Chapter 1

General introduction

Before you lies my thesis, the presentation of my research and findings as an “onderzoeker in opleiding” (PhD student) at the VU University in Amsterdam. This short introduction is specifically aimed at introducing the biology and techniques of my thesis to a more general audience. A detailed and specific introduction can be found at the start of each individual scientific chapter. So reader, off you go, into the world of the cell, a world that has surprised, shocked, and continues to amaze me every day.

1.1. Cell biology

1.1.1. DNA

Life on the planet is extremely diverse: plants, trees, mammals, reptiles, fish, and even bacteria are all shaped differently. Yet they share the same code of life: deoxyribonucleic acid, better known as DNA. DNA in its turn is built out of four fundamental ‘letters’: G, C, A, and T (the nucleotides: guanine, cytosine, adenine, thymine). And although we are able to read the words that are formed by these letters, the meaning of
1. General introduction

the genetic code remains a mystery. Almost all cells in our body contain a complete copy of our genetic code. To store this code, the letters are made into long strings called DNA strands. Two of those strands can pair, as each letter only fits with one other, namely: G–C and A–T. These double-stranded DNA (dsDNA) molecules adopt the well-known double–helical structure of DNA, see figure 1.1.

**Figure 1.1 DNA:**

Deoxyribonucleic acid. In the middle the bases that make up the base pairs: G–C and A–T. The gray band represents the backbone of the DNA which consists of sugars and phosphate groups. The 3D structure is the well known DNA double helix.

1.1.2. Proteins

DNA contains the recipe for making proteins, in a process called protein synthesis. Each protein is like a little machine that performs one or more specific tasks. Together they build, operate, and maintain the cell. Proteins that speed up chemical reactions are referred to as enzymes, and the molecule that the protein acts on is often called ‘substrate’. In the next sections I will describe the proteins and their specific function that you will find in my thesis.
1.1.3. Restriction enzymes – DNA looping

Restriction enzymes (RE, or restriction endonucleases) are proteins that defend the cell against viruses. Viruses attack cells by injecting them with their viral DNA. Restriction enzymes recognize and bind to a specific sequence on the viral DNA, called restriction or recognition site. After binding the RE cuts the DNA to prevent infection of the cell (figure 1.2). The discovery and isolation of the first restriction enzyme was awarded with a Nobel Prize in 1978. Today, 3,000 different restriction enzymes are discovered in both bacteria and Archaea\(^1\), humans have developed different systems to take care of viruses. They have since become one of the fundamental tools in biochemistry labs for modifying and manipulating DNA.

The restriction enzymes found in my thesis function by binding two recognition sites at the same time. As a result, the DNA between the two sites is trapped in a loop. Once the loop is formed, the restriction enzymes destroy the viral DNA by cleaving it (figure 1.2). The enzymes in my thesis do not require any fuel to function (fuel in the cell mostly comes in the form of ATP, adenosine triphosphate). Instead, they require a magnesium ion (Mg\(^{2+}\)) to cut the DNA. When calcium (Ca) is used instead of magnesium, they can bind to the restriction sites and form the DNA loop, but no cleaving can take place. It is the process of DNA loop formation and the stability of these loops, that I have studied in my thesis.

Why is the study of DNA looping interesting? Well, we know that DNA loops are involved in many important processes in the cell. This comes as no surprise as, anywhere proteins interact with multiple sites on the same DNA molecule, a DNA loop might occur. But it turns out that these DNA loops serve an actual function. For example: a DNA

\(^1\)Organisms that make up the fourth kingdom of life. They where long thought to be bacteria, but have clearly evolved on a separate path.
1. General introduction

Figure 1.2 Schematic representation of a Type II restriction enzyme cleaving DNA. The restriction enzyme associates with two recognition sites, thereby inducing a loop in the DNA. The DNA is subsequently cleaved to destroy it.

loop can stop other proteins from accessing certain parts of the DNA, thus acting as a regulator for protein production or DNA duplication. We find another common application of DNA loops in the folding of DNA, called condensation.

1.1.4. Architectural proteins – DNA condensation

Chapter 4 and 5 deal with DNA condensation. Throughout the domains of life, cells share a similar problem: they contain a huge amount of DNA compared to the size of the cell, making it difficult to accommodate the whole genome inside the cell. To take care of this problem, cells compress their DNA. For example, the length of the bacterial genome is around 1.5 mm, which occupies, if not compressed, around 200 μm³. However, a bacterium has a cell volume of about 2 μm³, and thus it needs to compact its DNA at least a factor of 100 in order for the DNA to fit in the cell. For human cells the problem is even worse. Our cells typically contain up to 2 meter of DNA, that not only needs to fit in the cell but in an even smaller compartment: the cell nucleus. To achieve this our cells have come up with several clever condensation methods to compress our DNA a factor of 150.000! And what is even more stunning is that the
cell is still able to read the information on the DNA, indicating that the compression and reading is an active and dynamic process.

Although during evolution each organism developed its own set of proteins, known as architectural proteins, to compact their genome. However, the actual method of compaction in each organism seems highly similar. Up until now all cells seem to use one or more of three well-established protein-induced compaction mechanisms. These are: holding two strands of the DNA together (DNA bridging, figure 1.3a), introducing sharp bends to fold the DNA molecule into a smaller space (DNA bending, figure 1.3b), and wrapping DNA tightly around a protein (DNA wrapping, figure 1.3c). Although in chapter 5, I propose a different mechanism that the architectural proteins seem to use in our cells.

Figure 1.3. Three well-known DNA compaction methods. Proteins help condensing the DNA in the cell; shown are the three compaction methods that are well-established in the scientific literature. a. DNA compaction by bridging two strands of DNA. b. Folding the DNA by inducing sharp bends. c. Wrapping the DNA tightly around a protein.
1. General introduction

1.1.5. Membrane–binding proteins – Vesicle fusion

Chapter 7 deals with proteins that bind to biological membranes. What are membranes and what is their function? Membranes basically consist of a layer of lipids, and form structures such as micelles and bilayers (see figure 1.4). They can separate cells from the outside world, or create separate compartments within a cell that allow for different local chemical environments within the cell. Usually a bilayer contains a vast amount of proteins that help to i. stabilize and shape the membrane; ii. transport cargo through the membrane; iii. sense the cell’s environment, and iv. physically connect the cell to the outside world.

Figure 1.4. Biological membranes in action. Biological membranes consist of a lipid layer mixed with proteins, for example a lipid vesicle carrying cargo (top) and a cell membrane (bottom). In the cell membrane we see a protein transporting salt ions across the membrane and other proteins fusing the vesicle to retrieve the cargo.

The interaction between two membranes (membrane–membrane interaction) plays an important role in our brain. Signals in the brain were
1.2. Single–molecule biophysics

long thought to be dependent on electrical signals only. However, in the early 20th century, it was discovered that neurons can communicate to each other and to non–neuronal cells by sending special chemical signals, called neurotransmitters, through a small space between these cells, the synaptic cleft. The neurotransmitters are packaged inside lipid vesicles, and diffuse towards the membrane of the receiving cell. The vesicles are captured and fused with the outer cell membrane of the receiving cell with help from the membrane proteins, releasing the neurotransmitters inside the cell (see figure 1.4).

1.2. Single–molecule biophysics

The proteins that I described in the previous section are all studied one at a time in this thesis. To help develop a feeling for how remarkable that is, let me calculate the number of water molecules in the (hypothetical) cup of tea the reader is drinking at the moment. Lets say a reasonably sized cup of tea is around 200 ml, thus containing 200 grams of water (1 Liter = 1 kg). Water weighs 18 gram per mole (1 mole means $6 \cdot 10^{23}$ molecules). Thus our cup of tea contains $\approx 7 \cdot 10^{24}$ molecules of H$_2$O that is 7,000,000,000,000,000,000,000! Although we work with much smaller volumes, and dilute the proteins that we study, beating those 24 zeros in order to isolate only one molecule is a big challenge.

Despite the big challenges, the single–molecule approach has proven to be a powerful tool in chemistry and physics. The first observations were done with electronic signals and electron microscopy. With the discovery of the green fluorescent protein, a protein that emits green light (awarded with a Nobel Prize in 2008), single biomolecules can now be visualized at work inside the cell (called in vivo), or in a test tube in the lab (called in vitro). A whole field of single–molecule biophysics has emerged, dedicated to explore the physical details of the cell one molecule at a time. In order to learn about the physical properties of
1. General introduction

1.2.1. Tethered particle motion

A very elegant technique to study a single protein looping a DNA molecule is tethered particle motion (TPM). In TPM a DNA molecule is attached to a glass surface on one end while the other end is attached to a small bead (figure 1.5a). How much of the bead is imaged on a CCD camera and tracked using a computer. The amount the bead can move depends on the length of the DNA (just like a balloon on a string wobbling in the wind). Thus, by measuring the motion of the bead, we learn about the length of the DNA. This is ideal for measuring DNA looping by restriction enzymes (see section 1.1.3), as the DNA loops formed by the restriction enzyme effectively shorten the DNA. This technique allows for the characterization of the state of the DNA, i.e. unlooped, looped, condensed, for multiple molecules at the same time.

1.2.2. Atomic force microscopy

With the word “microscopy” we usually think of a traditional light microscope, in which a sample is magnified using two lenses. However, “microscopy” is used in a much broader sense, simply meaning “to image”. One such very loose interpretations is the atomic force microscope, or AFM (figure 1.5b). In the AFM a sample is “imaged” using a sharp needle, called a tip, that feels the roughness of the surface, much like the needle of a record player feels the grooves inscribed on a record (see figure 1.5b). With the AFM, structures that are are impossible to see with
1.2. Single–molecule biophysics

Figure 1.5. Single–molecule techniques used in my thesis.  

a. Tethered particle motion  The motion of the bead (semi–circles) is related to the length of the DNA tether. If a proteins induces a loop in the DNA, the motion of the bead gets reduced from the black semi–circle to the gray semi–circle.

b. Atomic force microscopy  DNA incubated with proteins is imaged by very gently probing the surface with a very sharp needle (tip). When the tip moves over the DNA the tip gets slightly bent. This bending of the cantilever is measured using a laser and a sensor.

c. Laminar flow in the optical trap. Using a microfluidic flow cell creates flow channels that do not mix (there are no physical barriers involved). This allows for an easy buildup of an experiment: catching two beads in the traps; attaching a single–DNA molecule between the beads; probing the DNA in buffer; loading the protein onto the DNA; and probing the DNA again to observe the effect of the proteins on the DNA.

d. Combining optical trapping and fluorescence microscopy. Fluorescent markers are attached to our proteins of interest (top). By taking an image with a CCD camera, single–proteins can be seen on the DNA (bottom). Note that the DNA itself is not visible in this CCD image.
a normal light microscope can be imaged, for example DNA which is only a few nanometers thick. Although one can also manipulate single-molecules using an AFM, by gluing specific molecules to the tip, in my thesis the AFM is only used to image DNA and proteins.

1.2.3. Optical tweezers

One of the main tools I have been working with are optical tweezers. Optical tweezers use tightly focused laser light to catch and manipulate small particles, like a tractor beam. With this laser we attracts and ‘catches’ small polystyrene beads, which can then be moved in all dimensions by adjusting the position of the laser (using mirrors and smoke lenses). After catching two beads in two different laser beams, we attach a single DNA molecule between them (the DNA is chemically modified in order to stick to the beads). Next we introduce proteins, or for example, different chemicals and address questions such as: how much force do we need to apply in order to pull off the proteins from the DNA, or how did the mechanical state of the DNA change? These questions are answered by measuring the amount of force that we need to apply in order to stretch the DNA molecule, we’ll get back to this in a bit.

For easy buildup of our experiment, we use a flow cell consisting of multiple ‘streams’ of liquid (we call them channels). Each channel contains one ingredient that is needed for our experiments (beads, DNA, proteins, etc), and since the amounts of liquids are so little, the streams stay parallel and do not mix (this phenomena is called laminar flow, see figure 1.5c). Using this flow system we can rapidly drag the beads between the different channels, allowing for a fast and flexible build up of our experiments.

Unfortunately, these measurements only probe the ‘global’ changes of the proteins to the DNA. We are only feeling the collective changes
1.2. Single–molecule biophysics

of the proteins on the DNA, and not, directly observing where each protein is situated and what it is doing. To allow direct observation of the proteins we can attach a fluorescent molecule to them that will send out light when it is excited, much like paint or clothing glowing in the disco. Detecting the light of a single protein allows us to localize the protein with nanometer precision, see what it is doing, and count the number of proteins sitting on the DNA (figure 1.5d).

1.2.4. Physical properties of DNA

In cells, DNA is constantly twisted, bent and stretched by numerous proteins as it is copied, repaired, etc. Understanding these essential biological processes requires in–depth knowledge of the behaviour of DNA under tension. Over the past two decades, much effort has been devoted to obtaining a complete picture of the mechanics of DNA. For forces above 65 pN, the double stranded helical structure of DNA starts to break down, and separates into two single strands of DNA, this process is known as force–induced melting. Below these forces, the dsDNA can be described as a rather stiff flexible tube that is able to stretch (much like a garden hose).

Several so called ‘worm–like chain models’ mathematically describe the extension of the dsDNA as a function of force (figure 1.6b). In these models the DNA is described by several parameters, which are mentioned throughout my thesis: the flexibility of the DNA (persistence length), the total length along the double helix (contour length), the effective spring constant (stretching modulus), and the unwinding of the DNA helix as a function of the force.
1. General introduction

Figure 1.6. Probing the force needed to extend DNA, using an optical trap. a. Two beads that are chemically linked to a DNA molecule are trapped using double optical tweezers. By moving the two beads apart (thus controlling the extension of the DNA), a force is exerted on the DNA molecule. This allows us to construct a force–extension curve. b. Typical force extension curve of double–stranded DNA. As the DNA is extended the force on the molecule rises. The behaviour is modeled using the ‘worm–like chain’ models, allowing quantification of the physical parameters of the DNA.
1.3. Outline of this thesis

Here ends the general introduction of the topics covered in my thesis, I am happy to see you made it this far. As you might have noticed, the introduction covers several different techniques, and an even broader range of topics. Therefore, the scientific part of my thesis is split into three separate parts. Each part consists of two chapters, and is based on multiple articles (see page 205). Before I start with the science, let me point the reader to the acknowledgements on page 203.

In **part I** DNA looping is studied using tethered particle motion (TPM). Chapter 2 contains a method that allows characterization of the kinetics involved in protein–induced DNA looping, including protein–DNA association and dissociation. Using this method, the looping pathways of the type II restriction endonucleases SfiI, Eclk18I (chapter 2) and FokI (chapter 3) are revealed, and compared to the existing biochemical rates and models. In chapter 3, I also investigate the influence of the DNA’s bending and twisting rigidity on the protein–induced looping kinetics. To this end, loops with different topologies and torques are designed and used to test the ability of FokI to capture and stabilize these distinct DNA configurations.

**Part II** contains the study of two different DNA–condensing proteins, using various techniques such as atomic force microscopy, quadruple optical trapping, and dual optical trapping combined with fluorescence microscopy. The changes in the physical properties of the DNA upon protein binding are characterized and related to the enzymes’ biological functions. Furthermore, by constructing ligand binding curves from single–molecule measurements the DNA–binding kinetics of the proteins are obtained. In chapter 4, I study the role of Alba one of the most abundant proteins in Archaeal chromatin organization. I show that Alba is able to both stiffen the DNA and condense the DNA, de-
pending on the DNA–protein stoichiometry. The condensing mechanism of Alba is characterized, and the interplay between the two paralogs Alba1 and Alba2 is revealed. The data yield a structural model that explains the multi modal behaviour and regulation of the Alba proteins. In chapter 5, the DNA–protein dynamics of the human mitochondrial transcription factor A (TFAM), an abundant protein in our mitochondria, are studied. I show that TFAM induces DNA compaction by changing the DNA’s flexibility by local melting of the double helix. A molecular model is proposed that connects the revealed compaction method and found kinetics to its role as a transcription factor.

Part III is dedicated to two new approaches in single–molecule biophysics. Chapter 6 contains a novel single–molecule technique to apply forces to tethered biomolecules. I present a proof of concept for an optical pushing apparatus, which exerts an optical stretching force to DNA–tethers by ‘pushing’ the beads upwards using a collimated laser beam. The advantage of using a laser to apply forces to the molecules is the speed at which forces can be changed. Furthermore, the chapter contains an analytical model that is used to calibrate the force in the optical pushing. The model is also applicable to other techniques, such as tethered particles motion or magnetic tweezers. In chapter 7, I present a new method to investigate the interactions of single lipid–binding proteins, using a conventional optical tweezers setup. Using the optical tweezers, two lipid coated beads are brought into close contact to probe the lipid–protein interactions. The new method is used to measure the rupture forces of a membrane bridging protein: Doc2b. Doc2b is involved in the capture of lipid vesicles containing neurotransmitters in our nerve cells.