The DNA double helix challenged by force
This thesis was reviewed by:

prof.dr. Ulrich Bockelmann  
ESPCI CNRS UMR Gulliver, Paris

prof.dr. Fred C. MacKintosh  
VU University, Amsterdam

prof.dr. Antoine M. van Oijen  
University of Groningen

prof.dr. Helmut Schiessel  
Leiden University

prof.dr. Mark C. Williams  
Northeastern University, Boston

prof.dr. Gijs J. L. Wuite  
VU University, Amsterdam

This work is part of the research programme ATLAS, a European Commission-funded Marie Curie early stage training network. The research was performed in the section ‘Physics of Complex Systems’ at the VU University, Amsterdam.
promotor: prof. dr. G.J.L. Wuite
copromotor: dr. ir. E.J.G. Peterman
# Contents

1 Introduction 1

1.1 What is (so exciting about) life? 2

1.2 Cells 7

1.2.1 What makes cells alive? 8

1.3 DNA, the molecular carrier of life 13

1.3.1 From DNA to proteins (one-way road) 17

1.3.2 DNA replication 23

1.4 Cells tightly control their protein levels 25

1.4.1 Epigenetics 27

1.4.2 Chromatin 28

1.5 DNA as a physical object 30

1.5.1 DNA as a garden hose 30

1.6 Outline of this thesis 34

2 Combining optical tweezers and fluorescence microscopy 37

2.1 Introduction 38

2.2 Instrumentation 40

2.2.1 Optical Tweezers 41

2.2.2 Fluorescence microscopy 47

2.2.3 Microfluidics 50

2.3 Preparation of reagents 53

2.3.1 Terminally-labeled DNA molecules 53

2.3.2 Fluorescent labeling of DNA 54

2.3.3 Fluorescent labeling of proteins 55

2.4 Combining optical trapping, fluorescence microscopy and microfluidics 56

2.4.1 Example protocol II: Sequential isolation and visualization of a single DNA molecule using YOYO-1 56

2.4.2 Example protocol III: binding of RPA to ssDNA 58

2.4.3 Example protocol IV: binding or RAD51 to dsDNA 60

2.5 Conclusions 61

3 Visualizing the structure of overstretched DNA 63

3.1 Introduction 64

3.2 Results 64

3.2.1 DNA overstretching with unconstrained extremities 66
<table>
<thead>
<tr>
<th>CONTENTS</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>3.2.2 Overstretching torsionally constrained DNA</td>
<td>71</td>
</tr>
<tr>
<td>3.3 Discussion</td>
<td>73</td>
</tr>
<tr>
<td>3.4 Methods</td>
<td>75</td>
</tr>
<tr>
<td>4 Twist, stretch and melt: quantifying how DNA complies to tension</td>
<td>79</td>
</tr>
<tr>
<td>4.1 Introduction</td>
<td>80</td>
</tr>
<tr>
<td>4.2 Results</td>
<td>80</td>
</tr>
<tr>
<td>4.2.1 Unwinding and overwinding of DNA under tension</td>
<td>82</td>
</tr>
<tr>
<td>4.2.2 Force-induced DNA-strand unpeeling</td>
<td>85</td>
</tr>
<tr>
<td>4.2.3 Hysteresis during DNA strand reannealing</td>
<td>90</td>
</tr>
<tr>
<td>4.3 Discussion</td>
<td>92</td>
</tr>
<tr>
<td>4.4 Methods</td>
<td>93</td>
</tr>
<tr>
<td>5 Balance between two overstretching modes: melting in the interior vs. from the extremities</td>
<td>95</td>
</tr>
<tr>
<td>5.1 Introduction</td>
<td>96</td>
</tr>
<tr>
<td>5.2 Results</td>
<td>97</td>
</tr>
<tr>
<td>5.2.1 An alternative overstretching mechanism in the absence of unpeeling</td>
<td>97</td>
</tr>
<tr>
<td>5.2.2 Melting bubbles appear upon unpeeling arrest</td>
<td>99</td>
</tr>
<tr>
<td>5.2.3 Electrostatic shielding assists melting-bubble formation</td>
<td>100</td>
</tr>
<tr>
<td>5.2.4 Unpeeling stalls in GC-rich regions</td>
<td>102</td>
</tr>
<tr>
<td>5.2.5 Melting bubble formation initiates at AT-rich regions</td>
<td>103</td>
</tr>
<tr>
<td>5.2.6 Both melting modes differ significantly when approaching final strand separation</td>
<td>105</td>
</tr>
<tr>
<td>5.3 Discussion</td>
<td>106</td>
</tr>
<tr>
<td>5.4 Methods</td>
<td>108</td>
</tr>
<tr>
<td>6 Melting bubble statistics and dynamics of overstretched DNA</td>
<td>109</td>
</tr>
<tr>
<td>6.1 Introduction</td>
<td>110</td>
</tr>
<tr>
<td>6.2 Results</td>
<td>111</td>
</tr>
<tr>
<td>6.2.1 Monitoring the magnitude and the dynamics of fluctuations during force-induced bubble melting</td>
<td>111</td>
</tr>
<tr>
<td>6.2.2 Visualizing the number of bubbles</td>
<td>113</td>
</tr>
<tr>
<td>6.2.3 Dynamics of melting fluctuations at different stages of the overstretching transition</td>
<td>115</td>
</tr>
<tr>
<td>6.2.4 Modeling DNA fluctuations during DNA overstretching</td>
<td>115</td>
</tr>
<tr>
<td>6.3 Discussion</td>
<td>122</td>
</tr>
<tr>
<td>Section</td>
<td>Page</td>
</tr>
<tr>
<td>------------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>6.4 Methods</td>
<td>124</td>
</tr>
<tr>
<td>7 Single-stranded DNA templates for optical tweezers-experiments</td>
<td>127</td>
</tr>
<tr>
<td>7.1 Introduction</td>
<td>128</td>
</tr>
<tr>
<td>7.2 Materials and Methods</td>
<td>129</td>
</tr>
<tr>
<td>7.2.1 Biotinylation of two plasmid DNAs (pTR19-ASDS and pKYB1) on both ends of one strand</td>
<td>129</td>
</tr>
<tr>
<td>7.2.2 Optical Tweezers Setup</td>
<td>131</td>
</tr>
<tr>
<td>7.3 Results</td>
<td>132</td>
</tr>
<tr>
<td>7.3.1 dsDNA denatures at tensions exceeding the overstretching force</td>
<td>132</td>
</tr>
<tr>
<td>7.3.2 Dynamics of force-induced DNA denaturation</td>
<td>134</td>
</tr>
<tr>
<td>7.4 Discussion</td>
<td>137</td>
</tr>
<tr>
<td>List of publications</td>
<td>139</td>
</tr>
<tr>
<td>References</td>
<td>141</td>
</tr>
</tbody>
</table>
Introduction

Making the simple complicated is commonplace; making the complicated simple, awesomely simple, that’s creativity.

Charles Mingus

Charles Mingus, the famous jazz composer and bass player (1992-1979), is talking about music here, but actually the quote is far more general. Jazz musicians and scientists face some similar challenges. Both the scientist in his laboratory, the jazz musician during a gig, are exposed to significant complexity; nature itself for the scientist, or the seemingly effortless interplay of several skilled musicians during a jazz piece. Only the talent to break down all the input into small pieces, seeing structure and patterns in what happens around them enables both, the scientist and the jazz musician to function; channeling the surplus of input into a creative output.

Life and thus Biology is, by all means, intrinsically tied to complexity. Scientists studying life, mostly unaware of Charles Mingus’ thoughts, still follow his brain wave, reducing complexity, studying isolated systems, searching for structure, repetition and general patterns. Even though life comes in a multitude of different appearances, we can find a surprising amount of similarities when we zoom in, and look at how life is organized on the small scale. On the molecular level, organisms that look completely different on the large scale reveal very similar features. One of the most fundamental observations on the molecular level of life is the ubiquitous presence of DNA. Life and DNA are not separable, DNA is essential for life.

DNA will be in the focus in this thesis. Since DNA is a vital ingredient of life, I will first try to shed some light on its place in life in this introduction.
CHAPTER 1. INTRODUCTION

The first question that I want to address in my thesis is maybe the most fundamental one:

**Why is it so exciting to study life?**

Naturally, and in contrast to the rest of this thesis, I will not even attempt to provide an objective view on the first question in this introductory part, since I believe these questions can and should be only addressed in a personal way.

1.1 What is (so exciting about) life?

Life surrounds us every day. One of the first observations we make really early in our life is that life, and nature, is immensely diverse. An example of this is illustrated in figure 1.1. This diversity of life is one of the most stunning and fascinating facts that makes people go to the zoo, watch wildlife documentaries, and makes little children follow each beetle or butterfly they see. In other words, it naturally brings out the scientist, or the discoverer in us.

Life comes in a vast amount of different forms, and continues to surprise us in its creative ability to adapt to such an enormous variety of different environments. This talent to blend into almost every niche on this planet inspired Charles Darwin (1809 - 1882) to formulate his theory of ‘the origin of species’ [1]. Darwin was a keen observer and careful analyst. His early interest in science, greatly stimulated by the naturalist and explorer Alexander von Humboldt, lied in Geology. This turned into a great advantage for Darwin’s work, as it trained him to think in very long, geological time scales and anticipate dynamics of systems that initially seem stagnant from the human perspective. Most importantly, he was given the chance to witness the diversity of life. In his 5-years lasting expedition on board of the ‘Beagle’, he studied the flora and fauna on many islands, the Andes, the rain forests in Brazil and the endless prairies of Argentina. Here he found many fossils of already long extinct species, planting the seed for the idea of life as a dynamic entity. He observed that spatial separation due to a natural barrier like the Andes leads to a significantly different flora and fauna on the west and east flanks of this enormous mountain ridge, even though the climate is very similar. Isolated habitats like the pacific islands furthermore encouraged his thought that the so-called creation of life is not static. Darwin inevitably concluded that not only the surface of the earth is changing on very long time scales, life itself must do the same, even though the process is so slow that we have difficulties following its dynamics. Sometimes humans accelerate this process. We domesticated animals and plants like cows, sheep, dogs and many plants
1.1. WHAT IS (SO EXCITING ABOUT) LIFE?

Figure 1.1 – Aesthetic examples of the beauty and the diversity of life, represented through two ecosystems separated by approximately 200 meters, at the coast of the Red Sea, Egypt. [A]. A coral reef in the red sea, out of a diver’s perspective. Coral reefs are considered to be one of the 'hotspots' of life; a significant amount of marine life is concentrated into these relatively rare habitats. Striking is the diversification of life. The central organisms are coral polyps, that form a symbiosis with single celled organisms (Zooxanthellae from the eukaryotic kingdom of life), and build this beautiful reef structures from calcium carbonate. The coral reef is inhabited by a large diversity of crustaceans (like crabs and lobsters) and mollusks (slugs, snails and the less diverse squids, cuttlefish and octopus). These invertebrates are the major food source for some reef fish, to a large extend from the diverse family of wrasses (about 500 different species, ranging from several centimeters up to two meters for an adult humphead wrasse). Another class of fish that explores a different food source are butterflyfish and angelfish. These colorful, exotic fish feed on zooplankton, coral polyps, sea anemones, algae and so forth. Larger, carnivorous fish like sharks and barracudas mark the end of the food chain (not considering the impact of humans). Some sharks, like reef sharks, are permanent residents, while other nomads of the oceans, like the hammerhead shark, just occasionally pays a visit, sometimes only to receive some care by cleaner fish that remove parasites. [B]. A harsh environment, the desert at the Sinai Peninsula, chronically low on nutrients and water forms a totally different ecosystem. Here, the fundamental importance of a low turnover of water shaped life.

like wheat, corn and wine. Humans selected desirable features, isolated and bred them, and therefore accentuated these forms of life. Darwin was well aware of this: "a beautiful part of my theory, that domesticated races of organics are made by precisely same means as species - but latter far more perfectly and infinitely slower" [2].
Darwin was faced with two fundamental questions: where does this dynamics come from, and what laws does it follow? He was only able to answer part of the second question. He used the following ingredients for his approach:

- A high rate of reproduction which would lead to a constant growth of individual species, given unrestricted food and space.
- Since our earth is limited, species are under constant competition for food and habitats. Competition for bare survival leads to pressure.

Then he combined these considerations with an inherent plasticity of species. Over long time scales, species will stochastically evolve in every direction. Importantly, these drifts are non reversible. Combined with the environmental pressure, the often quoted 'struggle for survival', this leads to the 'survival of the fittest\(^1\). Nature itself decides what the most favorable adaptation is. Darwin did not need a controlling hand to explain the diversity of life, in his elegant approach, 'nature selects'. A dangerous thought at his time. He was convinced that even though there exists a law for the selection of favorable species, the intrinsic dynamics are stochastic. Humans, as a consequence, are nothing more than an accident. A thought that was not in line with all the dogmas of the church. Challenging these can be dangerous, as other revolutionary scientists like Nicholas Copernicus and Galileo Galilei had already experienced.

Filling some gaps in Darwin’s theory about the origin of species required the work of others. Hugo de Vries, a Dutch botanist (1848 - 1935) made two crucial contributions. He was the first to argue that species must have a physical object that contains the relevant information, the 'blueprint of life' (1889):

"According to our present conception of all nature, the wonderful phenomena of heredity must have a material basis, and this basis can be no other than the living protoplasm"

He named this object (pan) genes [3]. Furthermore he showed that genes can become modified. He demonstrated two distinct mechanisms how genes can change. First, and unaware of the groundbreaking but at that time widely ignored work of Gregor

---

\(^1\)This terminology does not come from Darwin himself. The British Philosopher Herbert Spencer coined this expression after reading 'The Origin of Species', and was thereby setting the foundation of an abomination of Darwins theory of evolution, the so called Social Darwinism. Today we know that cultural and biological evolution work fundamentally different. While Darwins idea of natural selection works well for Biology, cultural development follows the concepts of Jean-Baptiste Lamarck (1744-1829). What you actively achieve and develop, you pass on to the next generation. Neither the wheel nor the transistor is found in our genes.
1.1. WHAT IS (SO EXCITING ABOUT) LIFE?

Mendel (1822 - 1884); the first to find the laws of inheritance; he showed that sexual reproduction has a mechanism to alter the genetic code of the offspring. During his experiments, where he hybridized several different plant species, he found that the genetic code of the offspring represents a stochastic mix of the genetic information of the mother and the father. He was thereby among the first to demonstrate the powers of statistics in complex sciences such as genetics. Furthermore, De Vries also observed sudden, non-continuous changes in genes (or more precise, in the appearance, the phenotype), during his experiments on the evening primrose [4]. He was able to increase the frequency of these jumps using radioactivity and chemicals. Mutations, sudden jumps in genes, where therefore discovered shortly after Max Plank revolutionized physics by introducing ‘quantum jumps’ in the spectrum of black body radiation (1899). We now know that the stochastic inheritance of chromosomes during sexual reproduction and mutations are not the only ways that genes can change. Processes like gene drift or lateral gene transfer induce stochastic dynamics in the genetic code on a long timescale. The genetic code is a very dynamic entity.

From Darwin we learned that life originates from common ancestors, and the branching of the ‘tree of life’ imposed a structure in the relationship among the different forms of life. Some species are more related, some share almost no similarities. Our knowledge on all the species is very limited; we first focussed on species that are closely related, live in similar environments and that are relatively big, and thus easy to observe. We have categorized almost all mammals (∼4,200 species), birds (∼9,200 species), reptiles (∼6,400 species) and terrestrial plants (∼270,000 species) [5], while other groups challenge us with their astronomical number of species2 or with their small size and their inaccessible habitats (it is estimated that only about 1% of the about six million different microbial species are known [7]).

Looking at this astonishing diversity from a scientific point of view, is Biology not, confronted with such an enormous variety, a never-ending, tedious task? Since each organism looks different, where to start? And why should we, from the practical and medical perspective, study any other organism than the human?

Here comes the fascination of Biology. Life discloses patterns when looking at the molecular scale. As fundamentally different as the camel and the palm tree of figure 1.1B look from the outside, they use the same building blocks. As will be discussed in detail in the rest of this introduction, life among all species uses the same chem-

2Interestingly, the most diverse class of life are insects, with an estimate of six to ten million different species, which led to the famous quote of the pioneer geneticist John Burdon Sanderson Haldane (1892 - 1964) when he was asked what he can infer about the mind of the creator from his creation: “An inordinate fondness for beetles” [6])
ical building blocks, DNA and proteins. Furthermore, the specific tasks that these molecules perform in life seem highly conserved. Similar to LEGO™, life forges a gigantic amount of different appearances, represented by the astronomical number of different species, by modulating and rearranging these building blocks of life. Here lies, my personal view, a fascination of life to a scientist.

*If the astronomical complexity in life boils down to slight readjustments of the same molecular building blocks, it should be possible to understand life in its whole beauty by studying these building blocks in detail.*

This thought is pretty popular and the reason why we study each nut and bolt of life in detail. It embodies a bottom-up approach. This brain wave led to ideas like 'once we know the genetic information of human DNA, we know everything about the human'. Ten years after the Human Genome Project and Celera Genomics revealed the genetic code of humans, how much did we learn?

Francis Collins, leader of the Human Genome Project, commented:  "It’s hard to overstate the importance of reading our own instruction book and that’s what the Human Genome Project is all about." Now, ten years later, Craig Venter, the leading figure behind the efforts of Celera Genomics comments on the human genome, its role for medicine and false expectations: "And what else have I learned from my genome? Very little. We couldn’t even be certain from my genome what my eye color was. Isn’t that sad? Everyone was looking for miracle 'yes/no' answers in the genome." And furthermore: "Why did people think there were so many human genes? It’s because they thought there was going to be one gene for each human trait. And if you want to cure greed, you change the greed gene, right? Or the envy gene, which is probably far more dangerous. But it turns out that we’re pretty complex. If you want to find out why someone gets Alzheimer’s or cancer, then it is not enough to look at one gene. To do so, we have to have the whole picture."

It looks like that diversity cannot be found in a gene, we have to cast a closer look in the interactions between them. The details of all ingredients to life might play the secondary role. How they are wired together determines how life looks like. Of course this is an interpretation, and since we don’t understand life, this consideration is closer to philosophy than science. But it substantially enriches life in the eye of the scientist, and would add an additional dimension to the fascination of life:

*The beauty of life can be found in its design principle itself, or more explicitly, how self organization wires something immensely complex out of, when considered individually, rather simple ingredients.*
1.2. CELLS

DNA is not only the central subject of this thesis, it also sits right in the middle of this dilemma between the bottom-up and top-down approach. We know now how to read the code on DNA, but how can we decipher the meaning? Should we study each gene in detail, and build up from there, or is it wise to turn the focus away from the details, and look at how genes are wired? The answer here is not fully known, so I will cast the focus on what we understand: In particular I will discuss why DNA is necessary for the organization of life, give details about the molecule DNA and the basics of the current knowledge of how life reads and uses its genetic information.

1.2 Cells

How is life organized? One of the first answers on that question came from Robert Hooke in 1665. The British natural philosopher was intrigued by the work of Anthonie van Leeuwenhoek, who is commonly called the father of microscopy and the first microbiologist, and entertained and fascinated the world (or at least the class of higher educated people that had the time and money for such distraction) with his journeys to the microscopic world. Anthonie van Leeuwenhoek was a good businessman that promoted his work strategically intelligent (his technique to fabricate lenses, small glass spheres obtained by rapidly heating thin glass whiskers, was well hidden, while he made the public believe it involved skill and training in glass polishing). His microscopes used a single lens, in contrast to the design of the skilled lens makers Sacharias Jansen and Hans Lippershey that used a two-lens-compound design to fabricate the first telescope (Lippershey, Middenburg, The Netherlands, 1608) and microscope (Jansen, 1609, Middenburg, The Netherlands, 1609). Robert Hooke picked up this two lens design for his own microscope, and was the first one to report the presence of cells in living matter [8]. The word was given by Robert Hooke himself, and originates from Latin word *cellula* 'little room'. He initially discovered cells on cork (as shown in figure 1.2), but it turned out later that all life shares this feature. The finding that life is build up by cells was (and still is) very intriguing. It identifies cells as the autonomous unit of life. The diversity in life can thereby categorized by one guideline, the number of cells.

Life on earth is comprised, for a good share, out of unicellular organisms. The total number of unicellular organisms is astronomical ($4-6 \cdot 10^{30}$ organisms [9]. This colossal number is in the order of a billion times the diameter of our galaxy in meters). Hence, life as a single cell is a very successful concept. More importantly, this directly tells us that everything that we associate with life can be found in an individual cell.
CHAPTER 1. INTRODUCTION

Figure 1.2 – First scientific report of a cell. In 1665, Robert Hooke published this drawing of the cellular structure of cork. This image was taken from [8].

1.2.1 What makes cells alive?

Asking what is exciting about life is naturally linked to the question what makes cells alive. At the first glimpse, a cell is a bag containing lots of complex, big molecules. And inside, there is a lot of biochemistry going on, which mysteriously makes the cell alive. Without going into the strict definitions of life (that are under debate anyways), each cell, when considered alive, must have the three following features:

- A metabolism to provide energy
- A sensorial system and response mechanism that feels and reacts once a cell is triggered
- Some means of reproduction to multiply and create new, similar organisms

Almost all the biochemistry inside a cell is connected to one of these three requirements. Interestingly, without going into the detailed knowledge of molecular and cellular biology, these considerations, along with some very basic observations allow us to infer a lot about the conceptual design of life. The next section will focus on these implications.
Cells need energy

A cell needs a metabolism to maintain an ordered, organized structure. From an entropic point of view, the constant maintenance of such a structure requires energy. A cell must have some means to provide energy, and a sophisticated machinery to continuously preserve its state, and fight the decay. The mechanisms to produce energy are as diverse as life itself, but can be put in three classes. One class of organisms gets its energy directly from sunlight (‘phototrophic’) by, broadly speaking, merging water, light energy and carbon dioxide into oxygen and sugar. As advantageous as the sugar is for the organism itself, as advantageous is the oxygen for us, since these organisms formed the current chemical composition of the atmosphere, as we know it. Most organisms that use this source of energy live in the ocean. Predominantly they are bacteria, so-called cyanobacteria, and furthermore algae, phytoplankton, and of course plants, that inhabit the oceans as well as land.

In another variant of energy-production (‘lithotrophic’) involves harvesting inorganic, energy rich chemicals. These organisms can live in extreme environments. Some of these organisms require oxygen to conduct their chemical energy production (aerobic reactions), but there also exist species that are totally independent of oxygen in their metabolism (anaerobic), and survive extreme conditions, living deep underground, or, as one prominent example, at hot hydrothermal vents, at the depth of the ocean, withstanding temperatures of around 80 °C or more. Our knowledge of these species is very limited; we do not even have a good estimate how many of these species exist on our planet.

The last mechanism how life consumes energy involves consuming other organisms (‘organotropic’). Naturally, since organisms of that class are consumers, contrary to the producers of the other two classes, the total biomass of consuming forms of life must be small to keep a status quo.

The metabolism of a cell is responsible for converting what is consumed (light, energy rich chemicals, other organisms) into a 'fuel' that the cell can make use of, a process that is conceptually similar to the conversion from light, coal, oil or radioactive material into electrical current. Interestingly, the final product (the universal energy source of life) is the same for all species, a molecule called adenine triphosphate (ATP).
CHAPTER 1. INTRODUCTION

Cells need adaptation

In the current view of the origin of species, life as we know it was predominantly forged by the two factors, environmental stress like nutrition shortage, heat, cold or inaccessibility, and continuous competition between species and organisms for habitats and nutrients. Thus, only flexible or well-adapted species survive this constant struggle of life. This implies that life, to be competitive, must have elaborate sensory elements to continuously monitor the internal, cellular state (Am I hungry?) and the environment (is this here a nice place to be?). Even more importantly, this means that cells also continuously react, make decisions based on what they all sense.

Cells have a plan, which means that, broadly speaking, cells react to a certain environmental condition in a more or less homogeneous fashion. The first indication that cells adapt their protein levels according to the environmental conditions came from bacteria [11]. If the bacterium *Escherichia Coli* is exposed to the two food sources lactose and glucose, it always prefers to feed on glucose. Once all the glucose is consumed and only lactose is around, it changes its metabolism to feed on lactose. And this commitment is reversible. Upon addition of glucose, the bacterium ignores lactose, and instead eats its more favorite food.

Cells thus have a plan. More precisely, they do not have one plan; they have lots and lots of plans. And importantly, they use all their senses to decide which plan to execute. To be very flexible and successful, as a life form you need to have a plan for all possible situations you can encounter, even though the cell might not use each plan. A dramatic example of this can be seen from multicellular organisms. For example, nerve cells and tissue cells in humans all originate from the same mother cell, and even though they belong to the same organism, and have the same plan, they look completely different (figure 1.3); they only use a small subset of this plan, according to their function.

Thus, life is *flexible* and *rigid*, both in parallel; very adaptable and fast responding to changing conditions on the one hand, but still using an identical blueprint, such that successful concepts persist. Life meets these two very important criteria of flexibility and rigidity with ease. This suggests that life found a way to keep the two aspects, memory and adaptability conceptually separate, with a very distinct mechanism of interaction. This abstract consideration will receive detailed attention in section 1.3.1, where the focus is cast on the current knowledge of the inside of a cell.

---

3No rule without exception. Unicellular organisms seem to exploit stochasticity to generate fast phenotypic variation among a species to be well prepared for environmental changes. For an overview over the research of this fresh field, see [10].
Figure 1.3 – Cells can look (and function) fundamentally different, even though they carry the same genetic code. Significant research is devoted to understanding the functions of differentiated cells. On the left, a network of human neuron cells; the essence of the brain is shown. The right journal cover shows epithelial cells, the essential cell type of tissue. Both cells have a remarkable difference in shape and function, even though they carry exactly the same DNA.

Cells need a memory

Another demand for life is a memory that can be passed on from parents to their offspring. The identification that the genetic information is stored on chromosomes (see figure 1.4) was a result of the pioneering research of Thomas Hunt Morgan in 1910, who established the fruit fly Drosophila Melanogaster as one of the most important model organism of genetics, and was awarded with the Nobel Prize in Medicine in 1933.

In these days it was not clear how information is stored in the chromosome. But scientists anticipated that the chemical composition of chromosomes must have something to do with the information stored on it. The elucidation of the chemical composi-
CHAPTER 1. INTRODUCTION

Figure 1.4 – Chromosomes, the molecular carrier of heredity. In this microscopy image of garlic cells the chromosomes are stained to enhance the contrast. The upper cell is caught during cell division; the cell is distributing the already duplicated chromosomes evenly, such that each daughter cell has the same genetic information as the mother cell.

vation of chromosomes, along with their specific role took more time, though. It was known that the chromosome, just like the cell, consists predominantly out of two classes of molecules, amino acids and nucleotides. Both classes of molecules form bigger, more complex macromolecules, called proteins (build from amino acids) and DNA (deoxyribonucleic acid, composed out of nucleotides). Biochemists indentified that proteins consist out of a class of twenty different amino acids, while DNA is only built up by four different nucleotides. It was thus a surprise when the experiments of Oswald Avery, Colin MacLeod and Maclyn McCarty (1944) and Alfred Hershey and Martha Chase (1952) demonstrated that the genetic information is stored in DNA [15, 16], which uses a smaller ‘alphabet’. Finally the molecule of heredity was indentified, but its structure remained elusive. The breakthrough came from James D. Watson and Francis Crick in 1953. After seeing the X-Ray diffraction images that Rosalind Franklin obtained at King’s College London in 1952 (The famous ‘photo 51’), they proposed the well-known double-helical structure of DNA [17], which is sketched in figure 1.6.

was widely accepted back then. In his work, he showed that in DNA the number of Adenine and Thymine is same, and the number of Guanine is identical to the number of Cytosine [14]. But even this fact, the Cargaff rule, was misinterpreted by some scientists that claimed that the genetic code consists of only two letters, since A = T and G = C. The quest for the structure of DNA was buzzing, but the direction was unclear.
1.3 DNA, the molecular carrier of life

Each life form, from the simplest bacterium to the complex multicellular animals has one defined way of reading the information in DNA, in sharp contrast to the enormous variety between the languages of humans. But still there is a certain degree of relation between the language we use every day, and the information stored on DNA. Even though there has been significant effort, we are, however, far away from understanding this language.

In our language, there is a well defined structure, or hierarchy. Language starts with a limited amount of building blocks, the alphabet. A defined permutation of vocals and nouns form a word, the first element that has a defined contextual function, or meaning. However, information in a language is not transmitted via single words; it is a sentence, a collection of words in a distinct grammatical context that carries information.

Life uses a somewhat similar approach. DNA carries its information in its chemical composition. The fundamental building blocks of DNA (the alphabet of life) are four different nucleotides: Adenine, Thymine, Guanine and Cytosine, which are commonly abbreviated by A, T, G and C. These nucleotides are threaded, like beads on a string, in a defined sequence. The analogue of a word is called gene\(^5\), a sequence of several nucleotides (the size of a gene varies substantially, but is typically around a thousand nucleotides [18]). The context of a gene, (i.e. how genes form something related to a sentence) is this part of the language of life that is very challenging, and that we understand only in a very limited fashion. What we know about it, and what the difficulties are, will be briefly discussed in the next section. In this section, the focus will be on the DNA molecule, and its chemical composition.

DNA carries lots of genes. The simplest organisms (bacteria) still have about \(\sim 1000\) to 4000 genes (the organism with the least amount of genes, *Mycoplasma genitalium*, carries 470 genes [19]), while more complex organisms, like humans are estimated to have between 20,000 and 25,000 genes [20]. Therefore, DNA is a long molecule. Each human cell\(^6\) contains the whole human genome, about 2.85 billion nucleotides [20], distributed over 23 chromosome pairs. Each human cell contains DNA of the total length of, when stretched out, two meters, even though its diameter is tiny. Its total length is about 970 million times longer compared to the thickness of DNA. The

---

\(^5\)The nucleotide sequence of a gene encodes a protein, which represents the active component, it brings the plan encrypted in the gene into action. Therefore, more precisely, the gene represents the thought, the protein represents the spoken word.

\(^6\)As usual, Biology presents us an exception to the rule. Red blood cells do not have any DNA.
chemical composition and its structure is sketched in figure 1.5.

**Figure 1.5 – Chemical composition of DNA.** [A]. The four nucleotides that build up DNA. The three different modules, the deoxyribose sugar, the phosphate group and the base are highlighted for Cytosine. Also, for Cytosine, the number of each carbon, from 1’ to 5’ is highlighted. For each base, the atoms that form the hydrogen bonds between complementary bases are highlighted in black. [B]. The polynucleotide string that aligns different nucleotides in a sequence. The so called DNA backbone is highlighted in grey. The polarity of the DNA backbone is imposed by the sugar, and is termed by the 3’ and the 5’ end.

DNA is build out of four different nucleotides that share a common global structure (see figure 1.5). All four different nucleotides consist out of three components:

- A purine or pyrimidine base.
- A deoxyribose sugar molecule.
- A phosphate.
1.3. DNA, THE MOLECULAR CARRIER OF LIFE

All four nucleotides use the same sugar and phosphate group, and only differ in the base: DNA uses the two purine bases Adenine (A) and Guanine (G), and the two pyrimidine bases Cytosine (C) and Thymine (T), which are shown in figure 1.5A. The three components have a different function. While the deoxyribose sugar and the phosphate form the DNA backbone (the string), the four different nucleotides are attached to this backbone, and their sequence corresponds to the chemically stored information (figure 1.5B). DNA is composed out of two strands (figure 1.6). It uses the specific base-pairing interaction, mediated by hydrogen bonds, between complementary bases. Adenine forms a pair with Thymine, while Guanine pairs with Cytosine (figure 1.6A). Importantly, Adenine and Thymine are linked via two hydrogen bonds, while Guanine and Cytosine are hold together via three hydrogen bonds, directly implying that the GC base-pairing is stronger than the AT base-pairing. These two strands, and the interaction between the bases, give the DNA molecule its prominent double-helical structure.

Figure 1.6 – The physical conformation of DNA. [A]. Adenine and Thymine form two hydrogen bonds, while Guanine and Cytosine are connected via three hydrogen bonds. The base-pairing between the two complementary nucleotides imposes and angle; leading to the formation of a major and a minor groove. [B]. DNA possesses a double-helical three-dimensional structure.
Since DNA is the central theme of this thesis, I will present some more details on the structure, the chemical composition and the stability of this molecule.

1. The dimensions of the DNA molecule are well known. It has a diameter of 2 nanometers and the base pairs are separated by 0.34 nanometers. The DNA molecule is a helix, with one turn each 10.5 base pairs [17].

2. The alignment of the two strands is not symmetrical (due to the angle between the base paired nucleotides and the connection to the backbone, see figure 1.6A). DNA has a narrow groove and a wide groove, called minor and major groove. This is important for the sequence recognition of proteins. Dependent on the sequence, different atoms of the bases are exposed, and serve as a 'barcode' for the corresponding sequence, such that proteins can recognize different bases without breaking the base pairing.

3. The DNA strand has a polarity, introduced by the deoxyribose sugar. Each carbon of the sugar molecule is commonly labeled with a number (1’ to 5’). The phosphate is linked to this sugar via the 3’ and the 5’ carbon. Therefore the orientation of each strand is termed 3’- 5’ or 5’- 3’. Both strands of the DNA molecule are antiparallel, which means that if one strand has a 3’- 5’ orientation, the second strand is oriented in a 5’- 3’ manner.

4. The structural integrity of the DNA double helix is not given exclusively by the hydrogen bonds between complementary bases. A hydrogen bond is mediated by the attractive interaction of a hydrogen atom with an electronegative atom, such as, for DNA, nitrogen and oxygen. About half of the total bond stability between the two strands stems from base-stacking between adjacent base pairs, which is given by dipole-dipole and hydrophobic interactions [21]. Furthermore, the twist in the stacking of adjacent bases originates from hydrophobic interactions, since the non-polar bases preferentially burry themselves down into the heart of the DNA molecule to avoid contact with water. The last interaction that has an important impact on the structure and stability of DNA is of electrostatic nature. The phosphate groups of the DNA backbone carry two negative charges per base-pair, which makes DNA a strongly charged, self repelling polymer.

5. The stability of both base pairs is not equal. The bond strength of the G-C pairing is about 2.9k_BT, while the A-T base pair is rather weak, with about 1.3 k_BT [22]. In addition to this heterogeneity, the base-pairing strength, and
hence the double-helical structure is not rock solid. An individual base pair is thermally unstable, which means that the interaction cannot permanently hold two short DNA strands together in the presence of thermal motion and vibration, which excites about the same energy ($\sim k_B T$) than the base-pair strength ($\sim 1-3k_B T$) [22]. The DNA is only more or less stable for lengths exceeding about 20 base pairs. The DNA molecule at room temperature, and hence under natural conditions locally opens and closes (it 'breathes') stochastically, since it is continuously thermally bombarded by the surrounding water molecules [23]. This is an interesting point that should be emphasized here. The two individual strands of DNA, carrying the sequence, are built up via covalent bonds, the strongest molecular bond. The three-dimensional structure, which gives DNA its very unique properties, however stems from the complex interplay between three individually rather weak interactions mentioned above. Why it is actually advantageous to have such a weak binding strength that seems to just fulfill what is absolute necessary to maintain the three-dimensional structure of DNA, and not more, will become clear when the focus is cast on how proteins interact with DNA.

1.3.1 From DNA to proteins (one-way road)

As mentioned already, DNA and proteins are the dominant class of macromolecules inside cells. These two classes of molecules are not independent of each other; there is a very tightly regulated link between the two. As discussed in section 1.2.1, DNA contains genes. These genes encode the information for proteins, hence DNA possesses the blueprints of all proteins the cell could eventually need. Consequently, DNA does not resemble a book, the often used analogy to a Shakespeare is flawed. DNA is closer to a dictionary. We cannot read in it how old we will become, how tall children will grow, or if they will develop an extraordinary talent for music. Proteins are composed out of a chain of 20 different amino-acids \(^7\). This chain folds into a three-dimensional structure, and its function is crucially dependent on this structure. Proteins are the working horses of the cell; they put the plan written into DNA in action. They fulfill specific tasks in a large number of different processes that are vital.

\(^7\)The code is degenerate, since there are 64 nucleotide permutations, and only 20 amino acids. Interestingly, the code from the nucleotide sequence to a amino acid sequence is universal (almost; Mitochondria, membrane-enclosed substructures of a eukaryotic cell that have their own DNA and exhibit slight variations of the genetic code).
for the survival of the cell. Some proteins act as catalysts and break down nutrients during digestion, some proteins build up large scaffolds to give cells a distinct form, like nerve cells. Some proteins sense the chemical environment of the cell, while some help repairing the DNA once it got damaged, or duplicate DNA prior to cell division.

Central Dogma of Molecular Biology

DNA contains the information for proteins, but proteins are required to read and process this information. Also proteins are involved significantly in the decision when to produce which protein. Even though it looks pretty confusing on the first glance, there is a strict hierarchy in the process of protein production that is applied in an identical fashion for each life form, from humans to bacteria. This hierarchy was formulated by Francis Crick into the Central Dogma of Molecular Biology [24] (figure 1.7).

\[ \text{Figure 1.7 – Central Dogma of Molecular Biology} \]

It states directionality in the flow of the sequence information between DNA, RNA and proteins, represented by the arrows. By the time Francis Crick postulated this dogma [24], the molecular processes behind these arrows have been mostly unclear, and have been later on elucidated and categorized into: DNA replication and repair, transcription and translation.

It states a general rule for the transfer of information (encoded in a nucleotide sequence) between the fundamental biopolymers DNA, RNA and proteins. RNA

\[ ^8 \text{Some details of this interesting macromolecule, and its role in the cell are given in the following paragraph.} \]
1.3. DNA, THE MOLECULAR CARRIER OF LIFE

(ribonucleic acid, a biopolymer that strongly resembles a single strand of DNA, and only differs in the sugar molecule (Ribo instead of Deoxyribose) of the backbone) was identified as another central biopolymer that seemed to ubiquitously present in the cell.

Generally speaking, it formulates directionality in the information flow between the sequences of the biopolymers. This abstract consideration is graphically presented in figure 1.7. It states that the direction of information flow is a one-way road. The sequence information of the DNA (the gene) is transcribed into a defined sequence of a RNA molecule, which is in turn translated into a sequence of amino acids, and hence a protein. No sequence information of either RNA or of proteins is inserted into DNA. The protein production comprises three steps (see figure 1.8):

**Figure 1.8 – Two essential steps in protein synthesis.** [A]. Transcription. The protein RNA polymerase reads the sequence of a gene and generates a mirror image of this gene in form of a messenger RNA. [B]. This messenger RNA is used by the ribosome, which translates the mRNA sequence into a chain of amino acids, which, upon folding into a defined three dimensional structure, become a functional protein.

Only viruses are known to disturb this general flow. Some of these intruders found a way to insert their genetic information into the host genome, using a special protein called reverse transcriptase [25].
• Step 1: The cell decides that it requires a certain protein, and reads the sequence information. This task is performed by proteins that are involved in the decision process (transcription factors). This class of proteins guides another protein (RNA polymerase) whose function it is to read the DNA sequence, to the gene of the desired protein. *(Initiation of Transcription)*

• Step 2: RNA polymerase reads the DNA sequence of the gene, and in turn produces an RNA molecule (messenger RNA, or mRNA) with the complementary nucleotide sequence to the gene. *(Transcription)*

• Step 3: This mRNA is used by a molecular machine, called the ribosome. It translates the RNA code (polynucleotide chain) into the protein code (amino acid chain). *(Translation)*

Another class of processes sketched in the central dogma of molecular biology is fundamental in DNA maintenance. This, for example, involves DNA replication, where the genome is duplicated before cell division.

**Excursion for interested readers: RNA, or the struggles and glory of Francis Crick**

When stating the central dogma of molecular biology, Francis Crick has proven to be pretty bold. By that time scientists where still pretty puzzled about the mechanism by which a protein is produced from a gene. Back in the 1950s, the running theory, the Sequence Hypothesis, postulated that the sequence of nucleotides in DNA somehow dictates the order of the amino acids of a protein. However, in higher organisms the DNA resides in the nucleus, while it was known that protein synthesis takes place outside the nucleus, in the cellular cytoplasm [26]. Therefore it was clear that the process of protein synthesis needs an intermediate. RNA, a molecule that is chemically very much related to DNA, was a hot candidate as this intermediate. Indeed, the milestone work of the biochemists Jean Brachet and Torbjörn Caspersson around 1940 already gave strong indications that RNA plays a central role during protein synthesis [27, 28]. But as usual, the trouble seemed to be in the details. From the biochemical work during these days it became evident that RNA was not the working image of DNA. Biochemists found that on average, there is a sticking discrepancy between the base composition (they measured the average A-T versus G-C fraction) of DNA and RNA. After investigating the DNA and RNA base composition for lots of different bacteria, they found [29]:

"**The composition of deoxyribonucleic acid in different bacterial species [] was shown to vary widely (from extreme AT type to the highest GC-type); whereas that of ribonucleic**
acid from these species was shown to vary slightly, small differences being noted only between very diverse species"

Scientists at that time had a hard time to rationalize these findings. How can there be a universal relation between the sequence of DNA and that of RNA, when there seems to be no correlation between the average sequence composition of DNA and RNA inside a cell? The confusion back in the days is nicely summarized Francis Cricks report "The Present Position of the Coding Problem" [30]. There, Francis Crick offered six ways out of this dilemma, not without a bitter aftertaste: "[...] though in my view, they all, at the moment, appear unattractive."

1. **Only part of the DNA codes proteins.** DNA consists of parts with sequence information, and other parts that do not code for a protein.

2. **The DNA to RNA transcription mechanism varies.** Different species interpret the DNA code differently during the RNA synthesis.

3. **The DNA code is degenerate.** Most amino acids have several representations; there is a level of ambiguity in the transcription from DNA to RNA.

4. **The code is not universal.** Every organism has their own genetic language.

5. **The nucleic acid code has less than four letters.** Not all nucleotides in DNA encode information.

6. **The amino acid composition of proteins varies.** The variation in the nucleotide composition of the whole DNA molecule will be correlated with the variation in the amino acid composition of the encoded proteins.

With the current knowledge, we would be able to help Francis Crick out of his misery. Now we know that alternative (1) is true. Indeed, DNA is not exclusively built up with genes. But what is the function of these long, so called non-coding DNA stretches? Without going into detail, it becomes more and more evident that these DNA segments have a fundamental role in the regulation of gene expression (see chapter 1.4). Even though DNA and RNA are chemically very much related, it was clear really fast that their role in life is very different. While DNA is chemically very static (the long term memory of the cell, the molecular hard drive), RNA in cells is very dynamic and versatile (more like the molecular RAM). RNA is involved in a multitude of different cellular processes [31]. However, one very fundamental task of RNA is its role in the protein synthesis. The so called messenger RNAs carry the sequence information of DNA to the cytoplasm, where this code is translated into a
protein. This mRNA is the intermediate of protein synthesis, in accordance with the picture of the central dogma of molecular biology, postulated by Francis Crick. Now, it turns out that RNA continues to give Francis Crick and his work a hard time; its role cannot be tracked down easily. First, as already mentioned, RNA does not exclusively act as an intermediate of protein production. Actually, large scale investigations of the transcriptional state of DNA have revealed surprising facts. While only about 1.5% of the human genome contains protein-coding information, the majority of the DNA is transcriptionally active (∼60-80%) [32, 33]; hence RNA cannot be put in the picture of Francis Crick exclusively. There is now solid evidence that the so called non-coding RNA is functional, especially in the regulation of gene expression [31].

Furthermore, RNA directly challenged the central dogma of molecular biology. The trouble lies here. Messenger RNA encodes the information for a protein. But it turns out that for higher organisms (like humans), messenger RNA contains more. It consists out of regions that hold parts of the protein-information (so called exons), and regions with apparently no information on it (so called introns). An additional editing step occurs. After the transcription of messenger RNA, splicing takes place; the introns are cut out and removed, and the exons are linked together. Naturally, this puzzled scientists, what is such a mechanism good for? As it turned out, this process can edit proteins dependent on the demand of the cell. Alternative splicing routes where found, such that the same initial messenger RNA leads to the production of different proteins, dependent on how the exons are spliced together. It was found that different cell types of the same organism use this mechanism to produce different proteins [34]. How is this process regulated? Remember, in the light of Francis Crick’s dogma, no sequence information flow should occur backwards. This means that the control of splicing and its alternative routes should not lay in proteins hands. The RNA itself should encode what different splicing alternatives are present, and when to use which. Deciphering the code which governs RNA splicing turned out to be a tedious work. Nowadays there are strong indications that indeed, the splicing code resides on the RNA itself. However it seems to be hidden in subtle details [35], and to decipher it, huge datasets and strong computational power were needed. So it seems that the existence of RNA splicing did not shatter the fundamentals of the central dogma of molecular biology, but it came close.
1.3. DNA replication

DNA is the molecular representation of the memory of cells. Consequently, cells face a challenge when they decide to divide. The mother cell needs to pass on its memory to both daughter cells. It needs to have a 'DNA replication machine' that can duplicate the present DNA molecule without copy errors. A strong indication how this works was given by James D. Watson and Francis Crick when they elucidated the structure of DNA as a double helix of two complementary DNA strands. In their Nobel Prize honored study the stated:

"It has not escaped our notice that the specific pairing we have postulated immediately suggests a possible copying mechanism for the genetic material" [17].

What they implied was that duplicating DNA involves two steps. First the two strands of DNA are separated. To generate two duplicates of the mothers DNA, to each of the two DNA strands, the complementary strand needs to be fabricated.

DNA replication, the process of duplicating the mother DNA, involves several proteins. However, the two most fundamental proteins in this process can be directly linked to these two functions: Strand separation is conducted by DNA helicases, while DNA polymerases synthesize the complementary strand (see figure 1.9A,B). Therefore, the DNA copy machine consists out of one helicase that separates both strands of the DNA, and two DNA polymerases. As discussed before, the DNA strand is not symmetric along its backbone, the sugar imposes a polarity. This polarity is felt by both, helicases and the DNA polymerases, and imprints a one-way lane on DNA. Helicases involved in DNA replication move in a 5’ to 3’ direction, while DNA polymerases, during replication, move along the opposite orientation, in the 3’ to 5’ direction. Since both DNA strands are oriented in an antiparallel manner, the two DNA polymerases cannot both generate the complementary strand in a continuous fashion. Only one strand, the leading strand, is continuously replicated into a new DNA. The other strand, the lagging strand, is replicated in a discontinuous way (For a more detailed overview, see [36]). Another protein, a DNA primase, sporadically generates short starting points for a DNA polymerase, since this protein can only elongate the incomplete strand (see figure 1.9B); it requires a launch-pad in form of a double-helical structure. From there on, DNA polymerase will synthesize the missing strand by adding new nucleotides, one at a time, to the incomplete strand. This is no random process, DNA polymerase selects the 'right', complementary base with a great accuracy. Replicative DNA polymerases can possess an enormously low error rate, with only one mistake every one billion nucleotides [37]. They reach this astonishing
CHAPTER 1. INTRODUCTION

Figure 1.9 – The DNA replication machinery. [A]. The DNA double helix is opened up by a DNA helicase. [B]. The incomplete strand is fabricated, nucleotide by nucleotide, by a DNA polymerase. [C]. Two DNA polymerases and a DNA helicase form the DNA replication machinery. The polarity of DNA's strands imposes an asymmetry in the fabrication of both complementary strands. While the leading strand is generated in a continuous manner, the lagging strand is formed in a discontinuous fashion. The lagging strand polymerase requires a launch-pad, a so called primer, generated by DNA primase, another protein involved in DNA replication. The long stretches of single-stranded DNA during lagging strand replication are protected by single-stranded binding proteins.

accuracy in part by proofreading. If a mistake was actually made, DNA polymerases switch their direction, move back a little, excise the incorrect nucleotide, and then continue to generate the complementary DNA strand. The last fundamental class of proteins involved in DNA replication are so called single-stranded binding proteins.
1.4 CELLS TIGHTLY CONTROL THEIR PROTEIN LEVELS

These proteins protect the, compared to double-stranded DNA, chemically more unstable intermediate of DNA replication, single-stranded DNA, which is abundant during lagging strand replication (figure 1.9C).

1.4 Cells tightly control their protein levels

Where do the differences between organisms come from? Can they be found in their genes? Do very different organisms have very different genes / proteins? Broadly speaking, the answer is: very different organisms have different proteins, but not very different ones. Humans and mice for example have about 99% of their protein coding genes in common [31]. Wrapped up in a bit of a simplistic way, very fundamental proteins (for example proteins involved in transcription and translation) are remarkably similar between species, while more specific proteins differ. (Interestingly, the small differences in highly conserved proteins like the ribosome can be used to create a molecular measure of how related species are. For a more detailed overview, see [38]). So, where does differentiation in life come from? Actually, significant differentiation can already be found within one species.

![Figure 1.10](image)

**Figure 1.10** – Differences in the appearance (phenotype) can occur if different environmental conditions are present. The plant Arabidopsis, even though genetically identical (having the same genotype), looks different when it grows under different conditions. [A]. In the presence of light the stem stops growing and the leaves unfold to collects light. [B]. If grown in darkness, the plant continuous to grow its step and does not unfold its leaves. Images taken from [39].

An example for the different appearance (a different phenotype) of genetically identical organisms (having the same genotype) can be seen in figure 1.10. Apparently, the
genetic information stored on the DNA molecule is not rigidly linked to a well-defined manifestation. Phenotypes can have plasticity, which is another way for saying that life decides which form is advantageous. Thus, the language of life cannot be found on the DNA, the molecule containing all genes, alone. Since not the genes, but the encoded proteins fulfill the important tasks in a cell, the fundamental question is:

How does a cell control and regulate its own protein levels?

Here comes the challenge in understanding the language of life. Cells actually use a lot of different ways to tune their protein levels, which schematically depicted, in a general way, without focusing on the details, in figure 1.11.

![Figure 1.11](image)

**Figure 1.11 – Different regulation layers of gene expression.** The flow of sequence information (implying the shape and function of a protein) is restricted by the central dogma of molecular biology, but the regulation of the gene expression (implying the number of certain proteins in the cell) is very complex, cells use almost every possible regulatory layer (depicted by an arrow) for the adaptation of the cellular protein levels.

A very fundamental part of the ongoing research in Cellular and Molecular Biology is devoted to this regulation, such that every attempt to give an overview here would fail to the sheer complexity of this process and the incompleteness of our current knowledge. In line with Charles Mingus’ quote at the beginning of this introduction, here I do not want to focus on the details, but focus on the general features of life. For that consideration, it is sufficient to know that the protein levels are not random. There are ‘on-off’ switches that control the protein levels, and these switches are connected in a complex network. We begin to understand that even though the network controlling the cellular protein levels is very complex, it is far from being random [40], there seems to be a strict structure in the regulation of individual genes. How this
structure emerges becomes more and more clear from large-scale genomic investigations. One fundamental property of such control networks is their intrinsic dynamics over many generations. Transcription networks seem to be rather plastic, over time these gene regulation networks rewire, such that new, advantageous constellation are continuously explored \[41\]. But complex organisms have a lot of genes; for humans it was estimated that we have \(\sim 25,000\) genes. Interestingly, it turns out that gene regulation does not exclusively work in a one-by-one fashion. On top of the control of each individual gene, there is another layer of control for large clusters of genes, a ‘master-switch’ that is linked to the conformation of DNA itself, as will be discussed in the next section.

### 1.4.1 Epigenetics

Cells do not transcribe all genes. Some of them are only transcribed under specific external conditions (i.e. enzymes to digest specific nutrients), others are only required in a certain cell type (nerve cells do not need to produce enzymes for digestion), such that cell types of one species differ in their gene expression patterns, but not in their DNA content. The DNA molecule contains the plan for the whole organism, but each individual cell only needs to access a subset of this plan. Still, each cell possesses the full genome. The access to the genetic information in multicellular organisms is, however, strictly regulated, as can be seen in figure 1.12. Different cell types use different gene expression programs that they pass on to their daughter cells after division. Furthermore, from figure 1.12 it can be seen that many of the genes (in a multicellular organism) are not translated. They are silenced.

![Figure 1.12](image_taken_from_42.png)

**Figure 1.12 – Difference in the gene expression patterns of lung versus finger tissue cells.** Image taken from \[42\].
CHAPTER 1. INTRODUCTION

The layer of control that induces an inheritable gene expression pattern is called **epigenetics**. Instead of using a multitude of on and off switches to generate this silencing of big stretches of the genome, cells use a master switch. Cells modify the very nature of the conformational state of DNA to keep genes in the repressed / activated state. As discussed already, the central dogma of molecular biology states that the DNA sequence will not be modified, the genetic information stored on the DNA molecule always remains the same. But instead, the physical conformation of DNA is continuously altered and regulated, such that each cell only has access to the genes it needs to utilize for the particular function of this cell\(^\text{10}\).

Coming back to analogies, in our current understanding, the role of DNA for life is that of a thick dictionary that is not edited (Central Dogma of Molecular Biology). Certain pages are however glued together, while others are highlighted with bookmarks (Epigenetics). The book of life still has kept most of its seven seals.

### 1.4.2 Chromatin

To understand the molecular details of gene silencing, the focus has to be turned to how cells organize their genome, or generally speaking, how an enormously long DNA molecule fits into a very small cell.

DNA of higher organisms (*eukaryotic* species) is spooled around small protein complexes, called histones that resemble a wheel. These histones wrap DNA of a length of about 150 base pairs. This 'beads-on-a-string' structure itself folds into higher structures, as depicted in figure 1.13. This higher structure, called chromatin, is present in two different states. In one configuration, called heterochromatin, this structure is so tightly packed that effectively becomes genetically inaccessible; its information is not read, it is transcriptionally silenced. The other configuration, called euchromatin is less compact, which allows genes in this region to be transcribed on demand\(^\text{11}\). The chromatin state is not random, but subject to a very tight and specific regulation [45].

\(^{10}\)Biologists will not fully agree with this simplified view, and as usual they are right with pointing out the importance of the details. First of all, the variation in the chromatin state is a result of many molecular alterations of the DNA itself (not changing the code, however), like DNA methylation, scaffolds of non-coding, long RNA, or post-translational modifications on histones. These modifications act as signals for a class of proteins, chromatin remodeling complexes, that can displace histones. These modifications on the physical state of the DNA (*cis*-modifications) induce the higher order chromatin structure. Also feedback loops in networks of transcription factors can exhibit a robust bistability, and hence lock a cell into a stable transcription state. This layer of genetic expression control is called trans epigenetic signaling [43].

\(^{11}\)Recently, more detailed studies have shown strong indications that these two chromatin classes consist out of different subgroups. For the fruitfly *Drosophila Melanogaster*, five different states of chromatin have been identified [44].
1.4. CELLS TIGHTLY CONTROL THEIR PROTEIN LEVELS

It is this level of regulation that causes global differences of the gene expression of differentiated cells that can be seen in figure 1.12. The chromatin state seems to be the most fundamental gene regulation level. Inserting reporter genes into the silenced heterochromatin significantly affects the expression level of these genes [44], even if they still have their endogenous promoter.

![Image of heterochromatin and euchromatin]

**Figure 1.13 – Different states of chromatin.** DNA in higher organisms is condensed; it is wrapped around histones. This unit of DNA and histone, the nucleosome, is organized in some higher structure. In the heterochromatin state, the nucleosomes are tightly packed [46], DNA in this section is transcriptionally silent. The euchromatin state, in contrast, does not inhibit gene expression; the genetic information in these sections is available.

Coming back to the language of life, it cannot be found in the DNA with its genes, or the cellular concentration of different proteins alone. It is more a combination of the three, the DNA sequence that contains the plan of life, proteins that bring this plan into action and in addition the conformation of the DNA molecule that implies which genes of the whole plan each individual cell is allowed to read.

Taken together, one of the most important and conserved layer of gene regulation is directly linked to the physical state of the DNA molecule. Accordingly, understanding the chemistry of DNA (the sequence) is not enough, it is of fundamental importance to extend our knowledge to the physical properties of DNA as well.
CHAPTER 1. INTRODUCTION

1.5 DNA as a physical object

The interest in the physical properties of DNA is not new. On the contrary, famous physicists like Erwin Schrödinger were immensely fascinated by this "most essential part of the living cell" [47]. In his essay "What Is Life?" (1944), even without the knowledge of the structure of DNA, he devotes a lot of his thoughts to DNA:

"Let me anticipate [...] that the most essential part of the living cell, the chromosome fiber - may suitably be called an aperiodic crystal. In physics we have dealt hitherto only with periodic crystals. To a humble physicist's mind, these are very interesting and complicated objects; they constitute one of the most fascinating and complex material structures by which inanimate nature puzzles his wits. Yet, compared with the aperiodic crystal, they are rather plain and dull. The difference in structure is of the same kind as that between an ordinary wallpaper in which the same pattern is repeated again and again in regular periodicity and a masterpiece of embroidery, say a Raphael tapestry, which shows no dull repetition, but an elaborate, coherent, meaningful design traced by the great master."

In spite of all this excitement of one of the most influential physicists of the 20th century, he later touches upon an interesting point:

"Therefore it is a small wonder that the organic chemist has already made large and important contributions to the problem of life, whereas the physicist has made next to none."

This mismatch can at least in part be attributed to the difficulty. It is technically very challenging to wring out the physical aspects from this small but most fundamental molecule. Single-molecule techniques such as optical tweezers and atomic force spectroscopy, however, advanced in such way that they now make it possible to isolate and manipulate individual DNA molecules, subject them to mechanical stress and monitor their response. The introduction into 'single-molecule techniques', especially optical tweezers and fluorescence microscopy will be given later, in chapter 2. Here I want so summarize what these studies have revealed about the nature of DNA.

1.5.1 DNA as a garden hose

DNA possesses a remarkable stiffness for its minute diameter (∼ 2 nm). Speaking in analogs, the physical nature of DNA more resembles a garden hose than a shoe lace. Like every macroscopic counterpart, when left alone, DNA adopts a random
1.5. DNA AS A PHYSICAL OBJECT

coil conformation. But in sharp contrast to macroscopic strings like power chords and telephone cables, the natural environment of DNA is of aqueous nature. Water looks different for DNA than it does for us, not smooth (like a continuum) and very fluid. Water molecules itself are not significantly smaller than the diameter of DNA [48], there is not even a factor of ten in the difference of their extension and these molecules are not stationary; they are in constant, random (so called thermal) motion. Therefore, DNA feels a bombardment of the surrounding water, which leads to a nonstop change of the coiled structure of DNA; DNA’s conformation undergoes thermally induced fluctuations. But DNA does not flimsily surrender to this thermal motion and become a tight knot; its stiffness counteracts the trend to coil up firmly. This is conceptualized in a parameter called the persistence length \( L_p \). Formally, the persistence length describes the distance along the polymer \( |s - s'| \) over which the correlation of the orientation of its tangent vectors \( \vec{t}(s) \) is lost due to the thermal fluctuations (See figure 1.14). The correlation in orientation will, on average, decay exponentially:

\[
\langle \vec{t}(s_1) \vec{t}(s_2) \rangle = e^{-\frac{|s_1 - s_2|}{L_p}}
\]

Figure 1.14 – DNA in the worm-like chain model. The formal description of the bending rigidity of DNA considers over which distance (along the contour, or the s-coordinate) the correlation in the orientation of the tangent vectors \( \vec{t}(s) \) is lost. The longer this distance, called persistence length \( L_p \), the more rigid is this polymer. For DNA, the persistence length is typically \( \sim 50 \) nm, or 150 base-pairs. Hence, this schematic is not on scale, DNA is more rigid.
CHAPTER 1. INTRODUCTION

Then, the bending energy and elastic energy stored in a DNA stretched to an extension $x$ is given by:

$$E = \frac{1}{2} k_B T \int_0^{L_c} L_p \left( \frac{\partial^2 \gamma}{\partial s^2} \right)^2 ds - F x + \frac{1}{2} \frac{S}{L_c} x^2$$

(1.2)

Here, the length of the DNA molecule, the contour length, is described by $L_c$ and the spring constant (DNA is an extensible molecule) is given by the fraction of the stretching modulus $S$ and the contour length $L_c$. To obtain an analytical expression for the relation between the extension $x$ of DNA and the applied force $F$, the expression above is harmonically approximated. For sufficiently extended molecules this results in the following compact expression [49]:

$$x = L_c \left( \frac{1}{2} \sqrt{\frac{k_B T}{F L_p}} + \frac{1}{S F} \right)$$

(1.3)

This analytic description is very powerful. Just by extending a DNA molecule, while monitoring the resisting force, it grants access to two physical properties of DNA, namely the persistence length, which describes the bending flexibility, and the stretch modulus, which parameterizes the resistance against extension of the DNA backbone.

Figure 1.15 shows the first reported force-extension curve (1996), which represented a milestone in that time, and set a new standard in the isolation and manipulation of individual molecules. These experiments also revealed that DNA in its double-helical form cannot withstand all forces. At a force of 65 pN, force-extension measurements disclosed that DNA undergoes a structural transition. The macromolecule gains about 70% of its contour length without significant increase in tension. This observation did not come as a full surprise. While the two strands of DNA itself are build up exclusively via covalent bonds, which are very strong, the three-dimensional structure of DNA, the double-helix, is sustained by the interplay of three rather weak interactions; namely the hydrogen bonding of complementary bases, stacking interactions of adjacent bases and electrostatic repulsion of the phosphate group of the DNA backbone. Therefore the initial excitement of scientist, looking at these results, is easy to comprehend. These measurements should have the power to reveal rich detail on the complex interplay of the energies that hold DNA together and give this molecule its fundamentally important double-helical conformation. Furthermore these investigations should shed some light on how plastic the double-helical structure is. How does this molecule react when subjected to tension, what will happen to its double-helical structure when DNA feels mechanical stress?
1.5. DNA AS A PHYSICAL OBJECT

Figure 1.15 – First report of the overstretching transition for DNA. In their pioneering paper of 1996, Steve Smith and Carlos Bustamante show the complexity of the mechanical properties that an individual DNA molecule discloses when extended [50]. It enabled these researchers to directly extract the persistence length of DNA. Furthermore these measurements reveal that at a force of about 65 pN, the DNA molecule does not behave like an ordinary spring, it undergoes a structural rearrangement, referred to as 'DNA overstretching transition'. The details of this transition, which were under a heavy debate since 1996, will be illuminated in depth in the rest of this thesis.

DNA, however, did not give away all its secrets easily. Since this complex mechanical behavior was reported in 1996, there has been a vehement debate in the field about the new structure that DNA adopts when subjected to high tension (65pN). Details on this debate will be picked up and discussed in chapter 2 to 6.

Focusing again on the biological context; how cells condense their DNA in the highly compact, wound-up chromatin and proteins locally disrupt the double-helical structure to gain access to the encode sequence, the fundamental questions that stimulated the research of this thesis are:

**What are the mechanical properties of DNA?**

Or more specifically:

**How flexible is DNA, how does it comply to tension, and up to what mechanical tension can DNA sustain its biologically important double-helical structure?**
CHAPTER 1. INTRODUCTION

1.6 Outline of this thesis

Chapter 2
The second chapter focuses on the single-molecule techniques optical tweezers and fluorescence microscopy. It introduces the principle, and gives technical considerations for the construction of a combined optical trapping, fluorescence microscopy and microfluidics instrument. Finally, details about some general experiments involving DNA and fluorescent ligands that bind to either double-stranded DNA or single-stranded DNA are presented.

Chapter 3
This chapter focuses on the very cooperative structural transition that occurs when DNA is stretched beyond its limits, the so called overstretching transition. Despite numerous efforts in the past decade, single-molecule assays lacked the power to unequivocally resolve its structure. The powerful single-molecule methods that have already been discussed in chapter 2 helped to obtain new information on this structural transition. The work of this chapter demonstrates that close to physiological solvent conditions, double-stranded DNA when stretched beyond its limits, melts apart cooperatively by DNA strand unpeeling from the ends. The situation changes drastically when the DNA attachment is such that the two strands cannot physically separate. Under this confined situation, when unwinding of the double-helix is inhibited, the force required to melt the double-helical DNA structure nearly doubles. Furthermore, the nucleation behavior of DNA melting changes dramatically. In such cases tension breaks the hydrogen-bond base pairing of the weakest structures: adenine and thymine-rich regions; similar to thermal melting.

Chapter 4
This chapter gives new quantitative insight on the DNA elastics, and provides a quantitative understanding of the observations from chapter 3. Guided by high-resolution force-extension measurements, I first present a new mathematical model, which is motivated by the helical structure of DNA, the twistable worm-like chain model. This model describes how DNA reacts to tension by stretching, twisting and extending. Furthermore this chapter covers an energetical analysis of DNA overstretching, and rationalizes why it is favorable for DNA to give up its double-helical structure at a tension of around 65 pN. This process turns out to be fully determined by the elasticity of single and double-stranded DNA as well as the hybridization energy landscape of the DNA sequence. Upon constraining DNA unpeeling to only one individual unpeeling
Chapter 5
In this study I investigate the impact of the chemical environment on the elastic properties of DNA, and in particular how the double-helical structure denatures when mechanically stressed under various ionic solvent concentration. It is well known that modulating the salt concentration has a strong impact on the stability of DNA, due to its strong negative charge. In addition to modulating the overall stability of this macromolecule, varying the ionic solvent strength has more fundamental consequences. I show that there are two fundamentally different modes of DNA overstretching which critically depend on the solvent conditions. At low to intermediate ionic condition, high force causes the free strand of DNA to peel off, while at higher salt concentration, the formation of internal melting bubbles dominates. Furthermore I investigate the role of the DNA sequence in the competition between DNA unpeeling and melting bubble formation. There I find that melting bubbles preferentially nucleate in AT rich regions, which are structurally weak compared to GC rich regions, that form a more and more challenging energetic barrier to the unpeeling process as the salt concentration is increased.

Chapter 6
The fluctuations during force-induced DNA melting are subject of this chapter. I demonstrate that the internal denaturation state of overstretched DNA fluctuates significantly when DNA is overstretched under conditions where force-induced melting bubbles emerge. First I determine the nucleation energy of such bubbles by measuring their statistics at various stages of the overstretching transition. Then I present the fluctuation spectrum of a microbead in an optical trap, tehered to an overstretched DNA. Using a mechanical model for a fluctuating DNA, I dissect this spectrum of DNA from the movement of the microbead. I find melting bubble boundaries exhibit diffusive kinetics on the timescale of $\sim 2-5\mu s$ with a root mean square amplitude of $\sim 2-3$ base pairs. Furthermore I found indications that, in contrast to thermal melting, AT and GC base pairs denature in the same proportion over wide ranges of the overstretching transition, which can explain why the overstretching force remains widely constant for this transition.

Chapter 7
Because of technical difficulties, optical tweezers studies have to a large extend been
restricted to double-stranded DNA molecules, although interesting cellular processes involved in DNA repair and replication initiate on single-stranded DNA. Here I employ force-induced DNA denaturation using custom designed double-stranded DNA constructs as a means to efficiently generate single-stranded DNA templates. I first present a protocol to append biotin-labels to both ends of only one DNA strand and then demonstrate that forces just exceeding the overstretching force of $\sim 65$ pN are sufficient to denature DNA. As a last point, the time-scale over which this DNA denaturation occurs once the DNA is strained is investigated. This systematic study shows that a denaturation time in the order of seconds is necessary to melt DNA, which is furthermore strongly influenced by the ionic strength of the solvent.
Combining optical tweezers and fluorescence microscopy

**Abstract** — The technically challenging field of single-molecule biophysics has established itself in the last decade by granting access to detailed information about the fate of individual biomolecules, unattainable in traditional biochemical assays. The appeal of single-molecules methods lies in the directness of the information obtained from individual biomolecules. Technological improvements in single-molecule methods have made it possible to combine optical tweezers, fluorescence microscopy and microfluidic flow systems. Such a combination of techniques has opened new possibilities to study complex biochemical reactions on the single-molecule level. This chapter provides general considerations for the development of a combined optical trapping, fluorescence microscopy and microfluidics instrument, along with methods to solve technical issues that are critical for designing successful experiments. Finally, several experiments are demonstrated to illustrate the power of this combination of techniques.
CHAPTER 2. COMBINING OPTICAL TWEEZERS AND FLUORESCENCE MICROSCOPY

2.1 Introduction

Single-molecule detection methods have made numerous important contributions to biology. Although the range of single-molecule tools is continuously expanding, they can be categorized into two dominant approaches: (i) force detection and manipulation, and (ii) fluorescence imaging and spectroscopy. In this article we will focus on the technical aspects of combining these two approaches for studies of DNA-protein interactions on the single-molecule level.

Single-molecule manipulation techniques, such as magnetic and optical tweezers, are commonly used to study the mechanical properties of biopolymers, in particular DNA, and their interactions with other biomolecules. These interactions lead in many cases to changes in the mechanical properties of DNA (length, flexibility, elasticity) [51–54], granting access to detailed information on, e.g., the kinetics and mechanochemistry of proteins like the ribosome, RNA polymerase, DNA polymerase and equilibrium binding properties of DNA intercalators. With fluorescence microscopy, on the other hand, the location and dynamics of fluorescently labeled biomolecules can be measured more directly. In addition, single-molecule fluorescence microscopy techniques can provide information on the dynamics of chemical reactions and local conformational changes of biomolecules like DNAs and proteins [55]. Single-molecule fluorescence microscopy can be classified into two broad categories [56]: (i) confocal fluorescence microscopy, where an image is generated by scanning a colocalized excitation and detection focal spot over the sample, while a pinhole in front of the detector blocks out-of-focus light, and (ii) wide-field fluorescence microscopy, where a region of the specimen is excited and observed with an array detector. A successful method to reduce the fluorescence background in wide-field fluorescence microscopy is total internal reflection fluorescence (TIRF) microscopy, where the fluorescence excitation is limited to about one hundred nanometers from a glass-water interface.

Force and fluorescence-based approaches can provide highly complementary information [57]. However, the technical challenges involved have limited the number of successful combinations of these single-molecule techniques. Previous studies have incorporated both approaches in several ways. In a first approach, DNA combing (stretching DNA by non-specific adherence to a surface) and surface-tethering (attaching DNA specifically to a surface by either one or two extremities) have been used to study DNA-protein interaction with TIRF microscopy. This combination has been used to study the diffusion of a variety of fluorescent proteins along the DNA [58], including the DNA repair proteins Msh2-Msh6 [59], p53 [60], and the DNA polymerase processivity factor PCNA [61]. In combination with flow stretching, the surface-tethered
DNA approach has been used to visualize the dynamics of DNA replication (for phage T7 and E. coli) at the single-molecule level [62, 63]. In two pioneering studies it has been demonstrated that a wide-field fluorescence method like TIRF can be successfully combined with optical trapping. [64, 65]. The challenge of studying the fluorophores close to the surface has been addressed, in both cases, by attaching the molecule of interest to the surface. Lang et al. studied the disruption dynamics of a primer attached to a template that is anchored to the surface by single-molecule FRET, while Ishijima et al. attached myosin to the surface, and used two optical traps and a special design of their flowcell to bring an actin filament into contact with myosin.

In general, however, it is advantageous to position the DNA far away from the surface to avoid interactions of DNA or proteins with the glass and to reduce background fluorescence [66]. One way to achieve this is to attach DNA to one or two optically trapped beads and to position it tens of micrometers away from the surface. Wide-field epi-illumination can then be used for fluorescence excitation. When only one bead is attached to the DNA, buffer flow is required to stretch the DNA. This approach has been pioneered by Kowalczykowski et al. to study RecBCD, a bacterial helicase/nuclease. The translocation of this enzyme was monitored by visualizing the removal of an intercalating dye from the DNA [67], or alternatively by attaching a streptavidin-coated fluorescent bead to RecBCD [68]. This method, however, requires a continuous buffer flow that can interfere with the dynamics of the protein-DNA interaction [69]. This limitation is not present when both ends of the DNA are attached to beads that are controlled with either two optical traps or with an optical trap and a micropipette [70, 71]. Recent examples of combining fluorescent techniques with optical tweezers are the study of individual quantum-dot labeled restriction enzymes (such as EcoRV) interacting with double-stranded (ds)DNA [72], and an investigation of the interaction and dynamics of fluorescently-labeled Rad51 with dsDNA [57].

Here we provide background and practical information on how to perform experiments combining single-molecule fluorescence, optical trapping, and microfluidics. In section 2.2 we discuss the background of the experimental approach and provide a detailed description of the setup we use. In section 2.3 we describe several approaches for the functionalization of DNA and proteins. In section 2.4 we provide several exemplary experimental protocols illustrating how the combination of these techniques allows visualization of DNA and of proteins binding to DNA.
CHAPTER 2. COMBINING OPTICAL TWEEZERS AND FLUORESCENCE MICROSCOPY

2.2 Instrumentation

In this section we focus on the general principle and technical considerations of combining optical trapping, fluorescence microscopy, and microfluidics. In addition, we describe in detail the experimental implementation of such a combined approach based on the instrument we have built in our laboratory (figure 2.1). This instrument is constructed around a commercial inverted microscope (Eclipse TE2000-U, Nikon), which provides the mechanical stability for key optical elements, like the microscope objective. Moreover, it allows flexibility in the choice of the dichroic mirrors. The microscope is equipped with a stage riser kit such that two dichroic mirror turrets can be operated independently.

Figure 2.1 – Schematic of the combined dual trap and fluorescence microscope with an integrated flow system. Each module (optical trapping, fluorescence microscopy, microfluidics) is described in detail in the designated section. Following abbreviations are used: DM - dichroic mirror, EM - emission filter, IR - Infrared shortpass filter, PSD - position sensitive detector, LED - light emitting diode, (EM-)CCD - (electron-multiplied) charged coupled device.
2.2.1 Optical Tweezers

Principles of optical trapping

![Optical Tweezers Diagram](image)

**Figure 2.2 – Qualitative picture of the origin of the trapping force.** The deflection of photons caused by a difference in the index of refraction transfers a net momentum on the microbead, which causes a restoring force back to the focus of the laser beam. Two photon paths, $\vec{p}_1$ and $\vec{p}_2$, are compared. The momentum balance $d\vec{p} = d\vec{p}_i - d\vec{p}_f$ for two beam paths is depicted below, and shows that for a microbead that is slightly off the optical axis, a net restoring force, caused by $dp_x$, back to the center of the trap is generated.

Optical trapping is based on the transfer of momentum from photons to a transparent object immersed in a medium with a different refractive index. High photon-flux sources like lasers can provide sufficient momentum transfer to manipulate micron-sized polystyrene or silica spheres, as demonstrated by Ashkin in 1970 [73, 74]. A stable optical trap requires a potential minimum, such that a small excursion results in a restoring force back to the center of the trap. In order to understand this idea, it is illustrative to separate the effect of the momentum transfer of photons into two components: the gradient force, which attracts the microbead towards the focus (figure 2.2), and the scattering force, which acts in the direction of light propagation. In most cases the scattering force will dominate over the gradient force. If, however, the laser beam is tightly focused, the intensity gradient around the laser focus becomes steep. In such a situation, close to the tight focus, the gradient force component surpasses the scattering force, which leads to a stable optical trap [75] (figure 2.2). Using micron-sized beads, optical traps with a spring constant on the order of 250 pN/µm
CHAPTER 2. COMBINING OPTICAL TWEEZERS AND FLUORESCENCE MICROSCOPY

can be generated with a tightly focused laser beam of around one Watt. In this configuration, forces of 100 pN and more can be applied to the trapped bead [76]. Such forces are sufficient to study unfolding of proteins [77], characterize the stall forces of molecular motors, or induce structural rearrangements in DNA.

Light sources for optical tweezers

To create stable optical traps for biological studies, good quality laser sources are required with respect to beam-pointing, wavelength, and power stability. A major cause of beam pointing instabilities is mechanical instability of the laser cavity due to thermal drift. Many modern high-power lasers have active temperature control to increase stability. Beam pointing instabilities and intensity fluctuations can also be caused by feedback into the laser cavity due to back-reflections. This can be avoided by inclusion of a Faraday Isolator in the optical path, immediately after the laser (figure 2.1) (IO-5- HP, Optics For Research). A third cause of beam pointing instabilities is fluctuation in air density, caused by air flow. Air flow can result in local refractive index fluctuations, altering the direction of the laser beam. A straightforward way to suppress this unwanted effect is to enclose the laser beam and part of the optics in tubes and boxes. The choice in wavelength is mainly governed by the inflicted photodamage on biological material, which is minimal in the near-infrared in a spectral window between the absorption of water and proteins (∼800-1200 nm) [78, 79]. For optical traps that can generate forces in the range of 100 pN, continuous wave (cw) laser powers in the order of one Watt are required, which can be provided by various laser types. Fiber lasers can produce very high powers (tens of Watts). In addition, the use of a doped single-mode fiber as the active medium/waveguide results in excellent laser beam quality. Another class of lasers are diode pumped solid state lasers, like Nd:YAG and Nd:YVO₄ lasers. These lasers have a compact design, resulting in good (beam pointing) stability and can be obtained with output powers up to ∼10 W. In our instrument we use a 3 Watt Nd:YVO₄ laser (Ventus 1,064 nm 3 W cw, Laser Quantum).

Microscope objectives

A stable optical trap requires a tight focus, since photons entering the focus under a high angle to the optical axis (figure 2.2) contribute the most to the restoring force. Such a tight focus can be generated with a high numerical aperture (NA) objective. The NA of an objective describes the range of angles over which it transmits (or
2.2. INSTRUMENTATION

accepts) light and is related to the maximal angle \( \Theta_{\text{max}} \) as:

\[
NA = n \sin(\Theta_{\text{max}}).
\]  

Here \( n \) is the index of refraction of the immersion medium between the sample and the objective lens. For an oil immersion objective, a typical value of the \( NA \) is 1.4, corresponding to a maximum angle of about 130°. Oil immersion objectives suffer, however, from spherical aberrations, resulting in a severely distorted focus when trapping further than 10 micrometers away from the glass surface \([80]\). Water immersion objectives suffer far less from spherical aberrations and are better suited for trapping deep in an aqueous solution than oil immersion objectives, despite their lower numerical aperture (Typical number for high numerical apperture objectives: \( NA = 1.2 \)).

Another important aspect of an objective is its transmission for the infrared light used for trapping. The transmission of some objectives is very low at wavelengths above 950 nm, but many transmit around 70% \([81]\). In our setup, we generate the optical traps using a 60X, \( NA = 1.2 \) water immersion objective (Plan Apo, Nikon).

Generation and steering of a dual trap

In our setup two optical traps are produced by splitting the 1,064 nm laser into two beams with identical power using a polarizing beam splitter cube (10BC16PC.9, Newport). We expand each beam with a 1 : 2.67 telescope system, which ensures that both laser beams overfill the back focal plane of the objective. The beams are recombined using another polarizing beam splitter cube and coupled into the objective via a dichroic mirror (‘DM2’: 950dcsp, Chroma Tech Corp.) (figure 2.1).

A dual trap design is capable of reaching very high stability and resolution, as was demonstrated, in both theoretical and experimental studies \([82, 83]\). The dual trap design is very stable if the differential pathway between the two beams is minimized, reducing differences in air density fluctuations \([84]\). The effect of beam pointing instabilities in such a design can be suppressed by measuring the difference coordinate of both trapped particles \([83]\). In such a dual trap design, care has to be taken that parasitic depolarization effects due to highly curved surfaces in the objective do not lead to interference and crosstalk between the two traps, which would complicate accurate position detection \([85]\). Polarization rectification and frequency shifting of one laser can be used as an effective method to circumvent this problem \([85]\).

In order to manipulate trapped microspheres, a precise and fast method to laterally displace the optical traps is required. This displacement can be achieved by introducing an angular deflection of the direction of the laser beam in the backfocal plane.
of the objective, which translates into a change in the lateral position in the focal plane of the objective (figure 2.3). In our system, using a 60X Nikon objective with focal length 3.33 mm, an angular variation of 1 mrad of one of the two trapping laser beams in the back-focal plane results in a relative displacement of the two traps of around 3 µm.

Several methods can be used to achieve this angular deflection: translation of a moveable lens in a telescope system [86], Acousto-Optic Deflectors (AOD) [87, 88], electro-optic deflectors [89], or scanning mirrors [83]. All of these methods have their strengths and limitations. In an AOD, a standing ultrasound wave in a crystal generates a modifiable diffraction grating, resulting in a maximal deflection of about 15 mrad (corresponding to a displacement of ∼ 50 µm in the image plane, for our instrument). The scanning speed of the diffracted laser beam is in principle only restricted to the speed of sound in the crystal. On the downside, the limited diffraction efficiency leads to light losses of around 60%, and nonlinear response of the AOD leads to angular crosstalk [89]. Electro-optic deflectors (EOD) offer a fast scanning speed, and suffer less from transmission losses and nonlinear response than AODs. However, their scan range is limited to only ∼ 1.5 mrad, corresponding to a displacement of ∼ 5 µm in the image plane. Scanning (piezo or galvano) mirrors offer no light loss and deliver an intermediate scanning range and speed. A final method is to translate a telescope lens. This method is relatively simple to implement, causes no light loss and can result in a large scan range. A disadvantage is that scanning speeds are limited to tens of micrometers per second. In our instrument, we have chosen translation of a telescope lens as beam steering method: The first telescope in one beam path can be displaced in two directions using two linear actuators (T-LA28, Zaber Technologies Inc.), allowing positioning of that trap (' moveable trap ') in the sample chamber as illustrated in figure 2.1.

Figure 2.3 – Steering of a laser trap using a telescope system with the focal lengths $f_1$ and $f_2$. Lateral movement of the first telescope lens introduces an angle $\alpha$ in the beam path. This angular deflection in the back focal plane (BFP) of the objective translates into the displacement "d" in the focal plane.
Laser detection and optical trap calibration

For high-resolution force and distance detection with optical tweezers, accurate position detection of the trapped particles is imperative. The easiest way to achieve this is using back focal plane interferometry \cite{90}. In this method, the interference signal of the laser beam with light scattered from the microsphere is measured using a quadrant diode or position-sensitive detector, onto which the back-focal plane of the condenser is imaged \cite{81}. These detectors have a high temporal bandwidth, which is important for detector calibration (see below). In our setup, we use a DL-100-7PCBA (Pacific Silicon Sensor Inc.) position sensitive detector to measure displacements of a microsphere in the stationary trap, while blocking the light of the moveable trap with a polarizer (03 FPI 003, CVI Melles Griot) (figure 2.1). We align the angular orientation of this polarizer by minimizing the total photocurrent of the position sensitive detector for the moveable trap, while the stationary trap is disabled. Note that non-uniform rotation of the polarization of the laser beam on highly curved surfaces like the objective lens prevents full attenuation. A systematic investigation of this parasitic depolarization can be found in \cite{85}.

The signal from the position sensitive detector is digitized using a 24 bit A/D converter (NI-PCI-4474, National Instruments). To obtain values for displacement of the bead from the center of the trap and the corresponding force acting on it, the detector needs to be calibrated, and the trap stiffness determined. This is achieved by measuring the effect of the optical trap on the Brownian motion of the microbead. The Brownian motion of a microbead in an optical trap can be described analytically and depends on two parameters: the hydrodynamic friction coefficient $\gamma$ and the trap stiffness $k$. The equation of motion can be expressed in a Langevin equation:

$$\gamma \frac{dx}{dt} + kx = F(t) \tag{2.2}$$

The external, random force $F(t)$, which generates the Brownian motion, can be derived by the fluctuation dissipation theorem, which leads to the following spectral characteristic $|\tilde{F}(f)| = 4\gamma k_B T$. This implies that the random thermal force has an average value of zero and a constant power spectrum $S_{\text{therm}}(f) = |\tilde{F}(f)|^2$ (white noise). Here, $|\tilde{F}(f)|$ denotes the Fourier transform of the random thermal force $F(t)$. The power spectrum of the movement of the bead can be determined by solving eq. 2.2 in Fourier space, which leads to:

$$S_X(f) = \frac{k_B T}{\gamma \pi^2 (f_c^2 + f^2)} \tag{2.3}$$
with the corner frequency $f_c = k/(2\pi\gamma)$. Equation 2.3 has the shape of a Lorentzian function. The corner frequency of this power spectrum directly yields the stiffness of the optical trap, while the amplitude of the power spectrum of the uncalibrated detector can be used to convert the voltage signal into an actual spacial displacement. For an uncalibrated detector, the measured powerspectrum $S_v(f)$ has the units $[V^2/Hz]$. This powerspectrum $S_v(f)$ relates to the calibrated quantity $S_X(f)$ as follows:

$$S_X(f) = R \cdot S_v(f)$$ \hspace{1cm} (2.4)

where $R [m^2/V^2]$ represents the detector response. Knowledge of the asymptotic behavior for very low frequencies of equation 2.3 ($S_X(f) = k_B T/(\gamma \pi^2 f_c^2) = S_v^0 = \text{const.}$ for $f << f_c$) and the hydrodynamic drag coefficient $\gamma = 3\pi \eta d$ with $d$ the diameter of the particle and $\eta$ the viscosity, directly yields the detector response $R$:

$$R = \frac{k_B T}{\gamma \pi^2 S_v^0 f_c^2}$$ \hspace{1cm} (2.5)

Figure 2.4 shows a measured powerspectrum for a 1.87 $\mu$m polysterene bead with a corner frequency of 3,350 Hz, which corresponds to a trap stiffness of 340 pN/$\mu$m.

![Figure 2.4 – Powerspectrum of a microbead in an optical trap. A polysterene bead of a diameter of 1.87 $\mu$m is held in an optical trap, while the position sensitive detector signal is sampled at a frequency of 20 kHz. The graph shows the power spectrum of the detector signal at a laser power of 3.3 W, yielding a corner frequency of 3,350 Hz, and a trap stiffness of 340 pN/$\mu$m, obtained from the Lorentzian fit, depicted in grey. The asymptotic value $S_v^0$ for low frequencies can be used to convert the voltage signal of the detector into an actual displacement.](image-url)
2.2. INSTRUMENTATION

Calibration methods that rely on measuring the Brownian motion of an optically trapped bead require a high detector bandwidth to capture the influence of the optical trap on this Brownian motion (typical bandwidth: $\sim 10$ Hz - $10$ kHz). Hence care has to be taken in the choice of the position detector for the force measurements, since some detectors based on silicon are intrinsically limited to frequencies around 5 kHz at 1,064 nm. This is caused by the increasing absorption of silicon outside the depletion region due to the increased transparency of silicon in the infrared region. Fully depleted silicon and GaAs detectors, however, do not suffer from this effect, and have a bandwidth of 100 kHz [79]. The position sensitive detector we use (DL-100-7PCBA, Pacific Silicon Sensor Inc.) has a bandwidth exceeding 10 kHz at 1,064 nm [81].

Distance determination

In addition to the position of the microbead relative to the optical trap, we also measure the distance between the two microbeads using a bright-field microscope image, which is recorded using blue LED illumination (LXHL-NB98 Luxeon Star/O, Lumileds). To avoid spectral overlap with a fluorescence signal, the LED is band-pass filtered ('EM2': d440/20, Chroma) and coupled into the sample chamber with dichroic mirror 'DM1' (520dcrx, Chroma; figure 2.1). A bright-field image is taken with an analogue CCD camera ('CCD': 902K, Watec), using the dichroic mirror 'DM4' (z488rdc, Chroma), and is digitized using a frame grabber (IMAQ PCI-1409, National Instruments). The distance between the two beads is determined using a Labview program (National Instruments), applying template-directed pattern matching. As a template, an image of the bead in the stationary trap, taken at the beginning of the experiments, is used.

2.2.2 Fluorescence microscopy

Single-molecule fluorescence microscopy and background reduction

In our instrument we employ wide-field fluorescence microscopy to detect single molecules. This method does not involve scanning, like confocal laser scanning microscopy, and is therefore relatively simple to implement. It also does not have time gaps between sequential acquisitions as would be associated with scanning. A design goal of our instrument was to avoid working close to the glass surface. Consequently, we refrain from excitation using TIRF, but use wide-field epi-fluorescence illumination instead (figure 2.5). One potential complication of wide-field epi-fluorescence
microscopy is the significant background caused by unbound fluorophores in solution. We tackle this problem using a laminar microfluidics flow chamber. Suppressing the fluorescent background from unbound proteins or dyes can be achieved by rapidly transferring the DNA molecule held between optically trapped beads to an adjacent channel without unbound fluorophores (see section 2.4.2).

The combination of optical trapping and wide-field epi-fluorescence microscopy is technically challenging. Detecting the fluorescence emission of a few fluorophores in the presence of a significant light flux for optical trapping depends on the careful separation of trapping and fluorescence emission light with an enormous difference in light intensity. Typically, trapping light powers are on the order of Watts (which leads to intensities in the focus of around 100 MW/cm²) while epi-fluorescence excitation uses powers in the range of mW (≈ 100 W/cm²). In contrast, a typical fluorophore yields a photon flux of about $10^4$ to $10^5$ photons per second, which corresponds to fluorescence emission powers of Femtowatts. To separate this tiny fluorescence signal from the high background of excitation and trapping lasers, high-quality, multilayer thin-film optical filters are used with sharp transmission/rejection edges (≈ 10 nm) and strong suppression of light outside the transmission window (OD > 5). In addition to a band-pass emission filter, a shortpass filter is used to further suppress the infrared background. Consequently, the attenuation for the infrared trapping light and the fluorescence excitation light is $OD_{ir} > 10$ and $OD_{ex} > 5$, respectively. When combined with the intrinsically low quantum yield of EMCCD cameras in the infrared
(<10%), this leads to an almost undetectable infrared background.

The epi-fluorescence microscope is integrated in our optical tweezers setup as depicted in figure 2.1. First, for epi-fluorescence illumination, the excitation laser beam is expanded for wide-field excitation. Dependent on the fluorophores used, one of the two following laser / dichroic mirror / emission filter combinations is employed: (i) For dyes like eGFP, YOYO-1, and Alexa 488, 473 nm excitation light is generated by a Blues laser (25 mW cw, Cobolt AB) and coupled into the objective by a z473rdc dichroic mirror from Chroma (’DM3’). In this case the fluorescence emission light is band-pass filtered through a hq540/80m-2p filter from Chroma (’EM1’). (ii) For dyes like Alexa 555, Cy3, and Atto 550, a 532 nm laser (GCL-025-L, 25 mW cw, Crysta-Laser, Reno, NV) is coupled into the objective by a z532rdc dichroic mirror (Chroma). For this excitation wavelength, we use an hq575/50m emission filter (Chroma). For both configurations, we used an additional filter to suppress the infrared trapping laser light (’IR1’: E750SP, Chroma). This infrared short-pass filter is compatible with near-IR dyes like Cy5, such that this setup can detect fluorescent emission in the red, using an appropriate combination of excitation laser and filters. The filtered fluorescence emission light is imaged on an ultrasensitive electron multiplying CCD camera (’EM-CCD’: Cascade 512B, Princeton Instruments).

In a typical experiment (0.5 s integration time, ∼ 1.5 W trapping laser power) the average count rate in our instrument increases from 705 ± 1 (standard error) to 708 ± 1 (standard error) when the trapping laser is switched on and no fluorescence excitation light is present. This increase is negligible compared to the count rates of the fluorescent background (total signal including background ∼ 740) and the signal of the fluorescently labeled protein bound to single stranded DNA (total signal including background ± 950) for a typical low-excitation-power/low-bleaching-rate measurement.

Photobleaching

An irreversible effect that limits the observation time of a fluorophore is photobleaching. In many cases, photobleaching is caused by photo-oxidation. Molecular oxygen, excited into the reactive singlet state can oxidize the excited fluorophore. Up to a certain extend this can be suppressed by an oxygen scavenging system. It has been demonstrated that photobleaching is substantially enhanced in the proximity of optical traps [65]. Dyes in their excited state, even with a short excited state lifetime of ∼ ns, can absorb an infrared photon due to the enormous photon flux close to the traps. In this higher excitation state the dye can oxidize [91]. This effect can be partly
CHAPTER 2. COMBINING OPTICAL TWEEZERS AND FLUORESCENCE MICROSCOPY

overcome by the use anti-oxidants and by the choice of suitable fluorophores [91]. Other approaches include spatial [92] or temporal separation of trapping light and fluorescence excitation [93]. Typically, fluorophores in aqueous solution photobleach after emitting $10^5$-$10^6$ photons [91, 94]. Consequently, one cannot continuously follow fluorescent dyes over timescales longer than several minutes. A way to circumvent this problem is to use time-lapsed data acquisition, i.e., fluorescence excitation and data acquisition are only initiated for a short time interval, $\tau_1$, with a longer waiting interval $\tau_2$, where the sample is not excited [95]. The illumination time window $\tau_1$ is governed by illumination conditions, while the choice of $\tau_2$ requires some a priori knowledge of the timescale of the molecular process to be investigated, and needs to be iteratively optimized. We, however, present methods to visualize proteins or dyes that non-specifically bind to DNA. In such a situation, fluorescent images rely on the emission of several fluorophores (typically $\sim 10^{-100}$). Consequently, then it is possible to observe the DNA molecule significantly longer than the bleaching time of the individual fluorescent dyes.

2.2.3 Microfluidics

Single-molecule studies of DNA-protein interactions require precise temporal and spatial control over the chemical environment of the observed DNA molecule. Microfluidic flow cells provide relatively simple and versatile means to this end. Various types of microfluidic flow cells differ by the number of channels (single or multi-stream flow cell), the material used to construct the flow cell, and the fluid delivery system. The design, building techniques and applications of various flow cells has been reviewed recently by Brewer and Bianco [96]. For our experiments, we use a microfluidic system with multiple laminar channels. In such a flow chamber, the different input channels converge into a larger common channel where the experiment takes place (figure 2.6). In laminar flow, the adjacent streams of the different fluids only slowly mix upon merging of the channels [96]. This feature allows rapid switching of buffer conditions by simply moving the microscope stage in a direction perpendicular to the flow such that the optically trapped molecule(s) move into a different buffer stream. Moreover, it permits exclusion of fluorescent proteins from the detection area. The flow channels in our system are manually cut in a sheet of Parafilm [71], serving as a spacer, and sandwiched between a cover slip (24 X 60 mm #1) and a microscope slide (50 X 75 X 1 mm) with drilled holes allowing fluid influx. Before assembly, the glass slides and cover slips
2.2. INSTRUMENTATION

Figure 2.6 – Schematic of the laminar multichannel-flowsystem. The disposable sample chamber consists of a channel pattern cut out of parafilm and melted between a glass microscope slide with holes (emphasized by dashed lines) for fluid influx and a thin cover slip (shown on the right). The assembled flow chamber is placed between a holder and a top plate and is connected to incoming and outgoing tubes by a sample inlet plate.

are extensively cleaned as follows: (i) acetone washing, (ii) 70% (v/v) ethanol/water washing, (iii) extensive rinsing with deionized water, and (iv) plasma cleaning for 10 min using the plasma cleaner from Harrick Plasma, operating with air of a pressure of about 0.4 mbar. The flowcell is generated by applying heat (∼ 130°C for 15 min) to the glass-parafilm-coverslip sandwich by placing a heat block on top of the flow-cell, which is situated in a metallic mold. To bring the different liquids into the flow cell, we use a pressurized buffer container, as depicted in figure 2.7. To decrease the amount of photobleaching, we keep the pressure chamber in a nitrogen atmosphere, therefore reducing the concentration of dissolved oxygen. We control the pressure in our container by two solenoid valves, where the first is connected to a pressurized nitrogen line, while the other reduces the overpressure by releasing nitrogen out of the container. Generation of flow rates in the order of about 10 µm/sec typically requires an overpressure of ∼ 100mbar in our pressure chamber (figure 2.7).
Figure 2.7 – Schematic of the flow system. Sample solutions are stored in containers. Solenoid valves control the nitrogen pressure in the pressure chamber. Switches enable quick initiation / stop of each laminar flow channel in the sample chamber. In case of undesired air bubbles getting stuck in the sample chamber, a negative pressure can be manually applied by a syringe, assisting in removing air bubbles in the sample chamber.

In the following we will discuss several complications that can be encountered using such a flow-system.

1. The presence of air bubbles in the flow cell can cause flow perturbation due to local compression and expansion of the air. Pressure spikes can be high enough to displace the microspheres out of the optical trap. Increasing the flow rate for a short time, or flushing the flow chamber with 70% (v/v) ethanol is usually sufficient to remove air bubbles. In case higher flow rates are required to flush out air bubbles, the flow rate can be further increased by using a syringe to create a negative pressure at the exit of the flow chamber (figure 2.7). In this way the overpressure applied to the flow chamber can be kept low enough to not disrupt the flow chamber. Usually we restrict the maximum overpressure in the pressured container storing the reagents to a maximum of ~ 250 mbar.

2. In general, it is imperative to work with a clean flow cell, tubing and solutions, as small dust particles, bacteria or other contaminations can enter the optical trap and disturb the experiments. Therefore, all solutions need to be filter sterilized (0.22 µm) and flow cells are replaced typically after 1 to 2 weeks of
2.3. Preparation of reagents

In this section we describe strategies and protocols for the generation of DNA molecules possessing end labels that enable the linking to optically trapped beads. Furthermore we give an overview over dyes used to fluorescently stain individual DNA molecules. Finally, we present general considerations and a protocol to fluorescently label proteins.

2.3.1 Terminally-labeled DNA molecules

Choosing the DNA template for a single-molecule experiment depends on various constraints such as the specific nucleotide sequence required for a biological process, the location of the labeling (3’- versus 5’-end labeling), the shape (DNA forks, overhangs, dsDNA, single-stranded (ss)DNA), and the size of the DNA molecule (i.e., number of base pairs). The latter defines the maximum distance between the trapped beads, which is an important parameter in single-molecule experiments. If the DNA molecule is too short, the two traps are in close proximity, which can result in non-negligible interference and crosstalk between them [85]. Moreover, for the combination of trapping and fluorescence detection of proteins on DNA, the high photon flux in the vicinity of the traps can damage biological samples and enhance photobleaching of the fluorophores [91]. DNA molecules used in experiments combining optical trapping and fluorescence are typically between 5 and 50 kilobases (kb) in length.

A dsDNA molecule widely used in single molecule experiments is bacteriophage lambda DNA, a linear dsDNA of 48,502 base pairs (bp) with two complementary 12-nucleotide
single-stranded 5’-overhangs. In order to allow specific binding of the DNA to the streptavidin-coated polystyrene beads (1.87 µm diameter, Spherotech), the extremities of the DNA molecules are labeled with biotinylated nucleotides (see protocol below).

Example protocol I: biotinylation of lambda DNA

Biotinylated nucleotides are incorporated into the 5’-single-stranded overhangs of lambda DNA by Klenow DNA Polymerase exo−:

1. Lambda DNA (4 nM) is incubated with dTTP (100 µM), dGTP (100 µM), biotin-14-dATP (80 µM), biotin-14-dCTP (80 µM) and Klenow DNA Polymerase exo− (0.05 units/µl) at 37 °C for 30 min.

2. The enzymatic reaction is stopped by heating to 70 °C for 15 min.

3. The DNA molecules are purified with a GFX PCR DNA and Gel Band Purification Kit (GE Healthcare).

This labeling reaction results in a DNA molecule containing multiple biotin-labels at the 5’-end of each strand.

2.3.2 Fluorescent labeling of DNA

Fluorescent intercalating dyes are frequently used to directly visualize dsDNA. A number of dimeric cyanine intercalating dyes, covering the entire visible spectrum, are commercially available, for example TOTO-1, YOYO-1, and POPO-1 (Molecular Probes). These dyes strongly bind to dsDNA and show a 100 to 1,000-fold enhancement of their fluorescence quantum yield upon intercalating between the base pairs of nucleic acids. As a result, the background fluorescence from unbound dyes is extremely low, making these dyes ideal for imaging of individual DNA molecules [97]. A protocol describing the labeling of a DNA molecule during a single-molecule experiment by the cyanine dye YOYO-1 is provided in section 2.4.1. It should be noted that the intercalation of such dye molecules changes the structural properties of the DNA molecule, making it longer and stiffer [98], therefore potentially interfering with some biochemical processes. As an alternative to such dyes, fluorophores can be covalently attached to the DNA molecule. For instance, the combined use of Cy5-labeled nucleotides and Cy3-labeled primers in a polymerase chain reaction (PCR) amplification allows introduction of both internal and end-specific fluorescent labels in the DNA molecule [99]. However, this method can be used only for short and medium
2.3. PREPARATION OF REAGENTS

length DNA molecule (≤ 20 kb) due to DNA length limitations of PCR. A broad and detailed description of fluorescent probes for nucleic acids can be found in [100].

2.3.3 Fluorescent labeling of proteins

Methods to fluorescently label proteins for optical tweezers experiments can be divided broadly into two categories: (i) the covalent binding of small organic fluorophores to the protein and (ii) the molecular tagging by fusing the gene of a fluorescent protein to the coding sequence of the protein of interest. For both techniques, the ideal fluorophore has high photostability, high fluorescence quantum yield, a large absorption cross section, and shows little intensity fluctuation [101, 102]. Most small organic fluorophores with a high photostability, used for covalent labeling of proteins belong to the cyanine and rhodamine families of fluorophores, along with derivatives of fluorescein with an enhanced photostability, like Alexa 488. Among the cyanine dyes, Cy3, Cy5, Alexa 555, and Alexa 647 are extensively used because of their high photostability (reviewed in [103]). In the rhodamine family of fluorophores, tetramethyl rhodamine (TMR), rhodamine 6G, Texas Red, and some of the Alexa dyes (Alexa 546) are commonly used. The main targets for chemical labeling are the primary amine groups found on lysine residues and at the N-terminus of proteins. However, even if the N-terminus of the protein is preferentially modified by performing the labeling reaction at near neutral pH, such labeling often results in a heterogeneous population of proteins containing a variable number of fluorophores if the target protein contains multiple lysine residues. Alternative targets for labeling are the sulfhydryls of cysteine residues, which are less abundant than primary amines (see example protocol below). In case multiple or no sulfhydryls are available on the surface of the protein of interest, cysteine residues may be removed from or added to a protein by point mutations. Hence, cysteine labeling allows for more specific labeling than amine-based labeling.

The use of genetically fused fluorescent proteins such as green fluorescent protein (GFP) ensures a high labeling specificity and avoids protein modification and labeling steps. In addition, a wide spectrum of autofluorescent proteins with high fluorescence quantum yield is now available (for a detailed comparison of fluorescent proteins see [104]). However, the fluorescent proteins commonly used are not very photostable and frequently exhibit blinking behavior. Moreover, they are relatively large (∼27 kDa for GFP). This may cause steric hindrance that can potentially interfere with the function of the target protein. It is therefore essential to test for enzymatic activity after labeling the protein of interest.
CHAPTER 2. COMBINING OPTICAL TWEEZERS AND FLUORESCENCE MICROSCOPY

2.4 Combining optical trapping, fluorescence microscopy and microfluidics

Here we describe a protocol to isolate an individual DNA molecule and successively visualize it using DNA-specific dyes. In addition, we show example experiments focusing on the specific binding of proteins to single-stranded and double-stranded DNA. Furthermore, we demonstrate the power of the combination of optical trapping, fluorescence microscopy, and microfluidics for the reduction of fluorescent background.

2.4.1 Example protocol II: Sequential isolation and visualization of a single DNA molecule using YOYO-1

In the beginning of an experiment, the different components (beads, DNA, buffer, YOYO-1) are each introduced into a separate channel of the multichannel flow cell and the flow is initiated. In the flow cell the different solutions do not mix because of the laminar flow conditions (figure 2.8).

![Figure 2.8 – Principle of the multiple-channel laminar flow chamber. I. Bead channel: two beads are optically trapped. II. DNA channel: a single DNA molecule is caught between the two beads. III. Buffer channel: a force-extension measurement of the DNA molecule is performed. IV. Protein channel: the DNA molecule is incubated with the studied protein or dye.](image)

To isolate and visualize a single DNA molecule the following steps are taken:

1. Trapping of two beads. This is achieved by turning on the trapping laser in the flow channel containing streptavidin-coated polystyrene beads (stock concentration: 0.5% (w/v)) diluted 1/1,000 in buffer (a typical buffer for experiments on DNA would be 10 mM Tris HCl pH 7.8 and 50 - 150 mM NaCl). Two beads are usually trapped within a few seconds and their size and shape are
visually checked to ensure that only one bead is present in each trap and that no contamination is present.

2. ’Catching’ a single DNA molecule. This is accomplished, in the presence of flow, by moving one bead back and forth around the position where the end of the flow-stretched DNA is anticipated. The presence of a DNA molecule is confirmed when, upon moving the beads with the trap, a correlated movement of the other bead is observed. In a picomolar DNA concentration the DNA-bead attachment usually occurs within a few seconds.

3. Force-extension analysis of the trapped DNA. After ’catching’ a DNA molecule, the beads are moved into a channel containing only buffer. Using force-extension analysis we ensure that only one DNA molecule connects the two beads (figure 2.9). When dealing with a single DNA molecule, the force-extension curve shows some very distinct features such as the cooperative structural transition at 65 pN, which yields a lengthening of the DNA up to a contour length of 1.7 its unstretched value without significant increase of force [50]. Attachment of multiple DNA molecules raises this transition to multiples of 65 pN.

4. In order to visualize the DNA, it is stained with YOYO-1. To this end, the DNA in the buffer channel is first extended to about 90% of its contour length to reduce spatial fluctuations of the DNA molecule and allow for sharp imaging. Next, the DNA is moved to the channel containing 20 nM YOYO-1, 10 mM Tris HCl, pH 7.8, and 50 mM NaCl. After 5 to 10 seconds incubation, the sample is illuminated with the 473 nm excitation laser and a snapshot is taken (Figure 2.10). The bright fluorescence signal of the dye and the low background allows for a visualization of the DNA molecule, even in a flow channel containing unbound dye in solution. It should be noted that intercalators like YOYO-1 induce photocleavage of DNA. Moderate staining ratios (∼1 dye every 500 bp) and low illumination conditions can reduce this effect. Under such conditions, we can continuously observe DNA on the timescale of a minute, before photocleavage breaks the DNA molecule.
CHAPTER 2. COMBINING OPTICAL TWEEZERS AND FLUORESCENCE MICROSCOPY

Figure 2.9 – Elastic behavior of an individual lambda DNA molecule. The gray trace represents the worm-like chain model for the elastic properties of DNA with a persistence length of 55 nm and a stretching modulus of 1350 pN. The distance is normalized to the contour length $L_c$ of the DNA molecule. At 65 pN the DNA molecule undergoes a structural transition, called overstretching transition, which will receive detailed attention in this thesis.

Figure 2.10 – Fluorescent image of the intercalating dye YOYO-1 bound to lambda DNA. The measurement was taken ten seconds after the DNA was transferred to a buffer containing 20 nM YOYO-1. A small force ($\sim 5$ pN) suppresses fluctuations during the data acquisition time (0.5 sec). A 488 nm laser excites the dye with a power of $\sim 3$ mW in wide-field excitation, over an area with a diameter of $\sim 35$ µm.

2.4.2 Example protocol III: binding of RPA to ssDNA

Here we present an example illustrating the visualization of a DNA-protein interaction: the binding of human Replication Protein A (RPA) to ssDNA. RPA is a heterotrimeric ssDNA binding protein that plays essential roles in many aspects of nucleic acid metabolism, including DNA replication and repair. RPA was fused here to the enhanced green fluorescent protein (eGFP) [105]. The experiment is conducted as follows:

1. Two beads are optically trapped and a single dsDNA molecule is caught between the beads (see steps 1 and 2, section 2.4.1).
2. After ‘catching’ a DNA molecule, the beads are moved into the channel containing only buffer. Next the dsDNA molecule is denatured into single-stranded DNA by putting it under $\sim 70$ pN tension (see chapter 7 for details).

3. The traps are moved into the flow channel containing the fluorescent protein (20 nM RPA). After 10 seconds incubation of the ssDNA molecule with the fluorescent protein, the sample is illuminated with the excitation light (488 nm) and a snapshot is taken (figure 2.11A).

4. Alternatively, the traps are moved back to the buffer-only channel and a snapshot is taken in a buffer without fluorescent background (figure 2.11B). The comparison of figure 2.11A and B clearly illustrates the advantage of a rapid exchange of buffer conditions to drastically reduce the fluorescent background.

Figure 2.11 – Visualization and background comparison of the fluorescent ssDNA binding protein RPA bound to ssDNA. The GFP labeled protein RPA, bound to ssDNA under a tension of 6 pN, was observed in [A] the laminar channel containing 20 nM fluorescently labeled RPA or in [B] a buffer containing no unbound fluorophores.
2.4.3 Example protocol IV: binding or RAD51 to dsDNA

Another example demonstrating the visualization of DNA-protein interactions is the binding of fluorescently labeled RAD51 to dsDNA. RAD51 is an essential component of the eukaryotic homologous recombination system, a crucial mechanism for the maintenance of genome integrity. Like other recombinases, RAD51 forms nucleoprotein filaments on both ssDNA and dsDNA. The protein was labeled with Alexa Fluor 555 as described in [106]. In this assay, the DNA was pre-incubated with the fluorescently labeled RAD51 in a buffer containing 1 mM ATP, 30 mM KCl, and 2 mM CaCl₂. The latter strongly reduces dissociation of the protein from the DNA. To obtain a patch-wise coverage of the DNA molecule with RAD51 filaments, the stoichiometry of RAD51 proteins versus base-pairs was adjusted. The experiment is carried out as follows:

1. Two beads are optically trapped, and a single dsDNA molecule, precoated with several RAD51 filaments, is caught between the beads (see steps 1 and 2, section 2.4.1)

2. The DNA is transferred to the buffer channel and illuminated with the excitation light (532 nm). Here, the fluorescence image and the force exerted on the DNA molecule are recorded simultaneously.

3. The distance between the two traps is then gradually increased, thus exerting a tension on the DNA molecule.

Figure 2.12 – Patch-wise formation of fluorescently RAD51 nucleoprotein filaments on dsDNA. One DNA molecule is suspended between two optically trapped microspheres, a second molecule is tethered from the lower bead. By increasing the distance between the traps, tension can be applied to the suspended DNA in a controlled manner (image from [106])
The fluorescence imaging shows an alternation of bare DNA and protein coated DNA regions (figure 2.12), thus providing information on the mechanism of filament formation. At the same time, the force-extension measurements give information on the elastic properties of the heterogeneously coated DNA molecules. In addition, it can be seen that at low tensions, the DNA molecule significantly fluctuates during the camera integration time, causing a blur in the fluorescence image of figure 2.12.

2.5 Conclusions

In this chapter we have presented a detailed description of our instrument that combines optical trapping, fluorescence microscopy and microfluidics. This instrument grants access to highly complementary information on DNA-protein interactions from optical tweezers and fluorescence microscopy. For example, the fluorescence intensity provides for a direct way to determine the amount of protein bound to a DNA molecule. Spatial information of protein binding on an individual DNA molecule allows for studies of specific binding or cooperatively in binding and unbinding. Moreover, the incorporation of multiple excitation lasers combined with multiband emission filters enables the use of multiple color fluorescence detection, which permits observation of co-localization of different proteins. Additionally, dynamical processes like diffusion, association/dissociation, and directed motion can be directly monitored. Taken together, the incorporation of optical tweezers with fluorescence microscopy and microfluidics combines the most powerful single-molecule manipulation and visualization techniques, and allows for high data throughput due to the ability to rapidly switch the chemical environment. This combination is widely applicable to the study of a vast number of different DNA-protein interactions.
Abstract — Single-molecule manipulation studies have revealed that double-stranded DNA undergoes a structural transition when subjected to tension. At forces that depend on the attachment geometry of the DNA (65 pN or 110 pN), it elongates $\sim 1.7$-fold and its elastic properties change dramatically. The nature of this overstretched DNA has been under vivid debate. In one model, the DNA cooperatively unwinds, while base-pairing remains intact. In a competing model, the hydrogen bonds between base pairs break and two single DNA strands are formed, comparable to thermal DNA melting. Here, we resolve the structural basis of DNA overstretching under physiological salt concentrations (50 mM-150 mM NaCl) using a combination of fluorescence microscopy, optical tweezers and microfluidics. In DNA molecules undergoing the transition, we visualize double- and single-stranded segments using specific fluorescent labels. Our data directly demonstrate that overstretching comprises a gradual conversion from double-stranded to single-stranded DNA, irrespective of the attachment geometry. We found that these conversions favorably initiate from nicks or free DNA ends. These discontinuities in the phosphodiester backbone serve as energetically favorable nucleation points for melting. When both DNA strands are intact and no nicks or free ends are present, the overstretching force increases from 65 pN to 110 pN and melting initiates throughout the molecule, comparable to thermal melting. These results provide new insights in the thermodynamics and stability of DNA.
CHAPTER 3. VISUALIZING THE STRUCTURE OF OVERSTRETCHED DNA

3.1 Introduction

The elastic properties of DNA affect a wide variety of cellular processes, such as protein-induced DNA bending, twisting or looping, but also genome compaction. Ever since the first single-molecule stretching experiments were performed on double-stranded DNA (dsDNA) [107], many efforts have been made to understand its elastic behavior. A striking feature of dsDNA elasticity is the overstretching transition: at forces of around 65 pN DNA gains about 70% of contour length while the force only increases slightly (figure 3.1A) [50, 108]. A similar transition is observed when the DNA is torsionally constrained [109, 110], then occurring at a force of around 110 pN (figure 3.1B). Two qualitatively different models have been put forward for the molecular mechanism of the overstretching transition. The first one assumes that it comprises a gradual conversion to a double-stranded conformation, structurally different from Watson and Crick’s B-DNA [17], named S-DNA [108]. Its structure is usually depicted as a partially unwound ladder with base pairing intact (figure 3.1C). The second model assumes force-induced melting, causing the hydrogen bonds between the two strands to gradually break, yielding single-stranded DNA (ssDNA) similar to thermal melting [111, 112]. This latter interpretation is supported by molecular dynamics simulations [113] and thermodynamic modeling [114], explaining dependencies on pH, salt, temperature [111, 112, 115, 116], and ssDNA-specific ligands [117–119]. However, four major arguments have been put forward against the force-induced melting interpretation. First, it has been observed that the two DNA strands do not immediately separate when pulled beyond the overstretching transition at 65 pN (figure 3.1A) [50]. Second, using similar DNA constructs, experiments at higher forces have revealed an additional transition at 150-300 pN, resulting in strand separation [120]. Therefore, it was concluded that no melting occurs in the first transition at 65 pN [120]. Third, the elastic properties of DNA beyond the 65 pN overstretching transition appear to differ from ssDNA [121]. Fourth, studies of torsionally constrained DNA suggested that S-DNA has a helicity of ~35 base pairs per turn [109,110]. Unambiguous discrimination between the two models would require direct visualization of the structure of DNA during overstretching.

3.2 Results

In order to visualize the structural changes of an individual DNA undergoing the overstretching transition, we used the combination of optical tweezers, fluorescence
3.2. RESULTS

Figure 3.1 – The overstretching transition of dsDNA under tension. [A]. Typical force-extension curve of a 3’-3’ attached DNA, with free 5’ ends (inset). The elastic properties of DNA below the overstretching force of 65 pN can be described by the extensible worm-like chain model (WLC, grey line). At 65 pN, the DNA molecule undergoes the overstretching transition, during which the intrinsic contour length of the DNA increases from 100% to about 170%. [B]. In a 3’5’-5’3’ attachment geometry, where all four strand ends are linked to the microspheres (inset), the overstretching force raises to 110 pN. The elastic properties of this DNA construct is well described by the extensible worm-like chain model (WLC, grey line) up to forces of around 40 pN. [C]. Schematic representation of two models explaining the nature of the overstretching transition. In the first model, the transition is due to gradual cooperative unwinding of the DNA double helix resulting in a base-paired structure, "S-DNA", resembling a parallel ladder. In the second model, force-induced melting of the two strands causes the transition. This force-induced DNA denaturation can occur via two mechanisms, force-induced bubble melting or strand unpeeling.
microscopy and micro-fluidics, as discussed in chapter 2. Using two optically trapped microspheres linked to the ends of a double-stranded DNA we hold and extend the molecule.

3.2.1 DNA overstretching with unconstrained extremities

Nucleation behavior of DNA overstretching

![Figure 3.2](image)

**Figure 3.2 – Initial YOYO binding has negligible effect on DNA length.** Kymograph of the slowly increasing fluorescence signal when exposing a partially overstretched, unlabeled 3'-3' attached dsDNA molecule to a YOYO-containing buffer. The schematic on the top explains how the kymograph is built up. The dashed lines depict the trapped microspheres, which are observed by their autofluorescence. Initially no YOYO-binding is observed. With time, two dsDNA segments show up by the YOYO staining. In addition to the two ends of the DNA, apparently also a single-stranded nick within the DNA molecule served as nucleation sites for the overstretching transition. Note that a minor lengthening of the dsDNA segments for increasing YOYO labeling is observed. For this reason, the ds-/ssDNA fraction estimation is performed on the first frame where YOYO fluorescence appears.

In our initial set of experiments focusing on the overstretching transition occurring at 65 pN, we use lambda DNA that is linked to the microspheres with one of the strands on either side of the DNA (designated 3'-3' attached DNA, figures 3.1A). In order to get insight into the nucleation behavior, we set out to specifically visualize these double-stranded DNA segments. Therefore we transfer the overstretched DNA to a channel in our micro-fluidic flow chamber containing YOYO, a dsDNA-specific dye [122]. To prevent YOYO from influencing the overstretching transition, we kept the incubation time to the minimum required for visualization, resulting in at most one dye per 500 base pairs. In figure 3.2 we show the dynamics of YOYO binding upon transferring the DNA into a channel containing YOYO, captured by a kymo-
3.2. RESULTS

graph. Initially no YOYO-binding is observed. With time, two dsDNA segments show up by the YOYO staining. Note that a minor lengthening of the dsDNA segments for increasing YOYO labeling is observed. For this reason, the ds-/ssDNA fraction estimation is performed on the first frame where YOYO fluorescence appears. Furthermore, each individual DNA was used for one measurement only and then discarded.

Figure 3.3 – Staining of 3’-3’ attached dsDNA stretches with the intercalating dye YOYO shows that the DNA overstretching transition at 65 pN is a nucleation-limited, first-order phase transition. [A]. Fluorescence images of different, extended dsDNA molecules exposed to YOYO show a binary subdivision of the DNA in labeled and unlabeled segments. The vertical dashed lines highlight the locations of the optically trapped beads. Image 1 is taken at a force of 50 pN ($L/L_c < 1$), where the DNA is labeled along its full length. At higher extensions ($L/L_c > 1$), (overstretching transition) at a force of 65 pN, discrete unlabeled segments appear at the expense of labeled, double-stranded segments. Image 3 shows an unlabeled segment halfway, suggesting that another nucleation took place at a nick. Image 6 is taken beyond the transition at a force of 80 pN, yet shows that the two strands are still connected by a short YOYO-labeled segment. [B]. The fraction of dsDNA plotted as a function of DNA extension. This fraction was obtained from the length of YOYO-labeled segments in fluorescence images such as in A, assuming them to have the same length as B-form dsDNA at this low degree of labelling. The grey dashed line connecting the two grey points indicates the behavior expected for a first order phase transition from 100% dsDNA at the start of the overstretching transition ($L/L_c = 1$) to no dsDNA at the end ($L/L_c = 1.7$).

Figure 3.3A shows fluorescence images of six different DNA molecules extended as indicated and then exposed to YOYO. At extensions $L$ up to the contour length $L_c$, the molecule is stained along its full length (image 1). In contrast, at extensions beyond the contour length (i.e., in the overstretching transition) only a discrete fraction of the molecule is stained (images 2-6). This result indicates that overstretching at 65 pN involves a gradual conversion of double-stranded B-DNA into a conformation to
which YOYO cannot bind. This behavior was observed for monovalent salt concentrations from 5 mM to 150 mM, at which similar images and force-extension curves were obtained. Interestingly, the regions at the extremities, close to the beads remained unstained. Hence, we conclude that for the 3’-3’ attached DNA construct, the ends of dsDNA are favorable nucleation points for the transition. In a few cases, an additional unlabeled segment between two labeled dsDNA segments was observed (e.g., image 3), which may imply overstretching initiated here from a nick, a break in one of the two DNA strands. Image 6 was taken beyond the overstretching transition at a force of 80 pN, and, strikingly, it still contained a short segment that got labeled. Apparently, some dsDNA is still present beyond the transition. We further analyzed the images by calculating the relative amount of base pairs that were still in a dsDNA conformation, for each extension (figure 3.3B). For comparison, the expected behavior of a first-order phase transition is shown: a linear decrease of the dsDNA fraction (dashed line) [123]. The data lie close to this line, implying that the YOYO-labeled fraction is a continuous double-stranded segment. If local overstretched patches, too small to be optically resolved, would be present, one would have expected an overestimation of the dsDNA fraction, which is not apparent from figure 3.3B. Taken together, our data reveal that the overstretching transition at 65 pN observed for the 3’-3’ attached DNA construct is processive and initiates from DNA ends.

**Single-stranded binding proteins reveal the presence of unpeeled DNA during overstretching**

Labeling overstretched DNA with YOYO alone does not allow discrimination between ssDNA and S-DNA. To test whether ssDNA is actually generated, we incubated overstretched DNA with Alexa-555-labeled human mitochondrial single-stranded DNA-binding proteins (mtSSB), which binds with high specificity to ssDNA [124]. First we confirmed that mtSSB does not affect the overstretching transition (figure 3.4D), similar to the related bacteriophage T4 gene 32 protein [117, 118, 125]. Next, we recorded fluorescence images of overstretched 3’-3’ attached DNA, incubated with mtSSB. Figure 3.4A shows that binding of mtSSB results in the appearance of two isolated fluorescent spots. We interpret that these spots are due mtSSB bound to relaxed, melted ssDNA, which has a radius of gyration on the order of the diffraction limit for the DNA lengths involved here [126]. When the DNA is extended further, the intensity of the fluorescent spots increases (figure 3.4A and B). We assign this increase to the binding of more and more mtSSB to relaxed ssDNA at the interface between dsDNA and ssDNA. From these observations for 3’-3’ attached DNA, two con-
3.2. RESULTS

![Figure 3.4](image)

Figure 3.4 – The DNA overstretching transition at 65 pN is a melting transition. [A]. Three consecutive fluorescence images of overstretched 3’-3’ attached DNA exposed to Alexa-555 labeled mitochondrial single-stranded binding protein (mtSSB). MtSSB accumulates in two spots (white arrows) that brighten and translocate upon further extension. [B]. The intensity in the moving spot scales linearly with the distance moved from the bead, showing that ssDNA accumulates during overstretching. Insert: fraction of base-paired nucleotides as a function of the relative extension, \( L/L_c \). The fraction is obtained from the distance between the mtSSB spots, assuming dsDNA in between. [C]. Interpretation of images of fluorescent mtSSB bound to overstretched DNA. Overstretching initiates on unpeeled DNA extremities near the beads. MtSSB binds to the melted DNA strand that is not connected to the bead and consequently not under tension. [D]. The elastic properties of the 3’-3’ attached DNA are not influenced by the presence of 50 nM mtSSB. Black squares: Force-extension measurement in a buffer containing no mtSSB. Grey dots: Force-extension measurement in 50 nM mtSSB, yielding identical force-extension behavior in the forward direction (upward arrow). A more pronounced hysteresis is observed (downward arrow), indicating hindered reannealing. Triangles show a third stretching-relaxing cycle, demonstrating that this DNA did not entirely reanneal to dsDNA, presumably because of mtSSB binding to ssDNA exposed by melting during the second extension. DNA overstretching still occurs at 65 pN, showing that mtSSB does not actively change the equilibrium of the transition.

Conclusions can be drawn right away. First, during the overstretching transition at 65 pN ssDNA is formed, which is incompatible with the S-DNA model for overstretching.
Second, mtSSB does not bind to the partially melted strand that is under tension, as depicted in figure 3.4C. We independently determined that mtSSB cannot bind to ssDNA under tensions exceeding ~ 40 pN, consistent with the wrapped binding mode of mtSSB [127]. We next calculated the relative amount of dsDNA from the distance between the spots in these images. The inset in Figure 3.4B shows that our data is consistent with the double-stranded fraction of DNA decreasing linearly with relative extension, analogous to the results obtained with YOYO labeling (figure 3.3).

**Double-stranded DNA preferentially unpeels from ends or nicks**

To find out how the results with YOYO and mtSSB relate, we performed two-color fluorescence measurements using both fluorescent markers. We partly overstretched 3’-3’ attached DNA in the presence of mtSSB (like in figure 3.4A) and subsequently exposed it to a buffer containing YOYO, while keeping the extension fixed. Figures 3.5A and B show two fluorescence images of two different DNA molecules, obtained by exciting Alexa555-mtSSB or YOYO, respectively. The mtSSB spots, indicating partly melted ssDNA, coincide with the edges of a YOYO-labeled dsDNA segment (white arrows in figure 3.5A and B). Figure 3.5B shows a molecule in which the overstretching transition nucleated not only at its extremities like in figure 3.4C, but also from nicks in the dsDNA. If the central, non-stained segment shown in figure 3.5B was nick-free, both single-stranded DNA strands would be under high tension and no ssDNA would accumulate at the transition interface. As a final confirmation, we performed two-color fluorescence experiments using a different set of probes: the bis-intercalating dye POPO-3 as dsDNA label and eGFP-tagged Replication Protein A (RPA) as ssDNA label. In contrast to mtSSB, RPA binds to ssDNA without wrapping [128]. Consequently, we inferred that it would bind to both relaxed and stretched ssDNA. In independent experiments we confirmed that RPA binds to ssDNA at forces up to at least 70 pN. In two-color fluorescence images (figure 3.5C) of partly overstretched POPO-3 and RPA-labeled 3’-3’ attached DNA, three regions can be discerned: (i) a dsDNA region where only POPO-3 binds is flanked by (ii) two bright spots of RPA bound to relaxed ssDNA strands, which are in turn flanked by (iii) regions of RPA-bound, stretched ssDNA connected to the beads. This third region, ssDNA under tension, was not observed in the mtSSB experiments (figures 3.5A, B). To confirm that the fluorescent spots at the ssDNA-dsDNA interface indeed correspond to relaxed ssDNA, we used our microfluidic device to apply a gentle flow perpendicular to the overstretched DNA. As can be seen in figure 3.5D, the fluorescent spot is stretched out along the flow direction, as expected for a relaxed piece of ssDNA. These experimental results
show that the overstretching transition occurring at 65 pN for 3'-3' attached DNA, is a force-induced melting transition that nucleates from free ends or nicks.

3.2.2 Overstretching torsionally constrained DNA

To study the molecular mechanism of the overstretching transition in the absence of such favorable nucleation points, we designed a DNA construct that was linked to the beads via both strands on either side (designated 3'5'-5'3' attached DNA, figure...
CHAPTER 3. VISUALIZING THE STRUCTURE OF OVERSTRETCHED DNA

3.1B). We found that the overstretching plateau for this construct occurs at substantially higher forces, \( \sim 110 \) pN (figure 3.1B), in agreement with previous studies using similarly attached DNA constructs [109]. Exposing 3’5’-5’3’ attached DNA to a buffer containing RPA did not result in binding of RPA at tensions below the overstretching force of 110 pN (figure 3.6A). Homogeneous RPA binding was however clearly observed at forces above 110 pN (figure 3.6A), indicating that ssDNA is generated.

Figure 3.6 – The DNA overstretching transition at 110 pN is a melting transition. [A]. Images of two different 3’5’-5’3’ attached DNA molecules labeled with the ssDNA marker eGFP-RPA, one held at a tension of 95 pN (before the onset of the OS transition) and the other at 135 pN (beyond the transition). At force below the transition, no eGFP-RPA binding is observed. Beyond the transition, eGFP-RPA covers the DNA completely. [B]. Images of two different 3’5’-5’3’ attached DNA molecules at a tension of 110 pN labeled sequentially with eGFP-RPA (top panels), and dsDNA-intercalator POPO (bottom panels). eGFP-RPA binds throughout the DNA molecules and no unique nucleation point can be observed.

In contrast to the 3’-3’ attached DNA (figure 3.5), where binding of RPA resulted in three clearly distinguishable regions on the DNA, spatial separation in distinct regions was far less clear at a tension of 110 pN, within the overstretching transition (figure 3.6B). For the 3’5’-5’3’ attached DNA construct, it appears that nucleation of force-induced melting occurs throughout the DNA molecule. This result was confirmed by subsequent exposure of RPA-labeled, overstretched DNA to POPO-3 to visualize dsDNA (figure 3.6B). In these experiments, regions of dsDNA were resolved throughout the DNA molecule. The intensity distribution of POPO-3 along the DNA molecule appeared to be anticorrelated with that of RPA, suggesting spatial separation of regions of ssDNA and dsDNA. However, this spatial separation cannot be totally resolved, due to the limited resolution of our instrument (\( \sim 250 \) nm, 750 bp). Taken together, these experimental results, using a combination of multi-color fluorescence imaging, opti-
3.3 Discussion

Our experiments have confirmed the dependence of the elastic behavior of DNA on the details of its terminal attachment [109]: the overstretching transition occurs at different forces depending on whether a nick or free end is available (3’-3’ attached DNA) or not (3’5’-5’3’ attached DNA). We probed the structure of DNA in the overstretching transition for ionic solvent concentrations of 50 mM - 150 mM NaCl by taking snapshots after applying specific single-stranded and double-stranded DNA probes. For both attachment geometries the experiments demonstrated the occurrence of ss-DNA, confirming that DNA gradually melts during the overstretching transition. In our experiments on the overstretching transition at 65 pN of 3’-3’ attached DNA, we observed that overstretching initiates at the extremities of the DNA. On each DNA end, only one of the strands of the double helix is attached to the microsphere, such that each strand is attached to only one bead. Apparently dsDNA extremities form a favorable nucleation point, or in other words, the overstretching transition at 65 pN is cooperative force-induced ‘unpeeling’ [23]. In some cases, we observed additional nucleation within a DNA molecule (figure 3.3A, image 3 and figure 3.5B). Our experiments combining YOYO and mtSSB staining (figure 3.5B) revealed that this is caused by single-stranded nicks. The remarkable processivity of overstretching we observed along the heterogeneous sequence of thousands of base pairs has not been anticipated in earlier studies, where overstretching was suggested to initiate in small AT-rich domains throughout the DNA molecule [112, 115, 129]. Our observations rule out the large-scale occurrence of such interior single-stranded melting ‘bubbles’ in the ionic solvent concentration range of 50mM-150mM NaCl, when a free end is available for nucleation \(^1\). Under these conditions the energy penalty for a new nucleation bubble is larger than a gradual progress of the melting front from the ends or from a nick, regardless of the sequence. This can be understood by realizing that melting at these locations allows one of the newly formed ssDNA strands to relax and release elastic energy (for more details, see chapter 4).

The situation is quite different for 3’5’-5’3’ attached DNA, which lacks energetically preferred nucleation points on DNA ends. For this DNA, the nucleation of force-

\(^1\)The overstretching behaviour at higher salt will be studied in chapter 5.
induced melting more closely resembles thermal melting (figure 3.6B), which was found to initiate predominantly from small, interior (AT-rich) regions [130]. Indeed, we show that RPA and POPO-3 bind throughout the overstretched DNA (figure 3.6B). In addition, we measured that the 3’5’-5’3’ attached DNA overstretches at forces of 110 pN, in agreement with previous studies on torsionally and end-constrained DNA [109]. Notably, when the DNA is fully overstretched (figure 3.6A, right panel), RPA binds homogeneously along the DNA, in contrast to earlier predictions that the structure of fully overstretched torsionally constrained DNA is biphasic in order to conserve the linking number [109]. The most straightforward explanation of homogeneous RPA binding is that the new structure, generated during overstretching at 110 pN, consists of two single DNA strands lacking hydrogen bonds between the bases, wrapped around each other with a linking number close to that of relaxed ds-DNA. In line with this hypothesis, one could speculate that the DNA structure with 1 turn per ∼35 bases that is obtained when underwound torsionally constrained DNA was stretched to ∼50 pN [109] is also composed of two melted and wrapped ssDNA strands. To resolve how these experiments relate to our findings, further studies that combine fluorescence imaging, force control and torsional control of the DNA will be required.

As mentioned in the introduction, an argument against the interpretation of force-induced melting for 3’-3’ attached DNA molecules has been the reported stability of DNA beyond the transition at 65 pN. Our approach of direct visualization yielded a surprising observation that allows us to straightforwardly explain this paradox. As seen in image 6 of figure 3.3A, at forces beyond the overstretching transition a small, transiently stable YOYO-labeled and thus double-stranded segment remains. Apparently, it takes more time to overcome the higher energy barriers to melt the last hundreds of base pairs and the transition is no longer in equilibrium on experimental time scales [120]. This can explain why the DNA molecule can at least transiently remain intact at forces above 65 pN. A similar conclusion was drawn from thermodynamic modeling [114], which indicated that the DNA is held together by several small ds-DNA regions. Interestingly, however, we find that only a single non-melted region keeps the DNA strands connected. This region appears to be consistently positioned asymmetrically along the molecule (rather than in the center), which might suggest that the GC-rich half of lambda DNA is the last part to melt, as expected. In accordance with a small dsDNA stretch remaining, the elastic properties of DNA beyond the overstretching transition can be described by a linear combination of ssDNA and a single, slowly vanishing fraction of dsDNA, indicating out-of-equilibrium behavior.
3.4 METHODS

Figure 3.7 – Elastic response of DNA beyond the overstretching transition, when pulled on from either side of the same strand. Black triangles represent force extension measurements at an extension rate of 2 \( \mu \text{m/sec} \). Grey circles display a consecutive measurement on the same DNA, showing how it denatured entirely during the first extension. At 150 mM NaCl (panel [A]) the elastic response of DNA beyond the transition differs slightly from fully single-stranded DNA (grey circles), as can be more clearly seen in the insert. The black solid line represents the elastic response of a dsDNA-ssDNA hybrid, with a small fraction of 2% dsDNA that remains transiently stable up to forces between 110-150 pN. At 50 mM NaCl (panel [B]) this dsDNA segment is smaller (1%) and less stable, typically disappearing at forces below 100 pN.

In conclusion, we have unveiled that, independent of the details of strand attachment, DNA overstretching unambiguously comprises a gradual conversion of dsDNA to ss-DNA. We directly visualized dsDNA using intercalating dyes and ssDNA using single-stranded binding proteins, concurrent with force-extension measurements. Our interpretation therefore does not rely on thermodynamic, mechano-chemical or other models and assumptions. We anticipate that our results will provide a new basis for the fundamental understanding of the thermodynamics of DNA and DNA-protein interactions.

3.4 Methods

The combined fluorescence and dual-beam optical trapping instrument was slightly adapted from the one described in chapter 2, to enable dual-color fluorescence imaging. For fluorescence excitation of either YOYO or Alexa-555, the linearly polarized

\(^2\)See chapter 5 for some additional considerations.
light of the 473-nm laser or 532-nm laser, respectively, was coupled into the microscope using a triple band dichroic mirror (z488/532/633ndc (triple-band), Chroma). Fluorescence was band-pass filtered (z488/532/633m (triple-band), Chroma) and imaged onto the electron multiplying CCD camera. Camera readout was externally triggered using a TTL signal for time-lapsed data acquisition.

To generate the DNA construct with biotin labels on one of the strands on either side of the DNA (3’-3’ attached DNA), we used the protocol to terminally label lambda DNA, as described in chapter 2. The DNA construct that possesses biotin-labels at both strands on either end (3’5’-5’3’ attached DNA) was generated as follows. First a 3’ overhang was introduced into the 8,393 bp plasmid pKYB1. Biotin-labeled oligos (29 nt) were ligated to both ends, yielding a DNA construct with biotinylated 5’ overhangs on both ends. Both 3’ ends where biotin-labeled by incorporation of biotinylated dATP’s (NEB) along with unlabeled nucleotides, by Klenow exo- DNA polymerase. YOYO-1 and POPO-3 were purchased from Molecular Probes. For fluorescent labeling of human mitochondrial single-stranded DNA-binding protein, a cysteine residue was introduced at the C-terminal end by PCR. The PCR fragment was cloned into pBacPAK9 and used to generate recombinant Autographa californica nuclear polyhedrosis virus (BacPAK, Clontech). The protein was expressed in Spodoptera frugiperda (Sf9) cells and purified as described in [131], with an additional purification step (prior to the hydroxyapatite column) using a 1-ml HiTrap SP column (GE Healthcare). The purified mtSSB cysteine variant was stored in 10 mM KPO4 pH 7.2, 0.1 M NaCl and 10% glycerol, and labeled with maleimide Alexa-555 dye (Molecular Probes). Unreacted dye was removed from the sample with size-exclusion spin-columns (Sephadex G-25, GE Healthcare).

To obtain fluorescent RPA, a DNA fragment encoding a polyhistidine-tagged variant of the enhanced green fluorescent protein (eGFP) was inserted in a frame at the 3’ end of the cDNA encoding the large subunit of human RPA in the expression plasmid p11d-tRPA [132]. Fluorescent hRPA-eGFP was produced in E. coli and purified by chromatography through a Hitrap FF column followed by chromatography through Hitrap SP HP and Hitrap Q HP columns (GE Healthcare Life Science). The protein, in 200 mM KCl, 20 mM Tris pH 7.5, 1 mM DTT, 0.5 mM EDTA and 10% glycerol, was snap frozen in liquid nitrogen and stored at -80°C. The eGFP tagged hRPA bound X174 ssDNA with the same affinity as untagged hRPA.

DNA molecules were captured between two optically trapped beads (1.87 µm streptavidin coated polystyrene beads, Spherotech) using the multi-channel laminar flow cell and stretched by increasing the distance between the optical traps. Fluorescence and force-extension data were recorded in a synchronized manner. For experiments
involving YOYO labeling, we first partially overstretched DNA in a buffer without YOYO, before exposing it to a YOYO-containing buffer by a rapid and complete buffer exchange (∼0.5 sec). This ensured that the progress of the phase transition itself cannot be influenced. During this buffer exchange, fluorescence was recorded. Every partly overstretched DNA molecule was exposed to a 10 mM Tris buffer (pH 7.4-7.8) containing 10-50 nM YOYO and either 5, 50, or 150 mM NaCl, while fluorescence was recorded. From the small initial tension drop caused by YOYO-induced lengthening [133] we estimate that in this case the dye : base pair ratio does not exceed 1 : 500. DNA molecules were always discarded after being exposed to YOYO. All stretching experiments were performed in a 10 mM Tris buffer (pH 7.4-7.8) containing 50 mM NaCl. For experiments with Alexa-555 labeled mtSSB, a protein concentration of 50 nM was used. Fluorescence snapshots were taken in a flow channel without mtSSB to avoid background fluorescence caused by unbound protein. For two-color experiments, mtSSB-labeled DNA molecules were briefly exposed to a 10 nM YOYO buffer. Alternatively, we used eGFP-labeled RPA (20 nM) to label ssDNA segments of an overstretched DNA molecule (both, the 3'-3' attached DNA and the 3'-5',5'-3' attached DNA), followed by exposure to a 50-70 nM POPO solution to label dsDNA segments. In order to extend unpeeled ssDNA segments for the 3'-3' attached DNA, we then imposed a lateral flow to the DNA.
Twist, stretch and melt: quantifying how DNA complies to tension

Abstract — Central biological processes involve continuous mechanical manipulation of DNA. Substantial experimental and theoretical effort has been devoted to elucidating the internal changes of DNA subjected to mechanical stress. Still lacking, however, is a complete mathematical description of DNA elasticity over the whole range of forces at which double stranded DNA remains hybridized. Here we present new force-extension measurements with unprecedented detail, which allow us to propose a new model for DNA extension, the twistable worm-like chain, in which DNA is considered a helical, elastic entity that complies to tension by extending and twisting. In addition, we reveal hitherto unnoticed stick-slip dynamics during DNA overstretching at ~ 65 pN, caused by the loss of base-pairing interactions. An equilibrium thermodynamic model solely based on DNA-sequence and elasticity is presented, which captures the full complexity of this transition. These results offer a deeper quantitative insight in the physical properties of DNA and present a new standard description of DNA-mechanics.
4.1 Introduction

In cells, double-stranded (ds)DNA is constantly being twisted, bended and stretched by numerous proteins mediating genome compaction, gene regulation, expression and DNA repair [52, 82, 134, 135]. Consequently, for an in-depth understanding of DNA manipulation and maintenance, it is imperative to understand the elastic properties of dsDNA in great detail. Over the last two decades, much effort has been devoted to obtain a complete picture of DNA elasticity [50, 108, 136, 137]. The DNA extensibility and bending rigidity in a force range of 0 to \( \sim 35 \) pN have been successfully described by the extensible worm-like chain (WLC) model [49, 136]. In this model, DNA is regarded a homogeneous, extensible rod. Important aspects of the dsDNA double-helical structure, such as twist and base-pairing interactions, are not taken into account, which might be the reason for the WLC-model's failure to capture DNA elasticity at forces above 35 pN. Recently it has been demonstrated experimentally that DNA extension cannot be considered separately from twist, since these two elastic parameters are intricately coupled: DNA overwinds when extended by forces up to \( \sim 35 \) pN, while it unwinds at higher forces [138]. The functional form of the force dependence of the twist-stretch coupling is however still unknown. Moreover, it is unclear how force-induced over- or unwinding affects the elasticity of DNA. Another striking feature of DNA elasticity is the overstretching transition, occurring at 65 pN, resulting in a 70% increase of DNA length [50, 108]. In chapter 3 we have directly demonstrated that this transition is due to force-induced separation of the two DNA strands [105]. Since the DNA sequence has direct consequences for the interaction energy of the two hybridized DNA strands, a strong sequence dependence on the force-induced DNA unpeeling process is expected. So far, however, a clear effect of local sequence on the dynamics and energetics of force-induced DNA melting has remained elusive. In the current study, we set out to measure the elastic properties of DNA with unprecedented resolution in order to provide a complete mathematical model of DNA elasticity including the overstretching transition, taking into account the sequence composition and the force-dependent twist-stretch coupling.

4.2 Results

Quantification of the mechanical behavior of DNA in a well-controlled and reproducible fashion, requires a properly designed DNA template. As is discussed in chapter 3 [105], force-induced strand unpeeling initiates at nicks and free DNA ends at
4.2. RESULTS

Figure 4.1 – Force-induced strand unpeeling of DNA studied with fluorescence microscopy and force-spectroscopy [A]. Schematic depictions of the DNA construct used in this study undergoing unpeeling upon extension. Three ends of the DNA strands are connected to the optically trapped microspheres. [B]. Three fluorescence images of partially overstretched DNA molecules. After partial overstretching, the DNA molecules are exposed to 20 nM fluorescently labeled single-stranded binding protein eGFP-RPA in a buffer containing up to 150 mM NaCl (50 mM shown here). Only the left part of the DNA construct is fluorescent, indicating unpeeling has only occurred from the left side of the DNA. [C]. Force-extension measurement of a DNA construct like shown in panel A. In region I, a strong deviation from the extensible WLC (grey line) is evident. In region II, the overstretching transition, pronounced step-wise increases in contour length can be discerned.

65 pN. To address the sequence dependence of this process, we generated a DNA-construct without nicks and only one free DNA end, such that unpeeling can initiate exclusively at one location, and proceed in only one direction. This was achieved by attaching biotin tags to both strands on one end, and to only one of the strands on
the other end (see section 4.4) (figure 4.1A). The DNA-construct was then connected to two optically trapped, streptavidin-coated microspheres. To verify that unpeeling initiates only from one end, we stretched the DNA to about 65 pN and monitored the generation of single-stranded (ss)DNA with fluorescence microscopy using eGFP-tagged Replication Protein A (RPA) as a ssDNA reporter, (see chapter 3) [105]. For all DNA molecules tested (N=20) overstretching resulted in the formation of ssDNA in a single deterministic opening fork (figure 4.1B). We next recorded high-resolution force-extension curves of such DNA constructs. In figure 4.1C two features immediately catch the attention: (i) the DNA stretching behavior deviates substantially from the extensible WLC model at forces above ∼35 pN (figure 4.1C, region I), and (ii) at higher forces, above 60 pN, overstretching sets in, not as a smooth transition, but with a distinct pattern of saw-teeth (figure 4.1C region II). Below we address these two regimes in detail.

4.2.1 Unwinding and overwinding of DNA under tension

In regime I, the extensible WLC model fails to accurately predict the dsDNA extension at forces above 35 pN. Apparently, it lacks important aspects of DNA elasticity that become relevant at intermediate forces. We hypothesized that this could be the mechanical coupling between extension and twist of the double helix, which has been predicted [139,140] and experimentally demonstrated [138]. This coupling has been shown to depend on force in a complex way: at low forces DNA overwinds when stretched, at forces above 35 pN it underwinds [138]. Consequently, when the DNA is torsionally unconstrained, as is the case here, it will comply to an external force in two ways: by stretching, and by overwinding or unwinding. Only stretching is included in the extensible WLC model. Here, we extend this model by incorporating an enthalpic term that captures the elongation of DNA due to changes in twist.

Initially, we focus on the force regime where entropic bending fluctuations can be neglected (F > 15 pN, enthalpic regime). Then, the elastic energy stored in a DNA molecule ($E_{DNA}$) stretched by a force ($F$) beyond its contour length ($L_c$) can be written as [138–140]:

$$E_{DNA} = \frac{1}{2} C \frac{\theta^2}{L_c} + g(F) \frac{x}{L_c} + \frac{1}{2} S \frac{x^2}{L_c} - xF.$$

Here, the material parameters are represented by: $L_c$ the contour length, $C$ the twist rigidity, $g(F)$ the tension-dependent twist-stretch coupling, $S$ the stretch modulus [50,110,136]. The end-to-end distance of the molecule is given by $x$, and $\theta$
4.2. RESULTS

Figure 4.2 – DNA unwinding under tension quantified in the twistable worm-like chain (tWLC) model. [A]. Force dependence of the twist-stretch coupling, extracted from a typical force-extension measurement. Open circles and error bars represent mean and standard deviation obtained from binning 20 data points. The grey line represents a linear fit \( g(F) = g_0 + g_1 F \) with \( g_0 = -670 \pm 10 \) pN nm and \( g_1 = 18.4 \pm 0.2 \) nm; errors represent the standard errors of the fit). [B]. Comparison of the extensible WLC model with the tWLC model. Black dots: force-extension measurement. Light grey curve: fit of the tWLC in the force-interval of 1.5 pN to 62 pN, yielding \( S = 1490 \) pN, \( L_p = 33 \) nm, \( g_0 = -670 \) pN nm, and \( g_1 = 18.4 \) nm. (Ensemble average (N=7) and standard deviation: \( S = 1480 \pm 160 \) pN, \( L_p = 36 \pm 5 \) nm, \( g_0 = -600 \pm 40 \) pN nm, and \( g_1 = 15 \pm 2 \) nm). Dark grey curve: fit of the extensible WLC model in the force interval of 1.5 pN to 45 pN yielding \( S = 1750 \) pN, and \( L_p = 33 \) nm (Ensemble average (N=7) and standard deviation: \( S = 1510 \pm 40 \) pN, \( L_p = 44 \pm 10 \) nm). [C]. Change in twist of a 8.4 kb DNA construct as a function of tension. Black: measured data by Gore et al. [138]. Light grey line: force-induced changes in twist calculated from the fit in panel A, using equation 4.3.

denotes the winding angle relative to the Watson-Crick structure. Minimizing the enthalpic Hamiltonian (4.1) with respect to \( x \) and \( \theta \) yields:

\[
x_{enthalpic} = \left( \frac{C \cdot L_c}{-g(F)^2 + SC} \right) \cdot F, \quad (4.2)
\]
Since the worm-like chain model does not consider crosstalk between the enthalpic and entropic contributions, both can be addressed separately. Incorporation of equation 4.2 in the worm-like chain model yields the following relation between the extension and the force acting on the DNA, the twistable worm-like chain (tWLC) model:

\[ x = L_c \cdot \left( 1 - \sqrt{\frac{k_B T}{L_p \cdot F} + \frac{C}{-g(F)^2 + SC}} \cdot F \right) \quad (4.4) \]

In this equation, the properties of DNA are represented by five parameters: the contour length \( L_c \), the twist rigidity \( C \), the stretching modulus \( S \), the persistence length \( L_p \) \([50, 110, 136]\), and the twist-stretch coupling \( g(F) \) \([138, 141, 142]\). To obtain insight into the functional form of the force-dependence of the twist-stretch coupling, we solved equation 4.4 for \( g(F)^2 \):

\[ g(F)^2 = SC - CF \left( \frac{x}{L_c} - 1 + \sqrt{\frac{k_B T}{L_p \cdot F}} \right)^{-1} \quad (4.5) \]

Note that this treatment yields the absolute value of \( g(F) \). Information about the sign of \( g(F) \) comes from magnetic tweezer studies, from which it is known that \( g(F) \) is negative at low forces and changes sign at \( \sim 35 \) pN \([138]\). Using this equation, the force dependence of the twist-stretch coupling is directly obtained from the experimental force-extension curves (figure 4.2), taking values for the contour length \( L_c \) and twist rigidity \( C \) known from literature \([110]\), and values for the persistence length \( L_p \) and stretching modulus \( S \) obtained from an extensible WLC fit up to 35 pN (figure 4.2).

Figure 4.2A shows that the twist-stretch coupling depends linearly on force: \( g(F) = g_0 + g_1 \cdot F \). From linear fits to multiple traces (\( N=7 \)), we obtain average values of \( g_0 = -600 \pm 40 \) pN nm and \( g_1 = 15 \pm 2 \) nm (Error represents the standard deviation between different measurements). These values for \( g_0 \) and \( g_1 \) result in sign reversal of the twist-stretch coupling at \( \sim 35 \) pN. Incorporation of this functional form of \( g(F) \) into the tWLC-model results in a substantially improved description of DNA elasticity, up to the onset of the overstretching transition (figure 4.2B).

As a final consistency check, we evaluated whether this functional form of \( g(F) \) correctly captures dsDNA force-induced overwinding and unwinding behavior (equation 4.3), which has been measured directly before \([138]\). As shown in figure 4.2C,
our model of $g(F)$ agrees well with the data from Gore and coworkers [138], especially for forces above 25 pN. At forces below $\sim 20$ pN, we overestimate $g(F)$ [138, 142], since overwinding contributes only to a limited extend to DNA elasticity at these forces. Taken together, we have obtained the functional form of the force-dependence of the twist-stretch coupling of dsDNA over a wide force regime. From figure 4.2C it is evident that the total amount of unwinding due to tension is limited: only $\sim 3\%$ of the total twist of dsDNA is lost at 65 pN, the onset of the overstretching transition. This indicates that significant forces are required to induce DNA unwinding by lateral tension. In the cell, where DNA is continuously subject to thermal and protein-induced tension, this robustness of the helical structure might be important for reliable sequence recognition by minor or major-groove binding proteins.

### 4.2.2 Force-induced DNA-strand unpeeling

Figure 4.3 – Stick-slip dynamics are sequence-dependent and close to equilibrium. [A]. Stick-slip melting of three different dsDNA molecules with identical sequence shows reproducible patterns. The curves are offset for visibility. [B]. Dynamics of burst-wise melting (extension rate, 10 nm/sec; sampling frequency, 23 Hz). Bistable behavior is apparent close to the critical force of melting events. [C]. G-C content of the investigated pKYB1 DNA construct from $j = 0$ (first base that unpeels) to $j = 8398$ using a binning window of 30 base pairs.
At forces of \( \sim 65 \text{ pN} \), regime II in figure 4.2B, the experimental force-extension behaviour of dsDNA shows a plateau [50, 105] that is clearly not predicted by the tWLC model, the overstretching transition. During this transition, the force does not remain constant. Instead, at pulling speeds of 10 nm/sec, force-extension curves display a clear saw-tooth-like pattern (figure 4.3A), which suggests that force-induced unpeeling proceeds in a burst-like manner: tension rises until elastic energy is released by the simultaneous opening of a number of base pairs. After a succession of individual unpeeling bursts the DNA molecule fully denatures. Similar molecular "stick-slip" dynamics have been observed in single-molecule dsDNA unzipping experiments, in which tension is applied to both strands on one side of the dsDNA template [76, 143, 144]. In that configuration, a force of 15 pN is needed to separate the two strands, approximately four times lower than required in the DNA stretching configuration. The key difference between the two experimental geometries is the gain in contour length obtained per melted base pair. This gain is four times larger in the unzipping geometry \( (2 \cdot L_{bp,ssDNA} \approx 1 \text{ nm}) \) than in the overstretching geometry \( (L_{bp,ssDNA} - L_{bp,dsDNA} \approx 0.25 \text{ nm}) \) [114]. Moreover, it has been shown that in the unzipping geometry the stick-slip dynamics depend on the base-pairing energy and thus the DNA sequence [143, 144]. Similarly, we observed that in the overstretching geometry the complex stick-slip pattern is very similar for different DNA molecules with identical sequence (\( N = 17 \)) (figure 4.3A), indicating that this pattern is a direct reflection of the sequence-specific base-pairing energy landscape (see also figure 4.3C).

![Figure 4.4](image)

**Figure 4.4** – Only DNA constructs with a single unpeeling front show a deterministic equilibrium stick-slip pattern during DNA overstretching. [A]. The onset of DNA overstretching for 3 different lambda DNA molecules with two non-confined strands. Curves are vertically shifted for visibility. In contrast to measurements with only one unpeeling front (figure 4.3A) no reproducible patterns emerge. [B]. Four successive force-extension measurements on one individual pKYB1-DNA molecule, with variable extension rates. Curves are vertically shifted for visibility.
4.2. RESULTS

The situation is different when unpeeling is allowed to occur from two (or more) fronts. In that attachment geometry, stick-slip features are still observed (figure 4.4A). The patterns, however, vary substantially between experiments and molecules, due to the many alternative unpeeling pathways. Next, we examined whether the observed DNA unpeeling occurs close to equilibrium. We found that reproducible, deterministic stick-slip dynamics are observed at different extension velocities up to 20 nm/sec (figure 4.4B), as expected for a process close to equilibrium. In addition, similar to force-induced RNA hairpin [145, 146] and DNA unzipping experiments [76], we observed bistability close to the critical melting force (figures 4.3B, 4.5). This provides further evidence that unpeeling is a close-to-equilibrium process.

![Figure 4.5](image)

**Figure 4.5** – Bistability close to the critical force of individual DNA unpeeling events. **[A]**. Bistability in a force-extension measurement. Grey triangles depict the stick-slip melting behavior during unpeeling. Open circles represent force and distance for three fixed separations between the optical traps. **[B]**. Histogram of the force-levels for each fixed separation of the two optical traps. **[C]**. Temporal evolution of the force-level for each separation, sampled at 23 Hz.
CHAPTER 4. TWIST, STRETCH AND MELT: QUANTIFYING HOW DNA COMPLIES TO TENSION

To quantitatively describe the deterministic DNA-melting process during overstretching we adapted a statistical physics description of mechanical DNA unzipping [76, 144]. In this model, the state of the DNA molecule is characterized by the number of unpeeled base-pairs $j$, the length of the dsDNA segment $l_{ds}$, and the length of the ssDNA strand under tension $l_{ss}(F)$. For a given state of the DNA construct, the total free energy is calculated by summing the following three energy terms: (i) The free energy change corresponding to opening the double helix from basepair 1 to basepair $j$, $E_{bp}(j)$. This term depends on the DNA sequence and is determined using the nearest-neighbor model for base-pairing of Santa Lucia [22]. (ii) The elastic energy of the ssDNA fraction of $j$ bases stretched to a length $l_{ss}(F)$, $E_{ssDNA}(j, l_{ss})$. This term is calculated using a freely jointed chain model [50]. (iii) The elastic energy of the dsDNA segment of $j_{tot} - j$ base-pairs stretched to a length $l_{ds}$, $E_{dsDNA}(j_{tot} - j, l_{ds})$, which is calculated from the tWLC model introduced in equation 4.4. Using these energy terms, expectation values and variances of observables can be calculated using canonical ensemble averaging [147] over the $(j, l_{ss}, l_{ds})$ phase space, while applying the boundary condition $l_{ds} + l_{ss} + 2x + d_{bead} = d_{trap}$ (with $d_{bead}$ the diameter of the microspheres, $x$ their displacement from the trap centers and $d_{trap}$ the experimentally imposed distance between the trap centers). In figure 4.6A and 4.6B this model is compared to the experimental data, yielding agreement, not only on the force level at which unpeeling occurs, but also on the location and size of the sequence-specific stick-slip melting bursts. Note that our model using the base-pairing energy determined at zero force, predicts the overstretching force of $\sim 65$ pN remarkably well. This indicates that at 65 pN the base-pairing energy, with contributions from hydrogen-bonding and base stacking, is not significantly reduced by tension. A more detailed comparison between theoretically predicted and experimentally observed stalling locations of the unpeeling front is shown in figure 4.6C. In the experimental dwell-time histogram (12 independent force-extension measurements) we observe stalling at all 39 predicted sequence locations. Eight additional stalling locations are not captured by the model. In most cases, these stalling events appear prior to a large melting burst (e.g. at $j = \sim 4200$ bp, where the model predicts a burst of $\sim 570$ bp), which we interpret as a consequence of the equilibrium condition not being entirely fulfilled for large bursts. In figure 4.7 we give additional detail of the energetics of the unpeeling process, showing that the mechanical energy gain equals the denaturization penalty, as expected for an equilibrium process. Hence, overstretching can be quantitatively described using an equilibrium thermodynamic model without any fitting parameter for all but the largest melting events, with the DNA sequence and elasticity of dsDNA and ssDNA as only input parameters.
Figure 4.6 – Modelling of force-induced, sequence-specific unpeeling of dsDNA.

[A]. Force-extension curves of DNA molecules with identical sequence. Grey triangles, data (extension speed, 10 nm/sec); black line, equilibrium model. The insert shows the start of unpeeling. [B]. Comparison of measured (open dots) and calculated (grey line) force as function of unpeeling progress. Here the assumption is made that the mechanical properties of a partly unpeeled DNA molecule are a linear combination of those of dsDNA and ssDNA. [C]. Comparison of the measured (lower part of the graph) and calculated (upper panel) stalling locations of the unpeeling process. Lower panel: dwell time histogram obtained from 12 independent force-extension curves of different DNA molecules with identical sequence. Upper panel: variance of the position of the unpeeling front \( j \) versus its position as calculated from our equilibrium model. Grey vertical lines highlight the predicted stalling locations from this calculation (var(\( j \)) <40 bp). Stars indicate stalling locations that are not captured by the model.
CHAPTER 4. TWIST, STRETCH AND MELT: QUANTIFYING HOW DNA COMPLIES TO TENSION

Figure 4.7 – Energetics of individual unpeeling burst. [A]. Comparison of the energy penalty for melting a dsDNA patch (grey circles) and the mechanical energy gain due to the increased contour length (black triangles) displayed against the sequence location. The melting energy was calculated using the nearest neighbor model for base-pairing of Santa Lucia [22]. The mechanical energy gain was determined by calculating the mechanical energy difference between the initial and final dsDNA/ssDNA hybrid up to the critical melting force, all extracted from the force-extension measurement seen in figure 4.1C. For comparison the absolute value of the energies is displayed. [B]. The energy penalty for a DNA unpeeling event and the mechanical energy gain due to extending the DNA contour length displayed against the number of base pairs in a single melting burst. The graph shows that both energies match over almost the whole range of burst lengths observed (50 - 550 bp). This supports our interpretation that burst-wise, force-induced DNA unpeeling is a close to equilibrium process.

4.2.3 Hysteresis during DNA strand reannealing

Finally, we investigated the reverse process (figure 4.8), reannealing of the unpeeled DNA upon reduction of the distance between the two optical traps. Similar to force-induced unpeeling, reannealing occurs in a burst-wise fashion (figure 4.8A). In contrast to unpeeling, the reannealing pattern strongly varies between experiments. In addition, drops of the force to values lower than 40 pN are frequently observed during reannealing, even at velocities as low as 20 nm/sec. We hypothesized that the hysteresis observed is due to the formation of secondary structure in the unpeeled ssDNA strand. This secondary structure needs to be dissolved, which involves crossing an energy barrier facilitated by thermal fluctuations. The stochastic nature of this process results in a high pausing variability in the reannealing process, causing substantial drops in tension.
4.2. RESULTS

Figure 4.8 – Hysteresis during DNA reannealing. [A]. Two force-extension measurements of the pKYB1 DNA construct with three confined DNA ends. Black: force-induced DNA unpeeling; grey: subsequent strand reannealing. [B]. Average energy per base pair stored in the secondary structure of the unpeeled strand (see equation 4.6). Horizontal line: average base pairing energy of correctly paired nucleotides (2.1 k_BT). Grey lines highlight regions of delayed reannealing, as experimentally observed (panel C). [C]. Dwell-time histogram of the position of the double-stranded/single-stranded junction during strand reannealing (N=5). Locations of prolonged stalls in the reannealing process clearly show up (grey vertical lines).

In order to obtain a quantitative idea about the way the secondary structure formed in the unpeeled strand contributes to the observed hysteresis during strand reannealing, we calculated the energy of this secondary structure using MFOLD [148]. This program calculates the energetically most favorable hybridization state of a single DNA strand. We used the following parameters: temperature 20°C, ionic buffer strength 50 mM NaCl. We followed the increase in the energy of the secondary structure during unpeeling (G_{tor}(j)) in the 5’ to 3’ direction by increasing the length of the unpeeled strand in steps of 100 base pairs from j=0 to j_{final}. Reannealing the segment between j_1 and j_2 needs to overcome the energy barrier posed by the secondary structure in the unpeeled strand (figure 4.8B), which can be determined by:
\[
\Delta G(j_2 - j_1) = G_{tot}(j_0 \leftrightarrow j_2) - G_{tot}(j_0 \leftrightarrow j_1).
\]

These calculations predict regions with enhanced secondary-structure stability, which appear to coincide with locations of prolonged stalls in the reannealing process (figure 4.8B and C).

## 4.3 Discussion

Taken together, our high-precision measurements have enabled us to construct a mathematical model that describes force-induced melting of DNA and the impact of tension on DNA twisting. This model describes DNA elasticity over the full range of forces at which dsDNA can sustain a double-helical structure. Essential in this model is the inclusion of both, the helical structure and the sequence of DNA, two physical features that have a strong impact on protein-DNA interactions \[149–151\]. The twistable worm-like chain (tWLC) model can be applied to gain insight in the effect of proteins and other ligands on DNA twisting and melting. In particular, this allows quantification of the impact of protein binding on the intrinsic twist-stretch coupling of DNA \((g_0)\) and its force dependence \((g_1)\). It is very likely that these properties are altered by protein or ligand binding, just like persistence length and contour length \[152,153\].

Likewise, a sequence-dependent mathematical description of force-induced DNA unpeeling sheds new light on the DNA overstretching transition. We observed a remarkable continuity of the unpeeling process. It has been theoretically predicted that no force-induced unpeeling occurs at ionic buffer strengths above 25 mM NaCl, assuming average base-pair strength. In this theoretical study \[121\] it has been argued that AT-rich regions should unpeel, while GC-rich regions should not and consequently arrest the unpeeling process. As is clear from our data (figure 4.3C, figure 4.6B) even regions with 75% GC-content do not stall the unpeeling front\(^1\).

Furthermore, we did not need to include a kinetic barrier or second transition prior to final strand separation in our model to properly describe the unpeeling of DNA by a single melting front (figure 4.6B) \[154–156\]. This kinetic barrier was interpreted as large-scale rearrangements in the DNA conformation prior to the final melting transition. We show here that for our experimental conditions, which involve controlled strand-attachment and slow extension speeds, no such kinetic barrier is present. This is consistent with our interpretation of DNA overstretching as burst-wise progression

\(^1\)For more detail on the continuity of DNA unpeeling, and its dependence of salt, see chapter 5.
of a single unpeeling front. Previously, theoretical effort has been put into understanding the overstretching transition in an Ising-model, which attempts to capture the cooperatively of the transition by a next-neighbor interaction [157]. We show in figure 4.6B that the phase-transition does not follow a sigmoidal shape, characteristic for the Ising-model. In our view, the high cooperatively of the overstretching transition originates from the very limited number of nucleation points for DNA unpeeling. Furthermore this model allows quantification of local stabilizing or destabilizing effects of protein binding on dsDNA. Unzipping of DNA has been applied to determine the location and interaction energy of DNA-bound proteins [158]. Here we show that our force-induced unpeeling approach has the power to yield sequence-specific data, providing complementary information in an experimentally more accessible configuration. In conclusion, the deeper understanding of dsDNA elasticity and mechanical stability presented in this study provides a new handle to quantify changes in twist and melting behavior of dsDNA. This makes DNA force-extension measurements an even more powerful tool for the investigation of the thermodynamics and structural rearrangements associated with the activity of proteins acting on DNA.

4.4 Methods

All force-extension measurements were conducted under following conditions: 10 mM Tris-HCl, pH 7.8, 50 mM NaCl, 20 °C. The visualization of fluorescently labeled RPA was performed using a combined fluorescence and dual-beam optical trapping instrument, which has been described in detail in chapter 2. The fluorescent labeling of RPA, along with the experimental procedure of visualizing fluorescent RPA on DNA was performed as explained in [105]. The DNA construct that possesses biotin-labels at three ends of the strands was generated as follows. First, two 5’ overhangs were introduced into the circular 8,393 bp plasmid pKYB1 (NEB, Ipswich, USA) using the restriction enzyme EcoRI (Fermentas, Burlington, Canada). The two recessed 3’ ends of this linearized plasmid where biotin-labeled by incorporation of biotinylated dATPs (NEB, Ipswich, USA) along with unlabeled dTTPs (Invitrogen, Carlsbad, USA), by Klenow exo− DNA polymerase (Fermentas, Burlington, Canada). A 3’ overhang was introduced using KpnI (Fermentas, Burlington, Canada), which possesses a single restriction site in this construct. The 8356 bp fragment was selected and purified. A 29 bp primer possessing 4 biotinylated nucleotides at the 5’ end (5’-cTcTcTcT ctc ttc tct ctt ctc tt gtac-3’ ; biotin modifications are depicted with a capital letter) was annealed and ligated to the 3’ overhang of the 8356 bp fragment using T4 DNA Ligase (Fermentas,
Burlington, Canada). The resulting 25 bp 5’-overhang was filled in with biotinylated dATPs along with unlabeled dGTPs, using Klenow exo− DNA polymerase.
Balance between two overstretching modes: melting in the interior vs. from the extremities

Abstract — Long polymers like DNA, RNA and polypeptides play fundamental roles in Biology. Among these macromolecules, DNA possesses unique physical features: its braided helical spatial organization results in a remarkable stiffness. This structure originates from the complex interplay between non-covalent interactions: hydrogen-bonding of complementary bases, stacking of adjacent bases and the electrostatic repulsion of the negatively charged phosphates. Varying the ionic strength can therefore have a profound impact on the double-helical conformation, and under extreme conditions result in a fundamentally different structure [159, 160]. Single-molecule experiments revealed that also tension can force DNA into a different structure [50, 109, 110]. While it is known that the energetics of this overstretching transition are strongly influenced by solvent properties [112], it is to this day unclear how the structure of overstretched DNA is influenced. Here we show that there are two fundamentally different overstretching modes which critically depend on solvent conditions. At low ionic strength the process is dominated by DNA unpeeling (chapter 3,4), while at higher salt concentrations overstretching leads to the formation of internal melting bubbles. Which mode prevails depends on the local DNA-stability, determined by DNA-sequence and shielding of DNA-charge by ions. These results enhance the understanding of the conditions at which DNA loses its double-helical structure, a process that is fundamental in many genomic transactions, such as DNA replication and transcription.
CHAPTER 5. BALANCE BETWEEN TWO OVERSTRETCHING MODES: MELTING IN THE INTERIOR VS. FROM THE EXTREMITIES

5.1 Introduction

The physical properties of DNA are of fundamental importance for a vast amount of cellular processes like genome compaction, gene regulation, transcription, and genome replication [63, 82, 135]. A large family of DNA-associated proteins constantly regulates and adjusts the conformation of DNA, involving bending, twisting and local disruption of DNA's double-helical structure. In order to access the genetic information encoded in the base-sequence, these proteins need to locally open the base pairing of the two strands forming the DNA double helix. While it is well known that the two DNA strands can be separated by increasing the temperature to 60-90 °C [161], proteins follow another approach and generate local stress [151, 162]. More than a decade ago, single-molecule studies of DNA's compliance to tension revealed that the molecule undergoes a structural transition at forces of about 65 pN [50, 108]. Here, the DNA molecule gains ~70% in contour length, while the force remains relatively constant. The synergistic activity of molecular motors can generate significant tension, in the range of the mechanical limits of the double helix, for example during sister chromatid separation [163]. Up to this day, it is still debated what happens to DNA at the molecular level, at this force [114, 121]. Recent experiments have demonstrated that during DNA overstretching, the bases of DNA, which are not exposed to the surrounding solution in the double-helical structure, get chemically modified in the presence of Glyoxal [119], indicating that the base pairing between the two DNA strands are gradually broken. In accord with this observation, our experiments combining optical trapping and fluorescence microscopy (chapter 3 and 4) have directly visualized that, during overstretching, DNA melts via unpeeling from its extremities and nicks close to physiological ionic conditions [105]. The energetics of this unpeeling process could be quantified and modeled in the framework of an equilibrium model governed solely by the elastic properties of single and double-stranded DNA and the underlying sequence, as shown in chapter 4. Strand unpeeling, however, requires either a nick or a free DNA end, which is generally only transiently present inside a cell. How DNA complies to tension under conditions without such unpeeling nucleation sites is not well understood. In chapter 3 we presented experiments on DNA, which was linked with all four strand-ends to optically trapped microspheres. This study has shown that the single-stranded DNA-binding protein replication protein A (RPA) binds to such DNA during overstretching, indicating that also in the absence of free DNA ends or nicks, DNA gradually melts. It should however be noted that, under these conditions, the DNA is torsionally constrained, and thus cannot release its twist. This energetical drawback shifts the overstretching transition to 110
pN \[105, 109\]. So far, however, no experimental procedures have been reported that inhibit DNA strand unpeeling, while allowing relaxation of the torsional stress that builds up during overstretching. Single-molecule DNA overstretching studies focusing on the reverse process, relaxation of tension due to decreasing the end-to-end distance of the DNA, demonstrate pronounced hysteresis effects \[50, 164\]. Two markedly different behaviors were observed, depending on sequence, ionic strength and temperature. These different modes were interpreted as the fingerprint of two distinct mechanisms of overstretching which compete over a wide range of ambient conditions \[164\]. However, this evidence was obtained under the transient out-of-equilibrium conditions, as is generally the case for hysteretic DNA strand reannealing. The two modes have not yet been observed directly and under equilibrium conditions, during low-loading-rate DNA force-extension measurements. But most importantly, the molecular basis underlying both modes has remained elusive. Here, we set out to combine optical tweezers with fluorescence microscopy to investigate how torsionally relaxed DNA complies to tension when unpeeling is inhibited. We show directly, using DNA force-extension measurements under equilibrium conditions, that overstretching proceeds via either of two distinct mechanisms: internal melting bubble formation or unpeeling from free ends or nicks. By varying the ionic strength and studying sequence-specificity, we provide evidence that DNA stability, which is governed by the interplay of electrostatic strand repulsion and base pairing, is the parameter that determines the balance between these two overstretching mechanisms.

5.2 Results

5.2.1 An alternative overstretching mechanism in the absence of unpeeling

To study how DNA complies to forces in the absence of free strand ends or nicks, it is important to separate the effects of inhibited strand unpeeling from the impact of torsional constraint, since structural transitions of DNA involve significant changes in twist. To this end, we used double-stranded (ds) lambda DNA that was partially covered by multiple, short, Alexa 555-labeled RAD51 filaments, as depicted in figure 5.1A. RAD51 is a protein involved in homologous recombination \[165\] and forms helical filaments on both double-stranded and single-stranded DNA (ssDNA) \[166\].
CHAPTER 5. BALANCE BETWEEN TWO OVERSTRETCHING MODES: MELTING IN THE INTERIOR VS. FROM THE EXTREMITIES

These filaments remain stably bound to the DNA in the absence of ATP hydrolysis, which can be suppressed by replacing Mg\(^{2+}\) with Ca\(^{2+}\) [57, 167]. RAD51 nucleoprotein filaments bound to double-stranded DNA are not affected by tension, even when exceeding the overstretching force of \(\sim 65\text{pN}\) [106]. Consequently, strand unpeeling of uncoated dsDNA between RAD51 filaments is inhibited. Yet, torsion cannot build up, it directly releases via unhindered rotation around the single covalent bonds of the phosphodiester backbone or the bond between biotin and chemically modified nucleotide.

Figure 5.1 – RAD51 filaments inhibit DNA unpeeling, which induces internal melting bubbles. [A]. Schematic depiction of lambda DNA coated with two RAD51 filaments, along with a fluorescent image of a lambda DNA molecule, partially covered by Alexa 555-labeled RAD51 filaments. The white dashed lines highlight the horizontal location of the DNA extremities, linked to the microspheres. [B]. Force-extension measurement of DNA molecule in panel A. Grey line: fit of the worm-like chain model up to 35 pN. [C]. Fluorescent image of the gradual overstretching process on lambda DNA coated with two RAD51 nucleofilaments (highlighted by white arrows). Single-stranded DNA was visualized with the single-stranded binding protein RPA. Separate emission channels of image 4 are shown in the upper panel. Image 1-3 show strand unpeeling from the DNA end. In image 4, the DNA complies to tension via formation of internal melting bubbles.
In figure 5.1A we show the fluorescence image of such a partially RAD51-coated DNA molecule. The two DNA unpeeling fronts, generated during DNA overstretching [105] will progress from the extremities, and finally encounter RAD51 filaments in the beginning of the overstretching transition. We estimate from the fluorescence image shown in figure 5.1A that only ~12% of the DNA molecule is able to unpeel, while ~88% is confined in between RAD51 filaments. A force-extension measurement of this RAD51-DNA molecule is shown in figure 5.1B. In the initial part of the overstretching transition (dark grey), overstretching takes place via stick-slip wise extension of the DNA, which is characteristic for DNA unpeeling from the extremities [105]. We estimate from figure 5.1B that, as expected, about 9% of the RAD51-DNA construct unpeels in this stick-slip wise manner. When extended to a length of ~18 µm, the stick-slip pattern vanishes, apparently when both unpeeling fronts encounter the RAD51 filaments closest to both DNA ends. The force level then rises to ~70 pN, and remains constant (shaded in light grey in figure 5.1B). From this observation we can directly conclude that, even in a situation where the DNA cannot unpeel, the molecule overstretches. The characteristics of overstretching, however, differ considerably: the force does not exhibit fluctuations, but remains constant at a level slightly higher than when unpeeling is allowed. Taken together, these observations provide evidence for an alternative overstretching mode energetically close to DNA unpeeling, but with a fundamentally different mechanism.

5.2.2 Melting bubbles appear upon unpeeling arrest

In order to unravel the mechanism of the alternative overstretching mode, we probed whether ssDNA becomes exposed, similar to force-induced DNA unpeeling. To test this hypothesis we performed two-color fluorescence measurements in which we inhibited the unpeeling process with Alexa 555 labeled RAD51 filaments (similar to figure 5.1B) and monitored the appearance of ssDNA with eGFP-RPA (the single-stranded-DNA binding protein replication protein A) [105]. The experimental procedure was as follows: (i) we partially overstretched RAD51-DNA constructs, (ii) subsequently exposed the construct to a buffer containing eGFP-RPA, while keeping the extension fixed, (iii) moved the construct out of the RPA containing buffer and (iv) finally excited and detected Alexa555 and eGFP fluorescence successively. This procedure was repeated several times with increased extension of the construct. Images obtained from such an experiment are shown in figure 5.1C. Images 1 to 3 show that, initially, ssDNA appears at the DNA extremities, characteristic for unpeeling [105]. In image 3, the unpeeling fronts have reached both short RAD51 filaments. When
the construct is extended even further, image 4, RPA is observed to bind within the DNA section that is protected from unpeeling. In striking contrast to unpeeled DNA, RPA binds in this DNA section in a non-contiguous manner. Taken together, these results indicate that when unpeeling is inhibited, overstretching results in a new DNA conformation that is specifically targeted by the single-stranded DNA binding protein RPA, which indicates that DNA locally denatures. This new conformation is scattered along the DNA molecule in a qualitatively similar way compared to thermal melting which also initiates by the formation of numerous melting bubbles [168].

5.2.3 Electrostatic shielding assists melting-bubble formation

It is well established that thermal DNA melting strongly depends on solvent ionic strength. Dissolved ions shield the strong, negative charge (2e− per base pair) of DNA and thus decrease the electrostatic repulsion between the two DNA strands, stabilizing the double helix [169]. Using the same argumentation, it is expected that the two modes of DNA overstretching identified above are affected differently by ionic strength. In case of melting bubbles, the two denatured DNA strands remain in close proximity, in contrast to unpeeling, implying that electrostatic interactions might play a more significant role in bubble melting than in unpeeling. Free ions in an electrolyte screen free charges and decrease electrostatic interactions exponentially over a length scale called the Debye-length (28). At monovalent salt concentrations of ~ 25 mM, this Debye length is comparable to the DNA diameter of ~ 2 nm. For higher, more physiological salt concentrations, the screening length becomes much smaller than the DNA diameter. To test the effect of electrostatic shielding on both overstretching modes we performed DNA force-extension measurements over a wide range of monovalent and divalent salt concentrations. We used a special DNA construct that has only one unconstrained DNA end (figure 5.2, inset), such that unpeeling can only initiate from one end and progress in only one direction [170], (chapter 4). This allows us to follow the location of the melting front in the DNA sequence during force-extension measurements, which is essential for distinguishing potential effects of ionic strength from those of DNA sequence.

In figure 5.2, force-extension curves of the DNA construct with only one unconstrained DNA end at different ionic conditions are presented. The curves obtained at increasing monovalent salt concentrations clearly indicate that at low salt concentrations (< 200 mM NaCl) overstretching proceeds primarily via unpeeling, while at high salt concentrations (> 500 mM NaCl) melting bubble formation prevails. The
5.2. RESULTS

Figure 5.2 – Increased electrostatic shielding shifts the DNA overstretching mode from unpeeling to force-induced melting bubbles. [A]. Force-extension measurements of DNA molecules with identical sequence at different monovalent salt concentrations, ranging from 50 mM NaCl to 1 M NaCl. Insert: schematic of DNA attachment with only one unconstrained DNA-end. [B]. Divalent salt titration of the force-extension behavior of DNA molecules with the same sequence as in panel A. Monovalent salt concentration: 150 mM NaCl. The divalent salt was increased from 0 mM to 20 mM MgCl₂.

overstretching force increases with ionic strength, in agreement with previous studies [112, 115, 121, 171]. Figure 5.2B shows that the relative occurrence of both overstretching modes depends even more critically on divalent salt concentration. At 150 mM NaCl, MgCl₂ concentrations below ~ 2 mM show primarily unpeeling, while at concentrations above ~ 10 mM bubble melting has the upper hand. Taken together, these measurements clearly show that the increased stability of DNA at higher ionic strength results in a strong inclination of DNA to overstretch by formation of melting bubbles.
5.2.4 Unpeeling stalls in GC-rich regions

Figure 5.3 – GC rich regions stall the unpeeling process at high salt. [A]. Black: relative GC content of the first 4000 base pairs of the DNA construct investigated in figure 5.2 (binning window: 40 bp). Regions with a GC content of more than 65% are highlighted in light grey to guide the eye. Grey: histogram of all locations where the unpeeling process arrested. This location coincided with the position where the stick-slip wise force-induced melting process switched over into the more homogeneous force-induced bubble melting mode, which can be seen in figure 5.2. The histogram was binned with 30 bp (localization accuracy of the position of the unpeeling front). [B]. Comparison of the distribution of the GC-content of all events of unpeeling arrest (grey) with the GC content of the DNA molecule after binning with a window of 40 bp.

Figure 5.3 – GC rich regions stall the unpeeling process at high salt. [A]. Black: relative GC content of the first 4000 base pairs of the DNA construct investigated in figure 5.2 (binning window: 40 bp). Regions with a GC content of more than 65% are highlighted in light grey to guide the eye. Grey: histogram of all locations where the unpeeling process arrested. This location coincided with the position where the stick-slip wise force-induced melting process switched over into the more homogeneous force-induced bubble melting mode, which can be seen in figure 5.2. The histogram was binned with 30 bp (localization accuracy of the position of the unpeeling front). [B]. Comparison of the distribution of the GC-content of all events of unpeeling arrest (grey) with the GC content of the DNA molecule after binning with a window of 40 bp.

Within a wide range of ionic strengths, both DNA melting mechanisms, unpeeling or bubble formation, take place during overstretching. In most force-extension experiments under these conditions, we noted that overstretching started with unpeeling and abruptly switched to the generation of melting bubbles (figure 5.2A, B). To gain insight into these switching events, we investigated the local GC content of switching locations (over a binning window of 40 bp). Since the vast majority of arrests occurred in the beginning of the DNA overstretching transition, we focused
this analysis on the first half of the DNA molecule (~4000 bp). From the histogram in figure 5.3A it is evident that unpeeling arrest and switching to bubble formation almost exclusively occurs at DNA regions with high GC content and thus high stability. This observation is supported by figure 5.3B, where we compare the GC content of all locations where unpeeling arrested with the distribution of the GC content of the whole DNA molecule. A systematic bias for the inhibition of strand unpeeling towards a high GC content is clearly visible. Taken together, these observations show that strand unpeeling during DNA overstretching primarily arrests at locations in the DNA sequence where the high GC content presents an energy barrier for unpeeling to proceed that is difficult to surmount, thus shifting the equilibrium to melting bubble formation.

5.2.5 Melting bubble formation initiates at AT-rich regions

Unlike DNA unpeeling, which requires a strand nick or free end to nucleate, force-induced melting bubbles are able to initiate throughout the DNA molecule. We hypothesized that bubbles would preferentially form at regions of lower DNA stability, i.e. in AT-rich regions. To test this, we overstretched lambda DNA at an ionic strength at which unpeeling is inhibited and overstretching occurs exclusively by melting bubble formation (150 mM NaCl and 20 mM MgCl₂) (figure 5.2B). We extended lambda DNA to a length at which overstretching had just initiated and subsequently transferred the DNA molecule to a buffer containing eGFP-RPA, in order to visualize melting bubbles (see chapters 3,4) [105,170]. From images like those presented in figure 5.4A, we obtained the local melting probability by applying a threshold to the fluorescence intensity and interpreting regions along the DNA molecule with exceeding intensity as denatured. A melting probability map (figure 5.4B) was calculated by summing the normalized and thresholded line intensities of 22 DNA molecules. Figure 5.4B indicates that force-induced bubble formation has a significant preference to initiate in the middle of lambda DNA. In our experiments we do not determine the orientation of the DNA molecule. Consequently, the melting probability map presented in figure 5.4B represents an average over both possible orientations. In figure 5.4C we show, for comparison, the GC content of lambda DNA averaged over both orientations. When comparing the melting probability map with the GC content it is evident that bubble formation preferentially initiates at AT-rich regions. This strong anti-correlation between hybridization energy and melting probability provides additional evidence that bubble formation is indeed the result of losing base pairing interactions. This is in accordance with a study showing that the bases of DNA become
exposed to the solvent during overstretching [119]. Taken together, these results indicate that the nucleation behavior of force-induced melting-bubble formation not only qualitatively resembles thermal DNA melting [161], but also shows a corresponding sequence-dependence [168,172].

**Figure 5.4 – Bubbles nucleate at AT-rich regions.** [A]. Selection of fluorescence images of RPA binding to lambda DNA in the very beginning of the overstretching transition at 150 mM NaCl and 20 mM MgCl₂. [B]. Relative probability for the nucleation of a melting bubble, obtained from fluorescence measurements of the initiation of RPA binding, like in panel A, for 22 different lambda DNA molecules. (For details, see text) [C]. GC content of lambda DNA after averaging both orientations that are indistinguishable in the optical tweezers assay (bin window: 40 bp).
5.2.6 Both melting modes differ significantly when approaching final strand separation

Next we focused on the final phase of overstretching, the final separation of both DNA strands. In figure 5.5A we compare force-extension curves of the DNA construct of figure 5.2 under conditions where unpeeling occurs (low ionic strength) to conditions where bubble melting occurs (high ionic strength). When comparing the final phases of the overstretching transitions, a striking difference is noticed. Unpeeling progresses without interruption and single-stranded DNA is formed without apparent obstruction, as can be judged from the correspondence of the elastic response with that of single-stranded DNA. In case of force-induced bubble melting, the situation is quite different: a clear barrier is observed before final strand separation, in agreement with previous studies \[120, 121\]. As can be seen from figure 5.5B, final strand separation is not characterized by a single barrier crossing, but exhibits rupture-like characteristics, with several, semi-stable intermediate states, suggesting a complex energy landscape. It has been demonstrated before that final strand separation occurs under out-of-equilibrium conditions \[120, 155\]. Even at the low extension velocities applied (\(\sim 10 \text{ nm/sec}\)), the rupture-like characteristics did not show deterministic features, indicating that individual DNA molecules do not have a homogeneous structural composition. Consistent with the interpretation of the data presented in this study, we speculate that, in almost completely overstretched DNA, the strands are held together by a stochastic collection of only a few, relatively stable, GC-rich segments, apparently requiring additional elastic strain to denature. Before final strand separation, after abandoning the overstretching force plateau, the elastic response of DNA cannot be described by either an individual ssDNA strand, or two non-interacting ssDNA strands (figure 5.5B), in agreement with \[121\]. Apparently, two non-interacting ssDNA strands do not provide a realistic description, which is, however, not surprising for two reasons. First, it has been demonstrated that DNA retains twist in this kinetic regime \[109, 110\]. On average, the two strands are wound around each other with a rise of 30 to 40 bps. Second, at 65 pN the phosphodiester backbone of ssDNA is fully extended and, in the backbone, no entropic degrees of freedom are thermally excitable \[114\]. The unpaired nucleosides, however, have substantial conformational freedom, which results in a significant entropic contribution to the elastic properties of ssDNA \[114\]. In the case of two intertwined non base-paired ssDNA strands, the conformational freedom of the nucleosides might well be restricted sterically, which will then have a strong impact on the elasticity. Quantification of this effect will require further experimental and theoretical studies. Taken together, the different
kinetics of final strand separation between the two overstretching modes originates from their fundamentally different nature. Unpeeling proceeds deterministically and follows a geometrically imposed pathway, while force-induced bubble melting occurs throughout the DNA molecules, with the most stable regions as the last ones to denature, raising the energy barrier for final strand separation.

Figure 5.5 – Final strand separation differs between force-induced bubble melting and strand unpeeling. [A]. Comparison of the force-extension behavior of two DNA molecules with identical sequence at two different salt concentrations. Insert: DNA attachment geometry, similar to experiments in figure 5.2. At low ionic buffer strength (50 mM NaCl), DNA unpeels at ~65 pN via stick-slip dynamics [170]. The hysteresis shows that DNA completely reanneals upon reducing the bead separation. At 150 mM NaCl, 10 mM MgCl$_2$, the DNA overstretches via the force-induced bubble melting mode. A barrier prior to the final strand separation is evident. The elastic response of ssDNA is shown for comparison. [B]. Three different measurements of the final strand separation at 150 mM NaCl, 10 mM MgCl$_2$. For comparison, the elastic response of ssDNA is shown in black open triangles. From this curve the elastic properties of two non-interacting ssDNA molecules (grey triangles) was inferred by $F_{2ssDNA}(x) = 2 \cdot F_{ssDNA}(x)$.

5.3 Discussion

The experiments presented here directly and unambiguously show that DNA overstretching proceeds via a competition between two fundamentally different modes: strand unpeeling from DNA ends and the formation of melting bubbles throughout the DNA. The force-levels at which both processes occur are close, indicating that they are
energetically proximate. The force-extension curves of both modes are dramatically different, reflecting fundamentally distinct mechanisms. DNA-unpeeling progresses in a stick-slip wise manner even under equilibrium conditions [170] (chapter 4). In sharp contrast, overstretching in the internal melting bubble mode progresses without distinct features. Since force-induced bubble formation initiates at multiple locations throughout the DNA molecule (figure 5.1C), multiple junctions between double and single-stranded DNA are formed, creating numerous denaturation pathways that can be explored. Force-extension measurements under bubble-formation conditions thus reflect a temporal and spatial average of the dynamics of many melting bubbles, blurring sequence inhomogeneity, resulting in a smooth overstretching transition.

We have identified melting bubbles using fluorescent RPA, a single-stranded DNA binding protein [128]. It cannot be ruled out, on basis of our data, that RPA binding during the overstretching transition does not reflect binding to single-stranded DNA but to 'S-DNA' [50,121], a structure proposed for overstretched DNA, with both strands parallel and base pairing intact. Two notions, however, support our interpretation that melting bubbles and thus single-stranded DNA is formed during overstretching. First, we observed a clear anti-correlation between locations of increased stability of the DNA double helix (reflected in GC content) and AT-rich regions as preferred locations of force-induced melting bubble nucleation (figure 5.5B and C). It has been observed before that binding of RPA is highly dependent on the stability of DNA, and occurs at thermally induced, transient melting bubbles [173], triggered by destabilizing the double-helical structure at low ionic strength [174,175], or by applying torque to the DNA [176]. Second, clear stacking interactions between aromatic residues of RPA and uncomplexed bases of single-stranded DNA can be discerned in the RPA's crystal structure that contribute to its affinity for single-stranded DNA [128]. The main ssDNA binding domains, DBD-A and DBD-B are located on the RPA70 subunit of this heterotrimeric protein [177]. Each of these domains consists of an OB fold, where two beta sheets wrap around the single-stranded DNA, generating a ~ 1.7 nm narrow channel, smaller than the diameter of base-paired DNA. Taken together, these notions support our interpretation that DNA overstretching involves the rupture of Watson-Crick base pairing, as reported by RPA binding to overstretched DNA.

We find that monovalent and divalent salt act differently, more pronounced than expected on basis of valence, in agreement with the observation that divalent cations like Mg$^{2+}$ specifically condense to the phosphate backbone [178]. In addition to ionic interactions, which neutralize the negative charges of the DNA, divalent cations tend to be fully hydrated, causing the interaction with double-stranded DNA to be more sequence specific. Water molecules of the solvation shell can specifically interact with
CHAPTER 5. BALANCE BETWEEN TWO OVERSTRETCHING MODES: MELTING IN THE INTERIOR VS. FROM THE EXTREMITIES

the bases by accepting or donating hydrogen bonds. This has been shown to result in an enhanced affinity of Mg\(^{2+}\) for DNA with an increased GC content [178]. Mg\(^{2+}\) ions bound to GC base pairs have been demonstrated to stabilize the denatured state and thus shift the equilibrium to the melted configuration, without increasing the rate of base-pair opening. It was proposed that this mechanism involves the formation of hydrogen bonds between water in the primary solvation shell and unpaired bases [179]. Taken together, these previous studies highlight that specific interactions of Mg\(^{2+}\) with GC-rich regions in DNA tend to stabilize transiently formed melting bubbles. In our experiments, this might be the reason for the shift in balance from unpeeling to melting bubble formation as the preferred overstretching mode.

Our unique experimental approach combining optical tweezers with fluorescence microscopy, allowed us to shed new light on the DNA-overstretching transition. Our results allowed us to clarify the impact of DNA sequence on overstretching and in particular the distinct behavior of GC-rich and AT-rich regions in DNA overstretching. We found that, at high ionic strength, the stability of GC-rich regions is sufficient to inhibit unpeeling, at the favor of melting bubble formation in AT-rich regions. The two mechanisms of overstretching compete over a broad range of solvent conditions, including those occurring in the cell. Thus, in its natural environment, AT-rich regions will melt under high mechanical stress, while GC-rich regions will form a barrier against complete strand separation, which will facilitate the reformation of the double-helical structure once the external stress has vanished. These findings contribute significantly to our knowledge of the mechanical properties and stability of DNA, which are of fundamental importance to understand how proteins control the physical state of DNA.

5.4 Methods

The visualization of fluorescently labeled RPA and RAD51 was performed using a combined fluorescence and dual-beam optical trapping instrument, which has been described in detail, in chapter 2. The fluorescent labeling of RPA and RAD51, along with the experimental procedure of visualizing the fluorescent protein on DNA was performed as explained in chapter 2 and 3. To obtain partially covered lambda DNA molecules, we adjusted the stoichiometry of DNA and RAD51, as described in [106]. The DNA construct that possesses biotin-labels at three ends of the strands was generated as described in chapter 5, while lambda DNA was prepared for optical tweezers experiments according to chapter 2.
Melting bubble statistics and dynamics of overstretched DNA

Abstract — The thermodynamic stability of the Watson-Crick double-helical structure of DNA has received significant attention due to both its central role in life and its complex physical properties. While the binding energy of individual base pairs and their interactions with neighboring base pairs are relatively well studied, understanding the cooperative dynamics that occur during DNA melting still remains a challenge. This denaturation can be induced in two ways, by temperatures exceeding 60-90 °C [161], or by applying a mechanical tension of ~ 65 pN [50, 108]. In chapter 5 we demonstrated that under intermediate to high salt concentrations, both mechanisms show striking similarities: denaturation occurs via the formation of melting bubbles at structurally weak AT-rich regions. While it is well known that the thermal melting transition is accompanied with significant fluctuations of the DNA conformation [180], these have not been studied during the force-induced melting, or overstretching transition. Here we investigate the number of melting bubbles for individual DNA molecules undergoing the overstretching transition, and furthermore study the fluctuation dynamics of these melting bubbles. We find that force-induced bubble melting requires a nucleation energy of 17.8 k_BT, which limits the number of melting bubbles to below ten for lambda DNA with around 50,000 base pairs (bp). The boundaries of melting bubbles undergo diffusive motion on a typical time scale of around 2 to 5 µs with an amplitude of around 2 bp. Our observations indicate that in contrast to thermal melting, during force-induced melting, GC and AT base pairs, despite their different binding energy, melt in the same proportion. This can explain why the force-level remains widely constant during DNA overstretching.
6.1 Introduction

The double-helical structure of DNA is of fundamental importance to life since it protects the genetic information stored in the sequence of the four distinct bases by burying them in the middle of the double helix, protecting the most reactive part of the nucleotide [181]. The Watson-Crick structure is sustained by base-stacking interactions and hydrogen bonding of complementary bases, resulting in an individual binding strength of complementary bases that is rather low (on the order of \( \sim 1-3 \, k_B T \)) [22]. This characteristic makes DNA a very dynamic entity. Even at room temperature, significantly below the melting temperature of DNA, individual bases are thermally excited. Single base pairs locally denature and reanneal in a stochastic manner and DNA exhibits so-called 'breathing' fluctuations [182]. These fluctuations become sizeable when the temperature is increased [180] and the probability for cooperative opening of numerous base pairs becomes significant. Melting bubbles emerge and the DNA starts to denature, a highly cooperative process due to the high energy cost of generating a new melting bubble compared to extending an existing one [183].

Breathing fluctuations of DNA are relatively well studied. Spectroscopic studies using fluorescence correlation spectroscopy [182], circular dichroism [23] and NMR [173] have investigated the dynamics and magnitude of the conformational fluctuations well below the melting temperature of DNA. Here, transient bubbles in AT-rich regions relax on a time scale of around 50 \( \mu s \) [182], while individual bases, when thermally denatured, reanneal in about 0.1 \( \mu s \) [173]. It has further been shown that the breathing fluctuations at the junction between double-stranded and single-stranded DNA involve about two base pairs, probed for AT base pairs [23]. Significantly less is known, however, about the DNA fluctuations during melting, their intrinsic size and timescale is not known. Studies probing temperature fluctuations during DNA melting, originating from local entropy fluctuations in DNA, have provided some qualitative insight [180]. Even though this study convincingly disclosed the presence of significant conformational changes during melting, the experimental approach did not allow for a detailed quantification of the process.

In chapter 5, we demonstrated that applying tension denatures DNA in a qualitatively similar manner to thermal melting. Indeed, theoretical studies confirmed that increasing the tension on DNA lowers the melting temperature up to the point where, for a sufficient force, DNA melts at room temperature [114, 184]. This highly cooperative process, occurring at \( \sim 65 \, pN \), is referred to as the DNA overstretching transition [50, 108]. The molecular basis of this structural transition, however, was only
very recently resolved unequivocally (Chapter 3, 5). For intermediate to high ionic solvent concentrations, DNA gradually melts by the formation of several melting bubbles initiating at weakly bound AT-rich regions [185]. For lower salt concentrations (150 mM NaCl and below), DNA overstretches via strand unpeeling from the extremities [105]. So far, however, it has been unknown to what extend DNA exhibits melting fluctuations during the DNA overstretching transition.

Here we investigate the conformational dynamics during DNA overstretching by monitoring fluctuations in the total length of the molecule in a single-molecule, optical-tweezers assay. We observe these end-to-end length fluctuations only in the presence of force-induced melting bubbles. We measure the number of melting bubbles with the help of the fluorescently labeled single-stranded binding protein RPA and typically find less than 10 melting bubbles for a DNA construct of 48,502 base pairs, corresponding to a nucleation energy of a melting bubble of 17.8 k_BT. The length fluctuations originate from opening and reannealing of individual base pairs on the microsecond time scale. The variation in length of overstretched lambda DNA is around 0.4 nm^2, which corresponds to a fluctuation of only on average 1-2 base pairs on the boundary of individual melting bubbles.

6.2 Results

6.2.1 Monitoring the magnitude and the dynamics of fluctuations during force-induced bubble melting

To access both the dynamics and the amplitude of the melting fluctuations of DNA, we measured the end-to-end length of DNA in an optical tweezers assay with a high bandwidth (Nyquist sampling frequency 10 kHz). It was demonstrated before that fully denatured DNA is \(\sim 170\%\) longer [50]. Consequently, fluctuations in the denaturation state during melting will result in fluctuations in the contour length of the DNA molecule. We mapped the fluctuation of a microsphere attached to one DNA end (lambda phage DNA, 16.4 \(\mu\)m, 48,502 bp) held in two optical traps, in a one-point microrheology approach [186] (figure 6.1A). At room temperature, at DNA tensions below the overstretching force of 65 pN, we observed the characteristic Lorentzian power spectrum of the Brownian motion of a bead in a harmonic trap [90] (see figure 6.1B, upper panel). In sharp contrast, at tensions of \(\sim 65\) pN, where DNA denatures [105] and melting bubbles occur (chapter 5), we observed deviations from the Lorentzian behavior. These additional fluctuations occurred exclusively in the di-
rection of the DNA molecule (‘x-direction’), no additional fluctuations were observed in the perpendicular direction (‘y-direction’ figure 6.1B, lower panel).

Figure 6.1 – Mapping the conformational fluctuations of DNA during overstretching. [A]. Schematic of the experimental approach. The motion of one trapped microsphere is recorded using a position-sensitive detector in direction of the DNA molecule (‘x-direction’) and in the perpendicular direction (‘y-direction’). The insert shows a schematic of a dynamic melting bubble, causing length fluctuations in the DNA molecule. [B]. The power spectrum of the motion of the microsphere in x and y direction prior to DNA overstretching is shown in the upper panel. The lower panel shows the power spectrum, recorded with a tethered DNA molecule at $\sim 65$ pN, where the DNA undergoes the overstretching transition. [C]. Power spectrum of the length fluctuations of DNA, obtained by $S_{\text{DNA}}(f) = S_x(f) - S_y(f)$. A $f^{-2}$ behavior is evident, characteristic for diffusive dynamics.

To ensure that these additional low frequency fluctuations originate from the stochastic dynamics of melting bubbles, we performed the following controls. To rule out the possibility of interference effects from the second optical trap, we recorded time traces that we converted to power spectra for one microsphere in a laser trap, while the second trap was disabled. The power spectra under these conditions were identical to the ones recorded for a pair of microspheres held in both traps at a separation of about 10 $\mu$m, and disclosed no additional motion at lower frequencies. Connecting the two beads with a DNA molecule did not result in an altered power spectrum. Nei-
ther for low DNA tensions, where the DNA molecule fluctuates due to the Brownian motion, nor at tensions just below the overstretching transition (∼55 pN), did we observe additional fluctuations at lower frequencies. Also DNA molecules subjected to tensions beyond the overstretching force (80 pN, figure 6.3B) did not show deviations from the Lorentzian power spectrum. As a final check we recorded power spectra for DNA overstretched in a buffer at low ionic strength (50 mM NaCl). In this environment, DNA overstretches via a different mechanism: strand unpeeling from the extremities in a stick-slip wise manner (chapter 4). For all these conditions we did not observe the deviations from the Lorentzian behavior, which are very prominent in the direction parallel to the DNA axis, during force-induced bubble melting (figure 6.1B, C).

The observed power spectrum $S(f)$ can be dissected into contributions of the Brownian motion of the microsphere and the conformational changes of the overstretched DNA by $S(f) = S_{\text{bead}}(f) + S_{\text{DNA}}(f)$. In figure 6.1B, lower panel, we measured $S(f)$ (x-direction') and $S_{\text{bead}}(f)$ (y-direction') directly, granting access to the low frequency length fluctuations of overstretched DNA. In figure 6.1C we show the spectral density of the length fluctuations $S_{\text{DNA}}(f)$. A clear $1/f^2$ behavior for frequencies higher than 10 Hz can be seen, characteristic for diffusive dynamics. We interpret this as stochastic opening and closing of the Watson-Crick base pairs at the boundary of the melting bubble, which results in diffusive kinetics of the boundary. At lower frequencies the power spectrum rolls off, indicating that slow conformational changes are confined by the two optical traps. To quantitatively map the end-to-end length fluctuations to the dynamics of melting-bubbles, we next developed a mechanical model for DNA fluctuations and determined the number of melting bubbles at different stages of the overstretching transition.

6.2.2 Visualizing the number of bubbles

The fluctuations shown in figure 6.1 originate from the dynamics of several melting-bubble boundaries. The free energy of a boundary between double-stranded (ds)DNA and single-stranded (ss)DNA has been determined before in biochemical studies, and revealed a cooperativity factor $\sigma$ during thermal melting to be on the order of $5 \cdot 10^{-5}$ [183]. To obtain information on the number of melting bubbles during force-induced DNA melting, we used a custom single-molecule setup that combines optical tweezers with fluorescence microscopy (chapter 2). We visualized melting bubbles using the fluorescently labeled single-stranded binding protein RPA, like in chapter 2-5. Experiments were performed as follows: In a buffer with a salt concentration of
CHAPTER 6. MELTING BUBBLE STATISTICS AND DYNAMICS OF OVERSTRETCHED DNA

Figure 6.2 – Visualizing the number of melting bubbles during overstretching. [A]. Fluorescence images of overstretched lambda DNA. Melting bubbles are visualized with the fluorescent single-stranded binding protein eGFP-RPA. [B]. Statistics of the number of bubble-boundaries for increased progression of the overstretching transition. Grey circles: individual measurements. Black diamonds: Binned average and standard error. The fit (see text) yields 17.8 k_BT as the energy to nucleate a melting bubble.

150 mM NaCl, 20 mM MgCl₂, where DNA exclusively overstretches via force-induced bubble formation (chapter 5), we partially overstretched lambda DNA. Subsequently, we exposed the construct to a buffer containing eGFP-RPA, while keeping the extension fixed. After a typical incubation time of about 5 seconds, we moved the construct out of the eGFP-RPA containing buffer and finally excited and detected eGFP fluorescence. This procedure was repeated several times with increased extension of the construct. Images obtained from such an experiment are shown in figure 6.2A. We analyzed 125 measurements of melting-bubble numbers for 16 different lambda DNA molecules, recorded at different stages of the overstretching transition (figure 6.2B). As anticipated, the number of boundaries monotonically increased up until halfway through the transition, after which the number of bubbles gradually decreases by coalescence, until one large bubble remains close to final denaturation. The average number of boundaries remained below 10, indicating a significant cost for melting bubble formation. We captured the distribution by a model where the stability of the DNA molecule is considered homogeneous such that bubble nucleation has an equal probability at each position along the molecule. Then the entropic contribution can be determined by counting the number of permutations of ways k base pairs can be used to form n bubbles over N base pairs, while the bubble-nucleation energy is given by n·E_b. The maximum likelihood estimation for E_b, given the data in figure 6.2B yielded E_b = 17.8 ± 0.08 k_BT (see Methods).
6.2.3 Dynamics of melting fluctuations at different stages of the
overstretching transition

We recorded power spectra of DNA molecules at different stages in the overstretching transition, as indicated in figure 6.3A. In figure 6.3B we show the measurement of six successive power spectra, recorded for the same DNA molecule at increasing overstretching progress. Two observations can be made. First, the fluctuations due to conformational changes of the DNA shifted to lower frequencies with increasing progress of the overstretching transition. Second, these fluctuations vanished when the extension is increased beyond the coexistence region at $\sim 65$ pN. In the following section, we will rationalize these qualitative observations in terms of the physical properties of DNA with an analytical model for DNA melting fluctuations.

Figure 6.3 – Changes in the fluctuation spectrum of DNA upon gradual overstretching. [A]. Force-extension measurement normalized to the contour length of lambda DNA. The overstretching transition is parameterized such that 0 represents the onset of the transition, and 1 represents the end. Six extensions, at which the length fluctuations of DNA were measured, are highlighted. [B]. Six power spectra of the diffusive motion of a microsphere attached to a DNA molecule at the extensions indicated in panel A.

6.2.4 Modeling DNA fluctuations during DNA overstretching

We applied a mechanical model in order to quantify the length fluctuations of overstretched DNA. The approach is schematically depicted in figure 6.4A. Without tether, the movement of both microspheres in the stationary laser traps is fully described by the spring constant $k_T$ of the harmonic laser trap and the viscosity $\eta_T$ of the bead in
the surrounding liquid. The microspheres are connected via a DNA molecule, which is subjected to a tension $f$. For simplicity, the DNA molecule is modeled by two segments in series, both with a spring constant $2k_{\text{DNA}}$ resulting in a total spring constant $k_{\text{DNA}}$ of the molecule.

![Figure 6.4](image)

**Figure 6.4** – Details of the mechanical model of the conformational fluctuations of DNA. [A]. The optical traps are parameterized by the spring constant $k_T$ and the viscosity $\eta_T$ of the microsphere in the surrounding liquid. Both microspheres are subjected to a thermal force $\xi_T$. The DNA tether is characterized by $k_{\text{DNA}}$. The total length of the DNA is allowed to fluctuate under the stochastic, thermal force $\xi_b$ via the dash-pot with a viscosity of $\eta_b$. [B]. Maximum-likelihood fit of the mechanical model to the powerspectrum of the motion of a microsphere linked to an overstretched DNA tether.

Non-elastic conformational changes altering the contour length are modeled by introducing a dash pot as a dissipative element. This accounts for the conversion between DNA in a double-helical and denatured conformation at the boundaries of the melting bubble, as both phases possess a different length per base pair: 0.34 nm for double-stranded DNA and 0.58 nm for denatured DNA. Therefore, thermal opening and reannealing of base pairs at the bubble boundary will induce tension fluctuations. We consider two thermal forces, one on each bead ($\xi_{T_1}$ and $\xi_{T_2}$), and one stochastic tension source due to fluctuations in the denaturation state at the bubble boundaries (modeled by the stochastic force pair $\xi_b$ and $-\xi_b$, see figure 6.4A). Both have the following stochastic properties, given by the fluctuation-dissipation theorem:

\[
\langle \xi \rangle = 0,
\]
\[
\langle \xi_{T_i} \xi_{T_j} \rangle = 2\eta_T k_B T \delta_{ij},
\]
\[
\langle \xi_b \xi_b \rangle = 2\eta_b k_B T.
\]
Here the inverse boundary mobility $\eta_b$ is expected to scale linearly with the number of melting bubble boundaries. Note that the thermal force acting on the melting bubble is integrated such that an elongation / contraction of the bubble results in an anticorrelated movement of both beads (figure 6.4A). The force balance on each segment with respect to the left, stationary laser trap yields:

$$k_T x_1 + \eta_T \dot{x}_1 = f,$$
$$2 \cdot k_{DNA} \cdot x_2 = f + \xi_{T_1},$$
$$\eta_b \cdot \dot{x}_3 = f + \xi_{T_1} + \xi_b,$$
$$2 \cdot k_{DNA} \cdot x_4 = f + \xi_{T_1},$$
$$k_T x_5 + \eta_T \dot{x}_5 = f + \xi_{T_1} + \xi_{T_2}.$$

Since the traps are fixed, the distance between the focal centers is constant in time:

$$x_1 + x_2 + x_3 + x_4 + x_5 = \text{const}.$$

Solving these equations for $x_1$, we obtain the following expression for the spectral density:

$$\langle x_1(\omega)x_1(\omega) \rangle = \frac{k_B T \eta_T}{k_T^2 + \eta_T^2 \omega^2} + \frac{k_B T (k_{DNA}^2(2 \cdot \eta_b + \eta_T) + \eta_b^2 \eta_T \omega^2)}{4k_{DNA}^2 k_T \eta_b^2 \omega^2 + \eta_b^2 \omega^2(k_T^2 + \eta_T^2 \omega^2) + k_{DNA}^2(k_T^2 + (2 \eta_b + \eta_T)^2 \omega^2)}.$$ (6.1)

We performed maximum likelihood fits on 60 power spectra of the movement of a microsphere linked to overstretched lambda DNA (figure 6.4B). We first focused on the change of the elasticity of DNA during overstretching. In figure 6.5, the spring constant obtained for DNA molecules undergoing the overstretching transition is displayed. Despite the significant spread in the values, a reciprocal dependence can be discerned (see inset), as expected for two elastic phases in series, where one is converted into the other.

This result of figure 6.5 would suggest that non-overstretched lambda DNA at the overstretching force possesses a spring constant of $(23 \pm 3) \cdot 10^{-3}$ pN/nm, while the overstretched conformation has an elasticity of $(9 \pm 1) \cdot 10^{-3}$ pN/nm. Taken into account the non-linear nature of the spring constant of DNA due to entropic effects (at low tensions) and unwinding of the double helix (at high tension) (see chapter 4), we expected a spring constant of double-helical DNA of about $60 \cdot 10^{-3}$ pN/nm at a tension of 65 pN. The power spectrum thus shows only limited sensitivity to the actual value of the spring constant of DNA, and underestimates $k_{dsDNA}$ by about a
CHAPTER 6. MELTING BUBBLE STATISTICS AND DYNAMICS OF OVERSTRETCHED DNA

Figure 6.5 – Change of DNAs spring constant upon increased progress through the overstretching transition. Each triangle was obtained by fitting equation 6.1 to an individual measurement of a power spectrum in the presence of an overstretched DNA molecule. Black diamonds show the average and standard error. A reduction of the total spring constant is evident. A fit assuming a serial combination of springs is shown as a black line, which yields \((23 \pm 3) \cdot 10^{-3} \text{ pN/nm}\) for dsDNA and an elasticity of \((9 \pm 1) \cdot 10^{-3} \text{ pN/nm}\) for the overstretched state. These numbers are compared with the established values in the text. The inset shows the inverse spring constant for the same data set.

factor of 3. Nonetheless, the systematic reduction of the DNA spring constant during overstretching by about a factor of 2.5 can provide interesting insight. Determination of the elastic properties of overstretched DNA (during bubble-melting), consisting out of two intertwined, non-base-paired DNA strands has proven to been difficult, to this day. It was shown that this complex denatured conformation retains an average twist of about 40 bp \([109, 110]\). Furthermore, we showed in chapter 5 that neither one single-stranded DNA strand \((19 \cdot 10^{-3} \text{ pN/nm})\), nor two non-interacting ssDNA strands \((38 \cdot 10^{-3} \text{ pN/nm})\) \([50]\) act as a realistic model for the elasticity of denatured DNA. The observed reduction of the spring constant of overstretched DNA by a factor of about 2.5 together with the established spring constant of lambda DNA at 65 pN \((60 \cdot 10^{-3} \text{ pN/nm})\) would indicate a spring constant of overstretched DNA of about \(24 \cdot 10^{-3} \text{ pN/nm}\). Further experimental and theoretical studies are necessary to confirm the reduction of \(k_{\text{DNA}}\) during overstretching by about 2.5. To this day the quantification of the elastic properties of this DNA conformation has proven to be difficult, due to its transient stability at the very end of the DNA overstretching transition \([120]\). Further quantitative insight into the elasticity of overstretched DNA will help illuminating the molecular conformation and interaction of this phase,
which remains a challenge to this day.

Next, we focused on the changes in DNA length \((x_2 + x_3 + x_4)(\omega)\) (figure 6.4A) induced by dynamic melting bubble boundaries. To this end we calculated the change in the DNA-length fluctuation spectrum as we let the boundaries become mobile in our mechanical model. We integrated the difference (shaded grey in figure 6.6A) over all frequencies, for each of the 60 individual measurements. In figure 6.6B we show the resulting variance in DNA length at various overstretching states. These fluctuations originate from the local denaturation / reannealing of individual base pairs. Then the length change from the Watson-Crick, base-paired state \((L_{ds}=0.34 \text{ nm})\) to the melted state \((L_{ss}=0.58 \text{ nm})\) and \textit{vice versa}, is 0.25 nm per base pair. In figure 6.6B, we normalized the DNA length fluctuations to the number of opening / closing base pairs. Relating the fluctuation amplitude in the end-to-end length of DNA to a fluctuation amplitude of individual melting-bubble boundaries required knowledge of the number of dynamic boundaries. From figure 6.2 we only obtain the number of bubble boundaries. DNA, however, possesses a very heterogeneous binding energy landscape. GC base pairs have a binding strength of \(\sim 2.9 \text{ k_B T}\), while AT base pairs only have a binding strength of \(\sim 1.3 \text{ k_B T}\) [22]. Given this, within the manuscript we will analyze our data set for two scenarios, representing extremes in sequence sensitivity: (i) all boundaries show identical dynamics. (ii) Half of the boundaries are dynamic (for AT base pairs), half are static (for GC base pairs). In figure 6.6C, we show the variance and the root-mean-square amplitude (in base pairs) of the motion of melting bubble boundaries for both scenarios. This quantity shows only small variation with the progress of the overstretching transition, and remains between 2.5 and 1.5 base pairs for both scenarios, in agreement with earlier measurements of breathing fluctuations at a ss/dsDNA junction [23]. We observed a small reduction of the fluctuation amplitude for increased progress of the overstretching transition, which did not follow the same trend as the number of melting bubbles (figure 6.2B) and hence cannot be explained solely by a change in the number of fluctuating melting bubble boundaries, indicating that the kinetics of fluctuations change at increased progress of the overstretching transition.

To investigate what causes the slight reduction of the fluctuation amplitude we analyzed the diffusive kinetics of individual melting-bubble boundaries. Therefore we converted the viscosity of the boundary \(\eta_b\), obtained by maximum-likelihood fitting, into a diffusion constant using the Einstein relation [187]. In a one-dimensional random walk approach, the diffusion constant can be related to a stepping length \(\Delta x\) and an average hopping time \(\tau\) according to \(D=\Delta x^2/2\tau\). For the motion of a melting bubble, \(\Delta x\) is given by the length difference of dsDNA and ssDNA (0.25 nm). Figure 6.7A
CHAPTER 6. MELTING BUBBLE STATISTICS AND DYNAMICS OF OVERSTRETCHED DNA

Figure 6.6 – Quantifying the magnitude of the melting bubble-fluctuations. [A]. Comparison of the power spectrum of motion in the presence and absence of DNA fluctuations. The mechanical parameters for the black curve are similar to the fit shown in figure 6.4B. The grey curve shows the power spectrum with an infinite viscosity of the melting bubble boundary. The area shaded in grey represents contributions of DNA fluctuations. [B]. We integrated the power spectrum of the DNA length ($x_2 + x_3 + x_4$) after determining the mechanical parameters by maximum-likelihood fitting of equation 6.1. The average and standard error for five extensions is shown in black. Similar to panel A, we subtracted the contributions of the microsphere in the trap. Given the difference of the length of a base pair of dsDNA (0.34nm) and ssDNA (0.58nm) we normalized this expression to the root-mean-square amplitude in base pairs. [C]. Normalization of the DNA-length fluctuation of panel B to the fluctuation of individual melting-bubble boundaries. Two scenarios are displayed: (i) all bubble boundaries show the same dynamics (homogeneous DNA stability) and (ii) only half the boundaries are dynamic, the rest is static (DNA stability heterogeneous).

shows the diffusion constant and the hopping time of end-to-end length of the DNA molecule. Again, we normalized these measures to the motion of individual boundaries for the two scenarios (i) all boundaries are equally dynamic and (ii) only half of the boundaries show dynamics. In figure 6.7B we show the diffusion constant and
the hopping time of individual melting-bubble boundaries, which can be associated with the opening / closing rate of the base pair at the junction between dsDNA and ssDNA. Both quantities vary only moderately with progress of overstretching. We find a typical time scale for the melting / reannealing dynamics of individual base pairs during overstretching of around 2 µs or 5 µs, dependent on the scenario of the number of dynamic melting bubbles. This quantity varies only moderately (≈ 15%) with progress of the overstretching transition. Only close to the end of the overstretching transition, we did observe a reduction of the hopping dynamics of individual boundaries by about 30%, compared to the average.

Figure 6.7 – Characterization of the diffusive motion of melting bubble boundaries. [A]. Diffusion constant and hopping time of the DNA length fluctuations at different stages of the overstretching transition. The diffusion constant of panel A was converted in a hopping time of the melting bubble boundary via \( D = \frac{\Delta x^2}{2\tau} \), with \( \Delta x \) as the length difference of dsDNA and ssDNA (0.25 nm). In both panels we show the average and standard error in black. [B]. The average of the diffusion constant and the hopping time from panel A was normalized to the kinetics of an individual bubble boundary. Error bars: propagated standard error. Two different scenarios are considered, representing extremes in the DNA stability (homogeneous / heterogeneous).
6.3 Discussion

Here we have shown that DNA undergoes melting fluctuations during overstretching. We characterized this process by quantifying the number of melting bubbles as well as the diffusive dynamics of the boundary between ssDNA and dsDNA. For a salt concentration of 150 mM NaCl, 20 mM MgCl$_2$, we found a nucleation energy cost for the formation of a melting bubble of 17.8 k$_B$T. This number is significantly larger than the $\sim$10.5 k$_B$T reported for thermal melting [183]. This difference is, however, not unexpected. The formation of a melting bubble during thermal melting is accompanied by an entropy gain since single-stranded DNA is significantly more flexible, and will hence fluctuate substantially [184]. During force-induced melting-bubble formation, this entropic benefit is not present to the same extend since the conformational degrees of freedom of the DNA phosphate-backbone are suppressed at a force of $\sim$65 pN [114]. It will be interesting to see if and how the nucleation energy depends on the ionic strength of the buffer. In particular the impact of divalent cations like Mg$^{2+}$ is of interest, since these ions can condense on the negatively charged DNA backbone, and specifically interact with the bases, altering their stability [179].

Our mechanical model for the length fluctuations of DNA enabled us to map the dynamics of the force-induced melting process of DNA. We found evidence for diffusive kinetics of the melting-bubble boundaries, caused by the stochastic denaturation / reannealing of boundary base pairs. We used the distribution of melting bubbles, directly obtained via fluorescence microscopy, to relate the fluctuation in end-to-end length of DNA to the diffusive motion of individual melting bubble boundaries. Due to the heterogeneity in the binding energy landscape, it is likely that boundaries in AT-rich regions are more dynamic than the ones in GC-rich regions. Therefore we analyzed our data set for two scenarios, representing two plausible extremes. In one approach, all boundaries exhibit the same kinetics. In the contrary extreme, showing high sequence sensitivity, only half of the boundaries are dynamic, while the others remain static. For both scenarios we found a root-mean-square amplitude of the diffusive motion of between one and two base pairs, in agreement with [23]. We anticipate that the actual value will lie between the amplitudes for both scenarios that sketch the extremes of the impact of the sequence heterogeneity.

Given the observation that melting bubbles nucleate in AT-rich regions (chapter 5), it would be plausible to expect that fluctuations in the beginning of the overstretching transition are more significant than at the end, due to a 'GC-distillation' in the remaining dsDNA segment. Such a mechanism was shown for thermal melting, it was demonstrated that AT base pairs denature at lower temperature than GC base pairs.
6.3. DISCUSSION

Increasing the temperature during thermal melting then increases the GC content of the remaining base-paired DNA section [182, 188]. From figure 6.6B it can be seen that the fluctuation dynamics do not disclose sequential melting of first the AT base pairs, followed by GC base pairs at the later stage of the overstretching transition. We only observed a moderate change in this amplitude (∼ 20%) for increased progress of the overstretching transition, in contrast to the difference in binding energy between AT and GC base pairing of a factor of two. This suggests that irrespective of the stage of the overstretching transition, AT and GC base pairs are denatured in more or less the same proportion. This observation is plausible when considering the high bubble nucleation energy for force-induced bubble melting compared to thermal melting. It counteracts an effective distillation mechanism for the remaining part of the double-helical structure. Even though force-induced melting bubbles are formed in AT-rich regions, the AT / GC-content of the flanking dsDNA regions will on average be stochastic, with no bias for DNA stability. When the extension between the laser traps is now increased, predominantly the existing bubbles will grow through a stochastic AT / GC-content, given the high energy cost for nucleating a new one. The absence of an effective GC-distillation process can explain why the force level, during DNA overstretching, remains relatively constant, which has been an unresolved issue of the force-induced melting model.

We observed the typical time scale of melting-bubble boundaries fluctuations between ~ 2 - 5 µs (figure 6.7B). NMR studies on the lifetime of the denatured state of individual base pairs at room temperature found comparable time scales for the life-time of the denatured state [173, 179]. We conclude that at the boundary of the melting bubble, the base-pair lifetime, which is around 10−2 s without external force [173], greatly reduces to the microsecond time scale, such that stable melting bubbles can occur. Analogous to changes in the fluctuation amplitude, we find the base-pair-fluctuation time scale to only moderately change for increasing progress of the force-induced bubble melting transition. This supports the interpretation that no large-scale separation dependent on the sequence (weaker AT base pairs in the denatured state, more stable GC base pairs in the bound state) occurs during DNA overstretching. Only very close to the end of the overstretching transition, we do observe a decrease of the hopping time by about 30%. It will be interesting to see if this can be interpreted as a precursor of the out-of-equilibrium dynamics of the final strand separation [120].

Taken together, the temporal bandwidth along with the precise manipulation ability of optical tweezers enabled us to study the melting bubble dynamics during DNA overstretching quantitatively. We have shown that the dynamics of individual melting-bubble boundaries occur on the microsecond time scale, and involve a diffusive root-
mean-square amplitude of one to two base pairs. Furthermore we found indications that, in contrast to thermal melting, AT and GC base pairs denature in the same proportion over wide ranges of the overstretching transition, which can explain why the overstretching force remains constant for this transition. These results will improve the understanding of the mechanical stability of DNA, and furthermore enable a quantitative comparison of thermal and force-induced DNA melting.

### 6.4 Methods

All experiments were performed with a custom setup that combines dual optical tweezers with a fluorescence microscope, as described in detail in chapter 2. Fluorescent labeling of RPA, along with the experimental procedure to visualize the fluorescent protein on DNA was performed as explained in chapter 3. Lambda DNA was prepared for optical tweezers experiments according to chapter 2. We derived an analytic expression for the bubble statistics of overstretched DNA as follows. In equilibrium, the probability of having \( n \) bubbles along a tether with \( N \) base pairs, which has been pulled a fraction \( x \) through the overstretching transition, can be written as:

\[
P_{x,N}(n) \propto W(n, x, N) \cdot e^{-nE_b/k_BT}
\]

Here \( E_b \) is the cost of forming a bubble and \( W(n, x, N) \) represents the number of states consistent with \( x, n, \) and \( N \). The \( n \) bubbles can be positioned between the \((1-x)N\) double-stranded base pairs according to:

\[
\frac{(1-x)N!}{((1-x)N-n)!n!}
\]

Similarly, the \( n \) continuous double-stranded regions can be positioned between the single-stranded nucleotides as:

\[
\frac{xN!}{(xN-n)!n!}
\]

Here, terms of size 1 were consistently ignored for simplicity, which should be justified with respect to our experimental accuracy. With this we can write:

\[
P_{x,N}(n) \propto \frac{e^{-nE_b/k_BT}}{(n!)^2((1-x)N-n)!(xN-n)!},
\]

where we have dropped any term independent of the number of bubbles, since this will be set by normalization. Working in the limit where we have a small number of bubbles compared to both the number of single stranded and double-stranded
basepairs, we are justified in using Stirling’s approximation on the last two factors of the denominator. For small \( n/N \) this results in:

\[
P_{x,N}(n) \sim \frac{e^{-nE}}{(n!)^2 I_0(2e^{-nE/2})} \quad \text{with} \quad E = E_b/k_B T - \ln(Nx) - \ln(N(1-x))
\]

where the normalization is calculated by summing \( n \) from zero to infinity, thereby relying on that \( E_b \) is large enough to suppress the unphysical contributions from \( n > xN, (1-x)N \). \( I_n \) represents the modified Bessel function. From this we can calculate the maximum likelihood function used to estimate \( E_b \) given a set of experimental outcomes \( n_i \):

\[
m(E_b) = -\sum_i \ln(P_{x,N}(n_i)) \quad \text{(6.2)}
\]

The most likely nucleation energy \( E_b \), given a dataset of bubble numbers, can be found by maximizing 6.2:

\[
\frac{\partial m(E_b)}{\partial n} = 0.
\]
Single-stranded DNA templates for optical tweezers experiments

Abstract — A significant challenge in single-molecule studies of DNA-protein interaction is the generation of adequate DNA templates. Essential genomic transactions such as DNA damage repair and DNA replication take place on single-stranded DNA (ssDNA). An efficient protocol to generate ssDNA templates will make these processes more accessible to optical tweezers studies. However, no systematic study benchmarked approaches to generate ssDNA templates for single-molecule studies. In contrast to biochemical assays, single-molecule studies involving optical or magnetic tweezers often require DNA templates in the range of several thousand base pairs due to their dimensions in the micrometer-range. Only such lengths permit fine adjustment and hence sound control over the separation between DNA ends. This demand introduces a technical challenge and limits the application of conventional biochemical methods such as polymerase chain reaction (PCR). At the same time, the ability to apply tension to DNA-templates offers opportunities for the generation of ssDNA. Here we survey force-induced DNA denaturation as a means to efficiently generate ssDNA templates for optical tweezers experiments. We first present a protocol to append biotin labels to both ends of only one DNA strand. We then demonstrate that forces just exceeding the overstretching force of ~ 65 pN are sufficient to denature DNA and we determine the time scale of this force-induced melting mechanism. We find a denaturation time in the order of seconds, which is furthermore strongly influenced by the ionic strength of the solvent. We envision that this method to reliably create ssDNA templates will grant access to a new range of optical tweezers assays investigating fundamental biochemical processes.
CHAPTER 7. SINGLE-STRANDED DNA TEMPLATES FOR OPTICAL TWEEZERS-EXPERIMENTS

7.1 Introduction

Single-stranded DNA is a naturally occurring intermediate in many processes in genome maintenance, such as DNA replication, repair and recombination. During DNA replication, large stretches of ssDNA intermediates occur on the lagging strand [189] and are protected from nuclease digestion by single-stranded binding proteins (SSB). SSBs also prevent reannealing of the single strands, avoid the formation of secondary structures and promote the recruitment of other replication factors to the template [52, 63, 190, 191]. In homologous recombination, recombinase proteins of the RecA/RAD51 family [57, 192–194] assemble on ssDNA and catalyze strand exchange between homologous DNA molecules. In order to study these fundamental cellular processes on a single-molecule level, a reliable protocol for the preparation of ssDNA constructs is crucial.

Optical tweezers are a powerful and versatile single-molecule technique to study the mechanics of DNA [50, 144] and the interaction between DNA and proteins. This technique has been used to study DNA replication, transcription, and recombination, and significantly contributed to the understanding of the DNA maintenance by specific biochemical reactions [66, 195–197]. Optical tweezers can isolate and manipulate DNA, and allow the controlled study of genomic transactions on an individual DNA template. The use of optical tweezers however imposes different demands on the design of the assay. DNA constructs need to be specifically labeled at the extremities to attach them to the microspheres that can be manipulated via optical tweezers. Most importantly, using optical tweezers for the manipulation and precise detection of changes in the DNA conformation due to protein interactions sets demands on the length of the DNA. For a dual-tweezers geometry, polarization crosstalk between both traps can impede with a precise force detection for trap separations in the range of 4 μm and below [85]. For a surface-based assay, with one end of the DNA tethered to glass, volume exclusion effects and a lateral force-detection makes DNA constructs significantly smaller than the microbead diameter, typically ranging around 1 μm, unpractical. A DNA template that is commonly used is lambda phage DNA (48,502 bp, 16.4 μm) [50]. DNA constructs in the range of ~1 μm and shorter have however also been used [198, 199]. Bacterial plasmids in the range of ~5,000 to ~10,000 base pairs, offer a versatile template for the generation of appropriate DNA constructs for single-molecule studies.

To this date, most single-molecule studies involve dsDNA, while only a limited number of optical or magnetic tweezers studies investigated processes on ssDNA, as discussed below. This has three key reasons. First, ssDNA is intrinsically less stable than dsDNA.
Second, labeling of ssDNA ends with biotin or digoxygen for microsphere attachment is more challenging since there are no specific biochemical tools for ssDNA modification. Third, ssDNA can form relatively stable secondary structure and can crosslink with other ssDNA molecules at high concentrations. For these reasons, trapping and successively isolating a single ssDNA molecule with optical tweezers is demanding. An effective approach for circumventing these problems is to first attach a single dsDNA molecule between two optically trapped microspheres and to subsequently convert it into ssDNA. Previously, this has been achieved by destabilizing the DNA with a sodium hydroxide solution \[^{[50, 153, 200, 201]}\], by digesting one DNA strand with exonucleases \[^{[52, 63, 191, 202]}\], or by applying forces higher than the DNA overstretching force of 65 pN \[^{[120, 153, 200]}\], at which DNA undergoes the overstretching transition, which was discussed, in detail in chapter 3-6. In these chapters we showed that this transition is caused by the cooperative melting of the DNA molecule \[^{[105, 185]}\]. The details of this melting process highly depend on the nucleotide sequence and the ionic solvent concentration. At low salt DNA, unpeels from its extremities \[^{[105, 170]}\], while at higher salt, internal melting bubbles form \[^{[185]}\]. It was also shown that this two melting modes coexist over a broad range of ionic solvent conditions, and result in a fundamentally different behavior at the final strand separation. While DNA unpeeling results in unhindered DNA denaturation, the formation of melting bubbles induces a significant energy barrier prior to final strand separation \[^{[185]}\].

Here we systematically study force-induced DNA denaturation as a method for the generation of ssDNA in optical tweezers over a range of physiologically relevant ionic solvent concentrations. We describe a strategy for biotin-labeling of dsDNA from two different plasmids on both ends of one strand. These DNA molecules denature into ssDNA after applying forces above 65 pN. We found that force-induced denaturation occurs on the timescale of seconds, and that its dynamics depend on ionic strength and tension on the DNA molecule.

## 7.2 Materials and Methods

### 7.2.1 Biotinylation of two plasmid DNAs (pTR19-ASDS and pKYB1) on both ends of one strand

The two expression vectors, pTR19-ASDS (10,729 bp, kind gift of Dr. Y.J.M. Bollen) and pKYB1 (New England Biolabs) were biotinylated on both ends of the same strand with a similar approach: On one end, biotinylated nucleotides where incorporated.
Figure 7.1 – Strategy for biotin-labeling of different linearized DNA plasmids on both ends. [A]. The expression vector pTR19-ASDS was digested with EcoRI and the 5’-overhangs were filled in with biotinylated nucleotides by Klenow exo−. Next a 3’-overhang was introduced by an Apal digestion reaction. The 9,900 bp fragment was purified and a 29 bp oligonucleotide was ligated to the 3’-overhang. [B]. The plasmid pKYB1 was linearized with EcoRI. Klenow exo− inserted biotinylated nucleotides at both 3’ ends. Next a 3’-overhang was introduced by KpnI. The 8,356 bp fragment was purified and a 29 bp oligonucleotide was ligated to the 3’-overhang. The GC content of the resulting DNA constructs are displayed for comparison.

to a single-stranded overhang, on the other end of the same strand, a biotinylated oligonucleotide was ligated as depicted in figure 7.1A. To this end, both circular plasmids, pTR19-ASDS and pKYB1 were digested with EcoRI (Fermentas), creating two 4-nucleotide 5’-overhangs (3’-TTAA-5’). The complementary nucleotides biotin-14-dATP and dTTP (Invitrogen) were incorporated into these overhangs by Klenow DNA Polymerase exo− (Fermentas). Digestion and labeling were conducted simultaneously (1.5 hours at 37 ℃) in a reaction containing: plasmid DNA (pTR19-ASDS or pKYB1) (20 nM, in total 3 μg), biotin-14-dATP (66 μM), dTTP (66 μM), EcoRI (0.6 units/μl), 0.2 units/μl Klenow exo− DNA polymerase and 2X Tango™ buffer (Fermentas). After a purification step using the GFX PCR DNA and Gel Band Purification Kit (GE Healthcare), the DNA construct was digested to introduce a 3’ overhang. For pTR19-ASDS, the restriction digestion was done with with Apal (0.6 units/μl, 1X Buffer B, Fermentas), resulting in a 9,896 bp and a 824 bp fragment. For pKYB1, the digestion
was carried out with KpnI (Fermentas), (1 Fast Digest Unit/µl, 1X Fast Digest Buffer, Fermentas), yielding a 8356 bp and a 37 bp fragment. The 9,900 bp (pTR19-ASDS) fragment containing a 3’-overhang was separated and purified on an agarose gel using the QIAquick gel extraction kit (Qiagen). For the pKYB1 construct, we purified the 8356 bp fragment with the GFX PCR DNA and Gel Band Purification Kit. Next, a biotinylated oligo (Biolegio, 5’-cTcTcTcT ctc ttc tct ctt ctc tt ggcc-3’ for the pTR19-ASDS DNA, 5’-cTcTcTcT ctc ttc tct ctt ctc tt gtac-3’ for the pKYB1 construct; biotin modification depicted by a capital letter) was annealed to this construct (50X excess) in 1X T4 Ligase buffer (Fermentas) by heating the reaction mixture to 65 °C for 5 minutes, followed by slow cooling down to room temperature. Subsequently, the ligation reaction was initiated by adding T4 DNA ligase (60 Weiss units, 2 hours at room temperature). Finally, the DNA was purified with the GFX PCR DNA and Gel Band Purification Kit. The resulting DNA constructs are 9,900 bp or 8360 bp long with two biotinylated dATPs at the 3’-end, and a 25 nucleotide overhang at the 5’-end of the same strand, containing 4 biotinylated dTTPs.

7.2.2 Optical Tweezers Setup

A detailed description of the optical tweezers instrument can be found in chapter 2. In short, a Nd:YAG laser (3 W continuous wave, 1064 nm, Ventus 1064, Laser Quantum) was used to generate two optical traps. The laser beam was split into two orthogonally polarized beams using a polarizing beam splitter cube (10BC16PC.9, Newport). The two beams were expanded with a 1 : 2.67 telescope system. One beam was steered by laterally displacing a telescope lens. Two traps were produced with a high-numerical aperture water-immersion objective (Plan Apo 60X, numerical aperture = 1.20, Nikon). Two beads (1.87 µm, streptavidin coated, Spherotech) were held in these traps, and their microscope image was digitized using a frame grabber board (IMAQ PCI-1409, National Instruments). We determined the distance between the two beads using a Labview program (National Instruments), applying template directed pattern matching. We used a position sensitive detector (DL100-7PCBA Pacific Silicon Sensor Inc.) for force detection. Its signal was lowpass-filtered with a cutoff frequency of 145 Hz, and subsequently digitized using a 16 bit data-acquisition board (PCI 6229, National Instruments). The force calibration factor was obtained by recording the powerspectrum of the unfiltered position sensitive detector signal, as explained, in detail, in chapter 2. Force and distance were simultaneously recorded by a Labview program, operating at a sampling rate of 25 Hz. For fluorescence excitation of Alexa-555, the linearly polarized light of a 532 nm laser
7.3 Results

7.3.1 dsDNA denatures at tensions exceeding the overstretching force

The basis of our approach to convert dsDNA into ssDNA using optical tweezers is to link only one of the two strands of dsDNA to the microspheres in the optical tweezers and subsequently denature the dsDNA molecule. Here we systematically test the feasibility of force as a means to generate this denaturation. In chapter 5 we demonstrated that DNA overstretches via two fundamentally different mechanisms, dependent on the DNA sequence composition and the ionic strength of the solvent. Importantly, we showed that the final denaturation is severely influenced by the overstretching mode.

To evaluate if applying mechanical tension is a practical method to generate ssDNA, independent of the details of the DNA sequence and the ionic strength of the solvent, we addressed two questions: Does the dsDNA denaturation occur at moderate forces, i.e., just exceeding the overstretching force of $\sim 65$ pN, and if so, on what timescales does the DNA denature?

To address these questions, we first biotin-labeled two DNA plasmids with different sequences on both ends of the same strand (figure 7.1). These DNA molecules were attached to two optically trapped beads in an optical tweezers instrument and gradually exposed to forces slightly exceeding the overstretching force of 65 pN.

First, we examined if overstretching the pTR19-ASDS construct at 50 mM NaCl, an ionic strength that induces unpeeling, results in DNA denaturation. Similar to chapter 3,[105], we used the tension-sensitive, fluorescently labeled single-stranded binding protein mtSSB to visualize unpeeled single-stranded DNA. In figure 7.2A, we visualize the denaturation process during unpeeling. It can be observed that the double-stranded section gradually diminishes upon further extension. In figure 7.2A, image 4, only a small dsDNA section is retained. Increasing the extension further leads to complete denaturation. This was confirmed in a force extension experiment shown in
7.3. Results

Figure 7.2B, where the elastic properties of the DNA molecule before and after overstretching are compared.

![Figure 7.2B](image)

**Figure 7.2 — Visualization of DNA melting via strand unpeeling, at 50mM NaCl, for the pTR19 construct.** [A]. As shown in chapter 3 and 4, DNA unpeels when overstretched in intermediate and low salt concentrations. We use the tension sensitive single-stranded binding protein mtSSB [105] as a marker for unpeeled DNA (highlighted with a white arrow). We observe that the unpeeling fronts move away from the DNA ends. In image 4, both unpeeling fronts are closer than the diffraction limit of our fluorescent microscope (~ 250 nm). Upon further extension, the complementary strand melted off. [B]. We confirmed the complete denaturing of DNA by comparing the force-extension behavior before and after the melting process. As can be clearly seen, after the overstretching process shown in panel A, the DNA denatured into ssDNA.

Next we examined if the pKYB1 DNA construct, which possesses a different sequence, shows similar denaturation when applying tension. In figure 7.3A, we show the force-extension behavior of this DNA construct at 50mM NaCl. We indeed observe a stick-slip wise lengthening of the DNA at the overstretching force at a pulling rate of 10 nm/sec. For these extension rates, this stick-slip wise lengthening of the DNA molecule is an equilibrium process, and characteristic for strand unpeeling (chapter 4, 5) [170, 185]. At the end of the overstretching transition, just as visualized in figure 7.2A for the pTR19-ASDS DNA, the DNA molecule completely denatures without further obstruction. The elastic properties of the pKYB1 DNA immediately coincide with that of ssDNA, as can be seen from figure 7.3A. The situation is quite different for higher salt concentrations. In figure 7.3B we overstretch the pKYB1 DNA construct
in 500mM NaCl. Two observations are striking. First, the overstretching transition in the force-extension measurement does not show pronounced features, as is the case for DNA unpeeling. As was shown in chapter 5, this indicative for force-induced bubble melting \cite{185}. Concurrently with the force-induced bubble melting mode, we observe a barrier that needs to be surmounted prior to final strand separation. This barrier is of kinetic nature \cite{120,155}, which means that for high salt, final denaturation involves the stochastic crossing of one or several energy barriers. As a result, higher force or longer waiting times are needed in order to obtain ssDNA.

\begin{figure}
\centering
\includegraphics[width=\textwidth]{figure7.3}
\caption{Melting behavior of the pKYB1 DNA construct depends on the salt concentration. [A]. The DNA construct is denatured by force at an ionic strength of 50 mM NaCl. The black trace shows the overstretching of a dsDNA molecule. The stick-slip pattern at $\sim 65$ pN is characteristic for DNA strand unpeeling. The elasticity of single-stranded DNA is shown, as a reference, in grey. [B]. A force-extension measurement of dsDNA, at an ionic strength of 500 mM NaCl, is shown in black. The featureless overstretching transition at $\sim 65$ pN is indicative for the formation of force-induced bubble formation. Final DNA denaturation occurs after an additional energy barrier is surmounted. The elastic properties of ssDNA are shown in grey, as a reference.}
\end{figure}

7.3.2 Dynamics of force-induced DNA denaturation

Next we focused on the time scale of the final strand separation since we have seen that DNA, dependent on the ionic solvent conditions, does not immediately denature at the end of the overstretching transition. For these experiments we used the pTR19-ASDS (9,925 bp) DNA construct, which is biotinylated on both ends of one of the strands. We repetitively extended single DNA molecules to three lengths (figure
7.3. RESULTS

Figure 7.4 – Characterization of force-induced DNA melting. [A]. Force-extension measurement of dsDNA (black dots) and ssDNA (green triangles). To convert dsDNA into ssDNA, the molecule was exposed to cycles of successive, straining, reannealing and probing. [B]. Force-time trace of a force-induced DNA melting event at a ionic solvent strength of 50 mM NaCl. The extension is cycled between the straining, reannealing and probing regime, indicated in panel A. The force drop from 65 pN to 20 pN in the second probing interval (both intervals highlighted in grey) reveals that the conversion of dsDNA to ssDNA was completed within the prior 5 seconds straining interval. [C]. Details of the dynamics of the final strand separation for the pTR19-ASDS DNA extended to a force of 120 pN in 150 mM NaCl. In the straining regime, the force drops significantly, until, after the fourth cycle, a force niveau of ∼95 pN is reached. At this force the DNA fully melted to ssDNA, as can be seen in a force value of ∼20 pN in the last probing regime.

7.4A). First, the dsDNA was extended beyond overstretching to more than 0.58 nm per base pair (corresponding to tensions of 70-85 pN, straining regime). For a time $t_s$ (straining time) the dsDNA was allowed to melt. Second, the DNA was relaxed to a tension below 5 pN for a period of typically 1 to 2 seconds, allowing partially melted
DNA to reanneal completely [155]. After reannealing, the DNA has become either entirely double-stranded (when the melting process had not completed during the straining period), or single-stranded DNA (when melting had completed). Third, to probe the state of the DNA, we extended the molecule to \( \sim 0.47 \) nm per basepair and measured the tension on the DNA. At this extension the elastic response of dsDNA is \( \sim 65 \) pN, while that of ssDNA is about 20 pN because of its significantly longer contour length (figure 7.4A). We cycled multiple DNAs, one at a time, through such straining - reannealing - probing cycles until they melted completely into ssDNA, broke, or melted into a ssDNA/dsDNA hybrid. We attribute the latter two events to the presence of nicks: breaking occurs when the biotinylated strand is nicked, incomplete melting when the unbiotinylated strand is nicked and only a segment of this strand is released in solution during denaturation. We restricted our analysis to complete melting events, since it was our objective to produce stable and well-defined ssDNA.

Figure 7.4B shows two typical probing - straining - reannealing cycles at 50 mM NaCl. At the start of the experiment, the DNA was entirely double-stranded, as can be seen from the force value of 65 pN in the first probing interval. Within the next 5 seconds of straining, the molecule completely melted into ssDNA, as is evident from the tension of only 20 pN measured in the second probing interval. For each DNA molecule showing such a tension drop in the probing regime, we confirmed the complete denaturation into ssDNA by a complete force-extension measurement. The dynamics of DNA denaturation change significantly when raising the salt concentration. In figure 7.4C we show an experiment similar to that in figure 7.4B, however at 150 mM NaCl, and with a substantially higher mechanical strain (120 pN instead of 80 pN). In the straining regime, the force drops significantly, until, after the fourth cycle, a force level of \( \sim 95 \) pN is reached. Here, the DNA molecule finally denatured into ssDNA, as is evident from the force value of \( \sim 20 \) pN in the successive probing cycle.

To systematically explore the impact of the ionic solvent concentration, we repeated this experiment for many pTR19-ASDS DNA constructs (\( N = 121 \)), with a constant straining force of 75 pN at different salt concentrations. We quantified our results by the melting efficiency, which we defined as the number of dsDNA molecules melted into ssDNA divided by the sum of probing - straining - reannealing cycles they required to denature. This number is equal to one if all DNA molecules are converted to ssDNA in one cycle; less if it takes more attempts and zero if no melting occurs for this given straining time for all probed dsDNA molecules.

Figure 7.5 shows the melting efficiency of the pTR19-ASDS DNA construct for various straining times and ionic strengths. At 5 mM NaCl, a straining time of 1 second was sufficient to denature all probed DNA molecules. At the higher salt concentra-
7.4 Discussion

Here we have provided details and conditions for the force-induced generation of ssDNA templates. We demonstrated the practicality of this process for two different plasmid DNAs, both yielding ssDNA in the range of 10,000 base pairs or ~ 5 µm. This represents a convenient length scale for DNA manipulation with optical tweezers, to obtain a position accuracy in the nanometer range and below [89]. The use of plasmid DNA is convenient since transfecting *E. Coli* with the desired plasmid gives access to large quantities of DNA, required for the downstream labeling procedure. We showed that, subjected to a force of about 75 pN, within several seconds these
CHAPTER 7. SINGLE-STRANDED DNA TEMPLATES FOR OPTICAL TWEEZERS-EXPERIMENTS

DNA templates convert into ssDNA, irrespective of the details of the DNA sequence. This approach is successful over a broad range of salt concentrations, including physiological conditions. For increased salt concentrations, we found the melting time to increase moderately; from about 3 seconds at 50 mM NaCl to 5 seconds at 150 mM. This approach of generating ssDNA via force-induced melting possesses advantages compared to the use of chemicals like sodium hydroxide to destabilize DNA or exonuclease strand digestion. Most importantly, it allows for DNA denaturation without a buffer exchange procedure, which would involve the use of a complex microfluidics system. Care should be taken in assays involving reagents that substantionally stabilize dsDNA, like RecA/RAD51 filaments. Force-induced melting of dsDNA templates coated with RecA/RAD51 filaments will not be successful, due to their significant stability [106]. This obstruction can however be circumvented by denaturing DNA in the absence of DNA stabilizing reagents, and subsequently applying a buffer exchange.

In summary, we demonstrated that DNA template tension offers a distinct pathway for the generation of specific templates for single-molecule experiments. The protocol to generate ssDNA in an optical tweezers setup presented here will facilitate single-molecule experiments on large number of genomic transactions, like the assembly of RecA/RAD51 filaments, the stability and binding properties of single-stranded binding proteins and the dynamics of replicative DNA polymerases.
List of publications

   *Combining optical tweezers, single-molecule fluorescence microscopy, and microfluidics for studies of DNA-protein interactions.*

   *Unraveling the structure of DNA during overstretching by using multicolor, single-molecule fluorescence imaging.*

   *Twist, stretch and melt: quantifying how DNA complies to tension.*
   under Review at Nature Physics — Chapter 4.

   *Visualizing the balance between two DNA-overstretching modes: melting in the interior versus from the extremities.*
   Manuscript under preparation — Chapter 5.

   *Melting bubble statistics and dynamics of overstretched DNA.*
   Manuscript under preparation — Chapter 6.

   *Single-stranded DNA templates for optical tweezers-experiments.*
   Manuscript under preparation — Chapter 7.

∗ Joint first authors, † Joint last authors.


REFERENCES


REFERENCES


REFERENCES


REFERENCES


REFERENCES


REFERENCES


REFERENCES


REFERENCES


