Characterization of DNA-replication proteins and their molecular mechanisms: a team business

Botto, M.

Chapter 1

General Introduction

1.1 Organization and conservation of the Genome

The genome is the building block of life for every living organism. It contains all the genetic information needed to build that organism, allowing it to grow, develop and function. It is composed by DNA, a polymer consisting of two polynucleotide chains that coil around each other to form a double helix (Fig. 1). The entire human genome resides in 23 pairs of chromosomes within the nucleus of all our cells. Each chromosome contains hundreds to thousands of genes, which carry the instructions for making proteins.

To be able to develop and grow, every organism must replicate its entire genome whenever a cell divides so that each daughter cell obtains a complete set of genetic information herself.

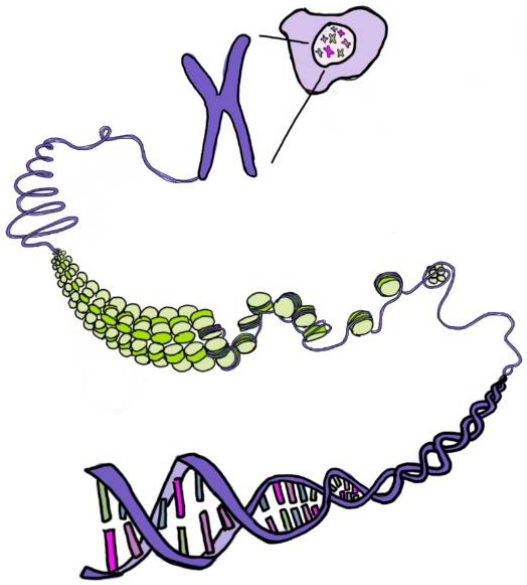

Figure 1. *The compaction of DNA. To fit all the DNA in the nucleus, it is compacted in multiple steps. In the first level of compaction, DNA is wrapped around nucleosomes to form a beads-on-a-string conformation. Subsequently, the nucleosomes interact to form chromatin fibers. Ultimately, during mitosis chromatin fibers can interact and condense to form chromosomes, the highest level of compaction.*

DNA is replicated to make almost perfect copies of itself, which is a remarkable task considering that in humans there are almost three billion base pairs of DNA to be copied.
It is vital that DNA is protected from errors or mutations that could otherwise have deleterious effect on the cell's development and growth. This is observed in some human diseases like cancer or neurodegenerative diseases that can be caused by the accumulation of errors in DNA resulting in the altered behavior of cells.
Nevertheless, changes in DNA are not always damaging. In fact, mutations are essential for evolution. Every genetic feature in every organism is the result of an initial accidental mutation. If a trait given by a new mutation is advantageous and helps the organism to survive and reproduce, the genetic variation is more likely to be passed on to the next generation. Over time, as generations of individuals with the trait continue to reproduce, the advantageous trait becomes increasingly common in a population, making the population different from the ancestral one.
Therefore, during DNA replication there needs to be a delicate balance between DNA preservation and insertion of mutations.

1.2 DNA replication

DNA replication is the biological process in which two identical copies of DNA are produced from one original DNA molecule. After the discovery of DNA structure by Watson and Crick[1], it was possible to unveil its replication mechanism. DNA was shown to be a right-handed helix consisting of two anti-parallel strands of deoxyribonucleotides (Fig. 2a). During DNA replication inside each cell the double stranded DNA is unwound and separated into two single strands. Each of the two DNA strands serves as a template for the formation of an entire new strand (Fig. 2b). The replication of each individual strand is performed by proteins called DNA polymerases that are molecules specifically dedicated to copying DNA by using deoxyribonucleoside triphosphates.
Because the genomic information for which the DNA codes has to be duplicated correctly for the next generation of healthy cells, the DNA polymerases act with a very high fidelity. They produce an error in the daughter strand with a frequency of 10^{-6} [2]. Daughter strand errors originate from the incorrect selection of base pair, leading to point mutations if not corrected. To avoid this, associated with the replicative polymerases there is a 3' to 5' replicative exonuclease (reviewed in [3]). This class of proteins excise the misincorporated nucleotide before the polymerase will resume DNA synthesis. If this process fails, a post replicative mismatch repair

system (MMR) is available to the cell to remove mistakes on the daughter strand after the polymerase has finished its synthesis task (reviewed in [4]).

Overall, DNA replication occurs in three coordinated steps: initiation, elongation and termination. The process begins at specific points in the DNA sequence known as origins of replication (ori) (reviewed in [5]). Here different proteins act to initiate DNA replication. In *E. coli* a protein called DnaA recognizes the origins of replication and starts to unwind them[6–9]. Next, DnaA interacts with the helicase DnaB and with the helicase loader DnaC by loading one DnaB-DnaC complex on each separated DNA strand[10–13]. Afterwards the interaction of the primase DnaG[14] with the helicase DnaB leads to the first primer formation, which is required for DNA elongation to start.

During elongation, while the original strands get separated, the two daughter strands are synthetized by the replicative DNA polymerases. Since the two parental strands run in an antiparallel direction, continuous DNA synthesis of the two new strands would require one strand to be synthetized in the 5' to 3' direction and the other in the 3' to 5' direction.

However DNA polymerases can synthetize DNA only in a 5' to 3' manner[15,16]. Hence one strand of DNA is synthetized in a continuous manner, the leading strand, while the other one, the lagging strand, is formed in short pieces of DNA called Okazaki fragments[17,18](Fig. 2b). Moreover, as DNA polymerases cannot start DNA synthesis *de novo,* every DNA fragment is started with an RNA primer previously synthetized by an RNA primase[19].

Termination of the DNA replication process requires the blocking of the progress of the replication fork. This occurs at special termination sites in the DNA sequence (reviewed in [20]) where a specific protein binds to the DNA and blocks the replication process. In *E. coli* this task is performed by the Tus protein that recognizes and binds to the specific Ter termination sites [21].

1.2.1 The *E. coli* replisome

DNA replication is conserved between species, and it is mostly studied in the *E. coli* system. The work in this thesis is focused on *E. coli*, therefore the discussion will continue only about the *E. coli* replisome.

One of the key players of bacterial DNA replication is the replicative DNA polymerase, that has first been identified and described in the 1970's[22–24]. In *E. coli* it is called the α subunit of the DNA polymerase III (Pol III) holoenzyme (Fig. 3), a complex of at least ten proteins that together orchestrate the synthesis of a new DNA strand[25].

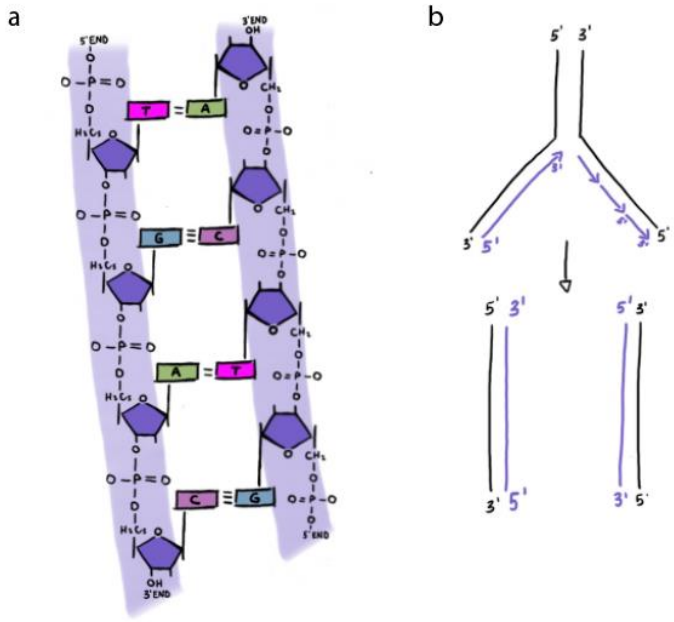

Figure 2. *The structure of DNA reveals its copying mechanism.* ***(a)*** *Schematic representation of the antiparallel double helical structure of DNA showing the sugar-phosphate backbone and complementary base pairing between bases.* ***(b)*** *Schematic representation of the DNA replication mechanism. In black are the two old DNA strands and in purple the newly synthetized ones.*

Pol IIIα, the *E. coli* replicative polymerase, is able to synthetizes DNA at speeds of up to 1000 base pairs (bp) per second with an error rate of 10^{-6} errors per base per round of replication[2,26]. To synthesize very long stretches of DNA it needs to associate with a DNA sliding clamp, called β, that helps the polymerase remain attached to the DNA, enhancing the processivity of the polymerase[27].The β clamp is a ring-shaped homo-dimeric protein[28]. Because of its structure, it needs to be loaded onto the DNA by the clamp loader complex[29], which is a pentameric complex that consists of three copies of τ and one δ and one δ' subunit[30]. The clamp loader complex is an ATPase that uses ATP hydrolysis to disrupt one β clamp dimer interface to break its ring structure to enclose it on DNA[29]. The complex plays an important role in coordinating the different subunits present in the holoenzyme by connecting the leading and lagging strand DNA polymerase to the DNA helicase, in addition to its clamp loading function (reviewed in[31]). Its subunit τ is connected to Pol IIIα and the helicase via separate domains[32,33]. By doing so, it ensures that the

individual activities performed by the holoenzyme are coupled together to achieve a greater efficiency in DNA synthesis.

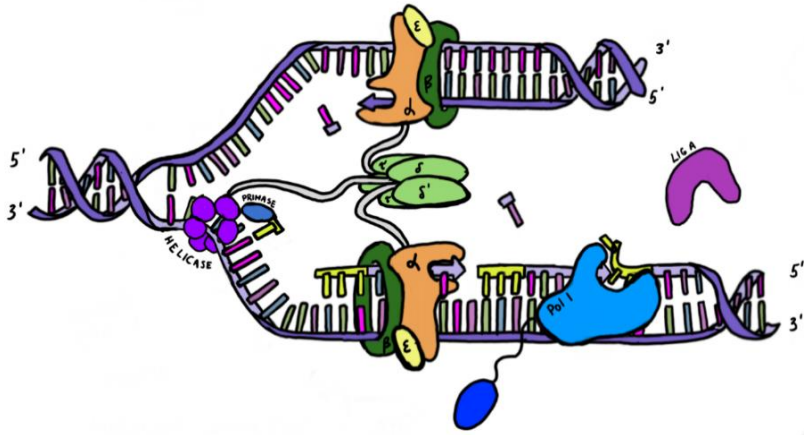

Figure 3. *DNA replication in E. coli. Scheme showing leading and lagging strand synthesis with the involved proteins.*

However, Pol IIIα can't synthesize polynucleotides *de novo*, i.e. linking two nucleotides to start a strand, and thus requires an already existing stretch of nucleotides from which it can elongate[34]. To overcome this problem, a primase called DnaG, which is attached to the helicase, will synthetize small stretches of ribonucleic acid (RNA), which serve as starting points for DNA synthesis performed by the DNA polymerase[35].

Finally attached to α is the replicative exonuclease, the ε subunit, which removes nucleotides that are misincorporated by Pol IIIα[25,36]. By doing so, it helps the polymerase to maintain its high fidelity by cleaving the newly inserted error from the newly synthetized strand.

1.2.2 The 'roadblock remover' exonuclease ε

Genome integrity is crucial to all forms of life and a mechanism to preserve it is vital. To avoid the insertion of mistakes into the DNA, replicative DNA polymerases are high fidelity enzymes. As a result, when they do make a mistake, they are not able to continue their DNA synthesis. Therefore, there are replicative exonucleases to remove the mistake and act like "roadblock removers".

Exonucleases remove an error as soon as it is inserted into the newly synthesized double stranded DNA by cleaving the phosphodiester bond of nucleic acids from a free 3' or 5' end.

In *E. coli* the replicative exonuclease ε is essential for the bacterium's survival[3,37]. It binds to the replicative polymerase Pol IIIα with which, together with θ, it forms the so-called Pol III core[25]. While α synthetizes DNA and ε removes any mistake made by the polymerase in a 3' to 5' direction, θ has no other known activity other than stabilizing ε[38,39].

When in complex with α, ε activity was shown to be 18 times higher on a mismatched termini compared to its activity in isolation[40]. On the other hand ε has been shown to stimulate Pol IIIα activity[41,42] and to help to achieve a high fidelity DNA synthesis. Therefore, the activity of the two proteins is closely intertwined and dependent one another.

In the rare occasions in which an incorrect nucleotide is incorporated during DNA synthesis, the distorted geometry of the mismatched base pair prevents further extension of the DNA strand[43]. Given the stall of Pol IIIα, the new error should be corrected to prevent its incorporation in the double stranded DNA. However, the exonuclease active site is ~60 Å away from the polymerase active site[44]. The comparison between the DNA synthesis conformation and the nucleotide removal conformation (Fig. 5) shows us a drastic conformational change of the DNA, that leads to the positioning of the 3' end of the new DNA strand in the exonuclease active site[45]. This dramatic change whereby the 3' end of the newly synthesized strand moves over 50 Å raises the question how the DNA moves from the polymerase active site to the exonuclease site when a mismatch is encountered. Chapter 2 describes our analysis of the movement of DNA during exonuclease activity, which reveals that the DNA is guided by a predefined path between the polymerase and exonuclease active site.

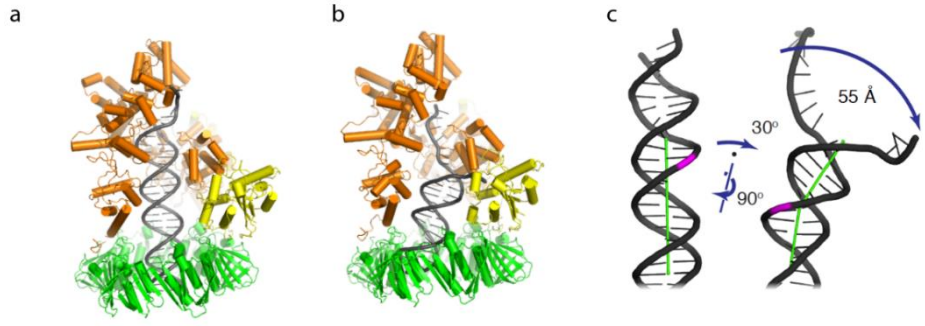

Figure 5. *Structures of Pol IIIα (in orange), exonuclease ε (in yellow) and β clamp (in green) in synthesis mode* ***(a)*** *and correction mode* ***(b)****.* ***(c)*** *Difference between the*

DNA conformation in the DNA synthesis and correction modes. Structures from Fernandez Leiro et al., NSMB, 2017 [45]

1.2.3 Lagging strand synthesis

As described above, because of the opposite directionality of the two DNA strands, the leading and lagging strand have to be synthetized in different ways. While the leading strand can be synthesized continuously, the lagging strand has to be synthesized in small segments called Okazaki fragments[17,18]. Each fragment consists of a short stretch of RNA (8-15 bp)[46] and a much longer stretch of DNA (between 1000 and 2000 bp in bacteria)[47].

Polymerase Pol IIIα starts its DNA synthesis activity from an RNA primer but it comes to a halt when it encounters the next one due to its intrinsically low DNA binding affinity[48–51]. However, the RNA primers embedded into the lagging strand have to be removed in order to have a continuous double stranded DNA.

DNA polymerase I (Pol I), is a polymerase that together with Pol IIIα works on the synthesis of the DNA lagging strand. Contrary to Pol IIIα, Pol I is able to continue DNA synthesis despite the presence of a block due to its ability of displacing and removing the RNA primers present on the DNA lagging strand[52,53]. Pol I was the first nucleotide polymerizing enzyme ever discovered in 1957[15]. Since then, it has been revealed that it performs four distinct enzymatic activities: DNA synthesis, 3' to 5' exonuclease activity, primer displacement and primer removal. To do so, Pol I contains multiples domains (Fig. 4). Three of these domains are referred to as thumb, fingers and palm domains that together perform DNA polymerase activity[54]. A fourth domain next to the palm domain contains the 3' - 5' exonuclease active site that removes incorrectly incorporated nucleotides[55–57]. Finally, a fifth domain contains the endonuclease activity that removes the single stranded DNA or RNA flap that is created by the polymerase during strand displacement activity[58,59].

Once the RNA primer has been removed by the Pol I strand displacement and endonuclease activity, a ligase, called LigA has to seal the nick left on the DNA in order to create a continuous lagging strand[60]. LigA catalyses the formation of a phosphodiester linkage between 5'-phosphoryl and 3'-hydroxyl groups in double-stranded DNA using NAD as a coenzyme and energy source[61].

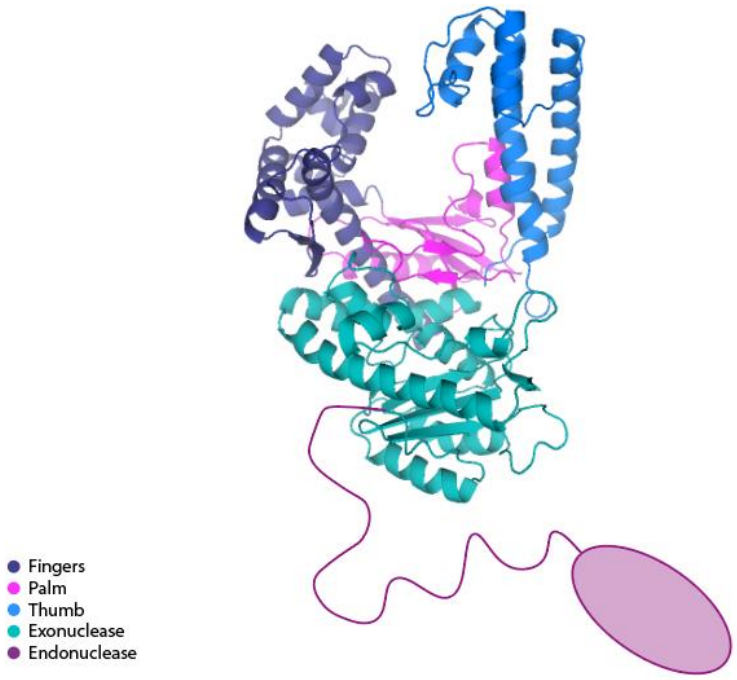

Figure 4. *Crystal structure of E. coli Pol I. The endonuclease is not present in the original structure because of its flexible nature. Structure from Ollis et al., Nature, 1985*[54].

Surprisingly, even though the activity of the individual proteins involved in the Okazaki fragment maturation has been described extensively, how they work together is still poorly understood. In Chapter 3, I will present our structural analysis of Pol I and our biochemical analysis of the coordination between Pol I, Pol IIIα, β clamp and the ligase LigA.

1.3 Drug resistance

While DNA replication avoids errors, at the same time errors are needed for an organism to evolve. For example, point mutations in the DNA can cause antibiotic drug resistance in bacteria by altering the target site of an inhibitor, thus enabling the bacterium to survive in the presence of drugs developed against them (reviewed in[62]). In recent years drug resistance became a worldwide concern (Fig. 6) and it has

forced the scientific community to try to urgently find new drug targets. Among the antibiotics already discovered, some inhibit RNA transcription, other RNA translation or DNA replication. Transcription and translation are inhibited directly by targeting the RNA polymerase or ribosome, respectively. DNA replication, in contrast, is inhibited indirectly through targeting of DNA gyrases[63]. Therefore, there are currently no antibiotics inhibiting DNA replication by directly targeting the replisome.

However, in recent years, several replisome proteins have been targeted to find novel inhibitors that may potentially evolve to become novel antibiotics. Some examples include the DNA helicase DnaB, the sliding clamp β, the RNA primase DnaG, the DNA ligase LigA and the replicative DNA polymerases Pol C and DnaE[64–66].

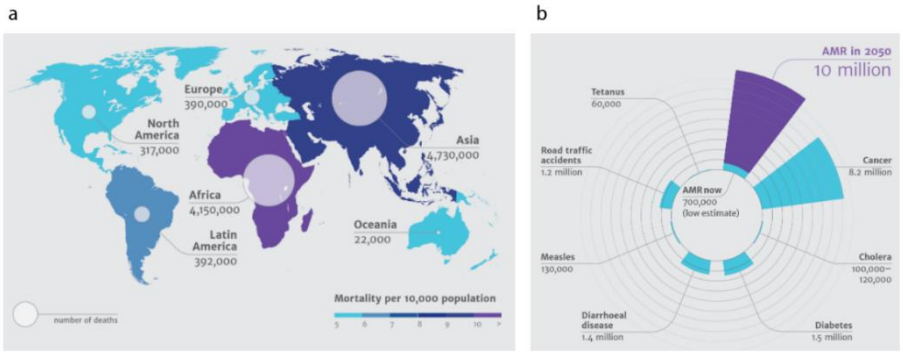

Figure 6. ***Drug resistance is becoming a major treat for humanity. (a)*** *Predicted deaths attributable to antibiotic resistance by 2050.* ***(b)*** *Deaths attributable to antibiotic resistance compared to other major causes of death. Data from 'Review on Antimicrobial resistance 2014'* [67].

1.3.1 DNA exonucleases as a drug target

The biological targets for antibiotics should be essential enzymes for bacterial growth and propagation but distinct enough from the ones in humans. In this way the inactivation of these enzymes doesn't have any consequence in humans[65]. New antibiotics with novel modes of action are required to combat the growing threat posed by drug resistant bacteria. Because of their fundamental role in DNA replication[57,68], exonucleases are potentially attractive targets for novel antibiotics. As every cell starts its growth with the replication of its genome, the inhibition of DNA replication is, quite literally, the most direct way to prevent the growth of a

new cell. It has been shown that when exonuclease activity is deleted in *E. coli*, *Salmonella typhimurium* or *M. tuberculosis*, the bacteria become compromised with strongly reduced growth rates and increased mutation rates[69–71].

Recently, novel inhibitors have been discovered that target directly the exonuclease domain of the *M. tuberculosis* replicative DNA polymerase DnaE1[72]. Inhibition of the *M. tuberculosis* DnaE1-exonuclease is especially attractive as it is structurally distinct from the human exonuclease and their active site consists of a deep pocket that is well suited for binding of inhibitors[73].

Therefore, to find novel inhibitors it would be a valuable tool to have a fast and efficient way to monitor the exonuclease activity in presence of possible inhibitors. However, up to now, no suitable exonuclease assays exist.

In Chapter 4 I introduce the 2-AP assay, a novel high throughput screening assay that will permit us to follow exonuclease activity in real time.

1.4 A single molecule approach to biology

The visualization and consequent interpretation of microscopic worlds has often fueled revolutionary discoveries in our journey throughout the understanding of nature. From the very first microscopes built in the 16th and 17th centuries to the modern Electron microscopes, scientists have always tried to push the limit of our understanding of the microscopic world. Following this desire, a large range of so-called single molecule techniques have been developed since the 1970s. They have since then contributed to our understanding of important biological processes like protein folding[74,75], molecular motors[76,77] and gene expression[78,79].

The advent of single molecule techniques have been revolutionary in the molecular biology field and it has changed the way we understand biological processes: instead of thinking in terms of group behaviors, we can now understand biological processes at single molecules level. For example, this is important to understand multi-step enzymatic reactions or ensembles with heterogeneous behaviors. In the first case by using single molecule techniques, we will be able to distinguish the sequential enzymatic steps involved in the process. In the second one we will be able to reveal a distribution of different behavioral states rather than a single state resulting from the average of all the different proteins behaviors (Fig. 7).

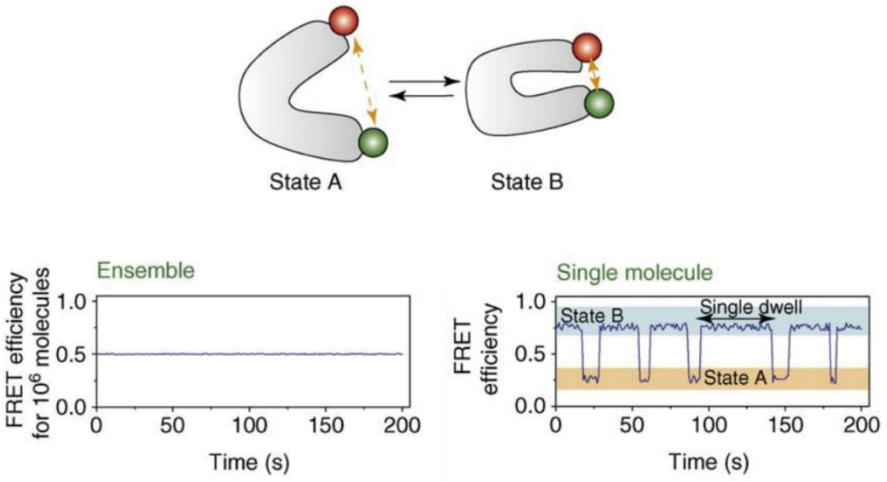

Figure 7. ***Single molecule methods reveal underlying protein dynamics****. In the ensemble experiment the two different protein conformations have been averaged out while the single molecule experiment reveals the transition between the two states. Figure taken from Kapanidis and Strick, 2009* [80]***.***

The existing single molecule techniques can be broadly divided into two main categories: force-based and fluorescent-based and the combination of the two. Protein dynamics on DNA have been studied using both force-based and fluorescence-based techniques, each with their advantages and detriments. By using a force-based technique, which include optical and magnetic tweezers, we would be able to manipulate single molecules to interrogate and alter the energy landscape of reactions, therefore directly examining elusive structural elements like transition states[81–83]. By using a fluorescence-based technique such as single molecule FRET and multi-color co-localization single molecule spectroscopy (CoSMoS), we would have the advantage of directly visualize the DNA and proteins in real time, monitoring their location, their structure and dynamics[84–87].

The *E. coli* replisome is composed of many different proteins that act as a team to catalyze multi-step enzymatic reactions. Even if its activity is known from decades (reviewed in [88]), many questions remain unanswered. Our final goal is to understand the molecular details of how the transition between different steps occur and how they are temporally organized. Specifically, we would like to study the organization of the different proteins that are part of the Okazaki maturation process. For this we decided to use a single molecule fluorescence technique, CoSMoS[89–91]. In

Chapter 5 I introduce the technique and show how I have implemented and employed it to initiate single molecules studies of the behavior of the polymerase Pol I and the ligase LigA on a DNA molecule.

References

1. Crick, F. & Watson, J. Molecular structures of nucleic acids. *Nature* (1953).
2. Kunkel, T. A. DNA Replication Fidelity. *J. Biol. Chem.* **279**, 16895–16898 (2004).
3. Shevelev, I. V. & Hübscher, U. The 3'-5' exonucleases. *Nat. Rev. Mol. Cell Biol.* **3**, 364–375 (2002).
4. Li, G. M. Mechanisms and functions of DNA mismatch repair. *Cell Res.* **18**, 85–98 (2008).
5. Leonard, A. C. & Mechali, M. DNA Replication Origins. *Cold Spring Harb. Perspect. Biol.* 1–17 (2013).
6. Fuller, R. S., Funnell, B. E. & Kornberg, A. The dnaA protein complex with the E. coli chromosomal replication origin (oriC) and other DNA sites. *Cell* **38**, 889–900 (1984).
7. Crooke, E., Thresher, R., Hwang, D. S., Griffith, J. & Kornberg, A. Replicatively active complexes of DnaA protein and the Escherichia coli chromosomal origin observed in the electron microscope. *Journal of Molecular Biology* **233**, 16–24 (1993).
8. Margulies, C. & Kaguni, J. M. Ordered and sequential binding of DnaA protein to oriC, the chromosomal origin of Escherichia coli. *J. Biol. Chem.* **271**, 17035–17040 (1996).
9. Weigel, C., Schmidt, A., Rückert, B., Lurz, R. & Messer, W. DnaA protein binding to individual DnaA boxes in the Escherichia coli replication origin, oriC. *EMBO J.* **16**, 6574–6583 (1997).
10. Marszalek, J. & Kaguni, J. M. DnaA protein directs the binding of DnaB protein in initiation of DNA replication in Escherichia coli. *J. Biol. Chem.* **269**, 4883–4890 (1994).
11. Sutton, M. D., Carr, K. M., Vicente, M. & Kaguni, J. M. Escherichia coli DnaA protein. The N-terminal domain and loading of DnaB helicase at the E. coli chromosomal origin. *J. Biol. Chem.* **273**, 34255–34262 (1998).
12. Seitz, H., Weigel, C. & Messer, W. The interaction domains of the DnaA and DnaB replication proteins of Escherichia coli. *Mol. Microbiol.* **37**, 1270–1279 (2000).
13. Abe, Y. *et al.* Structure and function of DnaA N-terminal domains: Specific sites and mechanisms in inter-DnaA interaction and in DnaB helicase loading on oriC. *J. Biol. Chem.* **282**, 17816–17827 (2007).
14. Tougu, K. & Marians, K. J. The interaction between helicase and primase sets the replication fork clock. *J. Biol. Chem.* **271**, 21398–21405 (1996).

15. Lehman I. R. ; Kornberg, A. Enzymatic synthesis of Deoxyribonucleic acid. *J. Biol. Chem.* **237**, 519–525 (1957).
16. Bessman, M. J., Lehman, I. R., Simms, E. S. & Kornberg, A. Enzymatic Synthesis of Deoxyribonucleic acid: General properties of the reaction. *J. Biol. Chem.* **237**, 519–525 (1957).
17. Sakabe, K. & Okazaki, R. A unique property of the replicating region of chromosomal DNA. *BBA Sect. Nucleic Acids Protein Synth.* **129**, 651–654 (1966).
18. Okazaki, R. Mechanism of DNA chain growth. *J. Mol. Biol.* **95**, 63–70 (1967).
19. Kuchta, R. D. & Stengel, G. Mechanism and evolution of DNA primases. *Biochim. Biophys. Acta - Proteins Proteomics* **1804**, 1180–1189 (2010).
20. Dewar, J. M. & Walter, J. C. Mechanisms of DNA replication termination. *Nat. Rev. Mol. Cell Biol.* **18**, 507–516 (2017).
21. Hill, T. M. & Marians, K. J. Escherichia coli Tus protein acts to arrest the progression of DNA replication forks in vitro. *Proc. Natl. Acad. Sci. U. S. A.* **87**, 2481–2485 (1990).
22. Kornberg, T. & Gefter, M. L. DNA synthesis in cell-free extracts of a DNA polymerase-defective mutant. *Biochem. Biophys. Res. Commun.* **40**, 1348–1355 (1970).
23. McHenry, C. & Kornberg, A. DNA polymerase III holoenzyme of Escherichia coli. Purification and resolution into subunits. *J. Biol. Chem.* **252**, 6478–6484 (1977).
24. McHenry, C. S. & Crow, W. DNA polymerase III of Escherichia coli. Purification and identification of subunits. *J. Biol. Chem.* **254**, 1748–1753 (1979).
25. McHenry, C. S. Bacterial replicases and related polymerases. *Curr. Opin. Chem. Biol.* **15**, 587–594 (2011).
26. Schaaper, R. M. Base selection, proofreading, and mismatch repair during DNA replication in Escherichia coli. *J. Biol. Chem.* **268**, 23762–23765 (1993).
27. Stukenberg, P. T., Studwell-Vaughan, P. S. & O'Donnell, M. Mechanism of the sliding β-clamp of DNA polymerase III holoenzyme. *J. Biol. Chem.* **266**, 11328–11334 (1991).
28. Oakley, A. J. *et al.* Flexibility revealed by the 1.85 Å crystal structure of the β sliding-clamp subunit of Escherichia coli DNA polymerase III. *Acta Crystallogr. - Sect. D Biol. Crystallogr.* **59**, 1192–1199 (2003).
29. O'Donnell, M., Jeruzalmi, D. & Kuriyan, J. Clamp loader structure predicts the architecture of DNA polymerase III holoenzyme and RFC. *Curr. Biol.* **11**, 935–946 (2001).

30. Pritchard, A. E. A novel assembly mechanism for the DNA polymerase III holoenzyme DnaX complex: association of deltadelta' with DnaX4 forms DnaX3deltadelta'. *EMBO J.* **19**, 6536–6545 (2000).
31. McHenry, C. S. DNA polymerase III holoenzyme of Escherichia coli: Components and function of a true replicative complex. *Mol. Cell. Biochem.* **66**, 71–85 (1985).
32. Gao, D. & McHenry, C. S. τ binds and organizes Escherichia coli replication proteins Through distinct domains. Partial proteolysis of terminally tagged τ to determine candidate domains and to assign domain V as the α binding domain. *J. Biol. Chem.* **276**, 4433–4440 (2001).
33. Gao, D. & McHenry, C. S. τ binds and organizes Escherichia coli replication proteins through distinct domains. Domain IV, located within the unique C terminus of τ, binds the replication fork helicase, DnaB. *J. Biol. Chem.* **276**, 4441–4446 (2001).
34. Kitani, T., Yoda, K. ya, Ogawa, T. & Okazaki, T. Evidence that discontinuous DNA replication in Escherichia coli is primed by approximately 10 to 12 residues of RNA starting with a purine. *J. Mol. Biol.* **184**, 45–52 (1985).
35. Bouche, J. P., Zechel, K. & Kornberg, A. dnaG gene product, a rifampicin resistant RNA polymerase, initiates the conversion of a single stranded coliphage DNA to its duplex replicative form. *J. Biol. Chem.* **250**, 5995–6001 (1975).
36. McHenry, C. S. Purification and characterization of DNA polymerase III'. Identification of τ as a subunit of the DNA polymerase III holoenzyme. *J. Biol. Chem.* **257**, 2657–2663 (1982).
37. Fijalkowska, I. J. & Schaaper, R. M. Mutants in the exo I motif of Escherichia coli dnaQ: Defective proofreading and inviability due to error catastrophe. *Proc. Natl. Acad. Sci. U. S. A.* **93**, 2856–2861 (1996).
38. Studwell-Vaughan, P. S. & O'Donnell, M. DNA polymerase III accessory proteins. V. θ encoded by holE. *J. Biol. Chem.* **268**, 11785–11791 (1993).
39. Taft-Benz, S. A. & Schaaper, R. M. The θ Subunit of Escherichia coli DNA Polymerase III: A Role in Stabilizing the ε Proofreading Subunit. *J. Bacteriol.* **186**, 2774–2780 (2004).
40. Maki, H. & Kornberg, A. Proofreading by DNA polymerase III of Escherichia coli depends on cooperative interaction of the polymerase and exonuclease subunits. *Proc. Natl. Acad. Sci. U. S. A.* **84**, 4389–4392 (1987).
41. Kim, D. R. & McHenry, C. S. Identification of the β-binding domain of the α subunit of Escherichia coli polymerase III holoenzyme. *J. Biol. Chem.* **271**, 20699–20704 (1996).

42. O'Donnell, M. & Studwell, P. S. Total reconstruction of DNA polymerase III holoenzyme reveals dual accessory protein clamps. *J. Biol. Chem.* **265**, 1179–1187 (1990).
43. Johnson, S. J. & Beese, L. S. Structures of mismatch replication errors observed in a DNA polymerase. *Cell* **116**, 1–14 (2004).
44. Fernandez-leiro, R., Conrad, J., Scheres, S. H. W. & Lamers, M. H. cryo-EM structures of the E . coli replicative DNA polymerase reveal its dynamic interactions with the DNA sliding clamp , exonuclease and t. *Elife* 1–16 (2015). doi:10.7554/eLife.11134
45. Fernandez-leiro, R., Conrad, J., Yang, J., Freund, S. M. V & Sjors, H. W. Self-correcting mismatches during high-fidelity DNA replication. *Nat. Struct. Mol. Biol.* **24**, 140–143 (2017).
46. Grosse, F. & Krauss, G. The primase activity of DNA polymerase α from calf thymus. *J. Biol. Chem.* **260**, 1881–1888 (1985).
47. Ogawa, T. & Okazaki, T. Discontinuous DNA replication. *Annu. Rev. Biochem.* (1980).
48. Wu, C. A., Zechner, E. L., Reems, J. A., McHenry, C. S. & Marians, K. J. Coordinated leading- and lagging-strand synthesis at the Escherichia coli DNA replication fork. V. Primase action regulates the cycle of Okazaki fragment synthesis. *J. Biol. Chem.* **267**, 4074–4083 (1992).
49. Stukenberg, P. T., Turner, J. & O'Donnell, M. An explanation for lagging strand replication: Polymerase hopping among DNA sliding clamps. *Cell* **78**, 877–887 (1994).
50. Georgescu, R. E. *et al.* Mechanism of polymerase collision release from sliding clamps on the lagging strand. *EMBO J.* **28**, 2981–2991 (2009).
51. Dohrmann, P. R., Manhart, C. M., Downey, C. D. & McHenry, C. S. The rate of polymerase release upon filling the gap between Okazaki fragments is inadequate to support cycling during lagging strand synthesis. *J. Mol. Biol.* **414**, 15–27 (2011).
52. Patel, P. H., Suzuki, M., Adman, E., Shinkai, A. & Loeb, L. A. Prokaryotic DNA polymerase I: Evolution, structure, and 'base flipping' mechanism for nucleotide selection. *J. Mol. Biol.* **308**, 823–837 (2001).
53. Singh, K., Srivastava, A., Patel, S. S. & Modak, M. J. Participation of the fingers subdomain of Escherichia coli DNA polymerase I in the strand displacement synthesis of DNA. *J. Biol. Chem.* **282**, 10594–10604 (2007).
54. Ollis, D. L., Kline, C. & Steitz, T. A. Domain of E. coli DNA polymerase I showing sequence homology to T7 DNA polymerase. *Nature* **313**, 818–819 (1985).

55. Brutlag, D., Atkinson, M. R., Setlow, P. & Kornberg, A. An active fragment of DNA polymerase produced by proteolytic cleavage. *Biochem. Biophys. Res. Commun.* **37**, 982–989 (1969).
56. Klenow, H. & Henningsen, I. Selective elimination of the exonuclease activity of the deoxyribonucleic acid polymerase from Escherichia coli B by limited proteolysis. *Proc. Natl. Acad. Sci. U. S. A.* **65**, 168–175 (1970).
57. Brutlag, D. Enzymatic Synthesis of Deoxyribonucleic acid: A proofreading function for the 3'to 5'exonuclease activity in deoxyribonucleic acid polymerases. *J. Biol. Chem.* **247**, 241–248 (1972).
58. Cozzarelli, N. R., Kelly, R. B. & Kornberg, A. Enzymic synthesis of DNA. XXXIII. Hydrolysis of a 5'-triphosphate-terminated polynucleotide in the active center of DNA polymerase. *J. Mol. Biol.* **45**, 513–531 (1969).
59. Xu, Y. *et al.* Biochemical and mutational studies of the 5'-3' exonuclease of DNA polymerase I of Escherichia coli. *J. Mol. Biol.* **268**, 284–302 (1997).
60. Olivera, B. M. & Lehman, I. R. Linkage of polynucleotides through phosphodiester bonds by an enzyme from Escherichia coli. *Proc. Natl. Acad. Sci. U. S. A.* **57**, 1426–1433 (1967).
61. Wilkinson, A. *et al.* Analysis of ligation and DNA binding by Escherichia coli DNA ligase (LigA). *Biochim. Biophys. Acta - Proteins Proteomics* **1749**, 113–122 (2005).
62. Christaki, E., Marcou, M. & Tofarides, A. Antimicrobial Resistance in Bacteria: Mechanisms, Evolution, and Persistence. *J. Mol. Evol.* **88**, 26–40 (2020).
63. Pham, T. D. M., Ziora, Z. M. & Blaskovich, M. A. T. Quinolone antibiotics. *Med. Chem. Commun.* **10**, 1719–1739 (2019).
64. Reiche, M. A., Warner, D. F. & Mizrahi, V. Targeting DNA replication and repair for the development of novel therapeutics against tuberculosis. *Front. Mol. Biosci.* **4**, 1–18 (2017).
65. Robinson, A., J. Causer, R. & E. Dixon, N. Architecture and Conservation of the Bacterial DNA Replication Machinery, an Underexploited Drug Target. *Curr. Drug Targets* **13**, 352–372 (2012).
66. Kaguni, J. M. The macromolecular machines that duplicate the escherichia coli chromosome as targets for drug discovery. *Antibiotics* **7**, 1–28 (2018).
67. Neill, J. O. '. Antimicrobial Resistance: Tackling a crisis for the health and wealth of nations. *Rev. Antimicrob. Resist.* (2014).
68. Bębenek, A. & Ziuzia-Graczyk, I. Fidelity of DNA replication—a matter of proofreading. *Curr. Genet.* **64**, 985–996 (2018).
69. Scheuermann, R. H. & Echols, H. A separate editing exonuclease for DNA

replication: The ε subunit of Escherichia coli DNA polymerase III holoenzyme. *Proc. Natl. Acad. Sci. U. S. A.* **81**, 7747–7751 (1984).

70. Lancy, E. D., Lifsics, M. R., Kehres, D. G. & Maurer, R. Isolation and characterization of mutants with deletions in dnaQ, the gene for the editing subunit of DNA polymerase III in Salmonella typhimurium. *J. Bacteriol.* **171**, 5572–5580 (1989).
71. Rock, J. M. *et al.* DNA replication fidelity in Mycobacterium tuberculosis is mediated by an ancestral prokaryotic proofreader. *Nat. Genet.* **47**, 677–681 (2015).
72. Nasir, N. & Kisker, C. Mechanistic insights into the enzymatic activity and inhibition of the replicative polymerase exonuclease domain from Mycobacterium tuberculosis. *DNA Repair (Amst).* **74**, 17–25 (2019).
73. Baños-mateos, S. *et al.* High-fidelity DNA replication in Mycobacterium tuberculosis relies on a trinuclear zinc center. *Nat. Commun.* (2017). doi:10.1038/s41467-017-00886-w
74. Kellermayer, M. S. Z., Smith, S. B., Granzier, H. L. & Bustamante, C. Folding-unfolding transitions in single titin molecules characterized with laser tweezers. *Science (80-.).* **276**, 1112–1116 (1997).
75. Rhoades, E., Gussakovsky, E. & Haran, G. Watching proteins fold one molecule at a time. *Proc. Natl. Acad. Sci. U. S. A.* **100**, 3197–3202 (2003).
76. Yildiz, A. *et al.* Myosin V walks hand-over-hand: Single fluorophore imaging with 1.5-nm localization. *Science (80-.).* **300**, 2061–2065 (2003).
77. Yildiz, A., Tomishige, M., Vale, R. D. & Selvin, P. R. Kinesin Walks Hand-Over-Hand. *Science (80-.).* **303**, 676–678 (2004).
78. Kapanidis, A. N. *et al.* Initial transcription by RNA polymerase proceeds through a DNA-scrunching mechanism. *Science (80-.).* **314**, 1144–1147 (2006).
79. Wen, J. Der *et al.* Following translation by single ribosomes one codon at a time. *Nature* **452**, 598–603 (2008).
80. Kapanidis, A. N. & Strick, T. Biology, one molecule at a time. *Trends Biochem. Sci.* **34**, 234–243 (2009).
81. Liphardt, J., Onoa, B., Smith, S. B., Tinoco I., J. & Bustamante, C. Reversible unfolding of single RNA molecules by mechanical force. *Science (80-.).* **292**, 733–737 (2001).
82. Hohng, S. & Ha, T. Fluorescence-forse spectroscopy maps two-dimensional reaction landscape of the Holliday Junction. *Science (80-.).* **1909**, 279–284 (2007).
83. Ainavarapu, S. R. K., Wiita, A. P., Dougan, L., Uggerud, E. & Fernandez, J. M.

Single-molecule force spectroscopy measurements of bond elongation during a bimolecular reaction. *J. Am. Chem. Soc.* **130**, 6479–6487 (2008).

84. Sonnleitner, A., Mannuzzu, L. M., Terakawa, S. & Isacoff, E. Y. Structural rearrangements in single ion channels detected optically in living cells. *Proc. Natl. Acad. Sci. U. S. A.* **99**, 12759–12764 (2002).
85. Diez, M. *et al.* Proton-powered subunit rotation in single membrane-bound F 0F1-ATP synthase. *Nat. Struct. Mol. Biol.* **11**, 135–141 (2004).
86. Tomishige, M., Stuurman, N. & Vale, R. D. Single-molecule observations of neck linker conformational changes in the kinesin motor protein. *Nat. Struct. Mol. Biol.* **13**, 887–894 (2006).
87. Andrecka, J. *et al.* Single-molecule tracking of mRNA exiting from RNA polymerase II. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 135–140 (2008).
88. O'Donnell, M. Replisome architecture and dynamics in Escherichia coli. *J. Biol. Chem.* **281**, 10653–10656 (2006).
89. Larson, J. *et al.* Design and construction of a multiwavelength, micromirror total internal reflectance fluorescence microscope. *Nat. Protoc.* **9**, 2317–2328 (2014).
90. van Oijen, A. M. Single-molecule approaches to characterizing kinetics of biomolecular interactions. *Curr. Opin. Biotechnol.* **22**, 75–80 (2011).
91. Friedman, L. J. & Gelles, J. Mechanism of transcription initiation at an activator-dependent promoter defined by single-molecule observation. *Cell* **148**, 679–689 (2012).