Chem Soc Rev

REVIEW ARTICLE

# DNA replication at the single-molecule level

Cite this: DOI: 10.1039/c3cs60391a

S. A. Stratmann and A. M. van Oijen*

A cell can be thought of as a highly sophisticated micro factory: in a pool of billions of molecules – metabolites, structural proteins, enzymes, oligonucleotides – multi-subunit complexes assemble to perform a large number of basic cellular tasks, such as DNA replication, RNA/protein synthesis or intracellular transport. By purifying single components and using them to reconstitute molecular processes in a test tube, researchers have gathered crucial knowledge about mechanistic, dynamic and structural properties of biochemical pathways. However, to sort this information into an accurate cellular road map, we need to understand reactions in their relevant context within the cellular hierarchy, which is at the individual molecule level within a crowded, cellular environment. Reactions occur in a stochastic fashion, have short-lived and not necessarily well-defined intermediates, and dynamically form functional entities. With the use of single-molecule techniques these steps can be followed and detailed kinetic information that otherwise would be hidden in ensemble averaging can be obtained. One of the first complex cellular tasks that have been studied at the single-molecule level is the replication of DNA. The replisome, the multi-protein machinery responsible for copying DNA, is built from a large number of proteins that function together in an intricate and efficient fashion allowing the complex to tolerate DNA damage, roadblocks or fluctuations in subunit concentration. In this review, we summarize advances in single-molecule studies, both *in vitro* and *in vivo*, that have contributed to our current knowledge of the mechanistic principles underlying DNA replication.

Received 1st November 2013

DOI: 10.1039/c3cs60391a

www.rsc.org/csr

## 1. Introduction

Life is as dynamic as its environment. Many key cellular processes cannot be described as outcomes from static associations of molecular components, but instead rely on an intricate spatial and temporal orchestration of many molecular players. For example, the conversion of chemical energy into mechanical work allows the transport of vesicles and molecules within the cytosol, along a membrane or between cells. On the single-molecule level, kinesins and other motor proteins move along the cytoskeletal filaments, transporter proteins shuffle metabolites between compartments, and multi-subunit complexes like replisomes, ribosomes, or the respiratory chain support an efficient maintenance and balancing of anabolism and catabolism.

Both fluorescence- and force-based single-molecule studies have provided fascinating new insights into some of these elaborate biological processes, such as cytoskeletal dynamics,[1–3] ATP synthesis,[4,5] RNA and DNA polymerization,[6–8] and viral packaging.[9,10] The more recent developments in live-cell single-molecule imaging allow us to record the cellular micro-management in real time, as has been demonstrated for example for transcription-factor dynamics,[11] protein-expression rates,[12] and signalling pathways[13] (reviewed to a greater detail in ref. 14).

*Zernike Institute for Advanced Materials, Centre for Synthetic Biology, University of Groningen, The Netherlands. E-mail: a.m.van.oijen@rug.nl*

What type of knowledge do we obtain from experiments monitoring individual molecules? Ensemble-averaging bulk assays provide information about the reaction rates of a pool of catalysts and, by synchronizing reactions, kinetic studies can reveal the first few transitions of a multi-step process. However, loss of synchronization due to the stochastic nature of chemical reactions will render it challenging to obtain kinetic parameters of short-lived intermediate states. Single-molecule studies capture the probabilities of reaction steps or conformational changes of an individual enzyme during any arbitrary point along a multi-step process and provide information on underlying heterogeneities in the dynamic behaviour of the population.[15,16] Watching individual reactions at work tells us not only about the stochasticity of consecutive pathways, but also about any temporal correlation: does an enzymatic reaction for example display non-markovian behaviour, *i.e.* are reaction steps affected by preceding paths?[17] One of the earliest single-molecule fluorescence studies demonstrated such a memory effect in a flavoenzyme: autocorrelations of on and off dwell times of the redox-cofactor FAD($H_2$) resolved heterogeneous kinetic rates, caused by conformational changes within the protein, that had been previously masked in bulk experiments.[16] Similarly, single-molecule analyses of the RecBCD helicase of *Escherichia coli* could decipher subpopulations or microstates of the enzymatic complex that differ in the velocity of DNA unwinding.[18] Here, conformational changes that

are adopted in the absence of the ligand/substrate ATP are ''memorized'' by the active RecBCD upon ATP addition and result in distinct rates of progression along the DNA template.

The actual chemical conversions in enzymatic reactions typically proceed on a sub-picosecond time scale. However, the limiting steps in catalysis are often the crossing of thermal activation barriers and the diffusive process necessary to mediate association of two reactants, which last orders of magnitude longer and are consequently the parameters to follow in single-molecule studies.[19,20] Typical fluorescence assays rely for example on visualizing a chromophore coupled to a molecule of interest and monitoring the appearance and disappearance of its signal as the labelled component is binding to and dissociating from a reaction partner molecule (Fig. 1). Binding lifetimes can be extracted and a probability distribution generated that contains the kinetic rates of the observed reaction. Several reviews on single-molecule enzymology provide excellent descriptions of enzymatic kinetics based on single-molecule reaction probabilities.[16,21]

In addition to successes in resolving single-protein kinetics, recent developments have focused on the visualization of protein dynamics and complex assemblies in real time, usually using fluorescence co-localization or fluorescence (Förster) resonance energy transfer (FRET) methods. These approaches have allowed, for example, the observation of the dimerization of EGF receptors in living cells, the complex formation of a reconstituted functional vesicle fusion construct of t- and v-SNARE proteins, or the Arp2/3-mediated branch formation on growing actin filaments.[13,22,23] Studies of the replisome, the machinery responsible for DNA replication, face the challenge of revealing the various and frequently transient interactions of the numerous enzymes that are involved.[24–26] The multi-component replisome is loaded on the DNA template in tight coordination with the cell cycle; it proceeds with a speed of up to thousand nucleotides per second (for certain bacterial systems), corrects wrongly incorporated nucleotides to an accuracy of about one mistake per $10^9$ nt and triggers repair processes upon detection of depurination, deamination, or pyrimidine-dimer formation.[27,28] Coordination of such a wide array of tasks, each on their own representing formidable molecular challenges, requires a finely tuned and balanced set of enzymatic activities. Building on the large base of knowledge we have on the individual components of the replication reaction, derived from many decades of genetic, biochemical and structural studies, single-molecule approaches represent a powerful approach to unravel the intricacies of how the various enzymatic activities at the replication fork are coordinated.

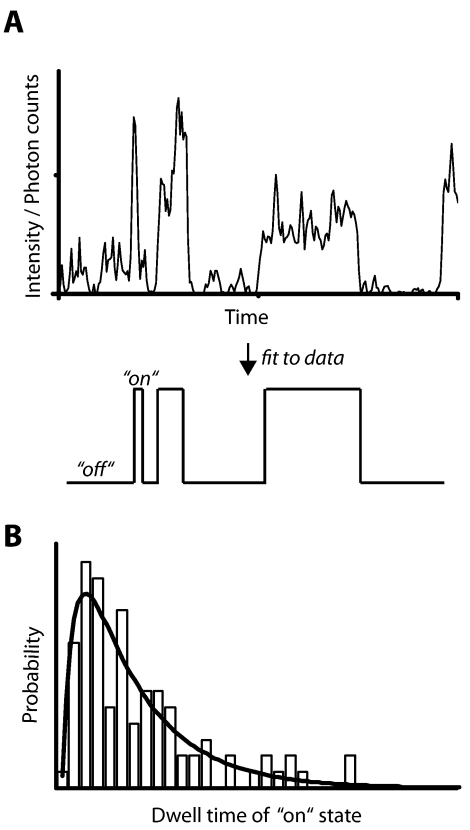



Fig. 1 Extraction of single-molecule kinetics from the observation of on and off times. These on- and off times can represent a variety of functional or structural transitions such as binding/unbinding, conformational transitions or chemical reactions. (A) On- and off times of an observed fluorescent emitter are recorded and the photon count per molecule is tracked over time and fitted to an appropriate function. (B) The time scales for on and off times are sorted in a distribution that provides the kinetic parameters of the individual reaction.

The process of replication needs to deal with a variety of molecular hurdles. For example, the antiparallel nature of the double-stranded DNA template imposes an asymmetry on the replication machinery, whose DNA polymerases can only synthesize in one particular direction. Besides the need for this asymmetric coordination, other obstacles have to be tackled, such as crowding effects and roadblocks caused by transcription-related processes and repair activities that take place simultaneously on the same DNA template. How exactly cells meet those challenges is a subject particularly well-suited for single-molecule studies – requiring methods to observe the spatiotemporal behaviour of individual molecules in a biologically relevant environment.

In this review, we describe recent developments in single-molecule research on the replisome *in vitro* and *in vivo*. First we highlight a section to the main technological developments in terms of microscope setups, design of fluorophores and labelling methods. Referring to the replication systems of the bacteriophages T7 and T4, of *Escherichia coli* (*E. coli*), and of eukaryotic cells, we guide the reader through the different aspects of important single-molecule studies that have contributed to a better understanding of the basic mechanics of DNA replication and organization.

## 2. Experimental strategies to image single molecules

Single-molecule techniques are typically categorized into two classes that we want to outline briefly: fluorescence microscopy allows the recording of the emitted photons of a fluorophore-labelled molecule of interest and is particularly applicable for detecting conformational changes within the protein of interest or its localization. Force-based measurement techniques, like atomic

force microscopy (AFM), magnetic tweezers, optical traps or flow-stretching setups, are useful in characterizing mechanical properties such as DNA topology or force exertion by motor proteins.

### 2.1 Getting proteins to shine

As early as the 1970s, it was demonstrated that single protein molecules labelled with a large number of dyes could be detected using an optical microscope:[29] Tomas Hirschfeld coupled roughly a hundred fluorescein dyes to a single antibody, swept a dilute solution of these constructs along a tightly focused laser beam, and observed bursts of fluorescence each corresponding to a single antibody. Not until two decades later, absorption and fluorescence measurements of single chromophores were successfully performed at cryogenic temperatures where absorption cross-sections are highest and photo-induced damage lowest.[30–32] Initially these studies were performed on doped molecular crystals, whereas later cryogenic single-molecule approaches were applied to study pigment–protein complexes.[33] Near-field scanning microscopy approaches demonstrated the feasibility to repeatedly image chromophores within biological samples at ambient temperature,[34] but were later joined by even more powerful and technically less-demanding far-field methods, mainly confocal and total-internal-reflection (TIRF) microscopy. Since then, great advances in high-sensitivity detection devices, in the engineering of photostable dyes and fluorescent proteins, and labelling strategies have pushed the sensitivity and resolution limits to a point where single molecules can be observed over timescales from milliseconds to minutes and down to spatial resolutions of a few nanometres. Furthermore, advances in live-cell imaging have enabled such experiments in a cellular context. Here, additional factors have to be considered in terms of cell viability (nutrients, $CO_2$, photodamage due to decomposition of fluorophores and radical release) and fluorophore choice (uptake, label selectivity and specificity). Even though these developments are relatively recent and many novel methods are still coming to fruition, single-molecule approaches are already revolutionizing the way mechanistic questions of biological systems are answered.

**2.1.1 Hardware technology.** The optical instrumentation required for single-molecule imaging and tracking can be roughly divided into two modes of operation: wide-field imaging and scanning confocal microscopy (Fig. 2). Both approaches have their advantages and need to be adapted to the actual question in consideration of both spatial and temporal resolution.

Wide-field imaging is a frequently used method to follow reactions at the single-molecule level in real time, *i.e.* to track particles and observe fast dynamics. In epifluorescence microscopy, a large sample volume is excited, limiting the signal-to-noise ratio in the region-of-interest. However, thin samples, either reconstituted isolated compounds or flat cells, can be analysed with single-molecule sensitivity, as shown for microtubule gliding on kinesins or live-cell protein expression.[35,36] Being proposed already in the 1950s but not fully developed until several decades later,[37,38] TIRF microscopy has proven to be exceptionally useful in improving signal detection. Here, at the coverglass/solution interface an evanescent field is induced that decays exponentially in the z plane and limits the excited volume to about 100 nm.

In confocal imaging, a diffraction-limited focus is positioned within the sample volume and scanned orthogonally to the optical axis.[39] The use of pinholes results in the selective detection of only in-focus fluorescence, while suppressing most out-of-focus background. In contrast to a TIRF setup, confocal microscopy allows the scanning of samples in three dimensions with a large penetration depth. However, the limiting factor is the scanning speed of the focal spot through the sample. Spinning-disk confocal setups employ a broad laser illumination that is focused by a large array of microlenses on a Nipkow disk, achieving high frame rates of up to 1000 frames per second.[40]

Non-linear two-photon techniques use optical sectioning as well, but here focussing relies on the probability of two-photon absorption, which is proportional to the square of the excitation intensity. A main advantage is that the required lower-energy wavelengths reduce photodamage of the fluorophores as well as scattering in tissue samples. Depths of several hundred micrometres are achievable with this method, as for example demonstrated in fascinating work on intact organs in living organisms.[41]

Technological developments in optical microscopy have contributed to a gradual improvement in spatial resolution, but with the size of the smallest resolvable structures still similar to the diffraction limit. The recent breakthroughs in super-resolution imaging, however, have allowed the imaging

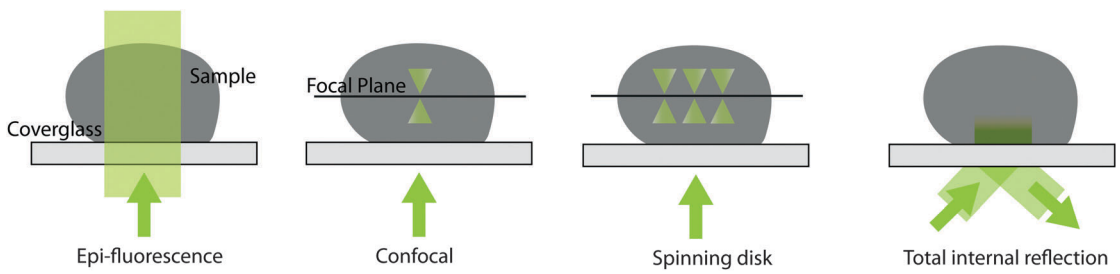



Fig. 2 Fluorescence microscopy designs frequently used in single-molecule studies. In epifluorescence microscopy, the light source illuminates the entire sample. In confocal microscopy, a pinhole is used to illuminate specifically the focal plane, thus reducing background fluorescence. By installing a Nipkow spinning disk, the sample is illuminated at multiple points within the focal plane simultaneously. In total-internal-reflection fluorescence (TIRF) microscopy, the incident laser is reflected from the coverglass surface, creating an exponentially decaying evanescent field on top of the surface, which reduces the thickness of the illuminated volume to about 100 nm.

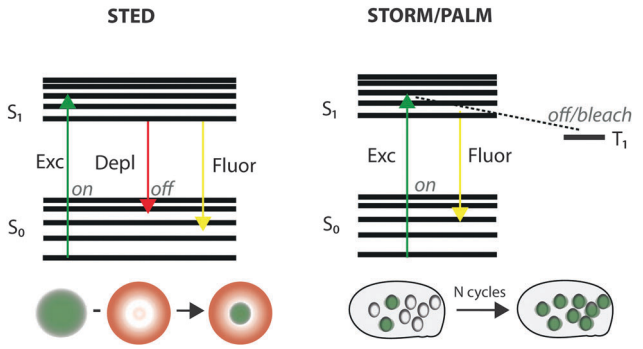



Fig. 3 Super-resolution techniques. In STED, fluorophores are excited to the $S_1$ state (green) and return to the ground state $S_0$ spontaneously while emitting photons (yellow). An intense red-shifted doughnut-shaped depletion laser beam (red) forces molecules into the ground state without them emitting fluorescence. As a result, only a sub-diffraction-limited area in the center of the depletion laser remains in the excited state and will be observable through the emission of a yellow fluorescent photon. A similar excitation geometry is used in ground state depletion (GSD) microscopy. However, instead of rendering the fluorophores around a point of interest nonfluorescent by depleting the fluorescent excited state, they are brought into a long-lived dark state. In PALM and STORM, molecules are switched on at low spatial densities, their positions determined with sub-diffraction-limited precision, and irreversibly photobleached. Repeating this procedure for a large number of molecules results in sub-diffraction-limited images (adapted from ref. 191).

of fluorescently labelled structures down to length scales that are an order of magnitude smaller than the diffraction limit. Super-resolution methods find their basis in the reduction of the point-spread function (PSF) in excitation, as in stimulated emission depletion (STED), ground-state depletion (GSD) or structured illumination microscopy (SIM), or in the modulation of the fluorophore's emission, as in photoactivated localization microscopy (PALM) and stochastic optical reconstruction microscopy (STORM) (Fig. 3). STED microscopy is based on the illumination of the sample with a doughnut-shaped beam profile. The excitation beam is narrowed by an overlaying ring-shaped longer-wavelength depletion beam, that forces the dyes into the ground state.[42] The higher the intensity of the depletion laser, the narrower the PSF becomes. In a similar design, but typically using only one wavelength, GSD brings the fluorophores to their lowest triplet dark state in the outer ring. An alternative approach to super-resolution imaging is enabled by the wide-field methods PALM and STORM, utilizing the stochastic activation of fluorophores that are photoactivatable or photoswitchable. The activation of a few fluorophores in the field of view allows each of them to be individually imaged and to be fit by a two-dimensional point-spread function and thus each of their centroid positions to be obtained with sub-diffraction-limited precision. It is the sum of several cycles of activation–centroid detection–bleaching/inactivation that leads to the reconstruction of the complete object of interest. The super-resolution techniques PALM and STORM have also played important roles recently in resolving intracellular dynamic processes at the single-molecule level (*e.g.* ref. 43 and 44).

**2.1.2 Fluorophore technology.** One of the major challenges in modern fluorescence microscopy is the engineering of appropriate dyes and the specific attachment to the biomolecule of choice. The properties required of chromophores for single-molecule imaging are demanding: the photostability in terms of lifetime and (absence of) blinking must be high, the conjugated molecular structure must be soluble and stable, and the fluorophore's dimensions and physicochemical properties should not interfere with protein conformations and function. For sub-nanometre tracking, high quantum yields and large Stokes shifts are especially important.

In general, three categories of probes can be differentiated: fluorescent proteins, organic dyes and quantum dots. Being genetically encoded as fusion constructs to the protein of interest, fluorescent proteins are labels with absolute specificity and represent a standard approach for *in vivo* imaging. Limitations are their photostability and brightness, as well as the bulkiness of the 25 kDa structure that potentially interferes with enzyme functionality. Organic dyes are significantly smaller and often display better photophysical properties. The commercial availability of dyes is enormous; brightness, stability, and solubility can be chosen with great flexibility. The major bottlenecks are the specificity end efficiency of the labelling chemistry and, for *in vivo* studies, the need for electroporation or alternative methods to introduce the dyes into the cell. Finally, quantum dots, fluorescent nanocrystals of 5–20 nm diameter, can be engineered in highly sophisticated ways, with extinction coefficients several times higher than those of organic dyes. Extremely high brightness and the resultant high signal-to-noise ratios allow nanometre-tracking of individual molecules.[45] The large size, however, can influence the mobility and conformational flexibility of the labelled protein.

*Fluorescent proteins.* Fluorescent proteins (FPs) consist of a rigid β-barrel composed of 11 β-sheets that surround a central α-helix containing the chromophore.[46] The naturally occurring variants have been extensively tuned in terms of brightness, emission range, photostability, monomeric character, and maturation rate.[47,48] Due to the strong autofluorescence of endogenous cellular fluorophores (flavins, NADH, amino acids) at wavelengths below 500 nm, the development of red-shifted FPs is one central consideration for *in vivo* imaging.

Newly developed classes of FPs with photoconvertible or photoswitchable chromophores allow super-resolution imaging even in a high-concentration environment, as only a limited fraction within the excitation field is switched on. Photoconvertible FPs such as Kaede, KikGR, Dendra and Eos are subject to a peptide-backbone cleavage step when illuminated with a 405 nm laser, leading to an enlargement of the conjugated system by an additional imidazole ring, which corresponds to a green-to-red shift in fluorescence.[48,49] The photoswitchable FP Dronpa has an excitation maximum at 503 nm, and can be switched off and on several times by strong 488 nm and weak 405 nm illumination, respectively. Alternatively, the green fluorescent Padron is switched on by blue excitation and off by UV light. The combination of those opposite switching behaviours allows two-color tracking in live cell imaging.[50] In terms of photochemistry, crystal structures of Dronpa suggest that the *cis–trans* isomerization and protonation of the chromophore are responsible for the different fluorescent states.[51]

Further progress in the design of fluorescent protein tags, especially far-red fluorescent as well as switchable probes, in combination with novel microscopy techniques will continue to provide powerful tools for *in vivo* imaging.

*Organic dyes and their coupling to proteins.* The main challenge in the use of organic dyes is a highly efficient and specific labelling reaction to the target protein. Several strategies exist for selective chemical tagging that can be basically subdivided into the introduction of a protein domain, a short peptide or a unique amino acid.[52]

A successful method to specifically couple an organic dye to a protein is the fusion to a target protein of an additional protein domain that itself binds the organic dye tightly and selectively. Prominent protein-domain fusion constructs are the commercially available dehalogenase and alkylguanosine transferase tags (HaloTag and SNAP tag, respectively). The HaloTag technology takes advantage of a self-labelling step of a 33 kDa-sized dehalogenase enzyme. The reaction catalysed by this enzyme consists of (1) a nucleophilic displacement of a halide ion from an alkane chain that is transferred to an aspartate residue, (2) histidine catalysed hydrolysis, finally regenerating the aspartate. Mutagenesis of the active-site histidine residue locks the dehalogenase in step 1, allowing specific labelling with a customized fluorescent alkane moiety.[53] The 20 kDa sized O6-alkylguanine-DNA alkyltransferase (hAGT) enzyme transfers an alkyl group from guanosine derivatives to its active site cysteine residue, allowing for the subsequent covalent coupling of alkyl-modified fluorophores.[54]

Smaller peptide tags are particularly advantageous when internal labelling positions are required. The Tsien lab developed a biarsenic tagging technology that depends on the high affinity of thiols to arsenic.[55,56] The probes 4′,5′-bis(1,3,2-dithioarsolan-2-yl)fluorescein (FlAsH) and the chemically similar resorufin-based ReAsH are non-fluorescent when bound to ethane dithiol (EDT), but fluoresce green and red, respectively, when a tetracysteine sequence CCXXCC replaces EDT. Another strategy relies on the incorporation of an aldehyde tagging peptide sequence LCTPSR into the target protein.[57] A co-expressed formylglycine-generating enzyme converts the cysteine's thiol group into an aldehyde that specifically reacts with hydrazide-functionalized molecules to produce a hydrazone. Other self-labelling tags are the hexa-histidine peptide or the Texas-red-binding aptamer, chelating with Ni–NTA-derivatized fluorophores or binding the Texas-red fluorophore with nano- to picomolar binding affinity.[58,59]

Cysteines are usually less abundant in proteins and due to their high reactivity towards maleimide thioesters they are a popular target for *in vitro* labelling. If cysteine mutagenesis is not favourable because of limitations related to protein functionality, the introduction of unnatural amino acids, as pioneered by the Schultz lab, represents an alternative approach. Co-expressed orthogonal tRNAs and aminoacyl tRNA synthetases incorporate a range of unnatural amino acids in response to amber stop codons or quadruplet codons.[60–63]

Despite the intrinsic bottleneck of selectivity in labelling, the advantage of organic dyes lies in the nearly unlimited options for fluorescence characteristics. Not only are dyes available that cover the entire spectral range, but also many fluorescent compounds have been developed with properties that can be externally modified by optical inputs. For example, caged chromophores can be activated by UV light, and several cyanine dyes can be coupled to construct activator–reporter FRET pairs.[64,65]

### 2.2 Trapping and pulling at individual DNA molecules

Force spectroscopy methods are frequently applied for characterizing mechanical properties of biomolecules at the single-molecule level, such as topological changes in DNA molecules or force exertion by individual motor proteins. In the context of studying DNA replication at the single-molecule level, such techniques are often used to stretch the DNA substrate and to probe the mechanical consequences of replication on the DNA (conversion between single- and double-stranded DNA,[8,66] change in supercoiling[67]), or to observe the motion of proteins along DNA.[68,69] Detailed reviews about the instrumental designs can be found elsewhere.[70–73]

In trapping techniques, one end of the biomolecule of choice is stably attached to a surface and the other one trapped with a magnetically or optically controlled bead or an AFM tip. Optical tweezers trap dielectric beads within a focused laser beam. The electromagnetic field polarizes the particle that is forced into the steep gradient at the focal spot. Spatial resolutions of down to 0.1 nm with sub-millisecond time resolutions are feasible by applying forces of about 0.1 to 100 pN.[70] Magnetic traps have a slightly lower spatial resolution of about 2 to 10 nm, can apply forces over a large range from pico- to nano-Newtons, and therefore are particularly useful in the measurement and manipulation of DNA topology. By attaching DNA on one end to a surface and on the other to a paramagnetic bead, the polymer is constrained and can be accurately controlled and placed in a particular topological conformation with defined twist and writhe.[74] Prominent topoisomerase experiments are performed on magnetically manipulated plectonemic DNA, as the ATP-dependent double-strand breaks remove two turns, thus changing the linking number by two.[75] Flow-stretching techniques rely on the hydrodynamic dragging of one-end anchored polymers in a microfluidic device. DNA-bead tethers are for example useful in tracking length changes of the molecule during the time course of replication[76] (Fig. 4).

Combining the strengths of fluorescence imaging and mechanical approaches, recent developments have allowed the observation of DNA-based single-molecule fluorescence while exerting well-defined stretching forces on the DNA template.[77,78] For example, Holliday-junction recombination events and conformational changes could be followed by creating FRET pairs within the four-stranded complex, tethered to an optical trap.[79] The angstrom resolution of FRET signals combined with sub-pN forces in the optical trap established a highly controlled system for controlling and following conformations of DNA structures. Such hybrid techniques, allowing both the tracking of fluorescent molecules and the detection of the chemomechanical reactions, hold tremendous power in understanding the many facets of multi-protein machineries acting on DNA.

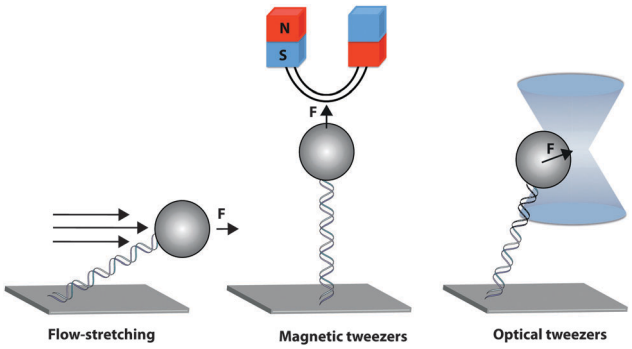


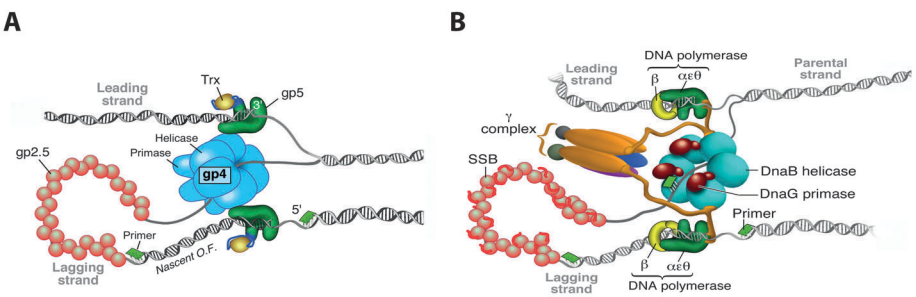

Fig. 4 Force manipulation setups. DNA molecules are attached on one side to the coverglass surface and coupled to a bead that is stretched in a hydrodynamic flow, or trapped magnetically or optically.

## 3. Replication machineries

Genomic DNA replication consists of three distinct phases: initiation, elongation and termination. The complexity of cell-cycle timing, its coupling to DNA synthesis, and in general the molecular details of DNA synthesis vary tremendously amongst the taxonomic domains. However, the main principles of the replication machinery are conserved: ring-structured replicative helicases encircle single-stranded DNA and couple the energy released from nucleotide hydrolysis to directional movement. The subsequent unwinding of the DNA provides a template for polymerases to synthesize the daughter strands by catalysing the coupling of an incoming nucleotide to the ribose 3′ hydroxyl group of the previously incorporated nucleotide. All known polymerases display this requirement of directionality: only DNA synthesis from the 5′ to the 3′ end allows for a continuation of synthesis (accompanied by the backwards removal of incorrectly incorporated nucleotides). With the antiparallel nature of double-stranded DNA, such a directional requirement for DNA synthesis results in a picture in which DNA is synthesized continuously on the so-called leading strand, with the leading-strand DNA polymerase acting in the same direction as the helicase is moving, and with the lagging-strand DNA polymerase polymerizing in a discontinuous fashion, giving rise to short stretches of DNA named Okazaki fragments (Fig. 5A and B).

A special class of polymerases known as primases synthesize short oligo-ribonucleotide primers that are used as starting template for the lagging-strand DNA polymerase. The timing of the enzymatic steps at the lagging strand, *i.e.* priming, utilization of the primer by the polymerase and its extension into an Okazaki fragment, is of importance for the orchestration of a coupled replication reaction: a process in which continuous synthesis on the leading strand is tightly coordinated with the discontinuous synthesis on the lagging strand.

Research on the replisome of the bacteriophage T4 initiated the idea of the trombone model that reconciles a symmetric replication fork, containing two DNA polymerases moving in the same direction, with the underlying asymmetry of the DNA template.[80] The formation of a looped structure in the lagging strand reorients the polymerase while synthesizing an Okazaki fragment, until a release event triggers the recycling of the polymerase to the next Okazaki fragment. The formation of a DNA loop and the close proximity of the lagging-strand DNA polymerase to the replisome




C

| Protein class | T7 | T4 | *E. Coli* | Eukaryotes |
|---|---|---|---|---|
| Polymerase | gp5 | gp43 | αεθ | α, δ and ε |
| Clamp/ processivity factor | thioredoxin | $gp45_3$ | β | PCNA |
| Clamp loader | - | $gp44_4/62$ | γ/τδδ′χψ | RFC |
| Helicase | $gp4_6$ | $gp41_6$ | $DnaB_6$ | MCM2-7/CMG |
| Helicase loader | - | gp59 | DnaC | ORC licensing complex |
| ssDNA binding protein | $gp2.5_2$ | gp32 | $SSB_4$ | RPA |
| Primase | $gp4_6$ | gp61 | DnaG | Primase |

Fig. 5 Replisome proteins and fork architecture in viruses, bacteria and eukaryotes. (A) The replication fork of bacteriophage T7. Two polymerases gp5, each associated with an *E. coli* thioredoxin molecule, bind to the hexameric helicase–primase gp4. The primase domain of gp4 synthesizes short ribonucleotide primers that are handed over to the lagging-strand polymerase for elongation into Okazaki fragments. The unwound single-stranded regions of the template DNA are covered by gp2.5 proteins (adapted from ref. 66). (B) The replication fork of *E. coli*. Two copies of the DNA polymerase holoenzyme are associated with β clamps on the leading and lagging strand. Three DnaG molecules associate with the DnaB helicase to synthesize primers on the lagging strand that is partly covered by SSB tetramers (adapted from ref. 192). (C) Comparison of the replisome components in phage, *E. coli* and eukaryotes.

result in a short travel distance to the next primer after completion of Okazaki fragment synthesis by the polymerase. The presence of replication loops is supported by several lines of evidence obtained from bacteriophage replication machineries generating Okazaki fragments of about 1000 to 2000 bp. In eukaryotes however, the much shorter Okazaki fragments (100 to 200 base pairs) make such a looping scenario less likely and certainly more difficult to observe.

In addition to the mechanistic demands placed on replication due to the antiparallel nature of duplex DNA, copying genomic stretches of DNA inside the cell comes with several other molecular challenges. Roadblocks such as nucleosomes need to be dealt with, the topology of the DNA needs to be controlled, and replication needs to be regulated and coordinated with other cellular activities such as DNA repair and recombination. As will be laid out in the remainder of this review, single-molecule biophysical techniques have begun to significantly contribute to our understanding of the molecular aspects of each of these processes. We will illustrate these efforts by starting with simple replication model systems, focusing on only the activities at the fork, followed by zooming out and considering the interplay of replication with topology, nucleosomes and the overall cell cycle.

### 3.1 Model systems for single-molecule studies

The main operating principles of the replisome are highly conserved across phages, bacteria and eukaryotes (Fig. 5), although the involved enzyme classes are structurally not necessarily homologous. Replication complexes that are well understood in terms of their composition, assembly and functioning are the ones of the bacteriophages T7 and T4, as well as that of *E. coli*. The much higher complexity of the eukaryotic replisome and of the cell-cycle checkpoints that regulate the start and progression of replication still requires further biochemical research in order to completely model the process of DNA duplication.[81,82] In the following sections, we will discuss briefly the biochemical properties of these systems, before focussing on single-molecule studies.

**3.1.1 Bacteriophage T7.** As one of the simplest replication machineries in terms of the number of proteins involved, the bacteriophage T7 replisome has proven to be a powerful platform to study the coordination of leading and lagging-strand synthesis, both at the ensemble and single-molecule level. Only four proteins (Fig. 5A) are needed to assemble a replication fork that proceeds with high processivity and stability, while also exhibiting remarkable dynamics in its interactions and composition. The DNA helicase–primase gene product 4, gp4, is responsible for both DNA unwinding and RNA primer deposition on the lagging strand. The N-terminal half of this bifunctional protein supports the primase activity. Faced away from the ds–ssDNA junction, the N-terminal zinc-binding domain (ZBD) scans the single-stranded lagging strand as it is extruded by the C-terminal helicase domain. After recognition of a signal sequence, a tetraribonucleotide primer is synthesized by the RNA-polymerase domain.[83] The ZBD remains associated with the primer and hands it off to the lagging-strand polymerase.[84] The C-terminal helicase domain of gp4 hydrolyses dTTP to translocate along ssDNA in 5′ to 3′ direction and displaces the complementary strand to unwind dsDNA. Gp4 exists as a hexamer as well as a heptamer in solution, but functions on ssDNA in its hexameric conformation.[85,86] As the T7 replisome lacks a helicase-loading protein in comparison to other systems (see below), it is hypothesized that the loss of one subunit facilitates the loading mechanism.[86] Alternatively or concomitantly, a loading site within the primase domain that interacts with the DNA may participate in the ring-opening mechanism required for loading on DNA.[86,87] The T7 DNA polymerase, a complex of gp5 with the *E. coli* thioredoxin protein as a processivity factor, synthesizes new DNA with one copy of the complex on the leading strand and one on the lagging strand. Gp5 on its own displays a processivity of only about 80 nt, but when bound to thioredoxin with a very low $K_d$ of 5 nM,[88] its binding lifetime to the primer–template, and thus its processivity, is increased ten-fold.[89,90] The activities of gp4 and gp5, unwinding and synthesis, are highly synergistic, so that a fully reconstituted T7 replisome achieves a processivity of >17 kbp in leading-strand synthesis, while Okazaki fragments are generated in the lagging-strand loops approximately every 1–2 kbp.[91] Finally, the ssDNA-binding protein gp2.5 binds and protects the transiently exposed single-stranded DNA on the lagging strand.[91,92] Beyond this classical ssDNA-binding role, gp2.5 is also important in mediating protein–protein interactions and regulating hand-off events at the replication fork.[93,94]

**3.1.2 Bacteriophage T4.** After its initial reconstitution *in vitro* by Alberts and coworkers in 1975,[95] the bacteriophage T4 replisome has been one of the most intensively studied replication systems. Detailed knowledge exists of the various protein structures, protein-interaction sites and enzyme kinetics, together forming an ideal basis for biophysical studies. A key property of the T4 system is its conceptual similarity to the replication systems of higher-order replisomes: like these, it contains ring-shaped clamp proteins that anchor the polymerases at the fork, clamp-loader proteins and helicase-loader proteins. The lower complexity, however, in terms of the total number of involved proteins or the regulation of replication initiation, has allowed the manipulation and study of its molecular mechanisms by single-molecule approaches.

The T4 replisome is composed of eight proteins (Fig. 5C), subdivided into the primosome (gp41 helicase, gp59 helicase loader, gp32 ssDNA-binding protein, gp41 primase) and the replicase/holoenzyme (gp43 polymerase, gp45 clamp, gp44/62 clamp loader).[96,97] The hexameric helicase loader has a high affinity for gp32-coated DNA segments at replication forks and coordinates the loading of the hexameric helicase.[98] Equimolar amounts of helicase loader and helicase were shown to be favourable for the helicase unwinding activity, pointing to a 1 : 1 binding stoichiometry,[99] analogous to the DnaB–DnaC complex in *E. coli*, as described below. The primase gp61 associates with the helicase on the lagging strand and synthesizes pentaribonucleotide primers to initiate Okazaki-fragment synthesis.[100,101] Reminiscent of the fused helicase–primase T7 gp4, gp61 shows maximal priming activity when present in a 6 : 1 molar ratio with the hexameric helicase.[102] As for T7,

most likely a primer hand-off mechanism to either the ssDNA-binding protein or the polymerase exists that prevents the primer from melting.[103] On both DNA strands, the polymerase gp43 associates with a trimeric sliding processivity clamp gp45 that prevents it from falling off the template and that is loaded by the gp44/62 clamp-loader complex.[104] This pentameric complex is required to break up the ring-shaped clamp in order to thread the double-stranded DNA through the clamp opening at the primer–template hybrid segment. The clamp-loader complex belongs to the class of AAA+ (ATPases Associated with diverse cellular Activities) proteins. However, in comparison to most other AAA+ enzymes that are hexameric, clamp loaders display one open interface instead, and form a spiral-like structure allowing access to the DNA-binding substrate. The loader works as a molecular switch: in its ATP-bound form it has a high affinity for the open homotrimeric clamp, but in its ADP bound or empty conformation this affinity is loosened.[104] Once fully assembled, the T4 replisome proceeds up to 20 kbp along the template with a velocity of about 250 nt $s^{-1}$.[105]

**3.1.3** ***Escherichia coli.*** A better understanding of not only replication elongation but also initiation and termination is made possible by the study of replication in single-cell model organisms such as *E. coli*. While still much simpler than the eukaryotic replication system, *E. coli* has to employ similar strategies in its ability to control the starting and ending of replication. Further, it also relies on efficient methods to deal with DNA lesions and resolve topological structures.

To initiate the formation of a replication fork, the initiator protein DnaA assembles at a unique origin containing a 245-bp long specific sequence, known as the oriC locus. The oriC consists of five 9 bp-DnaA boxes and three AT-rich 13-bp segments, the DNA-unwinding elements, that melt upon DnaA binding[106–109] (Fig. 6A). DnaA is a DNA-dependent AAA+ family ATPase that oligomerizes upon DNA binding and induces origin unwinding driven by ATP hydrolysis,[110–112] possibly *via* inducing locally negative supercoiling in the AT-rich segments. Histone-like proteins (HU/IHF) support the separation of the two strands at the replication fork by stabilizing DNA bending.[113] DnaA recruits the prepriming complex, composed of the hexameric helicase DnaB and its loader protein DnaC, along the unwound DNA region. DnaC interacts with both DnaB and DnaA and allows the helicase loading on both sides of the asymmetric replication bubble.[114] Like many other regulatory proteins, DnaC is a dual switch AAA+ protein – the ATP-bound form preferentially binds to ssDNA and inhibits DnaB unwinding activity. DnaB association triggers hydrolysis, and the formation of DnaC–ADP provides the starting signal for fork progression.[115]

During the initiation process, the polymerases are loaded at the replication fork to finally start elongation (Fig. 5B). *E. coli* expresses at least five different polymerases, specialized to support either replication (Pol III), Okazaki-fragment maturation (Pol I), repair (Pol I, II), or translesion synthesis (II, IV and V). The replicative Pol III is a multi-subunit complex, assembled from ten different proteins (Fig. 5C).[116] The core enzyme consists of the polymerase α, the 3′–5′ proofreading exonuclease ε, and θ, which stimulates the exonuclease activity. The holoenzyme includes the dimeric β clamp and the clamp-loader complex, either $\gamma\tau_2\delta\delta'\chi\psi$ or $\tau_3\delta\delta'\chi\psi$, with the τ subunits binding to α, dimerizing the core and thus being critical for dissociation of the polymerases.[117] The χ subunit binds to the single-stranded binding protein SSB on the lagging strand, and ψ bridges χ and γ. DnaG primase coordinates Okazaki fragment initiation

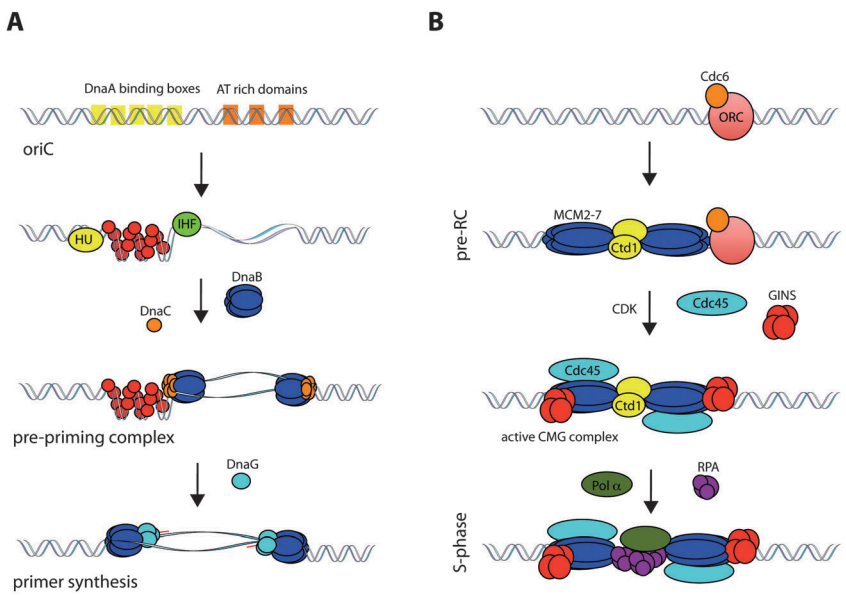



Fig. 6 Replication initiation. (A) *E coli*. DnaA oligomerizes at the oriC locus and recruits DnaB–C complexes to the unwound region. DnaG molecules associate with DnaB and prime synthesis of the daughter DNA strands. (B) Eukaryotes. The DNA-bound origin-recognition-complex (ORC) recruits Cdc6, Ctd1 and MCM2-7 to assemble the pre-replication complex (pre-RC). Upon MCM2-7 phosphorylation by CDKs and association with GINS and Cdc45, the active Cdc45–MCM2-7–GINS (CMG) complex unwinds the template DNA. RPA molecules protect the single-stranded region and Pol α primes the polymerization elongation reaction.

and interacts directly with DnaB at the replication fork.[118–120] The replication forks proceed around the circular chromosome until encountering each other again at the termination (ter) sites. Once there, they are sterically blocked by Tus proteins, that are tightly bound to the Ter sites and inhibit DnaB unwinding activity, finally resulting in the disassembly of the replisome.[121,122]

**3.1.4 Eukaryotic systems.** After having established the salient properties of phage and bacterial replisomes, an enormous amount of progress has recently been made in deciphering the molecular mechanisms underlying eukaryotic replication. However, in comparison to the previously described model systems, our understanding of eukaryotic replication is still far less complete. Not only the exact composition of the eukaryotic replication machinery remains unclear, but also the regulation of the replication reaction in terms of posttranslational modifications like PCNA ubiquitination or cell-cycle checkpoints is complicated and challenging to address with classical biochemical approaches.[123–125] Additionally, the details of the structural arrangement of chromosomes need to be considered as parameters that influence replication initiation and regulation. For example, histones have to be displaced during unwinding, but replaced onto the nascent DNA strands to preserve epigenetic information.[126]

Due to their size, each eukaryotic chromosome contains a large number of replication origins, onto which the replication initiation complexes assemble. To assure that each origin can act as a site of replication initiation maximally once per cell cycle, a licensing process starts in the G1 phase.[127,128] A pre-replication complex (pre-RC) is assembled on each of the origins in a process that is started by the binding of the hexameric origin-recognition complex ORC[129] (Fig. 6B). The ORC recruits first the cell division cycle proteins cdc6/cdc18 and cdt1, followed by the heterohexameric helicase MCM2-7. During the following S phase, these pre-RCs can be used as a platform to recruit polymerases, a primase, and numerous other replication factors to assemble a functional replisome. Once phosphorylated by cyclin-dependent protein kinases (Cdks),[130,131] the MCM2-7 hexamer associates with the cofactors cdc45 and GINS to form the actively unwinding CMG complex.[132–134] The heterotrimeric replication protein A (RPA) functions as ssDNA-binding protein, coating the lagging strand during fork progression.[135,136] During replication, the polymerases have to be switched according to their catalytic properties: the Pol α–primase complex synthesizes 7–10 nt long RNA primers and extends these by about 15 deoxynucleotides, before the pentameric replication factor C (RPC), analogous to the *E. coli* clamp-loader complex, displaces Pol α and hands the template over to the lagging-strand DNA polymerase Pol δ, whereas Pol ε most likely acts on the leading strand.[137] The trimeric PCNA (proliferating cell nuclear antigen) fulfils similar tasks to the bacterial β clamp, increasing the polymerase's processivity.[138,139]

Replication is regulated in accordance with the cell-cycle signalling. Cdks phosphorylate their target proteins, either activating them, like the MCM subunits, directly inactivating them, or labelling them for proteolytic degradation, like cdc6/cdc18.[140,141] In this way, secondary loading events at the origin sites are prevented and it is ensured that DNA is only copied once during every cell cycle.

As described above, the processes of replisome assembly and fork progression are highly dynamic, but tightly coordinated. Biochemical studies characterized the basic replication architecture, as well as enzymatic activities of the isolated components. This body of knowledge on function and structure has been critical to allow the single-molecule studies that we outline in the following sections.

### 3.2 Replication-fork assembly pathways

Replication initiation follows a concerted pathway that needs to result in the establishment of a complete fork, before allowing the polymerases to start the elongation process in a coupled manner. As described above, in T4 and higher-order replication systems, proteins with loading function participate in the assembly process, namely the helicase loader triggers a transient helicase opening to encircle the DNA template, and similarly the clamp loader positions the sliding clamps that increase polymerase processivity. With bulk assays such directed pathways are challenging to resolve. Instead, using single-molecule FRET microscopy, Benkovic and coworkers studied the T4 primosome assembly on forked DNA substrates. Fluorescently labelled replication proteins were loaded onto short artificial DNA forks attached to a coverglass surface, and imaged with total internal reflection fluorescence. Strong FRET signals could be observed between the donor–acceptor pair ssDNA-binding protein gp32–helicase loader gp59, indicating the formation of a tight complex at the single-to-double-strand DNA template fork. Adding gp41 helicase to the reaction did not interrupt this association, as long as no ATP substrate was present. However, active ATP hydrolysis by the helicase led to a displacement of the gp32–gp59 complex.[142] According to these data, it is the associated form of gp32 and gp59 at the fork that presents the landing platform for the helicase and triggers its unwinding activity, upon which the helicase loader gets released from the DNA. In a recent study, surface-attached forked DNA labelled with internal cyanine FRET dyes in the double-stranded region was used as a substrate for helicase–primase (primosome) complex formation, representing the next step in replication initiation.[143] Upon loading, the helicase migrated along the DNA template and opened the double helix, leading to dsDNA FRET fluctuations in the process of ''DNA breathing'' and ultimately to the loss of the FRET signal after complete unwinding. The addition of primase stimulated the unwinding activity, with the strongest effect in the case of a 1 : 6 primase : helicase subunit stoichiometry. However, without primase being present, the binding and processivity of the helicase were diminished, as well as when helicase and primase were pre-incubated before being loaded onto the DNA template. These results indicate that after successful helicase loading, helicase–DNA interactions are weak and that primase acts at the interface to stabilize the complex, supporting more than just priming activity. The authors proposed that a primase molecule bridges two helicase subunits, at the location where the NTP is binding. Subsequently, primase activates NTP hydrolysis by the helicase resulting in a transient release of the primase and a rotation of the helicase by one subunit towards the dsDNA fork.[100]

The primase–helicase interaction is crucial in the course of replication, as it determines the rate and coupling of leading and lagging-strand synthesis. The T4 primosome studies discussed above suggested that a single primase molecule per helicase hexamer is sufficient for the formation and stabilization of a primosome complex. However, the reconstitution of the complete replication fork is necessary to obtain information about the number of primase molecules within the replisome during a coupled replication reaction. Bulk assays gave lines of evidence for a multimeric primase organization within the replication fork of T4.[144] Stoichiometry measurements of the isolated *E. coli* DnaB hexameric helicase–DnaG primase also indicated the presence of several primases per helicase hexamer, namely three molecules,[120] suggesting a mechanistic need for a cooperatively functioning multimeric primase complex.

The visualization at the single-molecule level of the subsequent steps of loading of the gp43 DNA polymerase at the fork further increased our understanding of T4 fork assembly. By using single-molecule FRET imaging and reconstituting the primosome–holoenzyme assembly pathway on a forked DNA template *in vitro*, the mechanism of ordered association of fluorescently labelled enzymes was demonstrated that prevents premature replication initiation by the leading-strand polymerase before the helicase is loaded.[145] In the initially assembled complex of the helicase loader and the polymerase on a forked DNA substrate, the helicase loader locks the polymerase in an inactive state. The loading of the helicase likely disrupts this complex, displaces the loader and forms the functional leading-strand replisomal complex with the polymerase (and the sliding clamp) (Fig. 7). Taken together, these FRET-based studies provide a model of T4 assembly, which consists of (1) binding of the gp59 helicase loader on a DNA fork that is coated with the gp32 ssDNA-binding protein; (2) gp43/gp45 (polymerase/clamp) loading and interaction of the gp43 polymerase with the gp59 helicase loader, that blocks polymerization activity; (3) gp41 helicase loading and ATP-hydrolysis dependent disassembly of the gp59 helicase loader, (4) association of the gp61 primase with the gp41 helicase. These single-molecule experiments on T4 replication initiation are a beautiful example of how on/off switching of enzymatic activity can be accomplished by protein–protein interactions.

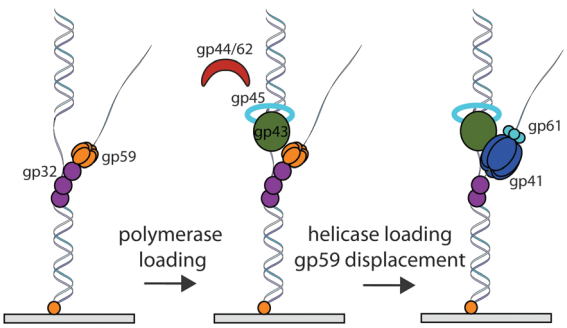



Fig. 7 T4 replication initiation. Benkovic and coworkers used forked DNA substrates coupled to a microscope coverglass surface to reconstitute the initiation pathway of T4.[142,145,193] Gp32 and gp59 bind to the fork structure and recruit the polymerase gp43 together with the clamp gp45 that is installed on the DNA strand by the loader complex gp44/62. Gp59 catalyses loading of the helicase gp61 and is displaced by the latter upon DNA unwinding. Primase molecules (gp61) associate with the helicase to form a stable primosome complex.

### 3.3 Leading and lagging-strand coordination

Achieving coordinated replication in an asymmetric polymerization configuration is a general requirement for all replisomes, but the question remains how the discontinuously acting lagging-strand polymerase can keep up with the leading-strand polymerase. Priming on the lagging strand inherently slows down lagging-strand synthesis due to the relatively slow rNTP polymerization kinetics[146,147] and the time needed to recruit a polymerase to this new primer. Different single-molecule assays for T7 and T4 replication provided distinct views on coordination mechanisms that we want to outline here.

The T7 replisome has served as an attractive model system to understand the functioning of the individual proteins as well as the behaviour of the whole replisome. Pioneering optical-trapping studies helped for example to understand the mechanochemical properties of the isolated gp5/thioredoxin DNA polymerase.[8,148] Here, it was demonstrated how DNA template tension induces switching between active polymerization and backtracking accompanied by the exonucleolytic removal of nucleotides. These experiments showed that the rate-limiting step in the catalytical cycle of T7 polymerase is force dependent, an observation that resulted in mechanistic insights into the orientation of the DNA in the polymerase active site.[149]

Tethered-bead experiments with 48.5-kb long lambda-phage DNA, anchored to a coverglass surface and stretched hydrodynamically in a flow cell, provided a direct read-out for replication of the fully reconstituted T7 replisome and addressed the mechanism with which leading-strand and lagging-strand syntheses are coupled (Fig. 8A).[66] Here, a bead was attached to the unreplicated, parental end of a forked DNA construct, the T7 replisome components loaded onto the fork, and leading-strand synthesis initiated. At the applied stretching force of $\sim$2 pN, ssDNA is more compact than dsDNA and conversion from double-stranded parental DNA into single-stranded lagging-strand DNA can be monitored by visualizing the gradual motion of the tethered bead towards the anchoring point as the total length of the DNA construct decreases. By comparing leading-strand synthesis traces in the presence or absence of primer synthesis, either *via* removal of the zinc-binding-domain of gp4 or *via* the omittance of ribonucleotides, the authors could demonstrate bead-stalling events of several seconds that were related to priming activity and that preceded lagging-strand loop formation and subsequent (fast) release. Consistent with these experiments is a model in which primase activity transiently halts the progression of the entire fork, preventing leading-strand synthesis from outpacing lagging-strand synthesis.

Another possible mechanism of coupling between the leading and lagging strand is the display of differential rates for the two polymerases at the fork.[150] High-resolution sequencing gels

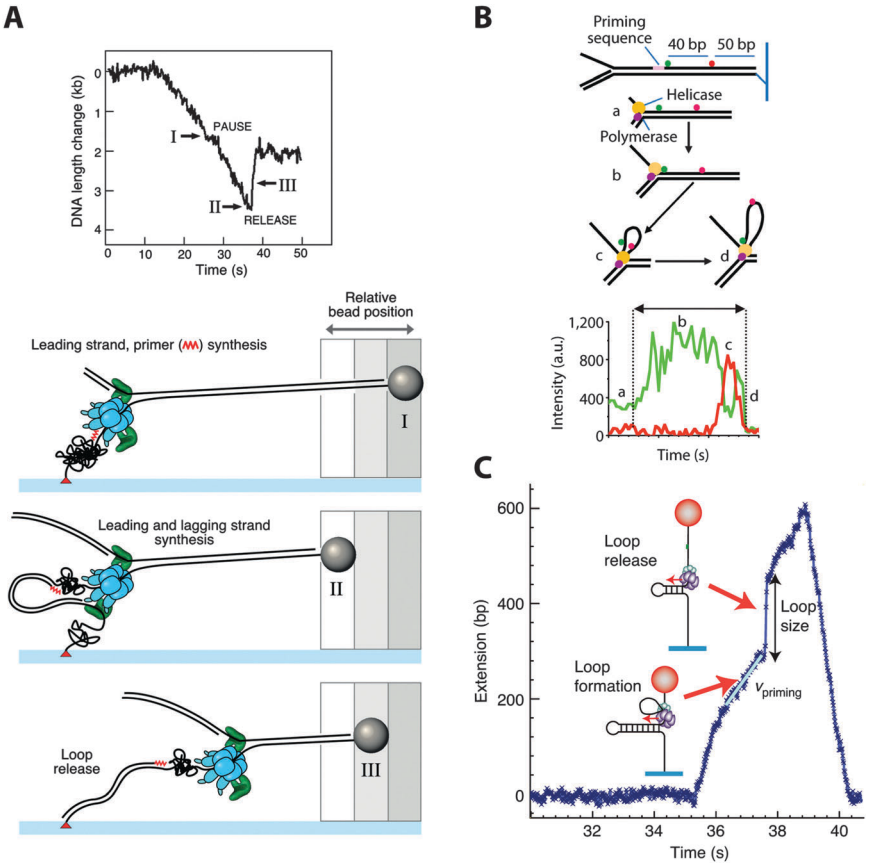



Fig. 8 Coordinated replication. (A) T7 replisome reconstitution in a hydrodynamic DNA flow-stretching assay. A bead attached to one end of the DNA template can be followed by bright-field microscopy and from its trajectory the replication kinetics are extracted. Pausing events in the single-molecule trajectories indicate primer synthesis reactions and lagging strand loop releases are visible as instantaneous DNA lengthening (adapted from ref. 66). (B) T7 replisome reconstitution in a single-molecule FRET assay. Cy3 (green)–Cy5 (red) FRET pairs are installed next to a priming sequence on a DNA template attached to a coverglass surface. Upon helicase/polymerase loading the Cy5 FRET signal increases, suggesting the formation of a priming loop (adapted from ref. 150). (C) T4 primosome–replisome reconstitution in a magnetic trap. A DNA hairpin structure is attached to a coverglass surface and a magnetic bead. The active unwinding by the helicase can be followed by measuring the extension of the DNA template. Priming loops and loop release events are extracted from the bead trajectories (adapted from ref. 151).

provided evidence for leading-strand synthesis not being delayed during priming, but rather, the leading-strand polymerase synthesizing slower than the lagging-strand polymerase. By positioning internal DNA FRET pairs next to the T7 priming sequence, the authors investigated the conformation of the lagging-strand template. They observed increasing FRET acceptor signals in the course of a loop formation event bringing the labelled DNA segments in close proximity to each other, and suggested these signals to be priming loops formed by the lagging-strand between the helicase and the primase domain of gp4 (Fig. 8B). In such a configuration, DNA synthesis can continue without interruption and primers can be synthesized concomitantly with DNA polymerization.

Where the T7 system is unique in that the primase and helicase functions are present in the same protein, the T4 replisome, as most other replication systems, utilizes two different proteins for these enzymatic activities. As described in the previous section, the T4-based DNA synthesis reaction can be reconstituted *in vitro* and magnetic tweezers have been used as a single-molecule approach to unravel the coordination between unwinding and priming.[151] In these studies, a DNA construct containing a hairpin structure was stretched and its extension measured while the helicase and primase were loaded onto the artificial fork (Fig. 8C). The setup allowed a distinction between helicase pausing, priming loop formation and primosome disassembly during primer synthesis, and both loop growth and primase displacement were detected. When additional replication proteins, ssDNA-binding protein, clamp and clamp loader, were applied, looping appeared more often than in the primosome-only complex, although disassembly remained predominant. As a comparative control, a fusion construct of primase and helicase exclusively primed in a looping configuration.

These different scenarios of pausing/looping/disassembly shown for T7 and T4 indicate a need for replication systems to adapt to their particular composition and structure, the available number of proteins at the fork and the relatively different needs for processivity and stability in replicating a phage genome of a few tens of kbp or a Mbp-long bacterial genome. The plasticity of the fork may permit all described

mechanisms interchangeably, thus being more robust towards any obstacles.

### 3.4 Polymerase dynamics

Besides an efficient coupling between leading and lagging-strand synthesis, a processive replication reaction requires a stable association of the polymerases within the replication fork. Since every new Okazaki fragment requires the recruitment of a lagging-strand polymerase, either DNA polymerases from solution need to associate with a newly synthesized primer or the lagging-strand DNA polymerase needs to be recycled efficiently to support the synthesis of multiple Okazaki fragments. Both polymerase exchange and recycling on the lagging strand are feasible, with the first scenario relying on sufficient protein concentrations around the fork so as to not impede the overall reaction kinetics.

In T7, the gp5/trx DNA polymerase was shown to employ two binding modes of different tightness to the gp4 helicase,[152,153] indicating the possibility of multiple distinct steps in the recruitment and utilization of polymerases at the replication fork. Considering gp4 to be hexameric within the replication fork, potentially the weak interaction site between the acidic C-terminal tail of gp4 and a basic patch within the thioredoxin-binding domain (TBD) on gp5[154] could result in a reservoir of polymerases being bound to the replisome. Interestingly, a similar electrostatic interaction could be found between the C terminus of the gp2.5 ssDNA-binding protein and the gp5 polymerase,[152] potentially further increasing the local concentration of polymerases around the replication fork. Such a local excess of polymerases would enable a rapid replacement of a polymerase after dissociation and thus would support a highly processive replication reaction. Ensemble-averaging dilution and competition experiments highlighted the dual behaviour resulting from polymerase switching and recycling: the processivity of the T7 replisome is not diminished by dilution, enforcing the hypothesis of efficient recycling.[155] On the other hand, by using a mutant gp5 that is resistant to inhibition by dideoxynucleotides in competition experiments, rapid exchange of polymerases was observed, as well.[153] Recently, a direct observation of exchange kinetics was feasible in a single-molecule study by tracking fluorescently labelled polymerases[69] (Fig. 9A and B). Here, a DNA substrate tagged with a fluorescent quantum dot at one site was anchored to a coverglass surface and the T7 leading-strand polymerase and helicase preassembled at the fork. The replication reaction could be followed by tracking the quantum dot moving towards the DNA attachment point. Upon addition of fluorescently labeled polymerases to the replication reaction, signals from those polymerases newly arriving at the replication fork were detected, suggesting that excess polymerases stayed associated with the replication fork for several tens of seconds, occupying the available docking sites on the gp4 helicase and ready to replace the synthesizing DNA polymerase.

The mechanism of DNA polymerase exchange and recycling is important in the context of lagging-strand synthesis and replication-loop release. Replication loops are formed in the lagging strand due to the fact that the lagging-strand DNA polymerase remains associated with the rest of the replisome while synthesizing new DNA in a direction that is opposite to the direction of movement of the rest of the replisome. A new loop is formed for every new Okazaki fragment that is synthesized and released before the initiation of the next one. Two pathways of loop release and Okazaki fragment initiation have been proposed.[156–158] In the collision model the polymerase is released upon encountering the 5′ end of the previous Okazaki fragments, thus resulting in the release of the replication loop. In the signalling model, the synthesis of a new primer triggers loop release, before the Okazaki fragment is finished (Fig. 9C). Using single-molecule approaches that rely on the length

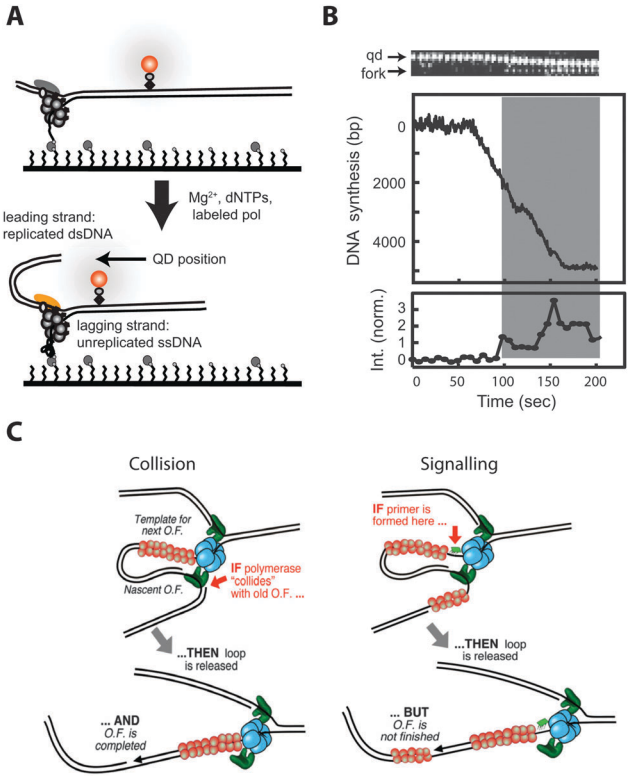



Fig. 9 Polymerase exchange dynamics of the T7 replisome. (A) Unlabeled leading-strand polymerase and helicase are preassembled on DNA. Upon initiation of the reaction, DNA synthesis occurs and is observed as shortening of the DNA by tracking the DNA-template anchored quantum dot (B, middle panel). Upon introduction of fluorescently labeled polymerases to the reaction, fluorescent spots appear at the position of the replication fork and remain there for several seconds (B, bottom panel). Lagging-strand synthesis is not taking place as ribonucleotides are excluded from the reaction (adapted from ref. 69). (C) Collision *vs.* signalling model of the T7 replisome. The hexameric gp4 (blue) translocates along the lagging strand while unwinding the DNA template and priming the Okazaki fragments (O.F.). The polymerases gp5 (green), complexed with thioredoxin, are bound to gp4 and synthesize the leading and lagging strands. Gp2.5 molecules (red) coat and protect the single-stranded DNA that has been extruded behind the helicase. In the collision model, the replication loop is released when the lagging-strand polymerase collides with the 5′ terminus of the previous Okazaki fragment. In the signalling model, the synthesis of a new primer triggers the release of the replication loop before the nascent Okazaki fragment is completed (adapted from ref. 194).

measurement of a single DNA molecule as it is being replicated, as depicted in Fig. 8A, the formation and release of such replication loops have been directly observed and their dynamic properties analysed.[76] These studies revealed that both models are operative during the T7 replication reaction and may serve jointly as a redundancy mechanism to ensure timely loop release. For both mechanisms, however, it is unclear whether the polymerase is recycled to initiate the synthesis of the next Okazaki fragment or whether it stays behind to fill the remaining gap in the previous Okazaki fragment (in the signalling mechanism) or simply dissociates in solution (in the collision mechanism). Combining the observation of DNA length changes during coordinated leading and lagging-strand replication while monitoring the arrival and departure of fluorescently labelled polymerases at the fork, as has been reported for leading-strand synthesis,[69] is an approach that likely will shed more light on these dynamic aspects of the replisome.

### 3.5 *In vivo* studies on the *E. coli* replisome

Extensive work has been done to decipher the interactions between partners within the bacterial replisome (further reviewed in ref. 28, 106 and 159–166). Here we focus on recent studies that particularly concentrate on the dynamics of replication in the context of the living cell. Recent *in vivo* single-molecule studies have provided considerable insight into the spatial and temporal properties of the bacterial replisome and the underlying protein dynamics.

**3.5.1 Exchange dynamics at the replication fork.** As discussed for both the bacteriophage T7 and T4 systems, the plasticity of the replisome is thought to play a critical role in the processive replication reaction. In *E. coli*, similar dynamics were recently demonstrated in terms of polymerase exchange. The clamp-loader protein in *E. coli* plays an important role in determining how many DNA polymerases are present at the fork. The *E. coli dna X* gene encodes two proteins that are present in the clamp loader, τ and γ. γ is a truncated version due to a translational frameshift and, in comparison to τ, does not bind to the Pol III core domain. As the clamp loader contains three copies of the DnaX protein, it was previously assumed that two of them are τ proteins that bind to two polymerase cores for leading and lagging strand synthesis. However, bulk-phase active-site titration analysis provided evidence of the presence of three polymerases in the replisome.[167] This observation was later confirmed by single-molecule *in vivo* imaging studies.[168] Here, several components of the replication fork, including the polymerase core subunits α and ε, were genetically tagged with YPet fluorophores and their intracellular locations as well as intensities tracked in a fluorescence microscope that allowed fast acquisition rates in combination with sufficient laser power (Fig. 10A). Stoichiometries of the different replisome components gave insights into the absolute number of proteins, showing that there exist three τ molecules per clamp loader, as well as three attached polymerase cores. Additionally, γ could be shown to be non-essential for cell growth, suggesting that likely in most forks only τ is present. Further *in vitro* reconstitution studies demonstrated the

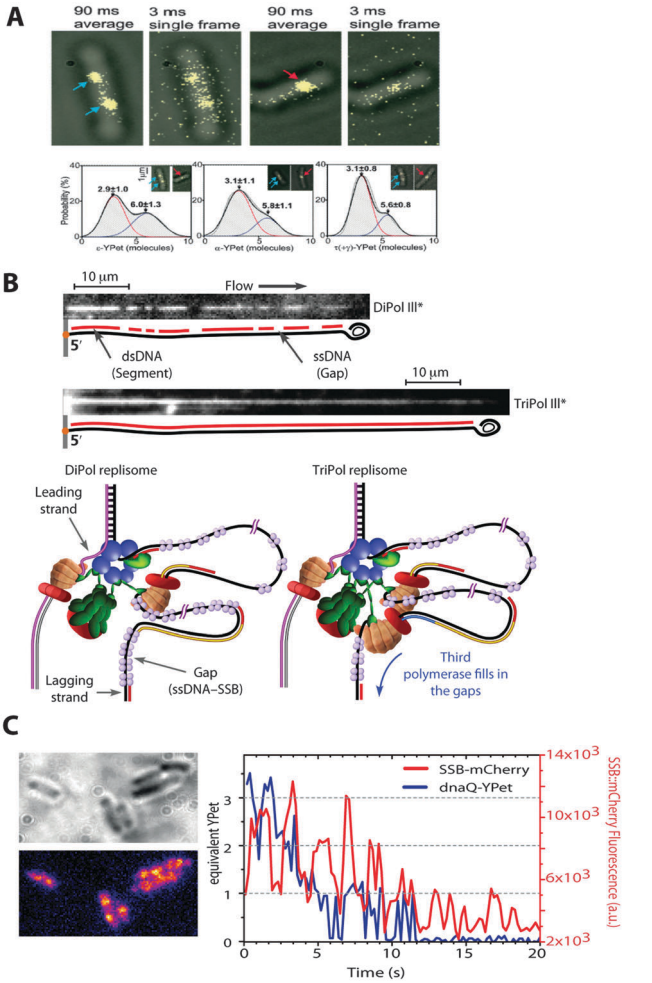



Fig. 10 (A) Polymerase stoichiometry of the *E. coli* replisome *in vitro* and *in vivo*. (A) Fluorescence microscopy of *E. coli* cells expressing ε-YPet. Two replication forks are detected either spatially separated (left) or together within the diffraction limit (right). Both the polymerase subunits α and ε as well as the clamp loader subunit τ are present in three copies in each replication fork (adapted from ref. 168). (B) *In vitro* reconstitution of di- and tri-polymerase replisomes. Introducing a third polymerase at the fork decreases the size of ssDNA gaps between Okazaki fragments, making the process of Okazaki fragment maturation more efficient in comparison to a replisome containing two polymerases (adapted from ref. 169). (C) Polymerase exchange kinetics *in vivo*. Fluorescent DnaQ (ε)-YPet proteins are localized within the confined replication fork regions in *E. coli* cells. Integrated fluorescence traces for individual replisomes are converted to absolute numbers of DnaQ copies and correlated to the time distribution of fluorescent SSB-mCherry molecules. Coincidental intensity fluctuations point to exchange events within Okazaki fragment initiation (adapted from ref. 170).

advantage of replisomes containing three polymerases in comparison to those containing only two: Okazaki fragments could be significantly better filled in, leading to a faster lagging-strand completion, and the processivity of the fork could be shown to be higher (Fig. 10B).[169] In this study, the authors point out that the presence of a third polymerase is likely to be beneficial for efficient primer capture, and that having the Pol III stay with a nascent Okazaki fragment after loop release prevents other polymerases of *E. coli* with less fidelity (Pol II, Pol IV)

from filling the remaining gap on the template. If there are three potential binding sites for the polymerase core complex, are the dynamics of polymerase exchange similar to the T7 system, as described above? In another *in vivo* study, both SSB and polymerase core ε subunits were fluorescently tagged and their signals correlated (Fig. 10C).[170] Correlating protein fluctuations at the fork suggested that for every Okazaki fragment a new polymerase associates on the lagging strand and, in addition, that it also proceeds faster than the leading-strand polymerase. An additional third polymerase in close vicinity would allow a rapid restart of synthesis after polymerase dissociation from the replisome upon loop release.

**3.5.2 Obstacles along the template.** Replication forks are likely to face various obstacles during their progression through large genomic stretches of DNA: lesions in the DNA and proteins bound to the DNA, such as transcribing RNA polymerases, may block the template and histone-like proteins introduce twist and bending of the DNA that may present hurdles for the replisome.

Those potential blockades might stall the replication fork and provoke a (partial) disassembly. Experimental approaches to characterize such stalling events involve for example UV irradiation or the use of thermosensitive replication mutants, as for DnaB.[171] With single-molecule detection sensitivity in live cells, the disassembly kinetics could be followed in real time after DnaB arrest by tracking the different replisome components encoded as fusion constructs to fluorescent proteins.[172] In this study, the authors observed not only remarkably distinct time scales of polymerase core dissociation, which may point to a stronger interaction of the leading-strand polymerase with the replisome than that of the lagging-strand polymerase, but also revealed the initiation of a fork rescue process with RecA filaments replacing SSB.

The double-helical structure of DNA itself also poses a barrier for replication. The torsional energy that builds up during replication or transcription gives rise to a number of topological challenges the replisome needs to deal with. The exact nature of topological deformation depends on how the replication fork proceeds. If the replisome is free to rotate, precatenanes behind the fork can be generated. Without rotation, positive supercoiling in front of the fork is the cost for unwinding activity.[173] In both cases, topoisomerases of type I and II are indispensable in dealing with the intertwined regions. A recent study addressed the torsional stress on the replisome in the context of coupled replication.[174] Here the authors argued that the dissociation mechanism of the lagging-strand polymerase and thus its exchange kinetics are important not only in the context of rapid primer capture, but also in terms of DNA stress release. The concept of the signalling model in which primer synthesis triggers loop release (Fig. 9C) could thus be extended to a torsion-dependent displacement of the lagging-strand polymerase.

**3.5.3 Chromosome structure.** Globally, replication is coordinated with the process of chromosome segregation. Structural maintenance of chromosome (SMC) proteins, in *E. coli* the MukBEF complex, act to organize the chromosome during the cell cycle. Recently, individual copies of the MukBEF complex could be imaged using fluorescence microscopy[175] (Fig. 11). Not only could the stoichiometry of the functional minimal unit of the complex be resolved by counting the fluorescently tagged components, but also the salient details of the ATP-dependent DNA-modelling pathway could be studied. Stable MukBEF complexes were found to accumulate at several discrete spots in the cell to form higher-order structures that are engaged in organizing the chromosome. Making use of ATPase-deficient mutant strains, the authors could show that in these spots the minimal MukBEF units exchange *via* ATP binding and hydrolysis cycles. Due to MukBEF multimerization, segments of DNA can be caught without losing contact with other ones, so that in sum, MukBEF works as a step-by-step remodeller, grabbing and bringing together distal segments of the chromosome (Fig. 11B).

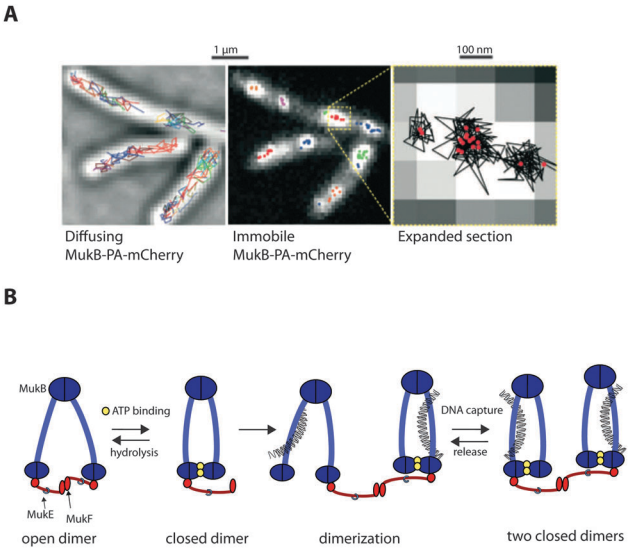



Fig. 11 MukBEF organization in *E. coli*. (A) Using live-cell PALM, diffusing and immobile MukB-PAmCherry molecules can be tracked and intensities integrated to estimate the number of proteins per spot. (B) "Rock-climber" model for the action of SMC complexes in cells. MukBEF complexes consist of dimers of dimeric 4 : 4 : 2 MukB : E : F molecules that associate on DNA. ATP hydrolysis triggers release from the DNA segment; however as the multimeric complexes exhibit several DNA binding sites, complete release from the chromosome is inhibited (adapted from ref. 175).

A recent study, based on pull-down and topology/superhelicity relaxation assays in bulk, looked into the interaction of MukB with ParC, a subunit of topoisomerase IV.[176] Here, MukB enhanced ParC-catalysed relaxation of negatively supercoiled and knotted DNA, exemplifying how multiple protein interactions and complexes are involved in the topological configuration of the chromosome that interacts with the replication progression.

### 3.6 Replication in the context of eukaryotic cell division

The eukaryotic genome is highly structured into chromatin, which is a critical requirement not only for proper cell division, but also for the storage of epigenetic information such as histone modifications. The regulation of replication, including aspects like the establishment of pre-replication complexes,

the timing of firing, the kinetics of replication elongation, termination, and numerous others, is critically linked to the chromosomal structure. Recent advances in single-molecule methodology have begun to address not only the mechanisms underlying the functioning of the replisome, but also its dynamic properties in the context of chromatinized DNA.

**3.6.1 The MCM licensing transformation.** Due to the large number of bases that have to be replicated in the eukaryotic cell, replication proteins are loaded on many sites on the chromosome. However, each of these replication origins needs to be tightly regulated so that it fires only once per cell cycle. The licensing process prevents that a single origin undergoes multiple replication start events during S phase. Misregulation can lead to under- or over-replication of chromosomal DNA, contributing to genome instability in cancer cells.[177] The eukaryotic replicative helicase MCM2-7 is a central component in the process of replication activation; however the molecular steps towards the formation of functional helicases and their regulation remain unclear.

A powerful single-molecule assay to observe eukaryotic replication has been developed recently that is based on cellular extracts from *Xenopus* egg cells.[178] The licensing process is recapitulated by loading the DNA substrate first with an extract fraction that initiates the ORC-dependent MCM2-7 helicase binding analogous to the processes taking place during the G1 phase of the cell cycle (Fig. 6B). Subsequently, an extract enriched in nuclear factors is applied that mimics the S phase, providing all replication proteins needed for successful replication-fork formation. By staining double-stranded DNA with intercalating dye and replication forks using the combination of dig-dUTP incorporation and fluorescent anti-dig labelling, the separation of the replisomes from the origin can be visualized (Fig. 12). In this way, it could be shown how the two sister replisomes starting from one origin uncouple during the course of DNA synthesis. The active eukaryotic replication process is also observable in real-time: in a recent study, a replication protein involved in the maturation of Okazaki fragments, Fen-1, was labelled with the photoswitchable mKikGr protein and introduced to the replication mixture of cell extracts at micromolar concentrations.[179] Where such a concentration would have been too high to allow for single-molecule imaging using conventional single-molecule fluorescence imaging methods, the authors applied a cycle of photoactivation, diffusion and excitation (PhADE) to allow the visualization of single Fen-1 proteins acting at the replication forks (Fig. 13).

The current model for eukaryotic replication initiation proposes the formation of inactive MCM2-7 hexamers that are loaded onto the DNA within the pre-replication complex.[180] Bulk assays as well as EM structures indicate that isolated MCM complexes encircle double-stranded DNA and can also slide along the duplex.[180,181] In contrast, other helicases, such as the bacterial and phage ones described in the previous sections, have been shown to proceed along single-stranded DNA while unwinding the template, presumably *via* a steric exclusion mechanism. As a result, two scenarios are commonly described for the active unwinding helicase: either MCM2-7 stays in its

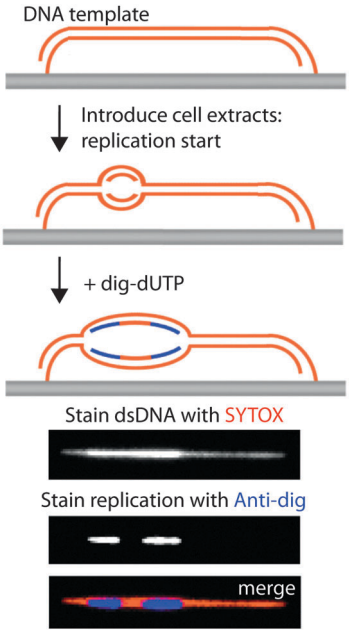



Fig. 12 Bidirectional eukaryotic replication. *Xenopus* cell extracts, containing the pre-replication licensing complex and the replication proteins, are applied to an immobilized DNA template. The nucleotide analogue dig-dUTP is added after replication has started, and is incorporated into the daughter DNA strands. Post-replication staining is achieved by using dsDNA intercalating SYTOX dye and anti-dig antibody. This approach allows the uncoupled sister replication forks to be resolved (adapted from ref. 195).

dsDNA-binding mode or its ring structure is triggered to be opened during the licensing process in order to encircle and translocate along ssDNA.[182] In order to examine the unwinding mechanism of MCM2-7 at the single-molecule level, strand-specific roadblocks were coupled to a lambda DNA template and fork progression was monitored.[183] Lagging-strand roadblocks, either biotin–streptavidin complexes or quantum dots, hardly stalled the proceeding replisomes, whereas leading-strand obstacles strongly arrested fork progression. These results show that the MCM2-7 helicase encircles ssDNA and translocates in a 3′–5′ direction and imply that during the G1/S phase transition the dsDNA-encircling MCM2-7 complex transitions into a ssDNA-encircling complex.[184]

**3.6.2 Nucleosomes as obstacles for DNA motors.** DNA is tightly packed in nucleosomal core particles (NCPs) consisting of an octamer of histone proteins, each containing two copies of H2A, H2B, H3, H4, and that is wrapped by a stretch of 147 base pairs of DNA.[185] Sliding of nucleosomes, posttranslational modifications (PTMs) of histones, and chromatin remodelling by chaperones and remodeller enzymes are processes that continuously influence the global structure of nucleosomally compacted DNA and play a key role in a broad range of regulatory mechanisms.[186]

The chromatin structure has to be preserved during replication, *i.e.* after the displacement of the parental histones during unwinding, nucleosomes need to be repositioned on the nascent DNA. Further, the epigenetic profile of histone modification,

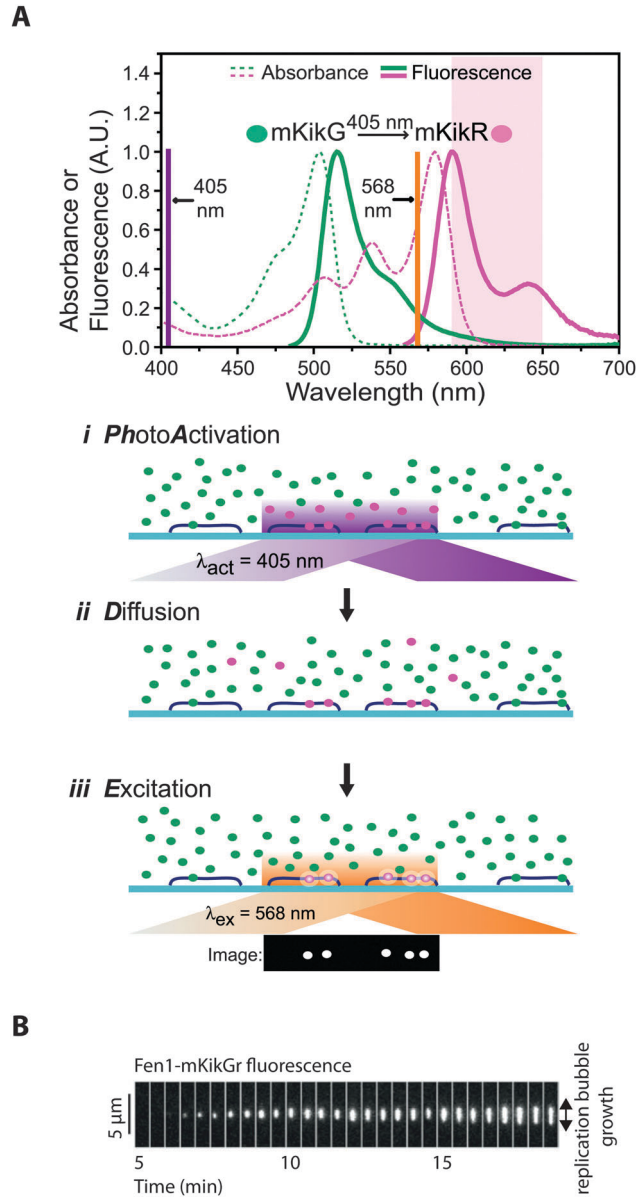



Fig. 13 The concept of PhADE (photoactivation, diffusion and excitation). (A) The photoswitchable mKikGR protein is activated by 405 nm light to transition from its green to red fluorescence state. In a TIRF setup, only those molecules within the evanescent field near the coverglass surface are activated. Waiting before imaging in the red channel until the activated proteins in solution have diffused to outside the evanescent field allows the selective visualization of only those proteins bound to the DNA. (B) In order to follow the eukaryotic replication fork progression in real time, the replication protein Fen1 is tagged with mKikGr protein and added to the *Xenopus* cell extract (see text for details). The kymograph recapitulates the growth and separation of the replication forks (adapted from ref. 179).

mainly acetylation and phosphorylation, has to be transmitted from the parental to the daughter DNA. Some hints about the underlying mechanisms were provided by a study that demonstrated that the MCM2 helicase subunit of yeast binds histones cooperatively with the FACT (facilitates chromatin transcription) complex, which acts as a histone chaperone.[126] As shown by immunoprecipitation, FACT is part of the replisome by interaction with MCM2, independent of concomitant histone binding. Although those multi-interaction processes are difficult to decipher, several single-molecule studies have begun to give significant insight into DNA-nucleosome dynamics (reviewed in greater detail elsewhere[187]). Using optical tweezers, the unzipping mechanism of single NCPs could be tracked and an energy landscape described for the distinct interaction domains, revealing that the dyad domain association is particularly strong.[187] In another recent study, magnetic tweezers as well as FRET microscopy were applied to distinguish the effects of post-transcriptional modifications depending on their exact positions. It appeared that especially near the dyad, post-transcriptional modifications speed up histone disassembly processes, whereas those localized near the DNA entry/exit site regulate the unwrapping pathway.[188] Similarly, by installing DNA FRET pairs on the template, the compaction as related to CpG island methylation could be quantified, pointing to a positive correlation between the amount of methylated sites and compaction.[189]

Further single-molecule studies of eukaryotic replication in a cellular context will likely give exciting insights into the complex processes involved in DNA metabolism. In combination with cell biological studies, theoretical modelling and mechanistic biochemical work, such single-molecule experiments will be necessary to not only provide fine-grained kinetic information on the various molecular processes but also enable a much more quantitative view of the molecular processes taking place inside living cells and as such play a role in bridging experimental and theoretical studies.

## 4. Conclusions and outlook

The statement "the whole is greater than the sum of its parts" applies to many biological systems, be it the complicated pathways of signal transduction, cytoskeletal reconstruction in cell division, or as described in this review, DNA replication. Initial single-molecule studies concentrated on the isolated DNA motors, quantifying the force generation and processivity of polymerases or unwinding activity of helicases, and laid the basis for studies of the entire replisome. Single-molecule experiments merely represent a category of approaches from the toolbox of biophysics and biochemistry. In combination with structural, biochemical, genetic and other approaches, they provide us with a tool that has proven to be powerful in revealing new features about the inner workings of individual motors and larger units that self-organize into multifunctional complexes. The understanding of the replisome, isolated or inside the living cell, gives us powerful information about complex cellular structures. The central theme, the plasticity of the replication fork in its ability to exchange components during replication, is remarkable considering the processivity and robustness of coupled replication.

Ultimately, a key goal is to construct a holistic picture of the cell and how it is built up from many molecular pathways. During the last few years, reductionist and bottom-up approaches to construct a synthetic cell have been explored and have already led

to intriguing results. The reconstitution of compartments or units of increasing intricacy is one key step towards recapitulating and understanding cellular complexity, and several single-molecule studies, many of them focussing on cytoskeletal systems, rely on this method. Highly related to cellular complexity is the crowded nature of the intracellular environment. Going forward, it is important to characterize and understand how biochemical pathways work not only in the context of highly purified and homogeneous systems, but also in the context of molecular crowding.[190] Single-molecule approaches have begun to significantly impact how complex biochemical pathways are studied and further integration with other tools and mechanistic insights will undoubtedly result in significant contributions to our understanding of the molecular principles underlying life.

## References


1 M. K. Gardner, M. Zanic, C. Gell, V. Bormuth and J. Howard, *Cell*, 2011, **147**, 1092–1103.
2 M. Tomishige, N. Stuurman and R. D. Vale, *Nat. Struct. Mol. Biol.*, 2006, **13**, 887–894.
3 R. D. Vale, T. Funatsu, D. W. Pierce, L. Romberg, Y. Harada and T. Yanagida, *Nature*, 1996, **380**, 451–453.
4 K. Adachi, K. Oiwa, M. Yoshida, T. Nishizaka and K. Kinosita Jr., *Nat. Commun.*, 2012, **3**, 1022.
5 H. Noji, R. Yasuda, M. Yoshida and K. Kinosita Jr., *Nature*, 1997, **386**, 299–302.
6 E. A. Abbondanzieri, W. J. Greenleaf, J. W. Shaevitz, R. Landick and S. M. Block, *Nature*, 2005, **438**, 460–465.
7 D. A. Schafer, J. Gelles, M. P. Sheetz and R. Landick, *Nature*, 1991, **352**, 444–448.
8 G. J. Wuite, S. B. Smith, M. Young, D. Keller and C. Bustamante, *Nature*, 2000, **404**, 103–106.
9 K. Aathavan, A. T. Politzer, A. Kaplan, J. R. Moffitt, Y. R. Chemla, S. Grimes, P. J. Jardine, D. L. Anderson and C. Bustamante, *Nature*, 2009, **461**, 669–673.
10 Y. R. Chemla, K. Aathavan, J. Michaelis, S. Grimes, P. J. Jardine, D. L. Anderson and C. Bustamante, *Cell*, 2005, **122**, 683–692.
11 J. Elf, G. W. Li and X. S. Xie, *Science*, 2007, **316**, 1191–1194.
12 Y. Taniguchi, P. J. Choi, G. W. Li, H. Chen, M. Babu, J. Hearn, A. Emili and X. S. Xie, *Science*, 2010, **329**, 533–538.
13 Y. Sako, S. Minoghchi and T. Yanagida, *Nat. Cell Biol.*, 2000, **2**, 168–172.
14 T. Xia, N. Li and X. Fang, *Annu. Rev. Phys. Chem.*, 2013, **64**, 459–480.
15 S. Chaudhury and O. A. Igoshin, *PLoS One*, 2010, **5**, e12364.
16 H. P. Lu, L. Xun and X. S. Xie, *Science*, 1998, **282**, 1877–1882.
17 L. Edman and R. Rigler, *Proc. Natl. Acad. Sci. U. S. A.*, 2000, **97**, 8266–8271.
18 B. Liu, R. J. Baskin and S. C. Kowalczykowski, *Nature*, 2013, **500**, 482–485.
19 S. C. Kou, B. J. Cherayil, W. Min, B. P. English and X. S. Xie, *J. Phys. Chem. B*, 2005, **109**, 19068–19081.
20 W. Min, B. P. English, G. Luo, B. J. Cherayil, S. C. Kou and X. S. Xie, *Acc. Chem. Res.*, 2005, **38**, 923–931.
21 Q. Chen, R. Groote, H. Schonherr and G. J. Vancso, *Chem. Soc. Rev.*, 2009, **38**, 2671–2683.
22 W. Bae, M. G. Choi, C. Hyeon, Y. K. Shin and T. Y. Yoon, *J. Am. Chem. Soc.*, 2013, **135**, 10254–10257.
23 B. A. Smith, K. Daugherty-Clarke, B. L. Goode and J. Gelles, *Proc. Natl. Acad. Sci. U. S. A.*, 2013, **110**, 1285–1290.
24 B. M. Alberts, *Philos. Trans. R. Soc., B*, 1987, **317**, 395–420.
25 S. J. Benkovic, A. M. Valentine and F. Salinas, *Annu. Rev. Biochem.*, 2001, **70**, 181–208.
26 N. Y. Yao and M. O'Donnell, *Mol. BioSyst.*, 2008, **4**, 1075–1084.
27 T. A. Kunkel and K. Bebenek, *Annu. Rev. Biochem.*, 2000, **69**, 497–529.
28 A. M. van Oijen, *Mol. BioSyst.*, 2007, **3**, 117–125.
29 T. Hirschfeld, *Appl. Opt.*, 1976, **15**, 2965–2966.
30 W. E. Moerner and L. Kador, *Phys. Rev. Lett.*, 1989, **62**, 2535–2538.
31 W. E. Moerner and W. P. Ambrose, *Phys. Rev. Lett.*, 1991, **66**, 1376.
32 M. Orrit and J. Bernard, *Phys. Rev. Lett.*, 1990, **65**, 2716–2719.
33 A. M. van Oijen, M. Ketelaars, J. Kohler, T. J. Aartsma and J. Schmidt, *Science*, 1999, **285**, 400–402.
34 E. Betzig and R. J. Chichester, *Science*, 1993, **262**, 1422–1425.
35 A. Yamada, T. Yamaga, H. Sakakibara, H. Nakayama and K. Oiwa, *J. Cell Sci.*, 1998, **111**(Pt 1), 93–98.
36 J. Yu, J. Xiao, X. Ren, K. Lao and X. S. Xie, *Science*, 2006, **311**, 1600–1603.
37 E. J. Ambrose, *Nature*, 1956, **178**, 1194.
38 D. Axelrod, *J. Cell Biol.*, 1981, **89**, 141–145.
39 E. H. Stelzer, I. Wacker and J. R. De Mey, *Semin. Cell Biol.*, 1991, **2**, 145–152.
40 A. Nakano, *Cell Struct. Funct.*, 2002, **27**, 349–355.
41 F. Helmchen and W. Denk, *Nat. Methods*, 2005, **2**, 932–940.
42 S. W. Hell and J. Wichmann, *Opt. Lett.*, 1994, **19**, 780–782.
43 F. V. Subach, G. H. Patterson, M. Renz, J. Lippincott-Schwartz and V. V. Verkhusha, *J. Am. Chem. Soc.*, 2010, **132**, 6481–6491.
44 S. Uphoff, R. Reyes-Lamothe, F. Garza de Leon, D. J. Sherratt and A. N. Kapanidis, *Proc. Natl. Acad. Sci. U. S. A.*, 2013, **110**, 8063–8068.
45 S. R. Nelson, M. Y. Ali and D. M. Warshaw, *Methods Mol. Biol.*, 2011, **778**, 111–121.
46 M. Ormo, A. B. Cubitt, K. Kallio, L. A. Gross, R. Y. Tsien and S. J. Remington, *Science*, 1996, **273**, 1392–1395.
47 G. J. Kremers, S. G. Gilbert, P. J. Cranfill, M. W. Davidson and D. W. Piston, *J. Cell Sci.*, 2011, **124**, 157–160.
48 N. C. Shaner, G. H. Patterson and M. W. Davidson, *J. Cell Sci.*, 2007, **120**, 4247–4260.
49 A. M. van Oijen, *Curr. Opin. Biotechnol.*, 2011, **22**, 75–80.
50 M. Andresen, A. C. Stiel, J. Folling, D. Wenzel, A. Schonle, A. Egner, C. Eggeling, S. W. Hell and S. Jakobs, *Nat. Biotechnol.*, 2008, **26**, 1035–1040.

51 M. Andresen, A. C. Stiel, S. Trowitzsch, G. Weber, C. Eggeling, M. C. Wahl, S. W. Hell and S. Jakobs, *Proc. Natl. Acad. Sci. U. S. A.*, 2007, **104**, 13005–13009.
52 M. Z. Lin and L. Wang, *Physiology*, 2008, **23**, 131–141.
53 G. V. Los, L. P. Encell, M. G. McDougall, D. D. Hartzell, N. Karassina, C. Zimprich, M. G. Wood, R. Learish, R. F. Ohana, M. Urh, D. Simpson, J. Mendez, K. Zimmerman, P. Otto, G. Vidugiris, J. Zhu, A. Darzins, D. H. Klaubert, R. F. Bulleit and K. V. Wood, *ACS Chem. Biol.*, 2008, **3**, 373–382.
54 A. Juillerat, T. Gronemeyer, A. Keppler, S. Gendreizig, H. Pick, H. Vogel and K. Johnsson, *Chem. Biol.*, 2003, **10**, 313–317.
55 B. A. Griffin, S. R. Adams and R. Y. Tsien, *Science*, 1998, **281**, 269–272.
56 S. R. Adams, R. E. Campbell, L. A. Gross, B. R. Martin, G. K. Walkup, Y. Yao, J. Llopis and R. Y. Tsien, *J. Am. Chem. Soc.*, 2002, **124**, 6063–6076.
57 X. Shi, Y. Jung, L. J. Lin, C. Liu, C. Wu, I. K. Cann and T. Ha, *Nat. Methods*, 2012, **9**, 499–503.
58 E. G. Guignet, R. Hovius and H. Vogel, *Nat. Biotechnol.*, 2004, **22**, 440–444.
59 K. M. Marks, M. Rosinov and G. P. Nolan, *Chem. Biol.*, 2004, **11**, 347–356.
60 C. J. Noren, S. J. Anthony-Cahill, M. C. Griffith and P. G. Schultz, *Science*, 1989, **244**, 182–188.
61 L. Wang, A. Brock, B. Herberich and P. G. Schultz, *Science*, 2001, **292**, 498–500.
62 J. C. Anderson, N. Wu, S. W. Santoro, V. Lakshman, D. S. King and P. G. Schultz, *Proc. Natl. Acad. Sci. U. S. A.*, 2004, **101**, 7566–7571.
63 L. Bossi and J. R. Roth, *Cell*, 1981, **25**, 489–496.
64 M. Bates, B. Huang and X. Zhuang, *Curr. Opin. Chem. Biol.*, 2008, **12**, 505–514.
65 M. Bates, B. Huang, G. T. Dempsey and X. Zhuang, *Science*, 2007, **317**, 1749–1753.
66 J. B. Lee, R. K. Hite, S. M. Hamdan, X. S. Xie, C. C. Richardson and A. M. van Oijen, *Nature*, 2006, **439**, 621–624.
67 T. Strick, J. Allemand, V. Croquette and D. Bensimon, *Prog. Biophys. Mol. Biol.*, 2000, **74**, 115–140.
68 C. M. Etson, S. M. Hamdan, C. C. Richardson and A. M. van Oijen, *Proc. Natl. Acad. Sci. U. S. A.*, 2010, **107**, 1900–1905.
69 J. J. Loparo, A. W. Kulczyk, C. C. Richardson and A. M. van Oijen, *Proc. Natl. Acad. Sci. U. S. A.*, 2011, **108**, 3584–3589.
70 K. C. Neuman and A. Nagy, *Nat. Methods*, 2008, **5**, 491–505.
71 I. De Vlaminck and C. Dekker, *Annu. Rev. Biophys.*, 2012, **41**, 453–472.
72 W. J. Greenleaf, M. T. Woodside and S. M. Block, *Annu. Rev. Biophys. Biomol. Struct.*, 2007, **36**, 171–190.
73 D. Dulin, J. Lipfert, M. C. Moolman and N. H. Dekker, *Nat. Rev. Genet.*, 2013, **14**, 9–22.
74 T. R. Strick, J. F. Allemand, D. Bensimon, A. Bensimon and V. Croquette, *Science*, 1996, **271**, 1835–1837.
75 T. R. Strick, V. Croquette and D. Bensimon, *Nature*, 2000, **404**, 901–904.
76 S. M. Hamdan, J. J. Loparo, M. Takahashi, C. C. Richardson and A. M. van Oijen, *Nature*, 2009, **457**, 336–339.
77 R. Zhou, M. Schlierf and T. Ha, *Methods Enzymol.*, 2010, **475**, 405–426.
78 P. Gross, G. Farge, E. J. Peterman and G. J. Wuite, *Methods Enzymol.*, 2010, **475**, 427–453.
79 S. Hohng, R. Zhou, M. K. Nahas, J. Yu, K. Schulten, D. M. Lilley and T. Ha, *Science*, 2007, **318**, 279–283.
80 B. M. Alberts, J. Barry, P. Bedinger, T. Formosa, C. V. Jongeneel and K. N. Kreuzer, *Cold Spring Harbor Symp. Quant. Biol.*, 1983, **47**(Pt 2), 655–668.
81 T. J. Kelly and G. W. Brown, *Annu. Rev. Biochem.*, 2000, **69**, 829–880.
82 H. Masai, S. Matsumoto, Z. You, N. Yoshizawa-Sugata and M. Oda, *Annu. Rev. Biochem.*, 2010, **79**, 89–130.
83 S. J. Lee and C. C. Richardson, *J. Biol. Chem.*, 2001, **276**, 49419–49426.
84 M. Kato, T. Ito, G. Wagner, C. C. Richardson and T. Ellenberger, *Mol. Cell*, 2003, **11**, 1349–1360.
85 E. A. Toth, Y. Li, M. R. Sawaya, Y. Cheng and T. Ellenberger, *Mol. Cell*, 2003, **12**, 1113–1123.
86 D. J. Crampton, M. Ohi, U. Qimron, T. Walz and C. C. Richardson, *J. Mol. Biol.*, 2006, **360**, 667–677.
87 P. Ahnert, K. M. Picha and S. S. Patel, *EMBO J.*, 2000, **19**, 3418–3427.
88 D. E. Johnson and C. C. Richardson, *J. Biol. Chem.*, 2003, **278**, 23762–23772.
89 S. Tabor, H. E. Huber and C. C. Richardson, *J. Biol. Chem.*, 1987, **262**, 16212–16223.
90 B. Akabayov, S. R. Akabayov, S. J. Lee, S. Tabor, A. W. Kulczyk and C. C. Richardson, *Proc. Natl. Acad. Sci. U. S. A.*, 2010, **107**, 15033–15038.
91 S. J. Lee and C. C. Richardson, *Curr. Opin. Chem. Biol.*, 2011, **15**, 580–586.
92 S. M. Hamdan and C. C. Richardson, *Annu. Rev. Biochem.*, 2009, **78**, 205–243.
93 S. Ghosh, S. M. Hamdan and C. C. Richardson, *J. Biol. Chem.*, 2010, **285**, 18103–18112.
94 S. Ghosh, B. Marintcheva, M. Takahashi and C. C. Richardson, *J. Biol. Chem.*, 2009, **284**, 30339–30349.
95 C. F. Morris, N. K. Sinha and B. M. Alberts, *Proc. Natl. Acad. Sci. U. S. A.*, 1975, **72**, 4800–4804.
96 N. K. Sinha, C. F. Morris and B. M. Alberts, *J. Biol. Chem.*, 1980, **255**, 4290–4293.
97 M. M. Spiering, S. W. Nelson and S. J. Benkovic, *Mol. BioSyst.*, 2008, **4**, 1070–1074.
98 S. R. Arumugam, T. H. Lee and S. J. Benkovic, *J. Biol. Chem.*, 2009, **284**, 29283–29289.
99 K. D. Raney, T. E. Carver and S. J. Benkovic, *J. Biol. Chem.*, 1996, **271**, 14074–14081.
100 D. Jose, S. E. Weitzel, D. Jing and P. H. von Hippel, *Proc. Natl. Acad. Sci. U. S. A.*, 2012, **109**, 13596–13601.
101 D. M. Hinton and N. G. Nossal, *J. Biol. Chem.*, 1987, **262**, 10873–10878.
102 A. M. Valentine, F. T. Ishmael, V. K. Shier and S. J. Benkovic, *Biochemistry*, 2001, **40**, 15074–15085.

103 S. W. Nelson, R. Kumar and S. J. Benkovic, *J. Biol. Chem.*, 2008, **283**, 22838–22846.
104 B. A. Kelch, D. L. Makino, M. O'Donnell and J. Kuriyan, *Science*, 2011, **334**, 1675–1680.
105 D. C. Mace and B. M. Alberts, *J. Mol. Biol.*, 1984, **177**, 295–311.
106 M. J. Davey, D. Jeruzalmi, J. Kuriyan and M. O'Donnell, *Nat. Rev. Mol. Cell Biol.*, 2002, **3**, 826–835.
107 J. P. Erzberger, M. M. Pirruccello and J. M. Berger, *EMBO J.*, 2002, **21**, 4763–4773.
108 K. E. Duderstadt and J. M. Berger, *Crit. Rev. Biochem. Mol. Biol.*, 2008, **43**, 163–187.
109 K. E. Duderstadt, K. Chuang and J. M. Berger, *Nature*, 2011, **478**, 209–213.
110 K. Sekimizu, D. Bramhill and A. Kornberg, *Cell*, 1987, **50**, 259–265.
111 B. Y. Yung, E. Crooke and A. Kornberg, *J. Biol. Chem.*, 1990, **265**, 1282–1285.
112 C. Margulies and J. M. Kaguni, *J. Biol. Chem.*, 1996, **271**, 17035–17040.
113 W. Messer, F. Blaesing, D. Jakimowicz, M. Krause, J. Majka, J. Nardmann, S. Schaper, H. Seitz, C. Speck, C. Weigel, G. Wegrzyn, M. Welzeck and J. Zakrzewska-Czerwinska, *Biochimie*, 2001, **83**, 5–12.
114 M. L. Mott, J. P. Erzberger, M. M. Coons and J. M. Berger, *Cell*, 2008, **135**, 623–634.
115 M. J. Davey, L. Fang, P. McInerney, R. E. Georgescu and M. O'Donnell, *EMBO J.*, 2002, **21**, 3148–3159.
116 Z. Kelman and M. O'Donnell, *Annu. Rev. Biochem.*, 1995, **64**, 171–200.
117 R. E. Georgescu, I. Kurth, N. Y. Yao, J. Stewart, O. Yurieva and M. O'Donnell, *EMBO J.*, 2009, **28**, 2981–2991.
118 L. Rowen and A. Kornberg, *J. Biol. Chem.*, 1978, **253**, 758–764.
119 Y. B. Lu, P. V. Ratnakar, B. K. Mohanty and D. Bastia, *Proc. Natl. Acad. Sci. U. S. A.*, 1996, **93**, 12902–12907.
120 A. V. Mitkova, S. M. Khopde and S. B. Biswas, *J. Biol. Chem.*, 2003, **278**, 52253–52261.
121 G. S. Khatri, T. MacAllister, P. R. Sista and D. Bastia, *Cell*, 1989, **59**, 667–674.
122 T. M. Hill and K. J. Marians, *Proc. Natl. Acad. Sci. U. S. A.*, 1990, **87**, 2481–2485.
123 J. T. Fox, K. Y. Lee and K. Myung, *FEBS Lett.*, 2011, **585**, 2780–2785.
124 K. Y. Lee and K. Myung, *Mol. Cells*, 2008, **26**, 5–11.
125 R. A. Sclafani and T. M. Holzen, *Annu. Rev. Genet.*, 2007, **41**, 237–280.
126 M. Foltman, C. Evrin, G. De Piccoli, R. C. Jones, R. D. Edmondson, Y. Katou, R. Nakato, K. Shirahige and K. Labib, *Cell Rep.*, 2013, **3**, 892–904.
127 C. Li and J. Jin, *Protein Cell*, 2010, **1**, 227–236.
128 L. N. Truong and X. Wu, *J. Mol. Cell Biol.*, 2011, **3**, 13–22.
129 M. Mechali, *Nat. Rev. Mol. Cell Biol.*, 2010, **11**, 728–738.
130 M. Lei and B. K. Tye, *J. Cell Sci.*, 2001, **114**, 1447–1454.
131 D. I. Lin, P. Aggarwal and J. A. Diehl, *Proc. Natl. Acad. Sci. U. S. A.*, 2008, **105**, 8079–8084.
132 S. E. Moyer, P. W. Lewis and M. R. Botchan, *Proc. Natl. Acad. Sci. U. S. A.*, 2006, **103**, 10236–10241.
133 N. Marinsek, E. R. Barry, K. S. Makarova, I. Dionne, E. V. Koonin and S. D. Bell, *EMBO Rep.*, 2006, **7**, 539–545.
134 Y. P. Chang, G. Wang, V. Bermudez, J. Hurwitz and X. S. Chen, *Proc. Natl. Acad. Sci. U. S. A.*, 2007, **104**, 12685–12690.
135 M. P. Longhese, P. Plevani and G. Lucchini, *Mol. Cell. Biol.*, 1994, **14**, 7884–7890.
136 A. Georgaki, B. Strack, V. Podust and U. Hubscher, *FEBS Lett.*, 1992, **308**, 240–244.
137 P. M. Burgers, *J. Biol. Chem.*, 2009, **284**, 4041–4045.
138 G. Prelich, C. K. Tan, M. Kostura, M. B. Mathews, A. G. So, K. M. Downey and B. Stillman, *Nature*, 1987, **326**, 517–520.
139 G. Prelich, M. Kostura, D. R. Marshak, M. B. Mathews and B. Stillman, *Nature*, 1987, **326**, 471–475.
140 B. O. Petersen, J. Lukas, C. S. Sorensen, J. Bartek and K. Helin, *EMBO J.*, 1999, **18**, 396–410.
141 X. H. Hua, H. Yan and J. Newport, *J. Cell Biol.*, 1997, **137**, 183–192.
142 Z. Zhang, M. M. Spiering, M. A. Trakselis, F. T. Ishmael, J. Xi, S. J. Benkovic and G. G. Hammes, *Proc. Natl. Acad. Sci. U. S. A.*, 2005, **102**, 3254–3259.
143 W. Lee, D. Jose, C. Phelps, A. H. Marcus and P. H. von Hippel, *Biochemistry*, 2013, **52**, 3157–3170.
144 J. Yang, J. Xi, Z. Zhuang and S. J. Benkovic, *J. Biol. Chem.*, 2005, **280**, 25416–25423.
145 J. Xi, Z. Zhang, Z. Zhuang, J. Yang, M. M. Spiering, G. G. Hammes and S. J. Benkovic, *Biochemistry*, 2005, **44**, 7747–7756.
146 R. J. Sheaff and R. D. Kuchta, *Biochemistry*, 1993, **32**, 3027–3037.
147 J. R. Swart and M. A. Griep, *Biochemistry*, 1995, **34**, 16097–16106.
148 B. Maier, D. Bensimon and V. Croquette, *Proc. Natl. Acad. Sci. U. S. A.*, 2000, **97**, 12002–12007.
149 A. Goel, M. D. Frank-Kamenetskii, T. Ellenberger and D. Herschbach, *Proc. Natl. Acad. Sci. U. S. A.*, 2001, **98**, 8485–8489.
150 M. Pandey, S. Syed, I. Donmez, G. Patel, T. Ha and S. S. Patel, *Nature*, 2009, **462**, 940–943.
151 M. Manosas, M. M. Spiering, Z. Zhuang, S. J. Benkovic and V. Croquette, *Nat. Chem. Biol.*, 2009, **5**, 904–912.
152 S. M. Hamdan, B. Marintcheva, T. Cook, S. J. Lee, S. Tabor and C. C. Richardson, *Proc. Natl. Acad. Sci. U. S. A.*, 2005, **102**, 5096–5101.
153 D. E. Johnson, M. Takahashi, S. M. Hamdan, S. J. Lee and C. C. Richardson, *Proc. Natl. Acad. Sci. U. S. A.*, 2007, **104**, 5312–5317.
154 S. J. Lee, B. Marintcheva, S. M. Hamdan and C. C. Richardson, *J. Biol. Chem.*, 2006, **281**, 25841–25849.
155 Z. Debyser, S. Tabor and C. C. Richardson, *Cell*, 1994, **77**, 157–166.
156 X. Li and K. J. Marians, *J. Biol. Chem.*, 2000, **275**, 34757–34765.
157 C. A. Wu, E. L. Zechner, J. A. Reems, C. S. McHenry and K. J. Marians, *J. Biol. Chem.*, 1992, **267**, 4074–4083.

158 J. Yang, S. W. Nelson and S. J. Benkovic, *Mol. Cell*, 2006, **21**, 153–164.
159 H. J. Geertsema and A. M. van Oijen, *Curr. Opin. Struct. Biol.*, 2013, **5**, 788–793.
160 A. Robinson and A. M. van Oijen, *Nat. Rev. Microbiol.*, 2013, **11**, 303–315.
161 R. E. Georgescu, N. Y. Yao and M. O'Donnell, *FEBS Lett.*, 2010, **584**, 2596–2605.
162 A. Johnson and M. O'Donnell, *Annu. Rev. Biochem.*, 2005, **74**, 283–315.
163 K. E. Duderstadt, R. Reyes-Lamothe, A. M. van Oijen and D. J. Sherratt, *Cold Spring Harbor Perspect. Biol.*, 2013, DOI: 10.1101/cshperspect.a010157.
164 G. Wijffels, B. Dalrymple, K. Kongsuwan and N. E. Dixon, *IUBMB Life*, 2005, **57**, 413–419.
165 P. M. Schaeffer, M. J. Headlam and N. E. Dixon, *IUBMB Life*, 2005, **57**, 5–12.
166 A. Robinson, R. J. Causer and N. E. Dixon, *Curr. Drug Targets*, 2012, **13**, 352–372.
167 P. McInerney, A. Johnson, F. Katz and M. O'Donnell, *Mol. Cell*, 2007, **27**, 527–538.
168 R. Reyes-Lamothe, D. J. Sherratt and M. C. Leake, *Science*, 2010, **328**, 498–501.
169 R. E. Georgescu, I. Kurth and M. E. O'Donnell, *Nat. Struct. Mol. Biol.*, 2012, **19**, 113–116.
170 G. Lia, B. Michel and J. F. Allemand, *Science*, 2012, **335**, 328–331.
171 J. J. Belle, A. Casey, C. T. Courcelle and J. Courcelle, *J. Bacteriol.*, 2007, **189**, 5452–5462.
172 G. Lia, A. Rigato, E. Long, C. Chagneau, M. Le Masson, J. F. Allemand and B. Michel, *Mol. Cell*, 2013, **49**, 547–557.
173 R. Bermejo, M. S. Lai and M. Foiani, *Mol. Cell*, 2012, **45**, 710–718.
174 I. Kurth, R. E. Georgescu and M. E. O'Donnell, *Nature*, 2013, **496**, 119–122.
175 A. Badrinarayanan, R. Reyes-Lamothe, S. Uphoff, M. C. Leake and D. J. Sherratt, *Science*, 2012, **338**, 528–531.
176 R. Hayama, S. Bahng, M. E. Karasu and K. J. Marians, *J. Biol. Chem.*, 2013, **288**, 7653–7661.
177 J. J. Blow and P. J. Gillespie, *Nat. Rev. Cancer*, 2008, **8**, 799–806.
178 H. Yardimci, A. B. Loveland, A. M. van Oijen and J. C. Walter, *Methods*, 2012, **57**, 179–186.
179 A. B. Loveland, S. Habuchi, J. C. Walter and A. M. van Oijen, *Nat. Methods*, 2012, **9**, 987–992.
180 C. Evrin, P. Clarke, J. Zech, R. Lurz, J. Sun, S. Uhle, H. Li, B. Stillman and C. Speck, *Proc. Natl. Acad. Sci. U. S. A.*, 2009, **106**, 20240–20245.
181 D. Remus, F. Beuron, G. Tolun, J. D. Griffith, E. P. Morris and J. F. Diffley, *Cell*, 2009, **139**, 719–730.
182 M. L. Bochman and A. Schwacha, *Microbiol. Mol. Biol. Rev.*, 2009, **73**, 652–683.
183 Y. V. Fu, H. Yardimci, D. T. Long, T. V. Ho, A. Guainazzi, V. P. Bermudez, J. Hurwitz, A. van Oijen, O. D. Scharer and J. C. Walter, *Cell*, 2011, **146**, 931–941.
184 A. Costa, I. Ilves, N. Tamberg, T. Petojevic, E. Nogales, M. R. Botchan and J. M. Berger, *Nat. Struct. Mol. Biol.*, 2011, **18**, 471–477.
185 K. Luger, A. W. Mader, R. K. Richmond, D. F. Sargent and T. J. Richmond, *Nature*, 1997, **389**, 251–260.
186 J. L. Killian, M. Li, M. Y. Sheinin and M. D. Wang, *Curr. Opin. Struct. Biol.*, 2012, **22**, 80–87.
187 M. A. Hall, A. Shundrovsky, L. Bai, R. M. Fulbright, J. T. Lis and M. D. Wang, *Nat. Struct. Mol. Biol.*, 2009, **16**, 124–129.
188 M. Simon, J. A. North, J. C. Shimko, R. A. Forties, M. B. Ferdinand, M. Manohar, M. Zhang, R. Fishel, J. J. Ottesen and M. G. Poirier, *Proc. Natl. Acad. Sci. U. S. A.*, 2011, **108**, 12711–12716.
189 J. S. Choy, S. Wei, J. Y. Lee, S. Tan, S. Chu and T. H. Lee, *J. Am. Chem. Soc.*, 2010, **132**, 1782–1783.
190 B. Akabayov, S. R. Akabayov, S. J. Lee, G. Wagner and C. C. Richardson, *Nat. Commun.*, 2013, **4**, 1615.
191 Y. Xu, T. J. Melia and D. K. Toomre, *Curr. Opin. Chem. Biol.*, 2011, **15**, 822–830.
192 N. A. Tanner, S. M. Hamdan, S. Jergic, K. V. Loscha, P. M. Schaeffer, N. E. Dixon and A. M. van Oijen, *Nat. Struct. Mol. Biol.*, 2008, **15**, 998.
193 R. D. Smiley, Z. Zhuang, S. J. Benkovic and G. G. Hammes, *Biochemistry*, 2006, **45**, 7990–7997.
194 S. M. Hamdan and A. M. van Oijen, *J. Biol. Chem.*, 2010, **285**, 18979–18983.
195 H. Yardimci, A. B. Loveland, S. Habuchi, A. M. van Oijen and J. C. Walter, *Mol. Cell*, 2010, **40**, 834–840.