ATOMIC FORCE MICROSCOPY CHARACTERIZATION OF DNA-BINDING PROTEINS INVOLVED IN THE REPAIR AND ORGANISATION OF DNA

Alejandro Martín González

Universidad Autonoma de Madrid

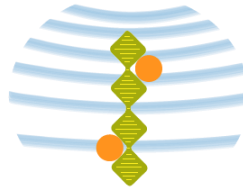
Thesis director:

Fernando Moreno Herrero

Centro Nacional de Biotecnologia

Departamento de estructura de macromoleculas

CNB-CSIC



Thesis title: Atomic Force Microscopy characterization of DNA-binding proteins involved in the repair and organisation of DNA.

Thesis tutor: Professor Julio Gomez Herrero, Condensed Matter physics department, Universidad Autonoma de Madrid.

The author of this thesis acknowdledges the financial support of the Spanish MINECO through a fellowship. Reference: BES-2015-071244

A mis padres y a mi hermano

"Nothing in life is to be feared, it is only to be understood. Now is the time to understand more, so that we may fear less." – Marie Sklodowska-Curie, first Nobel Laureate to be awarded a Nobel prize in two disciplines.

On a more relaxed context:

"In this house we follow the thermodynamic laws!" – Homer Jay Simpson

ABSTRACT

The invention of the Atomic Force Microscope played a fundamental role in the characterization of nanometric molecular complexes. The Atomic Force Microscope is a suitable research technique for the study of biomolecules one by one with the advantage to be an effective technique inside and outside a liquid medium. However, its use in biology has been limited by the availability of faster visualization techniques which do not need to deposit a sample on top of a flat surface. In this thesis I present the result obtained in the study of three biological processes from a biophysical point of view with Atomic Force Microscopes.

In the first project, three centromere-binding proteins are studied in the context of their plasmid partition systems. Plasmid partition systems consist of a group of proteins and DNA-sequences employed for the transmission and stabilization of plasmid molecules during cellular division. The three organisms, *B. subtilis*, *C. crescentus* and *C. botulinum*, have developed their partition systems differently. Regardless, these systems consist of a centromeric sequence near the origin of replication, a centromeric-binding protein and a NTPase motor protein. The formation of a segrosome complex, Centromeric-Binding protein bound to DNA, is showed by Atomic Force Microscopy images revealing a very different structure for the three biological organisms.

In the second project, two protein of the Homologous Recombination mechanism are studied. The first, CtIP, is a tetrameric protein with some studies pointing to a short-resection of DNA-ends activity. The second, Dna2, is a helicase/nuclease protein with a defined role in the long-resection of DNA-ends in Homologous Recombination. However, the binding-mode of these proteins to DNA molecules is not completely understood. Here, I present the data obtained on CtIP and Dna2 DNA-binding properties. Surprisingly, CtIP reliably binds to non-canonical DNA-ends and was found to join DNA molecules.

In the third project, the activity of the DNA mismatch repair mechanism in natural competent bacteria is studied. These organisms can introduce and incorporate exogenous DNA into its own DNA material. However, the DNA mismatch repair mechanism regulates and repairs the divergent regions between the two DNA strands. The deactivation of this mechanism increases mutation rate in bacteria. In this work, the role of mismatch repair proteins of competent *B. subtilis* bacteria is shown for the strand-exchange process.

RESUMEN

La invención de la Microscopía de Fuerzas Atómicas fue un hito crucial para la identificación de complejos moleculares manométricos. Este microscopio es una herramienta útil en la investigación de moléculas biológicas tanto en medios fisiológicos como al aire. No obstante, el uso de esta técnica para el estudio de moléculas biológicas ha sido limitado por la disponibilidad de otras técnicas de visualización más rápidas sin necesidad de depositar la muestra en una superficie plana. En este proyecto de tesis se presentan los datos obtenidos en el estudio de tres procesos biológicos con la Microscopía de Fuerzas Atómicas.

El primer proyecto que se presenta se estudia tres proteínas de unión a región centromerica en el marco de los sistemas de partición de plásmidos. Los sistemas de partición se conforman de proteínas y secuencias de ADN y su función es la de transmitir y estabilizar los plásmidos durante la división celular. Los tres organismos aquí presentados, *B. subtilis*, *C. crescentus* and *C. botulinum*; han evolucionado de manera que sus sistemas de partición difieran en su mecanismo. De todas maneras, los sistemas de partición están compuestos por una proteína de unión a región centromerica, una NTPasa proteína motora y una secuencia de ADN localizada cerca del origen de replicación. La formación de un complejo, denominado segrosoma y diferente para cada organismo, entre varios de estos elementos se muestra con imágenes adquiridas con una Microscopio de Fuerzas Atómicas.

En el segundo proyecto se presentan los datos obtenidos en el estudio de dos proteínas del proceso Recombinación Homóloga. CtIP es una proteína tetramerica de la que algunos estudios sugieren una actividad clave en la resección corta de finales rotos de ADN y Dna2 es una helicasa/nucleasa con un rol definido en la resección larga de finales rotos de ADN. Sin embargo, el modo de unión de estas proteínas a moléculas de ADN no está del todo desvelado. En esta tesis, se presentan los datos obtenidos en la interacción de ambas proteínas con ADN. Sorprendentemente, se desvela que CtIP puede unirse con facilidad a finales no-canónicos de ADN y puede realizar uniones entre moléculas de ADN.

En el último proyecto se presenta el rol del mecanismo de reparación de desajustes del ADN en bacterias naturalmente competentes. Estos organismos pueden introducir e incorporar ADN exógeno dentro de su propio ADN. En este caso el mecanismo de reparación de desajustes del ADN regula y repara las inserciones de ADN divergente. La desactivación de este mecanismo aumenta la ratio de mutaciones en bacterias. En esta tesis se ha estudiado el rol de este mecanismo en las bacterias naturalmente competentes de *B. subtilis*.

INDEX

LIST OF TABLES

LIST OF FIGURES

List of Abbreviations and Acronyms

AM	….	Amplitude Modulation
ATP	….	Adenine-Tri-Phosphate
bp	….	Base pairs
BSA	….	Bovine Serum Albumin
CBP	….	Centromere-binding Protein
C-tBP	….	C-terminal Binding Protein
DSB	….	Double Strand Break
dsDNA	….	Double-stranded DNA
EM	….	Electron Microscopy
FM	….	Frequency Modulation
GDP	….	Guanine-Di-Phosphate
GTP	….	Guanine-Tri-Phosphate
HGT	….	Horizontal Gene Transfer
HR	….	Homologous Recombination
IR	….	Ionizing Radiation
MMR	….	Mismatch Repair
MRN	….	Mre11-Rad50-Nbs1 complex of proteins
MT	….	Magnetic Tweezers
NHEJ	….	Non-Homologous End Joining
nt	….	nucleotides
OBD	….	Optical Beam Detector
RT	….	Room Temperature
SNOM	….	Scanning Near-field Optical Microscopy
SPM	….	Scanning Probe Microscope
SSB	….	Single Strand Break

ssDNA	….	Single-stranded DNA
STM	….	Scanning Tunneling Microscope
STORM	….	Stochastic Optical Reconstruction Microscopy
TIRF	….	Total-Internal Reflection Fluorescence
UHV	….	Ultra-high Vacuum
WT	….	Wild type

I. INTRODUCTION

The topic of this thesis is motivated on the study of several biological processes via the single molecule technique referred as Atomic Force Microscope. This first chapter is a broad introduction onto both, the research technique and the biological processes, for the reader to better understand the results and discussions provided on the following chapters.

"In science, everything is new because other people did not dare to try it"- Heinrich Rohrer, co-inventor of the Scanning Tunnelling Microscope

The writing of an introductory chapter is one of the most challenging aspects of the creation process of a scientific thesis. The introduction must present the necessary information for the reader to understand the motivation, background, results and most importantly, relevance of the scientific work performed by the doctoral student. In this case, this chapter will describe the Atomic Force Microscope, AFM, and two biological processes related to the maintenance and repair of the deoxyribonucleotide acid molecule or DNA. The field of study was carefully chosen considering the scientific background of the director and the tutor of this thesis, Fernando Moreno-Herrero and Julio Gomez-Herrero respectively.

To introduce the Atomic Force Microscope, I will describe the concepts of "Microscopy" and single molecule techniques, in which the AFM is included. The invention, evolution and future prospects of the AFM are further reviewed.

To continue with the introduction, I will present a description on the two biological mechanisms, plasmid partitioning and DNA repair, studied with an Atomic Force Microscope.

I.1 Microscopy or the art to observe the small

From the Cambridge's University Dictionary, "microscopy is the use of lenses to make very small objects look bigger". However, I would like to give a more personal definition for microscopy as "the art of observing microscopic particles or processes".

Science and art have been closely related since the first *Homo sapiens* tried to understand the events and conditions surrounding them. As humans we are terrified of the things we cannot describe. Science and art, thus, are means to reveal the unknown. In the case of microscopy, the study of "small objects" would help the society to understand the processes that are too small for the human eye to be observed. Separating science from art or eliminating one of them would be an error impeding a proper visualization or understanding of the world.

From Cambridge´s University Dictionary definition, a microscope would appear limited to the magnification provided by the use of lenses and, subsequently, a light source. However, in 1883, a scientific study described the limitation in resolution of a technique by the application of a light source as signal in magnification instruments [Abbe (1883)]. From this study, the use of a white light source would limit the resolution of an optical microscope to around 0.2 μm, or, more or less, half the wavelength of the light used. To overcome this limit, the scientific community developed "microscopy" techniques that are not based on the direct illumination and visualization of regions of interest. The 1986 Nobel

prize in physics was awarded to Ernst Ruska for the Electron Microscope and Gerd Binnig together with Heinrich Rohrer for the Scanning Tunnelling Microscope, STM. Both microscopy techniques are not based on the use of a direct ilumination source directed to the sample region and, thus, are better suited for the acquisition of images of small compounds or molecules than a typical optical microscope. The image acquisition process of an Electron Microscope is based on the scattering pattern of electrons bombarded to a surface [Knoll and Ruska (1932)]. On the other hand, the Scanning Tunnelling Microscope scans a conducting sharp tip along a conductive or semi-conductive surface mapping the electric current tunnelled through both of them [Binnig, et al. (1982)]. Alternatively, other microscopy techniques have been developed to overcome the diffraction limit based on light illumination, as the Stochastic Optical Reconstruction Microscopy or the Photoactivated Localization Microscopy, or based on other concepts such as the photoacoustic effect among many others [Betzig, et al. (2006); Rust, et al. (2006); Yao and Wang (2013)]. These microscopy techniques were developed on the basis of different scientific concepts and, thus, would require a whole different chapter for an in-depth description. Therefore, we will focus our attention into the STM considering it as the precursor of the Atomic Force Microscope.

In 1981, the first Scanning Probe Microscopy technique was invented, the Scanning Tunnelling Microscope It was applied to the study of metallic material surfaces [Figure I-1A] [Binnig, et al. (1982)]. The limitation of the STM was the requirement of an electric current between a tip and a surface, thus, requiring a conductive or semi-conductive surface. However, its creation paved the way to the development of other microscopy techniques derived from the concept of an interaction between a surface and a sharp tip. These techniques are referred as Scanning Probe Microscopies. The actual Scanning Probe Microscopies appeared from the popularization of the STM among the scientific community and the many advances around the technique. The developments were focused on cantilever microfabrication, fabrication of sharper tips and faster and more complex electronic controllers. These advances permitted a more precise movement and detection of the tip deflection by the interaction with a surface. This increase in precision allowed the detection of the atomic forces between a tip and a surface without the necessity of an applied electric current. The technique was named “Atomic Force Microscopy” [Binnig, et al. (1986)].

Nowadays, optical microscopes are applied to the study of living organisms, such as bacteria, under physiological conditions or to control the building process of small synthetic compounds, such as microchips, which are in the order of micrometres [Figure I-1 C]. In addition, the Scanning Probe techniques have been widely used for the study of individual particles or molecules [Figure I-1 A-B and

D]. These microscopy techniques popularized a new scientific field of study referred as single-molecule studies. The single-molecules studies are applied to the understanding of transient or rare molecular processes. These processes are susceptible of being masked by averaging under high molecular concentrations or under chaotic environments. The features of the single-molecule analysis made them a suitable choice for the study of biological mechanisms from a physical point of view. The results presented in the following chapters of this thesis are a part of the single-molecule studies field.

The next section will introduce the origins of the single-molecule studies and their relevance in this thesis.

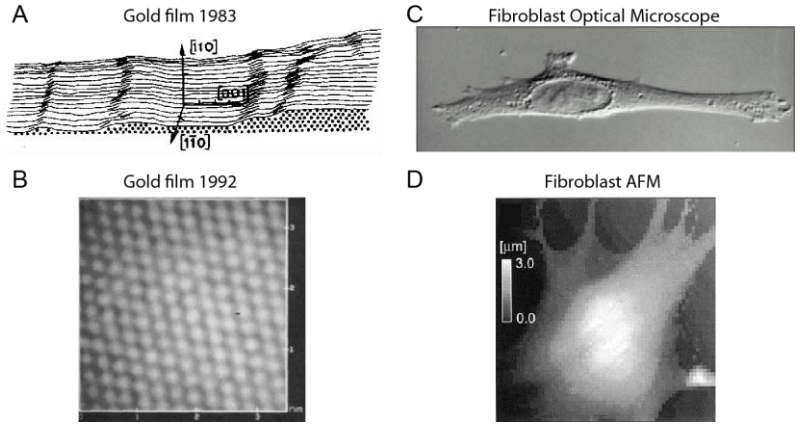


Figure I-1 Images of a gold surface obtained by STM at different years and images of a fibroblast with an optical microscope and AFM.

A) Topography of a gold surface at 300ºC. The distance between division sheets is of 1 nm with directions marked on top by three arrows. The dots indicate the position of the gold atoms. Image adapted from "Binnig and Rohrer (1983)". B) Topography of a gold film grown on mica. Image adapted from "Kim and Bard (1992)". C) Fibroblast cell under optical microscope. Image adapted from Uzman (2003). D) Topography of a fibroblast under AFM. Image adapted from "Haga, et al. (2000)".

I.2 The beginning of the single-molecule studies

Traditionally, scientific research is performed on large ensembles of molecules or molecular complexes. The manipulation of large concentrations or quantities of products makes the research processes easier to handle but harder to acquire for the researcher. One example from biochemistry is the production of highly-concentrated and pure proteins from cell cultures suitable for electrophoretic assays [Hellman and Fried (2007)]. These assays are valued methods for the analysis of molecular weights and molecular interactions for charged molecules.

The characteristics of this technique allows for the separation of the molecules of interest altogether according to their movement through a viscous gel by the circulation of an electric current [Smithies (1955)]. However, in the case of molecules with similar properties, the technique may not resolve them. In this situation, the biochemical assays can be complemented with other research techniques [O'Farrell (1975)]. Contrary to the many available bulk techniques, some scientific methods were developed to offer the possibility to analyse molecules or processes one-by-one. In 1976, the considered as first single-molecule technique, the patch-clamp, was invented by Erwin Neher and Bert Sakmann. This technique is able to measure the electric current passing through a single ion channel embedded on a cell membrane [Neher, et al. (1978)]. For their discovery, they received the 1991 Nobel Prize in physiology or medicine "concerning the function of single ion channels in cells" [Greger (1991)]. Today, the single-molecule studies have been popularized by the access to many different techniques. Therefore, the selection of the correct technique for the study of a molecule or process relies on the availability and suitability of the system.

Among the available techniques, the Atomic Force Microscope was found to be the best option for this thesis study. The AFM acquires topographic images of particles with sizes ranging from a couple of Angstroms (10^{-1} nm) to tens of micrometres (10^{4} nm) [Figure I-2]. The technique operates in the expected range for protein (1-100 nm) and DNA sizes (1 to $1x10^{4}$ nm). The advantage of the Atomic Force Microscopy among other single-molecule techniques is the possibility to acquire nano-scale images in physiological conditions. Ultimately, the decision also took into account other factors such as availability of expertise and access to an stablished and proficient Atomic Force Microscopy community.

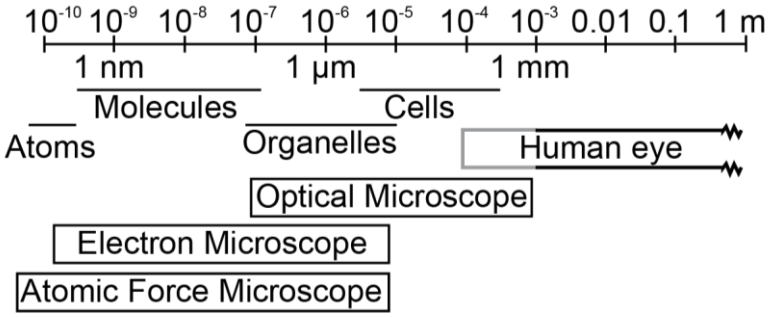


Figure I-2. Graphical representation for the resolution power of various microscopy techniques and model objects.

EM and AFM resolution in comparison to the human eye and light microscope in the context of biological and molecular objects. Image adapted from Kaufmann, et al. (2014) and Ode, et al. (2012).

I.3 Atomic Force Microscopy functionality for the study of biological processes

Biological processes are defined as the main vital reactions of an organism. Frequently, these processes are chemical reactions performed by proteins of the organism. During these processes the proteins interact with each other in a regulated and coordinated manner. However, the quantity of proteins in a human cell is estimated to be around $2x10^7$ proteins divided into, approximately, $1x10^4$ protein structures depending on the cells size, function and state [Beck, et al. (2011)]. The immense quantity of proteins inside a cell derives from their high specificity and the abundance of individual biological reactions. Consequently, the many individual proteins of the same organism are of very different chemical structure, shape and size. For example, a globular protein has a compacted 3D shape [Wu, et al. (1997)] [Figure I-3 A-B]. On the contrary, proteins with coiled-coil domains have elongated domains with rod-like shapes [Eeftens, et al. (2016); Wilkinson, et al. (2019)] [Figure I-3 C-D]. Hence, the single-molecule technique employed for the study of biomolecules must be adapted for the analysis of a variety of particles in dynamic roles.

The possibility to acquire images of particles or complex biological systems made Microscopy a suitable technique in this case. Nevertheless, the decision remained among the many Scanning Probe, Electron and Fluorescence microscopies. The Electron Microscope has seen an increase in scientific research publications through the recent years but the technique is not capable of acquiring dynamic information and requires a complex sample preparation [Kourkoutis, et al. (2012)]. Fluorescence microscopy techniques rely on the light absorption and emission of particles such as Green-Fluorescence Protein or Quantum Dots for the visualization of dynamic and static processes [Madariaga-Marcos, et al. (2019)]. Lastly, the Scanning Probe Microscopes scan a surface with a sharp tip. They were initially thought for the study of metallic materials but an intense development through the years saw a rise in popularity of the SPMs in the study of biological processes [Bustamante, et al. (2000);Ando, et al. (2001) ; Carrasco, et al. (2008)]. However, the Scanning Probe Microscopes require an interaction between the particle of interest and a surface which may modify the reactions outcome [Hansma (2001)]. Among the variations of the Scanning Probe Microscope (Atomic Force for the detection of particles, Kelvin Probe for the detection of surface potentials, Magnetic Force for detection of magnetic fields…) the Atomic Force Microscope is a valid technique for the study of biological processes together with the ample experience in its use at Fernando Moreno-Herrero´s research group.

Initially the Atomic Force Microscope images were acquired at a slow rate. The acquisition speed was sufficient for the study of stable materials but not for moving molecules or dynamic processes. The evolution of the technique through the years offered the possibility to obtain high-resolution static images or dynamic videos of molecules deposited onto a surface [Figure I-3] [Kodera, et al. (2010)].

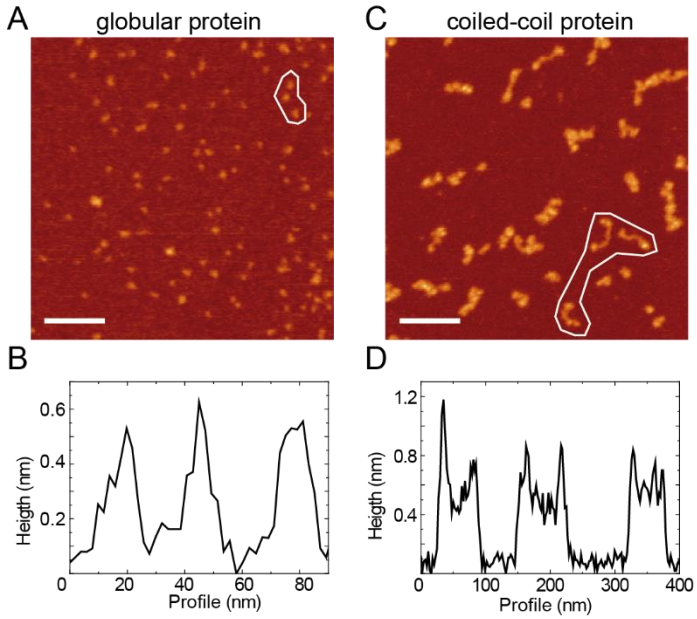


Figure I-3. AFM images of a globular protein and a coiled-coil protein with profiles of the proteins.

Comparison of two AFM images of a globular protein and an elongated protein with profiles along three selected molecules. A) AFM image of the globular protein TubR, scale bar is 100 nm. B) Profile along the three molecules enclosed in the image A) with horizontal sizes of 15, 12 and 13 nm and height on the order of 0.5-0.6 nm. C) AFM image of the elongated protein CtIP, scale bar is 100 nm. D) Profile along the three molecules enclosed in image C) with horizontal sizes of 63, 57 and 78 nm and height on the order of 0.4-1.2 nm.

I.4 The evolution of the Atomic Force Microscopy from its beginnings

In 1986, one of the inventors of the Scanning Tunnelling Microscope modified one of the STM systems to control the flexure of a cantilever with a sharp tip at the end when scanning a surface [Figure I-4 A]. In the beginnings, the most important components for the good resolution of the Atomic Force Microscope images were the long and flexible cantilevers acting as a spring and the sharp tip attached at the end of the cantilever. The surface resolution of an Atomic Force Microscope is set by the detection of the bending angle of the cantilever. Therefore, cantilevers

with a low spring constant (0.1 N/m) are more sensitive to surface features because the force required for bending the cantilever is lower. However, a stiffer cantilever with a higher resonance frequency decreases the sensitivity to the vibrational noise from a building, around 100 Hz. The solution was to employ a low-mass cantilever. With an "ultrasmall mass" cantilever the system was capable of measuring the movement of the cantilever a distance of 10^{-4} Å by a force as low as 10^{-18} N [Binnig, et al. (1986)]. The force was on the order of the atomic forces between single atoms, thus, the given name to this technique was Atomic Force Microscope. Before the AFM, the techniques suitable for the study of insulating surfaces were scarce; mainly, the Stylus Profilometer which had a horizontal resolution of 100 nm, a vertical resolution of 1 nm and a loading force of 10^{-2} -10^{-5} N [Williamson (1967)].

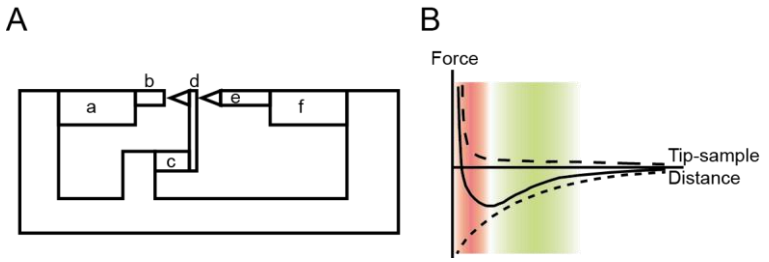


Figure I-4. First experimental setup for Atomic Force Microscopy and Lennard-Jones potential of forces.

A) Experimental setup by Binnig, et al. (1986) for the first AFM. a = piezo electric scanners and feedback for the AFM; b = sample; c = modulating piezo; d = AFM tip; e = STM tip; f = feedback for the STM system. B) Lennard-Jones potential of interaction forces between tip and sample. Solid line denotes the overall Lennard-Jones potential; dashed line for repulsive forces represent the short-range Coulomb forces and for attractive forces the van der Waals forces. The green regime corresponds to the tip-sample distance for which the dominating forces are attractive, referred as non-contact, and the red regime to the tip-sample distance with dominating repulsive forces, referred as contact.

To maintain a constant force on the surface of interest, an Atomic Force Microscope utilizes a feedback system. The feedback system controls the bending of the cantilever. A piezo-electric material connected to the opposite end of the tip-attached cantilever keeps the bending constant throughout the process of image acquisition [Binnig, et al. (1986)]. The resulting mode of operation was called "contact mode" as the AFM tip touches the surface. In this case, the "contact" regime refers to a tip-sample distance for which the dominating forces between tip and surface are repulsive [Figure I-4 B red regime].

In 1987, atomic resolution was achieved for graphite, a conductive material, and boron nitride, an insulator material, surfaces with microfabricated silicon tips in "contact" Atomic Force Microscopy [Petersen (1982); Binnig, et al. (1987)] [Figure

I-5 A]. Initially, the images acquired of graphite in air and Room Temperature conditions with these tips were compared to the ones acquired by a STM. However, the results were a topographic image for AFM and a combination of topography and electron density image for the STM [Binnig, et al. (1987)]. During this work the authors attempted to acquire atomic resolution data in the attractive regime of forces between tip and sample, known as “non-contact mode”, but were not successful [Figure I-4 B green regime]. To effectively obtain images in the attractive regime, they proposed to enclose an AFM into a Ultra-High Vacuum enclosure to increase the signal to noise ratio and decrease thermal drift [Binnig, et al. (1987)]. Around the same time, the group of Paul K. Hansma was able to reproduce the images with atomic resolution obtained for graphite inside a liquid medium [Marti, et al. (1987)].

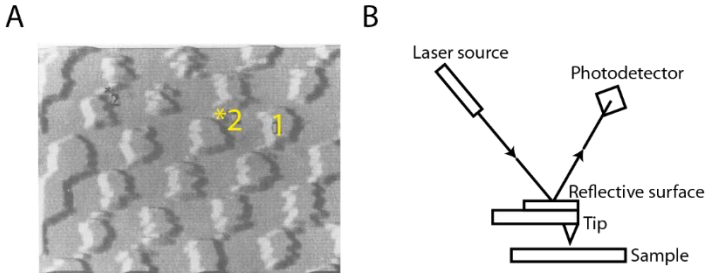


Figure I-5. Boron nitride image acquired by AFM and cantilever detection system with laser optics.

A) Atomic resolution image of the insulator material boron nitride acquired by Binnig, et al. (1987). 1 marks the position of strong spots and 2 marks the position of weak spots. B) Cantilever detection system with a laser source and a photodetector developed by Meyer and Amer (1988).

The use of contact mode for atomic resolution was controversial for the following years [Ohnesorge and Binnig (1993)].At the time, the theory and knowdledge about material surfaces suggested that the surfaces were not uniform due to the presence of defects [Pethica and Oliver (1987)]. However, the strong interaction between the tip and the sample could modify the shape of the tip and of the surface creating images of surfaces with no defects [Figure I-1 A-B]. To overcome this problem, the “non-contact” operation mode was developed [Martin, et al. (1987), Albrecht, et al. (1991)]. In this mode the cantilever is oscillated at its resonance frequency and the feedback signal is originated from the change on the oscillation frequency, referred as frequency-modulation, due to the tip-sample interaction. For Frequency Modulation-Atomic Force Microscopy, atomic resolution in Ultra-High Vacuum was not achieved until 1997 and in air until 2003 [Sugawara, et al. (1997); Seo, et al. (2003)].

Following in 1988, Gerhard Meyer and Nabil Amer substitued the Scanning Tunnelling Microscopy detection system on the Atomic Force Microscopy

cantilever for a novel optical system [Meyer and Amer (1988)] [Figure I-5B]. This detection system was demonstrated to recognize changes of 10^{-4} Å in a surface reliably [Amer (1983); Amer, et al. (1986)]. Additionally, this method simplified the AFM architecture contributing to the later popularization of the technique.

In 1991, the acquisition of high-quality images in Atomic Force Microscopes would, typically, take minutes. The bottleneck was the time needed for the movement of the piezo-electric materials and the response time of the feedback systems. To reduce the "*tedium out of imaging*" the idea of a high-speed AFM was starting to grow into the AFM scientific community [Barrett and Quate (1991)]. The first solution found was to acquire AFM images in constant-height mode and to scan 1 every 10 lines, in images of more or less 500 lines. In constant-height mode the piezo-electric material attached to the AFM tip does not move throughout the acquisition process to maintain the cantilever at the same height. Furthermore, by scanning 1 out of 10 lines the process was 10 times faster "*since most features of interest are larger than 10% of the image"*. The result was a scanning speed of 3 Hz or 3 images per second [Barrett and Quate (1991)].

In 1993, Digital Instruments™ developed an operational mode for image acquisition of soft samples in Atomic Force Microscopes. The new mode would operate between the repulsive and the attractive regimes by moving the tip between them by, theoretically, this would remove the effect of the compression and shear forces of the "contact" mode and the difficulty to operate on "non-contact" mode [Zhong, et al. (1993)]. "Tapping mode" was the name given to it. Similar to "non-contact" mode, the AFM cantilever is made to oscillate at its resonance frequency. However, the tip makes contact with the surface at each oscillation cycle. In the case of "tapping" mode, the feedback signal is originated from the change on the amplitude of oscillation, Amplitude-Modulation.

Advances in force control and image quality made the Atomic Force Microscopy a valuable technique for the study of soft-materials. Besides, acquisition of images for samples submerged in water or solution was relevant for the study of biomolecules. Therefore, the following years some research groups were focused on the development of AFM for the study of proteins or cells and later, with the "tapping" mode, of DNA-related processes [Drake, et al. (1989); Haberle, et al. (1992); Bezanilla, et al. (1994); Guthold, et al. (1994)].

From 1993 onwards, the main improvements for biomolecule imaging were focused on three different approaches:

1) To increase the scanning speed of the AFM to detect dynamic behaviour of molecules or reactions in their native environment [Radmacher, et al. (1994); Thomson, et al. (1996)].

2) To further decrease the lateral force exerted by the raster-scan of the tip along the surface [Anselmetti, et al. (1994); Hansma, et al. (1994)].
3) To perform force-spectroscopy analysis with AFM or combine an AFM system with other microscopy techniques for molecular recognition [Gumpp, et al. (2009); Hinterdorfer and Dufrêne (2006)].

A milestone was reached in 1996. Chemical etching allowed the fabrication of short cantilevers, being the smallest between 23 μm and 12 μm dimensions [Walters, et al. (1996)]. These cantilevers were fabricated from silicon nitride and were gold coated for an improved laser deflection. The sharp tetrahedral tip was located on one end. The short cantilevers are stiffer, thus, oscillating at a higher resonance frequency and decreasing the thermal noise. The higher resonance frequency is required for high-speed imaging as the tip must oscillate numerous times on top of each point of the sampled surface [Ando (2018)]. As a simple calculation, for a 1 Hz imaging speed of a 128x128 pixels per image, the oscillation frequency of the cantilever should be on the order of 30 kHz following Equation 1. Contemporary commercial cantilevers by Olympus have a typical resonance frequency in liquid medium of 25 kHz [Olympus Biolever Mini, BL-AC40TS-C2]. The drawback of the short cantilevers is the increased difficulty to operate the optical system for the detection of cantilever bending [Figure I-5 B].

$$T = \frac{2N^2}{f_c} \qquad \textbf{Equation 1}$$

Equation 1. Time required for image acquisition in an Atomic Force Microscope.
T represents time for image acquisition, N the number of pixels in an image and f_c the resonance frequency of the cantilever. The 2 arises from the time forward and backward scanning directions. Equation accounts only for one tap of the cantilever per pixel.

In 1998, specifically related to this thesis, a small company dedicated to the development, production and commercialization of Scanning Probe Microscopes started at the "*Universidad Autónoma de Madrid*". Nanotec Electronica S.L. was established by four researchers specialized in nanoscience studies with SPMs: Arturo M. Baró Vidal, Jaime Colchero Paetz, José M. Gómez Rodríguez and Julio Gómez Herrero. The company was involved in several research and developmental projects of hardware and software for their SPMs [Horcas, et al. (2007); Gimeno, et al. (2015)]. The motivation was the presence and growth of a strong SPM community in Spain [Pascual, et al. (1993); de Pablo, et al. (1998); Moreno-Herrero, et al. (2003)]. The Atomic Force Microscopes operated at the "Centro Nacional de Biotecnología-CSIC" during the work shown in this thesis were produced by Nanotec in 2008 and 2011 [Figure I-6].

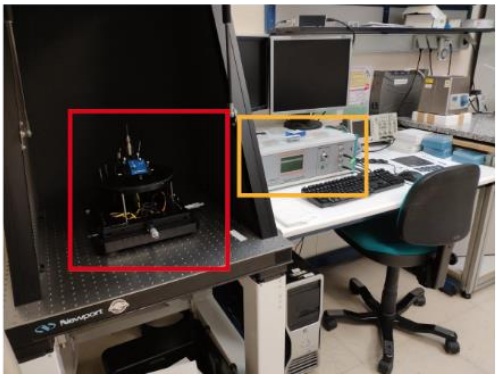


Figure I-6. Nanotec Electronica AFM employed during the thesis.
Cervantes model AFM system from Nanotec Electronica S.L. Red box highlights the position of the AFM mechanical parts and orange box for the electronic controllers.

In 1999, the first high-speed images of DNA molecules were acquired and presented in the article by Viani, et al. (1999). They worked with a brand new Atomic Force Microscope prototype with improved optics adapted for small cantilevers and a novel approach to apply the response signals. The Atomic Force Microscope was equipped with an Optical Beam Detector capable of signal detection for cantilevers of 3-5 μm wide and 10-14 μm long with a high resonance frequency, 100-200 kHz in liquid, and low spring-constant, 0.1-0.2 N/m, and a home-built RMS-to-DC for the amplification of the cantilever deflection signal to filter the feedback signal around the tapping frequency [Walters, et al. (1997)]. The whole system was capable of acquiring images for soft materials as DNA molecules at 1.7 s time without any major disturbance of the sample.

In 2001, a film acquired of the molecular motor myosin V at a rate of 10 Hz was a breakthrough for the popularization of High-Speed Atomic Force Microscopes [Ando, et al. (2001)] [Figure I-7 A]. Supplementary to the small cantilevers and Optical Beam Detectors, Toshio Ando developed a fast piezo-electric scanner and a fast-amplitude detector. The piezo-electric scanner was modified to reduce the drift and the response times of the scanner. The fast-amplitude detector allowed the system to detect changes in the oscillation amplitude each half oscillation cycle, contrary to the several oscillation cycles needed in Walters and others (1996). Nevertheless, the development of High-Speed and other AFM applications was still in motion.

I.5 High-Speed Atomic Force Microscopy

At the beginnings of the High-Speed AFM there were some difficulties related to the image acquisition of fragile biological samples [Ando (2018)]. Studies have

demonstrated how fragile molecules are affected by the interaction forces exerted by the AFM tip in different image acquisition methods [Moreno-Herrero, et al. (2004)]. Therefore, the High-Speed AFM kept evolving, much alike to a living organism, in the hands of many researchers. The same objectives were kept from the previously presented objectives for conventional AFM: to lower the acquisition time and to maintain a small force between the tip and the sample. The group of Toshio Ando was involved in many of the key developments for High-Speed AFM.

To maintain a low interaction force between the tip and the sample at High-Speed they developed a vibration damping technique for the Z-direction piezo-electric scanner and a new feedback controller combined with a compensator for drift in the excitation frequency of the cantilever [Kodera, et al. (2005); Kodera, et al. (2006)]. Altogether, the system could reach an acquisition time of 80 ms per frame as presented in the article Kodera, et al. (2010) with a less invasive imaging, allowing the user to acquire images for larger time periods [Figure I-7].

In “tapping” mode, the signals extracted from the cantilever bending are the frequency, amplitude and phase of the tip oscillation movement. The phase fluctuates in relation to the dissipation energy between the tip and the sample while the amplitude changes, mainly, with the tip-sample distance [Cleveland, et al. (1998)]. The frequency is maintained constant along the acquisition with a feedback system. Consequently, a faster phase-detector and a fast amplitude detector were developed. The phase-detector was enhanced by adapting a home-built detector with a combination of digital and analogic elements and an analogical phase signal output [Stark and Guckenberger (1999); Uchihashi, et al. (2006)]. The fast-amplitude detection system was reformed by the application of algorithms for amplitude and frequency demodulation by Fourier methods [Kokavecz, et al. (2006)].

These developments paved the way for in-depth studies of dynamic processes of biomolecules [Kodera, et al. (2010); Uchihashi, et al. (2011); Colom, et al. (2012)] [Figure I-7]. The high-speed AFM technique was fully stablished on the year 2008 [Ando, et al. (2008)].

A High-Speed AFM built by Toshio Ando’s group, similar to the AFM that acquired the Myosin V images, was used for a period of 3 months during this thesis to study the dynamics of various protein molecules showed in chapter III of this thesis [Figure I-7].

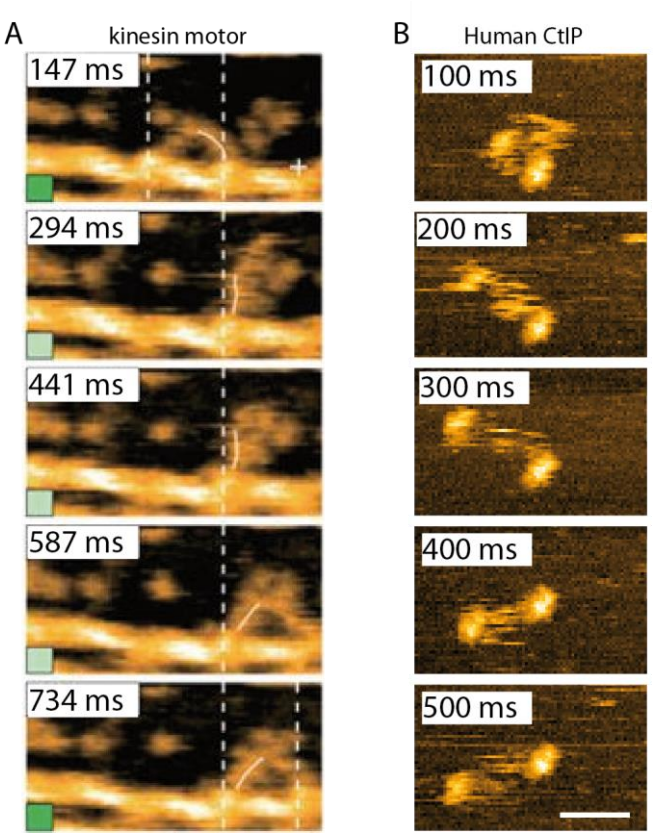


Figure I-7. High-speed AFM images and their times of acquisition for two proteins.
A) High-speed AFM images of myosin V motor sliding through an actin filament, images of 150x75 nm². Dark green squares signal images with Myosin V heads fully attached to the actin filament, light green signal images with, at least, one Myosin V head not attached to the acting filament. Figure adapted from Kodera and others (2010). B) High-speed AFM image of human CtIP moving on a mica surface, images of 100x65 nm². Scale bar is 30 nm.

I.6 Future prospects for Atomic Force Microscopy

The scientific community is curious by nature. The continuous seek for knowledge and research of the unknown drives the growth of the many scientific methods developed through the years. In this section, a brief summary on some future developments for the Atomic Force Microscopy technique elucidated on Ando (2018) are put into perspective.

I.6.1 Faster AFM

At the time of this thesis writing, the state-of-the-art acquisition time for low-invasive imaging on a 100x100 pixels and 100x100 nm^2 AFM video in “tapping” mode was 80 ms or, approximately, 12 Hz. However, biomolecules such as enzymes depend on much faster conformational changes for reaction catalysis [Kovermann, et al. (2015)]. The motions of enzymes are in the order of microseconds to milliseconds with shorter lived high-energy states responsible for the enzyme activity [Beach, et al. (2005)]. Fortunately, there is room for an increase in the Atomic Force Microscopy time resolution. With the actual architecture (Optical Beam Detectors, piezo-electric scanners, short and thin cantilevers, digital/analogical feedback controllers) the peak acquisition time would fall into the 6-10 ms range [Equation 1] [Ando (2018)]. A new revision of the AFM design would be necessary to further decrease the acquisition time. Nowadays, possibilities fall onto new cantilever deflection detector systems not based on laser optics [Sanii and Ashby (2010), Giessibl (2019)].

I.6.2 Hybrid AFM systems

Parallel to the faster AFM systems the resolution limit for AFM imaging has been improved to reach atomic resolution for hard materials and nanometre resolution for soft materials and biomolecules [de la Torre, et al. (2016); Ares, et al. (2016)]. However, one of the drawbacks of the AFM is the acquisition of topographic images without information on the surface’s chemical composition. Spectroscopy techniques have the advantage of particle recognition through its interaction with light. Recently, fluorescence techniques have made a leap into the super-resolution regime by overcoming the diffraction limit related to optical microscopy explained previously [Hell and Wichmann (1994); Klar, et al. (2001); Betzig, et al. (2006); Rust, et al. (2006)]. A drawback of these techniques are the stochastic occurrence of events (STochastic Optical Reconstruction Microscopy), necessary modifications for fluorescence molecules, time resolution and high-costs of these systems [Resch-Genger, et al. (2008); Suzuki, et al. (2007); Huang, et al. (2009)]. As a solution, a combination of high-resolution/high-speed AFM mounted on an optical microscope adapted for Total Internal Reflection Fluorescence (TIRF) or epifluorescence would benefit from the resolution and image acquisition of the AFM and the molecular detection of the fluorescence [Ambrose (1956); Webb and Brown (2013); Shaw, et al. (2006)].

It is also possible to combine both techniques into a Scanning Near-field Optical Microscope to further increase the resolution of the optical microscope by passing the light through the gap between the tip and the sample in the Atomic Force Microscope [Hecht, et al. (2000)]. For this architecture, the small space between

the AFM tip and sample acts as the aperture probe on the focusing spot of the light source.

I.6.3 Non-contact AFM for biomolecules in liquids

Achieving pure non-contact images inside a liquid medium remains one of the most challenging aspects of Atomic Force Microscopy. The signal obtained from the AFM tip operating at a few nanometres distance from the sample surface originates from the attractive Van der Waals forces [Hutter and Bechhoefer (1993)] [Figure I-4 B green regime]. Van der Waals forces are the dominating forces at the middle-interaction regime in AFM. In this regime the tip-sample interactions are weaker than in contact mode [Figure I-4 B red regime]. The lower interaction forces are high enough for high-sensitivity measurements on Frequency Modulation mode in an Ultra-High Vacuum AFM. In these conditions, high-resolution images were obtained without disrupting the sample for longer periods of time [Morita, et al. (1996)]. For air and liquid samples the non-contact mode has remained a challenge only attained for specific materials and conditions [Fukuma, et al. (2005)]. The lower forces between the tip and the sample make the non-contact mode a high-resolution/low-invasiveness image acquisition mode compared to tapping mode.

I.6.4 Nano-endoscopy

The study of dynamic movements and processes in AFM is limited by the necessary interaction between the tip and the sample. In liquid environment, only the molecules on the surface are imaged by the technique; molecules solved into the liquid are not visualized. Typical examples are the studies focused on the proteins of the outer cellular membrane [Muller and Engel (2007); Casuso, et al. (2012)]. Thus, *in vivo* studied of proteins, enzymes or nucleic acids localized in the interior of the cell membrane remains impossible. However, studies have proved that insertion of a long and thin pipette tip, 1-100 nm in diameter and 1-20 µm in length, into a cellular organism does not kill it while it is convenient for injection of material or spatial manipulation of cells [Obataya, et al. (2005); Yum, et al. (2010); Guillaume-Gentil, et al. (2014)]. The coupling of a long and thin tip to an AFM cantilever might permit the system to acquire small scale images of biomolecules supported on lipid bilayers inside a cell wall.

In conclusion, the Atomic Force Microscope has been developed since 1981 to be applied not only to stable surfaces but to dynamic and more fragile particles in different environments (Vacuum, air and liquid conditions) together with other Scanning Probe Microscopies.

To continue with the introduction, I will describe the biological processes to which the Atomic Force Microscope have been applied in this thesis.

I.7 The essential mechanisms for a living organism

By definition a biological process is an essential mechanism for the survival of a living organism [Montevil, et al. (2016); Mossio, et al. (2016)]. The purpose of these essential mechanisms is conserved between unicellular and multicellular living organisms. Mammals as *Homo sapiens* and bacteria as *Bacillus subtilis* need to grow, reproduce, adapt and organize themselves. The mechanisms responsible for these tasks are vital for the living organism.

The actual number of essential biological mechanisms in a living organism is unknown. Databases of the human reactome show more than 12000 different reactions for more than 2000 pathways are involved in the change or production of new molecules in human cells [Fabregat, et al. (2018)]. Nevertheless, the database only contains information for 53% of the supposed molecules codified in a typical human genome. Therefore, unknown pathways remain to be discovered or completed [Zerbino, et al. (2018)].

The biological processes related to the synthesis, packaging and repair mechanisms of the DNA molecules inside a cell are of importance for the evolution and transmission of genes between organisms. In this thesis I have studied four mechanisms from two DNA-related biological processes with Atomic Force Microscopy. Precisely, two DNA partition systems and two DNA break-repair mechanisms.

I.7.1 DNA partition

The genetic material of a cellular organism is organised into chromosomes and plasmids. Chromosomes are highly-packed long DNA molecules, up to $25x10^7$ nucleotides. They are visible under a light microscope at the cellular division phase by staining from where the name derives, chromo(Colour)-some(Body) [Burkholder and Weaver (1977)]. A chromosome contains the building and growth information of its organism.

Plasmids are smaller, circular, double-stranded DNA molecules with their own replication mechanisms, hence, they are independent molecules from chromosomes [Lederberg (1952)]. Plasmids often have a low-copy number inside the organism and are not that highly-packed. However, plasmids contain necessary information for the organism to adapt to its surroundings. For example, the antibiotic resistance genes are contained inside plasmids [Bennett (2008)].

The correct replication and transmission of plasmids during cellular division has an intrinsic role in evolution.

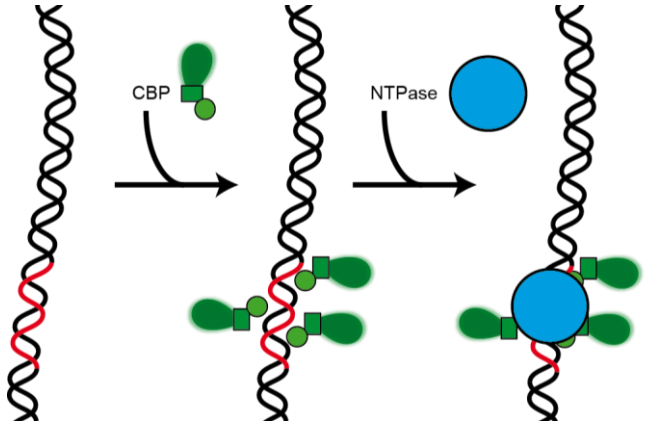


Figure I-8. The usual three elements of the plasmid partition systems.
Generalized plasmid partition system. DNA molecule in black with the centromere sequence in red. Centromere-binding protein in green. Molecular motor NTPase in blue. The specific sequence acts as anchoring point for the centromere-binding protein which recognizes and binds to the motor protein, in charge of the segregation of the DNA molecule

Previous to the cellular division phase, or mitosis, the genetic material is replicated (copied) [Takeda and Dutta (2005)]. Then, during the mitosis, chromosomes and plasmids are organized for a uniform distribution of genetic material among the resulting cellular offspring [Ruchaud, et al. (2007)]. Symmetric distribution of the genetic material among the progeny is key to survival and evolution [Potapova and Gorbsky (2017)].

Chromosomes were firstly observed around the 1870s but plasmids were not referenced in a scientific paper until 1952 [Fokin (2013); Lederberg (1952)]. The cause was that plasmids were vaguely defined as "any extrachromosomal hereditary determinant" and their process of maintenance and replication was unknown. Until the decade of 1980, researchers analysed the replication and segregation machinery of low-copy number plasmids [Abeles, et al. (1985), Martin, et al. (1987)]. Albeit different, the partition system of some low-copy plasmids has similarities to that of eukaryotic chromosomes [Oliva (2016)]. Previous studies show a model for partition systems composed of a *cis*-acting DNA sequence and two *trans*-acting proteins responsible for the regulation and activation of the process [Friedman and Austin (1988)] [Figure I-8]. The cis-acting DNA sequence is a 13-bp DNA repetition located nearby to the origin of replication [Mott and Berger (2007)]. This sequence is recognized by a DNA-binding protein, or Centromere Binding Protein, which binds to it. The bound protein forms a DNA-protein complex referred as “segrosome” [Hayes and Barilla

(2006)]. The segrosome complex is then anchored to an NTPase motor protein (NTPase refers to an enzyme with ATP or GTP hydrolisis activity) responsible for the transport of the plasmid molecule [Schumacher (2008)]. Among prokaryotes there are four classes of plasmid partition systems [Oliva (2016)]. Types I and II have been widely studied but remain unclear in some aspects. Type III and IV have been recently discovered and their mechanisms remain largely unknown mechanisms.

Distinctly, high-copy number plasmids rely on diffusion to facilitate the transfer process of genetic material during cellular division. A cellular organisms with several copies of a plasmid would naturally distribute them among the daughter cells without a specific segregation mechanism [Reyes-Lamothe, et al. (2014)].

I.7.2 DNA repair

The DNA molecule was once hypothesized to be a stable molecule that was rarely modified. However, it was later discovered that every day the DNA inside a human cell suffers up to $1x10^4$ lesions [Ohnishi, et al. (2009)]. In addition, a regular human body was estimated to contain around $3x10^{13}$ cells [Sender, et al. (2016)]. The values provides a number of $3x10^{17}$ DNA lesions on a human body per day. A lesser number of DNA lesions take place on bacterial organisms as their genetic material content is smaller than in mammals [Lindahl (1993)]. These lesions are produced by chemical reactions often caused by by-products of other biological reactions, as the regular metabolism, inside the organism or by exposure to high-energy radiation [Hoeijmakers (2009)].

The DNA molecule contained inside a chromosome is the largest molecule inside a cell. In a chromosome, not-resolved DNA lesions may cause the death of the organism [Hanawalt, et al. (1979)]. This molecule relies on the cellular repair mechanism for its maintenance as, typically, prokaryotic cells only contain one copy of each chromosome. For survival, every living organism has developed and evolved towards a different repair mechanisms. Currently, the repair mechanisms of many organisms are studied to understand the effect of damages and repair processes on DNA-related syndromes and disorders [Jackson and Bartek (2009); Biehs, et al. (2017); Wang, et al. (2018)].

The many classes of damages produced during a cellular lifetime gave rise to a large ensemble of specialized pathways and proteins for the necessary repair mechanisms [Gates (2009)]. Despite being enormous, the specialization of the different mechanisms made their study *in vitro* easier by protein picking. To simplify, I will introduce the DNA lesions of interest for the understanding of this thesis results.

I.7.2.1 DNA lesions

DNA lesions are responsible for the disruption of the chemical structure of the double-helix DNA molecule. The double-helix structure of the DNA was discovered in 1953 by James Watson and Francis Crick with support from X-ray diffraction data obtained by Rosalind Franklin and Maurice Wilkins [Watson and Crick (1953); Klug (1968)]. The DNA helixes are composed of deoxyriblonucleotide polymers. A deoxyribonucleotide monomer is formed by a nitrogenous base, a deoxyribose sugar and a phosphate group [Figure I-9]. The monomeric unit is shortly referred as "base". These bases are further divided depending on the chemical structure of the nitrogenous base (Adenine, Guanine, Cytosine, Thymine) [Figure I-9]. The polymerisation process of the deoxyribose sugar of a monomer to the phosphate group of another monomeric unit results in a "phosphate backbone".

The DNA molecule of a chromosome is made of the combination of two polymeric deoxyribonucleotide chains into a double-helix shape [Zerbino, et al. (2018)]. In humans, the length of one chromosomic chain ranges from $5x10^7$ to $25x10^7$ bases. The stabilization of both chains into a double-helix arise from the hydrogen bonds formation between nitrogenous bases of opposing chains. These hydrogen bonds are created between an adenine and a thymine base (two hydrogen bonds) or a guanine and a cytosine base (three hydrogen bonds) [Fonseca Guerra, et al. (2000)]. The difference on hydrogen bonds make the pair formed by guanine and cytosine more stable by the allocation of a higher number of bonds [Šponer, et al. (1996)].

Lesions on a DNA molecule alter or break any of the three elements of bases, the phosphate backbone or the union between bases in different strands. A further overview on the damage and repair mechanisms for the lesions known as "Double-strand breaks" (breakage of the phosphate backbone of both strands) and "Mismatches" (faulty pairing of a base with its corresponding pair on the other strand) follows.

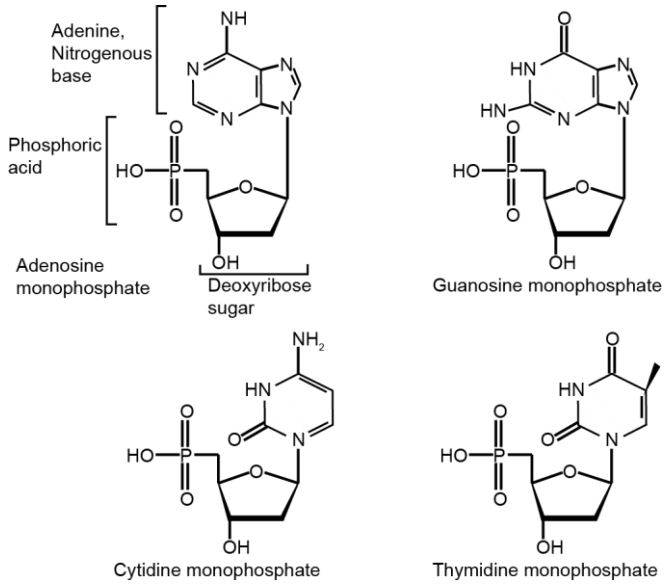


Figure I-9. The four DNA bases, Adenosine, Guanosine, Cytidine and Thymidine.
Chemical representation of the four deoxyribonucleotide acids. Lines represent bonds between atoms with P=Phosphor, N=Nitrogen, O=Oxygen and H=Hydrogen. Carbon is represented by a vertex for simplification.

I.7.2.1.1 DNA double-strand breaks

The bond between the deoxyribose of a DNA base and the phosphate group of the next DNA base is one of the strongest bonds in a DNA molecule [Lindahl (1993)]. Nonetheless, previous studies have shown the disrupting effect of external and internal agents on this bond [Kuzminov (1999); Adhikary, et al. (2012)]. One of the most dangerous DNA damages occurs when this bond is completely broken. Breakage of the phosphate backbone of a single DNA-strand is referred as Single-strand break, whereas the rupture of backbones on both DNA strands is referred as Double-strand break. Single-strand breaks are associated with increased genetic instability [Caldecott (2008)]. Double-Strand Breaks lead to a discontinuity on the DNA molecule and are associated with death of the organism and/or increased genetic instability, such as chromosomal translocation, a [Jeggo and Lobrich (2007)].

The rupture of a strong bond, such as the phosphate backbone bond, requires a high activation energy for the reaction to initiate. The main sources of high-energy radiation in humans are exposure to Sunlight and cosmic rays [NASA (1998); Cannan and Pederson (2016)]. Other sources of Double-Strand Breaks are exposure to endogenous chemical drugs and accumulation of damages near a susceptible area. Furthermore, Double-Strand Breaks appear as by-products of different cellular processes. For example, the manipulation and enzymatic

reactions during transcription and replication may lead to Single-Strand and Double-Strand Breaks.

The repair pathway for Double-Strand Breaks highly depends on the source of the damage and the live cycle of the cellular organism at the point of the DNA repair mechanism start. The cause of the damage may leave a “clean” or “dirty” lesion. A clean lesion would only have a deoxyribose to phosphate bond broken, whereas a dirty lesion may have other secondary bonds broken or molecules attached near the lesion. Dirty lesions slow, or even inhibit, the repair mechanisms. Accumulation of sequence damages or attached molecules impede the proper action of repair proteins at the damage location.

The two major pathways recognized for the repair of Double-Strand Breaks are Non-Homologous End-Joining and Homologous Recombination. Non-Homologous End-Joining is the main repair mechanism as it can operate on any phase of the cell cycle. Homologous Recombination, however, is usually limited to only 2 of the 4 to 5 cell cycles [Takata, et al. (1998)]. These two cell cycles are the synthesis (S) and second-growth (G2) phases. During these phases the cellular organisms have an identical copy of each chromosome facilitating the Homologous Recombination mechanism.

Studies on these repair mechanisms might elucidate the effect of drugs and radiation on the genetic material of various organisms and subsequently open the door for future therapies. An example of the importance of understanding these mechanism is the study of high-energy radiation and particle effect on an astronaut’s health while in outer-space [NASA (1998)]. A more in depth description of the Homologous Recombination will be presented at chapter III (Homologous Recombination) of this thesis.

I.7.2.1.2 DNA base mismatches

The hydrogen bonds responsible for the stabilization of the union between both DNA helixes is affected by faulty pairing of bases. The DNA bases are divided into purines and pyrimidines based on the chemical composition of the nitrogenous base. Adenine and guanine are purines (hexane and pentane bases, Figure I-9), whereas, cytosine and thymine are pyrimidines (hexane bases, Figure I-9). Inside a DNA double-helix, the bases are paired purine with pyrimidine; specifically, adenine with thymine and guanine with cytosine. The resultant hydrogen bonds make the guanine-cytosine couple stronger [Fonseca Guerra and others (2000)].

Mispairing of a base with its counterpart weakens the double-helix structure and is a known form of DNA base mismatch. Other modifications to the chemical

composition or loss of these bases on one DNA-strand trigger a mismatch. One of the main causes of mismatch lesions occurs in the process of DNA replication [Kunkel and Erie (2005)]. During replication, the polymerase in charge of copying the DNA has an error probability of 1 mismatch in 100.000 nucleotides inserted [Johnson, et al. (2000)]. For a human cell, this error probability translates into 120.000 mismatches after the process of DNA synthesis. The newly replicated DNA strand accumulates mismatches compared to the template or original strand. A mechanism in charge of "proofreading" the DNA localizes the mismatches by the difference between the newly synthesized and the template strand. The importance of this mismatch repair mechanism makes it a highly conserved process among prokaryotes and eukaryotes [Modrich (2016)]. When proteins from this mechanism are modified in detriment of their activity the organisms are more susceptible to the accumulation of mutations. However, the inactivation of the "proofreading" mechanism is thought to be involved in the evolution of living organisms [Denamur, et al. (2000)]. A further analysis on the role of proteins from the mismatch-repair mechanism is given in chapter IV (Horizontal-gene transfer).

I.8 Thesis objectives

The general objectives of the projects studied with Atomic Force Microscopy are:

- To study the role of the partition system proteins ParB, centromore-binding protein, in *Bacillus subtilis* and *Caulobacter crescentus* and of the proteins TubR, centromere-binding protein, and TubZ, motor protein, in *Clostridium botulinum*.
- To study the Homologous Recombination pathway throught the proteins CtIP from *Homo sapiens* and a homologous protein to *Homo sapiens* DNA2 from *Geobacillus stearothermophilus*.
- To study the role of the mismatch-repair pathway protein, MutS, in the integration of exogenous DNA in *Bacillus subtilis*.

II. PLASMID PARTITION SYSTEMS

A plasmid is a small, up to $2x10^5$ pairs of bases, circular DNA molecule found inside various living organisms. The genes codified in these molecules are not essential as those codified in chromosomes but provide some benefit for the cell under certain conditions. The cellular mechanisms responsible for the maintenance and organisation of these molecules are of importance due to their intrinsic role in the adaption and evolution of the organism.

"Because all of biology is connected, one can often make a breakthrough with an organism that exaggerates a particular phenomenon, and later explore the generality." – Thomas Cech, chemistry Nobel laureate 1989.

The Introduction has presented the necessary knowledge for a broad comprehension of this thesis results and discussions. However, essential biological processes have evolved differently on each type of organisms. The main purpose of mechanisms such as the partition system of a cell has remained constant while the molecules and intermediate steps have seen variations among organisms during evolution. To study and comprehend these systems one must understand the role of the many molecules involved. In this section of the thesis I will introduce the partition systems and the results obtained in this area with Atomic Force Microscopy to the reader.

II.1 Introduction

As described in section 7.1 in chapter I of this thesis; plasmid partitioning is the organisation of plasmidic molecules inside an organism for a uniform distribution of this genetic material among the resulting cellular offspring. Today, there are four different protein systems recognized for plasmid partitioning. The four systems are divided based on the molecular structure of the NTPase-motor molecule which, in turn, is translated into a difference in their mechanisms [Gerdes, et al. (2000); Guynet and de la Cruz (2011); Oliva (2016)] [Figure II-1]. The molecular differences among systems are summarized as:

-Type I partition systems have a NTP-ase protein classified as a Walker-type ATPase [Gerdes, et al. (2000)]. Type I systems are further divided according to the molecular structure of the Centromere-Binding Protein. In type Ia this protein has an HTH (helix-turn-helix) DNA-binding domain while in type Ib it has a RHH (Ribbon-Helix-Helix) DNA-binding domain [Schumacher (2012)].

-Type II systems have an actin-like ATPase or GTPase as the NTP-ase motor protein [Gerdes, et al. (2000), Schumacher (2012)].

-Type III systems have a tubulin-like GTPase as the NTP-ase motor protein [Oliva (2016)].

-Type IV systems do not have a NTPase motor protein. Nowadays, only one plasmid, R388, is known for having this partition system [Guynet and de la Cruz (2011)].

The first characterised structure of partition system proteins was that of the type II NTPase motor protein. The structural techniques were not suited for the characterisation of the high-flexibility Centromere-binding proteins [Schumacher (2012)]. Albeit different in molecular structure, the main purpose of the four partition systems is the same: the relocation and organisation of plasmidic content

during cell division. Continuous research on these cellular mechanisms revealed the relation between the plasmids present on an organism, the characteristics of the organisms and partition systems in charge of the segregation of the plasmids. To study and understand the partition system machinery on different organisms, the research project is focused on the study of two organisms with type Ia partition systems, *Bacillus subtilis* and *Caulobacter crescentus*, and on one with type III partition system, *Clostridium botulinum.*

B. subtilis is a widely studied bacterium due to its usage as an inmunostimulatory agent in the 1900s [Ciprandi, et al. (1986)]. The availability of this bacterium in laboratories made it a prominently studied organism. Today, *B. subtilis* is the most studied gram-positive organism and it is a model organism for the study of chromosome replication and differentiation [Pavlendova, et al. (2007)]. Due to similarities in chromosome and plasmid segregation systems, this organism is an ideal model for the study of plasmid partition proteins and further understanding of chromosome segregation.

C. crescentus cells have an uncommon differentiation between daughter cells. During division, the progeny of a *C. crescentus* cell is divided into “swarmer” and “stalked” cells [Huguenel and Newton (1982)]. This differentiation arises from a sequential modification of proteins and signals during cellular division [Hecht and Newton (1995)]. This organism is used as model for studies on the regulation of the cell cycles and cellular differentiation [Nierman, et al. (2001); Quon, et al. (1996)]. Opposite to *B. subtilis, C. crescentus* is a Gram-negative bacterium but, likewise, possesses a plasmid with a type Ia partition system.

Lastly, *C. botulinum* is mostly known for producing the botulinum toxin, which is a neurotoxin codified in the C-st plasmid of this organism. It is the deadliest toxin so far discovered [Peck (2009); Barash and Arnon (2014)]. The plasmid codifying this toxin, prophage c-st, segregates under a type III partition system. The majority of the Centromere-binding proteins are dimers in solution and form the DNA-binding domain at the interphase of both monomers. In contrast, the Centromere-Binding protein of the c-st plasmid is a stable monomer in solution [Martín-García, et al. (2018)]. The characteristics of the plasmid and the Centromere-binding protein make this a unique partition system.

The features and current knowledge on type Ia and III partition systems are further introduced in the following sections, respectively.

II.1.1 The type Ia partition system

Type Ia are the most abundant and studied partition systems. This wis due to the high availability of organisms with this partition system, mainly *Escherichia coli* and *Bacillus subtilis*, and to the ease to study a mechanism for small and low copy-number plasmids [Balzer, et al. (1992); Lewis and Errington (1997); Martin, et al. (1987); Gerdes, et al. (2010)]. The acquired knowledge on this system during the 1900s made it a model mechanism to understand the segregation of DNA in bacteria [Wang, et al. (2013)].

Type Ia systems are also referred as ParABS systems. ParABS refers to the names of the three essential elements of a partition system (except for partition system IV). *ParS* denotes a repeated sequence located near the origin of replication. In *B. subtilis*, the *parS* sequence is a 16 bp palindromic sequence (5'-TGTTNCACGTGAAACA-3') [Breier and Grossman (2007)]. The sequence is further recognized by a Centromeric-Binding protein, ParB. In *B. subtilis* and *C. crescentus,* ParB, also referred as Stage 0 sporulation protein j (Spo0j), is a DNA-binding protein with a Helix-Turn-Helix domain. Recent studies have shown specific and non-specific binding of ParB to DNA. The non-specificity of the binding is associated with DNA condensation during cellular division [Schumacher (2012); Taylor, et al. (2015)]. Lastly, ParA refers to a Walker-type motor protein activated by the hydrolisation of ATP [Walker, et al. (1982); Hwang, et al. (2013)]. The consumption of ATP by ParA proteins bound to the ParB-*parS* complex is responsible for the movement of the DNA during cellular division. The movement of the plasmid can be directed towards a cell-pole or to separate plasmids between cellular daughters [Figure II-1].

There are four basic working models proposed by the scientific community on the partition system mechanism [Guynet and de la Cruz (2011); Schumacher (2012] [Figure II-1]. These models describe the movement of the plasmid and the segrosome complex for a successful segregation of the DNA molecule connected to the action of the motor protein. In the “pulling” model, it is proposed that the polymeric motor protein is bound on one end to a cell pole and to the plasmid molecule on the other end; subsequent depolymerisation shortens the distance between the cell pole and the DNA molecule [Figure II-1 A]. In the “pushing” model, the polymeric motor protein is capped by two plasmid molecules; polymerisation of the motor protein, contrary to the “pulling” model, increases the distance between the two identical plasmid molecules [Figure II-1 B]. In the “tramming” model, the motor protein detects the position of the plasmid molecule and polymerises around the segrosome moving the DNA in the direction of a cell pole [Figure II-1 C]. In the last model, the “pilot-fish”, the segrosome complex

uses the cellular segregation machinery due to the absence of an active motor protein [Figure II-1 D].

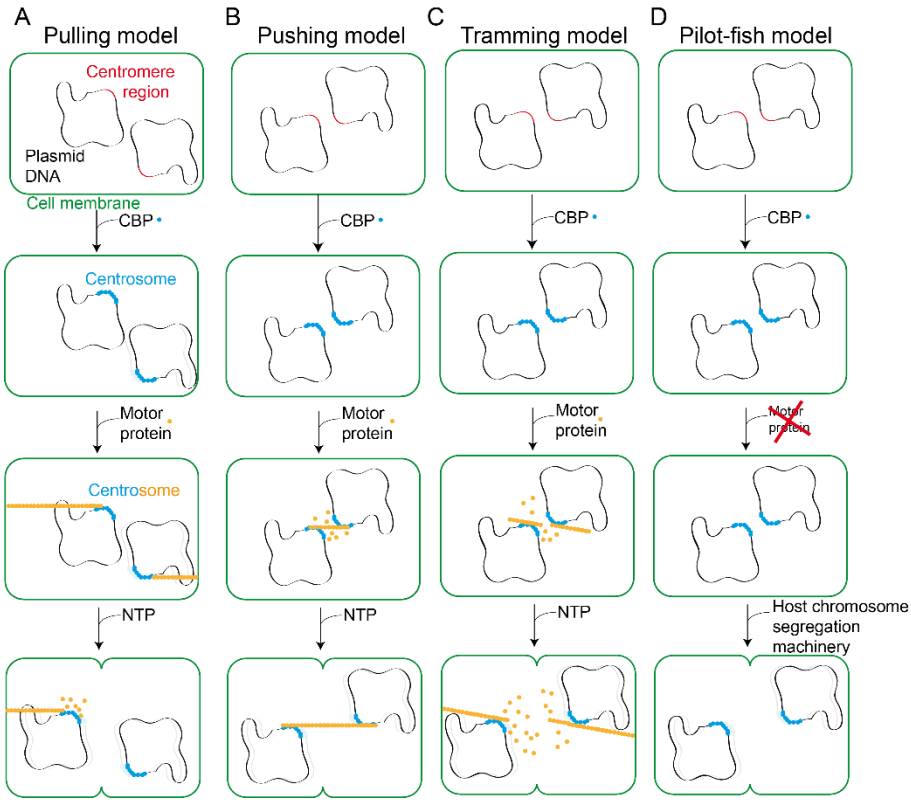


Figure II-1 Some proposed models for partition mechanism.

Models proposed for the mechanisms of various plasmid partitions systems. A) "Pulling" model. B) "Pushing". C) "Tramming" model. D) "Pilot-fish" model. Figure adapted from Guynet and de la Cruz (2011) and Schumacher (2012).

The general consensus in *B. subtilis* partition system mechanism is that of a "pulling" model with the concerted action of the three partition elements. The Centromere-Binding protein, ParB, attaches to the *par*S sequence to facilitate the spreading of other ParB proteins through the DNA molecule [Graham, et al. (2014)]. However, ParB is also capable of DNA-binding and spreading on DNA molecules without *parS* sequences [Taylor, et al. (2015)]. The resulting condensed ParB-DNA molecule binds to an ATP-bound ParA filament. The filament connects the ParB-DNA complex with a cell pole. Depolymerisation of the ParA filament by the interaction with the ParB-DNA complex moves the complex through the cell. The movement stops when each plasmid copy arrives to opposite cell poles [Gerdes, et al. (2010)] [Figure II-1A; Figure II-2C].

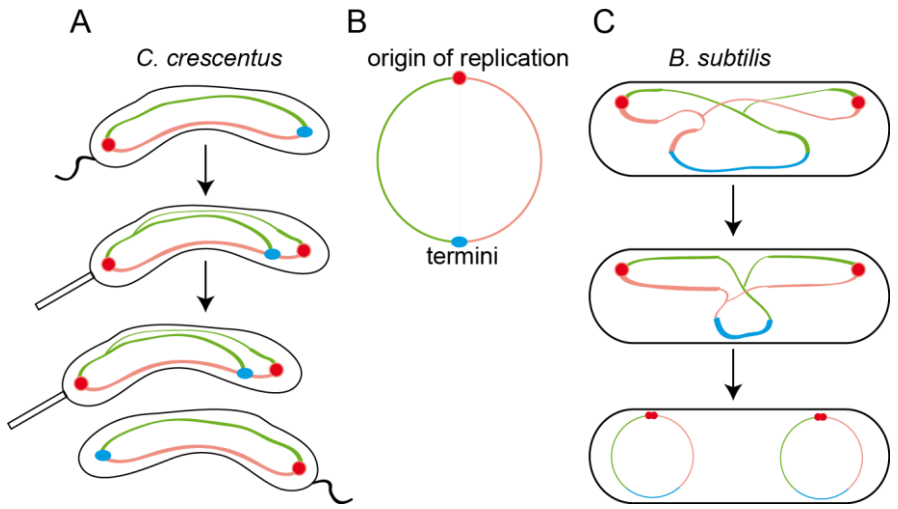


Figure II-2 Model for chromosome segregation in C. crescentus and B. subtilis.
A) Model for chromosome segregation in C. crescentus showing the difference in "swarmer" cells, flexible flagellum, and "stalker" cells, rigid flagellum. B) Schematic DNA structure with the origin of replication highlighted with a red dot and the terminal with a blue section. C) Model for chromosome segregation in B.subtilis. Image adapted from Badrinarayanan, et al. (2015).

In *C. crescentus*, contrary to *B. subtilis*, the understanding of this mechanism is limited. There is a recent study about the recognition of the *parS* sequence by the Centromere-Binding protein comparing it to that of *B. subtilis* [Tran, et al. (2018)]. Surprisingly, Tran, Stevenson et al. (2018) showed that the *parS* sequence of *C. crescentus* was similar to that of *B. subtilis*. The similarities made the *B. subtilis parS* sequence a suitable sequence for characterization of the Centromere-Binding protein. Biochemical studies showed a reliable binding of *C. crescentus* ParB to both sequences. This similarity between sequences allowed the study of the role of the Centromere-Binding protein using similar DNA molecules as the ones used for the *B. subtilis* protein.

II.1.2 Type III partition system

Type III partition systems were recently discovered by Tang, et al. (2007) inside *Bacillus thuringiensis*. Opposite to type I and type II, in type III systems the motor-protein is a GTPase protein similar to the eukaryotic protein tubulin [Gunning, et al. (2015)]. Tubulin is a filamentous protein present in the cellular cytoskeleton. The relevance of a tubulin homologue protein in prokaryotes increased the interest of the scientific community in studying the newly found system [Ni, et al. (2010); Aylett and Lowe (2012)]. The type III partition system elements are named TubRZC. Similar to type Ia partition systems, the three letters refer to the

Centromere-Binding, TubR, the GTPase tubulin-like motor protein, TubZ, and the DNA sequence recognised by the Centromere-Binding protein, *tubC*.

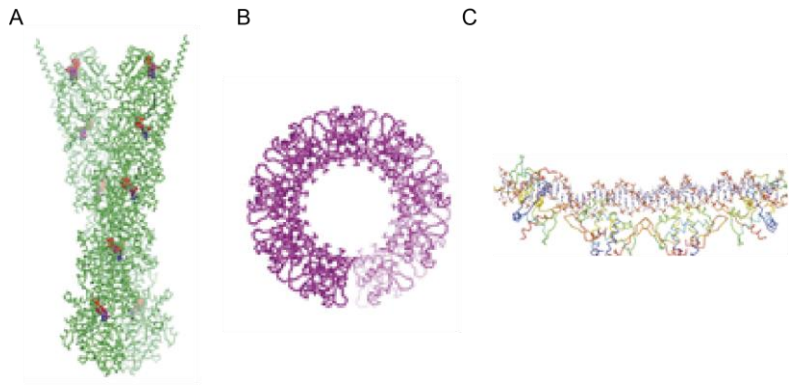


Figure II-3 Crystal structures for TubZ, TubR and the TubR-tubC complex.
A) Crystal structure of the TubZ-filament from Bacillus thuringiensis. B) Crystal structure of the TubR-complex from Bacillus megaterium. C) Crystal structure of TubR-tubC complex from Bacillus thuringiensis. Image adapted from Aylett and Lowe (2012).

Structural studies of the type III partition system proteins with Electron Microscopy and crystallography of *Bacillus thuringiensis* and *Bacillus megaterium* organisms found an active role of the DNA-binding protein TubR. The binding of TubR to DNA produces a bend on the complex structure around the *tubC* sequence [Aylett and Lowe (2012)] [Figure II-3C]. These bent structures are displayed as recognition positions for TubZ or TubZ polymeric filaments [Fuentes-Perez, et al. (2017)]. The bent DNA structure could act as the anchoring point for the polymeric filamentous motor-protein to start the partition mechanism. Nonetheless, the complete mechanism and the recognition of the different intermediate structures remains unknown.

II.2 Objectives

The aim of this project is to study and understand the role of the Centromere-Binding proteins of the partition systems in *B. subtilis* and *C. crescentus* and of the Centromere-Binding protein and motor protein in *C. botulinum*. The partition systems of *B. subtilis* and *C. crescentus* have been widely studied but the observation of the segrosome complexes remains undone. Furthermore, the project is complemented with the study of the less known partition system of *C. botulinum*.

The objectives were:

1) To observe the condensation complexes analysed in Taylor, et al. (2015) by Magnetic Tweezers of the *B. subtilis* partition segrosome.
2) To study the role of *C. crescentus* ParB in the formation of the segrosome complex and compare it to *B. subtilis* segrosome.
3) To study the role of the Centromere-Binding protein and the motor-protein in the formation of a segrosome complex in the type III partition system of *C. botulinum*.

II.3 Materials and Methods

II.3.1 *Bacillus subtilis* partition system project

II.3.1.1 ParB protein expression and purification

ParB was purified and subsequently biochemically analysed by Dr. Gemma Fisher from Prof. Mark S. Dillingham's laboratory (University of Bristol, UK) following the "Material and Methods" section on Taylor, et al. (2015).

II.3.1.2 DNA fabrication and purification

DNA molecules chosen for the analysis of ParB were prepared following the "Material and Methods" corresponding section on Fisher, et al. (2017) without the ligation step of digoxigenin or biotin handles necessary for Magnetic Tweezers.

II.3.1.3 Sample preparation and image acquisition for AFM

DNA-protein sample preparation for air acquisition

Samples for the study of the DNA-ParB interaction were initially incubated in buffer (50 mM Tris HCl pH 7.5 and 100 mM NaCl) at 1 nM DNA and 5 nM ParB concentrations. The resulting mixture (20 µl) was incubated for 5 minutes. Then, a solution of 7.5 mM $MgCl_2$ was added to the reaction. The final volume, ~21 µl, of the reaction was pipetted on a pre-cleaved mica surface. The solution deposited on mica was incubated for 30 seconds, washed three times with one millilitre of Milli-Q water and dried under constant flow of nitrogen air [Lyubchenko, et al. (2009)].

AFM image acquisition conditions

Images were obtained using a Nanotec Electronica S.L. Atomic Force Microscope using the amplitude modulation mode (*tapping*) and the WSxM freeware [Horcas, et al. (2007)] unless otherwise stated. Air sample images were obtained at room conditions of temperature and humidity at a rate of 1.48 lines per second and 512 lines per image and pixels per line, 346 second per image. The cantilevers employed were PPP-NCH (Nanosensors, Switzerland).

II.3.1.4 Image processing

In tapping mode, an Atomic Force Microscope registers the cantilever oscillation amplitude and the phase difference between the cantilever response and the applied oscillation. Scanning of a tip along a surface creates three images under modulated amplitude premises [Figure II-4A]. 1) A topographic image from the

surface movement in the *z* direction as to maintain a constant oscillation amplitude. 2) An amplitude image from the oscillation amplitude value at each scan point. 3) A phase image from the oscillation phase value.

The raw images presented in Figure II-4A are coloured numerical matrixes. Each pixel contains the value for the variables represented at that point. The numbers are given a colour on the basis of their value. The clearer the colour the higher the value [Z scale bar in Figure II-4 B3]. In topographic images, high spots on a surface would appear clear [DNA in Figure II-4A]. To correct them for a more precise and visual result the raw images are processed with the WSxM freeware [Horcas, Fernandez et al. (2007)].

Final topographic images are processed with three tools: 1) A basal plane to correct sample, cantilever or motor tilt; 2) A parabolic plane, to correct for the piezoelectric scanners movement and; 3) A *flatten* correction, to adjust all the scanned lines to a same basal level based on the image background [Gimeno, et al. (2015)] [Figure II-4B].

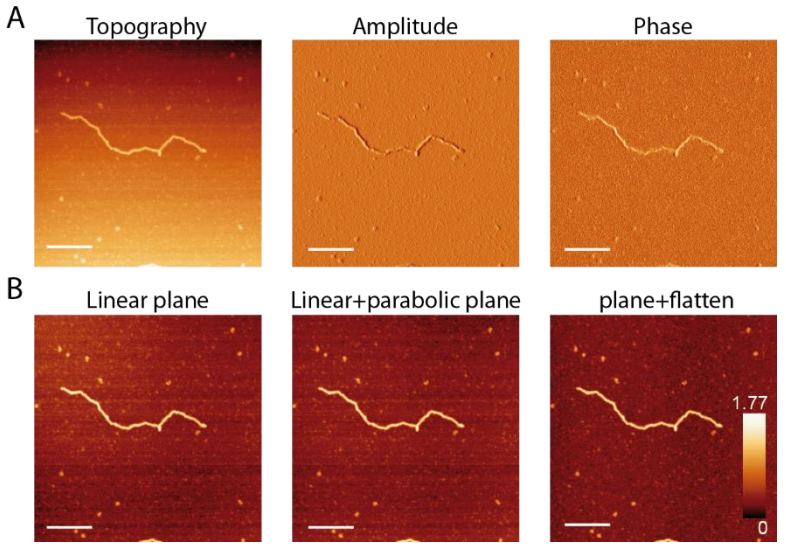


Figure II-4. Acquired images and image processing on an Atomic Force Microscope.
A) Three raw images acquired by an AFM operated on tapping mode: Topographic image, amplitude image and phase image. B) Topographic image corrected by: a basal linear plane, a linear and a parabolic plane and a linear and parabolic plane together with a flatten. Scale bar on images is 80 nm. Z colour scale is calibrated to a black colour at 0 nm black and a white colour at 1.77 nm. Z colour scale presented on image B.(plane+flatten) corresponding to the colour code of images on B section.

II.3.2 *Caulobacter crescentus* partition system project

II.3.2.1 ParB protein expression and purification

The protein ParB was purified by Adam Jalal from Dr. Tung Le's laboratory (John Innes Center, UK).

II.3.2.2 DNA fabrication and purification

DNA constructs were prepared by Dr. Clara Aicart-Ramos from Prof. Fernando Moreno-Herrero laboratory. The DNA molecules for AFM experiments consisted in a large insert containing one or two sets of 6 *parS* sites plus an additional *parS* site. The molecules were obtained by digestion with SphI and BamHI of the pET28a(+)-2xParS 2set (6ParS) plasmid. After digestion, the fragment was separated with electrophoresis on a 1% agarose gel and purified from the gel with precautions to avoid exposure of the substrates to intercalating agents as well as exposure to UV light. The gel was run with a side lane incorporating a small fraction of the preparation. That side lane together with the lane of markers were cut and separately stained with SYBR green, and the position of the band was marked in the LED transiluminator. The band in the gel was then excised based on the estimation of its position obtained from the stained side lanes and purified with QIAquick Gel Extraction Kit (QIAGEN).

II.3.2.3 Sample preparation and image acquisition for AFM

Samples for the study of the DNA-ParB interaction were initially incubated in buffer; 10 mM HEPES pH 7.5 and 75 mM NaCl; at a 1 nM DNA and 1 nM ParB. The resulting mixture volume, 20 µl, of DNA:ParB was incubated for 15 minutes. Then, a solution of 7.5 mM $MgCl_2$ was added to the reaction. The final volume, ~21 µl, of the reaction was pipetted on a pre-cleaved mica surface. The final deposited solution on the mica was incubated for 30 seconds, washed three times with one millilitre Milli-Q water and dried under constant flow of nitrogen air [Lyubchenko, et al. (2009)].

AFM image acquisition conditions

Images were obtained following the methods presented in chapter II.3.1.3.

Image processing was performed following the chapter II.3.1.4

II.3.2.4 Control samples

A proper base of control samples is key for the characterization of DNA-protein interaction processes. In this case, the DNA molecules were incubated with

protein and deposited onto a mica surface. Final reaction products together with protein and DNA were characterized on AFM images. As a control, initial DNA molecules with and without *parS* sequences were characterized on Atomic Force Microscopy images by contour length analysis. In this case, the three DNA molecules have 2124 bp (no *parS*), 2114 bp (13 *parS*) and 1490 bp (7 *parS*).

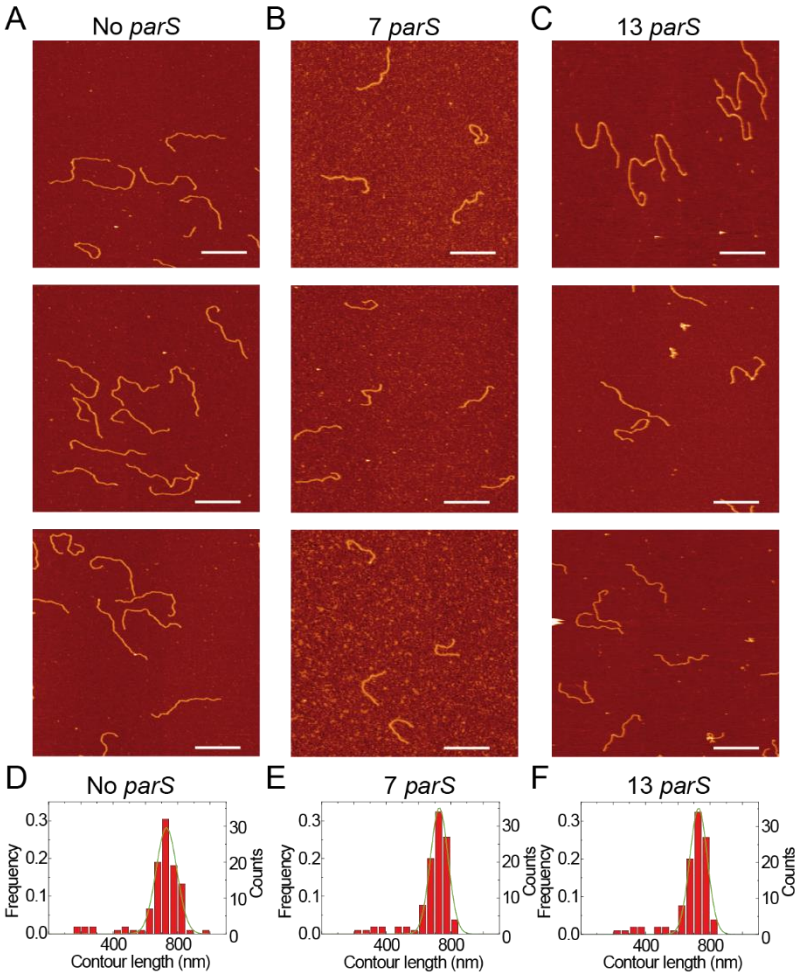


Figure II-5. DNA control samples for C. crescentus ParB project.

A) AFM images of the 2124 bp DNA without parS sequence. B) AFM images of the 1460 bp DNA with 7 parS sequences. C) AFM images of 2114 bp DNA with 13 parS sequences. Scale bars are 400 nm. D) Contour length histogram for the 2124 bp DNA with a Gaussian fitting centred at 730 nm. E) Contour length histogram for the 1460 bp DNA with a Gaussian fitting centred at 505 nm. F) Contour length histogram for the 2114 bp DNA with a Gaussian fitting centred at 729 nm.

Linear double-stranded DNA molecules were incubated and deposited onto a mica surface following the methods for DNA-Protein sample preparation without adding any protein. The control revealed a homogenous sample of linear polymeric rods deposited on a flat surface [Figure II-5]. As expected for a stochastic population distribution with independent variables (Central limit theorem), the population of contour length follows a normal distribution pattern with a unique peak centred around the AFM measured contour length [Sanchez-Sevilla, et al. (2002)]. The DNA length in base pairs can be inferred from the contour length measued in the AFM. Previous studies show a relation of 3.02-3.39 Å per bp for double-stranded DNA deposited with Mg^{2+} on a mica surface [Rivetti, et al. (1996); Rivetti and Codeluppi (2001); Fang, et al. (1999)]. The crystallographic length for B-DNA obtained by X-ray diffraction data is 3.4 Å per bp [Wing, et al. (1980)]. From our analysis the extracted contour lengths are 730 nm for the 2124 bp DNA, 729 nm for the 2114 bp DNA and 505 nm for the 1460 bp DNA [Figure II-5 D-E-F]. This leads an average rise per bp of 0.34±0.01 nm from the contour lengths of the control DNA molecules.

II.3.3 *Clostridium botulinum* partition system project

II.3.3.1 TubZ protein expression and purification

Wild type TubZ and the mutant TubZ^{T100A} were cloned and purified by Dr. Maria A. Oliva-Blanco as described in Oliva, et al. (2012).

II.3.3.2 TubR protein expression and purification

TubR was purified by Dr. Maria A. Oliva-Blanco as described in Martin-Garcia, et al. (2018).

II.3.3.3 DNA substrates fabrication and purification

DNA substrates are referred based on the sequence of the molecule from the 5' end: C is a 352 bp centromere region, P is the TubR promoter, R is the TubR gene, Z is the TubZ gene and NPD is a non-coding DNA sequence. The DNA molecules presented in this project are *tubC* (352 bp), CPRZ (866 bp), NPDC (907 bp), RCPR (977 bp) and a fragment of the c-st plasmid with 2010 bp. The molecules were obtained by PCR using small oligos and the pUC57 vector including the whole c-st partition operon (Shinegene). The NPDC fragment was made by ligation of a PCR product between the *tubC* molecule and the DNA related to the N-terminal domain of a bacterial cell division protein called FtsZ from *B. subtilis*. Hybridisation of complementary strands was conducted in a thermocycler by heating the mixture to 95°C and cooling it down to 10°C at 1°C/min. The correct hybridisation of these fragments was checked by polyacrylamide gel electrophoresis and Gel-Red (Biotium) staining. Two DNA molecules, AFM-C1 (508 bp) and c-st (2010 bp), not containing the centromere region were used as controls in AFM experiments. AFM-C1 was produced by PCR using appropriate oligos and the plasmid pSP73-ParB [Taylor, et al. (2015)]. AFM-C2 was produced by enzymatic digestion of a plasmid derived from the pET28a-ParB with *Bam*HI and *Sph*I.

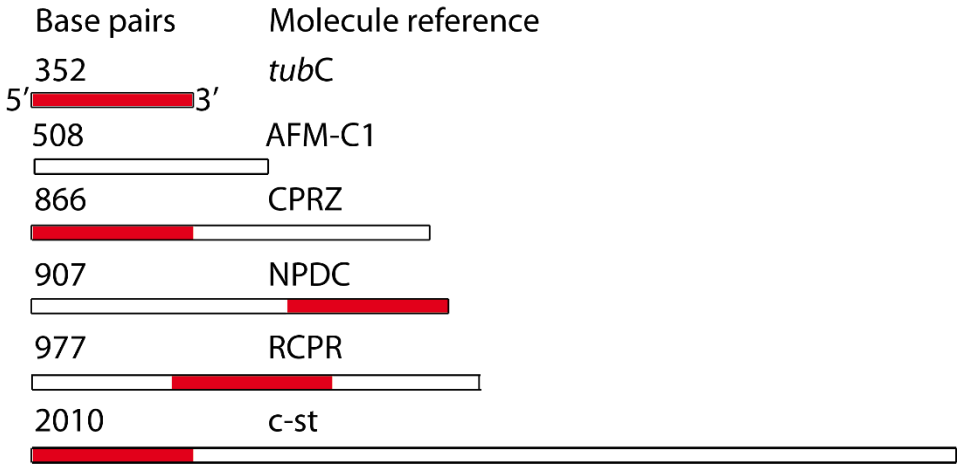


Figure II-6. Fbricated DNA molecules for Atomic Force Microscopy.
Fabricated DNA molecules for the TubRZC project. Red bar marks the position of the centromere sequence with 352 bp length. The only molecule with a tubC region centred is RCPR while AFM-C1 is the control molecule. The TubR promoter is a 16 bp region near the tubC sequence. The TubR is included in CPRZ, RCPR and c-st molecules while the TubZ gene is partly included in CPRZ (16 bp of 1677 bp) and c-st (1269 bp of 1677 bp).

II.3.3.4 Sample preparation and image acquisition for AFM

TubZ filament assembly and sample preparation for AFM

Proteins were diluted in PKE assembly buffer (50 mM Pipes/KOH [pH 6.5], 100 mM potassium acetate, 1 mM EDTA) with or without 200 g/L Ficoll-70 (prepared in PKE buffer), and polymerisation was initiated by the addition of nucleotide (1 mM GTP or 2 mM GDP) and 6 mM magnesium acetate. Polymer formation was quantified by sedimentation assays. The supernatants were carefully withdrawn and the pellets were resuspended in the same volume of buffer.

Filaments were deposited onto freshly cleaved mica surface pre-treated with 20 µl $MgCl_2$ (100 mM). After one minute of adsorption, the sample was gently washed with 3 ml Milli-Q water supplemented with 1 mM GTP and 6 mM magnesium acetate. GDP bound filaments were assembled as above, but with the addition of 20 µM wild type CbTubZ. Assembly time was set to 1 h and the mica was functionalized with 0.5% glutaraldehyde during the incubation time. GDP-bound filaments were adsorbed on the functionalized mica for 1 min and the sample was subsequently gently washed with 2 mL Milli-Q water supplemented with 2 mM GDP and 6 mM magnesium acetate. All AFM samples were dried by simple blotting.

TubR-DNA sample preparation for AFM

The TubR stock solution was first diluted in buffer to a final concentration of 1.71 µM. Different volumes of the solution were further diluted to the preferred concentration into a final volume of 20 µl containing 0.5, 1, or 2 nM DNA in 10 mM Tris–HCl pH 8.0, 75 mM KCl. The TubR–DNA mixture was incubated for 1 min and deposited onto freshly cleaved mica. The surface was thoroughly washed with 3 ml of Milli-Q water and dried under nitrogen air-flow [Lyubchenko, Shlyakhtenko et al. (2009)].

TubZ together with TubR-DNA sample preparation for AFM

TubZ and TubR-DNA samples were prepared separately as presented in the previous sections. Then, both samples were mixed and incubated together for 5 minutes before deposition onto a freshly cleaved and Mg^{2+}-pre-treated mica surface. Washing and drying steps were performed following the methods presented in the previous section.

AFM image acquisition conditions

Images were obtained following the methods presented in chapter II.3.1.3.

Image processing was performed following chapter II.3.1.4

II.3.3.5 TubR:DNA control samples

To characterize the binding of the centromere-binding protein TubR of *C. botulinum* to the centromere sequence *tubC,* we first incubated the protein with two DNA molecules without *tubC*. The samples were prepared with the two DNA molecules, 508 bp and 2114 bp (This molecules do not contain any of the sequences of the c-st plasmid introduced in Figure II-6), cover the same contour length and concentration range of the TubR:DNA tested with the DNA molecules with the *tubC* sequence. At one of the highest TubR:DNA ratios tested, the incubation of non-specific DNA with TubR protein produced no visible interaction as imaged by AFM [Figure II-7].

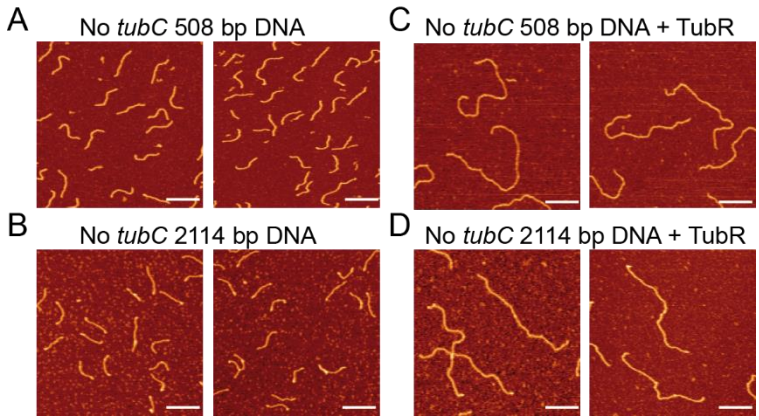


Figure II-7. TubR-DNA control samples imaged by AFM.
A) AFM representative images of the 508 bp DNA (2 nM) not incubated with TubR. B) AFM representative images of the 508 bp DNA incubated with TubR, 84 nM or 42 TubR:DNA ratio. C) AFM representative images of the 2114 bp DNA (0.5 nM) not incubated with TubR. D) AFM representative images of the 2114 bp DNA incubated with TubR, 21 nM or 42 TubR:DNA ratio. Scale bar is 200 nm.

As presented in section II.3.2.4, all the DNA substrates were deposited onto a mica surface following the sample preparation introduced the section without adding protein. The contour length of the DNA molecules was measured as a control for a proper preparation and characterization of the DNA molecules. Due to similarity of control images to the ones presented in Figure II-7 A and C and repetitive contour length analysis to the ones presented in Figure II-5 D-F, the data is not shown but is presented in Martin-Garcia, et al. (2018).

II.3.3.6 AFM image analysis for DNA and protein localization

Details of 240×240 nm^2 were selected from images of 2.5×2.5 μm^2 and processed (plane extraction, flattening, height selection and smooth filtering) using WSxM software [Horcas, Fernandez et al. (2007)]. A threshold height was applied to every AFM image. Data over 0.75 nm represented only the location of DNA molecules while data between 0.35 and 0.7 nm were considered as protein molecules, either TubZ or TubR. These two sets of heights were converted into plain text square matrixes of 49 points named as matrix A and matrix B. The correlation coefficient C_{AB} between matrixes was calculated following Equation 2,

$$C_{AB} = \frac{tr(AB^t)}{||A||||B||} \qquad \textbf{Equation 2}$$

Equation 2. Correlation coefficient between matrixes A and B.

II.3.3.7 TubR protein volume control

For TubR volume analysis, the protein was first diluted from stock in a buffer with 10 mM Tris–HCl pH 8.0 and 75 mM KCl to a final concentration of 8.5 nM. Then, the protein solution was supplemented with DNA and $MgCl_2$ to obtain a final concentration of 3.3 nM and 7.5 mM, respectively, and placed onto a freshly cleaved mica surface. After 30 s, the mica surface was washed with filtered-MilliQ water (Millipore, Billerica, MA) and blown dry in a gentle stream of nitrogen gas.

To obtain the protein volume we followed the procedure published in Fuentes-Perez, et al. (2012) using DNA as fiducial marker and obtained an average value of ~7 kDa [Figure II-8]. This value was in agreement with a monomeric CBP from a type III partition system [Ni, et al. (2010); Martin-Garcia, et al. (2018)].

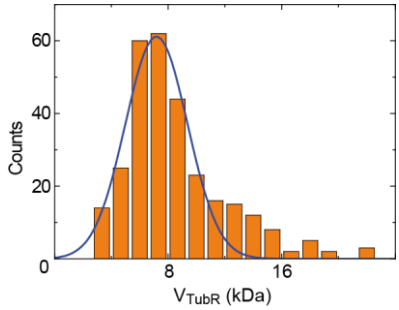


Figure II-8. Histogram of TubR volume on AFM images using DNA as fiducial marker. *Histogram of the TubR protein volume extracted from AFM images. The mean value corresponds to a protein volume of 7.8 kDa with a sample size of N=291.*

II.4 Results

These results were obtained in the study of the partition systems with Atomic Force Microscopy. The collaborator groups obtained relevant and complementary data with other biochemical analysis and techniques. In this thesis the data presented focuses on that obtained with the Atomic Force Microscope. However, the discussion relies mainly on the results presented here and on all the data obtained, including that acquired by collaborations.

II.4.1 Type Ia systems

II.4.1.1 *B. subtilis* ParB does not require the *parS* recognition domain to condense DNA

Joint projects between Mark S. Dillingham and Fernando Moreno-Herrero's laboratories have unravelled the role of the partition system protein *B. subtilis* ParB in DNA condensation and segregation [Taylor, et al. (2015)]. Moreover, recent studies have demonstrated the importance of the C-terminal domain of ParB for intermolecular binding and bridging. The C-terminal domain is crucial for the condensation of long DNA molecules. This was revealed with Magnetic Tweezers alone and combined with TIRF microscopy techniques [Fisher, et al. (2017); Madariaga-Marcos, et al. (2019)]. These techniques along with EMSA assays showed that Wild type ParB together with a *parS*-containing DNA produced two interaction complexes [Taylor, et al. (2015)]. At low ParB concentrations, the protein did recognize the centromere sequence, *parS*, and produced a DNA-ParB complex. However, at high ParB concentrations, a non-specific DNA-ParB complex was formed. To supplement these studies, I have imaged the *B. subtilis* ParB mutant proteins, R149G and R80A, together with DNA by Atomic Force Microscopy. The R149G mutation, located at the Helix-turn-Helix DNA-binding domain, was able to condense DNA at a lower rate compared to the wild-type protein [Autret, et al. (2001); Graham, et al. (2014); Chen, et al. (2015)]. The R80A mutation was previously described as a loss of function protein [Graham, et al. (2014)]. As expected, AFM images showed that the R149G variation was capable of DNA binding and condensing not only of one DNA molecule but several in a wool ball like conformation [Figure II-9 B]. The R80A variation showed no activity at low concentration but was capable of DNA binding at high concentrations [Figure II-9 C]. Both results were in agreement with previously proposed models for the ParB binding and bridging activities.

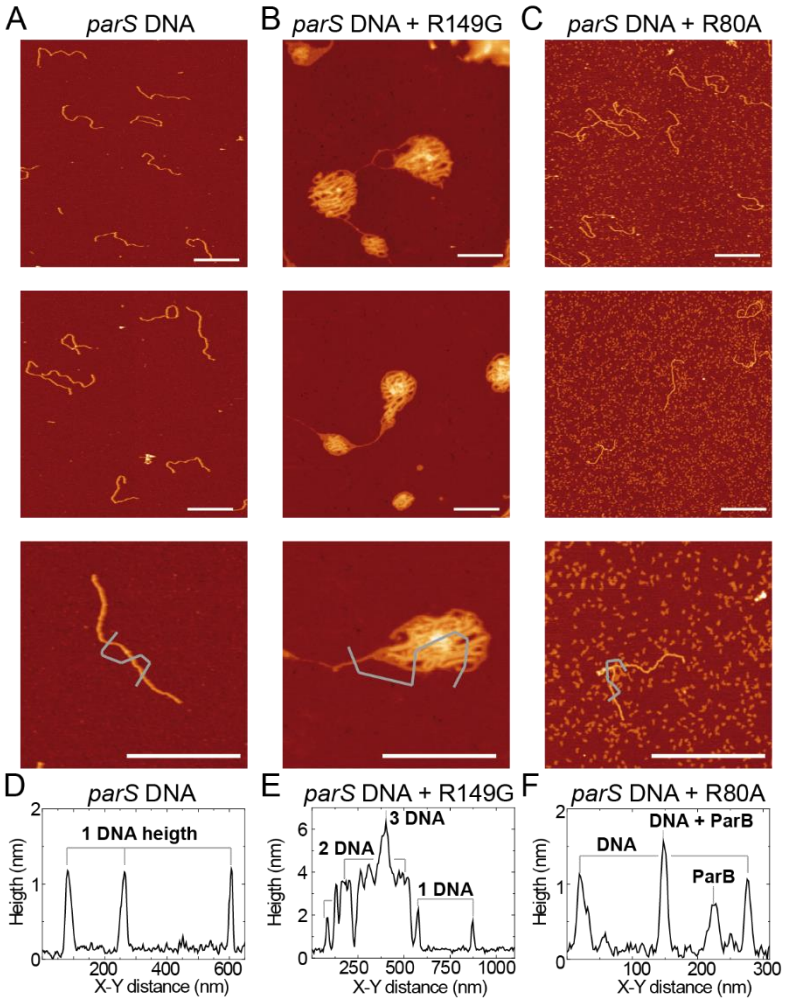


Figure II-9. AFM images and analysis of ParB-DNA complexes.

A) AFM images of a DNA molecule with parS sequence used for B. subtilis and C. crescentus ParB analysis. B) AFM images of R149G ParB and DNA complexes forming wool-ball-like structures. C) AFM images of R80A and DNA complexes. Scale bar of images is 500 nm. D) Profile along the grey line on A) with the typical height of a DNA molecule. E) Profile along the grey line on B) with the molecule height for the edge and centre of the "wool ball" compared to the typical DNA height. F) Profile along the grey line on C) with the molecule height compared to DNA height and protein height.

II.4.1.2 *C. crescentus* ParB binds to DNA producing secondary structures

In this project, we have studied the activity of *C. crescentus* ParB with *parS*-lacking and *parS*-containing DNA molecules by Atomic Force Microscopy. To understand the role of *C. crescentus* ParB we have used three DNA substrates with zero, seven or thirteen *B. subtilis parS* sequences distributed along the molecule. As previous studies have shown, *C. crescentus* ParB proteins are not-selective at *parS* recognition. *C. crescentus* ParB binds to *B. subtilis parS* sequence with high-fidelity [Tran, et al. (2018)]. This allowed the project to be carried out with DNA molecules containing *B. subtilis parS* used for the previously presented research. The employed DNA molecules have the *parS* sequences distributed in clusters of 16 bp and the sequence repeated six times. This 16 bp *parS* clusters are located in regions of 362 bp and due to the fabrication procedure of the DNA molecule it has a single *parS* located far from the clusters.

To image the segrosome complex of *C. crescentus,* I incubated *C. crescentus* ParB with three different DNA molecules. The three DNA molecules employed had 0, 7 and 13 *parS* sequences. The images showed several DNA molecules and ParB proteins deposited onto a mica surface. We found secondary structures not found on the control samples for the three DNA molecules incubated with ParB. These complexes accounted for 30% of the total number of molecules in the 0 *parS* DNA incubation and a 50% for the 7 and 13 *parS* [Figure II-5; Figure II-10 D-E]. To better characterize these secondary structures we divided the molecular species into 4 classes. These classes were divided as follows:

1) Linear DNA: DNA molecules elongated onto the mica surface; similar to the molecules seen in the control sample [Figure II-5].
2) Extruded DNA: DNA-protein complexes with a higher domain or domains. These domains were usually elongated perpendicularly to the expected DNA molecule direction.
3) Looped DNA: DNA-protein complexes with two separated DNA regions interacting and bridged together. In some cases, the loop was partially extruded (molecules with extrusions and loops are accounted into the looped species).
4) Complex structures: these structures may appear by the interaction of ParB proteins bound to various DNA molecules. They are unusually long molecules with an unusual shape.

Unexpectedly, the non-linear molecular species were present on the control samples. They accounted for a small population of 12-20 % of the total. In ParB-incubated samples the percentage rose to a 30%-56 % depending on the DNA molecule. For the 7 or 13 *parS* DNA, the non-linear molecules accounted for half of molecules in the field-of-views analysed. In addition, the presence of two *parS* clusters in the 13 *parS* DNA increased the population of looped DNA molecules

with respect to extruded DNA molecules compared to the 7 *parS* DNA molecule [Figure II-10E, blue population and red population].

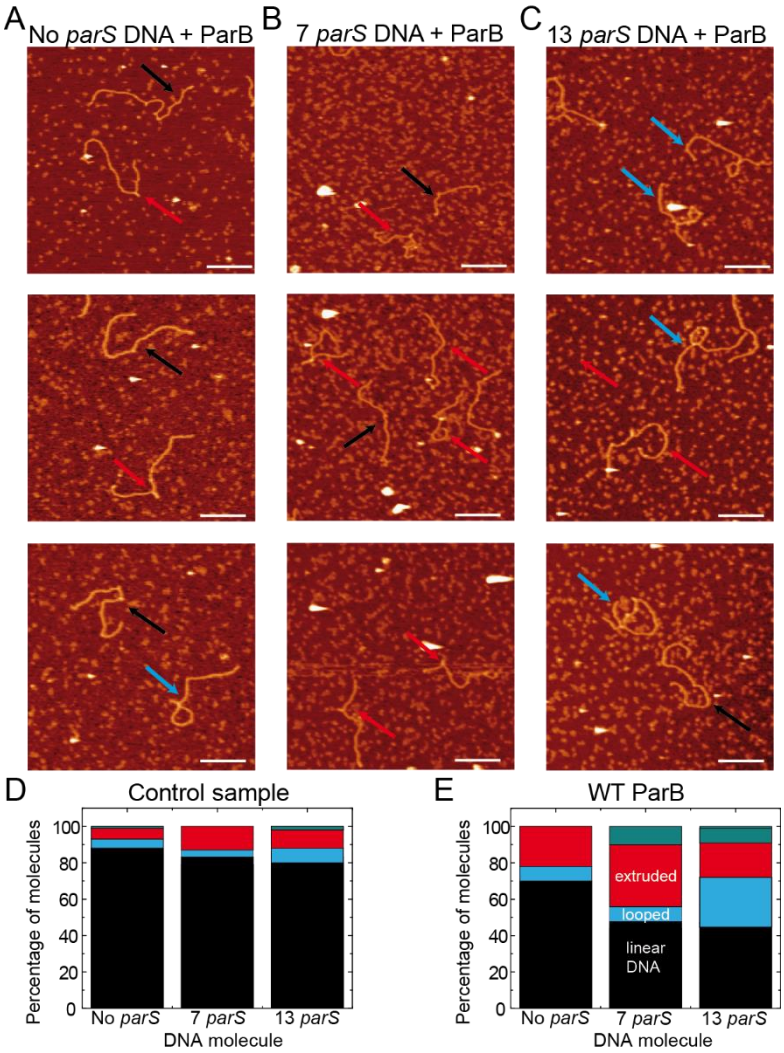


Figure II-10. AFM images and population graphs for the interaction of WT ParB with DNA. *A) AFM images of the no parS DNA incubated with Cc ParB. B) AFM images of the 7 parS DNA incubated with Cc ParB. C) AFM images of the 13 parS DNA incubated with Cc ParB. Black arrows show linear DNA molecules; red arrows show extruded DNA molecules and blue arrows show looped DNA molecules. Scale bar is 200 nm. D) Percentage of molecules for each of the four species in the control samples (black=linear DNA, blue=looped DNA, red=extruded DNA and green=complex structures), n=401 molecules analysed. E) Percentage of molecules for each of the four species in the WT Cc ParB samples. N=339 molecules analysed*

II.4.1.3 R104A and R163A mutations decreased ParB activity

To complement the study, the same analysis was performed with two mutants of *C. crescentus* ParB, R104A and R163A. The mutations R104A and R163A of *C.crescentus* ParB are similar to mutations R80A and R149G respectively for *B. subtilis* ParB. The mutation on the N-terminal, R104A, affects the dimerization domain while the mutation located at the C-terminal, R163A, disrupts the DNA-binding domain. However, the protein activity of the mutants was not expected to decrease dramatically as compared to *B. subtilis* case because *C. crescentus* ParB is not capable of DNA condensation but of a small DNA binding-effect. Expectedly, the mutations reduced the ratio of non-linear molecules in ParB-DNA samples [Figure II-11].

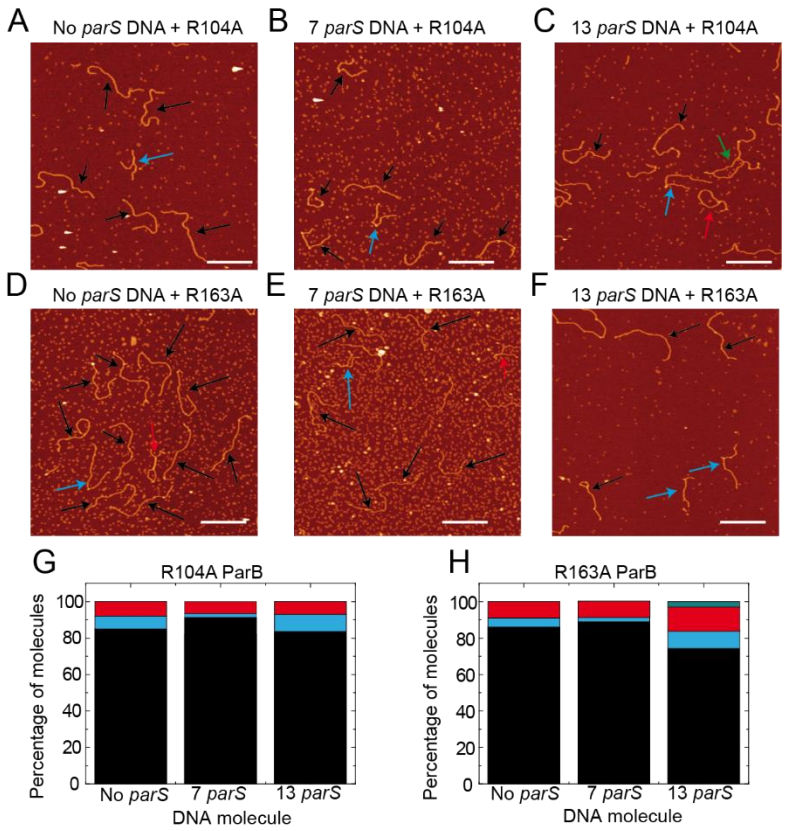


Figure II-11. AFM images and population graphs for R104A and R163A ParB mutants.
A) AFM image of the no parS DNA incubated with R104A. B) AFM image of the 7 parS DNA incubated with R104A. C) AFM image of the 13 parS DNA incubated with R104A. D) AFM image of the no parS DNA incubated with R163A. E) AFM image of the 7 parS DNA incubated with R163A. F) AFM image of the 13 parS DNA incubated with R163A. Black arrows show linear DNA molecules; red arrows show extruded DNA molecules; blue arrows show looped DNA molecules and green arrows show complex structures. Scale bar is 400 nm. G) Percentage of molecules for each of the four species in the R104A samples (black=linear DNA, red=extruded DNA, blue=looped DNA and green=complex structures). n=300 molecules analysed. H) Percentage of molecules for each of the four species in the R163A samples. n=300 molecules analysed.

The reaction products analysed with both mutants showed a population decrease of non-linear DNA species compared to the WT samples (from a 56% to a 24% of non-linear molecules). This effect suggested a lower DNA binding or bridging activity from these ParB mutants. The disruption of the dimerization domain or the DNA-binding domain reduced the appearance of secondary structures on linear DNA in the cases of *parS*-containing DNA molecules.

II.4.2 Type III systems

II.4.2.1 *C. botulinum* TubZ requires GTP for stable 4-stranded filament formation

Previous analysis of the tubulin-like protein TubZ with AFM at Fernando Moreno-Herrero's laboratory were focused on the filament and proto-filament formation of polymeric TubZ with GTP in *C. botulinum* [Fuentes-Perez, et al. (2017)]. However, other studies on *Bacillus thuringiensis* TubZ proposed a crucial change in filament structure during GTP nucleotide hydrolysis [Montabana and Agard (2014)]. In this project we aimed to understand the influence of nucleotide hydrolysis in the filament formation process of *C. botulinum* TubZ. Therefore, we applied the acquired knowledge in the TubZ filament formation process to analyse the T100A mutant of TubZ which disrupts the nucleotide hydrolysis domain impeding the hydrolisis of GTP by TubZ filaments [Fuentes-Perez, Nunez-Ramirez et al. (2017)]. In parallel, wild type TubZ was incubated with GDP instead of GTP to impede the hydrolysis process. As previously described, TubZ filaments were prone to depolymerisation [Fuentes-Perez, Nunez-Ramirez et al. (2017)]. Atomic Force Microscopy samples were not dried with Nitrogen gas but using a blotting method to avoid disturbing the deposited filaments [Section II.3.3.4]. This drying process resulted in salty mica surfaces in the Atomic Force Microscopy images but filaments were not depolymerised during surface adsorption. Furthermore, *in vitro* filament formation reactions were supplemented with Ficoll-70 to resemble a more crowded cytoplasmic medium. This experimental setting mimicked the medium where these filaments polymerise [Venturoli and Rippe (2005)]. Under these conditions, Atomic Force Microscopy images of Wild type TubZ GTP-filaments showed 4-stranded and 2-stranded filaments with a helical structure. The structure of the helix corresponds to a width of 7.7±0.7 nm, for 2-stranded filaments, and 13.7±2.2 nm, for 4-stranded filaments. The parameters were corrected for the tip-convolution effect [Fuentes-Perez, Nunez-Ramirez et al. (2017)]. In the same conditions, we found that GDP-filaments assemble into wider and thicker dimensions [Figure II-12 G-H]. GDP-filaments have 14±4 nm in width and appeared to be more rigid than GTP-bound ones. In contrast, we also found that the T100A filaments were similar to WT filaments [Figure II-12 D-E].

These results suggested a need for GTP-bound TubZ at the caps of the filaments. These caps control and maintain the formation of the polymeric filaments in a similar way as the microtubules are formed from other biological systems [Alushin, et al. (2014)]. However, in this case the filament disassembly was promoted by GTP hydrolisis followed by filament stiffening instead of by the production of a proto-filament curving as in microtubules. The stiffening of the filaments may provide the necessary forces to break the longitudinal contacts,

thus, weakening the contacts between filaments and supporting disassembly at the interface between GTP and GDP-bound TubZ monomers.

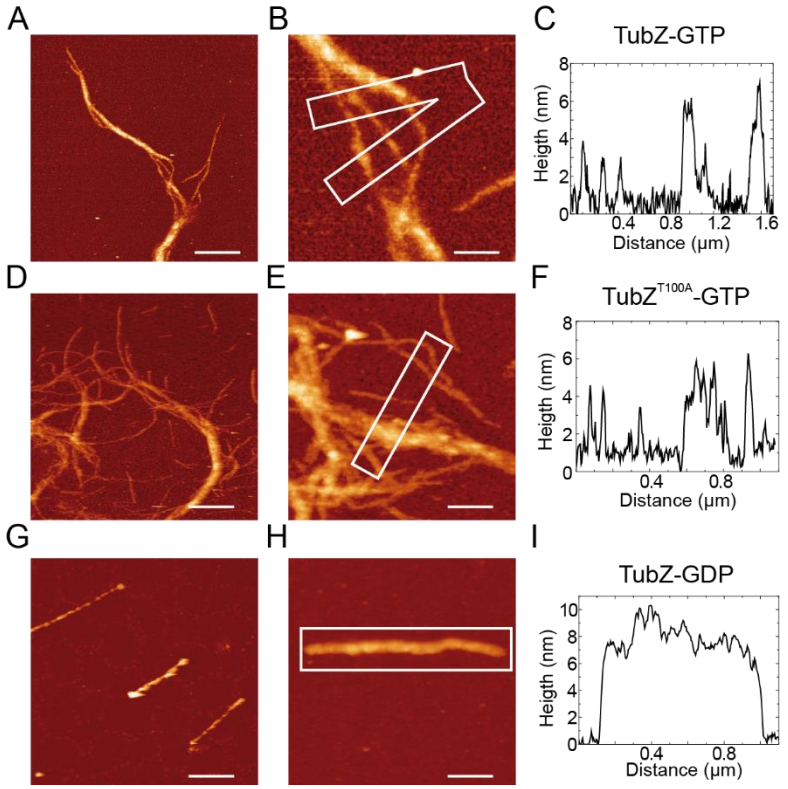


Figure II-12. AFM images and profiles of TubZ filaments.

A) AFM image of a WT TubZ-GTP filament with 2-stranded and 4-stranded structures shown. Scale bar 800 nm. B) AFM zoomed image of a WT TubZ GTP-filament with 2-stranded and 4-stranded structures shown. Scale bar is 200 nm. C) Profile along the squared region in B of three 2-stranded and two 4-standed filaments with different heights and widths. D) AFM image of various T100A TubZ GTP-filament with 2-stranded and 4-stranded structures shown. Scale bar 800 nm. E) AFM zoomed image of a T100A TubZ GTP-filament with 2-stranded and 4-stranded structures shown. Scale bar is 200 nm. F) Profile along the squared region in E of 3 2-stranded and 2 4-standed T100A filaments with different heights and widths. G) AFM image of various WT TubZ GDP-filament. Scale bar is 800 nm. H) AFM zoomed image of a WT TubZ GDP-filament. Scale bar 200 nm. I) Profile along the squared region in H, GDP-filaments with stable height corresponding to a helical structure but without 2- and 4- stranded filaments difference.

II.4.2.2 *C. botulinum* TubR forms a double-ring structure on a centromere DNA-sequence

TubR was described as an iteron-binding protein when it was first discovered in *Bacillus thuringiensis* [Tang, et al. (2007)]. To understand the role of the protein in the partition system of *C. botulinum,* TubR was incubated with a short DNA

molecule, 352 bp, containing the *tubC* region. For comparison, TubR was also incubated with another short DNA molecule, 500 bp, without a *tubC* region [Figure II-7]. Contrary to the control samples presented in section II.3.3.5, incubation samples prepared with *tubC*-containing DNA molecules revealed the presence of secondary structures on DNA molecules. The resulting secondary structures were in agreement with a previously proposed DNA-bending effect of TubR from other bacterial systems [Aylett and Lowe (2012)]. However, our Atomic Force Microscopy images showed a more pronounced effect. TubR was capable of selectively bending regions of ~400 bp on DNA molecules [Figure II-13 A]. This bending produced one or two curved regions often finished in a complete circular shape. The location of the bent structures correlated with the specific interaction region, *tubC* [Figure II-13 B]. From here on, the molecular species are classified as 1) linear double-stranded DNA molecules, 2) linear dsDNA molecules with one curved region, 3) with two curved regions, 4) linear dsDNA molecules with one looped region and 5) with two looped regions. To verify the role of the protein, we tested the bending effect at various TubR:DNA ratios: The lowest ratio tested corresponds to 4 TubR proteins per DNA molecule with 340 TubR proteins per DNA molecule as the highest. The lowest ratio shows that 50 % of molecules were typical linear dsDNA molecules (class 1) and 50 % were complex molecules (class 2, 3, 4, 5). The highest ratio shows a shift to more complex molecules: 10 % of class 1 molecules and 90 % of the other classes [Table 1]. The lack of secondary structures when the DNA did not contain the *tubC* region suggested that the DNA-protein complex, or segrosome, involved the formation of a DNA-bent structure, similar to other bacterial organisms showed in Figure II-7.

TubR:DNA ratios	Lowest 4	8	17	42	170	Highest 340
1) Linear DNA (%)	41	14	14	15	12	10
2) One curved region (%)	26	20	24	24	26	24
3) Two curved regions (%)	20	40	29	26	18	21
4) One looped regions (%)	8	8	12	13	17	18
5) Two looped regions (%)	5	18	21	22	27	27

Table 1. Percentage of molecules from each species at different TubR:DNA ratios.
The tested ratios were 4, 8, 17, 42, 170 and 340. Percentages were rounded to a total 100 % from an analysis of >2000 molecules.

Contour length analysis on DNA-bent molecules revealed an enclosure of 70 nm, or 206 bp, of DNA in looped structures in average [Wing, Drew et al. (1980)] [Figure II-13 D]. However, the DNA molecules prepared were 352 bp. Accommodation of two looped structures could be hindered by the shortness of the DNA molecule (5-20% of species were double-looped). To understand the effect of the sequence and the DNA size on the formation of the looped structures we repeated the analysis with longer DNA molecules.

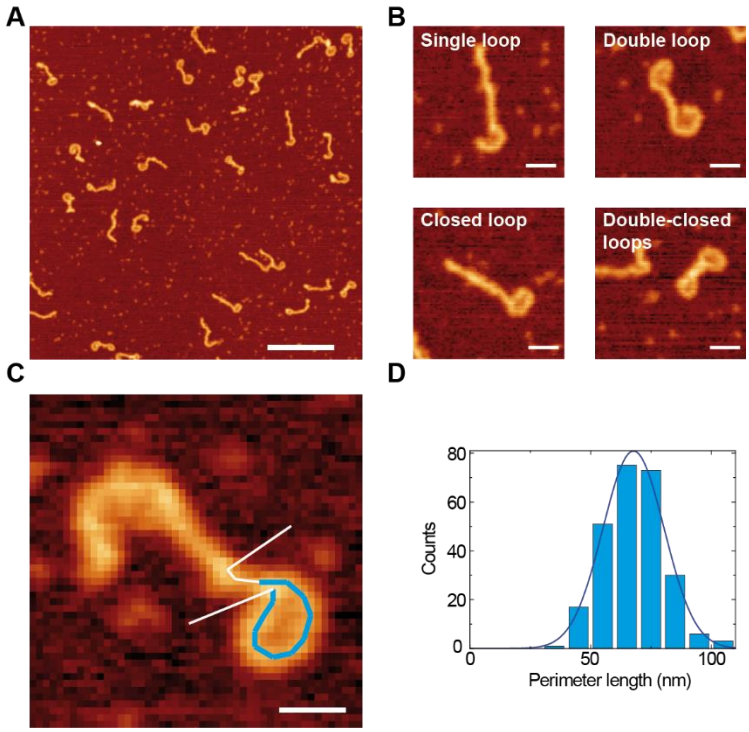


Figure II-13. AFM images and perimeter length analysis of the DNA secondary structures of TubR bound to tubC.

A) AFM image of a tubC sequence-containing DNA incubated with TubR. Scale bar is 200 nm. B) Secondary structures seen on AFM images divided into four classes bases on the number and shape of the structures: Class 1 with one open curved region, class 2 with two open curved regions, class 3 with one closed looped region and class 4 with two closed looped regions. Scale bars 30 nm. C) DNA-TubR molecule with a curved region and a looped region. Highlighted path signals the perimeter length of the looped region. Scale bar is 25 nm. D) Histogram of perimeter lengths for the 300 looped regions analysed.

II.4.2.3 The double-ring structure encloses the *tubC* region

To easily allocate two looped regions into a DNA substrate and correlate the position of these regions with the centromere sequence we produced four additional DNA molecules [Figure II-6]. Three of these molecules had around 900 bp and the other one had 2010 bp. A 866 bp (CPRZ) and 907 bp (NPDC) DNA molecules were constructed with the centromere sequence located at one of the DNA-ends; the difference was the presence (CPRZ) and absence (NPDC) of the TubR promoter region and the location of the *tubC* region at the 5' end (CPRZ) or 3' end (NPDC). A 977 bp molecule was created with the centromere sequence located on the middle region of the DNA molecule (RCPR). Lastly, a 2010 bp region of the complete c-st phage was linearized with the centromere sequence at the 5' end followed by the TubR promoter (C-ST) [Sakaguchi, et al. (2005)].

Surprisingly, the population of the five molecular species presented in Table 1 was similar among the new molecules [Table 2].

	tubC	**CPRZ**	**NPDC**	**RCPR**	**c-st**
Linear DNA (%)	12	11	9	10	10
One-Two curved regions (%)	44	40	38	38	37
One-Two looped regions (%)	44	49	53	52	53

Table 2. Population of species across different DNA substrates at 170 TubR:DNA ratio.
DNA substrates are in the following order: 352 bp DNA from section II.4.2.2; 866 bp or CPRZ DNA; 907 bp or NPDC DNA; 907 bp ot NPDC DNA; 977 bp or RCPR DNA and 2010 bp or C-ST DNA. N>1000 of molecules analysed for each DNA substrate at 170 TubR:DNA ratio.

The low percentage of linear DNA molecules across the DNA molecules tested and the small variation on the population of the secondary structured species is thought to be representative of final reaction products. Moreover, Atomic Force Microscopy images showed similar population percentages along molecules with one or two looped structures for the five employed DNA molecules [Figure II-14]. Furthermore, we incubated the CPRZ DNA molecule with TubR at 1700 and 3400 TubR:DNA ratios. At this high concentration of TubR the species had a ~10-40-50 % population divided into linear, curved and looped species respectively. The distribution is similar to that at lower ratios [Table 2].

These secondary structures were found at/or near the *tubC* sequence. In four of the five employed DNA molecules the DNA-bent structures where found at a DNA-end while for the RCPR DNA molecule the structures were found in the middle, in the case of the RCPR DNA the sequence *tubC* is located in the middle section of the DNA molecule contrary to the rest of the DNA molecules tested [Figure II-14 C]. This change of position of the centromere sequence did not seem to affect the population distribution [Table 2].

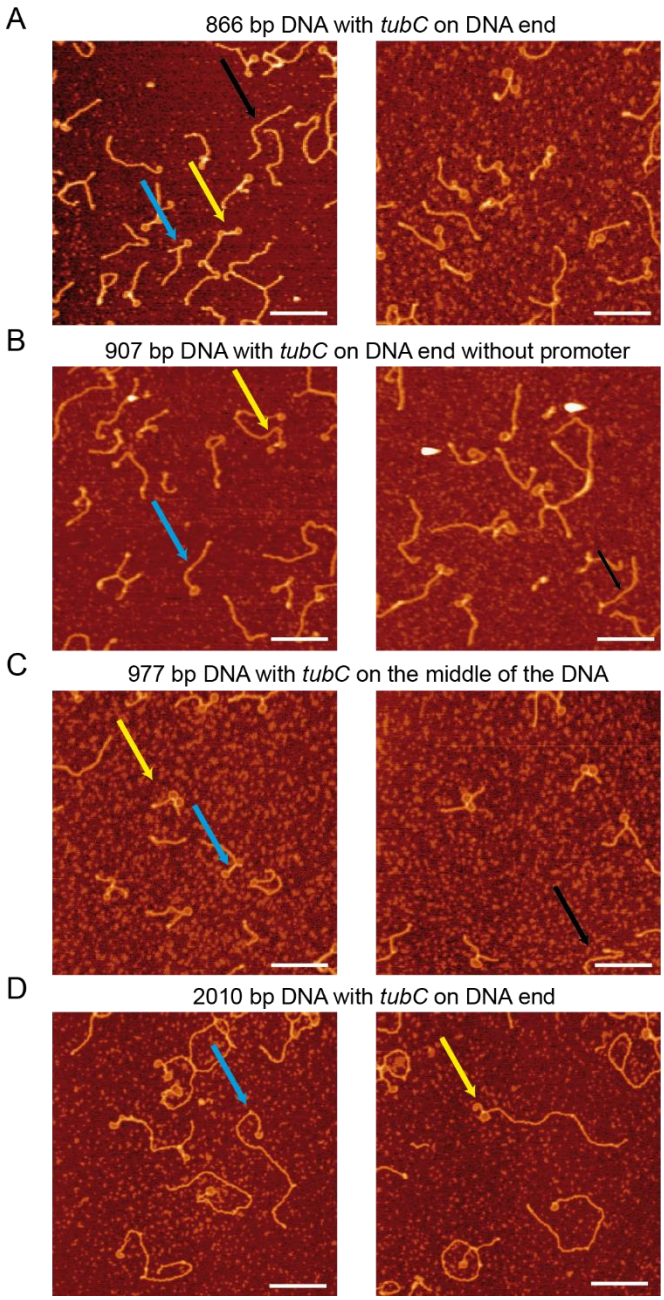


Figure II-14. AFM images of curved and looped regions on long DNA substrates.
A) AFM images of CPRZ DNA incubated with TubR. B) AFM images of NPDC DNA incubated with TubR. C) AFM images of RCPR DNA incubated with TubR. D) AFM images of c-st DNA incubated with TubR. Arrows indicate the position of linear (black), curved (blue) and looped (yellow) DNAs. Scale bar is 200 nm.

To recognize a possible auto-regulatory role of TubR by invasion of the adjacent promoter DNA region, the position of each looped structure was measured in double-looped molecules from the most distant DNA-end together with the distance between loops centres. For comparison, the DNA molecules analysed were the CPRZ and NPDC molecules. The CPRZ molecule contains the promoter region and the centromere sequence is located at the 5' DNA-end, while the NPDC molecule does not contain the promoter region and the centromere region is located solely at the 3'-end. In the CPRZ molecule, the loops were detected at 180 and 270 nm from the distant DNA-end [Figure II-6; Figure II-15 B]. In the NPDC molecule the loops were detected at 196 and 262 nm from the distant end [Figure II-15 D]. The distance between loops was bigger (90 nm) in the CPRZ double looped molecules than (66 nm) in the NPDC ones. Furthermore, the CPRZ molecules had one looped region possibly enclosing the part of the TubR promoter located at ~159 nm from the distant end of the *tubC* region. To verify the Atomic Force Microscopy results, further biochemical analysis were performed which confirmed the binding of TubR proteins to the promoter region [Martin-Garcia, et al. (2018)]. In addition, Electron Microscopy images of double-ring structures were obtained and analysed. The loops found in Electron Microscopy images have a loop diameter of 21.4±2.8 nm. The value is similar to the 68±25 nm perimeter length (or 21±8 nm in diameter for a perfect circle) values obtained by Atomic Force Microscopy [Martin-Garcia, et al. (2018)] [Figure II-13 D].

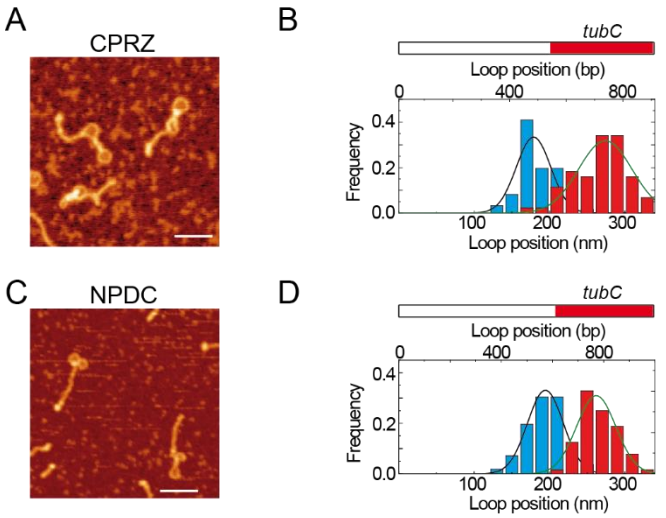


Figure II-15 Histograms of loop position for double-looped molecules.

A) AFM image of distant loops on CPRZ. Scale bar is 100 nm. B) Histogram of loop position for double-looped CPRZ molecules. Blue and red histograms correspond to the position of each of the two looped regions The white and red bar represents the position of tubC. C) AFM image of distant and adjacent loops on NPDC. Scale bar is 100 nm. D) Histogram of loop position for double-looped NPDC molecules. The white and red bar represents the position of tubC.

II.4.2.4 TubZ recognizes TubR-DNA looped structures

In type III partition systems from other organisms, the motor-protein, TubZ, is described to interact with the centromere-binding protein, TubR, by the C-terminal domain of both molecules [Ni, et al. (2010)]. It is described that the segrosome complex is key to the formation of a stable TubR-TubZ interaction [Aylett and Lowe (2012); Oliva, et al. (2012); Fink and Lowe (2015)]. Therefore, the role of the segrosome complex was expected to also be important for *C. botulinum* proteins. Image acquisition of TubZ filaments together with the looped structures of TubR-DNA complexes was challenging as the TubZ filaments depolymerised easily; thus, covering the mica with an excess of TubZ protein. The remaining filaments were two-to-three fold wider and higher than the TubR-DNA complexes [Figure II-12; Figure II-14]. To overcome this problem, both structures were preformed and, later, incubated together in a reaction tube for 5 minutes. The resulting reaction was deposited onto a mica surface for 5 minutes and then, washed and dried. With this method, the depolymerised filaments were detached from the mica surface only leaving the strongly TubZ-TubR-DNA interacting complexes. Equally, this deposition method was tested with TubZ-DNA samples to discard a possible effect of DNA molecules shielding TubZ molecules and leaving protein molecules attached to the mica surface only around DNA molecules. Atomic Force Microscopy images revealed an accumulation of protein molecules around DNA molecules only in the case of samples with the three partition elements [Figure II-16 B]. Additionally, to correlate the position of the protein molecules with the position of DNA molecules, Atomic Force Microscopy images were treated as matrixes. Chosen detailed Atomic Force Microscopy images of complex molecules with a size of 240x240 nm^2 were converted into matrixes of height values per coordinates X and Y. The matrices were analysed following the method presented in section II.3.3.6. The correlation coefficients obtained between “DNA” and “protein” matrices represent a possible protein-DNA association value, the higher the coefficient the higher the association. As expected, the results obtained were in agreement with the accumulation of protein in the Atomic Force Microscopy images of TubZ-TubR-DNA samples [Figure II-16 G]. But, the low coefficient for TubR-DNA samples was surprising. The DNA molecules are clearly modified by the presence of TubR in a sample. This suggested that a direct DNA-protein interaction as the one in TubR-DNA samples was not detected with this method as the protein is masked around the DNA molecule. For the TubZ-TubR-DNA sample, the protein TubZ surrounded the nearby DNA molecules not getting into direct contact with them.

The data obtained on the position of TubZ protein molecules together with the sizes of TubZ filaments (16 nm width) and TubR-DNA loops (21 nm diameter) were essential for the elaboration of a model for the type III partitions systems of *Clostridium botulinum* [Fuentes-Perez, et al. (2017); Martin-Garcia, et al. (2018)].

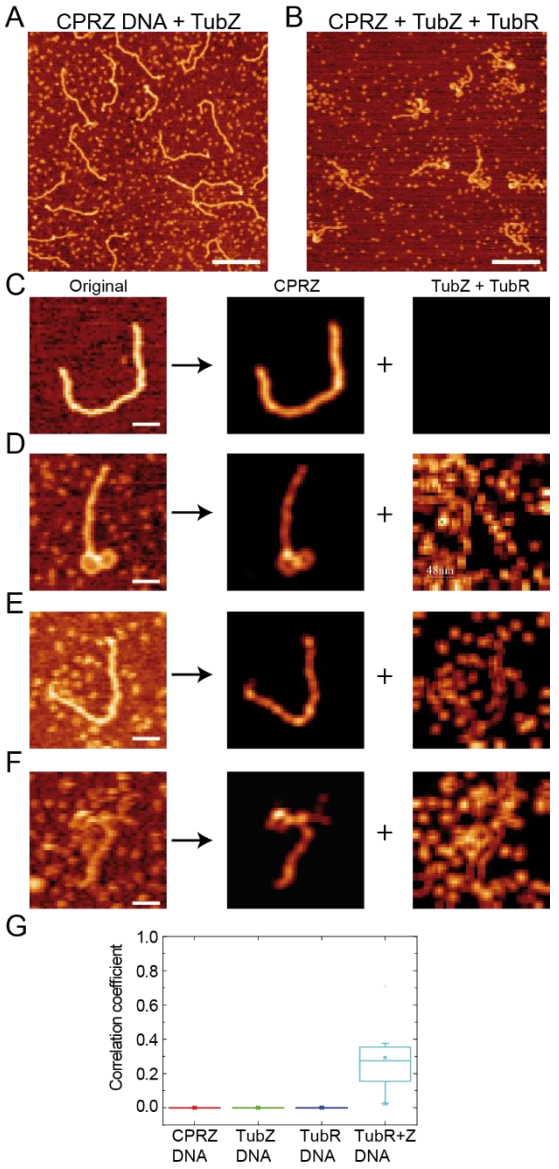


Figure II-16. AFM images, matrixes and correlation coefficients of TubZ, TubR and DNA samples.

A) AFM image of DNA incubated with TubZ protein. B) AFM image of DNA incubated with TubR and TubZ proteins. Scale bars are 200 nm. C) Zoomed image of a DNA molecule and the two matrixes resulting from height selection for DNA (left) and protein (right). D) Zoomed image of a TubR-DNA molecule and the two resulting matrixes. E) Zoomed image of a TubZ-DNA molecule and the two resulting matrixes. F) Zoomed image of a TubZ-TubR-DNA molecule and the two resulting matrixes. Scale bars are 48 nm. G) Correlation coefficient graph for the 4 cases analysed with N=43.

II.5 Discussion

In this chapter of the thesis we presented the studies about the centromere-binding protein, ParB, from the plasmid partitioning systems of *Bacillus subtilis* and *Caulobacter crescentus* together with the studies of the motor protein, TubZ, and Centromere-Binding protein, TubR, of *Clostridium botulinum*. In the case of *B. subtilis,* the ParB protein was thoroughly characterized in previous studies unsing MT but the structural characterisation of the segrosome complex using AFM was missing [Taylor, et al. (2015); Fisher, et al. (2017); Madariaga-Marcos, et al. (2018); Madariaga-Marcos, et al. (2019)]. Our AFM images of condensed DNA molecules increase the knowdledge on condensation effect of ParB by revealing the structure of the DNA-ParB condensates with the images showed in Figure II-9.

The ParB protein of the *C. crescentus* partition system was characterized for being a non-selective *parS* sequence recognition protein as was *B. subtilis* ParB [Taylor, et al. (2015); Tran, et al. (2018)]. To complement these studies and understand the role of ParB in *C. crescentus* organisms, we have studied the interaction complexes formed between ParB and DNA. The Atomic Force Microscopy images revealed an interaction of ParB with DNA, preferably with *parS* sequences, but not exclusive. The incubation of DNA with ParB formed different secondary structures on DNA molecules. Surprisingly, the protein was not capable of condensing DNA at the concentrations and conditions tested. The different molecular species showed in Figure II-10 had secondary structures on DNA molecules contrary to the condensed DNA showed in Figure II-9 for *B. subtilis*. This suggested a subtler effect onto DNA molecules mainly centred on the *parS* sequence. To supplement the data acquired by Atomic Force Microscopy, the system was studied with Magnetic Tweezers together with a mutant protein more prone to DNA condensation. In the future the main objective of this project will be to elaborate a reliable model for the role of ParB in this system and to understand the differences between type Ia partition systems such as the ones observed between *C. crescentus* and *B. subtilis* organisms.

In type III partition systems, the formation of TubZ filaments is known to rely on the generation of a helical protofilament and a complete filament composed of 2 to 4 protofilaments wrapped around each other in a helical shape [Aylett, et al. (2010)]. In Fuentes-Perez, et al. (2017) we have studied the role of the C-terminal domain of *C. botulinum* TubZ and the difference between GTP and GDP formed filaments. Altogether, the research project is completed with the proposed model on the role of the TubZ and TubR proteins during the partitioning process of the c-st plasmid [Figure II-17; Figure II-18]. The model gives an explanation on the polymerisation and stiffening processes of TubZ filaments based on the hydrolysis of GTP by the interaction between TubZ monomers. In this model, the

filaments are assembled by GTP binding, similar to tubulin protein [Buey, et al. (2006)]. Then, the interaction between TubZ monomers occurs on a rotated angle between monomers, inadequate for the hydrolysis fo GTP. Depending on the length of the C-terminal tail of TubZ monomers, the addition of one or several molecules is necessary for the tail-ends to make a proper contact. Contact between monomers is followed by the stabilization of the entire the C-tail via the charged residues [Figure II-17 A]. Once the polymer is correctly positioned and the catalytic loop reaches the terminal gamma phosphate domain, GTP hydrolysis takes place [Figure II-17 B]. The hydrolysis triggers the opening of the twist in the longitudinal subunit-subunit interface and rearranges the C-tail binding interactions with the surrounding molecules [Montabana and Agard (2014)] [Figure II-17 C]. The rearrangement and presence of GTP at the caps stabilizes the polymeric filament without stiffening it. However, the loss of the caps triggers a conformational change responsible for the weakening of the subunit-subunit interfaces which, in turn, leads to depolymerisation [Figure II-17 D]. The model is in general agreement with previously proposed treadmilling and dynamic instability models proposed for TubZ filaments of other organism.

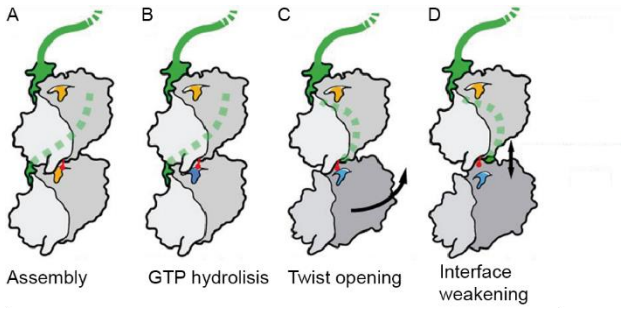


Figure II-17. Proposed model for TubZ filament formation and destabilization.

A) GTP-bound state induced the assembly, whereas the C-tail (green, and dashed line when behind a protein subunit) established specific contacts with upper subunits leading to the generation of a twist and allowing the formation of the canonical longitudinal interface. B) The catalytic residue (red) was correctly positioned and reached GTP (yellow), which was hydrolysed into GDP (blue). C) The enzymatic reaction induced changes in the bottom molecule, changing the C-tail interacting surface with the upper subunits and opening the twist between both molecules. D) The GDP-bound state resulted in weakening of the longitudinal interface, likely due to loss of contacts. C-tail truncated proteins did not establish the specific contacts with upper subunits within the filament, blocking the formation of the canonical longitudinal interface. Accordingly, GTP assembly leaded to state C filaments, while GDP-bound polymers were in state D. Model adapted from Fuentes-Perez, et al. (2017).

Differently to the other type Ia partition systems, the Centromere-Binding protein of type III systems was found to reliably bind to centromere DNA and form complex molecules by spreading rather than bridging [Aylett and Lowe (2012); Fisher, et al. (2017)]. Atomic Force Microscopy images showed the formation of

various secondary structures around the centromere region on DNA molecules of different sizes and sequences [Figure II-13, Figure II-14]. The sizes and distances between these bent-structures were in agreement with an auto-regulatory role of TubR proteins by blocking their promoter region while reliably detecting the position of the *tubC* region. In Martin-Garcia, et al. (2018), we proposed that these double-looped structures might be the typical segrosome complex architecture in Type III partition systems. This theory was supported by previous studies that described other centromere regions split into two boxes [Davis and Austin (1988)]. Together with additional data from Magnetic Tweezers, Electron Microscopy and other biochemical assays a model was suggested for the formation of the TubR-DNA segrosome complex and its role together with TubZ for plasmid partitioning in *C. botulinum*.

From footprinting assays presented in Martin-Garcia, et al. (2018); the recognition of a primary *tubC* sequence, "TTGAC", by TubR catalyses the binding of TubR to DNA. A secondary sequence, "$^{A}/_{T}$GAA", facilitates the spreading of TubR through the DNA molecule. As we have discovered the formation of the segrosome complex was only observed in DNA molecules containing portions of the *tubC* region [Figure II-7]. Atomic Force Microscopy together with Electron Microscopy images revealed a split segrosome structure into two curved shapes located inside the *tubC* region. The structures are often finished as loops or rings. Interestingly, the TubZ protein was capable of recognizing and accumulating around these regions. These results suggest the formation of a helical arrangement of TubR around DNA responsible for the bending of the DNA molecules. Lastly, the close correlation between the size of the looped structures (21 nm diameter) and the TubZ filaments (16 nm width) could complement the theory of the TubZ filament binding to the inner layer of the looped structures, similar to those described for type II partition systems [Gayathri, et al. (2012)].

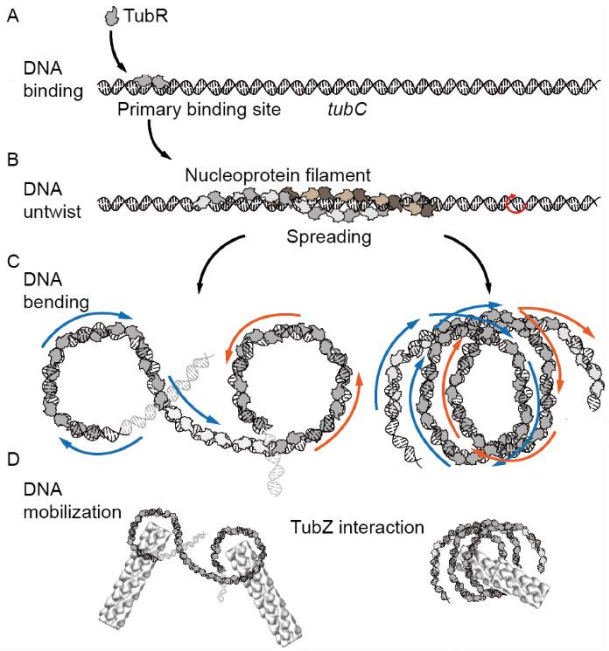


Figure II-18. Proposed wrapping and bending model for C. botulinum partition system. *A) TubR bound around DNA by wrapping one or alternatively two sides of tubC. B) TubR bending induced certain DNA twisting (red arrow). C) DNA wrapped by TubR bends at the stronger (red arrows) and weaker (blue arrows) binding sites generating two looped structures. D) Each ring could support one TubZ filament (left) for a synchronized DNA mobilization. Alternatively, a single TubZ filament could bind to the inter-ring region (right) favouring a simpler mechanism for DNA movement.*

III. HOMOLOGOUS RECOMBINATION

The chromosomic material of a cell contains the genetic information needed for its growth and development. It is a tightly-packed long DNA molecule. However, a DNA molecule is fragile. The different cellular organisms have developed mechanisms for DNA protection and repair. One of such repair mechanisms is known as Homologous Recombination.

"Biology is the study of complicated things that have the appearance of having been designed with a purpose" – Richard Dawkins, Professor for Public Understanding of Science at the University of Oxford.

One of the many concerns for space agencies and astronauts is the prolonged exposure to sunlight and cosmic rays in an out-of-atmosphere environment. The cellular organisms that evolved in the Earth have developed mechanisms for DNA protection in the conditions of an atmospheric environment. However, when these organisms are exposed to harsh conditions, such as the ones in space, their protection mechanisms might not be sufficient. In this chapter, we present the cellular mechanism known as Homologous Recombination and its importance in the DNA-repair processes of eukariotyc and prokariotic organisms.

III.1 Introduction

The Homologous Recombination process was firstly described in 1911 by scientists Thomas Hunt Morgan and William Bateson when studying the inheritance of correlated genes [Morgan (1911); Bateson (1907)]. Morgan, citing Bateson's work, realized that traits associated to the same gene where not inherited together. He proposed a "splitting" process during the strepsinema stage. The strepsinema stage refers to the cellular phase at which the chromosomes are paired in an X shape. This "splitting" process was more likely to affect traits located far from one another inside a gene. To this process he proposed that "rather than a random segregation of genes [Mendel (1865)] it exist an association of factors in genes", referred as a "recombination of genes". The "splitting" process was theorized by Morgan. It was not until 1931 that it was associated to the phases of the cell division in eukaryotes and thus to the inheritance in *Homo sapiens* [Creighton and McClintock (1931)] [Figure III-1].

As a "recombination of genes", Homologous Recombination refers to the specific interchange of genes between two identical chromosomes. Normally, cells maintain only one chromosomal copy to reduce occupied space [Schafer (1998)]. This limits the mechanism operational phases. To be operative, Homologous Recombination only takes place in the cellular cycle phases associated to abundance of chromosomal copies [Shibata and Jeggo (2014)]. These phases are the synthesis phase (S phase) and the second growth phase (G2 phase). The use of an identical copy allows to control the spreading of non-desired mutations. It is thought that Homologous Recombination is a less error-prone mechanism compared to other recombination processes [Guirouilh-Barbat, et al. (2014)]. The suppression of factors involved in the initiation of Homologous Recombination is associated with an increase in error-prone repair incidence [Tutt, et al. (2001)].

During S phase the cell synthesizes new DNA molecules, therefore creating additional copies of the chromosomes. The finalisation of the S phase leads to the second growth phase [Bertoli, et al. (2013)]. The synthesis process of a DNA molecule is one of the sources of degradation by the manipulation of the molecule

by the synthesis-related protein or by the exposure of the open DNA molecule to external agents [Aguilera and Gomez-Gonzalez (2008); Argueso, et al. (2008)]. The degradation caused in these steps may derive into breakage of the DNA double-helix in both strands. These lesions are known as Double-Strand-Breaks which are one of the most dangerous damages a DNA molecule can suffer [Jackson (2002)]. There are two main cellular mechanisms to repair the Double-Strand Breaks on DNA molecules: Non-Homologous End Joining and Homologous Recombination. Non-Homologous End Joining is capable of repairing the damages by copying a piece of similar DNA in any of the cell cycle phases [Takata, et al. (1998)]. Homologous Recombination, as stated, needs an exact copy and is a limited process.

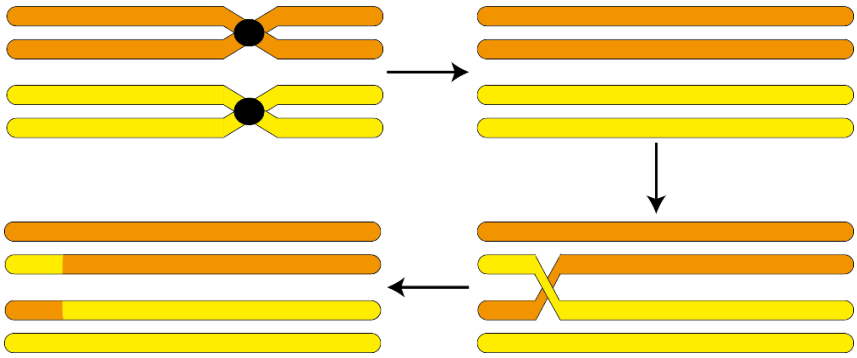


Figure III-1 Recombination of genes produced by Homologous Recombination during meiosis.

Graphical representation of the splitting process of two chromosomes, orange and yellow.

The regulation of both processes has been correlated to the presence of cycle-dependent kinase proteins. These proteins are responsible for the activation of other proteins by addition of phosphate groups to their structure in a cycle-dependent modulation [Shrivastav, et al. (2008)]. In the case of recombination mechanisms, phosphorylation of the proteins responsible for the first steps is thought to be crucial [Tougan, et al. (2010); Luo, et al. (2016)]. Some of these proteins take part in the detection and starting processes of the repair mechanism of Double-Strand Breaks. The activated proteins determine the proper break-repair mechanism for the DNA lesion.

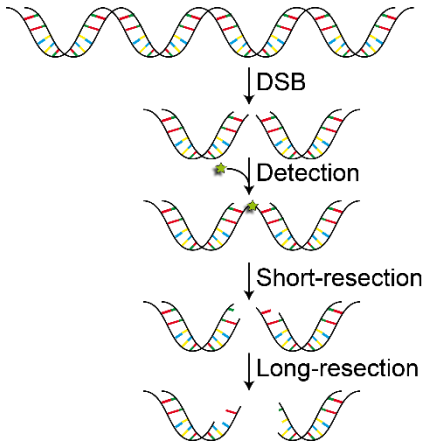


Figure III-2 Homologous Recombination regulation steps.
Schematic representation of the resection process during the signalling of DSBs divided into the 3 main steps: DNA-damage detection, short-resection and long-resection.

Previous studies have shown that, the first step for Homologous Recombination DNA-damage repair process consists in the resection, or degradation, of the damaged DNA [Cejka (2015)]. This step is usually further divided into three reactions: 1) detection and signalling of the Double-Strand Breaks position, 2) short resection of the damaged area, and 3) long resection of the damaged-DNA [Figure III-2]. This chapter of the thesis focuses on the study of two proteins involved on these last two steps: *Homo sapiens* CtIP that is a key protein in the short-resection reaction [Sartori, et al. (2007)], and *Geobacillus stearothermophilus* Dna2 that is a nuclease-helicase associated to the long-resection reaction [Paudyal, et al. (2017)].

III.1.1 *Homo sapiens* CtIP

The name of the protein CtIP comes from C-termini Binding and Interacting Protein [Schaeper, et al. (1998)]. Analogous proteins in other organisms for Homologous Recombination are Sae2 in budding yeast and Ctp1 in fission yeast [Forment, et al. (2015)]. Previous studies have shown an overall tetrameric structure in solution for CtIP [Davies, et al. (2015)]. The protein is frequently associated with the MRN protein complex, Mre11-Rad50-Nbs1. The whole complex is involved in Double-Strand Break recognition and binding [Ramirez-Lugo, et al. (2011); Williams, et al. (2009)]. Although the structure of CtIP has not been crystalized several studies have proposed that CtIP has a globular C-terminal domain and a coiled coil N-terminal domain. The N-terminal domain is recognized for being responsible of tetramer formation in a dimer of dimers fashion [Andres, et al. (2015)]. The central region of CtIP is a disordered domain

comprising aminoacids 145-769 [Figure III-3]. Mutations in the CtIP gene in *Homo sapiens* cause Jawad syndrome and Seckel2 syndrome [Qvist, et al. (2011)]. These syndromes are related to accelerated aging and mental disability. Other point-mutations of importance are the L27, located at the N-terminal, and R839, located at the C-terminal. [Davies, et al. (2015); Ferretti, et al. (2016)]. These mutations cause lack of formation of tetrameric structures (L27E) and disruption at the DNA recognition domain (R839A).

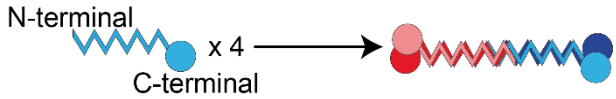


Figure III-3 CtIP monomer and tetramer.

Schematic representation of the CtIP monomer structure (left) and the overall resulting tetrameric structure (right). The circular domains represent the C-termini while the zig-zag sections represent the disordered middle and coiled-coil N-termini of the CtIP monomer.

The function of CtIP during Homologous Recombination is not completely understood. It is thought to play a key-role in directing the DNA-damage repair mechanism of choice to Homologous Recombination [Yun and Hiom (2009)]. It also upregulates the nuclease activity of the Mre11 subunit through the interaction with the MRN complex [Cannavo and Cejka (2014); Paull and Gellert (1999)]. Some authors propose that CtIP has endonuclease activity by itself [Lengsfeld, et al. (2007); Makharashvili, et al. (2014)]. However, the primary structure of CtIP does not show any resemblance to known nucleases. Despite the possible nuclease activity, the role of CtIP remains closely related to the MRN complex and the signalization of damaged DNA-ends for short-resection [Andres and Williams (2017); Aparicio, et al. (2016); Aparicio, et al. (2016); Quennet, et al. (2011); Paudyal, et al. (2017)]. Regardless of the scarce knowledge about CtIP and the interaction between CtIP and DNA, some studies have proposed models on the role and structure of CtIP [Forment, Jackson et al. (2015)] [Figure III-4].

Later on, successful short-resection of damaged DNA-ends leads to longer resection on the 5'-3' direction. This long-resection process is carried out by a group of specialized helicases and nucleases. One of these helicase-nucleases is DNA2 in *Homo sapiens*. The project of HR is complemented with preliminar results of a DNA2-like helicase/nuclease protein from *Geobacillus stearothermophilus*.

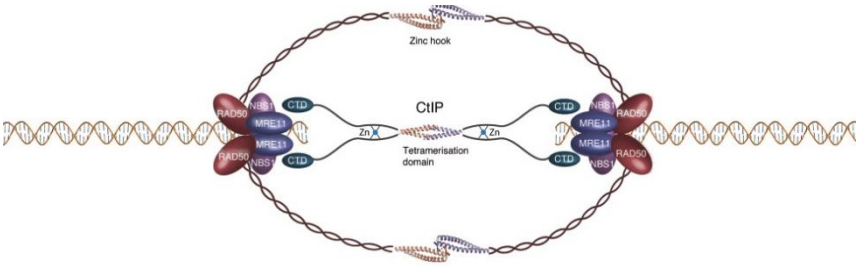


Figure III-4. Model for DNA-bridging by CtIP and the MRN complex at a DSB site.
Dimer of dimers model of CtIP structure. CtIP C-terminal domains bind to broken DNA ends in cooperation with the MRN complex. Coiled-coil domains are represented as interacting helical ribbons for CtIP and the MRN complex. Adapted from Forment, Jackson et al. (2015).

III.1.2 *Geobacillus stearothermophilus* DNA2-like protein

DNA2 refers to DNA Replication Helicase/Nuclease 2. It is an enzyme conserved in various organisms with the DNA damage-repair mechanism of Homologous Recombination, among them *Homo sapiens* [Duxin, et al. (2009)]. DNA2 is involved in the long-resection reaction of the Homologous Recombiantion mechanism [Daley, et al. (2017)] and its role is believed to be regulated by a DNA2-CtIP interplay. The structure has been discovered for various systems and the role of the several domains has been described for the helicase (opening of the DNA double-helix) and nuclease activities (cleaving of the DNA phosphate backbone) [Zhou, et al. (2015)]. Both functions may be associated to the long-resection of DNA broken molecules in a 5' to 3' direction [Pokharel and Campbell (2012)]. However, previous studies have proposed different mechanism of action for this protein: some articles suggest that DNA2 has nuclease activity but it is intrinsically devoid of helicase function, whereas, others suggested the collaboration among MRN, CtIP and DNA2 is responsible for the resection of clean DNA-ends [Kim and others (2006); Paudyal, et al. (2017)]. Mutations in *Homo sapiens* DNA2 are linked to progressive external ophthalmoplegia-6 and Seckel syndrome 8 [Ronchi, et al. (2013); Shaheen, et al. (2014)]. Other described mutations disable the nuclease and helicase activities located on the N-terminal and C-terminal respectively [Zhou, et al. (2015)].

Overall, the role of the DNA2 5'-3' nuclease activity during the resection process of the HR mechanism is poorly understood [Paudyal, et al. (2017)]. Some evidence of interplay between the nuclease and helicase domains suggests a key role of both domains [Pokharel and Campbell (2012)]. What it is known is that the creation of long 5' single-strand DNA overhangs by DNA2 signals the finalization of the initiation step of Homologous Recombination [Liao, et al. (2008)]. As the initiation step is finished by the action of DNA2, the DNA molecule is prepared for

the strand invasion of the damaged DNA by the homologous non-damage DNA molecule.

In this chapter we describe the results obtained with a bacterial helicase/nuclease (Bad) with domain homology to the eukaryotic DNA repair factor DNA2. Bad contains an Iron-Sulfur cluster and possesses 5'>3´helicase and single-stranded DNA translocase activity but is not involved in the HR mechanism as *G. stearothermophilus* does not employ this damage-repair process.

III.2 Objectives

The aim of this project is to study and understand the role of the protein CtIP. This protein takes part on the Homologous Recombination mechanism in *Homo sapiens.* To complement this study, a DNA2-like helicase/nuclease protein, Bad, from *G. stearothermophilus* is also visualized under Atomic Force Microscopy.

The objectives set were:

4) To study the protein conformation of *Homo sapiens* CtIP.
5) To understand the role of CtIP, without the MRN complex, in contact with DNA-ends related to Homologous Recombination.
6) To study the binding mechanism of the protein Bad in a reaction with a double-stranded DNA molecule similar to the DNA abducts created during Homologous Recombination.

III.3 Materials and Methods

III.3.1 CtIP project

III.3.1.1 CtIP protein expression and purification

The protein CtIP was purified and subsequently biochemically analysed by Dr. Oliver Wilkinson from Prof. Mark S. Dillingham's laboratory (University of Bristol, UK) following the "Material and Methods" corresponding section in Wilkinson, et al. (2019) [Figure III-5].

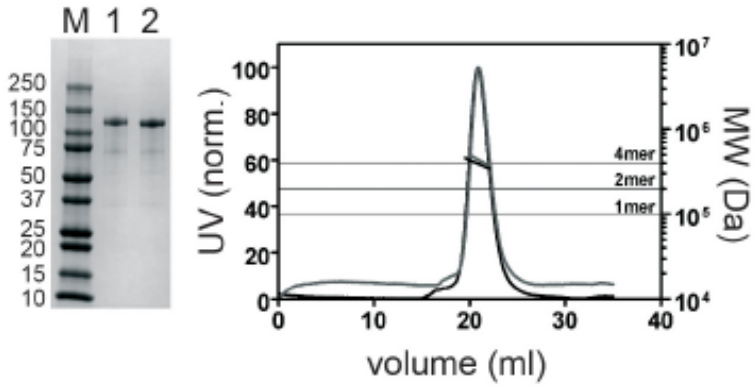


Figure III-5 WT CtIP oligomeric state analysis by SEC-MALS and SDS-PAGE.
SDS-PAGE and SEC-MALS analysis of wild type CtIP as prepared (lane 1, black trace) and following post-purification dephosphorylation (lane 2, grey trace). Horizontal lines on the SEC-MALS graph show the expected molecular weights for monomeric, dimeric and tetrameric CtIP species. Figure adapted from Wilkinson, et al. (2019).

Preparation of mutant (R839A, R100W, L27E, Jawad and Seckel) or dephosphorylated variants of the protein were performed by members of Prof. Mark S. Dillingham's laboratory.

III.3.1.2 DNA substrate fabrication and purification

DNA molecules chosen for the analysis of CtIP were prepared by Dr. Oliver Wilkinson from Prof. Mark S. Dillingham's laboratory (University of Bristol, UK) following the "Material and Methods" corresponding section in Wilkinson, et al. (2019).

III.3.1.3 Sample preparation and image acquisition for Atomic Force Microscopy

Protein sample preparation for air image acquisition

Samples prepared for the analysis of the CtIP protein conformations were incubated in 20 mM Tris-HCl pH 8.0, 200 mM NaCl and 1 mM DTT. The final CtIP concentration was 4 nM. The buffered solution containing CtIP was deposited onto a freshly cleaved mica surface. Then, to limit the interaction of CtIP with the mica surface the solution was washed after 30 seconds three times with 1 ml of Milli-Q water and, thereafter gently dried under a constant nitrogen air flow.

Protein sample preparation for liquid image acquisition

CtIP stock solution was diluted to 4 nM in 20 mM Tris-HCl pH 8.0 and 200 mM NaCl. A volume of 2 µl of the CtIP dilution was deposited onto a freshly cleaved mica and incubated for 15 minutes. Then, the solution was washed by exchanging the previous buffer with a low salt buffer; 20 mM Tris-HCl pH 8.0 and 30 mM KCl. The mica surface plus the protein solution was then moved onto a liquid cell with a buffered solution; 20 mM Tris-HCl pH 8.0 and 30 mM KCl.

DNA-protein sample preparation for air image acquisition

Samples for the study of the DNA-CtIP interaction were initially prepared in 100 mM HEPES pH 7.5 and 100 mM NaCl. Reactants were at a 10 nM concentration for DNA and $CtIP_4$ concentrations range from 0 to 250 nM. The titration was chosen for an in-depth analysis of CtIP concentration effect onto the DNA molecules. The resulting mixture volume, 2 µl, of DNA:$CtIP_4$ at ratios (1:0, 1:6, 1:13, 1:19 and 1:25) was incubated for 15 minutes. Simultaneously a freshly cleaved mica surface was pretreated with a solution of 7.5 mM $MgCl_2$ and washed three times with one millilitre of Milli-Q water. A volume of 18 µl of a low salt buffer with 4 mM HEPES pH 7.5 and 10 mM NaCl; was pipetted on the pretreated mica surface followed by 2 µl of the previously explained DNA-CtIP solution. This step decreases the final concentration of the sample and buffer components while allowing a high concentration incubation. This decreased the formation rate of protein or DNA aggregates due to low salt or low DNA/protein concentration [Fuentes-Perez, et al. (2012)].

The final deposited solutions were incubated for 30 seconds, washed three times with 1 ml of Milli-Q water and dried under constant flow of nitrogen air [Lyubchenko, et al. (2009)].

DNA-protein sample preparation for liquid image acquisition

CtIP stock solution was diluted to 4 nM in 20 mM Tris-HCl pH 8.0 and 200 mM NaCl while the DNA stock solution was diluted to 0.38 nM in 20 mM Tris-HCl pH 8.0, 30 mM KCl and 2 mM $MgCl_2$. One microliter of each solution was deposited onto a freshly cleaved mica to a total volume of 2 µl and incubated for 10 minutes. Then, the mixture was washed by exchanging the buffer three times with 20 mM Tris HCl pH 8.0 and 30 mM KCl. The washed sample was then moved onto a liquid cell with the previous stated buffer.

AFM image acquisition conditions

Images were obtained using a Nanotec Electronica S.L. Atomic Force Microscope using the amplitude modulation (*tapping*) mode unless otherwise stated. Air samples were imaged at room temperature and humidity at a rate of 1.48 lines per second and 512 lines per image and pixels per line, 346 second per image. Liquid samples were imaged at a higher rate, 4-540 lines per second, with a variable number of pixels per image, from 512x512 to 80x80 pixels. Cantilevers utilized were PPP-NCH (Nanosensors, Switzerland) for air image acquisition and BL-10-ACTS (Olympus, Japan) for liquid acquisition.

Image processing was performed following the method presented in chapter II.3.1.4.

III.3.2 Dna2 project

III.3.2.1 Identification, cloning and expression of Bad, a bacterial DNA2-like enzyme

A DNA2-like enzyme from *Geobacillus stearothermophilus* 10 (DSM accession number 13240, Taxonomy ID: 272567) was cloned by PCR from genomic DNA by Dr. Oliver Wilkinson from Prof. Mark S. Dillingham laboratory. The untagged gene was ligated into the pET28a vector (Novagen) for expression using the T7 promoter system. The entire sequence of the cloned gene, which has been annotated as a AAA+ ATPase, was identical to that reported in the G. stearothermophilus 10 genome (accession number WP_053413574). The gene encodes a 1245 amino acid protein with a molecular weight of 143 kDa and a theoretical extinction co-efficient of 195960 $M^{-1}cm^{-1}$. We note that a small number of other polypeptide sequences have been reported in the proteomes of *Geobacillus stearothermophilus* and closely-related strains which are virtually identical to Bad but which feature an additional 25 amino acids at the N-terminus (e.g. accession number ADU93794.1). Given that sequences of this type are in a

significant minority, and that all the longer sequences contain a methionine residue at position 26, it is likely that the start codon has been mis-assigned in these instances. We have cloned and expressed an example of the longer Bad polypeptide but this was found to be insoluble (data not shown).

III.3.2.2 Purification of *Geobacillus stearothermophilus* Bad

Bad was purified by Dr. Oliver Wilkinson from Prof. Mark S. Dillignham laboratory. For overexpression in *E. coli*, pET28a-Bad was transformed into BL21(DE3) cells. The cells were grown in LB to mid-log phase before induction with IPTG (1 mM) for 16h at 27°C. All buffers were degassed extensively prior to use in the preparation to prevent oxidation of the iron-sulphur cluster. The pellet (~10g) from 4L of bacterial culture was resuspended in 30mL lysis buffer (50mM Tris-HCl pH8.3, 1mM EDTA, 150mM NaCl, 10% glycerol, 5mM DTT, 1mM PMSF, Roche protease inhibitor cocktail) and sonicated on ice for a total of 1 minute in 5 second pulses with rests. After centrifugation at ~50,000g for 30 min at 4°C, ammonium sulphate was added slowly with stirring to the cleared lysate to a final concentration of 50% (w/v) at 4°C. The precipitated protein was then recovered to a pellet by centrifugation at 50,000g for 30 min at 4°C. This pellet was then resuspended in Buffer A (20mM Tris-HCl pH8.0, 1mM EDTA, 5% glycerol, 0.1mM PMSF, 5mM DTT, Roche protease inhibitor cocktail) up to a volume where the conductivity of the solution was 16mSv, and then loaded at 2mL/min onto a 5mL Heparin column equilibrated in Buffer B (20mM Tris-HCl pH8.0, 1mM EDTA, 100mM NaCl, 5% glycerol, 0.1mM PMSF, 5mM DTT). After washing for 5CV with Buffer B, a gradient was run over 20CV into Buffer C (20mM Tris-HCl pH8.0, 1mM EDTA, 1M NaCl, 5% glycerol, 0.1mM PMSF, 5mM DTT). The BAD-containing fractions were pooled and diluted with Buffer A to give a final salt concentration of ~170mM NaCl before loading at 2mL/min onto a 1mL MonoQ column pre-equilibrated in Buffer D (20mM Tris-HCl pH8.0, 100mM NaCl, 5mM DTT). After washing with 5CV Buffer D, the protein was eluted by running a gradient from Buffer D to 50% Buffer E (20mM Tris-HCl pH8.0, 1M NaCl, 5mM DTT) over 20CV. The most concentrated BAD-containing fractions were pooled and 1.5mL was injected onto a pre-equilibrated Superdex200 16/600 column in Buffer F (20mM Tris-HCl pH8.0, 300mM NaCl, 5mM DTT). The pool of the peak fractions was spin concentrated (Millipore, 50kDa cutoff) to a final concentration of ~15µM.

III.3.2.3 DNA molecule fabrication and purification

DNA molecules chosen for the analysis of Bad interaction with DNA were prepared based on molecules presented in Levikova, et al. (2013) and Luzzietti, et al. (2011) by Dr. Clara Aicart-Ramos or myself from Prof. Fernando Moreno-Herrero's laboratory.

The desired DNA molecules were obtained by linearization of the pNLrep plasmid (kindly gifted by Prof. Dr. Ralf Seidel), 6808 bp, with the restriction enzymes *Bsr*GI and *Bam*HI (New England BioLabs, USA) to a final length of 6610 bp. Introduction of a dephosphorilation step before inactivation of the restriction enzymes was necessary to avoid a later religation of the DNA fragment with itself [Luzzietti, et al. (2011)]. The linear piece was then purified by gel extraction and the purified fragment was subsequently digested with Nt. *Bbv*CI, a nicking endonuclease that cleaves only one strand of DNA on the double-stranded DNA substrate sequence. The nicking affects only one of the two strands resulting in the formation of 15–16 bases long fragments that after denaturation create a 63 bp-gap of single-stranded DNA in which desired oligonucleotides can be hybridised. The denaturation step of the DNA molecule occurs at 37 ºC for 2 hours, leaving a 63 bp-gap of single-strand DNA. Later, a poly-T tail oligo partially anneals into the gap leaving 37 non-complementary thymine nucleotides by heating at 80 ºC for 5 min and slowly cooling down to 30 ºC, leaving a 5' overhang of a single-strand poly-T tail DNA at 445 bp from the proximal end. A final step of ligation with T4 DNA ligase (New England Biolabs, USA) was performed.

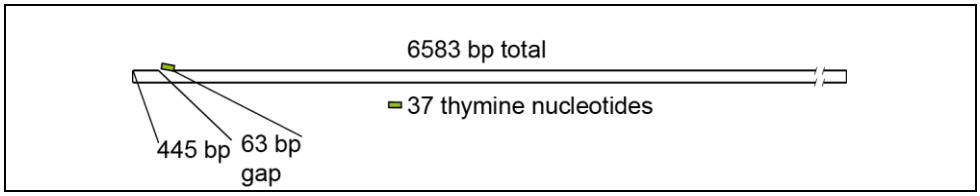


Figure III-6. DNA molecule employed in the Bad project.

III.3.2.4 Sample preparation and image acquisition for Atomic Force Microscopy

DNA-protein sample preparation for liquid image acquisition

Bad stock solution was diluted to a 2.5 nM concentration in 20 mM Tris-HCl pH 8.0, 300 mM NaCl and 1 mM DTT. The DNA stock solution was diluted to 0.76 nM in 20 mM Tris-HCl pH 8.0, 30 mM KCl and 2 mM $MgCl_2$. A volume of 20 µl composed of 10 µl of each solution was incubated with 4 mM ATP for 15 to 30 minutes. Then, 2 µl of the reaction were deposited onto a freshly cleaved mica for 10 minutes to stop the reaction. The final mixture was washed by exchanging the buffer three times with 20 mM Tris HCl pH 8.0, 30 mM KCl, 2 mM $MgCl_2$ and 4 mM ATP. The washed sample was then moved onto a liquid cell with the previous stated buffer.

AFM image acquisition conditions

Images were obtained using a home-built Atomic Force Microscope by Prof. Toshio Ando's group (Nano-Life Science Institute-Kanazawa University, Japan) using the amplitude modulation (*tapping*) mode. Sample images were acquired at a fast rate, 120-540 lines per second, with a variable number of pixels per image, from 150x150 to 80x80 pixels, and sizes of 400x400 nm^2 to 75x75 nm^2. Cantilevers employed were BL-10-ACTS or BL-7-ACTS (Olympus, Japan) for liquid image acquisition.

III.4 Results

III.4.1 Double-Strand Breaks short resection

III.4.1.1 *Homo sapiens* CtIP is a stable tetrameric protein

Biochemical analysis of purified wild type CtIP proteins by SEC-MALS revealed a single peak at an estimated weight of 391±23 kDa [Figure III-5]. This molecular weight is in agreement with a tetrameric protein state, as CtIP monomer has a predicted molecular weight of around 100 kDa [Qvist, Huertas et al. (2011)]. However, chromatographic step of CtIP protein purification showed an earlier than expected elution for a ~400 kDa protein [Wilkinson, et al. (2019)]. Elution of molecules on size-exclusion chromatographic columns depends on molecule weight and shape which, in this case, indicated an unusual shape for CtIP.

Atomic Force Microscopy images of purified CtIP molecules exhibited varying structures. Five species of molecules were classified based on their molecular shape [Figure III-7 A]. The class I is composed of molecules with a round shape. Class II is composed of more complex molecules with a round shape followed by a tail. Larger species were divided into classes III, IV and V based on the number of round shapes connected by a tail or linear structure. Class III has two round shapes interconnected. Class IV has three and class V has four or more connected round structures. A further volumetric analysis (explained in chapter II.3.3.7) was performed to confirm differences between classes [Figure III-7 C]. The largest molecular species, class V, had a volume corresponding to a ~380 kDa protein [Fuentes-Perez, et al. (2013)]. The weight is closely related to the estimated weight for CtIP tetrameric species. These molecules were similar to dumbbell-shaped particles revealed by Electron Microscopy or have a more open conformation [Figure III-7 B]. Classes I-II and classes III-IV were grouped together based on comparable volume distributions. These classes (I, II, III and IV) had lower volume distributions than class V. This is probably associated with lower oligomeric states produced by strong protein-surface interaction or by the necessary sample dilution for AFM image acquisition.

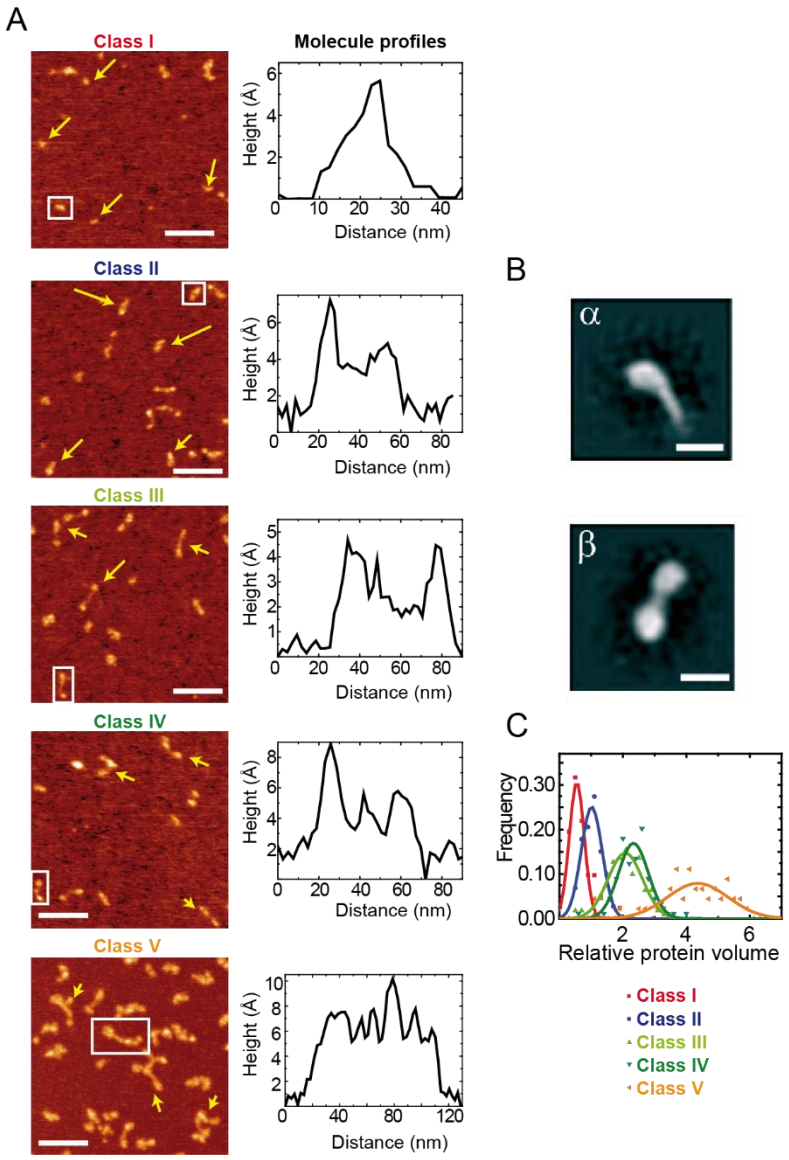


Figure III-7. Wild type CtIP conformation analysis by Atomic Force Microscopy and Electron Microscopy.

A) AFM images of purified WT CtIP divided into 5 classes. Images I-V show examples of proteins pointed by yellow arrows for each class with a representative profile of the molecule squared. Scale bar is 100 nm. B) 2D classification of WT CtIP by negative staining EM. Scale bar is 20 nm. C) Relative protein volume distribution for the 5 classes shown. A linear relation exists between relative volume and the protein molecular weight. A relative volume of 4.4 predicted a molecular weight of around 380 kDa [Fuentes-Perez, et al. (2013)]. Image adapted from Wilkinson, et al. (2019).

To supplement and verify the data obtained for the Wild type CtIP protein, additional analysis of two mutated CtIP variants were performed. A mutation on the aminoacid 839 (R839A), at the possible DNA-binding domain of CtIP, did retain the tetrameric state as revealed by SEC-MALS and was composed of the same class of species as revealed by Atomic Force Microscopy images [Andres, Appel et al. (2015)] [Figure III-8 A]. On the other hand, a mutation on the aminoacid 27 (L27E), at the tetramerisation domain [Davies, et al. (2015)], was predicted to have molecular weight around 270-190 kDa and revealed a different classification of species by Atomic Force Microscopy imaging [Figure III-8 B]. The molecular weight predicted by SEC-MALS and the subsequent disappearance of class V molecules in the L27E mutant suggested that the largest class of molecules was associated to the tetrameric state of the protein CtIP.

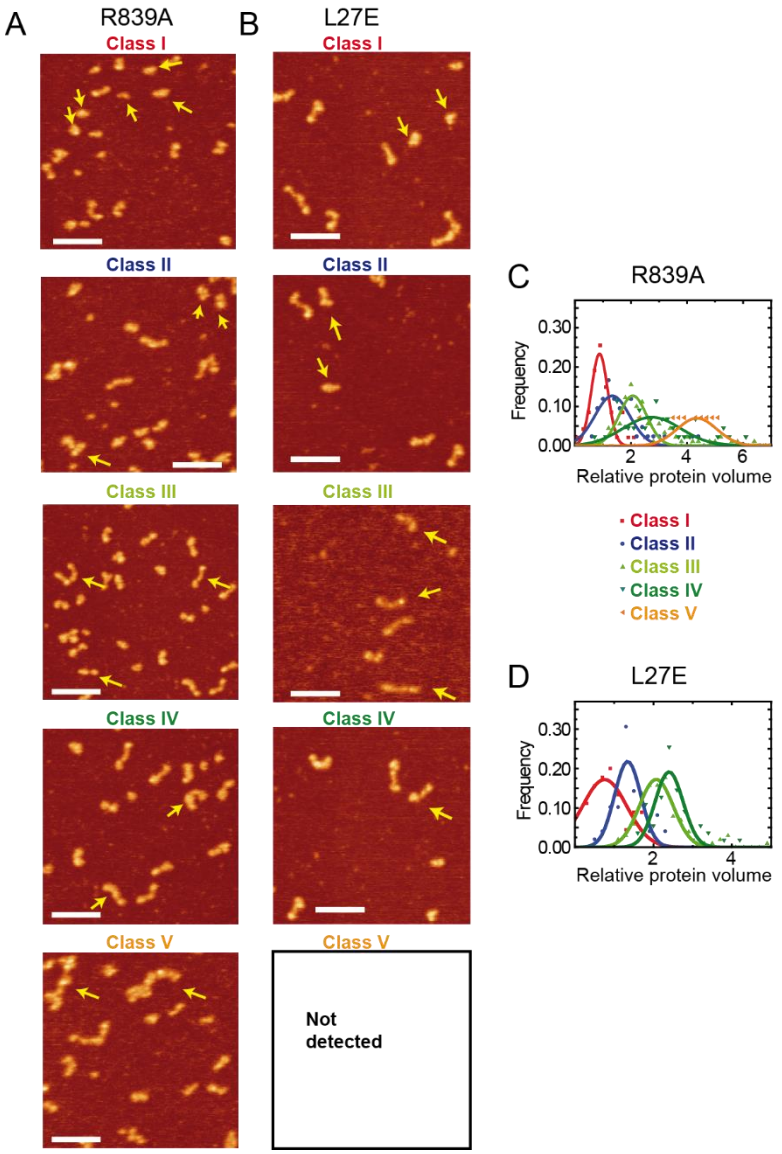


Figure III-8.Conformation analysis of R839A and L27E CtIP mutants by AFM.

A) AFM images of purified R839A CtIP divided into 5 classes. Images I-V show examples of proteins pointed by yellow arrows for each class. Scale bar is 100 nm. B) AFM images of purified L27E CtIP divided into 4 classes. Images I-IV show examples of proteins pointed by yellow arrows for each class. No molecules were detected with class V properties. Scale bar 100 nm. C) Relative protein volume distribution for the 5 classes shown for R839A CtIP. D) Relative protein volume distribution for the 4 classes shown for L27E CtIP. Image adapted from Wilkinson, et al. (2019).

III.4.1.2 CtIP is a flexible protein

Coiled-coil and disordered domains had been previously described as flexible domains by Atomic Force Microscopy [de Jager, et al. (2001)]. To complement the analysis on the oligomeric structure of the protein CtIP, flexibility-related analysis on the previously shown molecular classes was performed. In air-acquired images, protein molecules were anchored to a mica surface and, thus, dynamic behaviour was not present [Figure III-9 A]. However, contour length analysis of molecules from classes III and IV revealed an average of 30 nm length of the tailed domains with high variability between molecules [Figure III-9 B]. The high-variability of contour lengths is consistent with the presence of high flexibility coiled coil or disordered domains in the protein CtIP [Figure III-9 C-D].

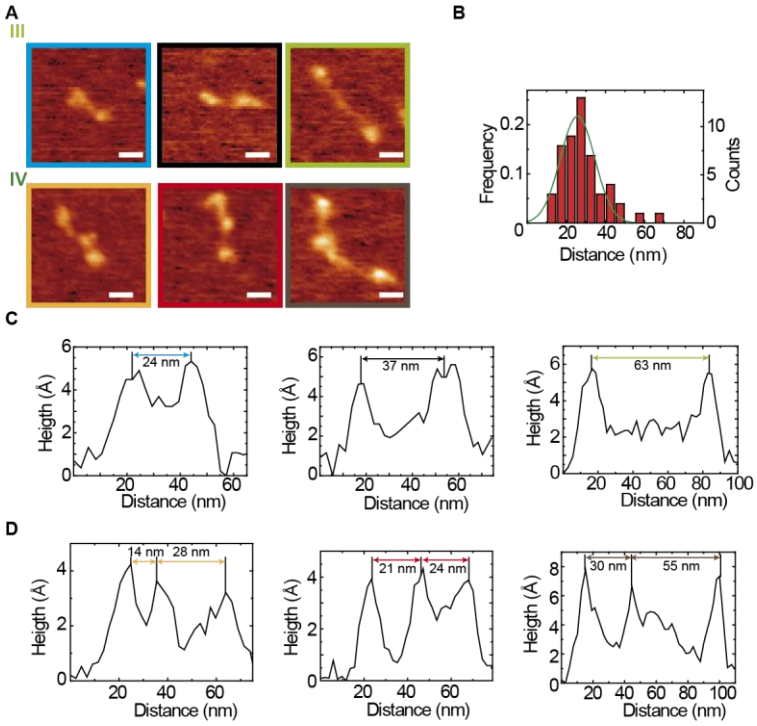


Figure III-9. AFM images and analysis on CtIP flexibility.

A) AFM images of molecules from class III and IV. Scale bar 20 nm. B) Contour length histogram for tailed domains. Peak at 26±18 nm. C) Profiles and contour lengths for class III molecules of A). D) Profiles and contour lengths for class IV molecules of A).

To study the dynamic behaviour of CtIP molecules, the protein was incubated and imaged in a liquid environment. Liquid medium samples were incubated in similar conditions to air dried samples without drying process. As expected, images acquired on these conditions showed a very high flexibility but also a fast movement and interaction with the surface [Figure III-10]. Due to the fast movement of the molecules the results had low resolution, the images were taken at a rate of 1 image per 10-15 seconds [Figure III-10 A-B]. To overcome the limits of the image acquisition time of the Nanotec S.L. Atomic Force Microscopy, a collaboration with the group of Prof. Toshio Ando at the Nano-Life Science Institute, Japan, was proposed to visualize the protein in a High-Speed Atomic Force Microscope. As described in the introductory section of this thesis, Professor Toshio Ando developed a High-Speed AFM for biological samples capable of image acquisition at a 10-15 Hz rate, 1 image per 4-6 seconds. There, CtIP protein molecules were imaged with one of the High-Speed Atomic Force Microscopes [Figure III-10 C-D]. The spatial resolution of the images acquired with Ando's HS-AFM proved to be similar to the resolution of the Nanotec S.L. AFM but at a much better time resolution. The dynamic behaviour of CtIP and fast exchange of proteins in close interaction with the surface made it challenging to obtain videos of moving proteins with a Nanotec AFM but the same protein was imaged with a High-Speed AFM at higher resolution. In Figure III-10, two CtIP molecules are showed in sequential images for a conventional AFM, left, and a High-Speed AFM, right. The time resolution of the HS-AFM allowed the acquisition of images without the "parachutting" effect and the protein movement is not masked by the movement of the tip along the image surface seen in the conventional AFM.

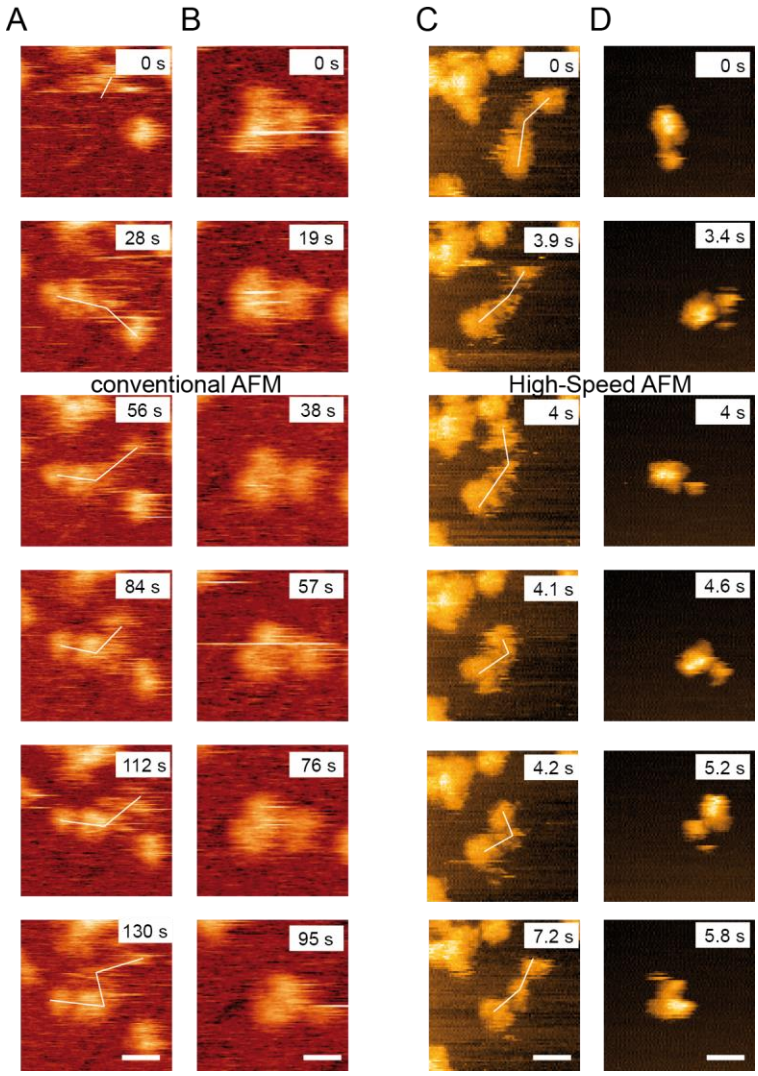


Figure III-10. Comparative CtIP images acquired by two AFMs inside liquid medium.
A) Images of a class III molecule acquired by a Nanotec AFM at 154x154 pixels and 28 s per image. Observe in images 1, 2, 5 and 6 the blurring regions corresponding to the disordered domain of CtIP. B) Images of a class II or III molecule acquired by a Nanotec AFM at 100x100 pixels and 19 s per image. C) Images of a class II-III molecule acquired by a High-Speed AFM at 100x100 pixels and 100 ms per image. D) Images of a class II molecule acquired by a High-Speed AFM at 100x100 pixels and 150 ms per image. Scale bar of images 20 nm.

The faster acquisition time revealed some aspects of CtIP dynamic behaviour which were consistent with the flexibility of the coiled-coils and disordered domains [Figure III-10 C-D]. The contour length of molecules measured by the Nanotec AFM was consistent with the predicted length of 30 nm per tailed domain

but the High-Speed AFM revealed a very fast exchange between "small" and "large" conformations of tailed domains [Figure III-9 B, Figure III-10 C-D, Figure III-11].

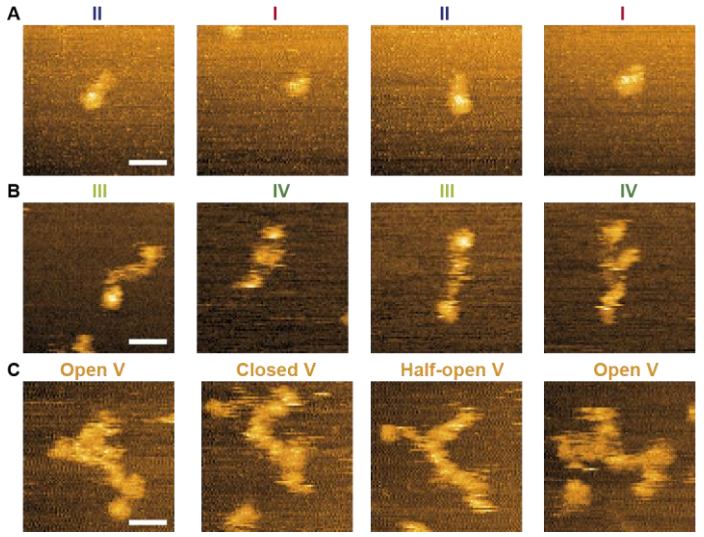


Figure III-11. Exchange of classes as seen by HS-AFM.

A) Fast-moving CtIP molecule switching conformation between classes I and II. B) Fast-moving CtIP molecule switching conformation between classes III and VI. C) Fast-moving CtIP molecule switching conformation between close and open class V. Images of 100x100 pixels and 100 ms per image. Scale bar is 20 nm.

The High-Speed AFM videos showed an interconnection between classes I and II and classes III and IV [Figure III-11 A-B]. These results were in agreement with the similar volume distribution of these classes [Figure III-7 C]. The largest class of species, class V, showed also dynamic behaviour with open, splayed globular domains, and closed, globular domains nearby, conformations [Figure III-11 C].

III.4.1.3 CtIP bridges non-canonical DNA-ends *in vitro*

The results obtained on the molecular state of the protein by AFM were useful for the understanding of the binding interaction between damaged-DNA and CtIP. To further extend this knowledge on CtIP interaction with DNA, a DNA molecule with non-canonical DNA-ends was incubated with CtIP and imaged under an AFM. To analyse the possible formation of CtIP-DNA complexes we engineered a small DNA molecule, 547 bp, with two short, 20 bp, single-stranded DNA overhangs attached on both 3' and 5' ends. The DNA molecule simulates a broken DNA piece generated by a Double-Strand Break with "dirty" ends. Before Atomic Force Microscopy analysis, EMSA assays showed that this DNA molecule interacts with CtIP [Wilkinson, et al. (2019)]. As control, we first imaged the DNA molecule

deposited onto a mica surface alone [Figure III-12 A]. The molecular size distribution fits to a single population of molecules with 195±22 nm in contour length [Figure III-12 A]. In agreement with the size of the DNA molecule (587 bp) in length DNA molecule [Wing, Drew et al. (1980)]. Next, we repeated the same experimental procedure in the presence of increasing ratios of DNA:Wild type $CtIP_4$. The ratios tested were 1:0, 1:6, 1:12, 1:19 and 1:25. Samples prepared under these ratios showed a combination of DNA, CtIP and complex molecules formed by the expected CtIP:DNA interaction. In addition, the AFM titration analysis revealed an increased number of long molecules in relation to the increase of the CtIP concentration. This effect is showed in Figure III-12 D as cumulative frequency of molecules at different DNA:$CtIP_4$ ratios. For the control sample, molecular complexes longer than 195±22 nm accounted only for the 4% of total molecules. At the 1:25 ratio of the DNA:CTIP samples, the percentage was increased to around 30%. Specifically, bridged molecules of DNA-CtIP (complex molecules with contour length being a multiple of 195±2σ nm, where σ refers to the standard error of the measure or in this case 11 nm), were not detected on the control samples but accounted for 7% in the Wild type CtIP samples [Figure III-12 D]. Note that analysis of contour lengths did not reflect the intramolecular binding events as these are single circularised molecules [Figure III-12 C]. Moreover, Atomic Force Microscopy images revealed intermolecular connections of molecules [Figure III-12 B, zoomed images]. However, the fraction of complex molecules observed was lower than that revealed by the electrophoresis assays performed by Dr. Oliver Wilkinson at Prof. Mark S. Dillingham laboratory because other biochemical analysis revealed an inhibitory effect to the DNA:CtIP interaction by Mg^{2+} ions which were present in AFM images to facilitate the deposition of DNA molecules onto a mica surface [Bezanilla, et al. (1995); Wilkinson, et al. (2019)]. The results, together with other biochemical data, suggest that the interaction of CtIP was not strictly limited to the DNA ends [Wilkinson, et al. (2019)]. To further prove this interaction effect of CtIP with DNA molecules, the R839A and L27E mutants and a dephosphorylated fraction of Wild type CtIP were incubated with DNA molecules.

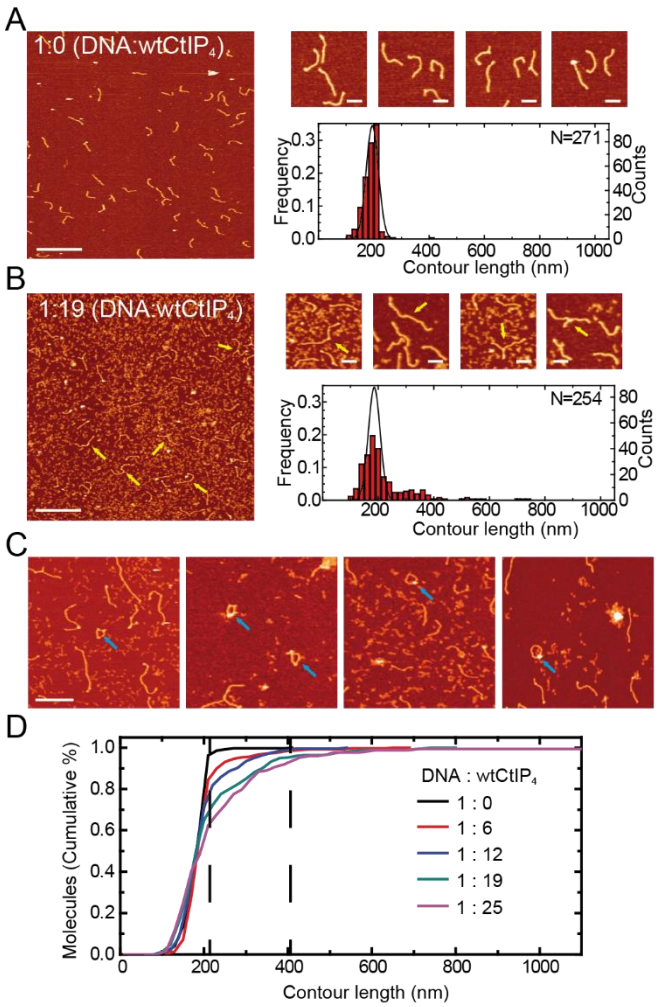


Figure III-12. Atomic Force Microscopy images and analysis of the DNA-CtIP interaction. *A) Representative AFM images of forked DNA molecules in the absence of CtIP. The contour length histogram shows a single Gaussian peak centred on a value equivalent to a single contour length. Scale bar is 500 nm for left image and 100 nm for zoomed examples. B) Representative AFM images of DNA-CtIP complex molecules. The contour length histogram shows multiple size populations. The solid line indicates the Gaussian distribution for the control sample population. Yellow arrows indicate observed complex molecules. Scale bar is 500 nm for left image and 100 nm for zoomed examples. C) Representative AFM images of intramolecular binding events. Blue arrows indicate position of circularized molecules. Scale bar is 250 nm. D) Cumulative graph for contour length molecule population. Increasing ratios of DNA:CtIP emphasized the effect of CtIP on the appearance of longer molecules. Dashed lines highlight the position of single and double contour length populations.*

III.4.1.4 Dephosphorylation of CtIP increases DNA-bridging

Phosphorylation processes and enzymes are associated to the activation of protein molecules for specific mechanism of the Synthesis and Second-growth phases of the cell cycle. Previous studies have shown an increase in DNA binding affinity for CtIP when treated with phosphatases [Anand, et al. (2016)]. Consequently, to analyse the effect of dephosphorylation on CtIP, Wild type CtIP was purified and treated with ʎ phosphatase. The previously presented images and analysis were repeated with this CtIP fraction (ʎCtIP). Atomic Force Microscopy image analysis shows an increase number of complex molecules in similar conditions to WT CtIP samples. The cumulative graph of the molecular contour length distribution shows an increase in molecular populations longer than the control sample [Figure III-13 C, left black dashed line]. Biochemical studies corroborated the effect [Wilkinson, et al. (2019)]. Dephosphorylation increased the affinity of CtIP for non-canonical DNA ends. The stoichiometry between DNA and $CtIP_4$ molecules was described as 0.56 $CtIP_4$ per DNA molecule or, approximately, one $CtIP_4$ per two DNA molecules from the anisotropy assays performed in Wilkinson, et al. (2019).

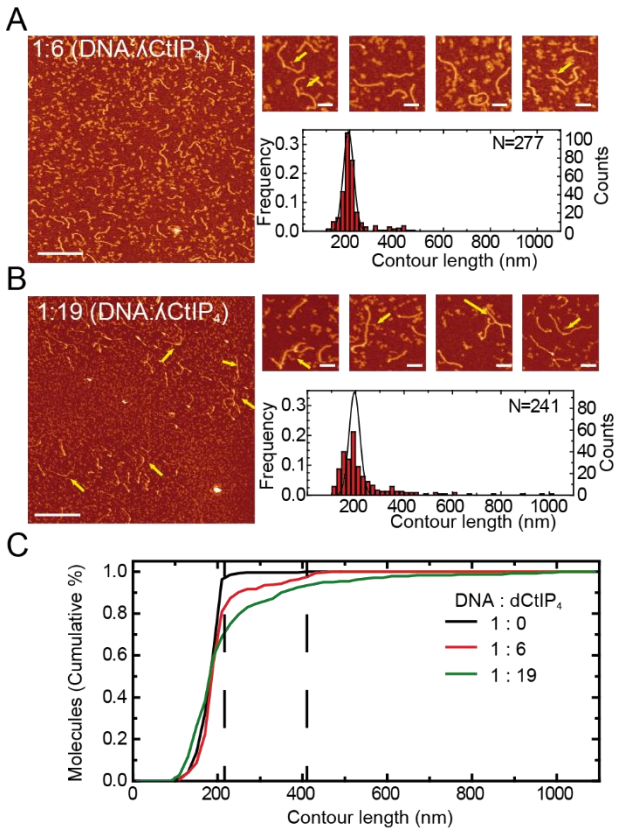


Figure III-13. Atomic Force Microscopy images and analysis of DNA-dephosphorylated CtIP interaction.

A) Representative AFM images of DNA-dephosphorylated CtIP complex molecules at 1:6 ratio. The contour length histogram shows multiple size populations. The solid line indicates the Gaussian distribution for the control sample population. Yellow arrows indicate observed complex molecules. Scale bar is 500 nm for left image and 100 nm for zoomed examples. B) Representative AFM images of DNA-dephosphorylated CtIP complex molecules at 1:19 ratio. The contour length histogram shows multiple size populations. The solid line indicates the Gaussian distribution for the control sample population. Yellow arrows indicate observed complex molecules. Scale bar is 500 nm for left image and 100 nm for zoomed examples. C) Cumulative graph for contour length molecule population. Dashed lines highlight the position of single and double contour length populations.

III.4.1.5 The N and C-terminal domains of CtIP are crucial for DNA-bridging

To understand the role of the N- and C-terminal domains of the protein we repeated the analysis with the mutations L27E and R839A. In Wilkinson, et al. (2019) and previous biochemical studies, the formation of a CtIP tetramere from the N-terminal connections is showed to be essential for the connection of damaged DNA-ends in Homologous Recombination [Davies, et al. (2015)]. In contrast, the role of the C-terminal domain is not completely understood. The globular domains have a Sae2 homologous DNA-binding domain composed of a Ribbon-Helix-Ribbon motif [Ferretti, et al. (2016)]. Analysis by Atomic Force Microscopy of the DNA-CtIP complexes present in a sample incubated either with L27E or R839A revealed a significant decrease of molecules longer than 195±22 nm [Figure III-14 A-B]. Somewhat surprising, the effect was similar between both mutants. The results showed a decrease from around 30% of complex molecules for Wild type CtIP to 10% for these mutants [Figure III-12 D, Figure III-14 C]. This result suggests a cooperative role of the N- and the C-terminal domains, therefore, the relevance of the formation of a stable tetrameric protein from the N-terminal with four active C-terminal DNA-binding domains.

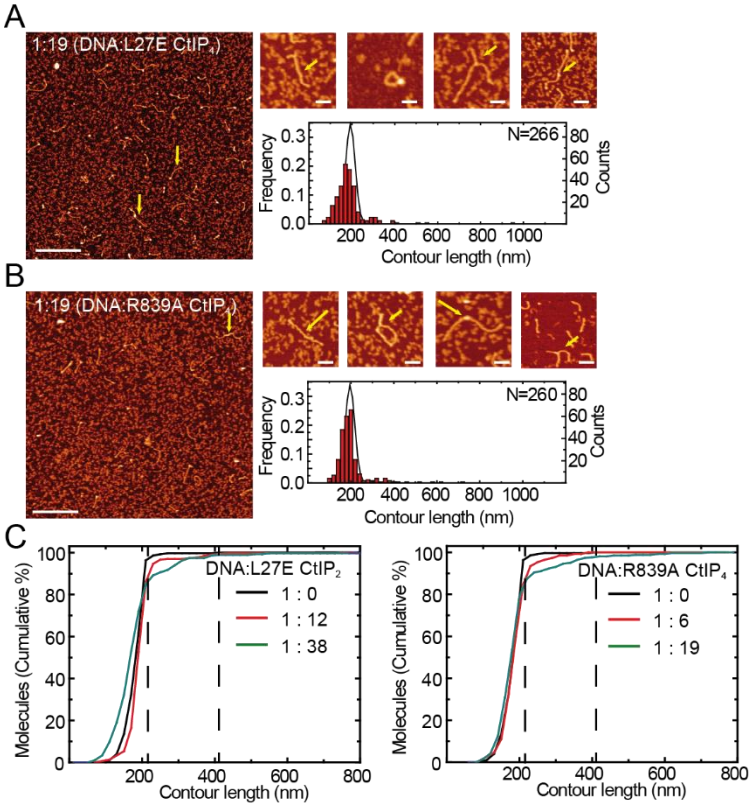


Figure III-14. AFM images and analysis of the mutant CtIP interacting with DNA.

A) Representative AFM images of DNA-L27E CtIP complex molecules at 1:19 ratio. The contour length histogram shows a decrease in complex molecules compared to WT CtIP. The solid line indicates the Gaussian distribution for the control sample population. Yellow arrows indicate observed complex molecules. Scale bar is 500 nm for left image and 100 nm for zoomed examples. B) Representative AFM images of DNA-R839A CtIP complex molecules at 1:19 ratio. The contour length histogram shows a similar decrease compared to the L27E. The solid line indicates the Gaussian distribution for the control sample population. Yellow arrows indicate observed complex molecules. Scale bar is 500 nm for left image and 100 nm for zoomed examples. C) Cumulative graphs of contour length per molecule population for L27E (left) and R839A (right). Dashed lines highlight the position of single and double contour length populations.

III.4.1.6 CtIP role in syndromes Seckel 2 and Jawad

The role of CtIP as a key protein for the first steps of Homologous Recombination is emphasised from its role in syndromes Seckel 2 and Jawad [Qvist, et al. (2011)]. These syndromes are related to infant deficiency in DNA-repair mechanisms and cause accelerated aging and mental retardation. Studies correlated CtIP gene variations with the appearance of these two syndromes [Qvist, et al. (2011)]. These mutations cause a premature transcription stop of the CtIP gene. The stop creates a truncated CtIP with less aminoacids on the C-terminal domain. In Seckel 2 syndrome the protein is expressed with 782 aminoacids instead of the 897 aminoacids of the Wild type CtIP. The Jawad syndrome CtIP protein contains 608 aminoacids. The reduction in C-terminal domain aminoacids results in a defective Double-Strand Break-repair processing protein. It is also associated to a defect in the ATR-dependent signalling process [Sartori, et al. (2007)]. To comprehend the role of these syndromes on DNA-bridging, the truncated versions of CtIP were produced. A third variant was tested, R100W, which is also described to be present in Seckel 2 syndrome patients [Shaheen, et al. (2014)].

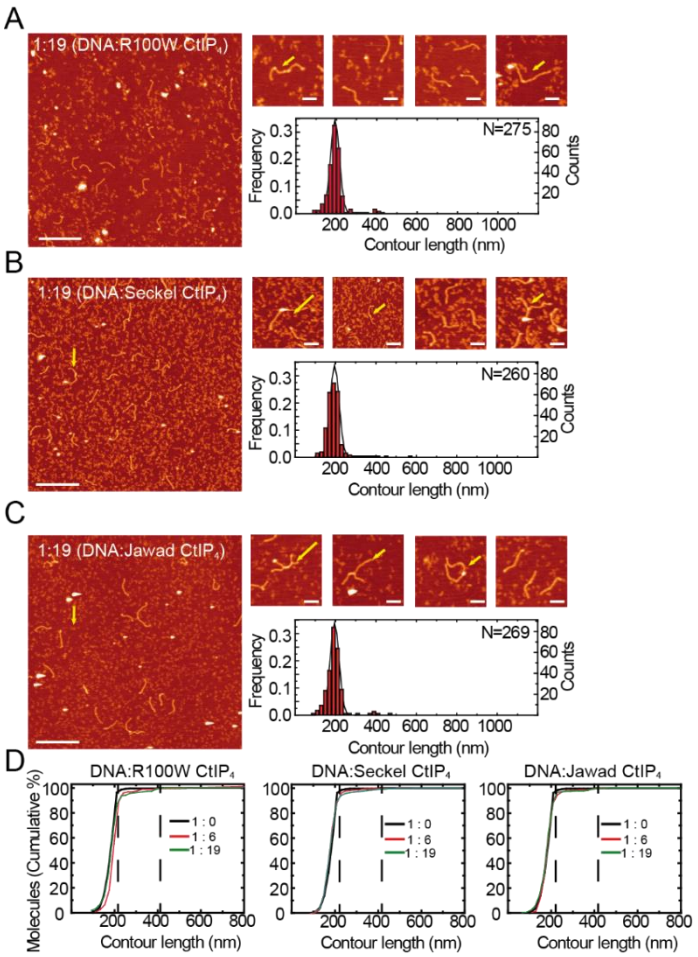


Figure III-15. Atomic Force Microscopy images and analysis of syndrome CtIP variants. *A) Representative AFM images of DNA-R100W CtIP complex molecules at 1:19 ratio. The contour length histogram shows a similar population for complex molecules compared to the R839A and L27E mutants. The solid line indicates the Gaussian distribution for the control sample population. Yellow arrows indicate observed complex molecules. Scale bar is 500 nm for left image and 100 nm for zoomed examples. B) Representative AFM images of DNA-Seckel 2 CtIP complex molecules at 1:19 ratio. The contour length histogram shows barely any population outside the control distribution. The solid line indicates the Gaussian distribution for the DNA sample population of the DNA sample. Yellow arrows indicate observed complex molecules. Scale bar is 500 nm for left image and 100 nm for zoomed examples. C) Representative AFM images of DNA-Jawad CtIP complex molecules at 1:19 ratio. The contour length histogram, similarly to the Seckel 2, shows barely any molecular species outside the control distribution. The solid line indicates the Gaussian distribution for the control sample population. Yellow arrows indicate observed complex molecules. Scale bar is 500 nm for left image and 100 nm for zoomed examples. D) Cumulative graphs of contour length per molecule population for R100W (left), Seckel 2 (middle) and Jawad (right). Dashed lines highlight the position of single and double contour length populations.*

The three mutations tested had very low affinity for DNA bridging as revealed by the cumulative graphs [Figure III-15 D]. The molecule population outside the control distribution for the R100W mutant had similar results to the previously showed R839A and L27E CtIP mutations. The other two syndrome proteins had poor bridging ability with very few molecules outside the control distribution. These results suggest an inactivation of the bridging interaction of CtIP in both syndromes which were also related to defective processing of Double-Strand Breaks [Sartori, et al. (2007)].

III.4.1.7 CtIP "slides" through double-strand DNA

The current model on the CtIP role on Homologous Recombination proposes a static collaboration between the MRN complex and CtIP to bind and bridge broken DNA ends [Forment, Jackson et al. (2015)]. Nevertheless, analysis of the dynamic behaviour of CtIP by High-Speed Atomic Force Microscopy in section III.4.1.2 showed the dynamics of a fast and flexible protein. The flexibility of CtIP may be of importance to stablish a non-rigid stable union between broken ends. To study the dynamic behaviour of the CtIP-DNA complexes, Wild type CtIP was incubated with a long, double-stranded DNA molecule with a small single-strand flap for CtIP binding. The samples were analysed with a High-Speed Atomic Force Microscope from Prof. Toshio Ando laboratory. The preparation of the sample was found to be troublesome due to frequent tip-sample interaction. However, videos were acquired of CtIP molecules "sliding" through a double-strand DNA molecule [Figure III-16]. In these videos it is revealed that CtIP did not need of free-DNA ends for binding to dsDNA. Once bound to DNA the protein is able to move along the molecule [Figure III-16]. This result was in agreement with intermolecular connection of molecules suggested in section III.4.1.3 as the protein did not necessarily bridge DNA molecules through their ends.

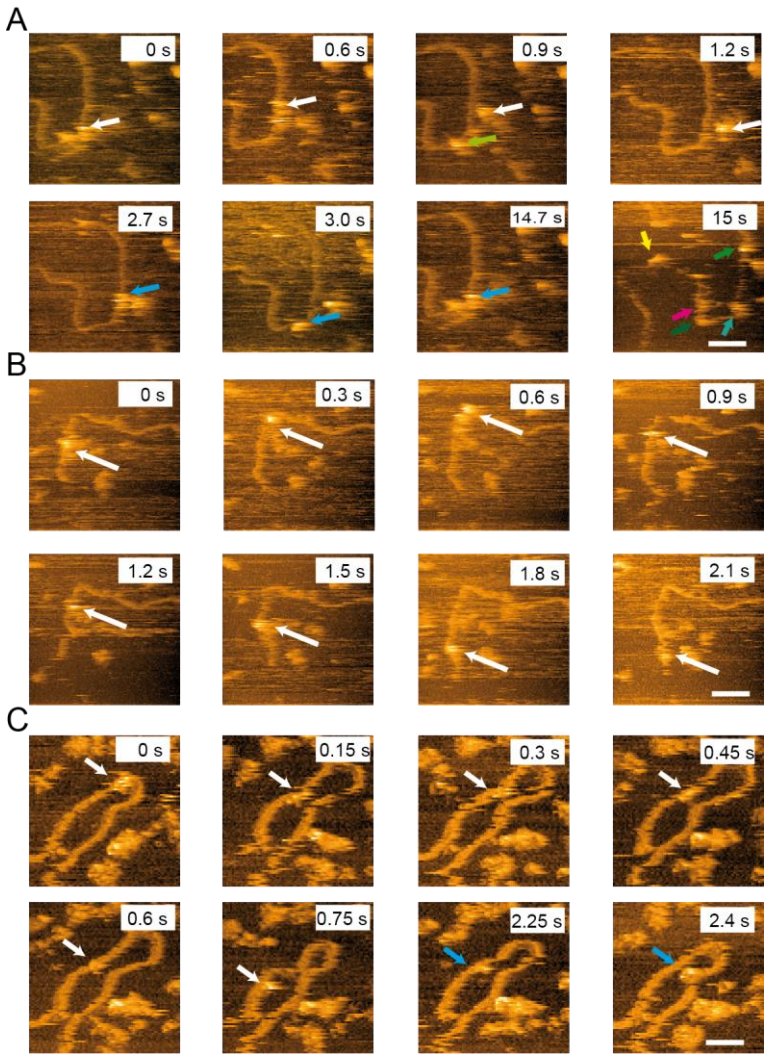


Figure III-16. CtIP molecules moving along a double-stranded DNA molecule.

A) Images of WT CtIP molecules moving along a dsDNA molecule anchored to a mica surface inside liquid medium. Arrows signal individual molecules "sliding" through DNA. Images acquired of 80x80 pixels at 300 ms per image. Scale bar is 80 nm. B) Images of L27E CtIP molecules moving along a dsDNA molecule. Arrow signals an individual molecule. Images acquired of 80x80 pixels at 300 ms per image. Scale bar is 80 nm. C) Images of WT CtIP moving and exchanging strands on a dsDNA molecule. Arrows signal individual molecules. Images acquired of 80x80 pixels at 150 ms per image. Scale bar is 40 nm.

The DNA molecule used in this section has blunt ends. In biochemical assays, this type of DNA ends were found to be less susceptible to CtIP binding [Wilkinson, et al. (2019)]. However, the exchange of CtIP proteins bound to DNA

was fast. Few molecules were captured trapped at a DNA-end. Surprisingly, it was frequent to find many CtIP protein molecules moving along a DNA molecule.

III.4.2 Double-Strand Breaks long resection

III.4.2.1 Bad remains attached to double-stranded DNA after the translocation had stopped

To further understand the Homologous Recombination mechanism, a DNA2-like helicase/nuclease was imaged under High-Speed AFM. Bad from *Geobacillus stearothermophilus* is a helicase/nuclease protein with 5'>3' helicase and ssDNA translocase with DNA2 homology in its structure and function as an ATP-dependent 5'-flap endonuclease [Balakrishnan, et al. (2010); Kim, et al. (2006)]. Other single molecule studies of this type of proteins have described a fast-paced unwinding mechanism of double-stranded DNA molecules with small 5' single-strand DNA overhangs. The activation of the nuclease or helicase domains are controlled by the concentrations of ATP and Mg^{2+}. In conditions of [ATP] <[Mg^{2+}], the DNA2-like proteins cleave ssDNA overhangs located in front of the protein, while, in conditions of [ATP]>[Mg^{2+}], they are able to unwind the DNA without detaching from it due to the cleaving of one strand. Nonetheless, a nuclease-dead mutant is necessary for the protein to completely unwind a DNA molecule [Levikova, et al. (2013)]. Without endonuclease activity the protein remains attached to the dsDNA throughout the unwinding mechanism. To better comprehend this mechanism we prepared a long dsDNA molecule with a ssDNA overhang. A reaction process between the DNA molecule and a nuclease-dead mutation of *Geobacillus stearothermophilus* Bad was imaged by High-Speed Atomic Force Microscope. As preliminary results, AFM videos showed Bad molecules anchoring to the DNA molecule at the expected position of the single-strand overhand in conditions of [ATP]<[Mg^{2+}] or helicase inactive [Figure III-17 A].

In addition, a DNA unwinding reaction was performed with Bad and the DNA molecule in conditions of [ATP]>[Mg^{2+}], after incubation, the reaction was stopped by deposition onto a mica surface. The HS-AFM videos showed long DNA molecules connected in the presence of protein [Figure III-17 B]. The unwound dsDNA is able to rehybridise behind the helicase by the absence of single-strand binding proteins, as opposed to the *in vivo* conditions which do not allow the rehybridation [Bae, et al. (2001)]. However, Bad remained bound to DNA molecules in spots far from the ssDNA overhang position stabilizing long DNA loops [Figure III-17 B].

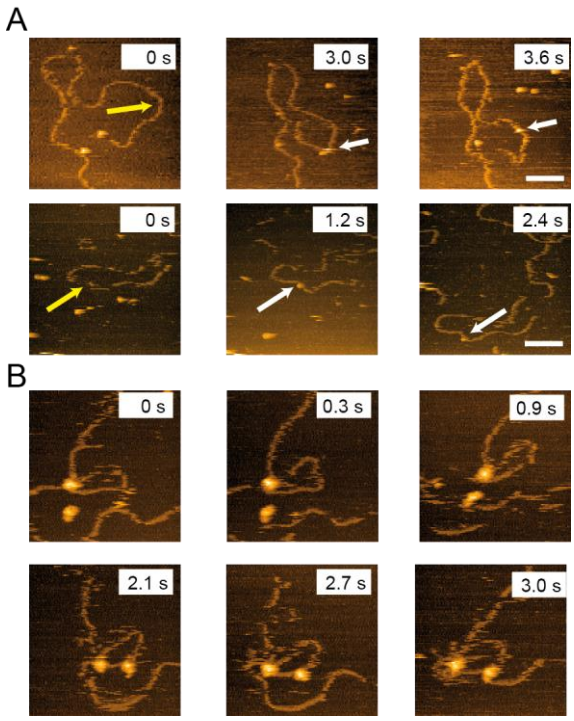


Figure III-17. Dna2 molecules imaged on double-strand DNA.

A) Dna2 molecules binding at the expected position of the single-strand overhand in the double-strand DNA in presence of [ATP]<[Mg^{2+}]. Yellow arrow signals the expected position of the ssDNA overhang (445 bp from one of the DNA ends); white arrows signal a Dna2 molecule bound to the DNA molecule. Upper and lower column represent different DNA molecules. Images acquired of 80x80 pixels at 150 ms per image. Scale bar is 80 nm. B) Two examples of Dna2 molecules bound to a dsDNA molecule and stabilizing a DNA-loop. Upper and lower column represent different DNA molecules. Images acquired of 80x80 pixels at 300 ms per image. Scale bar is 40 nm.

III.5 Discussion

In this study we have characterized the conformation of the *Homo sapiens* CtIP, and its DNA binding properties by Atomic Force Microscopy. Complementarily, data shown on this section from biochemical assasy were performed by Prof. Mark S. Dillingham laboratory members and Electron Microscopy images were obtained by Prof. Dale B. Wigley laboratory members [Wilkinson, et al. (2019)]. Atomic Force Microscopy images together with negative-stain Electron Microscopy and SEC-MALS revealed a CtIP tetrameric state with a "dumbbell" shape. Additionally, AFM and EM showed other secondary protein conformations with additional analysis revealing a volume relation to other oligomeric states differing from the stable tetrameric state [Figure III-7]. The tetrameric "dumbbell" shape of CtIP was in agreement with previous studies based on the hypothesis of a "dimer of dimers" conformation [Andres, et al. (2015); Davies, et al. (2015); Forment, et al. (2015)]. Supplementary analysis of CtIP conformational dynamics by High-Speed AFM showed a fast and dynamic protein [Figure III-11]. Altogether, we proposed a working model for the description of the structure of wild type CtIP [Figure III-18]. In this model, each of the 5 classes of CtIP molecules are connected to a CtIP oligomeric or conformation state. By volume distribution analysis the monomeric state was associated to CtIP classes I and II [Figure III-7], the dimeric state corresponds to classes III and IV and the tetrameric state is noted as class V. From the shape of the molecule and the expected domain arrangement from the "dimer of dimers" model each of the classes is given an overall conformation. For class I species, the disordered and N-terminal domains are represented as a wrapped complex around the globular domain. In class II, the protein is represented as an elongated domain with a tail. In class III, the two globular domains from the C-terminal globular domains are represented in contact with the disordered tails elongated. For class IV, the C-terminal globular domains are represented as separated domains with a connection through the disordered domain. In class V, the tetrameric state of the protein with is represented with a "dumbbell" or "splayed dumbbell" shape in concordance with a "dimer of dimers" model. The globular domains are grouped two by two in the "dumbbell" state or are represented separately in a "splayed dumbbell" conformation [Figure III-11].

Furthermore, Electrophoretic Mobility Shift Assays and fluorescence anisotropy assays found that the tetrameric state of CtIP was capable of tightly bind to DNA molecules (data not shown) [Wilkinson, et al. (2019)]. A crucial role of the C-terminal domains of the CtIP monomers is evidenced by the inability of the mutant R839A to bind to DNA despite retaining a comparable oligomeric state to the wild

type CtIP [Figure III-7; Figure III-8]. On the other hand, a mutation on the tetramerisation domain located at the N-terminal, L27E, was also found to be defective in DNA binding. In this case, this mutant does not form tetramers contrary to the wild type and R839A proteins [Figure III-7; Figure III-8].

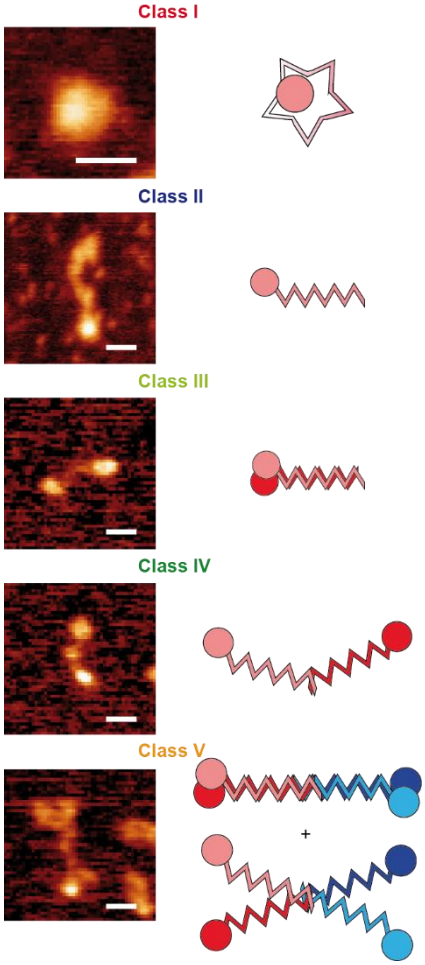


Figure III-18. Model proposed for CtIP conformation compared to Atomic Force Microscopy images.

A) AFM zoomed images of the 5 classes of CtIP molecules found. Scale bar is 25 nm. B) Model proposed for CtIP molecules found on AFM images. The globular C-terminal domain is represented as a circle while the disordered and coiled-coil N-terminal domains are represented as zig-zag lines.

Electrophoretic Mobility Shift Assays revealed an increase on wild type CtIP binding to DNA with non-canonical DNA-ends. The biggest increase was found for single-strand DNA forks or model nucleoprotein blocks (data not shown) [Wilkinson, et al. (2019)]. This effect suggests a preference for “complex” or “dirty”

ends formed in some metabolic processes such as those in topoisomerase poisoning [Aparicio, et al. (2016)]. However, the mechanism by which CtIP is able to process these DNA damages remains unknown. Some hypotheses for the working model of this mechanism suggest a direct binding of CtIP to the nucleo-protein or "dirty" nucleic acid conjugate or of a more dynamic behaviour by which CtIP binds and slides through double-strand DNA with the DNA-ends acting as exits. The first hypothesis seems unlikely as a direct interaction between the protein and a variable DNA-end blockage is improbable. The second hypothesis, however, is in agreement with the fast and dynamic behaviour of the protein as seen by High-Speed AFM [Figure III-16]. In this case, dissociation of CtIP from the double-strand DNA may be inhibited by the presence of bulky or blocked DNA-ends. The importance of CtIP interaction with other Double Strand Break-repair proteins such as the MRN complex and BRCA1-BARD1 remains to be tested.

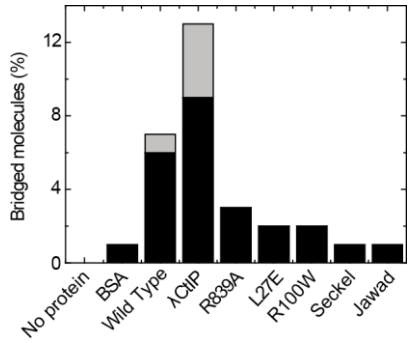


Figure III-19. Bar graph of bridge DNA molecules by the action of CtIP.

Bar graph representing the population of bridged molecules for wild type, dephosphorylised and mutated CtIP compared to a sample of the same DNA incubated with Bovine-Serum Albumin. The black bar of each sample represents the percentage of complex molecules with double contour length of the DNA-alone sample (range of contour lengths corresponding to 390±44 nm). Grey bar corresponds to triple, quadruple and quintuple contour lengths (range corresponding to 585±44 nm, 780±44 or 975±44 nm respectively).

Additionally, Atomic Force Microscopy analysis of DNA-CtIP molecular complexes revealed a bridging interaction between CtIP and double-strand DNA molecules with small single-strand DNA overhangs. In Figure III-19, the population of molecules composed of 2x DNA lengths (black bars) and of 3x, 4x or 5x DNA lengths (grey bars) is showed for all the CtIP variants tested in this project. The control sample with no protein and a sample prepared with Bovine-Serum Albumin, BSA, are also presented as negative control. As expected, the population of bridge molecules is the highest for the wild type and dephosphorylated CtIP experiments whereas mutations on either the C-, N-

terminal or syndrome-related CtIP molecules have marginal effect for bridging compared to the control cases. The mutations tested together with the flexibility analysis and the High-Speed AFM images of CtIP interactions with double-strand DNA suggest a dynamic CtIP-DNA interaction behaviour [Figure III-9; Figure III-16]. As such, the model proposed in this thesis is focused on the more dynamic "slide and capture" mechanism [Figure III-20].

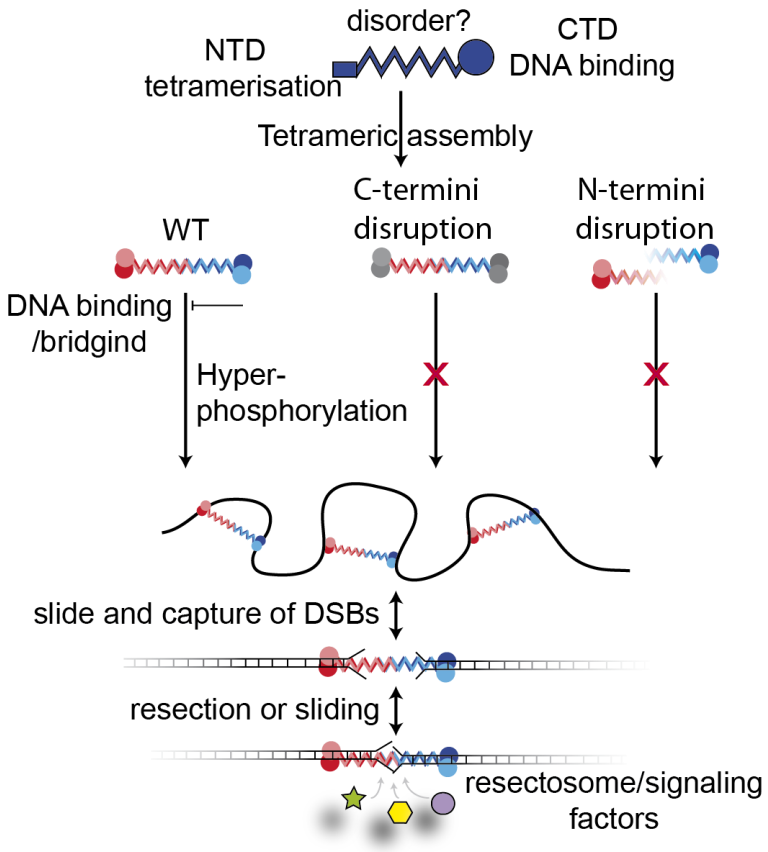


Figure III-20. Model proposed for DNA-bridging mechanism by CtIP.
The protein CtIP consists of a globular C-terminal, a disorder central and a coiled-coil N-terminal domains. Assembly of 4 CtIP molecules forms the active tetrameric structure with a "dumbbell" shape. Disruption of either the C- or N-terminal domain results in low DNA-binding affinity. Activation of CtIP by dephosphorylation results in an increase of the DNA-binding affinity. In this model of a "slide and capture" mechanism CtIP may bind the dsDNA and slide through the molecule up to DSB DNA-ends. Once on a broken end CtIP remains attached and may stabilize the union of other cofactor and DNA molecules necessary for HR repair mechanism.

Hereafter, the study of the role or cooperation of CtIP with other Homologous Recombination proteins and a more in-depth structural analysis of the C-terminal

domain may shed more light on these steps of the HR mechanism. This project will remain as an open research for further study of other HR proteins and mechanism related to CtIP. However, the results obtained had been of importance for the elaboration of a working model for CtIP.

One of the following projects would be to study and understand the role of Dna2 with and without CtIP in the long-resection step of Homologous Recombination. At the same time of the CtIP study some research experiments of *G. stearothermophilus* Dna2 were performed with Atomic Force Microscopy and High-Speed AFM. The data showed in chapter III.4.2 are the results obtained with *G. stearothermophilus* Dna2 and High-Speed AFM. The protein proved to be susceptible to the interaction with DNA deposited on a surface, either functionalized mica with cations, lipids, APTES or non-functionalized mica, thus, the opening of the double-strand DNA molecule by the helicase was not observed. However, incubation of Dna2 with the DNA molecule described in chapter III.3.2.3 and posterior deposition onto a mica surface functionalized with cations proved successful. The result suggests a strong attachment of the protein to the DNA molecules during the unwinding process as molecular complexes were found on the mica surface with clear connection spots. These molecular complexes can be interpreted as the result of the DNA rehybridisation behind Dna2 due to the lack of single-strand binding proteins [Figure III-21]. Nonetheless, the loops seen in the AFM images may be the result of the protein being capable of binding to both DNA strands instead of opening the DNA by binding to only one DNA strand. Thus far, the role of the helicase/nuclease Dna2 remains to be understood.

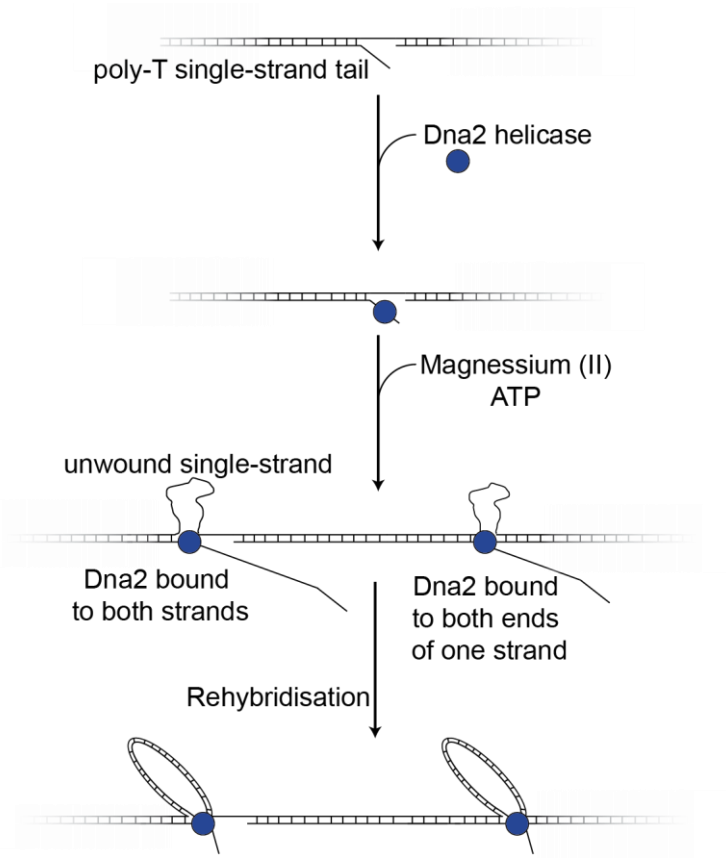


Figure III-21. Model proposed for the binding and unwinding of the Dna2 helicase to a DNA molecule with a small single-strand overhang.

The nuclease devoid mutant of Geobacillus stearothermophilus Dna2 unwinds DNA molecules with small single-strand DNA overhangs in conditions of higher ATP concentration than Magnesium. Disruption of the nuclease activity of Dna2 allows the protein to unwind DNA without detaching from it. The results are long molecular complexes with the protein attached to the DNA molecules at intersections of various double-strand tails. The binding mode of Dna2 to the DNA molecules may be a key feature for the rest of the Homologous Recombination repair proteins.

IV. DNA MISMATCH REPAIR

The transference of genes between different organism leads to genetic diversity associated to evolution. Typically, gene transference refers to the incorporation of genes from an organism into its offspring, known as vertical gene transfer. For example, Homologous Recombination is a functional mechanism for the integration of DNA in either vertical or horizontal gene transfer. Horizontal gene transfer indicates the incorporation of genes from a different organism. The Mismatch Repair is a mechanism for the regulation and control of the DNA integration in the vertical gene transfer process.

“Biology will relate every human gene to the genes of other animals and bacteria, to this great chain of being” – Walter Gilbert, 1980 Physics Nobel laureate

From the previous chapter, Homologous Recombination is a mechanism for DNA Double-strand Break-repair and is also associated to vertical gene transfer. During Homologous Recombination, the chromosomic material of an organism may undergo an exchange process linked to genetic diversity. In the case of vertical gene transfer, an exogenous DNA sequence is introduced into an organism and sometimes introduced into its chromosomic material. Vertical gene transfer is one of the main mechanisms for antibiotic resistance spreading in bacteria. Cellular organisms have developed "proofreading" mechanisms to regulate the content of their chromosomic material.

IV.1 Introduction

Diversity in nature plays the role of organism adaptation and survival [Nevo (2001)]. Genetic diversity can be obtained by means of gene transfer between non-related organisms in prokaryotes, such as bacteria. Horizontal gene transfer is of lesser importance in eukaryotes as the DNA transference regularly occurs from father to offspring [Nelson, et al. (1999); Doolittle (1999); Lan and Reeves (1996); Andersson (2005)]. Eukaryotes have developed specific mechanisms for Horizontal Gene Transfer and are more focused on sexual gene transfer between cells of the same species [Hamilton, et al. (1990)].

In bacteria, the process of Horizontal Gene Transfer was first discovered in 1928 by Frederick Griffith. He injected mice with a living strain of a non-virulent bacteria and a dead strain of a virulent variant. After a few hours, the mice died from the action of the bacteria. This result was the first evidence of gene transference between bacteria [Griffith (1928)]. In prokaryotes, Horizontal Gene Transfer is divided into three mechanisms: transformation, transduction and conjugation [Johnston, et al. (2014); Zinder and Lederberg (1952); Tatum and Lederberg (1947)].

Transformation is widely used in biochemistry laboratories [Figure IV-1 A]. The process allows a cell to uptake exogenous DNA from the surrounding environment and incorporate it. This mechanism is activated by the cellular organism in the nature under conditions of pressure or stress [Griffith (1928)]. In a laboratory, cells are modified suitable for transformation [Solomon and Grossman (1996)]. The perpetual transforming cells are known as competent cells. Incorporation of DNA during the transformation process is closely related to survival in bacteria. For example, the activation of the transformation mechanism in bacterial organism through antibiotics stress allows the introduction of exogenous DNA from the surrounding medium increasing the probability to incorporate plamids encoding the necessary antibiotic-resistance [Croucher, et al.

(2009)]. In scientific research projects, this mechanism is used for amplification of plasmid DNA molecules [Avery, et al. (1944)].

Transduction mechanism is defined as the addition of viral DNA into a cellular organism [Figure IV-1 B]. The viral particle is responsible of the insertion of exogenous DNA inside a host-cell membrane. Later, the host replication machinery would be responsible for the replication and transcription of the exogenous DNA [Ikeda and Tomizawa (1965)]. Previous studies have revealed the possibility of incorporating not only viral DNA but also bacterial DNA with this method. In some cases, viral particles include a small fraction of bacterial DNA from a previously infected organism and thus, transduced it altogether with the viral DNA [Brussow, et al. (2004)]. The discovery of the transformation mechanism opened the possibility for the introduction of exogenous DNA into a cellular organism in a laboratory. In this case, cells do not need to be modified as with competent cells. However, the exogenous DNA requires to be integrated inside a viral capsid for its incorporation inside the desired organism [Denning, et al. (2013)].

Conjugation refers to the exchange of genes between organisms by direct contact [Figure IV-1 C]. The pilus of one organism, the donor cell, connects it to the other cell, the recipient cell. The pilus allows the exchange of genetic material between the organisms [Ou and Anderson (1970)]. To transfer the genetic material, the donor cell unthreads and nicks part of its DNA thus keeping a single-strand DNA and donating the other single-strand DNA piece. With enough time, the donor cell may transfer part of its genetic material into the recipient cell [Adelberg and Pittard (1965)]. Then, the transcription machinery of the recipient cell incorporates the newly acquired DNA. This mechanism transfers the pilus genes into the recipient cell. It creates a new pilus converting the previously-recipient cell into a future donor cell [Ou and Anderson (1970)]. As with the previous mechanisms, conjugation has been implemented as a laboratory technique for the insertion of a chosen-piece of DNA into an organism. It has the advantage of not disrupting the membrane of the recipient cell and allowing the exchange of bigger DNA sequences [Heinemann and Sprague (1989); Yoon and Koob (2005)]. The conjugation mechanism has the drawback of requiring an existing donor organism with the chosen DNA already incorporated.

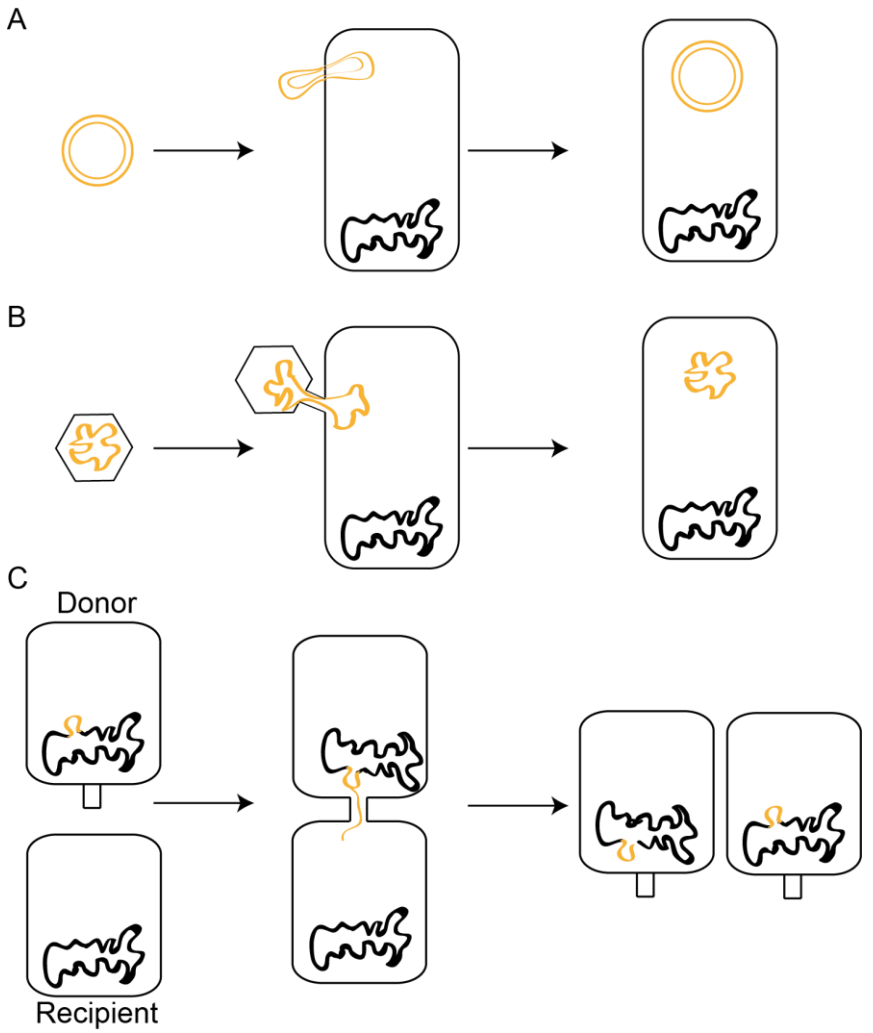


Figure IV-1 Schematic representation of Horizontal Gene Transfer mechanisms.
Schematic representation of the three mechanisms associated to HGT. A) Transformation process, a cell incorporates an exogenous plasmid from the surrounding environment. B) Transduction process, a cell is infected by a viral particle introducing exogenous DNA inside it. C) Conjugation process, a donor cell, holding a pilus, communicates with a recipient cell to exchange DNA.

In the majority of the cases, the insertion of the exogenous DNA into the already existing genomic material is not necessary. For example, plasmids can be directly replicated by the host's replication machinery without the need to incorporate them inside the preexisting chromosomic material [Actis, et al. (1999)] [Figure IV-1 A-B]. Alternatively, the genetic material gained during the Horizontal Gene Transfer process is integrated. For example, the insertion of the F factor transposable element through Homologous Recombination [Chumley, et al.

(1979)] [Figure IV-1 C]. In bacteria, to incorporate the DNA through Homologous Recombination the newly integrated single-strand DNA is first covered by a filament of RecA proteins [Galletto, et al. (2006)]. RecA filaments increase the helical pitch to facilitate the strand-invasion process of a sufficiently long, at least 30 nucleotides [Chen, et al. (2008); Majewski and Cohan (1999)] [Figure IV-2]. The invasive strand forms a D-loop around the homology sequence [Jayasena and Johnston (1993)]. Subsequently, a set of Homologous Recombination proteins exchange the DNA strands [Smith (2012); Dillingham and Kowalczykowski (2008)]. Homology-directed strand-invasion mechanism is thought to be regulated via a "proofreading" mechanism [De Vlaminck, et al. (2012)].

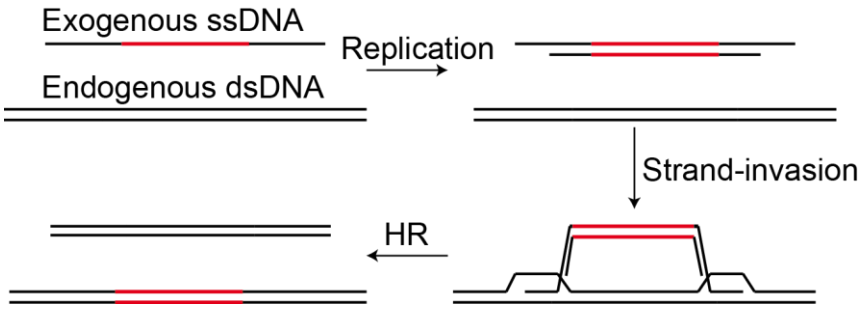


Figure IV-2 Schematic representation of exogenous DNA integration by Homologous Recombination.

Schematic representation of the integration process of exogenous DNA. The newly introduced exogenous DNA is replicated by the host's replication machinery; then, a small portion of ssDNA invades the endogenous DNA on a homology region forming two D-loop structures. Last, the HR mechanism repairs the junction introducing the heterologous region into the endogenous DNA. Image adapted from Kowalczykowski, et al. (1994) and Carrasco, et al. (2019).

At the moment, the "proofreading" mechanism is under debate. For *E. coli* the Mismatch Repair proteins degrade the non-homologous region of the invasive strand for further repair of the heterologous region [Lahue, et al. (1989), Viswanathan, et al. (2001)]. The activation of the Mismatch Repair mechanism decreases, up to 1000-fold, the genomic exchange between species [Rayssiguier, et al. (1989)]. However, natural chromosomal transformation with less than 15% divergence is marginally affected. This raised the question on "what is the effect of the Mismatch Repair proteins on naturally competent bacteria?".

Natural competent bacteria are organisms with the Horizontal Gene Transfer "transformation" mechanism naturally activated. *B. subtilis* organisms are susceptible to natural competence activation. To activate this mechanism, *B. subtilis* cells are depleted of glucose availability at the early stage of the stationary phase [Dubnau (1991)]. Under these conditions, *B. subtilis* cells are able to

uptake extracellular DNA. Once the exogenous DNA is introduced inside the cell wall, the recombination machinery internalizes the exogenous DNA. Studies have shown the importance of the protein RecA and other cofactors for successful integration of DNA molecules [Cox (1993); Carrasco, et al. (2016)]. RecA is a single-strand DNA binding protein. It is responsible for the localization and binding of DNA pieces into homology regions by strand invasion [Jayasena and Johnston (1993)]. However, complete internalization of DNA is thought to be regulated by the Mismatch Repair proteins depending on sequence divergence between exogenous and endogenous DNA.

This chapter results are based on the work published as Carrasco, et al. (2019).

IV.2 Objectives

This project aim was to study and understand the role of the protein MutS from the Mismatch Repair mechanism in *B. subtilis* organisms in the regulation of the integration of non-homologous DNA during the transformation mechanism.

The objectives set were:

7) To understand the activity of the *B. subtilis* recombination machinery with the integration of non-homologous DNA.
8) To comprehend the role of *B. subtilis* MutS with the integration of non-homologous DNA.
9) To produce a model on the transformation mechanism of *B. subtilis* incorporating the Mismatch repair proteins.

IV.3 Material and Methods

IV.3.1 Protein and DNA purification

The proteins RecA, RecO, SsbA, MutS and DNA molecules were purified and subsequently biochemically analysed by Dr. Begoña Carrasco and Dr. Ester Serrano from Prof. Juan Carlos Alonso's laboratory (Centro Nacional de Biotecnologia, Spain) following the "Material and Methods" corresponding section on Carrasco, et al. (2019).

IV.3.2 Sample preparation and image acquisition for Atomic Force Microscopy

DNA-Protein sample preparation for air image acquisition

Samples for the study of the incorporation of a circular single-strand DNA into a linear double-strand DNA were firstly prepared by Dr. Begoña Carrasco and Dr. Ester Serrano. The ssDNA, 10 mM in nucleotide concentration, was preincubated with SsbA and RecO in buffer A (50 mM Tris-HCl pH 7.5, 1 mM DTT, 50 mM NaCl, 10 mM MgOAc, 50 mg/ml BSA, 5 % glycerol) with an additional 5 mM ATP for 5 min at 37 ºC. Then, the dsDNA (20 mM in nt), RecA and MutS were added to the reaction and incubated for 60 min at 37 ºC. A (d)ATP regeneration system, 8 units/ml creatine phosphokinase and 8 mM phosphocreatine, was included in all recombination reactions. After incubation, samples were deproteinised and fractionated by 0.8% agarose gel electrophoresis with ethidium bromide [Ayora, et al. (2002)].

The purified and deproteinised DNA recombination reaction was further diluted in buffer A, 1000-fold. The diluted reaction was deposited onto freshly cleaved mica and incubated for 30 s. Then, the surface was thoroughly washed with 3 ml of Milli-Q water and dried under constant nitrogen air flow [Lyubchenko, et al. (2009)].

AFM image acquisition conditions

From chapter II.3.1.3: Images were obtained using a Nanotec Electronica S.L. Atomic Force Microscope using the amplitude modulation (*tapping*) mode unless stated otherwise. Air sample images were obtained at room temperature and humidity at a rate of 1.48 lines per second and 512 lines per image and pixels per line, 346 second per image. The cantilevers utilized were PPP-NCH (Nanosensors, Switzerland).

Image processing was performed following the method presented on chapter II.3.1.4.

IV.3.3 Control Samples

A base of control samples is key for the characterization of DNA-protein interaction processes. In this project, the Atomic Force Microscopy samples were purified and deproteinised. Thus, only final reaction products of single-strand DNA and double-strand DNA were characterized on AFM images. The reaction proteins RecA, RecO, MutS and SsbA are not present due to the deproteinisation step. As control, initial DNA reactants were characterized under the AFM by contour length, volume and molecular species analysis.

Linear double-strand DNA molecules were incubated and deposited onto a mica surface following the methods for DNA-Protein sample preparation without adding any protein. The control revealed a homogenous sample of linear polymeric rods deposited on a flat surface [Figure IV-3 A]. As expected for a stochastic population distribution with independent variables (Central limit theorem), the population of DNA contour lengths follows a Normal distribution pattern with a unique peak centred around the expected contour length [Sanchez-Sevilla, et al. (2002)]. The DNA length in base pairs can be inferred from the measured contour length. Previous studies show a relation of 3.02-3.39 Å per base-pair for dsDNA deposited with Mg^{2+} on a mica surface [Rivetti, et al. (1996); Rivetti and Codeluppi (2001); Fang, et al. (1999)]. The crystallographic length for B-DNA obtained by X-ray diffraction data is 3.4 Å per bp [Wing, et al. (1980)]. From the Atomic Force Microscopy analysis the extracted contour length for the 3276 bp DNA is 1040 nm, or around 3100 bp [Figure IV-4 A].

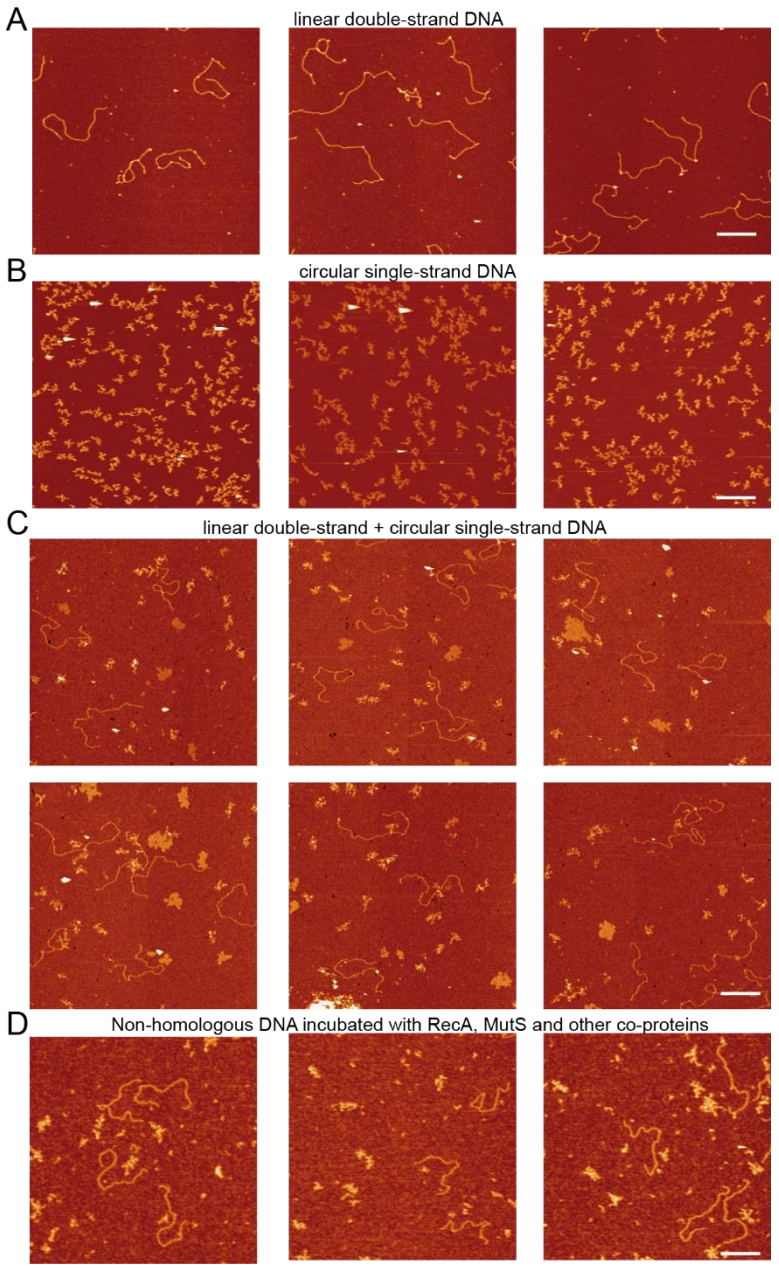


Figure IV-3. Images of Atomic Force Microscopy control samples.

A) Images of the 3276 base pairs linear dsDNA. B) Images of the 3353 base pairs circular ssDNA. C) Images of the homologous dsDNA and ssDNA sample. D) Images of the non-homologous dsDNA and ssDNA sample. Scale bar 400 nm.

Single-strand DNA circular molecules were incubated and deposited onto a mica surface following the methods for DNA-Protein sample preparation without adding any protein. The control showed a homogenous sample of condensed masses

deposited onto a mica surface [Figure IV-3 B]. Contrary to double-strand DNA, the single-strand DNA molecules are not elongated when deposited on a mica surface with Mg^{2+} as the molecule is more flexible [Hamon, et al. (2007)]. Contour length population could not be extracted from the Atomic Force Microscopy images. As solution, the volume of the ssDNA molecules is compared to the volume of a piece of dsDNA. The volume relation can be correlated to the ssDNA base-pair length [Fuentes-Perez, et al. (2013)] [Equation 3]. Applying Equation 3 to the resulting ssDNA volume showed in Figure IV-4 B results in 2900 base-pair content for a 3276 bp ssDNA.

$$\frac{V_{ssDNA}}{V_{dsDNA}} = K\ xL_{ssDNA} + C \quad \textbf{Equation 3}$$

Equation 3. Relation between the volume of a single-strand DNA molecule and the reference volume of a double-strand DNA molecule.

K=7.6±0.1 x10^{-3} nt $^{-1}$. C=0.31±0.09. L_{ssDNA} length in base-pairs of a ssDNA molecule. Equation from Fuentes-Perez, et al. (2013).

Next, a control sample was done by incubation of circular single-strand DNA with a homologous linear double-strand DNA following the methods for the DNA-protein sample preparation without adding any protein. Images acquired showed presence of linear rod-like molecules, dsDNA, and aggregated masses, ssDNA [Figure IV-3 C]. As expected, there was no interaction among both DNA species neither strand exchange events, appearance of circular dsDNA molecular.

Lastly, a control sample of non-homologous circular single-strand DNA and linear double-strand DNA was incubated following the methods for the DNA-protein sample preparation. Images obtained were similar to the other control samples [Figure IV-3 D]. As predicted, no strand exchange between non-homologous DNA molecules occurred in the presence of proteins.

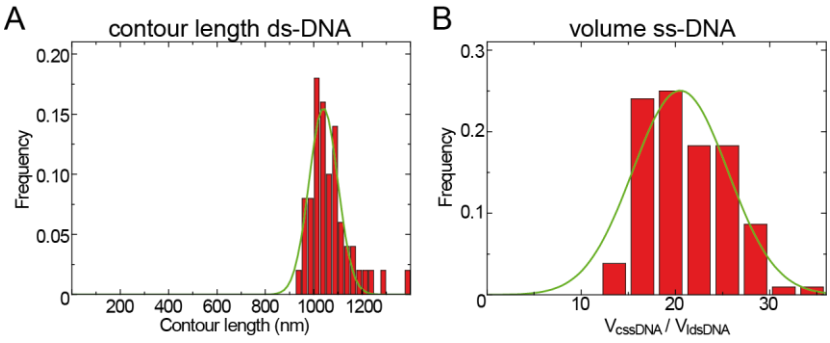


Figure IV-4. Histograms of molecule populations for the double-strand DNA and the single-strand DNA control samples.

A) Contour length histogram of dsDNA molecules centred at 1040±60 nm. B) Volume histogram of ssDNA with dsDNA volume as reference. Normal distribution centred at 21±1.

IV.4 Results

IV.4.1 A heterologous section in a single-strand DNA delays strand exchange

In absence of the Mismatch Repair proteins, RecA-mediated strand exchange regulates interspecies conjugation [Rayssiguier, Dohet et al. (1991)]. The currently establish mechanism proposes the formation of a RecA filament onto an exogenous single-strand DNA molecule. The RecA-DNA filament recognizes a homologous sequence on the organism double-strand DNA. In this model, the initial role of the protein RecA is to bind and stabilize an exogenous ssDNA piece. Then, the RecA-ssDNA filament invades a homology region of a dsDNA by promoting the opening and binding of the RecA-ssDNA to the homologous strand [Yadav, et al. (2014)].

In naturally competent bacteria, the RecA-mediated mechanism requires auxiliary proteins. The auxiliary proteins stabilize single-strand DNA intermediates and increase the binding rate and formation of the RecA filament. With sufficient time, incubation of RecA, the auxiliary proteins, circular ssDNA and linear dsDNA carry out strand exchange of homologous DNA molecules. The result of the strand exchange reaction is an accumulation of circular dsDNA, composed of a circular strand and a linear strand, and linear ssDNA molecules [Carrasco, Yadav et al. (2015)]. However, increasing sequence divergence decreases the rate of completed strand-exchange events [Carrasco, et al. (2016)]. The incomplete products are composed of an integrated part of the circular ssDNA into an otherwise linear dsDNA.

To corroborate the effect of the sequence divergence, a dsDNA with a 77 base-pair heterologous near the 3' end was produced. The 77 bp region is located 427 bp from the 3' end and contains 10 bp mismatches. Incomplete strand-exchange molecules require a 300-400 bp homology region. Hence, the 427 bp located between the 3' DNA end and the heterelogous region. For a 25 min reaction time, the Atomic Force Microscope images revealed the presence of ssDNA, linear dsDNA, circular dsDNA and intermediate molecules [Figure IV-5 A 1-4]. As *B. subtilis* RecA-mediated strand-exchange originates from the 3' or the 5' end independently, some aggregated molecules were found. The aggregated molecules are the result of two reaction events occurring on both ends of the same linear dsDNA molecule [Figure IV-5 A 5]. Nevertheless, longer reaction periods were reported to accumulate circular dsDNA molecules [Carrasco, et al. (2019)].

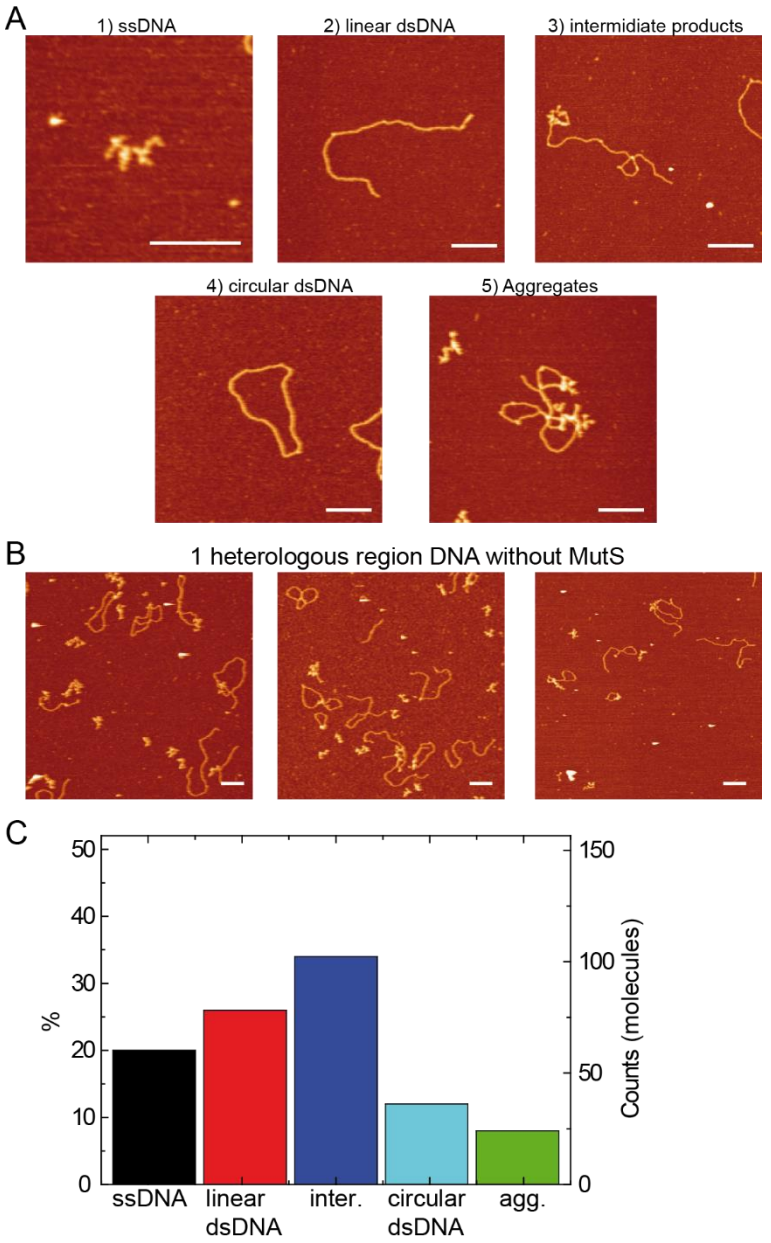


Figure IV-5. Quantification and analysis of reaction species from AFM.

A) Classification of DNA molecules from RecA-mediated strand-exchange reaction; 1) ssDNA, 2) linear dsDNA, 3) intermediate molecule with single-strand and double-strand DNA, 4) circular dsDNA and 5) DNA molecule formed with by aggregation. B) Big field-of-view images of strand-exchange reaction. C) Population graph for the 5 DNA species. AFM images scale bar 200 nm.

IV.4.2 MutS promotes strand exchange of RecA filaments in the presence of a heterologous section

In presence of the Mismatch Repair proteins, RecA-mediated strand-exchange of divergent DNA molecules is inhibited. In organisms as *E. coli*, presence of MutS and MutL proteins inhibits the integration of exogenous DNA [Worth, et al. (1994); Calmann, et al. (2005)]. MutS and MutL are early stage MMR mechanism proteins. These proteins are thought to detect mismatched regions signalling the defect for further repair.

Contrary to *E. coli* 3'-end strand invasion, RecA-mediated strand-exchange reaction initiates from the 3' or 5' DNA ends in *B. subtilis* [Carrasco, et al. (2016)]. The lack of directionality at the initiation step of the strand-exchange process in *B. subtilis* introduces an uncertainty for the Mismatch Repair proteins. Research studies on *B. subtilis* detected an influence of the Mismatch Repair proteins on the initiation of the strand-exchange mechanism of homologous DNA molecules. However, D-loop formation is not affected by neither MutS nor MutL [Tham, et al. (2016)].

The role of the "proofreading" proteins MutS and MutL remains unclear in *B. subtilis*. To understand the effect of *B. subtilis* MutS protein on the strand-exchange mechanism, the previously presented analysis was repeated in the presence of MutS. A heterologous double-strand DNA molecule was incubated with a single-strand DNA in the presence of the strand-exchange proteins and MutS. The Atomic Force Microscopy images show a strong occurrence of circular dsDNA molecules [Figure IV-6 A]. In comparison, the results from the same reaction without MutS protein reveal an increase in the amount of circular dsDNA. Without MutS, 12% of molecules are circular dsDNA. With MutS, 35% of the molecules analysed are circular dsDNA [Figure IV-6 B]. The increasing value of circular dsDNA and the associated decrease in intermediate species suggest an activation or catalysis of the strand-exchange reaction by MutS.

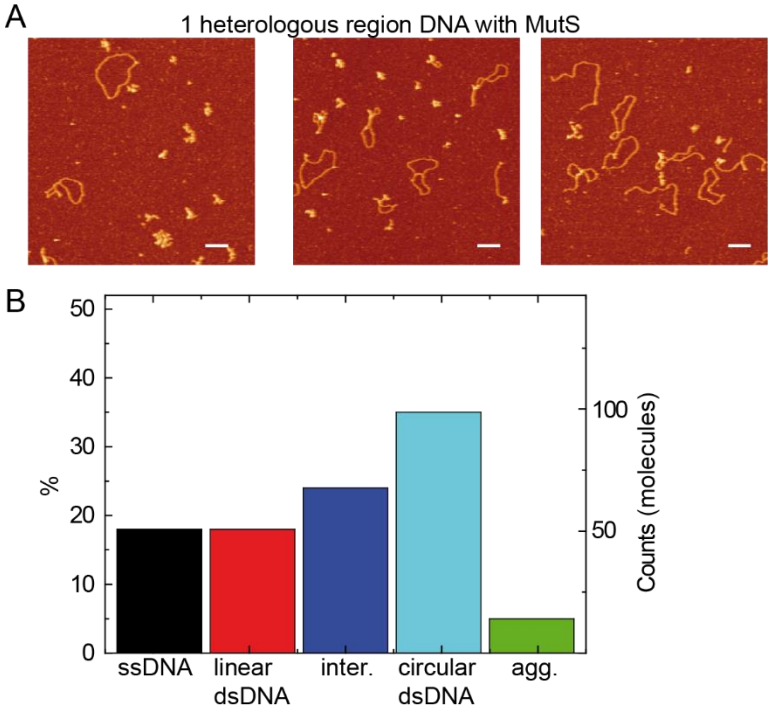


Figure IV-6. Image and quantification analysis of RecA-mediated strand-exchange with MutS protein.

A) AFM big field-of-view images of the final reaction products. B) Population graph for the 5 DNA species. AFM images scale bar 200 nm.

The hypothesis presented for the activation of the strand-exchange mechanism in presence of MutS were:

1) MutS impedes the interaction and subsequently stabilizes the RecA-DNA filaments around the heterologous section.
2) Co-purification of a nuclease protein together with MutS could act on the heterologous region removing it.
3) The MutS-DNA complex directs the RecA-DNA filament formation beyond the heterologous sequence.
4) MutS diffusion around the heterologous region might destabilize RecA-DNA filament formation on the 3' end and thus favouring the formation on the homologous 5' end.

IV.4.3 Heterologous sections on 3' and 5' ends block strand exchange

To assess the different hypothesis, a DNA molecule was purified with two heterelogous regions on both the 3' and 5'-ends. The heterology comprises two 77 base-pairs regions located at 427 and 3271 bp from the 3' end. The heterologous regions have 10 mismatches. The 5'-end heterology is located near the DNA-end impeding the formation of intermediate molecules at this position. To corroborate the correct purification and insertion of the two heterologous regions, a control sample without MutS was analysed. The Atomic Force Microscopy images reveal a decrease in circular DNA molecules. In contrast to the DNA-protein sample prepared with only one heterologous region, the percentage of circular DNA molecules decreases from 12% to 1% [Figure IV-7]. This result suggests a further decrease in strand-exchange activity in the presence of two heterologous regions.

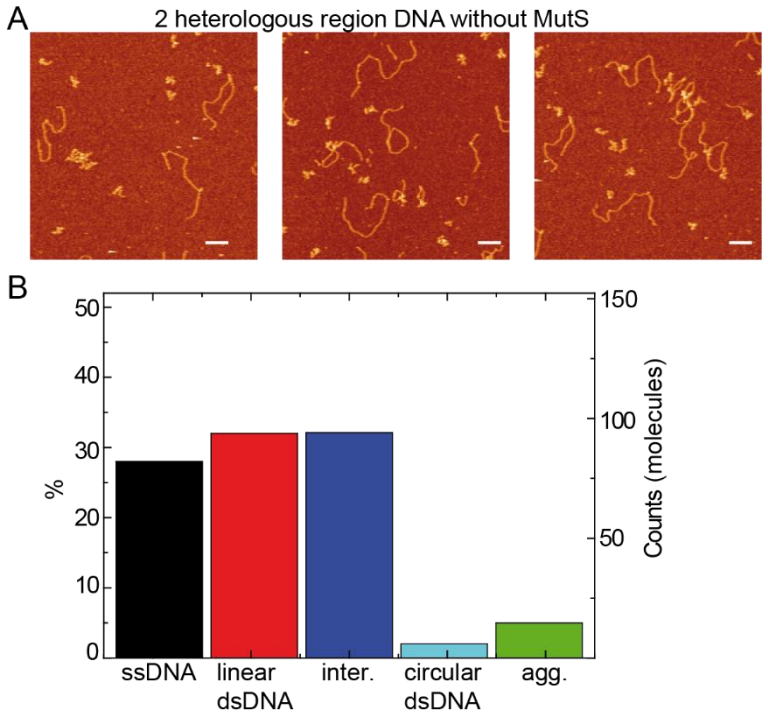


Figure IV-7. Image and quantification analysis of double heterology DNA strand-exchange. *A) AFM big field-of-view images of the final reaction products. B) Population graph for the 5 DNA species. AFM images scale bar 200 nm.*

The double heterologous region blocks the strand-exchange mechanism. However, in the presence of MutS the percentage of circular double-strand DNA molecules could increase, similar to Figure IV-6 B.

The analysis of the species population variation present in a double heterologous DNA-MutS sample compared to previous samples correlates with the different hypothesis in the following way:

1) The RecA-DNA filament would overcome the heterologous region on both the 3' and 5' ends resulting in similar levels of circular double-strand DNA molecules as in Figure IV-6 B.
2) The nuclease enzyme might act onto the 3' end, the 5' end or both resulting in an increase in either intermediate molecules or circular dsDNA molecules with respect to Figure IV-7 B.
3) RecA would bind to the single-strand DNA molecule beyond the heterelogous region with a similar effect to one presented in 1).
4) MutS diffusion on the 3' and 5' end would destabilize the formation of the RecA-DNA filaments. MutS diffusion blocks the strand-exchange reaction resulting in a similar or lower value of circular dsDNA with respect to Figure IV-7 B.

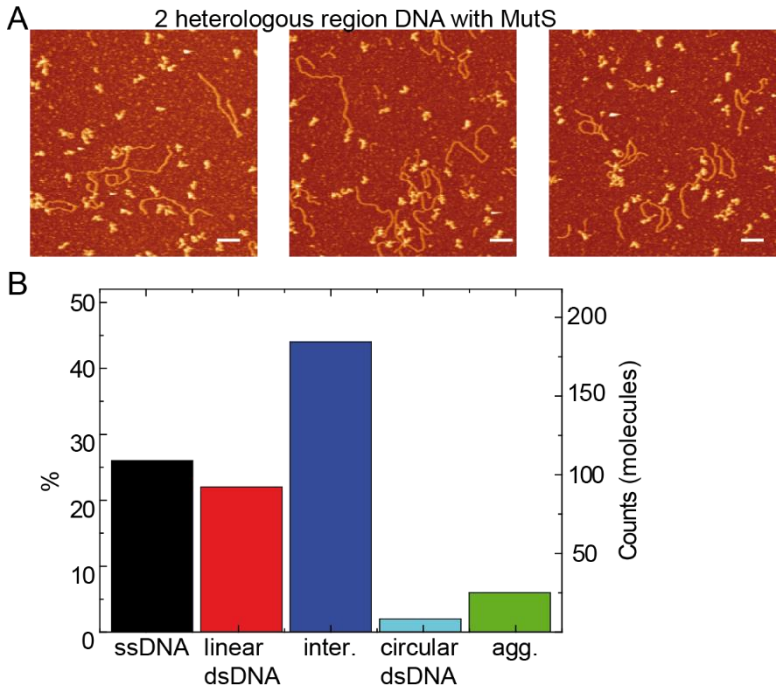


Figure IV-8. Image and quantification analysis of double heterology DNA strand-exchange with MutS.

A) AFM big field-of-view images of the final reaction products. B) Population graph for the 5 DNA species. AFM images scale bar 200 nm.

The Atomic Force Microscopy species population presented in Figure IV-8 B have a marginal difference for circular dsDNA population compared to Figure IV-7 B case. This suggests that the number of molecules that complete the strand-exchange process in these conditions is low. Therefore, double heterologous regions at the 3' and 5'-ends devoid the system of any strand-exchange activity in presence or absence of MutS. However, a slight increase in intermediate molecules is in concordance with a possible spontaneous branch migration increase.

To further validate this hypothesis, the shape and size of intermediate products was also evaluated. The intermediate molecules from the strand-exchange reaction are divided into 3 species. Early products have a joined intersection between the linear and the circular DNA of less than 427 bp, position were one heterology is located [Figure IV-9 A1]. Intermediate products have linear and circular DNA with similar occupancy [Figure IV-9 A2]. Near-final products have a circular double-strand DNA attached to a single-strand DNA [Figure IV-9 A3].

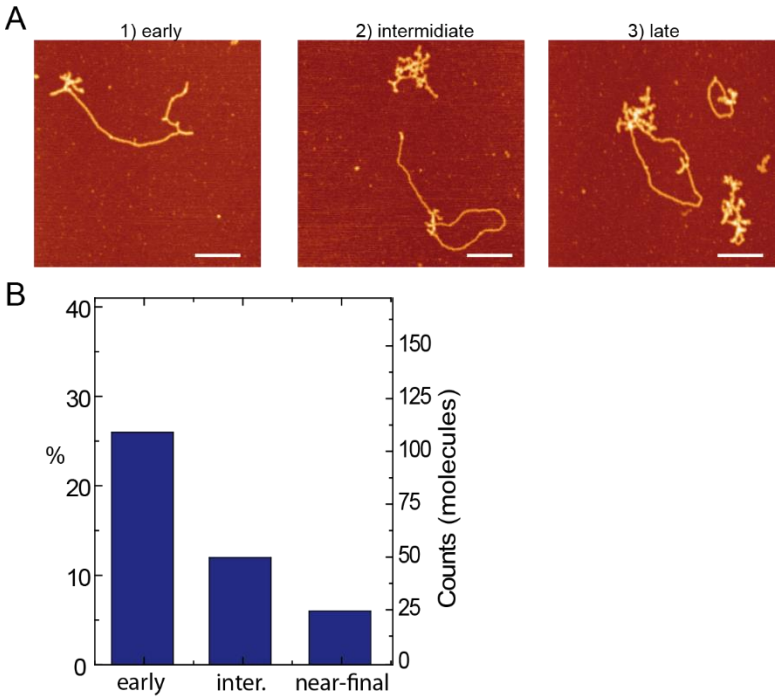


Figure IV-9. Atomic Force Microscopy images and quantification of intermediate products for the MutS-RecA reaction.

A) AFM images of intermediate products; 1) early stage linear dsDNA with less than 427 bp interaction region, 2) intermediate product with approximately half and half linear-circular dsDNA, and 3) near-final product of circular dsDNA with a piece of ssDNA bound. B) Quantification of intermediate species in a MutS RecA-mediated strand-exchange reaction. Scale bar 200 nm.

IV.5 Discussion

During this project, we have analysed the role of MutS in the strand-exchange mechanism of *B. subtilis in vitro*. RecA-mediated strand-exchange is known to have a decrease in efficiency with increasing sequence divergence. Moreover, MutS was expected to have an inhibitory role in heterologous DNA strand-exchange. In this chapter, we have observed the effect of MutS in the strand-exchange process between divergent DNA molecules with Atomic Force Microscopy. To understand the influence of MutS we have considered the final reaction products of a three-strand exchange process. The three-strand exchange reaction is performed either with a single or a double 77 bp heterologous region DNA molecule and *B. subtilis* purified proteins.

		25 min reaction percentage of species (%)				
Heterologous regions	MutS	Single-strand DNA	Linear double-strand DNA	Inter.	Circular double-strand DNA	Agg.
1 (3')	-	20	26	34	12	8
1 (3')	+	18	18	24	35	5
2 (3'+5')	-	28	32	32	2	5
2 (3'+5')	+	26	22	44	2	6

Table 3. Percentage of species quantified by Atomic Force Microscopy image.
Table representing the species quantified in percentage for the 4 different samples tested. Molecular species are: single strand-DNA, linear double-strand DNA, intermediate species, circular double-strand DNA and aggregates of molecules.

Without MutS, results reveal that a 77 bp region near the 3' end with ~16% sequence divergence delays the formation of final reaction products. In this case, a complete reaction took 60 min to reach the same recombination yield as a 20 min homologous strand-exchange. Atomic Force Microscopy images of a 25 min incubation reaction reveal a low percentage of final reaction products [Table 3]. This result suggests a possible slow formation of RecA-DNA filaments from the 5'-end with accumulation of early-stage intermediates assembled at the 3'-end.

Surprisingly, incorporation of MutS on the process increased the reaction yield in a 25 min reaction. Fast formation of final reaction products with low accumulation of intermediate molecules correlates with a modulating MutS activity around mismatched regions [Figure IV-6]. In Carrasco, et al. (2019), we have proposed that stabilization of mismatched regions by MutS destabilizes the assembly of the RecA-DNA filaments. Specifically in this situation, binding of MutS to the 77 bp heterologous region near the 3' end catalyses depolymerisation of the Rec-DNA filament at the 3'-end of the DNA molecule increasing the rate of polymerisation at the 5'-end [Figure IV-10].

To assess and confirm this hypothesis we acquired Atomic Force Microscopy images of similar strand-exchange reactions with a two 77 bp heterologous regions DNA. This DNA molecule contains mismatched regions near the 3'-end and at the 5'-end. Without MutS, the analysis of molecular species on AFM images shows low percentage of final products at 25 min [Table 3]. Biochemical analysis for 60 min reactions also show a low formation of final products. As expected, impossibility to bind through the 5' end and a small region at the 3' end blocks the invasion of the homologous region on the double-strand DNA by RecA. With MutS, the reaction yield was maintained. However, the presence of a higher percentage of intermediate species was unexpected [Table 3]. With further analysis on the intermediate species we proposed a spontaneous branch migration as solution. The assembly of, mainly, early-stage intermediate species initiates from the 3'-end and it is stabilized by interaction between the circular single-strand DNA and the linear dsDNA on a sequence of around 400 bp.

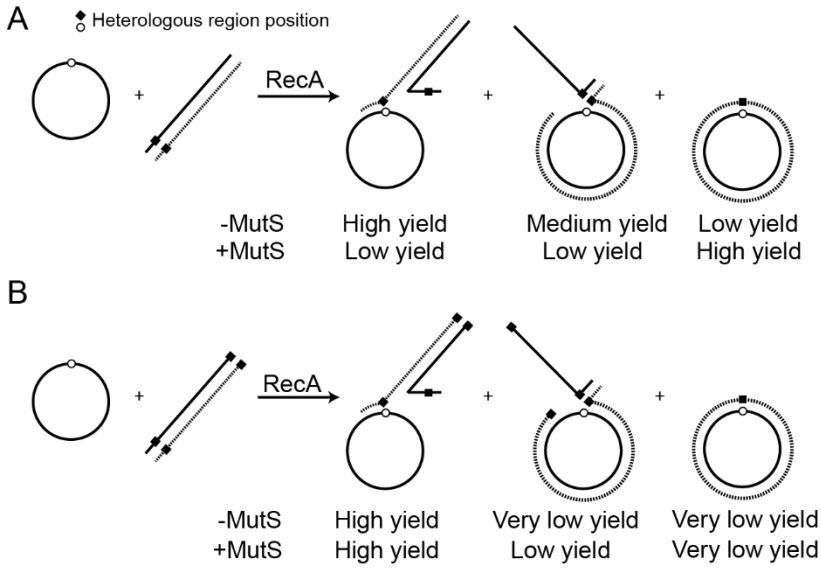


Figure IV-10. Proposed model for the RecA-mediated MutS modulated strand-exchange reaction in B. subtilis.

A) Model proposed and expected yields for the strand-exchange process of a three-strand exchange reaction with 1 heterologous region. B) Model proposed and expected yields for the strand-exchange process of a three-strand exchange reaction with 2 heterologous regions.

The Atomic Force Microscopy images, the population analysis of species and the biochemical analysis presented in Carrasco, et al. (2019) are adequate for the elaboration a model on the RecA-mediated MutS-modulated strand-exchange mechanism. In this model, a circular single-strand DNA molecule is first stabilized by the union to SsbA and RecO. The union of SsbA and RecO to the ssDNA supports the formation of a RecA-DNA filament. The RecA protein DNA binding domains along the RecA-DNA filament detect the homologous regions on a linear double-strand DNA [Kurumizaka, et al. (1996)]. The strand invasion step preferentially initiates from the 3'-end on the linear dsDNA with a lower efficiency for the 5' end. Mismatches on any of the linear DNA-ends stops the strand-exchange step. MutS modulates the assembly of the RecA-DNA filament by binding to the mismatched regions. With one mismatched region, MutS directs RecA beyond the heterology. However, two mismatched regions inhibit the reaction in the presence of MutS [Figure IV-10].

V. SUMMARY

The scientific content of this thesis is focused on the application of the single molecule technique referred as Atomic Force Microscope, however, the three biological processes presented are widely different. Nonetheless, these processes are interconnected by a strong connection in DNA modification. To ease the reader's tedium this section summarises some common or confusing points discussed through the thesis before the conclusions section.

"First, you have to finish"- Michael Schumacher, Seven-time Formula 1 World Champion with Benetton and Ferrari.

In this thesis we have applied the Atomic Force Microscope on the study of the role of some proteins involved in three biological processes related to DNA preservation and repair. In Figure V-1 the DNA molecules employed are summarised and categorized.

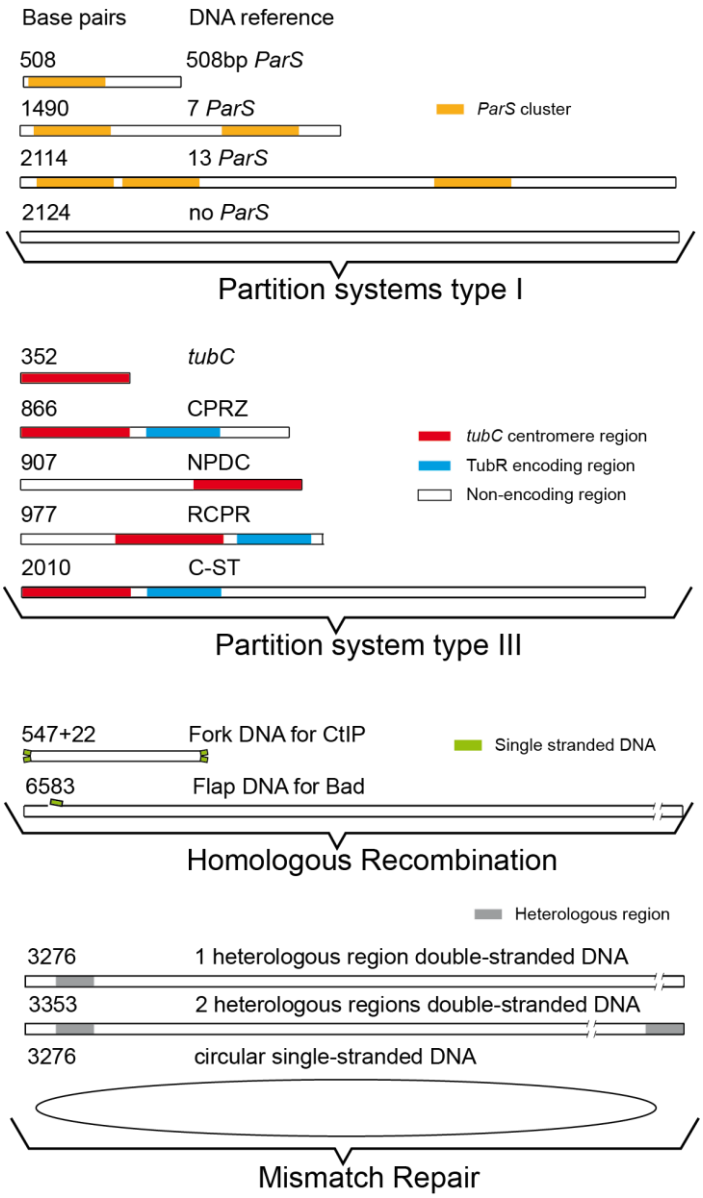


Figure V-1. DNA molecules presented in the thesis.

DNA molecules used for the Atomic Force Microscopy experiments presented in this thesis. The molecules are categorized based on the research project for which they were used. The molecules are not to scale and the regions signalled with coloured bars are estimated positions.

The partition systems are composed of a group of proteins and DNA sequences with the role of successfully segregate the DNA during cellular division. The study of the centromere-binding protein of three organisms, *B. subtilis, C. crescentus* and *C. botulinum*, revealed differences in the binding mechanism and subsequent segrosome structure required for the recognition of the DNA molecule by the motor protein. In addition, the Atomic Force Microscope was applied to study the polymerisation process of the motor protein from the *C. botulinum* organism.

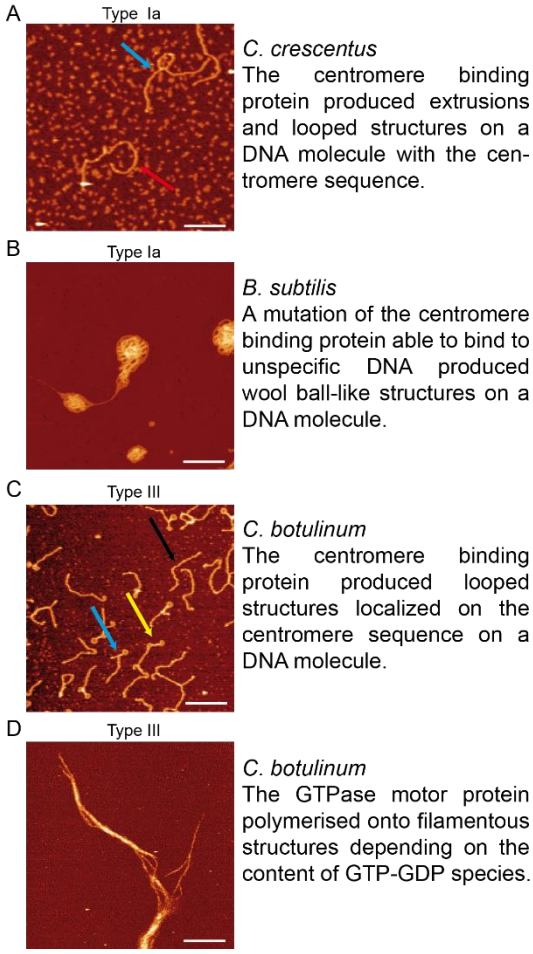


Figure V-2. Partition systems images obtained with Atomic Force Microscoopy.
Images obtained from the partition systems proteins and DNA molecules. A) Secondary structures observed in a sample of the centromere-binding protein of C. crescentus and DNA. Coloured arrows point the position of secondary structures. Scale bar 200 nm. B) Secondary structures observed in a sample of the centromere-binding protein of B. subtilis and DNA. Scale bar 500 nm. C) Secondary structures observed in a sample of the centromere-binding protein of C- botulinum and DNA. Coloured arrows point the position of secondary structures. Scale bar 200 nm. D) Filamentous structure formed by the polymerisation of the motor protein from C. botulinum organisms. Scale bar 800 nm.

The Homologous Recombination is a repair mechanism for the Double-Strand Breaks on DNA molecules. The role of the proteins responsible for the detection and signalling is thought to be crucial for a successful DNA repair. The study of the DNA-binding protein, CtIP, and helicase/nuclease protein, Dna2, from *H. sapiens* and *G. stearothermophilus* respectively shed light on the role of both protein and their mechanisms of action.

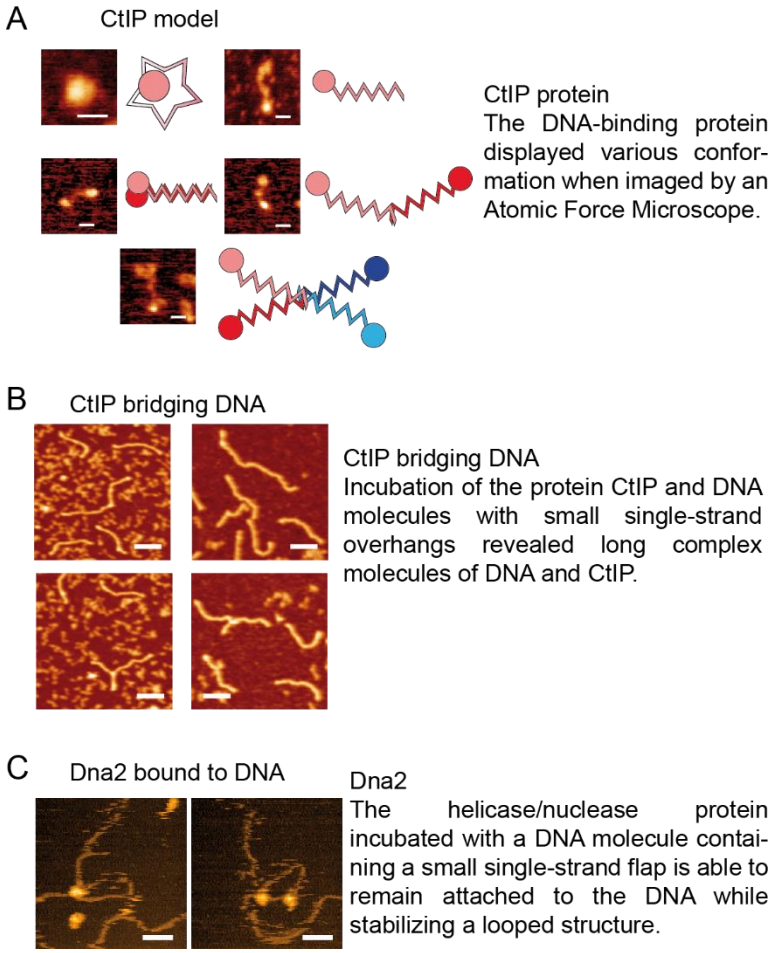


Figure V-3. Homologous Recombination images obtained with Atomic Force Microscopy. *Images obtained from the Homologous Recombination proteins CtIP and Dna2. A) Images acquired of CtIP proteins and the structural model. Scale bar 25 nm. B) Images of long DNA-CtIP complexes formed by a bridging of the DNA molecules through interaction with CtIP. Scale bar 100 nm. C) Images of Dna2 proteins tightly bound to DNA molecules and stabilizing loops possibly formed during the opening of the DNA double-helix. Scale bar 40 nm.*

The DNA mismatch repair is a mechanism for the detection and repair of erroneous bases inserted in a DNA molecule. However, the proteins of this mechanism are thought to be involved in the regulation of the mutation rate of an organism. The presence or absence of the protein MutS from *B. subtilis* in a DNA-strand exchange reaction regulated the insertion of the exogenous single-stranded DNA into the double-stranded DNA molecule.

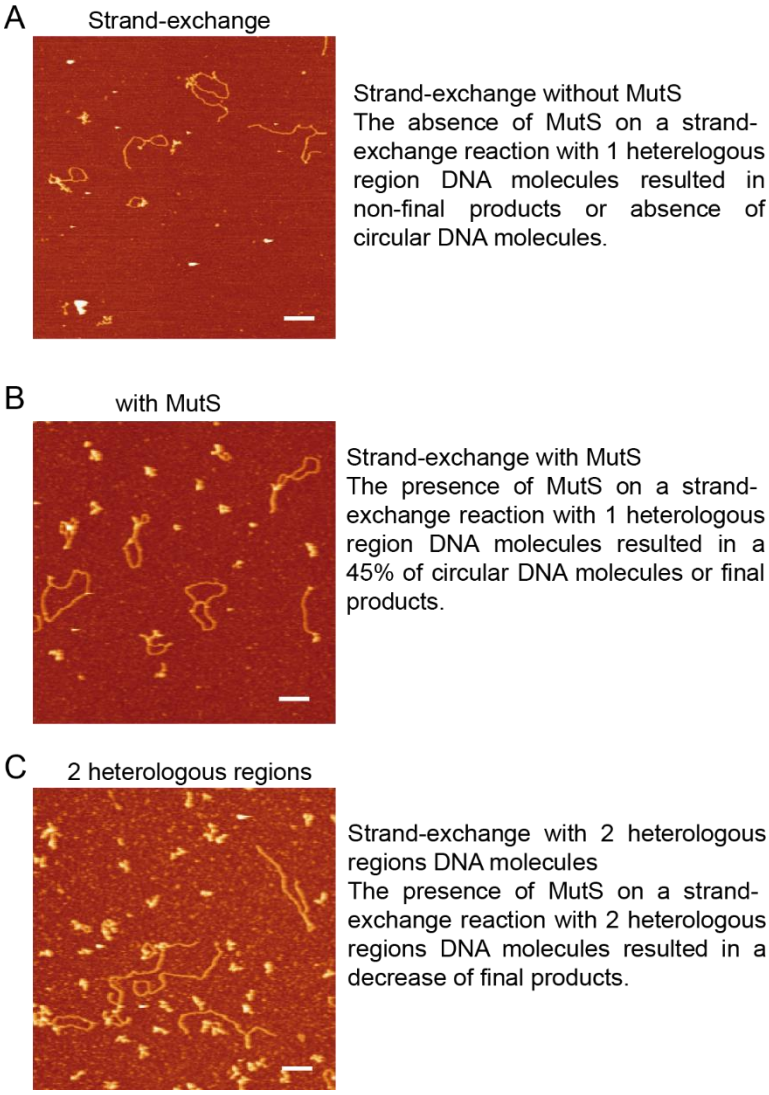


Figure V-4. DNA mismatch repair images obtained with Atomic Force Microscopy.
Images obtained of a strand-exchange reaction with Atomic Force Microscope. A) Non-circular DNA molecules product of the incubation of a linear double-stranded and a circular single-strand DNA molecule in presence of RecA and other co-proteins. Scale bar 200 nm. B) Similar reaction to A) in presence of MutS. Scale bar 200 nm. C) Similar reaction to B) but with more heterologous DNAs. Scale bar 200 nm.

VI. CONCLUSIONS

In this thesis I have applied the Atomic Force Microscope on the study of proteins involved in three biological processes of DNA preservation and repair. The processes are plasmid partitioning, homologous recombination and DNA mismatch repair. All the samples prepared and data acquired throughout the process of conclusion of this thesis were analysed in the best found conditions for the preservation and acquisition of reliable information. The main conclusions derived from the material presented are divided according to the three processes.

"一寸先は闇" – Japanese proverb used for reflecting the unkwnon character of the time to come.

Plasmid Partition Systems:

The plasmid partition projects were centered on the study of the interaction of a Centromere-Binding Protein of various organism with DNA molecules. The formation of a stable protein-DNA complex, segrosome, is described as being involved in the correct organisation and segregation of DNA during cellular division by subsequent recognition and binding to a motor-protein. In this study, we have characterized the formation of the segrosome complexes of *B. subtilis*, *C. crescentus* and *C. botulinum* by Atomic Force Microscope.

The summarized conclusions of the second chapter are:

- *B. subtilis* Centromere-Binding Protein, ParB, non-specific binding to DNA results in "wool ball-like" DNA condensed molecules. The organisation of DNA molecules by the interaction with ParB is not affected by the presence or absence of the centromere sequence. However, the mutation of a key amino-acid, R80A, responsible for protein dimerization impedes the DNA condensation.

- *C. crescentus* ParB interacts with DNA molecules forming secondary structures as a result. These secondary structures appear in DNA molecules with *B. subtilis* centromere sequence, *parS*. The results suggest a formation of a filament or a closed structure of DNA by protein multimerization similar to, but less active than, *B. subtilis* ParB.

- *C. botulinum* TubR interacts with DNA by bending the molecule into a one or two ring-like structures. The size and position of these structures is correlated to the location of the *C. botulinum* centromere sequence, *tubC*, and possibly the promoter region of TubR.

- *C. botulinum* motor-protein, TubZ, forms tubulin-like filament structures. The necessity of hydrolysable GTP molecules in solution and the role of the C-terminal domain of the protein revealed the polymerisation-depolymerisation mechanism of this protein. The TubZ filaments are stabilized by GTP-caps at the ends. Hydrolysis of the GTP lead to depolymerisation. This mechanism of polymer formation is kwon as "treadmilling".

- Additionally, the size and shape of the filaments suggest a direct interaction of the motor-protein with the ring-like structures of TubR with DNA. The structure formed by the union of the segrosome, TubR-DNA, with the filaments, TubZ, is responsible for the partitioning of plasmidic material inside the cell.

Homologous Recombination:

CtIP and Dna2 are two proteins involved in the initiation steps of the Homologous Recombination repair mechanism of DNA Double-Strand Breaks. The step consists of a short-resection of the damaged DNA-ends by the MRN complex in concordance with CtIP and a long-resection by Dna2. However, the mechanism of resection and cooperation between proteins was not known. Here we have studied the domain organisation and role of the *Homo sapiens* CtIP and of *Geobacillus stearothermophilus* Dna2 with Atomic Force Microscope.

The summarized conclusions of the third chapter are:

- Human protein CtIP forms a stable tetramer in solution with a "Dumbbell" shape. The characterization of CtIP by Atomic Force Microscope reveal the formation of monomers, dimers and tetramers with distinct shapes.
- Atomic Force Microscopy molecular species of CtIP displayed a flexible organisation with variable sizes. Further analysis with High-Speed Atomic Force Microscopy confirmed the flexibility of the "tail" domains with relation to the globular domains. Under High-Speed, the molecular species were seen as interchangeable species in some cases.
- Incubation of Human CtIP with non-canonical DNA molecules reveal the formation of DNA-protein complex structures. The presence of Wild-type or dephosphorylated CtIP increased the percentage of complex molecules while mutations on the C- or N-terminal domains decreased them.
- The DNA-CtIP complex structures are in concordance with DNA-joined molecules by the C-terminal domains of a CtIP tetramer. The formation and stabilization of the joined DNA-ends is a key step in the short-resection reaction.
- *Geobacillus stearothermophilus* Dna2 is a helicase/nuclease protein involved in the long-resection reaction of Homologous Recombination. The helicase activity of Dna2 may be regulated by the binding of Dna2 to both DNA-strands. The Dna2 proteins tightly bound to DNA molecules with looped structures nearby suggest that the rehybridisation of the single-stranded DNA happened but the protein did not detach from the DNA molecule possibly by being bound to both DNA strands.

DNA Mismatch Repair:

The “proofreading” mechanism of the DNA molecules in cellular organism is known for the detection and repair of mismatch lesions. However, natural competent bacteria can assimilate exogenous DNA molecules from the media surrounding them. The regulation and activity of the “proofreading” mechanisms in these cases is not totally understood. In this project we have studied the effect of the protein MutS of the Mismatch repair mechanism in the regulation of DNA assimilation in *B. subtilis*.

The summarized conclusions of the fourth chapter are:

- The formation of a stable complex formed of single-strand DNA and *B. subtilis* RecA is responsible for the assimilation of the exogenous single-strand DNA. The complex detects a homologous region around the double-strand DNA molecule of the bacteria and exchanges one of the strands. The resulting molecule has base mismatches at heterologous regions.
- The presence of *B. subtilis* MutS at the moment of the strand-exchange reaction destabilizes the formation of the RecA-ssDNA and dsDNA D-loop at the heterologous region. However, the presence of a minor heterologous region does not impede the integration of the exogenous ssDNA from other homologous region.
- In the presence of *B. subtilis* MutS and various heterologous regions the formation of the RecA-ssDNA filament is still disturbed at the heterologous regions. Thus, the strand-exchange reaction is blocked by the presence of MutS at various heterologous regions by the impossibility to form a D-loop at a homologous position between the RecA-ssDNA comples and the dsDNA.

VII. CONCLUSIONES

En este trabajo de tesis he tenido la oportunidad de estudiar tres procesos biológicos empleando el microscopio de fuerzas atómicas. Los procesos biológicos estudiados se centran en la reparación (recombinación homóloga y reparación de errores de emparejamiento) y mantenimiento de ADN (sistemas de partición, dentro de una célula). Todas las muestras preparadas y datos obtenidos durante este trabajo de tesis han sido analizados en las mejores condiciones para la conservación de las moléculas involucradas. Las conclusiones mostradas a continuación están divididas en concordancia con los tres procesos estudiados.

Sistemas de partición de plásmidos:

Los proyectos de investigación realizados en los sistemas de partición se han centrado en el estudio de la proteína de unión a región centromerica de cada de ellos. La formación de un complejo estable de ADN y proteína, segrosoma, es un paso clave en el proceso de organización y segregación del ADN durante la división celular. Este complejo es reconocido por la proteína motora y actúa como centro de segregación del ADN. En este capítulo he estudiado la formación del segrosoma en los organismos *B. subtilis*, *C. crescentus* y *C. botulinum* mediante un microscopio de fuerzas atómicas.

El resumen de las conclusiones obtenidas en el segundo capítulo de esta tesis es:

- La unión no especifica de la proteína de unión a región centromerica de *B. subtilis*, ParB, a una molécula de ADN da como resultados estructuras complejas con forma de "bola de lana". La formación de estas estructuras de interacción ADN-proteína es independiente de la presencia de la secuencia centromerica. Sin embargo, la mutación en el aminoácido 80, R80A, crucial para la formación de dímeros de proteína inhibe la condensación de ADN.

- ParB del organismo *C. crescentus* interacciona con moléculas de ADN formando estructuras secundarias como resultado. Estas estructuras aparecen con mayor frecuencia en moléculas de ADN con la región centromerica de *B. subtilis*. Estos resultados sugieren la formación de un filamento o estructura cerrada de ADN por multimerización de la proteína alrededor del ADN, similar a pero menos condensado que las estructuras formadas por ParB de *B. subtilis*.

- TubR de *C. botulinum* interacciona con ADN uniéndose y doblando la molécula en una o dos estructuras con forma de anillo. El tamaño y position de estas estructuras se correlaciona con la posición de la región centromerica de este organismo y posiblemente con la región promotora del gen de TubR.

- La proteína motora de *C. botulinum,* TubZ, forma estructuras poliméricas con forma filamentosa parecida a los filamentos de tubulina. El estudio del hidrólisis de GTP por el extremo C-terminal de la proteína revelo el proceso de formación de estos filamentos. Los polímeros están estabilizados por la presencia de tapones de GTP. Una vez que el GTP se hidroliza el filamento se desestabiliza y despolimeriza. Este mecanismo es conocido como "treadmilling".

- Por último, el tamaño y forma de los filamentos de TubZ se correlacionan con el tamaño y forma de las estructuras con forma de anillo formadas por TubR en una molécula de ADN. En conjunto, los datos sugieren una interacción directa de los filamentos de TubZ con los anillos de TubR-DNA responsables por la segregación de la molécula de ADN durante la división celular.

Recombinación Homóloga:

CtIP y Dna2 son dos proteínas involucradas en el inicio del mecanismo de reparación de roturas de doble-hebra de ADN conocido como Recombinación Homóloga. La Recombinación Homóloga empieza por la resección corta de los finales dañados por el complejo de proteínas MRN en cooperación con CtIP y, después, por la resección larga de los finales dañados por Dna2. El rol especifico de las proteínas en estos pasos no está del todo desvelado. En esta tesis he estudiado la conformación y rol de la proteína CtIP de *Homo sapiens* y el rol de la proteína Dna2 de *Geobacillus stearothermophilus* mediante la Microscopia de Fuerzas Atómicas.

El resumen de las conclusiones obtenidas en este capítulo es:

- La protein CtIP humana forma un tetramero estable en solución con forma de "Dumbbell". Las imágenes del Microscopio de Fuerzas Atómicas revelaron las conformaciones de otras formas oligomericas como los monómeros y los dímeros.
- Las especies moleculares de CtIP mostraron una gran flexibilidad con tamaño variable. Análisis de estas especies con un Microscopio de Fuerzas Atómicas de alta velocidad confirmo la flexibilidad de la proteína formada por dominios globulares unidos por estructuras flexibles.
- La unión de la proteína a ADN con extremos no-canónicos mostro la formación de estructuras complejas de ADN y proteína. Con la proteína salvaje o la variante defosforilada se aumentó la presencia de los complejos de ADN proteína comparado con las variantes mutadas en los extremos C- o N-.
- Los complejos de ADN y proteína sugieren una unión de la proteína por el extremo C- un ADN. El tetrámero de una proteína, por lo tanto, puede unir moléculas de ADN uniéndose por extremos C- a diferentes moléculas. La formación y estabilización de estas uniones de ADN es un paso clave de al Recombinación Homóloga.
- Dna2 de *Geobacillus stearothermophilus* es una helicasa/nucleasa involucrada en la resección larga de finales de ADN dañados. La actividad helicasa de Dna2 puede estar regulada por la unión de Dna2 a las dos hebras de una molécula de ADN. La visualización de moléculas de Dna2 fuertemente unidas a moléculas de doble-hebra de ADN y la presencia de estructuras circulares sugiere una unión de la proteína a las dos hebras del ADN.

Reparación de desajustes del AND:

El mecanismo de "corrección" de errores en organismos celulares está encargado de la detección y reparación de desajustes de la doble-hebra de ADN. Células naturalmente competentes pueden incorporar ADN exógeno a su genoma desde el medio que la envuelve. La regulación y actividad del mecanismo de corrección en estos casos no está del todo desvelado. En este proyecto se ha estudiado el rol de la proteína MutS de *B.subtilis* en la asimilación de ADN exógeno.

El resumen de las conclusiones obtenidas en este capítulo es:

- La formación de un complejo estable de ADN y RecA de *B. subtilis* es responsable por la asimilación del ADN exógeno de cadena sencilla. Este complejo detecta regiones de homología en un ADN de hebra doble e intercambia una hebra por la exógena. La molécula resultada tiene errores en las regiones desiguales.
- La presencia de MutS de *B. subtilis* durante el proceso antes descrito desestabiliza la unión del complejo RecA-AND al ADN de hebra doble en zonas desiguales. La presencia de regiones desiguales pequeñas no impide el intercambio completo de hebras.
- En presencia de varias regiones desiguales la uniñon del complejo RecA-ADN exogeno al ADN de hebra doble se ve perturbada y puede llegar a bloquear la integración del ADN exogeno.

VIII. FUTURE PERSPECTIVES

The research work performed in this thesis encompasses in three biological processes. Although, some of these studies were a continuation of from previous research projects performed at Fernando Moreno's laboratory (*Geobacillus* Dna2, *B.subtilis* ParB, *C. botulinum* TubZ), others were the beginning of new research lines (*Homo sapiens* CtIP, *C.crescentus* ParB, *B. subtilis* MutS, *C. botulinum* TubR). Thus, the research presented in this dissertation is not finished and may be completed from future research projects.

"Your future hasn't been writen yet, noones has, your future is whatever you make it. So make it a good one."- Dr. Emmet Brown, built a time-machine into a DMC Delorean

The research projects of *Geobacillus* Dna2, *B. subtilis* ParB and *C. botulinum* TubZ were continued from previous research studies performed by Dr. Carolina Carrasco, Dr. Cesar Lopez-Pastrana and Dr. María Eugenia Fuentes-Perez. The knowledge acquired by them was of importance for adapting the projects to the study with Atomic Force Microscope. Specially, *B. subtilis* ParB was my first project with the Atomic Force Microscope and other biochemical techniques such as EMSA and ATPase assays. The knowledge passed from Dr. Cesar Lopez and Dr. Gemma Fisher to myself paved the way for the plasmid partition system studies and served as a biological learning platform. The knowledge provided in this thesis may serve as platform for the continuation or initiation of other research studies or projects.

The Atomic Force Microscope is a proven single-molecule technique. However, the future of the technique must follow an ever-evolving state-of-the-art. As described in the Introduction chapter, the full potential of the Atomic Force Microscopes has not been yet attained. To complement the studies about biological processes showed in the thesis, a combination of the Atomic Force Microscopy with other Microscopy techniques with fluorescence or spectroscopy capabilities is the way to go for the identification of specific proteins in complex molecular complexes. As an alternative, to observe intermediate steps or fast reaction mechanisms of biological processes is to adapt the recently used Microscopes for faster acquisition time so to study these dynamic reactions.

The research projects of *Homo sapiens* CtIP, *C. crescentus* ParB, *B. subtilis* MutS, *C. botulinum* TubR are not completely finished. These biological processes are a complex set of reactions performed by specific groups of proteins. For CtIP, the future of the project is to study the role of CtIP when combined with the MRN complex in the short-resection or with Dna2 in the long-resection of DNA broken molecules. For *C. crescentus* ParB, the project can be complemented with a more in depth study on the C-terminal disparity with *B. subtilis* ParB and its role in DNA condensation. Lastly, the *C. botulinum* TubR was a complement to the study done for *C. botulinum* TubZ. However, the type III partition system has and additional co-protein, TubY, thought to be a regulatory protein of TubZ.

IX. ARTICLES PUBLISHED

During the period of research of this thesis work, part of the scientific results obtained were published in peer-reviewed scientific journals (Scientific Reports, Journal of Molecular Biology, Nucleic Acids Research, eLife and Frontiers in Microbiology) or are under reviewing conditions for a posterior publication with the following references:

M. E. Fuentes-Perez, R. Nunez-Ramirez, **A. Martin-Gonzalez**, D. Juan-Rodriguez, O. Llorca, F. Moreno-Herrero and M. A. Oliva (2017). "TubZ filament assembly dynamics requires the flexible C-terminal tail." Sci Rep 7: 43342.

R. Arroyo, **A. Martin-Gonzalez**, M. Echaide, A. Jain, W. H. Brondyk, J. Rosenbaum, F. Moreno-Herrero and J. Perez-Gil (2018). "Supramolecular Assembly of Human Pulmonary Surfactant Protein SP-D." J Mol Biol 430(10): 1495-1509.

B. Martin-Garcia*, **A. Martin-Gonzalez***, C. Carrasco, A. M. Hernandez-Arriaga, R. Ruiz-Quero, R. Diaz-Orejas, C. Aicart-Ramos, F. Moreno-Herrero and M. A. Oliva (2018). "The TubR-centromere complex adopts a double-ring segrosome structure in Type III partition systems." Nucleic Acids Res 46(11): 5704-5716.

O. J. Wilkinson, **A. Martin-Gonzalez**, H. Kang, S. J. Northall, D. B. Wigley, F. Moreno-Herrero and M. S. Dillingham (2019). "CtIP forms a tetrameric dumbbell-shaped particle which bridges complex DNA end structures for double-strand break repair." Elife 8.

B. Carrasco, E. Serrano, **A. Martín-González**, F. Moreno-Herrero and J. C. Alonso (2019). "Bacillus subtilis MutS Modulates RecA-Mediated DNA Strand Exchange Between Divergent DNA Sequences." Frontiers in Microbiology 10(237).

A. Marín-González, C.L. Pastrana, R. Bocanegra, **A. Martín-González**, J.G. Vilhena, R. Pérez, B. Ibarra, C. Aicart-Ramos, F. Moreno-Herrero. In-phase A-tracts modulate the multiscale mechanical properties of DNA. (In preparation).

*The authors of this paper wish it to be known that the first two authors should be regarded as first joint authors.

ACKNOWLEDGEMENTS

Hasta llegar a la escritura y defensa de una tesis, he tenido que aprender y nutrirme del ambiente que me rodea para desarrollar una investigación de calidad. En esta sección quiero agradecer a todas las personas que de alguna manera han formado parte de este proceso de aprendizaje. Nombrar a todas y cada una de las personas involucradas, sería una tarea casi tan ardua como escribir otra tesis. Por esto, si me olvido de alguien, por favor siéntase representado y tenga en cuenta que le estoy agradecido a cualquiera que leyera esta tesis en cualquier momento.

En primer lugar, deseo agradecerle todo la formación y trabajo realizado a mi director de tesis, Fernando Moreno Herrero. Gracias a él he podido llevar a cabo esta investigación de calidad y con unos medios y conocimientos al alcance de pocos. Además, su filosofía de trabajo y liderazgo son un ejemplo para cualquier investigador.

Después de a mi director de tesis me gustaría agradecerle a mi tutor de la tesis, Julio Gómez Herrero, hombre de ideas y maestro de la Microscopía de Fuerzas Atómicas. Su ayuda como contacto con la Universidad Autónoma de Madrid y disponibilidad para hablar sobre el AFM han sido de gran ayuda.

A mis mentores en la Microscopía de Fuerzas Atómicas, María Eugenia Fuentes y Pablo Ares, también les quiero agradecer su paciencia y enseñanzas. Sin ellos el AFM hubiera sido una montaña muy difícil de escalar.

A los colaboradores que nos han proporcionado información y material sobre diversos mecanismos biológicos:

-Mark Simon Dillingham, Professor at the University of Bristol, for the meetings and scientific talks shared. He has been one of the best collaborators backed by an amazing group of scientists (Gemma Fisher, Oliver Wilkinson, Sarah Northall). The CtIP, Dna2 and B. subtilis ParB projects are the result of this collaboration.

-María Ángela Oliva, investigadora del Centro de Investigaciones Biológicas, por su entusiasmo en el estudio del sistema de partición tipo III y por dejarnos estudiar sus proteínas TubR y TubZ.

-Juan Carlos Alonso, investigador del Centro Nacional de Biotecnología, y su grupo por su colaboración en el proyecto de estudio de las proteínas RecA, MutS y MutL.

-Tung Le, researcher at the John Innes Center, and his group for its support with the C. crescentus ParB project.

-Jesús Pérez Gil, Decano de la facultad de biología de la Universidad Complutense de Madrid, por ofrecerme la oportunidad de participar y enseñar a Raquel Arroyo a usar el AFM como parte de su tesis doctoral sobre la proteína humana SP-D.

-金沢大学と Nano-LSI の安藤先生に感謝したい。金沢大学で安藤先が私に HS−AFM を三ヶ月間使用する機会を与えてくれました。

A mis compañeros de laboratorio en el CNB les quiero agradecer su paciencia, conocimientos, tiempo y amistad. A mis vecinas de mesa y/o mamis científicas, Adriana Gil y María Teresa Arranz, en disposición similar al cerebro humano; a la izquierda el lado lógico y de la razón y a la derecha el lado creativo y artístico. A las expertas en pinzas y AFM que tanto me han ayudado cuando me quedaba atascado con el microscopio, Carolina Carrasco y Silvia Hormeño. A la bioquímica y encargada del "wet lab", Clara Aicart, por la fabricación de variadas moléculas de ADN y por su ayuda para entender todos los procesos biológicos que aquí están presentados. A mis compañeros de fechorías, Dr. Cesar, Dra. Julene, Alberto, Mikel y Roberta les tengo que agradecer las incontables horas de discusiones científicas y experiencias vividas (Como casi morirme en un scape room de un chungazo). De cada uno me llevo una pequeña parte para mí.

Por último, agradecer a los no-científicamente involucrados. A mis padres, Ernesto y María Teresa por su apoyo, educación y enseñanzas durante tantos años. A mi familia: abuelos (los cuatro), tíos (los diez), primos (los ocho) y sobrina (solo una); nombraros a todos llevaría un montón de tiempo así que daos por nombrados. No podía olvidar al último, Diego, mi hermano.

Y por otro lado a mis amigos ya sean de España, Hong Kong, Irlanda, Japón y a la cuadrilla en Vitoria. Ya tiene suficiente tela el aguantarme como para que me ponga a escribir ñoñerías sobre cada uno así que gracias por vuestra presencia.

X. REFERENCES

The knowledge required for the writing of the thesis and for the understanding of the many different mechanism and molecules involved was acquired through the reading and comprehension of many scientific articles, books, seminars and discussions with other scientists. In the references section, some part of this knowledge is reflected as articles, papers or book chapters that are directly mentioned in the text written in the thesis.

Abbe, E. (1883) XV.—The Relation of Aperture and Power in the Microscope (continued)*, *Journal of the Royal Microscopical Society*, **3**, 790-812.

Abeles, A.L., Friedman, S.A. and Austin, S.J. (1985) Partition of unit-copy miniplasmids to daughter cells. III. The DNA sequence and functional organization of the P1 partition region, *Journal of molecular biology*, **185**, 261-272.

Actis, L.A., Tolmasky, M.E. and Crosa, J.H. (1999) Bacterial plasmids: replication of extrachromosomal genetic elements encoding resistance to antimicrobial compounds, *Frontiers in bioscience : a journal and virtual library*, **4**, D43-62.

Adelberg, E.A. and Pittard, J. (1965) Chromosome Transfer in Bacterial Conjugation, *Bacteriological reviews*, **29**, 161-172.

Adhikary, A.*, et al.* (2012) Direct formation of the C5'-radical in the sugar-phosphate backbone of DNA by high-energy radiation, *The journal of physical chemistry. B*, **116**, 5900-5906.

Albrecht, T.R.*, et al.* (1991) Frequency modulation detection using high-Q cantilevers for enhanced force microscope sensitivity, *Journal of Applied Physics*, **69**, 668-673.

Alushin, G.M.*, et al.* (2014) High-resolution microtubule structures reveal the structural transitions in alphabeta-tubulin upon GTP hydrolysis, *Cell*, **157**, 1117-1129.

Ambrose, E.J. (1956) A surface contact microscope for the study of cell movements, *Nature*, **178**, 1194.

Amer, N.M. (1983) NEW APPROACHES TO PHOTOTHERMAL SPECTROSCOPY, *J. Phys. Colloques*, **44**, C6-185-C186-190.

Amer, N.M., Skumanich, A. and Ripple, D. (1986) Photothermal modulation of the gap distance in scanning tunneling microscopy, *Applied Physics Letters*, **49**, 137-139.

Anand, R.*, et al.* (2016) Phosphorylated CtIP Functions as a Co-factor of the MRE11-RAD50-NBS1 Endonuclease in DNA End Resection, *Molecular cell*, **64**, 940-950.

Andersson, J.O. (2005) Lateral gene transfer in eukaryotes, *Cellular and molecular life sciences : CMLS*, **62**, 1182-1197.

Ando, T. (2018) High-speed atomic force microscopy and its future prospects, *Biophysical reviews*, **10**, 285-292.

Ando, T.*, et al.* (2001) A high-speed atomic force microscope for studying biological macromolecules, *Proceedings of the National Academy of Sciences of the United States of America*, **98**, 12468-12472.

Ando, T.*, et al.* (2008) High-speed AFM and nano-visualization of biomolecular processes, *Pflugers Archiv : European journal of physiology*, **456**, 211-225.

Andres, S.N.*, et al.* (2015) Tetrameric Ctp1 coordinates DNA binding and DNA bridging in DNA double-strand-break repair, *Nature structural & molecular biology*, **22**, 158-166.

Andres, S.N. and Williams, R.S. (2017) CtIP/Ctp1/Sae2, molecular form fit for function, *DNA repair*, **56**, 109-117.

Anselmetti, D.*, et al.* (1994) Attractive-mode imaging of biological materials with dynamic force microscopy, *Nanotechnology*, **5**, 87-94.

Aparicio, T.*, et al.* (2016) MRN, CtIP, and BRCA1 mediate repair of topoisomerase II-DNA adducts, *The Journal of cell biology*, **212**, 399-408.

Ares, P.*, et al.* (2016) High resolution atomic force microscopy of double-stranded RNA, *Nanoscale*, **8**, 11818-11826.

Argueso, J.L.*, et al.* (2008) Double-strand breaks associated with repetitive DNA can reshape the genome, *Proceedings of the National Academy of Sciences of the United States of America*, **105**, 11845-11850.

Avery, O.T., Macleod, C.M. and McCarty, M. (1944) Studies on the Chemical Nature of the Substance Inducing Transformation of Pneumococcal Types : Induction of Transformation by a Desoxyribonucleic Acid Fraction Isolated from Pneumococcus Type Iii, *The Journal of experimental medicine*, **79**, 137-158.

Aylett, C.H. and Lowe, J. (2012) Superstructure of the centromeric complex of TubZRC plasmid partitioning systems, *Proceedings of the National Academy of Sciences of the United States of America*, **109**, 16522-16527.

Aylett, C.H.*, et al.* (2010) Filament structure of bacterial tubulin homologue TubZ, *Proceedings of the National Academy of Sciences of the United States of America*, **107**, 19766-19771.

Ayora, S.*, et al.* (2002) Homologous-pairing activity of the Bacillus subtilis bacteriophage SPP1 replication protein G35P, *The Journal of biological chemistry*, **277**, 35969-35979.

Badrinarayanan, A., Le, T.B. and Laub, M.T. (2015) Bacterial chromosome organization and segregation, *Annual review of cell and developmental biology*, **31**, 171-199.

Bae, S.H.*, et al.* (2001) RPA governs endonuclease switching during processing of Okazaki fragments in eukaryotes, *Nature*, **412**, 456-461.

Balakrishnan, L.*, et al.* (2010) Acetylation of Dna2 endonuclease/helicase and flap endonuclease 1 by p300 promotes DNA stability by creating long flap intermediates, *The Journal of biological chemistry*, **285**, 4398-4404.

Balzer, D.*, et al.* (1992) KorB protein of promiscuous plasmid RP4 recognizes inverted sequence repetitions in regions essential for conjugative plasmid transfer, *Nucleic Acids Res*, **20**, 1851-1858.

Barash, J.R. and Arnon, S.S. (2014) A novel strain of Clostridium botulinum that produces type B and type H botulinum toxins, *The Journal of infectious diseases*, **209**, 183-191.

Barrett, R.C. and Quate, C.F. (1991) High-speed, large-scale imaging with the atomic force microscope, *Journal of Vacuum Science & Technology B: Microelectronics and Nanometer Structures Processing, Measurement, and Phenomena*, **9**, 302-306.

Beach, H.*, et al.* (2005) Conservation of μs−ms Enzyme Motions in the Apo- and Substrate-Mimicked State, *Journal of the American Chemical Society*, **127**, 9167-9176.

Beck, M.*, et al.* (2011) The quantitative proteome of a human cell line, *Molecular systems biology*, **7**, 549.

Bennett, P.M. (2008) Plasmid encoded antibiotic resistance: acquisition and transfer of antibiotic resistance genes in bacteria, *British journal of pharmacology*, **153 Suppl 1**, S347-357.

Bertoli, C., Skotheim, J.M. and de Bruin, R.A. (2013) Control of cell cycle transcription during G1 and S phases, *Nature reviews. Molecular cell biology*, **14**, 518-528.

Betzig, E.*, et al.* (2006) Imaging intracellular fluorescent proteins at nanometer resolution, *Science*, **313**, 1642-1645.

Bezanilla, M.*, et al.* (1994) Motion and enzymatic degradation of DNA in the atomic force microscope, *Biophysical journal*, **67**, 2454-2459.

Bezanilla, M.*, et al.* (1995) Adsorption of DNA to Mica, Silylated Mica, and Minerals: Characterization by Atomic Force Microscopy, *Langmuir : the ACS journal of surfaces and colloids*, **11**, 655-659.

Biehs, R.*, et al.* (2017) DNA Double-Strand Break Resection Occurs during Non-homologous End Joining in G1 but Is Distinct from Resection during Homologous Recombination, *Molecular cell*, **65**, 671-684 e675.

Binnig, G.*, et al.* (1987) Atomic resolution with atomic force microscope, *Surface Science*, **189-190**, 1-6.

Binnig, G., Quate, C.F. and Gerber, C. (1986) Atomic force microscope, *Physical review letters*, **56**, 930-933.

Binnig, G. and Rohrer, H. (1983) Scanning tunneling microscopy, *Surface Science*, **126**, 236-244.

Buey, R.M., Diaz, J.F. and Andreu, J.M. (2006) The nucleotide switch of tubulin and microtubule assembly: a polymerization-driven structural change, *Biochemistry*, **45**, 5933-5938.

Burkholder, G.D. and Weaver, M.G. (1977) DNA-protein interactions and chromosome banding, *Experimental cell research*, **110**, 251-262.

Bustamante, C.*, et al.* (2000) Single-molecule studies of DNA mechanics, *Current opinion in structural biology*, **10**, 279-285.

Calmann, M.A., Evans, J.E. and Marinus, M.G. (2005) MutS inhibits RecA-mediated strand transfer with methylated DNA substrates, *Nucleic Acids Res*, **33**, 3591-3597.

Cannavo, E. and Cejka, P. (2014) Sae2 promotes dsDNA endonuclease activity within Mre11-Rad50-Xrs2 to resect DNA breaks, *Nature*, **514**, 122-125.

Carrasco, B.*, et al.* (2019) Bacillus subtilis MutS Modulates RecA-Mediated DNA Strand Exchange Between Divergent DNA Sequences, *Frontiers in microbiology*, **10**.

Carrasco, B.*, et al.* (2016) Chromosomal transformation in Bacillus subtilis is a non-polar recombination reaction, *Nucleic Acids Res*, **44**, 2754-2768.

Carrasco, C.*, et al.* (2008) Manipulation of the mechanical properties of a virus by protein engineering, *Proceedings of the National Academy of Sciences of the United States of America*, **105**, 4150-4155.

Casuso, I.*, et al.* (2012) Characterization of the motion of membrane proteins using high-speed atomic force microscopy, *Nature nanotechnology*, **7**, 525-529.

Cejka, P. (2015) DNA End Resection: Nucleases Team Up with the Right Partners to Initiate Homologous Recombination, *The Journal of biological chemistry*, **290**, 22931-22938.

Ciprandi, G.*, et al.* (1986) In vitro effects of Bacillus subtilis on the immune response, *Chemioterapia : international journal of the Mediterranean Society of Chemotherapy*, **5**, 404-407.

Cleveland, J.P.*, et al.* (1998) Energy dissipation in tapping-mode atomic force microscopy, *Applied Physics Letters*, **72**, 2613-2615.

Colom, A.*, et al.* (2012) High-speed atomic force microscopy: cooperative adhesion and dynamic equilibrium of junctional microdomain membrane proteins, *Journal of molecular biology*, **423**, 249-256.

Cox, M.M. (1993) Relating biochemistry to biology: how the recombinational repair function of RecA protein is manifested in its molecular properties, *BioEssays : news and reviews in molecular, cellular and developmental biology*, **15**, 617-623.

Creighton, H.B. and McClintock, B. (1931) A Correlation of Cytological and Genetical Crossing-Over in Zea Mays, *Proceedings of the National Academy of Sciences of the United States of America*, **17**, 492-497.

Croucher, N.J.*, et al.* (2009) Role of conjugative elements in the evolution of the multidrug-resistant pandemic clone Streptococcus pneumoniaeSpain23F ST81, *Journal of bacteriology*, **191**, 1480-1489.

Chen, B.W.*, et al.* (2015) Insights into ParB spreading from the complex structure of Spo0J and parS, *Proceedings of the National Academy of Sciences of the United States of America*, **112**, 6613-6618.

Chen, Z., Yang, H. and Pavletich, N.P. (2008) Mechanism of homologous recombination from the RecA-ssDNA/dsDNA structures, *Nature*, **453**, 489-484.

Chumley, F.G., Menzel, R. and Roth, J.R. (1979) Hfr formation directed by tn10, *Genetics*, **91**, 639-655.

Daley, J.M.*, et al.* (2017) Enhancement of BLM-DNA2-Mediated Long-Range DNA End Resection by CtIP, *Cell reports*, **21**, 324-332.

Davies, O.R.*, et al.* (2015) CtIP tetramer assembly is required for DNA-end resection and repair, *Nature structural & molecular biology*, **22**, 150-157.

Davis, M.A. and Austin, S.J. (1988) Recognition of the P1 plasmid centromere analog involves binding of the ParB protein and is modified by a specific host factor, *The EMBO journal*, **7**, 1881-1888.

de Jager, M.*, et al.* (2001) Human Rad50/Mre11 is a flexible complex that can tether DNA ends, *Molecular cell*, **8**, 1129-1135.

de la Torre, B.*, et al.* (2016) Atomic-Scale Variations of the Mechanical Response of 2D Materials Detected by Noncontact Atomic Force Microscopy, *Physical review letters*, **116**, 245502.

de Pablo, P.J.*, et al.* (1998) Jumping mode scanning force microscopy, *Applied Physics Letters*, **73**, 3300-3302.

De Vlaminck, I.*, et al.* (2012) Mechanism of homology recognition in DNA recombination from dual-molecule experiments, *Molecular cell*, **46**, 616-624.

Denamur, E.*, et al.* (2000) Evolutionary implications of the frequent horizontal transfer of mismatch repair genes, *Cell*, **103**, 711-721.

Denning, W.*, et al.* (2013) Optimization of the transductional efficiency of lentiviral vectors: effect of sera and polycations, *Molecular biotechnology*, **53**, 308-314.

Dillingham, M.S. and Kowalczykowski, S.C. (2008) RecBCD enzyme and the repair of double-stranded DNA breaks, *Microbiology and molecular biology reviews : MMBR*, **72**, 642-671, Table of Contents.

Doolittle, W.F. (1999) Lateral genomics, *Trends in cell biology*, **9**, M5-8.

Drake, B.*, et al.* (1989) Imaging crystals, polymers, and processes in water with the atomic force microscope, *Science*, **243**, 1586-1589.

Dubnau, D. (1991) Genetic competence in Bacillus subtilis, *Microbiological reviews*, **55**, 395-424.

Duxin, J.P.*, et al.* (2009) Human Dna2 is a nuclear and mitochondrial DNA maintenance protein, *Molecular and cellular biology*, **29**, 4274-4282.

Eeftens, J.M.*, et al.* (2016) Condensin Smc2-Smc4 Dimers Are Flexible and Dynamic, *Cell reports*, **14**, 1813-1818.

Fang, Y., Spisz, T.S. and Hoh, J.H. (1999) Ethanol-induced structural transitions of DNA on mica, *Nucleic Acids Res*, **27**, 1943-1949.

Ferretti, L.P.*, et al.* (2016) Cullin3-KLHL15 ubiquitin ligase mediates CtIP protein turnover to fine-tune DNA-end resection, *Nature communications*, **7**, 12628.

Fisher, G.L.*, et al.* (2017) The structural basis for dynamic DNA binding and bridging interactions which condense the bacterial centromere, *eLife*, **6**.

Fokin, S.I. (2013) Otto Bütschli (1848–1920): Where we will genuflect?, *Protistology*, **8**, 22-35.

Fonseca Guerra, C.*, et al.* (2000) Hydrogen Bonding in DNA Base Pairs: Reconciliation of Theory and Experiment, *Journal of the American Chemical Society*, **122**, 4117-4128.

Forment, J.V., Jackson, S.P. and Pellegrini, L. (2015) When two is not enough: a CtIP tetramer is required for DNA repair by Homologous Recombination, *Nucleus*, **6**, 344-348.

Friedman, S.A. and Austin, S.J. (1988) The P1 plasmid-partition system synthesizes two essential proteins from an autoregulated operon, *Plasmid*, **19**, 103-112.

Fuentes-Perez, M.E., Dillingham, M.S. and Moreno-Herrero, F. (2013) AFM volumetric methods for the characterization of proteins and nucleic acids, *Methods*, **60**, 113-121.

Fuentes-Perez, M.E.*, et al.* (2012) Using DNA as a fiducial marker to study SMC complex interactions with the atomic force microscope, *Biophysical journal*, **102**, 839-848.

Fuentes-Perez, M.E.*, et al.* (2017) TubZ filament assembly dynamics requires the flexible C-terminal tail, *Scientific reports*, **7**, 43342.

Fukuma, T.*, et al.* (2005) True molecular resolution in liquid by frequency-modulation atomic force microscopy, *Applied Physics Letters*, **86**, 193108.

Galletto, R.*, et al.* (2006) Direct observation of individual RecA filaments assembling on single DNA molecules, *Nature*, **443**, 875-878.

Gates, K.S. (2009) An overview of chemical processes that damage cellular DNA: spontaneous hydrolysis, alkylation, and reactions with radicals, *Chemical research in toxicology*, **22**, 1747-1760.

Gayathri, P.*, et al.* (2012) A bipolar spindle of antiparallel ParM filaments drives bacterial plasmid segregation, *Science*, **338**, 1334-1337.

Gerdes, K., Howard, M. and Szardenings, F. (2010) Pushing and pulling in prokaryotic DNA segregation, *Cell*, **141**, 927-942.

Gerdes, K., Moller-Jensen, J. and Bugge Jensen, R. (2000) Plasmid and chromosome partitioning: surprises from phylogeny, *Molecular microbiology*, **37**, 455-466.

Giessibl, F.J. (2019) The qPlus sensor, a powerful core for the atomic force microscope, *The Review of scientific instruments*, **90**, 011101.

Gimeno, A.*, et al.* (2015) 'Flatten plus': a recent implementation in WSxM for biological research, *Bioinformatics*, **31**, 2918-2920.

Graham, T.G.*, et al.* (2014) ParB spreading requires DNA bridging, *Genes & development*, **28**, 1228-1238.

Greger, R. (1991) [Nobel Prize for Medicine and Physiology 1991. Analysis of the function of single ion channel], *Deutsche medizinische Wochenschrift*, **116**, 1849-1851.

Griffith, F. (1928) The Significance of Pneumococcal Types, *The Journal of hygiene*, **27**, 113-159.

Guillaume-Gentil, O.*, et al.* (2014) Force-controlled manipulation of single cells: from AFM to FluidFM, *Trends in biotechnology*, **32**, 381-388.

Guirouilh-Barbat, J.*, et al.* (2014) Is homologous recombination really an error-free process?, *Frontiers in genetics*, **5**, 175.

Gumpp, H.*, et al.* (2009) Ultrastable combined atomic force and total internal reflection fluorescence microscope [corrected], *The Review of scientific instruments*, **80**, 063704.

Gunning, P.W.*, et al.* (2015) The evolution of compositionally and functionally distinct actin filaments, *Journal of cell science*, **128**, 2009-2019.

Guthold, M.*, et al.* (1994) Following the assembly of RNA polymerase-DNA complexes in aqueous solutions with the scanning force microscope, *Proceedings of the National Academy of Sciences of the United States of America*, **91**, 12927-12931.

Guynet, C. and de la Cruz, F. (2011) Plasmid segregation without partition, *Mobile genetic elements*, **1**, 236-241.

Haberle, W.*, et al.* (1992) In situ investigations of single living cells infected by viruses, *Ultramicroscopy*, **42-44 (Pt B)**, 1161-1167.

Haga, H.*, et al.* (2000) Elasticity mapping of living fibroblasts by AFM and immunofluorescence observation of the cytoskeleton, *Ultramicroscopy*, **82**, 253-258.

Hamilton, W.D., Axelrod, R. and Tanese, R. (1990) Sexual reproduction as an adaptation to resist parasites (a review), *Proceedings of the National Academy of Sciences of the United States of America*, **87**, 3566-3573.

Hamon, L.*, et al.* (2007) High-resolution AFM imaging of single-stranded DNA-binding (SSB) protein--DNA complexes, *Nucleic Acids Res*, **35**, e58.

Hanawalt, P.C.*, et al.* (1979) DNA repair in bacteria and mammalian cells, *Annual review of biochemistry*, **48**, 783-836.

Hansma, H.G. (2001) SURFACE BIOLOGY OF DNA BY ATOMIC FORCE MICROSCOPY, *Annual review of physical chemistry*, **52**, 71-92.

Hansma, P.K.*, et al.* (1994) Tapping mode atomic force microscopy in liquids, *Applied Physics Letters*, **64**, 1738-1740.

Hayes, F. and Barilla, D. (2006) The bacterial segrosome: a dynamic nucleoprotein machine for DNA trafficking and segregation, *Nature reviews. Microbiology*, **4**, 133-143.

Hecht, G.B. and Newton, A. (1995) Identification of a novel response regulator required for the swarmer-to-stalked-cell transition in Caulobacter crescentus, *Journal of bacteriology*, **177**, 6223-6229.

Heinemann, J.A. and Sprague, G.F., Jr. (1989) Bacterial conjugative plasmids mobilize DNA transfer between bacteria and yeast, *Nature*, **340**, 205-209.

Hell, S.W. and Wichmann, J. (1994) Breaking the diffraction resolution limit by stimulated emission: stimulated-emission-depletion fluorescence microscopy, *Optics letters*, **19**, 780-782.

Hellman, L.M. and Fried, M.G. (2007) Electrophoretic mobility shift assay (EMSA) for detecting protein-nucleic acid interactions, *Nature protocols*, **2**, 1849-1861.

Hinterdorfer, P. and Dufrêne, Y.F. (2006) Detection and localization of single molecular recognition events using atomic force microscopy, *Nature methods*, **3**, 347.

Hoeijmakers, J.H. (2009) DNA damage, aging, and cancer, *The New England journal of medicine*, **361**, 1475-1485.

Horcas, I.*, et al.* (2007) WSXM: a software for scanning probe microscopy and a tool for nanotechnology, *The Review of scientific instruments*, **78**, 013705.

Huguenel, E.D. and Newton, A. (1982) Localization of Surface Structures During Procaryotic Differentiation: Role of Cell Division in Caulobacter crescentus, *Differentiation*, **21**, 71-78.

Hutter, J.L. and Bechhoefer, J. (1993) Manipulation of van der Waals forces to improve image resolution in atomic-force microscopy, *Journal of Applied Physics*, **73**, 4123-4129.

Hwang, L.C.*, et al.* (2013) ParA-mediated plasmid partition driven by protein pattern self-organization, *The EMBO journal*, **32**, 1238-1249.

Ikeda, H. and Tomizawa, J.I. (1965) Transducing fragments in generalized transduction by phage P1. II. Association of DNA and protein in the fragments, *Journal of molecular biology*, **14**, 110-119.

Jackson, S.P. (2002) Sensing and repairing DNA double-strand breaks, *Carcinogenesis*, **23**, 687-696.

Jackson, S.P. and Bartek, J. (2009) The DNA-damage response in human biology and disease, *Nature*, **461**, 1071-1078.

Jayasena, V.K. and Johnston, B.H. (1993) Complement-stabilized D-loop. RecA-catalyzed stable pairing of linear DNA molecules at internal sites, *Journal of molecular biology*, **230**, 1015-1024.

Johnson, R.E.*, et al.* (2000) Fidelity of human DNA polymerase eta, *The Journal of biological chemistry*, **275**, 7447-7450.

Johnston, C.*, et al.* (2014) Bacterial transformation: distribution, shared mechanisms and divergent control, *Nature reviews. Microbiology*, **12**, 181-196.

Kaufmann, R., Hagen, C. and Grunewald, K. (2014) Fluorescence cryo-microscopy: current challenges and prospects, *Current opinion in chemical biology*, **20**, 86-91.

Kim, J.H.*, et al.* (2006) Isolation of human Dna2 endonuclease and characterization of its enzymatic properties, *Nucleic Acids Res*, **34**, 1854-1864.

Kim, Y.T. and Bard, A.J. (1992) Imaging and etching of self-assembled n-octadecanethiol layers on gold with the scanning tunneling microscope, *Langmuir : the ACS journal of surfaces and colloids*, **8**, 1096-1102.

Klar, T.A., Engel, E. and Hell, S.W. (2001) Breaking Abbe's diffraction resolution limit in fluorescence microscopy with stimulated emission depletion beams of various shapes, *Physical review. E, Statistical, nonlinear, and soft matter physics*, **64**, 066613.

Klug, A. (1968) Rosalind Franklin and the discovery of the structure of DNA, *Nature*, **219**, 808-810 passim.

Knoll, M. and Ruska, E. (1932) Das Elektronenmikroskop, *Zeitschrift für Physik*, **78**, 318-339.

Kodera, N., Sakashita, M. and Ando, T. (2006) Dynamic proportional-integral-differential controller for high-speed atomic force microscopy, *Review of Scientific Instruments*, **77**, 083704.

Kodera, N.*, et al.* (2010) Video imaging of walking myosin V by high-speed atomic force microscopy, *Nature*, **468**, 72-76.

Kodera, N., Yamashita, H. and Ando, T. (2005) Active damping of the scanner for high-speed atomic force microscopy, *Review of Scientific Instruments*, **76**, 053708.

Kokavecz, J.*, et al.* (2006) Novel amplitude and frequency demodulation algorithm for a virtual dynamic atomic force microscope, *Nanotechnology*, **17**, S173-177.

Kourkoutis, L.F., Plitzko, J.M. and Baumeister, W. (2012) Electron Microscopy of Biological Materials at the Nanometer Scale, *Annual Review of Materials Research*, **42**, 33-58.

Kovermann, M.*, et al.* (2015) Structural basis for catalytically restrictive dynamics of a high-energy enzyme state, *Nature communications*, **6**, 7644.

Kowalczykowski, S.C.*, et al.* (1994) Biochemistry of homologous recombination in Escherichia coli, *Microbiological reviews*, **58**, 401-465.

Kunkel, T.A. and Erie, D.A. (2005) DNA mismatch repair, *Annual review of biochemistry*, **74**, 681-710.

Kuzminov, A. (1999) Recombinational repair of DNA damage in Escherichia coli and bacteriophage lambda, *Microbiology and molecular biology reviews : MMBR*, **63**, 751-813, table of contents.

Lahue, R.S., Au, K.G. and Modrich, P. (1989) DNA mismatch correction in a defined system, *Science*, **245**, 160-164.

Lan, R. and Reeves, P.R. (1996) Gene transfer is a major factor in bacterial evolution, *Molecular biology and evolution*, **13**, 47-55.

Lederberg, J. (1952) Cell genetics and hereditary symbiosis, *Physiological reviews*, **32**, 403-430.

Lengsfeld, B.M.*, et al.* (2007) Sae2 is an endonuclease that processes hairpin DNA cooperatively with the Mre11/Rad50/Xrs2 complex, *Molecular cell*, **28**, 638-651.

Levikova, M.*, et al.* (2013) Nuclease activity of Saccharomyces cerevisiae Dna2 inhibits its potent DNA helicase activity, *Proceedings of the National Academy of Sciences of the United States of America*, **110**, E1992-2001.

Lewis, P.J. and Errington, J. (1997) Direct evidence for active segregation of oriC regions of the Bacillus subtilis chromosome and co-localization with the SpoOJ partitioning protein, *Molecular microbiology*, **25**, 945-954.

Liao, S., Toczylowski, T. and Yan, H. (2008) Identification of the Xenopus DNA2 protein as a major nuclease for the 5'->3' strand-specific processing of DNA ends, *Nucleic Acids Res*, **36**, 6091-6100.

Lindahl, T. (1993) Instability and decay of the primary structure of DNA, *Nature*, **362**, 709-715.

Luo, K.*, et al.* (2016) A phosphorylation-deubiquitination cascade regulates the BRCA2-RAD51 axis in homologous recombination, *Genes & development*, **30**, 2581-2595.

Luzzietti, N.*, et al.* (2011) Efficient preparation of internally modified single-molecule constructs using nicking enzymes, *Nucleic Acids Res*, **39**, e15.

Lyubchenko, Y.L., Shlyakhtenko, L.S. and Gall, A.A. (2009) Atomic force microscopy imaging and probing of DNA, proteins, and protein DNA complexes: silatrane surface chemistry, *Methods in molecular biology*, **543**, 337-351.

Madariaga-Marcos, J.*, et al.* (2018) Force determination in lateral magnetic tweezers combined with TIRF microscopy, *Nanoscale*, **10**, 4579-4590.

Madariaga-Marcos, J.*, et al.* (2019) ParB dynamics and the critical role of the CTD in DNA condensation unveiled by combined force-fluorescence measurements, *eLife*, **8**.

Majewski, J. and Cohan, F.M. (1999) DNA sequence similarity requirements for interspecific recombination in Bacillus, *Genetics*, **153**, 1525-1533.

Makharashvili, N.*, et al.* (2014) Catalytic and noncatalytic roles of the CtIP endonuclease in double-strand break end resection, *Molecular cell*, **54**, 1022-1033.

Marti, O., Drake, B. and Hansma, P.K. (1987) Atomic force microscopy of liquid-covered surfaces: Atomic resolution images, *Applied Physics Letters*, **51**, 484-486.

Martin-Garcia, B.*, et al.* (2018) The TubR-centromere complex adopts a double-ring segrosome structure in Type III partition systems, *Nucleic Acids Res*, **46**, 5704-5716.

Martín-García, B.*, et al.* (2018) The TubR–centromere complex adopts a double-ring segrosome structure in Type III partition systems, *Nucleic Acids Research*, **-**, - gky370.

Martin, K.A., Friedman, S.A. and Austin, S.J. (1987) Partition site of the P1 plasmid, *Proceedings of the National Academy of Sciences of the United States of America*, **84**, 8544-8547.

Martin, Y., Williams, C.C. and Wickramasinghe, H.K. (1987) Atomic force microscope–force mapping and profiling on a sub 100-Å scale, *Journal of Applied Physics*, **61**, 4723-4729.

Mendel, G. (1865) Versuche über Pflanzen-hybriden, ***Verhandlungen des naturforschenden Vereines in Brünn IV Band***, 3- 47.

Meyer, G. and Amer, N.M. (1988) Novel optical approach to atomic force microscopy, *Applied Physics Letters*, **53**, 1045-1047.

Modrich, P. (2016) Mechanisms in E. coli and Human Mismatch Repair (Nobel Lecture), *Angewandte Chemie*, **55**, 8490-8501.

Montabana, E.A. and Agard, D.A. (2014) Bacterial tubulin TubZ-Bt transitions between a two-stranded intermediate and a four-stranded filament upon GTP hydrolysis, *Proceedings of the National Academy of Sciences of the United States of America*, **111**, 3407-3412.

Montevil, M.*, et al.* (2016) Theoretical principles for biology: Variation, *Progress in biophysics and molecular biology*, **122**, 36-50.

Moreno-Herrero, F., Colchero, J. and Baro, A.M. (2003) DNA height in scanning force microscopy, *Ultramicroscopy*, **96**, 167-174.

Moreno-Herrero, F.*, et al.* (2004) Atomic force microscopy contact, tapping, and jumping modes for imaging biological samples in liquids, *Physical review. E, Statistical, nonlinear, and soft matter physics*, **69**, 031915.

Morgan, T.H. (1911) Random Segregation Versus Coupling in Mendelian Inheritance, *Science*, **34**, 384.

Morita, S.*, et al.* (1996) Contact and non-contact mode imaging by atomic force microscopy, *Thin Solid Films*, **273**, 138-142.

Mossio, M., Montevil, M. and Longo, G. (2016) Theoretical principles for biology: Organization, *Progress in biophysics and molecular biology*, **122**, 24-35.

Mott, M.L. and Berger, J.M. (2007) DNA replication initiation: mechanisms and regulation in bacteria, *Nature reviews. Microbiology*, **5**, 343-354.

Muller, D.J. and Engel, A. (2007) Atomic force microscopy and spectroscopy of native membrane proteins, *Nature protocols*, **2**, 2191-2197.

NASA (1998) Strategic program plan for space radiation health research.

Neher, E., Sakmann, B. and Steinbach, J.H. (1978) The extracellular patch clamp: a method for resolving currents through individual open channels in biological membranes, *Pflugers Archiv : European journal of physiology*, **375**, 219-228.

Nelson, K.E.*, et al.* (1999) Evidence for lateral gene transfer between Archaea and bacteria from genome sequence of Thermotoga maritima, *Nature*, **399**, 323-329.

Nevo, E. (2001) Evolution of genome-phenome diversity under environmental stress, *Proceedings of the National Academy of Sciences of the United States of America*, **98**, 6233-6240.

Ni, L.*, et al.* (2010) Plasmid protein TubR uses a distinct mode of HTH-DNA binding and recruits the prokaryotic tubulin homolog TubZ to effect DNA partition, *Proceedings of the National Academy of Sciences of the United States of America*, **107**, 11763-11768.

Nierman, W.C.*, et al.* (2001) Complete genome sequence of Caulobacter crescentus, *Proceedings of the National Academy of Sciences of the United States of America*, **98**, 4136-4141.

O'Farrell, P.H. (1975) High resolution two-dimensional electrophoresis of proteins, *The Journal of biological chemistry*, **250**, 4007-4021.

Obataya, I.*, et al.* (2005) Mechanical sensing of the penetration of various nanoneedles into a living cell using atomic force microscopy, *Biosensors and Bioelectronics*, **20**, 1652-1655.

Ode, H.*, et al.* (2012) Molecular dynamics simulation in virus research, *Frontiers in microbiology*, **3**, 258.

Ohnishi, T., Mori, E. and Takahashi, A. (2009) DNA double-strand breaks: their production, recognition, and repair in eukaryotes, *Mutation research*, **669**, 8-12.

Oliva, M.A. (2016) Segrosome Complex Formation during DNA Trafficking in Bacterial Cell Division, *Frontiers in molecular biosciences*, **3**, 51.

Oliva, M.A.*, et al.* (2012) Tubulin homolog TubZ in a phage-encoded partition system, *Proceedings of the National Academy of Sciences of the United States of America*, **109**, 7711-7716.

Ou, J.T. and Anderson, T.F. (1970) Role of pili in bacterial conjugation, *Journal of bacteriology*, **102**, 648-654.

Pascual, J.I.*, et al.* (1993) Quantum contact in gold nanostructures by scanning tunneling microscopy, *Physical review letters*, **71**, 1852-1855.

Paudyal, S.C.*, et al.* (2017) Dna2 initiates resection at clean DNA double-strand breaks, *Nucleic Acids Res*, **45**, 11766-11781.

Paull, T.T. and Gellert, M. (1999) Nbs1 potentiates ATP-driven DNA unwinding and endonuclease cleavage by the Mre11/Rad50 complex, *Genes & development*, **13**, 1276-1288.

Pavlendova, N., Muchova, K. and Barak, I. (2007) Chromosome segregation in Bacillus subtilis, *Folia microbiologica*, **52**, 563-572.

Peck, M.W. (2009) Biology and genomic analysis of Clostridium botulinum, *Advances in microbial physiology*, **55**, 183-265, 320.

Pethica, J.B. and Oliver, W.C. (1987) Tip Surface Interactions in STM and AFM, *Physica Scripta*, **T19A**, 61-66.

Pokharel, S. and Campbell, J.L. (2012) Cross talk between the nuclease and helicase activities of Dna2: role of an essential iron-sulfur cluster domain, *Nucleic Acids Res*, **40**, 7821-7830.

Potapova, T. and Gorbsky, G.J. (2017) The Consequences of Chromosome Segregation Errors in Mitosis and Meiosis, *Biology*, **6**.

Quennet, V.*, et al.* (2011) CtIP and MRN promote non-homologous end-joining of etoposide-induced DNA double-strand breaks in G1, *Nucleic Acids Res*, **39**, 2144-2152.

Quon, K.C., Marczynski, G.T. and Shapiro, L. (1996) Cell cycle control by an essential bacterial two-component signal transduction protein, *Cell*, **84**, 83-93.

Qvist, P.*, et al.* (2011) CtIP Mutations Cause Seckel and Jawad Syndromes, *PLoS genetics*, **7**, e1002310.

Radmacher, M.*, et al.* (1994) Direct observation of enzyme activity with the atomic force microscope, *Science*, **265**, 1577-1579.

Ramirez-Lugo, J.S.*, et al.* (2011) CtIP interacts with TopBP1 and Nbs1 in the response to double-stranded DNA breaks (DSBs) in Xenopus egg extracts, *Cell cycle*, **10**, 469-480.

Rayssiguier, C., Thaler, D.S. and Radman, M. (1989) The barrier to recombination between Escherichia coli and Salmonella typhimurium is disrupted in mismatch-repair mutants, *Nature*, **342**, 396-401.

Resch-Genger, U.*, et al.* (2008) Quantum dots versus organic dyes as fluorescent labels, *Nature methods*, **5**, 763.

Reyes-Lamothe, R.*, et al.* (2014) High-copy bacterial plasmids diffuse in the nucleoid-free space, replicate stochastically and are randomly partitioned at cell division, *Nucleic Acids Res*, **42**, 1042-1051.

Rivetti, C. and Codeluppi, S. (2001) Accurate length determination of DNA molecules visualized by atomic force microscopy: evidence for a partial B- to A-form transition on mica, *Ultramicroscopy*, **87**, 55-66.

Rivetti, C., Guthold, M. and Bustamante, C. (1996) Scanning force microscopy of DNA deposited onto mica: equilibration versus kinetic trapping studied by statistical polymer chain analysis, *Journal of molecular biology*, **264**, 919-932.

Ruchaud, S., Carmena, M. and Earnshaw, W.C. (2007) Chromosomal passengers: conducting cell division, *Nature reviews. Molecular cell biology*, **8**, 798-812.

Rust, M.J., Bates, M. and Zhuang, X. (2006) Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM), *Nature methods*, **3**, 793-795.

Sanchez-Sevilla, A.*, et al.* (2002) Accuracy of AFM measurements of the contour length of DNA fragments adsorbed on mica in air and in aqueous buffer, *Ultramicroscopy*, **92**, 151-158.

Sanii, B. and Ashby, P.D. (2010) High sensitivity deflection detection of nanowires, *Physical review letters*, **104**, 147203.

Sartori, A.A.*, et al.* (2007) Human CtIP promotes DNA end resection, *Nature*, **450**, 509-514.

Schaeper, U.*, et al.* (1998) Interaction between a cellular protein that binds to the C-terminal region of adenovirus E1A (CtBP) and a novel cellular protein is disrupted by E1A through a conserved PLDLS motif, *The Journal of biological chemistry*, **273**, 8549-8552.

Schafer, K.A. (1998) The cell cycle: a review, *Veterinary pathology*, **35**, 461-478.

Schumacher, M.A. (2008) Structural biology of plasmid partition: uncovering the molecular mechanisms of DNA segregation, *The Biochemical journal*, **412**, 1-18.

Schumacher, M.A. (2012) Bacterial plasmid partition machinery: a minimalist approach to survival, *Current opinion in structural biology*, **22**, 72-79.

Sender, R., Fuchs, S. and Milo, R. (2016) Revised Estimates for the Number of Human and Bacteria Cells in the Body, *PLoS biology*, **14**, e1002533.

Seo, Y., Choe, H. and Jhe, W. (2003) Atomic-resolution noncontact atomic force microscopy in air, *Applied Physics Letters*, **83**, 1860-1862.

Shaheen, R.*, et al.* (2014) Genomic analysis of primordial dwarfism reveals novel disease genes, *Genome research*, **24**, 291-299.

Shaw, J.E.*, et al.* (2006) Coupling evanescent-wave fluorescence imaging and spectroscopy with scanning probe microscopy: challenges and insights from TIRF–AFM, *Surface and Interface Analysis*, **38**, 1459-1471.

Shibata, A. and Jeggo, P.A. (2014) DNA double-strand break repair in a cellular context, *Clinical oncology*, **26**, 243-249.

Shrivastav, M., De Haro, L.P. and Nickoloff, J.A. (2008) Regulation of DNA double-strand break repair pathway choice, *Cell research*, **18**, 134-147.

Smith, G.R. (2012) How RecBCD enzyme and Chi promote DNA break repair and recombination: a molecular biologist's view, *Microbiology and molecular biology reviews : MMBR*, **76**, 217-228.

Smithies, O. (1955) Zone electrophoresis in starch gels: group variations in the serum proteins of normal human adults, *The Biochemical journal*, **61**, 629-641.

Solomon, J.M. and Grossman, A.D. (1996) Who's competent and when: regulation of natural genetic competence in bacteria, *Trends in genetics : TIG*, **12**, 150-155.

Šponer, J., Leszczynski, J. and Hobza, P. (1996) Structures and Energies of Hydrogen-Bonded DNA Base Pairs. A Nonempirical Study with Inclusion of Electron Correlation, *The Journal of Physical Chemistry*, **100**, 1965-1974.

Suzuki, T.*, et al.* (2007) Recent advances in fluorescent labeling techniques for fluorescence microscopy, *Acta histochemica et cytochemica*, **40**, 131-137.

Takata, M.*, et al.* (1998) Homologous recombination and non-homologous end-joining pathways of DNA double-strand break repair have overlapping roles in the maintenance of chromosomal integrity in vertebrate cells, *The EMBO journal*, **17**, 5497-5508.

Takeda, D.Y. and Dutta, A. (2005) DNA replication and progression through S phase, *Oncogene*, **24**, 2827.

Tang, M.*, et al.* (2007) Iteron-binding ORF157 and FtsZ-like ORF156 proteins encoded by pBtoxis play a role in its replication in Bacillus thuringiensis subsp. israelensis, *Journal of bacteriology*, **189**, 8053-8058.

Tatum, E.L. and Lederberg, J. (1947) Gene Recombination in the Bacterium Escherichia coli, *Journal of bacteriology*, **53**, 673-684.

Taylor, J.A.*, et al.* (2015) Specific and non-specific interactions of ParB with DNA: implications for chromosome segregation, *Nucleic Acids Res*, **43**, 719-731.

Tham, K.C., Kanaar, R. and Lebbink, J.H. (2016) Mismatch repair and homeologous recombination, *DNA repair*, **38**, 75-83.

Thomson, N.H.*, et al.* (1996) Protein tracking and detection of protein motion using atomic force microscopy, *Biophysical journal*, **70**, 2421-2431.

Tougan, T.*, et al.* (2010) The Mek1 phosphorylation cascade plays a role in meiotic recombination of Schizosaccharomyces pombe, *Cell cycle*, **9**, 4688-4702.

Tran, N.T.*, et al.* (2018) Permissive zones for the centromere-binding protein ParB on the Caulobacter crescentus chromosome, *Nucleic Acids Res*, **46**, 1196-1209.

Tutt, A.*, et al.* (2001) Mutation in Brca2 stimulates error-prone homology-directed repair of DNA double-strand breaks occurring between repeated sequences, *The EMBO journal*, **20**, 4704-4716.

Uchihashi, T., Ando, T. and Yamashita, H. (2006) Fast phase imaging in liquids using a rapid scan atomic force microscope, *Applied Physics Letters*, **89**, 213112.

Uchihashi, T.*, et al.* (2011) High-speed atomic force microscopy reveals rotary catalysis of rotorless F(1)-ATPase, *Science*, **333**, 755-758.

Uzman, A. (2003) Molecular biology of the cell (4th ed.): Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K., and Walter, P, *Biochemistry and Molecular Biology Education*, **31**, 212-214.

Venturoli, D. and Rippe, B. (2005) Ficoll and dextran vs. globular proteins as probes for testing glomerular permselectivity: effects of molecular size, shape, charge, and deformability, *American journal of physiology. Renal physiology*, **288**, F605-613.

Viani, M.B.*, et al.* (1999) Fast imaging and fast force spectroscopy of single biopolymers with a new atomic force microscope designed for small cantilevers, *Review of Scientific Instruments*, **70**, 4300-4303.

Viswanathan, M.*, et al.* (2001) Redundant exonuclease involvement in Escherichia coli methyl-directed mismatch repair, *The Journal of biological chemistry*, **276**, 31053-31058.

Walker, J.E.*, et al.* (1982) Distantly related sequences in the alpha- and beta-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold, *The EMBO journal*, **1**, 945-951.

Walters, D.A.*, et al.* (1996) Short cantilevers for atomic force microscopy, *Review of Scientific Instruments*, **67**, 3583-3590.

Walters, D.A.*, et al.* (1997) *Atomic force microscopy using small cantilevers*. Photonics West '97. SPIE.

Wang, J.L.*, et al.* (2018) Dissection of DNA double-strand-break repair using novel single-molecule forceps, *Nature structural & molecular biology*, **25**, 482-487.

Wang, X., Montero Llopis, P. and Rudner, D.Z. (2013) Organization and segregation of bacterial chromosomes, *Nature reviews. Genetics*, **14**, 191-203.

Wilkinson, O.J.*, et al.* (2019) CtIP forms a tetrameric dumbbell-shaped particle which bridges complex DNA end structures for double-strand break repair, *eLife*, **8**.

Williams, R.S.*, et al.* (2009) Nbs1 flexibly tethers Ctp1 and Mre11-Rad50 to coordinate DNA double-strand break processing and repair, *Cell*, **139**, 87-99.

Williamson, J.B.P. (1967) Paper 17: Microtopography of Surfaces, *Proceedings of the Institution of Mechanical Engineers, Conference Proceedings*, **182**, 21-30.

Wing, R.*, et al.* (1980) Crystal structure analysis of a complete turn of B-DNA, *Nature*, **287**, 755-758.

Worth, L., Jr.*, et al.* (1994) Mismatch repair proteins MutS and MutL inhibit RecA-catalyzed strand transfer between diverged DNAs, *Proceedings of the National Academy of Sciences of the United States of America*, **91**, 3238-3241.

Wu, X.*, et al.* (1997) Topography of ribosomes and initiation complexes from rat liver as revealed by atomic force microscopy, *Biological chemistry*, **378**, 363-372.

Yadav, T.*, et al.* (2014) Roles of Bacillus subtilis DprA and SsbA in RecA-mediated genetic recombination, *The Journal of biological chemistry*, **289**, 27640-27652.

Yao, J. and Wang, L.V. (2013) Photoacoustic Microscopy, *Laser & photonics reviews*, **7**.

Yoon, Y.G. and Koob, M.D. (2005) Transformation of isolated mammalian mitochondria by bacterial conjugation, *Nucleic Acids Res*, **33**, e139.

Yum, K., Wang, N. and Yu, M.F. (2010) Nanoneedle: a multifunctional tool for biological studies in living cells, *Nanoscale*, **2**, 363-372.

Yun, M.H. and Hiom, K. (2009) CtIP-BRCA1 modulates the choice of DNA double-strand-break repair pathway throughout the cell cycle, *Nature*, **459**, 460-463.

Zerbino, D.R.*, et al.* (2018) Ensembl 2018, *Nucleic Acids Res*, **46**, D754-D761.

Zhong, Q.*, et al.* (1993) Fractured polymer/silica fiber surface studied by tapping mode atomic force microscopy, *Surface Science Letters*, **290**, L688-L692.

Zhou, C., Pourmal, S. and Pavletich, N.P. (2015) Dna2 nuclease-helicase structure, mechanism and regulation by Rpa, *eLife*, **4**.

Zinder, N.D. and Lederberg, J. (1952) Genetic exchange in Salmonella, *Journal of bacteriology*, **64**, 679-699.