse along DNA, Related to Figure 4.20B
Displaced OCCM could switch to a diffusive mode while staying intact, as judged by the presence of ORC (green) and Cdt1-MCM (blue) signal, at the T7T site upon transcription termination.

Movie 16. Cohesin translocates rapidly along DNA at high salt, Related to Figure 5.4B
Cohesin (green) translocated rapidly on doubly-tethered DNA (gray) at high salt (0.5 M NaCl) in the absence of buffer flow and MCM DH (blue).

Movie 17. MCM DH constrains cohesin translocation on DNA at high salt, Related to Figure 5.4C
Origin-bound MCM DH (blue) was an efficient barrier as it fully constrained cohesin (green) translocation on doubly-tethered DNA (gray) in the absence of buffer flow at high salt (0.5 M NaCl).

Movie 18. Cohesin can occasionally bypass MCM DH on DNA at high salt, Related to Figure 5.5A
Origin-bound MCM DH (blue) was occasionally bypassed by translocating cohesin (green) on doubly-tethered DNA (gray) in the absence of buffer flow at high salt (0.5 M NaCl).

Movie 19. Cohesin translocates rapidly along DNA at low salt, Related to Figure 5.6B
Cohesin (green) translocated rapidly on doubly-tethered DNA (gray) at low salt (0.15 M NaCl) in the absence of buffer flow and MCM DH (blue).

Movie 20. MCM DH constrains cohesin translocation on DNA at low salt, Related to Figure 5.6E
Origin-bound MCM DH (blue) was an efficient barrier as it fully constrained cohesin (green) translocation on doubly-tethered DNA (gray) in the absence of buffer flow at low salt (0.15 M NaCl).

Movie 21. Cohesin can occasionally bypass MCM DH on DNA at low salt, Related to Figure 5.6F
Origin-bound MCM DH (blue) was occasionally bypassed by translocating cohesin (green) on doubly-tethered DNA (gray) in the absence of buffer flow at low salt (0.15 M NaCl).

Movie 22. Cohesin translocation frequently pauses at MCM-bound origins, Related to Figure 5.7D
Translocating cohesin (green) frequently pauses at origins bound by $MCM^{Mcm3\text{-}YDF}$ DH (blue) on doubly-tethered DNA (gray) in the absence of buffer flow at low salt (0.15 M NaCl)

A3 List of Figures

A4 List of Tables

A5 List of Abbreviations

ARS ..... *autonomously replicating sequence*
ATP ..... *adenosine triphosphate*
ATPγS ..... *adenosine-5'-o-(3-thio-triphosphate)*
bps ..... *base pairs*
BV ..... *bed volumes*
CI ..... *confidence interval*
CMG *Cdc45-Mcm2-7-GINS*
CoSMoS ..... *co-localization single-molecule spectroscopy*
CTCF ..... *CCCTC binding factor*
C-tier ..... *C-terminal tier*
CTP ..... *cytidine triphosphate*
CV ..... *column volumes*
DDK ..... *Dbf4-dependent kinase*
DH ..... *double-hexamer*
DM ..... *dichroic mirror*
DNA ..... *deoxyribonucleic acid*
dsDNA ..... *double-stranded DNA*
EM ..... *electron microscopy*
FEN1 ..... *Flap endonuclease 1*
FRET ..... *fluorescence resonance energy transfer*
GTP ..... *guanosine triphosphate*
HAWK ..... *HEAT repeat proteins associated with kleisins*
kb ..... *kilobases*
KCP ..... *kinetic change point*
KDE ..... *kernel density estimation*
KGlu ..... *potassium glutamate*
KOAc ..... *potassium acetate*
LB ..... *Luria-Bertani*
Mcm ..... *minichromosome maintenance*
MCM *Mcm2-7*
mmTIRF ..... *micromirror TIRF*
MO ..... *microscope objective*
MSD ..... *mean squared displacement*
MSE ..... *modified steric exclusion*
n ..... *number of observations*
N-tier ..... *N-terminal tier*
NTP ..... *nucleoside triphosphate*
OCCM *ORC-Cdc6-Cdt1-MCM*
OF ..... *Okazaki fragment*
ORC ..... *origin recognition complex*
oriC ..... *origin of replication in E. coli*
OSS ..... *oxygen scavenging system*
OZT ..... *oocyte-to-zygote transition*
PCA ..... *3,4-Dihydroxybenzoic acid*
PCD ..... *Protocatechuate 3,4-Dioxygenase*
PCNA ..... *proliferating cell nuclear antigen*

PCR ..... *polymerase chain reaction*
pN ..... *pico-Newton*
pre-RC ..... *pre-replicative complex*
RFC ..... *replication factor C*
RNA ..... *ribonucleic acid*
RNAP ..... *RNA polymerase*
RT ..... *room temperature*
S-CDK ..... *S phase cyclin-dependent kinase*
SCE ..... *side channel extrusion*
SD ..... *standard deviation*
SDS-PAGE ..... sodium dodecyl sulphate–polyacrylamide gel electrophoresis
SE ..... *steric exclusion*
SEM ..... *standard error of the mean*
SMC ..... *structural maintenance of chromosomes*
snHi-C ..... *single-nucleus Hi-C*
SOC ..... *Super Optimal broth with Catabolite repression*
ssDNA ..... *single-stranded DNA*
T7P ..... *T7 promoter*
T7T ..... *T7 terminator*
TAD ..... *topologically associating domain*
TIRF ..... *total internal reflection fluorescence*
TIRFM ..... *total internal reflectance fluorescence microscopy*
TLS ..... *translesion synthesis*
Tus ..... *terminus site-binding protein*
U ..... *units*
UTP ..... *uridine triphosphate*
v/v ..... *volume per volume*
w/v ..... *weight per volume*
WAPL ..... *Wings apart-like*
YP ..... *Yeast extract Peptone*
λ DNA ..... *DNA from bacteriophage λ*

Acknowledgments

When I started my PhD five years ago, I did not expect it to be such a challenging but also exciting journey. During this journey, my project but also I personally went through quite some ups and downs which makes me very proud to now have successfully completed my PhD thesis. This would not have been possible without all the people who joined me during my PhD journey and this extraordinary time, a big thanks to all of you!

First, I would like to thank Prof. Dr. Karl Duderstadt in particular. Thank you that you gave me the opportunity to conduct my thesis in your research group. It was very exciting to join as a “first generation” PhD student, set up the lab and it was inspiring to develop and work on the first projects which set the foundation and direction of your research group. Nevertheless, I always had the freedom to develop, explore and pursue my own ideas but you also had an open hear for any problems I encountered on my journey. The time in your lab made me the scientist that I am now at the end of my PhD, thank you a lot for your full trust during these years.

I would also like to thank all colleagues from the group and the Max Planck Institute of Biochemistry I interacted with in the last years. You created such a nice and vibrant working atmosphere in which I enjoyed working a lot. I especially would like to thank you, Barbara, for being the best colleague and friend from the first day I joined the lab and always being there for me. Sebastian, thank you for always helping and sharing your experience with me, it was great meeting you and becoming friends over the years. Geri, I am very glad that you joined the lab on floor below me, thank you for the awesome time inside and outside of the lab.

A special thanks goes to the International Max Planck Research School for Molecular Life Sciences for offering an excellent program with lots of workshops and networking events. Thank you Dr. Hans-Jörg Schäffer, Dr. Ingrid Wolf and Maximiliane Reif for your time and continuous effort to set up and maintain such an excellent graduate school. In this regard I also thank my TAC members Prof. Dr. Michael Sattler, Dr. Boris Pfander and Dr. Dirk Remus for continuous support and scientific input during the years.

Moreover, I would also like to thank all collaboration partners. I especially would like to thank Bart Dequeker, Dr. Kikuë Tachibana and Dr. Jan-Michael Peters for the great opportunity to join an exciting collaboration. In this regard, thank you Iain, Ulla and especially Renate for material and excellent technical assistance for this project.

Thanks to my family, especially my parents, and friends. You always believed in me and supported me on my way and I could not have achieved this without you.

Last but not least, thank you Stefan for being the best collaboration partner in science and most importantly in life. I am very happy to have met you and share my life with you.

References

Aberg, C., Duderstadt, K.E., and van Oijen, A.M. (2016). Stability versus exchange: a paradox in DNA replication. Nucleic Acids Res *44*, 4846-4854. doi:10.1093/nar/gkw296.

Abid Ali, F., Douglas, M.E., Locke, J., Pye, V.E., Nans, A., Diffley, J.F.X., and Costa, A. (2017). Cryo-EM structure of a licensed DNA replication origin. Nat Commun *8*, 2241. doi:10.1038/s41467-017-02389-0.

Abid Ali, F., Renault, L., Gannon, J., Gahlon, H.L., Kotecha, A., Zhou, J.C., Rueda, D., and Costa, A. (2016). Cryo-EM structures of the eukaryotic replicative helicase bound to a translocation substrate. Nat Commun *7*, 10708. doi:10.1038/ncomms10708.

Agarwal, R., and Duderstadt, K.E. (2020). Multiplex flow magnetic tweezers reveal rare enzymatic events with single molecule precision. Nat Commun *11*, 4714. doi:10.1038/s41467-020-18456-y.

Aitken, C.E., Marshall, R.A., and Puglisi, J.D. (2008). An oxygen scavenging system for improvement of dye stability in single-molecule fluorescence experiments. Biophys J *94*, 1826-1835. doi:10.1529/biophysj.107.117689.

Aladjem, M.I., Almouzni, G., Araki, H., Aravind, L., Bartek, J., Bastia, D., Bell, S.D., Bell, S.P., Bestor, T.H., Blow, J.J.*, et al.* (2006). DNA Replication and Human Disease (Cold Spring Harbor Laboratory Press). ISBN:978-087969766-2.

Alberts, B.M., Barry, J., Bedinger, P., Formosa, T., Jongeneel, C.V., and Kreuzer, K.N. (1983). Studies on DNA replication in the bacteriophage T4 in vitro system. Cold Spring Harb Symp Quant Biol *47 Pt 2*, 655-668. doi:10.1101/sqb.1983.047.01.077.

Altmann, R. (1889). Ueber Nucleinsäuren. In Archiv für Anatomie und Physiologie, Physiologische Abteilung (Leipzig), pp. 524-536.

Amin, A., Wu, R., Cheung, M.H., Scott, J.F., Wang, Z., Zhou, Z., Liu, C., Zhu, G., Wong, C.K., Yu, Z.*, et al.* (2020). An Essential and Cell-Cycle-Dependent ORC Dimerization Cycle Regulates Eukaryotic Chromosomal DNA Replication. Cell Rep *30*, 3323-3338 e3326. doi:10.1016/j.celrep.2020.02.046.

Aria, V., and Yeeles, J.T.P. (2019). Mechanism of Bidirectional Leading-Strand Synthesis Establishment at Eukaryotic DNA Replication Origins. Mol Cell *73*, 199-211 e110. doi:10.1016/j.molcel.2018.10.019.

Arias-Palomo, E., Puri, N., O'Shea Murray, V.L., Yan, Q., and Berger, J.M. (2019). Physical Basis for the Loading of a Bacterial Replicative Helicase onto DNA. Mol Cell *74*, 173-184 e174. doi:10.1016/j.molcel.2019.01.023.

Austin, R.J., Orr-Weaver, T.L., and Bell, S.P. (1999). Drosophila ORC specifically binds to ACE3, an origin of DNA replication control element. Genes Dev *13*, 2639-2649. doi:10.1101/gad.13.20.2639.

Avery, O.T., Macleod, C.M., and McCarty, M. (1944). Studies on the Chemical Nature of the Substance Inducing Transformation of Pneumococcal Types : Induction of Transformation by a Desoxyribonucleic Acid Fraction Isolated from Pneumococcus Type Iii. J Exp Med *79*, 137-158. doi:10.1084/jem.79.2.137.

Azmi, I.F., Watanabe, S., Maloney, M.F., Kang, S., Belsky, J.A., MacAlpine, D.M., Peterson, C.L., and Bell, S.P. (2017). Nucleosomes influence multiple steps during replication initiation. Elife *6*. doi:10.7554/eLife.22512.

Baretic, D., Jenkyn-Bedford, M., Aria, V., Cannone, G., Skehel, M., and Yeeles, J.T.P. (2020). Cryo-EM Structure of the Fork Protection Complex Bound to CMG at a Replication Fork. Mol Cell *78*, 926-940 e913. doi:10.1016/j.molcel.2020.04.012.

Bastié, N., Chapard, C., Dauban, L., Gadal, O., Beckouët, F., and Koszul, R. (2021). Smc3 acetylation, Pds5 and Scc2 control the translocase activity that establishes cohesin dependent chromatin loops. bioRxiv. doi:10.1101/2021.04.21.440823.

Beattie, T.R., Kapadia, N., Nicolas, E., Uphoff, S., Wollman, A.J., Leake, M.C., and Reyes-Lamothe, R. (2017). Frequent exchange of the DNA polymerase during bacterial chromosome replication. Elife *6*, e21763. doi:10.7554/eLife.21763.

Bell, S.P. (2017). Rethinking origin licensing. Elife *6*. doi:10.7554/eLife.24052.

Berg, O.G., Winter, R.B., and von Hippel, P.H. (1981). Diffusion-driven mechanisms of protein translocation on nucleic acids. 1. Models and theory. Biochemistry *20*, 6929-6948. doi:10.1021/bi00527a028.

Berghuis, B.A., Kober, M., van Laar, T., and Dekker, N.H. (2016). High-throughput, high-force probing of DNA-protein interactions with magnetic tweezers. Methods *105*, 90-98. doi:10.1016/j.ymeth.2016.03.025.

Bessman, M.J., Kornberg, A., Lehman, I.R., and Simms, E.S. (1956). Enzymic synthesis of deoxyribonucleic acid. Biochim Biophys Acta *21*, 197-198. doi:10.1016/0006-3002(56)90127-5.

Bianconi, E., Piovesan, A., Facchin, F., Beraudi, A., Casadei, R., Frabetti, F., Vitale, L., Pelleri, M.C., Tassani, S., Piva, F.*, et al.* (2013). An estimation of the number of cells in the human body. Ann Hum Biol *40*, 463-471. doi:10.3109/03014460.2013.807878.

Bienko, M., Green, C.M., Crosetto, N., Rudolf, F., Zapart, G., Coull, B., Kannouche, P., Wider, G., Peter, M., Lehmann, A.R.*, et al.* (2005). Ubiquitin-binding domains in Y-family polymerases regulate translesion synthesis. Science *310*, 1821-1824. doi:10.1126/science.1120615.

Birnboim, H.C., and Doly, J. (1979). A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucleic Acids Res *7*, 1513-1523. doi:10.1093/nar/7.6.1513.

Bisht, K.K., Daniloski, Z., and Smith, S. (2013). SA1 binds directly to DNA through its unique AT-hook to promote sister chromatid cohesion at telomeres. J Cell Sci *126*, 3493-3503. doi:10.1242/jcs.130872.

Blattner, F.R., Plunkett, G., 3rd, Bloch, C.A., Perna, N.T., Burland, V., Riley, M., Collado-Vides, J., Glasner, J.D., Rode, C.K., Mayhew, G.F.*, et al.* (1997). The complete genome sequence of Escherichia coli K-12. Science *277*, 1453-1462. doi:10.1126/science.277.5331.1453.

Bleichert, F., Botchan, M.R., and Berger, J.M. (2017). Mechanisms for initiating cellular DNA replication. Science *355*. doi:10.1126/science.aah6317.

Blow, J.J., and Laskey, R.A. (1988). A role for the nuclear envelope in controlling DNA replication within the cell cycle. Nature *332*, 546-548. doi:10.1038/332546a0.

Bonner, C.A., Hays, S., McEntee, K., and Goodman, M.F. (1990). DNA polymerase II is encoded by the DNA damage-inducible dinA gene of Escherichia coli. Proc Natl Acad Sci U S A *87*, 7663-7667. doi:10.1073/pnas.87.19.7663.

Bowers, J.L., Randell, J.C., Chen, S., and Bell, S.P. (2004). ATP hydrolysis by ORC catalyzes reiterative Mcm2-7 assembly at a defined origin of replication. Mol Cell *16*, 967-978. doi:10.1016/j.molcel.2004.11.038.

Bradford, M.M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal Biochem *72*, 248-254. doi:10.1006/abio.1976.9999.

Bramhill, D., and Kornberg, A. (1988). Duplex opening by dnaA protein at novel sequences in initiation of replication at the origin of the E. coli chromosome. Cell *52*, 743-755. doi:10.1016/0092-8674(88)90412-6.

Branzei, D., and Foiani, M. (2010). Maintaining genome stability at the replication fork. Nat Rev Mol Cell Biol *11*, 208-219. doi:10.1038/nrm2852.

Broach, J.R., Li, Y.Y., Feldman, J., Jayaram, M., Abraham, J., Nasmyth, K.A., and Hicks, J.B. (1983). Localization and sequence analysis of yeast origins of DNA replication. Cold Spring Harb Symp Quant Biol *47 Pt 2*, 1165-1173. doi:10.1101/sqb.1983.047.01.132.

Burgers, P.M.J., Gordenin, D., and Kunkel, T.A. (2016). Who Is Leading the Replication Fork, Pol epsilon or Pol delta? Mol Cell *61*, 492-493. doi:10.1016/j.molcel.2016.01.017.

Burgers, P.M.J., and Kunkel, T.A. (2017). Eukaryotic DNA Replication Fork. Annu Rev Biochem *86*, 417-

438. doi:10.1146/annurev-biochem-061516-044709.

Burkhart, R., Schulte, D., Hu, D., Musahl, C., Gohring, F., and Knippers, R. (1995). Interactions of human nuclear proteins P1Mcm3 and P1Cdc46. Eur J Biochem *228*, 431-438. doi:10.1111/J.1432-1033.1995.0431n.x.

Burnham, D.R., Kose, H.B., Hoyle, R.B., and Yardimci, H. (2019). The mechanism of DNA unwinding by the eukaryotic replicative helicase. Nat Commun *10*, 2159. doi:10.1038/s41467-019-09896-2.

Busslinger, G.A., Stocsits, R.R., van der Lelij, P., Axelsson, E., Tedeschi, A., Galjart, N., and Peters, J.M. (2017). Cohesin is positioned in mammalian genomes by transcription, CTCF and Wapl. Nature *544*, 503-507. doi:10.1038/nature22063.

Camdere, G., Guacci, V., Stricklin, J., and Koshland, D. (2015). The ATPases of cohesin interface with regulators to modulate cohesin-mediated DNA tethering. Elife *4*. doi:10.7554/eLife.11315.

Candelli, T., Gros, J., and Libri, D. (2018). Pervasive transcription fine-tunes replication origin activity. Elife *7*. doi:10.7554/eLife.40802.

Carbone, A., Fioretti, F.M., Fucci, L., Ausio, J., and Piscopo, M. (2012). High efficiency method to obtain supercoiled DNA with a commercial plasmid purification kit. Acta Biochim Pol *59*, 275-278. doi:10.18388/abp.2012_2151.

Carney, S.M., and Trakselis, M.A. (2016). The excluded DNA strand is SEW important for hexameric helicase unwinding. Methods *108*, 79-91. doi:10.1016/j.ymeth.2016.04.008.

Champasa, K., Blank, C., Friedman, L.J., Gelles, J., and Bell, S.P. (2019). A conserved Mcm4 motif is required for Mcm2-7 double-hexamer formation and origin DNA unwinding. Elife *8*. doi:10.7554/eLife.45538.

Chastain, P.D., 2nd, Makhov, A.M., Nossal, N.G., and Griffith, J. (2003). Architecture of the replication complex and DNA loops at the fork generated by the bacteriophage t4 proteins. J Biol Chem *278*, 21276-21285. doi:10.1074/jbc.M301573200.

Chen, I., Dorr, B.M., and Liu, D.R. (2011). A general strategy for the evolution of bond-forming enzymes using yeast display. Proc Natl Acad Sci U S A *108*, 11399-11404. doi:10.1073/pnas.1101046108.

Chen, J., Zhang, Z., Li, L., Chen, B.C., Revyakin, A., Hajj, B., Legant, W., Dahan, M., Lionnet, T., Betzig, E*., et al.* (2014). Single-molecule dynamics of enhanceosome assembly in embryonic stem cells. Cell *156*, 1274-1285. doi:10.1016/j.cell.2014.01.062.

Chilkova, O., Stenlund, P., Isoz, I., Stith, C.M., Grabowski, P., Lundstrom, E.B., Burgers, P.M., and Johansson, E. (2007). The eukaryotic leading and lagging strand DNA polymerases are loaded onto primer-ends via separate mechanisms but have comparable processivity in the presence of PCNA. Nucleic Acids Res *35*, 6588-6597. doi:10.1093/nar/gkm741.

Chiu, A., Revenkova, E., and Jessberger, R. (2004). DNA interaction and dimerization of eukaryotic SMC hinge domains. J Biol Chem *279*, 26233-26242. doi:10.1074/jbc.M402439200.

Cho, W.H., Kang, Y.H., An, Y.Y., Tappin, I., Hurwitz, J., and Lee, J.K. (2013). Human Tim-Tipin complex affects the biochemical properties of the replicative DNA helicase and DNA polymerases. Proc Natl Acad Sci U S A *110*, 2523-2527. doi:10.1073/pnas.1222494110.

Conti, C., Sacca, B., Herrick, J., Lalou, C., Pommier, Y., and Bensimon, A. (2007). Replication fork velocities at adjacent replication origins are coordinately modified during DNA replication in human cells. Mol Biol Cell *18*, 3059-3067. doi:10.1091/mbc.e06-08-0689.

Corn, J.E., Pease, P.J., Hura, G.L., and Berger, J.M. (2005). Crosstalk between primase subunits can act to regulate primer synthesis in trans. Mol Cell *20*, 391-401. doi:10.1016/j.molcel.2005.09.004.

Costa, A., Renault, L., Swuec, P., Petojevic, T., Pesavento, J.J., Ilves, I., MacLellan-Gibson, K., Fleck, R.A., Botchan, M.R., and Berger, J.M. (2014). DNA binding polarity, dimerization, and ATPase ring remodeling in the CMG helicase of the eukaryotic replisome. Elife *3*, e03273. doi:10.7554/eLife.03273.

Coster, G., and Diffley, J.F.X. (2017). Bidirectional eukaryotic DNA replication is established by quasi-symmetrical helicase loading. Science *357*, 314-318. doi:10.1126/science.aan0063.

Dauban, L., Montagne, R., Thierry, A., Lazar-Stefanita, L., Bastie, N., Gadal, O., Cournac, A., Koszul, R., and Beckouet, F. (2020). Regulation of Cohesin-Mediated Chromosome Folding by Eco1 and Other Partners. Mol Cell *77*, 1279-1293 e1274. doi:10.1016/j.molcel.2020.01.019.

Davidson, I.F., Bauer, B., Goetz, D., Tang, W., Wutz, G., and Peters, J.M. (2019). DNA loop extrusion by human cohesin. Science *366*, 1338-1345. doi:10.1126/science.aaz3418.

Davidson, I.F., Goetz, D., Zaczek, M.P., Molodtsov, M.I., Huis In 't Veld, P.J., Weissmann, F., Litos, G., Cisneros, D.A., Ocampo-Hafalla, M., Ladurner, R.*, et al.* (2016). Rapid movement and transcriptional re-localization of human cohesin on DNA. EMBO J *35*, 2671-2685. doi:10.15252/embj.201695402.

Davidson, I.F., and Peters, J.M. (2021). Genome folding through loop extrusion by SMC complexes. Nat Rev Mol Cell Biol. doi:10.1038/s41580-021-00349-7.

De Ioannes, P., Leon, V.A., Kuang, Z., Wang, M., Boeke, J.D., Hochwagen, A., and Armache, K.J. (2019). Structure and function of the Orc1 BAH-nucleosome complex. Nat Commun *10*, 2894. doi:10.1038/s41467-019-10609-y.

De Jesus-Kim, L., Friedman, L.J., Looke, M., Ramsoomair, C.K., Gelles, J., and Bell, S.P. (2021). DDK regulates replication initiation by controlling the multiplicity of Cdc45-GINS binding to Mcm2-7. Elife *10*. doi:10.7554/eLife.65471.

de Munnik, S.A., Hoefsloot, E.H., Roukema, J., Schoots, J., Knoers, N.V., Brunner, H.G., Jackson, A.P., and Bongers, E.M. (2015). Meier-Gorlin syndrome. Orphanet J Rare Dis *10*, 114. doi:10.1186/s13023-015-0322-x.

Debyser, Z., Tabor, S., and Richardson, C.C. (1994). Coordination of leading and lagging strand DNA synthesis at the replication fork of bacteriophage T7. Cell *77*, 157-166. doi:10.1016/0092-8674(94)90243-7.

Dequeker, B.J.H., Brandão, H.B., Scherr, M.J., Gassler, J., Powell, S., Gaspar, I., Flyamer, I.M., Tang, W., Stocsits, R., Davidson, I.F.*, et al.* (2020). MCM complexes are barriers that restrict cohesin-mediated loop extrusion. bioRxiv. doi:10.1101/2020.10.15.340356.

Devbhandari, S., Jiang, J., Kumar, C., Whitehouse, I., and Remus, D. (2017). Chromatin Constrains the Initiation and Elongation of DNA Replication. Mol Cell *65*, 131-141. doi:10.1016/j.molcel.2016.10.035.

Devbhandari, S., and Remus, D. (2020). Rad53 limits CMG helicase uncoupling from DNA synthesis at replication forks. Nat Struct Mol Biol *27*, 461-471. doi:10.1038/s41594-020-0407-7.

Dewar, J.M., and Walter, J.C. (2017). Mechanisms of DNA replication termination. Nat Rev Mol Cell Biol *18*, 507-516. doi:10.1038/nrm.2017.42.

Diffley, J.F., Cocker, J.H., Dowell, S.J., and Rowley, A. (1994). Two steps in the assembly of complexes at yeast replication origins in vivo. Cell *78*, 303-316. doi:10.1016/0092-8674(94)90299-2.

Dixon, J.R., Selvaraj, S., Yue, F., Kim, A., Li, Y., Shen, Y., Hu, M., Liu, J.S., and Ren, B. (2012). Topological domains in mammalian genomes identified by analysis of chromatin interactions. Nature *485*, 376-380. doi:10.1038/nature11082.

Dixon, N.E. (2009). DNA replication: prime-time looping. Nature *462*, 854-855. doi:10.1038/462854a.

Douglas, M.E., Ali, F.A., Costa, A., and Diffley, J.F.X. (2018). The mechanism of eukaryotic CMG helicase activation. Nature *555*, 265-268. doi:10.1038/nature25787.

Drury, L.S., Perkins, G., and Diffley, J.F. (2000). The cyclin-dependent kinase Cdc28p regulates distinct modes of Cdc6p proteolysis during the budding yeast cell cycle. Curr Biol *10*, 231-240. doi:10.1016/s0960-9822(00)00355-9.

Duderstadt, K.E., and Berger, J.M. (2013). A structural framework for replication origin opening by AAA+ initiation factors. Curr Opin Struct Biol *23*, 144-153. doi:10.1016/j.sbi.2012.11.012.

Duderstadt, K.E., Chuang, K., and Berger, J.M. (2011). DNA stretching by bacterial initiators promotes replication origin opening. Nature *478*, 209-213. doi:10.1038/nature10455.

Duderstadt, K.E., Geertsema, H.J., Stratmann, S.A., Punter, C.M., Kulczyk, A.W., Richardson, C.C., and van Oijen, A.M. (2016). Simultaneous Real-Time Imaging of Leading and Lagging Strand Synthesis Reveals the Coordination Dynamics of Single Replisomes. Mol Cell *64*, 1035-1047. doi:10.1016/j.molcel.2016.10.028.

Dulin, D., Lipfert, J., Moolman, M.C., and Dekker, N.H. (2013). Studying genomic processes at the single-molecule level: introducing the tools and applications. Nat Rev Genet *14*, 9-22. doi:10.1038/nrg3316.

Duzdevich, D., Warner, M.D., Ticau, S., Ivica, N.A., Bell, S.P., and Greene, E.C. (2015). The dynamics of eukaryotic replication initiation: origin specificity, licensing, and firing at the single-molecule level. Mol Cell *58*, 483-494. doi:10.1016/j.molcel.2015.03.017.

Edelstein, A., Amodaj, N., Hoover, K., Vale, R., and Stuurman, N. (2010). Computer control of microscopes using microManager. Curr Protoc Mol Biol *Chapter 14*, Unit14 20. doi:10.1002/0471142727.mb1420s92.

Edwards, M.C., Tutter, A.V., Cvetic, C., Gilbert, C.H., Prokhorova, T.A., and Walter, J.C. (2002). MCM2-7 complexes bind chromatin in a distributed pattern surrounding the origin recognition complex in Xenopus egg extracts. J Biol Chem *277*, 33049-33057. doi:10.1074/jbc.M204438200.

Eickhoff, P., Kose, H.B., Martino, F., Petojevic, T., Abid Ali, F., Locke, J., Tamberg, N., Nans, A., Berger, J.M., Botchan, M.R.*, et al.* (2019). Molecular Basis for ATP-Hydrolysis-Driven DNA Translocation by the CMG Helicase of the Eukaryotic Replisome. Cell Rep *28*, 2673-2688 e2678. doi:10.1016/j.celrep.2019.07.104.

Erickson, H.P. (2009). Size and shape of protein molecules at the nanometer level determined by sedimentation, gel filtration, and electron microscopy. Biol Proced Online *11*, 32-51. doi:10.1007/s12575-009-9008-x.

Evrin, C., Clarke, P., Zech, J., Lurz, R., Sun, J., Uhle, S., Li, H., Stillman, B., and Speck, C. (2009). A double-hexameric MCM2-7 complex is loaded onto origin DNA during licensing of eukaryotic DNA replication. Proc Natl Acad Sci U S A *106*, 20240-20245. doi:10.1073/pnas.0911500106.

Flower, A.M., and McHenry, C.S. (1990). The gamma subunit of DNA polymerase III holoenzyme of Escherichia coli is produced by ribosomal frameshifting. Proc Natl Acad Sci U S A *87*, 3713-3717. doi:10.1073/pnas.87.10.3713.

Flyamer, I.M., Gassler, J., Imakaev, M., Brandao, H.B., Ulianov, S.V., Abdennur, N., Razin, S.V., Mirny, L.A., and Tachibana-Konwalski, K. (2017). Single-nucleus Hi-C reveals unique chromatin reorganization at oocyte-to-zygote transition. Nature *544*, 110-114. doi:10.1038/nature21711.

Foss, E.J., Gatbonton-Schwager, T., Thiesen, A.H., Taylor, E., Soriano, R., Lao, U., MacAlpine, D.M., and Bedalov, A. (2019). Sir2 suppresses transcription-mediated displacement of Mcm2-7 replicative helicases at the ribosomal DNA repeats. PLoS Genet *15*, e1008138. doi:10.1371/journal.pgen.1008138.

Freiburger, L., Sonntag, M., Hennig, J., Li, J., Zou, P., and Sattler, M. (2015). Efficient segmental isotope labeling of multi-domain proteins using Sortase A. J Biomol NMR *63*, 1-8. doi:10.1007/s10858-015-9981-0.

Frick, D.N., and Richardson, C.C. (1999). Interaction of bacteriophage T7 gene 4 primase with its template recognition site. J Biol Chem *274*, 35889-35898. doi:10.1074/jbc.274.50.35889.

Friedberg, E.C., Walker, G.C., Siede, W., Wood, R.D., Schultz, R.A., and Ellenberger, T. (2006). DNA Repair and Mutagenesis, 2 edn (American Society of Microbiology). ISBN:978-1555813192.

Friedman, L.J., Chung, J., and Gelles, J. (2006). Viewing dynamic assembly of molecular complexes by multi-wavelength single-molecule fluorescence. Biophys J *91*, 1023-1031. doi:10.1529/biophysj.106.084004.

Frigola, J., He, J., Kinkelin, K., Pye, V.E., Renault, L., Douglas, M.E., Remus, D., Cherepanov, P., Costa, A., and Diffley, J.F.X. (2017). Cdt1 stabilizes an open MCM ring for helicase loading. Nat Commun *8*, 15720. doi:10.1038/ncomms15720.

Frigola, J., Remus, D., Mehanna, A., and Diffley, J.F. (2013). ATPase-dependent quality control of DNA replication origin licensing. Nature *495*, 339-343. doi:10.1038/nature11920.

Fu, Y.V., Yardimci, H., Long, D.T., Ho, T.V., Guainazzi, A., Bermudez, V.P., Hurwitz, J., van Oijen, A., Scharer, O.D., and Walter, J.C. (2011). Selective bypass of a lagging strand roadblock by the eukaryotic replicative DNA helicase. Cell *146*, 931-941. doi:10.1016/j.cell.2011.07.045.

Galburt, E.A., Grill, S.W., Wiedmann, A., Lubkowska, L., Choy, J., Nogales, E., Kashlev, M., and Bustamante, C. (2007). Backtracking determines the force sensitivity of RNAP II in a factor-dependent manner. Nature *446*, 820-823. doi:10.1038/nature05701.

Gambus, A., Jones, R.C., Sanchez-Diaz, A., Kanemaki, M., van Deursen, F., Edmondson, R.D., and Labib, K. (2006). GINS maintains association of Cdc45 with MCM in replisome progression complexes at eukaryotic DNA replication forks. Nat Cell Biol *8*, 358-366. doi:10.1038/ncb1382.

Gambus, A., van Deursen, F., Polychronopoulos, D., Foltman, M., Jones, R.C., Edmondson, R.D., Calzada, A., and Labib, K. (2009). A key role for Ctf4 in coupling the MCM2-7 helicase to DNA polymerase alpha within the eukaryotic replisome. EMBO J *28*, 2992-3004. doi:10.1038/emboj.2009.226.

Ganier, O., Prorok, P., Akerman, I., and Mechali, M. (2019). Metazoan DNA replication origins. Curr Opin Cell Biol *58*, 134-141. doi:10.1016/j.ceb.2019.03.003.

Ganji, M., Shaltiel, I.A., Bisht, S., Kim, E., Kalichava, A., Haering, C.H., and Dekker, C. (2018). Real-time imaging of DNA loop extrusion by condensin. Science *360*, 102-105. doi:10.1126/science.aar7831.

Gao, D., and McHenry, C.S. (2001). 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*, 4441-4446. doi:10.1074/jbc.M009830200.

Gao, Y., Cui, Y., Fox, T., Lin, S., Wang, H., de Val, N., Zhou, Z.H., and Yang, W. (2019). Structures and operating principles of the replisome. Science *363*. doi:10.1126/science.aav7003.

Gassler, J., Brandao, H.B., Imakaev, M., Flyamer, I.M., Ladstatter, S., Bickmore, W.A., Peters, J.M., Mirny, L.A., and Tachibana, K. (2017). A mechanism of cohesin-dependent loop extrusion organizes zygotic genome architecture. EMBO J *36*, 3600-3618. doi:10.15252/embj.201798083.

Geertsema, H.J., Kulczyk, A.W., Richardson, C.C., and van Oijen, A.M. (2014). Single-molecule studies of polymerase dynamics and stoichiometry at the bacteriophage T7 replication machinery. Proc Natl Acad Sci U S A *111*, 4073-4078. doi:10.1073/pnas.1402010111.

Georgescu, R., Yuan, Z., Bai, L., de Luna Almeida Santos, R., Sun, J., Zhang, D., Yurieva, O., Li, H., and O'Donnell, M.E. (2017). Structure of eukaryotic CMG helicase at a replication fork and implications to replisome architecture and origin initiation. Proc Natl Acad Sci U S A *114*, E697-E706. doi:10.1073/pnas.1620500114.

Georgescu, R.E., Kurth, I., and O'Donnell, M.E. (2011). Single-molecule studies reveal the function of a third polymerase in the replisome. Nat Struct Mol Biol *19*, 113-116. doi:10.1038/nsmb.2179.

Georgescu, R.E., Langston, L., Yao, N.Y., Yurieva, O., Zhang, D., Finkelstein, J., Agarwal, T., and O'Donnell, M.E. (2014a). Mechanism of asymmetric polymerase assembly at the eukaryotic replication fork. Nat Struct Mol Biol *21*, 664-670. doi:10.1038/nsmb.2851.

Georgescu, R.E., Yao, N., Indiani, C., Yurieva, O., and O'Donnell, M.E. (2014b). Replisome mechanics: lagging strand events that influence speed and processivity. Nucleic Acids Res *42*, 6497-6510. doi:10.1093/nar/gku257.

Givaty, O., and Levy, Y. (2009). Protein sliding along DNA: dynamics and structural characterization. J Mol Biol *385*, 1087-1097. doi:10.1016/j.jmb.2008.11.016.

Gore, J., Bryant, Z., Stone, M.D., Nollmann, M., Cozzarelli, N.R., and Bustamante, C. (2006). Mechanochemical analysis of DNA gyrase using rotor bead tracking. Nature *439*, 100-104. doi:10.1038/nature04319.

Goswami, P., Abid Ali, F., Douglas, M.E., Locke, J., Purkiss, A., Janska, A., Eickhoff, P., Early, A., Nans, A., Cheung, A.M.C., *et al.* (2018). Structure of DNA-CMG-Pol epsilon elucidates the roles of the non-catalytic polymerase modules in the eukaryotic replisome. Nat Commun *9*, 5061. doi:10.1038/s41467-018-07417-1.

Graham, B.W., Schauer, G.D., Leuba, S.H., and Trakselis, M.A. (2011). Steric exclusion and wrapping of the excluded DNA strand occurs along discrete external binding paths during MCM helicase unwinding. Nucleic Acids Res *39*, 6585-6595. doi:10.1093/nar/gkr345.

Graham, J.E., Marians, K.J., and Kowalczykowski, S.C. (2017). Independent and Stochastic Action of DNA Polymerases in the Replisome. Cell *169*, 1201-1213 e1217. doi:10.1016/j.cell.2017.05.041.

Green, R., and Rogers, E.J. (2013). Transformation of chemically competent E. coli. Methods Enzymol *529*, 329-336. doi:10.1016/B978-0-12-418687-3.00028-8.

Griffith, F. (1928). The Significance of Pneumococcal Types. J Hyg (Lond) *27*, 113-159. doi:10.1017/s0022172400031879.

Gros, J., Devbhandari, S., and Remus, D. (2014). Origin plasticity during budding yeast DNA replication in vitro. EMBO J *33*, 621-636. doi:10.1002/embj.201387278.

Gros, J., Kumar, C., Lynch, G., Yadav, T., Whitehouse, I., and Remus, D. (2015). Post-licensing Specification of Eukaryotic Replication Origins by Facilitated Mcm2-7 Sliding along DNA. Mol Cell *60*, 797-807. doi:10.1016/j.molcel.2015.10.022.

Gruszka, D.T., Xie, S., Kimura, H., and Yardimci, H. (2020). Single-molecule imaging reveals control of parental histone recycling by free histones during DNA replication. Sci Adv *6*. doi:10.1126/sciadv.abc0330.

Guilliam, T.A., and Yeeles, J.T.P. (2020). Reconstitution of translesion synthesis reveals a mechanism of eukaryotic DNA replication restart. Nat Struct Mol Biol *27*, 450-460. doi:10.1038/s41594-020-0418-4.

Haarhuis, J.H.I., van der Weide, R.H., Blomen, V.A., Yanez-Cuna, J.O., Amendola, M., van Ruiten, M.S., Krijger, P.H.L., Teunissen, H., Medema, R.H., van Steensel, B., *et al.* (2017). The Cohesin Release Factor WAPL Restricts Chromatin Loop Extension. Cell *169*, 693-707 e614. doi:10.1016/j.cell.2017.04.013.

Hacker, K.J., and Alberts, B.M. (1994). The rapid dissociation of the T4 DNA polymerase holoenzyme when stopped by a DNA hairpin helix. A model for polymerase release following the termination of each Okazaki fragment. J Biol Chem *269*, 24221-24228. doi:10.1016/S0021-9258(19)51071-7.

Haering, C.H., Farcas, A.M., Arumugam, P., Metson, J., and Nasmyth, K. (2008). The cohesin ring concatenates sister DNA molecules. Nature *454*, 297-301. doi:10.1038/nature07098.

Hamdan, S.M., Loparo, J.J., Takahashi, M., Richardson, C.C., and van Oijen, A.M. (2009). Dynamics of DNA replication loops reveal temporal control of lagging-strand synthesis. Nature *457*, 336-339. doi:10.1038/nature07512.

Hanahan, D., and Weinberg, R.A. (2011). Hallmarks of cancer: the next generation. Cell *144*, 646-674. doi:10.1016/j.cell.2011.02.013.

Harvey, K.J., and Newport, J. (2003). CpG methylation of DNA restricts prereplication complex assembly in Xenopus egg extracts. Mol Cell Biol *23*, 6769-6779. doi:10.1128/mcb.23.19.6769-6779.2003.

Hassler, M., Shaltiel, I.A., and Haering, C.H. (2018). Towards a Unified Model of SMC Complex Function. Curr Biol *28*, R1266-R1281. doi:10.1016/j.cub.2018.08.034.

He, B., Rong, M., Lyakhov, D., Gartenstein, H., Diaz, G., Castagna, R., McAllister, W.T., and Durbin, R.K. (1997). Rapid mutagenesis and purification of phage RNA polymerases. Protein Expr Purif *9*, 142-151. doi:10.1006/prep.1996.0663.

Heinz, S., Texari, L., Hayes, M.G.B., Urbanowski, M., Chang, M.W., Givarkes, N., Rialdi, A., White, K.M., Albrecht, R.A., Pache, L., *et al.* (2018). Transcription Elongation Can Affect Genome 3D Structure. Cell *174*, 1522-1536 e1522. doi:10.1016/j.cell.2018.07.047.

Hershey, A.D., and Chase, M. (1952). Independent functions of viral protein and nucleic acid in growth of

bacteriophage. J Gen Physiol *36*, 39-56. doi:10.1085/jgp.36.1.39.

Higashi, T.L., Pobegalov, G., Tang, M., Molodtsov, M.I., and Uhlmann, F. (2021). A Brownian ratchet model for DNA loop extrusion by the cohesin complex. Elife *10*. doi:10.7554/eLife.67530.

Hill, F.R., van Oijen, A.M., and Duderstadt, K.E. (2018). Detection of kinetic change points in piece-wise linear single molecule motion. J Chem Phys *148*, 123317. doi:10.1063/1.5009387.

Hill, J., Eickhoff, P., Drury, L.S., Costa, A., and Diffley, J.F.X. (2020). The eukaryotic replisome requires an additional helicase to disarm dormant replication origins. bioRxiv. doi:10.1101/2020.09.17.301366.

Hill, M. (2012). McGraw-Hill Encyclopedia of Science and Technology, 11th edn (New York: McGraw-Hill Education). ISBN:978-0071792738.

Hodges, C., Bintu, L., Lubkowska, L., Kashlev, M., and Bustamante, C. (2009). Nucleosomal fluctuations govern the transcription dynamics of RNA polymerase II. Science *325*, 626-628. doi:10.1126/science.1172926.

Hoege, C., Pfander, B., Moldovan, G.L., Pyrowolakis, G., and Jentsch, S. (2002). RAD6-dependent DNA repair is linked to modification of PCNA by ubiquitin and SUMO. Nature *419*, 135-141. doi:10.1038/nature00991.

Hons, M.T., Huis In 't Veld, P.J., Kaesler, J., Rombaut, P., Schleiffer, A., Herzog, F., Stark, H., and Peters, J.M. (2016). Topology and structure of an engineered human cohesin complex bound to Pds5B. Nat Commun *7*, 12523. doi:10.1038/ncomms12523.

Hoshina, S., Yura, K., Teranishi, H., Kiyasu, N., Tominaga, A., Kadoma, H., Nakatsuka, A., Kunichika, T., Obuse, C., and Waga, S. (2013). Human origin recognition complex binds preferentially to G-quadruplex-preferable RNA and single-stranded DNA. J Biol Chem *288*, 30161-30171. doi:10.1074/jbc.M113.492504.

Hsieh, T.H., Weiner, A., Lajoie, B., Dekker, J., Friedman, N., and Rando, O.J. (2015). Mapping Nucleosome Resolution Chromosome Folding in Yeast by Micro-C. Cell *162*, 108-119. doi:10.1016/j.cell.2015.05.048.

Hu, Y., Tareen, A., Sheu, Y.J., Ireland, W.T., Speck, C., Li, H., Joshua-Tor, L., Kinney, J.B., and Stillman, B. (2020). Evolution of DNA replication origin specification and gene silencing mechanisms. Nat Commun *11*, 5175. doi:10.1038/s41467-020-18964-x.

Huang, H., Stromme, C.B., Saredi, G., Hodl, M., Strandsby, A., Gonzalez-Aguilera, C., Chen, S., Groth, A., and Patel, D.J. (2015). A unique binding mode enables MCM2 to chaperone histones H3-H4 at replication forks. Nat Struct Mol Biol *22*, 618-626. doi:10.1038/nsmb.3055.

Huis in 't Veld, P.J., Herzog, F., Ladurner, R., Davidson, I.F., Piric, S., Kreidl, E., Bhaskara, V., Aebersold, R., and Peters, J.M. (2014). Characterization of a DNA exit gate in the human cohesin ring. Science *346*, 968-972. doi:10.1126/science.1256904.

Indiani, C., Langston, L.D., Yurieva, O., Goodman, M.F., and O'Donnell, M. (2009). Translesion DNA polymerases remodel the replisome and alter the speed of the replicative helicase. Proc Natl Acad Sci U S A *106*, 6031-6038. doi:10.1073/pnas.0901403106.

Itsathitphaisarn, O., Wing, R.A., Eliason, W.K., Wang, J., and Steitz, T.A. (2012). The hexameric helicase DnaB adopts a nonplanar conformation during translocation. Cell *151*, 267-277. doi:10.1016/j.cell.2012.09.014.

Jacob, F., Brenner, S., and Cuzin, F. (1963). On the regulation of DNA replication in bacteria. Cold Spring Harbor Symp Quant Biol *28*, 329-348. doi:10.1101/SQB.1963.028.01.048.

Jaremko, M.J., On, K.F., Thomas, D.R., Stillman, B., and Joshua-Tor, L. (2020). The dynamic nature of the human origin recognition complex revealed through five cryoEM structures. Elife *9*. doi:10.7554/eLife.58622.

Jeng, S.T., Gardner, J.F., and Gumport, R.I. (1990). Transcription termination by bacteriophage T7 RNA polymerase at rho-independent terminators. J Biol Chem *265*, 3823-3830. doi:10.1016/S0021-9258(19)39668-1.

Jensen, T.H., Jacquier, A., and Libri, D. (2013). Dealing with pervasive transcription. Mol Cell *52*, 473-484. doi:10.1016/j.molcel.2013.10.032.

Jergic, S., Ozawa, K., Williams, N.K., Su, X.C., Scott, D.D., Hamdan, S.M., Crowther, J.A., Otting, G., and Dixon, N.E. (2007). 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*, 2813-2824. doi:10.1093/nar/gkm079.

Johnson, D.S., Bai, L., Smith, B.Y., Patel, S.S., and Wang, M.D. (2007). Single-molecule studies reveal dynamics of DNA unwinding by the ring-shaped T7 helicase. Cell *129*, 1299-1309. doi:10.1016/j.cell.2007.04.038.

Johnson, R.E., Klassen, R., Prakash, L., and Prakash, S. (2015). A Major Role of DNA Polymerase delta in Replication of Both the Leading and Lagging DNA Strands. Mol Cell *59*, 163-175. doi:10.1016/j.molcel.2015.05.038.

Kang, Y.H., Farina, A., Bermudez, V.P., Tappin, I., Du, F., Galal, W.C., and Hurwitz, J. (2013). Interaction between human Ctf4 and the Cdc45/Mcm2-7/GINS (CMG) replicative helicase. Proc Natl Acad Sci U S A *110*, 19760-19765. doi:10.1073/pnas.1320202110.

Kapadia, N., El-Hajj, Z.W., Zheng, H., Beattie, T.R., Yu, A., and Reyes-Lamothe, R. (2020). Processive Activity of Replicative DNA Polymerases in the Replisome of Live Eukaryotic Cells. Mol Cell *80*, 114-126 e118. doi:10.1016/j.molcel.2020.08.014.

Kath, J.E., Jergic, S., Heltzel, J.M., Jacob, D.T., Dixon, N.E., Sutton, M.D., Walker, G.C., and Loparo, J.J. (2014). Polymerase exchange on single DNA molecules reveals processivity clamp control of translesion synthesis. Proc Natl Acad Sci U S A *111*, 7647-7652. doi:10.1073/pnas.1321076111.

Katou, Y., Kanoh, Y., Bando, M., Noguchi, H., Tanaka, H., Ashikari, T., Sugimoto, K., and Shirahige, K. (2003). S-phase checkpoint proteins Tof1 and Mrc1 form a stable replication-pausing complex. Nature *424*, 1078-1083. doi:10.1038/nature01900.

Kawakami, H., Muraoka, R., Ohashi, E., Kawabata, K., Kanamoto, S., Chichibu, T., Tsurimoto, T., and Katayama, T. (2019). Specific basic patch-dependent multimerization of Saccharomyces cerevisiae ORC on single-stranded DNA promotes ATP hydrolysis. Genes Cells *24*, 608-618. doi:10.1111/gtc.12710.

Kesti, T., Flick, K., Keranen, S., Syvaoja, J.E., and Wittenberg, C. (1999). DNA polymerase epsilon catalytic domains are dispensable for DNA replication, DNA repair, and cell viability. Mol Cell *3*, 679-685. doi:10.1016/s1097-2765(00)80361-5.

Kim, S., Dallmann, H.G., McHenry, C.S., and Marians, K.J. (1996a). Coupling of a replicative polymerase and helicase: a tau-DnaB interaction mediates rapid replication fork movement. Cell *84*, 643-650. doi:10.1016/s0092-8674(00)81039-9.

Kim, S., Dallmann, H.G., McHenry, C.S., and Marians, K.J. (1996b). tau couples the leading- and lagging-strand polymerases at the Escherichia coli DNA replication fork. J Biol Chem *271*, 21406-21412. doi:10.1074/jbc.271.35.21406.

Kim, Y., Shi, Z., Zhang, H., Finkelstein, I.J., and Yu, H. (2019). Human cohesin compacts DNA by loop extrusion. Science *366*, 1345-1349. doi:10.1126/science.aaz4475.

Knoot, C.J., Purpero, V.M., and Lipscomb, J.D. (2015). Crystal structures of alkylperoxo and anhydride intermediates in an intradiol ring-cleaving dioxygenase. Proc Natl Acad Sci U S A *112*, 388-393. doi:10.1073/pnas.1419118112.

Kochaniak, A.B., Habuchi, S., Loparo, J.J., Chang, D.J., Cimprich, K.A., Walter, J.C., and van Oijen, A.M. (2009). Proliferating cell nuclear antigen uses two distinct modes to move along DNA. J Biol Chem *284*, 17700-17710. doi:10.1074/jbc.M109.008706.

Kong, D., Coleman, T.R., and DePamphilis, M.L. (2003). Xenopus origin recognition complex (ORC) initiates DNA replication preferentially at sequences targeted by Schizosaccharomyces pombe ORC. EMBO J *22*, 3441-3450. doi:10.1093/emboj/cdg319.

Kornberg, A., and Baker, T.A. (1992). DNA Replication (W.H. Freeman). ISBN:978-1891389443.

Kornberg, R.D. (1974). Chromatin structure: a repeating unit of histones and DNA. Science *184*, 868-871. doi:10.1126/science.184.4139.868.

Kosak, S.T., Skok, J.A., Medina, K.L., Riblet, R., Le Beau, M.M., Fisher, A.G., and Singh, H. (2002). Subnuclear compartmentalization of immunoglobulin loci during lymphocyte development. Science *296*, 158-162. doi:10.1126/science.1068768.

Kose, H.B., Larsen, N.B., Duxin, J.P., and Yardimci, H. (2019). Dynamics of the Eukaryotic Replicative Helicase at Lagging-Strand Protein Barriers Support the Steric Exclusion Model. Cell Rep *26*, 2113-2125 e2116. doi:10.1016/j.celrep.2019.01.086.

Kossel, A. (1879). Ueber das Nuclein in der Hefe. Z Physiol Chem *3*, 284-291.

Kossel, A. (1891). Ueber die chemische Zusammensetzung der Zelle. DuBois-Reymond's Arch *181*, 181-186.

Kowalski, D., and Eddy, M.J. (1989). The DNA unwinding element: a novel, cis-acting component that facilitates opening of the Escherichia coli replication origin. EMBO J *8*, 4335-4344. doi:10.1002/j.1460-2075.1989.tb08620.x.

Kumar, C., and Remus, D. (2016). Eukaryotic replication origins: Strength in flexibility. Nucleus (Calcutta) *7*, 292-300. doi:10.1080/19491034.2016.1187353.

Kuo, A.J., Song, J., Cheung, P., Ishibe-Murakami, S., Yamazoe, S., Chen, J.K., Patel, D.J., and Gozani, O. (2012). The BAH domain of ORC1 links H4K20me2 to DNA replication licensing and Meier-Gorlin syndrome. Nature *484*, 115-119. doi:10.1038/nature10956.

Kurat, C.F., Yeeles, J.T.P., Patel, H., Early, A., and Diffley, J.F.X. (2017). Chromatin Controls DNA Replication Origin Selection, Lagging-Strand Synthesis, and Replication Fork Rates. Mol Cell *65*, 117-130. doi:10.1016/j.molcel.2016.11.016.

Ladstatter, S., and Tachibana-Konwalski, K. (2016). A Surveillance Mechanism Ensures Repair of DNA Lesions during Zygotic Reprogramming. Cell *167*, 1774-1787 e1713. doi:10.1016/j.cell.2016.11.009.

Lander, E.S., Linton, L.M., Birren, B., Nusbaum, C., Zody, M.C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W.*, et al.* (2001). Initial sequencing and analysis of the human genome. Nature *409*, 860-921. doi:10.1038/35057062.

Langston, L., and O'Donnell, M. (2017). Action of CMG with strand-specific DNA blocks supports an internal unwinding mode for the eukaryotic replicative helicase. Elife *6*, e23449. doi:10.7554/eLife.23449.

Langston, L.D., Mayle, R., Schauer, G.D., Yurieva, O., Zhang, D., Yao, N.Y., Georgescu, R.E., and O'Donnell, M.E. (2017). Mcm10 promotes rapid isomerization of CMG-DNA for replisome bypass of lagging strand DNA blocks. Elife *6*, e29118. doi:10.7554/eLife.29118.

Langston, L.D., Zhang, D., Yurieva, O., Georgescu, R.E., Finkelstein, J., Yao, N.Y., Indiani, C., and O'Donnell, M.E. (2014). CMG helicase and DNA polymerase epsilon form a functional 15-subunit holoenzyme for eukaryotic leading-strand DNA replication. Proc Natl Acad Sci U S A *111*, 15390-15395. doi:10.1073/pnas.1418334111.

Larson, J., Kirk, M., Drier, E.A., O'Brien, W., MacKay, J.F., Friedman, L.J., and Hoskins, A.A. (2014). Design and construction of a multiwavelength, micromirror total internal reflectance fluorescence microscope. Nat Protoc *9*, 2317-2328. doi:10.1038/nprot.2014.155.

Larson, M.H., Greenleaf, W.J., Landick, R., and Block, S.M. (2008). Applied force reveals mechanistic and energetic details of transcription termination. Cell *132*, 971-982. doi:10.1016/j.cell.2008.01.027.

Laughery, M.F., Hunter, T., Brown, A., Hoopes, J., Ostbye, T., Shumaker, T., and Wyrick, J.J. (2015). New vectors for simple and streamlined CRISPR-Cas9 genome editing in Saccharomyces cerevisiae. Yeast *32*, 711-720. doi:10.1002/yea.3098.

Lee, C.S.K., Cheung, M.F., Li, J., Zhao, Y., Lam, W.H., Ho, V., Rohs, R., Zhai, Y., Leung, D., and Tye, B.K. (2021). Humanizing the yeast origin recognition complex. Nat Commun *12*, 33. doi:10.1038/s41467-020-20277-y.

Lee, D.G., and Bell, S.P. (1997). Architecture of the yeast origin recognition complex bound to origins of DNA replication. Mol Cell Biol *17*, 7159-7168. doi:10.1128/mcb.17.12.7159.

Lee, D.G., Makhov, A.M., Klemm, R.D., Griffith, J.D., and Bell, S.P. (2000). Regulation of origin recognition complex conformation and ATPase activity: differential effects of single-stranded and double-stranded DNA binding. EMBO J *19*, 4774-4782. doi:10.1093/emboj/19.17.4774.

Lee, J.B., Hite, R.K., Hamdan, S.M., Xie, X.S., Richardson, C.C., and van Oijen, A.M. (2006). DNA primase acts as a molecular brake in DNA replication. Nature *439*, 621-624. doi:10.1038/nature04317.

Lee, S.J., and Richardson, C.C. (2002). Interaction of adjacent primase domains within the hexameric gene 4 helicase-primase of bacteriophage T7. Proc Natl Acad Sci U S A *99*, 12703-12708. doi:10.1073/pnas.202471499.

Lee, S.J., Syed, S., Enemark, E.J., Schuck, S., Stenlund, A., Ha, T., and Joshua-Tor, L. (2014). Dynamic look at DNA unwinding by a replicative helicase. Proc Natl Acad Sci U S A *111*, E827-835. doi:10.1073/pnas.1322254111.

Lei, M., Kawasaki, Y., and Tye, B.K. (1996). Physical interactions among Mcm proteins and effects of Mcm dosage on DNA replication in Saccharomyces cerevisiae. Mol Cell Biol *16*, 5081-5090. doi:10.1128/mcb.16.9.5081.

Lewis, J.S., Spenkelink, L.M., Jergic, S., Wood, E.A., Monachino, E., Horan, N.P., Duderstadt, K.E., Cox, M.M., Robinson, A., Dixon, N.E*., et al.* (2017a). Single-molecule visualization of fast polymerase turnover in the bacterial replisome. Elife *6*, e23932. doi:10.7554/eLife.23932.

Lewis, J.S., Spenkelink, L.M., Schauer, G.D., Hill, F.R., Georgescu, R.E., O'Donnell, M.E., and van Oijen, A.M. (2017b). 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*, 10630-10635. doi:10.1073/pnas.1711291114.

Lewis, J.S., Spenkelink, L.M., Schauer, G.D., Yurieva, O., Mueller, S.H., Natarajan, V., Kaur, G., Maher, C., Kay, C., O'Donnell, M.E*., et al.* (2020). Tunability of DNA Polymerase Stability during Eukaryotic DNA Replication. Mol Cell *77*, 17-25 e15. doi:10.1016/j.molcel.2019.10.005.

Li, H., and O'Donnell, M.E. (2018). The Eukaryotic CMG Helicase at the Replication Fork: Emerging Architecture Reveals an Unexpected Mechanism. Bioessays *40*. doi:10.1002/bies.201700208.

Li, N., Lam, W.H., Zhai, Y., Cheng, J., Cheng, E., Zhao, Y., Gao, N., and Tye, B.-K. (2018). Structure of the origin recognition complex bound to DNA replication origin. Nature *559*, 217-222. doi:10.1038/s41586-018-0293-x.

Li, N., Zhai, Y., Zhang, Y., Li, W., Yang, M., Lei, J., Tye, B.K., and Gao, N. (2015). Structure of the eukaryotic MCM complex at 3.8 A. Nature *524*, 186-191. doi:10.1038/nature14685.

Li, S., Wasserman, M.R., Yurieva, O., Bai, L., O'Donnell, M.E., and Liu, S. (2021). ORC binds and remodels nucleosomes to specify MCM loading onto DNA. bioRxiv. doi:10.1101/2021.08.17.456647.

Li, X., and Marians, K.J. (2000). Two distinct triggers for cycling of the lagging strand polymerase at the replication fork. J Biol Chem *275*, 34757-34765. doi:10.1074/jbc.M006556200.

Li, Y., Haarhuis, J.H.I., Sedeno Cacciatore, A., Oldenkamp, R., van Ruiten, M.S., Willems, L., Teunissen, H., Muir, K.W., de Wit, E., Rowland, B.D*., et al.* (2020). The structural basis for cohesin-CTCF-anchored loops. Nature *578*, 472-476. doi:10.1038/s41586-019-1910-z.

Lia, G., Michel, B., and Allemand, J.F. (2012). Polymerase exchange during Okazaki fragment synthesis observed in living cells. Science *335*, 328-331. doi:10.1126/science.1210400.

Liao, Y., Li, Y., Schroeder, J.W., Simmons, L.A., and Biteen, J.S. (2016). Single-Molecule DNA Polymerase

Dynamics at a Bacterial Replisome in Live Cells. Biophys J *111*, 2562-2569. doi:10.1016/j.bpj.2016.11.006.

Lieberman-Aiden, E., van Berkum, N.L., Williams, L., Imakaev, M., Ragoczy, T., Telling, A., Amit, I., Lajoie, B.R., Sabo, P.J., Dorschner, M.O.*, et al.* (2009). Comprehensive mapping of long-range interactions reveals folding principles of the human genome. Science *326*, 289-293. doi:10.1126/science.1181369.

Lionnet, T., Spiering, M.M., Benkovic, S.J., Bensimon, D., and Croquette, V. (2007). Real-time observation of bacteriophage T4 gp41 helicase reveals an unwinding mechanism. Proc Natl Acad Sci U S A *104*, 19790-19795. doi:10.1073/pnas.0709793104.

Liu, Y., Zhou, K., Zhang, N., Wei, H., Tan, Y.Z., Zhang, Z., Carragher, B., Potter, C.S., D'Arcy, S., and Luger, K. (2020). FACT caught in the act of manipulating the nucleosome. Nature *577*, 426-431. doi:10.1038/s41586-019-1820-0.

Looke, M., Reimand, J., Sedman, T., Sedman, J., Jarvinen, L., Varv, S., Peil, K., Kristjuhan, K., Vilo, J., and Kristjuhan, A. (2010). Relicensing of transcriptionally inactivated replication origins in budding yeast. J Biol Chem *285*, 40004-40011. doi:10.1074/jbc.M110.148924.

Loparo, J.J., Kulczyk, A.W., Richardson, C.C., and van Oijen, A.M. (2011). 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*, 3584-3589. doi:10.1073/pnas.1018824108.

Low, E., Chistol, G., Zaher, M.S., Kochenova, O.V., and Walter, J.C. (2020). The DNA replication fork suppresses CMG unloading from chromatin before termination. Genes Dev *34*, 1534-1545. doi:10.1101/gad.339739.120.

Lowary, P.T., and Widom, J. (1998). New DNA sequence rules for high affinity binding to histone octamer and sequence-directed nucleosome positioning. J Mol Biol *276*, 19-42. doi:10.1006/jmbi.1997.1494.

Luo, G., Wang, M., Konigsberg, W.H., and Xie, X.S. (2007). Single-molecule and ensemble fluorescence assays for a functionally important conformational change in T7 DNA polymerase. Proc Natl Acad Sci U S A *104*, 12610-12615. doi:10.1073/pnas.0700920104.

Lupianez, D.G., Kraft, K., Heinrich, V., Krawitz, P., Brancati, F., Klopocki, E., Horn, D., Kayserili, H., Opitz, J.M., Laxova, R.*, et al.* (2015). Disruptions of topological chromatin domains cause pathogenic rewiring of gene-enhancer interactions. Cell *161*, 1012-1025. doi:10.1016/j.cell.2015.04.004.

Macdonald, L.E., Durbin, R.K., Dunn, J.J., and McAllister, W.T. (1994). Characterization of two types of termination signal for bacteriophage T7 RNA polymerase. J Mol Biol *238*, 145-158. doi:10.1006/jmbi.1994.1277.

Macdonald, L.E., Zhou, Y., and McAllister, W.T. (1993). Termination and slippage by bacteriophage T7 RNA polymerase. J Mol Biol *232*, 1030-1047. doi:10.1006/jmbi.1993.1458.

Mackiewicz, P., Zakrzewska-Czerwinska, J., Zawilak, A., Dudek, M.R., and Cebrat, S. (2004). Where does bacterial replication start? Rules for predicting the oriC region. Nucleic Acids Res *32*, 3781-3791. doi:10.1093/nar/gkh699.

Mahbubani, H.M., Chong, J.P., Chevalier, S., Thommes, P., and Blow, J.J. (1997). Cell cycle regulation of the replication licensing system: involvement of a Cdk-dependent inhibitor. J Cell Biol *136*, 125-135. doi:10.1083/jcb.136.1.125.

Maier, B., Bensimon, D., and Croquette, V. (2000). Replication by a single DNA polymerase of a stretched single-stranded DNA. Proc Natl Acad Sci U S A *97*, 12002-12007. doi:10.1073/pnas.97.22.12002.

Mairhofer, J., Wittwer, A., Cserjan-Puschmann, M., and Striedner, G. (2015). Preventing T7 RNA polymerase read-through transcription-A synthetic termination signal capable of improving bioprocess stability. ACS Synth Biol *4*, 265-273. doi:10.1021/sb5000115.

Mangiameli, S.M., Merrikh, C.N., Wiggins, P.A., and Merrikh, H. (2017). Transcription leads to pervasive replisome instability in bacteria. Elife *6*, e19848. doi:10.7554/eLife.19848.

Manosas, M., Spiering, M.M., Zhuang, Z., Benkovic, S.J., and Croquette, V. (2009). Coupling DNA

unwinding activity with primer synthesis in the bacteriophage T4 primosome. Nat Chem Biol *5*, 904-912. doi:10.1038/nchembio.236.

Mantiero, D., Mackenzie, A., Donaldson, A., and Zegerman, P. (2011). Limiting replication initiation factors execute the temporal programme of origin firing in budding yeast. EMBO J *30*, 4805-4814. doi:10.1038/emboj.2011.404.

Marahrens, Y., and Stillman, B. (1992). A yeast chromosomal origin of DNA replication defined by multiple functional elements. Science *255*, 817-823. doi:10.1126/science.1536007.

Maric, M., Maculins, T., De Piccoli, G., and Labib, K. (2014). Cdc48 and a ubiquitin ligase drive disassembly of the CMG helicase at the end of DNA replication. Science *346*, 1253596. doi:10.1126/science.1253596.

Maric, M., Mukherjee, P., Tatham, M.H., Hay, R., and Labib, K. (2017). Ufd1-Npl4 Recruit Cdc48 for Disassembly of Ubiquitylated CMG Helicase at the End of Chromosome Replication. Cell Rep *18*, 3033-3042. doi:10.1016/j.celrep.2017.03.020.

McAllister, W.T., and Carter, A.D. (1980). Regulation of promoter selection by the bacteriophage T7 RNA polymerase in vitro. Nucleic Acids Res *8*, 4821-4837. doi:10.1093/nar/8.20.4821.

McDonald, J.P., Levine, A.S., and Woodgate, R. (1997). The Saccharomyces cerevisiae RAD30 gene, a homologue of Escherichia coli dinB and umuC, is DNA damage inducible and functions in a novel error-free postreplication repair mechanism. Genetics *147*, 1557-1568. doi:10.1093/genetics/147.4.1557.

McGarry, T.J., and Kirschner, M.W. (1998). Geminin, an inhibitor of DNA replication, is degraded during mitosis. Cell *93*, 1043-1053. doi:10.1016/s0092-8674(00)81209-x.

McGeoch, A.T., Trakselis, M.A., Laskey, R.A., and Bell, S.D. (2005). Organization of the archaeal MCM complex on DNA and implications for the helicase mechanism. Nat Struct Mol Biol *12*, 756-762. doi:10.1038/nsmb974.

McInerney, P., Johnson, A., Katz, F., and O'Donnell, M. (2007). Characterization of a triple DNA polymerase replisome. Mol Cell *27*, 527-538. doi:10.1016/j.molcel.2007.06.019.

McInerney, P., and O'Donnell, M. (2004). Functional uncoupling of twin polymerases: mechanism of polymerase dissociation from a lagging-strand block. J Biol Chem *279*, 21543-21551. doi:10.1074/jbc.M401649200.

Meselson, M., and Stahl, F.W. (1958). The Replication of DNA in Escherichia Coli. Proc Natl Acad Sci U S A *44*, 671-682. doi:10.1073/pnas.44.7.671.

Miescher, F. (1869). Letter I; to Wilhelm His; Tübingen, February 26th, 1869. W His, et al (Eds), Die Histochemischen und Physiologischen Arbeiten von Friedrich Miescher—Aus dem wissenschaftlichen Briefwechsel von F Miescher, vol 1, FCW Vogel, Leipzig (1869), 33-38.

Miescher, F. (1871). Ueber die chemische Zusammensetzung der Eiterzellen. Med-Chem Unters *4*, 441-460.

Miller, T.C.R., Locke, J., Greiwe, J.F., Diffley, J.F.X., and Costa, A. (2019). Mechanism of head-to-head MCM double-hexamer formation revealed by cryo-EM. Nature *575*, 704-710. doi:10.1038/s41586-019-1768-0.

Mirny, L., Slutsky, M., Wunderlich, Z., Tafvizi, A., Leith, J., and Kosmrlj, A. (2009). How a protein searches for its site on DNA: the mechanism of facilitated diffusion. Journal of Physics A: Mathematical and Theoretical *42*. doi:10.1088/1751-8113/42/43/434013.

Mischo, H.E., and Proudfoot, N.J. (2013). Disengaging polymerase: terminating RNA polymerase II transcription in budding yeast. Biochim Biophys Acta *1829*, 174-185. doi:10.1016/j.bbagrm.2012.10.003.

Moffitt, J.R., Chemla, Y.R., Smith, S.B., and Bustamante, C. (2008). Recent advances in optical tweezers. Annu Rev Biochem *77*, 205-228. doi:10.1146/annurev.biochem.77.043007.090225.

Monachino, E., Jergic, S., Lewis, J.S., Xu, Z.Q., Lo, A.T.Y., O'Shea, V.L., Berger, J.M., Dixon, N.E., and van

Oijen, A.M. (2020). A Primase-Induced Conformational Switch Controls the Stability of the Bacterial Replisome. Mol Cell *79*, 140-154 e147. doi:10.1016/j.molcel.2020.04.037.

Moolman, M.C., Krishnan, S.T., Kerssemakers, J.W., van den Berg, A., Tulinski, P., Depken, M., Reyes-Lamothe, R., Sherratt, D.J., and Dekker, N.H. (2014). Slow unloading leads to DNA-bound beta2-sliding clamp accumulation in live Escherichia coli cells. Nat Commun *5*, 5820. doi:10.1038/ncomms6820.

Murayama, Y., and Uhlmann, F. (2014). Biochemical reconstitution of topological DNA binding by the cohesin ring. Nature *505*, 367-371. doi:10.1038/nature12867.

Murayama, Y., and Uhlmann, F. (2015). DNA Entry into and Exit out of the Cohesin Ring by an Interlocking Gate Mechanism. Cell *163*, 1628-1640. doi:10.1016/j.cell.2015.11.030.

Myllykallio, H., Lopez, P., Lopez-Garcia, P., Heilig, R., Saurin, W., Zivanovic, Y., Philippe, H., and Forterre, P. (2000). Bacterial mode of replication with eukaryotic-like machinery in a hyperthermophilic archaeon. Science *288*, 2212-2215. doi:10.1126/science.288.5474.2212.

Nelson, D.L., and Cox, M.M. (2005). Lehninger Principles of Biochemistry (W. H. Freeman). ISBN:978-1429234146.

Nelson, S.W., Kumar, R., and Benkovic, S.J. (2008). RNA primer handoff in bacteriophage T4 DNA replication: the role of single-stranded DNA-binding protein and polymerase accessory proteins. J Biol Chem *283*, 22838-22846. doi:10.1074/jbc.M802762200.

Neuman, K.C., and Nagy, A. (2008). Single-molecule force spectroscopy: optical tweezers, magnetic tweezers and atomic force microscopy. Nat Methods *5*, 491-505. doi:10.1038/nmeth.1218.

Nguyen, V.Q., Co, C., Irie, K., and Li, J.J. (2000). Clb/Cdc28 kinases promote nuclear export of the replication initiator proteins Mcm2-7. Curr Biol *10*, 195-205. doi:10.1016/s0960-9822(00)00337-7.

Nishana, M., Ha, C., Rodriguez-Hernaez, J., Ranjbaran, A., Chio, E., Nora, E.P., Badri, S.B., Kloetgen, A., Bruneau, B.G., Tsirigos, A.*, et al.* (2020). Defining the relative and combined contribution of CTCF and CTCFL to genomic regulation. Genome Biol *21*, 108. doi:10.1186/s13059-020-02024-0.

Noguchi, Y., Yuan, Z., Bai, L., Schneider, S., Zhao, G., Stillman, B., Speck, C., and Li, H. (2017). Cryo-EM structure of Mcm2-7 double hexamer on DNA suggests a lagging-strand DNA extrusion model. Proc Natl Acad Sci U S A *114*, E9529-E9538. doi:10.1073/pnas.1712537114.

Nora, E.P., Caccianini, L., Fudenberg, G., So, K., Kameswaran, V., Nagle, A., Uebersohn, A., Hajj, B., Saux, A.L., Coulon, A.*, et al.* (2020). Molecular basis of CTCF binding polarity in genome folding. Nat Commun *11*, 5612. doi:10.1038/s41467-020-19283-x.

Nora, E.P., Goloborodko, A., Valton, A.L., Gibcus, J.H., Uebersohn, A., Abdennur, N., Dekker, J., Mirny, L.A., and Bruneau, B.G. (2017). Targeted Degradation of CTCF Decouples Local Insulation of Chromosome Domains from Genomic Compartmentalization. Cell *169*, 930-944 e922. doi:10.1016/j.cell.2017.05.004.

Nora, E.P., Lajoie, B.R., Schulz, E.G., Giorgetti, L., Okamoto, I., Servant, N., Piolot, T., van Berkum, N.L., Meisig, J., Sedat, J.*, et al.* (2012). Spatial partitioning of the regulatory landscape of the X-inactivation centre. Nature *485*, 381-385. doi:10.1038/nature11049.

O'Neill, T.E., Smith, J.G., and Bradbury, E.M. (1993). Histone octamer dissociation is not required for transcript elongation through arrays of nucleosome cores by phage T7 RNA polymerase in vitro. Proc Natl Acad Sci U S A *90*, 6203-6207. doi:10.1073/pnas.90.13.6203.

Oakley, J.L., Strothkamp, R.E., Sarris, A.H., and Coleman, J.E. (1979). T7 RNA polymerase: promoter structure and polymerase binding. Biochemistry *18*, 528-537. doi:10.1021/bi00570a023.

Okazaki, R., Okazaki, T., Sakabe, K., Sugimoto, K., and Sugino, A. (1968). Mechanism of DNA chain growth. I. Possible discontinuity and unusual secondary structure of newly synthesized chains. Proc Natl Acad Sci U S A *59*, 598-605. doi:10.1073/pnas.59.2.598.

On, K.F., Beuron, F., Frith, D., Snijders, A.P., Morris, E.P., and Diffley, J.F. (2014). Prereplicative complexes assembled in vitro support origin-dependent and independent DNA replication. EMBO J *33*, 605-620.

doi:10.1002/embj.201387369.

Pandey, M., Syed, S., Donmez, I., Patel, G., Ha, T., and Patel, S.S. (2009). Coordinating DNA replication by means of priming loop and differential synthesis rate. Nature *462*, 940-943. doi:10.1038/nature08611.

Parelho, V., Hadjur, S., Spivakov, M., Leleu, M., Sauer, S., Gregson, H.C., Jarmuz, A., Canzonetta, C., Webster, Z., Nesterova, T.*, et al.* (2008). Cohesins functionally associate with CTCF on mammalian chromosome arms. Cell *132*, 422-433. doi:10.1016/j.cell.2008.01.011.

Park, K., Debyser, Z., Tabor, S., Richardson, C.C., and Griffith, J.D. (1998). Formation of a DNA loop at the replication fork generated by bacteriophage T7 replication proteins. J Biol Chem *273*, 5260-5270. doi:10.1074/jbc.273.9.5260.

Pelechano, V., Wei, W., and Steinmetz, L.M. (2013). Extensive transcriptional heterogeneity revealed by isoform profiling. Nature *497*, 127-131. doi:10.1038/nature12121.

Perez-Arnaiz, P., Bruck, I., Colbert, M.K., and Kaplan, D.L. (2017). An intact Mcm10 coiled-coil interaction surface is important for origin melting, helicase assembly and the recruitment of Pol-alpha to Mcm2-7. Nucleic Acids Res *45*, 7261-7275. doi:10.1093/nar/gkx438.

Petela, N.J., Gligoris, T.G., Metson, J., Lee, B.G., Voulgaris, M., Hu, B., Kikuchi, S., Chapard, C., Chen, W., Rajendra, E.*, et al.* (2018). Scc2 Is a Potent Activator of Cohesin's ATPase that Promotes Loading by Binding Scc1 without Pds5. Mol Cell *70*, 1134-1148 e1137. doi:10.1016/j.molcel.2018.05.022.

Petryk, N., Dalby, M., Wenger, A., Stromme, C.B., Strandsby, A., Andersson, R., and Groth, A. (2018). MCM2 promotes symmetric inheritance of modified histones during DNA replication. Science *361*, 1389-1392. doi:10.1126/science.aau0294.

Pope, B.D., Ryba, T., Dileep, V., Yue, F., Wu, W., Denas, O., Vera, D.L., Wang, Y., Hansen, R.S., Canfield, T.K.*, et al.* (2014). Topologically associating domains are stable units of replication-timing regulation. Nature *515*, 402-405. doi:10.1038/nature13986.

Powell, S.K., MacAlpine, H.K., Prinz, J.A., Li, Y., Belsky, J.A., and MacAlpine, D.M. (2015). Dynamic loading and redistribution of the Mcm2-7 helicase complex through the cell cycle. EMBO J *34*, 531-543. doi:10.15252/embj.201488307.

Pradhan, B., Barth, R., Kim, E., Davidson, I.F., Bauer, B., van Laar, T., Yang, W., Ryu, J.-K., van der Torre, J., Peters, J.-M.*, et al.* (2021). SMC complexes can traverse physical roadblocks bigger than their ring size. bioRxiv. doi:10.1101/2021.07.15.452501.

Prioleau, M.N., and MacAlpine, D.M. (2016). DNA replication origins-where do we begin? Genes Dev *30*, 1683-1697. doi:10.1101/gad.285114.116.

Pugacheva, E.M., Kubo, N., Loukinov, D., Tajmul, M., Kang, S., Kovalchuk, A.L., Strunnikov, A.V., Zentner, G.E., Ren, B., and Lobanenkov, V.V. (2020). CTCF mediates chromatin looping via N-terminal domain-dependent cohesin retention. Proc Natl Acad Sci U S A *117*, 2020-2031. doi:10.1073/pnas.1911708117.

Purcell, E.M. (1977). Life at Low Reynolds-Number. Am J Phys *45*, 3-11. doi:10.1119/1.10903.

Rausch, C., Weber, P., Prorok, P., Horl, D., Maiser, A., Lehmkuhl, A., Chagin, V.O., Casas-Delucchi, C.S., Leonhardt, H., and Cardoso, M.C. (2020). Developmental differences in genome replication program and origin activation. Nucleic Acids Res *48*, 12751-12777. doi:10.1093/nar/gkaa1124.

Remus, D., Beuron, F., Tolun, G., Griffith, J.D., Morris, E.P., and Diffley, J.F. (2009). Concerted loading of Mcm2-7 double hexamers around DNA during DNA replication origin licensing. Cell *139*, 719-730. doi:10.1016/j.cell.2009.10.015.

Reyes-Lamothe, R., Sherratt, D.J., and Leake, M.C. (2010). Stoichiometry and architecture of active DNA replication machinery in Escherichia coli. Science *328*, 498-501. doi:10.1126/science.1185757.

Ribeck, N., and Saleh, O.A. (2008). Multiplexed single-molecule measurements with magnetic tweezers. Rev Sci Instrum *79*, 094301. doi:10.1063/1.2981687.

Rothenberg, E., Trakselis, M.A., Bell, S.D., and Ha, T. (2007). MCM forked substrate specificity involves dynamic interaction with the 5'-tail. J Biol Chem *282*, 34229-34234. doi:10.1074/jbc.M706300200.

Rubio, E.D., Reiss, D.J., Welcsh, P.L., Disteche, C.M., Filippova, G.N., Baliga, N.S., Aebersold, R., Ranish, J.A., and Krumm, A. (2008). CTCF physically links cohesin to chromatin. Proc Natl Acad Sci U S A *105*, 8309-8314. doi:10.1073/pnas.0801273105.

Rudolph, C.J., Upton, A.L., Stockum, A., Nieduszynski, C.A., and Lloyd, R.G. (2013). Avoiding chromosome pathology when replication forks collide. Nature *500*, 608-611. doi:10.1038/nature12312.

Schauer, G.D., Spenkelink, L.M., Lewis, J.S., Yurieva, O., Mueller, S.H., van Oijen, A.M., and O'Donnell, M.E. (2020). Replisome bypass of a protein-based R-loop block by Pif1. Proc Natl Acad Sci U S A *117*, 30354-30361. doi:10.1073/pnas.2020189117.

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B.*, et al.* (2012). Fiji: an open-source platform for biological-image analysis. Nat Methods *9*, 676-682. doi:10.1038/nmeth.2019.

Schmidt, J.M., and Bleichert, F. (2020). Structural mechanism for replication origin binding and remodeling by a metazoan origin recognition complex and its co-loader Cdc6. Nat Commun *11*, 4263. doi:10.1038/s41467-020-18067-7.

Schneider, C.A., Rasband, W.S., and Eliceiri, K.W. (2012). NIH Image to ImageJ: 25 years of image analysis. Nat Methods *9*, 671-675. doi:10.1038/nmeth.2089.

Schurr, J.M. (1979). The one-dimensional diffusion coefficient of proteins absorbed on DNA. Hydrodynamic considerations. Biophys Chem *9*, 413-414. doi:10.1016/0301-4622(75)80057-3.

Sedlackova, H., Rask, M.B., Gupta, R., Choudhary, C., Somyajit, K., and Lukas, J. (2020). Equilibrium between nascent and parental MCM proteins protects replicating genomes. Nature *587*, 297-302. doi:10.1038/s41586-020-2842-3.

Selvin, P.R., and Ha, T. (2008). Single-Molecule Techniques: A Laboratory Manual (Cold Spring Harbor Laboratory Press). ISBN:978-087969775-4.

Senavirathne, G., Lopez, M.A., Jr., Messer, R., Fishel, R., and Yoder, K.E. (2018). Expression and purification of nuclease-free protocatechuate 3,4-dioxygenase for prolonged single-molecule fluorescence imaging. Anal Biochem *556*, 78-84. doi:10.1016/j.ab.2018.06.016.

Sengupta, S., van Deursen, F., de Piccoli, G., and Labib, K. (2013). Dpb2 integrates the leading-strand DNA polymerase into the eukaryotic replisome. Curr Biol *23*, 543-552. doi:10.1016/j.cub.2013.02.011.

Shen, Z., Sathyan, K.M., Geng, Y., Zheng, R., Chakraborty, A., Freeman, B., Wang, F., Prasanth, K.V., and Prasanth, S.G. (2010). A WD-repeat protein stabilizes ORC binding to chromatin. Mol Cell *40*, 99-111. doi:10.1016/j.molcel.2010.09.021.

Shi, Z., Gao, H., Bai, X.C., and Yu, H. (2020). Cryo-EM structure of the human cohesin-NIPBL-DNA complex. Science *368*, 1454-1459. doi:10.1126/science.abb0981.

Shibata, E., and Dutta, A. (2020). A human cancer cell line initiates DNA replication normally in the absence of ORC5 and ORC2 proteins. J Biol Chem *295*, 16949-16959. doi:10.1074/jbc.RA120.015450.

Shibata, E., Kiran, M., Shibata, Y., Singh, S., Kiran, S., and Dutta, A. (2016). Two subunits of human ORC are dispensable for DNA replication and proliferation. Elife *5*. doi:10.7554/eLife.19084.

Silva, M.C.C., Powell, S., Ladstatter, S., Gassler, J., Stocsits, R., Tedeschi, A., Peters, J.M., and Tachibana, K. (2020). Wapl releases Scc1-cohesin and regulates chromosome structure and segregation in mouse oocytes. J Cell Biol *219*. doi:10.1083/jcb.201906100.

Simon, A.C., Zhou, J.C., Perera, R.L., van Deursen, F., Evrin, C., Ivanova, M.E., Kilkenny, M.L., Renault, L., Kjaer, S., Matak-Vinkovic, D.*, et al.* (2014). A Ctf4 trimer couples the CMG helicase to DNA polymerase alpha in the eukaryotic replisome. Nature *510*, 293-297. doi:10.1038/nature13234.

Singer, M.F. (1968). 1968 Nobel Laureate in Medicine or Physiology. Science *162*, 433-436. doi:10.1126/science.162.3852.433.

Skoneczna, A., McIntyre, J., Skoneczny, M., Policinska, Z., and Sledziewska-Gojska, E. (2007). Polymerase eta is a short-lived, proteasomally degraded protein that is temporarily stabilized following UV irradiation in Saccharomyces cerevisiae. J Mol Biol *366*, 1074-1086. doi:10.1016/j.jmb.2006.11.093.

Smiley, R.D., Zhuang, Z., Benkovic, S.J., and Hammes, G.G. (2006). Single-molecule investigation of the T4 bacteriophage DNA polymerase holoenzyme: multiple pathways of holoenzyme formation. Biochemistry *45*, 7990-7997. doi:10.1021/bi0603322.

Soh, Y.M., Burmann, F., Shin, H.C., Oda, T., Jin, K.S., Toseland, C.P., Kim, C., Lee, H., Kim, S.J., Kong, M.S*., et al.* (2015). Molecular basis for SMC rod formation and its dissolution upon DNA binding. Mol Cell *57*, 290-303. doi:10.1016/j.molcel.2014.11.023.

Sousa, R., Chung, Y.J., Rose, J.P., and Wang, B.C. (1993). Crystal structure of bacteriophage T7 RNA polymerase at 3.3 A resolution. Nature *364*, 593-599. doi:10.1038/364593a0.

Sparks, J.L., Chistol, G., Gao, A.O., Raschle, M., Larsen, N.B., Mann, M., Duxin, J.P., and Walter, J.C. (2019). The CMG Helicase Bypasses DNA-Protein Cross-Links to Facilitate Their Repair. Cell *176*, 167-181 e121. doi:10.1016/j.cell.2018.10.053.

Stedman, W., Kang, H., Lin, S., Kissil, J.L., Bartolomei, M.S., and Lieberman, P.M. (2008). Cohesins localize with CTCF at the KSHV latency control region and at cellular c-myc and H19/Igf2 insulators. EMBO J *27*, 654-666. doi:10.1038/emboj.2008.1.

Stigler, J., Camdere, G.O., Koshland, D.E., and Greene, E.C. (2016). Single-Molecule Imaging Reveals a Collapsed Conformational State for DNA-Bound Cohesin. Cell Rep *15*, 988-998. doi:10.1016/j.celrep.2016.04.003.

Stinchcomb, D.T., Struhl, K., and Davis, R.W. (1979). Isolation and characterisation of a yeast chromosomal replicator. Nature *282*, 39-43. doi:10.1038/282039a0.

Stracy, M., Schweizer, J., Sherratt, D.J., Kapanidis, A.N., Uphoff, S., and Lesterlin, C. (2021). Transient non-specific DNA binding dominates the target search of bacterial DNA-binding proteins. Mol Cell *81*, 1499-1514 e1496. doi:10.1016/j.molcel.2021.01.039.

Strick, T.R., Allemand, J.F., Bensimon, D., and Croquette, V. (1998). Behavior of supercoiled DNA. Biophys J *74*, 2016-2028. doi:10.1016/S0006-3495(98)77908-1.

Studitsky, V.M., Clark, D.J., and Felsenfeld, G. (1994). A histone octamer can step around a transcribing polymerase without leaving the template. Cell *76*, 371-382. doi:10.1016/0092-8674(94)90343-3.

Studitsky, V.M., Kassavetis, G.A., Geiduschek, E.P., and Felsenfeld, G. (1997). Mechanism of transcription through the nucleosome by eukaryotic RNA polymerase. Science *278*, 1960-1963. doi:10.1126/science.278.5345.1960.

Stukenberg, P.T., Turner, J., and O'Donnell, M. (1994). An explanation for lagging strand replication: polymerase hopping among DNA sliding clamps. Cell *78*, 877-887. doi:10.1016/s0092-8674(94)90662-9.

Sun, B., Johnson, D.S., Patel, G., Smith, B.Y., Pandey, M., Patel, S.S., and Wang, M.D. (2011). ATP-induced helicase slippage reveals highly coordinated subunits. Nature *478*, 132-135. doi:10.1038/nature10409.

Sun, J., Shi, Y., Georgescu, R.E., Yuan, Z., Chait, B.T., Li, H., and O'Donnell, M.E. (2015). The architecture of a eukaryotic replisome. Nat Struct Mol Biol *22*, 976-982. doi:10.1038/nsmb.3113.

Svoboda, K., and Block, S.M. (1994). Biological applications of optical forces. Annu Rev Biophys Biomol Struct *23*, 247-285. doi:10.1146/annurev.bb.23.060194.001335.

Tafvizi, A., Huang, F., Leith, J.S., Fersht, A.R., Mirny, L.A., and van Oijen, A.M. (2008). Tumor suppressor p53 slides on DNA with low friction and high stability. Biophys J *95*, L01-03. doi:10.1529/biophysj.108.134122.

Tahirov, T.H., Temiakov, D., Anikin, M., Patlan, V., McAllister, W.T., Vassylyev, D.G., and Yokoyama, S. (2002). Structure of a T7 RNA polymerase elongation complex at 2.9 A resolution. Nature *420*, 43-50. doi:10.1038/nature01129.

Tanaka, S., Halter, D., Livingstone-Zatchej, M., Reszel, B., and Thoma, F. (1994). Transcription through the yeast origin of replication ARS1 ends at the ABFI binding site and affects extrachromosomal maintenance of minichromosomes. Nucleic Acids Res *22*, 3904-3910. doi:10.1093/nar/22.19.3904.

Tang, M., Shen, X., Frank, E.G., O'Donnell, M., Woodgate, R., and Goodman, M.F. (1999). UmuD'(2)C is an error-prone DNA polymerase, Escherichia coli pol V. Proc Natl Acad Sci U S A *96*, 8919-8924. doi:10.1073/pnas.96.16.8919.

Tanner, N.A., Hamdan, S.M., Jergic, S., Loscha, K.V., Schaeffer, P.M., Dixon, N.E., and van Oijen, A.M. (2008). Single-molecule studies of fork dynamics in Escherichia coli DNA replication. Nat Struct Mol Biol *15*, 998. doi:10.1038/nsmb0908-998a.

Tanner, N.A., Loparo, J.J., Hamdan, S.M., Jergic, S., Dixon, N.E., and van Oijen, A.M. (2009). Real-time single-molecule observation of rolling-circle DNA replication. Nucleic Acids Res *37*, e27. doi:10.1093/nar/gkp006.

Tanner, N.A., Tolun, G., Loparo, J.J., Jergic, S., Griffith, J.D., Dixon, N.E., and van Oijen, A.M. (2011). E. coli DNA replication in the absence of free beta clamps. EMBO J *30*, 1830-1840. doi:10.1038/emboj.2011.84.

Tanner, N.A., and van Oijen, A.M. (2009). Single-molecule observation of prokaryotic DNA replication. Methods Mol Biol *521*, 397-410. doi:10.1007/978-1-60327-815-7_22.

Taylor, M.R.G., and Yeeles, J.T.P. (2018). The Initial Response of a Eukaryotic Replisome to DNA Damage. Mol Cell *70*, 1067-1080 e1012. doi:10.1016/j.molcel.2018.04.022.

Theis, J.F., and Newlon, C.S. (1997). The ARS309 chromosomal replicator of Saccharomyces cerevisiae depends on an exceptional ARS consensus sequence. Proc Natl Acad Sci U S A *94*, 10786-10791. doi:10.1073/pnas.94.20.10786.

Thomen, P., Lopez, P.J., Bockelmann, U., Guillerez, J., Dreyfus, M., and Heslot, F. (2008). T7 RNA polymerase studied by force measurements varying cofactor concentration. Biophys J *95*, 2423-2433. doi:10.1529/biophysj.107.125096.

Ticau, S., Friedman, L.J., Champasa, K., Correa, I.R., Jr., Gelles, J., and Bell, S.P. (2017). Mechanism and timing of Mcm2-7 ring closure during DNA replication origin licensing. Nat Struct Mol Biol *24*, 309-315. doi:10.1038/nsmb.3375.

Ticau, S., Friedman, L.J., Ivica, N.A., Gelles, J., and Bell, S.P. (2015). Single-molecule studies of origin licensing reveal mechanisms ensuring bidirectional helicase loading. Cell *161*, 513-525. doi:10.1016/j.cell.2015.03.012.

Tsimring, L.S. (2014). Noise in biology. Rep Prog Phys *77*, 026601. doi:10.1088/0034-4885/77/2/026601.

van Oijen, A.M., Blainey, P.C., Crampton, D.J., Richardson, C.C., Ellenberger, T., and Xie, X.S. (2003). Single-molecule kinetics of lambda exonuclease reveal base dependence and dynamic disorder. Science *301*, 1235-1238. doi:10.1126/science.1084387.

Vaziri, C., Saxena, S., Jeon, Y., Lee, C., Murata, K., Machida, Y., Wagle, N., Hwang, D.S., and Dutta, A. (2003). A p53-dependent checkpoint pathway prevents rereplication. Mol Cell *11*, 997-1008. doi:10.1016/s1097-2765(03)00099-6.

Venter, J.C., Adams, M.D., Myers, E.W., Li, P.W., Mural, R.J., Sutton, G.G., Smith, H.O., Yandell, M., Evans, C.A., Holt, R.A., *et al.* (2001). The sequence of the human genome. Science *291*, 1304-1351. doi:10.1126/science.1058040.

Vian, L., Pekowska, A., Rao, S.S.P., Kieffer-Kwon, K.R., Jung, S., Baranello, L., Huang, S.C., El Khattabi, L., Dose, M., Pruett, N., *et al.* (2018). The Energetics and Physiological Impact of Cohesin Extrusion. Cell *173*, 1165-1178 e1120. doi:10.1016/j.cell.2018.03.072.

Wagner, J., Gruz, P., Kim, S.R., Yamada, M., Matsui, K., Fuchs, R.P., and Nohmi, T. (1999). The dinB gene encodes a novel E. coli DNA polymerase, DNA pol IV, involved in mutagenesis. Mol Cell *4*, 281-286. doi:10.1016/s1097-2765(00)80376-7.

Wasserman, M.R., Schauer, G.D., O'Donnell, M.E., and Liu, S. (2019). Replication Fork Activation Is Enabled by a Single-Stranded DNA Gate in CMG Helicase. Cell *178*, 600-611 e616. doi:10.1016/j.cell.2019.06.032.

Watson, J.D., and Crick, F.H. (1953). Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. Nature *171*, 737-738. doi:10.1038/171737a0.

Weigel, C., and Seitz, H. (2006). Bacteriophage replication modules. FEMS Microbiol Rev *30*, 321-381. doi:10.1111/j.1574-6976.2006.00015.x.

Weinberg, R.A. (2014). The Biology of Cancer, 2nd edn (Garland Publishing Inc). ISBN:978-0815342205.

Wendt, K.S., Yoshida, K., Itoh, T., Bando, M., Koch, B., Schirghuber, E., Tsutsumi, S., Nagae, G., Ishihara, K., Mishiro, T.*, et al.* (2008). Cohesin mediates transcriptional insulation by CCCTC-binding factor. Nature *451*, 796-801. doi:10.1038/nature06634.

Wohlschlegel, J.A., Dwyer, B.T., Dhar, S.K., Cvetic, C., Walter, J.C., and Dutta, A. (2000). Inhibition of eukaryotic DNA replication by geminin binding to Cdt1. Science *290*, 2309-2312. doi:10.1126/science.290.5500.2309.

Worthington, A.S., and Burkart, M.D. (2006). One-pot chemo-enzymatic synthesis of reporter-modified proteins. Org Biomol Chem *4*, 44-46. doi:10.1039/b512735a.

Wu, C.A., Zechner, E.L., Reems, J.A., McHenry, C.S., and Marians, K.J. (1992). Coordinated leading- and lagging-strand synthesis at the Escherichia coli DNA replication fork. V. Primase action regulates the cycle of Okazaki fragment synthesis. J Biol Chem *267*, 4074-4083. doi:10.1016/S0021-9258(19)50632-9.

Wuite, G.J., Smith, S.B., Young, M., Keller, D., and Bustamante, C. (2000). Single-molecule studies of the effect of template tension on T7 DNA polymerase activity. Nature *404*, 103-106. doi:10.1038/35003614.

Wutz, G., Ladurner, R., St Hilaire, B.G., Stocsits, R.R., Nagasaka, K., Pignard, B., Sanborn, A., Tang, W., Varnai, C., Ivanov, M.P.*, et al.* (2020). ESCO1 and CTCF enable formation of long chromatin loops by protecting cohesin(STAG1) from WAPL. Elife *9*. doi:10.7554/eLife.52091.

Wutz, G., Varnai, C., Nagasaka, K., Cisneros, D.A., Stocsits, R.R., Tang, W., Schoenfelder, S., Jessberger, G., Muhar, M., Hossain, M.J.*, et al.* (2017). Topologically associating domains and chromatin loops depend on cohesin and are regulated by CTCF, WAPL, and PDS5 proteins. EMBO J *36*, 3573-3599. doi:10.15252/embj.201798004.

Xi, J., Zhang, Z., Zhuang, Z., Yang, J., Spiering, M.M., Hammes, G.G., and Benkovic, S.J. (2005). Interaction between the T4 helicase loading protein (gp59) and the DNA polymerase (gp43): unlocking of the gp59-gp43-DNA complex to initiate assembly of a fully functional replisome. Biochemistry *44*, 7747-7756. doi:10.1021/bi047296w.

Yao, N.Y., Georgescu, R.E., Finkelstein, J., and O'Donnell, M.E. (2009). Single-molecule analysis reveals that the lagging strand increases replisome processivity but slows replication fork progression. Proc Natl Acad Sci U S A *106*, 13236-13241. doi:10.1073/pnas.0906157106.

Yardimci, H., Loveland, A.B., van Oijen, A.M., and Walter, J.C. (2012). Single-molecule analysis of DNA replication in Xenopus egg extracts. Methods *57*, 179-186. doi:10.1016/j.ymeth.2012.03.033.

Yatskevich, S., Rhodes, J., and Nasmyth, K. (2019). Organization of Chromosomal DNA by SMC Complexes. Annu Rev Genet *53*, 445-482. doi:10.1146/annurev-genet-112618-043633.

Yeeles, J.T., Deegan, T.D., Janska, A., Early, A., and Diffley, J.F. (2015). Regulated eukaryotic DNA replication origin firing with purified proteins. Nature *519*, 431-435. doi:10.1038/nature14285.

Yeeles, J.T., and Marians, K.J. (2011). The Escherichia coli replisome is inherently DNA damage tolerant. Science *334*, 235-238. doi:10.1126/science.1209111.

Yeeles, J.T., and Marians, K.J. (2013). Dynamics of leading-strand lesion skipping by the replisome. Mol Cell *52*, 855-865. doi:10.1016/j.molcel.2013.10.020.

Yeeles, J.T.P., Janska, A., Early, A., and Diffley, J.F.X. (2017). How the Eukaryotic Replisome Achieves Rapid and Efficient DNA Replication. Mol Cell *65*, 105-116. doi:10.1016/j.molcel.2016.11.017.

Yin, J., Lin, A.J., Golan, D.E., and Walsh, C.T. (2006). Site-specific protein labeling by Sfp phosphopantetheinyl transferase. Nat Protoc *1*, 280-285. doi:10.1038/nprot.2006.43.

Yin, Y.W., and Steitz, T.A. (2002). Structural basis for the transition from initiation to elongation transcription in T7 RNA polymerase. Science *298*, 1387-1395. doi:10.1126/science.1077464.

Yuan, Z., Bai, L., Sun, J., Georgescu, R., Liu, J., O'Donnell, M.E., and Li, H. (2016). Structure of the eukaryotic replicative CMG helicase suggests a pumpjack motion for translocation. Nat Struct Mol Biol *23*, 217-224. doi:10.1038/nsmb.3170.

Yuan, Z., Riera, A., Bai, L., Sun, J., Nandi, S., Spanos, C., Chen, Z.A., Barbon, M., Rappsilber, J., Stillman, B.*, et al.* (2017). Structural basis of Mcm2-7 replicative helicase loading by ORC-Cdc6 and Cdt1. Nat Struct Mol Biol *24*, 316-324. doi:10.1038/nsmb.3372.

Yuzhakov, A., Kelman, Z., and O'Donnell, M. (1999). Trading places on DNA-a three-point switch underlies primer handoff from primase to the replicative DNA polymerase. Cell *96*, 153-163. doi:10.1016/S0092-8674(00)80968-X.

Yuzhakov, A., Turner, J., and O'Donnell, M. (1996). Replisome assembly reveals the basis for asymmetric function in leading and lagging strand replication. Cell *86*, 877-886. doi:10.1016/s0092-8674(00)80163-4.

Zhai, Y., Cheng, E., Wu, H., Li, N., Yung, P.Y., Gao, N., and Tye, B.K. (2017). Open-ringed structure of the Cdt1-Mcm2-7 complex as a precursor of the MCM double hexamer. Nat Struct Mol Biol *24*, 300-308. doi:10.1038/nsmb.3374.

Zhang, Z., Spiering, M.M., Trakselis, M.A., Ishmael, F.T., Xi, J., Benkovic, S.J., and Hammes, G.G. (2005). Assembly of the bacteriophage T4 primosome: single-molecule and ensemble studies. Proc Natl Acad Sci U S A *102*, 3254-3259. doi:10.1073/pnas.0500327102.

Zhou, J.C., Janska, A., Goswami, P., Renault, L., Abid Ali, F., Kotecha, A., Diffley, J.F.X., and Costa, A. (2017). CMG-Pol epsilon dynamics suggests a mechanism for the establishment of leading-strand synthesis in the eukaryotic replisome. Proc Natl Acad Sci U S A *114*, 4141-4146. doi:10.1073/pnas.1700530114.

Zhu, W., and Depamphilis, M.L. (2009). Selective killing of cancer cells by suppression of geminin activity. Cancer Res *69*, 4870-4877. doi:10.1158/0008-5472.CAN-08-4559.