UNIVERSITÀ DEGLI STUDI DI TRIESTE

Graduate school in Nanotechnology

PhD Thesis

Atomic force microscopy studies of DNA binding properties of the archaeal MCM complex

AMNA ABDALLA MOHAMMED KHALID

XXVIII CICLO - ANNO ACCADEMICO 2014 / 2015

UNIVERSITÀ DEGLI STUDI DI TRIESTE

XXVIII CICLO DEL DOTTORATO DI RICERCA IN

NANOTECNOLOGIE

Atomic force microscopy studies of DNA binding properties of the archaeal MCM complex

Settore scientifico-disciplinare: FIS/03

DOTTORANDO

AMNA ABDALLA MOHAMMED KHALID

COORDINATORE

PROF. LUCIA PASQUATO

SUPERVISORE DI TESI

DR. LOREDANA CASALIS

CO-SUPERVISORE

DR. SILVIA ONESTI

TUTORE

DR. PIETRO PERISSE

ANNO ACCADEMICO 2014 / 2015

Abstract

The MCM (mini-chromosome mainteinance) protein complex has a key role in the replication machinery of Eukaryotes and Archaea. Beside the main unwinding role, this helicase is also supposed to act as one of the essential element licensing replication to ensure that each segment of the genome is replicated only once per cycle.

Six homologous MCM proteins belonging to the AAA+ ATPase superfamily, known as MCM2-7, are forming in Eukaryotes a hexameric hetero-complex that is supposed to "load" onto a single DNA strand and to use energy from ATP (Adenosine triphosphate) hydrolysis to unzip the DNA double helix. In Archaea, there is a single MCM gene, coding for a single MCM protein. Therefore the protein complex is made by a homohexamer with six equivalent subunits. The DNA loading mechanism for double-strand unzipping is similar to the Eukaryotes case.

The scope of this thesis work is to investigate the structural details of the interaction between DNA and the MCM complex in near physiological condition, by means of Atomic Force Microscopy (AFM) single molecule imaging. For that, we used archaeal MCM from *Methanothermobacter Thermautotrophicus* as a model system. There are structural and biochemical evidences that MCM interacts with DNA in two ways: the canonical “loaded” mode where the MCM protein complex is encircling the DNA for unwinding, and an “associated” mode where the DNA is supposed to be wrapped around the external part of the proper ring structure. In this second configuration, less studied than the loaded one, the binding interaction between MCM and DNA might play a role in licencing/initiating the replication process.

By means of AFM, we studied the conformational changes induced by the protein complex on blunt-ends, double-stranded (ds) DNA filaments of different lengths. AFM experiments were carried out in two different conditions: in air to understand the static conformations of DNA-protein complexes; in liquid to follow the interaction dynamics. We first optimized the protocol (surface treatment, buffer conditions) for AFM imaging in air to obtain high resolution images of surface-equilibrated DNA molecules before and after the interaction with the protein complex. From statistical analysis of AFM images, we localized the protein complexes along the isolated dsDNA sequences and calculated DNA contour length and

bending angle before and after the interaction with the protein complex. We discriminated between proteins with DNA wrapped around, calling them “associated”, and the ones interacting with DNA without inducing any bending, calling them “loaded”. To confirm this topographical assignment, we tested two mutants: the NβH complex, which presents a mutation in the central hole of the hexamer, inhibiting DNA loading; the ΔsA complex, which presents a mutation in the external part, preventing DNA association. In the case of NβH, only associated DNA was observed; in the case of ΔsA, only loaded complexes were found, to prove the soundness of our assumptions. Moreover, the total number of DNA bound complexes decreased from 84% to 20% from MCM to ΔsA, proving that association is involved in favouring the replicative helicase loading, and initiating the double-helix unwinding. Finally, we found a DNA compaction of about 13 nm for wild type MCM and ΔsA mutant. In the case of NβH mutant, the compaction is of about 18nm, and comes together with a bending angle increase of about 16°, strongly supporting the “association” model.

Finally, we studied the dynamics of DNA-MCM complex interaction, in the presence of ATP, via moderately fast (few seconds/images) AFM imaging in liquid. After ATP loading, we observed a change of topographic height of the DNA strands, consistent with the formation of ssDNA, a sign of DNA unwinding by the MCM complex. This effect was unexpected, since there are no biochemical evidences in literature of efficient unwinding of MCM complex on blunt-ended DNA. Different structural types of DNA substrates (e.g. with forks, bubbles, sticky ends, etc.) are currently under investigation.

Acknowledgements

I would like to thank my supervisors Loredana Casalis and Silvia Onesti for their guidance and support during the course of my doctoral work. Especially, Loredana for the many opportunities to attend scientific meetings and for encouraging my independent thought, I will remain grateful for the opportunity they offered me to join their research groups. I would like to thank Pietro P., who trained me to work with Atomic Force Microscopy (AFM) and most of the reported measurements would have not possible without him.
This work could have not been possible without the materials from group of Silvia Onesti. Especially I would like to thank Barbara Medagli for her constant support.

I thank all the present and past members of my group in Nanoinnovation lab at Elettra for their help, support and all the funny moments that we spent together.
I would also like to thank Iwan Schaap and his group members Mitja Platen, who hosted me for three months and help me to perform my experimental work using fast AFM.
I thank all my teachers, who have transferred their valuable knowledge to enlighten my mind. Also I thank my best friends Omima, Emtinan, Nehad, Monica and John, whose words made me strong, and provided confidence to pursue my research work. And many thanks to all my friends here and everywhere over the world.

Thanks to everybody in ICTP, especially Sandro Scandolo, Koutou Mabilo and Sabina Emili.

Finally, my research would have not gone smoothly without my family. I thank my parents, my sister and brothers for all that have given me since my childhood and even now, their love, prayers and sacrifices that have brought me up to this level. I wish them good health and happiness for all their life.

This thesis is dedicated to my parents.

TABLE OF CONTENTS

List of abbreviations

AAA+: ATPases Associated with various cellular Activities

AFM: Atomic Force Microscopy

ATP: Adenosine Triphosphate

bp: base pairs

Cdc45: Cell division cycle 45

Cdc48: Cell division cycle 48

Cdc6: Cell division cycle 6

CDK: Cyclin Dependent Kinase

Cdt1: Chromatin licensing and DNA replication factor 1

CMG: Cdc45-MCM2-7-GINS

DDK: Dbf4-Dependent Cdc7 Kinase

DNA: Deoxyribonucleic Acid

dsDNA: double stranded DNA

EDTA: Ethylenediaminetetraacetic Acid

EM: Electron Microscopy

GINS: Go-Ichi-Nii-San (5-1-2-3 in Japanese)

H2I: Helix 2 Insert

HEPES: 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid

His: Histidine

MCM: Minichromosome-maintenance protein

Mka: Methanopyrus Kandleri

*Mth*MCM: *Methanothermobacter thermautotrophicus* MCM

NβH: N-terminal β-Hairpin

nt: nucleotides

ORC: Origin Recognition Complex

PCR: Polymerase Chain Reaction

Pfu: *Pyrococcus furiosus*

pre-RC: pre-Replication Complex

sA: subdomain A

sB: subdomain B

sC: subdomain C

WH: Winged Helix

ΔsA: Subdomain-A deletion mutant

1 Introduction

The process of DNA replication is at the core of cell growth and proliferation. An intricate and dynamic network of protein complexes assembles on specific DNA sites under tight cell-cycle control, to generate bidirectional replication forks that make a copy of the entire DNA, to pass onto daughter cells. As protein-DNA interactions can have topological implications, biochemical analysis are often not sufficient for their detailed characterization, which also requires biophysical and structural tools. Among those, atomic force microscopy (AFM) has proven very useful to visualise and characterize the interaction between proteins and nucleic acid. Here we carried out biophysical studies using AFM, to visualize protein-DNA interaction at single molecule level under physiological environment, as a function of different interaction parameters.

Understanding at the molecular level the mechanisms that govern DNA replication in proliferating cells is fundamental to understand diseases connected to genomic instabilities, such as genetic disease and cancer.

1.1 DNA replication

DNA replication is the process required to copy the cellular genome to ensure the transmission of the genetic information to the next generation. It is a complex series of events that requires a large set of proteins and enzymes as well as energy from ATP (adenosine triphosphate) hydrolysis. A key step for DNA replication to take place is the breaking of hydrogen bonds between bases of the two antiparallel strands (Fig. 1.1); this is catalysed by an important class of enzymes called helicases.

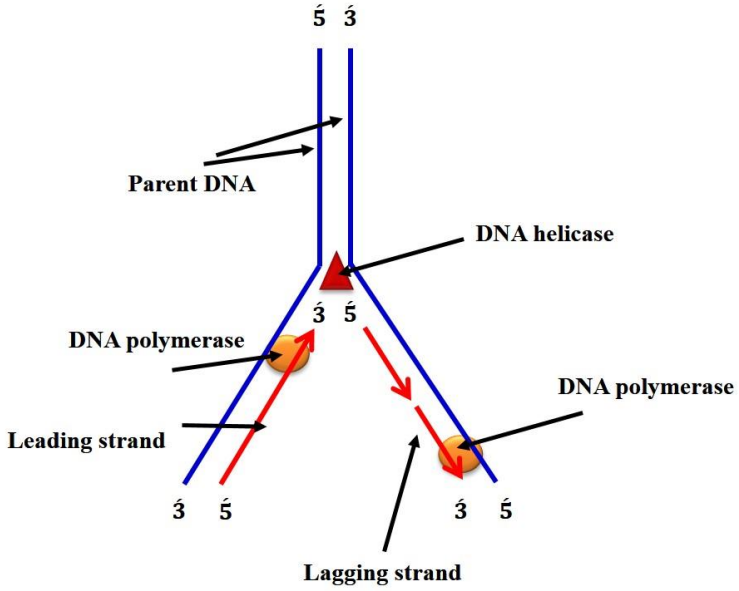


Figure 1.1: DNA replication. The double helix is unwound and each strand acts as a template for the next strand. Bases are matched to synthesize the new partner strands.

▪ Initiation of DNA replication in Eukaryotes

The unwinding of the double strands DNA begins at specific sites on the DNA called the origins of replication. From the origin two replisomes with opposite polarity are assembled and two replication forks proceed, in a bidirectional manner. The specific structure of the replication origin varies somewhat from species to species but all share some common characteristic such as high A-T content.

Prokaryotic and eukaryotic origins of replication show significant differences:

- Most bacteria have a single circular molecule of DNA, and typically only a single origin of replication per circular chromosome.
- Archaea, (prokaryotic organisms that are classified as a third domain of life distinct from bacteria), can have one or a few origin(s) to initiate replication of their circular chromosomes.
- Due to the large size of their genomes, as well as the presence of more than one chromosome, eukaryotic cells need to initiate DNA replication at multiple origins which are switched on according to a very tightly regulated program to make sure that each origin is used once and only once per cell cycle; moreover, sometimes origins fire with different timing. The activity of these origins must therefore be tightly coordinated so

that the entire genome is accurately replicated and no segment in unreplicated or re-replicated.

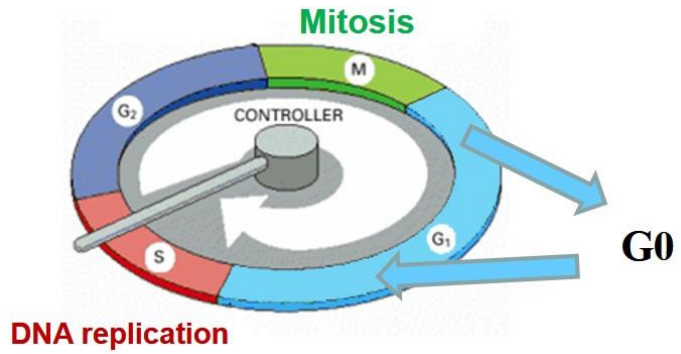


Figure 1.2: A diagram showing the main elements of the eukaryotic cell cycle. A growth phase, G1, is followed by DNA replication or S phase. After another growth phase, G2, the cell separates its chromosomes into two sets (mitosis) and divides into two new cells (cytokinesis). Most cells in an adult organism do not divide and are stalled G0, a quiescent state.

Cell division requires a set of events in the cell that occur to produce two daughter cells, following the cycle shown in (Fig. 1.2).

During the G1 phase, the cell increase in size, produce RNA and synthesize protein. Toward the end of this phase there is an important cell cycle control (G1 Checkpoint), to ensure that everything is ready for DNA synthesis. DNA replication occurs during the S (synthesis) phase, which is then followed by a second gab or growth phase (G2 phase). There is another control checkpoint (G2 Checkpoint) at the end of this gap, to determine if the cell can now proceed to enter M (mitosis) and divide. All of the cell's energy is focused on the complex and orderly division into two similar daughter cells. Mitosis is much shorter than interphase, lasting perhaps only one to two hours. As in both G1 and G2, there is a Checkpoint in the middle of mitosis (Metaphase Checkpoint) that ensures the cell is ready to complete cell division.

Most of the differentiated cells are not in the cell cycle and are stalled in a quiescent state called G0.

- *Control mechanisms during replication:*

Eukaryotic cells prevent re-replication of DNA by having a distinct stage that occurs during the G1 and G2 phase. Two different types of control mechanisms have been identified. A “positive” control mechanism ensures that replication origins are “licensed” for replication only during the G1 phase of the cycle (Blow et al., 1988; Bochman et al., 2008; Nishitani et

al., 2002 and Sclafani et al., 2007,). Origin licensing leads to the recruitment of the replicative helicase, the MCM (minichromosome maintenance) protein complex, to DNA origin via a complex network of interactions that includes the origin recognition complex (ORC), Cdc6 and Cdt1, during the G1 phase. In higher eukaryotes a negative control mechanism also exist: the replication factor "geminin" prevents the assembly of MCM proteins to DNA by binding and inhibiting the DNA replication factor Cdt1, which is essential for the licensing reaction. As geminin is expressed during the S and G2 phases of the cell cycle and is degraded during M phase (Mitosis), it further ensures that no reloading of MCM can occur in the S/G2 phases.

- *Pre-replication complex:*

In eukaryotes, the initiators ORC complex bind to the replication origin, and recruit other proteins to form the pre-replication complex (pre-RC) in the G1 phase of the cell cycle (Fig. 1.2). Pre-RC formation involves assembly of a number of proteins including the origin recognition complexes (ORC) itself, Cdc6, Cdt1, and the DNA helicase MCM 2–7 (Fig. 1.3) (Evrin et al., 2009 and Remus et al., 2009b). This will act as a platform for the further loading of other proteins and the triggering of series of events, including the switching of the helicase activity, the melting of the origin, the binding of DNA polymerases and other various factors to the chromatin to start DNA replication.

Figure 1.3: Schematic representation of the pre-replication complex (pre-RC).

1.2 MCM protein complex

The MCM (Minichromosome Maintenance) proteins are a family of polypeptides, ranging from 600 to 1200 amino acid residues, originally discovered and named as factors that support minichromosome maintenance in yeast (Forsburg, 2004, Maiorano et al., 2006; Costa & Onesti, 2009). Six of these proteins (called MCM2, MCM3, MCM4, MCM5, MCM6, MCM7) are conserved in all eukaryotes and form a complex, whereas two other polypetides (MCM8 and MCM9) are present only in some organisms, do not interact with the MCM2-7 complex and have a distinct cellular role.

A lot of the initial biochical and structural analysis of the MCM proteins has been carried out using the simplified archaeal orthologues (Miller & Enemark, 2015). In most Archaea, the helicase is a single MCM protein, forming a homohexamer that is related to the eukaryotic MCM2-7. The complex has a processive ATP-dependent helicase activity in the 3'→ 5' direction.

Whereas the archaeal MCM complex has an intrinsic helicase activity, the eukaryotic complex is inactive or an extremely weak helicase on its own. Although the MCM2-7 complex is indeed the helicase motor, the eukaryotic MCM helicase activity requires two additional components: the Cdc45 replication factor and the GINS tetrameric complex. All together these constitute the CMG (Cdc45-MCM-GINS) complex that acts as the replicative helicase assembly. Archaea do possess orthologues of both GINS and Cdc45, but their exact role in DNA replication has not yet been established (Onesti et al, 2013; Krastanova et al. 2012).

1.2.1 The AAA+ ATPase superfamily

The MCM helicase belongs to the general division of P-loop NTPases, in which there are two functional motifs that are necessary for ATP binding and hydrolysis: the Walker A, known as phosphate-binding loop, or "P loop" and the Walker B motifs. Nevertheless, the motor core of the helicase consists of a particular fold known as ASCE (Additional Strand Catalytic E) (Thomsen et al., 2008) that characterized by an additional strand between the Walker A and Walker B motifs, and by the presence of a catalytic glutamate (Leipe et al., 2003; Medagli et al., 2013).

Within the ASCE division, the MCM proteins are a distinct subgroup of the large AAA+ ATPase superfamily (ATPases Associated with various cellular Activities), which generally form large ATP-dependent complexes. They often assemble as higher order oligomers, with a strong preference for hexameric rings (Erzberger et al., 2006; Ogura et al., 2001).

The ATP binding site is usually consist of residues spanning neighbouring subunits and therefore are located at the interface between the subunits shown in (Fig. 1.4). This particular arrangement can drive a change in the architecture of the complex and be transformed into mechanical work caused by ATP binding.

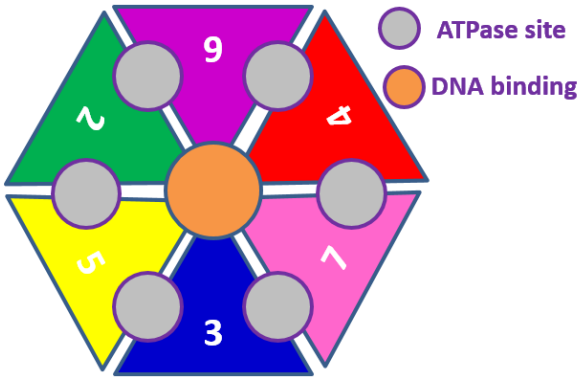


Figure 1.4: Schematic representation of the MCM2-7 hexamer structure. Six MCM subunitsThe DNA is expected to pass through the central hole (in orange) and the ATP binding site at the interface between subunits (in grey).

1.2.2 The general organization of MCM protein

MCM proteins are conserved between Archaea and Eukarya, with most Archaea possessing a single copy forming homomeric rings, while eukaryotes have at least six different paralogs which are known as MCM2-7 that assemble into a heterohexameric complex.

Due to its complexity, most of the early studies on the MCM complex have been carried out on the homologous archaeal complex. The similarity between eukaryotic replication proteins and those found in archaea (Edgell et al. 1997) has led to the archaeal MCM complexes being adopted as simplified model systems for studying eukaryotic DNA replication (Miller & Enemark, 2015). More recently both biochemical studies and structural studies using single particle cryo-electron microscopy have shed light on some of the eukaryotic complexes from yeast and Drosophila (Ali et al., 2016; Li et al., 2016; Sun et al., 2015; Yuan et al., 2016).

A canonical MCM protein can be divided into three domains (Figure 1.5): the N-terminal, AAA+ and C-terminal domains: the AAA+ domain is the catalytic core of the enzyme, whereas the N- and C-terminal domains are less conserved.

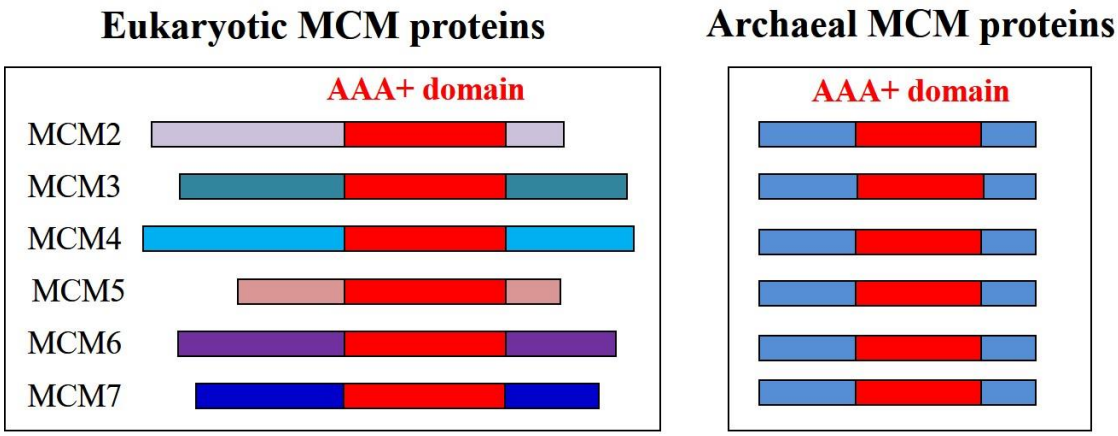


Figure 1.5: Schematic representation of the primary structure of typical eukaryotic and archaeal MCM complex, divided to three domains: the AAA+ domain (in red) contains a Walker A, Walker B and arginine finger motifs typical of AAA+ ATPases. The N-terminal domain (on the left of AAA+ domain) contains a β-hairpin (NβH), capable of interacting with DNA. The C-terminal domain (in the right of AAA+ domain) contains WH domain.

- *N-terminal domain*

The N-terminal domain does not have catalytic activity but was shown to be important for hexamer formation, DNA binding and enzyme regulation (Barry et al., 2007; Jenkinson and Chong, 2006; Kasiviswanathan et al., 2004; Liu et al., 2008; Pucci et al., 2007).

The N-term domain, is capable of binding both ssDNA (single-stranded DNA) and dsDNA (with a preference for ssDNA) is strong determinant for hexamerization (Fig 1.6). It is able to influence the processivity of the helicase and its polarity and most likely acts as a brake to ensure that the helicase activity is tightly regulated (Barry et al., 2007; Jenkinson et al., 2006; Kasiviswanathan et al., 2004; Liu et al., 2008 and Pucci et al., 2007).

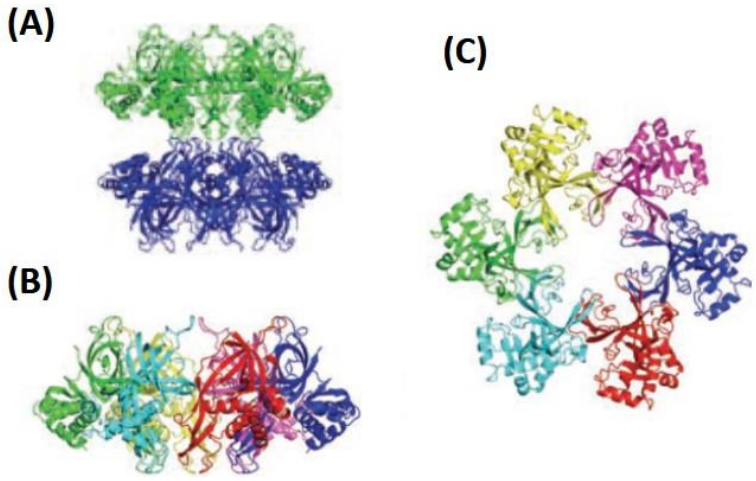


Figure 1.6: The crystal structure of the N-terminal domain of M. thermautotrophicus MCM (Fletcher et al., 2003). A) The side view of the double hexameric configuration is shown with the two hexamers highlighted in green and blue, respectively. B) The side view and C) The top view of the MthMCM Nterminal hexamer are color-coded according to the different subunit. The structure contains a channel which is large enough to fit dsDNA. (**Adapted from Costa et al., 2009**).

A number of crystal structures are available for the archaeal N-terminal domains, either is a single hexamer, double hexamer or helical configuration (reviewed in Costa and Onesti, 2009; Miller et al., 2015). The N-term is composed by three subdomains (sA, sB and sC). Subdomain A is located in the peripheral belt of the MCM ring and is involved in the regulation of the helicase activity. Subdomain B contains the Zn motif that contributes to DNA binding and to the formation of a double-hexameric structure. Subdomain C is responsible for the hexamerization and contains a β-hairpin that projects towards the centre of the channel and is involved in DNA binding (Kasiviswanathan et al., 2004). Two different configurations have been observed for subdomain A: a "close configuration" where the domain is adjacent to subdomain B (Fletcher et al. 2003, Liu et al. 2008; Brewster at al., 2008) and an "open configuration" where the subdomain swings away from subdomain B (Jenkinson et al., 2009; Miller et al., 2014). This conformational transition may be instrumental for the initial interaction of the MCM complex with dsDNA (Costa et al, 2008a & 2008b).

The recent structure of the N-terminal domain of *Pyrococcus furiosus* MCM (N-*Pfu*MCM) showed for the first time the hexameric ring interacting with a short ssDNA (Froelich et al., 2014). Upon this interaction the N-terminal ring becomes asymmetric, generating an elliptical pore that accommodates two short ssDNA stretches, which are likely to be melting intermediates.

- *AAA+ domain*

The core of the MCM helicase is the central AAA+ domain, which is responsible for the catalytic activity and contains two specific insertions (the PS1BH and h2i loops) within the AAA+ fold with a critical role in DNA unwinding (Jenkinson et al., 2006; McGeoch et al., 2005).

Three crystal structures provide some atomic information on the domain architecture; of these structures two show only a monomer (Brewster et al., 2008; Bae et al., 2009) and are therefore not very useful to understand the catalytic mechanism that occur at the interface between the subunits. A more recent crystallographic structure of a chimeric MCM helicase, in which the N-terminal domain from *S. solfataricus* is fused to the AAA+ domain from *P. furiousus* and bound to ADP, showed the hexameric ring and hence provides information on the structural determinants of the active site (Miller et al., 2014). Both the PS1BH and h2i are projected into the channel, providing a number of positively charged and aromatic residues that can interact with the DNA backbone and bases.

- *C-terminal domain*

The C-terminal domain is in general less conserved: sequence analysis of this domain of archaeal MCM proteins shows the presence of a winged-helix (WH) domain (Aravind et al. 1999). The study of the C-terminal domain of human Mcm6 using high resolution NMR structure confirmed the predicted winged helix fold (Wei et al., 2010).

In archaea, the deletion of the C-terminal domain showed an increase in helicase and ATPase activity, indicating its role of controlling the helicase activity (Jenkinson et al., 2006). The NMR analysis of the C-terminus of *Mth*MCM and *Sso*MCM confirmed the presence of a winged helix fold in Archaea (Wiedemann et al., 2013).

1.2.3 Structural analysis of the MCM helicases assembly

The MCM complex can assume a large number of conformations, as visualised by electron microscopy and crystallography, including single hexamer or heptamer, double hexamer/heptamer structure with as head-to-head configuration (Fig. 1.7), and a variety of open rings and/or helical assemblies.

- *MCM double rings*

Whereas in most species the unbound proteins forms a single ring, the *Mth*MCM protein is unusual among MCMs in forming stable double rings in solution (Chong et al., 2000; Kelman et al., 1999 and Shechter et al., 2000). Most of the proteins assemble as double rings when bound to double-stranded DNA (Costa et al., 2006a&b; Remus et al., 2009).

Figure 1.7: Electron microscopy three dimensional reconstruction of a double hexamer of MthMCM from a negatively stained sample (Gomez-Llorente et al., 2005): the diameter of the hexameric ring is about 13 nm, whereas the overall length is about 20 nm (**Adapted from Costa et al., 2009**).

Most EM studies of archaeal proteins exploit the symmetric nature of the prokaryotic complex to use average in order to improve the signal/noise in the map.

More recent EM studies of the asymmetric eukaryotic MCM showed that the double hexamer structure is arranged in a tilted and twisted conformation (Evrin et al., 2009; Sun et al., 2014; Li et al., 2015, show in Fig. 1.8).

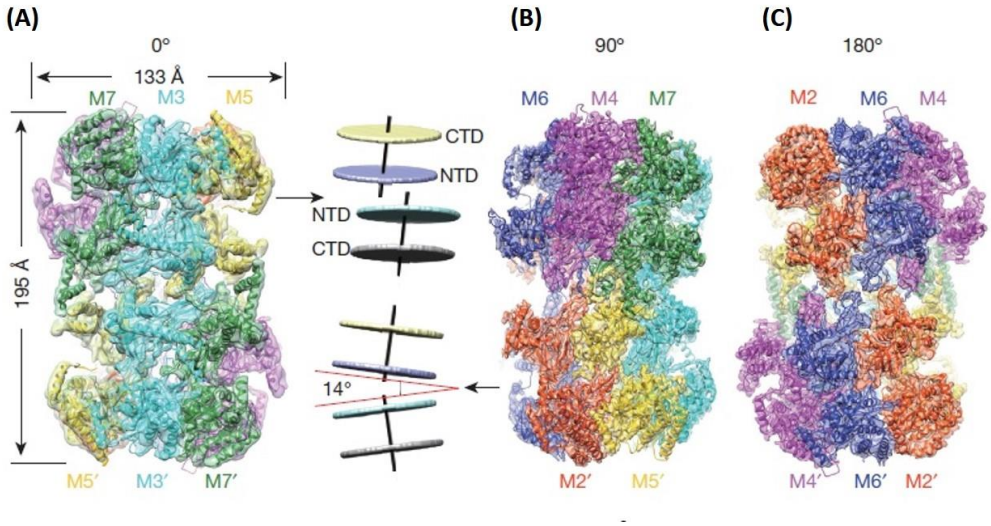


Figure 1.8: The organization of the MCM2–7double hexamer. Side-views of the cryo-EM density map super imposed with the atomic model. Unsharpened map, A) is displayed from the two-fold axis, and sharpened map (B and C) displayed with indicated rotations relative to A along the cylinder axis. The side panels of A and B illustrate the tilted arrangement of the two single hexamers. (**Adapted from Li et al., 2015**).

Using Electron Microscope (EM), the Mcm2–7 double hexamers bound to 1kb linear DNA as well circular DNA has been studied (Fig. 1.9). These observations are consistent with DNA passing through the central channel of Mcm2–7 double hexamers (Remus et al., 2009a).

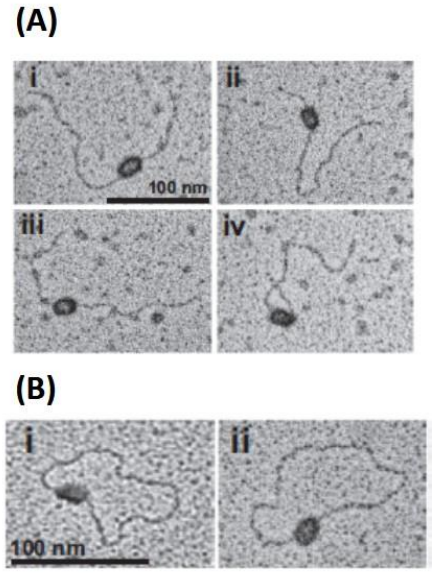


Figure 1.9: The Mcm2–7 double hexamers Can Slide on DNA. A) Mcm2–7 double hexamers bound to linear 1 kb ARS305-containing DNA visualized after tungsten and B) rotary shadow casting. Mcm2–7 double hexamers bound to circular 1 kb ARS305-containing DNA visualized. (**Adapted from Remus et al., 2009a**).

- *MCM single rings*

The loading process of the MCM helicase into the DNA requires the opening of the MCM2-7 complex to allow the DNA to insert and then the ring closes around the DNA. However, the studies of *S. cerevisiae* MCM2-7 In vitro, revealed that pre-incubation of the complex with ATP or a poorly hydrolysable analogue impaired the binding to circular ssDNA while for the linear DNA had negligible effects on the ability of the complex to enter in the central channel, which indicating the existence of a "gate" in the ring that is closed in the presence of ATP (Bochman and Schwacha, 2008). Furthermore, the mutations in Walker A motif of Mcm5 or the arginine finger of Mcm2 showed the impairment, indicating that the gate could be localised at the interface between these two subunits (Bochman et al., 2008). The Mcm2-7 single ring is not very stable .in the absence of nucleic-acid substrates and/or cofactor, but a couple of low resolution EM structures of the Drosophila MCM2-7 (Costa et al., 2011) show the presence of various forms of open rings (Fig. 1.10) and confirm the position of the gate between the MCM2 and MCM5 subunits.

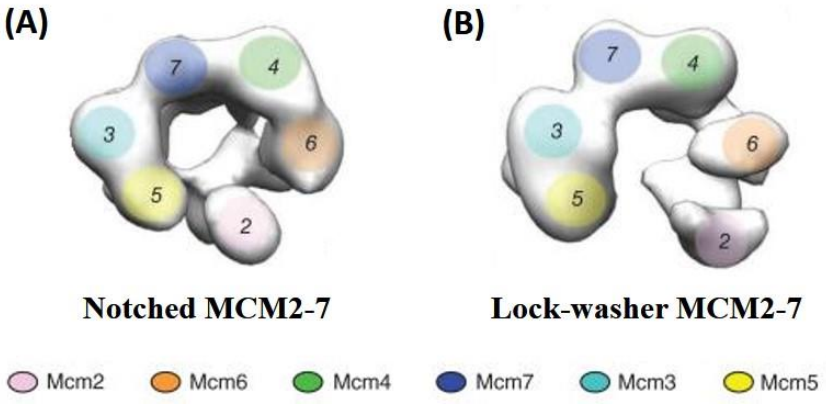


Figure 1.10: The apo-form of *Drosophila* MCM2-7 forms a gapped ring **(adapted from Costa et al., 2011).** (A) 3D EM reconstruction of the notched ring viewed from the AAA+ face. (B) 3D structure of the lock-washer ring viewed from the AAA+. See text for details.

- *Different modes of MCM-DNA interactions*

The variety of structures that have been observed for the MCM complex indicates an intrinsic polymorphism that relates to the multiple roles that this complex assumes within the cell cycle. This variability is partly dependent on the presence of substrates (nucleotides, ssDNA, dsDNA) and co-factors (Cdt1, GINS, Cdc45, etc..) but it also reflects a large degree

of plasticity which is liklely to be important for both the loading, the activation and progression of the MCM helicase.

One structure of *Mth*MCM obtained by cryo-electron microscopy in the presence of long dsDNA segments (5000 bp) revealed an unexpected interaction with dsDNA. Initial micrographs showed a protein-DNA interaction network where the DNA made a kink when encountering a protein complex (Fig1.11). A cryo-electron microscopy reconstruction indicated that the DNA, instead of threading through the hexameric ring central channel, wrapped around the external periphery of the ring (Costa et al., 2008 - EMBO J, Fig 1.12). The structure also suggested the swinging of subdomain A to accommodate the dsDNA, with a conformational change similar to that predicted based on biochemical and structural data (Fletcher et al., 2003) and then confirmed by EM and crystallographic structures (Jenkinson et al., 2009; Miller et al., 2014). The EM data is supported by biochemical data, confirming that the deletion of subdomain A strongly impairs binding of MthMCM to dsDNA.

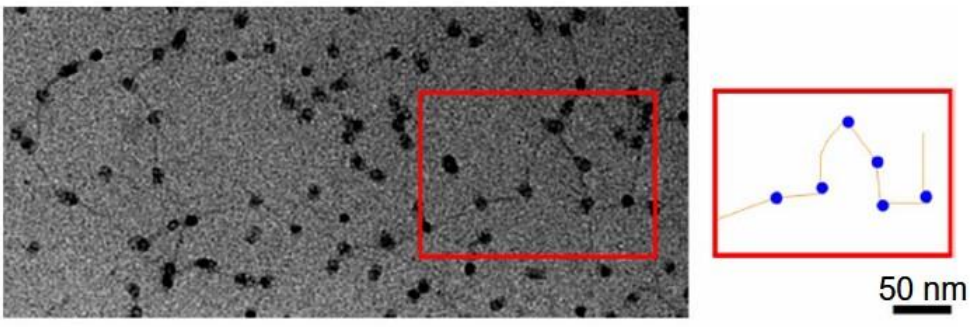


Figure 1.11: Micrograph of MCM proteins in complex with a 5600 bp double stranded DNA fragments using Cryo-Electron Microscopy (**Adapted from Alessandro costa's thesis**).

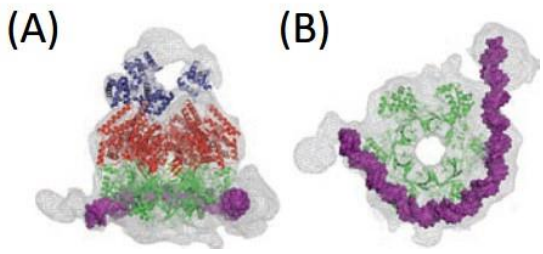


Figure 1.12: Cryo-EM 3D reconstruction of MthMCM and long dsDNA. A) side and B) bottom view of the EM electron density map, (N-term: green, AAA+ : red, C-term: blue, dsDNA: orange). A novel type of DNA-MCM interaction, with the dsDNA wrapping around N-terminal region of the ring. (**Adapted from Costa et al., 2008a**).

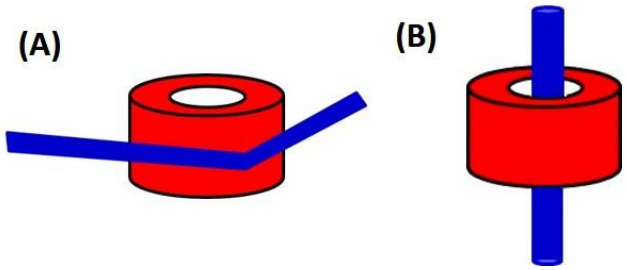


Figure 1.13: The two models of the DNA-MCM interaction that have be described in literature. A) Associated MCM, the DNA wraps around the external part of the structure has been visualized in the Cryo-EM 3D reconstruction of archaeal MCM bound to long double stranded DNA segments (Costa et al., 2008a) and B) Loaded MCM, the DNA is threaded through the hexameric ring channel, with the protein ready to act as helicase (Sun et al., 2014).

It has been suggested that this structure ay provide a structural model for the initial association of MCM to dsDNA. Indeed biochemical data indicates that in S. cerevisiae one can distinguish two mode of interactions between MCM2-7 and dsDNA: these have been described as “associated” protein, which is sensitive to high salt washes, and “loaded” protein that is stable in the presence of high salt (Bowers et al., 2004). The interaction between dsDNA and MCM described in this paper may provide a model for the “association” mode (Fig. 1.13). Similar two-step mechanisms have been observed in other systems, such as the Rho helicase/termination factor, another AAA+ protein (Costa et al. 2008 - Biocghem. Transactions).

1.2.4 The activation of the MCM helicase activity.

The exact mechanism of action of MCM protein is not yet established. Of the various models proposed the one that has more data supporting it is the steric-exclusion model. In this model, the MCM ring circles the ssDNA and walks along it, excluding the other strand.

One of the peculiar characteristic of the eukaryotic helicase, with respect to the viral, bacterial and archaeal counterparts, is in fact that the protein is loaded onto DNA in an active form and needs to be activated at a particular moment in the cell cycle. Indeed MCM is recruited to chromatin via the pre-RC in the G1 phase of the cycle, sits there in an inactive form, and then its helicase activity is finally activated when entering the S phase. The exact detailed molecular mechanisms governing this switch, and the consequent establishing of the replication fork, are still not completely understood. They involve a set of

phosphorylation events, and the ordered assembly of a large number of protein factors; of those Cdc45 and GINS remain associated with the MCM helicase, forming the CMG complex) and are essential for the helicase activity

Genetic studies have provided a wealth of information on the role of the MCM complex during the initiation step of DNA replication. In particular, DDK (Dbf4-dependent kinase)-mediated phosphorylation of various subunits of the MCM2–MCM7 complex is believed to trigger the switch which starts DNA replication (Hoang et al., 2007), together with the S-CDK mediated phosphorylation of other replication factors, that directs the assembly of the pre-initiation factors, that leads to recruitment of Cdc45 and GINS, as well as the DNA polymerases (Fig 1.14).

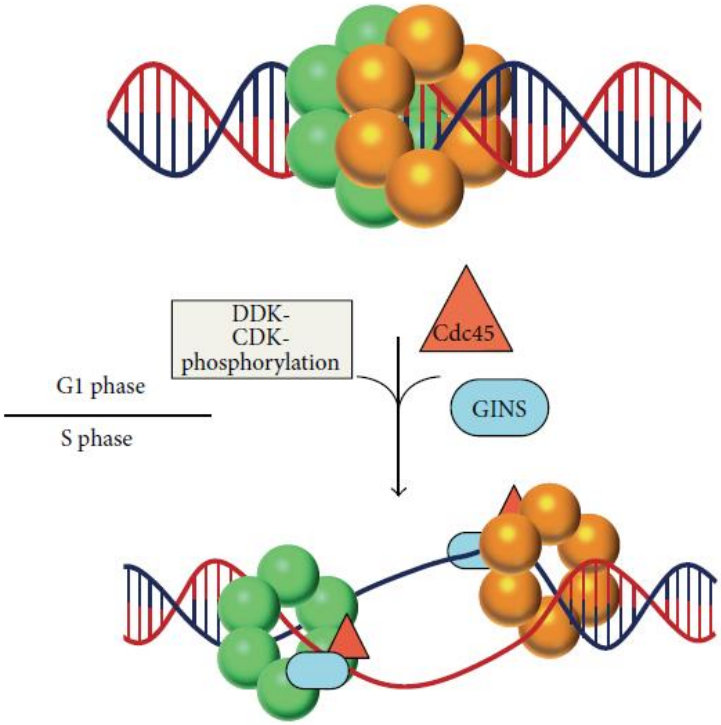


Figure 1.14: The MCM2–7 activation and organization. MCM2–7 rings are loaded as inactive double hexamers at origins of replication by the Origin Recognition Complex (ORC), Cdc6, and Cdt1 (not shown). In a cell-cycle dependent fashion, theDbf4-dependent Cdc7 kinase (DDK) and cyclin-dependent kinases (CDKs) drive the association ofCdc45 (red triangle) and the GINS complex (blue oval) with the phosphorylated Mcm2–7 ring to yield the active replicative helicase complex, termed the CMG complex (Cdc45-Mcm2–7-GINS). (**Adapted from Miller et al., 2015**).

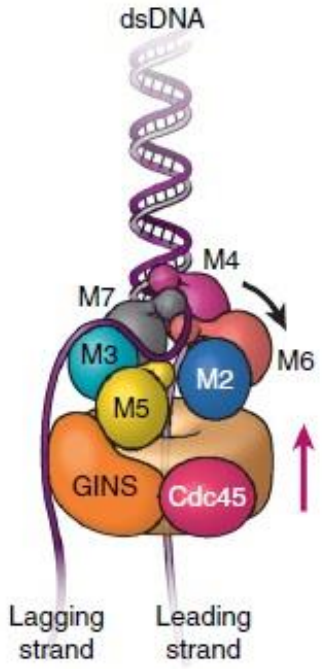


Figure 1.15: The pre-replication complex loaded into ssDNA, such as the leading strand (light purple) goes through the Mcm2–7 ring while the lagging strand (dark purple) is excluded to the outside. (**Adapted from Yuan et al., 2016**).

Archaeal proteins are constitutively active helicase. The MCM complex hydrolyses ATP have been stimulated by the addition of either ssDNA or dsDNA. For example, *Sso*MCM prefers substrates that contain both ssDNA and dsDNA, such as forks and bubbles (Pucci et al., 2004; Rothenberg et al., 2007), while *Mth*MCM slightly prefers ssDNA over dsDNA (Sakakibara et al., 2008). The helicase activity of archaeal complexes show a robust 3'→5' direction, that is dependent on ATP and Mg2+ and which demonstrates very little DNA substrate specificity: blunt, singly tailed, and forked substrates are all similarly unwound (Sakakibara et al, 2009; Miller eta l., 2015).

In vitro, for long time it was thought that the purified MCM2-7 did not display any helicase activity, with only a weak helicase activity found associated with the MCM 4/6/7 sub-complex, which possessed a weak but detectable ATP-dependent, DNA-unwinding activity in 3'→ 5' direction (Ishimi, 1997; Lee et al., 2001 and You et al., 2002). More care over the purification procedures have shown that the eukaryotic MCM2-7 did have a weak helicase

activity when purified in the absence of chloride ions, as smaller anions tend to inhibit the activity (Bochman & Schwacha, 2009).

1.2.5 Structure and function of CMG complex

Although MCM does constitute the "motor" of the helicase, the active helicase in DNA replication is indeed the CMG complex, which include Cdc45, MCM2-7 and the tetrameric GINS complex. The CMG complex translocates along the leading strand in 3'→5' direction to unwind the DNA double helix (Ilves et al 2010; Fu et al., 2011; Moyer et al., 2006).

The first EM model of *Drosophila* CMG revealed that MCM2-7 within the complex adopts a notched ring conformation, previously shown also by the apoMCM2-7 (Fig. 1.17). The notch was further narrowed by incubation with ADP•BeF3. Moreover, the additional density on the side of the ring, identified as GINS/Cdc45, embraced Mcm3, Mcm2 and 5, bridging the gap among the latters. These observations suggested that the enhanced ATPase and helicase activity measured for the CMG complex with respect to MCM2-7, could be due to the closure of the gate, enabling a proper repositioning of the ATPase active sites, suggesting that the CMG complex may initially assemble on the double hexamer, prior to the separation upon origin melting and fork progression.

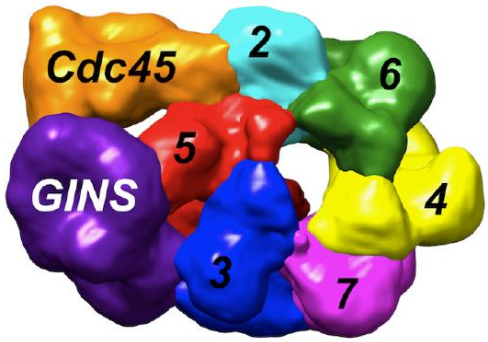


Figure 1.17: The EM structure of the *Drosophila* recombinant CMG (*Dm*CMG), showing the six member of MCM2-7 in complex with Cdc45 and GINS, were seen to bind across the Mcm2/Mcm5 gate. (**Adapted from Medagli et al.,**).

A higher resolution map (18 Å) for the CMG with a 3' tailed DNA and non-hydrolysable ATP analogue ATPγS was obtained allowing a more accurate fitting of the subunits (Costa

et al., 2014). The AAA+ ATPase tier adopts a clear right-handed spiral conformation, while the N-terminal tier remains almost planar, supporting the hypothesis that the N-terminal domains could coordinate and restrain the movements of the associated C-terminal regions. GINS/Cdc45 may play a role in stabilising this spiral configuration, which could have important implications for DNA unwinding.

More recently, a higher resolution map (7.4Å) of the Drosophila melanogaster CMG complex was determined upon incubation with a model replication fork in the presence of non-hydrolysable ATP analogue ATPγS (Abid et al., 2016) and showed that the CMG helicase trapped on a DNA fork (Fig. 1.18).

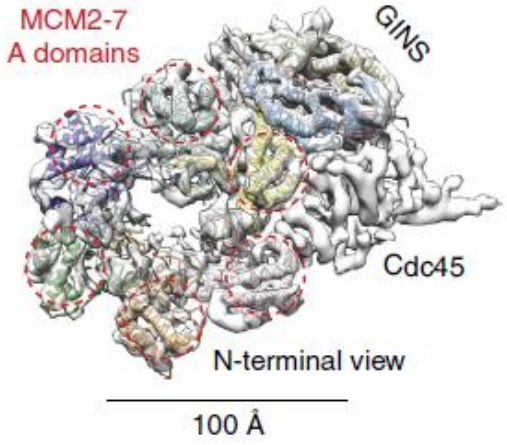


Figure 1.18: CMG helicase structure at subnanometre resolution. (a) Resolution density (7.4Å) map of the CMG viewed from the MCM N-terminal face, without or with docked MCM and GINS atomic structures. The electron density in the middle of the ring corresponds to the bound DNA (**Adapted from Ali et al., 2016**).

1.3 Summary

The MCM helicase is the key player in a complex and ordered network of events leading to the unwinding of dsDNA ahead of the replication fork. As such it is subjected to a very tight and accurate regulation, and needs to interact with a myriad of substrates and replication factors. The identification of the essential events and their timing is complicated in the eukaryotic Mcm2–7 complex because the intrinsic complexity of the asymmetric system. For this reason, archaeal MCM proteins have been used as useful models for elucidating the essential features of the complex.

Atomic resolution crystal structures and a variety of structures from electron microscopy (ranging from low to quasi-atomic resolution) will continue to reveal mechanistically important conformational states of the MCM complex, including how MCM hexamers specifically respond to binding different nucleotides, oligonucleotides, and protein interaction partners. Despite the recent progresses many questions are still open.

In order to answer the many open question and to better understand the cellular role and mechanism of MCM complex, other complementary techniques are required to give a comprehensive picture. Among those is atomic force microscopy. Although AFM does not reach single atomic resolution, it can visualize protein-DNA interaction in detail, at resolution of few nanometers in physiological environment, as a function of different interaction parameters.

Many replication proteins are abundant in transforming cancer cell lines when compared to normal cells, making them, in principle, potential biomarkers for cancer detection and prognosis. MCM proteins have been found to be overexpressed in a variety of tumors; most results show that they are more sensitive and specific markers than the conventional proliferative markers Ki-67 and PCNA (rev. in Giaginis et al., 2010). More recently many reports have shown that also Cdc45 and GINS are very promising candidates for novel proliferation markers and potential drug targets. Moreover, defects in MCM and GINS subunit themselves can cause cancer (Shima et al, 2007; Dutta et al, 2007). A detailed knowledge of the structure and function of the MCM and CMG complexes, and their interaction with nucleic acids is therefore a pre-requisite for the full exploitation of their potential in cancer diagnostic and therapy.

1.4 My project

I carried out a biophysical study to dissect the interaction of an archaeal MCM complex with blunt-ended DNA at a single molecule level with nanometer resolution using Atomic Force Microscopy (AFM). I have used the archaeal MCM from M. thermoautotrophicum (*Mth*MCM), which is capable to bind both single-stranded DNA (ssDNA) and double stranded DNA (dsDNA) (Sakakibara et al., 2008). I started my work using blunt-ended dsDNA of different lengths to visualize the interaction and to discriminate between the loaded and associated configurations.

I performed experiments using two different conditions: imaging in air, to understand the static conformations of DNA-MCM complexes via statistical analysis; imaging in liquid, to visualize the dynamic interactions of DNA-MCM complexes in the presence of ATP.

First, I optimized the protocol for AFM imaging in air to obtain clear and high resolution images of DNA molecules before and after the interaction with protein complex. Then I tested two MCM mutants: a mutant was designed to disrupt the N-terminal β-hairpin that has been shown to be important in binding the DNA in the central channel (NβH mutant); a second mutant lacked subdomain A, which is supposed to be involved in the initial loose association with DNA.

Next, I studied the dynamic of the DNA-MCM interaction using the same DNA fragment. I fueled the MCM helicase with ATP in a solution containing DNA molecules and then I incubated a drop of the solution on a mica surface using the optimized protocol for imaging in air. This allowed us to visualize the occurred unwinding analysing the effects of MCM-DNA interaction on the DNA fragments deposited on the surface.

Finally, I optimized the buffer condition for AFM imaging in liquid. I again investigated the DNA and MCM-DNA interaction using fast AFM. As well I studied the dynamic of the interaction in the presence of ATP.

2.1 DNA and protein samples preparation

2.1.1 *Mth*MCM expression and purification protocol (mutants)

*Mth*MCM was cloned into the pET21b expression vector to express a C-terminal His-tagged protein with an uncleavable tag (vector kindly provided by J.P. Gautier, Columbia University). The clones to produce the protein lacking subdomain A and the NBH mutants were a gift from Z. Kelman (University of Maryland). All the proteins were expressed in E.coli BL21 cells.

2.1.2 DNA substrate preparation

The DNA substrates of various length 250bp, 495bp, 807bp, 1027bp, 1200bp, 1350bp have been obtained using Polymerase chain reaction (PCR). PCR is normally used to amplify DNA fragments in the range of 100bp to 10kb (Sambrook et al., 2001). The protocol of reaction requires several components and reagents:

- A DNA template containing the DNA region to be amplified.
- Two primers that are the complementary to the 3' ends of each strand of the DNA target.
- DNA polymerase
- Deoxynucleoside triphosphates (dNTPs)
- Buffer condition suitable for selected DNA polymerase.
- Divalent cations, generally Mg^{+2} is used.
- MilliQ water.

The choice of the DNA polymerase depends on the DNA template. Usually, in order to obtain accurate replication of a desired DNA template, a DNA polymerase with high fidelity is used. In this study, we used Taq polymerase and Phusion polymerase and reaction buffer provided with them.

For the PCR condition, we used touchdown method, by which the primers will avoid amplifying nonspecific sequences.

The PCR product was obtained after 25 cycles using the thermocycling conditions reported in Table 2.1. Next, we purified the PCR product using PCR purification kit (qiagen).

Table 2.1: The thermocycling conditions have been used.

STEP	TEMPERATURE	TIME
Initial Denaturation	98°C	30 seconds
Denaturation	98°C	30 seconds
Annealing	60°C	30 seconds
Elongation	72°C	1 minute and 30 seconds
Final Extension	72°C	5 minutes
Hold	10°C	

We then checked the length of the DNA fragments amplified using Agarose gel electrophoresis shown in Fig. 2.1. It is the most effective method to determine the length of the DNA fragments of varying sizes, ranging from 100 bp to 25 kb. As a result, we obtain DNA fragments with appropriated length.

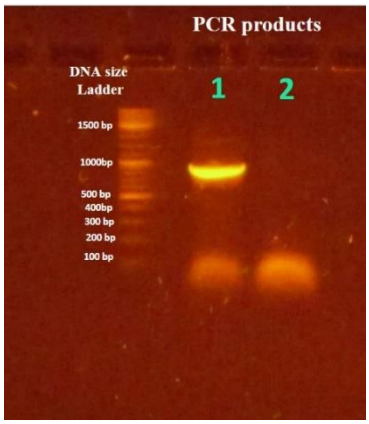


Figure 2.1: Agarose gel electrophoresis of PCR products “not purified yet”, there is a clear band around 900 bp for the PCR product 1, and there is no band for the PCR product 2, which is mean that the PCR product 1 is well produce while 2 is failure and that could be due to missing one of the reaction component.

Note:

The concentrations of the proteins and the DNA fragments were controlled using Nanodrop 2000c. It is a micro-volume spectrophotometer routinely used to check in a small amount of material (microliters) and in small amount of time (few seconds) the concentration of the DNA, RNA and proteins. It is an absorbance based method, the output of Nanodrop measurement is the spectral data and purity ratios that provide the concentration and indicate the sample purity.

2.2 Atomic Force Microscopy (AFM)

AFM is a Scanning Probe Microscopy (SPM) technique introduced in early 1986 by Binnig and colleagues to overcome the limitations of another SPM technique, Scanning Tunnelling Microscopy (STM), which can be used to image conducting materials and therefore hardly applicable to the analysis of surface attached biomolecules, like DNA, RNA and proteins.

As in all SPM, a probe is moved across a surface sample and the morphology of the sample can be reconstructed by exploiting a distance dependence of an observable derived from interaction between the tip and the sample. Whereas in STM the observable between the probe (a metallic, nanometre-sized tip) and the sample is the tunnelling current, limiting then the use to conductive or semi-conductive materials, the basic idea of the AFM is to measure the local attractive or repulsive forces (Van der Waals force, chemical bonding, mechanical friction, electrostatic charge or magnetic interaction) between the tip and the sample (Fig. 2.2). To measure the interaction force, Binnig and colleagues combined the STM principle and the stylus profilometer, using an ultra-small probe tip at the end of a cantilever. One year later, Wickramsinghe and his colleagues developed an AFM setup with a vibrating cantilever, which used an optical lever mechanism to detect the oscillation of the cantilever. This system is the most common snow in use in standard AFM setups (as described in the following section).

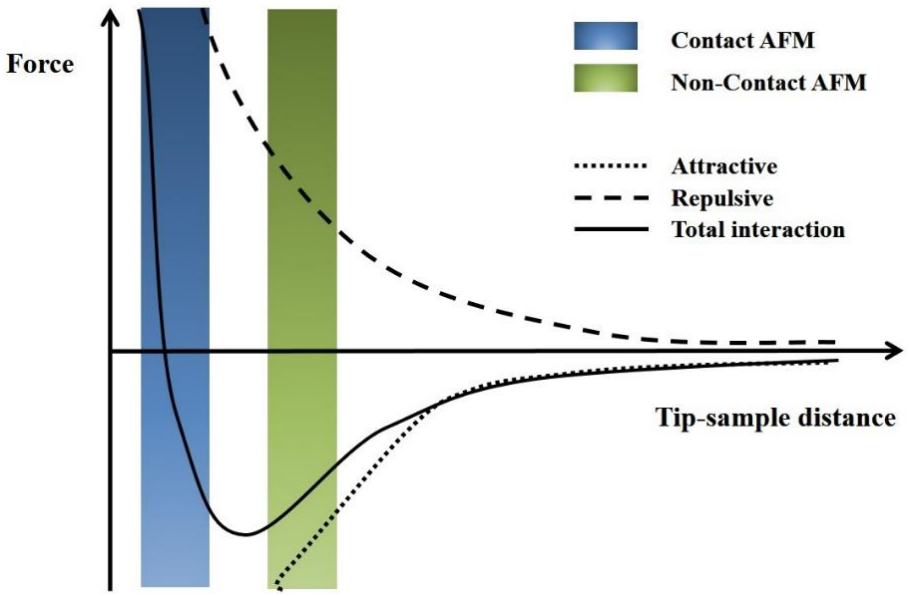


Figure 2.2: Force as function of tip-sample distance

2.2.1 AFM setup

An AFM is composed by the cantilever, a force-signal amplifier, the piezo scanner and the optical detection system, which includes a laser and a 4-quad photo-detector to detect the signal reflected by the back of the cantilever (Fig. 2.3). The position of the laser spot is measured by comparing the signals from the different sections of the detector. Most AFMs use a photodiode that is made of four quadrants, so that the laser spot position can be calculated in two directions, by comparing the signals. The vertical deflection (measuring the interaction force) can be calculated by comparing the amount of signal from the "top" and "bottom" halves of the detector. The lateral twisting of the cantilever can also be calculated by comparing the "left" and "right" halves of the detector.

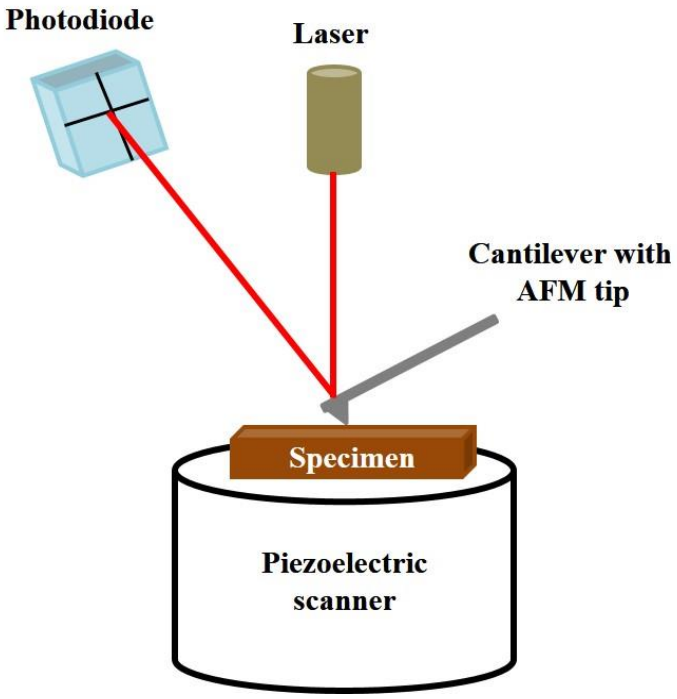


Figure 2.3: Detection of the signal

The cantilever is typically made of silicon or silicon nitride with a sharp tip (probe) radius of curvature of few nanometres at its end that is used to scan the sample surface. Different cantilevers lengths, shapes and materials provide a variety of spring constants (N/m) and resonant frequencies (Hz);

- The resonant frequency is a natural frequency of vibration determined by the physical parameters of the cantilever.
- The force sensitivity of the cantilever is inversely proportional to its spring constant; the lower the spring constant the more sensitive the AFM.

When the tip is brought into proximity of a sample surface, forces between the tip and the sample lead to a deflection of the cantilever according to Hooke's law (Cappella et al., 1999). We can use AFM for detecting the morphology, the stiffness and the frictional forces of any kind of material, in almost any kind of environment.

Two basic instrument designs are possible for AFM imaging:

- The cantilever can be attached to the piezo-scanner for scanning the tip across the sample surface.
- The sample can be scanned across the probe while the tip is held in fixed position.

2.2.2 Imaging modes:

The dominant interactions at short probe-sample distances in the AFM are Van der Waals (VdW) attractive interactions and coulombic repulsive forces. Other long-range interactions (i.e. capillary, electrostatic, and magnetic) are significant further away from the surface, and are therefore usually neglected. In Figure 2.2, the interaction potential as a function of the tip-sample distance is reported. Two main interaction regimes are evidenced in the plot: the one dominated by the repulsive forces and the one dominated by the attractive forces.

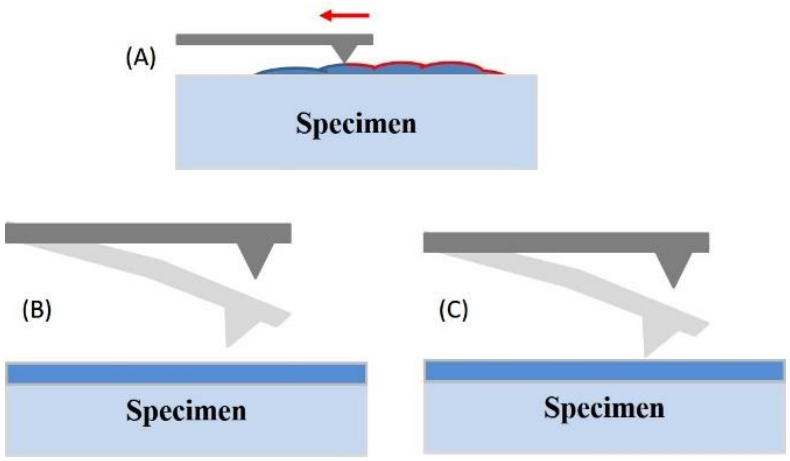


Figure 2.4: AFM imaging modes: A) Contact mode, B) Non-contact mode and C) Tapping Mode.

AFM operation is usually divided to two categories, according to the tip motion (Static and dynamic mode) and the distance from the surface (Contact and Non-Contact modes).

The choice of the operating mode is done as a function of the characteristics of the specific substrate that we would like to investigate and on the type of measurement: soft matter as biomolecules on surfaces are normally measured in the attractive regime, while frictional forces can be measured in repulsive regime only.

2.2.2.1 Static mode

The probe is moved across the sample surface and the static deflections of the cantilever due to the interaction forces are recorded to reconstruct the topography. If the tip-sample distance is kept in the repulsive regime, the operational mode of the AFM is commonly known as contact mode. In contact mode the tip is "dragged" across the surface of the sample and the contours of the surface are measured either using the deflection of the cantilever directly or, more commonly, using the feedback signal required to keep the cantilever at a constant position (Fig. 2.4A). In this operation mode, the measured forces are the highest possible, and it is possible to achieve atomic resolution (Müller et al., 1999). However, on soft samples this mode is destructive due to high shearing forces acting during scanning.

2.2.2.2 Dynamic mode

In the dynamic mode the cantilever is oscillated at its resonant frequency (frequency modulation) or just above (amplitude modulation).

For AM-AFM, the oscillation amplitude and frequency are kept constant during the imaging process. When the tip approaches the sample, elastic and inelastic interactions case changes in both amplitude and the phase signal that used by the feedback loop to map the surface morphology.

In the case of FM-AFM, the cantilever is subjected to a controlled positive feedback such that it oscillates at the resonance frequency with constant amplitude. And here the tip-sample interactions cause frequency shifts of the cantilever resonance.

The choice of the optimal amplitude of oscillation depends on many parameters, such as the selected cantilever and sample roughness. If the sample is flat, small amplitude values (few

nm) can be used, but if it is rough then the amplitude should be larger (tens to hundreds of nanometres) in order to follow the sample morphology without damaging the tip. To avoid instabilities, is better to use the lowest input amplitude possible. Another relevant parameter is the scanning speed that can be limited by the response time of the sensor and the length of the cantilever.

Hereafter we will describe the main dynamic modes: non-contact mode (Fig. 2.4B), semi-contact mode (Fig. 2.4C) and jumping mode.

- *Non-contact mode*

The probe is oscillating with small amplitudes (few nm) to keep the tip-sample distance always in the attractive regime (Gross et al., 2009). Using a feedback loop it is possible to monitor changes in the amplitude due to attractive van der Waals forces and to reconstruct the surface topography. Very low forces are exerted on the surface (few pN), but due to the further tip-sample distance the resolution is not optimal, apart from some special cases (i.e. atomic resolution in liquid (Möller et al., 1999).

- *Semi-contact Mode*

The cantilever is oscillated at its resonant frequency, with an amplitude adjusted so that at each oscillation the probe “taps” lightly the surface once, contacting the surface at the bottom of its swing (Geisse, 2009). By maintaining constant oscillation amplitude, a constant tip-sample interaction is maintained and an image of the surface is obtained. Semi-contact mode has increasingly become more broadly used for biological studies, since allows to increase the spatial resolution while reducing the repulsive interactions.

In fact, this method of "tapping" lessens the damage done to the surface and the tip compared to contact mode. Additionally, there are significant less lateral forces in the tip-sample interaction in tapping mode.

This makes semi-contact mode very indicated for biological applications and several works have been reported on the use of this technique for the imaging of very soft and fragile samples, (Rivetti et al. 1996, Shlyakhtenko et al., 2010, Wegrzyn et al., 2014, Rivetti et al., 2015). Also, incorporated with phase imaging, the tapping mode AFM can be used to analyse the different components of the surface.

- *Jumping Mode (JM)*

From an operational point of view Jumping Mode can be seen as an intermediate between contact and dynamic modes. In JM a sequence of force versus distance curves is acquired at each point of the surface (de Pablo et al., 1998). For every cycle, the tip is first brought in contact with the surface while the feedback keeps the cantilever deflection at the set point value. Then turning off the feedback, the tip is withdrawn from the surface. In order to avoid lateral forces, the tip is moved to the next lateral point only at maximum tip-sample separation, i.e. at zero force level. Once in the new position, the tip is approached again into contact with the surface to start the new cycle. This minimization of lateral forces allows for a better control of tip-sample distance and of the applied forces, making this technique particularly suited for studying biomolecules in liquid (Moreno-Herrero et al., 2003 & 2004). Since for every point force versus distance curves, separation and approach steps need to be performed, the techniques is intrinsically slower than other scanning modes.

2.2.3The resolution in AFM:

- *Vertical resolution:*

It is defined as the minimum height variation on the sample that the machine can detect and its value is limited by any source of noise in the relative vibrations of the tip above the surface (Acoustic noise, thermal vibration and floor vibration, detector and feedback loop system's noise). In the case of AM-AFM the signal is the amplitude of the oscillation. Standardly, the vertical noise δh is defined as the ratio between the noise in the amplitude signal A and the slope of the amplitude with respect to the average distance of the tip-sample "z":

$$\delta h = \frac{\delta A}{\left|\frac{\partial A}{\partial z}\right|} \qquad 1$$

In the case of FM-AFM, low vertical noise is obtained for low noise frequency measurement ($\delta\Delta f$) and a steep slope of the frequency shift as feedback for topography image. It is given by the noise in the imaging signal and the slope of the imaging signal with respect to z (Morita et al., 2012):

$$\delta z = \frac{\delta \Delta f}{\left|\frac{\partial \Delta f}{\partial z}\right|} \qquad 2$$

The tip-sample distance is measured indirectly through the frequency.

The noise introduced in the signal by the detection method is not the only factor that affects the vertical resolution. There is also the derivative of the feedback signal with respect to the vertical displacement, which depends on the tip-sample interaction.

- *Lateral resolution:*

It depends on the tip geometry. The high resolution of the image requires sharp tip. In the figure, we report the schematic view of surface with two objects close each other imaged with two tips of different curvature radius (Fig. 2.5). Due to convolution effects between the objects and the tip in the case of a sharp tip we are able to clearly distinguish the two objects on the surface (B), while for a blunt tip the two objects are hardly distinguishable.

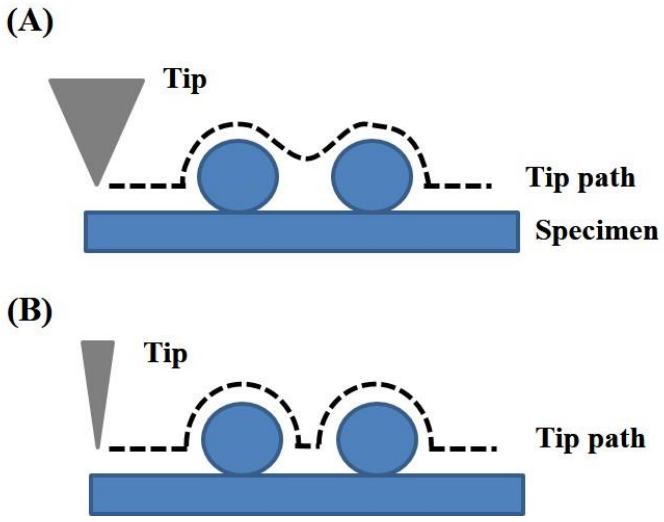


Figure 2.5: Lateral resolution, the image B will have a higher resolution compare to image A, due to the tip radius “tip B is sharper than tip A”.

Mathematically, the lateral separation “*l*” between two sharp features imaged using tip with radius R depends on the height difference ΔH between the adjacent features and the vertical noise δh (equation 1), is;

$$l = \sqrt{2R}\left(\sqrt{\delta h} + \sqrt{\delta h + \Delta H}\right) \quad 3$$

This equation is obtained by assuming that $l > \sqrt{2R\ \Delta H}$. Actually this equation shows the nonlinear nature of the image formation in AFM. Moreover, the lateral resolution is not only depending on intrinsic parameters of the measurement, such as tip radius and the vertical resolution, but also on the height difference between adjacent features (sample's property).

The resolution perpendicular to the surface mainly depends on the sensitivity of the cantilever and the piezo and is of the order 0.1 nm. The resolution that can be achieved using a local probe technique depends on the interaction distance, the decay length of the interaction and on the probe size. As the interaction force is the sum of a number of forces with different distance dependencies, the decay length of the interaction depends on the tip-sample distance. At a long distance, i.e. in the range of several nanometers, electrostatic interactions dominate the tip-sample interaction.

On the other hand, the localization of the interaction, that is the origin of the high resolution, is of the order: $\sqrt{R + {}^{D}/_{k}}$, where D is the effective interaction distance, R the effective probe size, 1/k the decay length of the interaction (Rohrer, 1994).

With respect to air, liquid media can damp vibrations leading to reduced acoustic noise from the background. However, the reduction of the cantilever Q-factor reduces the resolution of the instrument when the AFM is operated in non-contact mode. In buffer solution, moreover, the pH and ionic strength of the imaging media can be adjusted to balance the Van der Waals and electrostatic interactions between the tip and the sample.

2.2.4 Direct visualization of dynamic protein-DNA interaction

Atomic Force Microscopy (AFM) conferred the possibility to image biological samples under physiological conditions with nanometer resolution. However the use of AFM to study processes between active individual molecules became a challenging goal (Hansma et al., 1995, Han et al., 1997).

In order to achieve good quality AFM data in physiological conditions, high requirements are made on the sample preparation, the imaging parameters and overall the AFM system.

In this study both AFM parameters and the buffer conditions allowing imaging of interactions between single MCM complex and DNA with and without ATP are discussed.

2.2.5 High speed visualization

The possibility to operate an AFM in buffer solution, allows investigation of single biological molecules in their natural environment at nanometer resolution. The molecular processes visualization requires a time resolution comparable with the time constant of the process. For most of the commercially available AFMs, the time of a single image of soft biological sample, is in the order of few minutes, which is much slower than most of the biological processes (Van Noort et al., 1999).

Ando and his colleagues started the development of AFM instruments that allow imaging biological samples on the millisecond timescale (Van Noort et al., 1999, Ando et al., 2001, Fantner et al., 2006, Picco et al., 2007). Indeed, through the use of fast piezo scanners and small cantilevers with a low spring constant and a high resonant frequency, they were able to visualize and to study the molecular dynamic behaviour of various biomolecules. .

In this study we used fast AFM “Nanotec Electrónica Cervantes”, that allowed us to record images with scan rate of 4-20 seconds per frame. Indeed, it is still not the range of the dynamics of biological processes but was sufficient to observe the unwinding process for 807bp DNA due to the activation of the MCM complex.

2.2.6 Substrates for DNA-proteins deposition

Since AFM does not require the substrate to be conductive, the choice of substrates is almost unlimited (taking into acount the surface roughness). Both glass cover slips and cleaved mica sheets have been used to adsorb various macromolecules and organelles with varying degrees of success.

The most commonly used substrate is mica (Pashley et al., 1981). Mica is a non-conducting layered material. It is cheap and can easily be cleaved, usually with a pin or a cellotape, to produce clean, atomically flat surfaces up to even millimetres in size. The commonest form of mica is Muscovite KAl2 (OH) 2AlSi3O10 (Brewster et al., 2010).

Mica has been successfully used in numberless studies especially for AFM imaging of double stranded DNA and DNA-protein complexes, protein arrays, and densely packed proteins. Although the mechanism by which macromolecules absorb to this substrate still remains poorly understood, a large number of protein samples adhere tightly to this surface (Kearsey et al., 2003).

2.2.7 Sample preparation

Structural characterization of DNA-protein complexes with AFM is only reliable when the deposition process itself does not affect the conformation of the DNA and of the complexes (Bustamante et al., 1996). For imaging with AFM it is necessary to deposit the DNA-protein complexes on a flat, two-dimensional substrate. Mica is most commonly used as a substrate for DNA deposition (Pashley et al., 1981) as described before.

To study DNA-protein interaction on a surface we need an interaction with the surface strong enough that DNA and the protein complex is not pushed away by the tip during imaging, and at the same time weak enough that DNA conformation is not affected by the presence of the surface. In this way, 3D DNA deposited from solution can equilibrate on the surface to find its most favourable 2D conformation. The DNA is negatively charged in solution due to sugar-phosphate backbone.

In this study, we tested two different substrates: mica treated with Poly-L-ornithine, a synthetic amino acid chain that is positively charged having one hydrobromide group per unit of ornithine (Fig. 2.6) and freshly cleaved mica with a buffer solution containing divalent ions (e.g. Mg^{+2} and Ni^{+2}). Divalent ion selection is highly correlated with the affinity

purification of protein complexes. Most of the protein complex used in this study are in fact recombinant proteins to which a histidine tag is added for purification. His-tag is known to have high affinity for Ni^{+2}. Thus, a buffer containing nickel might affect the binding of DNA-protein complex to mica surface.

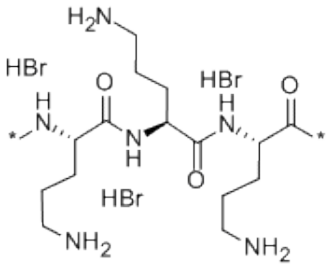


Figure 2.6: Poly-L-Ornithine structure

To prepare mica treated with Poly-L-ornithine we incubated 20 µl of Poly-L-ornithine with 0.05µg/ml concentration on freshly cleaved mica for one minute then rinsed with milliQ water and gently dried using nitrogen gas. In the case of divalent ions, they were just included in the buffer solution.

Table 2.2: The buffer solutions that have been used for imaging in air and/or in liquid (pH: 6.8).

The name	Buffer contains
Buffer A	9mM MgCl2, 90mM NaCl, 30mM HEPES and 5% Glycerol
Buffer B	5mM MgCl2, 150mM NaCl and 30mM HEPES.
Buffer C	1nM $NiCl_2$ and 40mM HEPES
Buffer D	2nM $NiCl_2$ and 40mM HEPES

2.2.7.1 AFM imaging in air

Images were performed in air at room temperature with an MFP3-3D (Asylum Research) and a SOLVER_Pro NT-MDT available in our lab in Trieste, and with a Cypher (Asylum Research) made available at the Physics Department of the University of Göttingen, Germany.

The Cypher allows the use of small cantilevers (less than 10 µm) for fast imaging. Moreover, it is has blue Drive photo-thermal excitation, which produces almost ideal cantilever response. All the images were taken in tapping mode. I normally used NSG30 (NT-MDT) tips (typical resonant frequency of 320 kHz and spring constant of 40 N/m) and NSG03 (NT-MDT) tips (90 kHz and 1.74 N/m).

For the experiments performed in Germany with Cypher, I used AC240 tips with spring constants 2 N/m and resonant frequency 70 kHz.

The samples can be stored in nitrogen box. The samples, once prepared, can be imaged many times and after imaging they are stored as described. Their shelf life is more than a week.

- *DNA sample:*

DNA was diluted to a concentration of 2 nM in a buffer A or buffer B. A 20µl droplet of DNA solution was then incubated on freshly cleaved mica or mica treated with Poly-L-ornithine for one minute, rinsed with 5 ml milliQ water to remove unbound DNA molecules and then dry the sample using nitrogen gas.

- *DNA-protein complex sample:*

DNA is diluted to a concentration of 2 nM in buffer A, and different concentrations of the protein complex (10nM and 20nM). Solutions were incubated at room temperature for 30 minutes. A 20µl droplet of DNA–protein complex solution was deposited onto freshly cleaved mica or mica treated with poly-L-ornithine for one minute, rinse with 5 ml milliQ water to remove unbound molecules and then dry it using nitrogen gas.

2.2.7.2 AFM imaging in liquid

First, using our facilities at Nanoinnovation lab at Elettra, we used MFP3-3D (Asylum Research) operating in tapping mode, to perform measurements at room temperature in liquid. I used Olympus-RC800PSA tip (typical resonant frequency of 69 kHz and spring constant is 0.39 N/m).

For fast imaging, we used the Nanotec Electrónica Cervantes (available in the lab of Dr. Iwan Schaap, Göttingen, Germany). It is very compact layout and stable compared to other AFM instruments (acoustic noise, thermal vibration and floor vibration), and it offers high

resolution with scan speed of few second per frame. Therefore, it can provide the highest resolution scans to study the dynamic of single DNA and/or proteins in liquid.

Images were continuously acquired in tapping and/or jumping modes, reproducibly without visible damage for at least 30 minutes, using a scan size of 1 x 1 μm^2, 512 x 512 pixels, 20s per image. I used AC40TS from Olympus with resonance frequency 25 kHz in water and small spring constant of around 0.1 N/m. The scan frequency was typically 4-8 kHz.

The protocol of imaging in liquid using Asylum AFM:

DNA is diluted to a concentration of 2 nM in a buffer C or buffer D. A 20µl droplet of DNA solution incubated on freshly cleaved mica for ten minutes, rinse with 5 ml buffer C to remove unbound molecules and then put 40µl and process with imaging in this buffer.

The protocol of imaging in liquid using Nanotec AFM:

- Prepare the solution of the DNA/ DNA-protein complex sample using 20mM HEPES and (20-25) mM NaCl.
- Deposit (20-40) µL of the solution on freshly cleaved mica for five minutes.
- Add 2 µL of 10mM $MgCl_2$ for five minutes.
- Add 5 µL of 100mM $MgCl_2$ for five minutes.
- Add 20 µL of 20mM HEPES.
- Add 0.7 µL of 100mM $NiCl_2$.

Image processing and analysis

All the AFM images were processed using 2D *single molecule* software (Roiter et al., 2005) and WSxM (Horcas et al., 2007) software and then graphed with Igor Pro. AFM images are mostly displayed using a colour or grey scale, in which dark is used for low parts and bright for the high parts. Where the DNA contour length (L) and the end-to-end distance (R) measurements were performed as described in Ref. (Rivetti et al., 1996).

2.3 Single-particle analysis

The chain statistic and conformational analysis of DNA molecules deposited into mica surface as well the DNA-surface interaction have been reported below.

2.3.1 Molecular deposition kinetics

During deposition, the molecules transfer from solution (3D) to the substrate (2D) losing one degree of freedom that definitely will reduce the number of possible configurations. To better understand the mechanism through which DNA molecules are transported from the deposition drop to the surface, the DNA adsorption was determined as a function of the deposition time.

The fraction of DNA molecules bound to the surface at any time t is approximately given by;

$$\frac{n_{f(t)}}{n_0} = \sqrt{\frac{4D}{\pi}}\ \sqrt{t} \qquad 4$$

Where $n_{f(t)}$ is the number of molecules on the surface (molecules/ cm^2), n_0 is the number of molecules in solution at t=0 (molecules/cm^3) and D is the diffusion coefficient of the molecules.

On the other hand, the diffusion coefficients of DNA at reference temperature (20°C) (Rivetti et al., 1998, and Refs therein), is given by:

$$D_0 = \frac{8.218*10^{-8}* M^{0.445} + 0.0146}{M} \qquad 5$$

Where M is the molecular mass of the DNA, given by (Brown, 1991):

M.W. of dsDNA = (number of nucleotides x 607.4) + 157.9 6

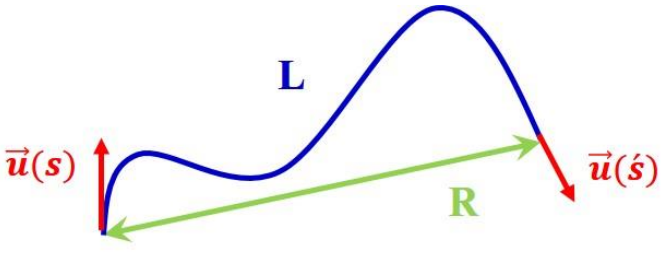
Replacing equation (6) in (5), the diffusion coefficient of 1000bp DNA is:

$$\mathbf{D_0 = 3.314 * 10^{-8}\ cm^2/s}$$

Measuring the number of DNA molecules per image in different time and fitting the data using equation 4, we can obtain the value of the diffusion coefficient D. if the diffusion coefficient obtained from AFM images is comparable to the semi-theoretical value calculated above then the diffusion controlled and therefor it is irreversible process.

2.3.2 Chain statistic

We used the Worm-like chain (WLC) model, also known as the Kratky-Porod model that is used to describe the entropic elasticity of long polymer molecules. The WLC model, treats the polymer as relatively stiff rod made up of a homogenous elastic material. Here we are treating DNA as a polymer as shown in Fig. 2.7.



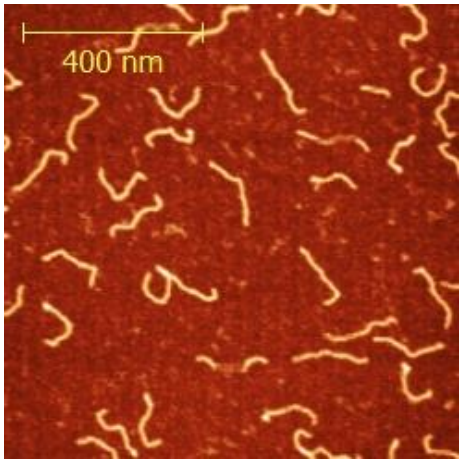


Figure 2.7: DNA 495 bp imaged in air

L = Contour length of a polymer chain is its length at maximum physically possible extension.

R = End-to-end distance, is defined as the distance between the first and the last points of each DNA path.

And $\vec{u}(s), \vec{u}(s')$= Unit vectors tangent at positions ($s\ and\ s'$).

In 2D, one of the important statistical parameters result of WLC model, the average directional correlation "cos (θ)" between two segments of the polymer which decreases exponentially with their separation, can be written as;

$$\langle \cos(\theta) \rangle_{2D} = e^{-\frac{|\acute{s} - s|}{2P}} \qquad 7$$

The stiffness of the chain is described by its persistence length (P), i.e. the decay length through which the initial orientation of the molecule persists.
Using equation 1, the mean square end-to-end distance $\langle R^2 \rangle$ of a worm-like chain of length L and persistence length P (Rivetti et al., 1998), can be written as;

$$\langle R^2 \rangle = \int_0^L ds \int_0^L d\acute{s}\ \langle \vec{u}(s).\vec{u}(\acute{s}) \rangle$$

$$= \int_0^L ds \int_0^L \langle \cos(\theta(s)) - \theta(\acute{s}) \rangle\, d\acute{s} \qquad 8$$

Replacing equation (7) in (8) and the integration gives:

$$\langle R^2 \rangle_{2D} = 4PL\left(1 - \frac{2P}{L}\left(1 - e^{-\frac{L}{2P}}\right)\right) \qquad 9$$

Therefore the end-to-end distance $\langle R^2 \rangle$ and the contour length L are correlated by the persistent length P. Measuring $\langle R^2 \rangle$ and L from single molecule AFM imaging for different DNA lengths, and then fitting the data using equation (9) we can distinguish between DNA molecules that have equilibrated on the surface and the one that have been strongly bound to the surface "trapped", according to the value found for P. The persistence length of free double stranded DNA is in fact known to be in the range of 50 nm (Bustamante et al., 1994). If from the fitting we find a value consistent with this, then we can say that the DNA is freely equilibrate on the surface.

2.4 Protein-DNA studies using AFM

AFM has unique ability to provide structural and functional information on biomolecular interactions, such as protein-DNA interactions at the level of the single molecules, at high resolution in air and under physiological environment as described before. In literature there are several studies of protein-DNA interaction using AFM. The group of Carlos Bustamante has pioneered this field (Bustamante et al., 1996). Major recent contributions are coming from the groups of Claudio Rivetti (Doniselli et al., 2015) and Luda S. Shlyakhtenko (Lyubchenko et al., 2014).

Rivetti and Bustamante in their first review on scanning force microscopy of protein-nucleic acid complexes (Bustamante & Rivetti, 1996) already pointed out that the study of protein-DNA interaction by AFM is reliable only if the deposition process does not affect the DNA-protein complex and if the complexes are allowed to equilibrate on the surface during deposition.

This condition still remains valid after 20 years of work and in their paper the authors underlined the importance of choosing the proper substrate and the proper conditions of immobilization to extract reliable information from the analysis of AFM images of DNA-protein complexes.

Among the choice of the substrates, is therefore important to avoid kinetic trapping effects. In this respect, freshly cleaved mica is usually preferable with respect to glow discharged mica. In fact, the glow discharged mica caused strong surface interactions, while they obtained freely equilibrated DNA molecules for the case of freshly cleaved mica. Moreover

the ionic strength, salt type and the composition of the buffer used for the deposition is also crucial for the correct immobilization of the complexes.

Rivetti, Sushko and co-worker (Sushko et al., 2006) built a simple model for the adsorption of polyelectrolytes (such as DNA) on a like-charged surface (such as mica). In this model they studied the effect of the salt type and concentration on the adsorption of DNA molecules on mica surface, using AFM imaging in air. They tested the effect of NiCl2 or MgCl2 in the absence and in the presence of NaCl. As a result, they found out that $MgCl_2$ concentration needed to be adjusted also with respect to the relative concentration of NaCl, since both salts contribute to the attractive and repulsive forces between DNA and the mica surface.

Once optimized the deposition protocol, it is possible to study the protein-DNA interaction for different protein complexes, determining the bending angle of the DNA in the presence of the protein and its relation with specific and non-specific binding. By analyzing the extracted volume of the individual molecules involved it was possible to extract molecular weight of the multiprotein-nucleic acid assemblies.

Recently, Rivetti's group focused (Doniselli et al., 2015) on the mechanisms of guanosine tetraphosphate (ppGpp) and DksA protein on Escherichia coli RNA polymerase–promoter (RNAP) complex, using two long DNA templates, contain the rrnB P1 or the rrnA P1 promoters. They imaged RPo (polymerase/open promoter) complexes formed at different promoters in the absence and in the presence of one or both modulators (ppGpp and DksA) on freshly cleaved mica using buffer containing MgCl2, using AFM imaging in air. As a result, they observed promoter-specific and non-specific complexes that were distinguished by mapping the relative position of the RNAP along the DNA template, by measuring the DNA contour length. Further, they obtained a statistically meaningful value for the number of complexes with an RNAP specifically bound at the promoter site and the total number of DNA molecules (bound and unbound) under different conditions known as the promoter occupancy.

Shlyakhtenko, Lyubchenko and co-workers (Lyubchenko et al., 2014) studied the specificity of single stranded DNA-binding proteins (SSBs) to its target (ssDNA). Using AFM imaging in air, they engineered a hybrid DNA construct with a ssDNA conjugated to a DNA duplex, and studied the DNA-SSB complex under different ionic conditions. They optimized the protocol suitable for DNA-SSB interaction using mica functionalized with aminopropyl silatrane (APS). In such work they treated the mica surface to immobilize properly the protein-DNA complex. They used two different DNA substrates: a tail-DNA hybrid, where ssDNA is at the end of the construct; a gap-DNA substrate, where the ssDNA

was in between two duplexes. Duplex in their case was playing the role of a marker. In their study, SSB was found to specifically bind to ssDNA tracts, only at high ionic strength or in the presence of Mg^{2+} cations. In the absence of Mg^{2+} cations and under low-salt conditions, the SSB was binding also duplexes, capable of binding DNA duplexes. Moreover, they observed a critical role of Mg^{2+} cations in the specific SSB-DNA binding, regardless the ionic strength (Shlyakhtenko et al., 2012). From this study, we can say that the ionic strength is a critical parameter in the protein-DNA interaction.

In another work, Sun, Lyubchenko and co-workers (Sun et al., 2014) studied the remodelling of RecG helicase (a key player in stalled replication fork rescue) at the DNA fork assisted by SSBs. Here they used the same treated mica (APS-mica) surface as before and a fork DNA substrate. From single molecule analysis, they first studied the interaction of DNA with SSB or RecG alone. Then they loaded the complex, being able to distinguish the two proteins on the DNA strength, and highlighting novel mechanistic insights on the complex interaction involving SSB inducing RecG remodelling. From AFM images analysis, they made the assumption that SSB promoted translocation of RecG along the parental duplex DNA: it binds to the ssDNA arm of the fork, inducing the detachment of the ssDNA-binding RecG protein domain, allowing the translocation of the entire RecG along the dsDNA sequence. Also in this study, the interaction was studied in different buffers such as buffer contains Mg^{2+} cations, and translocation was studied in liquid, in the presence of the ATP (Adenosine Triphosphate) and in the presence of ADP (Adenosine diphosphate). This work inspired us to to study the dynamic interaction of the MCM complex with DNA. Also, it stimulated the idea of studying the effect of all the other proteins of the replication complex, on MCM activity.

Regarding the protein-DNA interaction imaging in liquid: several studies have been performed (Jan Knappe & Szabolcs Soeroes 2010; Lyubchenko et al., 2014; Shlyakhtenko et al., 2013) using different protocols, and fast AFM imaging. For example, Lyubchenko and co-workers studied the EcoRII protein translocation process along the DNA "looped structure" with high-speed AFM, five seconds per frame. They performed these experiments using mica treated with APS. They recorded series of consecutive frames observing the sliding of the protein along the DNA (Lyubchenko et al., 2014).

Following similar schemes to the ones reviewed before, I concentrated on different AFM imaging conditions for studying MCM-DNA interaction. First, I used Solver pro and Asylum MFP 3D for imaging the interaction in air to visualize non-dynamic details of MCM-DNA

binding. Then, I moved to imaging in liquid, loading ATP in the buffer, in order to study the interaction dynamics. For that, I needed a fast AFM and so I moved to the group of Iwan Schaap, where I used Asylum Cypher and Nanotec Inc. machines, which allowed for imaging in air and in liquid, at few seconds per frame.

In spite of the many AFM works reported on DNA-protein interaction, each systems requires its own conditions to work properly. Therefore I dedicated long part of my work to optimize buffer and surface conditions for the specific protein complex, MCM, I was investigating.

3 Surface Optimization

In this chapter, the results obtained on surface optimization for DNA molecules equilibration will be discussed. All the experiments carried out during my PhD work were performed using various length of non-specific blunt-ended DNA sequence, ranging from 250 to 1350bp.

AFM is a surface scanning technique and as such any molecule to be investigated must be deposited onto a suitable substrate for imaging. In this study we used mica, an atomically flat surface (roughness: 0.1 nm), which use as a reference for measuring the topography of DNAs and proteins deposited on it. DNA and mica are both negatively charged in buffer solution, and their interaction needs to be mediated by opposite charges, either absorbed on the surface (e.g. NH^{2+} groups) or present in solution (divalent cations).

Surface optimization has been performed both in air and in liquid. Imaging in air, in fact, has been used in the course of this thesis work to visualize DNA-protein static interaction, with MCM non-active as a molecular motor (i.e. without ATP). Imaging in liquid, instead, has been used to study the dynamics of the interaction, in the presence of ATP.

3.1 Imaging in air

To immobilize DNA onto the mica surface, we tested two different strategies: mica treated with poly-L-ornithine and freshly cleaved mica as described in chapter two.

3.1.1 Mica treated with Poly-L-ornithine

We started imaging two different DNA fragments of 250 bp and 1000 bp (Fig. 3.1) on Poly-L-ornithine treated mica, following the protocols described in chapter two. The measurement performed using MFP-3D operated in tapping mode, NSG30 (NT-MDT) tip, 4 minutes per frame. Stiff cantilevers are typically used, as tips can get stuck in the water contamination layer. Once we optimized the buffer condition for imaging in air we used a soft cantilever.

Analysing 208 DNA molecules, we measured the DNA contour length (L) and end-to-end distance (R) from recorded images, as described in chapter two. We reported the results by plotting the mean square end-to-end distance $\langle R^2 \rangle$as function of the contour length (L) and we fitted the data collected from AFM images using equation 9 (Fig. 3.1E).

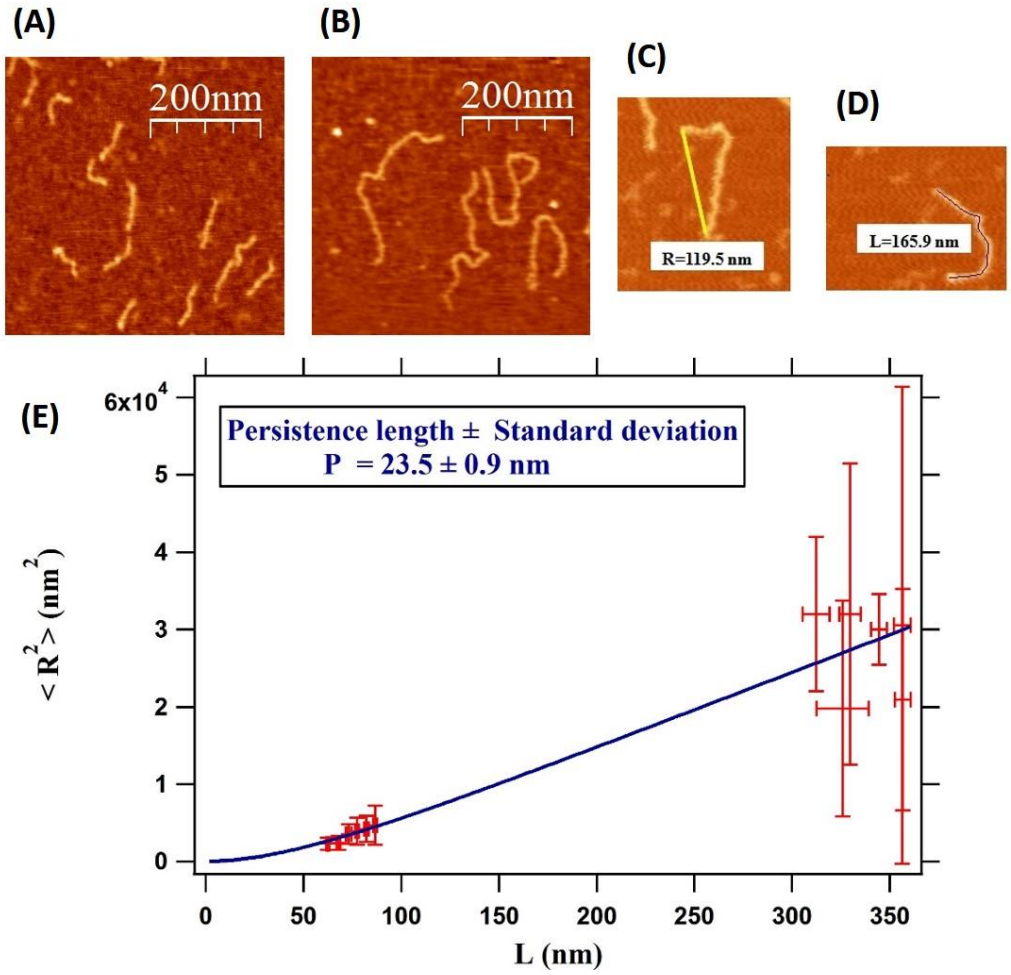


Figure 3.1: DNA molecules on mica treated with Poly-L-ornithine. A) AFM image of DNA 250 bp, B) AFM image of DNA 1000 bp, C) The line drawn to measure the end-to-end distance (yellow line), D) The tangent drawn to measure the DNA contour length (think black line), both C and D the measurement in nanometer that we measured for each DNA molecule, and E) The mean square end-to-end distance $\langle R^2 \rangle$ as a function of contour length L. The red points are the data collected from AFM images and the blue line is the fitting curve using equation (9).

For the long DNA fragment (1000 bp) the variability of the measured end-to-end distance is high (error bar that we measured is: 64%). The length of this fragment in fact corresponds to about 340 nm, much bigger than the persistence length of the double stranded DNA, which is about 50 nm (Bockelmann et al., 1998; Bustamante at al., 1994; Chi et al., 2013). It is therefore expected that this flexible DNA fragments will adopt on the surface different conformations, as observed in Fig. 3.1B, and as a consequence generate a broad distribution of end-to-end distances, which in turn will lead to large errors.

From this fitting, we extract a value for the persistence length of 23.5± 0.9 nm (Fig. 3.1E) that is much smaller than the expected one, indicating the strong interaction of the DNA

molecules with the Poly-L-ornithine treated mica surface. Indeed, the concentration of poly-L-ornithine was varied, in order to tune the number of NH^{2+} pinning centers. Data reported here are relative to the lowest possible concentration that allowed to obtain DNA molecules bound to the surface. These results indicated that the Poly-L-ornithine surface strongly influenced the DNA binding and conformations, as can be confirmed by DNA deposition kinetics studies described in the next paragraph.

3.1.2 DNA deposition kinetics

Rivetti, Bustamante and co-workers demonstrated that in the case of DNA deposition on mica at low salt concentration, the deposition process was governed only by diffusion (Lang et al., 1968). As described in chapter 2.3.1 the fraction of molecules adsorbed at a given time on the surface is a function of the square root of the deposition time through a coefficient which is connected to the diffusion coefficient (equation 4). To test whether with Poly-L-ornithine we were working in diffusion-limited conditions, we deposited the DNA molecules on Poly-L-ornithine treated mica and on freshly cleaved mica, incubated for 10, 30, 60 and 90 seconds. From the analysis of AFM images we derived the density of DNA molecules (per cm^2) bound to the surface. Increasing the deposition time, we observed the increasing of the number of DNA molecules per image. We then plot the ratio between surface-bound DNA density and the total number of molecules per cm^3 as a function of time, and fitted the data using equation 4. This equation is valid if: first, the DNA molecules are irreversibly adsorbed to the surface; convection current does not contribute to the transport of the molecules to the surface; the deposition drop (contains the DNA molecules) is not significantly evaporated during the time of deposition (Rivetti et al., 1996).

We performed the measurement for the DNA 1000 bp on mica treated with poly-L-ornithine. The semi-theoretical diffusion coefficient depends on the the molecular mass of the DNA that we calculated using equation 5 described in chapter two (session, 2.3.1), we obtained a value equal to $3.314 * 10^{-8}$ cm^2/s.

From the analysis of our data with equation 4, we extrapolated a diffusion coefficient value of $(4.9 \pm 0.7) * 10^{-7}$ cm^2/s that is one order of magnitude higher than the semi-theoretical value described before (Fig. 3.2B), demonstrating that the deposition process on Poly-L-ornithine is not diffusion controlled.

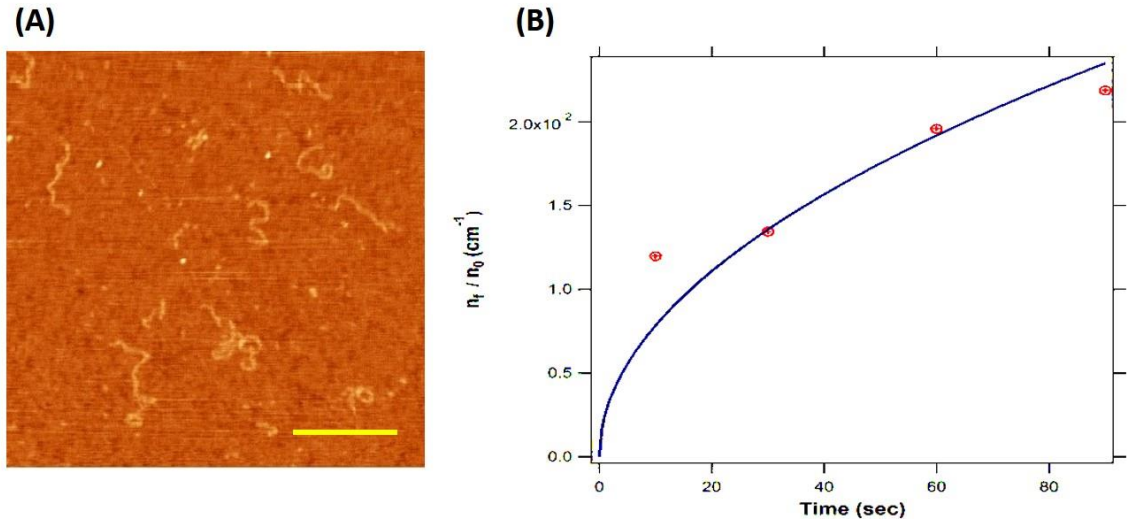


Figure 3.2: Mica treated with poly-L-ornithine, the DNA incubated using buffer B for 60 sec, A) AFM images of DNA 1000 bp and B) The ratio between the number of DNA molecules per cm^2 and the total number of molecules per cm^3 as a function of time is plotted. A diffusion coefficient D = $(4.9 \pm 0.7)*10^{-7}$ cm^2/s is derived. The blue line is the fitting curve using equation (4). The bare scale is 250 nm.

Reasonably, the density of NH^{2+} pinning centres is too high on this functionalized mica surface preventing the dsDNA molecules to equilibrate on the surface. Therefore, this strategy is not suitable to study DNA or DNA-protein interaction.

We repeated the same experiments on freshly cleaved mica using buffer B, obtaining a diffusion coefficient value of $(2.2 \pm 0.2)*10^{-8}$ cm^2/s that is close to the semi-theoretical value and which therefore points to molecular equilibration on the surface (Fig. 3.3B).

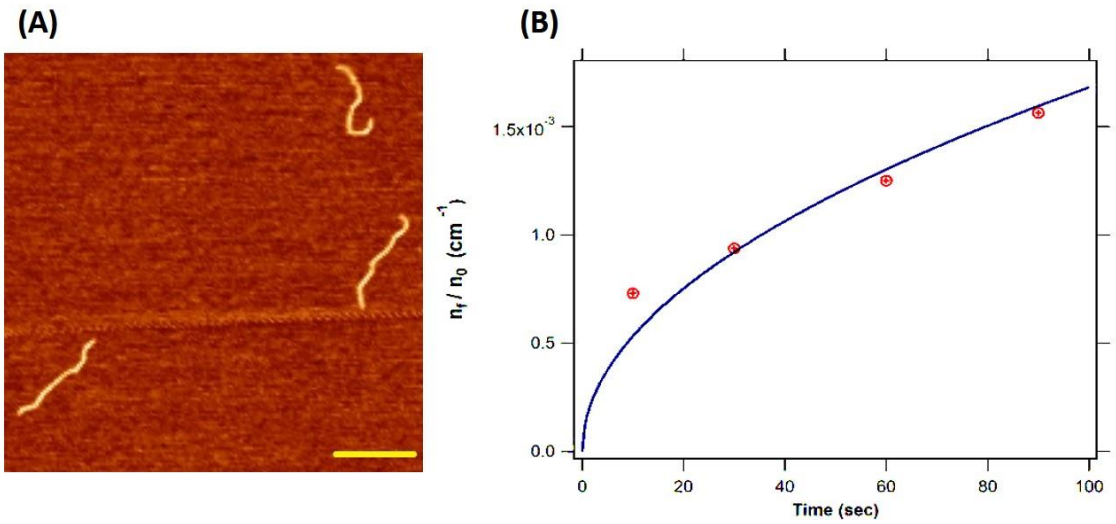


Figure 3.3: Freshly cleaved mica, the DNA incubated using buffer B for 60 sec, A) AFM image of DNA 1000 bp and B) The ratio between the number of DNA molecules per cm^2 and the total number of molecules per cm^3 as a function of time. D =$(2.2 \pm 0.2) * 10^{-8}\ cm^2/s$. The red points are the data we collected from AFM images and the green line is the fitting curve using equation (4). The scale bar is 250nm.

Note:

From the results we obtained from diffusion kinetics studies for the two different mica surfaces, we concluded that of the two conditions explored here, only freshly cleaved mica guaranteed a diffusion controlled deposition process in which the molecules are irreversibly adsorbed onto the substrate, in a freely equilibrated conformation. From this moment on, all the experiments performed in this thesis work have been done on freshly cleaved mica.

3.1.3 Freshly cleaved mica

Similarly to what was done on Poly-L-ornithine, we tested two DNA fragments of 250 bp and 1200 bp (Fig. 3.4) on freshly cleaved mica using two different buffer conditions, buffer A and B. The experiments performed using MFP-3D operated in tapping mode, NSG30 (NT-MDT) tip, 4 minutes per frame.

Therewith, we measured the DNA contour length (L) and end-to-end distance (R) and plot the mean square end-to-end distance as function of the contour length (L), as already done. Results for freshly cleaved mica in buffer B (5mM MgCl2, 15mM NaCl and 30mM HEPES) are shown in Fig. 3.4C. By fitting these data, we obtained a persistence length of 42.8±6.6 nm, in quite good agreement with the literature (Bockelmann et al., 1998; Bustamante at al., 1994; Chi et al., 2013).

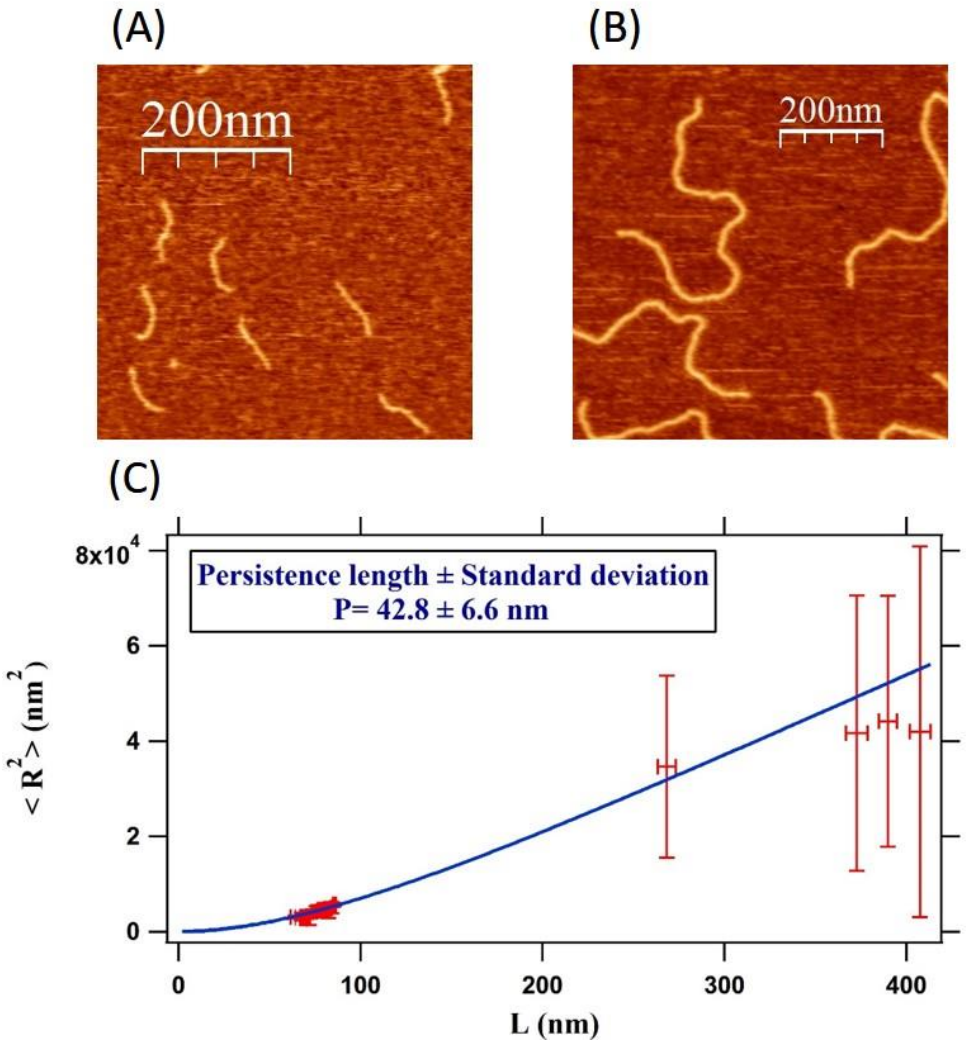


Figure 3.4: DNA molecules on freshly cleaved mica using buffer B. A) DNA 250 bp, B) DNA 1200 bp and C) the mean square end-to-end distance $\langle R^2 \rangle$ as function of contour length L. The red points are the data we collected from AFM images and the blue line is the fitting curve using equation (9).

Again the error bar for the long DNA fragment (1200 bp) is quite huge, as discussed in the previous session.

Thus, we can conclude here that the molecules are irreversibly adsorbed onto the substrate (Fig. 3.3B) and that the persistence length measured from AFM images is in very good agreement with that of worm-like chain polymers at equilibrium in two dimensions (Hamon et al., 2007; Rivetti et al., 1996). Therefore freshly cleaved mica is a suitable substrate to study DNA and protein-DNA interaction which will be addressed in the following chapters.

Furthermore, we optimized the buffer condition by adding Glycerol (Buffer A: 9 mM MgCl2, 90 mM NaCl, 30 mM HEPES and 5% of Glycerol) which is known to enhance protein stability, and related solutes.

We tested here three different DNA fragments, with lengths of 495 bp, 807 bp and 1200 bp, as shown in Fig. 3.6. We obtained a DNA average height of 0.6 nm for the different DNA fragments. This height is very low compare to the expected value, which is 2 nm (Moreno-Herreo et al., 2003), due to the layer of adsorbed salt on the surface that embeds the DNA molecules.

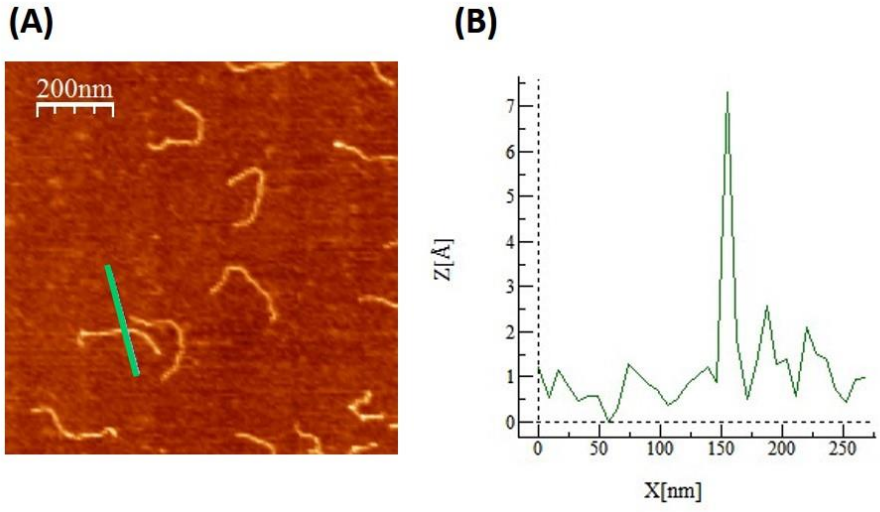


Figure 3.5: DNA imaged in air using buffer A, A) AFM image of DNA 807 bp and B) the DNA height profile corresponding to the molecule highlighted in green in panel A: the DNA has height of 0.7 nm, much smaller than the real width of the DNA double helix (2nm).

In these new conditions, we obtained a persistence length value of 45.0±0.3 nm (Fig. 3.6D) that is close to the expected value (Bockelmann et al., 1998; Bustamante at al., 1994; Chi et al., 2013), confirming that the presence of glycerol is not affecting DNA binding and binding kinetics.

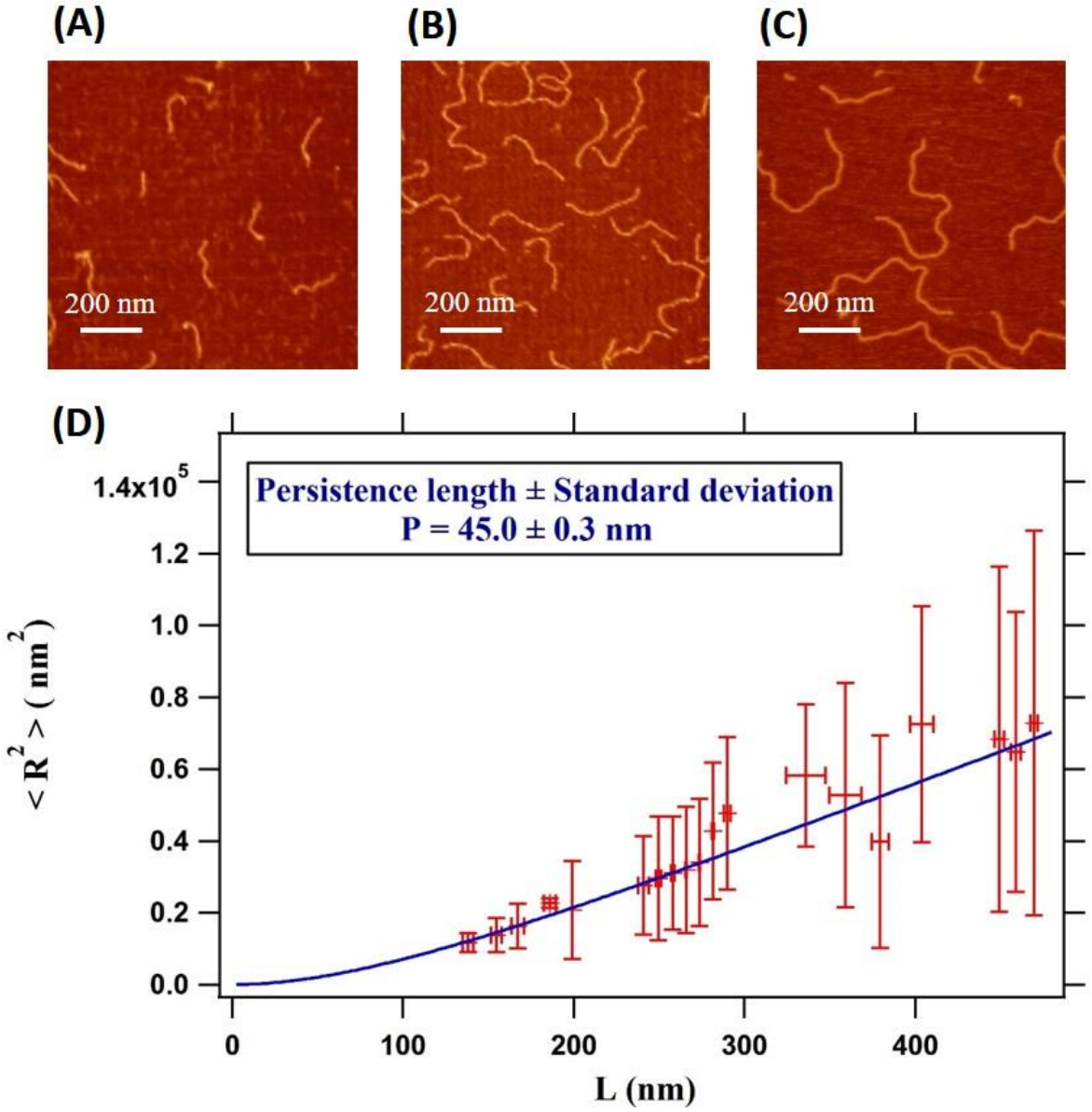


Figure 3.6: DNA molecules on freshly cleaved mica using buffer A. A) DNA 495 bp, B) DNA 807 bp, C) DNA 1200 bp and D) the mean square end-to-end distance $\langle R^2 \rangle$ as function of contour length L. The red points are the data collected from AFM images and the blue line is the fitting curve using equation (9).

Therefore, from now on we will use these buffer conditions for protein-DNA interaction, and imaging in air.

3.2 Imaging in liquid

To image DNA molecules on freshly cleaved mica in liquid, we first used the same buffer A that we optimized for imaging in air (Section 3.1). Since we didn't observe any DNA molecules bound to the surface in these conditions, we assumed that the concentration of $MgCl_2$ within the volume of the liquid cell was not enough to trap the DNA under liquid measurement. Therefore, we increased $MgCl_2$ concentration from 5 mM to 10 mM and then to 20 mM. In both cases we observed few DNA molecules that were not stable during scanning. In order to overcome this limitation, we tried other divalent ions, such as Ni^{2+} which have been extensively used for imaging in air (Herrero-Galàn et al., 2013) as well as in liquid (Pyne et al., 2014).

We optimized the experimental conditions and designed a protocol using buffer C (contains: 1 mM $NiCl_2$) for both immobilizing and imaging the DNA sample. Using MFP-3D (Asylum Research), we recorded good resolution images as shown in Fig. 3.7, with scan rate of 60 seconds per frame, using an Olympus-RC800PSA tip with spring constant of 0.39 N/m.

The measured DNA height was 2 nm, in very good agreement the expected (Moreno-Herreoet al., 2003). On the other hand, we observed that the 1027 bp DNA molecules were moving during scanning, as shown in Fig. 3.7. This result, indicated that the DNA is not completely trapped and more ions required to keep the DNA molecules stable during imaging under liquid conditions.

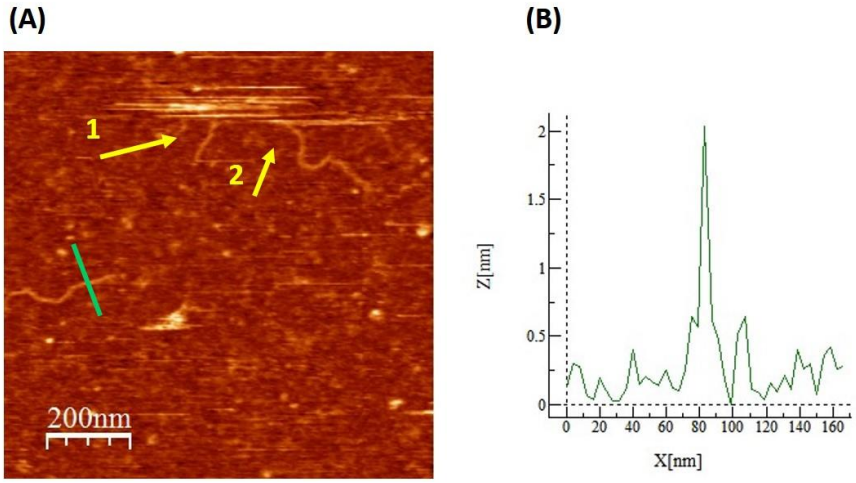


Figure 3.7: DNA 1027 bp imaged in liquid using buffer D, 60 sec per frame. A) The number 1 and 2 point the DNA molecules that were moving during scanning. B) The height profile for the selected molecule is 2 nm which in good agreement with the diameter of double stranded DNA (Moreno-Herreoet al., 2003).

We therefore decided to slightly increase the concentration of $NiCl_2$ to 2 mM (buffer D) and perform again the experiments. In this new buffer we obtained good resolution and stable DNA molecules, imaged at 60 sec per frame. For the height profile of the DNA, we got an average height of 1.6 nm (Fig. 3.8). This molecular height is comparable to the real width of the double stranded DNA (2 nm).

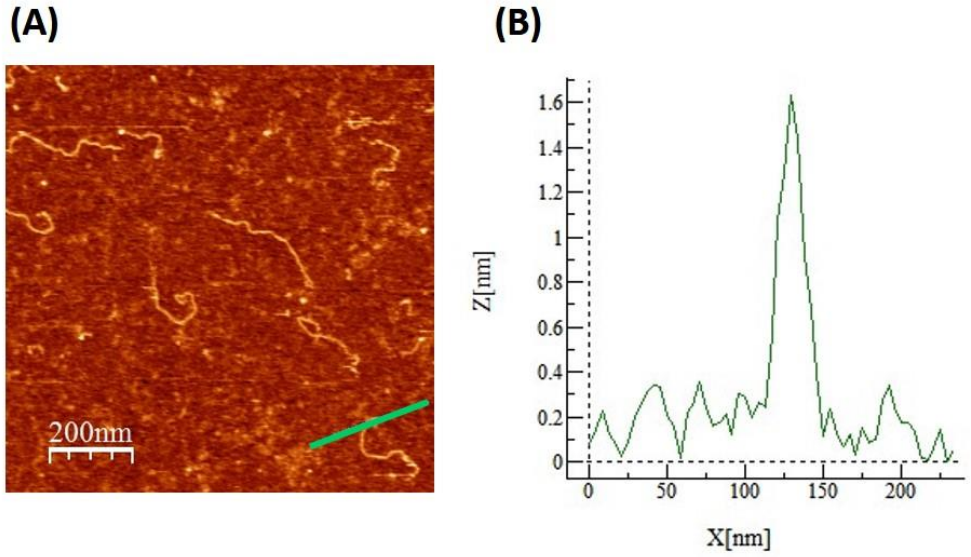


Figure 3.8: DNA imaged in liquid using buffer D, A) DNA 1027 bp and B) the DNA height profile, the DNA has height of 1.6 nm.

Good resolution and DNA height is then highly depending on the buffer solution and the sample preparation that we successfully achieved, beside of course the type of tip used for imaging.

For studying the interactions dynamics however, the experiments require stable DNA binding but not irreversible trapping as we have obtained for imaging in air. Also, fast imaging is an important requirement, which provides an increase in the resolution, due to the low forces applied to the sample and a reduction in scan time (Rico et al., 2013). Towards this aim we used a fast Nanotec AFM, which offers high resolution and scan speed of one second by frame. With buffer D, however, we were not capable to perform fast imaging. Moreover, adding too much $NiCl_2$ to the buffer solution was not recommendable, since it can affect the protein-DNA interaction (Shlyakhtenko et al., 2012).

Imaging in liquid, using buffer containing $NiCl_2$ have been optimized to study DNA-nucleosomes on mica (Jan Knappe & Szabolcs Soeroes 2010). Following their protocol, we

re-optimized the buffer condition. The new buffer contains (1 mM of NiCl2 and 7.4 mM of MgCl2) that provides stable DNA molecules for at least 30 minutes.

We performed experiments using this new protocol, recording good resolution images at 4-8 seconds per frame as the examples in Fig. 3.9 and Fig. 3.10, using an AC40TS Olympus tip with a spring constant 0.1 N/m. The DNA molecules were stable for at least 30 minutes. Moreover, this machine offers the possibility of injecting new buffer, protein complex and/or ATP solution in the open liquid cell during scanning. We prepared the sample according to the protocol described in chapter two (2.2.7.2).Five minutes after incubation, we started imaging: the same area was imaged for three consecutive times, at 4 seconds per frame (an example is shown in Fig. 3.9; the arrow shows the part of the DNA that is slightly moving).

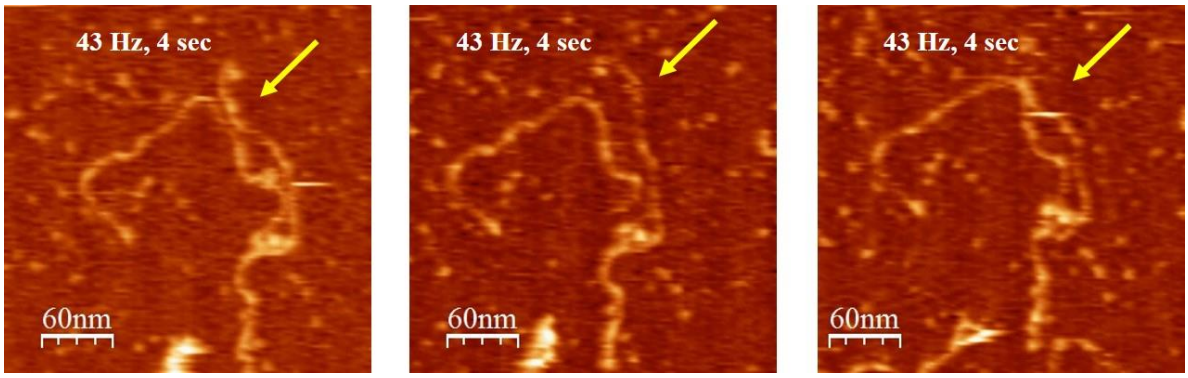


Figure 3.9: DNA 807 bp imaged in tapping mode, with Nanotec AFM using an Olympus AC40TS tip, in liquid. The series images after 5 minutes form incubation, 4 seconds per frame. 64X64 pixels.

After two minutes, we imaged again the same area, as shown in Fig. 3.10, for other three times, at 8 seconds per frame for the first two images and the 6 seconds per frame for the last one. The arrow shows the part of the DNA that is slightly moving.

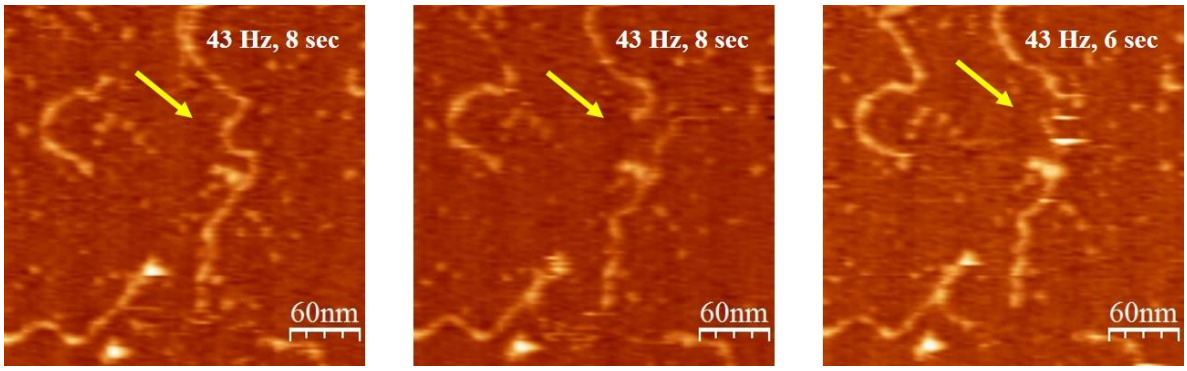


Figure 3.10: DNA 807 bp imaged in tapping mode, with Nanotec AFM using an Olympus AC40TS tip, in liquid. It is the same sample showed in Fig. 3.9, series images after 7 minutes form incubation, 6 or 8 seconds per frame. The DNA molecules seem stable under fast imaging. The first and the second image from the left, contains 128X128 pixels. And the image in the right, 64X64 pixels.

In conclusion, we demonstrated to be able to visualize DNA molecules in physiological condition, at high resolution, with few second/frame imaging. These are suitable condition for protein-DNA interaction, where a sub-seconds timescale might be required to visualize the conformation and the dynamic of the interaction. In the next two chapters I discuss how I used this protocol for imaging in liquid, in order to visualize the protein-DNA interaction and possibly their dynamic under physiological condition.

4 Protein-DNA interaction

As discussed in section 1.2.3, there is some evidence from both biochemical studies (Bowers et al., 2004) and structural analysis (Costa et al., 2008) that there are two distinct modality of interaction between MCM and dsDNA. The aim of the work presented in this chapter is to confirm this and highlight the structural differences between MCM loaded on DNA, with the DNA filament passing through the complex central hole, and loosely associated to it.

In this chapter, the results obtained DNA interacting with the Archaeal MCM protein complex and its mutants (ΔsA complex, lacking subdomain A, and NβH, lacking key positively residues in the central channel), will be discussed. We studied these interactions in the absence of ATP, using the protocols optimized for imaging in air and in liquid show in Chapter three. The choice of mutants was based on previous work which has established that the "association mode" require subdomain A and the "loading" requires an integral N-terminal β-hairpin (NβH).

We optimized the surface and obtained optimal buffer conditions for imaging in air and in liquid using various DNA fragment lengths. From now one we used 807 bp DNA, which is long enough to study the protein-DNA interaction presented in this chapter as well the interaction dynamic will be discussed in the next chapter. Moreover, it is short enough to not form knots that will make then difficult DNA conformation studies.

4.1 Imaging in air

After having optimized the imaging condition in air using freshly cleaved mica, the next step was to study DNA protein interactions. For that, we carried out our studies on 807 bp DNA fragments and Archaeal MCM protein complex and its mutants.

In the experiments we used different protein concentrations. We used the DNA concentration of 2 nM and then we varied protein concentration 2 nM, 10 nM and 20 nM. All the images shown in this chapter were selected from the experiments performed using 10 nM protein concentration. We selected this concentration because it provided optimal results that allowed us to visualize the interaction with a reasonable number of molecules per image.

4.1.1 DNA-MCM interaction

- *Imaging of DNA-MCM complex:*

We imaged the interaction between DNA and MCM complex (1:5) using the buffer A conditions for imaging in air as described in chapter three, at room temperature. In image Fig. 4.1, we can see that among the visualized DNA molecules, only few of them have one MCM complex bound to it. This complex appears with an average height of 2.5 nm, and is prevalently bound to the end of the DNA filament.

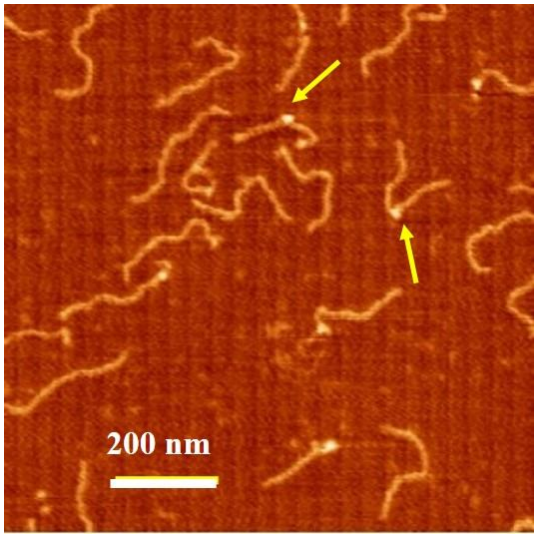


Figure 4.1: The MCM proteins in complex with 807 bp DNA molecule (1:5) on freshly cleaved mica.

Furthermore, in the same sample we observed that DNA molecules have MCM complex bound in different positions along the DNA, as shown in Fig. 4.2 pointed by yellow arrows. 15% of the total number of the MCM complex was bound to the DNA as shown in Fig. 4.2. And we observed only few of MCM complexes bound to the surface and in some case there are no proteins bound to the surface as you can see in Fig. 4.2. This indicates that the rest of MCM complexes were in the solution and they washed away during deposition process. Moreover, we observed that there is only one MCM complex per strand. These observations were independent on protein complex concentration.

- *Height measurements:*

From the images, the measured height profiles are on average 0.4 nm for DNA molecules and of 1.0 nm for the MCM complex (Fig. 4.2), These values are much smaller than the one available from the literature, around two nanometre for the double stranded DNA (Moreno-Herreoet al., 2003) and ten nanometre for the MCM single hexamer (Costa & Onesti., 2009).

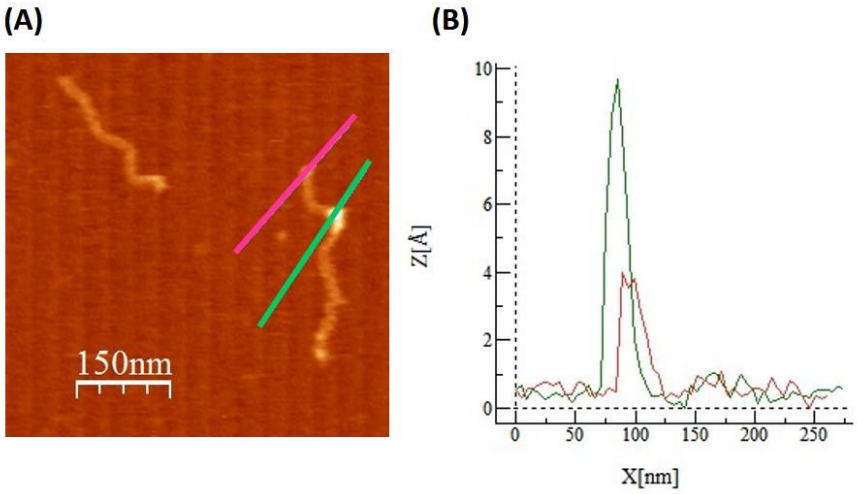


Figure 4.2: DNA-MCM complex (1:5) on freshly cleaved mica, A) DNA 807 bp interacted MCM complex and B) The height measurements of the DNA and MCM complex. The peaks values, 0.4 nm for the DNA and 1.0 nm for the MCM complex. Both heights are very small compare to their real heights (Moreno-Herreoet al., 2003).

The low height measured is due to the layer of adsorbed salt on the surface that embeds the DNA molecules. Heights achieved are reasonable since we carried out this measurement in air (Moreno-Herreoet al., 2003).

- *Models of interaction:*

Analysing more than 100 images, we observed different modalities for the interaction and we divided them into to two categories. The first category represents the binding of the protein complex on the DNA without affecting its conformation (Fig. 4.3A). We made the assumption that when no change in DNA conformation is observed, the complex is considered to be “loaded” to DNA. In fact, recalling Fig. 4.3A, in this configuration MCM is ready to work as a molecular motor: DNA is passing inside the central hole of the complex, and when fuelled with ATP, the complex would change conformation, walking along the strand to unzip it.

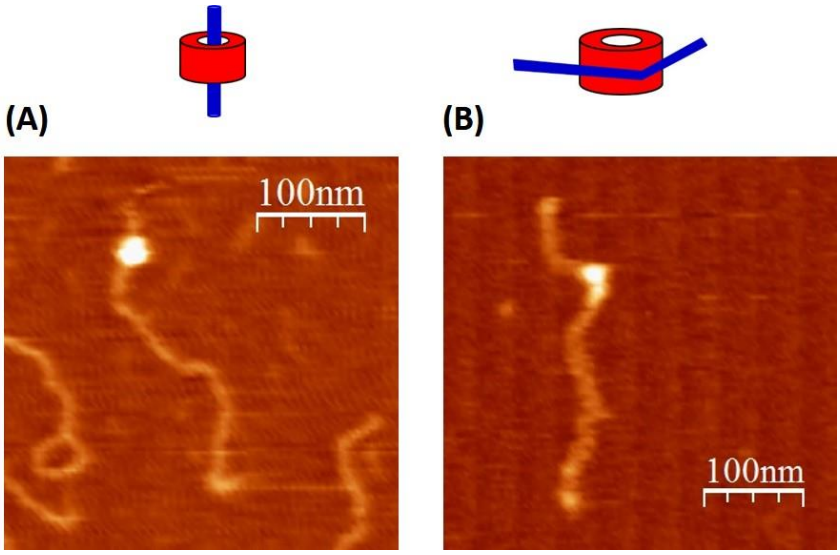


Figure 4.3: The possible interactions of MCM-DNA complex (1:5), A) Loaded MCM, the DNA passing through the central hole of the MCM. Here the MCM bound to the end of the DNA fragments. The free scale is 100 nm. B) Associated MCM, the DNA wrapping around the external part of the structure. In the top of each image there is the representative sketch for corresponding its interaction.

The second category refers to the interacting complexes for which the binding of the protein complex to the DNA induces a bending (kink) of the DNA. We made the assumption that MCM complexes belonging to this category are in the "associated" configuration (Fig. 4.3B): DNA is wrapping around the external part of the MCM (Costa et al., 2008) inducing a DNA bending. This is compatible with the biochemical observations as well the studies performed on DNA-MCM interaction using cryo-electron microscopy (cryo-EM), as described in chapter one. By the fact that we can distinguish these two classes of interaction, we can conclude that the buffer condition we used in this study, is not affecting the DNA-protein interaction.

- *Statistical analysis:*

We analysed 1517 DNA molecules, finding 84% of the DNA molecules deposited on the surface interacting with the MCM complex. According to the definition we made for loaded and associated configurations, we calculating the percentage of loaded and associated MCM on the basis of kink formation on the DNA, out of the total number of MCM bound to the DNA molecules, as reported in table 4.1 for two experimental conditions: DNA:MCM (1:5) and (1:10).

Table 4.1: The percentage of loaded and associated configuration of DNA-MCM complex.

DNA : MCM DNA concentration = 2nM	Loaded %	Associated %
1:5	89.6 ± 1.6	10.4 ± 1.6
1:10	88.2 ± 0.7	11.8 ± 0.7

We found out that in both cases around 90% of the MCM complexes are loaded onto the DNA and only 10% of them is associated to the DNA. We observed that most of the loaded configuration, the MCM complex is bound to the end of the strand (Fig. 4.1). To bind to DNA in fact, some of the areas of contact between adjacent monomers should change configuration allowing DNA to enter into the central hole (see Fig. 4.3A). For this to happen, ATP is needed. The fact that we observe DNA loading even without supplying ATP, is related to the fact that in cells MCM is actively loaded onto the DNA via a complex network of loading factors that are expected to open up the MCM hexamer to thread the DNA through. In the absence of those factors, the MCM ring does not open and loading probably occurs only via the free DNA ends. It is also connected to the inability of the protein to "walk" along the DNA in the absence of ATP (discussed in chapter one). While for the associated configuration, we observed only 10% of the MCM associated to the DNA, with the DNA wrapping around the MCM complex (see Fig. 4.3B). That could be due to the fact that the associated mode is less stable compare to the loaded mode. Further, could be refer to the hypothesis that MCM complex associates to the DNA and then loads into it: since we pre-incubated the DNA-protein complex in solution for 30 minutes it is possible that the protein changes from associated to loaded configuration (more details see chapter one).

4.1.2 DNA-ΔsA interaction

- *Imaging DNA-Δ*sA* *complex:*

In order to justify the assumption we made to distinguish loaded/associated MCM based on AFM topography imaging of DNA conformational changes, we tested specific MCM mutants. In particular, we started with the ΔsA complex, described in Chapter one, which has the mutation in the N-terminal "subdomain A", is located in the external part of the MCM complex structure. Because of this mutation, ΔsA can only load into the DNA. We imaged the interaction between DNA and the ΔsA complex in ratio (1:5) using the same buffer conditions used for imaging in air, as described in chapter three, at room temperature. Only 6% (see table 4.2) of the proteins interacting with DNA are in associated conditions, versus the 10 % measured for wild-type MCM. Also in this case, we observed that the proteins are preferably bound to the end of the DNA fragment (Fig. 4.4) as for the wild-type MCM complex described in the previous session.

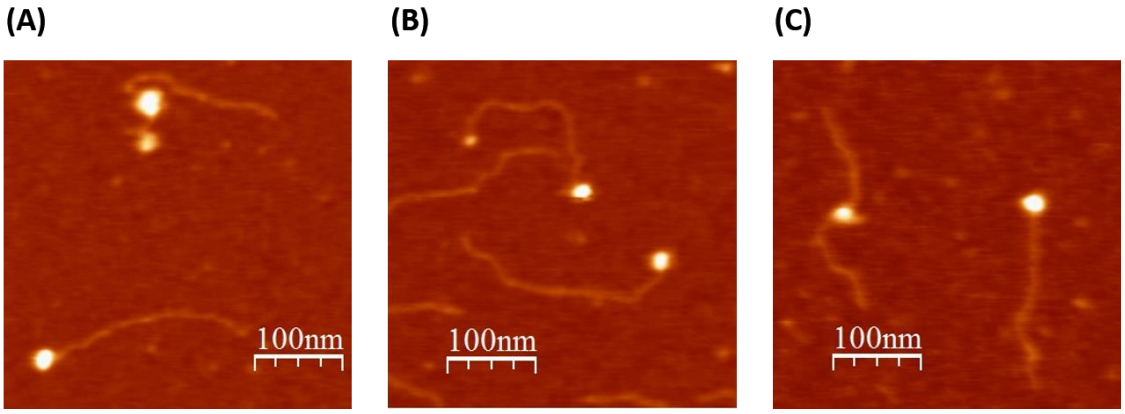


Figure 4.4: DNA-ΔsA complex on freshly cleaved mica. Most of the ΔsA complexes bound to the end of the DNA.

- *Statistical analysis:*

A very streaky result is that out of the 1566 DNA molecules analysed in this experiment, only 20% of the DNA was found interacting with the ΔsA complex, against the 84% observed in the case of wild-type MCM. The biological implication of this observation is that the subdomain A has a critical role in initiating the interaction between the protein complex and DNA molecule. Reducing the "associate" binding capability of the complex, as in the case of the ΔsA mutant, affects strongly also it's "loading" capability. This is also

proved by the fact that most of the ΔsA complexes are found at the very end of the DNA fragments, the "easy" site from which to load the strand.

Table 4.2: The percentage of loaded and associated configuration of DNA-ΔsA complex.

DNA : ΔsA DNA concentration = 2nM	Loaded %	Associated %
1:5	94.0 ± 2.0	6.0 ± 2.0
1:10	97.7± 0.7	2.3 ± 0.7

We repeated these experiments for a DNA: ΔsA ratio (1:10). We found that in this case the percentage of associated DNA reduces to few percentages, as can be seen in table 4.2. While the percentage of the DNA bound to ΔsA, we obtained a number double compared to the result obtained for 1:5 ratio.

- *Height measurements:*

The measured average height profiles were of 0.7 nm for DNA molecules and of 3.5 nm for the ΔsA complex (Fig. 4.2), are still smaller than the values found in the literature value for both the double stranded DNA (Moreno-Herreoet al., 2003) and the ΔsA complex (Costa & Onesti, 2009). However, in this case the AFM images as the one of Fig. 4.5 were acquired using a Cypher that mounted soft cantilevers (AC240TS Olympus tip). Therefore, we achieved heights of the DNA and ΔsA complex that are better than the heights we obtained in the study of DNA-MCM interaction with Asylum using stiffer cantilever, but still we have small height due to the tip-sample forces for imaging in air.

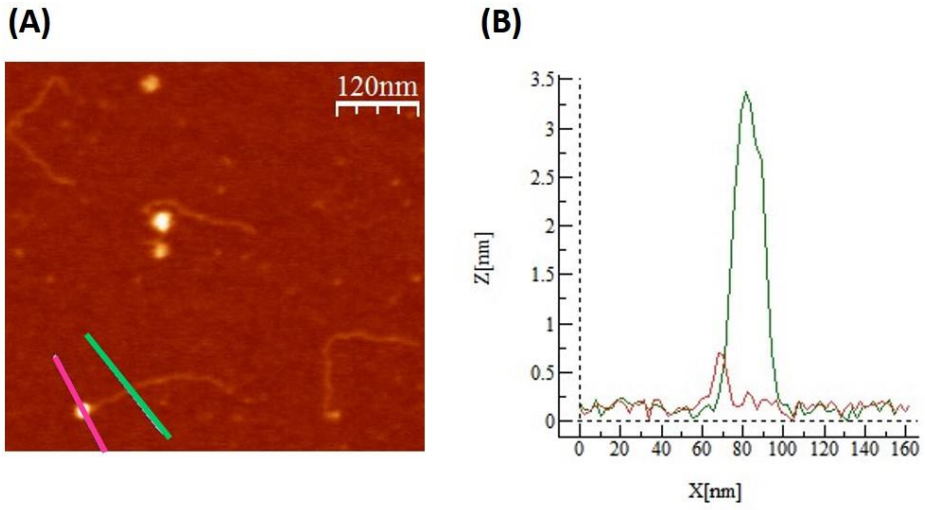


Figure 4.5: DNA-ΔsA complex, A) DNA 807 bp interacted ΔsA complex and B) the height profile of the DNA (0.7 nm) and the ΔsA complex (3.5 nm).

4.1.3 DNA-NβH interaction

MCM proteins are predicted to bind both ssDNA and dsDNA within the central channel. A number of loops have been implicated in this interaction; among those are the N-terminal β-hairpin loops that include positively charged residues. Biochemical data show that mutation of these loops abolish DNA binding. We imaged the interaction between DNA and an MCM protein where the key NβH residues have been mutated (NβH complex) using the buffer A condition for imaging in air as described in chapter three, at room temperature.

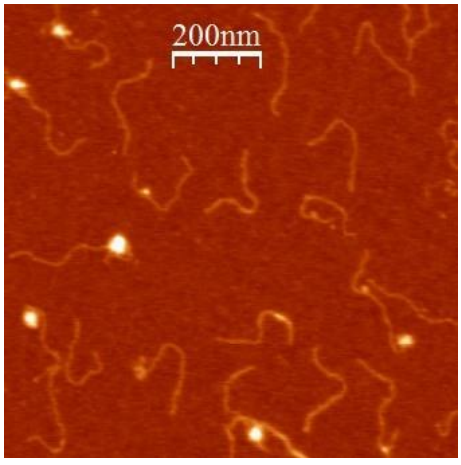


Figure 4.6: The NβH proteins in complex with 807 bp DNA molecule on freshly cleaved mica. In the image, the protein complexes associated to the DNA.

We observed, as expected, that the all the NβH proteins interacting with DNA are only in associated state, as shown in Fig. 4.6, with the DNA wrapping around the protein complex.

We observed a much lower degree of binding to the DNA, with 82% of the nucleic acid molecules free (compare to 16% for the wild-type MCM complex), and indeed we can see that the proteins are mostly associated to the DNA, as shown in Fig. 4.6.

- *Statistical analysis:*

We analysed 1634 DNA molecules, 18% of the DNA interacted with the NβH complex, due to the instability of the interaction and because the associated configuration is less stable than the loaded configuration as we observed for the case of wild type MCM complex interacted with the DNA.

The percentage of associated configuration from the total NβH complex bound to the DNA molecules is almost 100% both in the case of DNA:NβH (1:5) and 1:10, as reported in table 4.3.

Table 4.3: The percentage of loaded and associated configuration of DNA-NβH complex.

DNA : NβH DNA concentration = 2nM	Loaded %	Associated %
1:5	0.5 ± 0.5	99.5 ± 0.5
1:10	0.0 ± 0.0	100 ± 0.0

- *Height measurements:*

We obtained height profiles for the DNA molecules of 0.7 nm and for the NβH complex of 5.0 nm (Fig. 4.7), similar to the case of ΔsA experiments, still smaller than the values reported in the literature for the double stranded DNA (Moreno-Herreoet al., 2003) and NβH complex (Costa & Onesti, 2009), due the tip-sample force in air.

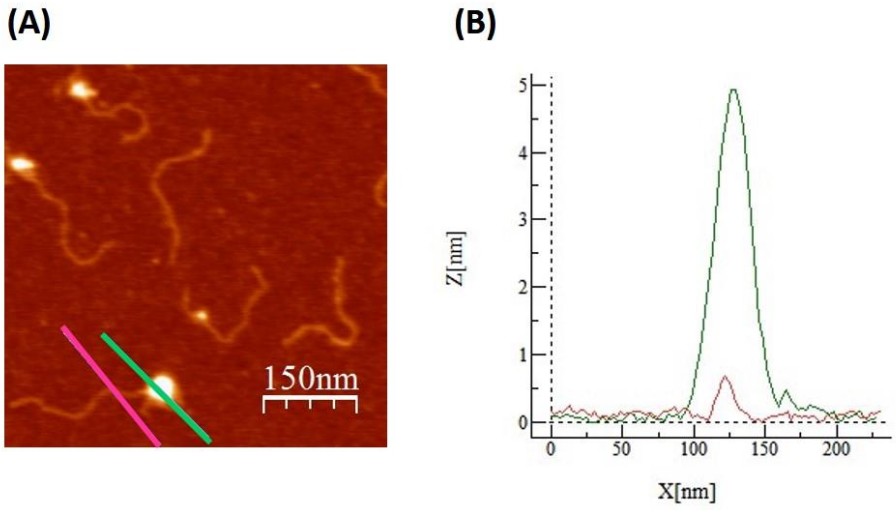


Figure 4.7: DNA-NβH complex, A) DNA 807 bp interacted NβH complex and B) the height profile of the DNA (0.7 nm) and the NβH complex (5.0 nm).

- *Unexpected interaction:*

We imaged the interaction between DNA and NβH complex in ratio (1:5) using the buffer solution A, imaging in air. As we can see from (Fig. 4.8B), in few cases two MCM complexes are associated to the same DNA strand. Moreover, unexpectedly (Fig. 4.9), we observed one MCM complex is associated with more than one DNA molecules.

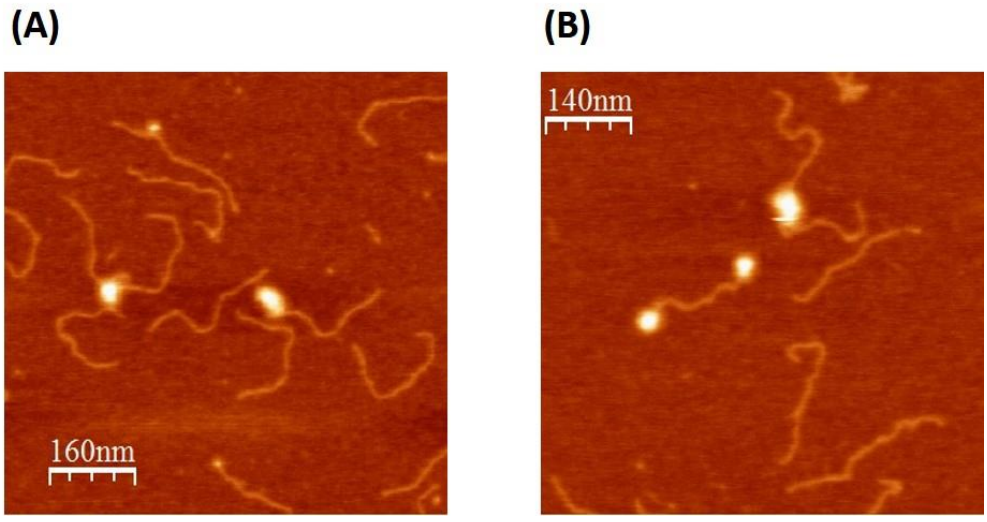


Figure 4.8: The NβH proteins in complex with 807 bp DNA molecule. In the image, the NβH complexes associated to the DNA. A) The protein complex associated to one DNA molecule as well two DNA molecules bound to the same protein complex. B) Two protein complex associated to the same DNA molecule and again two DNA molecules bound to the same protein complex.

These results indicate that probably the MCM complexes shown here are not single hexamer. More likely, the MCM complex here is a double hexamer as described in Chapter one.

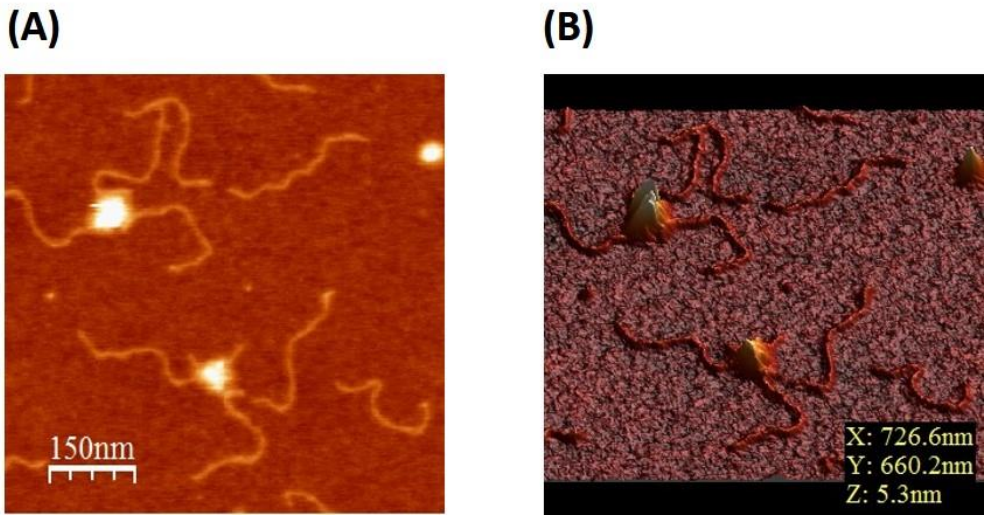


Figure 4.9: The NβH proteins in complex with 807 bp DNA molecule. In the image, A) two or three DNA molecules bound to the same protein complex. B) The 3D of the same image in A.

In fact in the associated configuration, when the DNA is wrapped around the external part of the protein structure, the DNA is bound to the N-terminal "subdomain A" of a single hexameric MCM (Costa et al., 2008). Therefore, if more than one DNA can bind to same MCM complex, a superstructure of the MCM single hexamer complex should be involved (see chapter one).

- *DNA-NβH interaction in the presence of EDTA*

From the crystal structure of MCM it is known that a zinc motif mediates the head-to-head double hexamer formation (Fletcher et al., 2003 & 2005). Thus, to test the double hexamer hypothesis, we added 1 mM EDTA (Ethylenediaminetetraacetic acid), known for the high affinity for zinc, to the buffer solution A (9 mM MgCl2, 90 mM NaCl, 30 mM HEPES and 5% of Glycerol), in the presence of the NβH mutant, to prevent double hexamer formation. From all the AFM images, as the one of Fig. 4.10, we observe that in these conditions DNA is interacting at most with one single mutant complex. This proves indirectly the contribution of double-hexameric structures of the protein complex to the interaction with DNA sequences.

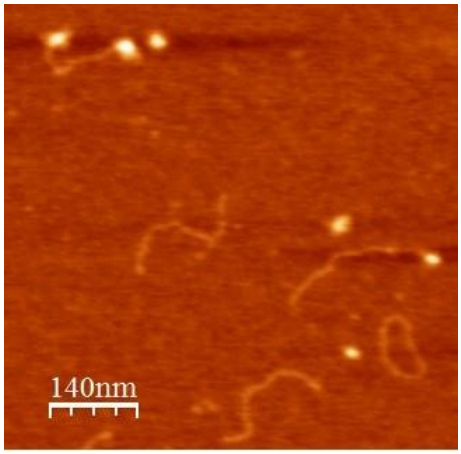


Figure 4.10: The NβH proteins in complex with 807 bp DNA molecule on freshly cleaved mica and we added 1mM of EDTA to the buffer solution.

We observed that 10% of the DNA molecules had a single MCM complex, 5% had two MCM complexes bound and 85% were free of proteins, similar to the result obtained for MCM-DNA interaction in the absent of EDTA. This indicates that the presence of EDTA doesn't affect the DNA binding affinity.

On the other hand, we observed 70% of the MCM complexes bound to the surface therefore the EDTA increased the affinity of the MCM complex bound to surface compared to the result obtained in the previous paragraph.

- *Height measurements:*

We measured height profiles from AFM topographic images also in the presence of EDTA. While the height of DNA molecule stays the same as in the previous paragraph, 0.5 nm, the one of NβH was 2.5 nm, half of the value measured in the absence of EDTA (Fig. 4.11). This measurement is supporting the idea of the double-hexamer formation explained in the previous paragraph, in the hypothesis that the axis of the double hexamer is oriented perpendicularly to the surface.

From the height analysis, all the MCM complexes measured in presence of EDTA resulted to be single hexamers, with an height value between 2.5 and 3 nm. This observation demonstrated the effect of EDTA on the MCM complex, this supporting the fact that the zinc motif contributes to the formation of a double hexameric structure.

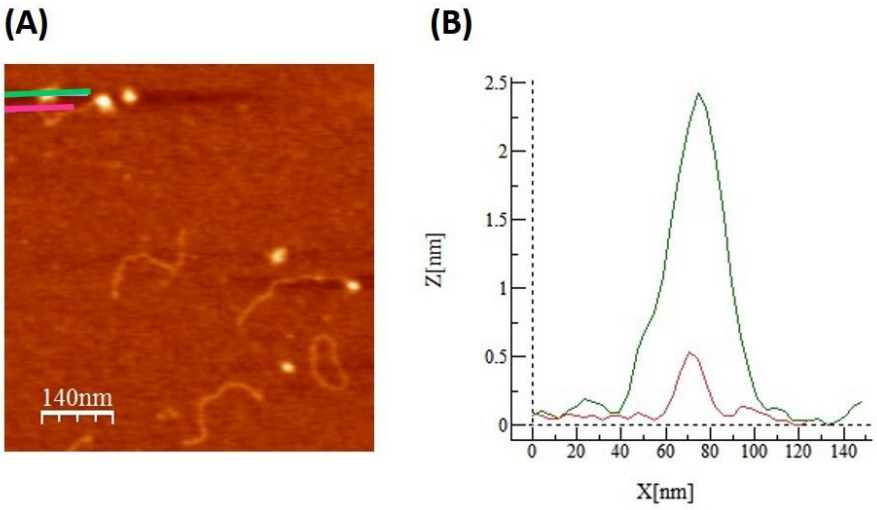


Figure 4.11: DNA-NβH complex, A) DNA 807 bp interacted NβH complex and B) the height profile of the DNA (0.7 nm) and the NβH complex (4.7 nm).

Note:

The variation in height measured for DNA filaments and protein complexes throughout this chapter, was mostly due to the variations in the forces applied on the sample by the tip during imaging. Applying high forces leads to low height profile of the biological sample.

The experimental results shown in this chapter were obtained using two different AFM instruments the Solver Pro and the Cypher using different cantilevers (stiff and soft). In the beginning we used stiff cantilever since the goal was to optimize the buffer condition, then the second step was to use soft cantilever to minimize the interaction forces.

Indeed, applying a large force to the sample, often increases imaging resolution, but also causes some problem, such as lower tip life since the tip gets contaminated/ broken quickly. Therefore, it is important to make a compromise between low force to avoid sample damaging/compression and high force to get good resolution.

4.1.4 DNA contour length analysis

In order to get more details on DNA-protein interaction, we performed a careful analysis of DNA contour length, before and after the interaction with the protein. Here is reported the analysis of data relative to 807 bp DNA fragments, whose estimated contour length L is 274.4 nm. We plot our data as contour length frequencies distribution, as in the histograms shown in Fig. 4.14.

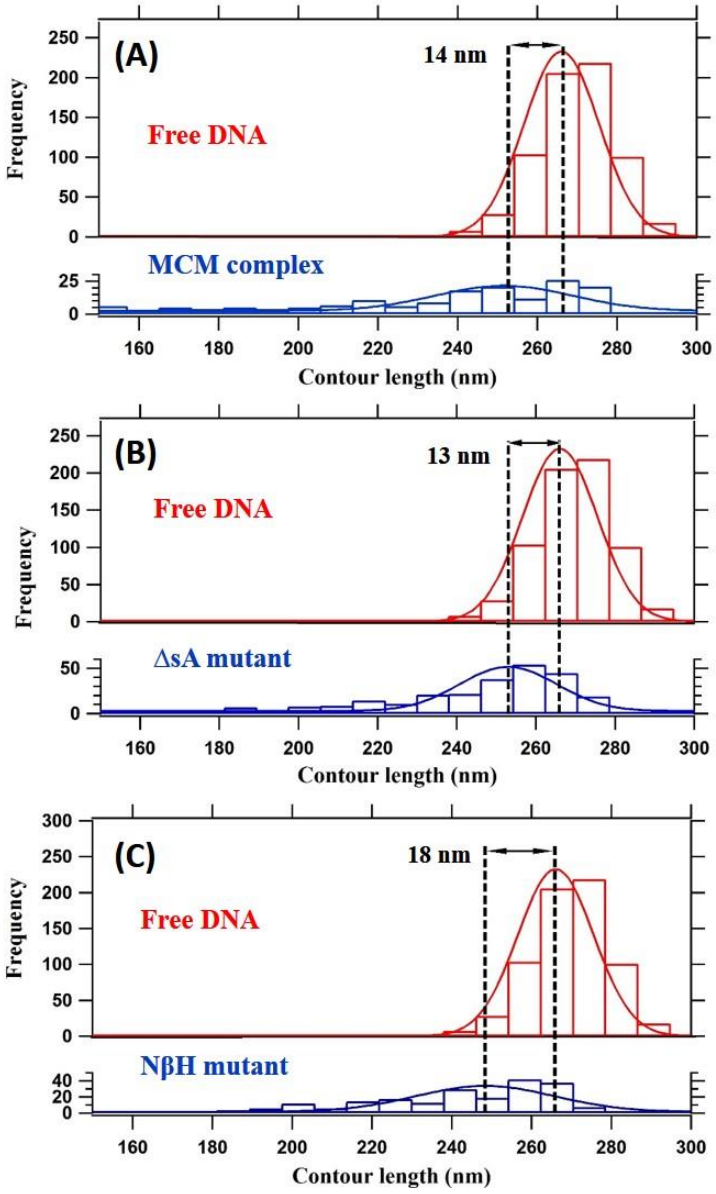


Figure 4.12: Histogram showing the 807 bp DNA contour length distributions of DNA (red distribution) and protein-DNA 807 bp complex (blue distribution). A) DNA-MCM complex. B) DNA-ΔsA complex. And C) DNA-NβH complex. Histograms of all the DNA contour length distributions were constructed with bin size of 8nm.

As a result, we found the following mean values ± the standard error (SE), of the measured L: for the free DNA L = 266±0.5 nm; for DNA-MCM interaction complex L = 252±1.6 nm; for DNA-ΔsA complex L = 253±1.1 nm and for DNA-NβH complex L = 248±1.6 nm. We considered in this analysis only DNA molecules with one protein bound (more than 95% of the cases for all conditions).

In the case of free DNA, we measured an average L which is 8 nm shorter than the estimated length. This could be due to the digitization process, in which the exact contour of the original DNA molecule is lost. The accuracy of DNA molecule contour length measurement, requires in fact high resolution images and a proper method to compute the contour length "2D single molecule software". In the case of DNA-MCM complex and DNA-ΔsA complex, we observed the same DNA compaction of about 13 nm, which refer to DNA compaction in loading configuration because the double hexamer MCM complex has twisted central channel (Li et al., 2015) that cannot be recognized from AFM image, which can make a reduction of 2 nm. On the other hand the salt concentration may affect the DNA contour length especially for the case protein complex bound to the DNA strand. All this can affect the contour length as we measured.

Finally, for the DNA-NβH complex a DNA compaction of about 18 nm is measured, corresponding to the wrapping of DNA around the protein complex in associated interaction mode.

4.1.5 Bending angle analysis

A statistical analysis of the protein-induced DNA kinks based on the bending angle (β) shown in Fig. 4.13. The analysis has been performed for free DNA molecules, DNA-MCM complex and DNA-NβH complex

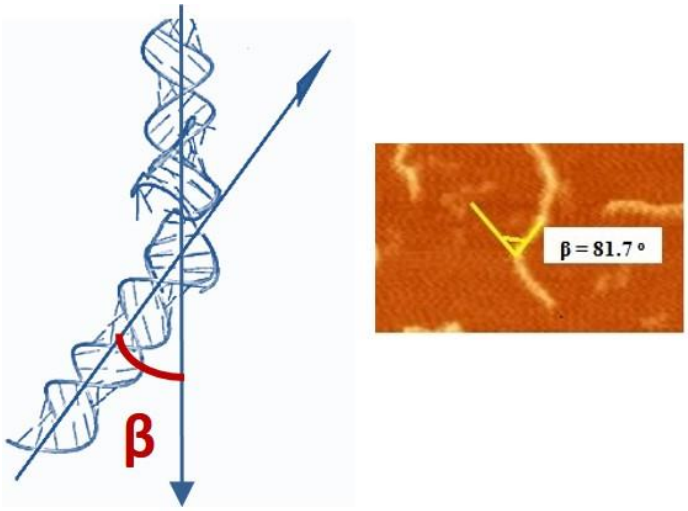


Figure 4.13: The bending angle (β) shown on the DNA molecule on the left. And on the right, the example of measuring the bending angle from AFM image.

For the DNA-MCM complex, the distribution is centred around 100°, shifted by 16° comparing to the free DNA distribution (Fig. 4.14A). This result is identical to the result obtained using cryo-Electron microscopy (Cryo-EM) by Alessandro Costa. He studied the wild-type *Mth*MCM (*Methanothermobacter thermautotrophicus* MCM) protein in complex with a 5.6 kbp double stranded DNA fragment and he found out that the average bending angle is 101°, with standard deviation 39° (Costa et al., 2008).

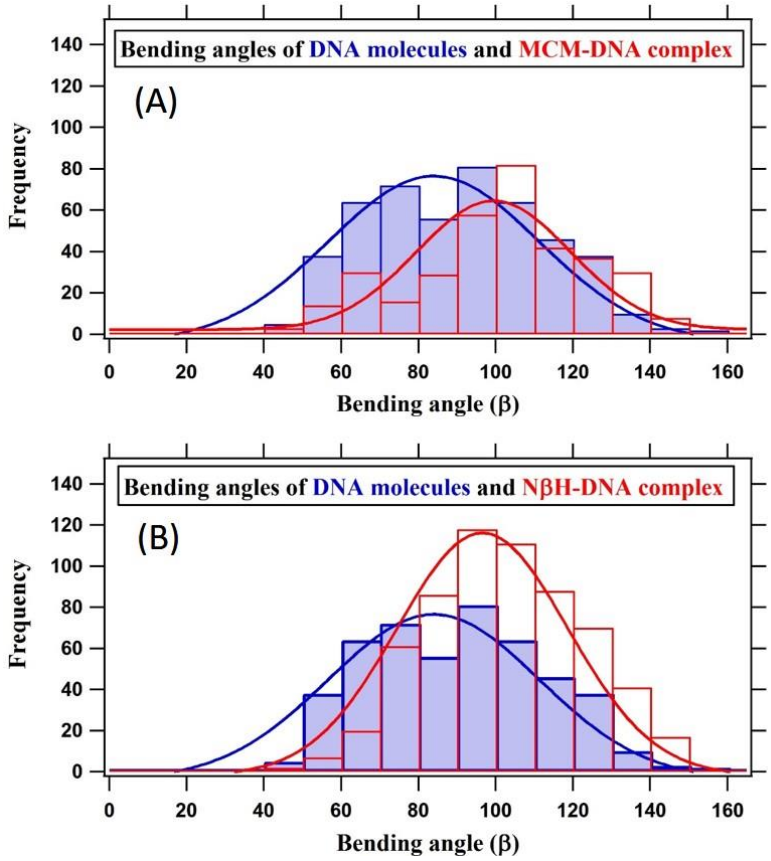


Figure 4.14: Histograms of the bending angles characterising the free DNA (blue distribution), the average angle is 84°, with a standard deviation of 39° and protein-mediated DNA kinks (red distribution), A) DNA-MCM complex, the average angle is 100°, with a standard deviation of 28°, and B) NBH-DNA complex, the average angle is 97°, with a standard deviation of 31°

On the other hand, for DNA-NβH complex, the distribution centred around 97° (Fig. 4.14B), which is close to the DNA-MCM distribution. Both MCM complex or NβH mutant are increasing the bending angle of the DNA fragment.

These results support the idea of a DNA conformational change upon interaction with protein complexes in the associated configuration.

4.2 Imaging in liquid

In this study we used the MCM complex with cleaved histidine tag, therefore here we could have chosen buffer conditions containing Ni^{+2} as described in chapter two. We carried out imaging in liquid using the Nanotec AFM which provided us with high resolution imaging with a scan rate of few seconds. In the image (Fig. 4.15A), few of the shown DNA molecules have one MCM complex bound to the end of the strand and the rest is free DNA.

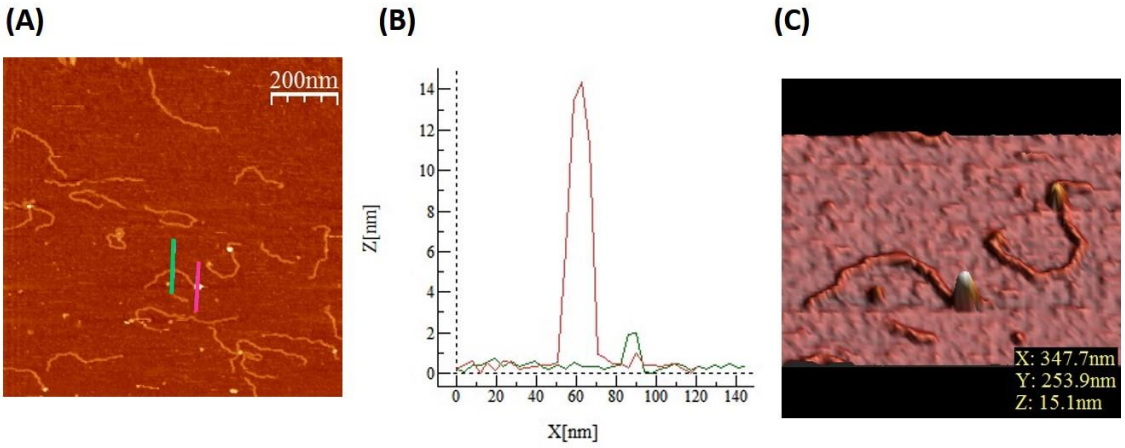


Figure 4.15: DNA-MCM complex (1:3) on freshly cleaved mica in liquid, using tapping mode with scan rate 20 seconds, A) DNA 807 bp interacted MCM complex and B) the height measurements of the DNA and MCM complex. The peaks values, 2.0 nm for the DNA and 14 nm for the MCM complex. C) The 3D of the same image in A.

From the AFM height measurements, we obtained DNA height profiles of 2.0 nm and MCM complexes profiles of 14 nm (Fig. 4.15B) (Gomez-Llorente et al., 2005), in very good agreement with the expected heights. Therefore imaging in liquid, allows to obtain the real heights of biomolecules. As a result, we were able to visualize the DNA-MCM interaction, through high resolution and fast imaging in physiological environment.

In the next chapter we will show few outcomes of this optimization.

5 Interaction dynamic

The main objective of this chapter is to highlight the mechanism by which the MCM complex unwinding the double stranded DNA. Prior to this step, we had to optimize the buffer condition suitable for studying with AFM DNA and its interaction with the protein complex, both in air and in liquid conditions. We then studied the DNA-protein complex interaction in static conditions, understanding DNA conformation in these interactions. We then distinguished from AFM images two types of interactions: the loaded and associated configurations that have been then confirmed from the studied we carried out on two different MCM mutants (ΔsA and NβH complex) with mutation on the external and internal part of the MCM complex, respectively. We also visualized the interaction at high resolution in liquid using fast AFM. All those results were discussed in the chapters three and four.

Now we will describe the DNA-MCM interaction dynamics, by adding ATP to the buffer solution. All the interaction experiments, included the ones presented in this chapter, were performed using blunt-ended DNA fragments of 807 bp, as a starting point. We used DNA 807 bp because it is the suitable length for the protein-DNA interaction as I described in the previous chapter.

5.1 Imaging in air

We started with imaging in air, diluting the DNA and MCM complex in the buffer A, leaving them interacting for 30 minutes at 50°C (archaeal MCM is stable and well active at this temperature) and then we added the ATP and incubated the solution on freshly cleaved mica. To stop the interaction, we rinsed the sample using milliQ water. We prepared for each experiment, 100 µL solution of the DNA-MCM and then we added ATP to just 50% of the solution, to better grasp the effect, if any, of DNA unwinding by the protein complex.

- *Imaging of DNA-MCM complex:*

As first step, we imaged the DNA-MCM complex in the absence of ATP in air. We incubated 20 µL of the DNA-MCM solution on freshly cleaved mica for one minute then we rinsed and dried as described in chapter two. As a result, we observed few of the MCM protein complexes bound to the DNA molecules and to the surface, as shown in Fig. 5.1.

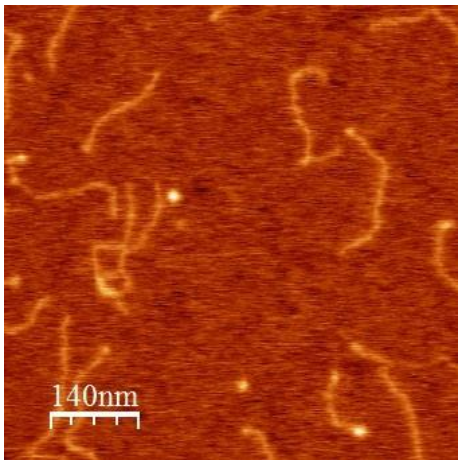


Figure 5.1: The MCM proteins in complex with 807 bp DNA molecule (1:3) on freshly cleaved mica in air with Cypher AFM operating in tapping mode.

Next, to study the interaction dynamics, we added ATP. Using the other 50% of the same solution prepared for the previous experiment, we added ATP in concentration of 2µM. After one minute of interaction we incubated the solution on freshly cleaved mica for another one minute and then stopped the interaction by rinsing with milliQ water. The experiments were performed using the Cypher AFM operated in tapping mode. Typical images obtained after DNA-MCM interaction in the presence of ATP, are the ones reported in Fig. 5.2.

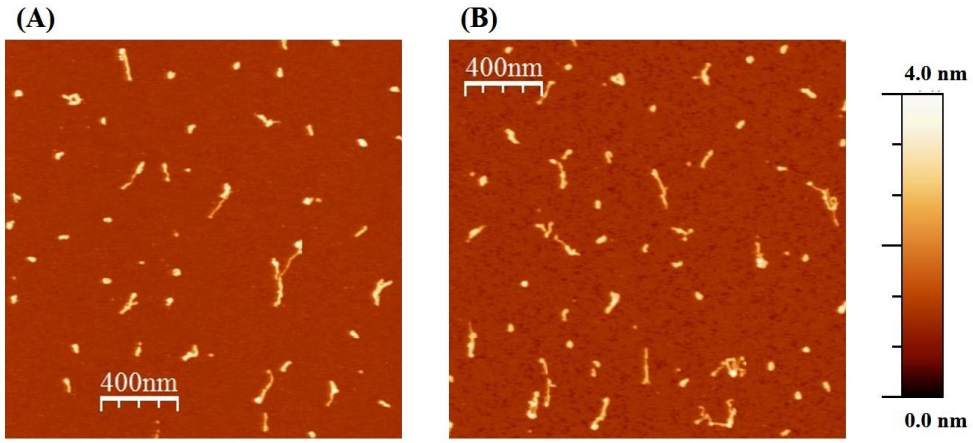


Figure 5.2: The interaction between 807bp DNA and MCM complex (1:3) in the presence of 2µM of ATP, imaging in air with Cypher AFM operated in tapping mode, 40 seconds per fame.

The profile of the DNA molecules in Fig. 5.3 is different from previous experiments. Here, we see for the first time, with high resolution, that part of the strand has been unwound by the action of MCM.

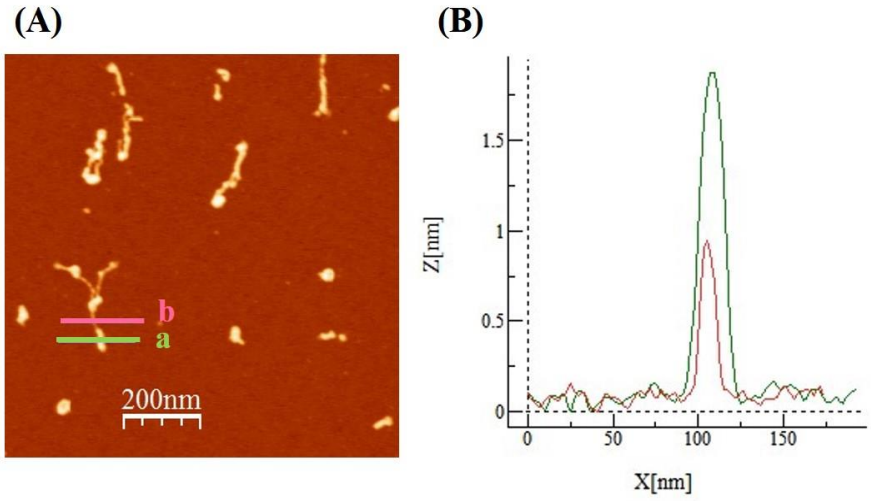


Figure 5.3: DNA-MCM complex (1:3) interacted with 2 µM of ATP, imaging in air, A) DNA 807 bp interacted MCM complex and ATP, and B) The height measurements of two different positions on the DNA. The peaks values, 1.8 nm for the green peak and 0.9 nm for the pink peak.

Height profiles of selected molecules are reported in Fig. 5.3: at position "a" the height value is 1.8 nm while at position "b" the value is 0.9 nm. The results is consistent with the co-presence of double stranded DNA "b" and supercoiled single stranded DNA "a" on the same fragment.

The unwound piece of ssDNA is not stable since the buffer used is optimized for dsDNA only. Therefore, it does not equilibrate on the surface, rather is expected to be coiled, due to the high flexibility of ssDNA, as the feature indicated by the green arrow in Fig. 5.4.

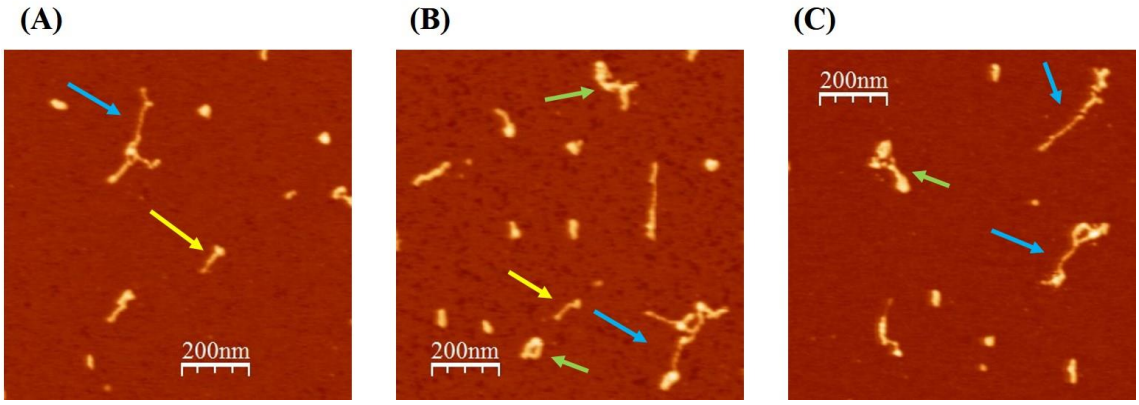


Figure 5.4: The interaction between 807 bp DNA and MCM complex (1:3) in the presence of 2 µM of ATP, imaging in air with Cypher AFM operated in tapping mode, 40 seconds per fame.

The molecules observed in the Figure. 5.4, have different structures that could be referred to many different DNA configurations. We observe small single stranded supercoiled DNA molecules (green arrow), mixed ss-and ds-DNA fragments (light blue arrow), small pieces (around 150 nm in length) of dsDNA (yellow arrow) and then the MCM complex bound to the DNA and/or to the surface.

From images and height profiles, we can assert that, DNA unwinding by MCM protein complex can occur also on blunt-ends DNA, although, in the literature there are no biochemical evidences in literature of efficient unwinding of MCM complex on blunt-ended DNA. Further confirmation of this might come from the study of this interaction under physiological environment.

5.2 Imaging in liquid

To further demonstrate that DNA unwinding by MCM occurs on our DNA substrates, as shown in Fig. 5.3, we studied the unwinding process powered by ATP hydrolysis with AFM imaging in liquid. Using the protocol optimized for imaging in liquid with Nanotec AFM, we first prepared sample for DNA-MCM interaction as described in chapter two. We started imaging the DNA-MCM complex interaction in liquid, without ATP, as shown in Fig. 5.5A.

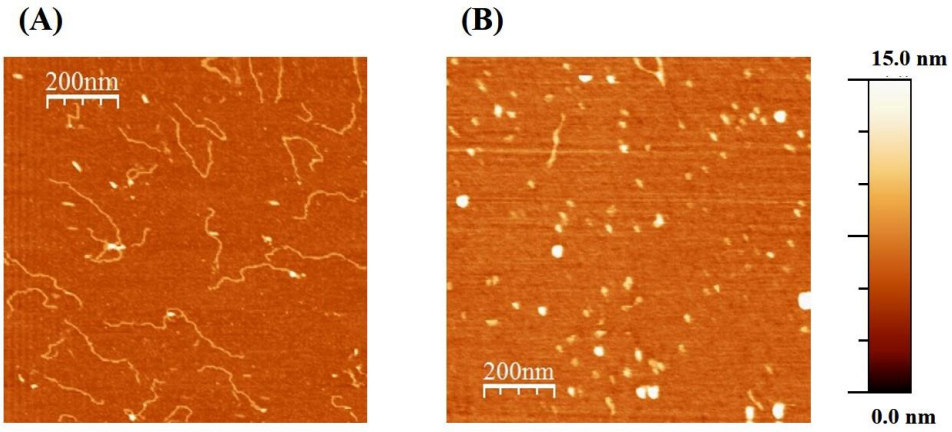


Figure 5.5: The MCM proteins in complex with 807 bp DNA molecule (1:3) on freshly cleaved mica, with Nanotec AFM. A) The interaction in the absence of the ATP, the image operated in jumping mode, 40 seconds per frame. B) The interaction in the presence of the ATP, the image operated in tapping mode, 20 second per frame.

Next, in the same experiment we lifted up the tip from the surface, to inject an aliquot of ATP and we approached the surface again by the tip and immediately started imaging. A representative image is shown in Fig. 5.5B in which we observed small pieces of DNA and MCM complex bound to the surface. This very different outcome took place in few seconds, and is consistent with the unwinding process of the double stranded DNA in physiological environment.

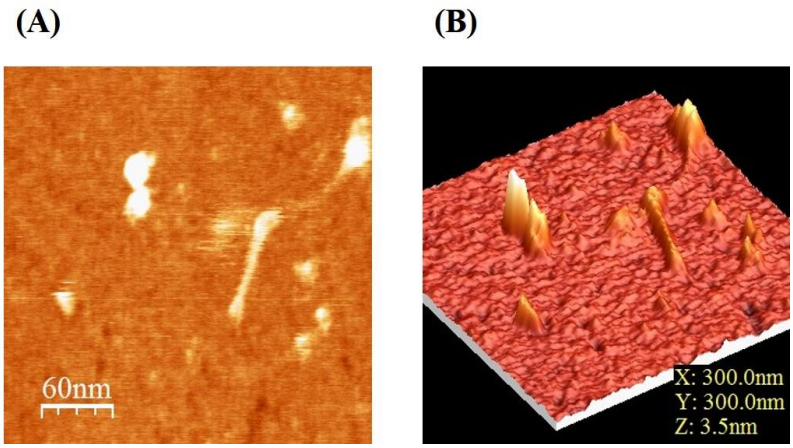


Figure 5.6: The MCM proteins in complex with 807 bp DNA molecule in the presence of 2 µM ATP. In the image, A) a piece of DNA and protein complexes bound to the same surface. B) The 3D of the same image in A.

On the other hand, from the 3D representation, it is very clear that even the small pieces of DNA (around 150 nm) can have a portion of ssDNA as shown in Fig. 5.6B. Based on these observations we suggested that the MCM protein may have the ability to unwind dsDNA fragment.

Note:

In the experiments performed in liquid we didn't observe the coiled ssDNA or the supercoiled ssDNA seen from the images in air as the one in Fig. 5.2. This is due to the presence of the liquid. Since the supercoiled DNA pieces are just weakly bound to the surface, since the buffer condition for imaging in liquid was optimized for the double stranded DNA only, they start floating in solution and therefore cannot be imaged by the AFM tip.

In conclusion, the choice of appropriate conditions allowed us to study the DNA-protein interaction in air and in liquid, at high resolution. Moreover, those condition allowed the

study of the interaction dynamics “in the presence of the ATP”, with high resolution and fast imaging AFM (few seconds per image).

These preliminary observations seem to confirm the effect of unwinding on blunt-ended DNA, already observed in air, adding another challenge to the MCM complex paradox. Further experiments are required to confirm this prediction, as for instance the use of ATP analogue (AMP-PNP) that cannot be hydrolysed, but can trap the protein complex in a structure closely related to the ATP-bound state. Second, we plan to test other types of DNA molecules (fork, bubble, etc.) using the same imaging conditions in air and in liquid, to comment on the efficiency of unwinding of different substrates by MCM.

Conclusions and future plan

Understanding the DNA-MCM protein complex is crucial to unravel the mechanism of DNA replication, which is at the core of cell proliferation. Although the MCM proteins have been characterized over the years both structurally and biochemically, unanswered questions remain. For example, how the helicase binds and unwinds the double stranded DNA is still unclear.

Based on structural and biochemical evidences, there are two models of MCM-DNA interaction that have been described: the canonical “loaded” mode where the MCM protein complex is encircling the DNA for unwinding, and an “associated” mode where the DNA is supposed to be wrapped around the external part of the proper ring structure.

The main goal of this thesis work, is to confirm and support the model structure of MCM-dsDNA interaction and its dynamics in near physiological condition, by means Atomic Force Microscopy (AFM) single molecule imaging. We studied archaeal MCM from *Methanothermobacter thermautotrophicus* as simplified model system, interacting with blunt-end, double-stranded DNA fragments of different lengths.

- *Protein-DNA interaction:*

To study the topological determinants of the interaction of the MCM complex with dsDNA, we employed AFM to image MCM complex bound to the DNA in the absence and in the presence of ATP. By AFM mean, the work was carried out in two different conditions: in air to understand the static conformations of DNA-protein complexes; in liquid to follow the dynamics of the interaction.

First, we optimized the protocol for AFM imaging in air to obtain clear and high resolution images of surface-equilibrated DNA molecules before and after the interaction with protein complex. Direct visualization of the MCM complexes bound to the DNA, allowed us to distinguish between loaded and associated configurations by means of an accurate analysis of AFM topographic images. The presence of bending angle induced by the MCM protein has been attributed to the “associated” configuration, whereas when the interaction of the MCM complex with DNA did not produce any bending, it was assigned to the loaded configuration. This assignment has been tested through the study of the separate effects of mutations in the inner hole (NβH mutant) or in the outer ring (ΔsA mutant) of the MCM on the DNA conformation. Indeed, previous works established that the "association mode"

requires subdomain A and the "loading" requires an integral N-terminal β-hairpin (NβH). As a result, the NβH mutant interacting with DNA, is expected only to associate to DNA, while the ΔsA mutant is expected to interact in loaded configuration with DNA. Our AFM analysis strikingly confirmed the preference of associated configuration of the NβH and the presence of only loaded complexes with the ΔsA mutant. These results confirmed the essential role of subdomain A in the association step (Costa et al., 2008). Moreover, the total number of DNA bound complexes decreased from 84% to 20% from MCM to ΔsA, suggesting that association is involved in favouring the replicative helicase loading, and initiating the double-helix unwinding (Ali et al., 2016), as it has been discussed in chapter four.

Along with the topographical analysis, I performed a statistical analysis of DNA contour length and binding angle to highlight the effects of MCM complex and its mutants upon interaction with DNA on the physical properties of the system. I studied DNA contour length of a DNA fragment of 807 bp, with and without the protein complex, finding out that the presence of a single protein complex bound to the DNA reduced the DNA contour length of about 13 nm for wild type MCM and ΔsA mutant, which could be due to the fact that the double hexamer MCM complex has twisted central channel (Li et al., 2015), which can reduce the length by 2 nm. This twisted feature cannot be recognized from AFM image. Therefore the measured length was shorter than the exact length. In the case of NβH mutant, the DNA contour length is on average 18 nm shorter than bare DNA: this effect can be explained by the DNA wrapping around the protein complex (associated interaction mode).

Subsequently, we studied the bending angle of MCM complex and NβH complex interacting with 807 bp DNA sequences. We observed that the bending angle increased of 16° compared to bare DNA, strongly supporting the “association” model that has been observed using cryo-EM by A. Costa (Costa et al., 2008).

Those results showing the conformational changes of DNA molecules upon interaction with protein complex due to the associated configuration, confirm the results of previous electron-microscopy studies and biochemical prediction.

- *Interaction dynamic:*

Since we observed that the MCM complexes bound to the blunt-ended DNA fragment, we decided to study the dynamic of the interaction using the same DNA fragment. We fueled the MCM helicase with ATP in solution contain 807 bp DNA molecules and then incubated

the solution on mica surface using the imaging condition of the experiments described so far. We stopped the interaction by rinsing and drying the sample, and then we performed imaging in air with Cypher AFM as described in chapter five. As a result, after the incubation of DNA with MCM complex and ATP, from height profile AFM image along the DNA we observed the presence of supercoiled single stranded DNA (ssDNA) and mixed ss-dsDNA deposited on the surface. This can be the sign of unwinding of the double stranded DNA by MCM complex. This effect was unexpected, since most of the biochemical evidence in the literature suggests that MCM requires a fork or emifork-like substrate. However the first characterization of the *Mth*MCM does report a weak activity towards a blunt-ended substrate (Shechter et al., 2000).

To follow the dynamic we studied the protein-DNA interaction, in the presence of ATP, via moderately fast (few seconds/images) AFM imaging in liquid. When we added ATP, we observed a change of topographic height of the DNA strands, consistent with the formation of ssDNA, indicating the unwinding of the double stranded DNA. These preliminary observations seem to confirm the effect of unwinding on blunt-ended DNA, already observed in air, adding another piece in the puzzle of DNA-MCM complex interaction.

Based on these observations we suggested that the MCM protein may have the ability to unwind dsDNA fragment and further experiments are required to confirm this hypothesis. For example we are planning to test other types of DNA molecules (fork, bubble, etc.) as well as MCM mutants that are unable to unwind DNA, using the same imaging conditions in air and in liquid.

Future AFM observations should then provide more insights into the dynamical properties of MCM protein complex. Moreover, we would like to study MCM-DNA interaction by adding other protein factors involved replication process (such as GINS and Cdc45) that may help to understand better the MCM mechanism.

References

Ali F. A., Renault, L. Gannon, J. Gahlon, H. L. Kotecha, A. Zhou, J. C. Rueda, D. & Costa, A. (2016) Cryo-EM structures of the eukaryotic replicative helicase bound to a translocation substrate. NATURE COMMUNICATIONS. 7:10708.

Ando, T., Kodera, N., Takai, E., Maruyama, D., Saito, K., and Toda, A. (2001) A high speed atomic force microscope for studying biological macromolecules. Proc. Natl. Acad. Sci. U.S.A 98, 12468-12472.

Aravind L., Koonin E. V. (1999) DNA-binding proteins and evolution of transcription regulation in the archaea. Nucleic Acids Res. 27(23):4658-70.

Ares, P. Fuentes-Perez, M. E. Herrero-Galán, E. Valpuesta, J. M. Gil, A. Gomez-Herrero, J. and Moreno-Herrero, F. (2016) High resolution atomic force microscopy of double-stranded RNA. Nanoscale Advance Article.

Bae B, Chen YH, Costa A, Onesti S, Brunzelle JS, Lin Y, Cann IK and Nair SK. (2009). Insights into the architecture of the replicative helicase from the structure of an archaeal MCM homolog. Structure 17:211–22.

Barry ER, McGeoch AT, Z and Bell SD. (2007). Archaeal MCM has separable processivity, substrate choice and helicase domains. Nucleic Acids Res 35:988–998.

Binnig, G.; Quate, C. F.; Gerber, Ch. (1986). "Atomic-Force Microscope". Physical Review Letters 56 (9): 930–933.

Blow, J.J. and Laskey, R.A. (1988) A role for the nuclear envelope in controlling DNA replication within the cell cycle. Nature 332, 546–548

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

Bochman,M.L. & Schwacha,A. (2007) Differences in the single-stranded DNA binding activities of MCM2–7 and MCM467: MCM2 and MCM5 define a slow ATP-dependent step. Biol. Chem., 282, 33795–33804.

Bockelmann, U. Essevaz-Roulet, B. Heslot, F. (1998) DNA strand separation studied by single molecule force measurements. Phys. Rev. E, 58, pp. 2386-2394.

Bowers JL, Randell JCW, Chen S et al (2004) ATP hydrolysis by ORC catalyzes reiterative Mcm2-7 assembly at a defined origin of replication. Mol Cell 16:967–978.

Brewster, A. S. & Chen, X. S. (2010) Insights into MCM functional mechanism: lessons learned from the archaeal MCM complex. Crit Rev Biochem Mol Biol., 45(3), 243–256.

Brewster, A. S. Wang G, Yu X, Greenleaf WB, Carazo JM, Tjajadi M, Klein MG, Chen XS. (2008) Crystal structure of a near-full-length archaeal MCM: functional insights for an AAA+ hexameric helicase. Proc Natl Acad Sci U S A. 23; 105(51):20191-6.

Bustamante, C. & Rivetti, C. (1996) Visualizing protein-nucleic acid interactions on a large scale with the scanning force microscopy. Annu. Rev. Biophys. Biomol. Struct. 25, 395–429.

Bustamante, C. Erie, D. A. & Keller, D. (1994) Biochemical and structural applications of scanning force microscopy. Curr. Opin. Struct. Biol. 4.750-760.

Bustamante, C., Marko, J. F., Siggia. E. D. And Smith, S. B. (1994) Entropic elasticity of l-phage DNA. Science. 265:1599.

Bustamante, C., Rivetti, C. & Keller, D. J. (1997) Scanning force microscopy under aqueous solutionsCurr. Opin. Struct. Biol. 7, 709–716.

Cappella, B; Dietler, G (1999). Force-distance curves by atomic-force microscopy. Surface Science Reports 34 (1–3): 1–104.

Chi, Q. Wang, G. Jiang, J. (2013) The persistence length and length per base of single-stranded DNA obtained from fluorescence correlation spectroscopy measurements using mean field theory. Physica A 392:1072–1079.

Chong, J.P., Hayashi,M.K., Simon,M.N., Xu,R.M. and Stillman,B. (2000) A double-hexamer archaeal minichromosome maintenance protein is an ATP-dependent DNA helicase. Proc. Natl. Acad. Sci. U.S.A., 97, 1530–1535.

Colton, R. J., Baselt, D. R., Dufreˆne, Y. F., Green, J. B. & Lee, G. U. (1997) Scanning probe microscopy. Curr. Opin. Chem. Biol. 1, 370–377.

Costa A and Onesti S. (2008) The MCM complex: (just) a replicative helicase? Biochem Soc Trans 36:136–140.

Costa A. and Onesti S. (2009) Structural biology of MCM helicases. Mol. Biol. 44, 326-342.

Costa, A. Hood, I. V. & Berger, J. M. (2013) Mechanisms for Initiating Cellular DNA Replication. Annu. Rev. Biochem. 82. 25-54.

Costa, A. Pape, T. van Heel, M. Brick, P. Patwardhan, A. & Onesti, S. (2006) Structural studies of the archaeal MCM complex in different functional states. J. Struct. Biol. 156, 210–219.

Costa, A. Pape, T. van Heel, M. Brick, P. Patwardhan, A. & Onesti S. (2006) Structural basis of the Methanothermobacter thermautotrophicus MCM helicase activity. Nucleic Acids Res. 34(20):5829–38

Costa, A. van Duinen, G. Medagli, B. Chong, J. Sakakibara, N. Kelman, Z, Nair, SK. Patwardhan, A. & Onesti, S. (2008) Cryo-electron microscopy reveals a novel DNA-binding site on the MCM helicase. EMBO J 27:2250–2258.

Costa,A., Ilves,I., Tamberg,N., Petojevic,T., Nogales,E., Botchan,M.R. and Berger,J.M. (2011) The structural basis for MCM2–7 helicase activation by GINS and Cdc45. Nat. Struct. Mol. Biol., 18, 471–477.

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

Czajkowsky, D. M. & Shao, Z. (1998) Submolecular resolution of single macromolecules with atomic force microscopy. FEBS Lett. 430, 51–54.

Deegan, T. D. and Diffley, J. F. (2016) MCM: one ring to rule them all. Structural Biology, 37:145–151.

Doniselli, N. Rodriguez-Aliaga, P. Amidani, D. Bardales, J. A. Bustamante, C. Guerra, D. G. & Rivetti1, C. (2015) New insights into the regulatory mechanisms of ppGpp and DksA on Escherichia coli RNA polymerase–promoter complex. Nucleic Acids Res., 43, 5249–5262.

Drake, B. Prater, C. B. Weisenhorn, A. L. Gould, S. A. Albrecht, T. R. Quate, C. F. Cannell, D. S. Hansma, H. G. And Hansma, P. K. (1989) Imaging crystals, polymers, and processes in water with atomic force microscope. Science 243:n969-975.

Edgell D.R., and Doolittle W.F. (1997) Archaea and the origin(s) of DNA replication proteins. Cell 89, 995–998

Erzberger, J. P. and Berger, J. M. (2006) Evolutionary relationships and structural mechanisms of AAA+ proteins. Annu Rev Biophys Biomol Struct. 35:93-114.

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

Fantner, G. E., Schitter, G., Kindt, J. H., Ivanov, T., Ivanova, K., Patel, R., et al. (2006) Components for high speed atomic force microscopy. Ultramicroscopy 106, 881-887.

Ferdos Abid Ali and Alessandro Costa (2016). The MCM Helicase Motor of the Eukaryotic Replisome. YJMB I-64978.

Fletcher, R. J. Bishop, B. E. Leon, R. P. Sclafani, R.A. Ogata, C. M. & Chen, X. S. (2003) The structure and function of MCM from archaeal M. thermoautotrophicum. Nat. Struct. Biol. 10, 160–167.

Forsburg SL. (2004) Eukaryotic MCM proteins: beyond replication initiation. Microbiol Mol Biol Rev 68:109–131, table of contents.

Froelich C. A. Kang S, Epling LB, Bell SP, Enemark EJ. (2014) A conserved MCM single-stranded DNA binding element is essential for replication initiation. Elife. 3:e01993.

Fu YV, Yardimci H, Long DT, Ho TV, Guainazzi A, Bermudez VP, Hurwitz J, van Oijen A, Schärer OD, Walter JC. (2011) Selective bypass of a lagging strand roadblock by the eukaryotic replicative DNA helicase. Cell. 146(6):931-41.

Geisse, Nicholas A. (2009) AFM and combined optical techniques.Materials Today 12 (7–8): 40–45.

Gomez-Llorente,Y., Fletcher,R.J., Chen,X.S., Carazo,J.M. & San Martin,C. (2005) Polymorphism and double hexamer structure in the archaeal minichromosome maintenance (MCM) helicase from Methanobacterium thermoautotrophicum. J. Biol. Chem., 280, 40909–40915.

Gross, L.; Mohn, F.; Moll, N.; Liljeroth, P.; Meyer, G. (2009) The Chemical Structure of a Molecule Resolved by Atomic-Force Microscopy. Science 325 (5944): 1110–1114.

Günterodt. (1994). Biological materials studied with dynamic force microscopy. J. Vac. Sci.

Guthold, M. Falvo, M. Matthews, W. G. Paulson, S. Mullin, J. Lord, S. Erie, D. Washburn, S. Superfine, R. Brooks, F. P. Jr. and Taylor, R. M. (1999) Investigation and modification of molecular structures with tha nanoManipulator. J. Mol Graph Model 17: 187-197.

Guthold, M., M. Bezanilla, D. A. Erie, B. Jenkins, H.G. Hansma, and C. Bustamante. 1994. Following the assembly of RNA polymerase-DNA complexes in aqueous solutions with the scanning force microscope. Proc. Natl. Acad. Sci USA 91: 12927-12931.

Hamon, L.Pastré, D. Dupaigne, P. Breton, C. L. Cam, E. L. & Piétrement, O.(2007) High-resolution AFM imaging of single-stranded DNA-binding (SSB) protein—DNA complexes. Nucleic Acids Res. 35, 8.

Hansma, H. G. & Oroudjev, E. (2010) Atomic Force Microscopy of Biomaterials, Mica, and the Origins of Life. Microscopy Today, 6, 16-20, pp 16-20.

Hansma, H. G. 1995. Atomic force microscopy of biomolecules. J. Vac. Sci. Technol. B. 14: 1390-1395.

Hansma, H. G. Golan, R. Hsieh, W. Daubendiek, S. L. & Kool, E. T. (1999) Polymerase Activities and RNA Structures in the Atomic Force Microscope. Struct. Biol. 127, 240–247.

Hansma, H. G., and D. E. Laney. 1996. DNA binding correlates with cationic radius: assay by atomic force microscopy. Biophys. J. 70: 1933-1939.

Hansma, H. G., Bezanilla, M., Nudler, E., Hansma, P. K., Hoh, J., Kashlev, M., Firouz, N. &Smith, B. L. (1998) Left-handed orientation of histidine-tagged RNA polymerase complexes imaged by atomic force microscopy .Probe Microscopy 1, 127–134.

Herrero-Galan, E. Fuentes-Perez, M. E. Carrasco, C Valpuesta, J. M. Carrascosa, J. L. Moreno-Herrero, F. and Arias-Gonzalez, J. R. (2013) Mechanical Identities of RNA and DNA Double Helices Unveiled at the Single-Molecule Level Am. Chem. Soc. 135, 122−131.

Hoang M.L., Leon R.P., PessoaBrandao L., Hunt S., Raghuraman M.K., Fangman W.L., Brewer B.J., Sclafani R.A. (2007) Structural changes in Mcm5 protein bypass Cdc7–Dbf4 function and reduce replication origin efficiency in S. cerevisiae. Mol. Cell. Biol. 27:7594 7602.

Horcas I, Fernández R, Gómez-Rodríguez JM, Colchero J, Gómez-Herrero J, Baro AM., Horcas, I. Fernández, R. Gómez-Rodríguez, J. M. Colchero, J. Gómez-Herrero, J. Baro, A.

M. (2007) WSXM: a software for scanning probe microscopy and a tool for nanotechnology. Rev Sci Instrum. 78(1):013705.

Ilves, I. Petojevic,T. Pesavento, J. J. & Botchan, M. R. (2010) Activation of the MCM2–7 helicase by association with Cdc45 and GINS proteins. Mol. Cell, 37, 247–258.

Ishimi Y. (1997). A DNA helicase activity is associated with an MCM4, -6, and -7protein complex. J Biol Chem 272:24508–24513.

Jenkinson, E. R. & Chong, J. P. (2006) Minichromosome maintenance helicase activity is controlled by N- and C-terminal motifs and requires the ATPase domain helix-2 insert. Proc. Natl. Acad. Sci. U.S.A., 103, 7613–7618.

Jenkinson, E. R. Costa, A. Leech A. P. Patwardhan A. Onesti S, Chong JP. (2009) Mutations in subdomain B of the minichromosome maintenance (MCM) helicase affect DNA binding and modulate conformational transitions. J Biol Chem. 284(9):5654-61

Jingchuan Sun1, Yi Shi, Roxana E Georgescu, Zuanning Yuan, Brian T Chait, Huilin Li & Michael E O'Donnell (2015)The architecture of a eukaryotic replisome. nature structural & molecular biology . 22; 12.

Kasiviswanathan R., Shin J.H., Kelman Z. (2005) Interactions between the archaeal Cdc6 and MCM proteins modulate their biochemical properties. Nucleic Acids Res. 33:4940–4950.

Kasiviswanathan, R. Shin, J. H. Melamud, E. & Kelman, Z. (2004) Biochemical characterization of the Methanothermobacter thermautotrophicus minichromosome maintenance (MCM) helicase N-terminal domains.J Biol Chem 279:28358–28366.

Kearsey, S. E. & Cotterill, S. (2003) Enigmatic variations: divergent modes of regulating eukaryotic DNA replication .Molecular Cell. 12, 1067–1075.

Kelman, Z., Lee, J. K. & Hurwitz, J. (1999). The single minichromosome maintenance protein of Methanobacterium thermoautotrophicum DeltaH contains DNA helicase activity. Proc. Natl Acad. Sci. USA, 96, 14783–14788.

Krastanova I., Sannino V., Amenitsch H., Gileadi O., Pisani F. M., Onesti S. (2012). Structural and functional insights into the DNA replication factor Cdc45 reveal an evolutionary relationship to the DHH family of phosphoesterases. J. Biol. Chem. 287, 4121–4128 10.1074/jbc.M111.285395

Lee, J. K. and Hurwitz, J. (2001). Processive DNA helicase activity of the minichromosome maintenance proteins 4, 6, and 7 complex requires forked DNA structures. Proc Natl Acad Sci USA 98:54–59.

Leipe, D.D. Koonin, E. V. Aravin, L. (2003) Evolution and classification of P-loop kinases and related proteins. J Mol Biol. 333: 781-815.

Li, N. Zhai, Y. Zhang, Y. Li, W. Yang, M. Lei, J. et al., (2015) Structure of the eukaryotic MCM complex at 3.8 Å. Nature.

Liu, W. Pucci, B. Rossi, M. Pisani, F. M. and Ladenstein, R. (2008) Structural analysis of the Sulfolobus solfataricus MCM protein N-terminal domain. Nucleic Acids Res 36:3235–3243.

Lyubchenko, Y. L. Shlyakhtenko, R. Harrington, P. Oden, S. Lindsay, (1993) Atomic force microscopy of long DNA: imaging in air and under water. Proc. Natl. Acad. Sci. USA 90. 2137-2140.

Lyubchenko, Y. L. Gall, A. A. and Shlyakhtenko, L. S. (2014) Visualization of DNA and Protein-DNA Complexes with Atomic Force Microscopy. Methods Mol Biol. 1117: 367–384.

Maiorano D, Lutzmann M and Mechali M. (2006). MCM proteins and DNAreplication. Curr Opin Cell Biol 18:130–136.

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

Medagli, B & Onesti, S. (2013) Structure and Mechanism of Hexameric Helicases. . Spies, M. (ed.), DNA Helicases and DNA Motor Proteins, Advances in Experimental Medicine and Biology, Springer Science +Business Media New York. 75-95.

Miller, J. M. & Enemark, E. J. (2015) Archaeal MCM Proteins as an Analog for the Eukaryotic Mcm2–7 Helicase to Reveal Essential Features of Structure and Function. ARCHAEA.

Miller, J. M., Arachea, B. T., Epling, L. B. and Enemark,E.J. (2014) Analysis of the crystal structure of an active MCM hexamer. Elife, 3, e03433.

Möller, C. Allen, M. Elings, V. Engel, A. Daniel, J. Müller. Tapping-Mode Atomic Force Microscopy Produces Faithful High-Resolution Images of Protein Surfaces. Biophys. J. 77, 1150 ~1999.

Moreno-Herrero, F. Colchero, J. & Baro, A. M. (2003) DNA height in scanning force microscopy. Ultramicroscopy. 96.167-174.

Moreno-Herrero, F. Colchero, J. Gómez-Herrero, J. & Baró, A. M. (2004) Atomic force microscopy contact, tapping, and jumping modes for imaging biological samples in liquids. Phys. Rev. 69, 031915.

Moreno-Herrero, F. de Pablo, P. J. Álvarez, M. Colchero, J. Gómez-Herrero, J. & Baró, A. M. (2003) Jumping mode scanning force microscopy: a suitable technique for imaging DNA in liquids. Appl. Surf. Sci. 210, 22–26.

Morita, S. Wiesendanger, R and Meyer, E. (2012) Noncontact Atomic Force Microscopy, Volume 1.

Moyer SE[1], Lewis PW, Botchan MR. (2006) Isolation of the Cdc45/Mcm2-7/GINS (CMG) complex, a candidate for the eukaryotic DNA replication fork helicase. Proc Natl Acad Sci U S A.103(27):10236-41.

Müller D. J., Fontiadis D., Scheuring, S. Müller, S. A. and Engel, A. (1999) Biophys. J. 76, 1101.

Nishitani, H. & Lygerou, Z. (2002) Control of DNA replication licensing in a cell cycle .Genes to Cells. 7, 523–534.

Ogura, T. and Wilkinson, A. J. (2001) AAA+ superfamily ATPases: common structure--diverse function. Genes Cells. 6(7):575-97.

Onesti S. and MacNeill S.A. (2013) Structure and evolutionary origins of the CMG complex. *Chromosoma*. 122, 47-53.

Pashley R. M. (1981) Hydration forces between mica surfaces in a ueous electrolyte solutions. J. Colloi2 Interface Scz. 80:153-62.

Picco, L. M., Bozec, L., Ulcinas, A., Engledew, D., Antognozzi, M., Horton, M., et al. (2007) Breaking the speed limit with atomic force microscopy. Nanotechnology 18, 044030.

Pietrasanta, L. I. Thrower, D. Hsien, W. Rao, S. Stemmann, O. Lechner, J. Carbon, J. & Hansma, H. (1999) Probing the Saccharomyces cerevisiae centromeric DNA (CEN DNA)–binding factor 3 (CBF3) kinetochore complex by using atomic force microscopy. Natl. Acad. Sci. 96, 3757–3762.

Pucci B, De Felice M, Rocco M, Esposito F, De Falco M, Esposito L, Rossi M and Pisani FM. 2007. Modular organization of the Sulfolobus solfataricus mini-chromosome maintenance protein. J Biol Chem 282:12574–12582.

Pucci,B., De Felice,M., Rossi,M., Onesti,S. and Pisani,F.M. (2004) Amino acids of the Sulfolobus solfataricus mini-chromosome maintenance-like DNA helicase involved in DNA binding/remodeling. J. Biol. Chem., 279, 49222–49228.

References

Remus, D. Diffley, J. F. Eukaryotic (2009) DNA replication control: Lock and load, then fire, Curr. Opin. Cell Biol. 21:771–777.

Remus,D. Beuron,F. Tolun,G. Griffith,J.D. Morris,E.P. & Diffley,J.F. (2009a) Concerted loading of Mcm2–7 double hexamers around DNA during DNA replication origin licensing. Cell, 139, 719–730.

Rivetti C, Walker C, Bustamante C. (1998) Polymer chain statistics and conformational analysis of DNA molecules with bends or sections of different flexibility. J MOL BIOL. 280(1):41-59.

Rivetti, C. Guthold, M. & Bustamante, C. (1996) Scanning force microscopy of DNA deposited onto mica: EquilibrationversusKinetic trapping studied by statistical polymer chain analysis. J. Mol. Biol. 264. 919-932.

Rohrer, H. (1994). Scanning tunneling microscopy: a surface tool and beyond. Surface science 299: 956-964.

Rothenberg E[1], Trakselis MA, Bell SD, Ha T. (2007) MCM forked substrate specificity involves dynamic interaction with the 5'-tail. J Biol Chem. 282(47):34229-34.

Rothenberg, E. Trakselis, M. A. Bell, S. D. and Ha, H. (2007) MCM Forked Substrate Specificity Involves Dynamic Interaction with the 5′-Tail* J Biol Chem, 282, 34229-34234.

Sakakibara N, Kasiviswanathan R, Melamud E, Han M, Schwarz FP, Kelman Z (2008) Coupling of DNA binding and helicase activity is mediated by a conserved loop in the MCM protein. Nucleic Acids Res.36(4):1309-20.

Sakakibara N, Schwarz FP, Kelman Z. (2009) ATP hydrolysis and DNA binding confer thermostability on the MCM helicase. Biochemistry. 48(11):2330-9.

Santos, S. V. Barcons, H. K. Christenson, J. Font, N. H. Thomson, (2011) The intrinsic resolution limit in the atomic force microscope: implications for heights of nano-scale features. PLoS One .6 (8) .e23821.

Sclafani, R. A. & Holzen, T. M. (2007) Cell Cycle Regulation of DNA Replication. Annu Rev Genet. 41: 237–280.

Shechter, D. F., Ying, C. Y. & Gautier, J. (2000). The intrinsic DNA helicase activity of Methanobacterium thermoautotrophicum delta H minichromosome maintenance protein. J. Biol. Chem. 275, 15049–15059.

Shlyakhtenko, L. S. Gall, A. A. and Lyubchenko, Y. L. (2013) Mica Functionalization for Imaging of DNA and Protein-DNA Complexes with Atomic Force Microscopy. Methods Mol Biol. 931: 10.1007/978-1-62703-056-4_14.

Shlyakhtenko, L. S. Lushnikov, A. Y. Li, M. Lackey, L. Harris, R. S. & Lyubchenko, Y. L. (2011) Atomic Force Microscopy Studies Provide Direct Evidence for Dimerization of the HIV Restriction Factor APOBEC3G. J. BIOL. CHEM. 286, 3387–3395.

Shlyakhtenko, L. S. Lushnikov, A. Y. Miyagi, A. and Lyubchenko, Y. L. (2012) Specificity of Binding of Single-Stranded DNA-Binding Protein to Its Target. Biochemistry, 51: 1500-1509

Sun Z, Tan HY, Bianco PR, Lyubchenko YL. (2015) Remodeling of RecG Helicase at the DNA Replication Fork by SSB Protein. Sci Rep. 5:9625.

Sun, J., Evrin,C., Samel,S.A., Fernandez-Cid,A., Riera,A., Kawakami,H., Stillman,B., Speck,C. and Li,H. (2013) Cryo-EM structure of a helicase loading intermediate containing ORC-Cdc6-Cdt1-MCM2–7 bound to DNA. Nat. Struct. Mol. Biol., 20, 944–951.

Sun, J., Fernandez-Cid,A., Riera,A., Tognetti,S., Yuan,Z., Stillman,B., Speck,C. and Li,H. (2014) Structural and mechanistic insights into Mcm2–7 double-hexamer assembly and function. Genes Dev., 28, 2291–2303.

Sushko, M. L. Shluger, A. L. and Rivetti, C. (2006) Simple Model for DNA Adsorption onto a Mica Surface in 1:1 and 2:1 Electrolyte Solutions. Langmuir, Vol. 22, No. 18.

Thomsen, N. D. and Berger, J. M., (2008) Structural frameworks for considering microbial protein- and nucleic acid-dependent motorATPases. MolecularMicrobiology, vol. 69,no. 5, pp. 1071– 1090.

Thomson, N. H., Fritz, M., Radmacher, M., Cleveland, J. P., Schmidt, C. F., and Hansma, P. K. (1996) Protein tracking and detection of protein motion using atomic force microscopy. Biophys. J. 70, 2421-2431.

Van Noort, S. J. T., K.O.van der Werf, A.P.M. Eker,C. Wyman, B.G. de Grooth, N.F. van Hulst, and J. Greve. (1998). Direct visualisation of dynamic protein-DNA interactions with a dedicated Atomic Force Microscope. Biophys. J. 74: 2840-2849.

Van Noort, S. J., van Der Werf, K. O., de Grooth, B. G., and Greve, J. (1999) High speed atomic force microscopy of biomolecules by image tracking. Biophys. J. 77, 2295-2303.

Wegrzyn, K. Fuentes-Perez, M. E. Bury, K. Rajewska, M. Moreno-Herrero, F. & Konieczny, I. (2015) Sequence-specific interactions of Rep proteins with ssDNA in the AT-rich region of the plasmid replication origin. Nucleic Acids Research, 42 (12). 7807–7818.

Wei Z, Liu C, Wu X, Xu N, Zhou B, Liang C, Zhu G. (2010) Characterization and structure determination of the Cdt1 binding domain of human minichromosome maintenance (Mcm) 6. J Biol Chem. 285(17):12469-73.

Wiedemann C1, Ohlenschläger O, Medagli B, Onesti S, Görlach M. (2013) ^{1}H, ^{15}N, and ^{13}C chemical shift assignments for the winged helix domains of two archeal MCM C-termini. Biomol NMR Assign. 8(2):357-60.

You Z, Ishimi Y, Masai H, Hanaoka F. (2002) Roles of Mcm7 and Mcm4 subunits in the DNA helicase activity of the mouse Mcm4/6/7 complex. J Biol Chem.;277(45):42471-9.

Yuan, Z. Bai, L. Sun, J. Georgescu, R. Liu, J. O'Donnell, M. E. and Li, H. (2016) Structure of the eukaryotic replicative CMG helicase suggests a pumpjack motion for translocation. Nature Structural & Molecular Biology.