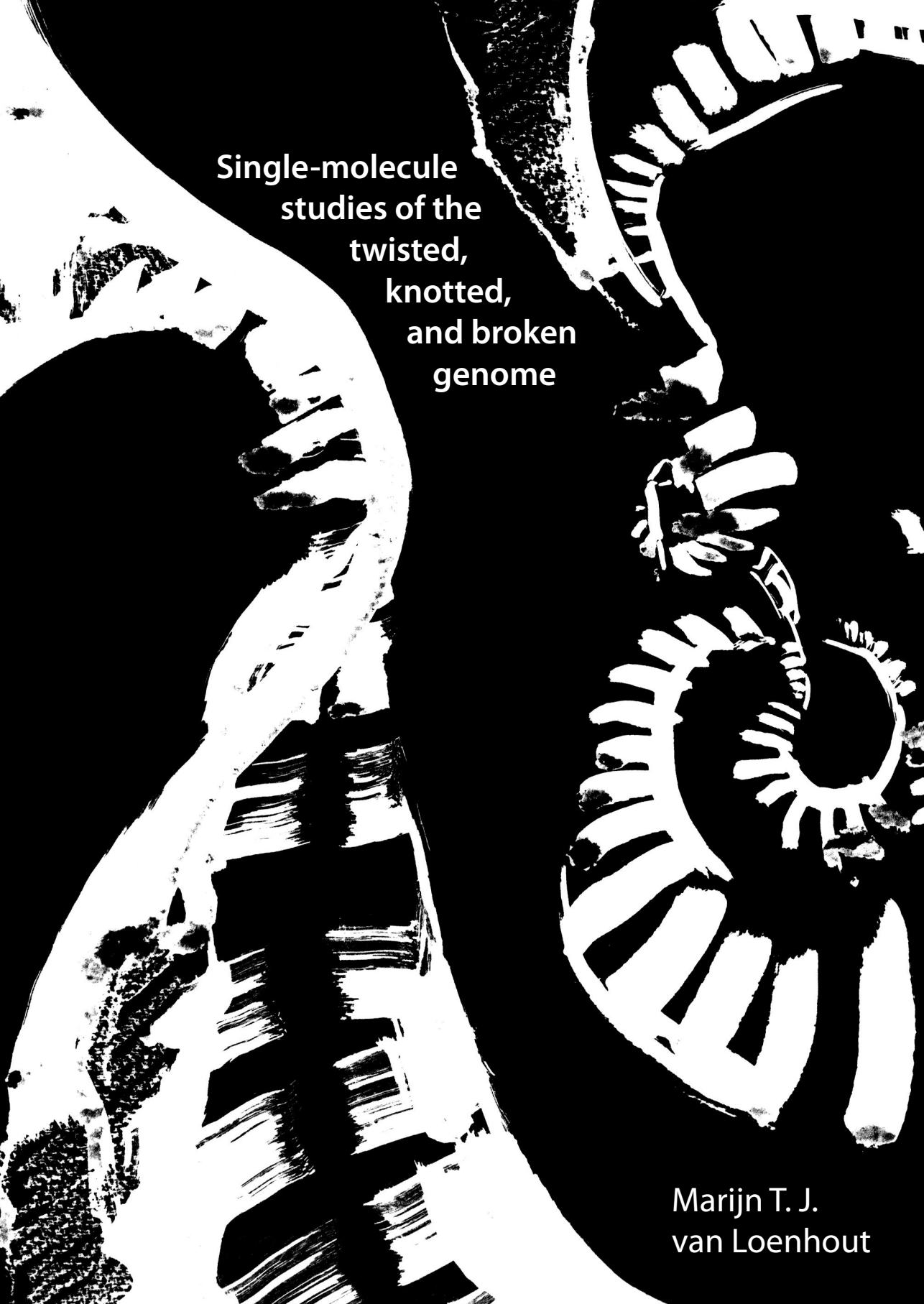


Single-molecule studies of the twisted, knotted, and broken genome

Marijn T. J. van Loenhout

Single-molecule studies of the twisted, knotted, and broken genome

Proefschrift

ter verkrijging van de graad van doctor
aan de Technische Universiteit Delft;
op gezag van de Rector Magnificus prof. ir. K.C.A.M. Luyben;
voorzitter van het College voor Promoties,
in het openbaar te verdedigen op maandag 5 november 2012 om 15:00 uur
door

Marinus Theodorus Johannes
VAN LOENHOUT

ingenieur in de technische natuurkunde
geboren te Nijmegen.

Dit proefschrift is goedgekeurd door de promotor:
Prof. dr. C. Dekker

Samenstelling van de promotiecommissie:

Rector Magnificus	Voorzitter
Prof. dr. C. Dekker	Technische Universiteit Delft, promotor
Prof. dr. C. L. Wyman,	Erasmus Universiteit, Rotterdam
Prof. dr. N. H. Dekker	Technische Universiteit Delft
Prof. dr. R. Bruinsma	University of California Los Angeles, VS
Dr. Ir. E. J. G. Peterman,	Vrije Universiteit Amsterdam
Dr. S. Neukirch	Université Pierre et Marie Curie, Paris, France
Dr. C. H. Häring	EMBL Heidelberg, Germany
Prof. dr. S.J. Tans	Technische Universiteit Delft, reservelid

Keywords: DNA, magnetic tweezers, optical tweezers, supercoiling, homologous recombination

Published by: Marijn van Loenhout
Cover image: Jacob Kerssemakers
Printed by: Uitgeverij BOXPress

The production of this thesis is financially supported by Delft University of Technology

An electronic version of this thesis is available at:
http://www.library.tudelft.nl/dissertations/

Casimir PhD series, Delft-Leiden 2012-28
ISBN: 978-90-8593-135-5

Preface

Biology presents perhaps the greatest unknown, but at the same time it is the closest by. Every day, cells use the laws of physics to do the most remarkable things. But they hide it well, and most processes are too small to even see with a microscope. As a biophysicist, the challenge lies in thinking out tricks and building some fancy instruments to figure out what is happening. But instead of neatly playing by the laws of physics, biology has its own quirks. Our largest problem: everything sticks. So we spent our days preventing the experimental samples from turning into one big bioblob and even when things do not stick, our relationship with these molecules tends to be fleeting. Like ephemera, they often disappear before we pull or shine light on them.

Is this then the reason it took years for this PhD to be finished? Well no. Simply put, I had a great time and wanted to finish it with a bang. Cees and Nynke turned MB into an entire Bionanoscience department, which, to a PhD student, is like the scientific ideal from back in the days when scientist still wore top hats: A collection of great people that know their science and their beers. Technicians that are so enthusiastic that they almost do the experiment for you. Travels to historically established universities like Heidelberg, Oxford, Barcelona, and Franeker. Talks by great scientists, almost removing the need to go to a mundane conference (unless it's in a sunny place of course). Plentiful resources available that make it possible to build and do whatever the experiment requires. But now it is time to move on, into the real world where there is a tough academic competition, where grants that include the word 'valorization' need to be written, and where a pinstripe suit is often more valued than fundamental knowledge or a nicely machined tool. I'm looking forward to it! But first I want to thank some people for the wonderful time I had.

First and foremost, I'd like to thank Cees, my promoter and supervisor. Your enthusiasm and excitement was inspiring. You always supported me and allowed me to choose my own research path. Thank you for that, it has been a great and sometimes challenging experience. You set up a whole new department with many interesting people, and I am very grateful to have been a part of that, meet talented people and have fun with them. Your involvement in BN also created a lot of freedom for me. I chose, for example, to create an algorithm that tracks the positions of beads. Didn't thousands of researcher try the same? Would I be able to make an improvement, would it be significant, and could I convince referees of that as well? Those were questions I didn't ask myself. I simply went ahead, maybe it was arrogance, lack of foresight, but in any case it I learned a lot on the way. That is perhaps also the nature of science, often unpredictable and opportunistic. Four years ago we tried to see plectonemes in a DNA molecule, but these experiments completely failed as all DNA molecules broke in two after only taking a single image. Three years later we had a new idea. I told Mark, a bachelor student, we're going to try this for two weeks, if it works you have a project, otherwise we'll have to think of something else. It turned out we had a project!

Cees, you enjoy the big and bold questions and stimulated me and Iwijn to build a machine that moves laser beams with the velocity of sound to handle tiny beads and molecules. All that effort was made to address this one question: How does a RecA filament search? It took over two years with many students and technicians involved and it worked out! Thank you Cees, it has been a great time.

If you want to study biology as a physicist, you better make sure you make some biologist friends. I've been fortunate enough to work with some of the best, from both Rotterdam and Heidelberg. I'll start in Rotterdam and I want to thank Claire and Roland for doing research together, great writing advice, and for their willingness to always share their in depth knowledge of DNA repair. As you say: "There is nothing special about single-molecule techniques, you just use these techniques when the question asks for them." Thanks for having us on the team. Tommy, it was great working and having a beer with you after starting our PhD's almost at the same time. You not only purified 'super' active RAD51 (is that supposed to be good?), but also had a real interest in tweezers and other physics stuff. Thanks for all the good times and the research we did together. Humberto, with your friendly optimism you fooled me with this tricky RAD54 protein; that turned out to be a nut too hard to crack in a magnetic tweezers instrument. Keep up the good work and the smiles. Eri, in the midst of many things happening in the field, you put some huge effort in MRN, I'm sure you'll put it together in the end. Indriati, all the best. I'll miss your cheerful voice asking me why I was still measuring these plectonemes every time you came to visit.

Heidelberg is a bit further away so fortunately Christian Häring visited us and talked about a protein that forms a ring to keep our chromosomes together, thereby enabling cells to divide and life to go on. I was fascinated. Would it be possible to

measure this? A single protein between two chromosomes? It seemed impossible and intriguing at the same time. One could (should?) ask oneself if this is a wise topic to pick in an academic world, where getting papers out fast is a must. Thanks to Cees' enthusiasm and some of the best collaborators I could have wished for (Christian and Sara), we took a big step, and with just a 'bit' more work we may really get know the secrets of cohesin. So how do you define a good collaborator? Let me give an example: I sent an email to Christian asking him about the status of an experiment and I received no response. So I asked him again, no response… One week later, I did get a message. "Sorry for the delay, we wanted to make sure everything works and had to do some additional experiments, we now understand everything and here are the results." Sara, it was great working with you and I always felt very welcome in Heidelberg. You are a great person and a super experimentalist, thanks.

So now on to you guys here in Delft. Thijn, thanks for introducing me in the single-molecule world. You showed me right from the start that it is not about the molecules, but about the people. You always had time for me and everyone alike. I enjoyed working with you, opinionated and truly interested in biology as you were meanwhile juggling apples and mink on the side. Rifka, first, perhaps an apology is appropriate for suggesting that one experiment that might just work… Your practical realism and my overoptimistic view of things did not quite match that first time, but I've grown to appreciate you a lot, thanks and good luck. Iwijn, many thanks, I don't think I could have had anyone better or more fun to work with. Watching the Tour de France, placing a mirror for the first time, building an impossible optical tweezers setup, good times! I learned a lot from how you pushed experiments to the limit and kept going when things were difficult. You organized an awesome group trip to Barcelona. On the bus ride back from the Pyrenees, sick and suffering from a hangover, competitive as you are, there was only one thing to do: empty the wallets of your colleagues during a game of poker. I knew your game by then; we split the cash. Hugo, it was short but good, and stem cells are the future, with that I fully agree. Gautam, thanks for being the friendly guy you are. Allard, speak out loud, otherwise everyone in this Bionanoscience department will forget what a resonance is. Dan, thanks for reading all my stuff and being my favourite procrastination. I'm still surprised you haven't hit me yet, after playing the most annoying Christmas songs when I was fed up with doing experiments, or dropping in every day to talk about the experiments *you* should do instead of writing my own thesis. You probably didn't hit me because you are 'proper' British. You like it that way and demand that scientists explain things properly, and if there is ever a factor of 2π missing, you will spot it. But be careful; you may not want a 'proper' job, but soon enough you'll realize you're a proper scientist.

I've had some great help from masters and bachelors student during my PhD and I want to thank them all: Kiona, confident and inseparable from Zeno, I'll remember your song. Zeno, next time you'll be swimming. Johan, you were part of the team getting those beads to move with a laser, good times. Edwin, you picked a hard project but did well, thanks. Stephanie, you learned a lot with a smile. Roeland, pulling on beads was never so chilled out. Benedetta, a lab journal full of pink flowers and hearts when the experiments worked, and matlab scripts with names I'd better not translate from Italian, thanks so much. Mark, you picked a project that could have failed after the first week, but it didn't and you shared the excitement. Dorine, equipment issues slowed you down but not your presentation, thanks. Jort, yes experiments can be hard, so enjoy the results.
I especially want to thank the supporting staff. I couldn't have done these experiments without them. First and foremost, I want to thank Suzan, you are a biolab genius and taught me a lot about molecular biology. You were involved in all the experiments described in this thesis, always friendly and thinking of new ways to make our crazy constructs. Bronwen, thanks for your help, you were always super quick, but take it from me, you need not apologize when a construct is not ready the next day,—try asking us for results in a day. Jelle, you make things in the workshop that are actually not possible, you do know that, do you? But not only that, every year you bring a room full of roses for all of us. These result in jealous looks from girls and women on the street and guys that remark: "So you must have something to make up for." Indeed I have. Jelle thanks for all your beautifully machined parts and keen interest. Dimitri, thanks, you built us a beautiful black box and a bunch of little amazingly crafted parts. Jaap, thanks for making all equipment run smoothly and making sure it does even if there is a power outing or anything else. Jacob, you are like a siren, luring me away from experiments with creative thoughts. Thank you for delaying my defense with at least 6 months with discussions on acousto-optic deflectors and bead tracking. I enjoyed it back then and still do. You are not only a scientist but also an artist, as can be seen by the wonderful drawings in this thesis that balance dark and light. Thank you, Elsemarieke, Liset, Leonie, Emmylou, Dijana, and Amanda, you make everything run smooth with a smile, thanks a lot.

Looking back, my roommates were very important during my PhD. They showed me how science is done, helped with any question I had, and just made it all fun. So thank you all, I've been very lucky. Let's start from the beginning: Marcel, straightforward as can be, you inspired me with your drive to do that one experiment and measure a single-molecule. I didn't beat you on equation length, but I did get close to your concept of '*total science*'. Iddo, not only were you way above my skill level at table football, you also made science seem easy and effortless. Igor, you showed that the impossible is possible; who in his sane mind would ever believe that a protein behaves vastly different in a glass tube compared to a plastic tube! I'm sorry we can't offer you a place in downtown San Francisco but maybe you can make it to Vancouver soon,

party on. Michiel, you didn't share an office with me, if we do it again, you should, thanks for the music. Sjoerd-Jan, you're a fun guy and it was great to see you do what you do best, physics, cleanroom stuff and now… coffee. Felix, thanks for making fun of us single-molecule DNA pullers, but remember you were once one yourself. You have become the graph master and convert swimming bacteria into art. But it's not about the nice pictures, you truly want to explain why bacteria rule the world, keep up the enthusiasm. Tim, it was fun having you around and seeing you explore Europe fully, but watch out when you go back they'll probably won't let a commy back in (and thanks to Meghan for all the wonderful cookies).
But there are off course more people in the lab that I won't forget. Jan, I'm not sure what I'll miss most: drinking a beer together, your unlimited knowledge about Stanford, Nobel laureates, and academic curiosities, or your thorough research and writing skills. Good luck, and make that Lipfert lab a science *magnet* and a good place for everyone. Calin, thanks for all that crazy shit you keep track of. David, thanks for impressing me with the most single-molecule data ever taken and always bringing back some goods stories after a night of drinking. Greg, chemists must be special, you managed to reincarnate Stefan and made Juan believe his hair really turned orange, keep up the magic. Bertus, Chirlmin, Christophe, thanks for your critical reading and best of luck expanding your labs. And then there are a few more who I will not forget, thank you all: Laiko, Tom, Koen, Stefan, Fabai, Maarten, Bojk, Simon, Helmut, Marc, Serge D., Serge L., Ilja, Jetty, Katja, Xander, Francesco, Marie-Eve, Elio, Ann, Christine, Jaan, Zohreh, Matthew, Andrew, Peter, Juan, Charl, Magnus, Ralph, Martin, Edgar, Adam, Pradyumna, and Bernhard.

When Cees mentioned in a meeting we need a replacement for Marijn there was some laughter. But sure enough, a few weeks later, a guy named Marijn showed up. That didn't work of course so if it's not in the name, perhaps it's in stars. Tessa, you picked the right birthday and also figured out how it works within a week: Be 100% focused with these tempered single-molecules, because 99% of a molecule is nothing more than a broken molecule. I'm sure you'll do well, but as much as the results, enjoy how you get there.

There is of course also life outside of the lab and I want to thank my friends for all the good times. Doing a PhD meant that I now thought I had a proper excuse to ignore your emails, but it seems that I have missed out on some great trips, parties and bike rides. However, since I'm super fit now again from commuting to Gouda by bike, who wants to go for a ride, or a beer? I'm very happy to have two of you, Diederik and Guus, as my paranimfen, and I want to thank you all for being true friends.

Family is special, what would I be without them? Especially after forgetting my bag with Cees' last comments on this thesis in a train headed for Groningen. A quick call to Jasper, and he picked it up from the train two hours later, added a kind smile to Nelleke and my bag with the comments was safe at home again the next day, thank you both. Thanks to my parents, Jos and Joke, who carefully read every

word of this thesis and found surprisingly little spelling mistakes for a son that was 'seriously' challenged with spelling problems at elementary school. Thank you for you unconditional and continuous support.

Finally I want to thank Nelleke, you showed me how it is to have a real job and you supported me with my 'hobby' for these past years. You challenged the yellow monsters of the Dutch railroads day after day, year after year, while I kept saying: "the thesis will be done soon…", well here it finally is. You taught me to enjoy and live the here and now, instead of only dreaming about a future that always moves away. I look forward to our future, and to today, to make it exciting and simply pleasant every day.

Marijn van Loenhout, October 2012

Contents

Part 1—Twist

Part 2—Knots

Part 3—Breaks

CHAPTER 1

Introduction

Our genome contains our hereditary information stored in long DNA polymers. This information is encoded by the sequence of two polynucleotide chains that together form the DNA double helix. The mechanical properties and thread-like nature of the DNA influence almost all DNA processes and the proteins that interact with it. Each of our cells, only tens of micrometers in size, has to handle approximately 2 meters of DNA. This is a huge task as DNA must be replicated and segregated for cell division, transcribed and translated into proteins, recombined to generate diversity, and repaired if damage occurs. DNA and the proteins that interact with it have evolved into a remarkably fine-tuned system that uses its physical properties to handle and process the information encoded in the genome.

As first noted by Watson and Crick[1], the double helix structure of DNA enables efficient replication of the DNA by separating the strands and synthesizing two new strands complementary to each of the original strands. The double helix must be unwound to separate the strands, resulting in waves of torsion that transmit from the replication fork. This additional twisting or 'supercoiling' of the DNA affects it in two ways. First, it may destabilize the double helix and form special structures which proteins may recognize and interact with. Second, it can form intertwined loops known as plectonemes which change the conformation of the genome and bring distant regions together. In part 1 of this thesis we describe measurements that reveal the dynamics and structure of supercoiled DNA.

Once replication is complete, the DNA must be segregated between the daughter cells and each must receive exactly one copy of each chromosome. If the DNA was simply a random coil, the chromosomes would get entangled making it sheer impossible to separate them. Cells therefore condense replicated chromosomes and link them together resulting in the familiar X-shaped pair of chromatids. In part 2 of this thesis we address the mechanical properties of cohesin which provides this link between the two chromatids.

The stability of our genome is under constant threat by a plethora of processes and agents that may damage or modify our DNA. These can be both endogenous, e.g. replication errors, mechanical stress, free radicals, as well as exogenous, e.g.

ultraviolet light, X-rays, or mutagenic chemicals. A variety of lesions can occur including UV-photoproducts, oxidized bases, base mismatches, insertions, deletions, intra- and interstrand crosslinks, and single- and double-stranded breaks. To repair these lesions cells have evolved a range of repair pathways. Double-stranded breaks are among the most severe lesions and may be repaired by either non-homologous end-joining or homologous recombination. In the part 3 of this thesis, we discuss key processes and proteins involved in the repair of double-stranded breaks by homologous recombination.

To understand the molecular mechanisms and underlying processes of the genome a number of sensitive techniques have been developed over the past decades, e.g. optical and magnetic tweezers, atomic force microscopy, and single-molecule fluorescence. These techniques visualize and probe individual molecules and are aptly named single-molecule techniques. In this thesis we develop and use such techniques to elucidate the biophysical principles that underlie biology at the nanoscale. First, we briefly review these techniques.

Magnetic tweezers

Just like a compass needle responds and orients to the magnetic field of the earth small magnetic particles may be manipulated with an external magnet. Crick & Huges[2] first used this technique to probe the viscoelastic properties of the cytoplasm, thereby providing the first demonstration of magnetic tweezers. Later this technique was extended by Smith *et al.* and Strick *et al.* to stretch[3] and coil[4] individual DNA molecules tethered between a magnetic bead and a flow cell surface. These first experiments form the basis of what now has become a widely used technique known as single-molecule magnetic tweezers.

Magnetic tweezers typically tether a DNA molecule of several kilobases, microns in length, between a micron-sized superparamagnetic bead and the surface of a glass flow cell (Figure 1.1A). A pair of magnets positioned above the flow cell is used to exert a force on the DNA tether, which ranges from several femtoNewtons to several tens of picoNewtons. Analysis of the Brownian fluctuations of the bead can be used to calculate the applied stretching force. Rotating the magnets will also rotate the bead, and in this way it is possible to coil the DNA molecule. The bead position is monitored by video microscopy with a typical resolution of a only few nanometer in three dimensions. Recent improvements in digital camera technology have further extended the capabilities of magnetic tweezers and even smaller displacements may now be detected at frame rates in the kHz range[5]. The higher number of pixels in modern cameras allows to monitor multiple DNA tethers simultaneously, thereby greatly increasing the throughput of single-molecule experiments. Magnetic tweezers are typically used to monitor length changes of the DNA tether. These provide information on the DNA itself and on protein-induced changes in the physical properties of the DNA. In this thesis we additionally use the sideways displacements of the magnetic

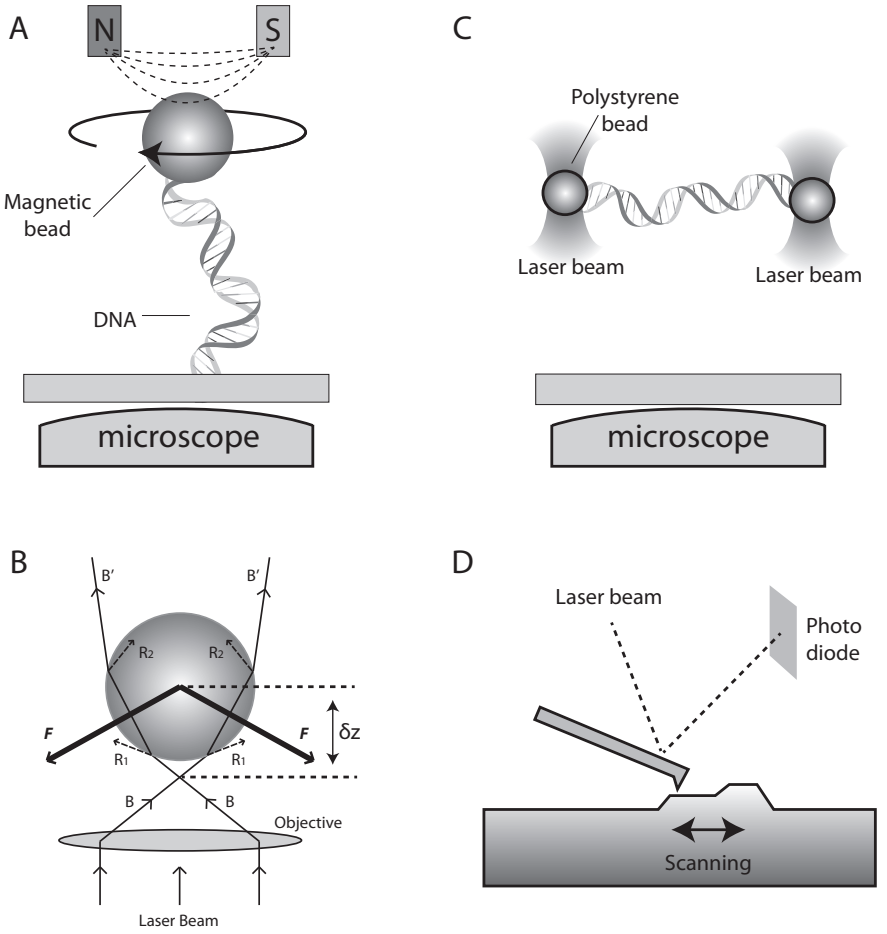


Figure 1.1: Single-molecule techniques. (**A**) In a magnetic tweezers setup, a DNA molecule is tethered between the surface of a flow cell and a magnetic bead. A pair of external magnets is used to stretch and twist the molecule. (**B**) Optical tweezers hold a polystyrene bead in the focus of a laser beam and can manipulate the bead in three dimensions. By attaching a DNA molecule between two beads held with optical tweezers, it is possible to measure its length and exert forces on it. (**C**) Working principle of optical tweezers. A micron-sized bead experiences a force directed towards the center of the optical trap due to the refraction of the light at its surface. Adapted from Ref. 6. (**D**) Schematic illustration of the working principle of an AFM.

bead as a sensitive probe for the interactions between two DNA molecules. We also combine fluorescence microscopy with magnetic tweezers to directly visualize the structure and dynamics of a supercoiled DNA tether.

Optical tweezers

In 1968, Arthur Askin demonstrated that a strongly focused laser is capable of holding small dielectric particles stably in three dimensions[6]. These so-called optical tweezers can move particles around and exert picoNewton forces on them. The working principle of optical tweezers is based on refraction (Figure 1.1C). A strongly focused laser beam will refract at the interface between a dielectric particle and the surrounding medium. Since light has momentum such a change in direction will result in a force that is directed to the center of the focus, thereby effectively 'trapping' the particle. For small displacement from the trap center, the force increases linearly, i.e. $F = k_{trap}\delta z$, where k_{trap} is the trap stiffness and δz the displacement from the trap center. Optical tweezers have found many applications in the biological sciences and have, for example, been used to reveal the individual steps of molecular motors moving along DNA molecules or microtubules[7-10]. In these experiments individual DNA molecules or motor proteins are attached to micron-sized beads which may then be manipulated and monitored with the laser beam. In this thesis, we use optical tweezers to pull on and manipulate DNA molecules and measure the interactions between them (Figure 1.1B). The main advantage compared to magnetic tweezers was the ability to move individual beads in three dimensions allowing us to bring DNA molecules together and even loop one molecule around another.

Atomic force microscopy

In 1981, Binnig and Rohrer invented the scanning tunneling microscope (STM)[11], which was the first of a family of scanning probe microscopes that are able to image surfaces at the nanometer scale by scanning a fine probe over it. Scanning microscopes are not limited by the wavelength of light or electrons, allowing them to resolve even single atoms. In 1986, the atomic force microscope (AFM) was developed by Binnig, Quate and Gerber[12] which has become one of the most widely used tools for imaging at the nanoscale. An AFM uses a flexible cantilever with a sharp tip that is brought into close proximity of the sample surface (Figure 1.1D). Forces between the tip and the sample lead to a deflection of the cantilever. When the tip is scanned over a surface, height differences will result in deflection changes of the cantilever, similar to a how a gramophone reads a record. The cantilever deflection is measured by a laser spot that is reflected from the top of the cantilever onto a four-quadrant photodiode. An AFM can image even while immersed in liquid, making it especially suited to image biological samples in buffer. DNA and proteins are typically imaged

by immobilizing them via electrostatic interactions on an atomically flat mica surface. The high resolution of the AFM makes it possible to visualize the detailed structure and conformation of individual DNA molecules and proteins.

DNA structure and physical properties

The structure of DNA is key to understanding how it works. The double helix contains two long polymers chains that are composed of three components: phosphates, sugars, and bases (Figure 1.2). A regular alteration of phosphates and sugars forms the backbone of the DNA molecule running on the outside of the double helix. The phosphates are negatively charged and DNA molecules therefore repel each other and can almost frictionless slide along each other. Proteins that interact with DNA often have a positive charge, helping them bind DNA via electrostatic interactions. The bases, which are not readily soluble in water, are stacked and tucked away on the inside of the double helix limiting their exposure to the aqueous surroundings. A monomer unit of a single DNA strand thus consists of a phosphate, sugar, and base and is called a nucleotide.

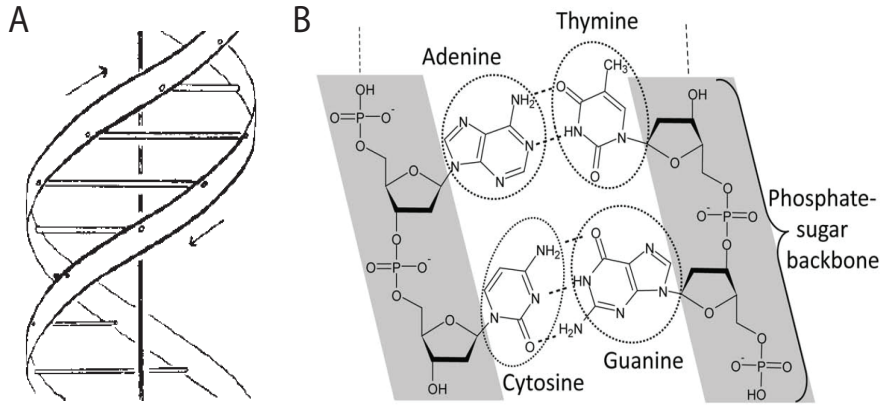


Figure 1.2: Structure of DNA. (**A**) Schematic representation of the double helix, illustration by Watson and Crick[1]. (**B**) Chemical structure of A-T and C-G base-pairs in DNA. The gray shaded area indicates the sugar-phosphate backbone image from Ref. 35.

There are four different types of nucleotides in DNA: adenine, thymine, cytosine, and guanine. The two chains of the DNA double helix run in opposite direction and are held together by the interactions between these nucleotides. Adenine always pairs with thymine via two hydrogen bonds, and cytosine pairs with guanine via three hydrogen bonds. Due to this base pairing, both strands are each other's complement. This base pairing also holds the key to understanding how cells copy their genetic information. Cells do this by enzymes that simply spit the strands and then use each

of the old strands as a template for synthesizing a complementary new strand. The sequence of the nucleotides forms the genetic code and certain stretches called genes may be transcribed into proteins. But there is more to the DNA sequences: it also encodes where proteins bind. Where nucleosomes, small proteins spools that wrap DNA around them, localize, and where special DNA structures such as DNA bends are likely to occur.

To understand how DNA behaves on larger distances, it is important to consider its mechanical properties. DNA molecules are generally very long and may be up to centimeters in length, but the helix is only 2 nanometer wide, making them truly very long slender threads. There are about 10 nucleotides per turn of the helix and each base pair takes up 0.34 nanometer along the chain. The DNA molecules used for the experiments described in this thesis are several to tens of kilobase pairs in length corresponding to a length of several micrometers. DNA resists bending on short distances, but on larger distances it is constantly perturbed by the thermal fluctuations of the surrounding liquid molecules. Due to these random thermal fluctuations a DNA molecule will coil up in a random coil that is much smaller than when it is stretched out to its full contour length. The thermal fluctuations set the typical length scale on which a DNA molecule may be considered flexible. This distance is called the persistence length and it is approximately equal to 50 nanometer or 150 base pairs. Bending a DNA molecule on distance much smaller than the persistence length will cost considerable energy, whereas bending it for longer lengths happens spontaneously.

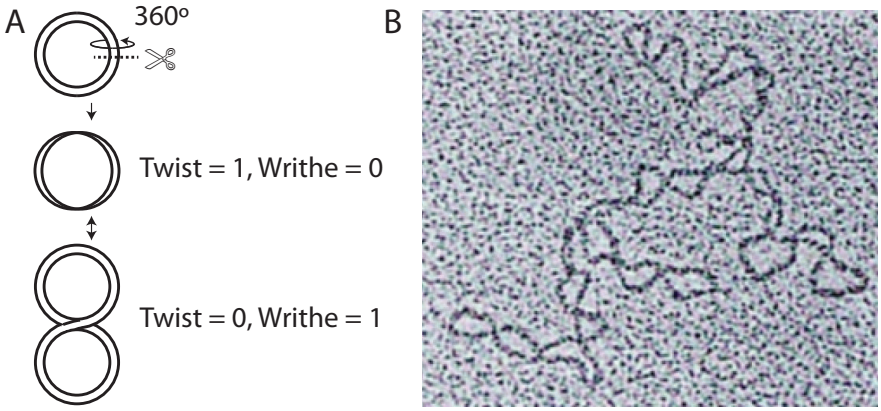


Figure 1.3: Supercoiling of DNA. (**A**) a circular closed plasmid can be torsionally relaxed (top image) or supercoiled leading to twist (middle image) or writhe (bottom image). (**B**) Electron micrograph of a supercoiled plasmid from *E .coli,* image from Ref. 14.

Because DNA is helical, like a screw, any polymerase that moves along the DNA must either rotate around the DNA or it will rotate the DNA. Experiments have shown evidence for the latter, i.e., often these polymerases remain stable and the DNA itself rotates[13]. If the DNA would be very short its free ends could simply spin. However, DNA

molecules are very long and often attached to proteins and larger structures. There is therefore a buildup of torque in the DNA which becomes supercoiled, forming intertwined DNA loops called plectonemes. Let's consider what happens to a circular DNA molecule to understand the effects of torque (Figure 1.3A). In the absence of torque such a molecule will be a relaxed circle. However, if we would cut it open, rotate the ends one time, and close it again, there will be a torque in the molecule. This can have two effects (I) the DNA helix will become homogeneously over- or underwound, which is called twist, or (II) the DNA axis will cross or loop over itself, which is called writhe. Together twist and writhe are referred to as supercoiling. In living cells, DNA is supercoiled as may be seen in images of plasmids from *E. coli*[14] (Figure 1.3B). Supercoiling has two effects, first, twist may change the structure or destabilize the double helix, making the bases more accessible to proteins. Second, writhe can bring distant DNA regions together, thereby facilitating recombination processes. It is therefore important to understand supercoiling and its dynamics as it influences nearly all cellular processes involving the genome[15].

Cohesin, keeping it together during cell division

Cell division presents perhaps the most dramatic and fascinating event in the life cycle of a cell. Cells must not only copy their genome but also properly segregate their chromosomes and ensure that each daughter cell receives exactly one copy. Key to this process is the protein complex cohesin, that links replicated chromosomes (called sister chromatids) together[16]. In 1882, Water Flemming was one of the first to observe cell division in detail by staining dividing salamander cells[17]. He named this process 'mitosis' after the Ancient Greek μίτος meaning thread (Figure 1.4). In the dividing cells, he observed thread-like structures which he called chromatin, after their ability to be stained with dyes. Mitosis takes up only a short time span in the cell cycle and cells spend most of the time in interphase where they grow and replicate their genome. Below we briefly discuss the different stages of mitosis.

Once the genome has been replicated, mitosis starts with prophase. In this phase, cells condense their chromatin into discrete chromosomes. These replicated chromosomes consist of two copies called chromatids, bound together at the centromere by the cohesin protein complex. At the same time, microtubules, rod-like polymers that form the cytoskeleton, nucleate from two centrosomes, creating the mitotic spindle. In the next step, prometaphase, the nuclear membrane disintegrates and kinetochores form. Kinetochores are proteins structures that connect microtubules to the centromere of each chromatid. In metaphase, the microtubules attach to the kinetochores and they begin to pull each of the chromatids to one side of the cell. Cohesin must withstand these forces of the pulling mitotic spindle to prevent the chromatids from separating prematurely. The cell must now ensure that both daughter cells receive one of each of the chromatids. This is achieved by a tug-of-war of the microtubules attached to the kinetochores. These forces align the chromatids in the center of the cell

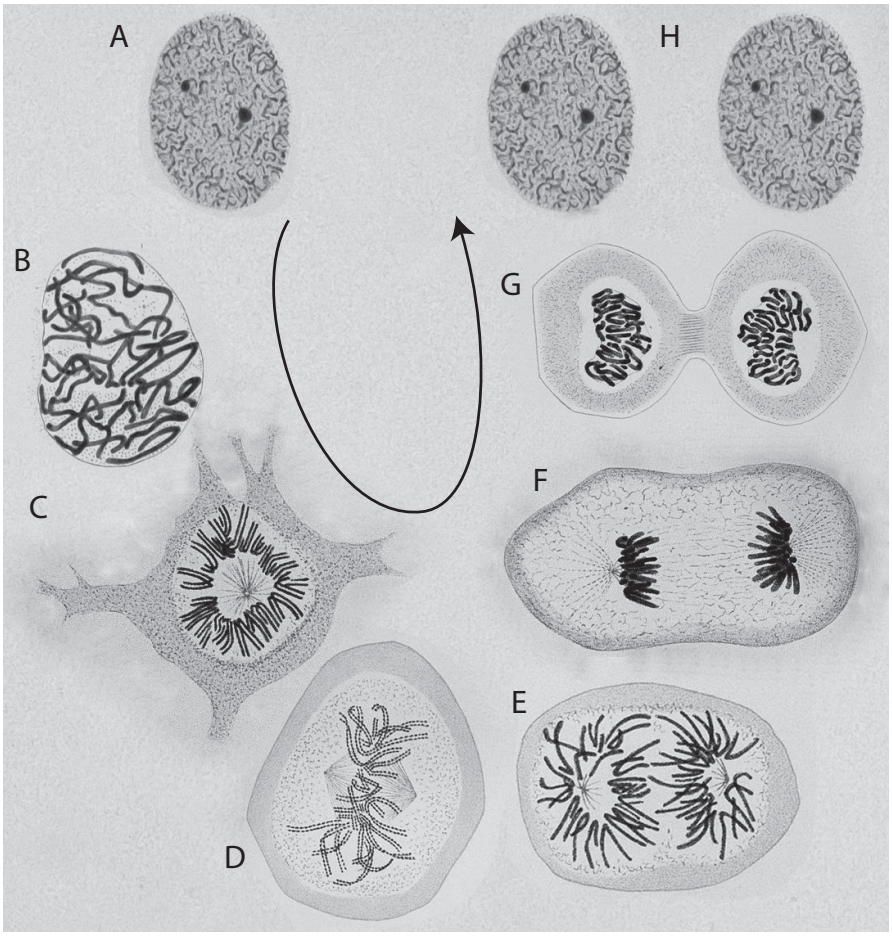


Figure 1.4: Mitosis as drawn by Walter Flemming[17]. (**A**) Interpase, (**B**) Prophase, (**C**) Prometaphase, (**D**) Metaphase, (**E**) Anaphase, (**F**) Telophase, (**G**) Cytokinesis, (**H**) Two new daughter cells.

in a configuration called bi-orientation. Cells verify proper bi-orientation by checking if there is centromere tension at each chromosome, i.e. microtubules are pulling from both sides. If there is no tension the spindle checkpoint is activated and the cell halts cell division. If all chromosomes are finally aligned, the cell enters anaphase. A protein, appropriately named separase, cleaves the cohesin complex separating the sister chromatids that now become the daughter chromosomes. The separated chromosomes are pulled apart by the kinetochore microtubules to opposing sides of the cell and the centrosomes move further toward the ends of the cell by elongating polar microtubules. In telophase, finally the cell elongates even more, new nuclear

membranes form, and the chromosomes start to decondense back into chromatin. To complete cell division, cytokinesis takes place in which the two daughter cells fully separate by pinching of their cell membrane in the middle.

Homologous recombination and double-stranded breaks

Double-stranded breaks are highly toxic lesion that may lead to mutations in genes, translocations, and loss of heterozygosity. If double-stranded breaks remain unrepaired, these can result in cell dysfunctions, cell death, or even lead to cancer[18]. Double-stranded breaks arise most frequently during replication as other lesions that may be present in the DNA become double-stranded breaks when the replication fork encounters them[19,20]. Double-stranded breaks may be repaired by non-homologous end joining, or by homologous recombination. We focus on the latter here. Homologous recombination (Figure 1.5) may be separated into three major steps: (i) pre-synapsis, where the break is recognized, a single-stranded DNA overhang is created, and a nucleoproteins filament forms, synapsis; (ii) where the search for homology takes place and, once found, a joint molecule is formed, and post-synapsis; (iii) where the broken strands are repaired and the intermediate DNA structure is resolved. Below we discuss the process in an eukaryotic cell step by step.

In pre-synapsis, the first proteins to arrive at the break are MRE11, RAD51, and NBS1 (the MRN complex)[21]. This protein complex signals the break and tethers the broken ends, preventing them from diffusing away[22,23]. Next a single-stranded DNA overhang is created. Although MRN has exonucleolytic activity, i.e. it can remove nucleotides from a DNA end, it can only do so in the wrong direction, MRN does use its nuclease activity to clean up the DNA ends in case of additional damage or stuck proteins[24]. There are several nucleases that may perform the resectioning to create the single-stranded overhand: CtIP interacts with MRN and promotes initial resection[25]. Subsequently a nuclease with a 5' to 3' activity (likely of the RecQ family) provides processive end-resectioning to form a long single-stranded overhang[26]. This single-stranded DNA will be quickly covered by the replication protein A (RPA) that binds with high affinity.

In synapsis, a recombinase protein RAD51 (or RecA in *E. coli*) polymerizes on the single-stranded DNA, displacing RPA, to form a helical filament. Initially a small RAD51 cluster nucleates on the DNA, which subsequently extends[27]. In this RAD51 filament, the single-stranded DNA is extended 1.5 times compared to the regular B-form DNA structure[28,29]. The extension is non-uniform and base triplets are formed with an approximately B-form DNA spacing and large gap in between. The RAD51 filament formation process is tightly regulated and may be enhanced by recombination mediators such as BRCA2[30,31]. Once formed, the RAD51 filament searches the double-strand template for a homologous sequence. This remarkable search process, which

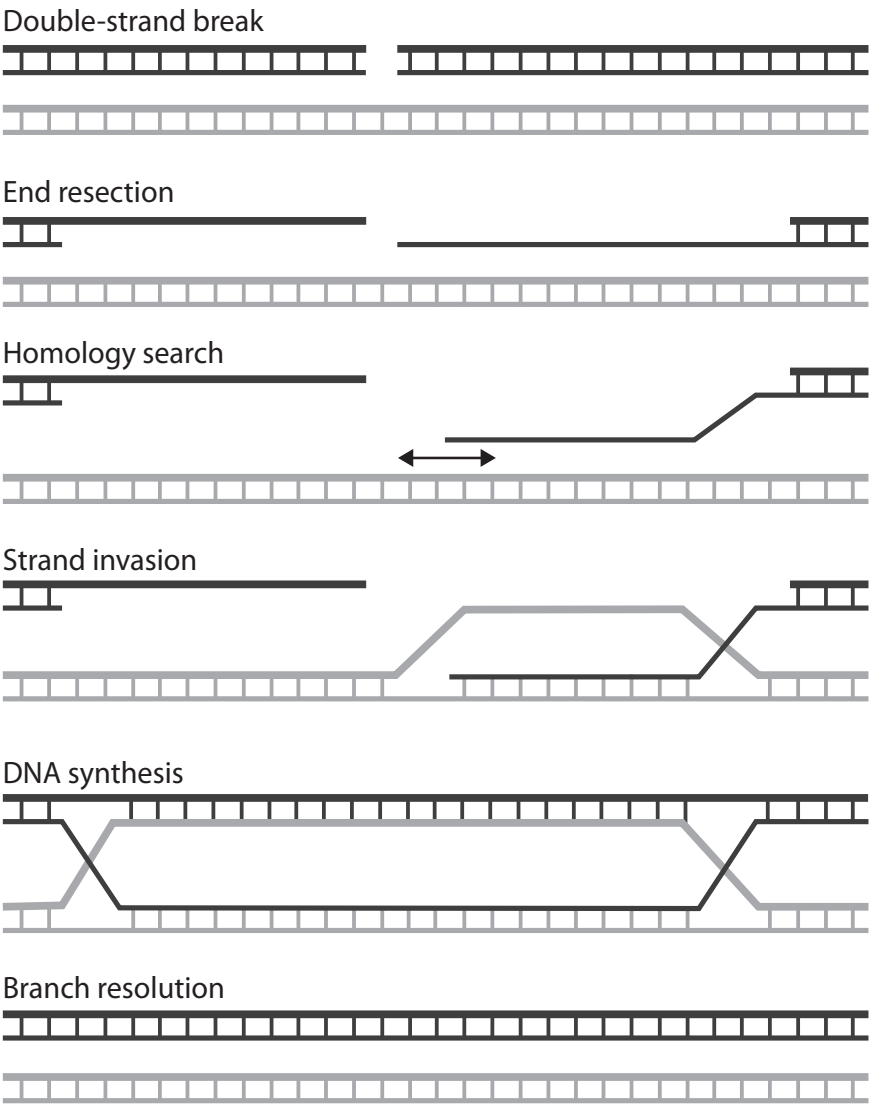


Figure 1.5: Double-strand-break repair by homologous recombination. After a double-strand break is detected, the ends are resected to create single-stranded overhangs. On the single-stranded DNA, a nucleoprotein filament forms that searches for homology. Once homology is found, the nucleoproteins filament invades the template strand, creating a joint molecule. Subsequently DNA synthesis takes place to extend the single-stranded DNA and recover the lost bases. Finally the branched structure is resolved to obtain two intact molecules without the loss of any genetic information.

may involve searching the 3 billion bases pairs of the human genome, is strongly enhanced if the template DNA is negatively supercoiled[32,33]. The search process itself requires no energy but relies on thermal fluctuations to search through the genome. The conformation of the 1.5 times extended nucleoprotein filament is conserved in all three kingdoms of life, which suggest that it is essential in the recombination process. Possibly the stretched conformation enables the high-fidelity thermodynamic search, i.e. not requiring ATP hydrolysis, by modulating the binding energy between the nucleoprotein filament and the double-stranded DNA template. Once homology is found, the RAD51 filament catalyzes strand invasion and exchange, leading to the formation of a joint molecule.

In post-synapsis, the RAD51 needs to disassemble so that polymerases can extend the invading strand to recover the sequence lost in the break. There are several different pathways by which the break can be resolved[34]. If there are two DNA ends present, the main pathways are synthesis-dependent strand annealing and double-strand break repair (Figure 1.5).

Outline of this thesis

This dissertation describes experimental techniques and measurements aimed at understanding the physical properties of supercoiled DNA, the proteins that structure DNA, and the mechanisms of DNA repair by homologous recombination. The work is ultimately motivated by the goal to understand the fundamental mechanistic processes that interact with and shape the genome. To do so, not only experiments were performed, but also sensitive measurement techniques were developed that allow us to directly probe single DNA molecules and proteins to obtain information on their conformations and interactions. This thesis is divided in four parts: Tools, Twist, Knots, and Breaks. Although only the 'Tools' section directly refers to the experimental methods, all other sections required the development of new techniques. The outline of this thesis is as follows:

In **Chapter 2** we introduce a new three-dimensional tracking routine for non-diffraction-limited particles that enables tracking with nanometer accuracy in a wide field of view thereby enabling a substantial increase the number of beads, and thus the number of molecules, that can be observed. The increased performance of the proposed algorithm uniquely enables to extract accurate data for the stiffness (persistence length) and end-to-end distance of more than 100 DNA tethers in a single experiment. A user-friendly software implementation of the algorithm was made available and this is now used by a number of other researchers.

In **Chapter 3** we present images that for the first time directly show the dynamics of individual plectonemes. We observe that multiple plectonemes can be present and that their number depends on applied stretching force and ionic strength.

Plectonemes move along DNA by diffusion or by a fast hopping process that facilitates very rapid long-range plectoneme displacement by nucleating a new plectoneme at a distant position. The hopping process presents a not previously seen mode of movement that allows long-distance reorganization of the conformation of the genome on a millisecond time scale.

We then move on to investigate DNA entanglements. In **Chapter 4** we describe how a novel combined magnetic and optical tweezers apparatus allows us to create a DNA loop that enables sensitive label-free detection of the positions and binding strength of proteins localized on DNA. The DNA loop acts as a scanning probe that can determine the relative distances between DNA-bound EcoRI proteins with 14 ± 8 nm resolution along a λ DNA molecule. We find a consistent offset between back and forwards scans of 35 ± 15 nm for the detected protein positions, which corresponds to the size of the DNA loop and is in agreement with theoretical estimates.

Next, in **Chapter 5** we discuss an experimental assay to measure the mechanical strength of a single cohesin protein complex. The mechanical strength and stability of cohesin is directly related to its function during cell division. Cohesin entraps sister chromatids and withstands the forces that align them by pulling them towards opposing spindle poles. The isolation procedure for cohesin-concatenated circular minichromosomes resulted in a high yield of minichromosomes in a near-native state, i.e. DNA-bound proteins and nucleosomes were still present. Initial single-molecule force spectroscopy experiments are presented as well as micrographs of concatenated chromosomes obtained by atomic force microscopy.

In the final part of this thesis we present experiments aimed at obtaining a mechanistic insight in the interactions of proteins and DNA during several stages of homologous recombination. In **Chapter 6** we describe how the MRE11-RAD50-NBS1 (MRN) complex, which recognizes, signals, and processes double-strand breaks is affected by nucleotide cofactors. Electrophoretic mobility shift assays and atomic force microscopy reveal that DNA is released from the complex upon nucleotide binding and hydrolysis. DNA release from the complex triggers a large conformational change with the appearance of an open arrangement of the RAD50 coiled coils. These results support a regulatory role for nucleotide binding, which not only affects affinity for DNA, but may also trigger changes in MRN intercomplex interactions.

We then go on to directly probe the structure and kinetics of RecA interactions with its biologically most relevant substrate, long single-stranded DNA molecules (**Chapter 7**). RecA-single-stranded DNA filaments are the catalytic core of homologous recombination in *E. coli,* and perform the search for homology and the subsequent strand invasion and exchange. We find that RecA ATPase activity is required for the formation of long continuous filaments on single-stranded DNA. The disassembly rates of RecA from single-stranded DNA decrease with applied stretching force, corresponding to a mechanism where protein-induced stretching of the DNA aids in the

disassembly. Finally, we show that RecA-single-stranded-DNA filaments can reversibly interconvert between an extended, ATP-bound, and a compressed, ADP-bound state.

In **Chapter 8** we analyze the effect of the BRCA2 CTRD domain on RAD51 filaments by an ensemble of single-molecule techniques. RAD51 is the human homolog of RecA and forms a similar helical filament on single-stranded DNA. The C-terminal RAD51 interaction domain (CTRD) of the breast-cancer tumor suppressor BRCA2 directly interacts with RAD51. We show that at high concentration, the CTRD entangles RAD51 filaments and reduces RAD51 filament formation in a concentration-dependent manner. We conclude that the CTRD does not reduce RAD51 dissociation kinetics, but instead hinders filament formation on double-stranded DNA. The CTRD mode of action is most likely sequestration of multiple RAD51 molecules, thereby rendering them inactive for filament formation on double-stranded DNA.

Finally, in **Chapter 9** we present experiments that address the remarkable mechanism of homology recognition by RecA. We built an unique combined optical and magnetic tweezers apparatus to directly probe the two-molecule interactions between a DNA molecule and a DNA-RecA filament. We find that recognition requires opening of the helix and is strongly promoted by torsional unwinding stress. Recognition is achieved upon binding of both strands of the incoming dsDNA to each of two ssDNA-binding sites in the filament. The data indicate a physical picture for homology recognition in which the fidelity of the search process is governed by the distance between the DNA binding sites.

References

1 Watson, J. D. & Crick, F. H. C. Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. *Nature* **171**, 737-738, (1953).

2 Crick, F. H. C. & Hughes, A. F. W. The Physical Properties of Cytoplasm - a Study by Means of the Magnetic Particle Method .1. Experimental. *Exp. Cell Res.* **1**, 37-80, (1950).

3 Smith, S. B., Finzi, L. & Bustamante, C. Direct mechanical measurements of the elasticity of single DNA-molecules by using magnetic beads. *Science* **258**, 1122-1126, (1992).

4 Strick, T. R., Allemand, J. F., Bensimon, D., Bensimon, A. & Croquette, V. The elasticity of a single supercoiled DNA molecule. *Science* **271**, 1835-1837, (1996).

5 Otto, O. *et al.* Real-time particle tracking at 10,000 fps using optical fiber illumination. *Opt. Express* **18**, 22722-22733, (2010).

6 Ashkin, A., Dziedzic, J. M., Bjorkholm, J. E. & Chu, S. Observation of a single-beam gradient force optical trap for dielectric particles. *Opt. Lett* **11**, 288, (1986).

7 Greenleaf, W. J., Woodside, M. T. & Block, S. M. High-resolution, single-molecule measurements of biomolecular motion. *Annu. Rev. Biophys. Biomol. Struct.* **36**, 171-190, (2007).

8 Svoboda, K., Schmidt, C. F., Schnapp, B. J. & Block, S. M. Direct observation of kinesin stepping by optical trapping interferometry. *Nature* **365**, 721-727, (1993).

9 Yin, H. *et al.* Transcription against an applied force. *Science* **270**, 1653-1657, (1995).

10 Smith, D. E. *et al.* The bacteriophage phi 29 portal motor can package DNA against a large internal force. *Nature* **413**, 748-752, (2001).

11 Binnig, G. & Rohrer, H. Scanning Tunneling Microscopy. *Helv. Phys. Acta* **55**, 726-735, (1982).

12 Binnig, G., Quate, C. F. & Gerber, C. Atomic Force Microscope. *Phys. Rev. Lett.* **56**, 930-933, (1986).

13 Liu, L. F. & Wang, J. C. Supercoiling of the DNA template during transcription. *Proc. Natl. Acad. Sci. USA* **84**, 7024-7027, (1987).

14 Boles, T. C., White, J. H. & Cozzarelli, N. R. Structure of plectonemically supercoiled DNA. *J. Mol. Biol.* **213**, 931-951, (1990).

15 Roca, J. The torsional state of DNA within the chromosome. *Chromosoma* **120**, 323-334, (2011).

16 Uhlmann, F., Lottspeich, F. & Nasmyth, K. Sister-chromatid separation at anaphase onset is promoted by cleavage of the cohesin subunit Scc1. *Nature* **400**, 37-42, (1999).

17 Flemming, W. *Zellsubstanz, Kern und Zelltheilung*. (Leipzig: Verlag von F.C.W. Vogel, 1882).

18 Hoeijmakers, J. H. Genome maintenance mechanisms for preventing cancer. *Nature* **411**, 366-374, (2001).

19 Wyman, C. & Kanaar, R. DNA double-strand break repair: all's well that ends well. *Annu. Rev. Genet.* **40**, 363-383, (2006).

20 Cox, M. M. *et al.* The importance of repairing stalled replication forks. *Nature* **404**, 37-41, (2000).

21 Lisby, M., Barlow, J. H., Burgess, R. C. & Rothstein, R. Choreography of the DNA damage response: Spatiotemporal relationships among checkpoint and repair proteins. *Cell* **118**, 699-713, (2004).

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

23 D'Amours, D. & Jackson, S. P. The Mre11 complex: at the crossroads of dna repair and checkpoint signalling. *Nat. Rev. Mol. Cell Biol.* **3**, 317-327, (2002).

24 Kinoshita, E., van der Linden, E., Sanchez, H. & Wyman, C. RAD50, an SMC family member with multiple roles in DNA break repair: how does ATP affect function? *Chromosome Res.* **17**, 277-288, (2009).

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

26 Bernstein, K. A., Gangloff, S. & Rothstein, R. The RecQ DNA helicases in DNA repair. *Annu. Rev. Genet.* **44**, 393-417, (2010).

27 van der Heijden, T. *et al.* Real-time assembly and disassembly of human RAD51 filaments on individual DNA molecules. *Nucleic Acids Res.* **35**, 5646-5657, (2007).

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

29 Reymer, A., Frykholm, K., Morimatsu, K., Takahashi, M. & Norden, B. Structure of human Rad51 protein filament from molecular modeling and site-specific linear dichroism spectroscopy. *Proc. Natl. Acad. Sci. USA* **106**, 13248-13253, (2009).

30 Carreira, A. *et al.* The BRC repeats of BRCA2 modulate the DNA-binding selectivity of RAD51. *Cell* **136**, 1032-1043, (2009).

31 Jensen, R. B., Carreira, A. & Kowalczykowski, S. C. Purified human BRCA2 stimulates RAD51-mediated recombination. *Nature* **467**, 678-683, (2010).

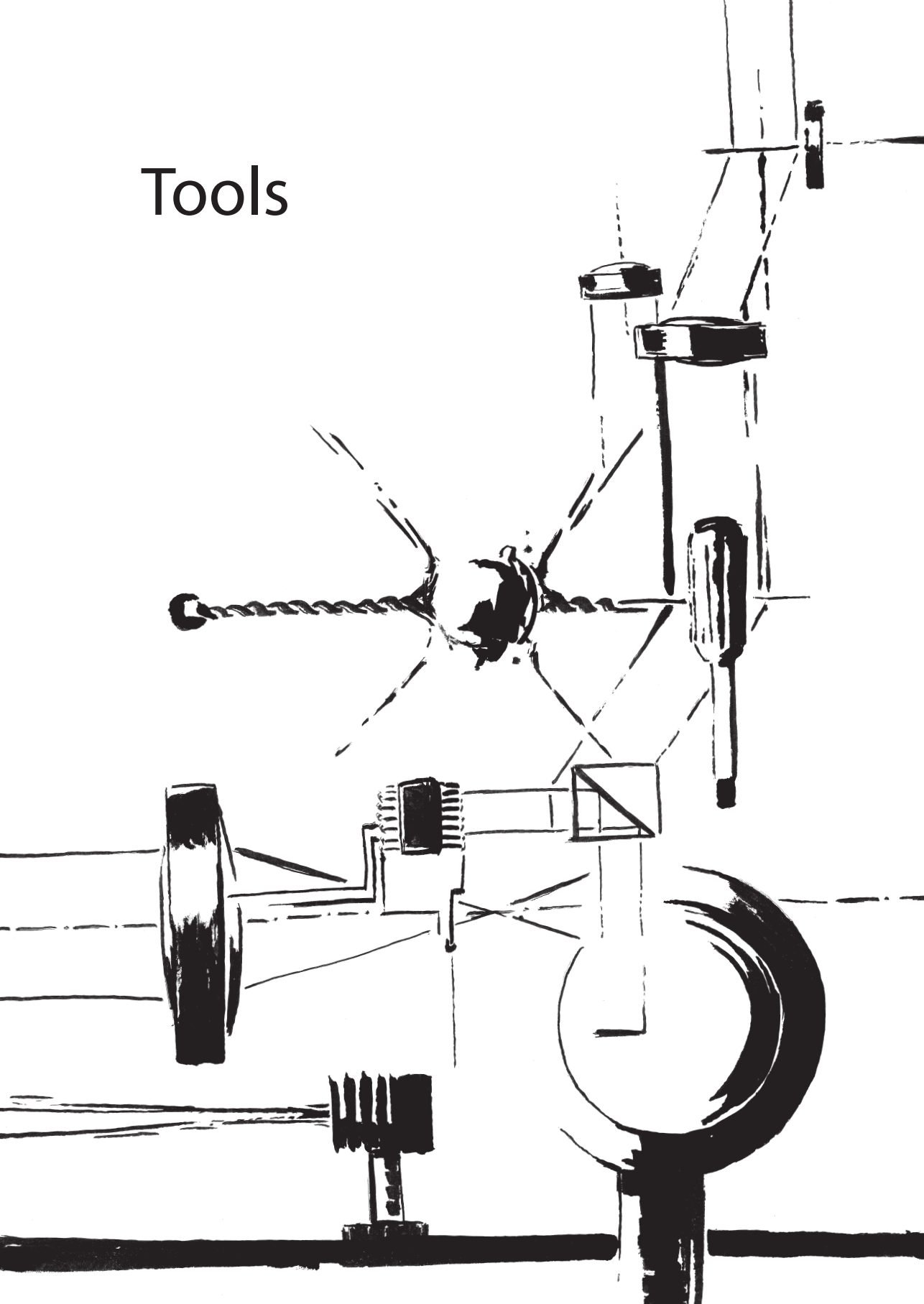
32 Voloshin, O. N. & Camerini-Otero, R. D. Synaptic complex revisited; a homologous recombinase flips and switches bases. *Mol. Cell* **15**, 846-847, (2004).

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

34 Paques, F. & Haber, J. E. Multiple pathways of recombination induced by double-strand breaks in Saccharomyces cerevisiae. *Microbiol. Mol. Biol. Rev.* **63**, 349-404, (1999).

35 van den Hout, M. *Forcing DNA and RNA through artificial nanopores*, PhD thesis, Delft University of Technology, (2010).

Tools



Part 0—Tools

Special nanoscopic tools are necessary to access the molecular world of DNA and proteins, as these molecules are simply too small to hold or see. Fortunately DNA is a long-chain molecule and orders of magnitude longer than most molecules or proteins. This thread-like nature of DNA allows us to use it as a tether between larger objects that we can more easily hold, manipulate, and detect. To make such a tether the DNA molecule must first be attached to a macroscopic object. Here biology provides a solution in the form of antibodies, proteins that tightly clamp a specific small molecule called an antigen. By adding antigens to the ends of the DNA and coupling the antibodies to macroscopic objects we create links that spontaneously snap into place. Most methods attach the DNA to a small bead which can be either held in the focus of a laser beam or which may be pulled on by a magnet. These methods are called optical and magnetic tweezers, respectively, referring to the traditional use of tweezers to hold small things. The size of the bead is also important as bigger beads are easier to work with but smaller beads respond faster and allow us to see smaller movements. Practically this results in bead sizes in the 0.5–2 μm range. These beads allow us to measure the extension and the forces acting on the DNA tethers. Optical detection is generally used as it can easily couple into a liquid environment and a high resolution can be obtained with sensitive cameras and photodiodes. Magnetic and optical tweezers can routinely measure single molecules but there is one major drawback: a single molecule in not necessarily representative for the whole ensemble. It is therefore essential to measure on as many individual molecules as possible and increase the throughput of these methods to obtain statistically relevant data. In this part of the thesis we present one way of increasing the throughput in magnetic tweezers by increasing the number of beads, and hence the number of tethers, we can see at once. We improved the tracking method to allow high-resolution tracking even at low magnification. A tool is, however, only a good tool if it is used. We therefore made this software easy and freely available to use to help more people in discovering the bio-molecular world.

CHAPTER 2

Non-bias-limited tracking of spherical particles, enabling nanometer resolution at low magnification

We present a three-dimensional tracking routine for non-diffraction-limited particles, which significantly reduces pixel bias. Our technique allows for increased resolution compared to previous methods especially at low magnification or at high signal-to-noise ratio. This enables tracking with nanometer accuracy in a wide field of view and tracking of many particles. To reduce bias induced by pixelation the tracking algorithm uses interpolation of the image on a circular grid to determine the *x*-, *y*-, and *z*-position. We evaluate the proposed algorithm by tracking simulated images and compare it to well-known center-of-mass and cross-correlation methods. The final resolution of the new method improves up to an order of magnitude in three dimensions compared to conventional tracking methods. We show that errors in x,y-tracking can seriously affect *z*-tracking if interpolation is not used. We validate our results with experimental data obtained for conditions matching those used in the simulations. Finally we show that the increased performance of the proposed algorithm uniquely enables it to extract accurate data for the persistence length and end-to-end distance of 107 DNA tethers in a single experiment.

This chapter has been published as: Marijn T.J. van Loenhout, Jacob Kerssemakers , Iwijn De Vlaminck, and Cees Dekker. Non-Bias-Limited Tracking of Spherical Particles, Enabling Nanometer Resolution at Low Magnification. *Biophysical Journal* **102**, 2362 (2012).

Introduction

Tracking of individual micron-sized particles by video microscopy has numerous applications in biophysics, particularly in magnetic and optical tweezers[1]. One of the main advantages of video microscopy lies in its flexibility, ease of implementation, and ability to track a large number of objects[2]. Apart from the camera and image system hardware the method is entirely based on software and requires no special hardware like laser-based detection schemes often employed in optical tweezers[3-6]. Software can therefore easily be reused and transferred between different applications. Ongoing advances in digital cameras and computing hardware continue to increase the applicability of video microscopy. Megapixel-size cameras now allow the monitoring of thousands of objects and fast acquisition rates allow cameras to be used for particle tracking in the kHz range[7].

The final resolution of any tracking method depends on the accuracy, which is a measure of the systematic error or bias, and the precision, which is a measure of the statistical error or scatter. With decreasing magnification, the area mapped onto a single pixel increases and the pixel bias, related to mapping of the intensity on the discrete pixel grid, becomes more important than the error due to scatter. High signal-to-noise ratios are commonly achieved when tracking non-diffraction-limited particles and these often reduce the statistical error well below the error due to bias. There is thus a need for tracking methods that account for this bias such that the bias is reduced to a level below the statistical error or scatter.

The procedure for tracking single fluorophores or other diffraction-limited objects by fitting with the point-spread function of the microscope or an approximation with a two-dimensional Gaussian was recently shown to be capable of achieving the theoretical minimum uncertainty[8,9]. However, for objects much larger than the wavelength of light ($d > \lambda$) or for significantly defocussed objects, the intensity profile is no longer accurately described by a Gaussian. A different tracking method is therefore required when dealing with such images. A large number of methods have been developed which can be roughly divided into two categories: firstly, methods that perform a direct calculation of the sub-pixel location (center of mass, Gaussian fit)[10,11]. Secondly, methods that compare the image with another image (be it a predefined kernel, a mirror image or a previous image) and perform a subsequent fitting step to achieve sub-pixel resolution. These methods include cross-correlation and the sum of absolute differences[12,13]. Several of these tracking methods were compared previously [14] and two-dimensional cross-correlation proved to be the preferred method for tracking non-diffraction-limited objects. However, a significant contribution of pixel bias to the final resolution remains for the cross-correlation method, which can severely limit the final tracking resolution.

We may understand the pixel-bias error to be due to a loss of information when an image is sampled on a discrete grid (Figure 2.1A and B)[15]. Information on the distribution of intensity within a single pixel is replaced by the mean intensity at the pixel center. If the spatial frequency spectrum of the image exceeds half the sampling

frequency information is lost according to the Nyquist theorem. For example, the centroid of the original image therefore need not be equal to the centroid of the sampled image. This is illustrated in Figure 2.1C where a Gaussian intensity profile is sampled by three pixels. The result of the sampling is an error in the position of the centroid calculated using the sampled data with respect to the true centroid position. This error or pixel bias can be substantial and is periodic in the pixel grid (Figure 2.1D). Kernel-based tracking methods may suffer in a similar way from pixel bias. To obtain the final position these methods require a sub-pixel fitting step, where the exact shape of the peak is partly lost in the sampling process, leading to a possible error in its determination[16].

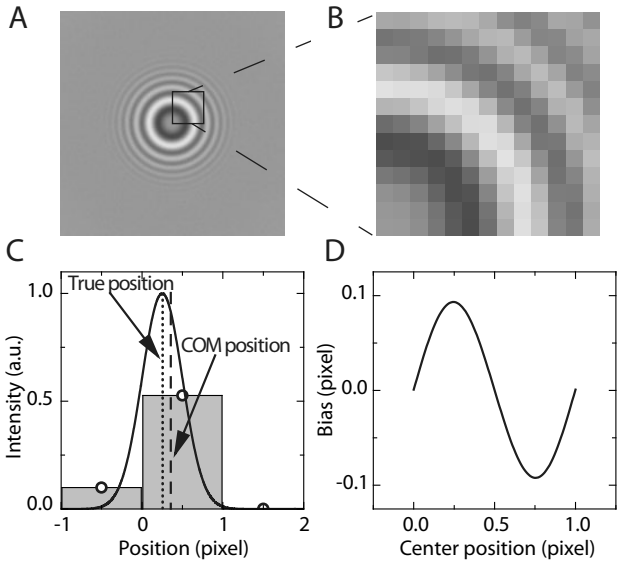


Figure 2.1: The origins of pixel bias. (**A**) Diffraction pattern of a bead. (**B**) Zoom in on the diffraction pattern showing the artifacts created by the sampling process. (**C**) A Gaussian intensity profile (solid line) is sampled by three pixels. The information of the intensity distribution within a pixel is lost by the sampling process and is replaced by a single intensity at the center of the pixel (circles). The centroid of the pixel intensities (dashed line) can have an error or pixel bias compared to the true center position of the Gaussian (dotted line). (**D**) The pixel bias shows a periodic variation when displacing the Gaussian profile across the pixel grid.

Pixel bias can be reduced by increasing the magnification and thereby sampling well above the Nyquist frequency. This is however not always possible nor desirable as it reduces the field of view and may also decrease contrast. If the exact intensity profile of the tracked object is known, as is the case for diffraction-limited objects, this

information can be used to correct for the information loss due to sampling and thus eliminate bias. If however the exact intensity profile is not known, the error due to bias will depend on the specifics of the algorithm.

Here we describe and evaluate a new method for tracking micron-sized beads, using an algorithm specifically aimed at reducing the pixel bias, thereby increasing the final tracking resolution. Our quadrant-interpolation (QI) algorithm makes use of the circular geometry of the diffraction pattern to resample the image on a circular grid. The resampling procedure reduces bias in two ways. First, it allows for a degree of upsampling which increases the spatial sampling frequency thereby reducing bias. Second, the resampling on the polar grid effectively removes the integer spacing of the pixel grid and thus averages out the errors due to pixel bias. The performance of the algorithm is evaluated by tracking computer-generated images derived from experimental images and comparing with two commonly used methods: center of mass (COM)[14,17] and cross-correlation (XCOR)[12,13,18]. We evaluate these different methods for scatter, bias, and resolution at different magnifications as well as the influence of the signal-to-noise ratio. We validate the improved performance of the presented algorithm for experimental conditions closely matching the simulations. The ability of the new algorithm to accurately track at low magnification uniquely enables highly parallel single-molecule experiments. We show that we were able to obtain accurate persistence length and end-to-end distance data for 107 DNA molecules in a single multiplexed magnetic tweezers experiment. The QI algorithm thus allows to vastly increase the throughput of single-molecule experiments and thereby enables the study of rare events and the acquisition of large statistical datasets from individual experimental runs.

Quadrant-Interpolation Algorithm

XY-tracking

The proposed QI algorithm uses the circular geometry of the diffraction pattern to strongly suppress bias and was implemented in Labview (National Instruments) software; for details see: 'Tracking software', at the end of this manuscript.The QI algorithm consists of four basic steps to determine the *x,y*-position:

I. In the first step, the center of the particle is roughly identified by a background-corrected center-of-mass calculation,(COM, Supplementary information Algorithm S2.1), followed by a 1D-cross-correlation step in the *x* and *y*-directions (XCOR, Algorithm S2.2). The accuracy of this first step is not crucial and it can be substituted by any method capable of locating the center of the particle to within approximately 1 pixel (Figure 2.2A, red cross).

II. The second step uses this first ($x_{est,}$ y_{est})-position to calculate a radial profile of the intensity of the particle for each quadrant on a circular grid, where points are spaced by δr and δθ, in radial and angular dimension respectively. The values at the circular grid positions are calculated by bi-linear interpolation of the four

nearest-neighbor pixel values, a single radial intensity profile is created representing the quadrant (Figure 2.2A). The grid is generally chosen to oversample in the radial direction compared to the pixel size, i.e. δr < pixel unit and is calculated for each quadrant of the image (i.e. top left, top right, bottom right, bottom left), yielding four intensity profiles $q_{TL}(r)...q_{BL}(r)$. These 4 radial profiles will be used to determine both the x,y as well as the z-position.

III. In the third step, relative shifts in the radial profiles are used to extract a correction to the initial (x_{est}, y_{est})-position. The sum of the right profiles, $q_R=q_{TR+BR}$, is concatenated with the left equivalent $q_L=q_{TL+BL}$, thus creating an intensity profile, I_x, representing the horizontal shift, where $||$ denotes the concatenation (Figure 2.2B)

$$I_x(r) = q_L(-r) \,||\, q_R(r). \tag{2.1}$$

The resulting intensity profile $I_x(r)$ is now cross-correlated by Fourier transform with its mirror profile $I_x(-r)$,

$$X_{xx'}(r) = \mathrm{IFFT}(\mathrm{FFT}(I_x(r)) * \overline{\mathrm{FFT}(I_{x'}(-r))}), \tag{2.2}$$

where IFFT and FFT are respectively the inverse and regular fast Fourier transforms, and $\overline{\mathrm{FFT}(I_{x'}(-r))}$ denotes the complex conjugate. The resulting cross-correlation function $X_{xx'}(r)$ is a curve with spacing δr and a peak at the position corresponding to twice the displacement d_r of the particle from the center of the circular grid. A similar routine is followed in the y-direction by processing $q_T=q_{TL+TR}$ and $q_B=q_{BL+BR}$.

IV. The final step in the algorithm determines the sub-bin peak position with a 5-point parabolic fit to the correlation function, yielding a shift d_r (Figure 2.2C). For corrections smaller than one pixel unit, i.e. small compared to the radial pattern fringe spacing, d_r relates to a truly x-directed pattern shift d_x via a simple geometrical correction: $d_x = d_r/(\pi/2)$, which accounts for the summation of the profiles, and their cosine-projected displacements, over a hemisphere. The tracked position thus finally is: $x = x_{est}-d_x$. The y-position is calculated analogously.

The second step of the algorithm, where radial intensity profiles are calculated, allows a choice of the parameters of the polar grid, $\delta\theta$ and δr. In general the final resolution increases with a higher sampling density in the radial direction (i.e. smaller δr). However, empirically we have found that the final resolution does not increase further when δr is reduced below ≈1/3 of a pixel unit. The angular spacing $\delta\theta$ is chosen to sample all pixels, i.e. $\delta\theta \cdot r_{max}$ = pixel size, where r_{max} is the maximum radius used to construct the radial intensity profile. The final resolution does not improve if $\delta\theta$ is decreased further. However, $\delta\theta$ may be increased considerably without severe loss in tracking resolution. This enables one to trade off accuracy for speed by sampling on a wider spaced grid. The execution speed of the algorithm, when sampling all pixels,

is sufficient for most applications and in its current implementation is approximately 0.01 s for a 128x128 pixel image on an 2GHz Intel core2 CPU.

Z-tracking algorithms

Methods for tracking the axial or *z*-position of a particle generally rely on first creating a lookup table of radial profiles by shifting the objective in known steps[13,19]. Here we employ the following method: an image stack is acquired by changing the focus of the objective, typically in 200 steps of 100 nm. Subsequently a *z*-lookup table is constructed by calculating the radial profile at each focus position. For the COM method the pixel intensities were directly mapped into radial bins with 1 pixel spacing (Figure 2.2D). For the XCOR method the pixel intensities were distributed by linear interpolation into radial bins with 1 pixel spacing. For the QI tracker, the radial profiles, calculated using bilinear interpolation for all four quadrants with a δr of ⅓ pixel, were averaged resulting in a much smoother lookup table (Figure 2.2E).

To determine the *z*-position for given image, the χ^2-difference of the radial profile with the planes of the *z*-lookup table was calculated. The minimum of the χ^2-difference corresponds to the best matching radial profile and therefore to the *z*-position of the object. To achieve a resolution smaller than the step size of the lookup table the final *z*-position was obtained by a parabolic fit of this minimum. When using an oil-immersion objective, the refractive index of the immersion medium is normally not equal to the refractive index of the medium containing the tracked object. Therefore, a displacement of the objective, used to make the *z*-lookup table, is not equivalent to a displacement of the object during a measurement[20]. We use an experimentally determined linear correction factor, (0.80 for our system, see Materials and Methods), to correct for this scaling of the *z*-position.

Materials and Methods

Experimental setup

Images, for the experimental data and the artificial image construction, were acquired with a Falcon 1.4M100XDR (Dalsa) camera in 8 bit mode using a Nikon CFI Plan Apo 50X oil immersion (NA 0.9) objective and a collimated green LED as light source. A 300 mm tube lens was used to obtain an effective magnification of 75X. Images were saved in tiff format and processed offline using Labview (National Instruments software). A sample of fixed beads was prepared by immobilizing 2.8 μm magnetic beads (Dynal M270, Invitrogen) on a glass cover slide by baking at 180°C for 10 minutes and a flow cell was constructed by sandwiching a second cover slide on top using a parafilm spacer. The z-correction factor was determined experimentally by creating a flow cell with a 50 μm gap, which allowed to focus on both the top and bottom surfaces. This flow cell was filled with both water and immersion oil separated by a thin parafilm barrier. Subsequently the required displacement of the objective, δz, to achieve focus at the top and bottom surfaces in both immersion oil and water was measured. The

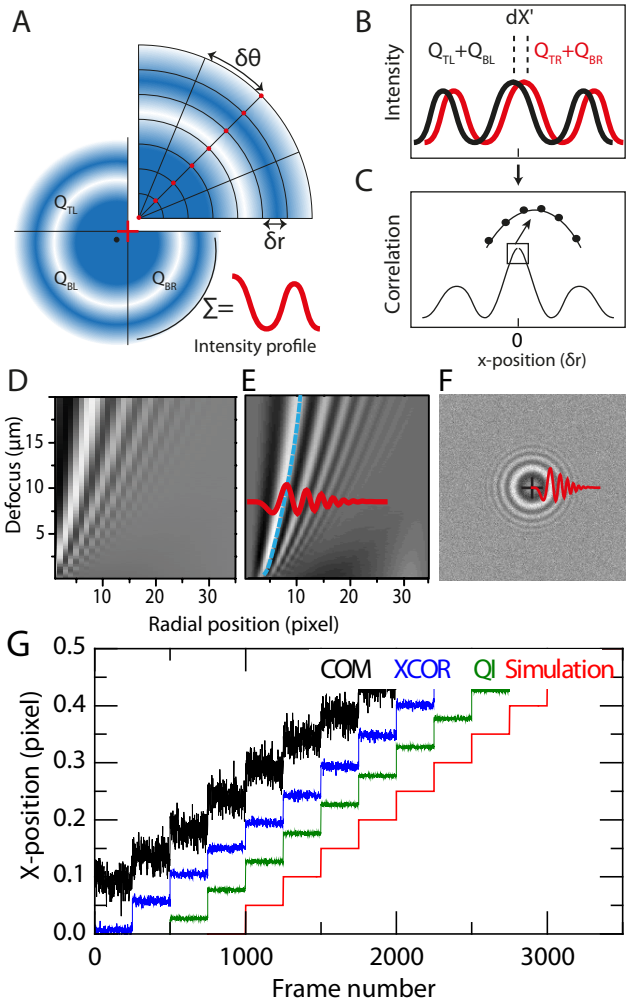


Figure 2.2: Schematic representation of the QI tracking algorithm. (**A**) First an approximate center position is obtained (red cross) by a different tracking method, e.g. center of mass. Second radial intensity profiles are constructed by bi-linear interpolation on a circular grid with spacing *δθ, δr*. For each quadrant the intensity profiles are summed to obtain a radial intensity profile for each quadrant (red curve). (**B**) Intensity profiles of left/right quadrants are concatenated to represent the pattern in *x*-direction (red curve) subsequently a cross-correlation with its mirror image (black curve) is performed to obtain the center position of the pattern. (**C**) The peak in the cross-correlation is fit to obtain final *x*-position. (**D**) *Z*-lookup table of images comparable to Figure 2.1A made without interpolation. (**E**) *Z*-lookup table made by summing the radial intensity profiles of four quadrants obtained by interpolation on a circular grid. The brightest fringe was fit by a parabolic function (blue

dashed line). Using this parabolic fit, all radial intensity profiles were scaled in the radial direction and averaged to create a single generic curve (red curve). (**F**) Artificial image generated using the generic curve. Poissonian noise was added to obtain different SNRs (**G**) Tracking result of simulated images that were regularly displaced by 0.1pixel in *x*-direction at 30x magnification and SNR of 30, offset vertically for clarity, COM (black), XCOR (blue), QI (green), and actual simulated position (red).

correction factor was then simply given by $\delta z_{oil} / \delta z_{water} = 0.8$ for our setup. We note that this correction factor is significantly different from the small angle approximation based on the index of refraction of the media: $n_{water} / n_{oil} = 1.33/1.51 = 0.88$, which is not valid for objectives with high numerical aperture.

Evaluation procedure and artificial image construction

Tracking algorithms are often evaluated by tracking a stationary object, where the fluctuations in the tracked position are used to estimate the accuracy of the algorithm. Measurements using immobilized beads or markers, however provide a too optimistic value of the resolution as they do not take the systematic errors or pixel bias into account. To measure the true resolution of a tracking algorithm, one must compare the tracked position with its actual position, which is experimentally complicated, but easily achieved with computer-generated images. We therefore artificially generated images simulating a stepwise movement of a particle over the image grid, similar to Cheezum *et al.*[14].

We used experimentally acquired images to construct artificial images. The advantage of this *a posteriori* approach is that a close match to experimental images is guaranteed and no information of the imaging system is required. First images of a 2.8 μM magnetic bead were recorded at 75x magnification for 200 different focal positions separated 100 nm, by stepwise moving the objective in *z*. Next, these images were used to calculate radial intensity profiles at each *z*-position, resulting in a *z*-lookup table (Figure 2.2E). Typically, a regularly expanding pattern of fringes is visible, which we approximate by the radial expansion of a single intensity profile function (Figure 2.2E, red curve). A parabolic fit of the brightest fringe was used to map the monotonous expansion of the fringes as a function of defocus (Figure 2.2E, blue dashed line). By scaling the radial intensity profiles at each defocus position, an average 'generic' radial profile was constructed, which, together with the parabolic fit of the brightest fringe, now suffices to reconstruct a noise-free pattern at any translational (*x,y*) and defocus (*z*) position. We thus have generated a perfect, noise-free image of a bead, henceforth referred to as reference pattern. To simulate the effect of shot noise Poissonian distributed noise was finally added to the reference pattern (Figure 2.2F).

Next, we generated an image series representing a movement of 2 pixels across the camera image grid in 40 steps. The z-position of the bead was set at 10 μm above the focal plane, corresponding to the middle of the lookup table. At each step 250 different

images were created by adding shot noise to the initial generated image. The signal-to-noise ratio (*SNR*) was subsequently determined from the generated images as:

$$SNR = \frac{S}{4\sigma} - 1, \tag{2.3}$$

where the signal range, S is defined as the difference between the maximum and minimum intensity of the image, $S = (I_{max}\text{-}I_{min})$. The noise is determined from a region of the image containing no signal, where 4σ sets a 95% confidence bound. The factor of 1 is subtracted to account for the apparent increase in the measured signal range due the addition of noise. Images at different magnifications were obtained by mapping the reference pattern on different pixel grids relative to the 75x magnification of the initial experimental images. Subsequently these images were tracked with the proposed QI method as well as with COM and COR algorithms (Figure 2.2G, offset for clarity). The scatter and bias were independently determined. Scatter was calculated as the standard deviation of the tracked position at each step. Bias was calculated as the deviation of the determined mean position from the simulated position, minus the overall error in mean position for all steps. Subtracting the overall error in mean position removes any non-position dependent difference between the simulation and tracked position, which is in most practical cases irrelevant as it will be equal for all tracked objects in a field of view. This approach allows us to determine the bias and scatter separately as a function of the noise, magnification, and z-position of the image.

Results

Using the procedure outlined above, we evaluated the influence of magnification, SNR and crosstalk on tracking of non-diffraction-limited particles. We compare the newly developed QI algorithm to the well-established COM and XCOR methods. In the results below, the standard deviation, σ, is shown in units of pixels at different magnifications for scatter and bias. An example of the scatter and bias at 75x magnification is shown in Figure 2.3B. The total resolution was subsequently calculated from the bias, σ_{bias} and scatter, $\sigma_{scatter}$ at each magnification as:

$$\sigma_{total} = \sqrt{\sigma_{bias}^2 + \sigma_{scatter}^2} \tag{2.4}$$

and is shown in nanometer for the given magnification based on a 7.4 μm pixel size, matching the pixel size of the camera used to obtain the experimental images. Unless stated otherwise, all results were evaluated at a SNR of 30.

The experimentally obtainable maximum signal to noise level for an 8-bit camera can be estimated to be around 30, based on the signal range and expected shot noise. The maximum peak-to-peak signal range, *S*, without clipping, will be approximately half the bit depth i.e. *S*= 128 for the 8-bit camera. Assuming a pixel-well depth of 100,000e for a typical camera the shot noise calculated for the median intensity of 50,000e will be $\sigma = 256/\sqrt{50000} \approx 1.1$ bit, which results in a signal-to-noise ratio of SNR $= /(4\sigma) \approx 28$. Higher SNRs can be obtained by using cameras with a larger pixel-well depth and using cameras capable of working at more than 8-bit. However, in most cases shot noise will be limiting as it is generally larger than 0.5 bit.

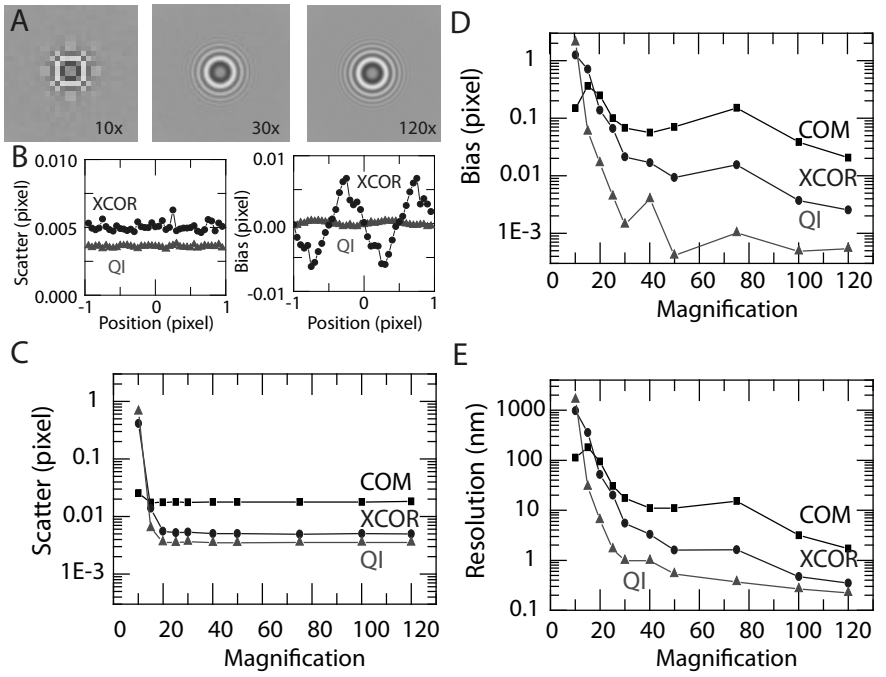


Figure 2.3: Scatter, bias, and resolution in the *x*-direction for simulated data as a function of magnification at signal-to-noise ratio of 30. For the COM (squares), XCOR (circles), and QI (triangles) methods. (**A**) Simulated images at different magnifications. (**B**) Example of scatter and bias at 75x magnification for a displacement across two pixels. (**C**) Scatter as a function of magnification. (**D**) Bias as a function of magnification. The error due to pixel bias was comparable to scatter for COM and XCOR, but significantly reduced for the QI tracker. (**E**) Resolution in nm for a 7.5 µm pixel size as a function of magnification.

XY-tracking

The results for in-plane tracking show that the stochastic error or scatter remains nearly constant for all three algorithms (Figure 2.3C). Only at magnifications below 20x, the scatter increased markedly for both the xcor and qi tracker, which likely results from undersampling of the special pattern. Above 15x magnification the qi tracker has improved performance compared to both com and xcor algorithms.

The bias performance of all algorithms is more sensitive to magnification than scatter and there are larger differences in tracking error between the tracking algorithms (Figure 2.3D). As expected, bias decreases with magnification as the image is spread out over more pixels reducing the effect of pixel borders. The qi algorithm shows a fast decrease in bias reaching a low-bias region above 50x magnification. Bias for the com and xcor methods reduces more slowly with magnification and never reaches the performance of the qi method. Importantly, the bias for the com and xcor algorithms is larger than the scatter for most magnifications, thereby limiting the total resolution (Figure 2.3E). In the absence of bias, the resolution of all tracking algorithms would improve linearly with magnification, as scatter measured in pixels stays constant. However the bias significantly affects performance, especially for the com an xcor tracking routines. The qi algorithm shows superior performance because it is not bias limited and resolution indeed scales with magnification from 30x upwards, where scatter is the main source of error.

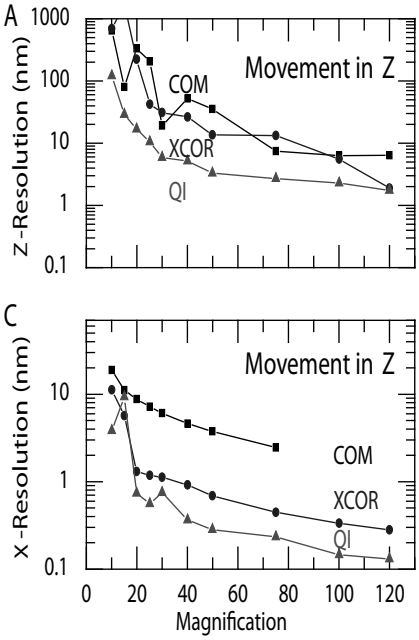
Z-tracking

Tracking resolution in the *z*-direction is generally less than the *x,y*-resolution as the expansion of the diffraction pattern changes only slowly with *z*-position (Figure 2.2D, E). Our system has approximately 1 nm radial expansion for a 7 nm *z*-displacement, as calculated from the local slope of the fringe pattern in the *z*-lookup table. The resolution in *z* was evaluated by a simulated displacement over 2 μm in 100 steps at constant *x,y*-position. The result is shown in Figure 2.4A, the qi algorithm achieves a resolution below 10 nm for magnifications above 25x. The com and xcor algorithms perform less well which is due to propagation of *x,y*-tracking errors. Using the more accurate *x,y*-positions obtained from the qi tracker as input for the com and xcor based *z*-tracking algorithms showed improved performance comparable to that of the qi *z*-tracking methods (data not shown), which underlines the importance of accurate *x,y*-tracking for accurate *z*-tracking.

To evaluate the crosstalk between in plane tracking and *z*-tracking, we evaluated the *z*-resolution while displacing the bead image in *x*-direction at constant *z*-position (Figure 2.4B). The qi algorithm again shows the best performance at all magnifications. The com non-interpolating z-algorithm performs more than an order of magnitude worse than the interpolating qi and xcor methods. This very large error is due to the non-interpolating *z*-tracking algorithm used for the com tracker, which simply maps the intensity of a pixel to 1 pixel wide bins in the radial intensity profile. The xcor *z*-tracker on the other hand uses linear interpolation to distribute the intensity

between two radial bins and performs significantly better. If the non-interpolating COM *z*-tracking method was used in the XCOR algorithm, performance decreased notably becoming comparable to that of the COM (data not shown). Interpolation of the radial profile thus enables much better suppression of in plane movements, irrespective of the tracking algorithm used.

Finally, we also evaluated the sensitivity of the *x*-resolution with respect to movements in z (Figure 2.4C). The crosstalk in this case is small compared to the error for in-plane movements. Therefore the *x,y*-resolution will generally only have a minor contribution from crosstalk due to *z*-movements. At very low magnifications, both the QI and XCOR method show a strong degradation of resolution as the diffraction pattern is not correctly sampled anymore.



D

		Move X	Move Y	Move Z	Total
X-Resolution (nm)	QI	0.4	0.2	0.2	0.5
	XCOR	1.6	0.4	0.4	1.7
	COM	15.2	1.8	2.4	15.5
Z-Resolution (nm)	QI	1.4	1.4	2.7	3.3
	XCOR	1.6	1.6	6.7	7.8
	COM	35.1	35.1	7.5	50.2

Figure 2.4: Axial resolution and crosstalk between *x* and *z*-position as a function of magnification for simulated data at signal-to-noise ratio of 30. COM (squares), XCOR (circles) QI (triangles). (**A**) *Z*-resolution determined for a 2 µm *z*-movement of the bead. (**B**) Crosstalk in *z*-tracking for a stepwise displacement over 2 pixels in the *x*-direction for the tracked bead. Displacements in *x* results in significant performance loss for the non-interpolating COM tracking method. (**C**) Crosstalk in *x*-direction for a simulated bead movement of 2 µm in *z*. (**D**) Crosstalk in tracking resolution in nm at 75X magnification and a signal-to-noise ratio of 30. Columns show the direction of movement for the tracked bead. The last column shows the total resolution calculated as, $\delta x_{total} = \sqrt{\delta x^2_{moveX} + \delta x^2_{moveY} + \delta x^2_{moveZ}}$ or

$$\delta z_{total} = \sqrt{\delta z^2_{moveX} + \delta z^2_{moveY} + \delta z^2_{moveZ}} .$$

The above results demonstrate that the resolution in tracking of an object moving in three dimensions can be dramatically worsened by crosstalk present between movements in *x,y* or *z*-direction. To gain further insight in the effects of crosstalk, we calculated the resolution in individual *x,y,* and *z*-directions as well as the total resolution at 75x magnification with a SNR of 30 (Figure 2.4D). This shows that the QI tracker has the lowest level of absolute crosstalk, XCOR and COM have higher levels of absolute crosstalk in *x,y*-direction. Most notable is the very severe crosstalk between *x,y*-movements and *z*-tracking, resulting in a significantly reduced performance for the non-interpolating COM *z*-algorithm.

Signal-to-noise ratio

To determine the performance of the tracking methods as a function of SNR, we evaluated the resolution at SNRs between 2 and 100, as defined in equation at a constant magnification of 75x. As expected, the scatter decreases monotonically with increasing SNR (Figure 2.5A, closed symbols). At low SNR the QI method performs better compared to XCOR, which is explained by the fact that it samples all pixels, whereas XCOR only samples a subset resulting in higher effective noise. The bias (Figure 2.5A, open symbols) is influenced in a more complex way by the SNR: COM and QI show a decrease in bias with increasing SNR. The bias for the XCOR is however independent of the SNR. Importantly, the bias for the COM and XCOR methods is larger than the scatter for most SNR, and bias thus dominates the total resolution (Figure 2.5B).

To gain insight in the general effects of magnification and SNR on tracking resolution, we developed a simple analytical expression of the achievable resolution which is based on the signal amplitude and SNR of an image (Method S2.3). This prediction (grey line Figure 2.5B) reproduces the SNR behavior of the QI tracker within a factor of two for SNR below 40. For higher SNRs the resolution of the QI levels off and does not match the prediction anymore. Next, we evaluated the SNR response for *z*-tracking (Figure 2.5C). *Z*-resolution again improves with the SNR for all methods. The QI algorithm performs best for all SNR, whereas the non-interpolating *z*-algorithm of the COM tracker has a severely reduced resolution. We extended the resolution prediction to the *z*-direction by scaling the predicted *x*-resolution with a factor of 7, equal to the relative expansion of the diffraction pattern in radial direction for a given focus displacement. This simple scaling indeed approximates the behavior of the *z*-resolution of the QI tracker (Figure 2.5C gray line).

Experimental validation

To validate the simulation approach and the performance of the QI tracker, we performed an experiment closely matching the simulations. The *z*-resolution was measured for a *z*-displacement at different SNRs. To create a *z*-displacement the objective

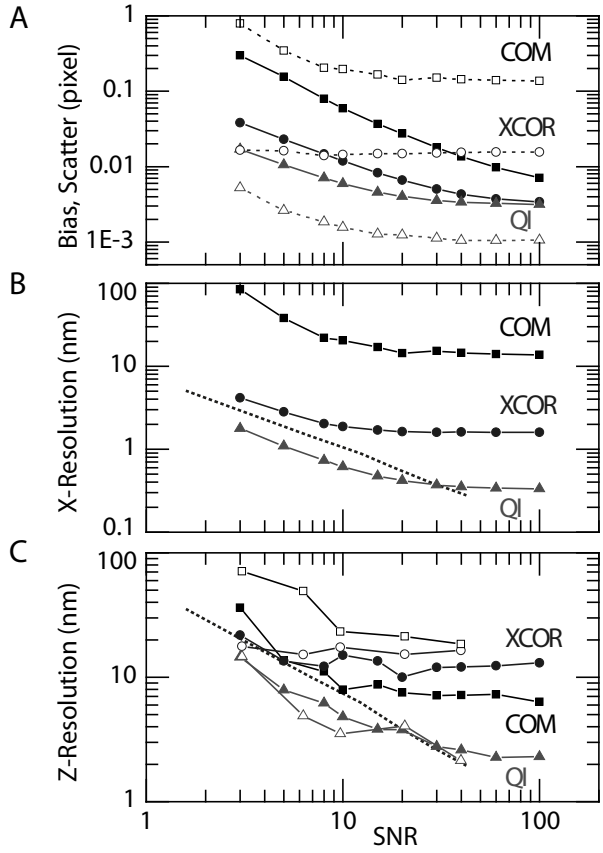


Figure 2.5: Influence of the signal-to-noise ratio at 75x magnification. (**A**) Scatter (solid lines/symbols) and bias (dashed lines, open symbols) for simulated data in the *x*-direction as a function of SNR, for the COM (squares), XCOR (circles), and QI (triangles) algorithms. Scatter decreased monotonically with increasing SNR for all algorithms. The bias for both the COM and XCOR methods was larger than their scatter for most SNRs thus negatively dominating the final resolution. The bias of the QI algorithm was well below scatter for all SNRs, resulting in a superior total resolution. (**B**) *x*-resolution as function of SNR for simulated data. Bias heavily affects performance of the COM and XCOR algorithms. The resolution prediction calculated for the experimental images (dotted line, Method S2.3) agrees with the simulation results within a factor of two over the whole range of SNRs. (**C**) *Z*-resolution as function of signal to noise ratio for simulated (closed symbols) and experimental data (open symbols). Experimental results for both QI and XCOR algorithm agree well with the simulation results. The resolution prediction calculated for the experimental images (dotted line, Method S2.3) matches well with the experimental data.

was moved in 20 steps of 100 nm matching the 2 μm total displacement used in the simulations. 2.8 μM magnetic beads, identical to those used to generate the reference image for the simulations, were immobilized on a glass cover slide and tracking resolution of the different algorithms was evaluated at 75x magnification as a function of SNR. The SNR was set in the range between 3 and 40 by changing the illumination intensity. To remove experimental drift a single bead was used as a reference. The tracked position of this 'reference bead' was subsequently subtracted for each time point from the positions of the other beads (n=4). The resolution was calculated as the standard deviation of the reference-subtracted bead positions divided by $\sqrt{2}$, to account for the noise added by subtracting the reference bead[21]. These experimental results provide a lower limit for the resolution and closely match the simulation results for the QI, and XCOR algorithm (Figure 2.5C, green and blue open symbols). The experimental resolution of the COM algorithm (Figure 2.5C, black open symbols) shows markedly worse performance than the results from the simulations. This can be explained by the fact that the COM algorithm is very sensitive to disturbances such as uneven background illumination or asymmetries in the tracked beads. Both of these effects are present in the experimental data but not in the simulated, thereby decreasing the observed experimental tracking resolution. The simple resolution estimate outlined in Supporting Material Method S2.3 (Figure 2.5C, grey line) matches both the experimental and simulated data within a factor of two. The above results show that under near identical conditions the z-resolution of the experimental data match closely to the simulation results, validating the evaluation approach based on simulations.

The QI algorithm enables highly parallel magnetic tweezers

To demonstrate the experimental relevance of the proposed QI algorithm we performed measurements at 25x magnification in multiplexed magnetic tweezers. Targeted DNA tethering was used to tether a large number of 1 μm paramagnetic beads to 7.3 kb DNA molecules in a 10 mM Tris buffer (pH 7.5) (Figure 2.6A)[22]. Using this technique we were able to identify 245 individual beads in the 300 × 400 μm field of view. A measurement procedure optimized to efficiently handle a large number of the DNA tethers in a single experimental run was used obtain force extension characteristics[22-24]. The COM, XCOR and QI algorithms were used in parallel to obtain tracking data during these measurements and evaluate the performance of each algorithm. The tracking results for a force extension measurement of a single DNA tether reveal that only the QI algorithm is able to accurately track the bead position throughout this force extension measurement (Figure 2.6B). The COM algorithm loses tracking completely for most data points. Whereas the XCOR algorithm manages to track the bead for most positions but suffers from large discrete steps in the tracked position resulting in large errors and an overestimation of the DNA length.

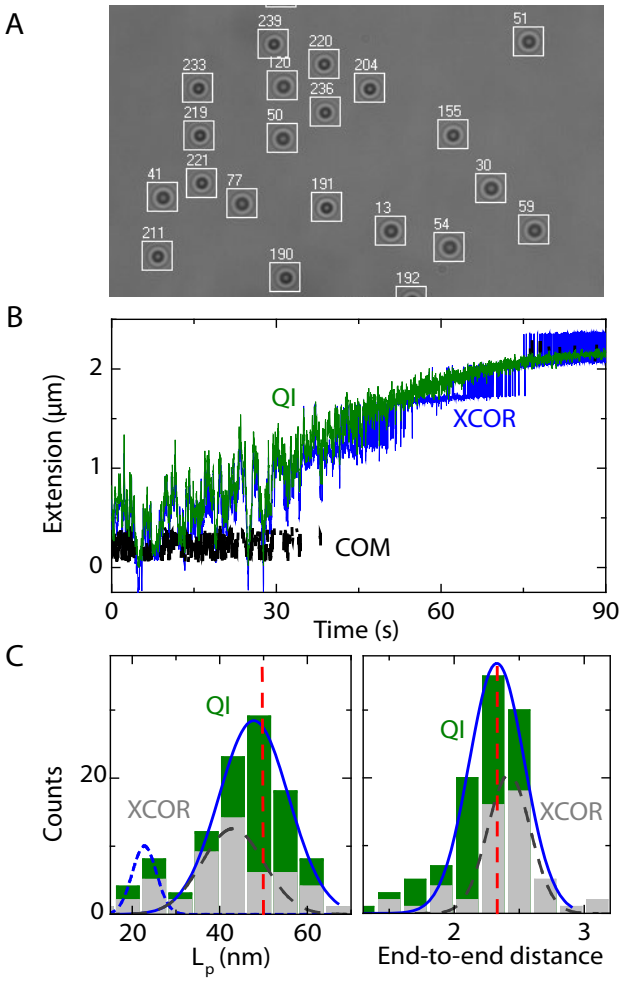


Figure 2.6: Tracking in highly parallel multiplexed magnetic tweezers at 25x magnification. (**A**) Part of a field of view in a multiplexed magnetic tweezers setup showing multiple beads at low resolution, in the total field of view 245 beads were identified. (**B**) Tracking results for the three different trackers of an experimental DNA force extension curve. The force applied on the bead was changed from approximately 20 fN at $t = 0$ s, to 3 pN at $t = 90$ s. The large errors of the COM and XCOR algorithms show their failure at low magnification. (**C**) Persistence lengths and end-to-end distances extracted from analysis of beads tracked with the QI (green bars) and XCOR algorithms (grey bars). Two peaks are found for the persistence length corresponding to single and double tethered DNA molecules (blue solid and dashed line QI, black dashed line XCOR). The persistence lengths and end-to-end distances determined by the QI algorithm are close to the expected values (red dotted lines).

To evaluate the quality of the data obtained by the different tracking algorithms we performed an analysis of the end-to-end length of 7.3 kb DNA molecules and their persistence length (L_p), a measure for the length scale over which orientational fluctuations decay. Out of all the beads tracked the QI algorithm yielded 107 DNA tethers which could be fit by the wormlike chain model[25,26]. Only 51 tethers tracked by XCOR algorithm could be fit and none of the tethers tracked by the COM algorithm provided useful data. The results for the obtained L_p and end-to-end distance of the DNA molecules are shown in Figure 2.6C. The histogram of L_p for the QI tracker shows two peaks, one at 48 nm and one at 23 nm, corresponding to beads tethered by a single DNA molecule and beads tethered by two molecules[27]. The measured L_p obtained by the QI algorithm is in good agreement with previous work[25] which found that $L_p \approx 50$ nm and is modestly dependent on salt in the range of 30 mM to 150 mM Na^+. The histogram for the persistence length determined from the data of the XCOR algorithm showed a peak at 43 nm considerably lower than the expected value of 50 nm.

A histogram of the measured end-to-end distance of the DNA molecules is shown in Figure 2.6C. The end-to-end distance at a nominal force of 1.8 pN was obtained by measuring the bead height for 25 s at a nominal force of 1.8 pN. The peak found for the QI algorithm at 2.33 ± 0.18 μm is close to the expected length = 2.26 ± 0.005 μm (at 1.8 pN, 25 s measurement time). The peak for the XCOR algorithm was located at a distance of 2.42 ± 0.14 μm, longer than the expected end-to-end distance. The increase in end-to-end distance determined by XCOR algorithm can be understood by looking at the tracked bead positions in Figure 2.6B. The tracking errors made by the XCOR algorithm lead to an underestimation of the lowest positions and an overestimation of the largest extensions, resulting in a longer measured end-to-end distance.

Discussion

In this work, we have developed a novel algorithm for tracking of non-diffraction limited objects in 3D. We have demonstrated superior performance of this newly developed QI algorithm compared to the COM and XCOR algorithms. The QI algorithm especially leads to a reduction of bias and reduced cross talk between *x,y* and *z*-motions. The properties of the QI algorithm allow it to function at high precision and at low magnification. The QI algorithm enables large scale multiplexed single-molecule measurements on DNA tethers where other algorithms fail.

The difference in bias performance between the algorithms depends on how they deal with the information loss due to sampling. The point spread function (PSF) of the microscope sets an upper limit for the spatial frequency components of the image. For a perfect aberration-free objective lens, the PSF in the focal plane takes the form of an Airy disk. The Airy disk has a diameter of $d_a = 1.22\lambda/NA = 705$ nm, where $\lambda = 520$ nm and NA=0.9 in our setup. Fourier analysis of the Airy pattern shows that 99% of the information is contained in spatial frequencies above 350 nm. Applying the Nyquist criterion to the Airy pattern gives a minimum required sampling resolution of 350

nm/2=175 nm which corresponds to 42x magnification in our setup. This shows that the QI tracking method makes effective use of all the information in the sampled image as bias does not improve for magnifications above 50x. The COM and XCOR algorithms on the other hand do not suppress bias well and require magnifications well above the Nyquist criterion to reduce bias.

A notable feature for all trackers was the sharp decrease in resolution at magnifications below 20x. This loss of tracking is also related to the spatial information content of the image. In this case however, it is linked to the fringes in the diffraction pattern, which account for the main information component in the image. As the fringe spacing approaches the pixel size a severe loss of information occurs resulting in major tracking errors. The fringe spacing between the main and second fringe in the generated images was approximately 1.3 μm which corresponds to a sampling limit at a magnification of 12x, matching well with the observed loss in resolution for magnifications below 20x (Figure 2.3E).

The resolution in z is considerably lower than the x-resolution, due to the fact that the diffraction pattern only expands slowly with movements in z. Indeed the analytical expression for the achievable resolution (Method S3) correctly captures the z-resolution by simply scaling the predicted x-resolution by a factor equal to the relative radial expansion of the diffraction pattern for an given axial displacement of the particle (Figure 2.4C). To increase z-resolution, it would thus be necessary to make the diffraction pattern more sensitive to changes in z. This can be achieved by simply increasing magnification or the numerical aperture of the objective, moreover interference-based measurements are able to greatly increase the sensitivity to z-displacements and have indeed successfully been applied to increase z-resolution[18,28].

The tracking resolution for particles moving in multiple dimensions can suffer from crosstalk. We showed that especially z-tracking can be significantly affected by crosstalk of in-plane movements (Figure 2.4B) as in-plane tracking errors will broaden and deform the radial profile. By using interpolation to create the radial intensity profile the QI and XCOR methods suppress pixelation deformations resulting in a much improved z-resolution. In experimental conditions, the effects of crosstalk may easily go unnoticed; tracking fixed beads will not reveal the errors due to crosstalk, but particles undergoing Brownian fluctuations will exhibit strong crosstalk. Simulations are thus essential to verify the performance of algorithms.

Single-molecule measurements are often not limited by resolution but suffer from the number of events that can be recorded in a single run. Multiplexing of the measurements is therefore highly desirable and several approaches have been published[21,29]. Reducing magnification allows more beads to be imaged and tracked. It is therefore interesting to compare the tradeoff between tracking resolution and magnification. The tracking resolution scales approximately linearly with magnification but the number of beads scales as the square of the inverse resolution. Decreasing the magnification thus provides a means to track many more beads at a limited loss of resolution. The Nyquist frequency however sets a limit to the minimal magnification that is

allowed before a major loss in resolution occurs. Our results show that at 30x magnification a sub-1nm resolution for *x,y*, and a sub-10nm resolution for the *z*-direction is possible. These resolutions, only achieved with the QI tracker, are more than sufficient for many single-molecule experiments. Decreasing the magnification to 30x allows 11 times more beads/area to be tracked compared to a 100x magnification.

Several additional sources of error exist that were not considered in this work. These include camera errors and pixel non-linearity which has been shown to influence tracking of single fluorophores[30]. Further experimental sources of noise include mechanical and acoustical vibrations, as well as uneven illumination and background. Particles may also be non-spherical. The QI tracking algorithm is however robust to such deformations of the circular geometry if these remain well below the fringe spacing in the diffraction pattern. With proper care it should however be possible to reduce these noise sources and obtain results similar to the simulations, as indeed show in our experimental verification data.

We have optimized the settings for each algorithm and believe that the results are representative for the general performance of each method. The clear performance trends increase our understanding of the underlying mechanisms. Many different algorithms and variations are used in the literature and results may differ depending on specific parameters. Claims in literature about the obtained resolution should therefore always be treated carefully as they often provide only an indication of performance under specific conditions, e.g. for static objects where crosstalk or bias may go unnoticed.

We showed that under many circumstances the tracking resolution of non-diffraction-limited particles is limited by bias and crosstalk. Simulations are essential to understand these effects, and provide a means to evaluate and select a tracking algorithm with superior performance. We have shown that multiplexed magnetic tweezers using the QI algorithm were able to provide vastly more data, obtaining a full histogram of L_p and end-to-end distances in a single experimental run. The QI tracker not only allowed us to obtain data for more than twice the number of DNA tethers compared to the XCOR algorithm, but more importantly the superior tracking resolution of the QI algorithm proved essential to obtain correct physical properties for the measured DNA molecules. We highly recommend the use of the QI algorithm for tracking non-diffraction limited spherical particles as it operates with very low bias and crosstalk and thereby achieves a higher final resolution in three dimensions. It is especially useful in conditions that are normally affected by bias, i.e., when tracking multiple beads at low magnification or when working at high SNR.

Tracking software

The QI tracking method was implemented in Labview (National Instruments) software. The base tracking algorithm as well as a standalone program capable of loading and tracking images is available for download from the authors website: http://ceesdekkerlab.tudelft.nl/download.

References

1 Neuman, K. C. & Nagy, A. Single-molecule force spectroscopy: optical tweezers, magnetic tweezers and atomic force microscopy. *Nat. Methods* **5**, 491-505, (2008).

2 Crocker, J. C. & Hoffman, B. D. in *Cell Mechanics* Vol. 83 *Methods in Cell Biology* 141-178 (Elsevier Academic Press Inc, 2007).

3 Ghislain, L. P., Switz, N. A. & Webb, W. W. Measurement of small forces using and optical trap. *Rev. Sci. Instrum.* **65**, 2762-2768, (1994).

4 Simmons, R. M., Finer, J. T., Chu, S. & Spudich, J. A. Quantitative measurements of force and displacement using an optical trap. *Biophys. J.* **70**, 1813-1822, (1996).

5 Rohrbach, A. & Stelzer, E. H. K. Three-dimensional position detection of optically trapped dielectric particles. *J. Appl. Phys.* **91**, 5474-5488, (2002).

6 Huisstede, J. H. G., van der Werf, K. O., Bennink, M. L. & Subramaniam, V. Force detection in optical tweezers using backscattered light. *Opt. Express* **13**, 1113-1123, (2005).

7 Otto, O. *et al.* Real-time particle tracking at 10,000 fps using optical fiber illumination. *Opt. Express* **18**, 22722-22733, (2010).

8 Smith, C. S., Joseph, N., Rieger, B. & Lidke, K. A. Fast, single-molecule localization that achieves theoretically minimum uncertainty. *Nat. Methods* **7**, 373-U352, (2010).

9 Flyvbjerg, H., Mortensen, K. I., Churchman, L. S. & Spudich, J. A. Optimized localization analysis for single-molecule tracking and super-resolution microscopy. *Nat. Methods* **7**, 377-U359, (2010).

10 Brown, L. G. A survey of image registration techniques. *Computing Surveys* **24**, 325-376, (1992).

11 Berglund, A. J., McMahon, M. D., McClelland, J. J. & Liddle, J. A. Fast, bias-free algorithm for tracking single particles with variable size and shape. *Opt. Express* **16**, 14064-14075, (2008).

12 Gelles, J., Schnapp, B. J. & Sheetz, M. P. Tracking kinesin-driven movements with nanometre-scale precision. *Nature* **331**, 450-453, (1988).

13 Gosse, C. & Croquette, V. Magnetic tweezers: Micromanipulation and force measurement at the molecular level. *Biophys. J.* **82**, 3314-3329, (2002).

14 Cheezum, M. K., Walker, W. F. & Guilford, W. H. Quantitative comparison of algorithms for tracking single fluorescent particles. *Biophys. J.* **81**, 2378-2388, (2001).

15 Alexander, B. F. & Ng, K. C. Elimination of systematic-error in subpixel accuracy centroid estimation. *Opt. Eng.* **30**, 1320-1331, (1991).

16 Dejong, P. G. M., Arts, T., Hoeks, A. P. G. & Reneman, R. S. Determination of Tissue Motion Velocity by Correlation Interpolation of Pulsed Ultrasonic Echo Signals. *Ultrasonic Imaging* **12**, 84-98, (1990).

17 Carter, B. C., Shubeita, G. T. & Gross, S. P. Tracking single particles: a user-friendly quantitative evaluation. *Phys. Biol.* **2**, 60-72, (2005).

18 Kim, K. & Saleh, O. A. A high-resolution magnetic tweezer for single-molecule measurements. *Nucleic Acids Res.* **37**, (2009).

19 Zhang, Z. P. & Menq, C. H. Three-dimensional particle tracking with subnanometer resolution using off-focus images. *Appl. Optics* **47**, 2361-2370, (2008).

20 Hell, S., Reiner, G., Cremer, C. & Stelzer, E. H. K. Aberrations in Confocal Fluorescence Microscopy Induced by Mismatches in Refractive-Index. *J Microsc-Oxford* **169**, 391-405, (1993).

21 Ribeck, N. & Saleh, O. A. Multiplexed single-molecule measurements with magnetic tweezers. *Rev. Sci. Instrum.* **79**, 6, (2008).

22 De Vlaminck, I. *et al.* Highly Parallel Magnetic Tweezers by Targeted DNA Tethering. *Nano Lett.* **11**, 5489-5493, (2011).

23 Kruithof, M., Chien, F., de Jager, M. & van Noort, J. Subpiconewton dynamic force spectroscopy using magnetic tweezers. *Biophys. J.* **94**, 2343-2348, (2008).

24 De Vlaminck, I., Henighan, T., van Loenhout, M. T. J., Burnham, D. R. & Dekker, C., (unpublished work).

25 Marko, J. F. & Siggia, E. D. Stretching DNA. *Macromolecules* **28**, 8759-8770, (1995).

26 Bustamante, C., Marko, J. F., Siggia, E. D. & Smith, S. Entropic Elasticity of Lambda-Phage DNA. *Science* **265**, 1599-1600, (1994).

27 Charvin, G., Vologodskii, A., Bensimon, D. & Croquette, V. Braiding DNA: Experiments, simulations, and models. *Biophys. J.* **88**, 4124-4136, (2005).

28 Shtengel, G. *et al.* Interferometric fluorescent super-resolution microscopy resolves 3D cellular ultrastructure. *Proc. Natl. Acad. Sci. USA* **106**, 3125-3130, (2009).

29 Ramanathan, S. P. *et al.* Type III restriction enzymes communicate in 1D without looping between their target sites. *Proc. Natl. Acad. Sci. USA* **106**, 1748-1753, (2009).

30 Pertsinidis, A., Zhang, Y. X. & Chu, S. Subnanometre single-molecule localization, registration and distance measurements. *Nature* **466**, 647-U611, (2010).

Supplementary information

S2.1: Center of mass (COM)

Centroid or Center-of-mass tracking is commonly used in tracking and serves as a reference algorithm[1-3]. The COM algorithm is computationally efficient, but is very sensitive to a non-zero-background in the image, which can be especially prominent in bright-field imaging microscopy. A first step, therefore, is to subtract the image median, I_{med}, and take the absolute value $I_{ij'} = |I_{ij} - I_{med}|$. Then, the center-off-mass x-position is given by:

$$X_{com} = \sum_{i=1}^{n}\sum_{j=1}^{n}\left(i \bullet I_{ij'}\right) \Big/ \sum_{i=1}^{n}\sum_{j=1}^{n} I_{ij'} \qquad \text{(S2.1)}$$

were i is an integer pixel x-coordinate in an n by n image. The y-position is calculated analogously.

S2.2: Cross-correlation (XCOR)

Cross-correlation is generally the preferred method for tracking a non-diffraction-limited object and is used here as a second reference algorithm[2,4-6]. This method is based on selecting a band of pixel rows (for the x-direction) from the image and cross correlating its average intensity profile with its own mirror profile. In this work, we average a band of width $0.2n$ around the image center line, for an n by n image:

$$P(i) = \frac{1}{0.2 \cdot n} \sum_{j=0.4n}^{0.6n} I_{ij}. \qquad \text{(S2.2)}$$

For computational efficiency the cross-correlation C_{xx} of *P(i)* with its mirror image *P(-i)* is performed in Fourier space

$$C_{xx} = \mathrm{IFFT}(\mathrm{FFT}(P(i)) * \overline{\mathrm{FFT}(P(-i))}). \qquad \text{(S2.3)}$$

If the pattern is off the symmetry axis by a distance δ_x, the correlation curve will exhibit a peak at position $2\delta_x$. Following common practice, we perform a 5-point parabolic fit around this peak to obtain a sub-pixel position, $2\delta_x$. The relative center position finally is $x = 2\delta_x/2$.

S2.3: A simple estimate of resolution

Here we describe a simple analytical expression to estimate the resolution of positional information in an image, irrespective of how this information is extracted. Methods to estimate this accuracy have been proposed, either by simulating the full optical path, or by determining the Cramer-Rao lower bound of a specific image with respect to translational motion[8,31,32]. While the first method requires careful optical evaluation of the imaging system (which is often difficult to achieve), the latter requires elaborate analysis especially when the shape of the object is not a priori known.

Here we propose a simple alternative method that allows a quick, order-of magnitude estimate of the obtainable resolution which only requires a single image. Such a quick estimate can be performed online while tuning an optical system, thus allowing easy optimization of illumination levels, magnification, etc.

We assume an image consists of $n \cdot n = N$ pixels, mapping out an image pattern and background noise. Next, we note that, for any tracking algorithm, the part of the image with the steepest intensity gradient will contribute most to the accuracy, where, by contrast, areas with the weakest gradient will contribute little information on any translation. Likewise, noise at the position of steep intensity gradients will perturb the tracking most. Thus, instead of considering all pixels we simplify our problem by only considering the 'contributing pixels', N_c, defined by setting a threshold to detect the regions of high intensity gradient. Moreover, we assume that a one-pixel translation will change the value of a contributing pixel over the signal range *S,* defined as the difference between maximum and minimum intensity of the pattern. A sub-pixel motion of δ (in pixel units) will analogously cause a change of $\delta \cdot S$. Note that, although in doing this we oversimplify the information content of the image, the effects of ignoring both noise and signal contributions outside the steep parts will partly cancel each other out.

The smallest detectable change in intensity will be on the order the background noise σ. If we take a 95 percent confidence level, the smallest detectable motion is then simply $4\sigma/S$. Since we have N_c of such measurements, we arrive at a simple estimate for the image resolution δ_x,

$$\delta_x = \frac{1}{S}\sqrt{\frac{(4\sigma)^2 + 0.5}{N_c}}, \qquad \text{(S2.4)}$$

where a least-significant-bit error of 0.5 bit was added. In practice, representative values of the various parameters should be obtained from an experimental image. We assume a typical image pattern, having continuous, noisy edges with the pattern of interest roughly centered. The noise relevant for *x*-directed motion is taken by differentiating the two top and two bottom pixel rows yielding a standard deviation σ, for the background intensity as a measure for the noise:

$$\sigma \approx \left\langle std(P_i(j) - P_i(j-1)) \right\rangle / (2), \qquad \text{(S2.5)}$$

where *std* denotes the standard deviation. The average signal range is simply estimated as the difference of the image maximum and minimum intensity, $S=(I_{max}-I_{min})$. To determine the number of edges in the *x*-direction, we apply a threshold to the image. Using equation (2.1), we may evaluate and compare various images intended for tracking, such as the typical defocusing rings of beads used in magnetic tweezers experiments.

We note that this approach not only allows estimating translational precision, but also potentially for motions in the *z*-direction (defocusing changes) and rotation (angular changes). For such motions, it is necessary to identify the average change of pixel pattern upon one unit of motion (a defocus step, or a degree of rotation) and the relevant edges.

Supplementary references

1 Carter, B. C., Shubeita, G. T. & Gross, S. P. Tracking single particles: a user-friendly quantitative evaluation. *Phys Biol* **2**, 60-72, (2005).

2 Cheezum, M. K., Walker, W. F. & Guilford, W. H. Quantitative comparison of algorithms for tracking single fluorescent particles. *Biophys. J.* **81**, 2378-2388, (2001).

3 Berglund, A. J., McMahon, M. D., McClelland, J. J. & Liddle, J. A. Fast, bias-free algorithm for tracking single particles with variable size and shape. *Opt. Express* **16**, 14064-14075, (2008).

4 Gelles, J., Schnapp, B. J. & Sheetz, M. P. Tracking kinesin-driven movements with nanometre-scale precision. *Nature* **331**, 450-453, (1988).

5 Otto, O. *et al.* Real-time particle tracking at 10,000 fps using optical fiber illumination. *Opt. Express* **18**, 22722-22733, (2010).

6 Gosse, C. & Croquette, V. Magnetic tweezers: Micromanipulation and force measurement at the molecular level. *Biophys. J.* **82**, 3314-3329, (2002).

Twist

Part 1—Twist

DNA is twisted. Not only are the two strands that make up a DNA molecule twisted to form a helix but in cells the DNA helix itself is also additionally twisted on a larger scale to form loops that compact the genome. This supercoiling affects DNA in a remarkable way. Like a bottle with a cap we may either twist it to open it, allowing its contents out, or twist the cap on locking its contents inside. Similarly a DNA molecule may be negatively supercoiled to open up and make its information accessible, as in this state the single bases that make up the DNA more easily flip out to be read and processed by proteins. If the DNA is positively supercoiled, however, 'the cap goes on' and the bases become more tightly locked, protecting the information inside and limiting the access of proteins. Because supercoiling is so important, cells tightly regulate it using proteins called topoisomerases and gyrases that can decrease or increase the amount of supercoiling. The second way by which DNA supercoiling affects cells is by the creation of larger loops, called plectonemes, which change the conformation of the genome and brings distant DNA regions together. Supercoiling results from all the proteins that move along the DNA helix, thereby generating waves of over- and underwinding. In this part of the thesis we describe measurements that for the first time directly show how the loops or plectonemes present in a supercoiled DNA molecule move and jump along the DNA. These measurements provide insight into how cells may deal with the waves of supercoiling generated by moving proteins and also show how cells may uses plectonemes to quickly rearrange the spatial conformation of their genome.

CHAPTER 3

Dynamics of DNA supercoils

DNA in cells exhibits a supercoiled state where the double helix is additionally twisted to form extended intertwined loops called plectonemes. Although supercoiling is vital to many cellular processes, its dynamics remain elusive. Here we directly visualize the dynamics of individual plectonemes. We observe that multiple plectonemes can be present and that their number depends on applied stretching force and ionic strength. Plectonemes moved along DNA by diffusion or, unexpectedly, by a fast hopping process that facilitated very rapid (< 20 milliseconds) long-range plectoneme displacement by nucleating a new plectoneme at a distant position. The observations directly reveal the dynamics of plectonemes and identify a new mode of movement that allows long-distance reorganization of the conformation of the genome on a millisecond time scale.

This chapter has been published as: M. T. J. van Loenhout, M. V. Grunt, and C. Dekker. Dynamics of DNA supercoils. *Science* **338**, 94 (2012).

Introduction

Supercoiling and changes in the supercoiling state are ubiquitous in cellular DNA and affect virtually all genomic processes[1]. Proteins moving along the helical path of the DNA, for example, generate torsional stress, which produces twist (the over- or under-winding of the DNA double helix around its axis) and writhe (the coiling of the duplex axis around itself). Supercoiling affects the cell because it alters the conformation of the genome on two basic levels. First, supercoiling can induce local changes in the DNA structure such as a locally destabilized or deformed duplex, which subsequently affect transcription or trigger protein binding[2-4]. Second, supercoiling can induce global changes in the conformation of the genome, which brings distant DNA sequences together thereby facilitating DNA compaction and site-specific recombination[5,6]. Genomic DNA is organized in topological domains of 10-100 kilobases (kb) which isolate topological changes from neighboring regions[1]. Within these regions, torsion can rapidly transmit allowing for long-range communication between distant genomic locations[7]. To understand the cellular processes that are affected by supercoiling, it is essential to know its dynamics, of which virtually nothing is known. Do plectonemes move along DNA, and if so, by which process, and on what time scale? We address these questions at the single-molecule level.

The dynamics of supercoiled DNA has been difficult to address experimentally. Static images obtained by electron and atomic force microscopy showed that supercoiled plasmids are plectonemic[8,9]. These images could not capture the dynamics as DNA needs to be strongly immobilized for imaging. Measurements of site-specific recombination have provided indirect information on the speed of juxtaposition of DNA positions, but these results did not agree with theoretical predictions, leaving many unanswered questions on the dynamics on the diffusion speed of plectonemes[6,10]. Single-molecule magnetic tweezers have proven to be an ideal platform to study DNA mechanics as they allow to twist and apply a stretching force to a single DNA molecule[11]. In the traditional implementation this technique is limited, however, as it only measures the DNA end-to-end distance and not the position of a plectoneme along the DNA tether.

Methods

To visualize the dynamics of plectonemes directly along a single DNA molecule, we developed a magnetic tweezers apparatus (Figure 3.1A) that pulls a fluorescently labeled DNA molecule sideways and visualizes it along its length using epi-fluorescence. Each experiment started by creating plectonemes in a DNA molecule that was torsionally constrained between the flow cell surface and a magnetic bead by coiling it with a pair of magnets positioned above the DNA tether. The DNA molecule was positively supercoiled to a degree where 25 % of the DNA contour length was put into a plectonemic state (Figure 3.1B, Supplementary information Table S3.1). Subsequently, an additional

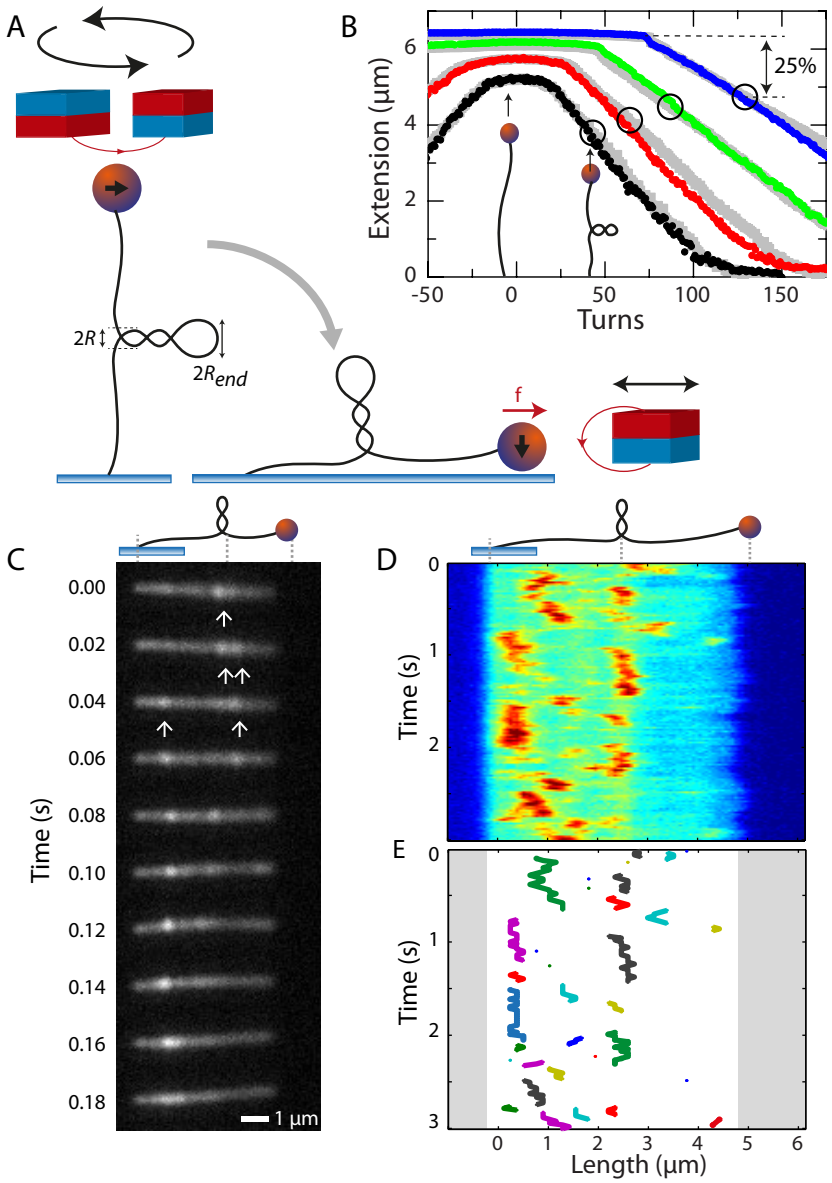


Figure 3.1: Visualization of plectonemes by fluorescence with side-pulling magnetic tweezers. (**A**) A DNA molecule is supercoiled by rotating a pair of magnets and is subsequently pulled sideways with an additional magnet into the focal plane of a high numerical aperture objective. (**B**) DNA rotation curves of a Cy3-labeled 21 kb molecule in a 300 mM NaCl buffer at various forces (black to blue: 0.4, 0.8, 1.6. 3.2 pN) and a non-labeled molecule (gray points). Experiments were performed at conditions where 25 % of the DNA contour

length was put in a plectonemic state (open circles). (**C**) Images at consecutive 20 ms intervals of a supercoiled DNA molecule showing plectonemes (arrows) at a force of 0.8 pN in a buffer containing 150 mM NaCl. (**D**) Kymograph of the supercoiled DNA molecule shown in panel D constructed by summation of the pixel intensities perpendicular to the stretching direction of the molecule, plotting the intensity (color scale) versus position and time. Plectonemes are visible as high-intensity regions in the kymograph. (**E**) Same kymograph as in panel D after background subtraction and peak fitting (Figure S3.3–S3.5). Individual plectonemes identified in consecutive frames are indicated with different colors.

magnet was brought near to side of the flow cell and the top magnets were removed, thereby pulling the DNA tether sideways at modest stretching forces (0.4–3.2 pN), which are readily generated by, for example polymerases *in vivo*[12]. All measurements were performed on 21-kb DNA molecules, similar in length to the topological domains observed in genomic DNA[1]. The DNA molecules were covalently labeled with Cy3 dyes attached by a long (~3 nm) linker at random positions along their length resulting in a low labeling density of ~1/25 bp. An efficient oxygen scavenging system[13] allowed us to monitor supercoiled DNA molecules for several seconds before photo-induced nicking occurs, which released all torsional stress. Cy3 labeling did not affect the mechanical properties of the DNA, as shown by the rotation curves of labeled molecules which overlap with those of unlabeled molecules (Figure 3.1B, Figure S3.1)[14].

Results

We observed individual plectonemes in supercoiled DNA molecules in images acquired with 20-ms time resolution. As illustrated in Figure 3.1C and Movie S3.1[14], plectonemes appear in the images as bright spots along the DNA, as the local DNA density is higher in the plectonemes. These bright spots were near-diffraction-limited with a typical spot size of ~500 nm. The spots were not present in non-supercoiled DNA and disappeared instantly if a DNA molecule became nicked (Figure S3.2)[14].

We observed that multiple plectonemes were present that appeared and disappeared and that moved along the DNA. The dynamics of plectonemes were analyzed by converting the time series of images to kymographs, which plot the intensity profile along the DNA position versus time (Figure 3.1D). Within the 2s time scale of the kymograph in Figure 3.1D, we observed many events where plectonemes nucleated, moved, and disappeared some time later. A fitting routine was developed to count and extract their position over time (Figure S3.3–S3.5)[14]. A typical result is shown in Figure 3.1E where individual plectonemes are marked by colored lines.

We find that the number of plectonemes in the DNA varies substantially with ionic strength and applied force. Experiments were performed for a range of ionic strengths ([NaCl] = 20 to 300 mM) and forces (f = 0.4 to 3.2 pN) (Figure 3.2A, Movie S3.2)[14]. Plectonemes were found to be very dynamic at low forces and quickly moved between positions along the DNA. At higher forces, the dynamics become restricted and single

plectonemes remain at the same position for long periods in time (several seconds), giving rise to almost static high intensity bands in the kymographs. The number of plectonemes present at any given time varied from a single one for high-force and high-salt conditions to about three for low-force and low-salt conditions (Figure 3.2B, C). The mean size of individual plectonemes varied between ~1.7 and ~5.3 kb, which is consistent with the ~2.3 kb observed by electron microscopy[8] and predictions of Monte Carlo simulations[15]. Our results confirm the general trends described in two recent theoretical studies by Emanuel *et al.*[16] and Marko *et al.*[17], which predict the presence of multiple plectonemes for low force (< 0.5 pN), or low salt concentrations (< 50 mM). At higher forces and increased salt concentrations (≥ 1pN, > 50 mM) Marko *et al.*[17] predict the presence of only a single plectonemic domain in a 10 kb DNA molecule, whereas we observed that a few plectonemes could still be present under these conditions.

The number of plectonemes present in a DNA molecule will be set by the free energy balance between the change in enthalpy and the change in entropy upon the formation of an additional plectoneme[15]. Entropy will favor the presence of multiple plectonemes as they can occupy multiple positions along the molecule and distribute the writhe between them[16,17]. The energy cost required to bend the DNA in the plectoneme, however, will favor a single plectoneme. The structure of a plectoneme can be simplified to consist of an intertwined section and an end loop (Figure 3.1A). The formation of an end loop with size $2R_{end}$ is energetically more costly than extending the intertwined region, making it unfavorable to form multiple plectonemes[18]. Surprisingly, the data showed that multiple plectonemes were present in DNA molecules.

The observed number of plectonemes decreased for higher ionic strengths (Figure 3.2B, C), which may be understood by considering the plectoneme structure. Increasing ionic strength will screen the electrostatic repulsion between the highly charged DNA backbones and leads to a reduction of the plectoneme radius (R) compared to the end loop radius (R_{end}) which is set by the mechanical bending of the DNA[19,20]. At high ionic strength, $R < R_{end}$ which favors a single plectoneme, but at low ionic strength R increases and becomes approximately equal to R_{end}, reducing the free energy penalty for forming additional plectonemes. Indeed, this response was experimentally observed as the number of plectonemes increased with decreasing ionic strength (Figure 3.2B, C). The applied stretching force did not have a strong influence on the number of plectonemes (Figure 3.2B). In contrast to ionic strength, an applied stretching force would reduce both R and R_{end} making the free energy penalty for the formation of an additional plectoneme rather insensitive to force.

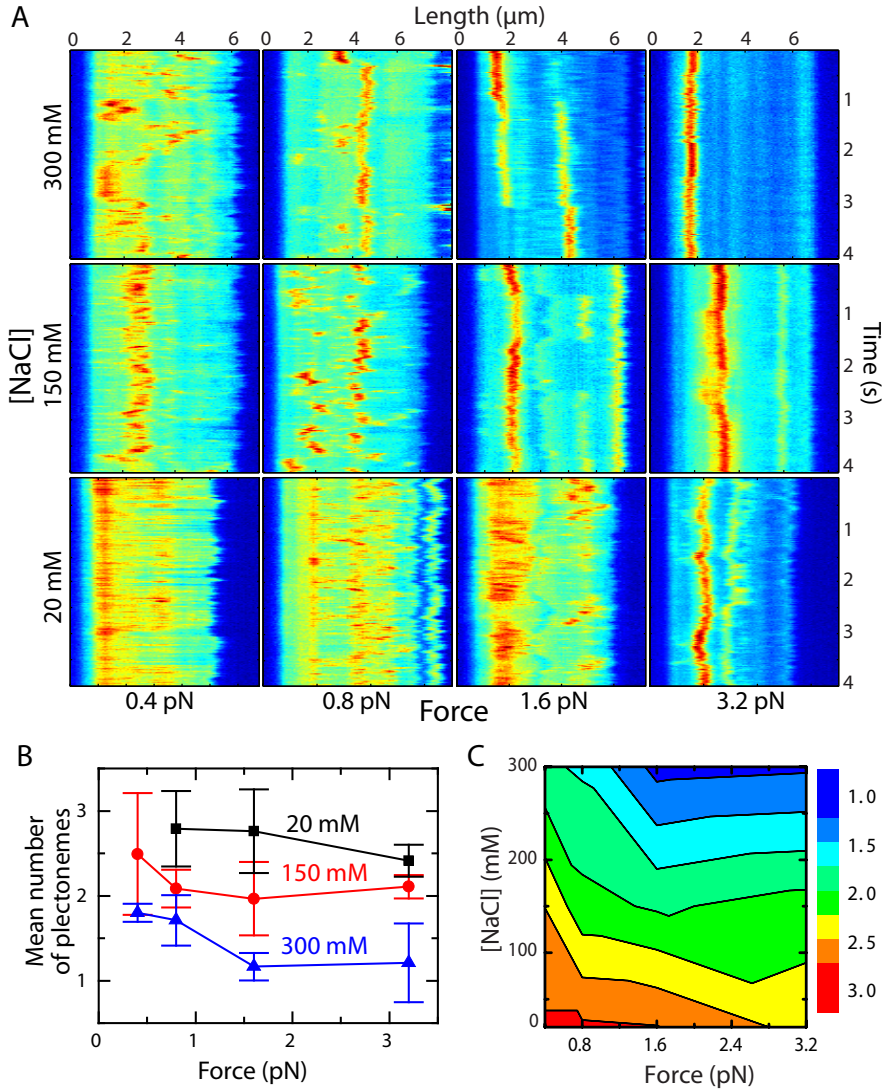


Figure 3.2: The number of plectonemes and their dynamics depend on the applied stretching force and ionic strength. (**A**) Kymographs of DNA molecules for 12 different stretching force / ionic strength combinations (also available as Movie S3.2). (**B**) The mean number of plectonemes increases with decreasing ionic strength (n ≥ 4 DNA molecules per data point, error bars indicate the standard deviation (**C**) Phase diagram showing the dependence of the number of plectonemes on applied stretching force and ionic strength.

We now turn to the dynamics of plectonemes. Here, unexpectedly, we observed two different types of motion. First, diffusive motion where a plectoneme randomly moved along a DNA molecule (Figure 3.3A–C). Second, hopping where a plectoneme suddenly shrank or disappeared while simultaneously a new plectoneme nucleated at a different location (Figure 3.4A–C). We first focus on diffusion. To quantify the diffusive behavior of plectonemes, we tracked the position of individual plectonemes with an extended lifetime (≥ 0.3 s) (Figure 3.3B and Figure S3.6)[14]. An example of the diffusional tracks for 24 plectonemes is shown in Figure 3.3D. To extract the plectoneme diffusion constant (D), we calculated the mean squared displacement (MSD), $\langle \Delta x^2(t) \rangle$, where Δx is the plectoneme displacement along the DNA and $\langle\rangle$ denotes the time average (Text S3.1, S3.2)[14]. As Figure 3.3E shows, we observe a near-linear relation of the MSD with time, $\langle \Delta x^2(t) \rangle = 2Dt$, which is characteristic for Brownian motion, indicating that plectonemes indeed move along DNA by one-dimensional diffusion, similar to the diffusion of knots in DNA[21].

Diffusion of a plectoneme requires not only the sideways motion of the plectoneme, but also the slithering motion of the DNA within the supercoil (Figure 3.3C)[22]. Remarkably, the diffusional constants that were experimentally observed were substantially lower than those predicted for the hydrodynamic drag of plectonemes (Figure 3.3F; Text S3.3)[14] and showed a rapid decrease with applied stretching force (Figure 3.3E, F). Surface effects caused by the proximity of the DNA molecule to the surface cannot explain such a large reduction in diffusion speed as they result only in a very small increase in the viscous drag (Text S3.4)[14].

The unexpectedly slow diffusion of plectonemes can originate from different microscopic causes, which in there most general form, may be represented by the presence of a rugged energy landscape that the plectoneme must navigate while moving along the DNA molecule. This energy landscape can derive from sequence-dependent mechanical properties such as the intrinsic curvature and bendability of the DNA that create local energy barriers for diffusion[23]. Indeed, we observed that plectonemes are not fully randomly distributed along a DNA molecule, but have some preference for certain positions along the DNA (Figure 3.2A).

We can quantitatively estimate the effect of a rugged energy potential on the diffusive motion of a plectoneme by expressing the actual diffusion constant D as the product of the hydrodynamic diffusion constant D_h and a retardation factor $F(\varepsilon)$. If we assume that the fluctuating part of the potential energy landscape along the DNA molecule obeys a Gaussian distribution, then $F(\varepsilon) = \exp[-(\varepsilon/k_bT)^2]$, where ε denotes the root-mean-square (rms) fluctuations in the energy potential[24]. As the bending energy for looping DNA scales as the square root of the applied stretching force[25], we propose a rugged-energy model with $\varepsilon(f) \propto \sqrt{f}$. Indeed, this provides an excellent fit (green line Figure 3.3F) to the experimental data compared to the simple hydrodynamic model (black line in Figure 3.3F; Figure S3.7)[14].

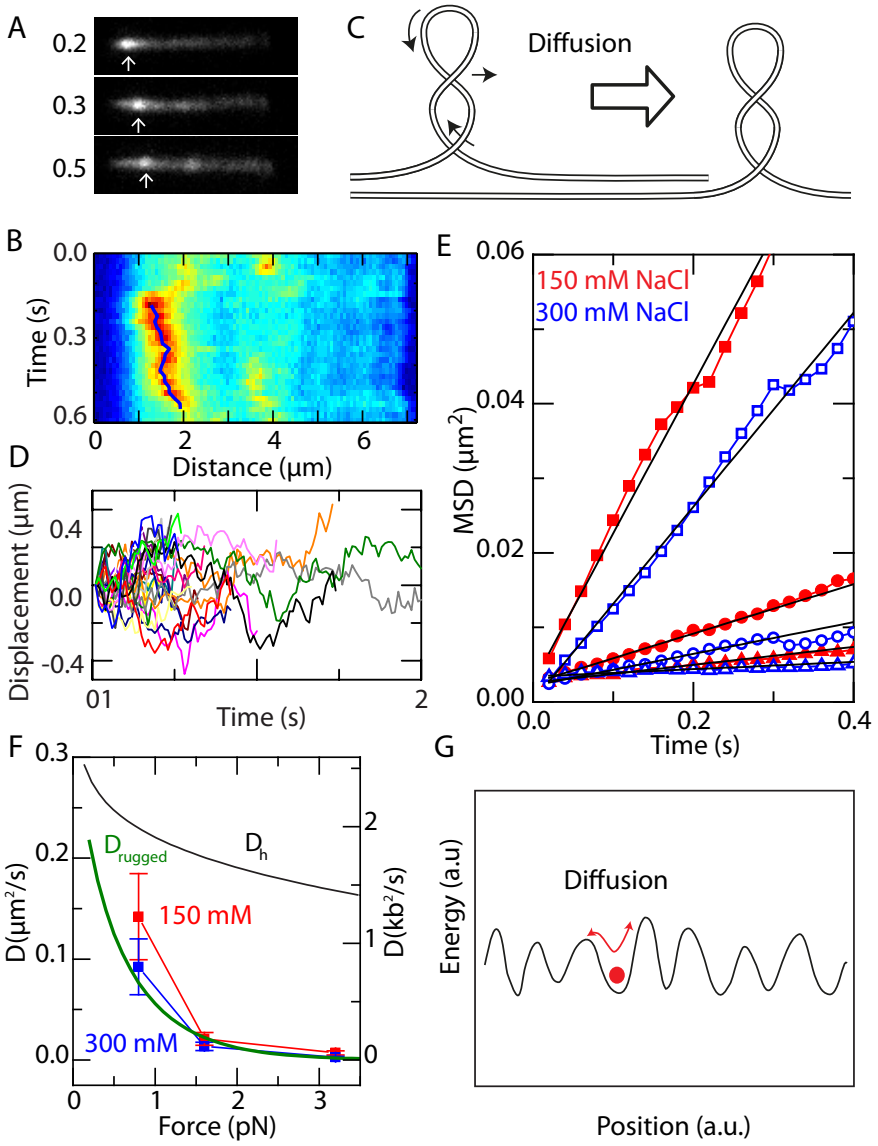


Figure 3.3: Plectoneme diffusion along a DNA molecule. (**A**) Images and (**B**) kymograph of a plectoneme diffusing along a DNA molecule (150 mM NaCl, 0.8 pN). Solid line denotes the diffusional track of the plectoneme as obtained by a fit to the intensity in the kymograph (blue line, Figure S3.6). (**C**) Diffusion of a plectoneme involves the sideways movement of the DNA but also a slithering motion within the plectoneme. (**D**) Example of diffusion traces (**E**) Mean squared displacement (MSD) of plectoneme positions at 0.8 pN

(squares), 1.6 pN (circles) and 3.2 pN (triangles) stretching force. Lines are linear fits to the MSD (Text S3.1). (**F**) Diffusion constants, corrected for the DNA extension at the experimentally probed forces (Text S3.1), obtained from linear fits of panel e. The observed diffusion constants decrease are much lower than those predicted from a hydrodynamic diffusion model (black line, Text S3.3). A model incorporating a rugged energy landscape (green line, Figure S3.7) fits the 300 mM data well. (**G**) Schematic diagram of plectoneme diffusion in a rugged energy landscape along a DNA molecule.

We found values for ε of 1 $k_B T$ at 150 mM and 0.8 pN to 2 $k_B T$ at 300 mM and 3.2 pN, which represent only a small modulation compared to the bending energies of the end loop which are 14 and 28 $k_B T$ respectively[25]. Importantly, the values of ε are close to the thermal energy, allowing for diffusion to occur, albeit at a reduced speed.

The second mechanism whereby plectonemes moved along DNA is hopping, where one plectoneme suddenly shrank or vanished and a new one simultaneously nucleated at a different position along the DNA (Figure 3.4A–C). These hopping events were generally fast, occurring abruptly within the 20 ms of our single-frame time resolution, and spanned large distances that could not be explained by a diffusion mechanism (Figure 3.4A, B; Text S3.5)[14]. The total length of DNA in plectonemes was constant at 25% in our experiments, so shrinking of a plectoneme necessarily resulted in the growing or nucleation of a new plectoneme and the observed nucleation rate (0.3–22 s^{-1}) is equivalent to the hopping rate. This rate decreased with increasing force and ionic strength[18] (Figure 3.4D, Text S3.6)[14], while the mean plectoneme lifetime (0.1–10 s) increased strongly with force and ionic strength (Figure 3.4E). The distribution of plectoneme lifetimes showed a strong decrease with time (Figure 3.4F), i.e. most plectonemes are short lived (< 0.1 s).

We can explain the distribution of lifetimes by considering what happens after nucleation of a plectoneme. It can either grow by absorbing more writhe, or shrink by releasing writhe. The growing and shrinking of a plectoneme can be described by a random walk, and the lifetime of the plectoneme is then set by the first return to the origin of the walk, i.e. the nucleation point. The probability of this first return at time t is given by $P(t) = \binom{t}{t/2} \frac{1}{(t-1)2^t}$, where $\binom{t}{t/2}$ denotes the binomial coefficient. Indeed, this random-walk model described the experimental data well (blue lines in Figure 3.4F; Text S3.7)[14]. A similar scaling of the lifetimes was observed for all experimental conditions (Figure S3.8)[14].

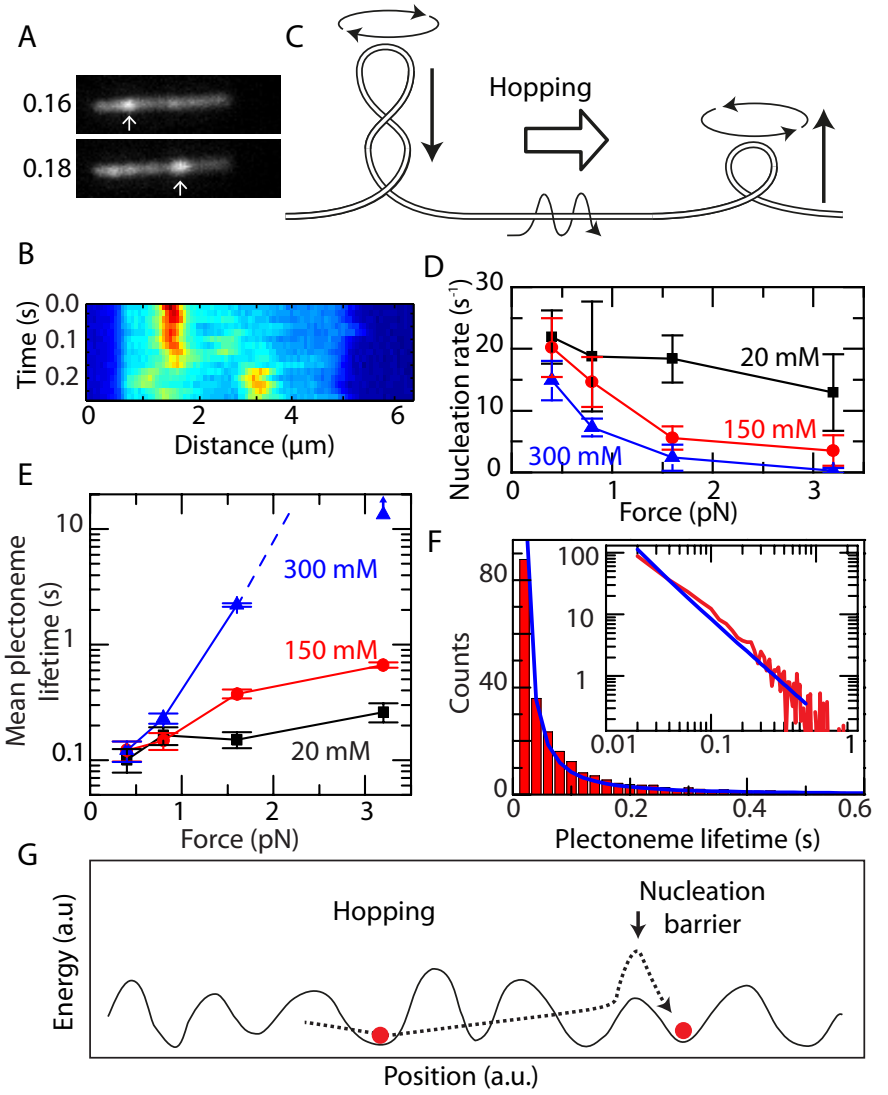


Figure 3.4: Plectoneme hopping along a DNA molecule. (**A**) Images and (**B**) kymograph of a hopping event (150 mM NaCl, 0.8 pN). (**C**) Hopping of a plectoneme involves the nucleation of a new loop at a different location and the transfer of writhe by rotation and sideways motion of the intermediate DNA. (**D**) Nucleation rate of plectonemes as a function of force at different ionic strengths (Text S3.6, n ≥ 4 DNA molecules for each condition; error bars indicate the standard deviation). (**E**) Mean plectoneme lifetime versus force (n ≥ 4 DNA molecules for each condition; error bars indicate the standard deviation, the data point at 3.2 pN, 300 mM NaCl only provides a lower bound as nicking of the DNA molecule limited

the observed lifetimes). (**F**) Histogram of the plectoneme lifetimes at 150 mM NaCl, 0.8 pN; inset shows the same data on a loglog scale. A model for the first-return-to-the-origin of a random walk for the same number of events agrees well with the experimental data (blue lines). (**G**) Schematic diagram showing that hopping can bypass roughness in the energy landscape that retards diffusion along the molecule

Hopping allows plectonemes to move over large distances. We observed maximum hopping distances of up to 15 kb (~5 μm) in our 21-kb DNA molecules (Figure S3.9)[14]. Interestingly, hopping can occur over these large distances irrespective of the rugged energy landscape in between, as the nucleation at the new location is independent of any energy barriers associated with the DNA in between the shrinking plectoneme and the nucleation spot (Figure 3.4G). Hopping is merely restricted by the energy required for nucleation of a new plectoneme and the subsequent transfer of writhe to the new plectoneme by the rotation of the intermediate DNA. The nucleation barrier depends both on applied stretching force and ionic strength, predicting lower nucleation rates at high force and high ionic strength[18], which is indeed the behavior we observe (Figure 3.4D).

The viscous drag associated with the rotation and movement of the intermediate DNA during hopping (Figure 3.4C) grew only slowly with distance and was much smaller than the drag for the diffusion of a plectoneme. Surprisingly, thermal fluctuations led to a hopping distance that scaled linearly with time, in contrast to the distance for diffusive motion, which scaled as the square root of time (Text S3.8)[14]. The lower drag allowed plectonemes to relocate by hopping over many kilobases in a fraction of a second, which would be impossible for diffusion of a plectoneme along the DNA molecule.

The observed dynamics of DNA supercoils reveal how plectonemes change the DNA conformation. Multiple plectonemes were found to be present in a supercoiled DNA molecule under applied force, with a typical density of 1 plectoneme per 10 kb. Diffusion of plectonemes was strongly dependent on applied stretching force, suggesting that it is retarded by local inhomogeneities in the DNA mechanical properties. In contrast, hopping of plectonemes results in a fast long-range rearrangement of the DNA conformation, which may explain the fast search times for site juxtaposition of two distant DNA regions. Hopping can also aid to recruit a plectoneme to a DNA sequence that exhibits inherent curvature or to a site of protein-induced DNA bending. Such a mechanism will allow to change the conformation of the genome at millisecond timescale, thereby triggering protein binding or influencing gene expression.

References

1 Roca, J. The torsional state of DNA within the chromosome. *Chromosoma* **120**, 323-334, (2011).

2 De Vlaminck, I. *et al.* Torsional regulation of hRPA-induced unwinding of double-stranded DNA. *Nucleic Acids Res.* **38**, 4133-4142, (2010).

3 Liu, L. F. & Wang, J. C. Supercoiling of the DNA-template during transcription. *Proc. Natl. Acad. Sci. USA* **84**, 7024-7027, (1987).

4 Gartenberg, M. R. & Wang, J. C. Positive supercoiling of DNA greatly diminishes mRNA synthesis in yeast. *Proc. Natl. Acad. Sci. USA* **89**, 11461-11465, (1992).

5 Kimura, K. & Hirano, T. ATP-dependent positive supercoiling of DNA by 13S condensin: A biochemical implication for chromosome condensation. *Cell* **90**, 625-634, (1997).

6 Parker, C. N. & Halford, S. E. Dynamics of long-range interactions on DNA - the speed of synapsis during site-specific recombination by resolvase. *Cell* **66**, 781-791, (1991).

7 Kouzine, F., Sanford, S., Elisha-Feil, Z. & Levens, D. The functional response of upstream DNA to dynamic supercoiling in vivo. *Nat. Struct. Mol. Biol.* **15**, 146-154, (2008).

8 Boles, T. C., White, J. H. & Cozzarelli, N. R. Structure of plectonemically supercoiled DNA. *J. Mol. Biol.* **213**, 931-951, (1990).

9 Zuccheri, G., Dame, R. T., Aquila, M., Muzzalupo, I. & Samori, B. Conformational fluctuations of supercoiled DNA molecules observed in real time with a scanning force microscope. *Appl. Phys. A-Mater. Sci. Process.* **66**, S585-S589, (1998).

10 Oram, M., Marko, J. F. & Halford, S. E. Communications between distant sites on supercoiled DNA from non-exponential kinetics for DNA synapsis by resolvase. *J. Mol. Biol.* **270**, 396-412, (1997).

11 Strick, T. R., Allemand, J.-F., Bensimon, D., Bensimon, A. & Croquette, V. The elasticity of a single supercoiled DNA molecule. *Science* **271**, 1835-1837, (1996).

12 Wang, M. D. *et al.* Force and velocity measured for single molecules of RNA polymerase. *Science* **282**, 902-907, (1998).

13 Aitken, C. E., Marshall, R. A. & Pulglisi, J. D. An oxygen scavenging system for improvement of dye stability in single-molecule fluorescence experiments. *Biophys. J.* **94**, 1826-1835, (2008).

14 Materials and methods are available as supplementary materials on *Science* Online.

15 Vologodskii, A. V., Levene, S. D., Klenin, K. V., Frankkamenetskii, M. & Cozzarelli, N. R. Conformational and thermodynamic properties of supercoiled DNA. *J. Mol. Biol.* **227**, 1224-1243, (1992).

16 Emanuel, M., Lanzani, G. & Schiessel, H. Multi-plectoneme phase of double-stranded DNA under torsion. *arXiv:1204.1324v2*, (2012).

17 Marko, J. F. & Neukirch, S. Competition between curls and plectonemes near the buckling transition of stretched supercoiled DNA. *Phys. Rev. E* **85**, 011908, (2012).

18 Brutzer, H., Luzzietti, N., Klaue, D. & Seidel, R. Energetics at the DNA supercoiling transition. *Biophys. J.* **98**, 1267-1276, (2010).

19 Stigter, D. Interactions of highly charged colloidal cylinders with applications to double-stranded DNA. *Biopolymers* **16**, 1435-1448, (1977).

20 Neukirch, S. & Marko, J. F. Analytical description of extension, torque, and supercoiling radius of a stretched twisted DNA. *Phys. Rev. Lett.* **106**, 138104, (2011).

21 Bao, X. R., Lee, H. J. & Quake, S. R. Behavior of complex knots in single DNA molecules. *Phys. Rev. Lett.* **91**, 265506, (2003).

22 Marko, J. F. & Siggia, E. D. Fluctuations and supercoiling of DNA. *Science* **265**, 506-508, (1994).

23 Laundon, C. H. & Griffith, J. D. Curved helix segments can uniquely orient the topology of supertwisted DNA. *Cell* **52**, 545-549, (1988).

24 Zwanzig, R. Diffusion in a rough potential. *Proc. Natl. Acad. Sci. USA* **85**, 2029-2030, (1988).

25 Strick, T. R. *et al.* Stretching of macromolecules and proteins. *Rep. Prog. Phys.* **66**, 1, (2003).

Acknowledgments

We thank M. Emanuel and H. Schiessel for discussions on plectoneme energetics, I. De Vlaminck for discussions, D. van der Vlies for experimental assistance, S. Hage and B. Cross for DNA constructs, and J. van der Does for mechanical modifications to the magnetic tweezers. This work has been supported in part by the ERC research grant NanoforBio (#247072).

Supplementary materials and methods

DNA constructs

A 20678 bp pSupercos1 plasmid was made by removal of the MluI fragment from pSupercos1 (Stratagene) and insertion of two lambda fragments. Plasmid DNA was isolated by midiprep (Qiagen) and restricted with XhoI and NotI resulting in a 20666 bp fragment. Biotin and digoxygenin labeled handles were prepared by PCR on a pbluescriptIISK+ template with a taq polymerase (PCR core system I, Promega) and the addition of 2 µl of Biotin-16-2'deoxy-uridine-5'- triphosphate (bio-dUTP,Roche), or 2 µl digoxigenin-11-2'-deoxy-uridine-5'- triphosphate (dig-dUTP,Roche) to the nucleotide mixture respectively. The biotin PCR product was digested with XhoI resulting in 554 bp and 684 bp fragments. The digoxygenin product was digested with NotI resulting in 624 bp and 614 bp fragments. Labeled handles were purified with a nucleospin extractII kit (Machery Nagel), combined with the restricted plasmid DNA and ligated with T4 DNA ligase overnight at 4˚C and subsequently purified by phenol extraction.

DNA labeling

Purified DNA constructs were Cy3 labeled with label-it reagent (Mirus Bio)[26]. Typical reaction conditions used 10 µl DNA construct (250 ng/µl) and 5 µl Label-it reagent in a total volume of 50 µl for 30 min. at 37 ˚C. Excess dye was removed by spin column filtration (illustra ProbeQuant G-50). Labeling density was measured to be 25 ± 5 bp/Cy3 as determined from the DNA and dye adsorption ratio at 260 nm and 550 nm respectively with a spectrophotometer (NanoDrop). For spectrophotometric detection DNA was purified with a nucleospin extractII kit (Machery Nagel) to avoid increased UV-adsorption due to G50 spin columns residues.

Magnetic tweezers

Experiments were performed in square glass capillaries 0.7 × 0.7 × 50 mm (Vitro-Com) coated with a 0.1 % polystyrene in toluene solution and functionalized with anti-digoxygenin Fab fragments[27]. DNA was incubated with 1 µm streptavidin magnetic beads (New England Biolabs) for 10 minutes and flushed into the capillary. The imaging buffer contained 20, 150 or 300 mM NaCl, depending on indicated salt concentrations, 20 mM Tris-HCl pH 7.5, 0.01 % Tween and an oxygen scavenging system[13] consisting of 2.5 mM protocatechuic acid, 50 nM protocatechuate-3,4-dioxygenase, 1 mM Trolox and 100 µM N-Propyl gallate (all from Sigma-Aldrich). A modified microscope (Zeiss Axiovert 200M) equipped with motorized stages was used as magnetic tweezers. Each experiment was started by locating coilable DNA tethers and measuring their rotation curves. Subsequently the molecule was pulled sideways and the position of the side magnet was set to achieve the desired force by measuring the power spectrum of the bead's fluctuations. Next the molecule was pulled upwards again and the desired number of turns were applied (table S1). Finally the molecule was pulled sideways again and the length and force were verified before starting to record images under epi-fluorescent illumination by a 532 nm laser. Images were acquired with 20-ms time resolution using an Andor Neo camera. While the side-pulling geometry brings the DNA molecule within close distance of the surface, we observed no sticking interactions and the DNA molecules were able to move freely (Figure S3.10). Plectonemes also did not show any significant preference for locating at the DNA attachment points to either the surface or the magnetic bead.

Image analysis

Images were analyzed in custom software (Figure S3.3-S3.5, Matlab Mathworks). A fitting routine was developed to count and extract their position over time (Figure S3.3–S3.5). The fitting routine involved background subtraction and peak fitting of the kymograph and achieved a minimum detectable plectoneme size of ~0.5 kb (Figure S3.11).

Supplementary text

Text S3.1: Determination of the diffusion constant

The plectoneme diffusion constant D was determined from the *MSD* given by,

$$MSD(n,N) = \frac{\sum_{i=1}^{N-n} (x_{i+n} - x_i)^2}{N-n}, \qquad \text{(S3.1)}$$

where N is the total number of frames for a single diffusional tract, n the frame interval (time) over which the MSD is calculated, and where the fitted positions of the plectoneme are represented by x_i. The observed diffusion coefficient D was obtained by a weighted-linear-least-squares fit of the *MSD* versus time. The slope of the

MSD was fit up to 0.3 s (1 to 15 frames) and weighted by the expected uncertainty[28] $\delta\mathrm{MSD}(n)=\pm[2n/3(N-n)]^{1/2}$ at frame number n. In order to reduce the inherent statistical variance in the determination of D from individual diffusion trajectories, we averaged the *MSDs* from all plectonemes recorded under the same experimental conditions before fitting.

Plectonemes do not necessarily have a fixed size as the writhe may distribute among several plectonemes in a single DNA molecule in time. As the plectoneme size changes, the diffusion constant of the plectoneme will change accordingly. We therefore only analyzed the diffusional tracts of plectonemes that contained at least 70 % of the total amount of DNA that was in a plectonemic state. The main error in the determination of D now stems from the uncertainty in the size of the plectoneme. We assume an error of 30 % in the plectoneme size, matching the 70 % size criterion used to obtain the diffusional tracts (Fig. S3.6), resulting in an error of $\delta D \approx 0.3\, D$.

The observed *MSD* represents the displacement of the plectoneme in linear space. This is not equal to the displacement of the plectoneme along the contour length of the DNA, as the extension of the DNA, $z=l/l_c$, under an applied stretching force f and torque Γ is shorter than its contour length, l_c. The diffusion constant obtained from the fit of the *MSD* was converted to a diffusion constant along the contour length of the DNA by correcting for this incomplete stretching with a model derived by Moroz and Nelson[29]:

$$z(f,\Gamma)=1-\frac{1}{2}\left[\frac{l_p f}{k_B T}-\left(\frac{\Gamma}{2k_B T}\right)^2-\frac{1}{32}\right]^{-1/2}, \qquad \text{(S3.2)}$$

where l_p denotes the persistence length, k_B is the Boltzmann constant, and T the temperature. We use l_p = 50 nm and T=297 *K*. Since the torque only represents a small correction to the extension at forces between 0.8 and 3.2 pN, we neglect it here and set Γ=0. The diffusion constant along the contour length of the DNA is now given by:

$$D = D_{obs}/z(f,0). \qquad \text{(S3.3)}$$

Text S3.2: Effects of thermal fluctuation in the DNA tether on the observed diffusion of a plectoneme

Changes in the position of a plectoneme are not only due to its diffusion along the DNA, but also result from thermal fluctuations of the DNA tether itself. These fluctuations could potentially result in an overestimate of the observed diffusion constant. Here, however we assess this effect both theoretically and experimentally and we show that its contribution to the observed diffusion rates is minimal.

Both the bead and DNA experience a frequency-independent thermal force from the aqueous environment with noise spectral density $s_F(f)=4\gamma k_B T$, where γ is the hydrodynamic drag coefficient, T is the temperature, and k_B is the Boltzmann constant.

The micron-sized bead has a much larger viscous drag than the DNA and will therefore dominate the low frequency noise displacements that we observe in our 50 Hz bandwidth. By calculating bead displacements following De Vlaminck *et al.*[30] we can thus provide an upper bound for the DNA fluctuations. DNA close to the bead will experience positional fluctuations that are almost equal to those of the bead while these fluctuations will reduce to zero at the flow cell attachment point. The thermal force gives rise to the following power spectral density of the bead position in x, the direction of the applied magnetic force[31]

$$S_x(f) = \frac{k_B T}{\gamma \pi^2 (f_c^2 + f^2)}, \qquad \text{(S3.4)}$$

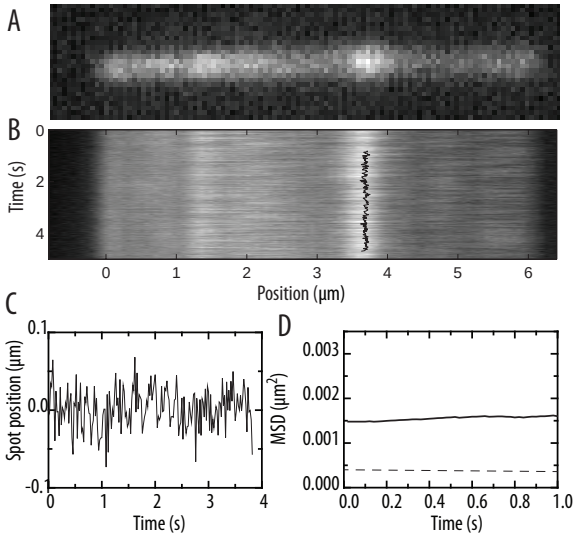
where $f_c = k_{tether}/2\pi\gamma$ is the mechanical response frequency of the DNA-bead tether, and k_{tether} the force- and length-dependent stiffness of the DNA molecule calculated from the worm like chain model. The standard deviation of the bead fluctuations, $\delta x_{thermal}$, along the contour of the DNA molecule then reads[30]:

$$\delta x_{thermal} = \left[\int_0^{B_{eq}} S_x(f) df \right]^{\frac{1}{2}} = \left[\frac{2 k_B T}{\pi \kappa_{tether}} \arctan\left(\frac{B_{eq}}{f_c} \right) \right]^{\frac{1}{2}}, \qquad \text{(S3.5)}$$

where B_{eq} is defined as the equivalent noise bandwidth, which accounts for the low-pass filter effect due to the finite frame rate of the camera. We can now calculate the bead fluctuations for our 21 kb DNA tether at the experimental forces of 0.8, 1.6, and 3.2 pN, which result in a $\delta x_{thermal}$ of 45, 20, and 7.6 nm, respectively. The mean squared displacement (MSD) is then simply given by the variance of the thermal fluctuations which are $2.0*10^{-3}$ μm² $4.0*10^{-4}$ μm² $5.7 \cdot 10^{-5}$ μm², respectively. The MSDs of the bead, which set an upper bound for the DNA fluctuations, are thus much smaller than the experimentally observed MSDs for diffusing plectonemes, as is evident from comparison to Figure 3.3E in the main text.

We can also verify these small thermal DNA fluctuations experimentally. Panel A of the figure below shows a torsionally unconstrained molecule at 1.6 pN applied stretching force, which contains an anomalous bright spot, i.e. this molecule did not contain any plectonemes but did have a (rare) locally bright region. We can track this position following the method outlined in Supplementary Figure S3.6. Parameters of the tracking routine were relaxed to allow the tracking of this spot which had a lower intensity compared to that of plectonemes in the diffusion analysis. The kymograph and tracking shown in panel B and C below confirm that the bright spot remains stationary with only small positional fluctuations. The tracked positions of the bright spot contain fluctuations due to DNA thermal motion, but also due to tracking errors. The MSD calculated for the tracked positions (panel D) is virtually constant in time confirming that the spot does not diffuse but fluctuates around a stationary position. The observed MSD of $1.5 \cdot 10^{-3}$ μm² is higher than the theoretical estimate at 1.6 pN (blue

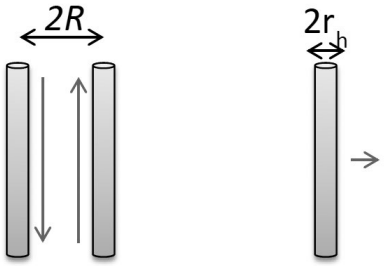
dashed line in panel D), likely due to the additional contribution of tracking errors to the experimentally determined MSD.



Thermal DNA fluctuations in a 21 kb DNA molecule at 1.6 pN. A, Image of an torsionally unconstrained DNA molecule containing a bright spot due to a (rare) inhomogeneous dye distribution. **B**, Kymograph of the molecule shown in panel A. The blue line shows the tracked position of the bright spot. **C**, Tracked position of the bright spot versus time. Fluctuations are observed to be on the order of 50 nm. **D**, MSD of the tracked spot (red line). A near-constant behavior is seen, in contrast to the linear increase that is expected and observed for diffusion (cf. Figure 3.3E in the Main text). The experimental fluctuations, which are due to both intrinsic thermal fluctuations and experimental accuracy in the determination of the spot position, are observed to be higher that the theoretical estimate for thermal fluctuations at 1.6 pN (blue dashed line).

Text S3.3: A simple model for the hydrodynamic drag of a plectoneme

The hydrodynamic drag of a plectoneme can be simplified to consist of two components: First the slithering motion of the DNA within the plectoneme, and second, the perpendicular motion of the whole plectoneme in the direction of the diffusion.



We express the drag coefficient $\mu=k_BT/D$ for these components below in equations S3.6 and S3.7. The first contribution reads[22]:

$$\mu_{slithering} = \frac{2\pi\eta L}{\log(R/r_h)}, \tag{S3.6}$$

here μ is calculated for the motion of two parallel cylinders of hydrodynamic radius r_h =1 nm, spacing 2R, and length L/2, where η is the viscosity of water and L the contour length of the DNA within the plectoneme. We take R to be the plectoneme radius, see equation S3.7 below. The second contribution reads[32]:

$$\mu_{\perp} = \frac{4\pi\eta L}{\log(l_{seg}/r_h) - 0.1}, \tag{S3.7}$$

which gives the ratio of drag force to velocity for sideways movement of a thin cylinder through water, where l_{seg} represents the length of the cylinder over which it can be considered straight. We set this segment length to the experimentally determined length of 59 nm in a single turn of the plectoneme at 150 mM and 0.8 pN. The result is, however, not very sensitive to the exact choice of l_{seg}, as it appears in the logarithm. To address the salt and force dependence of D, we calculate the plectoneme radius R, which depends on the effective charge ν of the DNA. Following the theory developed by Neukirch and Marko[20]:

$$\nu = \frac{1}{b}\frac{1}{\gamma(L_B, b, \kappa_D r_h)}\frac{1}{\kappa_D r_h K_1(\kappa_D r_h)}, \tag{S3.8}$$

where b = 0.17 nm is half of the 0.34 nm spacing of successive base pairs along DNA, L_B is the Bjerrum length in water, κ_D^{-1} is the Debye length, $K_n(x)$ the n^{th} modified Bessel function of the second kind, and parameter γ was interpolated from Table III of Stigter[33] for 20, 150 and 300 mM NaCl and T = 297 K and L_B = 0.7 nm. To lowest order, the plectoneme radius R is now given by:

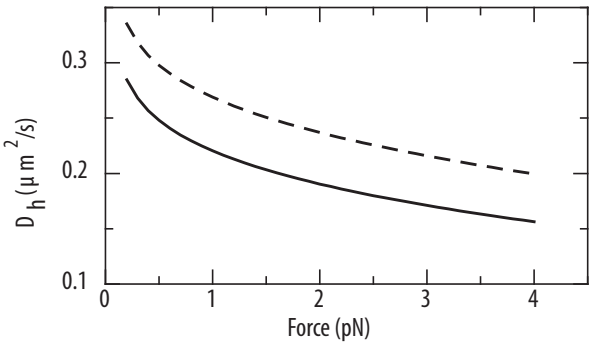
$$R \simeq (2\kappa_D)^{-1} \log(\sqrt{9\pi/8}\nu^2 L_B k_B T / g(f)), \quad \text{(S3.9)}$$

where *g(f)* is, to first approximation, the free energy per unit length of the untwisted stretched DNA molecule[34]:

$$g(f) \simeq f - k_B T\sqrt{f/A}, \quad \text{(S3.10)}$$

with A the bending stiffness of DNA. The hydrodynamic diffusion constant D_h a function of force and ionic strength was finally calculated as:

$$D_h = \left(\frac{\mu_{slithering}}{k_B T} + \frac{\mu_{\perp}}{k_B T} \right)^{-1}, \quad \text{(S3.11)}$$



Hydrodynamic diffusion constant of a plectoneme as a function of force. The hydrodynamic diffusion constant Dh was calculated from Eq. S11, for 150 mM NaCl (dashed line) and 300 mM NaCl (solid line). The plectoneme size was held constant at 1.76 μm and a writhe of 30 turns. The diffusion constant decreases for increasing ionic strength and increasing force. Both of these effects reduce the plectoneme radius R and thereby increase the drag of the slithering motion.

Text S3.4: Surface effects only result in a small increases in the viscous drag experienced by a moving supercoil

The logarithmically decaying flow field of a diffusing plectoneme interacts with the surface, resulting in an increase in viscous drag. The distance of the 21 kb DNA molecule varies from 0 μm at its attachment point to the flow cell to 0.5 μm at its attachment to the 1 μm diameter magnetic bead. The viscous drag coefficient of an infinite cylinder moving parallel to a surface is given by[35] $\mu_{surface}=4\pi\eta L/\log(2b/r_h)$, where b is the distance to the surface. We can compare this drag coefficient to the hydrodynamic drag of the plectoneme in the absence of a surface. If we evaluate the drag at b = 50nm, i.e. for a distance corresponding to 10 % of the contour length of the DNA

from the attachment point, then 90 % of the DNA molecule will be further away from the surface and experience less surface-induced drag. The total drag coefficient for the plectoneme in free solution is given by $\mu_{plectoneme} = \mu_{slithering} + \mu_{\perp}$. Taking a plectoneme radius of R = 2.6 nm, representative for 0.8 pN and 150 mM NaCl, we find that the surface-induced drag coefficient $\mu_{surface}$ at 50 nm from the surface only represents 10 % of the drag of the plectoneme in free solution. For distances further away from the surface, i.e. for 90 % of the molecule the effect will be even smaller. This was confirmed experimentally as there was no obvious difference in diffusion speed between plectonemes located close to the flow cell attachment point and those located midway or even close to the bead.

Surface effects may however play a role in localization of plectonemes along the DNA molecule. In our experiments we only rarely observed plectonemes at the DNA ends. This may be due to the fact that the surfaces (bead and flow cell) restrict the available space for plectoneme movements and create an excluded volume. Such an exclude volume will exert an entropic pressure which forces the plectoneme inward, away from the DNA ends. Surfaces could thus act as soft reflecting boundaries to diffusing plectonemes, which is in line with our observations.

Text S3.5: Hopping of plectonemes over long distances cannot be explained by diffusion

Fast diffusing plectonemes could lead to false identification of nucleation events. Here we estimate this quantitatively. The probability of such large diffusional steps can be calculated from their observed diffusion constants. The mean length l_{dif} of a Brownian motion step during time Δt=δt is given by l_{dif}^2=2$D\Delta t$. If we assume a maximum diffusion constant of 0.5 μm²/s, well above both the predicted theoretical diffusion constants and more than a factor of three above those observed experimentally (Figure S3.7), and given our time resolution Δt=20 ms, the mean displacement is calculated to be l_{dif} = 0.14 μm. This length is also equal to the standard deviation σ of the diffusion distances for Δt =20 ms. The probability of finding a diffusional step larger than 8 pixels, i.e. 520 nm = 3.7 σ, in a single frame is now given by $erfc(3.7/\sqrt{2}) = 2.4 \cdot 10^{-4}$, where *erfc* is the complementary error function. This probability is equivalent to a rate of 0.01 falsely detected nucleation events per second. The experimentally obtained nucleation rates were found to be in the range of 0.4–22 s^{-1}, well above the rate for large diffusional steps. Hopping events therefore cannot be attributed to fast diffusion events but are real nucleation events of a plectoneme at a different position.

Text S3.6: Hopping of plectonemes

To quantify the nucleation rate of plectonemes, and related to that the hopping rate, we used the plectoneme detection method outlined in Figure S3.3–S3.5. A nucleation event was counted as such if it occurred a distance of at least 8 pixel (520 nm) away from a plectoneme in the previous frame. Plectonemes appearing at distances

shorter than 8 pixels were counted as a diffusion event, i.e. a continuation of the nearest preexisting plectoneme.

Text S3.7: Experimental factors influence the observed distribution of plectoneme lifetimes

The observed distribution of plectoneme lifetimes shows a power law decay $P(t) \sim t^{\alpha}$ with an exponent close to -1.5 (see main text Figure 3.4F). Such behavior may be described by the first return to the origin for a random walk. The probability, $P(t)$, for this process reads[36]:

$$P(t) = \binom{t}{t/2} \frac{1}{(t-1)2^t}. \qquad \text{(S3.12)}$$

It has been shown[37] that this first-return-to-the-origin random-walk probability results in a power-law scaling $P(t) \sim t\alpha$ with an exponent of α = -1.5. Experimentally, power law scaling is challenging to identify because it requires a dataset spanning several orders of magnitude[38,39]. Although this random walk model visually agrees well with the histogram of plectoneme lifetimes (Figure 3.4F main text), a more sensitive analysis using the cumulative distribution of plectoneme lifetimes shows that the experimental data considerably deviates from a true power law, see figure below.

There are several experimental factors that may cause the experimentally observed lifetimes to deviate from the theoretical random walk model even if this is the underlying mechanism. Three main differences from the theoretical model are:

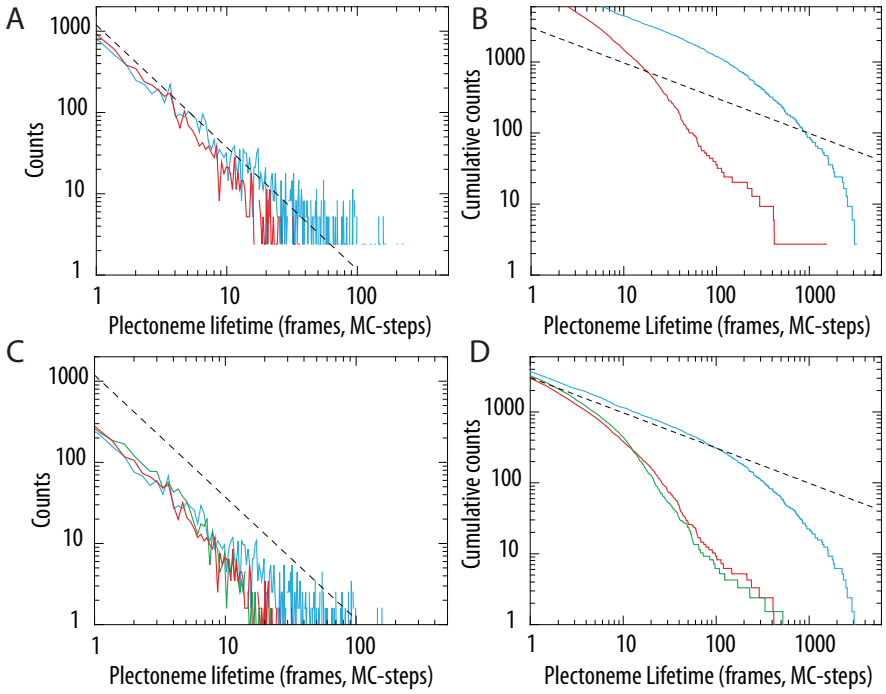
1. Experiments have a finite observation time, due to photo-induced nicking of the DNA.
2. The plectoneme has a finite maximum size, in contrast to the infinite size of the random walk model.
3. Experimental data include errors in plectoneme detection.

To test the effects of these experimental effects, a Monte Carlo (MC) simulation-based approach was employed to determine the first return to the origin of a random walk on a one dimensional lattice for time points $t = 2n$ where n is a MC step. All three deviations of the theoretical model could be accounted for by the MC simulation as follows:

1. The actual measured experimental observation times were used in the MC simulation.
2. The finite plectoneme size was accounted for by including a reflecting boundary condition for the random walk model. The size of the simulation lattice was set equal to the number of turns in the plectoneme, i.e. 30 steps, matching the maximal 30 turns in the plectoneme at 150 mM NaCl 0.8 pN.
3. A reduction in plectoneme lifetime due to errors in the plectoneme detection method was accounted for by including an error of 3 % for the identification of plectonemes, i.e. a plectoneme had a 97 % chance of being detected in a single

frame. In most cases, this is a reasonable number considering the noise in the experimental data and the involved fitting procedure (supplementary Figure S3.3–S3.5). If the plectoneme size is large, however, such an error seems excessive, the error rate was therefore set to 0 % if the size of the plectoneme was more than 50 % of the total DNA that was in a plectonemic state.

The results of the MC simulation incorporating the above effects are shown in the figure below. By incorporating all the expected experimental differences from the random-walk model, an excellent agreement between the MC simulation and the experimental data is obtained.

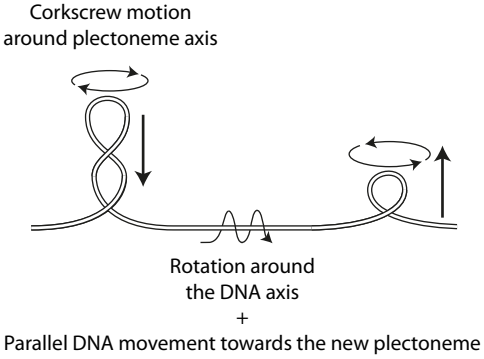


A modified random walk model accounting for experimental conditions accurately captures the observed scaling of plectoneme lifetimes. (**A)** The histogram of experimentally observed plectoneme lifetimes at 150 mM NaCl 0.8 pN (red line) shows an approximately linear behavior on a loglog plot with slope of -1.5 indicating near-power-law scaling and matching a model describing the first return to the origin for a random walk (black dashed line). A MC simulation of the random-walk model for the experimental number of events and accounting for the actual finite observation times due to nicking of DNA molecule (blue line, Supplementary text S3.7) deviates considerably from the experimental data. (**B)** Cumulative distribution of plectoneme lifetimes, created by summation starting

at the longest observed lifetime. While the histogram of plectoneme lifetimes (panel A) is rather insensitive to small deviations from a true power law scaling behavior, a cumulative distribution offers a far more sensitive way to represent the data[38,39]. The cumulative distribution shows that the experimental data (red line) deviates considerably from the expected scaling of the ideal random walk model (black dashed line). The MC simulation using the actual experimental measurement times (blue line) differs from both the theoretical model and the experimental observed distribution. (**C)** Histogram comparing the experimentally observed plectoneme lifetimes with a MC simulation including all experimental effects. The MC simulation including all experimental effects (green line, Supplementary text S3.7) provides a close match to the experimental data (red line). (**D)** Cumulative distribution comparing the observed plectoneme lifetimes with MC simulation results. Both the random walk model (dashed black line) and the MC simulation using experimental measurement times (blue line) differ considerably from the experimental data (red line). The MC simulation including all experimental effects (green line) closely matches the experimental data, red line. These results show that experimental conditions significantly affect the distribution of plectoneme lifetimes and that the experimentally observed distribution is closely reproduced when these effects are taken into account. The proposed random walk model therefore is a good candidate to predict the observed scaling as it presents both a valid physical model for growing and shrinking plectonemes, and, when accounting for experimental factors, accurately describes the observed distribution of lifetimes.

S3.8: The expected hop distance for plectonemes scales, to first order, linearly with time in contrast to the diffusion distance which scales as √t

Hopping involves the corkscrew motion of the growing and shrinking plectonemes around their axes as well as the rotation and sideways movement of the DNA in between the old and new plectoneme. We now quantify the drag associated with hopping by taking these three drag coefficients into account.



The corkscrew motion results in a rotation of the end loop of the plectoneme. We approximate the drag coefficient for this motion by the drag coefficient for perpendicular movement of the DNA through water, $\mu_{\perp}$, as given in equation S3.7. We set the

length of moving DNA equal to twice the contour length of DNA in a single turn of the plectoneme to account for the rotation of both the shrinking and growing plectoneme.

The second contribution to the drag for hopping results from the sideways movement of the intermediate DNA and this drag coefficient is given by[32],

$$\mu_{\parallel} = \frac{2\pi\eta L_{hop}}{\log(l_{ps}/r_h) - 1.2}, \qquad \text{(S3.12)}$$

where L_{hop} is the hopping distance and l_{ps} represents the length over which the DNA may be approximated by a straight cylinder. Following Nelson[40] we set l_{ps} equal to the structural persistence length of DNA, i.e. l_{ps} = 130 nm. An applied stretching force will tend to increase the length over which the DNA can be considered to be straight, this effect will, however, be minor as l_{ps} appears in the logarithm.

The third contribution to the hopping drag results from the rotation of the DNA in between the growing and shrinking plectoneme around its axis. The drag coefficient for axial rotation is given by[41]

$$\mu_{rot} = 4\pi\eta L_{hop} r_h^2. \qquad \text{(S3.14)}$$

To compare the hopping distance with the diffusion distance, we evaluate the typical energy involved in diffusion and calculate the hopping distance for this typical energy. The diffusion distance is given by: $L_{dif} = \sqrt{\left\langle x_{diff}^2 \right\rangle} = \sqrt{2Dt} = \sqrt{2\frac{k_B T}{\mu} t}$, which can be rearranged to $L_{dif}\mu v_{dif} = 2k_B T$, where v_{dif} is the typical diffusion velocity $v_{dif} = L_{dif}/t$. We now calculate the typical hop distance for this $2k_B T$ energy scale from the energy balance,

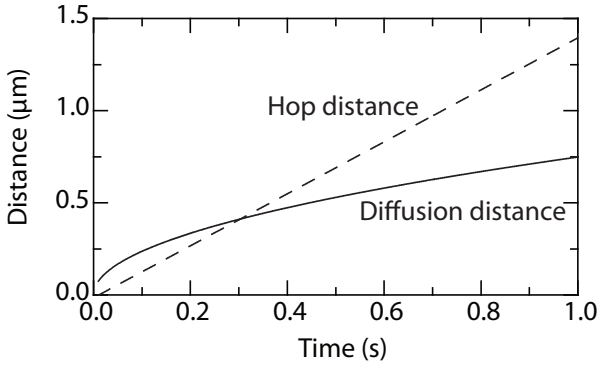
$$2k_B T = \mu_{\perp}(2q)R^2\omega^2 t + \mu_{rot}(L_{hop})\omega^2 t + \mu_{\parallel}(L_{hop}){v_{hop}}^2 t, \qquad \text{(S3.15)}$$

where q is the contour length of DNA in a single turn of the plectoneme, N the number of turns in the plectoneme, $\omega = 2\pi N/t$ the angular frequency of the spinning DNA, and $v_{hop} = L_{hop}/t$ the hop velocity. Solving equation S15 for L_{hop} we find:

$$L_{hop} = \left(2k_B T - \frac{(2\pi RN)^2}{t}\frac{8\pi\eta q}{\log\left(l_{seg}/r_h\right) - 0.1}\right)\frac{t}{N^2 2\pi\eta}\left(8\pi^2 r_h^2 + \frac{q^2}{\log(l_{ps}/r_h) - 1.2}\right)^{-1}. \qquad \text{(S3.16)}$$

Remarkably, equation S3.16 shows that the typical hop distance depends, to first order, linearly on time, in contrast to the expected diffusion distance which scales as $\sqrt{t}$. The linear time dependence of the hop distance allows for much larger displacements of plectonemes by a hopping compared to diffusion. The drag due to the corkscrew motion results in an energy penalty, making diffusion the preferred mode of transport for writhe over short distances. Also noteworthy is the fact that the hop distance scales

as the inverse square of the number of turns in the plectoneme, N, thereby allowing small plectonemes to hop over even larger distances in the same amount of time.



The theoretically expected hop distance for plectonemes scales, to first order, linearly with time in contrast to the expected diffusion distance which scales as $\sqrt{t}$. Diffusion distance l_{dif} (solid line) versus time was calculated as $l_{dif} = \sqrt{2Dt}$ for a diffusion constant of D = 0.28 µm²s⁻¹ representative for a plectoneme of L = 1.76 µm at 0.8 pN and 150 mM NaCl. The distance over which a plectoneme can typically hop within a certain time period (dashed line) scales linear with time making it the preferred mode of movement for large distances. The typical hopping distance was calculated for identical parameters as the diffusion i.e. for a plectoneme size of L = 1.76 µm at 0.8 pN and 150 mM NaCl.

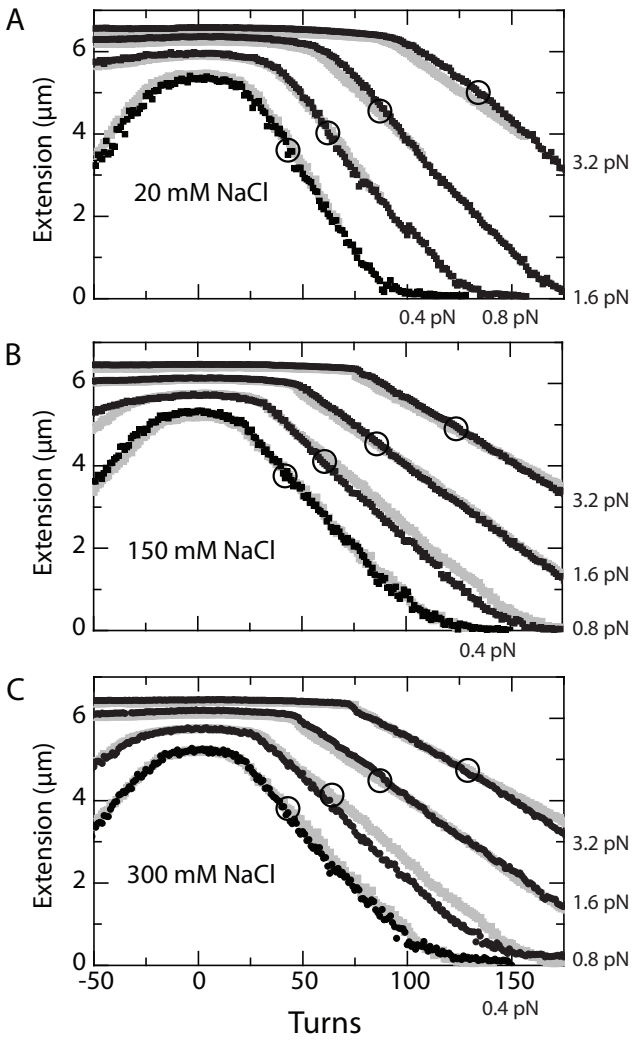


Figure S3.1: Rotation curves of cy3-labeled 21 kb DNA molecules closely match those of unlabeled molecules. (**A**)Rotation curves of cy3-labeled DNA molecules (black points) in a buffer containing 20 mM NaCl for an applied stretching force of 0.4, 0.8, 1.6, and 3.2 pN (bottom to top). The data of the labeled molecules closely match those of unlabeled molecules (grey dots). All plectoneme visualization experiments were performed at conditions where 25 % of the contour length of the DNA molecule was in a plectonemic state, indicated by the open circles. (**B**) Rotation curves similar to panel A at 150 mM NaCl. (**C**) Rotation curves similar to panel A at 300 mM NaCl.

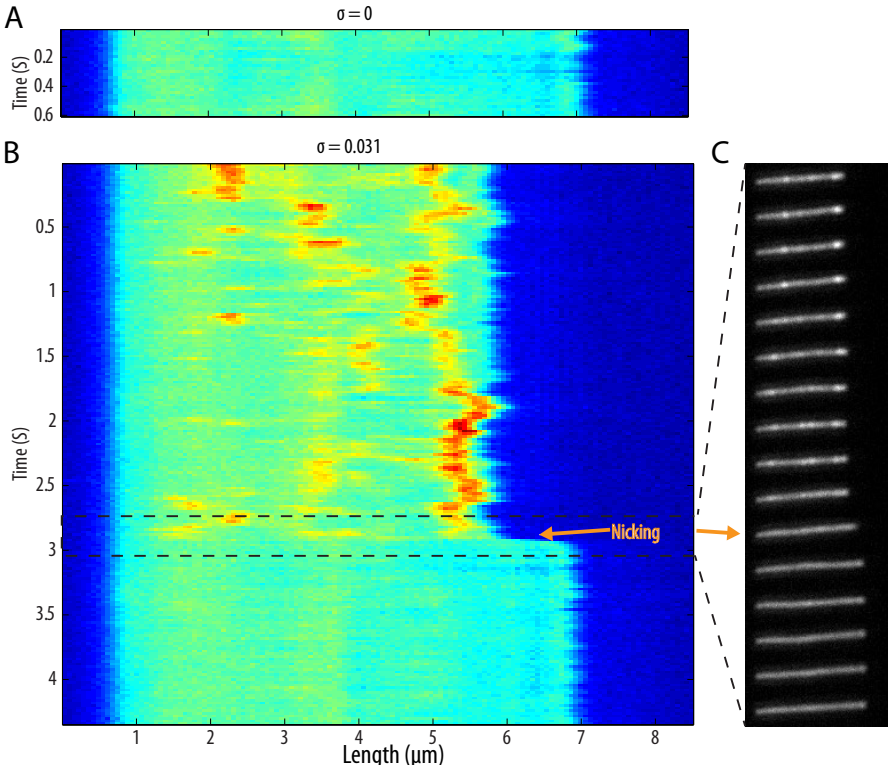


Figure S3.2: Plectonemes are only present in supercoiled DNA molecules. (**A**) Kymograph of a 21 kb DNA molecule at 0.8 pN 150 mM NaCl in the absence of supercoiling shows a homogeneous intensity along the molecule. (**B**) The same DNA molecule as in panel A after applying 61 positive turns, equivalent to a supercoiling density of σ = 0.031. The end-to-end distance of the DNA tether is reduced and plectonemes appear along the molecule as dynamic bright spots. At t = 2.9 s the molecule nicks due to photo damage, and it immediately extends and regains its pre-coiling end-to-end distance (compare to panel A). At the same time, the plectonemes disappear as is evident from the absence of bright spots along the molecule. (**C**) Image series showing the nicking event in detail; the time interval between the images shown is 20 ms.

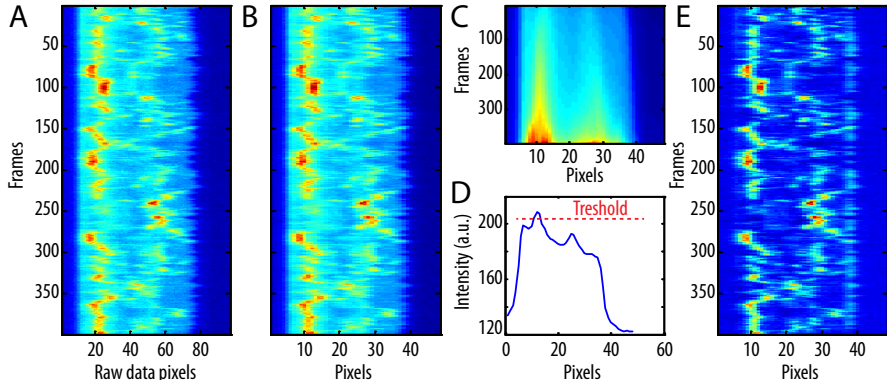
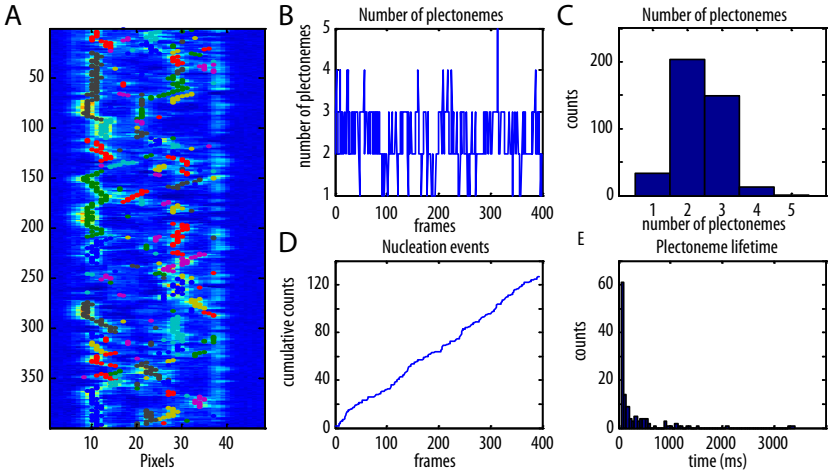


Figure S3.3: Background subtraction method for localizing plectonemes. (**A**) Kymograph of a supercoiled DNA molecule where each line is constructed by summation of the pixel intensity perpendicular to the DNA molecule. (**B**) In a first step the kymograph is down sampled by a factor of two by averaging pixel pairs along the *x*-axis, thereby increasing the signal-to-noise ratio. (**C**) For each *x*-position the intensity of the kymograph is reordered from low to high along the time-axis. This procedure separates the high intensity spots, when a plectoneme is present at a given location, from the intensity in the absence of a plectoneme at that *x*-position. (**D**) Mean intensity of the 10 % lowest values of the kymograph, i.e. the mean of the top 10 % rows of panel C. This provides the profile of the background intensity of the fluorescence along the DNA molecule. If plectonemes are very immobile, i.e. when they are present for more than 90 % of the time at a given *x*-position, peaks will still be present in this profile. To remove these peaks, a single manual threshold level was applied where appropriate, thereby limiting the maximum value of the profile. This profile will function now as a background intensity curve in the next step. (**E**) Kymograph created by subtracting the curve from panel D from the kymograph of panel B. This results in a kymograph which highlights the position of the plectonemes and which has a value close to zero for all positions where no plectonemes are present.



Supplementary Figure S3.4: Plectoneme identification by peak fitting. A, Individual identified plectonemes shown as different colored dots overlaid on the kymograph of Figure S3.3E. Plectonemes were connected between consecutive frames if their position changed less than 8 pixels from frame-to-frame, (equivalent to 520 nm), which is roughly equal to the ~500 nm observed spot size of the plectonemes. Plectonemes were identified by a peak-identification algorithm (findpeaks, Matlab Mathworks) applied to the intensity along the *x*-direction in the background-corrected kymographs (e.g. Supplementary Figure S3.3E). The peak threshold was typically set to 130 % of the fluorescence intensity level of the DNA in the absence of plectonemes and was manually optimized for each molecule. For comparison, the standard deviation of intensity fluctuations in the absence of plectonemes was normally around 5–7 % of the total intensity. The threshold level of 30 % above the background intensity therefore corresponds to a ~5σ confidence level resulting in a low number of falsely identified plectonemes (cf. Figure S3.5B). **B,** The number of plectonemes present as a function of frame number obtained by the peak-fitting procedure from the kymograph of panel A. **C,** Histogram of the number of plectonemes obtained from the data in panel B. **D,** Cumulative counts of nucleation events as a function of frame number. **E,** Histogram of the detected plectoneme lifetimes.

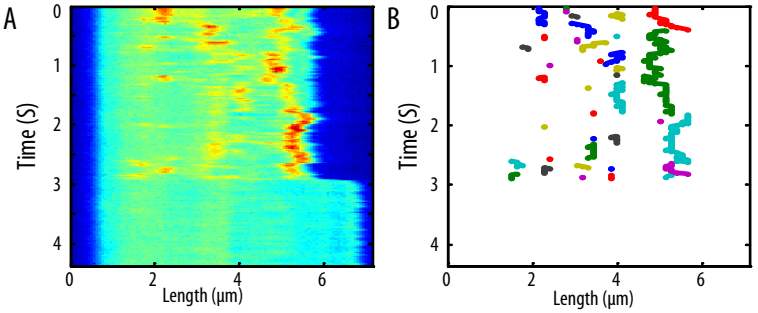


Figure S3.5: Plectonemes are only detected in supercoiled DNA molecules. (**A**) Kymograph of a supercoiled DNA molecule that nicks due to photo damage at time t = 2.9 (data of Figure S3.2B). (**B**) Individual plectonemes identified in the kymograph of panel A, shown as colored lines. Plectonemes are readily identified in the supercoiled molecule, but not in the relaxed molecule: after nicking, no plectonemes are detected anymore, confirming the accuracy of the plectoneme detection method.

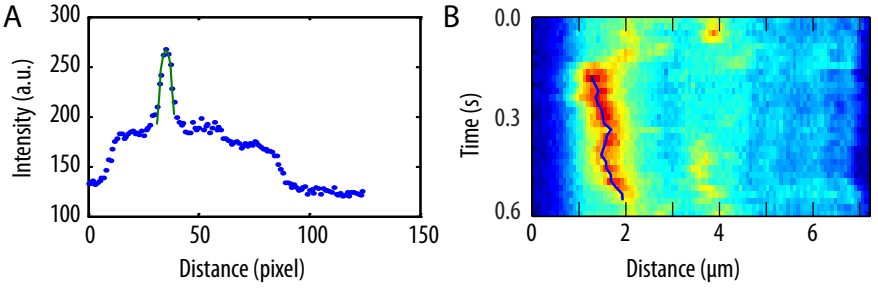


Figure S3.6: Fitting the position of a diffusing plectoneme. (**A**) Plectoneme sub-pixel positions were determined by a 9-point parabolic fit (green line) of a single line in a kymograph (blue points). We fit only plectonemes with an integrated intensity greater than 70 % of the total intensity contained in the plectonemes of the molecule, i.e. the plectoneme had to contain at least 70 % of the DNA that was in a plectonemic state. (**B**) Fitting the plectoneme position for each line in the kymograph results in a diffusional tract showing the center position of the plectoneme in time (blue line).

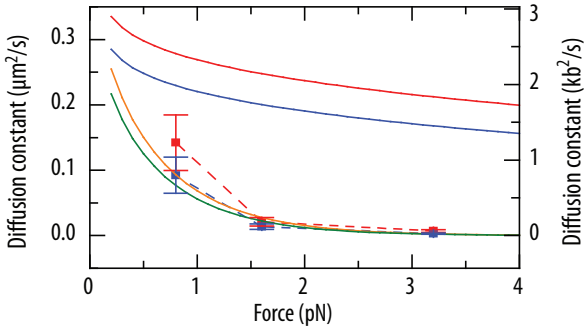


Figure S3.7: The effect of a rugged energy landscape on the diffusion of a plectoneme. The experimentally observed diffusion constant decreases rapidly with applied stretching force (red squares 150 mM; blue squares 300 mM NaCl; error bars denote the estimated error in the plectoneme size). The experimentally observed diffusion constants are considerably below those predicted for a simple hydrodynamic model (red and blue lines for 150 and 300 mM NaCl respectively, see Supplementary text S3.3). A model incorporating a rugged energy landscape for diffusion along the DNA (orange and green line for 150 and 300 mM NaCl, respectively, for description see main text) reduces the diffusion constants significantly and describes the observed data much better than the hydrodynamic model.

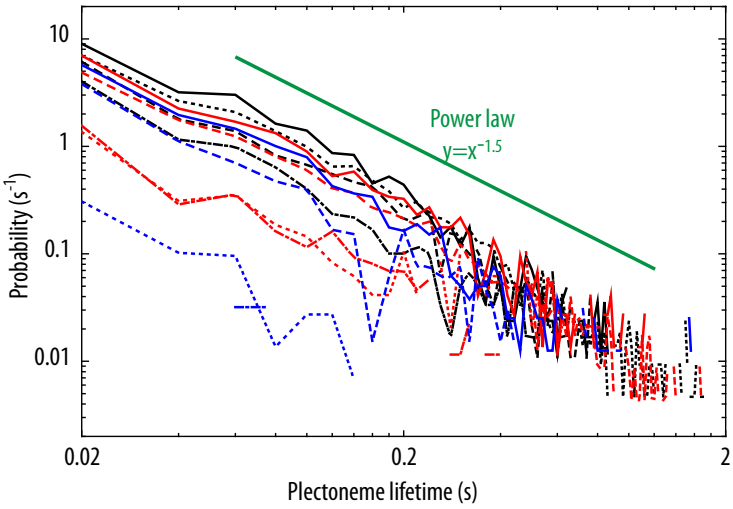


Figure S3.8: The distribution of plectoneme lifetimes shows the same scaling behavior for all experimental conditions. The probability of plectoneme lifetimes in buffers containing 20, 150 and 300 mM NaCl (black, red and blue lines respectively) under applied stretching forces of 0.4, 0.8, 1.6 and 3.2 pN (solid, dashed, dotted and dash-dotted lines respectively) shows identical scaling behavior for all conditions. The predicted power-law scaling is indicated by the green line.

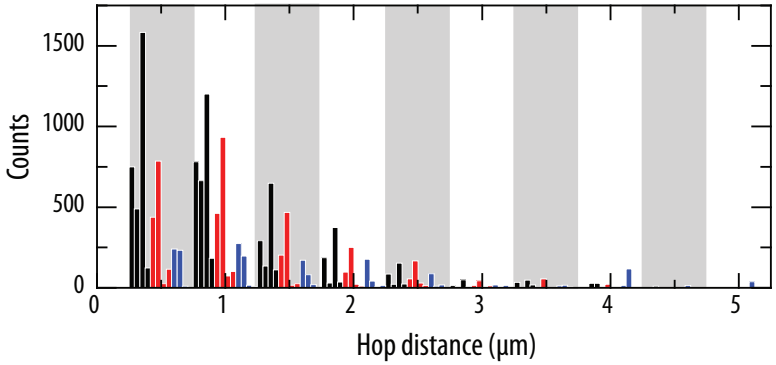


Figure S3.9: Distribution of hop distances for plectonemes. Plectonemes are able to hop over large distances along a DNA molecule, in the range of 0.5 μm to 5 μm. Hop distances are shown for salt concentrations of 0.2, 150, and 300 mM NaCl, (black, red, and blue bars), and forces 0.4, 0.8, 1.6, and 3.2 pN (from left to right for each salt concentration), gray and white areas indicate bins. Hopping occurs more frequently for shorter distances. This can be attributed to the fact that multiple plectonemes are present at any given time for most conditions and the fact that the maximum distance is limited by the end-to-end distance, (± 5 μm), of the supercoiled DNA molecule.

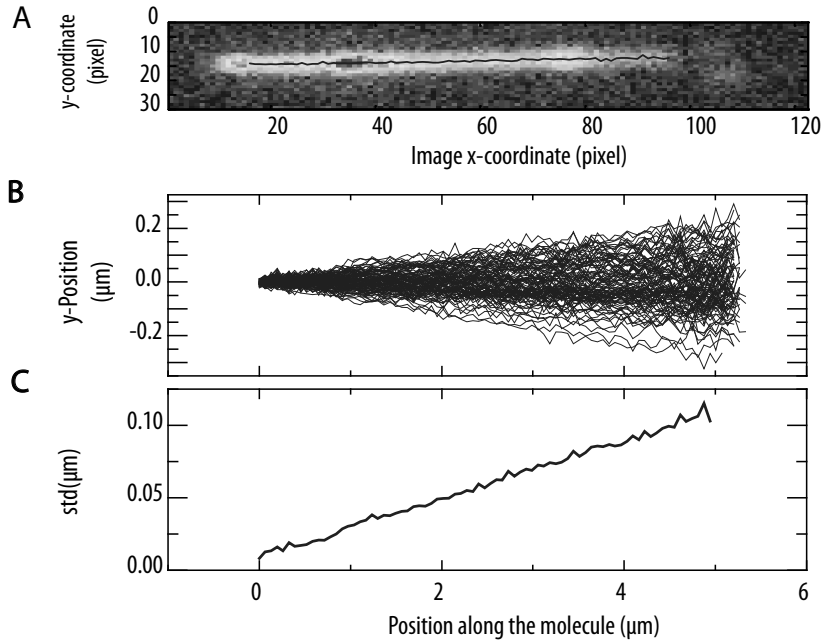


Figure S3.10: Tethered DNA molecules move freely and show no interactions with the surface of the flow cell. (**A**) Image of a supercoiled DNA molecule. The center position of the DNA molecule along its length (black line) was determined by a Gaussian fit of the intensities in the *y*-direction. No spatial digression were observed in the fitted positions, the plectonemes merely appear as near-diffraction-limited spots of higher intensity. Experimental conditions for the molecule shown are 300 mM NaCl at 3.2 pN. (**B**) Fitted DNA positions for a time series of images. Each line represents the fit to an single image such as in panel A. The DNA molecule freely swivels around its attachment point on the surface (see also Movies S1–S2), here located at position (12,14) in panel A. The DNA is stretched to a virtually straight line and has a 5.3 µm end-to-end distance between its attachment point and the fluctuating bead. (**C**) The standard deviation (std) of the detected DNA positions in the *y*-direction increases linearly along the molecule, confirming the absence of surface interactions at specific locations. If present, surface interactions would impede the movement of molecule and this would lead to regions with a reduced standard deviation.

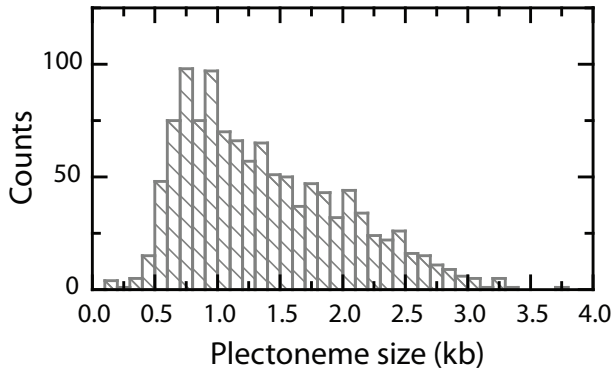


Figure S3.11: Size distribution of detected plectonemes. Plectoneme sizes were determined from the ratio of the intensity for identified plectonemes to the total intensity of the fluorescently labeled DNA molecule. The plectoneme intensity was calculated as the integrated intensity of 9 pixels centered on the detected plectoneme position. The histogram shows that the minimum detectable plectoneme size for the 130 % threshold level used in the peak-fitting procedure is approximately 0.5 kb.

Applied number of turns for ionic strengths and applied stretching forces

	0.4 pN	0.8 pN	1.6 pN	3.2 pN
20 mM	43	62	87	134
150 mM	42	61	86	124
300 mM	43	64	87	129

Table S3.1: Measurements were performed under conditions where 25 % of the molecule was in a positively supercoiled plectonemic state. The table shows the number of applied turns for the respective applied stretching force and ionic strength at which experiments were performed.

Movie S3.1: Plectoneme dynamics within a fluorescently labeled 21 kb DNA molecule. The movie corresponds to the molecule shown in Figure 3.1C and D in the main text at 0.8 pN applied stretching force and 150 mM NaCl. Plectonemes are clearly visible as bright spots which appear and disappear and move along the DNA molecule. Experimental images were recorded at 50 Hz, the movie is slowed down to 0.5X real-time i.e. 25 Hz, scale bar in the lower left hand corner corresponds to 1 µm.

Movie S3.2: Plectoneme dynamics for all probed experimental conditions. The movie corresponds to the kymographs shown in Figure 3.2A in the main text. Rows correspond to ionic strength, from top to bottom 300, 150 and 20 mM NaCl; columns correspond to

force, from left to right 0.4, 0.8, 1.6 and 3.2 pN. Plectonemes are clearly visible as bright spots which appear and disappear and move along the DNA molecules. The number and dynamics of the plectonemes change with applied stretching force and salt concentration. Experimental images were recorded at 50 Hz, the movie is slowed down to 0.5X real-time i.e. 25 Hz, Scale bar in the lower left hand corner corresponds to 1 μm.

Supplementary references

26 Slattum, P. S. *et al.* Efficient in vitro and in vivo expression of covalently modified plasmid DNA. *Mol. Ther.* **8**, 255-263, (2003).

27 Revyakin, A., Ebright, R. H. & Strick, T. R. Single-molecule DNA nanomanipulation: Improved resolution through use of shorter DNA fragments. *Nature Methods* **2**, 127-138, (2005).

28 Qian, H., Sheetz, M. P. & Elson, E. L. Single-Particle Tracking - Analysis of Diffusion and Flow in 2-Dimensional Systems. *Biophys. J.* **60**, 910-921, (1991).

29 Moroz, J. D. & Nelson, P. Entropic elasticity of twist-storing polymers. *Macromolecules* **31**, 6333-6347, (1998).

30 De Vlaminck, I. & Dekker, C. Recent advances in magnetic tweezers. *Annual review of biophysics* **41**, 453-472, (2012).

31 Gittes, F. & Schmidt, C. F. Thermal noise limitations on micromechanical experiments. *Eur Biophys J Biophy* **27**, 75-81, (1998).

32 Broersma, S. Viscous Force Constant for a Closed Cylinder. *J. Chem. Phys.* **32**, 1632-1635, (1960).

33 Dirk, S. The charged colloidal cylinder with a gouy double layer. *Journal of Colloid and Interface Science* **53**, 296-306, (1975).

34 Marko, J. F. & Siggia, E. D. Stretching DNA. *Macromolecules* **28**, 8759-8770, (1995).

35 Takaisi, Y. The forces on a circular cylinder moving with low speeds in a semi-infinite viscous liquid bounded by a plane wall. *J. Phys. Soc. Jpn.* **10**, 407-415, (1955).

36 Grinstead, C. M., Snell, J.Laurie. *Introduction to Probability*. 2nd edn, (American Mathematical Society 1997).

37 Sornette, D. *Critical Phenomena in Natural Sciences.* 2nd edn, (Springer, 2006).

38 Clauset, A., Shalizi, C. R. & Newman, M. E. J. Power-Law Distributions in Empirical Data. *Siam Rev* **51**, 661-703, (2009).

39 Stumpf, M. P. H. & Porter, M. A. Critical Truths About Power Laws. *Science* **335**, 665-666, (2012).

40 Nelson, P. Transport of torsional stress in DNA. *Proc. Natl. Acad. Sci. USA* **96**, 14342-14347, (1999).

41 Levinthal, C. & Crane, H. R. On the Unwinding of DNA. *Proc. Natl. Acad. Sci. USA* **42**, 436-438, (1956).

Knots

Part 2—Knots

We all use knots to hold and bundle things together or attach a rope to an object. A good knot is one that is strong but which also may be easily untied. A long rope may also become knotted due to random entanglement, but this is more of a nuisance than a proper knot. Biology seems to agree on this, since randomly entangled knots are extremely rare in proteins and DNA. Proteins, like DNA molecules, are made up of a long chain of subunits, amino acids, which fold up to produce a blob-like protein. Remarkably, knots occur orders of magnitude less frequent than expected for simple random folding of such a chain. DNA poses a different problem: since DNA molecules are very long, it is cumbersome to remove a knot by threading one end back through the knot to remove it. There is however a different solution: one can cut a rope at the place of the knot, remove the knot, and splice the rope back together. Indeed this is what cells do, and special proteins, called topoisomerases, have evolved that cut the DNA, pass it behind a crossing DNA strand, and then reattach the severed ends. Remarkably, these proteins are thus able to remove nearly all knots from DNA by only sensing the local crossing angles of the knotted DNA. So, are there no knots in cells? Knots are very useful to hold things together if they can be tied and untied as required. A class of proteins called Structural Maintenance of Chromosomes (SMC) exist which performs this task and hold DNA together in a structured manner. SMC proteins may form loops linking DNA or chromosomes together and have hooks at their ends which they use to engage each other. In this part of the thesis, we report on two different kind of 'knots'. First we describe a method to sense the position of DNA-bound proteins using an artificially created loop of DNA, made by looping one DNA molecule around another using optical and magnetic tweezers. Second, we report initial results aimed at measuring the strength of the SMC protein complex cohesin that provides the essential link between sister chromatids during cell division by enclosing them in a protein loop.

CHAPTER 4

Localization of DNA-bound proteins by scanning a DNA molecule using hybrid magnetic and optical tweezers

The functional state of the genome is determined by its interactions with proteins that bind, modify, and move along the DNA. We developed a combined magnetic and optical tweezes apparatus that allows for the sensitive label-free detection of the positions and binding strength of proteins localized on DNA. A DNA loop, which acts as a scanning probe, is created by looping an optically trapped DNA tether around the magnetic bead tether. We are able to determine the relative distances between DNA-bound EcoRI proteins with 14 ± 8 nm resolution along a λ DNA molecule. We find a consistent offset between back and forwards scans of 35 ± 15 nm for the detected protein positions, which corresponds to the size of the DNA loop and is in agreement with theoretical estimates. At higher applied stretching forces the scanning loop is able to remove bound proteins from the DNA. The use of magnetic tweezers in this assay not only allows the facile preparation of many single-molecule tethers which can be scanned one after the other, but it also allows for direct control of the supercoiling state of the DNA molecule making it uniquely suitable to address the effects of torque on protein-DNA interactions.

Marijn T.J. van Loenhout, Iwijn De Vlaminck, Benedetta Flebus, Johan den Blanken, Ludovit Zweiffel, Koen Hooning, Jacob Kerssemakers, Susanne Hage, and Cees Dekker. *To be submitted.*

Introduction

DNA is the center of action in cells: proteins bind to specific sequences, RNA-polymerases move along and transcribe genes, DNA is modified and wrapped around nucleosomes. Together the actions and locations of these proteins determine how genetic information is used in a cell[1]. There is thus an evident need for techniques which are able to localize DNA-bound proteins and probe their interactions. Single-molecule experiments, which allow for precise control and detection of individual DNA molecules and proteins, have made it possible to determine many of the intrinsic properties of DNA and associated proteins[2,3]. Electron microscopy and atomic force microscopy (AFM) allow for the direct visualization of DNA bound but require the immobilization of DNA and proteins on a surface[4-6]. Optical tweezers have been used to monitor the movement of single proteins along DNA in buffer, but rely on labeling of the proteins for optical or mechanical detection[7]. Inspired by the work of Noom *et al.*[8] we developed a method that allows label-free high-accuracy detection of proteins bound to DNA by the use of a scanning loop formed by one DNA molecule that is looped around another[9]. A simplified scheme depicting the loop formed between the two DNA molecules that are held in optical and magnetic tweezers is shown in Figure 4.1A. When the loop is scanned, by moving the horizontal DNA molecule with optical tweezers, a DNA-bound protein will act as a friction barrier. Upon encountering the bound protein, the sliding loop will be halted the protein, and the magnetic bead will be displaced, thereby indicating the position of the protein (Figure 4.1D).

The combined use of magnetic and optical tweezers to create and manipulate the DNA loop has several advantages over using only optical tweezers: it allows for the facile measurement of multiple molecules, it has a much higher force resolution, and it creates new possibilities to probe the influence of DNA supercoiling and its influence on DNA-protein interactions[10]. To demonstrate the functionality of this scanning technique, we used it here to determine the position of EcoRI proteins bound to DNA. At low applied stretching forces we observe that the loop gets stuck for a short period of time upon encountering a protein and then passes over, whereas, at higher stretching forces, the force exerted by the scanning loop on the protein can remove it from the DNA.

Results

We designed and built a combined magnetic and dual-beam optical tweezers instrument that allows manipulation and detection of interaction forces of an optically trapped DNA molecule looped around a second DNA molecule that is tethered in a magnetic tweezers configuration (Figure 4.2)[9]. Magnetic tweezers were used to create a vertical single-molecule DNA tether between the flow cell surface and a magnetic bead by exerting a pulling force via a magnet positioned above the flow cell[11]. Two optical traps were generated by splitting a beam into two orthogonally polarized

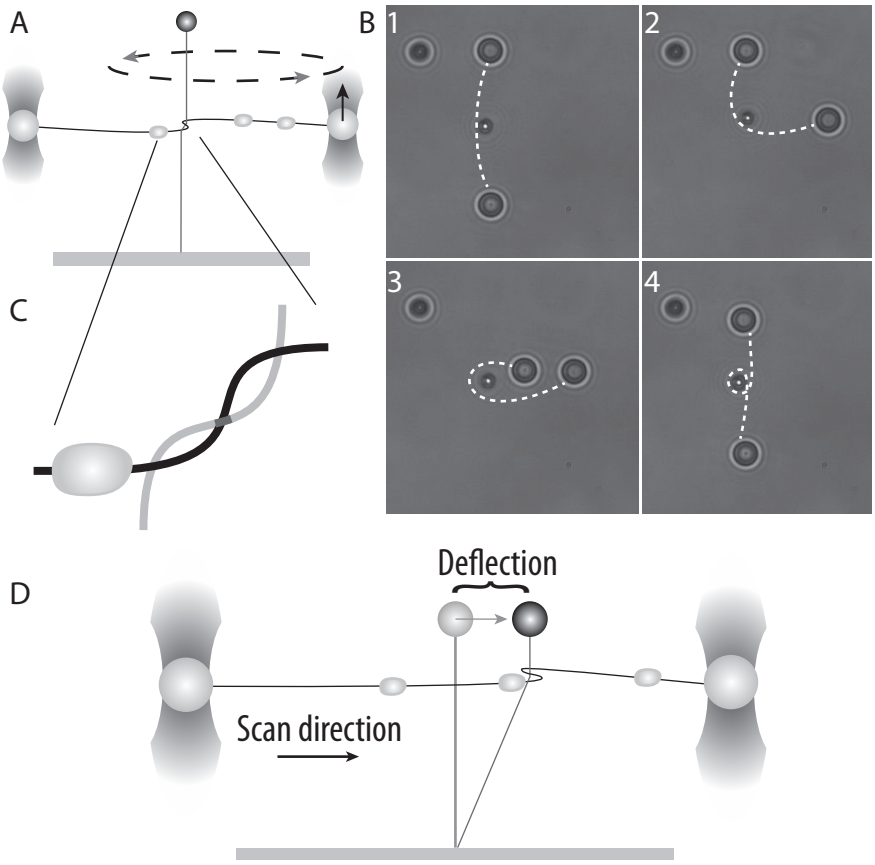


Figure 4.1: Detection of DNA-bound proteins using a scanning DNA loop and magnetic and optical tweezers. (**A**) A DNA loop is created by moving the optically trapped beads around the magnetic-bead DNA tether. (**B**) Image series showing a top view of the formation of the DNA loop. The loop is made by rotating the beads trapped by optical tweezers around the magnetic bead which is located in the center of the image. The position of the DNA molecule is indicated by the white dashed line. The bead in the upper left corner functions as a reference and is stuck to the surface of the flow cell. (**C**) Zoom in on the intertwined geometry of the DNA loop showing a bound protein. (**D**) The loop is scanned in the horizontal direction by moving both optically trapped beads in concert. Upon encountering a bound protein the DNA cannot slide through the loop anymore and the magnetic bead tether is deflected.

beams, which could be independently steered in *x*- and *y*-direction by acousto-optic deflectors (AODs). Video microscopy was used to simultaneously detect the position of both the magnetic and optically trapped beads with nanometer resolution in three dimensions. The position of the flow cell surface was monitored by the use of

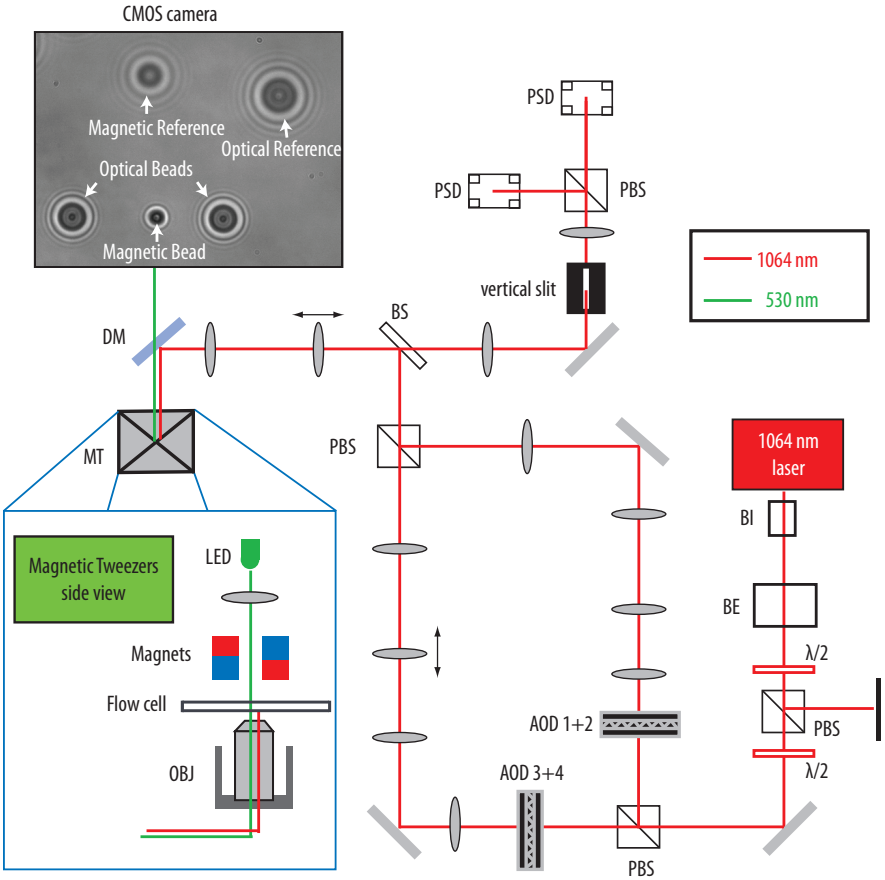


Figure 4.2. Schematic illustration of the hybrid magnetic and optical tweezers setup. A 4 W 1064nm laser beam passes first trough a beam isolator (BI) and is expanded by a beam expanded (BE). Two half-wave plates (λ/2) and a polarizing beamsplitter (PBS) are used to control the power in both polarizations of the beam. The beam is subsequently split and sent through a pair of acousto-optic-deflectors (AODs) to steer the beam. Lenses are used to further expand the beam and transfer the conjugate plane of the AODs to the back focal plane of the objective. Two lenses are mounted on translation stages to allow adjustment of the position of the optical traps in the axial direction. 10% of the back reflected light is collected with a plate beamsplitter (BS) and directed onto a separate position sensitive detector (PSD) for each trap. A vertical slit is used to enhance the signal to noise ratio of the detection by blocking reflections from the surfaces of the flow cell.

a reference bead that was stuck to the surface. Force calibration of the optical trap stiffness was performed using backscattered light detection (Figure 4.3)[12]. A laminar flow cell with multiple parallel flows was used to step-by-step assemble the DNA tethers and create separated buffer environments[9,13,14]. The flow cell was mounted on a

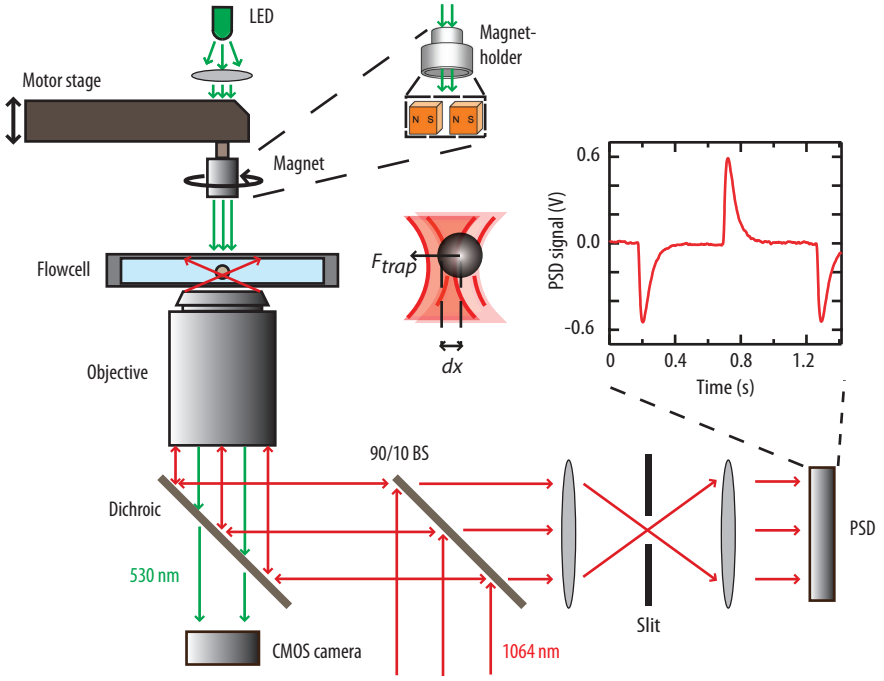


Figure 4.3: Magnetic tweezers and detection configuration. A magnet holder containing two neodymium magnets is mounted on a motorized stage to enable rotation and up-and-down movement of the magnets. A LED is used to image both the magnetic bead and optically trapped beads on a CMOS camera at 50 Hz or 100 Hz. The displacement of the optical beads from the laser focus is detected in back reflection. A 90/10 plate beam splitter is used to split off the back-reflected light and a slit is used as a spatial filter to block reflections from the flow cell surfaces. Trap stiffness was calibrated by using a square-wave method where the optical traps are quickly displaced and the return of the bead to the equilibrium position is monitored (see PSD signal top right).

translation stage to enable controlled movement of the optically trapped beads into the different laminar flow channels.

Each experiment starts by introducing magnetic beads into the flow cell and tethering each via a single DNA molecule in a side channel of the flow cell. This procedure creates many (>100) DNA tethers and is similar to the procedure commonly used in magnetic tweezers experiments[11]. To create a single-molecule DNA tether between the optically trapped beads a step-by-step assembly procedure was employed, using laminar flow channels (Figure 4.4). First, two streptavidin-functionalized polystyrene beads are caught in two optical traps (Figure 4.4C). Next, these beads are moved to a channel containing biotin end-labeled λ DNA molecules of 16 μm (48 kb) length. The formation of a DNA tether between the beads is monitored by moving one bead

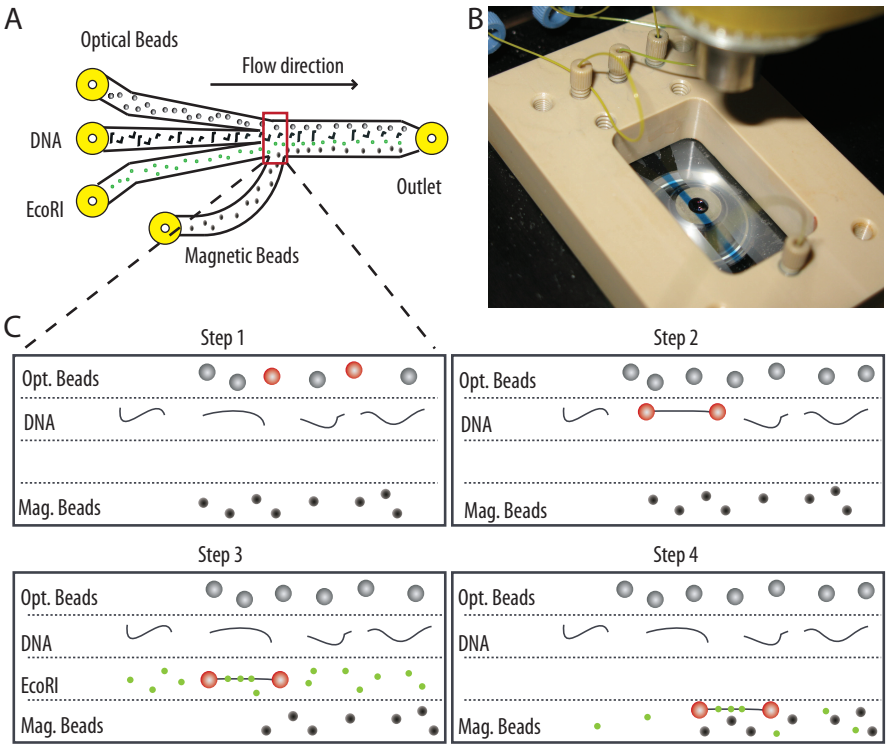


Figure 4.4: Step by step assembly of DNA tethers using a laminar flow system. (**A**) Schematic of the laminar flow cell showing four different inlets, that are combined into a central channel. Magnetic beads are flushed in via a side channel that exits perpendicular to the main flow direction. This allows for a very low flow rate during experiments, while preventing diffusion into this channel. (**B**) Photograph of a flow cell using three channels, the central channel contains blue dye, showing clear separation by laminar flow from the other two flow lanes. (**C**) Illustration of the step-by-step assembly procedure to perform the scanning loop experiments. In step 1 two beads are caught by optical tweezers. In step 2 a biotin-labeled DNA molecule is caught between the streptavidin-coated beads. In step 3 the presence of a single DNA molecule is confirmed by force extension analysis and EcoRI proteins are allowed to bind to the tethered DNA molecule. In step 4 the DNA tether is brought within close distance of a magnetic bead tether and a loop is made. A low concentration of EcoRI proteins is present in this channel to assure that proteins remain bound to the DNA.

to and from the other bead until a force on the non-moving bead is detected. The two-bead DNA tether is moved to a channel without DNA and force-extension analysis is used to ensure the presence of a single DNA molecule. To perform the interaction experiments, the optically trapped molecule is moved within close proximity of the

magnetic bead tether and a loop is made by moving one of the optically trapped beads around the magnetic DNA tether (Figure 4.1B).

The DNA loop acts as a scanning probe, enabling the label-free detection of DNA-bound proteins. After making a loop around the 4 μm (12 kb) long magnetic bead tether, as illustrated in Figure 4.1A–B, we position the optically trapped beads in line with the magnetic bead, but at a lower *z*-position, thus creating a crossed DNA configuration at a height approximately halfway between the magnetic bead and flow cell surface, i.e. 2 μm below the bead. The precise geometry of the DNA at the loop (cf. Figure 4.1C) is set by the bending stiffness of the DNA and the tensions applied on the magnetic and optically trapped DNA molecules. Unless otherwise stated the tension in both molecules in the experiments reported below was set in the range of 12–16pN to create a symmetric structure.

The displacement of the magnetic bead acts as a sensitive probe to detect the interactions between the two DNA molecules. Any friction present in the sliding loop while scanning will displace the magnetic bead tether sideways, and consequently the magnetic bead (Figure 4.1D). In the absence of DNA-binding proteins we observed no interactions while scanning the λ DNA molecule (data not shown), indicating the absence of friction between DNA molecules as also observed previously[8,9]. The absence of friction is likely due to electrostatic repulsion between the negatively charged DNA backbones, which prevent direct mechanical contact[15].

To demonstrate the ability of the setup to detect proteins along a DNA molecule, we performed measurements in the presence of the restriction enzyme EcoRI under noncleaving conditions. The EcoRI restriction enzyme binds specifically to its recognition sequence GCTT but does not cleave the DNA in the presence of Ca^{2+} ions[16]. The λ DNA molecule contains five recognition sites along its length which can function as site-specific markers to validate the detection method. Figure 4.5 shows an example of two consecutive forward and backward scans of a λ DNA molecule in the presence of 50 nM EcoRI, made by horizontally moving the optically trapped beads at a scan rate of 1 μm/s. We clearly identify three spikes in both the forward and backward scans. The other two EcoRI positions were located outside of the scan range and therefore could not be detected. The spikes result from increased friction and are present in both the forward and the backward scans. We also occasionally observed spikes at other locations along the DNA which most likely correspond to nonspecifically bound proteins along the DNA. These spikes however often disappeared after a single scan (data not shown), indicating that the proteins had been dislodged from the DNA. Figure 4.6A shows superimposed experimental data from 7 consecutive forward and backward scans. This plot clearly shows that the spikes originate from three locations on the λ DNA molecule and that these positions match with the positions expected from the DNA sequence. To determine the positions of the bound proteins we calculated the intersection of a linear fits to individual spikes and the baseline (Figure 4.6A, blue dashed line, and green line respectively). The above method resulted in an accuracy of the position differences between bound EcoRI proteins as compared to

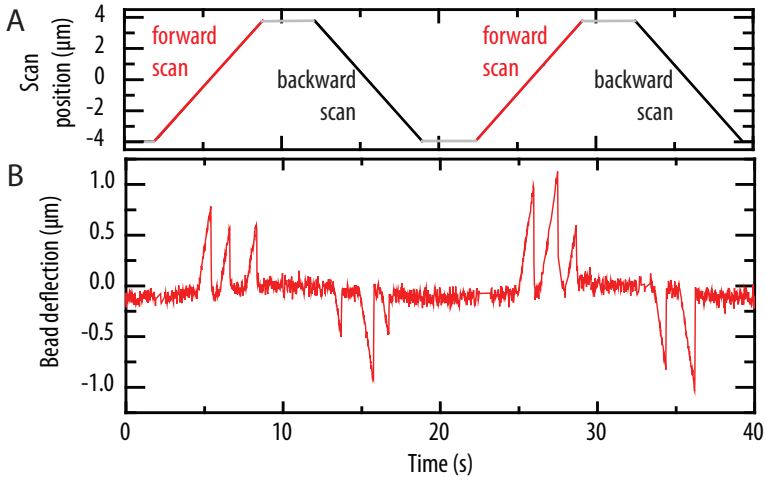


Figure 4.5: DNA-bound proteins are detected by scanning a DNA loop. (**A**) Scan position calculated as the mean position of the two optically trapped beads (red line forward scans, and black line backward scans). (**B**) Magnetic bead deflection showing three spikes in forward and backward scans indicating the presence of DNA-bound proteins.

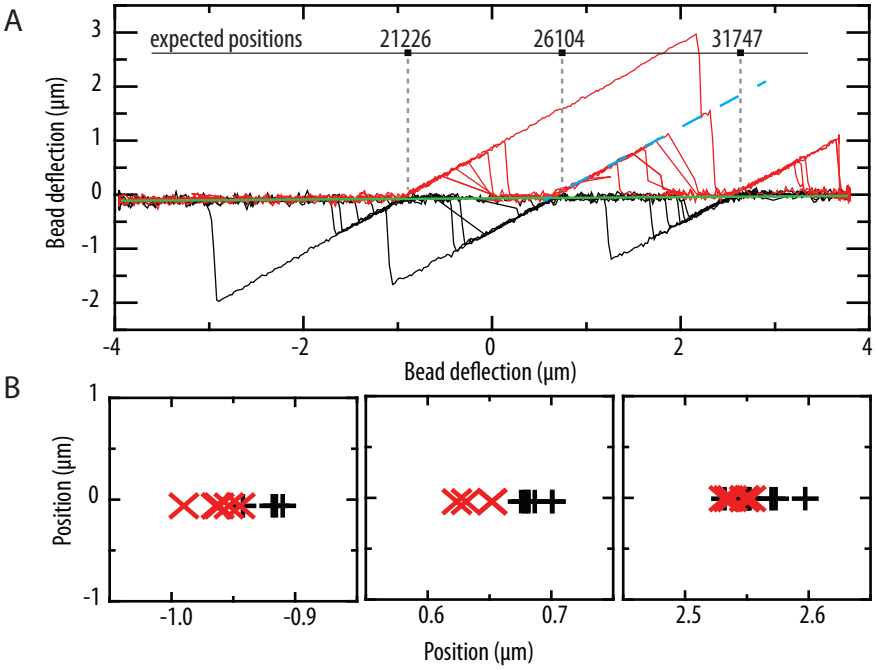


Figure 4.6: Detected spikes correspond to DNA-bound EcoRI proteins. (**A**) Seven consecutive scans superimposed (red lines forward scans, black lines backward scans). Pro-

tein positions (grey dashed lines) were determined from the intersect of linear fits to the baseline (green line) and individual identified spikes (blue dashed line). Expected position passed on the DNA sequence are indicated at the top. (**B**) Protein position determined from forward (red crosses) and backward (black crosses) show that EcoRI positions are systematically detected to the left in forward scans compared to the position detected in the backward scans.

the positions predicted by DNA sequence of 14 ± 8 nm, (the error denotes standard error of the mean, n=60 spikes obtained from 4 different λ DNA molecules were fit). The error in the absolute positions of the EcoRI proteins along the λ DNA molecule was approximately 60 nm, which is expected to be due to differences in size of the optically trapped beads.

Surprisingly we found that there was a consistent offset of the positions detected in forward and backward scans. Figure 4.6B shows a separate analysis of forward (red points) and backward (black points) for the data shown in Figure 4.6A. There is a consistent offset of 35 ± 15 nm between the detected positions in forward and backward scans. This may be understood by considering the geometry of the scanning loop (Figure 4.1C). The probe DNA will start to displace the magnetic bead tether as soon as a protein encounters the loop, i.e. the detected position will not be in the middle of the DNA loop but at the point where the protein first encounters the loop. For scans in the reverse direction the detected interaction point will be at the other side of the scanning loop. The offset between forward and reverse scans is therefore a measure of the size of the probing loop. We can estimate the size of the loop by a simple model accounting for the mechanical and electrostatic properties of the DNA. Assuming a circular shape of the loop, its diameter will be: $D \approx \sqrt{2 l_p k_B T / f}$, where l_p is the persistence length of DNA of around 50 nm and, f, the applied tension[3]. This gives a loop diameter of 5–7 nm for tensions in the range of 10–20 pN. The total length of DNA in the loop will then be 16–22 nm and accounting for the size of the EcoRI protein of ≈ 3 nm we expect an offset of 19–25 nm, which is close to the determined value of 35 ± 15 nm. The slightly larger measured loop size may be accounted for by the fact that the experimental DNA geometry likely differs substantially form an ideal circular loop (Figure 4.1C).

To understand the forces acting on the DNA loop leading to the displacement of the magnetic bead, we performed numerical calculations of the loop and the magnetic bead positions for different scan positions. The loop was considered free sliding along the DNA, but the protein was not allowed to pass through the loop. At each position all forces acting on the loop, i.e. those of the two DNA molecules and their geometry were solved iteratively and the stretching of the DNA molecules was accounted for by the WLC model[17]. The local geometry of the forces acting at the loop pushes it upward as the magnetic bead is displaced (Figure 4.7A–B). The near linear displacement of the magnetic bead observed in the experiments was accurately reproduced by the numerical calculations as shown by the overlap of the experimental

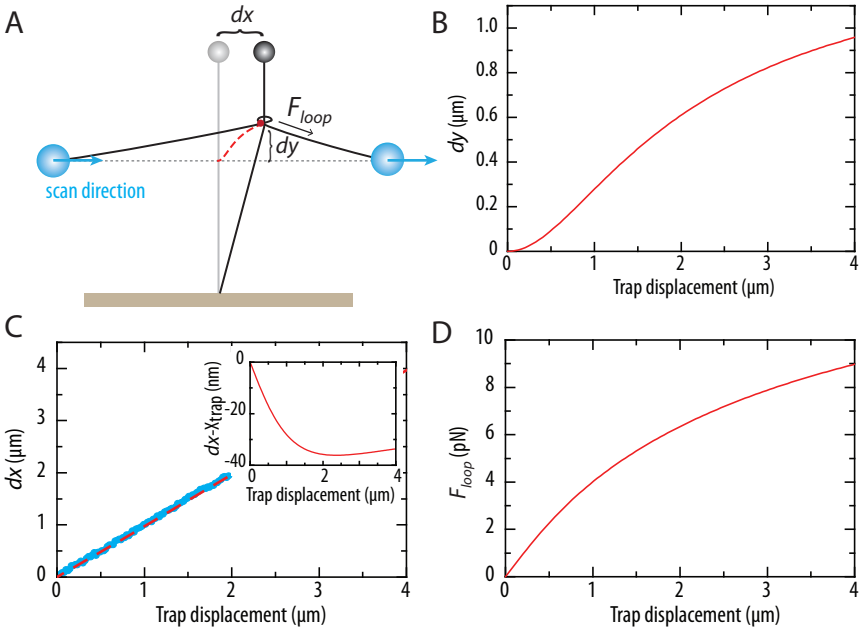


Figure 4.7: Calculated position and forces acting on the scanning DNA loop. (**A**) Schematic diagram, not to scale, showing the displacement of the DNA loop as the optically trapped beads are scanned to the right. Forces and positions were calculated iteratively for a 48 kb DNA molecule, at an initial loop height of 2 µm above the surface, initial tensions were 10 pN for both DNA molecules, and optical trap stiffness was set to 100 pN/µm. The elasticity of the DNA was modeled by the worm-like-chain model. As the loop encounters a bound protein or other obstacle during the scan, the DNA will no longer slide through the loop but displace the magnetic bead by *dx*. Due to the direction of forces acting on the loop it will move to the right and upward following the track depicted by the red dotted line. (**B**) Calculated upward movement, *dy*, of the bead as a function of trap displacement after encountering a protein. (**C**) Movement of the magnetic bead in the scan direction after the loop encounters a bound protein (blue points experimental data, red dashed line calculation). The displacement of the magnetic bead is almost equal to that of the scanning traps, inset shows the difference between the magnetic bead position, *dx*, and trap position, *Xtrap*, note the nm scale. (**D**) Calculated force acting on a protein blocking the DNA loop.

data and calculation results (Figure 4.7C, blue points, and red dashed line respectively). The force acting on the loop increase sublinear for very large bead displacements (Fig 4.7D). For small displacements the force scales approximately linear with a magnetic-bead displacement of ~4 pN / µm, making the magnetic bead a sensitive probe for the exerted force. The force detection limit of the experimental scanning method therefore is ≈ 0.1 pN/√Hz, taking a 4σ confidence bound and considering the standard deviation of the thermal fluctuations of the magnetic bead of σ = 0.05 µm.

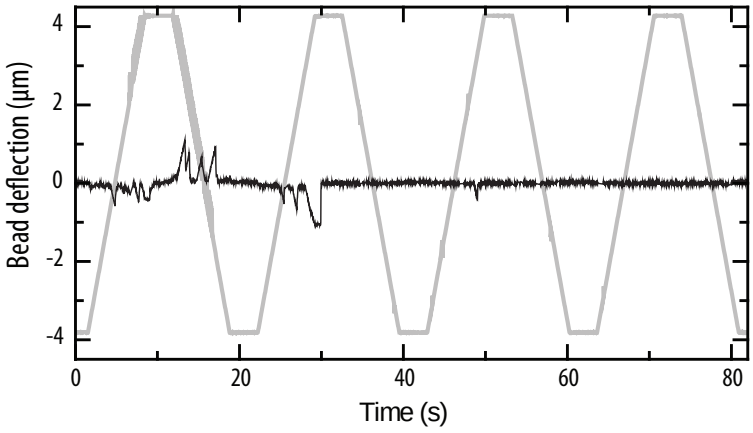


Figure 4.8: Three EcoRI proteins are detected and removed by the scanning DNA loop. The magnetic bead deflection (red line) shows that the proteins are removed after the third scan at 20 pN applied stretching force for both DNA molecules. At $t = 50$ s a rebinding event is observed, but this protein is also immediately dislodged. Scans are indicated by the gray line representing the center position of the scanning DNA, calculated as the mean of the two optically trapped-bead positions.

To determine if the force that is applied by the loop on the bound EcoRI proteins could pull them off the DNA, we performed experiments at different forces. At moderate stretching forces of ~12 pN for both magnetic and optical tethers, proteins generally remained bound, and were observed for many consecutive scans (see Figure 4.6A). At forces above 20 pN applied on both the magnetic and optical tweezers tethers proteins were, however, easily dislodged and disappeared after a single or several scans. Figure 4.8 shows such a trace where 3 proteins are detected in the first few scans but they are removed shortly after and a frictionless signal is recorded as is expected for DNA-DNA interactions.

Discussion

We have developed a new method which combines magnetic and optical tweezers to localize DNA-bound proteins. The use of magnetic tweezers in this assay presents several advantages compared to the elegant method described by Dame *et al.*[18] and Noom *et al.*[8] using on 4 optical traps to create a DNA loop. The use of 4 optical traps in their method is however not easily adopted to tether two DNA molecules of different sequence or length, limiting its ability to address DNA sequence related biological questions. It is also an experimentally complex method, which is likely responsible for its limited adoption in more experiments. The combination of magnetic tweezers in addition to optical tweezers, described here, creates a robust platform to prepare

a large number >100 of tethered DNA molecules in advance. This allows one to scan multiple molecules sequentially without having to capture new beads with optical tweezers. The combination of magnetic and optical also facilitates the use of different DNA molecules as the two DNA tethers are assembled independently. Importantly, magnetic tweezers allow to apply torque on the tethered DNA molecule and our method can thus be extended to induce supercoils or modulate and probe the binding affinity of proteins in response to torque in the magnetic-bead DNA tether. Magnetic tweezers also are capable of detecting very low forces (10 fN) as their detection limit is principally set by thermal noise.

Several other techniques have been developed to localize DNA-bound proteins. EM allows for direct visualization of proteins and DNA but requires the samples to be frozen or deposited on a surface[19]. The best established scanning probe technique is atomic force microscopy (AFM), which uses a scanning tip to image or manipulate DNA and proteins immobilized on a surface[5,6]. In contrast to AFM, the method described here holds the DNA molecules free in buffer solution and away from the (potentially highly charged) surface. Importantly the forces applied by the scanning DNA molecule act along the DNA contour closely resembling *in vivo* processes, where DNA molecules are expected to slide across each other and motor proteins move along and remodel DNA-bound proteins. More recently solid-state nanopores have also been used to identify local structures on DNA molecules[20]. These rely on the detection of a change in ionic current as a DNA molecule passes through a nanometer-sized hole and partially blocks it. The method is in principle very simple but the high speed of translocation makes it difficult to detect individual bound proteins, and the force exerted on the bound proteins is ill defined[21].

The experiments described here used a horizontal scanning configuration, where the optically trapped DNA molecule was moved from side-to-side. This is however not the only mode of operation as the optical trapped tether can also be displaced vertically to scan along the magnetically trapped DNA molecule. Unlike the horizontal scanning procedure this method is asymmetric since on side of the DNA molecule is attached to surface and the other to the magnetic bead. If the loop is scanned in the upward direction the applied force upon encountering a bound protein will increase gradually as due to the stretching of the DNA between the flow cell surface and the scanning loop. If the loop is scanned towards the surface and encounters a DNA-bound protein it will pull the magnetic bead down. The force applied to the loop will in this case be nearly constant and equal to the force applied on the magnetic bead. This scanning direction therefore provides a simple method to apply constant force on DNA-bound proteins and monitor their stability. By rotating the magnets it is possible to introduce torsion in the DNA molecule tethered to the magnetic bead. Scanning in the vertical direction can thus be used to directly probe the interactions of proteins bound to supercoiled DNA.

Many biological processes rely on proteins diffusing or actively moving along DNA. These moving proteins may encounter DNA-bound proteins that can stop or slow their progress. For example, crowded regions occur at highly transcribed genes, where tens of RNA-polymerases move along the same tract and can get stuck in 'traffic jams'[22]. These moving proteins will have different speeds and they will therefore collide and interact[23,24]. The label-free properties of the proposed method and the ability to use the DNA loop as a controlled roadblock will allow the study of these poorly understood interactions. We envision that the developed method using a scanning-DNA loop and the combination of magnetic and optical tweezers will not only prove useful for the study static bound proteins, but can also be used to count or track moving proteins.

Materials and methods

Hybrid magnetic and optical tweezers

A combined dual-beam optical tweezers and magnetic tweezers instrument was designed and built to perform the scanning experiments. Figure 4.2 shows a detailed outline of the setup. The optical traps are generated with a 1,064 nm laser (diode-pumped, solid-state Nd:YAG, Coherent Compass 1064-4000M), isolated against back-reflections by a Faraday isolator and expanded by a beam expander. A combination of a half-wave plate and a polarizing beam splitter allows tuning the total laser power. A second half-wave plate is used to control the polarization of the beam. Two independent traps are generated by a polarizing-beam splitter cube. Both traps can be independently steered using two-axis acousto-optic deflectors (AODs, DTSXY-250, AA Opto-Electronic). The RF-frequency electrical drive signals for the AODs are generated using a home-built four channel Direct Digital Synthesizer (DDS, AD9959, Analog Devices) and subsequently amplified to 1 W (ZHL-5W-1, mini-circuits)[25]. A field programmable gate array (FPGA, NI PXI-7853R) equipped with analog-to-digital converters is used to interface with the DDS allowing fast beam steering and feedback[26]. In both beam paths, we implemented two telescope systems, a 1:1 followed by a 2:3 telescope, before recombining both beams with a polarizing beam splitter cube. The first lens in the 2:3 telescope system could be displaced axially in one of the paths using a computer-controlled translation stage (8MT167-25, Standa) to move one of the traps in axial direction. A final telescope equipped with a computer-controlled translation stage (8MT167-25, Standa) allows the axial position of both traps to be controlled and expands the beam to the appropriate size before coupling the beam into a water-immersion objective (Nikon CFI PLAN APO VC 60X WI). The position of the external magnet on top of the flow cell makes transmission-based detection of the laser beam impossible. Instead, detection of the bead displacement of the optical traps is performed by detecting the back-reflected signal[12]. Backscattered light is collected by the objective and separated using a 90-10 plate beam splitter and directed onto a position-sensitive detector (PSD, DL100-7PCBA3, Pacific Silicon Sensor).

The high-bandwidth detector allows performing trap-stiffness calibrations through a power-spectrum analysis. The magnetic tweezers consist of an external magnet placed above the flow cell. External motors (M-126.PD2 and M-150.PD, Physik Instrumente) allow positioning and rotating the external magnets and thereby to stretch and twist the magnetic-bead DNA tethers in the flow cell. Positional tracking of two optically trapped polystyrene beads (streptavidin coated, 2.1 μm, Spherotech), the superparamagnetic bead (streptaviding coated, 1 μm, Dynabeads MyOne, Invitrogen), and two fiducial markers (one superparamagnetic bead and one polystyrene bead placed on the bottom of the flow-cell) is performed using video-microscopy (at 50 Hz or 100 Hz). Forces applied on the laser-trapped beads were extracted from the position difference of the set trap position and the video-microscopy-based bead positions using the trap stiffness determined from back-reflection calibrations.

Buffers and DNA constructs

All measurements were carried out at 22 °C and were performed in a buffer of 10 mM Tris (pH 7.4), 100 mM NaCl, 0.125 mM $CaCl_2$. λ DNA 48 kB construct: The 12 base pair overhangs on the λ-DNA are filled by a Klenow polymerase for 2 h at 37°using biotin-labeled cytosines and regular ATP, GTP and TTP nucleotides. The construct was purified by phenol extraction, ethanol precipitated and the pellet was finally dissolved in 10 mM Tris (pH 8.0), 1mM EDTA, 1% ethanol.

DNA Magnetic tweezers construct

A 12kb magnetic tweezers construct was prepared by PCR and subsequent ligation of a digoxygenin labelled handle as follows: A 11940 kb fragment was prepared by pcr on a λ DNA template. A 1238 bp digoxigenin-labeled fragment was prepared by PCR on a pbluescriptIISK+ template a standard PCR reactions was performed, except 2 μl of digoxigenin-11-2'-deoxy-uridine-5'- triphosphate (dig-dUTP, Roche) is added. PCR products were purified with nucleospin extract II kit (Machery Nagel). Both fragments were cut with XhoI restriction enzyme giving to 2 fragments after digestion (554 and 684 bp) for the digoxygenin labeled handle and a 11926 bp fragment from the λ DNA pcr. The digestions were purified with a nucleospin extract II kit (Machery Nagel), mixed and ligated with T4 DNA ligase the labeled ends were added in 10 molar excess. To purify the construct the ligation was phenol extracted and ethanol precipitated. The pellet is dissolved in 10 mM Tris (pH 8.0), 1mM EDTA, 1% ethanol.

References

1 Goldberg, A. D., Allis, C. D. & Bernstein, E. Epigenetics: A landscape takes shape. *Cell* **128**, 635-638, (2007).

2 Greenleaf, W. J., Woodside, M. T. & Block, S. M. High-resolution, single-molecule measurements of biomolecular motion. *Annu. Rev. Biophys. Biomol. Struct.* **36**, 171-190, (2007).

3 Strick, T. R. *et al.* Stretching of macromolecules and proteins. *Rep. Prog. Phys.* **66**, 1-45, (2003).

4 Frank, J. Single-particle imaging of macromolecules by cryo-electron microscopy. *Annu. Rev. Biophys. Biomol. Struct.* **31**, 303-319, (2002).

5 Lindsay, S. M. *et al.* Stm and Afm Images of Nucleosome DNA under Water. *J. Biomol. Struct. Dyn.* **7**, 279-287, (1989).

6 Binnig, G., Quate, C. F. & Gerber, C. Atomic Force Microscope. *Phys. Rev. Lett.* **56**, 930-933, (1986).

7 Moffitt, J. R., Chemla, Y. R., Smith, S. B. & Bustamante, C. Recent advances in optical tweezers. *Annu. Rev. Biochem.* **77**, 205-228, (2008).

8 Noom, M. C., van den Broek, B., van Mameren, J. & Wuite, G. J. L. Visualizing single DNA-bound proteins using DNA as a scanning probe. *Nat. Methods* **4**, 1031-1036, (2007).

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

10 De Vlaminck, I. & Dekker, C. Recent advances in magnetic tweezers. *Annu. Rev. Biophys.* **41**, 453-472, (2012).

11 Strick, T. R., Allemand, J. F., Bensimon, D., Bensimon, A. & Croquette, V. The elasticity of a single supercoiled DNA molecule. *Science* **271**, 1835-1837, (1996).

12 Huisstede, J. H. G., van der Werf, K. O., Bennink, M. L. & Subramaniam, V. Force detection in optical tweezers using backscattered light. *Opt. Express* **13**, 1113-1123, (2005).

13 Brewer, L. R., Corzett, M. & Balhorn, R. Protamine-induced condensation and decondensation of the same DNA molecule. *Science* **286**, 120-123, (1999).

14 Wuite, G. J. L., Davenport, R. J., Rappaport, A. & Bustamante, C. An integrated laser trap/flow control video microscope for the study of single biomolecules. *Biophys. J.* **79**, 1155-1167, (2000).

15 Stigter, D. Interactions of highly charged colloidal cylinders with applications to double-stranded DNA. *Biopolymers* **16**, 1435-1448, (1977).

16 Vipond, I. B., Baldwin, G. S. & Halford, S. E. Divalent metal-ions at the active-sites of the EcoRV and EcoRI restriction endonucleases. *Biochemistry* **34**, 697-704, (1995).

17 Bouchiat, C. *et al.* Estimating the persistence length of a worm-like chain molecule from force-extension measurements. *Biophys. J.* **76**, 409-413, (1999).

18 Dame, R. T., Noom, M. C. & Wuite, G. J. L. Bacterial chromatin organization by H-NS protein unravelled using dual DNA manipulation. *Nature* **444**, 387-390, (2006).

19 Taylor, K. A. & Glaeser, R. M. Electron-diffraction of frozen, hydrated protein crystals. *Science* **186**, 1036-1037, (1974).

20 Kowalczyk, S. W., Hall, A. R. & Dekker, C. Detection of local protein structures along DNA using solid-state nanopores. *Nano Lett.* **10**, 324-328, (2010).

21 Spiering, A., Getfert, S., Sischka, A., Reimann, P. & Anselmetti, D. Nanopore translocation dynamics of a single DNA-bound protein. *Nano Lett.* **11**, 2978-2982, (2011).

22 Klumpp, S. & Hwa, T. Stochasticity and traffic jams in the transcription of ribosomal RNA: Intriguing role of termination and antitermination. *Proc. Natl. Acad. Sci. USA* **105**, 18159-18164, (2008).

23 Tripathi, T. & Chowdhury, D. Interacting RNA polymerase motors on a DNA track: Effects of traffic congestion and intrinsic noise on RNA synthesis. *Phys. Rev. E* **77**, (2008).

24 Galburt, E. A. *et al.* Backtracking determines the force sensitivity of RNAP II in a factor-dependent manner. *Nature* **446**, 820-823, (2007).

25 Vucinic, D. & Sejnowski, T. J. A Compact Multiphoton 3D Imaging System for Recording Fast Neuronal Activity. *Plos One* **2**, (2007).

26 Wallin, A. E., Ojala, H., Haeggstrom, E. & Tuma, R. Stiffer optical tweezers through real-time feedback control. *Appl. Phys. Let.* **92**, (2008).

CHAPTER 5

Determining the mechanical strength of a single cohesin complex

Cohesin provides the essential linkage between sister chromatids by entrapping their DNA within a tripartite ring. During cell division, cohesin must withstand the forces that align chromosomes by pulling them towards the opposing spindle poles via microtubules. Experimentally it has not been possible to determine the mechanical strength of the cohesin complex or to even visualize how sister chromatids are linked by cohesin. Here we report our progress on developing an assay that can directly measure the strength of a single cohesin complex. Cohesin-concatenated circular minichromosomes were isolated from yeast by a magnetic bead affinity separation via tetO repeats present in the DNA of the artificially constructed minichromosomes. An enzymatic nicking and ligation strategy was used to covalently attach DNA linkers that enable us to create minichromosome-cohesin-minichromosome tethers between an optically trapped bead and the surface of a flow cell. The isolation procedure resulted in a high yield of minichromosomes in a near native state, i.e., DNA-bound proteins and nucleosomes were still present. First single-molecule force-spectroscopy experiments as well as images of concatenated chromosomes obtained by atomic force microscopy are presented.

Marijn T.J. van Loenhout, Sara Cuylen, Suzanne Hage, Allard Katan, Iwijn De Vlaminck, Christian Haering, and Cees Dekker. *To be submitted.*

Introduction

Cell division presents a tremendous task as the genome not only needs to be replicated but it also must be equally divided to ensure that each daughter cell receives one complete copy. During cell division the genome is organized and compacted by plectonemes, nucleosomes, and proteins that pack the DNA into chromosomes of only a few micrometer in length as first observed by Flemming[1] (Figure 5.1A). How does a cell ensure that identical sister chromatids are paired and organized in space such that one chromatid moves to one cell pole while the other moves to the opposite? Sister chromatid cohesion is key to this problem and essential for successful division[2,3]. The protein complex cohesin provides this sister chromatid cohesion by linking the two chromatids together at the centromere. Cohesin must withstand the dynamic tug-of-war by kinetochore-attached microtubules (Figure 5.1B) that results in bi-orientation and proper chromatid division[4,5]. Once all chromosomes are properly bi-orientated, separase cleaves the cohesin ring and the sister chromosomes move to the opposing poles, marking the transition into anaphase.

The cohesin complex consists of four subunits; an hetero-dimer of Smc1 and Smc3 proteins, a Scc1 protein that closes the tripartite ring, and Scc3 that associates with Scc1 (Figure 5.1B). Smc1 and Smc3 are members of the structure and maintenance of chromosomes (SMC) family of proteins and are composed of 50-nm-long antiparallel coiled coils creating a rod-like protein with a hinge domain at one end and an ATP-nucleotide-binding domain (NBD) at the other. The NDBs of each of the two SMC subunits tightly associate with low nanomolar K_d when ATP is sandwiched in between them[6]. Scc1 closes the ring connecting Smc1 and Smc3 forming a huge protein structure approximate 40 nm in diameter, that can easily encompass two strands of DNA containing nucleosomes[6,7]. ATP binding and hydrolysis by Smc1 and Smc3 are essential *in vivo* but their exact role still remains unclear. Several models for the concatenation of chromatids by cohesin have been put forward the available experimental evidence points to a model where the DNA of both chromatids is entrapped within a single cohesin ring[8-10].

Central to proper bi-orientation of the chromosomes during cell division is the physical tension established between sister chromatids by kinetochores and attached microtubules that pull toward the spindle poles. Nicklas[11,12] first recognized that this tension is necessary and directly demonstrated this by artificially creating tension inside a dividing cell using a microneedle. The remarkable property that kinetochores-microtubule connections are stable under tension, but unstable if tension is absent creates an error-correcting mechanism which properly positions the chromosomes on the metaphase plate[4,5].

Cohesin complexes thus must withstand the mitotic spindle forces. Surprisingly, a single cohesin complex can withstand these forces and ensure successful chromatid division as shown by Haering *et al.*[8], using small artificially constructed minichromosomes in yeast. Perhaps even more remarkable is the fact that the entrapment between chromosomes provided by cohesin can be functional over forty years: human oocytes

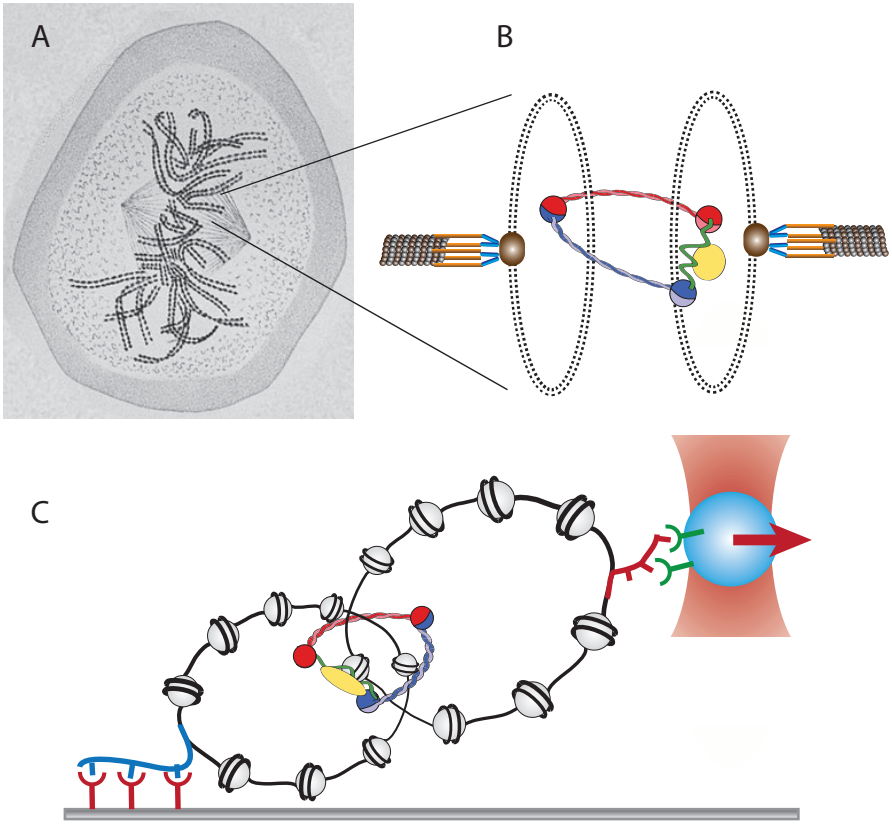


Figure 5.1: Cohesin forms the essential link during mitosis to sustain tension between sister chromosomes. (**A**) Drawing of a salamander cell during mitosis by Flemming in 1882[1]. Clearly visible are the condensed chromosomes aligned in the middle of the cell on the metaphase plate. Microtubules attach at the centromere to the chromosomes and extend to the left and right to the spindle pole bodies. (**B**) Zoom in, showing two circular minichromosomes (black) that are held together by a cohesin complex consisting of Smc1 and Smc3 subunits (red and blue) that are linked together by Scc1 (green) to close the ring. Microtubules attach via the kinetochore to the centromere on the chromosome and pull on the sister chromosomes to create the essential tension for successful cell division. (**C**) Proposed experimental assay to test the strength of cohesin complexes. Two circular minichromosomes linked by a cohesin complex are tethered between a surface and a polystyrene bead held by optical tweezers. This experimental geometry closely mimics the geometry encountered *in vivo*, The minichromosomes are attached via functionalized linkers (blue and red) that attach to antibodies on the flow cell and bead surface.

are formed in the fetus and then arrest with cohesed chromosomes until they eventually develop into a short lived egg. Cohesin deterioration seems to be directly linked to loss of chromatid pairing in these arrested oocytes and can result in aneuploidy or chromosome missegregation[13].

Direct measurements of the mechanical strength and stability of the cohesin complex are missing although these properties are central to its function. We here report our progress to develop an experimental assay that directly probes its rupture force by pulling on a single cohesin ring that connects two circular minichromosomes with optical tweezers (Figure 5.1C), thus creating a geometry that closely mimics that encountered *in vivo*. Images of the nature of the entrapment of chromosomes by cohesin are also missing likely resulting from the fact that cohesin can only be loaded onto DNA *in vivo*, which makes it impossible to investigate cohesin-concatenated DNA molecules by overexpressing and purifying cohesin complexes and *in vitro* assembling them onto DNA. To overcome these hurdles, we developed a highly efficient method for attachment of functionalized linkers and purification of cohesin-concatenated minichromosomes under mild conditions, thereby preserving cohesin and DNA bound proteins. We will first describe the purification method and covalent attachment of linkers, which are essential to perform force spectroscopy experiments. We obtain high yields that allows us to image concatenated minichromosomes by atomic force microscopy (AFM). Finally we report on the first measurements using a custom-built optical-tweezers device[14] to pull on cohesin-concatenated minichromosomes and determine its breaking strength.

Results

Affinity-based isolation of cohesin-bound minichromosomes

A method for the isolation of concatenated minichromosomes and the covalent attachment of functionalized linkers for optical tweezers was developed as outlined step-by-step in Figure 5.2. A circular minichromosome of 6 kb in length was introduced in yeast. We used an affinity-based purification using the interaction between Tet repressor and its operator, tetO, to isolate minichromosomes via magnetic beads. Subsequently biotin- and digoxygenin-functionalized oligomers were ligated to a single-stranded DNA (ssDNA) gap created by nicking the minichromosomes at two locations with nicking enzymes[15-19]. Finally, the functionalized minichromosomes were released by disrupting the TetR-tetO interaction. The method is described in more detail below.

Yeast cells containing the minichromosomes and expressing TetR-PK6 and were grown until midlog phase and arrested in mitosis using nocodazole, which interferes with microtubule polymerization. This ensures that all chromosomes are fully replicated and that microtubule-kinetochores connections are absent. Yeast cells were then washed, spheroplasted by treatment with lyticase, and lysed by the addition of a Triton X-100 buffer. The lysate was centrifuged at 12,000 g for 5 minutes to separate

the minichromosomes from the larger host chromosomal DNA and cell debris. The minichromosomes where now isolated from the supernatant using a tandem array of 21 tetO operators present in their sequence. Protein-G magnetic beads with anti-PK antibody were incubated overnight with the supernatant to bind the minichromosomes via the binding of Protein-G to anti-PK antibody to TetR-Pk6 to tetO. This procedure resulted in a high yield of minichromosomes (Figure 5.2A).

In the next steps, we covalently attached functionalized oligomers to the minichromosomes. The attachment of the minichromosomes to the magnetic beads allows for convenient handling and buffer exchange. Minichromosomes where incubated with nicking enzymes Nb.BbvCI and Nt.BspQI to remove a 9 nucleotide short ssDNA fragment (TTT TTT TGC). Nicking was efficient with efficiencies of 57 % and 94 % for Nb.BbvCI and Nt.BspQI respectively resulting in an ~54 % overall efficiency in creating the ssDNA fragment (Figure 5.2B). The nicking enzymes where then removed by washing the beads.

To attach functionalized linkers for force spectroscopy pulling experiments two 50 nt ssDNA oligomers with the 9 nt complementary sequence at its end and containing either three digoxygenin or three biotin labels and were added to the bead-bound minichromosomes. Due to its short length the 9 nt fragment is only weakly bound to its complementary sequence on the minichromosomes (melting temperature < 20° C) and it is therefore easily replaced by the 50 nt ssDNA added in high concentration (1 μM)[18]. The 50 nt ssDNA fragment was ligated in the gap using T4 DNA ligase at 16° C creating a stable covalent linkage. We were unable to test the ligation efficiency in the minichromosomes, but a test using plasmid DNA containing the same 9 nt nicking sequence showed that no covalently closed plasmid could be reformed in the presence of the 50 nt ssDNA, indicating efficient replacement of the 9 nt fragment (Figure 5.2C). The two cohesed minichromosomes are fully identical, but two different functional groups are required, as otherwise both minichromosomes would bind only to the bead or flow cell surface. By using a 1:1 mixture of functionalized ssDNA oligos containing either three biotin or three digoxygenin groups, we expect that 50 % of the minichromosome-cohesin-minichromosome complexes will have biotin labels in one minichromosome and digoxygenin labels in the other.

The functionalized minichromosomes were now released from the magnetic bead by addition of anhydrotetracycline, which releases TetR from the tetO repeats present in the DNA sequence of the minichromosomes, resulting in a > 50 % release (data not shown). This procedure allowed us to obtain a concentrated minichromosome solution with and estimated concentration of 1 pM, see discussion section. We characterized the cohesed minichromosomes by AFM and force-spectroscopy rupture experiments to determine the breaking strength of the cohesin complex.

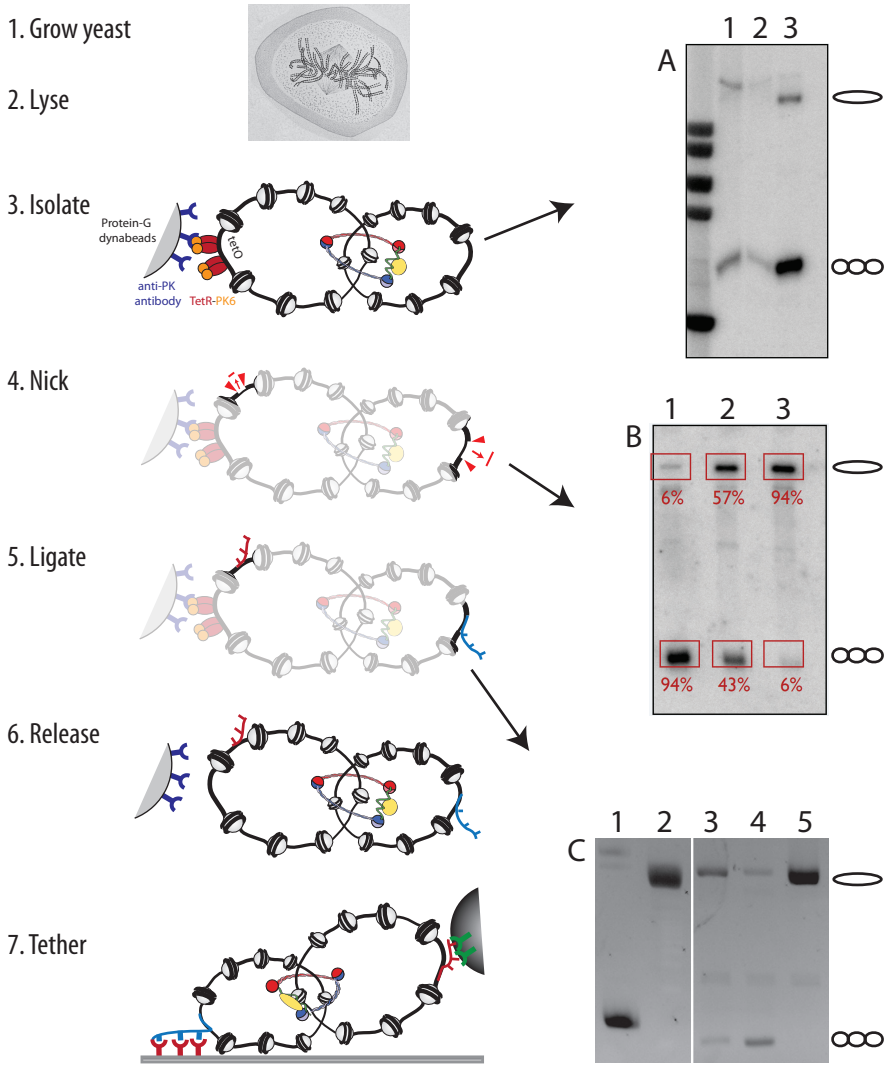


Figure 5.2: Purification of cohesed minichromosomes and attachment of functionalized linkers to allow tethering between a bead and a flow cell surface. The step-by-step procedure (outlined with accompanying cartoons) shows the procedure to obtain the final minichromosome-cohesin-minichromosome tethers. (**1**) Yeast cells are grown and arrested with nocodazole. Subsequently cells are spheroplasted and lysed (**2**). The lysate is centrifuged to obtain the chromosome fraction. (**3**) This fraction is incubated with protein-G magnetic beads which bind the minichromosomes via a PK-antibody to TetR-PK6 proteins, that attach to the tetO repeats present on the minichromosomes. The cohesed minichromosomes are now bound to the magnetic beads allowing for efficient handling

in the next steps. (**4**) A 9 nt fragment is removed from the minichromosomes by nicking with two nicking enzymes (Nt.BspQI and Nb.BbvCI). (**5**) 50 nt functionalized oligos containing 3 biotin or 3 digoxygenin groups are hybridized and ligated in the 9 nt gap that was created after nicking. (**6**) Minichromosomes are released from the magnetic beads by incubation with anhydrotetracycline, that releases the TetR from the tetO repeats. (**7**) Minichromosomes are incubated with antidigoxygenin polystyrene beads and introduced in a streptavidin-functionalized flow cell to create tethers. (**A**) Gel showing the isolation efficiency. Both nicked and supercoiled minichromosomes fractions are detected as indicated by the symbols on the right. (lane 1)The input fraction of minichromosomes in the lysate. (lane 2) The unbound fraction after incubation with protein-G beads. (lane 3) A 5X concentrate of the bound fraction eluted from the beads using SDS. (**B**) Nicking efficiency of minichromosomes. (lane 1) 94 % of the minichromosomes bound to the magnetic beads were in a supercoiled state. (lane 2) Incubation with Nb.BbvCI results in 57 % nicked minichromosomes. (lane 3) Nt.BspQI is more efficient giving 94 % nicked minichromosomes. (**C**) Ligation efficiency of a plasmid containing the same nicking sequence as the minichromosomes. (lane1) The plasmid is predominantly in a supercoiled state. (lane 2) Incubation with Nt.BspQI and Nb.BbvCI results in the non-supercoiled nicked form. (lane 3) Incubation of the nicked plasmid with T4 ligase recovers a small proportion of supercoiled plasmid. (lane 4) Incubation of the nicked plasmid in the presence of 1 μM 9 nt oligomers with a complementary sequence to the nicked region increases the fraction of supercoiled plasmid as the additional complementary oligomers are ligated in the gap. (lane 5) Incubation of the nicked plasmids in the presence of a 50 nt fragment with a complementary sequence to the 9 nt nicked region prevents the formation of supercoiled plasmid. This indicated efficient ligation of the 50 nt oligomer in the gap.

AFM imaging of cohesin bound minichromosomes.

We imaged minichromosomes isolated by the procedure described above to determine if they were still cohesed after the purification and to see if it would be possible to determine the nature of their entrapment by cohesin. Figure 5.3 shows AFM micrographs of the minichromosomes deposited on mica and imaged in air. Most of the minichromosomes had several globular features present along their length, which could be nucleosomes or DNA-bound proteins. We were successful in locating several DNA molecules that had a total contour length of approximately twice the 2 μm expected for a single minichromosome, indicating that these were likely a couple of two cohesed minichromosomes or catenated DNA circles (Figure 5.3B,C). In these first experiments only a limited number of micrographs showing minichromosomes were recorded and we were unable to identify the cohesin complexes concatenating these minichromosomes. These first images of cohesed chromosomes do however provide a clear basis for future experiments.

Force spectroscopy of minichromosome-cohesin-minichromosome tethers

To determine the mechanical stability or cohesin complexes under force, we used a custom-built optical-tweezers setup14 to pull on minichromosome-cohesin-minichromosome tethers. The experimental geometry closely resembles that encountered *in vivo*, where pulling forces are applied on both chromatids via the kinetochore (cf.

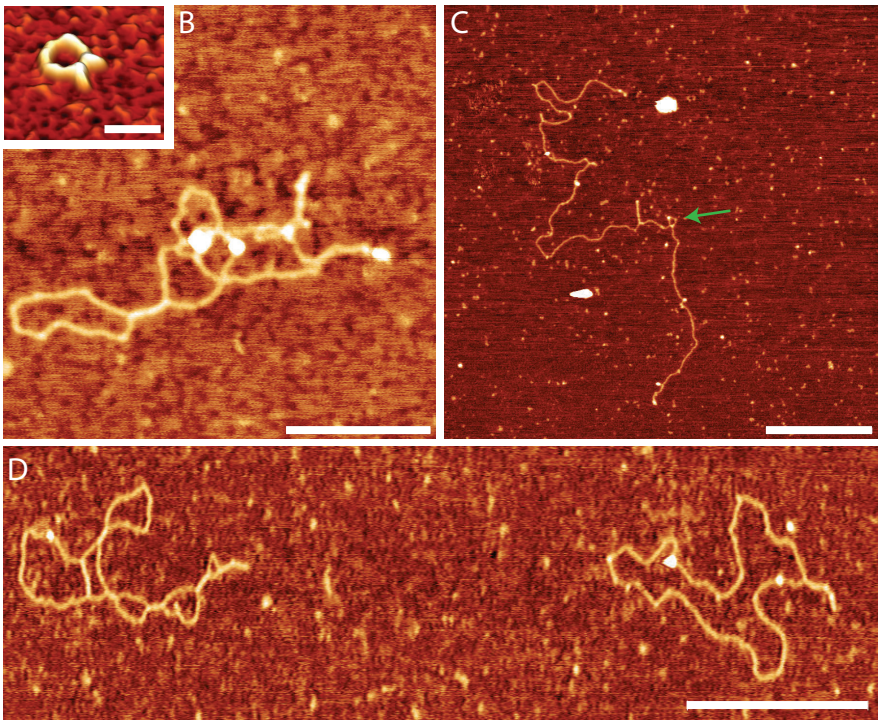


Figure 5.3: AFM images of cohesed minichromosomes. (**A**) AFM image showing a ring-like structure with features expected for a cohesin complex (scale bar 50 nm). (**B**)AFM image of cohesed minichromosomes deposited on mica. Two minichromosomes are visible which are possibly linked by a cohesin complex. Globular features are present on the minichromosomes which could correspond to nucleosomes, *z*-scale 0–2 nm scale bars 500 nm. (**C**) Image showing two minichromosomes which are possibly tethered by a cohesin complex at the position indicated by the green arrow, *z*-scale 0–2 nm scale bar 500 nm. (**D**) AFM image showing two minichromosomes in the same field of view. Globular features are visible which could correspond to nucleosomes.

Fig 5.1B and C). To perform the force-spectroscopy experiments, we tethered the minichromosomes and subsequently verified the length of these tethers (Figure 5.4). Each experiment started by incubating the purified minichromosome solution with 2 µm diameter antidigoxygenin-coated polystyrene beads. These beads were then flushed into a flow cell that was surface passivated with polyethylene glycol (PEG) and had streptavidin proteins linked covalently to the surface via PEG-biotin linkers. The beads were allowed to settle and form minichromosome tethers between the beads and the surface. Tethered beads were visually identified and the length of the tether was determined by moving the bead in the x- and y-direction using optical tweezers at low laser intensity (Figure 5.4B). The reduced laser intensity limited the maximum

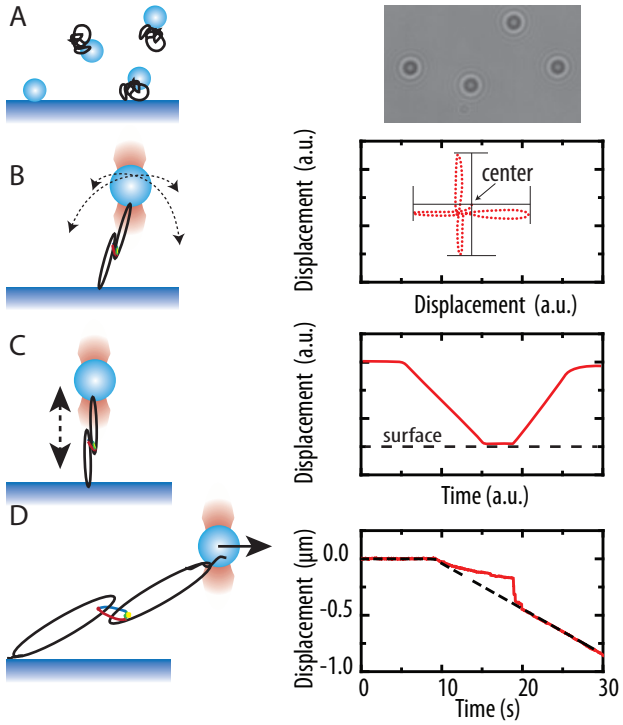


Figure 5.4: Experimental assay to test the strength of the cohesin complex. (**A**) Polystyrene beads with bound minichromosomes were allowed to settle and attach to the bottom of the flow cell via digoxygenin antibodies. (**B**) The length of the tether was verified and the center of the attachment point was determined by moving the bead at low trap stiffness in the *xy*-plane. The center of the attachment point was calculated from the maximum excursions where the bead was pulled out of the trap in *x*- and *y*-direction. (**C**) The position of the flow cell surface was determined by moving the bead downward until it hit the surface. (**D**) Rupture experiment. The bead was moved sideways at a height of 1 μm above the flow cell surface to determine the breaking strength of the minichromosome-cohesin-minichromosome tether.

amount of force that could be applied to the tethers thereby preventing premature breakage. If the tethers had a length that could not be explained by simple bead sticking (i.e. >1 μm), we proceeded by determining the position of the flow cell surface by moving the bead downward until it hit the surface (Figure 5.4C). To determine the breaking strength of the tether the bead was now positioned 1 μm above the center of the attachment point and moved horizontally away from the attachment point at a constant height until the tether broke (Figure 5.4D). During this displacement the position of the bead and its deflection from the optical-trap center were monitored using video microscopy. Figure 5.4D illustrates such a pulling experiment.

An overview of initial results of the force-spectroscopy rupture experiments is presented in Figure 5.5. The trap position and the actual bead position of a typical pulling experiment are shown in Figure 5.5B. Initially, the bead follows the trap, but as the minichromosome tether is extended, a tension develops which deflects the bead from the trap position, until finally the tether breaks and the bead returns to the position of the moving trap. For small deviations from the trap position, the deflection corresponds linearly to the force applied on the tether (Figure 5.5C). Rupture events can be seen as a sudden decrease of the force. Force curves for 7 different tethers (Figure 5.5C–I) often show that multiple rupture events are observed. Such events could result from the unwrapping of nucleosomes on the DNA or from other conformational changes in the tether that increase its length. The breakage may be a signature of the rupture of the cohesin complex, but further experiments are needed to establish this. Most rupture forces were in the range of 40–80 pN, two rupture events were seen to occur at substantially higher forces (Fig.5 G and I), these forces are expected to be inaccurate as these correspond to displacements far outside the linear range of the optical trap. These higher observed forces could potentially also correspond to DNA-DNA catenations that were topologically linked without cohesin.

Discussion and outlook

We developed a method to isolate and attach functional linkers to circular minichromosomes that were linked together by cohesin. Preliminary AFM images of cohesed minichromosomes were recorded and we performed first rupture experiments to determine the mechanical strength of the cohesin complex. Recently, two other affinity based methods have been developed to immobilize protein complexes from cell lysate for single-molecule fluorescence methods, but these did not enable covalent attachment of linkers[20,21]. Isolating cohesed minichromosomes for force spectroscopy measurements presents a substantial challenge as only a single pair cohesed minichromosomes is present in the cell compared to proteins that may be overexpressed. Moreover, no link in the tether should be weaker than the cohesin complex to be certain that the measured rupture events correspond to the breakage of cohesin. We therefore used a nicking and ligation method to covalently link handles that contained multiple functional groups to the minichromosomes.

The purification method uses mild conditions throughout, i.e. physiological pH, modest salt concentrations (50–200 mM NaCl), and temperatures ≤ 37°C. These benign conditions should not only leave the cohesin complex intact, but will also prevent degradation or loss of other DNA bound proteins. The AFM micrographs of the minichromosomes indeed confirm the presence of DNA-bound proteins, as globular feature were observed along the DNA molecules that could be nucleosomes or other DNA-bound proteins. The force-spectroscopy pulling experiments also showed multiple rupture events, which suggest the presence of nucleosomes on the DNA. The method that we present here will therefore be applicable to many other experiments

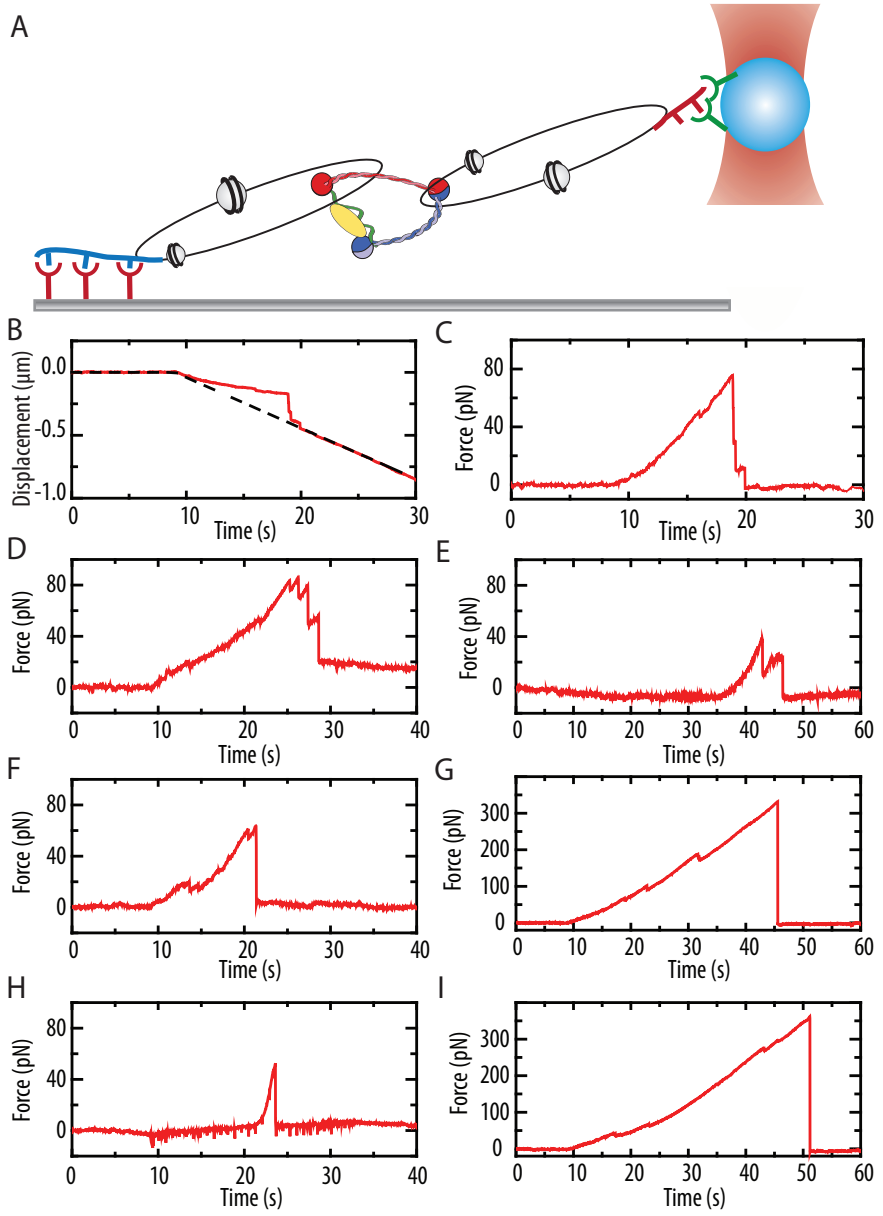


Figure 5.5: Force spectroscopy measurements of minichromosome-cohesin-minichromosome tethers. (**A**) Schematic drawing (not to scale) showing the experimental assay. Note that nucleosomes may still be present on the minichromosomes. (**B**) Displacement of the optically trapped bead during the experiment (red line). The bead is displaced com-

pared to the position of the moving trap (black dashed line) by the tension in the tether until it breaks at t ≈ 19 s. (**C**) Force corresponding to the trace shown in panel B. (**D–I**) Examples of experimentally obtained force curves. Note that multiple rupture events often are visible. Recorded rupture force values above 100 pN are inaccurate due to non-linearity of the trap stiffness for such large forces/displacements.

where chromosomes must be labeled and isolated in a near-native protein-bound state.

There have been very few reports of molecular scale imaging or single-molecule experiments of cohesed chromosomes. The complexity and limited yield of available purification methods likely is responsible for the lack of such experiments. Our method resulted in yields that allowed AFM and force-spectroscopy experimetns. We will now estimate how many cohesed chromosomes can be obtained by our purification method and discuss the consequences for single-molecule experiments. A starting culture of 2 l yeast cells at OD 0.6 contains approximately $1{\cdot}10^{10}$ yeast cells and an equal number of cohesed minichromosomes[22]. The multiple purification steps result in losses of around 50 % for each step and we estimate the final recovery to be approximately 5 %. After purification we thus obtain roughly $5{\cdot}10^{8}$ minichromosomes in a volume of 600 µl, corresponding to a concentration of 1.3 pM. We now compare this number to the number of molecules normally used in single-molecule experiments, e.g. magnetic tweezers and AFM. A magnetic tweezers experiments starts with the creation of DNA tethers, similarly to the pulling experiments described here, and uses approximately $1{\cdot}10^{7}$ beads and an equal number of DNA molecules. The total number of purified minichromosomes would thus allow for approximately 50 experiments that tether beads in a flow cell.

We can also compare the minichromosome concentration to that used in AFM experiments. A typical AFM experiment, where several DNA molecules are visible in a 100 µm2 scan area, uses a concentration of approximately 0.1 ng/µl DNA, corresponding to a 25 pM concentration of 6 kb DNA molecules. The 1.3 pM minichromosomes concentration should therefore just make it possible to find a minichromosome after scanning a few 100 μm^2 areas. This matches our experimental observations as we found a few minichromosomes after scanning many 100 μm^2 areas.

Our pulling experiments suggest the presence of minichromosome-cohesin-minichromosome tethers and their rupture forces during force spectroscopy experiments. Most recorded rupture forces where in the range of 40-80 pN (cf. Figure 5.5). However, more experiments will be required to assess the various contributions and determine the mechanical stability of the cohesin complex. We can for now only speculate on the expected rupture forces of cohesin and its implications for cell division. It is expected that cohesin is at least stable under applied forces in excess of 30–40 pN as such forces are generated *in vivo*, by for example polymerases. Breakage of cohesin at these forces would therefore seriously endanger chromatid cohesion. An upper bound for the rupture force of cohesin is harder to estimate. DNA overstretches at 65 pN resulting in structural changes and making it more prone to damage. However,

after reduction of the force, DNA will quickly return to its normal state. *In vivo* data of cells containing double centromeric chromosomes indicated that mitotic forces are strong enough to induce DNA breakage[23], which requires forces well in excess of 65 pN. If cohesin ruptures at forces around 65 pN this could therefore lead to premature loss of cohesion due to these mitotic spindle forces.

We are optimistic that the developed methods will make it possible to determine the rupture force of the cohesin complex and identify which interface within the cohesin ring breaks first, using our previously developed cohesin mutants and crosslinking strategies[8]. Such data on the interaction strength of each of the interfaces of cohesin ring will greatly aid in understanding the functional role of cohesin. We recommend that fluorescent labels are incorporated into the ssDNA linkers to create visual markers that can be used to unambiguously determine the presence of two minichromosomes and confirm rupture of the cohesin complex.

The demonstrated feasibility of AFM imaging (Figure 5.3) of the cohesin-minichromosome structures should enable experiments that can discriminate between the proposed models for chromosome entrapment by cohesin. AFM imaging could also provide new insights in the interaction of cohesin complexes with polymerases that are expected to push cohesin in front of them while moving along the DNA[24], or address the other functions of cohesin such as the stabilization of large chromatin loops during interphase through interactions with site-specific transcription factors[25].

Acknowledgments

We thank Roeland van Elsas and Johan den Blanken for assistance in developing the experimental setup and Jutta Metz for assistance with the minichromosome purification.

References

1 Flemming, W. *Zellsubstanz, Kern und Zelltheilung*. (Verlag von F.C.W. Vogel, 1882).

2 Guacci, V., Koshland, D. & Strunnikov, A. A direct link between sister chromatid cohesion and chromosome condensation revealed through the analysis of MCD1 in S-cerevisiae. *Cell* **91**, 47-57, (1997).

3 Michaelis, C., Ciosk, R. & Nasmyth, K. Cohesins: Chromosomal proteins that prevent premature separation of sister chromatids. *Cell* **91**, 35-45, (1997).

4 Ault, J. G. & Nicklas, R. B. Tension, Microtubule Rearrangements, and the Proper Distribution of Chromosomes in Mitosis. *Chromosoma* **98**, 33-39, (1989).

5 Tanaka, T. U., Fuchs, J., Loidl, J. & Nasmyth, K. Cohesin ensures bipolar attachment of microtubules to sister centromeres and resists their precocious separation. *Nat. Cell Biol.* **2**, 492-499, (2000).

6 Haering, C. H., Lowe, J., Hochwagen, A. & Nasmyth, K. Molecular architecture of SMC proteins and the yeast cohesin complex. *Mol. Cell* **9**, 773-788, (2002).

7 Gruber, S., Haering, C. H. & Nasmyth, K. Chromosomal cohesin forms a ring. *Cell* **112**, 765-777, (2003).

8 Haering, C. H., Farcas, A. M., Arumugam, P., Metson, J. & Nasmyth, K. The cohesin ring concatenates sister DNA molecules. *Nature* **454**, 297-U219, (2008).

9 Ivanov, D. & Nasmyth, K. A physical assay for sister chromatid cohesion in vitro. *Mol. Cell* **27**, 300-310, (2007).

10 Ivanov, D. & Nasmyth, K. A topological interaction between cohesin rings and a circular minichromosome. *Cell* **122**, 849-860, (2005).

11 Nicklas, R. B. How cells get the right chromosomes. *Science* **275**, 632-637, (1997).

12 Nicklas, R. B. Chromosome Micromanipulation .2. Induced Reorientation and Experimental Control of Segregation in Meiosis. *Chromosoma* **21**, 17-&, (1967).

13 Jessberger, R. Age-related aneuploidy through cohesion exhaustion. *EMBO Rep* **13**, 539-546, (2012).

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

15 Luzzietti, N., Knappe, S., Richter, I. & Seidel, R. Nicking enzyme-based internal labeling of DNA at multiple loci. *Nat. Protoc.* **7**, 643-653, (2012).

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

17 Ang, W. H., Brown, W. W. & Lippard, S. J. Preparation of Mammalian Expression Vectors Incorporating Site-Specifically Platinated-DNA Lesions. *Bioconjugate Chem.* **20**, 1058-1063, (2009).

18 Kuhn, H. & Frank-Kamenetskii, M. D. Labeling of unique sequences in double-stranded DNA at sites of vicinal nicks generated by nicking endonucleases. *Nucleic Acids Res.* **36**, (2008).

19 Wang, H. X. & Hays, J. B. Simple and rapid preparation of gapped plasmid DNA for incorporation of oligomers containing specific DNA lesions. *Mol. Biotechnol.* **19**, 133-140, (2001).

20 Yeom, K. H. *et al.* Single-molecule approach to immunoprecipitated protein complexes: insights into miRNA uridylation. *Embo Reports* **12**, 690-696, (2011).

21 Jain, A. *et al.* Probing cellular protein complexes using single-molecule pull-down. *Nature* **473**, 484-U322, (2011).

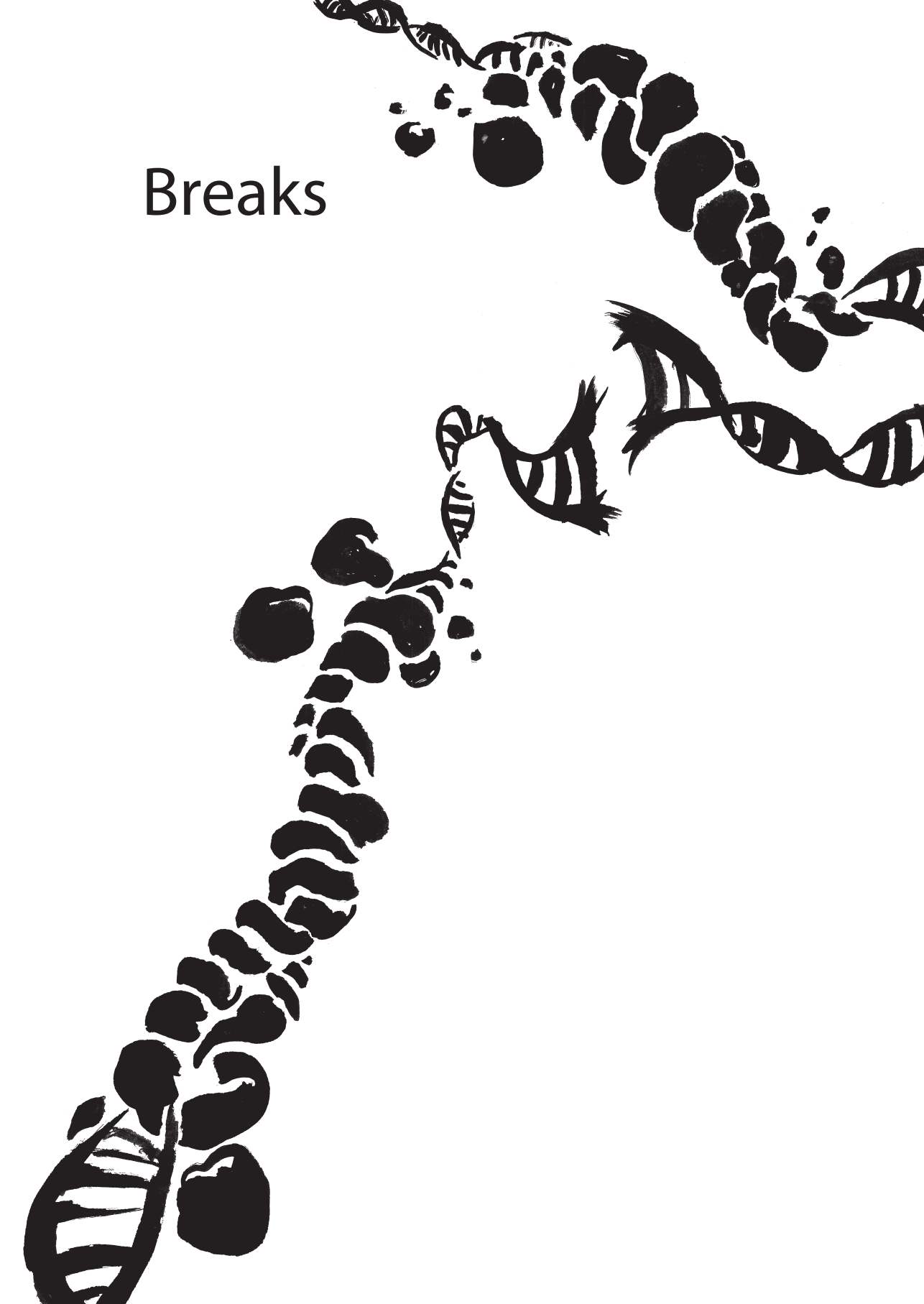
22 Day, A., Schneider, C. & Schneider, B. L. *Methods in molecular biology*. 55-76 (2004).

23 Brock, J. A. & Bloom, K. A chromosome breakage assay to monitor mitotic forces in budding yeast. *Journal of cell science* **107 (Pt 4)**, 891-902, (1994).

24 Lengronne, A. *et al.* Cohesin relocation from sites of chromosomal loading to places of convergent transcription. *Nature* **430**, 573-578, (2004).

25 Schmidt, D. *et al.* A CTCF-independent role for cohesin in tissue-specific transcription. *Genome Res.* **20**, 578-588, (2010).

Breaks



Part 3—Breaks

Things that are heavily used eventually break down. The same holds for DNA, which is intensely used and constantly transcribed, to create new proteins, and replicated, to form new DNA for daughter cells. DNA breakage poses a serious problem as the genetic information is essential to the cell and breaks in the DNA may lead to cell death or cancer. Breaks can result from mechanical stress or radiation, similar to a rope that may break under force or after prolonged exposure to sunlight. Many of the proteins involved in DNA repair were identified by exposing cells to radiation, and these are appropriately named RAD#. Storing the DNA for safekeeping only, without using it, is generally not an option and breaks therefore occur and must be repaired. The simplest option for break repair is to stick the two broken ends back together. Cells indeed use this method, but it has the major drawback that information may be lost at the site of the break. All forms of life therefore also use a second mechanism, which guarantees accurate repair by using a spare copy with the identical, or homologous, DNA sequence. We humans have 23 pairs of chromosomes, 1 copy of each from our father and mother, so if one of these breaks, there is a spare copy that may be used for repair by homologous recombination. In this part of the thesis we explore several stages of this important repair mechanism. After a double-strand DNA break occurs it must first be detected, and this task is performed by the MRE11-RAD50 (MR) protein complex. MR not only detects and prepares the break for repair but it also keeps the two broken ends together. Next, the DNA ends are processed to have an overhang of single-stranded DNA, i.e., one of the strands making up the DNA double helix is removed from the DNA end. On this single-stranded piece of DNA, a helical protein filament forms, made up of RecA or RAD51 proteins. In humans, the BRCA2 protein, named after its direct link to breast cancer, facilitates and controls the assembly of the RAD51 filament. This RAD51 or RecA filament now performs the remarkable task of searching through all the millions, or even billions, of bases to find the homologous sequence, that can be used for repair. Double-strand break repair presents one of the wonders of molecular biology. The fact that it is intimately tied to cancer makes understanding it even more so important. Non-functional repair may result in the formation of cancer cells, paradoxically, once developed these cancer cells themselves become reliant on DNA repair mechanisms to survive and adapt.

CHAPTER 6

ATP triggers DNA release resulting in conformational changes in MRE11-RAD50-NBS1 complexes

The MRE11-RAD50-NBS1 (MRN) complex recognizes, signals, and processes double-strand breaks (DSBs), which are among the most genotoxic lesions in DNA. MRN acts as molecular machine that responds to DNA binding with a large conformational change. The interplay between ATP binding/hydrolysis and DNA binding remains, however poorly understood. Here we determine the functional response of MRN to nucleotide binding and ATPase activity on DNA binding. Electrophoretic mobility shift assays and atomic force microscopy reveal that DNA is released from the complex upon nucleotide binding and hydrolysis. DNA release from the complex triggers a large conformational change with the appearance of an open arrangement of the RAD50 coiled coils. Our results support a regulatory role for nucleotide binding, which not only affects affinity for DNA, but may also triggers changes in MRN intercomplex interactions.

Marijn T.J. van Loenhout, Eri Kinoshita, Edwin Rijgersberg, Eddy van der Linden, Cees Dekker, and Claire L. Wyman

Introduction

Double-strand breaks (DSBs) are amongst the most severe forms of DNA damage, and their repair is essential to cell survival[1]. DSBs can arise accidentally during normal cell metabolism, but can also be induced by exogenous stresses or in programmed genome rearrangements[2]. To ensure timely and precise repair of DSBs, a class of highly conserved proteins has evolved. DSB repair proceeds predominantly by either homologous recombination (HR) or by nonhomologous end joining (NHEJ)[3,4]. HR provides an error-free repair pathway by using the sister chromatid as a template. NHEJ, in contrast, relies on direct ligation after limited and potentially mutagenic end processing[5].

We now first review the functions of the MRE11-RAD50-NBS1 (MRN) complex and specifically address the effects of ATP binding and hydrolysis. MRN has an essential role in the early stages of DSB repair and is central to many processes involving DNA end processing including HR, NHEJ, microhomology-mediated end-joining, meiosis, and telomere maintenance[6-8]. MRN recognizes DSBs, signals for cell cycle arrest, and subsequently processes DNA ends for repair[9]. HR starts with the recognition and end processing of the broken DNA ends by MRN. Subsequently a single-stranded DNA (ssDNA) overhang is created by resection in the 5'–3' direction. A RAD50 nucleoprotein filament forms on this ssDNA overhang, which searches and performs strand invasion of the homologous duplex DNA[10]. The multiple functions of MRN are reflected in several enzymatic activities including: endonuclease and 3'–5'-exonuclease activity, ATP hydrolysis, and checkpoint signaling by kinase ataxia telangiectasia mutated (ATM) activation[11-13]. MRN also has important structural functions in DSB repair, most strikingly DNA binding and tethering of DNA ends through interactions between the end parts of the RAD50 coiled coils[14-18]. MRN thus acts as a multicomponent molecular machine whose activities are modulated by binding and release of substrates that induce changes in its architecture.

The core MR complex consists of a dimer of the endo- and exonuclease MRE11 and two RAD50 polypeptides (Figure 6.1). The MRE11 and RAD50 nucleotide-binding domains form a globular structure that has DNA binding activity and ATP-stimulated nuclease activities[19,20]. Two RAD50 coiled-coil domains protrude from the globular head structure and can form inter and intra complex connections via zinc hook domains located at the ends of the coiled coils[14-18]. NBS1 binds to this core complex with a $M_2R_2N_2$ or $M_2R_2N_4$ stoichiometry and helps to organize repair foci, and is also involved in ATM activation[21]. The complex thus contains several dimerization interfaces, and changes at any of these interfaces potentially influences interactions at the others. Binding of DNA or nucleotide substrates is also expected to change the interface interactions and complex architecture, and to modulate the biological function.

RAD50 belongs to the structural maintenance of chromosomes (SMC) protein family, which is involved in higher-order chromosome organization and dynamics[22]. SMC proteins have a Walker A and B nucleotide (NTP)-binding motif at their N- and C-terminus respectively placing them in the group of ATP-binding cassette

Figure 6.1: Architectural features of the of the MR complex. The schematic representation shows two RAD50 coiled-coils that interact at their N- and C-terminal globular domains via Walker A- and B-type ATPase domains (A and B respectively). The region in between de globular domains forms an extended coiled-coil structure with a zinc hook at its apex (CXXC) that can self-interact. DNA binding in the globular head domain results in a conformational change of the coiled coils to a parallel, extended configuration. MRE11 (labeled M) binds to the base the coiled-coil regions.

(ABC) ATPases[23,24]. ATP binds between a protein dimer where Walker A and B motifs from one subunit contact the highly conserved signature motif from another. ATP binding and hydrolysis are generally linked to architectural changes in SMC proteins. SMC dimers such as cohesin and condensing, for example, undergo association and dissociation of the ATPase heads following ATP binding and hydrolysis[25,26].

Although potential roles for ATP hydrolysis have been described, it is not yet clear how the ATPase cycle modulates functions of the MRN complex. Structural studies show that ATP binding to the RAD50-ATP-binding domains induces a rotation of the MRE11 helix-loop-helix and RAD50 coiled coil domains creating a clamp like conformation[27-30]. Remarkably however, AFM observations showed no noticeable difference in the arrangement of the coiled coils in the presence of different nucleotide cofactors[15]. A mechanism was also proposed where ATP hydrolysis forms a switch between the exo- and endo-nuclease states of MR complex[31].

We previously showed by AFM imaging that DNA binding to the MR complex has a dramatic effect on the RAD50 coiled-coil conformation[15]. The flexibility of the coiled coils allows the Zn hook domains at both ends of a single complex to interact with each other. Upon DNA binding to the MR globular domains, the coiled coils adopt a parallel configuration preventing intracomplex interactions and favoring inter-complex interactions essential for DNA tethering. It is however unclear how ATP affects DNA binding to the MR(N) complex, as both ATP-dependent[32,33] and ATP-independent DNA binding[14,16,19] have been observed. AFM images revealed large clusters of MR

formed on DNA substrates without nucleotide cofactors, whereas ATP increased oligomer binding to substrates with 3' overhangs compared to blunt ends or those with 5' overhangs[14,34]. The capability of MR(N) to form oligomers on DNA indicates another hierarchical level at which function can be controlled by modulating the ability of intercomplex interactions and possibly DNA tethering.

The interplay between DNA binding and ATP binding and hydrolysis by MRN is poorly understood leaving many questions on how substrate binding influences conformational changes and biochemical activity. To address these questions we use electrophoretic mobility shift assays (EMSA) to determine the influence of nucleotide cofactors on DNA binding by MRN. We also used AFM imaging where the conformational change of the RAD50 coiled coils is and obvious read out for DNA binding allowing us to probe the structural changes of MRN upon nucleotide binding and ATP hydrolysis in real time.

Materials and methods

Protein purification of MR and MRN complexes

Human MR and MRN complexes were produced in adherent culture by co-infection of Sf9s cells with baculovirusses expressing C-terminally 6-histidinge tagged RAD50, untagged MRE11 (MR) and untagged NBS1 (MRN). Constructs for viruses were a generous gift from T. Paull and M. Gellert. Cells were harvested after 72 hours and purified as described previously[16]. Protein concentration was verified by a Bradford assay and subsequently aliquoted, flash frozen and stored at -80°C until further use.

ATPase assays

ATPase activity of the protein preparations was determined by thin layer chromatography (TLC). Purified human MR and MRN complexes (74 nM) were mixed with radiolabeled [γ^{32}P]ATP supplemented with non-labeled ATP to concentrations of 1, 10, 20, 50, and 100 μM in a total volume of 20 μL binding buffer (25 mM Tris-HCl, pH 7.5, 110 mM NaCl, 5 mM $MgCl_2$, 5% glycerol, 1 mM DTT, 0.1% Tween, and 2% PEG-6000). The reactions were incubated at 37°C for up to 120 minutes and stopped every 30 minutes by adding 125 mM EDTA. Samples were analyzed by TLC (Merck TLC plates) run in 0.7 M K_2HPO_4 0.4 M Boric acid and quantified by phosphor imaging.

Electrophoretic mobility shift assays (EMSAs)

DNA binding experiments were performed by incubating the indicated substrate concentrations (see figure captions) for 5 minutes at 25°C with MR(N) in binding buffer (25 mM Tris-HCl, pH 7.5, 110 mM NaCl, 5 mM $MgCl_2$, 5% glycerol, 1 mM DTT, 0.1% Tween, and 2% PEG-6000) in a total volume of 20 μl. Assays addressing the effect of nucleotide cofactors on DNA binding by MR(N) were performed by adding 2 μl of 1–25 mM AMP-PNP or ATP in a final volume of 20 μl and incubating for an additional 5 minutes at 25°C. The reaction products were separated on 4.8% non-denaturing

polyacrylamide gels for reactions with bp range substrates and on 0.7 % agarose gels for substrates in the kb range (400 bp – 3 kb) in 0.5 x TBE buffer. Alexa fluor 532 labeled DNA was visualized with a 532 nm laser scanner (Typhoon 9200) using a 555-20 nm bandpass filter. Unlabeled substrates were stained with ethidium bromide and scanned similarly. Data was analyzed from at least three separate experiments and quantified using ImageQuant software.

AFM in air

MRN was incubated with DNA at the indicated concentrations (see figure captions) in protein storage buffer (20 mM Tris-HCl, pH 8.0, 100 mM NaCl, 10% glycerol, 1 mM DTT, and 0.1% Tween 20) for 1 minute at 25°C. Experiments testing the effect of nucleotides had 1 mM ATP added and were incubated for 1 minute at 25°C prior to deposition. The samples were deposited on freshly cleaved mica for 1 minute, rinsed with distilled water and dried under compressed air. Samples were imaged at 1µm × 1µm can size with at Nanoscope IV using Silicon tips (Nanoprobes , Digital Instruments).

AFM in liquid

MR(N) complex at a final concentration of 3.5 nM and DNA, 35 nM 96 bp oligonucleotides at a final concentration of 35 nM, were incubated 5 minutes at 23°C in protein storage buffer (20 mM Tris-HCl, pH 8.0, 100 mM NaCl, 10% glycerol, 1 mM DTT, and 0.1% Tween 20) supplemented with 2 mM ATP or AMP-PNP were indicated. Samples were deposited onto freshly cleaved mica and immediately mounted in a liquid cell (Veeco). Subsequently the liquid cell was filled by injecting 60 µl of imaging buffer (25 mM Tris-HCl pH 8.0, 75 mM KCl and 3 mM $MgCl_2$). The samples were imaged in tapping mode using Veeco MSNL probes with a spring constant of 0.1 N/m at a drive frequency around 9 kHz using a Nanoscope III AFM (Digital Instruments). Typically a first image could be recorded within 5 minutes of depositing MR(N) complexes on the mica. Experiments addressing the effect of nucleotides in real-time used the same procedure; however, now after 15 minutes of imaging, the buffer was exchanged for one containing 2 mM ATP (or AMP-PNP). To exchange the buffer the AFM probe was retracted 20 µm above the mica surface and 150 µl of imaging buffer (containing 2 mM of ATP where indicated) was flushed through the flowcell. Immediately after the buffer exchange the probe was approached again to the surface and imaging was continued. Images were processed by flattening to remove background slope, and protein complexes were counted manually.

Results

MR(N) preferentially binds single-stranded DNA but not double-stranded DNA ends

The recognition and binding to DNA ends by MRN is an essential first step in DSB repair. To test if DNA-end binding is an intrinsic property of the MRN complex, we determined its binding affinity to various substrates. EMSAs were used to determine the concentration at which 50 % of the DNA was bound by MR for four different double-stranded DNA substrates: supercoiled 3 kb plasmid, 3 kb linear dsDNA, 3 kb nicked plasmid DNA and 400 bp linear dsDNA (Figure 6.2A, C). Remarkably, we find that MR has no clear preference for DNA ends as the 50 % binding concentrations were respectively 20, 30, 35 and 35 nM.

To see whether MR has a preference for single-stranded DNA or substrates containing a ssDNA overhang, we prepared 90 nt ssDNA, junction DNA (45 bp + 45 nt 3'-overhang, or 45 bp + 45 nt 5'-overhang) and 90 bp dsDNA substrates. These short substrates have the advantage of minimizing protein-protein interactions that may occur for substrates in the kb range. EMSAs were performed at a constant DNA concentration of 1 nM while titrating the protein concentration (Figure 6.2B, D). MR bound to all substrates in the following order of preference: ssDNA bound best, then junction DNA (the 3'- or the 5'-overhang), and then the dsDNA. The protein concentrations at which 50 % of the DNA was bound were 1, 2, 2, and 4 nM respectively. Similar results showing a preference for ssDNA were obtained for longer DNA substrates: M13 circular ssDNA, 1.5 kb dsDNA with a 500 nt 3'-overhang, 1.5 kb linear dsDNA, and 3 kb linear dsDNA (Figure 6.2C). These substrates had 50 % binding concentrations of 10, 10, 30, and 40 nM respectively. MR thus consistently binds with an approximately 3 to 4-fold preference to ssDNA over dsDNA, irrespective of the substrate length. The EMSAs described above were repeated for MRN with similar results (data not shown). To gain insight in the binding mode of MRN to DNA, AFM images of 1.5 kb linear dsDNA with a 500 nt 3'-overhang in the presence of MRN were recorded (Figure 6.3A). These clearly show that MRN is bound to the 3'-overhang and confirm the preference for ssDNA over dsDNA.

ATP and AMP-PNP induce MRN to release DNA

To determine the effect of nucleotide cofactors on the DNA-binding by MRN, we performed AFM experiments and EMSAs in the presence and absence of ATP or the non-hydrolyzable ATP analog AMP-PNP. MR and MRN slowly hydrolyzed ATP at a measured rate of 0.044 and 0.048 min^{-1} $complex^{-1}$ respectively, typical for these complexes[34]. AFM images of MRN incubated with 3 kb linear dsDNA in the absence of ATP show large oligomeric aggregates that tether several DNA molecules (Figure 6.4A). Adding 1 mM ATP induced dissociation of the individual MRN complexes from the aggregates, resulting in many individual MRN complexes localized close to the large aggregates (Figure 6.4B). The same phenomenon was observed upon the addition of AMP-PNP (data not shown). To determine if the dissociation from the aggregates, seen by AFM,

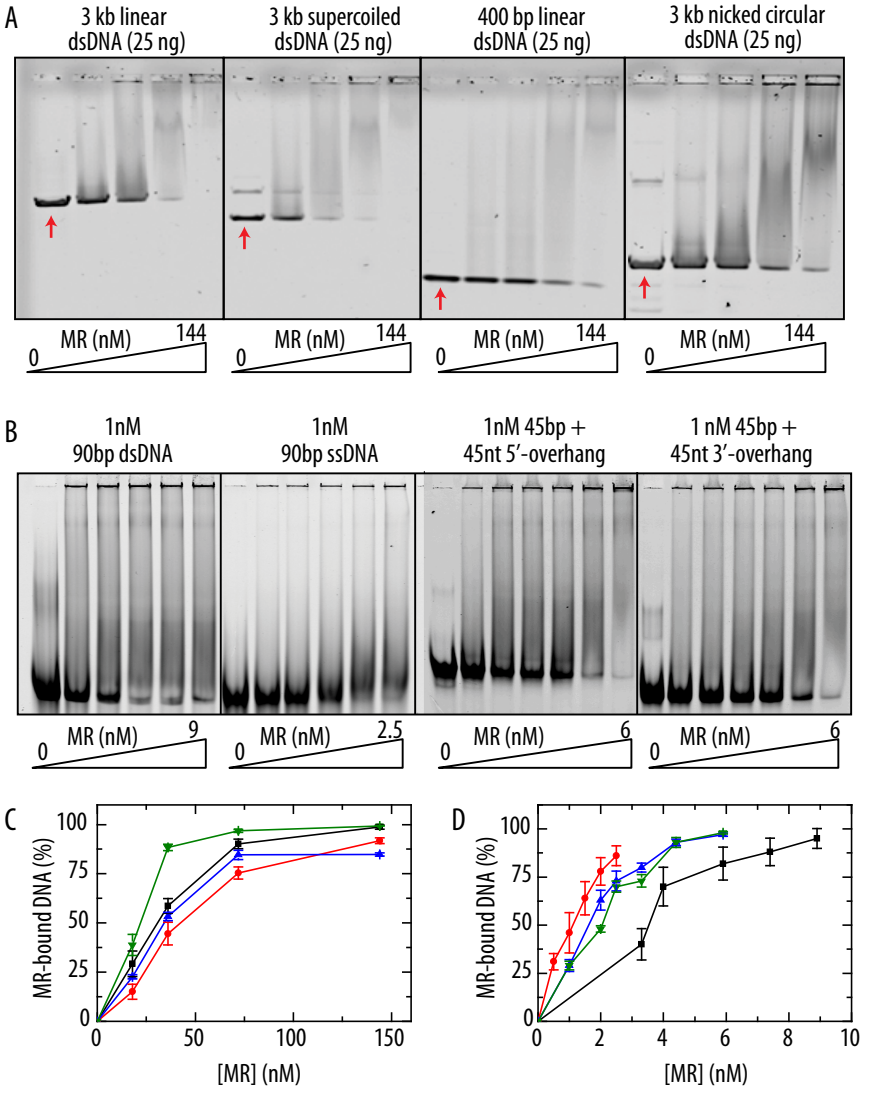


Figure 6.2: MR has no preference for DNA ends but binds preferentially to single-stranded DNA. (**A**) EMSAs showing disappearance of unbound DNA (red arrow) and increasing shift for increasing MR concentrations indicating DNA binding. (**B**) EMSAs showing MR binding to short DNA fragments, MR binds most efficiently to ssDNA or overhang DNA, note the different MR concentrations. (**C**) Graphical representation of the EMSAs shown in panel A. 3 kb linear dsDNA (black), 0.4 kb linear dsDNA (red), 3 kb nicked circular DNA (blue), and 3 kb supercoiled plasmid DNA (green). The MR-bound DNA fraction was calculated from the intensity ratio between the unbound DNA fraction in the absence and presence of MR. (**D**)

Graphical representation of the EMSAs shown in panel B, for: dsDNA 90 bp (black), ssDNA 90 nt (red), 5′-overhang DNA 45 bp dsDNA+ 45 nt ssDNA (blue), and 3′-overhang DNA 45 bp dsDNA+ 45 nt ssDNA (green).

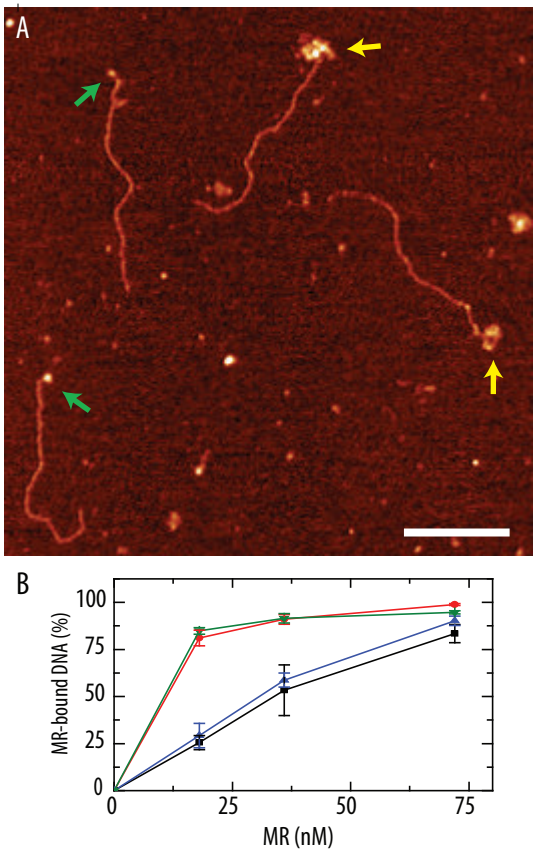


Figure 6.3: binds preferentially to single-stranded or single-stranded-DNA overhangs. (**A**) 3 nM MRN incubated with 1 nM 1.5 kb linear DNA fragment with a 500 nt 3′-overhang incubated for 10 minutes and deposited on Mica. MRN is seen bound to the ssDNA tails to some of the molecules (yellow arrows). Unbound ssDNA overhangs are visible as small blobs (green arrows) due to their short persistence length. Scale bar indicates 200 nm image z-scale corresponds to 3 nm from brown to white. (**B**) Results from EMSAs (data not shown) for 1.5 kb linear dsDNA (black), 3 kb linear dsDNA (blue), 1.5 kb linear dsDNA with 3′-overhang (red), and M13 circular ssDNA (green). The MR-bound DNA fraction was calculated from EMSA intensity ratio between the unbound DNA fraction in the absence and presence of MR (data not shown).

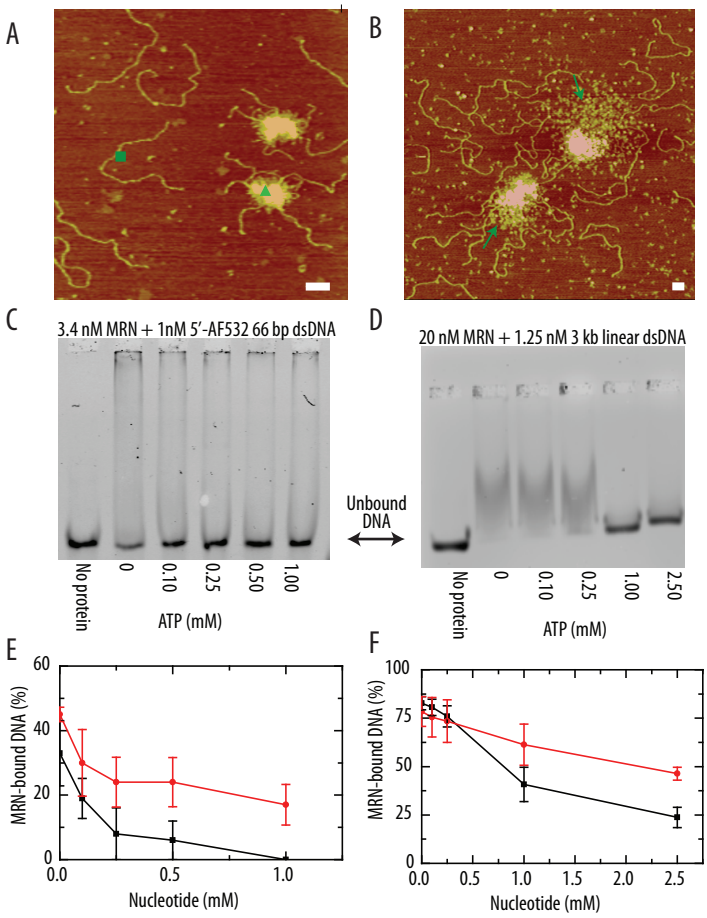


Figure 6.4: MRN release from DNA in the presence of ATP or AMP-PNP. (**A**) AFM images acquired in air of MRN complexes and linear 3kb dsDNA fragments. 2.5 nM MRN was incubated with 0.2 nM DNA for 1 minute at 25°C in the absence of nucleotide cofactors and subsequently deposited on mica and imaged. Both unbound single DNA molecules were visible (green square) as well as large DNA protein DNA complexes that tethered multiple DNA molecules (green triangle). The scale bar is 100 nm, color represents height form 0 to 3 nm (brown to white). (**B**) AFM images of MRN complexes and linear 3kb dsDNA fragments. 2.5 nM MRN was incubated with 0.2 nM DNA for 1 minute at 25°C and subsequently 1 mM AMP-PNP was added and incubated for 1 more minute at 25°C. The addition of AMP-PNP resulted in dissociation of MRN complexes from the large aggregates forming a cloud of single MRN complexes around the larger DNA bound aggregates (green arrows). The scale bar is 100 nm, color represents height form 0 to 3 nm (brown to white). (**C**) EMSA showing ATP dependent release of MRN from 66 bp dsDNA. (**D**) EMSA showing ATP dependent release of MRN from 3 kb dsDNA. (**E**) Quantification of MRN release from 66bp dsDNA for increasing concentra-

tion of ATP (red) and AMP-PNP (black). At 1 mM AMP-PNP all DNA is released from MRN. The MR-bound DNA fraction was calculated from the intensity ratio between the unbound DNA fraction in the absence and presence of MRN. (**F**) Quantification of MRN release from 3 kb dsDNA for increasing concentrations of ATP (red) and AMP-PNP (black). Release from longer substrates is less efficient than from the short 66 bp oligonucleotides.

was induced by DNA release, we performed EMSAs (Figure 6.4C, E) using a constant concentration of 1 nM 66bp dsDNA and 3.4 nM MRN, a concentration where approximately 50 % of the DNA is bound. Varying concentrations (0–1 mM) of ATP and AMP-PNP were added to test if the addition of nucleotides would trigger release of bound DNA. This is indeed observed, see Figure 6.4C, E. AMP-PNP was more efficient at triggering DNA release and full release was observed for 1 mM AMP-PNP, as seen by the recovery of free DNA (Figure 6.4E). The effect of nucleotide cofactors was also tested for longer 3 kb linear dsDNA molecules, which again resulted in DNA release, but to a lesser degree compared to the 66 bp substrates (Figure 6.4D, F).

DNA release may be prevented or reduced due to an increased stability of oligomeric complexes of DNA bound MRN. To address this question experiments were performed for different MRN concentrations (0–16 nM) at constant nucleotide and DNA concentrations (0.5 mM AMP-PNP, 1 nM 66 nt dsDNA). Indeed here we only observed DNA release for concentrations below 8 nM MRN indicating that MRN-MRN interactions may be involved (Figure 6.5A, B). To determine if oligomers were present, we imaged reaction mixtures with 4 nM and 16 nM MRN by AFM to visualize the protein DNA complexes (data not shown). At 4 nM, MRN was present in dimeric and low multimeric complexes while at 16 nM the protein was mostly in larger oligomeric complexes.

To determine if the dissociation from the aggregates was induced by DNA release, we performed AFM measurements using 96 bp dsDNA molecules. DNA release from MRN induces a conformational change of the RAD50 coiled coils, which change from a parallel, DNA-bound, conformation to and open conformation (Figure 6.6A, B)[15]. AFM imaging of MR, MRN, and RN complexes with and without DNA showed that this behavior was consistent for all complex compositions (Figure 6.6C, D). In the absence of DNA the complexes were in a mostly open conformation (>80 %), whereas after incubation with DNA the complexes were in a mostly (>80 %) closed conformation.

To directly visualize the effect of nucleotide cofactors on the MR(N) conformation in real time we performed in situ imaging with buffer exchange to change nucleotide conditions. Previous AFM experiments detected no influence of ATP on the conformation of the RAD50 coiled coils[15], which is surprising considering the large rotation of the MRE11 and RAD50 coiled-coil domains observed in X-ray crystallography structural studies of isolated domains[27-30]. If nucleotide cofactors induce a release of DNA, a change of the MRN complex to an open conformation is expected. Indeed, we observed an increase in the number of open complexes, indicating DNA release, upon changing to a buffer containing ATP (Figure 6.7). We did not observed this change in conformation when performing the buffer exchange with one containing

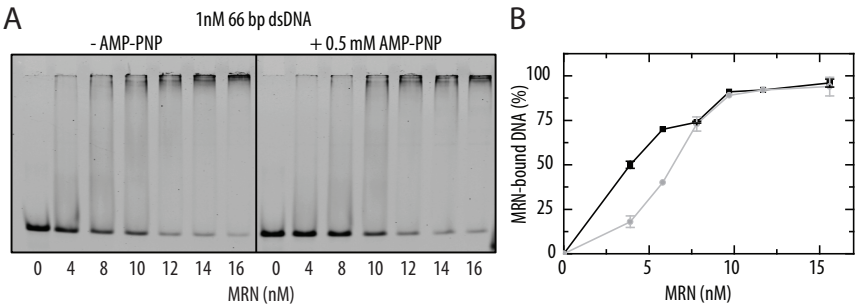


Figure 6.5: MRN releases DNA in the presence of 0.5 mM AMP, but only at low MRN concentrations. (**A**) EMSAs showing DNA binding and release for different MRN concentrations, addition of 0.5 mM AMP-PNP results in a release of bound DNA at MRN concentrations below 8 nM. (**B**) Quantification of MRN bound DNA without AMP-PNP (black) and with 0.5 mM AMP-PNP (gray).

no ATP. Buffer exchange reduces the DNA concentration in solution by a factor of 10, making DNA rebinding after release unlikely in these experiments. Our results are thus consistent with a model where nucleotide binding acts as a switch for the MRN complex to release DNA and our data show that this DNA release resulted in a conformational change of the MRN complex. The presence of oligomeric complexes, however, stabilizes MRN complexes on DNA, in the presence of nucleotide cofactors that would normally result in DNA release.

Discussion

MRN act as a multifunctional molecular machine but the molecular mechanisms by which its parts communicate to orchestrate its multiple actions are poorly understood. DNA binding is central to its functions of signaling and processing at DSBs and tethering of DNA ends. We report here that MR(N) showed no preference for binding to dsDNA ends compared to internal positions on dsDNA in agreement with previous work[12,19,20,35]. It thus remains an open question how MRN is able to sense and clear protein bound DNA ends *in vivo*[36,37]. MRN did however show a preference for ssDNA and ssDNA overhangs, which are an intermediate substrate, present at most, if not all, DSBs, and which require processing by MRN to complete repair.

AFM images showed that large protein clusters are formed by MRN that likely stabilize DNA binding (Figure 6.4). Such clusters may be essential to the functions of MRN in recruiting other proteins to the DSB and tethering DNA ends. If MRN would only bind efficiently to DNA ends then only a single complex could be bound at the DSB end site providing just a few interaction sites for partners. MRN clusters, on the other hand, are able to provide a large scaffolding area and stable tethering via many

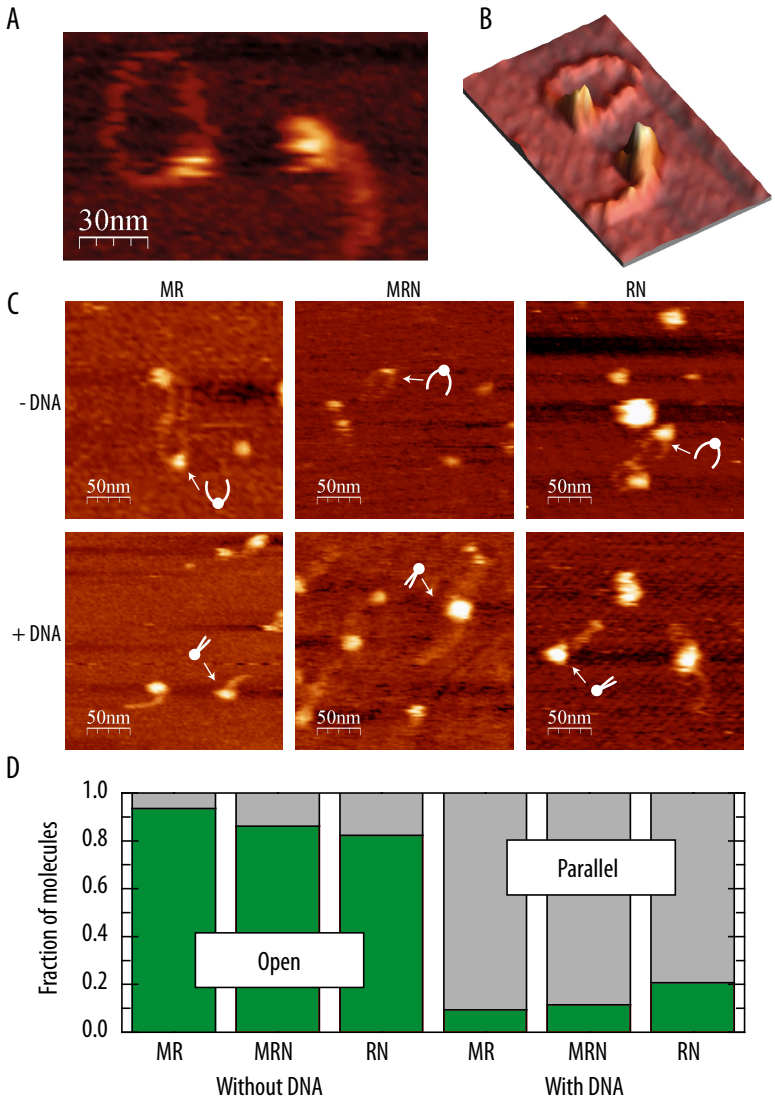


Figure 6.6: RAD50 coiled coils change to a parallel conformation upon DNA binding. (**A** and **B**) MR coiled coils were either in an open conformation, where the two coils leave the DNA binding head under an angle, or in a parallel conformation where the two RAD50 coiled coils align parallel. (**C**) AFM images of MR, MRN and RN complexes acquired in liquid. 3.5 nM protein complexes where incubated with or without 35 nM 96 bp dsDNA for 10 minutes at 25°C and deposited on mica. All complex compositions showed predominantly an open coiled-coil conformation (indicated by the white cartoons) in the absence of DNA (top row). In the presence of DNA the coiled coils adopted a parallel conformation (bot-

tom row). Height scale 0–3 nm (brown to white) for all images **D**, Fraction of molecules in either a parallel or open conformation determined from AFM images in liquid (n > 60 for each condition).

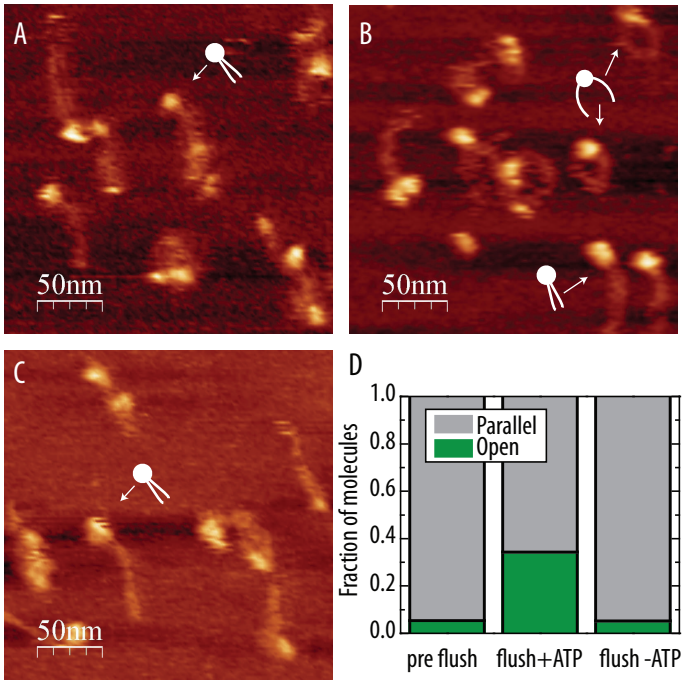


Figure 6.7: MR complexes release DNA in the presence of 1 mM ATP resulting in a conformational change of the RAD50 coiled coils. (**A**) MR complexes, imaged in liquid after incubation with 96 bp dsDNA, were predominantly in a parallel coiled-coil conformation, indicated by the white cartoon. (**B**) MR complexes adopted an open coiled-coil conformation after changing the buffer to one containing 1 mM ATP. (**C**) MR complexes maintained an parallel coiled-coil conformation after changing to a buffer containing no ATP. (**D**) Fraction of MR complexes in parallel and open coiled-coil conformation before and after exchanging the buffer for one containing ATP or no ATP. The fraction of complexes in a parallel coiled-coil conformation decreased after changing to a buffer containing 1 mM ATP, indicating DNA release (n > 200 for each condition). The fraction of molecules in a parallel coiled-coil conformation stayed constant after changing to a buffer containing no ATP indicating that DNA release and its resulting conformational change of the coiled coils were induced by ATP binding.

of the RAD50 coiled coils. AFM images indeed confirm such a tethering mechanism where multiple DNA molecules are bound to MRN clusters (Figure 6.4).

In the presence of ATP or AMP-PNP we observed the release of MRN from DNA both in AFM and EMSAs (Figure 6.4–6.7). As in previous work, we did not see an effect of nucleotide cofactor on the RAD50 coiled coil conformation[15]. Human MRN complexes are found not to undergo conformational change to a parallel coiled-coil configuration when bound to a nucleotide cofactor, but only adopt a parallel configuration when bound to DNA. By direct AFM imaging of MRN complexes in liquid, we now show that the RAD50 coiled coils undergo a conformational change upon exchange to a buffer containing ATP indicating DNA release (Figure 6.7). Nucleotide binding is thus able to change the affinity and thereby acts as a switch to release DNA. This experimental observation appear to contrast recent structural studies, of the Thermotoga maritime protein, that describe a large rotation in the Mre11-Rad50 globular domain upon nucleotide binding[28] inducing a more compact structure which could act as a clamp around the DNA, thereby stabilizing MRN DNA interaction. Small-angle X-ray scattering and Forster resonance energy transfer data of MR complexes in solution however show that, although the MR head domains become more compact upon nucleotide binding, a sizable fraction remains in an open conformation[28,31]. The exact nature of the 'clamp' conformations therefore remains unclear as both higher[28] and lower[31] DNA affinities have been reported for the nucleotide bound state. Additionally Lammens *et al.*[28] were unable to crosslink the clamped protein around DNA as would be expected 'clamp' geometry. Lim *et al.*[30] showed that ATP binding closes the MR complex, thereby preventing DNA binding to the Mre11 nuclease domains. Differences in DNA binding by MR(N) may therefore be related to the accessibility of the Mre11 nuclease domains and not result from a clamp like conformation around the DNA.

From a mechanistic point of view, MRN tethering, mediated by a parallel RAD50 coiled coil conformation, should be linked to DNA binding and not to ATP binding, as tethering is only functional when MR(N) is bound to DNA. The modest ATPase activity of MR(N) below 1 ATP min^{-1} $complex^{-1}$ suggest a regulatory role, where ATP binding and hydrolysis act as a conformational switch to regulate MR(N) endo- and exonuclease activity and provide a turnover mechanism for release of MR(N) from DNA. Taken together, our results show that the DNA binding by MRN has a direct functional link to its coiled coil conformation. ATP binding does not influence MRN coiled coil conformation but acts as an allosteric switch for MRN to an ATP-bound state that has a lower affinity for DNA.

References

1 Mills, K. D., Ferguson, D. O. & Alt, F. W. The role of DNA breaks in genomic instability and tumorigenesis. *Immunol. Rev.* **194**, 77-95, (2003).

2 Kanaar, R., Wyman, C. & Rothstein, R. Quality control of DNA break metabolism: in the 'end', it's a good thing. *EMBO J.* **27**, 581-588, (2008).

3 Hoeijmakers, J. H. J. Genome maintenance mechanisms for preventing cancer. *Nature* **411**, 366-374, (2001).

4 Harper, J. W. & Elledge, S. J. The DNA damage response: Ten years after. *Mol. Cell* **28**, 739-745, (2007).

5 Mladenov, E. & Iliakis, G. Induction and repair of DNA double strand breaks: The increasing spectrum of non-homologous end joining pathways. *Mutat. Res.-Fundam. Mol. Mech. Mutag.* **711**, 61-72, (2011).

6 Borde, V. The multiple roles of the Mre11 complex for meiotic recombination. *Chromosome Res.* **15**, 551-563, (2007).

7 Rass, E. *et al.* Role of Mre11 in chromosomal nonhomologous end joining in mammalian cells. *Nat. Struct. Mol. Biol.* **16**, 819-U838, (2009).

8 Williams, R. S., Williams, J. S. & Tainer, J. A. Mre11-Rad50-Nbs1 is a keystone complex connecting DNA repair machinery, double-strand break signaling, and the chromatin template. *Biochemistry and Cell Biology-Biochimie Et Biologie Cellulaire* **85**, 509-520, (2007).

9 Kinoshita, E., van der Linden, E., Sanchez, H. & Wyman, C. RAD50, an SMC family member with multiple roles in DNA break repair: how does ATP affect function? *Chromosome Res.* **17**, 277-288, (2009).

10 Filippo, J. S., Sung, P. & Klein, H. Mechanism of eukaryotic homologous recombination. *Annu. Rev. Biochem.* **77**, 229-257, (2008).

11 Hopkins, B. B. & Paull, T. T. The P-furiosus Mre11/Rad50 Complex Promotes 5 ' Strand Resection at a DNA Double-Strand Break. *Cell* **135**, 250-260, (2008).

12 Paull, T. T. & Gellert, M. The 3 ' to 5 ' exonuclease activity of Mre11 facilitates repair of DNA double-strand breaks. *Mol. Cell* **1**, 969-979, (1998).

13 Lee, J. H. & Paull, T. T. ATM activation by DNA double-strand breaks through the Mre11-Rad50-Nbs1 complex. *Science* **308**, 551-554, (2005).

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

15 Moreno-Herrero, F. *et al.* Mesoscale conformational changes in the DNA-repair complex Rad50/Mre11/Nbs1 upon binding DNA. *Nature* **437**, 440-443, (2005).

16 van der Linden, E., Sanchez, H., Kinoshita, E., Kanaar, R. & Wyman, C. RAD50 and NBS1 form a stable complex functional in DNA binding and tethering. *Nucleic Acids Res.* **37**, 1580-1588, (2009).

17 Hopfner, K. P. *et al.* The Rad50 zinc-hook is a structure joining Mre11 complexes in DNA recombination and repair. *Nature* **418**, 562-566, (2002).

18 van Noort, J. *et al.* Inhomogeneous flexibility of an intra-molecular coiled-coil structure in Rad50 measured with AFM. *Biophys. J.* **84**, 158a-158a, (2003).

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

20 Trujillo, K. M. & Sung, P. DNA structure-specific nuclease activities in the Saccharomyces cerevisiae Rad50 center dot Mre11 complex. *J. Biol. Chem.* **276**, 35458-35464, (2001).

21 Stracker, T. H., Morales, M., Couto, S. S., Hussein, H. & Petrini, J. H. J. The carboxy terminus of NBS1 is required for induction of apoptosis by the MRE11 complex. *Nature* **447**, 218-U217, (2007).

22 Strunnikov, A. V. & Jessberger, R. Structural maintenance of chromosomes (SMC) proteins - Conserved molecular properties for multiple biological functions. *Eur. J. Biochem.* **263**, 6-13, (1999).

23 Hopfner, K. P. *et al.* Structural biology of Rad50 ATPase: ATP-driven conformational control in DNA double-strand break repair and the ABC-ATPase superfamily. *Cell* **101**, 789-800, (2000).

24 Rees, D. C., Johnson, E. & Lewinson, O. ABC transporters: the power to change. *Nat. Rev. Mol. Cell Biol.* **10**, 218-227, (2009).

25 Cuylen, S., Metz, J. & Haering, C. H. Condensin structures chromosomal DNA through topological links. *Nat. Struct. Mol. Biol.* **18**, 894-U852, (2011).

26 Nasmyth, K. & Haering, C. H. Cohesin: Its Roles and Mechanisms. *Annu. Rev. Genet.* **43**, 525-558, (2009).

27 Williams, G. J. *et al.* ABC ATPase signature helices in Rad50 link nucleotide state to Mre11 interface for DNA repair (vol 18, pg 423, 2011). *Nat. Struct. Mol. Biol.* **18**, 1084-1084, (2011).

28 Lammens, K. *et al.* The Mre11:Rad50 Structure Shows an ATP-Dependent Molecular Clamp in DNA Double-Strand Break Repair. *Cell* **145**, 54-66, (2011).

29 Schiller, C. B. *et al.* Structure of Mre11–Nbs1 complex yields insights into ataxia-telangiectasia–like disease mutations and DNA damage signaling. *Nat. Struct. Mol. Biol.* **19**, 693-700, (2012).

30 Lim, H. S., Kim, J. S., Park, Y. B., Gwon, G. H. & Cho, Y. Crystal structure of the Mre11-Rad50-ATP gamma S complex: understanding the interplay between Mre11 and Rad50. *Genes Dev.* **25**, 1091-1104, (2011).

31 Majka, J., Alford, B., Ausio, J., Finn, R. M. & McMurray, C. T. ATP Hydrolysis by RAD50 Protein Switches MRE11 Enzyme from Endonuclease to Exonuclease. *J. Biol. Chem.* **287**, 2328-2341, (2012).

32 Lee, J. H. *et al.* Regulation of Mre11/Rad50 by Nbs1 - Effects on nucleotide-dependent DNA binding and association with ataxia-telangiectasia-like disorder mutant complexes. *J. Biol. Chem.* **278**, 45171-45181, (2003).

33 Moncalian, G. *et al.* The Rad50 signature motif: Essential to ATP binding and biological function. *J. Mol. Biol.* **335**, 937-951, (2004).

34 de Jager, M., Wyman, C., van Gent, D. C. & Kanaar, R. DNA end-binding specificity of human Rad50/Mre11 is influenced by ATP. *Nucleic Acids Res.* **30**, 4425-4431, (2002).

35 Trujillo, K. M. *et al.* Yeast Xrs2 binds DNA and helps target Rad50 and Mre11 to DNA ends. *J. Biol. Chem.* **278**, 48957-48964, (2003).

36 Lobachev, K. S., Gordenin, D. A. & Resnick, M. A. The Mre11 complex is required for repair of hairpin-capped double-strand breaks and prevention of chromosome rearrangements. *Cell* **108**, 183-193, (2002).

37 Neale, M. J., Pan, J. & Keeney, S. Endonucleolytic processing of covalent protein-linked DNA double-strand breaks. *Nature* **436**, 1053-1057, (2005).

CHAPTER 7

Dynamics of RecA filaments on single-stranded DNA

RecA, the key protein in homologous recombination, performs its actions as a helical filament on single-stranded DNA (ssDNA). ATP hydrolysis makes the RecA-ssDNA filament dynamic and is essential for successful recombination. RecA has been studied extensively by single-molecule techniques on dsDNA. Here we directly probe the structure and kinetics of RecA interaction with its biologically most relevant substrate, long ssDNA molecules. We find that RecA ATPase activity is required for the formation of long continuous filaments on ssDNA. These filaments are found to both nucleate and extend with a multimeric unit. Disassembly rates of RecA from ssDNA decrease with applied stretching force, corresponding to a mechanism where protein-induced stretching of the ssDNA aids in the disassembly. Finally, we show that RecA-ssDNA filaments can reversibly interconvert between an extended, ATP-bound, and a compressed, ADP-bound state. Taken together our results demonstrate that ATP hydrolysis has a major influence on the structure and state of RecA filaments on ssDNA.

This chapter is adapted from: Nucleic Acids Research: Marijn T. J. van Loenhout, Thijn van der Heijden, Roland Kanaar, Claire Wyman, Cees Dekker. Dynamics of RecA filaments on single-stranded DNA. Nucleic Acids Res., 37, 4089-4099, (2009).

Introduction

RecA-single-stranded DNA (ssDNA) filaments are the catalytic core of homologous recombination (HR) in *E. coli* and RecA has become the prototypical member of a family of proteins ubiquitous to all organisms including the human RAD51 and DMC1 recombinases[1-3]. The RecA-ssDNA filament must perform multiple actions during recombinational DNA repair: it first has to assemble on ssDNA and search for homologous duplex DNA, subsequently the RecA-ssDNA filament invades and exchanges strands with the homologous duplex DNA, and finally it dissociates to allow for further processing by polymerases. The repair of DNA double-stranded breaks by HR is vital to the cells' survival and therefore all these steps have to be robust enough to deal with heterology or other obstacles that may hinder the process. ATP hydrolysis by RecA is essential for successful HR but many questions on how ATP hydrolysis influences the molecular structure and dynamics of RecA-ssDNA filaments remain.

The assembly of RecA filaments on ssDNA is an essential early step in HR and requires both a nucleotide cofactor (i.e. ATP) and a divalent cation (i.e. Mg^{2+}) bound at the monomer-monomer interfaces[4,5]. Filament assembly has been studied in detail on double-stranded DNA (dsDNA) where the consensus is that filament formation is dominated by a slow nucleation step, requiring a nucleation cluster of 5 or 6 monomers, followed by rapid filament extension resulting in long continuous filaments[6-9]. On the biologically more relevant ssDNA substrate, nucleation also requires a multimeric unit and is orders of magnitude faster[9,10]. Filament extension on ssDNA is less cooperative, possibly resulting in the formation of short filament patches as observed for RAD5111-13. The 3' end is typically favored for addition of RecA monomers and disassembly occurs from the 5' end resulting in a net 5' to 3' assembly direction on ssDNA[9,14]. Due to the 5' to 3' directionality, the 3' end is more likely to be covered with RecA resulting in more efficient pairing reactions at the 3' end *in vitro*[15,16].

Disassembly of the RecA filament also is an important step in the overall recombination process. The nucleoprotein filament needs to be dismantled after strand exchange to clear the path for DNA synthesis and the completion of HR. Recent experiments suggest that strand exchange and filament disassembly are coupled, underlining the importance of disassembly kinetics[17]. Disassembly of RecA is coupled to ATP hydrolysis at the monomer-monomer interface and proceeds from filament ends on both ss- and dsDNA. Since the monomer-monomer interfaces are presumably identical throughout the filament, the asymmetry in assembly and disassembly rates can only be sustained at the expense of external energy[18]. The recently resolved structure of RecA complexed with ssDNA indicates that ATP binding and ATP hydrolysis indeed mediate the binding and release of RecA from DNA through allosteric coupling[5]. ATP hydrolysis is, however, not the only driving factor for dissociation, as the mechanical interactions of a recombinase filament with its stretched DNA substrates also facilitate disassembly[19]. Although pairing reactions proceed readily when RecA disassembly is blocked by non- or slowly hydrolysable ATP analogs, e.g. ATPγs, the delicate balance

between assembly and disassembly of the RecA-ssDNA filament is likely important for strand exchange and homology search[20,21].

RecA forms a right-handed helical filament on ssDNA in the presence of ATP, extending it by 50 % compared to B-form dsDNA. DNA is bound within the filament with a stoichiometry of three nucleotides (nt) or base pairs (bp) per RecA monomer and 6 monomers make up one helical turn[5,22-24]. Single-stranded DNA is thus considerably extended and restricted in its conformational freedom within the RecA filament. ATP hydrolysis occurs throughout the filament and is not restricted to the disassembly end[25]. The hydrolytic cycles of individual monomers are uncoupled on ssDNA in contrast to RecA bound to dsDNA where hydrolytic cycles of RecA monomers cooperatively influence each other[26-28]. The structure of the ssDNA-RecA filament is also affected by the bound nucleotide cofactor and in the presence of ADP a compressed filament forms[29].

ATP hydrolysis not only drives the turnover of RecA on DNA, but it is also linked to several other functions: the bypass of long stretches of heterology, fork regression, and four-strand-exchange reactions[30]. All these processes seem to require the generation of torsion in the pairing strands and several models to generate this torsion have been proposed. The simplest models link the generation of torsion to a redistribution of RecA monomers on DNA[31,32]. The accompanying change in helical pitch of the dsDNA, between its canonical B-form of 10.5 base pairs per turn to 18.5 base pairs per turn when bound by RecA, could generate the required torsion. A second model postulates the presence of concerted motor action of RecA monomers to actively rotate the homologous strand around the filament[1]. A third model links the generation of torsion to a conformational change of the RecA filament upon ATP hydrolysis[33].

To describe RecA filament dynamics and the mechanical changes of the filament coupled to ATP hydrolysis we measured RecA filament assembly and disassembly on single ssDNA molecules. We find that ATP hydrolysis by RecA was essential for the formation of long uninterrupted filaments on ssDNA. By fitting filament assembly data to Monte Carlo (MC) simulations, we conclude that both nucleation and extension of these filaments occurred with multimers. RecA disassembly is facilitated by the energy stored in the bound and stretched ssDNA. Dissociation from ssDNA is more sensitive to force than dissociation of recombinases from dsDNA. Interestingly, we show that the mechanical properties of RecA-ssDNA filaments change drastically upon ATP hydrolysis and that this process is fast and reversible, implying a rapid exchange of ADP from the filament.

Materials and methods

DNA construct

A ssDNA construct was prepared as described previously[13], additionally a 7.3 kb construct was prepared analogously starting from a λ-DNA template with primer sequences AACTCAGCTCACCGTCGAACA and 5' biotinylated, GACGCAGGGGACCTGCAG and digested with PspOMI.

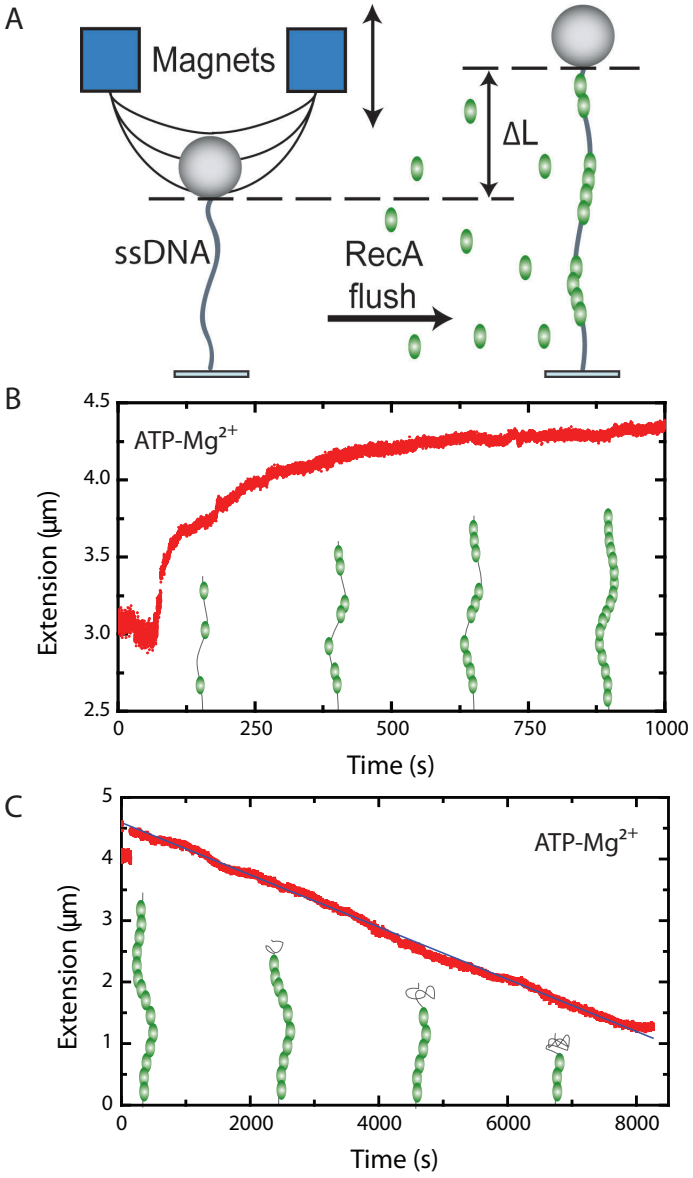


Figure 7.1: RecA filament formation and dissociation on ssDNA as measured by magnetic tweezers. (**A**) Schematic drawing of the magnetic tweezers setup. (**B**) A 8.6 kb ssDNA molecule is extended by 48% compared to B-form DNA at a stretching force of 6 pN in the presence of 1 µM RecA, 1 mM ATP and 5 mM Mg^{2+}. (**C**) Linear disassembly profile of a 8.6 kb RecA-ssDNA filament recorded at 0.5 pN.

Magnetic tweezers assay

A magnetic tweezers setup was used in these experiments as described[34]. By using image processing, 5-nm position accuracy of the bead was obtained in all three dimensions[35]. To exclude the effect of thermal drift, all positions were measured relative to a 3.2 μm polystyrene bead (Bang Laboratories, Carmel, IN) fixed to the bottom of the flow cell. Polystyrene beads as well as DNA constructs carrying a magnetic bead at one end were anchored to the bottom of a flow cell as described elsewhere[35]. The force-extension curve of single DNA molecules was measured. After conformation of the correct contour and persistence lengths, experiments were started by addition of RecA. All measurements were carried out at 25 °C.

RecA/DNA reactions

The flow-cell final volume was approximately 100 μl. All reactions were done in 25 mM Tris-HCl (pH 7.5), 5 mM $MgCl_2$ or $CaCl_2$, and 1 mM DTT. RecA was purchased from New England Biolabs. RecA and ATP (final concentrations 1 μM RecA, unless stated otherwise, and 1 mM ATP) were added into the flow cell. Interaction of RecA with the tethered DNA molecule was monitored through measurement of the height of the magnetic bead.

Monte Carlo simulations

We follow the protocol described in Van der Heijden *et al.*[13,36].

Results

ATP hydrolysis is essential to form long continuous RecA-ssDNA filaments

Our magnetic-tweezers single-molecule technique allows monitoring long (~8 kb) individual DNA molecules. The DNA is bound at one end to a magnetic bead and at the other end to the surface of a flow cell. A force is applied by a pair of external magnets (Figure 7.1A). Video microscopy is used to monitor the bead position, and thus the DNA end-to-end distance, in real time. The assembly and disassembly of RecA filaments is monitored by changes in the end-to-end distance of the DNA tether. RecA binding increases tether length as it extends and stiffens ssDNA. Length changes of the tethered DNA can thus be correlated to changes in RecA filament coverage of the ssDNA molecule. Filament assembly was recorded at forces around 6 pN to minimize the influence of secondary structure in the ssDNA on the assembly profile. At 6 pN, the tether length of a ssDNA molecule is approximately equal to the length of a B-form dsDNA molecule with the same number of nucleotides. When a complete RecA filament forms, the tether is expected to extend to 1.5 times the B-form length. Filament disassembly was monitored at forces below 0.5 pN, unless stated otherwise. Below 0.5 pN, the end-to-end distance of ssDNA is negligible compared to the stiff RecA filament. A decrease in the end-to-end distance can therefore be directly correlated to dissociation of the filament.

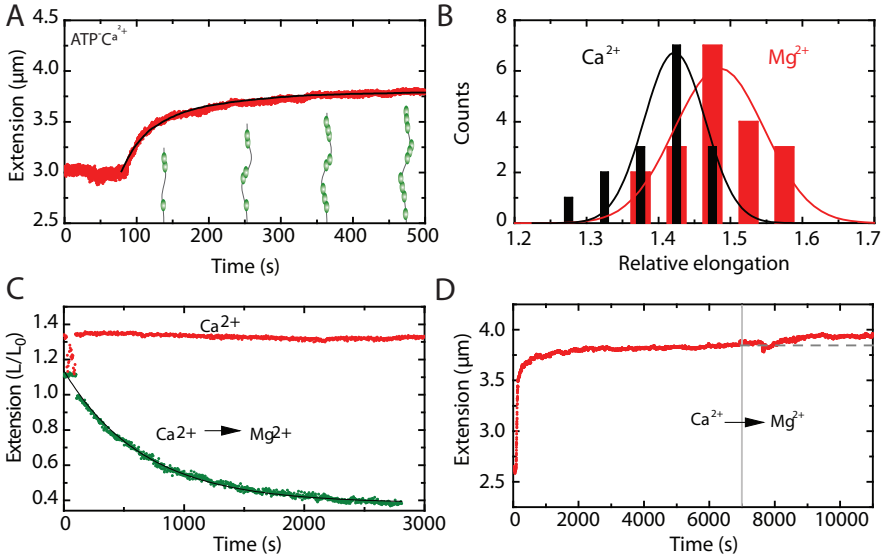


Figure 7.2: RecA-ssDNA filaments in conditions suppressing ATP hydrolysis. (**A**) Filaments formed on a 8.6 kb ssDNA molecule showed a similar binding behavior at a stretching force of 6 pN in a 5 mM Ca^{2+} buffer to that in the presence of Mg^{2+}. The black line is a fit from MC simulations, adopting hexamers for nucleation and extension units, used to determine the nucleation and extension rate. (**B**) Histogram of RecA-ssDNA tether length at a stretching force of 6 pN in the presence of either 5 mM Mg^{2+} (red) or 5 mM Ca^{2+} (black). Solid lines are a fit to a normal distribution. (**C**) RecA disassembly at a stretching force of 0.5 pN from a 8.6 kb ssDNA was blocked in the presence of 5 mM Ca^{2+} (red points), but readily occurred when Ca^{2+} was replaced by Mg^{2+} (green points) The filament dissociation fits well to a single exponential decay (black line). (**D**) A 7.3 kb RecA-ssDNA filament assembled in the presence of Ca^{2+} at a stretching force of 6 pN showed a further increase in end-to-end distance when the buffer was exchanged for one containing 1 µM RecA and 5 mM Mg^{2+}. The gray dashed line is a guide to the eye.

Filament assembly was started by introducing 1 µM RecA in the presence of 1mM ATP and 5 mM Mg^{2+} into the flow cell. Immediately after buffer flow was stopped, DNA end-to-end distance started to increase, consistent with the formation of DNA-RecA nucleoprotein filament(s) that extend and stiffen the DNA tether (Figure 7.1B). In contrast to the binding profile seen on dsDNA (Figure S7.1)[7], the profile is strongly non-linear, indicative of a substantially less cooperative assembly. Cooperativity can be defined as the ratio between the rate of filament extension and nucleation. At high ratios (> 106), the binding profile displays a linear increase, whereas at low ratios the increase is sigmoidal (Figure S7.2)[36]. A low cooperativity will result in incomplete DNA coverage because different nucleation positions can be out of register with each other, leading to gaps of at least 1 or 2 nt considering the 3 nt binding site of RecA.

Multimeric binding, where the unit of filament nucleation or extension consist of more than 1 RecA monomer, will further increase the size of these gaps leading to even less complete coverage. Experimentally we found that the final end-to-end distance of the RecA filaments corresponded to a 48.0 ± 1.4% (N = 19) elongation compared to B-form DNA of the same nucleotide length. The measured elongation is in excellent agreement with electron microscopy and other studies of RecA filaments on ssDNA[5,33]. The elongation obtained thus corresponds to a fully covered ssDNA molecule. We note that this is a very surprising result because the shape of the binding profile would suggest multiple nucleation events leading to gaps in the filament as reported for RAD51[13].

Dissociation measurements also indicated the absence of gaps in the final filament as the obtained profiles were linear, indicative of dissociation from a single filament (Figure 7.1C). After assembly, dissociation of RecA from ssDNA was initiated by exchanging the buffer in the flow cell for one containing 1 mM ATP and 5 mM Mg^{2+} but no RecA. This allows continuous ATP hydrolysis and therefore disassembly of the bound RecA filament. The measured disassembly rate of 0.83 ± 0.16 monomers/s (N = 10) compares well to the bulk value of 0.91 monomers per second[37].

To understand the interplay between ATP hydrolysis driven RecA binding and dissociation during assembly, we directly measured filament assembly under conditions where RecA ATP hydrolysis was suppressed. In the presence of 5 mM Ca^{2+}, RecA binds to ssDNA but ATP hydrolysis is strongly attenuated[4]. Dissociation of RecA during filament assembly can thus be inhibited by replacing Mg^{2+} with Ca^{2+}. In the presence of Ca^{2+} and ATP, we observed a non-linear growth profile similar to the data for Mg^{2+} (Figure 7.2A). The end-to-end distance, however, of the RecA-ssDNA filaments did not reach a full 50% elongation compared to B-form DNA, but instead a significantly lower elongation of 38 ± 2% (N = 16) (Figure 7.2B). Based on the ratio of measured to expected elongation, this 38 ± 2% (N = 16) corresponds to a fractional coverage of 79 ± 7% of the DNA. The end-to-end distance of the DNA tether did not decrease upon changing the buffer for one containing 1 mM ATP and 5 mM Ca^{2+} but lacking RecA (Figure 7.2C). This shows that filament disassembly was indeed inhibited, yielding stable filaments when ATP hydrolysis is suppressed. The exchange of Ca^{2+} for Mg^{2+}, by changing to a buffer containing ATP and Mg^{2+}, led however to a decrease in the end-to-end distance of the molecule as the filament dissociated (Figure 7.2C). The disassembly profile was now exponential, indicative of dissociation from the ends of multiple filament patches. The shorter end-to-end distance and presence of gaps in the filament assembled in the presence of Ca^{2+} could also result from secondary structure in the ssDNA. To test if secondary structure influenced these measurements we assembled filaments at a force of 20 pN, which effectively removes all secondary structure in the ssDNA[38]. The dissociation behavior of these filaments was the same as observed for filaments assembled at low force (Figure S7.3). The marked difference between the RecA-ssDNA filaments formed in the presence of Mg^{2+} and Ca^{2+} remained essentially the same (cf. Figures 7.1C and 7.2C). This shows that secondary structure

does not play a major role in preventing the formation of a continuous RecA filament on ssDNA in the presence of Ca^{2+}.

The above results indicate that filament formation by random nucleation of RecA on ssDNA creates multiple filament patches that, however, anneal into one continuous filament via ATP- hydrolysis-dependent rearrangements. This implies that a patchy RecA filament, with bare tracts of ssDNA, formed when ATP hydrolysis is suppressed, is able to anneal into one continuous filament when ATP hydrolysis is again allowed. This was indeed observed: A filament that had reached its saturation extension in ATP-hydrolysis suppressing conditions (Ca^{2+}, ATP, and RecA) showed a further increase in extension after switching to a buffer that allowed for ATP hydrolysis (Mg^{2+}, ATP and RecA) (Figure 7.2D). The final extension of the filaments reached 50 ± 3% (N = 7) corresponding to ssDNA fully covered by RecA.

RecA nucleates and extends filaments with multimers

RecA filament assembly involves a nucleation followed by filament extension. Nucleation has been reported to occur with 5 or 6 monomers whereas a single monomer was suggested as the functional unit for filament extension[9,10]. A monomeric unit for filament extension compared to a multimeric nucleation unit should give rise to a concentration dependence of the cooperativity. This was not supported by our experiments as the shape of the binding profiles is conserved for different concentrations (Figure S7.4), implying that the nucleation and extension unit are the same.

In order to quantitatively understand RecA filament formation, we modeled the nucleoprotein filament formation with MC simulations (for details see ref. 13, 36). The simulations involve four parameters to describe the RecA filament formation: nucleation rate, cooperativity, nucleation unit (monomers or multimers) and extension unit (again either monomers or multimers). We first simulated filament assembly curves to obtain the final coverage of the molecule for different nucleation and extension units at different cooperativity (Figure 7.3). Higher cooperativity resulted in a higher final coverage as fewer gaps were formed. Larger extension units resulted in a lower final coverage as the minimum gap size between filament patches increases. The final occupancy of 79 ± 7% obtained from our measurements in the presence of Ca^{2+} and ATP corresponds to the regime expected for both nucleation and extension with a multimeric unit and is inconsistent with a monomeric extension unit.

Because disassembly is inhibited in the presence of ATP and Ca^{2+}, we neglected disassembly to simulate growth profiles obtained in the presence of Ca^{2+}. Nucleation rates were obtained from experimental growth profiles by fitting MC simulations with two free fit parameters (nucleation rate and cooperativity), adopting hexamers for nucleation and extension units. A fit to an experimental growth profile, based on the MC simulations, is shown in Figure 7.2A. The growth profiles in the presence of ATP and Mg^{2+} (Figure 7.1B) can be modeled using the nucleation rate and cooperativity determined for conditions where ATP hydrolysis was suppressed, and the dissociation rate that was independently obtained from the linear disassembly

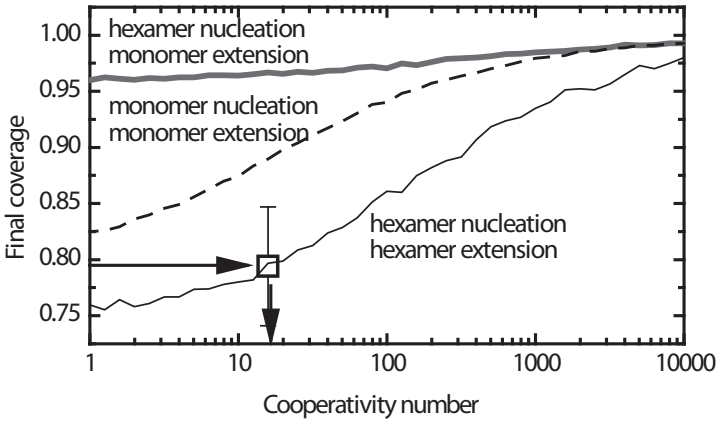


Figure 7.3: Final DNA coverage of a RecA filament is determined by the cooperativity as well as the nucleation and extension unit. Monte Carlo simulations of the saturation coverage of RecA on ssDNA in the absence of dissociation. Our measured final coverage of 79 % ± 7 (black box) in the presence of Ca^{2+} is in the range of nucleation and extension by hexamers.

of complete filaments in the presence of ATP and Mg^{2+}. These simulations always resulted in multiple filament patches and 100% coverage was never obtained. Inclusion of the dissociation reduced the gaps due to multimeric nucleation and extension units to 1 or 2 nt, but it did not completely remove these gaps. Thus accounting for assembly and disassembly alone in our simulations was not sufficient to describe the observed formation of one continuous filament in the presence of ATP and Mg^{2+}. An additional element of the filament dynamics is needed to account for the data, which is examined in the discussion section.

RecA dissociation from ssDNA is force dependent

Exerting a force may change the mechanical interactions between RecA and its DNA substrate. Stretching forces applied to a dsDNA molecule enhance binding and slow down dissociation of nucleoprotein filaments[6,39,40]. Upon dissociation of RecA in the absence of any force exerted on ssDNA, the end-to-end distance will decrease as the ssDNA behaves as an entropic spring. The free energy related to stretching 3 nt of ssDNA to 1.5 times the B-form DNA length, approximately 3 k_BT, could very well contribute to the disassembly process. To test if an applied force influences RecA dissociation from ssDNA, dissociation was initiated as before. Subsequently, the filament was subjected to a constant force in the range of ($16 \geq F \geq 0.1$ pN) during a 270 second interval, while disassembly proceeded. The change in filament length was monitored by periodically adjusting the force to 0.5 pN (Figure 7.4A). RecA-ssDNA filaments subjected to a high force ($F \geq 8$ pN) showed almost no change in end-to-end distance. At lower forces however, the end-to-end distance of the RecA-ssDNA

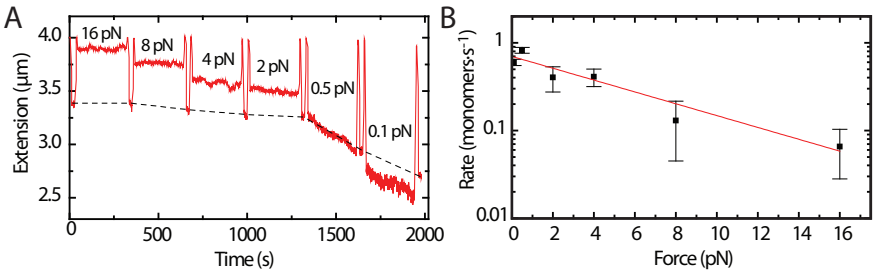


Figure 7.4: Stretching ssDNA strongly affects the disassembly of RecA-ssDNA filaments. (**A**)Dissociation at constant force intervals between 16 pN and 0.1 pN. The RecA-ssDNA filament length was determined by temporarily shifting the force to 0.5 pN after each interval. The black dotted line is a guide to the eye showing the disassembly. (**B**) The RecA dissociation rate is force dependent. Fitting the force dependence with a single exponential $k(F) = k_o \exp(-dF/k_BT)$ (red line) gives a distance to the energy barrier of 0.6 ± 0.2 nm. Data in b are means of about 5 different time traces error bars show s.e.m.

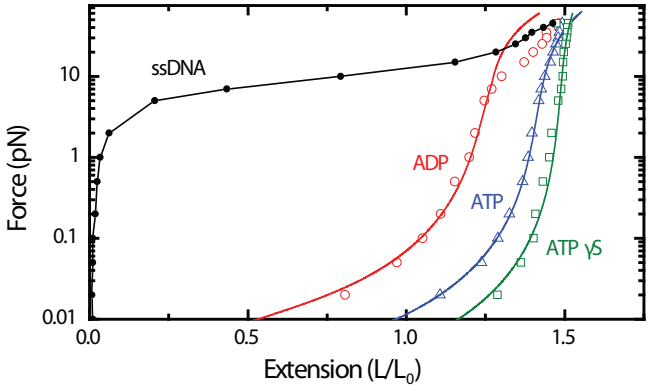


Figure 7.5: RecA-ssDNA filament reversibly interconvert between an extended ATP-bound and a compressed ADP-bound state. Real time measurement of extension changes of a RecA-ssDNA filament at a stretching force of 3 pN in response to different nucleotide cofactors. Flushing with a buffer lacking ATP at $t = 280$ s resulted in a decrease in extension of a RecA-ssDNA molecule (red circles). Reintroduction of ATP (1 mM) resulted in a rapid length increase at $t = 480$ s. The initial extension was however not fully recovered likely due to dissociation of RecA. This procedure was repeated two more times before a buffer containing ATPγS was flushed in at $t = 1230$ s. This resulted in a fast extension surpassing that of the previous ATP-bound filament state. The calculated response based Michaelis-Menten kinetics accurately predicts the observed behavior (black dotted line) (see Figure S7. 5). (b) Force extension behavior of RecA-ssDNA filaments depends on the nucleotide cofactor bound. RecA-ssDNA filaments complexed with ATPγS (green squares) or ATP (blue triangles) are well described by an extensible worm-like chain (solid lines). RecA-ssDNA filaments converted to the ADP-bound state (red circles) are not shorter and more flexible

than filaments in the ATP or ATPγS form. The force extension of bare ssDNA (black circles) is shown for comparison. No fit to the WLC model was made for ssDNA as it behaves qualitatively different.

tether decreased considerably, indicating disassembly of RecA from the DNA. The profiles observed during the different force intervals were all linear confirming that disassembly occurs from one single filament.

A strong decrease in the disassembly rate k is observed with increasing force (Figure 7.4B). Disassembly rates in monomers per second were calculated as the change in tether length at 0.5 pN divided by 1.5 nm, the length of three bases within a single RecA monomer. The disassembly rate had an exponential dependence on force; $k(F)=k_o \exp(-dF/k_BT)$, where k_o is the disassembly rate at zero force, k_BT the thermal energy, F the applied force, and d the distance to the energy barrier along the relevant reaction coordinate Thus, the energy barrier for dissociation is apparently raised by subjecting the RecA-bound DNA to force. The distance to the energy barrier from the force dependency was 0.6 ± 0.2 nm, corresponding to approximately half the length of the three nucleotides held fixed in the RecA-ssDNA filament. This agrees well with RecA dissociation from ssDNA as single monomers.

RecA-ssDNA filaments reversibly switch between extended and collapsed states during ATP hydrolysis

In the RecA-ssDNA filament, a nucleotide cofactor is bound at the important interface between RecA monomers. Changes to this interface are able to influence the structure and conformation of the filament, and modulate the filament's function in the recombination process[11,41,42]. To directly determine the influence of nucleotide cofactors on the mechanical state of the filament, we probed the length of the RecA-ssDNA filament, in real time, while exchanging buffers containing different nucleotide cofactors. First, a RecA-ssDNA filament was assembled in the presence of Mg^{2+} and ATP as described above. Subsequently the buffer was exchanged multiple times under continuous flow to induce exchange of the bound nucleotide cofactor. The force during this continuous flow experiment was approximately 3 pN, calibrated by a separate measurement based on the force-extension behavior of ssDNA.

The RecA-ssDNA filament changed length reversibly in response to the presence of cofactor (Figure 7.5A). The initial RecA-ssDNA filament, formed in the presence of 5 mM Mg^{2+} and 1 mM ATP, had an extension corresponding to a fully covered RecA-ssDNA molecule. Upon exchanging the buffer for one without ATP and RecA, the filament converted to a shorter state. In this state, the bound ATP is expected to be hydrolyzed to ADP, and we will henceforth refer to it as the ADP-bound state. Changing the buffer back to one containing 1 mM ATP but no RecA resulted in extension of the filament. The initial length of the filament was not fully recovered possibly because some disassembly occurred during the conversion to the ADP-bound state. This procedure was then repeated a second and third time on the same molecule showing a similar behavior indicating that the conversion between the ATP- and

ADP-bound states was reversible. Finally, the buffer was exchanged for one containing 1 mM ATPγS, but still lacking RecA. This resulted in an increase in length of the RecA-ssDNA filament that now surpassed that of the ATP-bound state of the previous cycle.

The transition from the elongated ATP-bound state to the shorter ADP-bound state, upon ATP depletion, was much slower than the fast increase in length observed when ATP was flushed in. This difference in response results from the interplay between the high affinity of RecA for ATP and the time-dependend concentration changes in the flow cell during buffer exchange. The K_m of RecA-ssDNA filaments for ATP is in the range of 20-50 μM[43]. Only when the concentration of ATP is reduced to the micromolar range will the RecA-ssDNA filament start to accumulate an ADP-bound state. To quantitatively understand this process we employed a finite-element simulation with a commercial computer program (Comsol Multiphysics) to model the time-dependent concentration in our flow cell (Figure S7.5). This allowed us to predict the time response of the RecA-ssDNA filament under the condition that ATP binding and ADP release are much faster than the changes in ATP concentration. The response was calculated as $\Theta = [ATP] / (K_m + [ATP])$, where Θ is the fraction of the RecA-ssDNA filament in the ATP-bound state. This calculated response, using a Km of 40 μM, can be scaled to an extension of the RecA-ssDNA filament.This matches well with the measured data (black dotted line in Figure 7.5A).

The mechanical properties of the RecA-ssDNA filament in the presence of different nucleotide cofactors were determined by probing its force-extension (Figure 7.5B). The initial RecA-ssDNA filament, formed in the presence of 5 mM Mg^{2+} and 1 mM ATP, had a force-extension curve corresponding to a stiff and extended filament. Upon exchanging the buffer for one without ATP and RecA, the filament converted to a shorter and more flexible ADP-bound state. Finally the RecA-ssDNA filament in the presence of 1 mM ATPγS was stiffer and more extended than the ATP-bound state. The mechanical properties of RecA-ssDNA filaments can be described in terms of persistence length and contour length by a worm-like chain (WLC) model. At higher forces, the elastic behavior of the filaments deviates from the WLC and is better described by an extension to the WLC model incorporating enthalpic stretching[44-46]. The extensible WLC model including a 7th order correction term proposed by Bouchiat *et al.* accurately describes our data for the ATP and ATPγS-bound filaments for forces ranging from 0.02 pN to 40 pN (Figure 7.5B)[46]. The data for the ADP-bound filaments, however, showed an increase in extensibility for forces above 7 pN and deviated substantially from the extensible WLC model. Force-extension data of ADP-bound filaments were therefore fitted only up to 7 pN. At higher forces, the more extensible behavior of the ADP filament approached that of bare ssDNA. The mechanical properties extracted from the extensible WLC model are summarized in Table I. RecA-ssDNA filaments in the ADP-bound state had a contour length that is approximately 6 % shorter than that of the ATP-bound filaments. The ADP-bound filaments were also considerably more flexible and this resulted in even larger differences in end-to-end distance at low forces. Indeed, at 0.01 pN the difference in end-to-end distance between the ADP

and ATP-bound filament was almost 50%. The mechanical differences between the ATP and ADP states of the RecA-ssDNA filament are particularly interesting as these two states continuously interconvert upon ATP hydrolysis in a RecA-ssDNA filament.

The rate of buffer exchange was increased for the force extension measurements to reduce the effects of disassembly of the RecA-ssDNA filament. Under those conditions complete exchange of the flow cell contents required just tens of seconds. The contribution of disassembly in these force-extension measurements is therefore expected to be very small as ADP-bound filaments showed considerable stiffening and an extension surpassing that of the ATP-bound form when a buffer containing C, but no RecA was flushed in. More importantly, subsequent addition of a buffer containing both ATPγS and RecA did not further increase the length of the filaments, indicating that there were no bare ssDNA sites available for RecA assembly (Figure S7.6). This confirmed that the change in filament properties was induced solely by the different nucleotide cofactors bound and not due to RecA disassembly.

Discussion

We have shown that the ATPase activity of RecA in the presence of Mg^{2+} is essential for the formation of a continuous filament over several kilobases on ssDNA. RecA polymerizes with a low cooperativity on ssDNA and substantial rearrangement of RecA monomers needs to take place to remove the gaps created by nucleation at multiple sites along the ssDNA. Modeling filament formation with MC simulations involving assembly and disassembly only cannot account for the annealing of these gaps. The gaps in the filament persist even if the cooperativity, assembly or disassembly rates are changed by orders of magnitude (Figure S7.7).

To explain the formation of continuous filaments, we therefore suggest a model where RecA monomers are able to transfer from one filament end to the neighboring filament in a unidirectional fashion (Figure 7.6C). This fits all our data well for a transfer rate of 2 monomers/s per filament end (Figure S7.7B). Several mechanisms can be envisioned for the transfer of these monomers between the ends of filament patches: monomers could diffuse or hop between the ends. Monomers could also transfer when ends collide and loosely stack end-to-end due to the very flexible ssDNA joints in between[47,48]. Upon dissociation of a RecA monomer the local concentration is very high facilitating rebinding at the neighboring filament patch. A somewhat similar redistribution of proteins was also suggested for single-stranded binding proteins along a ssDNA molecule[49].

Filaments can however also appear to be continuous but actually consist of trains of stacked filament patches as observed in electron microscopy [47,48]. In this case any excess of unbound ssDNA could be looped out at the interface between two stacked patches. A stacking mechanism is however not able to explain the differences in dissociation profile observed between filaments formed in the presence of Mg^{2+} and Ca^{2+}. Disassembly curves of filaments formed in the presence of Ca^{2+} indicate that

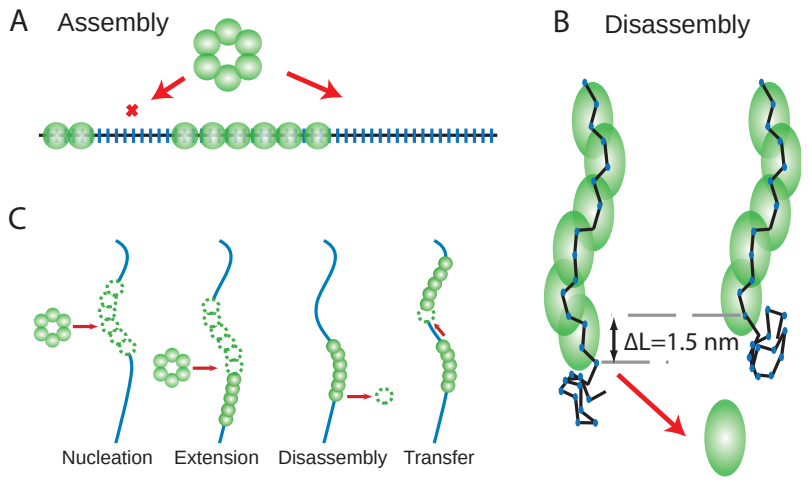


Figure 7.6: RecA assembly and disassembly mechanisms. (**A**) Nucleation as well as filament extension on ssDNA by a multimeric binding unit leaves gaps of considerable size in the filament. (**B**) Dissociation of RecA monomer from ssDNA is aided by the stretching energy stored in the ssDNA. Upon dissociation the total end-to-end distance decreases by approximately 1.5 nm, depending on the stretching force applied (c) Mechanisms considered in the Monte Carlo simulations describing RecA interactions with ssDNA.

blocking ATP hydrolysis results in the formation of multiple filament patches containing stretches of bare ssDNA, inconsistent with a strong filament stacking interaction (Figure 7.2C).

Free RecA can exist in solution predominantly in multimeric form with a preferred stoichiometry between 6 and 12 monomers[50]. The RecA assembly state in solution may have an effect on the preferred binding unit during RecA assembly. Our data indicates that both filament nucleation and filament extension proceeds by RecA multimers, which would leave gaps of up to 17 nucleotides in the filament when ATP hydrolysis is blocked (Figure 7.6A). This is in apparent contrast to a report by Joo *et al.* based on FRET measurements that suggested extension occurs by single monomers[9]. Several factors may explain these differences. Our assay uses ssDNA of several kilonucleotides in length, while the assay employed by Joo *et al.* uses very short DNA molecules (20–60 nts) with the reporter probes spaced only 13 nt apart. The length of the employed oligonucleotides has the potential to limit the ability to identify multimeric binding events or influence the RecA interaction with the ssDNA[43]. To determine dynamics at the 3'-extending end Joo *et al.* exploit their observation that, in their conditions, in the presence of ATPγS RecA-ssDNA filaments appear unstable. Currently, it is unclear how to interpret this observation, given that numerous others assays found that RecA-ssDNA filaments are stable in the presence of ATPγS over extended periods of time[11,51]. Finally, dynamics at the extending end were measured

at a very low RecA concentration of 8 nM, which may influence the stoichiometry of free RecA in solution[50].

We note that our conclusion that both filament nucleation and extension occur by multimers is derived from two independent types of data: Firstly, the final extension of the RecA-ssDNA filament under conditions suppressing disassembly indicates the presence of bare ssDNA gaps larger than the 3 nt RecA binding site. Secondly, the shape of the filament assembly profiles is conserved for different RecA concentrations (Figure S7.4), indicating that the cooperativity is independent of RecA concentration and therefore that the nucleation and extension unit are of equal size.

In a RecA filament, ssDNA is extended and its conformational freedom is severely restricted. The energy stored in this system is released upon dissociation and we determined the distance to the energy barrier to be 0.6 ± 0.2 nm. This distance is approximately half way between the 1.5 nm to which the ssDNA is extended by a RecA monomer and the negligible end-to-end distance for bare ssDNA. The sub-monomer distance to the energy barrier therefore supports a model where RecA dissociates as a single monomer at a time releasing three bases (Figure 7.6B). The 0.6 ± 0.2 nm distance to the energy barrier for dissociation from ssDNA is significantly larger than the 0.27 ± 0.04 nm reported for dissociation of RAD51 from dsDNA, or the 0.26 nm reported for polymerization of RecA on dsDNA[6,40]. On dsDNA, however, three base pairs, with an end-to-end distance of 1.5 nm in the filament, are released upon disassembly of a recombinase corresponding to a length decrease of 0.5 nm, which is in line with a distance to the energy barrier of around 0.27 nm. The differences in length change of the substrate, 1.5 nm for ssDNA or 0.5 nm for dsDNA, thus readily explains the differences in the distance to the energy barrier. We therefore believe that the larger distance reported here is not recombinase specific but substrate specific and expect the energy barrier for other recombinases on ssDNA to be similar to that determined for RecA here. The larger distance to the energy barrier on ssDNA compared with dsDNA gives rise to a stronger dependence of the dissociation rate on the applied force. On ssDNA, dissociation is already strongly reduced at forces around 10 pN, whereas a force of 48 pN was required to stall dissociation for a RAD51 filament bound to dsDNA[40]. These results indicate that the stability of nucleoprotein filaments may be regulated by forces applied to the DNA. During HR any forces existing between the invading strand and the template strand would thus tend to stabilize the RecA filament on the invading ssDNA. The forces required are on the order of a few pN, indeed the range that can be easily generated by polymerases or helicases.

In the RecA-ssDNA filament, a nucleotide cofactor is bound at the structurally important interface between monomers. Here we show that differences in the bound nucleotide cofactor are able to directly and reversibly change the extension and mechanical properties of the RecA-ssDNA filament. This conversion is a direct response of the RecA-ssDNA filament to ATP concentration and is well explained by Michaelis-Menten kinetics (Figure 7.5A).

Direct conversion between extended and compressed filament states reported here and previously[52], is at odds with studies that found a stoichiometry of 5 instead of 3 nucleotides per monomer for the compressed ADP-bound filament[29]. We expect that an active filament bound to nucleotide triplets maintains its stoichiometry upon ATP hydrolysis as adjacent RecA-bound nucleotides are unavailable. The reported ADP-bound filament with a stoichiometry of 5 may however result from de novo filament formation in the absence of ATP. Maintaining a stoichiometry of 3 nucleotides per monomer during changes in filament extension is consistent with the recent structure of RecA-DNA complexes where changes in the extension of one phosphate bond at the interface of two monomers could account for changes in filament length without rearrangement of the bound nucleotide triplet[5]. Force-extension measurements also indicate that interface between monomers in the ADP-bound state is different form the ATP-bound state as the force extension behavior of the filament becomes more extensible above 7 pN and resembles that of ssDNA (Figure 7.5B).

The mechanical parameters determined from force-extension data of ATP-bound RecA-ssDNA filaments are in good agreement with values reported in literature[6]. However, the persistence length that we determined for the ATPγS filament is much larger; 2.1 ± 0.1 μm compared to the 923 ± 46 nm previously reported[6]. We attribute this difference to the different procedure used to assemble the RecA-ssDNA filaments. We converted a fully continuous RecA-ssDNA filament that was preassembled in the presence of ATP-Mg^{2+} to an ATPγS filament by exchange of the nucleotide cofactor. Hegner *et al.* assembled filaments directly in the presence of ATPγS. Our measurements show that blocking of ATP hydrolysis will result in the formation of multiple filament patches. Such a patchy filament will have a lower apparent persistence length than the fully continuous RecA-ssDNA filament measured in our experiments.

Changes in the mechanical state of the filament may aid in the recombination process by modifying the affinity of the RecA-ssDNA filament for its homologous target or by inducing forces between the pairing strands. There are a number of processes unique to RecA; such as the bypass of long heterologous regions during strand exchange and the migration of DNA branches that require the generation of torque[30]. We suggest that ATP-hydrolysis-driven local contractions and extensions of the RecA-DNA filament likely also result in a change in pitch of the filament and consequently could generate the required torque. These conformational changes in the filament are expected to involve multiple RecA monomers as the ATP hydrolysis cycles of neighboring RecA monomers become coupled in contact with dsDNA during strand exchange[27,28,53]. It will be very interesting to see if the dynamics of the RecA-ssDNA filament can be directly linked to a mechanistic picture of nucleoprotein filaments interactions with its partners in the different steps of HR.

Acknowledgments

We thank Peter Veenhuizen, Susanne Hage, and Ya-Hui Chien for technical assistance and Iwijn De Vlaminck for critically reading of the manuscript.

Funding

This work was supported by the Biomolecular Physics program of the Dutch organization for Fundamental Research of Matter (FOM), grants from the Netherlands Organization for Scientific Research (NWO), the Netherlands Genomics Initiative/ NWO, European Commission STREP project BioNano-Switch, the Integrated Projects Molecular Imaging and DNA Repair (512113) and a National Cancer Institute-NIH USA program project.

References

1 Lusetti, S. L. & Cox, M. M. The bacterial RecA protein and the recombinational DNA repair of stalled replication forks. *Annu. Rev. Biochem.* **71**, 71-100, (2002).

2 Symington, L. S. Role of RAD52 epistasis group genes in homologous recombination and double-strand break repair. *Microbiol. Mol. Biol. Rev.* **66**, 630-670, (2002).

3 Cromie, G. A., Connelly, J. C. & Leach, D. R. F. Recombination at double-strand breaks and DNA ends: Conserved mechanisms from phage to humans. *Mol. Cell* **8**, 1163-1174, (2001).

4 Menetski, J. P., Varghese, A. & Kowalczykowski, S. C. Properties of the high-affinity single-stranded-DNA binding state of the Escherichia-coli RecA protein. *Biochemistry* **27**, 1205-1212, (1988).

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

6 Hegner, M., Smith, S. B. & Bustamante, C. Polymerization and mechanical properties of single RecA-DNA filaments. *Proc. Natl. Acad. Sci. U. S. A.* **96**, 10109-10114, (1999).

7 van der Heijden, T. *et al.* Torque-limited RecA polymerization on dsDNA. *Nucleic Acids Res.* **33**, 2099-2105, (2005).

8 Galletto, R., Amitani, I., Baskin, R. J. & Kowalczykowski, S. C. Direct observation of individual RecA filaments assembling on single DNA molecules. *Nature* **443**, 875-878, (2006).

9 Joo, C. *et al.* Real-time observation of RecA filament dynamics with single monomer resolution. *Cell* **126**, 515-527, (2006).

10 De Zutter, S. K. & Knight, K. L. The hRad51 and RecA proteins show significant differences in cooperative binding to single-stranded DNA. *J. Mol. Biol.* **293**, 769-780, (1999).

11 Menetski, J. P. & Kowalczykowski, S. C. Interaction of Reca protein with single-stranded-DNA - quantitative aspects of binding-affinity modulation by nucleotide cofactors. *J. Mol. Biol.* **181**, 281-295, (1985).

12 Modesti, M. *et al.* Fluorescent human RAD51 reveals multiple nucleation sites and filament segments tightly associated along a single DNA molecule. *Structure* **15**, 599-609, (2007).

13 van der Heijden, T. *et al.* Real-time assembly and disassembly of human RAD51 filaments on individual DNA molecules. *Nucleic Acids Res.* **35**, 5646-5657, (2007).

14 Register, J. C. & Griffith, J. The direction of RecA protein assembly onto single-strand DNA is the same as the direction of strand assimilation during strand exchange. *J. Biol. Chem.* **260**, 12308-12312, (1985).

15 Cox, M. M. & Lehman, I. R. Directionality and polarity in RecA protein-promoted branch migration. *Proc. Natl. Acad. Sci. U. S. A.* **78**, 6018-6022, (1981).

16 Dutreix, M., Rao, B. J. & Radding, C. M. The effects on strand exchange of 5' versus 3' ends of single-stranded-DNA in RecA nucleoprotein filaments. *J. Mol. Biol.* **219**, 645-654, (1991).

17 van der Heijden, T. *et al.* Homologous recombination in real time: DNA strand exchange by RecA. *Mol. Cell* **30**, 530-538, (2008).

18 Wegner, A. Head to tail polymerization of actin. *J. Mol. Biol.* **108**, 139-150, (1976).

19 van Mameren, J. *et al.* Counting RAD51 proteins disassembling from nucleoprotein filaments under tension. *Nature* **457**, 745-748, (2009).

20 Rice, K. P., Eggler, A. L., Sung, P. & Cox, M. M. DNA pairing and strand exchange by the Escherichia coli RecA and yeast Rad51 proteins without ATP hydrolysis - On the importance of not getting stuck. *J. Biol. Chem.* **276**, 38570-38581, (2001).

21 Kidane, D. & Graumann, P. L. Dynamic formation of RecA filaments at DNA double strand break repair centers in live cells. *J. Cell Biol.* **170**, 357-366, (2005).

22 Stasiak, A. & Dicapua, E. The helicity of DNA in complexes with RecA protein. *Nature* **299**, 185-186, (1982).

23 Egelman, E. H. & Stasiak, A. Structure of helical RecA-DNA complexes - complexes formed in the presence of ATP-Gamma-S or ATP. *J. Mol. Biol.* **191**, 677-697, (1986).

24 Dicapua, E., Schnarr, M., Ruigrok, R. W. H., Lindner, P. & Timmins, P. A. Complexes of RecA protein in solution - a study by small-angle neutron-scattering. *J. Mol. Biol.* **214**, 557-570, (1990).

25 Brenner, S. L. *et al.* RecA protein-promoted ATP hydrolysis occurs throughout RecA nucleoprotein filaments. *J. Biol. Chem.* **262**, 4011-4016, (1987).

26 Yu, X. & Egelman, E. H. Direct visualization of dynamics and co-operative conformational changes within RecA filaments that appear to be associated with the hydrolysis of adenosine 5'-O-(3-thiotriphosphate). *J. Mol. Biol.* **225**, 193-216, (1992).

27 Cox, J. M., Tsodikov, O. V. & Cox, M. M. Organized unidirectional waves of ATP hydrolysis within a RecA filament. *PLoS Biol.* **3**, 231-243, (2005).

28 Shan, Q. & Cox, M. M. RecA protein dynamics in the interior of RecA nucleoprotein filaments. *J. Mol. Biol.* **257**, 756-774, (1996).

29 Yu, X. & Egelman, E. H. Structural data suggest that the active and inactive forms of the RecA filament are not simply interconvertible. *J. Mol. Biol.* **227**, 334-346, (1992).

30 Cox, M. M. Motoring along with the bacterial RecA protein. *Nat. Rev. Mol. Cell Biol.* **8**, 127-138, (2007).

31 Howard-Flanders, P., West, S. C. & Stasiak, A. Role of RecA protein spiral filaments in genetic-recombination. *Nature* **309**, 215-220, (1984).

32 Kowalczykowski, S. C. & Krupp, R. A. DNA-strand exchange promoted by RecA protein in the absence of ATP - Implications for the mechanism of energy transduction in protein-promoted nucleic-acid transactions. *Proc. Natl. Acad. Sci. U. S. A.* **92**, 3478-3482, (1995).

33 Dunn, K., Chrysogelos, S. & Griffith, J. Electron-Microscopic visualization of RecA-DNA filaments - Evidence for a cyclic extension of duplex DNA. *Cell* **28**, 757-765, (1982).

34 Strick, T. R., Allemand, J. F., Bensimon, D. & Croquette, V. Behavior of supercoiled DNA. *Biophys. J.* **74**, 2016-2028, (1998).

35 van Noort, J., Verbrugge, S., Goosen, N., Dekker, C. & Dame, R. T. Dual architectural roles of HU: Formation of flexible hinges and rigid filaments. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 6969-6974, (2004).

36 van der Heijden, T. & Dekker, C. Monte Carlo simulations of protein assembly, disassembly, and linear motion on DNA. *Biophys. J.* **95**, 4560-4569, (2008).

37 Shan, Q., Bork, J. M., Webb, B. L., Inman, R. B. & Cox, M. M. RecA protein filaments: End-dependent dissociation from ssDNA and stabilization by RecO and RecR proteins. *J. Mol. Biol.* **265**, 519-540, (1997).

38 Montanari, A. & Mezard, M. Hairpin formation and elongation of biomolecules. *Phys. Rev. Lett.* **86**, 2178-2181, (2001).

39 Bennink, M. L. *et al.* Single-molecule manipulation of double-stranded DNA using optical tweezers: Interaction studies of DNA with RecA and YOYO-1. *Cytometry* **36**, 200-208, (1999).

40 van Mameren, J. *et al.* Counting RAD51 proteins disassembling from nucleoprotein filaments under tension. *Nature*, in press, (2008).

41 Zaitsev, E. N. & Kowalczykowski, S. C. A novel pairing process promoted by Escherichia coli RecA protein: inverse DNA and RNA strand exchange. *Genes Dev.* **14**, 740-749, (2000).

42 De Zutter, J. K., Forget, A. L., Logan, K. M. & Knight, K. L. Phe217 regulates the transfer of allosteric information across the subunit interface of the RecA protein filament. *Structure* **9**, 47-55, (2001).

43 Bianco, P. R. & Weinstock, G. M. Interaction of the RecA protein of Escherichia coli with single-stranded oligodeoxyribonucleotides. *Nucleic Acids Res.* **24**, 4933-4939, (1996).

44 Odijk, T. Stiff chains and filaments under tension. *Macromolecules* **28**, 7016-7018, (1995).

45 Wang, M. D., Yin, H., Landick, R., Gelles, J. & Block, S. M. Stretching DNA with optical tweezers. *Biophys. J.* **72**, 1335-1346, (1997).

46 Bouchiat, C. *et al.* Estimating the persistence length of a worm-like chain molecule from force-extension measurements. *Biophys. J.* **76**, 409-413, (1999).

47 Register, J. C. & Griffith, J. RecA protein filaments can juxtapose DNA ends - An activity that may reflect a function in DNA-repair. *Proc. Natl. Acad. Sci. U. S. A.* **83**, 624-628, (1986).

48 Kiianitsa, K. & Stasiak, A. Helical repeat of DNA in the region of homologous pairing. *Proc. Natl. Acad. Sci. U. S. A.* **94**, 7837-7840, (1997).

49 Kowalczykowski, S. C., Lonberg, N., Newport, J. W., Paul, L. S. & Vonhippel, P. H. On the thermodynamics and kinetics of the cooperative binding of bacteriophage T4-coded Gene-32 (helix destabilizing) protein to nucleic-acid lattices. *Biophys. J.* **32**, 403-418, (1980).

50 Brenner, S. L., Zlotnick, A. & Stafford, W. F. RecA protein self-assembly .II. Analytical equilibrium ultracentrifugation studies of the entropy-driven self-association of RecA. *J. Mol. Biol.* **216**, 949-964, (1990).

51 Arenson, T. A., Tsodikov, O. V. & Cox, M. M. Quantitative analysis of the kinetics of end-dependent disassembly of RecA filaments from ssDNA. *J. Mol. Biol.* **288**, 391-401, (1999).

52 Nishinaka, T., Doi, Y., Hara, R. & Yashima, E. Elastic behavior of RecA-DNA helical filaments. *J. Mol. Biol.* **370**, 837-845, (2007).

53 VanLoock, M. S. *et al.* ATP-Mediated conformational changes in the RecA filament. *Structure* **11**, 187-196, (2003).

Supplementary Information

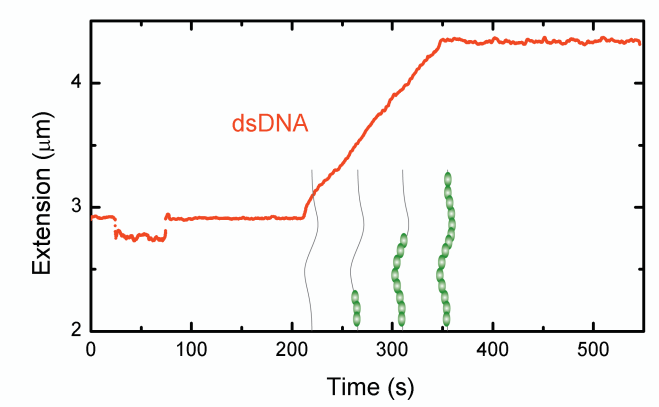


Figure S7.1: RecA polymerization on dsDNA. After flushing in 1 μM RecA, a pause is observed after which a filament nucleates and extends with a linear growth profile corresponding to the extension of a single RecA-dsDNA filament.

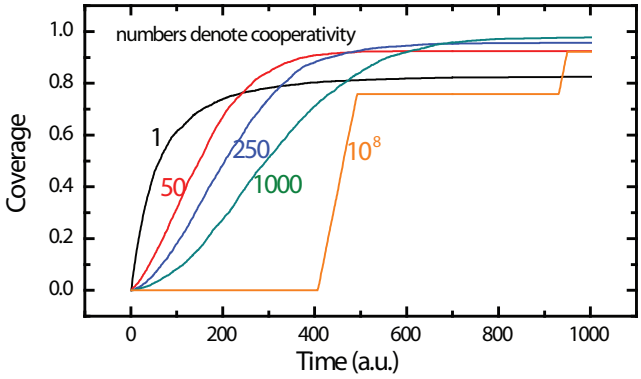


Figure S7.2: Shape of the binding profile changes from exponential for a cooperativity of 1 (black line) to sigmoidal for higher cooperativity (red, blue, green). Very large cooperativity leads to linear growth profiles (orange), the stepwise behavior in this curve is due to the unidirectional growth used in these simulations.

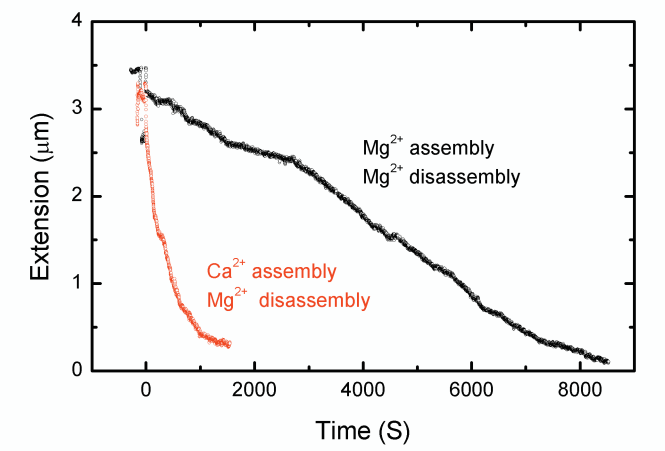


Figure S7.3: RecA disassembly after assembly at a high stretching force (20 pN) prevents the formation of secondary structure in ssDNA. An exponential disassembly profile is observed for filaments assembled in Ca^{2+} buffer indicating the presence of gaps (red curve). A filament formed in the presence of Mg^{2+} at a stretching force of 20 pN shows a linear dissociation profile indicative of dissociation from a single continuous RecA-ssDNA filament.

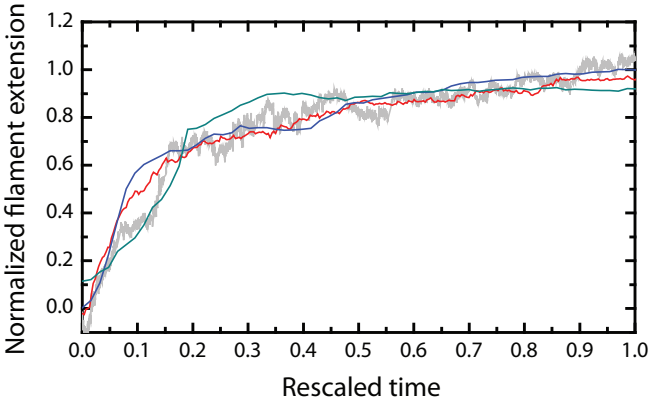


Figure S7.4: Binding profiles at different RecA concentrations: [RecA]=0.1 μM (gray), 0.4 μM (red), 0.7 μM (green), and 1.0 μM (blue). The time axis for all four traces were rescaled by the inverse of the nucleation rate determined from Monte Carlo simulation fits. The shape of the binding profile at different RecA concentrations is conserved confirming that filament nucleation and extension occur with the same binding unit.

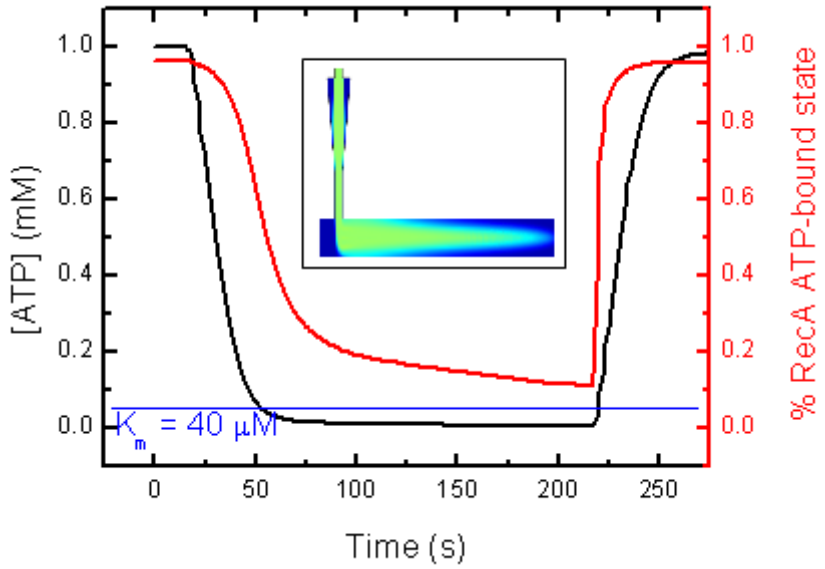


Figure S7.5: Time dependent ATP concentrations and ATP binding by RecA. Time dependent ATP concentrations were calculated using a commercial computer program Comsol multiphysics (black line). Inset shows the flow cell geometry (not to scale) at $t = 15$ s. The fraction of RecA in the ATP-bound state, Θ was calculated as $\Theta = [ATP] / (K_m + [ATP])$, where K_m was taken as 40 μM.

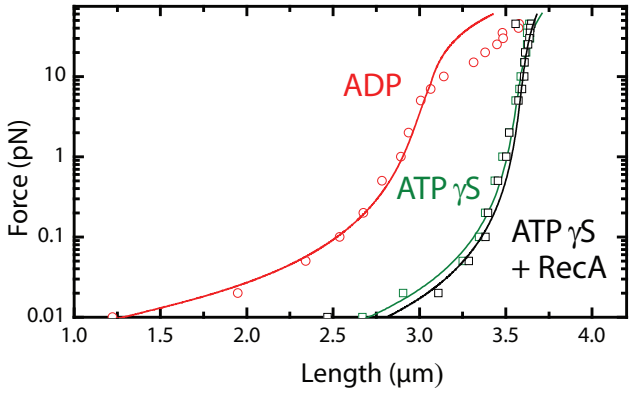


Figure S7.6: A force-extension curve for a RecA-ssDNA filament in the ADP-bound state showed a marked shift when ATPγS was added (red to green curve). Subsequently changing the buffer for one containing ATPγS and RecA did not noticeably change the force extension curve. This proves that no appreciable disassembly of RecA from the ADP filament takes place which would have resulted in a length increase when RecA was added (black).

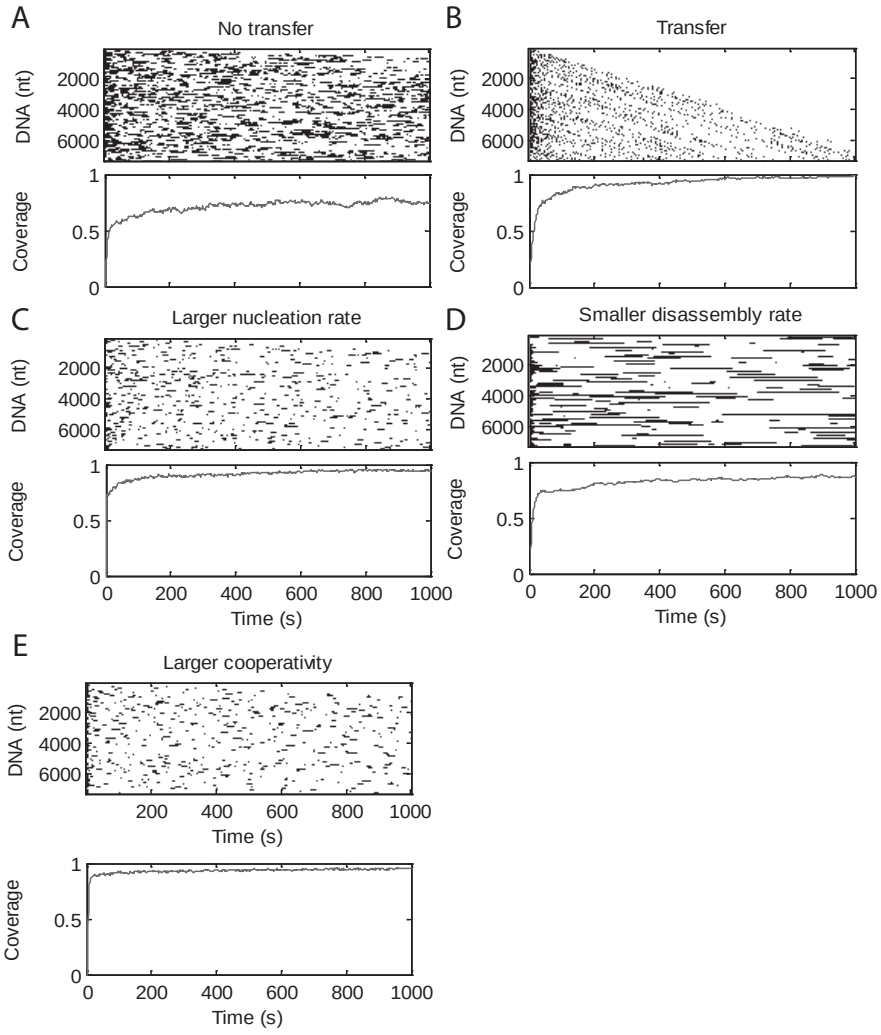


Figure S7.7: Monte Carlo simulations of RecA filament formation. Upper panels show kymographs of the simulated RecA-ssDNA filament. The evolution of a 7.3 kb ssDNA molecule, depicted on the *y*-axis, is shown with time. Gray areas show bound protein; at $t = 0$ no RecA is bound to the ssDNA, as time progresses RecA rapidly nucleates and a filament forms. The lower panels show the coverage of the molecule. Simulation parameters were obtained from experimental data as described in the main text. The table below list the simulation parameters. The modified parameters are shown in bold. (**A**) Monte Carlo simulation of RecA-ssDNA filament formation excluding transfer. Many gaps still remain after 10000s when transfer of monomers is not taken into account. Including transfer (**B**) yields a continuous filament. Monte Carlo simulation of RecA-ssDNA filament formation exclud-

ing transfer but with a 10 fold increased nucleation rate (**C**). Many gaps still remain after 1000s when transfer of monomers is not taken into account. Monte Carlo simulation of RecA-ssDNA filament formation excluding transfer but with a 10 fold reduced disassembly rate (**D**). Many gaps still remain after 1000s when transfer of monomers is not taken into account. Monte Carlo simulation of RecA-ssDNA filament formation excluding transfer but with a 10 fold increased cooperativity (**E**). Many gaps still remain after 1000s when transfer of monomers is not taken into account

Table S7.1: Simulation parameters for Figure S7.7

	A	B	C	D	E
Nucleation unit (monomers)	6	6	6	6	6
Extension unit (monomers)	6	6	6	6	6
Binding size (nt)	3	3	3	3	3
Nucleation rate ($s^{-1}nt^{-1}$)	0.006	0.006	0.06	0.006	0.006
Cooperativity	15	15	15	15	150
Disassembly (mon/s/filament)	0.3	0.3	0.3	0.03	0.03
Transfer (mon/s/filament)	0	2	0	0	0

CHAPTER 8

Effect of the BRCA2 CTRD domain on RAD51 filaments analyzed by an ensemble of single molecule techniques

Homologous recombination is essential for the preservation of genome stability, thereby preventing cancer. The recombination protein RAD51 drives DNA strand exchange, which requires the assembly, rearrangement and disassembly of a RAD51 filament on DNA, coupled to ATP binding and hydrolysis. This process is facilitated and controlled by recombination mediators and accessory factors. Here, we have employed a range of single molecule techniques to determine the influence of the C-terminal RAD51 interaction domain (CTRD) of the breast cancer tumor suppressor BRCA2 on intrinsic aspects of RAD51-DNA interactions. We show that at high concentration the CTRD entangles RAD51 filaments and reduces RAD51 filament formation in a concentration dependent manner. It does not affect the rate of filament disassembly measured as the loss of fluorescent signal due to intrinsic RAD51 protein dissociation from dsDNA. We conclude that, outside the context of the full-length protein, the CTRD does not reduce RAD51 dissociation kinetics, but instead hinders filament formation on double-stranded DNA. The CTRDs mode of action is most likely sequestration of multiple RAD51 molecules thereby rendering them inactive for filament formation on double-stranded DNA.

This chapter has been published as: J. T. Holthausen, M. T. J. van Loenhout, H. Sanchez, D. Ristic, S. E. van Rossum-Fikkert, M. Modesti, C. Dekker, R. Kanaar and C. Wyman. Effect of the BRCA2 CTRD domain on RAD51 filaments analyzed by an ensemble of single molecule techniques. Nucleic Acids Res. 39, 6558 (2011).

Introduction

Double-stranded DNA breaks (DSBs) are severe lesions that can result in chromosomal rearrangements leading to cellular senescence or the onset and progression of cancers[1,2]. In S and G2 phases of the cell cycle, homologous recombination (HR) provides a pathway for the faithful repair of such lesions[3]. HR can be divided in three steps: pre-synapsis, synapsis and post-synapsis[4]. Pre-synapsis consist of the recognition of a DSB, and the resection of one strand leaving tailed DNA that is bound by RAD51 to form a nucleoprotein filament[5,6]. During synapsis this nucleoprotein filament interacts with the intact sister chromatid (homology search), and strand invasion occurs at the site of homology resulting in joint molecule formation and strand exchange. Post-synapsis is defined as the events leading to the recovery of lost information by polymerase-mediated DNA resynthesis and the resolution of branched DNA structures resulting in two intact chromatids[7,8].

Biochemical assays have defined RAD51 recombinase as the catalyst of synapsis. It forms a helical filament on single-stranded DNA (ssDNA) that can drive the DNA transactions required for HR *in vitro*. The RAD51 nucleoprotein filament will invade a homologous double-stranded DNA (dsDNA) and form a joint molecule. To catalyze strand exchange the RAD51 nucleoprotein filament undergoes dynamic rearrangements that result in DNA strands exchanging base-paired partners leading to the formation of a filament on heteroduplex DNA – one strand being the invading ssDNA while the other is the complementary strand of the template DNA. The next step in repairing a DNA break requires the invading strand to prime DNA extension by a polymerase. This stage requires RAD51 removal from the heteroduplex DNA[9-11], and these dissociation reactions have been addressed in previous single molecule studies[12-15].

The actions of RAD51 are facilitated by recombination mediators and controlled by accessory factors[16,17]. One such mediator is the BRCA2 protein which is essential for efficient HR in mammalian cells[16,18]. The role that BRCA2 plays in maintaining genome stability has been attributed to its ability to interact directly with RAD51[19,20]. Biochemical studies with BRCA2 peptides have defined interactions with RAD51 via a series of eight BRC repeats and a C-terminal domain, designated CTRD (C-terminal RAD51 interaction domain; previously referred to as TR2[21,22]). This domain contains a cyclin-dependent kinase phosphorylation site that modulates RAD51 binding[21,23]. The BRC domains, best characterized by BRC4, which has the highest affinity for RAD51, interact with RAD51 at the polymerization domain and can thereby disrupt RAD51 filaments at concentrations equal to or higher than the RAD51 concentration[22-25]. The CTRD of BRCA2 interacts with multimeric RAD51 and can inhibit the destabilization of RAD51 filaments caused by the BRC repeats[22,24,25]. In avian cells the CTRD of BRCA2 influences the persistence of RAD51 in local nuclear accumulations in a manner consistent with stabilizing RAD51 filaments[26]. To determine whether the CTRD of BRCA2 might provide an effect opposite to the BRC repeats, by helping to stabilize RAD51 filaments on dsDNA, we used an array of single-molecule techniques together with ensemble studies. First we visualized the effect of the CTRD on RAD51 dissociation

from dsDNA. We observed individual filaments formed with fluorescent RAD51 and quantified their disassembly in time by fluorescent microscopy in a microfluidic flow system[12]. We then analyzed the effect of the CTRD on RAD51 nucleoprotein filament structure by optical fluorescence microscopy, by scanning force microscopy imaging[27], and a combination of both techniques[28]. We then analyzed the effect of the CTRD on RAD51 filament assembly on dsDNA, following the extension of individual dsDNA molecules by RAD51 binding, in magnetic tweezers[13,29].

Results

The initial aim was to measure the dissociation kinetics of RAD51 in the absence and presence of the CTRD of BRCA2 to determine whether the CTRD played a role in filament stabilization.

Filament dissociation

RAD51 filaments were formed by incubating Alexa Fluor 488 labeled protein with 48-kb lambda phage dsDNA molecules in the presence of ATP and $CaCl_2$, conditions that stabilize RAD51 nucleoprotein filaments[12,27,30]. The lambda phage dsDNA molecules were biotinylated at one end in order to tether the nucleoprotein filaments on the neutravidin-coated surface of a flow cell, where they were extended by buffer flow and observed in a fluorescent microscope. ATPase dependent dissociation of RAD51 from dsDNA was measured as a loss of fluorescent signal over time, corrected for background. Control traces in Ca^{2+}/ATP buffer in absence or presence of a synthetic CTRD peptide (Figure 8.1 hollow circle and triangle, respectively) showed that the RAD51 complex was stable over time and that the CTRD did not affect the photostability of the fluorophore. Switching buffer to one containing ATP and $MgCl_2$ triggered RAD51 dissociation from dsDNA[12,14]. Stabilization of RAD51 filaments by the CTRD was expected to slow the kinetics for RAD51 dissociation[23]. As a reference, RAD51 filaments were assembled in absence of the CTRD and RAD51 filament disassembly was measured over time (Figure 8.1, solid circles). When the CTRD was added at a ratio of RAD51:CTRD = 15:1 to pre-formed filaments, there was no detectable effect of the CTRD on RAD51 dissociation from dsDNA (Figure 8.1, solid triangles). The dissociation data overlap which showed that the rate is the same and implies that there is no difference in the way dissociation takes place. A phosphorylated CTRD peptide (P-CTRD), which does not interact with RAD51, also did not affect the rate of RAD51 dissociation significantly (Figure 8.1, crosses).

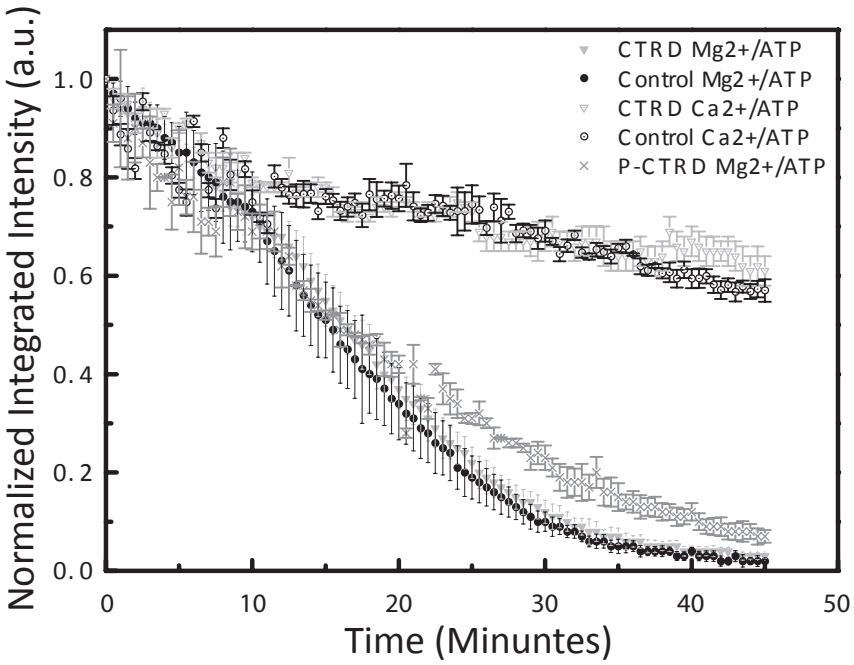


Figure 8.1: The effect of the CTRD on RAD51 disassembly from individual dsDNA molecules in a flow cell. RAD51 disassembly from dsDNA was measured as loss of fluorescence signal. The CTRD did not significantly influence the dissociation rate when ATP hydrolysis was triggered by Mg^{2+}/ATP. The Ca^{2+}/ATP control filaments were also not affected by the presence of the CTRD. The non-interacting P-CTRD also did not influence the dissociation rate significantly. The error bars represent the standard error.

Confirming CTRD interaction with RAD51-DNA complexes

To demonstrate that the CTRD peptide can interact with RAD51-dsDNA complexes under our reaction conditions, we employed electrophoretic mobility shift assays (EMSA). Using conditions identical to those in the flow cell before switching to Mg^{2+}/ATP buffer (i.e. ATPase suppressing Ca^{2+}/ATP buffer) complex formation between RAD51-dsDNA and the CTRD was assessed using agarose gels (Figure 8.2). In the flow cell experiment the ratio RAD51:CTRD was ~ 15:1. Therefore the amount of the CTRD was varied in a range encompassing the ratio in the flow cell. A higher amount of RAD51 was needed to fully shift the short dsDNA (see Figure S8.1) and therefore the CTRD concentrations were adjusted to keep the ratios of RAD51:CTRD similar. The CTRD concentrations used gave RAD51:CTRD ratios of 22:1 (45 nM CTRD), 11:1 (91 nM CTRD) and 7:1 (137 nM CTRD). Figure 8.2 shows a representative gel where titration of the CTRD had a minor effect on dsDNA migration in the controls (lanes 2-4), but showed clear interaction with the RAD51-dsDNA complexes by "super-shifting" them

in a concentration dependent manner (lanes 6-8). To show specificity, the EMSA assay was also performed with P-CTRD, which includes a phopsphate group at the cyclin-dependent kinase phosphorylation site (S3291E), that does not interact with RAD51[21]. This non-interacting P-CTRD failed to form ternary complexes with the RAD51-dsDNA complex (Figure S8.2). The assays confirm that the CTRD interacts with RAD51-dsDNA complexes at a similar ratio of RAD51:CTRD and in the same buffer conditions that were used in the flow cell.

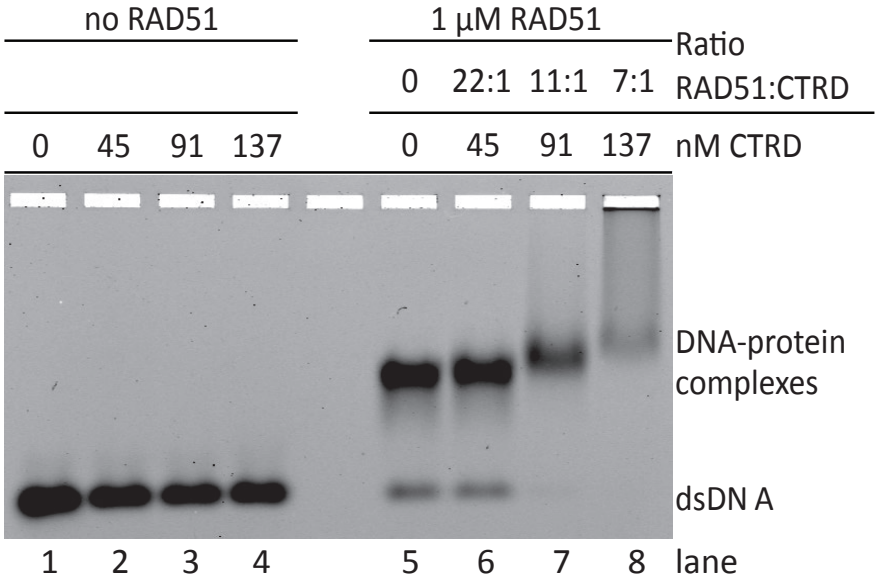


Figure 8.2: CTRD interaction with RAD51 filaments. Electrophoretic mobility shift assays (EMSA) were preformed with fluorescently end-labeled dsDNA and RAD51 (1 µM), CTRD or a combination of both using agarose gels to separate unbound from bound DNA. Reaction mixtures contained the indicated concentrations of the CTRD. Interaction of the CTRD with the RAD51–dsDNA complexes is evident in lanes 6–8 as 'super-shifting' of the complex in a CTRD concentration dependent manner.

CTRD-induced entanglement of RAD51-DNA complexes

Previous biochemical assays indicate that the BRCA2 CTRD stabilizes RAD51 filaments[23]. These experiments, however, utilized excess of the CTRD with respect to RAD51. To approach a comparable ratio of RAD51:CTRD, we increased the CTRD concentration 10-fold with respect to the conditions used for the experiments shown in Figure 8.1. However, distinctly separated, flow-stretched filaments were not observed, even at a RAD51:CTRD ratio of 1.5:1. Instead round, bright fluorescent signals were observed that presumably represented CTRD-induced entanglement of RAD51-dsDNA complexes (data not shown). A portion of these RAD51/CTRD-dsDNA complexes was deposited for imaging using combined fluorescence and scanning force microscopy

(SFM). Fluorescence detection allows locating and identifying the labeled RAD51 and SFM provides nm resolution structural information of the protein-DNA complexes (Figure 8.3). The fluorescent image (Figure 8.3A) shows an overlay of a representative complex of RAD51 protein and fluorescent polystyrene beads. The polystyrene beads serve to align the combined fluorescent picture (Figure 8.3A) with the topographic image (Figure 8.3B) obtained by SFM[28]. Analysis of the SFM image (Figure 8.3F) revealed that it was indeed an entanglement of (partial) RAD51 filaments since the height of the structure corresponded to single RAD51 filaments on dsDNA, while unbound dsDNA was also observed (compare Figure 8.3, panel F with E and H, respectively; see Figure 8.3D for height traces of the corresponding cross sections). Entanglement was due to the presence of the CTRD because in a control experiment using the same concentration of non-interacting P-CTRD no aggregation was observed (Figure 8.3C). Instead regular RAD51 filaments on dsDNA could be seen (compare Figure 8.3G with 3E, corresponding height traces in Figure 8.3D). In flow cell experiments without CTRD about 15 ± 6 filaments on average could be seen in a 40 µm by 40 µm field (n = 95). However, at CTRD concentrations above 45 nM about 1.5 ± 0.7 intense, round fluorescent signals were observed in the same size field (n = 9). Since the same amount of DNA and RAD51 were used in both filament formation reactions this suggests that one such aggregate consists on average of approximately 10 filaments.

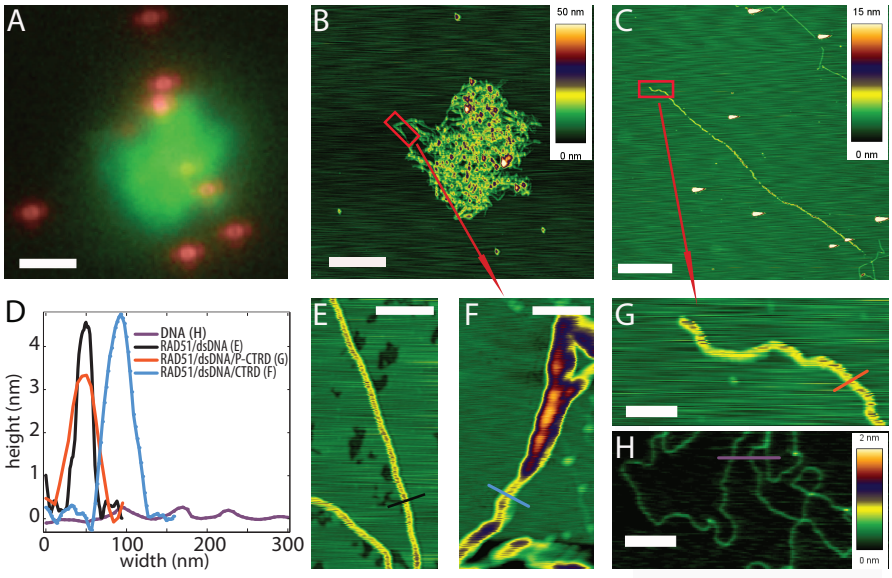


Figure 8.3: CTRD-induced RAD51 filament entanglement. Fluorescent (**A**) and SFM (**B, C, E**) images of RAD51 filaments on phage I dsDNA. Filaments were either assembled in the presence of 45 nM CTRD (A, B, F), 45 nM of phosphorylated control peptide P-CTRD (C and G) or without peptide (E). (**A**) Image of combined fluorescence signals from RAD51 and fluorosphere markers. The bright green object in the center of the image in panel A was

scanned for nanometer topography. (**B**) The topographic image reveals entangled RAD51 filaments. (**F**) A higher resolution picture of the red boxed area in panel B shows stretches of single filament (cyan cross section) and partial filaments revealing DNA stretches not bound by RAD51. (**D**) Plot of height and width of the cross sections taken in scan F (blue trace) and G (red trace) comparing them to control depositions of RAD51 filaments in absence of peptide (panel E, black trace) and dsDNA alone (panel H, magenta trace). Images A–C are 10 µm by 10 µm, the white bar is 2 µm. Images E–H are 1 µm by 0.5 µm, the white bar 0.2 µm. The height in the topography scans is indicated with color as shown by the scale bars to the right of each image. The height scale in panels E–G corresponds to the one in panel C.

Filaments formed in the presence of CTRD are qualitatively distinct

In the RAD51 dissociation experiments in the flow cell (such as plotted in Figure 8.1) the CTRD was added after RAD51 filament formation and incubated for an additional 15 minutes at 37°C. Examples of such filaments are shown in the upper left picture in Figure 8.4 (Figure 8.4A). These filaments appear similar as filaments with RAD51 alone[12]; the fluorescence covers most of the DNA contour with some gaps. In previous experiments the CTRD influenced filaments when pre-incubated with RAD51 for 15 minutes at 37°C before addition of DNA[21,23]. Therefore we prepared filaments after pre-incubation of the CTRD with RAD51. Strikingly, RAD51 filaments were distinct from the filaments formed in the absence of the CTRD (compare Figure 8.4B to 4A). The filaments formed after pre-incubation appeared to be less complete, consisting of shorter filament patches covering less DNA (Figure 8.4B). Thus a change in the order of addition revealed that a pre-incubation of the CTRD with RAD51 is interfering with extensive filament formation. The presence of the CTRD during filament formation results in incomplete, "patchy" filaments, which could not be analyzed reliably in dissociation experiments. Therefore we decided to analyze these filaments using SFM.

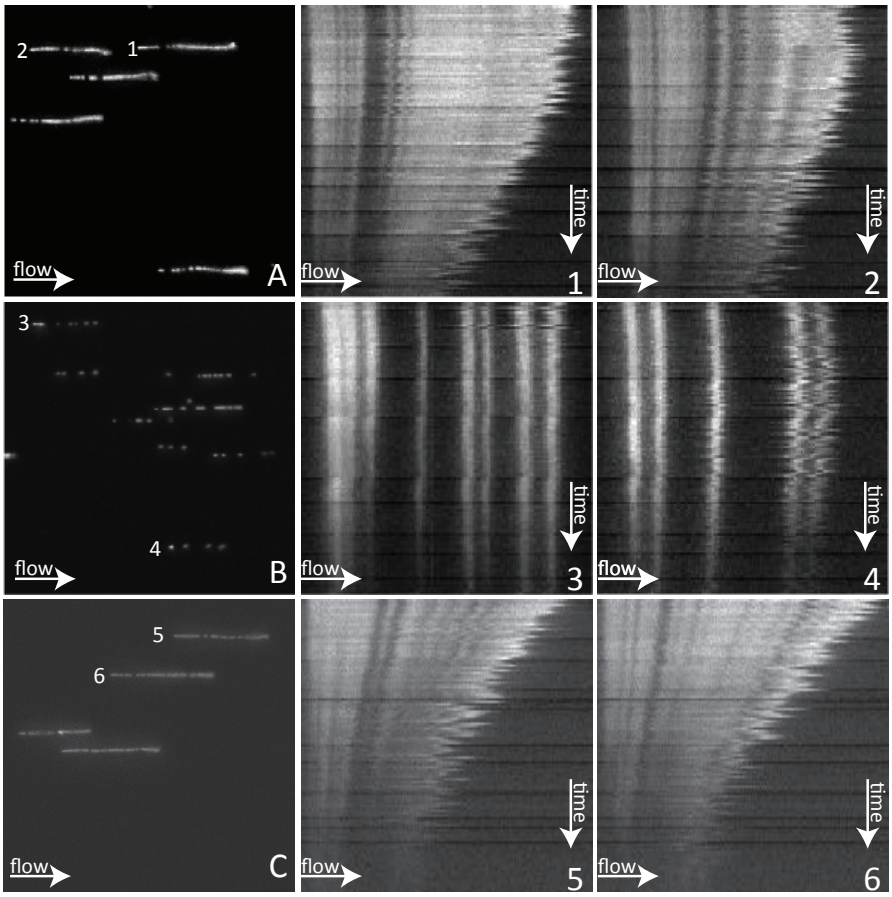


Figure 8.4: Appearance of RAD51 filaments in flow cell experiments. RAD51 nucleoprotein filaments (**A–C**) and their corresponding kymographs showing disassembly over time after ATP hydrolysis has been triggered (A 1/2, B 3/4, C 5/6). Fluorescent signal was acquired as described previously for flow cell experiments[12]. The images on the left (A–C) show 40 µm by 40 µm images of RAD51 filaments on dsDNA inside the flow cell before ATP hydrolysis was triggered. Adjacent to images A, B and C are the example kymographs of the filaments numbered accordingly. The kymographs are pictures horizontally displaying a filament, with its anchor point on the left, and vertically displaying disassembly of RAD51 over time. Each pixel line towards the bottom of the kymograph represents a 20 s step for a total time of 45 min. The upper panels show the typical appearance of RAD51 filaments. The middle panels show filaments formed after pre-incubation of RAD51 and the CTRD peptide. These filaments appear patchy as evidenced by stretches of bare DNA and less protein bound per DNA molecule. The lower panels show filaments formed after pre-incubation with the phosphorylated control peptide (P-CTRD). These filaments appear similar to the ones without peptide as they are not patchy.

SFM analysis of filament length in presence of the CTRD

To assess the effect of the CTRD on filament formation and structure, nucleoprotein filaments formed after pre-incubation of RAD51 and the CTRD, at a ratio of 5:1, on 3-kb dsDNA were analyzed by SFM[27]. The images did not show a strong difference in appearance of the control filaments and those formed in the presence of the CTRD (data not shown). As RAD51 polymerization onto DNA stretches it up to 50% over regular B form length, we compared the contour lengths of RAD51 filaments without and with the CTRD (RAD51:CTRD = 5:1). The average elongation of filaments was not much decreased by the presence of the CTRD (1.4 ± 0.1 μm versus 1.29 ± 0.09 μm). However, the distribution of filament lengths was clearly skewed towards shorter filaments in the presence of the CTRD (Figure 8.5). A reduction in contour lengths could be attributed to reduced RAD51 polymerization onto DNA. Filament formation is a dynamic process consisting of nucleation and extension (growth) and the CTRD could specifically influence one or the other of these processes.

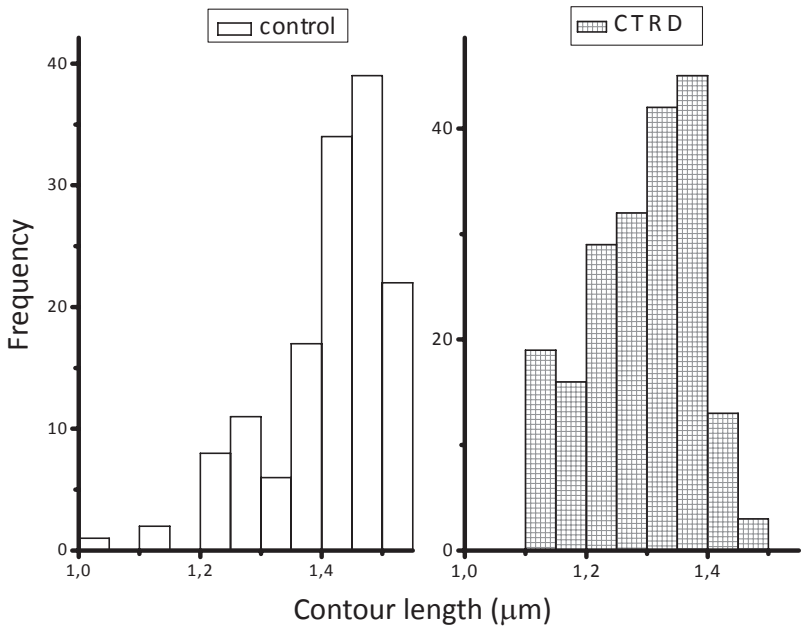


Figure 8.5: Contour length of RAD51 filaments formed in the presence and absence of the CTRD. The contour length of RAD51 filaments, from images obtained by SFM, were measured and plotted in histograms for control filaments (left) and filaments formed in the presence of the CTRD (right). The y-axis indicates the number of filaments, while the contour length in micrometer with a binning step of 50 nm is plotted along the x-axis. The left panel displays the histogram for the control filaments; the right panel for filaments formed after pre-incubation of RAD51 with CTRD at a ratio of 5:1.

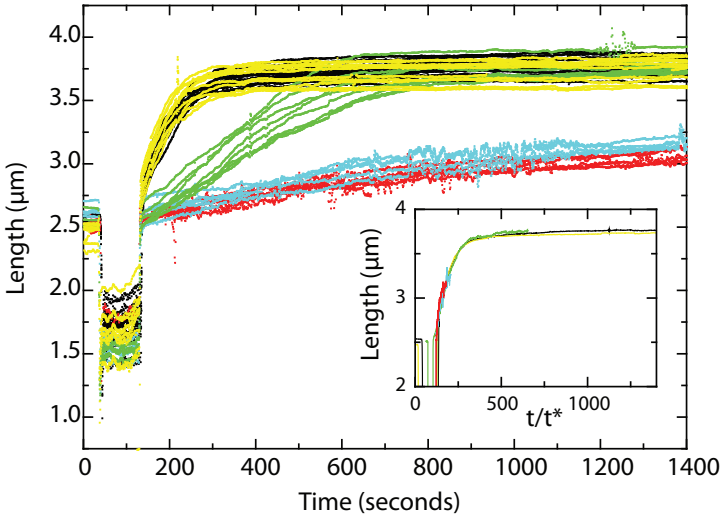


Figure 8.6: RAD51 filament assembly in real time. RAD51 filament assembly was measured as changes in the length of a DNA, tethered in magnetic tweezers, over time. Different CTRD concentrations are indicated by color: [CTRD] = 0 (black), [CTRD] = 2 nM (yellow), [CTRD] = 6nM (green), [CTRD] = 12 nM (cyan) and [CTRD] = 60 nM (red). Filament assembly was measured on several individual DNA molecules (n = 5–12) for each CTRD concentration. RAD51 and the CTRD were pre-incubated for 15 min at 37 °C. RAD51 concentration was constant at 150nM in all experiments. The inset shows the average growth profile for each concentration of the CTRD rescaled by t/t*, where t*=1/g and g is the nucleation rate determined from fitting to Monte Carlo simulations.

The CTRD decreases RAD51 filament assembly rate in magnetic tweezers measurements

RAD51 filament assembly was followed on individual dsDNA molecules (7.3 kb) in magnetic tweezers[13,31]. The DNA molecules were tethered between a magnetic bead and the surface of a flow cell. A force of 7 ± 2 pN was applied by a pair of magnets and video microscopy was used to measure the end-to-end distance of the DNA tether. Due to the known extension of dsDNA upon RAD51 binding an increase in tether length can be directly correlated to the polymerization of RAD51 onto dsDNA. Assembly of a RAD51 filament in the absence of the CTRD resulted in a DNA tether length increase from 2.54 ± 0.04 µm to 3.76 ± 0.07 µm corresponding to a 1.48-fold extension (Figure 8.6, black curves), similar to previous observations[13]. RAD51 filament assembly was however very sensitive to pre-incubation of RAD51 with the CTRD in a concentration dependent manner. At a ratio of RAD51:CTRD = 12.5:1 filament assembly was already markedly reduced (Figure 8.6, cyan curves).

The assembly profiles obtained by magnetic tweezers experiments were analyzed by fitting to Monte Carlo simulations, assuming a binding size of 3 bp per RAD51 monomer and filament nucleation and extension by RAD51 pentamers as described

previously[13,29,31]. Nucleation is the dominant step in filament formation. For the simulations the extension rate was linked to the nucleation rate by setting the cooperativity number ω to a value of 100 in line with results of fits with cooperativity as a free independent fit parameter and previous results[29]. The Monte Carlo fitting procedure yielded nucleation rates of RAD51 at different CTRD concentrations (Figure 8.7). The validity of this approach for extracting filament assembly rates was demonstrated by plotting averaged assembly curves for each CTRD concentration against a rescaled time axis (inset Figure 8.6). The time axis was normalized based on rates found after fitting to Monte Carlo simulations. After normalizing, the filament assembly profiles at different CTRD concentrations overlapped (inset Figure 8.6), confirming that the shape of the assembly curves is similar and that these curves differ in only one parameter, namely the assembly rate. Filament assembly was strongly decreased by the presence of the CTRD in a concentration dependent manner. The influence of the CTRD on filament assembly rates indicated that the amount of RAD51 available for filament assembly was decreased in the presence of the CTRD. When RAD51 nucleation rates are plotted against CTRD concentration (Figure 8.7), the power-law shape of the fit indicates that the CTRD interacts with multimeric RAD51, calculated to be on average 3-4 RAD51 monomers, (see Materials and Methods) consistent with previous observations[23,25].

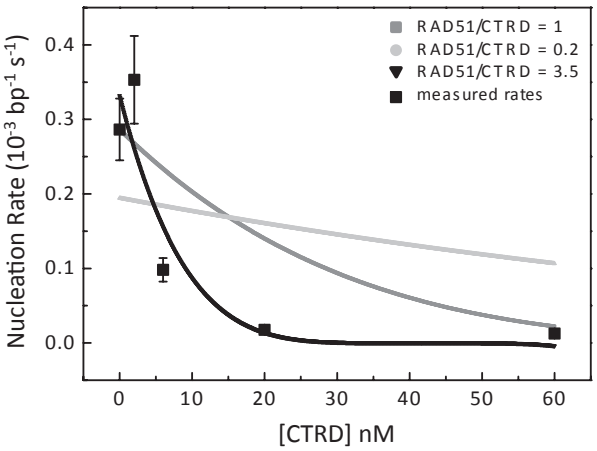


Figure 8.7: Effect of the CTRD on filament assembly rate. The nucleation rates of RAD51 filaments were extracted from fitting Monte Carlo simulations to the data from Figure 8.6 at a constant cooperativity number of w=100. Error bars indicate the standard deviations in the nucleation rates found by fitting the assembly profiles of Figure 8.6 at each CTRD concentration. A model used to fit nucleation rates (see 'Materials and Methods' section) resulted in a best fit for the number of RAD51 monomers interacting with CTRD of 3.5 ± 0.4 (black trace). For comparison a model assuming a CTRD interaction with monomeric RAD51 (dark grey trace), and a model assuming an interaction of monomeric RAD51 with multiple CTRDs (light gray trace) is shown.

Discussion

Here we have addressed the effect of the CTRD domain of BRCA2 on the key entity in HR, the RAD51 filament. RAD51 nucleoprotein filaments are dynamic structures undergoing constant rearrangements coupled to ATP hydrolysis. Either suppressing ATP hydrolysis or mechanical protein-protein interactions could stabilize RAD51 filaments. The CTRD of BRCA2 was proposed to stabilize RAD51 filaments by stabilization of protein-protein interactions, acting as a bridge over the RAD51 monomer interface in one model[23,25]. In the context of full-length BRCA2 the CTRD domain is expected to have a role modulating RAD51 activity based on observations that phosphorylation of S3291 of CTRD by CDKs abolishes RAD51 binding[21,26]. However, we observe no effect of the isolated CTRD domain on filament disassembly rates by direct visualization of RAD51 dissociation from lambda phage dsDNA. Although this result suggests the model based on CTRD bridging RAD51 monomers in a filament to prevent access of BRC-repeats might not hold up it is not contradictory to previous work showing that RAD51 filament disassembly by BRC-repeat peptides is inhibited by the CTRD[22–25]. That observation could also be due to CTRD induced RAD51 filament entanglement or aggregation as discussed below. Direct stabilization by altering inherent RAD51 dissociation was not previously tested[22–25]. Since RAD51 dissociation in our set-up is dependent on ATP hydrolysis[12,14], we can conclude that the presence of the CTRD (at a ratio of RAD51:CTRD = 15:1) does not suppress the ATPase activity of RAD51, nor is its interaction with the filament strong enough to slow dissociation in a flow cell set-up.

The absence of an effect of the CTRD peptide on RAD51 filament stability could have been due to absence of interaction with the filaments formed in our conditions. Although the amount of peptide available precluded adding it to the buffer flow during dissociation the persistence of CTRD-induced filament entanglement in the same flow conditions argues that the peptide does not simply dissociate. The EMSA experiments show that the CTRD interacts with RAD51-dsDNA complexes in a concentration-dependent manner in the same starting reaction conditions used in the flow cell i.e. Ca^{2+}/ATP. Because these gel assays required RAD51 at higher concentrations than in the flow cell, the concentration of the CTRD was also adjusted so that the ratio of RAD51:CTRD was 22:1, 11:1 and 7:1, encompassing the 15:1 ratio used in the flow cell. Thus, while the CTRD interacts with RAD51 filaments, it does not influence filament disassembly rates, suggesting that filament stability is not affected.

Several lines of evidence indicate that at higher concentration, or ratios to RAD51, CTRD causes entanglement of RAD51 filaments. Flow cell experiments were attempted at a higher protein concentration, such as required for the EMSA. At CTRD concentration of 45 nM, the lowest concentration showing an effect in the EMSA and corresponding to a final ratio of RAD51:CTRD of 1.5:1 in the flow cell, filaments could not be analyzed by flow stretching. The observation of round, intense, fluorescent signals, which sometimes untangled upon ATP hydrolysis, indicated that the long RAD51 filaments on lambda phage dsDNA may have aggregated in these conditions. Combined fluorescence and SFM imaging revealed that the large, brightly fluorescent

objects are indeed CTRD-induced entangled RAD51 filaments that include stretches of naked DNA (Figure 8.3). This aggregation of RAD51-DNA complexes at higher CTRD concentration would have a detrimental effect on interactions of other proteins with the filament structure. Indeed, CTRD-induced filament aggregation has also been noted in other studies[23]. This could explain the previously observed protective effect of the CTRD on filaments exposed to high concentrations of BRC4 that will otherwise disrupt filaments[23,25]. The correlation between stronger CTRD:RAD51 interaction and longer persistence of local nuclear accumulations of RAD51 is also consistent with CTRD-induced filament aggregation or entanglement[26].

By contrast, the CTRD-RAD51 interaction did inhibit filament assembly. Interaction of RAD51 and the CTRD before adding dsDNA had a strong effect on filament appearance, even at a final CTRD concentration of approximately 5 nM (ratio RAD51:CTRD = 291:1). The nucleoprotein filament formed had more frequent and larger gaps of naked DNA between protein-bound patches (compare Figure 8.4 upper and lower panel), an appearance we refer to as "patchy". Furthermore, the concentrations used here were equivalent to the ones in the disassembly experiments, indicating that the lack of an effect on filament disassembly in the flow cell was not due to concentrations of the CTRD insufficient for interaction with RAD51. RAD51filaments, formed on 3 kbp-long dsDNA, in the presence of the CTRD are also somewhat shorter than control filaments (Figure 8.5). The average DNA extension by filament pre-incubated with the CTRD is only about 70% that of the control filaments (29% versus 40%). In the magnetic tweezers assay, RAD51 filament assembly was very sensitive to pre-incubation with the CTRD. Rates progressively decreased with increasing CTRD concentration (Figure 8.6). The shape of the assembly curves reflects the interplay between nucleation and extension events in RAD51 DNA-binding to assemble a filament[13]. The average assembly curves for each CTRD concentration have the same shape, implying that the CTRD has not changed the way RAD51 extends dsDNA as reflected in the ratio of nucleation to extension. These data all support the idea that the CTRD inhibits RAD51 filament assembly by effectively reducing the active RAD51 concentration.

Alternatively, the CTRD could specifically cap nascent filament patches and block further extension. This would produce filaments with different appearance, shorter protein patches with a length correlating to the CTRD concentration, and would require more nucleation events to cover the DNA. However, filaments formed after pre-incubating RAD51 and the CTRD at a 5:1 ratio appeared mostly regular, without obvious gaps or multiple kinks that would indicate reduced extension events. Additionally, filament capping by the CTRD is expected to inhibit dissociation and to change the ratio of nucleation to extension in filament assembly kinetics. None of these effects were observed, thus the CTRD interaction with RAD51 apparently does not cap filaments.

The RAD51 assembly reactions show a power-law dependence of the nucleation rate on CTRD concentration. This was used to determine the stoichiometry of RAD51 affected by the CTRD. Assuming a nucleation unit of 5 RAD51 monomers[13,29]

the power-law fit had a minimum for least squares error i = 3.5 ± 0.4, indicating that one CTRD reacts on average with 3-4 RAD51 monomers. This is in accordance with the observation that the CTRD binds multimeric RAD51 in solution[23,25] and suggests that the mode of the CTRD action in this experiment is to sequester RAD51 from filament assembly.

Inactivation of 3-4 RAD51 monomers for filament formation by CTRD binding can explain altered filament appearance as well as reduced assembly rates. The alternative model, that extending RAD51 filament patches are capped in RecX-like fashion is unlikely[32-34]. Overall, our data are in accordance with recent work on the full-length BRCA2[35-37] and previous studies with BRC peptides[22-25,38-40]. Those studies show that BRCA2 and its fragments interact with RAD51-dsDNA complexes and that BRCA2 modulates RAD51 DNA-binding such that filaments preferentially form on ssDNA over dsDNA. Here we show that not only BRC repeats[23,38-40] but also the CTRD of BRCA2 interferes with RAD51 filament formation on dsDNA. These two peptide portions of BRCA2 differ in details of how they affect RAD51 filaments. Stimulation of RAD51 binding to ssDNA specifically occurs when ATP hydrolysis is possible[38,39]. In contrast, the effect we report of the CTRD on RAD51 filaments does not require ATP hydrolysis, as reactions occur in conditions where RAD51's ATPase activity is suppressed. The interaction stoichiometry of the two domains is also different. The BRC repeat forms a complex with RAD51 at a 1:1 ratio[24,41,42]. The CTRD interacts with RAD51 multimers[23,24], and we show it affects filaments formation at substoichiometric amounts. How these different interaction phenomena are manifested in the context of full-length BRCA2 is an important question that can now be addressed. The CTRD domain could work in concert with the BRC domains modulating DNA binding towards ssDNA.

Acknowledgments

We thank Drs. Tina Thorslund and Stephen West for the generous gift of CTRD peptides and for insightful comments on the manuscript.

Funding

This work was supported by grants from the Netherlands Organization for Scientific Research TOP to R.K., VICI 700.56.441 to C.W.]; the Netherlands Genomics Initiative; National Cancer InstitituteUSA [SBDR 5PO1CA092584]; Association for International Cancer Research [09-0633 to M.M]; Agence Nationale de la Recherche [RADORDER to M.M.]; and Marie Curie Intra European Fellowship [FP7-221069 to H.S.].

Materials & methods

Protein production

Wild-type and variants of human RAD51 were purified by similar procedures, essentially as previously described[12]. Briefly, RAD51 was expressed in bacteria, precipitated by ammonium sulfate, subsequently dialyzed and further purified by Heparin, gel filtration and anion exchange chromatography. The CTRD peptides were as described previously[21,22].

Protein labeling

RAD51 was Alexa Fluor 488 labeled on a specific cysteine residue using maleimidine chemistry and checked for activity as previously described[12]. The degree of labeling was 0.75 fluorophores per RAD51.

Surface tethering and visualization of fluorescent filaments in buffer

Flow cells, constructed from #1 glass cover slips (Menzel-Glaser) separated by a double-layer Parafilm in a custom holder, were prepared as follows: Neutravidin (Pierce) was introduced at 1 mg/ml and allowed to interact for 30 min. After removal of excess Neutravidin, the flow cell was washed and blocked with 2 mg/ml acetylated BSA, 2 mg/ml α-casein, 10 mM DTT, 50 mM Tris-HCl (pH 7.5), 30 mM KCl, and 10% glycerol. Lambda phage dsDNA was biotinylated at one end and labeled with digoxygenin at the other by annealing biotinylated oligonucleotides to CosL (5'-P-AAG TCG CCG CCC dR-bio dR-bio-bioTEG-3') and 5'-P-GGG CGG CGdRdig CCT CGG CGC CCG GCC GCG dTdigAA ACG CGG CCG GGC GCC GG-3') to CosR as described[43]. Filaments were assembled in a 20 µl reaction mixture containing 91 pM (molecules) of lambda phage dsDNA, 1.4 µM RAD51, 50 mM Tris-HCl (pH 7.5), 1 mM ATP, 2 mM $CaCl_2$, 1 mM DTT, and 60 mM KCl. Reaction mixtures were incubated at 37°C for 30 min. When indicated 4.8 nM CTRD peptide was added and the reaction mixtures incubated for another 15 min (RAD51:CTRD of 291:1). Alternatively, CTRD and RAD51 were incubated for 15 min before addition of the other reaction components. All reaction mixtures were diluted by addition of 380 µl $CaCl_2$/ATP buffer (50 mM Tris-HCl (pH 7.5), 1 mM ATP, 2 mM $CaCl_2$, 1 mM DTT, and 60 mM KCl), and injected into a flow cell. For the CTRD reactions the $CaCl_2$/ATP buffer also included 4.8 nM CTRD, changing the RAD51:CTRD ratio to 15:1. The flow was stopped for about 10 minutes to allow interaction of the biotynilated DNA with the Neutravidin surface. Unbound filaments were flushed by flow of $CaCl_2$/ATP buffer, including 4.8 nM CTRD if present in assembly reaction. Dissociation was triggered by switching to Mg^{2+}/ATP buffer (50 mM Tris-HCl (pH 7.5), 1 mM ATP, 10 mM $MgCl_2$, 1 mM DTT, and 60 mM KCl). Hydrodynamic flow was controlled with a precision pump (Harvard Apparatus). Dynamic visualization of filaments was performed with a Nikon 60 X or 100 X (NA 1.45) TIRF objective in a Nikon TE2000U inverted microscope equipped with a Cascade 512B CCD camera (Princeton Instruments) driven by Metamorph software (Molecular Devices). Excitation was performed with a mercury arc lamp. Intensity measurements were obtained

defining regions of interest around the construct contour length in all planes, tracking them over time and correcting for background.

DNA constructs for magnetic tweezers

A 7.3 kb dsDNA construct was prepared as described previously(33).

Combined fluorescence and scanning force microscopy

RAD51 filaments were assembled as described above. After 30 min incubation at 37°C, the CTRD or a phosphorylated CTRD (P-CTRD) was added to a final concentration of 45 nM (RAD51:CTRD of 31:1). Reaction mixtures were diluted 10 fold in 50 mM Tris-HCl (pH 7.5), 1 mM ATP, 2 mM $CaCl_2$, 1 mM DTT, 60 mM KCl, and 45 nM CTRD for flow cell experiments, changing the RAD51:CTRD to 1.5:1. 10 µl aliquots of the dilution were supplemented with 3 pM red fluorescent (580/605) 40 nm diameter beads (FluoSpheres® microspheres from Invitrogen) and deposited on a freshly cleaved Mica. Fluorescent images were obtained with the set-up described above and correlated with topographic images obtained with a NanoWizard II scanner (JPK instruments) as described[28].

Electrophoretic mobility shift assays

Reactions were performed in a final volume of 20 µl containing 5'-end Alexa Fluor 532-labeled 66bp ds DNA[44] at 66 nM (nucleotides), 1 µM RAD51, 50 mM Tris-HCl (pH 7.5), 1 mM DTT, 60 mM KCl, 2 mM $CaCl_2$, and 1 mM ATP[30]. After 5 min incubation at 37°C, the CTRD peptide was added to the indicated concentrations in a 2 µl volume and incubations were continued for a further 15 min at 37°C. Reaction mixtures were then supplemented with 3 µl 30% glycerol and 24 µl were fractionated by 0.8% agarose gel electrophoresis in 0.5X Tris-Borate for 2.5 hrs at 60 V at 4°C. Gels were analyzed using a T9600 typhoon scanner exciting the dye-coupled DNA with a 532 nm laser and detecting emission intensity using a the 555 nm BP 20 filter at 800 V PMT, 3 mm focal plane. Images obtained were analyzed with ImageQuant version 5.2 (Molecular Dynamics) or Image J.

Scanning force microscopy

Nucleoprotein filaments were formed in 10 µl reaction mixtures containing 7.5 µM 3-kb dsDNA (concentration in bp), 2.5 µM human RAD51, 25 mM HEPES–KOH (pH 7.5), 5 mM $CaCl_2$, 2 mM ATP and 30 mM KCl. Reaction mixtures were incubated at 37°C for 1 hr and then placed on ice. When indicated, 500 nM CTRD was added to RAD51 for 15 min at 37 °C before addition of the other components (RAD51:TR2 = 5:1). Aliquots of the reaction mixtures were then diluted 15-fold in 10 mM HEPES–KOH (pH 7.5) and 10 mM MgCl2 and deposited on freshly cleaved mica, scanned and analyzed as previously described[27].

Magnetic tweezers assay

The magnetic tweezers set-up used in these experiments was described previously[45]. By using image processing, 10-nm position accuracy of the bead was obtained in all three dimensions. To exclude the effect of thermal drift, all positions were measured relative to beads fixed to the bottom of the flow cell. DNA constructs carrying a magnetic bead at one end were anchored to the bottom of a flow cell. Experiments were started by addition of RAD51 or RAD51 pre-incubated with the CTRD at the indicated concentrations, in 50 mM NaCl, 50 mM Tris-HCl pH 7.5, 2 mM $CaCl_2$, 1 mM ATP. All measurements were carried out at 25 °C.

The nucleation rate dependence on the CTRD concentrations was fitted with a model that takes RAD51 nucleation into account. RAD51 nucleation requires multiple RAD51 monomers according to $nRAD51+DNA \rightarrow (RAD51)n-DNA$, where values of n are reported between 2.7 and 4.3, indicating that nucleation requires 3 to 5 RAD51 monomers[13,15]. This model leads to a power-law dependence of the nucleation rate according to $R_{nuc} \propto [RAD51]^n$. We propose a model where the CTRD can interact with multiple RAD51 monomers to reduce the concentration of active RAD51 according to $[RAD51]_{active} = [RAD51]-i[CTRD]$, where i is the number of RAD51 monomers that a CTRD peptide binds.

References

1 Hoeijmakers, J. H. J. Genome maintenance mechanisms for preventing cancer. *Nature* **411**, 366-374, (2001).

2 West, S. C. Molecular views of recombination proteins and their control. *Nat. Rev. Mol. Cell Biol.* **4**, 435-445, (2003).

3 Wyman, C. & Kanaar, R. DNA double-strand break repair: All's well that ends well. *Annu. Rev. Genet.* **40**, 363-383, (2006).

4 Wyman, C., Ristic, D. & Kanaar, R. Homologous recombination-mediated double-strand break repair. *DNA Repair* **3**, 827-833, (2004).

5 Longhese, M. P., Bonetti, D., Manfrini, N. & Clerici, M. Mechanisms and regulation of DNA end resection. *EMBO J.* **29**, 2864-2874, (2010).

6 Mimitou, E. P. & Symington, L. S. DNA end resection: Many nucleases make light work. *DNA Repair* **8**, 983-995, (2009).

7 Rass, U. *et al.* Mechanism of Holliday junction resolution by the human GEN1 protein. *Genes Dev.* **24**, 1559-1569, (2010).

8 Svendsen, J. M. & Harper, J. W. GEN1/Yen1 and the SLX4 complex: solutions to the problem of Holliday junction resolution. *Genes Dev.* **24**, 521-536, (2010).

9 Filippo, J. S., Sung, P. & Klein, H. Mechanism of eukaryotic homologous recombination. *Annu. Rev. Biochem.* **77**, 229-257, (2008).

10 Li, X. & Heyer, W. D. RAD54 controls access to the invading 3-OH end after RAD51-mediated DNA strand invasion in homologous recombination in Saccharomyces cerevisiae. *Nucleic Acids Res.* **37**, 638-646, (2009).

11 Li, X., Stith, C. M., Burgers, P. M. & Heyer, W. D. PCNA Is Required for Initiation of Recombination-Associated DNA Synthesis by DNA Polymerase delta. *Mol. Cell* **36**, 704-713, (2009).

12 Modesti, M. *et al.* Fluorescent human RAD51 reveals multiple nucleation sites and filament segments tightly associated along a single DNA molecule. *Structure* **15**, 599-609, (2007).

13 van der Heijden, T. *et al.* Real-time assembly and disassembly of human RAD51 filaments on individual DNA molecules. *Nucleic Acids Res.* **35**, 5646-5657, (2007).

14 van Mameren, J. *et al.* Counting RAD51 proteins disassembling from nucleoprotein filaments under tension. *Nature* **457**, 745-748, (2009).

15 Hilario, J., Amitani, I., Baskin, R. J. & Kowalczykowski, S. C. Direct imaging of human Rad51 nucleoprotein dynamics on individual DNA molecules. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 361-368, (2009).

16 Yuan, S. S. F. *et al.* BRCA2 is required for ionizing radiation-induced assembly of rad51 complex in vivo. *Cancer Res.* **59**, 3547-3551, (1999).

17 Moynahan, M. E., Pierce, A. J. & Jasin, M. BRCA2 is required for homology-directed repair of chromosomal breaks. *Mol. Cell* **7**, 263-272, (2001).

18 Yu, V. P. C. C. *et al.* Gross chromosomal rearrangements and genetic exchange between nonhomologous chromosomes following BRCA2 inactivation. *Genes Dev.* **14**, 1400-1406, (2000).

19 Venkitaraman, A. R. Linking the Cellular Functions of BRCA Genes to Cancer Pathogenesis and Treatment. *Annual Review of Pathology-Mechanisms of Disease* **4**, 461-487, (2009).

20 Thorslund, T. & West, S. C. BRCA2: a universal recombinase regulator. *Oncogene* **26**, 7720-7730, (2007).

21 Esashi, F. *et al.* CDK-dependent phosphorylation of BRCA2 as a regulatory mechanism for recombinational repair. *Nature* **434**, 598-604, (2005).

22 Galkin, V. E. *et al.* BRCA2 BRC motifs bind RAD51-DNA filaments. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 8537-8542, (2005).

23 Esashi, F., Galkin, V. E., Yu, X., Egelman, E. H. & West, S. C. Stabilization of RAD51 nucleoprotein filaments by the C-terminal region of BRCA2. *Nat. Struct. Mol. Biol.* **14**, 468-474, (2007).

24 Davies, O. R. & Pellegrini, L. Interaction with the BRCA2 C terminus protects RAD51-DNA filaments from disassembly by BRC repeats. *Nat. Struct. Mol. Biol.* **14**, 475-483, (2007).

25 Petalcorin, M. I. R., Galkin, V. E., Yu, X., Egelman, E. H. & Boulton, S. J. Stabilization of RAD-51-DNA filaments via an interaction domain in Caenorhabditis elegans BRCA2. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 8299-8304, (2007).

26 Ayoub, N. *et al.* The carboxyl terminus of Brca2 links the disassembly of Rad51 complexes to mitotic entry. *Curr. Biol.* **19**, 1075-1085, (2009).

27 Ristic, D. *et al.* Human Rad51 filaments on double- and single-stranded DNA: correlating regular and irregular forms with recombination function. *Nucleic Acids Res.* **33**, 3292-3302, (2005).

28 Sanchez, H., Kanaar, R. & Wyman, C. Molecular recognition of DNA-protein complexes: A straightforward method combining scanning force and fluorescence microscopy. *Ultramicroscopy* **110**, 844-851, (2010).

29 van der Heijden, T. & Dekker, C. Monte Carlo Simulations of Protein Assembly, Disassembly, and Linear Motion on DNA. *Biophys. J.* **95**, 4560-4569, (2008).

30 Bugreev, D. V. & Mazin, A. V. Ca2+ activates human homologous recombination protein Rad51 by modulating its ATPase activity. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 9988-9993, (2004).

31 van Loenhout, M. T. J., van der Heijden, T., Kanaar, R., Wyman, C. & Dekker, C. Dynamics of RecA filaments on single-stranded DNA. *Nucleic Acids Res.* **37**, 4089-4099, (2009).

32 Cox, M. M. Regulation of bacterial RecA protein function. *Crit. Rev. Biochem. Mol. Biol.* **42**, 41-63, (2007).

33 Drees, J. C., Lusetti, S. L., Chitteni-Pattu, S., Inman, R. B. & Cox, M. M. A RecA filament capping mechanism for RecX protein. *Mol. Cell* **15**, 789-798, (2004).

34 Ragone, S., Maman, J. D., Furnham, N. & Pellegrini, L. Structural basis for inhibition of homologous recombination by the RecX protein. *EMBO J.* **27**, 2259-2269, (2008).

35 Jensen, R. B., Carreira, A. & Kowalczykowski, S. C. Purified human BRCA2 stimulates RAD51-mediated recombination. *Nature* **467**, 678-U662, (2010).

36 Liu, J., Doty, T., Gibson, B. & Heyer, W. D. Human BRCA2 protein promotes RAD51 filament formation on RPA-covered single-stranded DNA. *Nat. Struct. Mol. Biol.* **17**, 1260-1262, (2010).

37 Thorslund, T. *et al.* The breast cancer tumor suppressor BRCA2 promotes the specific targeting of RAD51 to single-stranded DNA. *Nat. Struct. Mol. Biol.* **17**, 1263-1265, (2010).

38 Carreira, A. *et al.* The BRC Repeats of BRCA2 Modulate the DNA-Binding Selectivity of RAD51. *Cell* **136**, 1032-1043, (2009).

39 Shivji, M. K. K. *et al.* A region of human BRCA2 containing multiple BRC repeats promotes RAD51-mediated strand exchange. *Nucleic Acids Res.* **34**, 4000-4011, (2006).

40 Shivji, M. K. K. *et al.* The BRC repeats of human BRCA2 differentially regulate RAD51 binding on single- versus double-stranded DNA to stimulate strand exchange. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 13254-13259, (2009).

41 Pellegrini, L. *et al.* Insights into DNA recombination from the structure of a RAD51-BRCA2 complex. *Nature* **420**, 287-293, (2002).

42 Shin, D. S. *et al.* Full-length archaeal Rad51 structure and mutants: mechanisms for RAD51 assembly and control by BRCA2. *EMBO J.* **22**, 4566-4576, (2003).

43 Graneli, A., Yeykal, C. C., Prasad, T. K. & Greene, E. C. Organized arrays of individual DNA molecules tethered to supported lipid bilayers. *Langmuir* **22**, 292-299, (2006).

44 van der Linden, E., Sanchez, H., Kinoshita, E., Kanaar, R. & Wyman, C. RAD50 and NBS1 form a stable complex functional in DNA binding and tethering. *Nucleic Acids Res.* **37**, 1580-1588, (2009).

45 Strick, T. R., Allemand, J. F., Bensimon, D. & Croquette, V. Behavior of supercoiled DNA. *Biophys. J.* **74**, 2016-2028, (1998).

Supplementary information

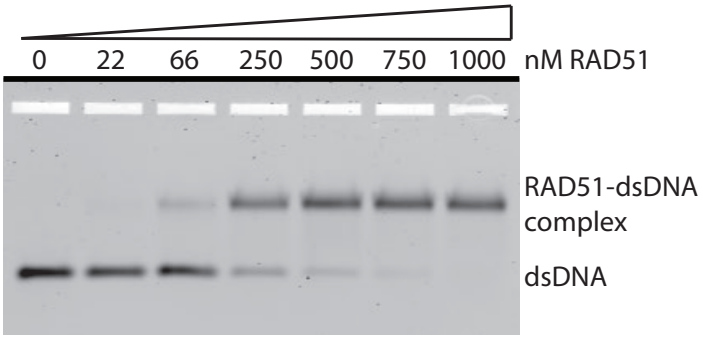


Figure 8.S1: RAD51 interaction with dsDNA. Electrophoretic mobility shift assays (EMSA) were preformed with fluorescently end-labeled dsDNA (66 nM bp) and RAD51 at the indicated concentrations, using 0.8% agarose gels to separate unbound from bound DNA. Reaction mixtures were in Ca^{2+}/ATP conditions that suppress RAD51 ATPase activity. RAD51-dsDNA complex formation is dependent RAD51 concentration. A complete shift is obtained at 1000 nM RAD51.

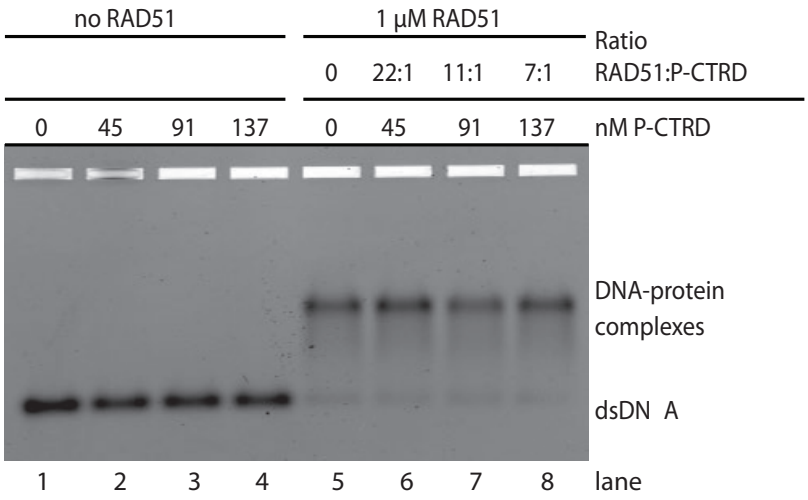


Figure 8.S2: P-CTRD interaction with RAD51 filaments. Electrophoretic mobility shift assays (EMSA) were preformed with fluorescently end-labeled dsDNA and RAD51 (1 μM), P-CTRD or a combination of both using agarose gels to separate unbound from bound DNA. Reaction mixtures contained the indicated concentrations of the P-CTRD. Lanes 6-8 show that the P-CTRD does not interact with the RAD51-dsDNA complexes nor with dsDNA (lanes 2-4).

CHAPTER 9

Mechanism of homology recognition in DNA recombination from dual-molecule experiments

In *E. coli* homologous recombination, a filament of RecA protein formed on DNA searches and pairs a homologous sequence within a second DNA molecule with remarkable speed and fidelity. Here, we directly probe the strength of the two-molecule interactions involved in homology search and recognition using dual-molecule manipulation, combining magnetic and optical tweezers. We find that the filament's secondary DNA-binding site interacts with a single strand of the incoming double-stranded DNA during homology sampling. Recognition requires opening of the helix and is strongly promoted by unwinding torsional stress. Recognition is achieved upon binding of both strands of the incoming dsDNA to each of two ssDNA-binding sites in the filament. The data indicate a physical picture for homology recognition in which the fidelity of the search process is governed by the distance between the DNA binding sites.

This chapter has been published as: Iwijn De Vlaminck, Marijn T.J. van Loenhout, Ludovit Zweifel, Johan den Blanken, Koen Hooning, Susanne Hage, Jacob Kerssemakers, and Cees Dekker. Mechanism of Homology Recognition in DNA Recombination from Dual-Molecule Experiments. *Molecular Cell* 46, 616 (2012).

Introduction

RecA from *E. coli* is the prototype of a family of recombinases essential in double-stranded (ds)DNA break repair and recombination[1]. RecA forms a nucleoprotein filament on a single-stranded (ss)DNA that searches and pairs a homologous sequence within another, double-stranded DNA molecule[2]. The search for homology conducted by the RecA filament poses a formidable challenge from a kinetic and thermodynamic point of view[3]. The RecA filament is able to detect a short homologous sequence of DNA embedded in genomic-length dsDNA[4]. The search process is completed within the timescale set by the cell's life cycle, implying an impressive $>10^3\ s^{-1}$ base-sampling frequency[5]. Homology search is little affected by a large background of heterologous DNA[6], suggesting that the filament is somehow able to cope with the abundantly present short sequences in the genome that display heterology or partial homology

With these characteristics, the RecA homology search is an example of a general class of target-localization processes commonly encountered in molecular biology in which a recognizing molecule finds a specific target among many look-alikes in a noisy background[7,8]. Recently, it was proposed that conformational proofreading is used as a general strategy in target localization[7]. Here, a conformational mismatch between the target-bound and unbound states improves the selectivity of the process by necessitating the recognizing molecule to access an energetically unfavorable intermediate state during proofreading. For the RecA homology search, this raises questions about how the structure of the minimally stable target-bound state is defined and which intermediate states are transiently accessible during proofreading.

Proofreading or homology sampling occurs via Watson-Crick-type base pairing of the bases of the incoming dsDNA with the bases of the ssDNA in the primary binding site at the center of the filament[9,10]. The filament has a secondary DNA-binding site (SBS) that mediates the homology sampling reaction (see inset in Figure 9.1A). Two different mechanistic models for the function of the SBS during the homology sampling reaction have been proposed in which the SBS binds the incoming dsDNA in structurally distinct intermediate states. The first model assumes that the SBS destabilizes the incoming dsDNA upon binding of dsDNA into an extended and underwound conformation, thereby facilitating base sampling[8,11-15]. A second model assumes that the SBS binds to one of the strands of the dsDNA, leaving the second strand available for base sampling[9,10,16]. Here, it is not clear whether the filament has an active role in dsDNA helix opening or whether homology recognition relies on intrinsic DNA-breathing dynamics[10,16]. Homology sampling and recognition constitute difficult-to-isolate intermediate steps in a complex pathway that includes RecA filament formation, initial homologous alignment, strand exchange and strand displacement[1,2]. Complexes formed at sites of heterology and early recognition products are inherently short-lived and therefore difficult to investigate[17]. Consequently, the structure of the minimally stable recognition product and the role of the SBS in homology sampling have proven challenging to resolve.

Here, we isolate and study binding interactions relevant in homology recognition using a dual-molecule manipulation technique (Figure 9.1). A combination of magnetic-tweezers and optical-tweezers-based single-molecule manipulation allows bringing two distinct molecules into local contact to investigate the strength and frequency of occurrence of binding interactions. We directly probe and compare interactions of RecA filaments formed on ssDNA or dsDNA with ssDNA and (non) supercoiled dsDNA, using the four different experimental configurations depicted in Figure 9.1B (panels I–IV). We present measurements of the probability and strength of intermolecular binding as well as measurements of friction arising between filaments and DNA during sliding of one molecule across the other.

We find that the affinity of the SBS for dsDNA is too weak to account for local stretching or unwinding of dsDNA during homology sampling. We confirm that the SBS has a strong preference for ssDNA[18], but we find that its affinity for ssDNA is too weak to stably trap ssDNA bubbles in dsDNA. We furthermore find homologous pairing is strongly promoted by negative supercoiling of the incoming dsDNA. All in all, the data indicate a model for recognition in which the SBS of the filament binds to a single strand of the incoming dsDNA during homology recognition. When both strands of the incoming dsDNA bind to each of the two independent ssDNA-binding sites in the filament, the SBS and the sequence-specific ssDNA in the core of the filament, a stable joint molecule can be formed and homology recognition can be achieved. We propose that the fidelity of the recognition process is governed by the distance between both DNA-binding sites.

Results

Dual-molecule manipulation assay

The dual-molecule molecule technique introduced here is a combination of dual-bead optical tweezers and magnetic tweezers. Figure 9.1C shows a schematic of the setup and its main components (see Supplemental Information for a more detailed description). Two independently steerable optical traps are generated in the sample. A DNA molecule or RecA filament is end-attached between two surface-functionalized polystyrene beads held in the optical traps. A stretching force can be applied to this molecule through independent movement of the traps. Simultaneous movement of the traps allows 3D-manipulation of this molecule in the proximity of a second molecule that is tethered in between the bottom of a flow cell and a paramagnetic bead in a magnetic tweezers configuration. Positioning and rotation of an external magnet pair allows stretching and coiling of this second molecule. The magnetic tweezers thus permit studying the effects of twist and torque on the intermolecular binding. The position of five beads, two optically trapped beads, the paramagnetic bead and two fiducial markers, are measured in real time during the experiments using video-microscopy (50–100 Hz, 0.5 nm accuracy per frame). The distances between the molecules are accordingly deduced. Intermolecular binding forces are

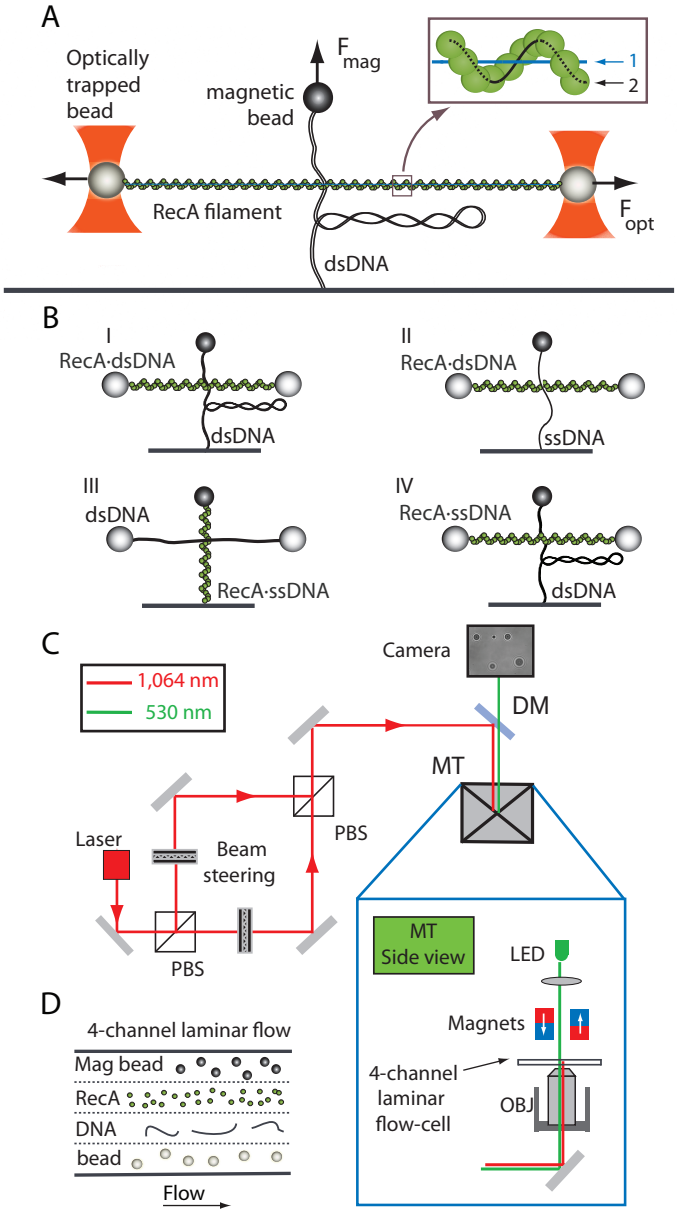


Figure 9.1: Experimental approach. (**A**) Side view of dual-molecule assay where a RecA filament that is held in the optical tweezers interacts with a coiled dsDNA that is tethered in a magnetic-tweezers configuration. Inset: schematic of the RecA filament represented with two DNA-binding sites, a primary site (1) and secondary DNA binding site (2). (**B**) The

four different experimental configurations I-IV used in this work. (**C**) Schematic outline of the setup indicating the main components (for a more detailed description see Fig. S9.1). Two independently steerable optical traps are generated using a 1,064 nm laser system. The beams are split and recombined using a polarizing beam splitter (PBS) and focused using a high numerical aperture objective (OBJ) in the volume of a 4-channel laminar flow cell. Positioning and rotation of an external magnet pair allows stretching and coiling of a molecule tethered in the magnetic tweezers configuration. (**D**) Top view of the 4-channel laminar flow-system used for the assembly of the molecular constructs. See also Figure S9.1.

determined on the basis of lateral deflections of the superparamagnetic bead (see Experimental Procedures). The superparamagnetic bead of the magnetic tweezers serves as a sensitive force probe, with a force-resolution that is only limited by the thermal force noise acting on the bead (~10fN/√Hz; see Supplemental Information). The force probe furthermore has ideal force-clamp characteristics[19]. A four-channel laminar-flow cell[20,21] was used for the independent, stepwise assembly of DNA molecules and RecA filaments tethered in the magnetic tweezers and optical tweezers, prior to interaction experiments. The flow system allows interaction experiments to take place under buffer conditions that can be chosen independently from the DNA and filament assembly buffers (see Supplemental Information). RecA assembly reactions were followed to completion and the mechanical properties of the various single-molecule constructs were characterized prior to interaction experiments (see Figures S9.1D– S9.1G).

Strength of binding of the secondary binding site to ssDNA and dsDNA

We first examine the interactions of dsDNA or ssDNA molecules tethered in the magnetic tweezers with a RecA·dsDNA filament held in the optical tweezers (configuration I and II in Figure 9.1B). Since the bases of the double stranded DNA in the primary site within the RecA·dsDNA filament are inaccessible, interactions of incoming DNA with the filament are restricted to the SBS. The probability and strength of binding of dsDNA and ssDNA to the SBS are first investigated using a push-probe experimental approach where a contact between both molecules is first established by pushing the filament against the DNA and then disrupted by withdrawing to probe the strength of the intermolecular bond (see Figure 9.2A, and Movie S1).

We observe a strong binding of ssDNA to the SBS. The data in Figure 9.2B, for example, show a force of 2.9 pN required to disrupt the intermolecular bond. The strength of binding of ssDNA to the SBS of a RecA·dsDNA filament was quantified through an analysis of the statistical distribution of bond-rupture forces[22,23]. The rupture-force distribution is a function of the force ramp-rate, dF/dt (Figure S9.2). Figure 9.2D shows a rupture-force probability histogram, $p_F(F)$, acquired from multiple ($n = 84$) push-probe measurements as in Figure 9.2B, at a fixed force-ramp speed, $dF/dt = 25$ pN/s. $p_F(F)$ can be transformed into the force dependence of the bond

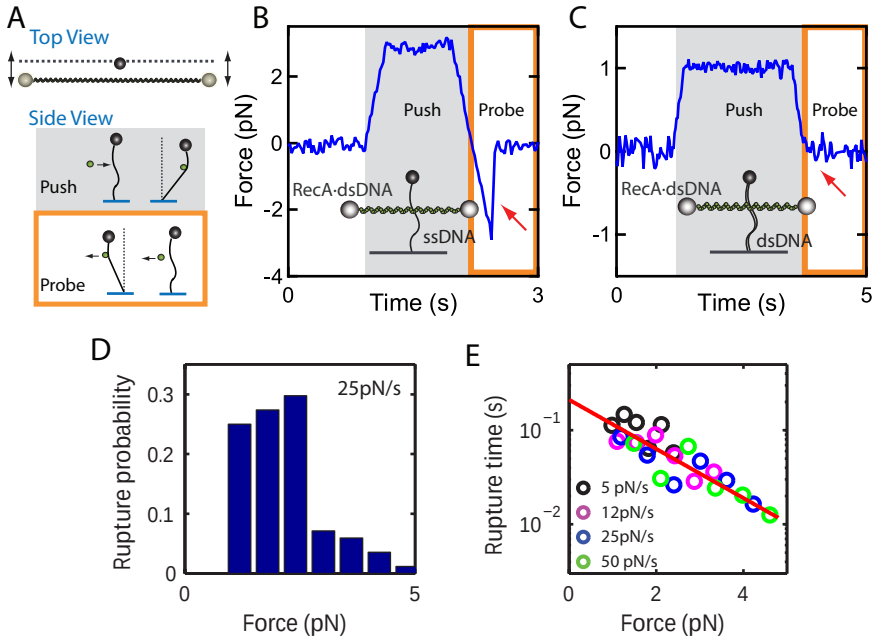


Figure 9.2: The secondary binding site binds ssDNA but does not bind dsDNA. (**A**) Schematic shows a top view and a side view (along the filament) of the push and probe phases of a push-probe experiment. Forces were obtained from the deflection of the magnetic bead (see SI). (**B**) Push-probe experiments reveal binding of ssDNA to the secondary binding site of a RecA·dsDNA filament. A force of 2.4 pN is required to disrupt the intermolecular bond (red arrow). F_{mag} = 15 pN. (**C**) Absence of binding of dsDNA (red arrow). F_{mag} = 2 pN. (D) Rupture-force histogram from multiple push-probe measurements (n = 84) as in panel B (force ramp rate, dF/dt = 25 pN/s). (**E**) Force-dependence of the bond lifetime of ssDNA bound to the SBS. The bond lifetime was obtained by converting data from rupture-force histograms measured at different ramp speeds using Equation 9.1. The red line is a fit of Bell's formula to the data, $\tau(F) = \tau_o \exp(-Fx_b/kT)$, where τ_o is the zero-force bond lifetime, x_b is a distance to the transition state, k_b is the Botzmann constant and T is the temperature, τ_o =0.2 s and x_b = 2.5 nm. See also Figure S9.2.

lifetime τ(F) using[22]: $\tau(F) = \int_F^\infty \left(p(f) / \left((dF/dt)\, p(F) \right) \right) df$. The data for $\tau(F)$ acquired at different dF/dt (5 pN/s to 50 pN/s) collapse onto a single master curve (Figure 9.2E) indicating that the rupture-force kinetics measurable at constant force behave as a single exponential[22]. Figure 9.2E shows a fit to the data of Bell's formula, $\tau(F)=\tau_o \exp(-Fx_b/k_bT)$, where τ_o is the zero-force bond lifetime, x_b is a distance to the transition state, k_b is the Boltzmann constant and T is the temperature. (τ_o = 0.2 s, x_b

= 2.5 nm). To directly compare the free energy of binding of ssDNA to the SBS to the free energy of formation of a B-form DNA duplex, we have furthermore investigated interactions of the SBS with dsDNA as function of the degree of mechanical underwinding (see below).

In contrast to these strong ssDNA-SBS binding interactions, the dsDNA-SBS experiments show no sign of dsDNA binding to the SBS within the force and time resolution of our instrument (Figures 9.2C and S9.2A). This was confirmed in more than 400 interaction experiments with applied forces on the magnetic bead, F_{mag}, in the range F_{mag} = 0.6 - 3 pN and a supercoil density, σ, applied to the dsDNA in the range σ = -0.05 to +0.02. (Supercoil density is defined as $\sigma = (L - L_o)/L_o$, where L_o and L are the linking numbers of the relaxed and supercoiled DNA, respectively[19]. The data confirm that the SBS has a strong preference for ssDNA over dsDNA[18,24]. Given the force resolution of the dual-molecule technique used here and the maximum forces measured for ssDNA-SBS interactions (>5 pN; see Figure S9.4), we determine that ssDNA-SBS interaction forces are at least 2 orders of magnitude stronger than dsDNA-SBS interaction forces.

To test whether these weak dsDNA-SBS interactions are a consequence of the presence of dsDNA in the primary site of the filament, interactions were also probed between dsDNA and RecA filaments formed on ssDNA. Figure 9.3B shows measurements of intermolecular friction that arises during sliding of a dsDNA molecule back and forth across a RecA·ssDNA filament that is held in the magnetic tweezers (configuration III; see schematic in Figure 9.3A and Movie S2). Before sliding, a contact between both molecules was established by pushing the dsDNA against the RecA filament (Figure S9.4A). The data in Figure 9.3B show that the force measured during sliding remains essentially unchanged to the level of the thermal force noise as measured before the start of the sliding movement. These data indicate that dsDNA-SBS interactions are indeed weak, independent of whether ssDNA or dsDNA is present in the primary site. Pushprobe measurements with molecules tethered in the same configuration further confirm this observation (see Figure S9.3C). The previous experiments were carried out in the presence of ATP-γ-s, a poorly hydrolysable analog of ATP. In additional experiments, the effect of RecA-driven ATP-hydrolysis was explored. No evidence of dsDNA-filament interactions was found, independent of whether or not ATP-hydrolysis can occur (Figure S9.3A). Lastly, our torque-sensitive dual-molecule experiments do not provide evidence for extensive (>1–2 turns) local unwinding of dsDNA at a nonspecific dsDNA-filament contact point (Figure S9.3B), such as was suggested on the basis of bulk experiments[14,25].

SBS-ssDNA interactions are too weak to allow stable trapping of a ssDNA bubble within dsDNA

To directly compare the free energy of binding of ssDNA to the SBS, ΔG_{SBS}, to the free energy of formation of a B-form DNA duplex, ΔG_B, we studied interactions of underwound dsDNA with the SBS of a RecA filament (configuration I). To this end, a dsDNA molecule tethered in the magnetic tweezers configuration is mechanically

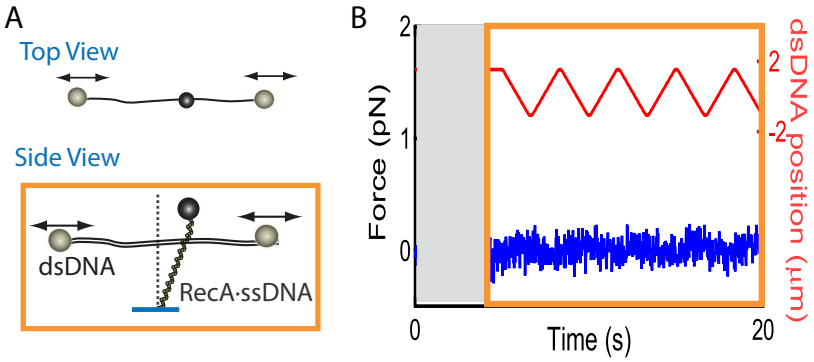


Figure 9.3: Weak interactions between dsDNA and RecA·ssDNA. (**A**) Schematic of side view and top view of the assay for measuring intermolecular friction. Force is measured on the magnetic bead while sliding of a dsDNA across a RecA·ssDNA filament. (**B**) Force obtained from the deflection of the magnetic bead (blue trace) and position of the center of the dsDNA tethered in dual-bead optical tweezers before and during the sliding motion (red trace). The force trace indicates friction-less sliding. F_{mag} = 15 pN, *Fopt* = 20 pN. See also Figure S9.3.

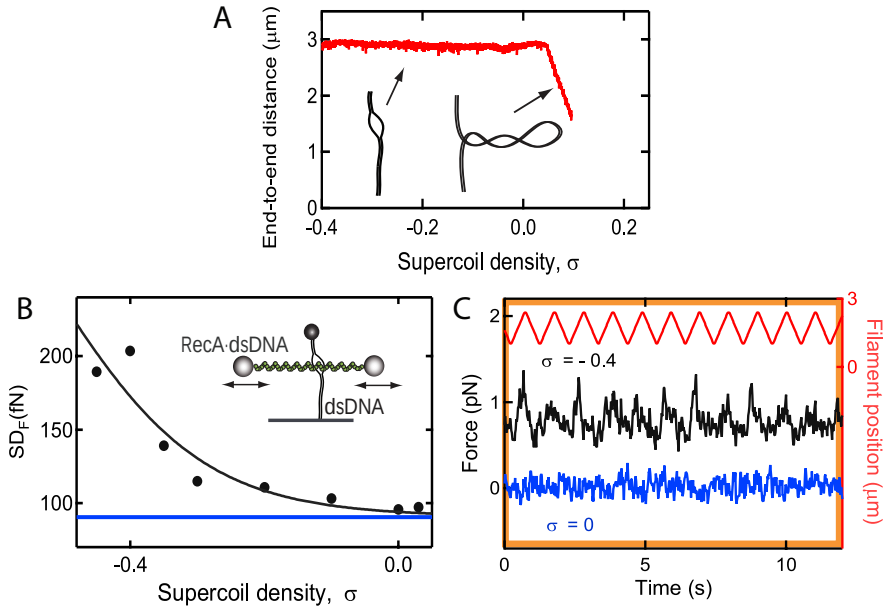


Figure 9.4: The affinity of the SBS for ssDNA is too weak to allow stable trapping of ssDNA bubbles. (**A**) End-to-end distance vs. supercoil density for dsDNA (F_{mag} =3.5 pN). The curve is highly asymmetric with weak coiling-number dependence of the end-to-end dis-

tance for $\sigma < 0$. For $\sigma > 0$ and beyond a critical supercoil density, plectonemes are formed upon coiling, leading to a continuously decreasing end-to-end distance with further coiling. For $\sigma < 0$, formation of plectonemes is energetically unfavorable, and coiling leads to mechanical denaturation of the dsDNA[19]. (**B**) Measurement of friction during sliding (cf. schematic Figure 9.3A) of a RecA·dsDNA filament across a mechanically underwound dsDNA (inset in B). Blue trace: friction-less sliding for non-supercoiled dsDNA ($\sigma = 0$). Black trace: friction appears for strongly underwound dsDNA ($\sigma = -0.40$ offset for clarity). $F_{mag} = 2.5$ pN. (**C**) Standard deviation (*SD*) of the measured force, SD_F, (bandwidth 50 Hz) as a function of supercoil density compared to the expected thermal force noise, SD_t, acting on a 1 micron bead: $SD_t = 91$ fN (bandwidth 50 Hz, blue line).

coiled while stretched with a stretching force $F_{mag} > 1$ pN. Figure 9.4A shows a measurement of the end-to-end distance of a dsDNA as a function of supercoil density at $F_{mag} = 3.5$ pN. As reported previously[26], the curve is highly asymmetric with virtually no coiling dependence of the end-to-end distance for $\sigma < 0$. For $\sigma > 0$, plectonemes are formed upon coiling, leading to a continuously decreasing end-to-end distance with further coiling. The reason that such plectonemes are not formed for $\sigma < 0$, is that mechanical denaturation of the dsDNA is energetically more favorable[19,27] have provided evidence for DNA denaturation in undertwisted dsDNA through incubation of the DNA with reagents specific for unpaired bases. It was furthermore shown that segments of ssDNA are accessible for binding of single-stranded DNA-binding proteins[28] in mechanically stretched ($F_{mag} > 0.5$ pN) and underwound dsDNA.

Here, we probe binding of transiently exposed ssDNA segments to the SBS of the filament as function of the degree of uncoiling. Figure 9.4B shows data of friction during sliding of a RecA· dsDNA ($\sigma = 0$), and across strongly underwound dsDNA ($\sigma = -0.4$, $F_{mag} = 2.5$ pN). At $\sigma = 0$, frictionless sliding is observed, i.e., the force measured on the paramagnetic bead equals the thermal noise force. The data for $\sigma = -0.4$, however, display stick-slip behavior during sliding, indicative of intermittent intermolecular (un)binding. Figure S9.4D,E show zoom-ins of stick-slip events. The standard deviation of the measured force, SD_F, as plotted in Figure 9.4C as a function of σ, increases beyond the thermal noise force for strong underwinding ($\sigma < -0.2$), indicating increasingly stronger intermolecular interactions at higher levels of underwinding. Figure 9.5C (red points) shows the probability of binding of underwound dsDNA to the SBS of a RecA filament as measured using the push-probe experimental mode. The probability of binding increases with increasing levels of underwinding ($\sigma < -0.4$). Intermolecular binding at high levels of underwinding can be explained by binding of the SBS to ssDNA that is transiently accessible upon local denaturation of the dsDNA.

An analysis of the statistical distribution of rupture forces, as in Figures 9.2 and S9.2, shows that the strength of binding interactions with the SBS are very similar for ssDNA and underwound dsDNA (Figure S9.4G). This observation further corroborates the interpretation of intermolecular binding during interactions between mechanically underwound dsDNA and a RecA filament as binding of a ssDNA segment that is transiently exposed upon local denaturation of the dsDNA.

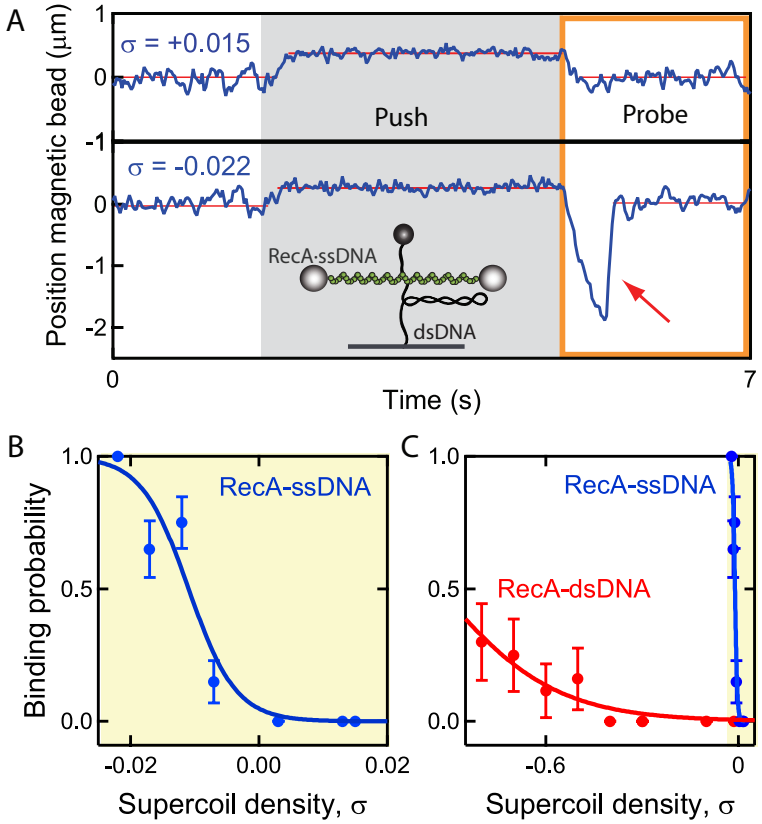


Figure 9.5: Homologous pairing is strongly enhanced by negative supercoiling. (**A**) Push-probe experiments (see schematic Figure 9.2A) reveal binding (red arrow) of a RecA·ssDNA filament to homologous, negatively supercoiled dsDNA (bottom) and absence of binding to positively supercoiled dsDNA (top). F_{mag} = 0.6 pN, F_{opt} = 20 pN. (**B**) Binding probability as function of supercoil density. Underwinding of dsDNA strongly promotes the efficiency of pairing. Number of events at each σ, n = 20. (**C**) Comparison of binding probability as function of supercoil density for RecA·dsDNA (red, F_{mag}=2.4 pN, F_{opt}=20 pN, n=10) and RecA·ssDNA filaments (blue, data from panel b). The error bars are calculated as the standard error for a binomial distribution $(P(1-P)/n)^{½}$, where P is the probability of joint-molecule formation. See also Figure S9.4.

The requirement for very high levels of underwinding (>20%) indicates that at more moderate levels of negative supercoiling, binding of ssDNA to the SBS is outcompeted by rebinding to the nearby complementary ssDNA in the denaturation bubble. This shows that ssDNA binding to the SBS is weaker than binding of ssDNA to complementary ssDNA, i.e. $|\Delta G_B| > |\Delta G_{SBS}|$. In other words, the affinity of the SBS for ssDNA is in itself too weak to stably trap ssDNA bubbles in dsDNA.

Homology recognition probability is strongly enhanced by DNA unwinding

To address a potential role for intrinsic DNA-breathing dynamics during base sampling[16], we tested whether the probability of homology recognition is enhanced by negative supercoiling of the dsDNA, which is known to strongly promote the frequency of occurrence and lifetime of DNA-breathing bubbles[29,30]. For these experiments, RecA is assembled on a 20 kb ssDNA that is formed upon mechanical overstretching of a dsDNA molecule (see Figure S9.5) in the dual-bead optical tweezers (configuration IV), leaving the possibility to twist a homologous dsDNA tethered in the magnetic tweezers. The experimental results (traces in Figure 9.5A) show a remarkably strong dependence of intermolecular binding on the supercoil density of the dsDNA. At σ = +0.015, no intermolecular binding is observed in a push-probe experiment, whereas at σ = -0.022 a stable joint molecule is formed. Figure 9.5B shows the probability of joint-molecule formation as a function of σ. It is clear from these data that slight underwinding of the dsDNA strongly stimulates joint-molecule formation, i.e. binding is not observed at positive supercoiling ($\sigma > 0$) whereas 100% probability of binding is observed for $\sigma <$ -0.015 ($n = 20$). For a RecA·dsDNA filament, where base-pairing interactions with the DNA in the primary site are excluded and homologous pairing cannot take place, binding events are only observed at much higher levels of negative supercoiling, $\sigma <$ -0.4 (red data in Figure 9.5C). The strong dependence of the recognition probability on negative supercoiling points to a model for recognition in which an early step involves the spontaneous breathing dynamics of the donor duplex (Figure 9.6A)[25].

Discussion

Transient intermediates and stable recognition product

The combined dual-molecule data provide insight into conformations of intermediates of the search and recognition reactions and into the energetics involved in the formation of a stable product of homology recognition. Our measurements showed that ssDNA-SBS interaction forces are at least two orders of magnitude stronger than dsDNA-SBS interaction forces. Since we also find that the affinity of the SBS for ssDNA is weaker than the free energy related to the formation of a B-form DNA duplex, $|\Delta G_B| < |\Delta G_{SBS}|$, we conclude that the affinity of the SBS for dsDNA is too weak to account for local stretching or underwinding of dsDNA during homology sampling. A mechanistic model such as given in Figure 9.6A is thus preferred, in which the SBS interacts with only a single strand of the incoming dsDNA during homology sampling. Figure 9.6A shows a reaction pathway from an initial state (panel I), via transient intermediates (panels II–III) to the product of the recognition reaction (panel IV). Short-lived intermediates formed in dsDNA-SBS interactions (panel V, Figure 9.6A) are unstable transients that do not lead to stable product formation. Since the RecA filament predominantly collides and interacts with heterologous dsDNA during the

search process, it is beneficial for the speed of the search that these interactions are indeed weak and short-lived[2,15,31].

On the basis of our dual-molecule data, we can make specific estimates about the energetics of the reaction and the conformation of the stable product of recognition. From the structure of complementary ssDNA bound to the ssDNA in the primary site of the filament, it is known that the incoming complementary ssDNA interacts only weakly with the RecA monomers and that the newly formed heteroduplex is stabilized primarily by Watson-Crick-type base pairing[9]. The heteroduplex formed within the primary site of the filament adopts a stretched and underwound conformation, resulting in a free energy of binding, $|\Delta G_{HD}|$, which is weaker than the energy of complementary pairing in a nondistorted B-DNA conformation, i.e., $|\Delta G_{HD}| < |\Delta G_{B}|$. Formation of a stable product thus implies that the loss of binding energy in this less stable conformation, $|\Delta G_{B}| - |\Delta G_{HD}|$, is compensated by the energy of binding of the other strand of the incoming dsDNA to the SBS, i.e., :

$$|\Delta G_{B}| - |\Delta G_{HD}| < |\Delta G_{SBS}|. \qquad (9.1)$$

Another inequality is derived from our dual-molecule interaction experiments between underwound dsDNA and the SBS of the filament. Here, the data indicated that ssDNA binding to the SBS is weaker than binding of ssDNA to complementary ssDNA, i.e., $|\Delta G_{B}| > |\Delta G_{SBS}|$. Since the affinity of the SBS for ssDNA thus is too weak to stably trap ssDNA bubbles in dsDNA, heterologous contacts, where binding is restricted to the SBS, are unstable (panel II, Figure 9.6A). Combining the above deduced inequalities, $|\Delta G_{B}| > |\Delta G_{SBS}|$ and $|\Delta G_{B}| > |\Delta G_{HD}|$, with Equation 9.1, we thus conclude that:

$$|\Delta G_{SBS}| + |\Delta G_{HD}| > |\Delta G_{B}| > |\Delta G_{SBS}|, |\Delta G_{HD}|. \qquad (9.2)$$

Stable binding and product formation thus occur exclusively when both strands of the incoming dsDNA bind to each of the two independent ssDNA-binding sites of the filament, the sequence-specific ssDNA in the core of the filament and the sequence-nonspecific SBS (Figure S9.6).

The inequalities in Equation 9.2 place tight constraints on the possible values of the binding energies of the two ssDNA-binding sites in the filament. By applying the signal-detection formalism developed by Savir and Tlusty[8] to the present model for homology recognition, we furthermore find that the ability of RecA to discriminate homologous and heterologous sequences is optimal when $|\Delta G_{SBS}|$ and $|\Delta G_{HD}|$ are interdependent, viz. $|\Delta G_{SBS}| = |\Delta G_{B}| - 0.5\,|\Delta G_{HD}|$ (see Equation S9.5 and Figure S9.6). The condition of optimal recognition thus further restricts the possible free energy values.

Mechanistic insight into the minimal length required for homologous pairing

A remarkable consequence is derived from the above-described model and the large spatial separation(~25 A°) between the primary binding site at the central axis of the filament and the second ssDNA-binding site of the filament[9]. As a result of this large

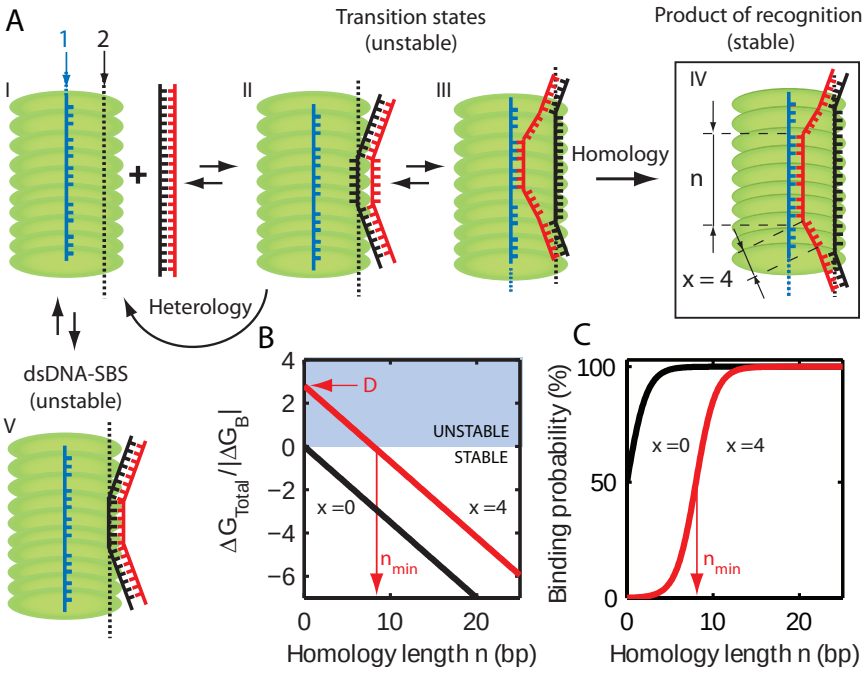


Figure 9.6: Mechanism of homology recognition. (**A**) Mechanistic model for homology search and recognition. The RecA filament (green) is represented with two binding sites (panel I, indicated 1, 2). A stable joint molecule is formed and homology recognition is achieved when both strands of the dsDNA bind to both ssDNA-binding sites in the RecA filament (panel IV). A variety of transition states are accessible along the pathway towards recognition (see examples in panels II-III). Interactions limited to the SBS are unstable and short-lived (panel II). dsDNA-filament interactions are short-lived as well (panel V). For non-homologous interactions, base paring with the primary site cannot take place, and the transition state (II) will collapse back to the initial state (I) (**B**) The reported, large distance between the two binding sites[35] prevents simultaneous binding to both sites at the edges of the pairing region. This geometrical constraint leads to an energy cost, Δ, that needs to be overcome by the gain in energy in the doubly-paired region. This in turn leads to a minimal number of homologous base pairs, n_{min}, required for stable binding to occur. The graph shows the total free energy per bp versus homology length, $\Delta G_{Total} = -(n + 2x)|\Delta G_{SBS}| - n|\Delta G_{HD}| + (n + 2x)|\Delta G_B|$. Introduction of the energy cost Δ shifts the point where $\Delta G_{Total} = 0$ from $n = 0$ to $n = n_{min}$. Parameter values used in the model: $|\Delta G_{HD}| = 0.7|\Delta G_B|$, $|\Delta G_{SBS}| = 0.65|\Delta G_B|$, see SI. (**C**) Binding probability $p(\Delta G_{Total}) = 1/(1+\exp(\Delta G_{Total}/k_bT))$, for ΔG_{Total} in panel B. $|\Delta G_B| = 2.5\ k_bT$, with k_b the Boltzmann constant, and T the temperature. The finite distance x between the binding sites is seen to have a strong influence. See also Figure S9.5.

separation, the strands of the incoming dsDNA can not bind both sites simultaneously over the entire region of pairing. Some bases of one of the strands necessarily remain unpaired at the edges of the pairing region (Figure 9.6A, panel IV). Since binding to only one of the ssDNA-binding sites does not lead to stable pairing (Equation 9.2), these two edge regions give rise to an energy cost, Δ, (see Figure 9.6B) that needs to be compensated for by the gain in binding energy in the central region. This in turn leads to a minimum length of homology, n_{min}, for which stable pairing can occur (Figures 9.6B and 9.6C), which is dependent on the distance between both ssDNA-binding sites. Interestingly, under conditions where binding is optimal in the signal-detection formalism, n_{min} simply reads $n_{min} = 2x$, with x the number of unpaired bases (Equations S9.6 and S9.7 and Figure S9.6). For RecA, Hsieh *et al.*[4] reported a minimal homology length for stable pairing of 8 bp, corresponding to $x = 4$ bp, which fits very well with the reported distance between the postulated location of the SBS and the primary site at the filament axis (~25 A°)[9]. Hsieh *et al.* exploited the inability of a restriction endonuclease to cleave the duplex within a paired region to probe the homology-length dependence of the probability of homologous pairing. The above-described model directly fits the data by Hsieh *et al.* remarkably well (see Figure S9.6H). The analysis furthermore shows that the data by Hsieh *et al.* is not properly described with a model that does not take into account a finite distance between both binding sites in the filament (black line in Figure 9.6C).

Such a minimal homology length, n_{min}, explains the ability of RecA to avoid long-term pairing to short (<8 bp) sequences that at random exhibit partial homology and that are abundantly present in the *E. coli* genome (Figure S9.6). The fidelity of the reaction is thus governed by the physical distance between the SBS and the primary binding site.

dsDNA is the active search entity

The experiments in Figure 9.5 directly show that initial homologous pairing does not require a free end of the filament but can occur at any site along a filament. As a result, the search process can be accelerated by a mechanism of parallel search, where homology is sampled at multiple filament-dsDNA contact sites simultaneously[32]. The model described in Figure 9.6 is akin to a conformational proofreading scheme[7,8] where, interestingly, the dsDNA, and not the RecA filament, is the active, recognizing search entity. A large conformational mismatch exists between the target-bound and unbound states of the dsDNA. The target-bound state is accessed via energetically unfavorable intermediate states, as discussed above. The conformational mismatch improves the selectivity of the recognition reaction.

Dynamics of helix opening during homology sampling

In our model, formation of a stable recognition product requires opening of the helix over a distance longer than $n_{min} + 2x$ (Figure 9.6A, panel IV), and the question remains how such long-range helix opening is achieved during homology sampling. Although

spontaneous formation of large ssDNA bubbles is energetically unfavorable, thermally activated, long-range opening of the dsDNA helix can take place due to the low energy cost related to the extension of a shorter-range intermediate bubble[33,34], in particular for AT-rich sequences[10,29]. ssDNA contacts to one or both of the ssDNA-binding sites in the filament (with free energy of binding values restricted by the inequalities in Equation 9.2) further lower the extension energy, leading to longer bubble sizes and lifetimes (Figure 9.6A, example intermediates in panels II–III)[29].

Negative supercoiling promotes helix breathing, explaining the data that showed that homologous pairing is sensitively dependent on negative coiling of the incoming duplex. The strong sensitivity of homologous pairing on supercoiling provides a means for torsional regulation of the reaction, as was similarly proposed for other DNA-metabolic pathways[34].

Given the structural and functional similarity of bacterial and eukaryotic recombinases[1], we anticipate that our results and conclusions also qualitatively apply to the mechanism of homology search and recognition in eukaryotic systems. The dual-molecule technique that was introduced here provides the possibility to interact distinct DNA substrates with great spatial control and allows sensitively probing the effects of torsional stress on intermolecular binding reactions. The technique will be applicable in the study of a wide range of protein-mediated DNA-DNA interactions.

Experimental procedures

Buffer conditions

All measurements were carried out at 22°C and were performed in a buffer of 20 mM Tris (pH 6.9), 10 mM NaCl, 13 mM MgCl2, 100 mM DTT. Buffers were filtered (0.22 mm Millipore™ GV filter, PVDF membrane). RecA protein was purchased with New England Biolabs. The experiments were performed in the presence of a poorly hydrolysable ATP-analog (ATP-γ-s) unless otherwise indicated. Experiments with RecA-ssDNA filaments were performed with a low concentration of RecA (15 nM) in the background to counteract spontaneous RecA disassembly (under ATP and ATP-γ-s conditions).

Force extraction

Interaction forces in dual-molecule experiments were deduced from the measured in-plane deflection of the magnetic bead along the direction of movement, Δx. Δx is converted into a corresponding rupture force, $F_{x,rupture}$ using:

$$F_{rupture} = \frac{F_{mag}}{h}\Delta x, \tag{9.3}$$

where *h* is the height of the contact point measured from the bottom of the flow cell. The height and position of the bead can be measured with high accuracy, $\delta x = 0.5$ nm, $\delta z \approx 3$ nm per frame. Given the flexibility of the molecular constructs, there is an uncertainty in the height of the contact point of 100-200 nm (5-10%). The paramagnetic beads used in this work varies with 5-10%, force calibrations measurements are subject to a similar level of variability.

Supplementary information

Supplemental Information includes six figures, Supplemental Experimental Procedures, Supplemental References, and two movies and can be found with this article online at doi:10.1016/j.molcel.2012.03.029.

Acknowledgments

We thank Claire Wyman, Roland Kanaar, and Stephen Kowalczykowski for discussions. This work was supported by a DNA-in-action grant from the "Stichting voor Fundamenteel Onderzoek der Materie (FOM)," which is financially supported by the "Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NWO)." Author contributions: I.D.V., M.T.J.v.L., and C.D. designed the experiment. I.D.V., M.T.J.v.L., K.H., J.d.B., and J. K. built the setup; S.H. prepared the DNA constructs; I.D.V. and L.Z. performed the experiments and analyzed data; and I.D.V. and C.D. wrote the manuscript. All authors discussed the results and commented on the manuscript.

References

1 Kowalczykowski, S. C., Dixon, D. A., Eggleston, A. K., Lauder, S. D. & Rehrauer, W. M. Biochemistry of homologous recombination in Escherichia-Coli. *Microbiol. Rev.* **58**, 401-465, (1994).

2 Radding, C. M. Helical Interactions in homologous pairing and strand exchange driven by RecA protein. *J. Biol. Chem.* **266**, 5355-5358, (1991).

3 Barzel, A. & Kupiec, M. Finding a match: how do homologous sequences get together for recombination? *Nat. Rev. Genet.* **9**, 27-37, (2008).

4 Hsieh, P., Cameriniotero, C. S. & Cameriniotero, R. D. The synapsis event in the homologous pairing of DNAs - RecA recognizes and pairs less than one helical repeat of DNA. *Proc. Natl. Acad. Sci. U. S. A.* **89**, 6492-6496, (1992).

5 Camerini-Otero, R. D. & Hsieh, P. Parallel DNA triplexes, homologous recombination, and other homology-dependent DNA interactions. *Cell* **73**, 217-223, (1993).

6 Honigberg, S. M., Rao, B. J. & Radding, C. M. Ability of RecA protein to promote a search for rare sequences in duplex DNA. *Proc. Natl. Acad. Sci. U. S. A.* **83**, 9586-9590, (1986).

7 Savir, Y. & Tlusty, T. Conformational proofreading: The impact of conformational changes on the specificity of molecular recognition. *Plos One* **2**, (2007).

8 Savir, Y. & Tlusty, T. RecA-mediated homology search as a nearly optimal signal detection system. *Mol. Cell* **40**, 388-396, (2010).

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

10 Folta-Stogniew, E., O'Malley, S., Gupta, R., Anderson, K. S. & Radding, C. M. Exchange of DNA base pairs that coincides with recognition of homology promoted by E-coli RecA protein. *Mol. Cell* **15**, 965-975, (2004).

11 Danilowicz, C. *et al.* RecA homology search is promoted by mechanical stress along the scanned duplex DNA. *Nucleic Acids Res.* **40**, 1717-1727, (2012).

12 Dorfman, K. D., Fulconis, R., Dutreix, M. & Viovy, J. L. Model of RecA-mediated homologous recognition. *Phys. Rev. Lett.* **93**, (2004).

13 Mani, A., Braslavsky, I., Arbel-Goren, R. & Stavans, J. Caught in the act: the lifetime of synaptic intermediates during the search for homology on DNA. *Nucleic Acids Res.* **38**, 2036-2043, (2010).

14 Rould, E., Muniyappa, K. & Radding, C. M. Unwinding of heterologous DNA by RecA protein during the search for homologous sequences. *J. Mol. Biol.* **226**, 127-139, (1992).

15 Sagi, D., Tlusty, T. & Stavans, J. High fidelity of RecA-catalyzed recombination: a watchdog of genetic diversity. *Nucleic Acids Res.* **34**, 5021-5031, (2006).

16 Voloshin, O. N. & Camerini-Otero, R. D. Synaptic complex revisited: A homologous recombinase flips and switches bases. *Mol. Cell* **15**, 846-847, (2004).

17 Müller, B., Koller, T. & Stasiak, A. Characterization of the DNA-binding activity of atable RecA-DNA complexes - Interaction between the 2 DNA-binding sites within RecA helical filaments. *J. Mol. Biol.* **212**, 97-112, (1990).

18 Mazin, A. V. & Kowalczykowski, S. C. The function of the secondary DNA-binding site of RecA protein during DNA strand exchange. *EMBO J.* **17**, 1161-1168, (1998).

19 Strick, T. R. *et al.* Stretching of macromolecules and proteins. *Rep. Prog. Phys.* **66**, 1-45, (2003).

20 Noom, M. C., van den Broek, B., van Mameren, J. & Wuite, G. J. L. Visualizing single DNA-bound proteins using DNA as a scanning probe. *Nat. Methods* **4**, 1031-1036, (2007).

21 Wuite, G. J. L., Davenport, R. J., Rappaport, A. & Bustamante, C. An integrated laser trap/flow control video microscope for the study of single biomolecules. *Biophys. J.* **79**, 1155-1167, (2000).

22 Dudko, O. K., Hummer, G. & Szabo, A. Theory, analysis, and interpretation of single-molecule force spectroscopy experiments. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 15755-15760, (2008).

23 Evans, E. Probing the relation between force - Lifetime - and chemistry in single molecular bonds. *Annu. Rev. Biophys. Biomol. Struct.* **30**, 105-128, (2001).

24 Mazin, A. V. & Kowalczykowski, S. C. The specificity of the secondary DNA binding site of RecA protein defines its role in DNA strand exchange. *Proc. Natl. Acad. Sci. U. S. A.* **93**, 10673-10678, (1996).

25 Wong, B. C., Chiu, S. K. & Chow, S. A. The role of negative superhelicity and length of homology in the formation of paranemic joints promoted by RecA protein. *J. Biol. Chem.* **273**, 12120-12127, (1998).

26 Strick, T. R., Allemand, J. F., Bensimon, D., Bensimon, A. & Croquette, V. The elasticity of a single supercoiled DNA molecule. *Science* **271**, 1835-1837, (1996).

27 Allemand, J. F., Bensimon, D., Lavery, R. & Croquette, V. Stretched and overwound DNA forms a Pauling-like structure with exposed bases. *Proc. Natl. Acad. Sci. U. S. A.* **95**, 14152-14157, (1998).

28 De Vlaminck, I. *et al.* Torsional regulation of hRPA-induced unwinding of double-stranded DNA. *Nucleic Acids Res.* **38**, 4133-4142, (2010).

29 Jeon, J. H., Adamcik, J., Dietler, G. & Metzler, R. Supercoiling Induces denaturation bubbles in circular DNA. *Phys. Rev. Lett.* **105**, (2010).

30 Jeon, J. H. & Sung, W. How topological constraints facilitate growth and stability of bubbles in DNA. *Biophys. J.* **95**, 3600-3605, (2008).

31 Yancey-Wrona, J. E. & Camerini-Otero, R. D. The search for DNA homology does not limit stable homologous pairing promoted by RecA protein. *Current biology : CB* **5**, 1149-1158, (1995).

32 Adzuma, K. No sliding during homology search by RecA protein. *J. Biol. Chem.* **273**, 31565-31573, (1998).

33 Altan-Bonnet, G., Libchaber, A. & Krichevsky, O. Bubble dynamics in double-stranded DNA. *Phys. Rev. Lett.* **90**, (2003).

34 Choi, C. H. *et al.* DNA dynamically directs its own transcription initiation. *Nucleic Acids Res.* **32**, 1584-1590, (2004).

35 Chen, Z., Yang, H. & Pavletich, N. P. Mechanism of homologous recombination from the RecA-ssDNA/dsDNA structures. *Nature* **453**, 489-494, (2008).

Supplementary Information

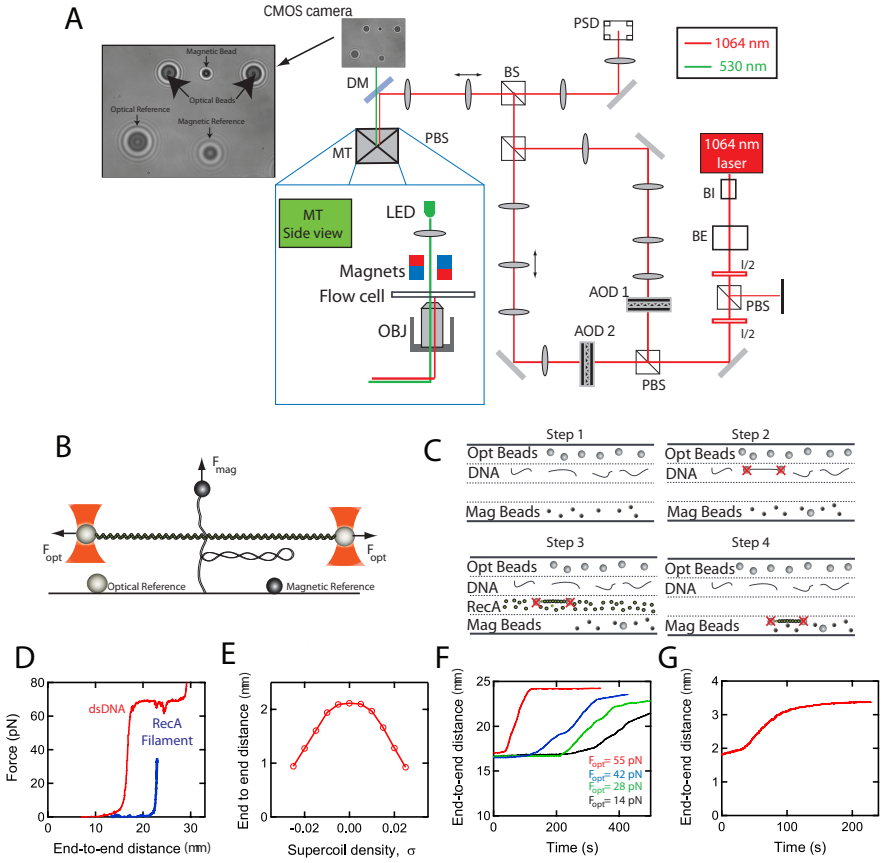


Figure S9.1: (A) Schematic outline of the experimental setup. For a description, see the extended experimental procedures section. **(B)** Schematic of one of four dual-molecule

configurations used in this assay, with F_{opt} the force applied to the filament held in the dual-bead optical tweezers and F_{mag} the force exerted on the DNA substrate tethered in the magnetic tweezers. Two fiducial markers are used, providing a reference for the *xyz*-position of the magnetic and optical beads during the experiments. **(C)** Schematic representation of various substeps in protocol followed for the preparation of DNA substrates and filaments in RecA dsDNA-DNA interaction experiments. Red crosses denote the position of the optical traps. **(D, E)** DNA substrates and filaments are characterized prior to interaction experiments to assure integrity of the substrates and to assure only a single molecule is tethered. Panel **(D)** shows measurements of the mechanical response of a single dsDNA and RecA filament tethered in the dual-bead optical tweezers configuration. The contour length of the filament is a factor 1.45 longer than the contour length of dsDNA[36]. The plateau in the force response of the dsDNA at a stretching force of ~65 pN, corresponds to the overstretching transition, a phase transition during which the molecule can be stretched without further increase in applied force[37,38]. This property is exploited to form a ssDNA tethered in the dual-bead optical tweezers (see also Figure S9.5). The occurrence of a overstretching plateau at a force of ~ 65 pN was used to verify that only a single molecule was tethered in the dual-bead optical tweezers. **(E)** End-to-end distance of a dsDNA molecule tethered in the magnetic tweezers configuration as function of supercoil density. The coiling behavior of dsDNA was used to distinguish single and multiple tethers[39] and to assure the integrity of the molecular substrate. **(F,G)** The assembly reactions of RecA on DNA are recorded in real-time to assure completion of filament assembly. **(F)** End-to-end distance of dsDNA as function of time during filament assembly at constant F_{opt} (F_{opt} = 14 pN, F_{opt} = 28 pN, F_{opt} = 42 pN, F_{opt} = 55 pN). For details, see the extended experimental procedures section. The reaction rate increases for higher F_{opt} as observed previously[40]. **(G)** End-to-end distance of ssDNA tethered in the magnetic tweezers configuration as function of time during filament formation, F_{mag} = 10 pN (for details see the extended experimental procedures section).

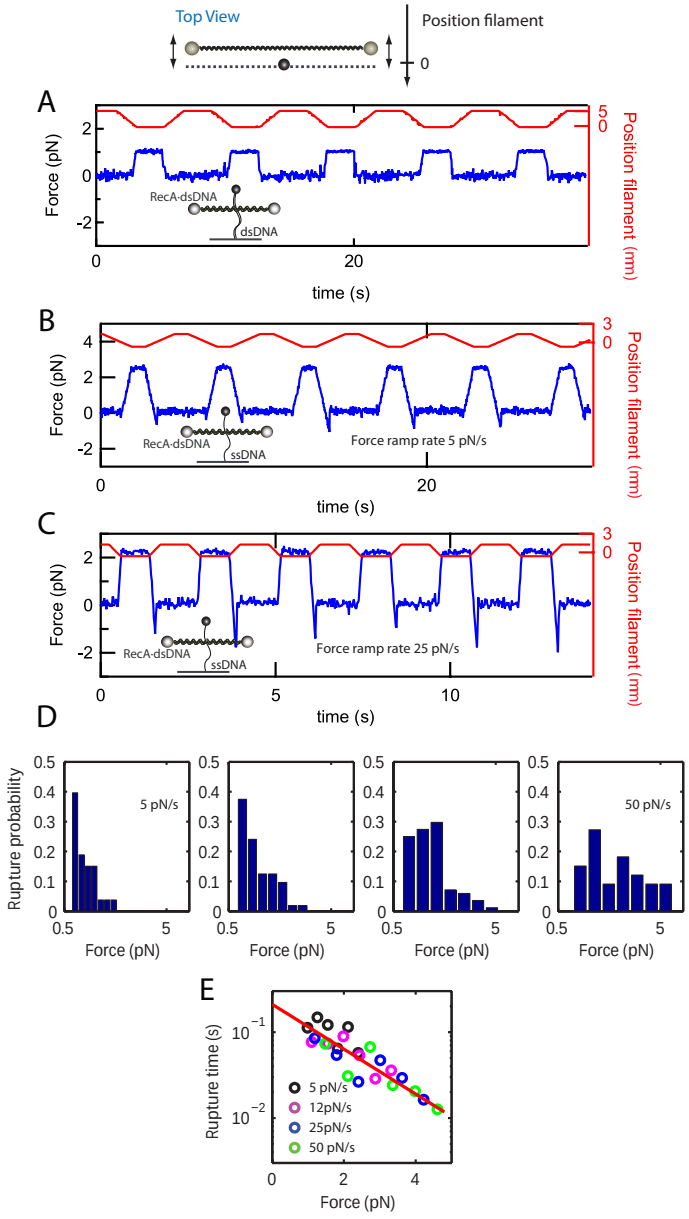


Figure S9.2: (A) No evidence of binding of dsDNA to the SBS of a RecA-dsDNA filament F_{mag} = 2 pN, F_{opt} = 20 pN. The same behavior was observed in more than 400 push-probe trials for applied forces, F_{mag}, in the range F_{mag} = 0.6-3 pN and a supercoil density, σ in the range σ = -0.05 to +0.02. **(B-C)** Consistent binding of ssDNA to the SBS of a RecA-dsDNA filament in a push-probe experiment for different force ramp rates, dF/dt = 5pN/s (panel B) and

25 pN/s (panel C). F_{mag} = 15 pN, F_{opt} = 20 pN. The distribution of rupture force is a function of *dF/dt*. For higher dF/dt, the average rupture force is seen to be higher. A statistical analysis of the distribution of the rupture forces allows quantification of the binding strength[23]. **(D)** Seven-bin histograms of the rupture-force probability, $p_r(F)$, for *dF/dt* = 5 pN/s (n = 53), *dF/dt* = 12 pN/s (n = 104), *dF/dt* = 25 pN/s (n = 84) and *dF/dt* = 50 pN/s (n = 33) acquired from push-probe measurements as in panels B and C. **(E)** The force dependence of the bond lifetime τ(F) measurable at a constant force can be obtained from the rupture force probability histograms in D using Eq. (9.1) in the main text[22]. The data for τ as function of *F* acquired for different *dF/dt* collapse onto a single master curve indicating that the constant-force kinetics behave as a single exponential[22]. The force dependence of the lifetime was fit (red line) using $\tau(F) = \tau_o exp(-Fx_b/kT)$, where τ_o is the zero-force bond lifetime, x_b is a distance to the transition state, k_b is the Boltzmann constant and *T* is the temperature. We extract τ_o = 0.2 s and x_b = 2.5 nm.

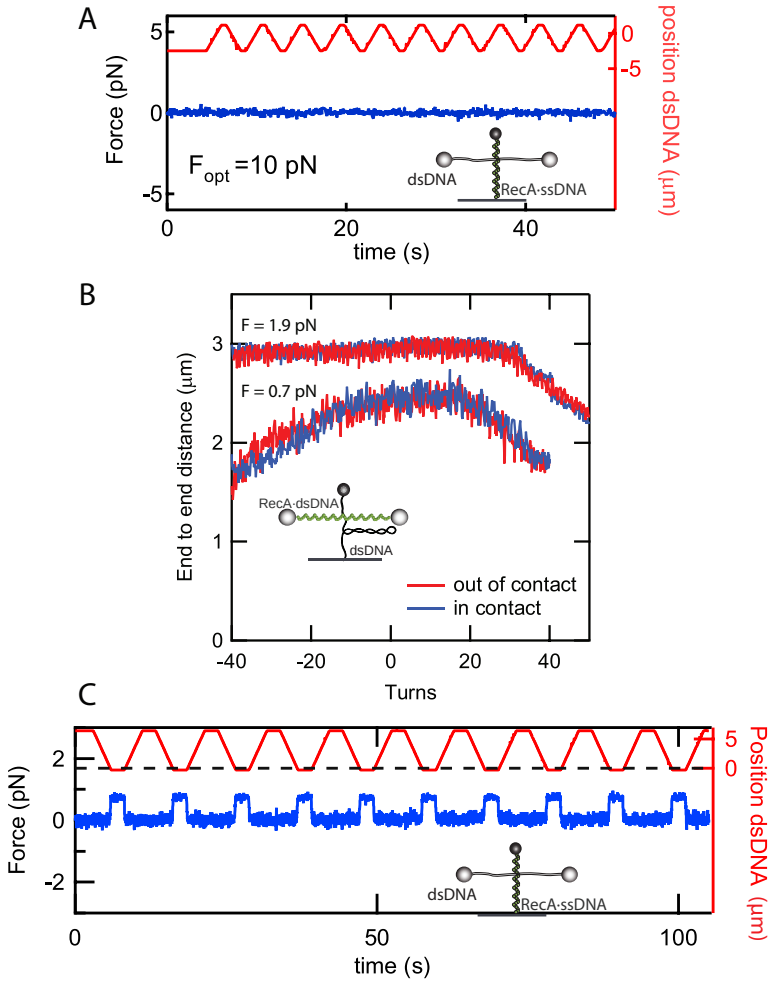


Figure S9.3: (A) Friction-less sliding of a dsDNA (F_{mag}=15 pN) across a RecA-ssDNA filament under conditions of ATP hydrolysis and a stretching force of F_{opt}=10 pN. (B) End-to-end distance of DNA molecules as a function of the number of magnet turns. dsDNA in contact (blue) and out-of-contact (red) with a RecA-dsDNA filament at F_{mag} = 1.9 pN and F_{mag} = 0.7 pN. Since the linking number of the dsDNA is constant in this assay, enzyme-induced changes in twist are compensated by a change in the number of plectonemes stored in the molecule[41]. E.g. local unwinding (overwinding) of one turn of dsDNA gives rise to removal (addition) of one plectoneme. The curves in B overlap at low and high force. Given the resolution in measurements of bead height in this assay (100-150 nm) and the size of a single plectoneme (40 – 60 nm, depending F_{mag}) we conclude that there is no sign of local unwinding greater than 2-3 turns of heterologous dsDNA in contact with a RecA filament. **(C)** No evidence of binding of dsDNA to the SBS of a RecA-ssDNA filament in a push-probe experimental approach, F_{mag} = 20 pN, F_{opt} = 22 pN.

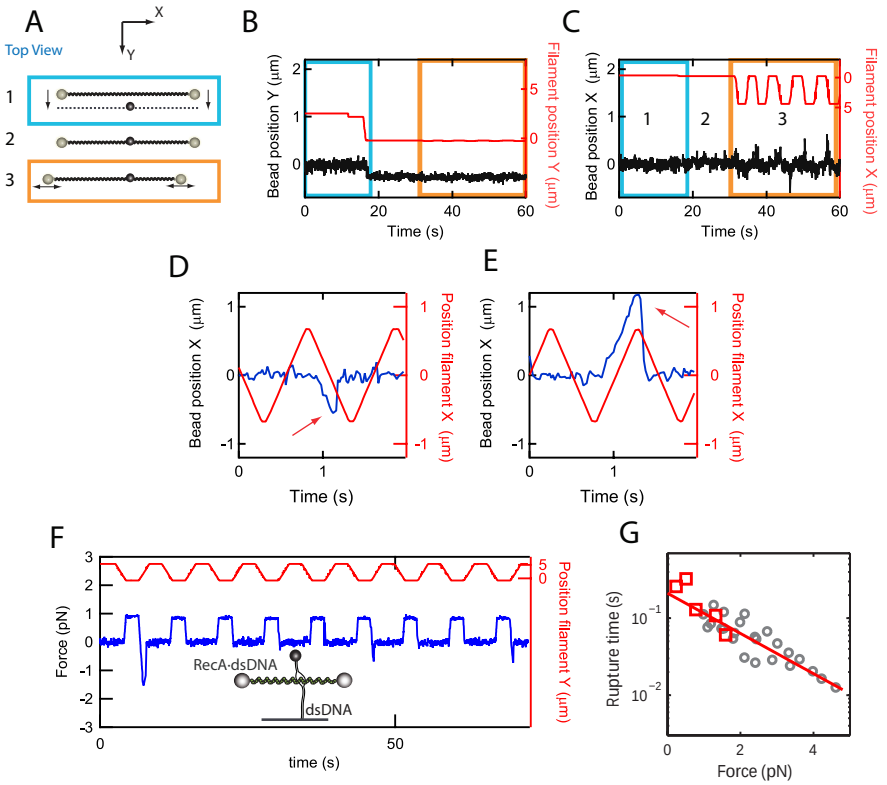


Figure S9.4: (A) Sequence of movements before and during measurements of intermolecular friction during sliding (as in Figure 9.4). The schematic shows the example of sliding of a RecA•dsDNA filament across mechanically underwound dsDNA. The filament held in the dual-bead optical tweezers is initially not in contact with the dsDNA molecule tethered in the magnetic tweezers configuration (panel 1). Contact is established between both molecules by movement of the filament along the y-axis (panel 2). The sliding movement parallel to the *x*-axis is initiated after contact is established (panel 3). **(B-C)** Measured magnetic bead and filament positions before and during sliding of a RecA·dsDNA filament across a mechanically underwound dsDNA ($\sigma = -0.4$) as in Figure 9.4 ($F_{mag} = 2.5$ pN). Contact is confirmed by monitoring the magnetic bead position along the Y direction. Bead movements in the *x*-direction display stick-slip behavior during sliding, indicative of intermittent intermolecular (un-)binding. **(D,E)** Zoom-in of stick-slip behavior. The data show intermolecular binding and subsequent release upon which registry between the drive signal (the moving filament) and the response signal is lost (arrows). **(F)** Push-probe experiment reveals binding of a RecA·dsDNA filament to a mechanically underwound dsDNA ($dF/dt = 3.3$ pN/s, $F_{mag} = 2.4$ pN, $\sigma = -0.5$). **(G)** A force spectroscopy analysis, as in Figure 9.2 and Figure S9.2, was applied to quantify the force-dependence of the lifetime, $\tau(F)$, for the intermolecular bonds formed in this configuration, red squares ($\sigma = -0.5$, $n = 19$). Comparison of $\tau(F)$ measured for interactions between a RecA·dsDNA filament and mechanically underwound

dsDNA with τ(F) values measured for interactions of a RecA·dsDNA filament with ssDNA (grey markers, data from figure S9.2). It is clear from this graph that the bond lifetime and its force dependence measured in these different configurations is very similar. This observation provides further support for the interpretation of the intermolecular binding observed in interactions between mechanically underwound dsDNA and RecA·dsDNA filaments as binding of a ssDNA segment that is transiently exposed upon local denaturation of the dsDNA to the SBS of the filament (panels B-E and Figure 9.4).

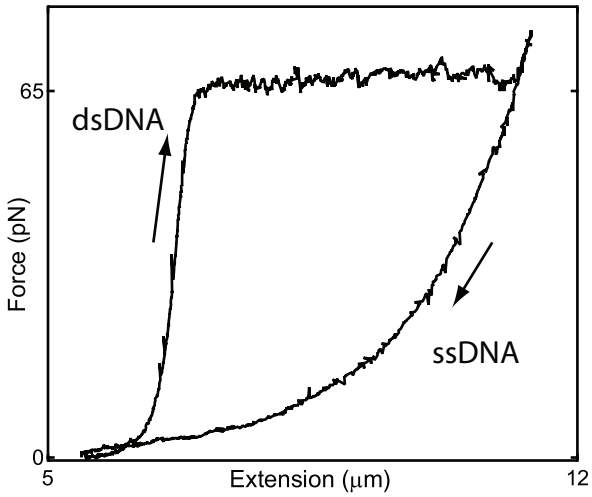


Figure S9.5: Mechanical overstretching of a dsDNA molecule to form ssDNA. At a stretching force of 65 pN, dsDNA undergoes a phase transition referred to as the overstretching transition, during which the molecule can be stretched (length increase with factor of up to 1.7) without a further increase in applied force[37,38]. In a 20 mM MES (pH 6.25), 100 mM KCl. buffer, the overstretching transition is accompanied by a gradual and cooperative transition from dsDNA to ssDNA, as one strand of the DNA unpeels form the other (6). After overstretching, a stable ssDNA molecule remains.

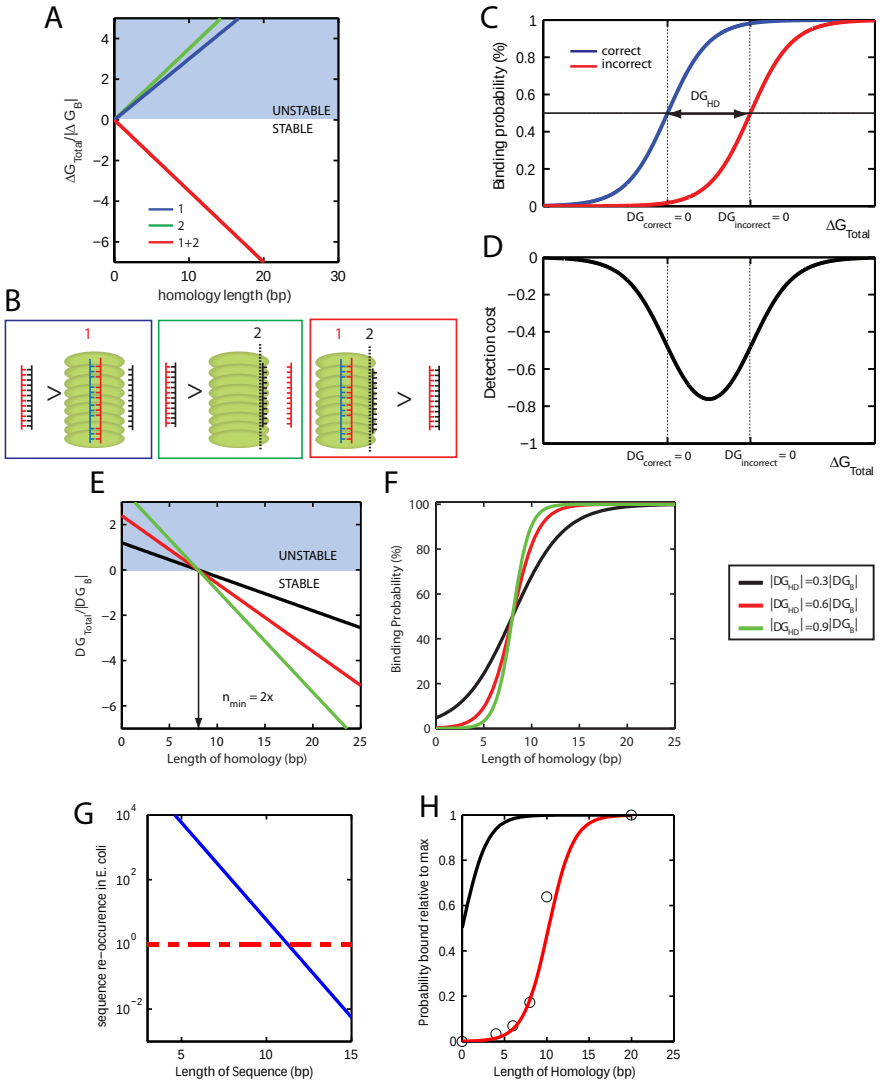


Figure S9.6: (A) and (B) Energetics involved in the model. Total free energy per bp, ΔG_{Total}, relative to the free energy of binding in a B-DNA conformation, ΔG_B, is plotted as function of the number of homologous base pairs, n. $\Delta G_{Total}=-n(|\Delta G_{SBS}|+|\Delta G_{HD}|-|\Delta G_B|)$ (red line), where ΔG_{SBS} is the free energy of binding of ssDNA to the SBS per basepair, and ΔG_{HD} is the free energy of binding per basepair of ssDNA to the ssDNA in the primary site of the filament. Also plotted are the free energies related to binding to only one of both single stranded DNA binding sites: for binding to only the single stranded DNA in the primary site: $\Delta G_{Total} = -n(|\Delta G_{HD}|-|\Delta G_B|)$, (blue line), and for binding to only the SBS: $\Delta G_{Total}=-n(|\Delta G_{SBS}|-|\Delta G_B|)$ (green line)). A stable bond is only formed when both ssDNA binding sites are accessible ($|\Delta G_{SBS}| <$

$|\Delta G_B|$, $|\Delta G_{HD}| < |\Delta G_B|$ but $|\Delta G_{SBS}| + |\Delta G_{HD}| > |\Delta G_B|$). In the graph, we chose $|\Delta G_{HD}| = 0.7|\Delta G_B|$ and assumed optimal conditions of binding, i.e. $|\Delta G_{SBS}| = |\Delta G_B| - |\Delta G_{HD}|/2 = 0.65|\Delta G_B|$ (see figure S9.10). **(C)** Binding probability, $p = 1/(1+\exp(-|\Delta G_{Total}|/k_bT)$, as function of total binding free energy, ΔG_{Total}, with $|\Delta G_{Total}| = |\Delta G_{SBS}|+|\Delta G_{HD}|-|\Delta G_B|$, upon binding a correct sequence and incorrect sequence. **(D)** The detection cost[8], $C = p_{incorrect} - p_{correct}$, is minimal for $|\Delta G_{Total}| = |\Delta G_{HD}|/2$, corresponding to $|\Delta G_{SBS}| = |\Delta G_B| - |\Delta G_{HD}|/2$. **(E)** Total free energy per bp relative to the free energy of binding in a B-form DNA conformation, $\Delta G_{total}/|\Delta G_B|$, as a function of the length of homology in our model. A distance between both binding sites $x = 4$ bp was assumed, consistent with structural data[9]. Under conditions of optimal binding, the minimal length for stable pairing ($n_{min} = 2x$) is *not* a function of ΔG_{HD}. $\Delta G_{total}/|\Delta G_B|$ was plotted for different values of $|\Delta G_{HD}| =$ 0.3, 0.6 and $0.9|\Delta G_B|$, all yielding $n_{min} = 8$. **(F)** Binding probabilities, $p(\Delta G_{Total}) = 1/(1+\exp(\Delta G_{Total}/k_bT))$ for ΔG_{total} in the left panel. $|\Delta G_B| = 2.5\ k_bT$, with k_b the Boltzmann constant, and T the temperature. **(G)** Average number of re-occurrences of a random sequence in the *E. coli* genome as a function of the length of the random sequence, calculated as $6*10^6/(4)^L$, with L the length of the sequence (blue line). The graph illustrates the importance of a mechanism during the search for homology that allows the RecA filament to avoid short, partially homologous sequences. Sequences longer than 11 bp are unique. Red line indicates an average sequence occurrence equal to 1. **(H)** Fit of the above-described model to data on the homology-length dependence of homologous pairing acquired by Hsieh *et al.*[4]. These authors monitored the homology-length dependence of the efficiency of homologous pairing, exploiting the inability of a restriction endonuclease to cleave the duplex within a paired region. The red line is a fit of the above-described model to the data considering $|\Delta G_{HD}| = 0.6|\Delta G_B|$ and optimal binding conditions. We find a best fit for $|\Delta G_B| = 2.26\ k_bT$ per bp, in line with literature values for the free energy of B-form DNA duplex formation[42], and $x =$ 5 bp, corresponding to a minimal homology length of 10bp, slightly larger than the 8 bp mentioned by Hsieh *et al.*[4]. The black line shows the binding probability considering $x = 0$ bp and demonstrates that a model does not take into account a finite distance between both binding sites does not describe the data well.

Protocols for preparation DNA substrates

Optical tweezers construct 1: 48 kB dsDNA (configuration I, configuration II, configuration III, Figure S9.4)

The starting material was 35 µl lambda DNA (645 ng/µl, Promega). 4 µl bio-dCTP (0,4 mM, Invitrogen), and 4 µl bio-dATP (0,4 mM, Invitrogen), 0,5 µl dGTP (10mM, Promega), 0,5 µl dTTP (10mM, Promega), 4 µl Klenow (-exo 5U/µl, NEB), and 7 µl Neb2 buf (10X, NEB) were added to 18 µl Milli-Q water (MQ) and the solution was subsequently incubated for 2 h at 37°C. To purify the construct, the ligation was phenol extracted (phase lock tubes (VWR)) and ethanol precipitated (2 ½ V of 100% EtOH, and 1/10 V of 3M NaAc)). The pellet was dissolved in 10 mM Tris (pH 8.0), 1mM EDTA, 1% ethanol.

Optical tweezers construct 2: 20 kB dsDNA (configuration IV)

Preparation started with a PCR reaction on a lambda DNA template where the following PCR primers (Biolegio) were used: forward (fw) primer5': CGTGCGAACTCTAGATGAATTTCTGAAAGAGTTACCCCTCTAAGTAATGAGG, reverse (rev) primer 5':

T_bCT$_b$GGAAT$_b$TGGGCAGAAGAAAACTGTCGATGCAGCCAAAATTTGTGGCGG (T_b in rev primer = biotin label). PCR mix; 10 µl HercII buf (5x, Stratagene), 5 µl fw primer (10 µM), 5 µl rev primer (10 µM), 2 µl dNTP's (10 mM, Promega), 1 µl lambda DNA (50 ng/ µl, Promega), 1 µl Herc II fusion (Stratagene), 2,5 µl DMSO (Stratagene), 23,5 µl MQ. The PCR product was purified with nucleospin extract II kit (Machery Nagel). The fragment was subsequently digested with XbaI restriction enzyme (20 U/µl, NEB) and incubated for 1h at 37°C. After digestion, the length of the fragment was 20083 bp. The fragment was subsequently purified with a G25 column (GE Healthcare) and used in a Klenow fill-in reaction (Mixture: 80 µl DNA, 4 µl Klenow (-exo 5 U/ul, NEB), 0,2 µl dGTP (10 mM, Promega), 2 µl bio-dUTP (1 nmol/µl, Roche), 5 ul bio-dATP (0,4 mM Invitrogen), 5 µl bio-dCTP (0,4mM Invitrogen) 11 µl NEB2 buffer (NEB), 2,8 µl MQ). The mixture was incubated for 1 h at 37°C followed by a heat-inactivation step (20 min at 75°C). To purify the construct, the ligation was phenol extracted and ethanol precipitated. The pellet was dissolved in 10 mM Tris (pH 8.0), 1mM EDTA, 1% ethanol.

dsDNA Magnetic tweezers construct 1 (configuration IV)

Preparation started with a fresh miniprep isolation (Quikpure kit, Machery Nagel) of plasmid GL120 (in-house collection). This is a cloning vector with a 5 kb lambda fragment cloned-in with in-fusion recombination (Clontech): fw primer: 5' ACACTCATCTTGGATTCTCGGACGAGTGTTCAGTAATGAACC, rev primer 5' GAATCGGGTTCTCGATTACGTCCGTCACGTTCACGCATCAGG. The 5 kb lambda fragment was homologous to the 20 kb optical tweezers construct 2. The plasmid GL120 was cut with XbaI, BamHI, and AflII restriction enzymes (length of fragment: 10118 bp). The digestion product was purified with the nucleospin extract II kit (machery Nagel). Assembly of digoxigenin ends: PCR fragment (1238 bp). PCR template pbluescriptIISK+. Forward primer 5'GACCGAGATAGGGTTGAGTG 3', Reverse primer 5' CAGGGTCGGAACAGGAGAGC 3'. The PCR reactions were performed with a taq polymerase (PCR core system I, Promega). Standard PCR reactions were performed, except 2 µl of digoxigenin-11-2'-deoxy-uridine-5'- triphosphate (dig-dUTP, Roche) is added. The PCR fragment was digested with BamHI (605 – 633 bp).

Biotinylated: PCR fragment (1238 bp). PCR template pbluescriptIISK+. Forward primer 5'GACCGAGATAGGGTTGAGTG 3', reverse primer 5' CAGGGTCGGAACAGGAGAGC 3'. The PCR reactions were performed with a taq polymerase (PCR core system I, Promega). Standard PCR reactions were performed, except 2 µl of Biotin-16-2'deoxy-uridine-5'- triphosphate (bio-dUTP 50 nmol, Roche) is added. PCR fragment was digested with XbaI (617 – 621 bp). The digestions were purified with the nucleospin extraction kit (Machery Nagel). The dig-labeled and bio-labeled ends were subsequently ligated to the 10118 bp fragment (XbaI, BamHI overhang). The ligation was performed with T4 DNA ligase (NEB). The labeled ends were added in 10 molar excess. To purify the construct the ligation was phenol extracted and ethanol precipitated. The dsDNA can be mechanically separated to form ssDNA using the procedure described in Figure S9.6.

dsDNA Magnetic tweezers construct 2 (configuration I)
Preparation started with a fresh miniprep isolation (Quikpure kit, Machery Nagel) of plasmid GL003 (in-house collection). Plasmid GL003 was vector pBluescriptIISK+ (Stratagene) + 3 different lambda pieces cloned-in. The plasmid was cut with XbaI and SacI (1954 – 10162). Subsequently, the 10162 bp fragment was gel extracted with a nucleospin extract II kit (machery Nagel). Digoxigenin ends: PCR fragment (1238 bp). PCR template pbluescriptIISK+. Forward primer 5'GACCGAGATAGGGTTGAGTG 3', Reverse primer 5' CAGGGTCGGAACAGGAGAGC 3'. The PCR reactions were performed with a taq polymerase (PCR core system I, Promega). Standard PCR reactions were performed, except 2 µl of digoxigenin-11-2'-deoxy-uridine-5'- triphosphate (dig-dUTP, Roche) was added. PCR fragment is digested with SacI (593 – 645 bp). Biotinylated: PCR fragment (1238 bp). PCR template was pbluescriptIISK+. Forward primer: 5'GACCGAGATAGGGTTGAGTG 3', Reverse primer 5' CAGGGTCGGAACAGGAGAGC 3'. The PCR reactions were performed with a taq polymerase. Standard PCR reactions were performed, except 2 µl of Biotin-16-2'deoxy-uridine-5'- triphosphate (bio-dUTP 50 nmol, Roche) is added. The PCR fragment was digested with XbaI restriction enzyme (617 and 621 bp). Both digestions of the labeled ends were purified with the nucleospin extraction kit. The 10162 bp fragment (XbaI, SacI overhang) was ligated to the dig-labeled end (SacI overhang) and bio-labeled end (XbaI overhang). The ligation was performed with T4 DNA ligase. The labeled ends were added in 10 molar excess. To purify the construct the ligation is phenol extracted and ethanol precipitated. The pellet was dissolved in 10 mM Tris (pH 8.0), 1mM EDTA, 1% ethanol.

ssDNA Magnetic tweezers construct 3 (configuration II)
The preparation started with a PCR reaction of a 7329 kb fragment (template DNA: lambda) using following primers: reverse primer 5' DIG AACTCAGCTCACCGTCGAACA, rorward primer 5' BIO GACGCAGGGGACCTGCAG. The PCR mix is: 10 µl HercII buf (5x, Stratagene), 1 µl rev primer (10 uM), 1 µl fw primer (10 uM), 1 µl dNTP's (10 mM, Promega), 2 µl lambda DNA (10 ng/ µl, Promega), 0,5 µl Herc II fusion, 34,5 µl MQ. A PCR reaction of digoxigenin-labeled end (1238 bp) followed (PCR template pbluescriptIISK+). forward primer: 5'GACCGAGATAGGGTTGAGTG 3', reverse primer 5' CAGGGTCGGAACAGGAGAGC 3'. The PCR reactions were performed with a taq polymerase (PCR core system I, Promega). Standard PCR reactions were performed, except 2 µl of digoxigenin-11-2'-deoxy-uridine-5'- triphosphate (dig-dUTP, Roche) was added. Both PCR products were purified with nucleospin extract II kit (Machery Nagel). Both fragments were cut with PspOMI (20 U/ µl, NEB). 7329 bp fragment x PspOMI, fragment loses the 5' dig label (final length: 7279 bp). Dig end: x PspOMI with 2 fragments after digestion (lengths: 545 and 693 bp). The digestions were purified with a nucleospin extract II kit (Machery Nagel). The 7279 bp fragment was ligated to the dig ends (PspOMI overhang). The ligation was performed with T4 DNA ligase. The labeled ends were added in 10 molar excess. To purify the construct the ligation was phenol extracted and ethanol precipitated. The pellet is dissolved in 10 mM Tris (pH

8.0), 1mM EDTA, 1% ethanol. The dsDNA was converted in a ssDNA using the chemical denaturation method described below.

ssDNA Magnetic tweezers construct 4 (Supplementary Figure 9.5)
The preparation start with a PCR reaction of a 11940 kb fragment (template: lambda DNA) using following primers: forward primer 5' BIO CTCATGCTCACAGTCTGAGCGGTTCAACAGG, reverse primer 5' AACGCTTCACTCGAGGCGTTTTTCGTTATGTATAAATAAGGAGCACACC. The PCR mix was; 10 µl HercII buf (5x, Stratagene), 1 µl Forward primer (10 uM), 1 µl rev primer (10 uM), 2 µl dNTP's (10 mM, Promega), 1 µl lambda DNA (50 ng/ul, Promega), 1 µl Herc II fusion (Stratagene), 34 µl MQ. A PCR reaction of a digoxigenin-labeled fragment follows (1238 bp). PCR template was pbluescriptIISK+: forward primer 5'GACCGAGATAGGGTTGAGTG 3', reverse primer 5' CAGGGTCGGAACAGGAGAGC 3'. The PCR reactions were performed with a taq polymerase. Standard PCR reactions were performed, except 2 µl of digoxigenin-11-2'-deoxy-uridine-5'- triphosphate (dig-dUTP, Roche) was added. Both PCR products were purified with nucleospin extract II kit (Machery Nagel). Both fragments were cut with XhoI restriction enzyme leading to 2 fragments after digestion (554 and 684 bp). The digestions were purified with a nucleospin extract II kit. The 11926 bp fragment was ligated to the dig-labeled end fragments (XhoI overhang). The ligation was performed with T4 DNA ligase. The labeled ends were added in 10 molar excess. To purify the construct the ligation was phenol extracted and ethanol precipitated. The pellet was dissolved in 10 mM Tris (pH 8.0), 1mM EDTA, 1% ethanol. The dsDNA was converted in a ssDNA using the chemical denaturation method described below.

Biophysical Preparation of Molecular constructs
Four different dual-molecule experimental configurations were used in this study. A detailed description of the biochemical and biophysical preparation of these configurations is presented here.

We refer to Fig. S9.1C for a schematic of the various substeps in the assembly protocol of a RecA·dsDNA filament tethered in dual-bead optical tweezers. Prior to interaction experiments, RecA filaments and DNA substrates were characterized through measurement of the force-extension response, or in the case of coilable dsDNA molecules, by means of a measurement of the rotation response of the molecules. For examples of such measurements see Fig. S9.1D and E.

Preparation of a RecA-dsDNA filament tethered in dual-bead optical tweezers
Streptavidin-coated polystyrene beads are captured in two independent optical traps in a bead-containing laminar-flow channel. The polystyrene beads are automatically selected for size using video-microscopy and moved into a laminar-flow channel that contains dsDNA. The DNA was biotinylated at both ends and tethered between the two polystyrene beads. The DNA construct was moved into a separate laminar-flow channel and the force-distance characteristics were analyzed and compared to the

expected behavior for a single tethered molecule. Subsequently, the construct is held in a RecA-containing buffer, and RecA assembly is performed under constant-force conditions. The buffer conditions for RecA assembly were: 20 mM Tris (pH 6.9), 50 mM NaCl, 1 mM $MgCl_2$, 100 mM DTT, 2.6 μM RecA, 0.1 mM ATP-γ-s.

Preparation of a RecA-ssDNA filament tethered in dual-bead optical tweezers
First, a similar procedure is used for the preparation of a RecA-dsDNA filament. Here, a 20 kb dsDNA was tethered in between the optical trapped beads. Before RecA binding, however, the dsDNA is converted in ssDNA by mechanical separation of one of the strands of the dsDNA in a following buffer: 20 mM MES, pH 6.25, 100 mM KCl (see Figure S9.5). This yields a ssDNA on which subsequently RecA is assembled in the following RecA assembly buffer: 20 mM Tris (pH 6.9), 50 mM NaCl, 1 mM $MgCl_2$, 100 mM DTT, 2.6 μM RecA, 0.1 mM ATP-γ-s.

Preparation of a dsDNA tethered in magnetic tweezers
Streptavidin-coated, superparamagnetic beads were incubated with dsDNA. The DNA was biotin-labeled at one end and digoxigenin-labeled on the other end. After DNA-binding, the bead-DNA constructs were incubated with biotinylated BSA to passivate the surface of the superparamagnetic beads and to prevent spurious DNA-binding during experiments. The bottom of the flow-cell was functionalized with anti-digoxigenin and passivated with BSA. After surface immobilization of the bead-DNA assemblies, the DNA molecule was rotationally constrained.

Preparation of ssDNA tethered in magnetic tweezers
A dsDNA was bound to superparamagnetic beads (bead size 1.0 μm) using the same procedure as described above. After DNA incubation, the dsDNA is chemically denatured to form ssDNA in NaOH (100 mM). The beads are subsequently washed to remove the free ssDNA. The surface of the superparamagnetic beads was passivated with biotinylated BSA.

Preparation of a RecA-ssDNA filament tethered in magnetic tweezers
After surface immobilization of a 12 kb ssDNA (following the above outlined procedure), RecA filaments were formed on ssDNA after assembly of RecA in following buffer: 20 mM Tris (pH 6.9), 50 mM NaCl, 1 mM $MgCl_2$, 100 mM DTT, 2.6 μM RecA, 0.1 mM ATP-γ-s. Experiments with RecA-ssDNA filaments were performed with a low concentration of RecA (15 nM) in the background to counteract spontaneous RecA disassembly (under ATP and ATP-γ-s conditions).

Thermal noise force and force resolution
The Langevin force noise, δf_n, acting on the supermagnetic bead with radius, R, is calculated using:

$$\delta f_n = \sqrt{4k_bT(6\pi\eta R)\Delta f}\,, \tag{S9.1}$$

where η is the viscosity of water (10^{-3} Pa.s), Δf is the measurement bandwidth, and k_b Boltzmann's constant. For a magnetic bead with R = 0.5 μm as force probe, the force resolution of the dual molecule interaction technique used in this work is ~13 fN/√Hz.

Optimal cost of detection

In a recent paper by Savir *et al.*, the problem of homology recognition in recombination was formulated as a signal-detection problem. This analysis allowed these authors to extract the binding energy that optimizes the capability of RecA to discriminate homologous and heterologous sequences[8]. Here we apply a similar analysis to the model for recognition in which ssDNA binds to the SBS of the RecA filament in the early phase of recognition. We extract the optimal free energy of binding of ssDNA to the SBS that minimizes the detection cost, C, which is defined as:

$$C = p_{incorrect} - p_{correct}. \tag{S9.2}$$

Where $p_{correct}$ is the probability of binding to a fully homologous sequence and $p_{incorrect}$ is the probability of binding to a fully heterologous sequence:

$$p_{correct} = \frac{1}{(1+\exp(-|\Delta G_{correct}|/k_bT))}, \tag{S9.3}$$

$$p_{incorrect} = \frac{1}{(1+\exp(-|\Delta G_{incorrect}|/k_bT))}. \tag{S9.4}$$

Here, $\Delta G_{correct}$ and $\Delta G_{incorrect}$ are the free energies of binding associated with binding a homologous or heterologous DNA sequence respectively. From our model (see main text) we derive: $|\Delta G_{correct}| = |\Delta G_{HD}| + |\Delta G_{SBS}| - |\Delta G_B|$ *and,* $|\Delta G_{incorrect}| = |\Delta G_{SBS}| - |\Delta G_B|$. The detection cost is minimized for $|\Delta G_{optimal}| = |\Delta G_{HD}|/2 = |\Delta G_{HD}| + |\Delta G_{SBS}| - |\Delta G_B|$ (see Figure S9.6 C,D). We therefore conclude that the optimal free energy of binding of ssDNA to the SBS, $\Delta G_{SBS,optimal}$, is related to ΔG_{HD} and ΔG_B:

$$|\Delta G_{SBS,optimal}| = |\Delta G_B| - \frac{1}{2}|\Delta G_{HD}|. \tag{S9.5}$$

Minimum length of homology

The reported, large distance between both ssDNA binding sites in the filament[9] prevents simultaneous binding of both single DNA strands of the incoming dsDNA at the edges of the pairing region. We define x as the length of unpaired region at a single edge, see Figure 9.6A, main text. The unpaired region leads to an energy deficit of

$2x(|\Delta G_{SBS}| - |\Delta G_B|)$. This deficit needs to be overcome by n bases in the fully paired region, with a corresponding energy gain: $|\Delta G_{SBS}| + |\Delta G_{HD}| - |\Delta G_B|$ per base pair. The consequence is a mechanical constraint on the minimum length of base pairing, n_{min}, for which stable pairing can occur (total free energy of binding, $\Delta G_{Total} = 0$).

$$n_{\min} = 2x \frac{\left(|\Delta G_B| - |\Delta G_{SBS}|\right)}{\left(|\Delta G_{SBS}| + |\Delta G_{HD}| - |\Delta G_B|\right)}. \quad \text{(S9.6)}$$

In case of the optimal binding condition, n_{min} simply reduces to (using equation S9.5):

$$n_{\min} = 2x. \quad \text{(S9.7)}$$

Figure S9.6F shows the free energy of binding and the binding probability for different values of $|\Delta G_{HD}| = 0.3, 0.6$ and $0.9|\Delta G_B|$.

Supplementary references

36 Galletto, R., Amitani, I., Baskin, R. J. & Kowalczykowski, S. C. Direct observation of individual RecA filaments assembling on single DNA molecules. *Nature* **443**, 875-878, (2006).

37 Smith, S. B., Cui, Y. J. & Bustamante, C. Overstretching B-DNA: The elastic response of individual double-stranded and single-stranded DNA molecules. *Science* **271**, 795-799, (1996).

38 van Mameren, J. *et al.* Unraveling the structure of DNA during overstretching by using multicolor, single-molecule fluorescence imaging. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 18231-18236, (2009).

39 Charvin, G., Vologodskii, A., Bensimon, D. & Croquette, V. Braiding DNA: experiments, simulations, and models. *Biophys. J.* **88**, 4124-4136, (2005).

40 Leger, J. F., Robert, J., Bourdieu, L., Chatenay, D. & Marko, J. F. RecA binding to a single double-stranded DNA molecule: a possible role of DNA conformational fluctuations. *Proc. Natl. Acad. Sci. U. S. A.* **95**, 12295-12299, (1998).

41 van der Heijden, T. *et al.* Homologous recombination in real time: DNA strand exchange by RecA. *Mol. Cell* **30**, 530-538, (2008).

42 SantaLucia, J., Jr. A unified view of polymer, dumbbell, and oligonucleotide DNA nearest-neighbor thermodynamics. *Proc. Natl. Acad. Sci. U. S. A.* **95**, 1460-1465, (1998).

Summary

This thesis describes a series of single-molecule experiments aimed at understanding the physical properties of DNA itself and the proteins that interact with it, for example during DNA double-strand-break repair. DNA is long a serpentine-like molecule made up of two chains of nucleic acids that form a double helix. The sequence of nucleic acids encodes the genetic information necessary for all living organisms to function and propagate. The helical rope-like structure of DNA lies at the basis of many cellular processes as cells need to organize their DNA, duplicate it, repair it, and regulate the torsion that is generated when the double helix is wound or unwound.

To understand the physical mechanisms of these processes, we developed and applied sensitive techniques that allowed us to directly probe the conformation and interactions of individual DNA molecules and proteins. By attaching small micrometer-sized beads to DNA molecules, we can manipulate and exert forces on them using either magnetic tweezers, that can rotate and pull on a bead with a magnet, or optical tweezers which use a highly focused laser beam to manipulate the beads in three dimensions. Although magnetic and optical tweezers are powerful tools to probe molecular mechanisms and interactions, they provide limited information on molecular structure and conformation. To directly image the structure of proteins and DNA we employed two additional techniques. Fluorescence microscopy was used to visualize the dynamics and conformation of twisted DNA molecules, and atomic force microscopy (AFM), which uses a sharp scanning probe similar to a gramophone needle, was used to image surface-immobilized DNA and proteins. Below, I provide a short summary of the four parts of the thesis which address: experimental tools, and then twist, knots, and breaks in DNA.

In **Part 0** of this thesis an improved tracking algorithm is presented that allows for increased throughput in single-molecule experiments. In contrast to most bulk techniques, single-molecule measurements are not limited by ensemble averaging, and this allows for improved understanding molecular mechanisms and their dynamics. But measuring only a single molecule is not possible to understand the distributions and fluctuations of molecular properties. It is therefore essential to develop techniques that efficiently measure many individual molecules.

In **Chapter 2** a technique is presented, which greatly increases the throughput in magnetic tweezers measurements. This three-dimensional tracking routine for non-diffraction-limited particles has a substantially increased resolution compared to previous methods, especially at low magnification or at high signal-to-noise ratio. The described algorithm enables tracking with nanometer accuracy in a wide field of view, thereby enabling an increase in the number of beads, and thus the number of molecules, that can be observed. Magnetic tweezers measure the length of individual DNA molecules by tethering them between a magnetic bead and the surface of a flow cell. A pair of external magnets allows to twist and exert a force on the DNA tethers. Video microscopy is used to extract the position of the beads in three dimensions with sub-pixel resolution from digital images. The pixilation of the image introduces a bias in the tracked bead position that becomes large at low magnifications. By using interpolation of the image on a circular grid, we were able to improve the tracking resolution by up to an order of magnitude in three dimensions compared to traditional tracking methods. The increased performance of the proposed algorithm uniquely enables it to extract accurate data for the stiffness (persistence length) and end-to-end distance of more than 100 DNA tethers in a single experiment.

In **part 1** experiments are described that directly reveal the dynamics and conformation of a twisted DNA molecule. Magnetic tweezers can twist a DNA molecule by rotating the magnets that exert a force on the magnetic-bead tether. This additional twisting of the DNA double helix can lead to the formation of extended intertwined loops called plectonemes. In living cells DNA is also 'supercoiled' in this way as the proteins that unwind the DNA and move along its helical path generate torsion and twist. Control of supercoiling and supercoil dynamics is vital to many cellular processes. While the static equilibrium conformation of supercoiled DNA has been extensively studied, its dynamical behavior has remained elusive.

In **Chapter 3** experiments are described that visualize the dynamics of individual plectonemes in supercoiled 21-kilobase-pair (kb) long DNA molecules using magnetic tweezers and fluorescence microscopy. We observe that multiple plectonemes can be present and that their number depends on applied stretching force and ionic strength. Plectonemes are found to move along DNA by either slow diffusion or, unexpectedly, by a fast hopping process. Diffusion proceeds by Brownian motion at a speed lower than that predicted by hydrodynamic theory, which may be explained by plectoneme diffusion along a rugged energy landscape induced by the DNA sequence. In contrast, hopping, the second mode of motion, facilitates very rapid (< 20 ms) plectoneme displacement by nucleating a new plectoneme at a distant position that may be many kb away from pre-existing plectonemes. These observations directly reveal the dynamics of plectonemes and identify a new mode of movement that is likely to significantly impact cellular physiology by allowing long-distance reorganization of the global conformation of the genome on a millisecond timescale.

Part 2 of this thesis focuses on two different kind of DNA entanglements; first, one created manually between two DNA molecules and second, the entrapment of two chromosomes by the protein complex cohesin.

Chapter 4 shows how a DNA loop, created by winding one molecule DNA around another, can be used as a scanning probe to detect DNA-bound proteins. We developed a combined magnetic and optical tweezes apparatus that allows for the sensitive label-free detection of the positions and binding strength of proteins localized on DNA. A DNA loop was created by looping an optically trapped DNA tether around a magnetic bead tether. We are able to determine the relative distances between DNA-bound EcoRI proteins with 14 ± 8 nm resolution along a λ-DNA molecule. We find a consistent offset between back and forwards scans of 35 ± 15 nm for the detected protein positions, which corresponds to the size of the DNA loop and is in agreement with theoretical estimates. At higher applied stretching forces, the scanning loop is able to remove bound proteins from the DNA. The use of magnetic tweezers in this assay not only allows the facile preparation of many single-molecule tethers, which can be scanned one after the other, but it also allows for direct control of the supercoiling state of the DNA molecule, making it uniquely suitable to address the effects of torque on protein-DNA interactions.

In **Chapter 5** we report our progress on developing an assay that can directly measure the strength of a single cohesin complex. Cohesin provides the essential link between sister chromatids during cell division by entrapping their DNA within a tripartite protein ring. Cohesin must withstand the forces of the mitotic spindle that align chromosomes by pulling them towards the opposing poles via microtubules. Experimentally it has not been possible to determine the mechanical strength of the cohesin complex or to even visualize the entrapment of sister chromatids by cohesin. To determine its mechanical strength cohesin-concatenated circular minichromosomes were isolated from yeast by magnetic-bead-affinity separation via tetO repeats present in the DNA of the artificially constructed minichromosomes. An enzymatic nicking and ligation strategy was used to covalently attach DNA linkers that enable us to create minichromosome-cohesin-minichromosome tethers between an optically trapped bead and the surface of a flow cell. The isolation procedure resulted in a high yield of minichromosomes in a near native state, i.e. DNA-bound proteins and nucleosomes were still present. Initial single-molecule force spectroscopy experiments as well as micrographs of concatenated chromosomes obtained by atomic force microscopy are presented.

Part 3 describes four single-molecule experiments that address the molecular mechanisms of double-strand break repair by homologous recombination. Double-strand breaks are some of the most severe DNA lesions and all forms of life have evolved a mechanism that uses identical, or homologous, sequences to ensure faithful repair and genome stability. Mutations in the genes encoding for proteins involved in homologous recombination, such as BRCA2, can lead to the development of cancer. It is

therefore very important to understand the mechanism of homologous recombination in detail. Homologous recombination starts with the recognition of the break and subsequent resectioning of the DNA end to create a single-strand DNA overhang. In humans the MRE11-RAD50-NBS1 (MRN) complex recognizes, signals, and processes double-strand breaks for resectioning. MRN acts as molecular machine that responds to DNA binding with a large conformational change. MRN uses ATP, the molecular fuel of living organisms, to perform its actions. The interplay between ATP binding/hydrolysis and DNA binding remains, however poorly understood.

Chapter 6 addresses the functional response of MRN to nucleotide binding and hydrolysis on DNA binding. Electrophoretic-mobility-shift assays and atomic-force microscopy reveal that DNA is released from the complex upon nucleotide binding and hydrolysis. DNA release from the MRN complex triggers a large conformational change with the appearance of an open arrangement of the RAD50 coiled coils. Our results support a regulatory role for nucleotide binding, which not only affects the affinity for DNA, but may also trigger changes in the MRN intercomplex interactions.

In the next step of homologous recombination a helical nucleoprotein filament forms on the single-stranded DNA (ssDNA) overhang at the DNA end. In *E. Coli* this filament is formed by RecA. ATP hydrolysis, essential for successful recombination, makes the RecA-ssDNA filament dynamic. In **Chapter 7** we directly probe the structure and kinetics of RecA interaction on its biologically relevant substrate, long ssDNA molecules. We find that RecA ATPase activity is required for the formation of long continuous filaments on ssDNA. These filaments are found to both nucleate and extend with a multimeric unit. Disassembly rates of RecA from ssDNA decrease with applied stretching force, corresponding to a mechanism where protein-induced stretching of the ssDNA aids in the disassembly. Finally, we show that RecA-ssDNA filaments can reversibly interconvert between an extended, ATP-bound, and a compressed, ADP-bound state. Taken together, our results demonstrate that ATP hydrolysis has a major influence on the structure and state of RecA filaments on ssDNA.

In mammalian cells, a similar nucleoprotein filament is formed on ssDNA by RAD51. The actions of RAD51 are facilitated by the breast-cancer-tumor suppressor BRCA2, which is essential for efficient homologous recombination. Biochemical studies with BRCA2 peptides have defined interactions with RAD51 via a series of eight BRC repeats and a C-terminal domain, designated CTRD (C-terminal RAD51 interaction domain). In **Chapter 8**, we describe experiments using a range of single-molecule techniques to determine the influence of the CTRD on intrinsic aspects of RAD51-DNA interactions. At high concentration, the CTRD entangles RAD51 filaments and reduces RAD51 filament formation in a concentration-dependent manner. CTRD does not affect the rate of filament disassembly measured as the loss of fluorescent signal due to intrinsic RAD51 protein dissociation from dsDNA. We conclude that, outside the context of the full-length protein the CTRD does not reduce RAD51 dissociation kinetics, but instead hinders filament formation on double-stranded DNA. The CTRDs

mode of action is most likely the sequestration of multiple RAD51 molecules, thereby rendering them inactive for filament formation on double-stranded DNA.

After the RecA or RAD51 nucleoprotein filament has formed on the ssDNA overhang it performs a search for a homologous sequence, which is truly remarkable both from a kinetic and thermodynamic point of view. In *E. Coli*, the search process is completed within the timescale set by the cell's life cycle, implying an impressive $>10^3$ s^{-1} base-sampling frequency to find the target. Homology search is little affected by a large background of heterologous DNA, suggesting that the filament is somehow able to cope with the abundantly present short sequences in the genome that display heterology or partial homology. In **Chapter 9** we describe experiments that address this remarkable search mechanism by directly probing the strength of the two-molecule interactions involved in homology search and recognition using dual-molecule manipulation, combining magnetic and optical tweezers. We find that the RecA filament's secondary DNA-binding site interacts with a single strand of the incoming double-stranded DNA during homology sampling. Recognition requires opening of the helix and is strongly promoted by unwinding torsional stress. Recognition is achieved upon binding of both strands of the incoming dsDNA to each of two ssDNA-binding sites in the filament. The data indicate a physical picture for homology recognition in which the fidelity of the search process is governed by the distance between the DNA binding sites.

The results presented in this thesis demonstrate that, by developing and applying sensitive techniques, it is possible to unravel the physical and mechanistic process that shape and repair our genome.

Marijn van Loenhout
October 2012

Samenvatting

Dit proefschrift beschrijft de resultaten van reeks experimenten die gericht waren op de fysische eigenschappen van DNA en de wisselwerking tussen eiwitten en DNA bijvoorbeeld tijdens de reparatie van dubbelstrengs breuken. DNA is een lang molecuul dat bestaat uit twee ketens nucleïnezuren die samen een dubbele helix vormen. De sequentie van nucleïnezuren bevat de informatie, die alle levende organisme nodig hebben om te functioneren en zich te vermenigvuldigen. De draadachtige helix structuur van het DNA molecuul ligt aan de basis van vele cellulaire processen, aangezien cellen hun DNA moeten ordenen, verdubbelen, repareren en de torsie moeten reguleren die ontstaat als de dubbele helix zich windt of ontwindt.

Om de fysische mechanismen van deze processen te doorgronden hebben we gevoelige technieken ontwikkeld en toegepast die het mogelijk maakten de structuur en interacties van individuele DNA moleculen en eiwitten vast te leggen. Door kleine (micrometer grote) balletjes aan DNA moleculen te bevestigen konden we deze manipuleren en er krachten op uitoefenen. We hebben hiervoor zowel een magnetisch pincet gebruikt, waarmee we een magnetisch balletje kunnen roteren en eraan kunnen trekken, als ook een optisch pincet, waarmee we een plastic balletje in drie dimensies kunnen bewegen. Alhoewel magnetische en optische pincetten zeer krachtige technieken zijn om moleculaire mechanismen en interacties te onderzoeken geven ze slechts beperkte informatie over de moleculaire structuur. We hebben daarom nog twee andere technieken gebruikt om de vorm en structuur van eiwitten en DNA direct in beeld te brengen. Met behulp van fluorescentiemicroscopie werd de structuur en dynamica van opgedraaide DNA moleculen afgebeeld. Met een atomaire kracht microscoop, die werkt met een tastende naald analoog aan een platenspeler, werden DNA en eiwitten op een oppervlak in beeld gebracht. Hieronder geef ik een korte samenvatting van de vier delen van dit proefschrift: experimentele technieken, en dan vervolgens draaiing, knopen en breuken in DNA.

Deel 0 van dit proefschrift beschrijft een verbeterd positiebepalingsalgoritme dat het mogelijk maakt een groot aantal individuele moleculen tegelijkertijd te meten, om zo de doorvoer van experimenten te vergroten. In tegenstelling tot de meeste bulk technieken, zijn enkel-molecuul experimenten niet beperkt tot alleen het bepalen van een gemiddelde. Het is daarom mogelijk een veel nauwkeuriger begrip van

moleculaire mechanismen en hun dynamica te krijgen. Het verrichten van metingen aan één enkel molecuul is echter onvoldoende om de distributies en fluctuaties van moleculaire eigenschappen te kunnen begrijpen. Het is daarom van essentieel belang om technieken te ontwikkelen die efficiënt de eigenschappen van vele individuele moleculen kunnen bepalen.

Hoofdstuk 2 beschrijft een techniek, die in staat is de doorvoer van magnetische pincet experimenten aanzienlijk te vergroten. Deze driedimensionale positiebepalingsmethode voor niet diffractie-gelimiteerde deeltjes heeft een substantieel hogere resolutie in vergelijk met eerdere methoden, in het bijzonder voor lage vergrotingen of bij hoge signaal-ruisverhoudingen. Het algoritme maakt het daarom mogelijk posities te bepalen met nanometer nauwkeurigheid in een groot gezichtsveld en maakt het zo mogelijk een groter aantal magnetische balletjes en daarmee een groter aantal moleculen te observeren. Een magnetische pincet meet de lengte van individuele DNA moleculen door deze op te spannen tussen een magnetisch balletje en het oppervlak van een vloeistofkamer. Een tweetal externe magneten maakt het mogelijk individuele DNA moleculen op te winden en er een kracht op uit te oefenen. Videomicroscopie is gebruikt om in drie dimensies de positie met een resolutie beter dan een enkele pixel te bepalen op basis van digitale beelden van de magnetische balletjes. Het rooster van pixels in de digitale beelden veroorzaakt een afwijking in de bepaalde positie die sterk toeneemt bij lage vergrotingen. Door gebruik te maken van interpolatie van het beeld op een cirkelvormig rooster was het mogenlijk de lokalisatienauwkeurigheid in drie dimensies tot een orde van grootte te verbeteren ten opzichte van conventionele methodes. De verbeterde prestaties van het algoritme stonden toe nauwkeurige data voor de stijfheid (persistentielengte) en eind-tot-eind lengte te verkrijgen voor meer dan 100 opgespannen DNA moleculen in een enkel experiment.

Deel 1 beschrijft experimenten die dynamica en structuur van een opgedraaid DNA molecuul tonen. Met een magnetische pincet zijn DNA moleculen opgedraaid door de magneten te draaien. Deze extra draaiing bovenop de reeds aanwezige draaiing in de dubbele helix structuur van DNA heet 'supercoiling', en kan leiden tot het ontstaan van DNA lussen die plectonemes worden genoemd. In levende cellen is DNA vrijwel altijd supercoiled doordat de eiwitten langs de dubbele helix bewegen en DNA ontwinden, torsie en daarmee supercoiling genereren. Controle over de mate en dynamica van supercoiling is van essentieel belang voor vele cellulaire processen. Alhoewel de evenwichtstoestand van supercoiled DNA uitgebreid is onderzocht, is er vrijwel niets bekend over de dynamica.

Hoofdstuk 3 beschrijft experimenten die de dynamica van plectonemen laten zien met behulp van een magnetisch pincet en fluorescentiemicroscopie in supercoiled DNA moleculen. We zagen dat meerdere plectonemen aanwezig kunnen zijn en dat hun aantal afhangt van de aangelegde kracht en ionenconcentratie. Plectonemen bewogen langs het DNA door middel van trage diffusie of, onverwacht, door een snel sprongmechanisme. Diffusie vindt plaats door Brownse beweging, met een snelheid

die lager ligt dan verwacht op basis van hydrodynamische theorie. De trage diffusie kan verklaard worden door een model gebaseerd op de verplaatsing van plectonemen over een ruw energie landschap dat veroorzaakt wordt door de heterogeniteit in de DNA sequentie. Het sprongmechanisme daarentegen maakt zeer snelle (< 20 ms) verplaatsingen van plectonemen mogelijk door een nieuw plectoneem te laten ontstaan op een positie die vele kilobaseparen verwijderd kan zijn van reeds bestaande plectonemen. Deze observaties tonen direct de dynamica van plectonemen en identificeren een nieuw verplaatsingsmechanisme dat een belangrijke invloed kan hebben op de cel fysiologie door reorganisatie van de algehele structuur van het genoom mogelijk te maken op een milliseconde tijdschaal.

Deel 2 van dit proefschrift richt zich op twee types DNA verstrengelingen; ten eerste, een handmatig aangelegde verstrengeling tussen twee DNA moleculen en ten tweede, de verbinding van twee chromosomen door het eiwitcomplex cohesin.

Hoofdstuk 4 beschrijft hoe een DNA lus, gemaakt door één DNA molecuul om een ander DNA molecuul te winden, gebruikt kan worden als tastinstrument om DNA-gebonden eiwitten te detecteren. Hiertoe hebben we een gecombineerde magnetisch en optisch pincetopstelling gemaakt, die gevoelige label-vrije detectie van de positie en bindingssterkte van DNA gebonden eiwitten mogelijk maakt. Een DNA lus werd gemaakt door een DNA molecuul (vastgehouden met optische pincet) rond een tweede DNA molecuul te draaien, dat opgespannen werd met een magnetisch pincet. Hiermee konden relatieve posities tussen DNA-gebonden EcoRI eiwitten langs een λ-DNA molecuul met 14 ± 9 nm nauwkeurigheid bepaald worden. Er was een consistent verschil van 35 nm tussen de gedetecteerde eiwitpositie voor voorwaartse en terugwaartse scanbewegingen, dit verschil wordt veroorzaakt door de grootte van de DNA lus en is in overeenstemming met theoretische inschattingen. Bij hogere aangelegde krachten is de verplaatsende lus in staat om gebonden eiwitten van het DNA te verwijderen. Het gebruik van een magnetisch pincet in deze techniek maakt het niet alleen mogelijk om gemakkelijk vele DNA moleculen van te voren op te spannen en deze achtereenvolgens te onderzoeken, maar geeft ook directe controle over de mate van supercoiling van het DNA. Deze methode is daarom bij uitstek geschikt is om de invloed van torsie op eiwit-DNA interacties te bepalen.

In **Hoofdstuk 5** bespreken we een experimentele methode om de mechanische sterkte van een enkel cohesin complex te bepalen. Cohesin vormt de essentiële mechanishe verbinding tussen zuster-chromatiden door deze in een drieledige eiwitring te omsluiten. Tijdens de celdeling moet deze ring de krachten van de mitotische bundels weerstaan, welke de chromosomen oplijnen door ze naar tegenoverliggende polen te trekken. Het is vooralsnog experimenteel niet mogelijk gebleken de mechanische sterkte van cohesin te bepalen, of om zelfs maar de wijze van omsluiting van twee zuster-chromatiden te visualiseren. Om de mechanische sterkte van door een cohesin ring verbonden circulaire minichromosomen te bepalen hebben we deze geïsoleerd uit gistcellen door middel van een magnetische balletjes affiniteitsscheiding.

De affiniteitsscheiding maakte gebruik van tetO herhalingen die aanwezig zijn in het DNA van de kunstmatig samengestelde minichromosomen. Een enzymatische nicking en ligatie procedure werd vervolgens gebruikt om DNA linkers covalent te verbinden met de minichromosomen. Deze linkers maken het mogelijk om via een polystyreen balletje met een optische pincet te trekken aan de door cohesin verbonden DNA moleculen totdat er een breuk optreed. De gebruikte isolatieprocedure had in een hoge opbrengst van minichromosomen welke in nog vrijwel natuurlijke staat verkeerden, met gebonden eiwitten en nucleosomen. We tonen de eerste krachtspectroscopie experimenten op individuele moleculen en atomaire kracht microscopieplaatjes van door cohesin verbonden chromosomen.

Deel 3 beschrijft vier enkel molecuulexperimenten die ingaan op de moleculaire mechanisme van dubbelstrengs breukreparatie via homologe recombinatie. De breuk van beide DNA strengen is één van de meest ernstige vormen van schade die optreden in DNA. Alle levensvormen beschikken daarom over een mechanisme dat gebruik maakt van additionele identieke, ofwel homologe, DNA sequenties om een foutloze reparatie en de stabiliteit van het genoom te waarborgen. Mutaties in genen die coderen voor de eiwitten die betrokken zijn bij homologe recombinatie, zoals bijvoorbeeld BRCA2, kunnen leiden tot het ontstaan van kanker. Een goed begrip van alle processen die een rol spelen tijdens homologe recombinatie is daarom van groot belang. Homologe recombinatie begint met het herkennen van de breuk en vervolgens het gedeeltelijk afbreken van het DNA einde om een enkelstrengs DNA uiteinde te creëren. Bij mensen herkent het MRE11-RAD50-NBS1 (MRN) complex de dubbelstrengs breuken, en dit complex bereidt vervolgens de DNA uiteinden voor op de omvorming naar enkelstrengs DNA. MRN functioneert als een moleculaire machine, die reageert op DNA binding met een grote verandering van vorm. MRN gebruikt ATP, de moleculaire energie drager van levende organismen, om zijn acties uit te voeren. Er is echter nog veel onduidelijk over de wisselwerking tussen ATP-binding, ATP-hydrolyse en DNA binding van MRN.

Hoofdstuk 6 richt zich op de functionele respons van MRN op nucleotide binding, hydrolyse en DNA binding. Gelelektroforese en atomaire krachtmicroscopie experimenten laten zien dat het MRN complex DNA loslaat na nucleotide binding en hydrolyse. DNA binding aan het MRN complex leidt tot een grote conformatieverandering en het verschijnen van een open rangschikking van de RAD50 'coiled-coils'. Onze resultaten ondersteunen een controlerende rol voor nucleotide binding, die niet alleen de affiniteit voor DNA beïnvloedt maar ook veranderingen in MRN intercomplex interacties kan uitlokken.

In de volgende stap van homologe recombinatie vormt zich een helix-vormig nucleoproteine filament op het enkelstrengs DNA uiteinde. In *E. Coli* wordt dit filament gevormd door RecA. ATP hydrolyse maakt het RecA filament dynamisch en is van essentieel belang voor recombinatie.

Hoofdstuk 7 beschrijft onderzoek naar de structuur en dynamica van RecA op het biologisch meest relevante substraat, lange enkelstrengs DNA moleculen. De ATPase activiteit van RecA blijkt noodzakelijk te zijn om lange continue filamenten op enkelstrengs DNA te vormen. De metingen geven ook aan dat deze filamenten ontstaan en groeien door middel van RecA multimeren. De snelheid waarmee RecA van enkelstrengs DNA loslaat neemt af met toenemende aangelegde kracht op het DNA. Dit kan verklaard worden door een mechanisme waarbij dissociatie wordt bevorderd door het uitrekken van het enkelstrengs DNA door RecA. Tenslotte laten we zien dat RecA filamenten reversibel kunnen overgaan tussen een uitgestrekte ATP-gebonden toestand en een gecomprimeerde ADP-gebonden toestand. Tezamen demonstreren deze resultaten dat ATP-hydrolyse een grote invloed heeft op de structuur en toestand van RecA filamenten op enkelstrengs DNA.

In zoogdiercellen wordt een vergelijkbaar filament gevormd op enkelstrengs DNA door RAD51. De acties van RAD51 worden gefaciliteerd door de borstkankertumor onderdrukker BRCA2, welke essentieel is voor efficiënte homologe recombinatie. Biochemische studies met BRCA2 peptiden hebben aangetoond dat interacties met RAD51 plaats vinden via een reeks van acht BRC-herhalingen en een C-terminus domein, aangeduid als CTRD (C-terminus RAD51 interactie domein).

Hoofdstuk 8 beschrijft een reeks van enkel molecuul experimenten, die tot doel hebben de invloed van het CTRD op de intrinsieke RAD51-DNA interacties te bepalen. Deze experimenten laten zien dat het CTRD bij hoge concentraties de RAD51 filamenten verstikt en filamentvorming reduceert in een concentratieafhankelijke manier. Het CTRD heeft echter geen invloed op de dissociatiesnelheid van RAD51 filamenten, bepaald als de afname van fluorescentie na dissociatie van RAD51 eiwitten. We concluderen dat, buiten de context van het volledige BRCA2 eiwit, het CRTD geen invloed heeft op dissociatiekinetica maar wel de filamentvorming op dubbelstrengs DNA bemoeilijkt. De rol van het CRTD is hoogst waarschijnlijk het binden en buiten werking stellen van meerdere RAD51 moleculen, waardoor deze geen filament meer kunnen vormen op dubbelstrengs DNA.

Het RecA of RAD51 filament gevormd op enkelstrengs DNA verricht, een zowel vanuit kinetisch als thermodynamisch oogpunt, buitengewoon opmerkelijk zoekproces naar homologe sequenties. In *E. Coli* voltrekt het zoekproces zich binnen de tijdschaal van de cel cyclus, wat een indrukwekkende $>10^3\ s^{-1}$ base-sampling frequentie inhoud om het doel te vinden. Deze zoektocht naar homologie wordt nauwelijks beïnvloed door een grote achtergrond van heteroloog DNA, wat impliceert dat het filament om kan gaan met de ruimschoots aanwezige korte heterologe of gedeeltelijke homologe sequenties!

Hoofdstuk 9 beschrijft experimenten naar dit uitzonderlijke zoekmechanisme. We bepaalden rechtstreeks de sterkte van de twee-molecuul interacties die betrokken zijn bij de zoektocht naar homologie door twee moleculen te manipuleren met een gecombineerd optische en magnetisch pincet. De experimenten laten zien dat de secundaire bindingsplaats van RecA een wisselwerking kent met één enkele streng

van het inkomende dubbelstrengs DNA molecuul tijdens het testen van homologie. Herkenning vereist het openen van de helix en wordt in hoge mate versterkt door negatieve torsie die het DNA molecuul ontwindt. Herkenning is succesvol als beide strengen van het inkomende dubbelstrengs DNA molecuul aan elk van de twee enkelstrengs bindingsplaatsen van het filament binden. De data ondersteunen een fysisch model waarbij de betrouwbaarheid van homologieherkenning wordt bepaald door de afstand tussen de twee DNA bindingsplaatsen.

Dit proefschrift laat hiermee zien dat het, door middel van het ontwikkelen en toepassen van zeer gevoelige technieken, mogelijk is de fysische en mechanistische processen te ontrafelen die het genoom vorm geven en repareren.

Marijn van Loenhout
oktober 2012

Curriculum Vitae

27-10-1981	Born Nijmegen, The Netherlands
1994–2000	Pre-university education Stedelijk Gymnasium Nijmegen, The Netherlands
2000–2006	M.Sc. Applied Physics (with distinction) University of Twente, The Netherlands
2005–2005	Internship Huck group Melville Laboratory for Polymer Synthesis University of Cambridge, United Kingdom
2007–2012	Ph.D research under supervision of Prof. dr. C. Dekker *'Single-molecule studies of the twisted, knotted, and broken genome.'* Delft University of Technology, The Netherlands

List of publications

1. Dynamics of DNA supercoils.
   M. T. J. van Loenhout, M. V. Grunt, and C. Dekker
   Science **338**, 94 (2012).

2. *Magnetic Forces and DNA Mechanics in Multiplexed Magnetic Tweezers.*
   I. De Vlaminck, T. Henighan, M. T. J. van Loenhout, D. R. Burnham, and C. Dekker
   PLoS ONE **7**, e41432 (2012).

3. *Non-Bias-Limited Tracking of Spherical Particles, Enabling Nanometer Resolution at Low Magnification.*
   M. T. J. van Loenhout, J. W. J. Kerssemakers, I. De Vlaminck, and C. Dekker
   Biophysical Journal **102**, 2362 (2012).

4. *Mechanism of Homology Recognition in DNA Recombination from Dual-Molecule Experiments.*
   I. De Vlaminck, M. T. J. van Loenhout, L. Zweifel, J. den Blanken, K. Hooning, S. Hage, J. Kerssemakers and C. Dekker
   Molecular Cell **46**, 616 (2012).

5. *Highly Parallel Magnetic Tweezers by Targeted DNA Tethering.*
   I. De Vlaminck, T. Henighan, M. T. J. van Loenhout, I. Pfeiffer, J. Huijts, J. W. J. Kerssemakers, A. J. Katan, A. van Langen-Suurling, E. van der Drift, C. Wyman and C. Dekker
   Nano Lett **11**, 5489 (2011).

6. *Effect of the BRCA2 CTRD domain on RAD51 filaments analyzed by an ensemble of single molecule techniques.*
   J. T. Holthausen, M. T. J. van Loenhout, H. Sanchez, D. Ristic, S. E. van Rossum-Fikkert, M. Modesti, C. Dekker, R. Kanaar and C. Wyman
   Nucleic Acids Res **39**, 6558 (2011).

7. *Torsional regulation of hRPA-induced unwinding of double-stranded DNA.*
   I. De Vlaminck, I. Vidic, M. T. J. van Loenhout, R. Kanaar, J. H. G. Lebbink and C. Dekker
   Nucleic Acids Res **38**, 4133 (2010).

8. *Dynamics of RecA filaments on single-stranded DNA.*
M. T. J. van Loenhout, T. van der Heijden, R. Kanaar, C. Wyman, and C. Dekker
Nucleic Acids Res **37**, 4089 (2009).

9. *Hydrodynamic flow induced anisotropy in colloid adsorption.*
M. T. J. van Loenhout, E. S. Kooij, H. Wormeester, and B. Poelsema
Colloid Surface A **342**, 46 (2009).

Casimir PhD series 2012-28
ISBN: 978-90-8593-135-5