

Topoisomerase and the unwinding of stressed DNA

Proefschrift

ter verkrijging van de graad van doctor
aan de Technische Universiteit Delft,
op gezag van de Rector Magnificus prof. dr. ir. J.T. Fokkema,
voorzitter van het College voor Promoties,
in het openbaar te verdedigen op maandag 25 juni 2007 om 15.00 uur

door

**Daniel Alexander
KOSTER**

doctorandus in de Scheikunde
geboren te Amsterdam

Dit proefschrift is goedgekeurd door de promotor:

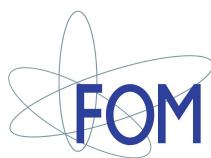
Prof. dr. C. Dekker

Toegevoegd promotor:

Dr. N.H. Dekker

Samenstelling van de promotiecommissie:

Rector Magnificus	voorzitter
Prof. dr. C. Dekker	Technische Universiteit Delft, promotor
Dr. N.H. Dekker	Technische Universiteit Delft, toegevoegd promotor
Prof. H. Schiessel	Universiteit Leiden
Prof. dr. ir. J. E. Mooij	Technische Universiteit Delft
Prof. dr. J. H. de Winde	Technische Universiteit Delft
Dr. M.A. Bjornsti	St. Jude Children's Research Hospital, Memphis, TN, VS
Dr. D. Bensimon	École Normale Supérieure, Paris, Frankrijk



Keywords: DNA-protein interactions, topoisomerase, camptothecin, topotecan, single-molecule, nanofabrication.

Published by: Daniel Koster

Cover design by: TU Delft/Tremani

Printed by: Gildeprint Drukkerijen B.V.

An electronic version of this thesis, including color figures, is available at:
<http://www.library.tudelft.nl/dissertations/>

Casimir PhD Series, Delft-Leiden, 2007-07

ISBN: 978-90-8593-031-0

Copyright © 2007 by Daniel Koster

Preface

Dear readers,

There are a lot of people that I'd very much like to thank. First, the members of my committee. I thank you very much for the effort and time you have invested and it is an honor that you were willing to take part. Then, my supervisors, Nynke Dekker and Cees Dekker. Nynke, I'll start with you. It's been a great learning experience to be your (first!) PhD student. I remember the first time we met, a few months before I started in Delft. There you were, sitting in a chair, with your legs curled up in some weird topological shape indicative of the research we would do together. To me, it seemed like a particularly painful pose, but since you were smiling, I figured your legs would be, yes, freely jointed. I experienced what it meant to "go into physics". In the process, you have always had the patience to allow me to learn, both by actually teaching me hands-on to work with the tweezers and more indirectly, by being a most generous provider of "learning experiences", such as talks and conferences abroad. This has been very stimulating and great fun and I thank you for that. I wish you all the best with the new research endeavors you will undertake, gradually transitioning from single-molecule to single-complex studies, perhaps even to the single-cell level.

Cees, you epitomize the quality to form a judgement in the first few minutes, or maybe even the first few seconds. Throughout the years, my perception of your personality has evolved. It started from the obvious: a highly charismatic and dynamic person that quickly and intuitively forms his opinion and is often right, and sometimes wrong. In the initial phases of the work, it has been a continuing effort to involve and interest you. I have happily made this effort, leading to very interesting, constructive and pleasant discussions that I and the work have benefited greatly from. Your writing, both scientific and non-scientific is exceptionally sharp and I've learnt much in this aspect as well. You are not the kind of professor that is never around - on the contrary, getting a hold of you for a

discussion is usually a matter of minutes, and sometimes a few hours, rather than days. That is wonderful and I thank you for that. As time went by, I got to know and appreciate your thoughtful (even spiritual) sides, through your involvement in discussions about the relation between science and religion, the fundamental value of human life and the limits there ought to be set for the technology we humans create. I do not always share your views of religion, though they are not strange for me and I respect them. It has been a real privilege to get to know you and work with you.

Much of what I have learnt during the PhD comes from interactions with my fellow lab mates. Throwing ideas at each other, often vague, not always well thought through, not always to the point, we shape each others thinking. At MB, the group of people is amazing. I thank all of you! I think that we, the PhD students, are very close and I literally want to thank all of you. In particular Koen whose relentless quest for control experiments is amazing and inspiring. Diego, it was nice to get to know you, together with Susan. Ale, I like you a lot, you're lucky to have such cool and warm parents at the same time. Take care, my dear. Uli (uhh, Herr Prof. Dr. Keyser, PhD), thanks for being so critical, grumpy and fun! You're great and I wish you all the best. Irene and Benny - take care and thank you. Dr. Seideltje, your giggle still resonates in my head - all the best. Bernhard, take care - you have such broad knowledge, thanks for your expert help on power spectrum renormalizations. Aurélien "Watson" Crut, it was a nice experience to work with you directly. I've learnt a lot from you. Suzanne, you are so good at what you're doing, thank you very much for synthesizing molecules and all the expert advice in the biolab. I wish you all the best! Sjoerd-Jan, with your new name you will undoubtedly fully integrate into Dutch society; besides a great Chinese, you might even become a great Dutchman. Don't sleep in the lab too much though. Serge, I think you're a very talented group leader, doing wonderful experiments. All the best in the future, together with Marcel and Claudia. Freek, our one and only heavy metal scientist, good luck with the hydrogenase. Douwe, all the best with your PhD in Germany. Jaan, I'm very curious what will happen with your evolution experiments. Take care. Ya-Hui, it's great you came to MB, I wish you and your family well. Martin, take care - all the best, I'm very curious what you will end up doing in the future - I'm sure you'll be good at it. It's been great to get to know you! Fabian, thanks for coming to MB and thanks for all your good work with the triple helix. I wish you all the best in your PhD! Elisa, you were my first MSc student - I've learnt a lot from you. You produced a lot of high quality data and I thank you for that! You can be very proud of your achievements. Arnold, was nice to get to know you and good luck with your career. Michiel, good luck with your challenging experiments

and all the best with your future. Igor, I really hope you'll get the RNA to work the way you want soon - you deserve it. All the best, wherever you may go in the future. Elsemarieke, thanks for being there, it was very helpful and all the best! Christine, your sharp voice has often made me aware of the equality of men and women. Thank you a lot for all your useful help in the clean room and for your good company. Thijn, thanks for being so honest and outspoken - I'm sure you'll continue to care for all your small, sweet and innocent little animals as much as you can. Marcel, I hope you'll get to the single-ion limit soon. That would be so nice! Liset, our financial expert, and expert on what project code I should use. Thank you so much and all the best in the future. Jaap, the evaporator is very happy that you're here and so are we! All the best. Marijn, you just embarked on your PhD route. Enjoy it, you have great experiments ahead of you. Frank, thanks for being such an easy subject of a movie. We didn't have to exaggerate a lot. Ralph, you are doing a great PhD. I'll remember "our" smokes together. Good luck with the future, wherever it may take you. Iddo, my ears still hurt from the "klinkers" you scored against me at the table soccer table. They were physically and psychologically very painful. Jelle, the speed and accuracy with which you can machine just about anything is amazing! Yuval, thanks for your advice on wave propagation and in general for being a nice partner for discussions. All the best with your new group at Bar-Ilan; we'll meet soon.

I'd also like to thank the collaborators I have been very fortunate to work with. Stewart, the collaboration on the vaccinia topo started with you. Your mechanistic insight into topo is phenomenal, including your knowledge and description (!) of the topo literature. Too bad you couldn't participate in the committee. Mary-Ann, combining single-molecule and *in vivo* work is very challenging, both technically and culturally. Most of the time, we spent discussing what it is exactly that you have done and exactly what it is that we have done. It required a lot of patience as well as self-reflectance on our diverse backgrounds. We have done so in a very constructive manner and I appreciate your part in this greatly. I've learnt a great deal from you and I'm convinced that it was worth the effort. We have shown our communities the power of combining two worlds that are only seemingly different. Good luck to you, and to Komar "cold-room" Palle with all your endeavors and your career! Vincent, I thank you for all of your good advice, both technically and conceptually. You have taught me a lot. David, you are one of the most inspiring persons I've met in my life: you have an admirable combination of scientific depth and breadth. You have helped me conceptually and practically in the topo experiments, and introduced me to systems biology. It has been a great privilege to get to know you and I'm looking much forward to more encounters in the future.

Before entering the clean room, I knew basically nothing about clean room work. I thank the clean-room staff for teaching me to work in this very special environment. Marc, thanks for being there to help with the semi-dead chlorine etcher - without you it would never have become in a usable state. Marco, the evaporator thanks you for always answering the phone when it calls for your help. I thank you too for all the support and training you gave me. Anja, thanks for helping me out with the e-beam pattern generator. Your knowledge and experience writing with the smallest beams was very helpful. Also thanks for teaching me the tricks of NEB resist. Roel, thanks for admitting you stole my ZEP resist. Also thanks for your numerous and helpful advice in surface chemistry and for cleaning of FuSi.

I've been fortunate to have been surrounded by many good friends and family. Let me just say here that you have shown me how much it is worth to have you around me. You have supported me a great deal, especially in the past year that was not at all easy for me. My next stop is the Weizmann Institute in Israel and I'm confident that keeping in touch will be easy. My new home will be always open to you. I'm looking forward to keep on sharing all aspects of our lives together in the future.

Daniel Koster
Delft, June 2007

Contents

1	Introduction	7
1.1	The single-molecule paradigm	8
1.2	Proteins and their structure	9
1.3	Enzyme catalysis	12
1.4	The structure of double-stranded DNA	13
1.5	Cellular processes induce DNA supercoiling	15
1.6	Topoisomerases remove DNA supercoiling	17
1.7	Topoisomerase IB	19
1.8	Overview of this dissertation	21
	References	23
2	Magnetic tweezers	27
2.1	Measuring forces in real space	28
2.2	Measuring forces in Fourier space	29
2.3	Temporal and spatial resolution	34
	References	36
3	Fast dynamics of supercoiled DNA revealed by single-molecule experiments	37
3.1	Introduction	38
3.2	Results	40
3.3	Discussion	47
3.4	Materials and Methods	52
3.5	Acknowledgements	53
3.6	Supplementary Information	54
	References	56

4	Friction and torque govern the relaxation of DNA supercoils by eukaryotic topoisomerase IB	59
4.1	Introduction	60
4.2	Topoisomerase IB uncoils DNA in a step-wise fashion	60
4.3	Topoisomerase hinders the uncoiling	63
4.4	Torque drive DNA rotation over a rugged energy landscape	65
4.5	Acknowledgements	67
4.6	Materials and methods	67
	References	69
5	Multiple events on single molecules: unbiased estimation in single-molecule biophysics	71
5.1	Introduction	72
5.2	Topoisomerase IB steps are subjected to a global constraint	73
5.3	Maximum likelihood and domain constraints	75
5.4	Numerical simulation and the consequences of ignoring global constraints	78
5.5	Application of the method	82
5.6	Processivity measurements on limited substrate	83
5.7	Transitions in FRET efficiency	83
5.8	Concluding remarks	85
5.9	Acknowledgements	85
5.10	Supplemental Materials	85
	References	92
6	Antitumor drugs impede DNA uncoiling by Topoisomerase I	95
6.1	Introduction	96
6.2	Drug binding hinders DNA uncoiling	97
6.3	Drug binding traps Top1B on the DNA	100
6.4	Implications for <i>in vivo</i> transcription and replication	104
6.5	Camptothecin resistant Top1B fails to accumulate positive supercoils	109
6.6	Methods	112
	6.6.1 DNA constructs	112
	6.6.2 Enzyme and buffers	112
	6.6.3 Strains and plasmids	112
	6.6.4 <i>In vivo</i> assays and 2-D gel electrophoresis	112
6.7	Acknowledgements	113
6.8	Supplementary Materials	114
	References	128

7 Atomic force microscopy shows that vaccinia topoisomerase IB generates filaments on DNA	131
7.1 Introduction	132
7.2 Materials and methods	133
7.3 Results	137
7.4 Discussion	143
7.5 Acknowledgements	147
References	147
8 Coated nanofabricated structures for the single-molecule study of DNA polymerizing enzymes	151
8.1 Single-molecule fluorescence at high concentrations. Why?	152
8.2 Setup of the experiment	153
8.3 Finite-element simulations of zero-mode waveguides	155
8.4 Nanofabrication of zero-mode waveguides	159
8.5 Characterization of waveguides	159
8.6 PEG coating of glass surfaces	160
8.7 Biotinylated PEG specifically binds streptavidin on a surface . . .	162
8.8 DNA polymerase activity on PEG surface	163
8.9 Acknowledgements	164
References	166
Summary	167
Samenvatting	171
Curriculum Vitae	175
List of Publications	177

Chapter 1

Introduction

This chapter starts by introducing the rationale for performing biophysics experiments at the level of single biomolecules. We then describe the biological context in which the experimental work described in this dissertation is performed. We introduce proteins and highlight the vast diversity in their structure and the myriad of functions that they have in the cell. We then explain how enzymes can accelerate chemical reactions, making complex and intrinsically slow chemical reactions possible at biologically relevant timescales. Subsequently, we focus on the topological dilemmas that the double-stranded and helical structure of DNA raises during ubiquitous cellular processes such as DNA replication and transcription. Finally, we describe the enzyme toolkit that the cell has at its disposal to solve topological dilemmas, which includes enzymes called topoisomerases and specifically focus on one member of the topoisomerase family, topoisomerase IB, and discuss which relevant topics we choose to investigate.

When studying biology, we are amazed at the level of complexity we observe in cells and at the ordered sub-structure within them, and we are fascinated by the exotic biomolecular machines that perform the work required for cell survival and proliferation. By studying biomolecules we learn to appreciate the complexity in cells as well as the unifying principles that govern their behavior. We learn about ourselves, about our present, past and future. We learn about our present by observing the structure of those biomolecules and the mechanism by which they operate in our bodies today. We learn about our past by comparing these structures and mechanism between different organisms and find common features that shed light on our evolution. By understanding how we function today, we can shape our tomorrow so as to live healthier lives by preventing or curing disease.

This dissertation contains experimental studies of biomolecules at the level of the individual molecule and lies in the field of single-molecule biophysics. The field has evolved significantly in the past 20 years up to the point these studies now have sufficient resolution to address detailed and quantitative questions on the mechanism of operation of single biomolecules. The experiments have been made possible because of the development of new experimental strategies or tools such as the mechanical manipulation of single DNA molecules or the development of sensitive CCD cameras capable of detecting fluorescence light emitted from single labeled molecules.

1.1 The single-molecule paradigm

It has become possible to measure on and manipulate single isolated biomolecules and follow their behavior in time, either in isolation, or in interaction with other biomolecules [1]. As the experimental techniques used are highly sensitive, one can gain much insight into these biomolecules. For example, magnetic tweezers were used to progressively extend a linear dsDNA molecule while simultaneously measuring the force it takes to do so [2, 3]. The precision of the measurement was so high that one was able to rigorously test existing models that describe force-extension behavior of polymers. With a well-established theoretical description of polymer stretching in hand, one was able to accurately quantify and interpret the effects that proteins or enzymes have on polymers such as DNA.

In general, the controlled application and monitoring of force applied to biomolecules is biologically relevant as these molecules are subjected to and themselves apply forces in their natural environment. The force range that one can expect biomolecules to exert can be roughly estimated from the energy that is available to them due to thermal agitation. Per degree of freedom, each molecule

has $(1/2)k_B T \sim 2\text{pN}\cdot\text{nm}$ of thermal energy at its disposal. They typically displace themselves or deform partner molecules at length scale comparable to their own length scales, *i.e.* of order 1 nm. The forces required will thus be of order pN. Indeed, RNA polymerase induces pN forces as it moves over the DNA [4], kinesin exerts pN forces as it steps over microtubules [5] and microtubules exert pN forces to segregate chromosomes lined up in the equatorial plane during mitosis [6]. If these biomolecule were not able to generate pN forces, but rather, say, only fN forces, they would not be able to perform their tasks, as they would not be able to withstand the forces of Brownian motion. They are not able to exert forces that are much higher (say nN) because these forces typically break the covalent bonds that the molecules are made of. Thus, pN seems a reasonable force range for biomolecules to exert. It would therefore be of particular use that an experimental apparatus used to study these biomolecules is capable to exert forces on these biomolecules that are in the pN range.

Aside from the ability to exert biologically relevant forces to biomolecules and monitor their response, there is a more fundamental motivation for studying biomolecules in their single-molecule limit, which is that molecules in principle behave stochastically [7]. When observing an ensemble of molecules, one thus necessarily observes their ensemble-averaged properties. Although virtually everything we know about biology, medicine, physics, chemistry, indeed, all we know about our world was inferred from techniques that are not single-molecule techniques, there are significant advantages to removing the masking effect of ensemble averaging. Observing single molecules is a natural way to do so, as synchronizing a population consisting of many billions of molecules is practically impossible.

Finally, it is important to realize that when molecules interact with one another in a living cell, they do so at the level of individual molecules. Thus, studying single molecules constitutes a biologically relevant regime. It may represent a "bottom-up" approach of biology with the hope that by understanding the interactions between biomolecules in their single molecule limit, one will be able to extrapolate the understanding to the cell in its entirety.

1.2 Proteins and their structure

A protein is a long polymer that consists of amino acids. A single amino acid contains an amino group and a carboxyl group linked together by a carbon atom. Also linked to the carbon atom is a side-group, that contains one of 20 possible chemical residues. Fig. 1.1a shows three such amino acids (designated aa₁, aa₂

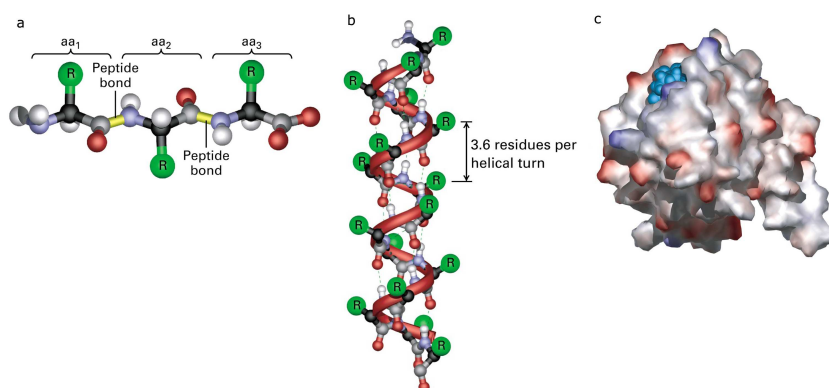


Figure 1.1: Protein structure. (a) Primary structure of a protein, consisting of amino acids linked together by peptide bonds. Proteins are typically made up of a polypeptide polymer with a length of between 50 and 2000 amino acids. (b) The backbone of proteins can form internal hydrogen bonds (dashed lines) yielding a stable structure called the secondary structure. (c) Proteins fold into a distinct 3D structure, called tertiary structure. One way to display the 3D structure of proteins is by showing the surface of proteins that is accessible for solvent molecules, as is shown here. Image adapted from [8].

and aa₃) connected by two peptide bonds. The sequence of these residues is called the protein's primary structure. Fig. 1.1b shows a protein's secondary structure, a stable structure that forms due to hydrogen bonding between the many amide groups and carboxylic groups of the protein backbone. Proteins finally have a distinct three-dimensional structure that plays a crucial role in their function and is shown in Fig. 1.1c.

Proteins are typically large macromolecules with an extremely large diversity in shapes that carry out a wide array of functions in a cell. Fig. 1.2 displays a number of proteins to scale. The top panel of Fig. 1.2 shows proteins or enzymes traversing a membrane. For example, shown in the top panel in ?? is the ATP synthase enzyme, capable of synthesizing adenine triphosphate (ATP), the energy currency of the cell. A gradient of protons that is built up across the membrane drives the flow of protons through the ATP synthase enzyme. This flow is then used to rotate part of the complex, allowing for the synthesis of ATP. Other proteins, such as cytochrome c oxidase also traverse the membrane and are instrumental in building up the proton gradient. In the middle panel of Fig. 1.2, a double-stranded DNA molecule is displayed with proteins and enzymes bound to it. For example, RNA polymerase is capable of using the information stored in the DNA to synthesize a messenger RNA molecule (mRNA) that is used for the synthesis of proteins. Other proteins mechanically deform DNA, such as

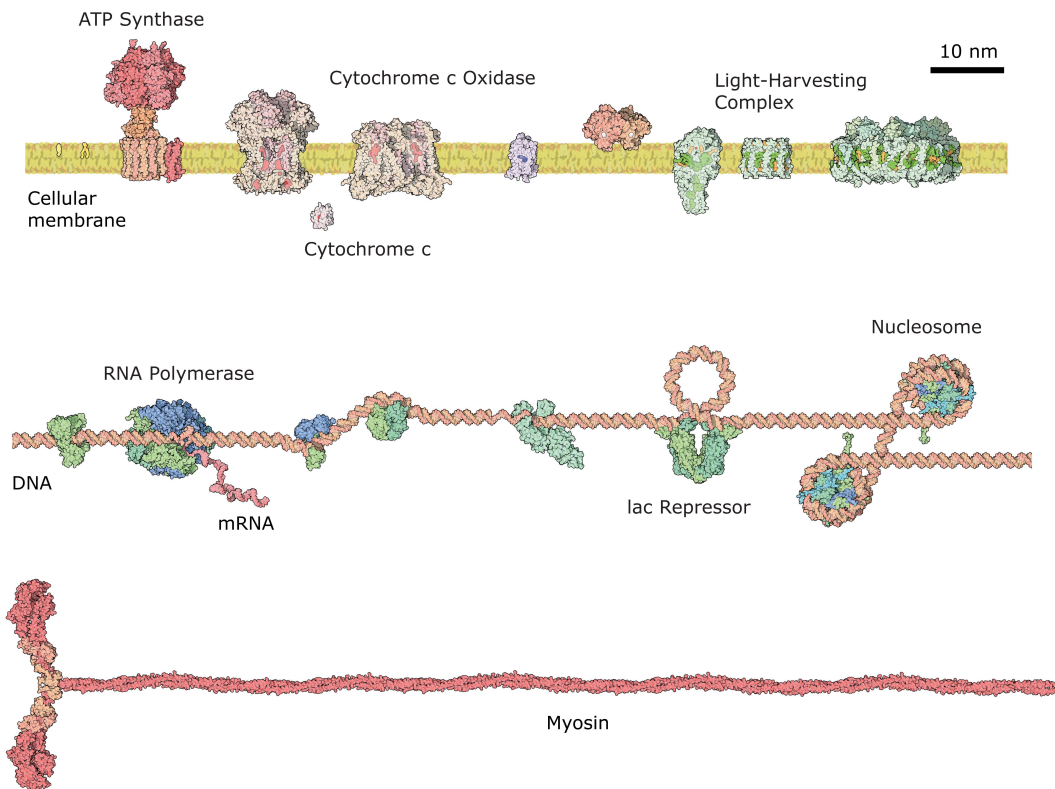


Figure 1.2: An exhibition of proteins and enzymes drawn to scale. The text describes examples of the myriad of functions that proteins perform in the cell, such as the synthesis of ATP (ATP synthase), the packaging of DNA into the cell nucleus (nucleosomes), regulation of transcription (lac repressor) or the generation of a mRNA copy based on a DNA template that is used for the synthesis of proteins. Figure adapted from RSCB protein data bank.

the lac repressor protein, that grabs on to the DNA at a specific sequence and introduces a loop in the DNA. As such, they prevent the binding and action of RNA polymerase at that specific DNA locus and are involved in the regulation of gene expression. Nucleosomes are proteins that wrap DNA around them and thus help in achieving the feat of packaging the meter-long DNA present in each cell into the small cell volume, which for many cells is on the order of a μm^3 . On the bottom panel of Fig. 1.2, Myosin is shown, which is a molecular motor that allows the cell to move.

1.3 Enzyme catalysis

Enzymes greatly accelerate (often million-fold) chemical reactions and thereby facilitate the complex chemistry in the cell. They do so in a highly specific manner: each enzyme is optimized to accelerate a specific chemical reaction and consequently there are many such enzymes present in a cell. The three-dimensional structure of enzymes is key to their ability to catalyze chemical reactions, *i.e.* to accelerate reaction rates. One description of the mechanism by which enzymes achieve this feat is in terms of an "induced fit" mechanism. Imagine an enzyme E whose substrate S is converted enzymatically to a product P. For the conversion to take place, S must pass a state, called the transition state, that is the most energetically unfavorable state, as depicted in Fig. 1.3(a) and denoted $S^\ddagger$. Enzymes increase the reaction rate by lowering the energetic activation barrier of the chemical reaction, which is equal to the free energy difference between S and its transition state and is denoted $\Delta G^\ddagger$. In the first step of enzyme catalysis, S is held in the binding pocket of the enzyme, which decreases the entropy and thus increases slightly the free energy of interaction, as shown by the shoulder (designated with a star) in the interaction energy graph in Fig. 1.3(a), middle sketch. Most of the energetically favorable interactions between S and E do not form when E interacts directly with S, but rather when E deforms slightly and S deforms as well. To use the analogy of a key that fits to a lock, one would say that the key only fits into the lock when both the key and the lock slightly change shape. The many weak interactions between the transition state and the enzyme cavity result in a decrease in the system's total free energy, as shown schematically in the large dip in the middle energy plot in Fig. 1.3. The sum of all the energetic terms sketched in Fig. 1.3(a) is sketched in Fig. 1.3(b) and exhibits the decreased barrier height. With the barrier height decreased due to the enzyme, a thermal fluctuation will drive the chemical change from $S^\ddagger$ to P within a biologically reasonable timeframe. As P does not fit to the enzyme cavity perfectly either, P unbinds, triggering a shape change of E back to its original shape. As enzymes return to their original shape and all the energy borrowed from the local enzyme-substrate interactions must be returned, enzymes are true catalysts that can be re-used over and over again. Therefore, enzymes cannot change the overall ΔG of the reaction and thus do not change the direction of the reaction.

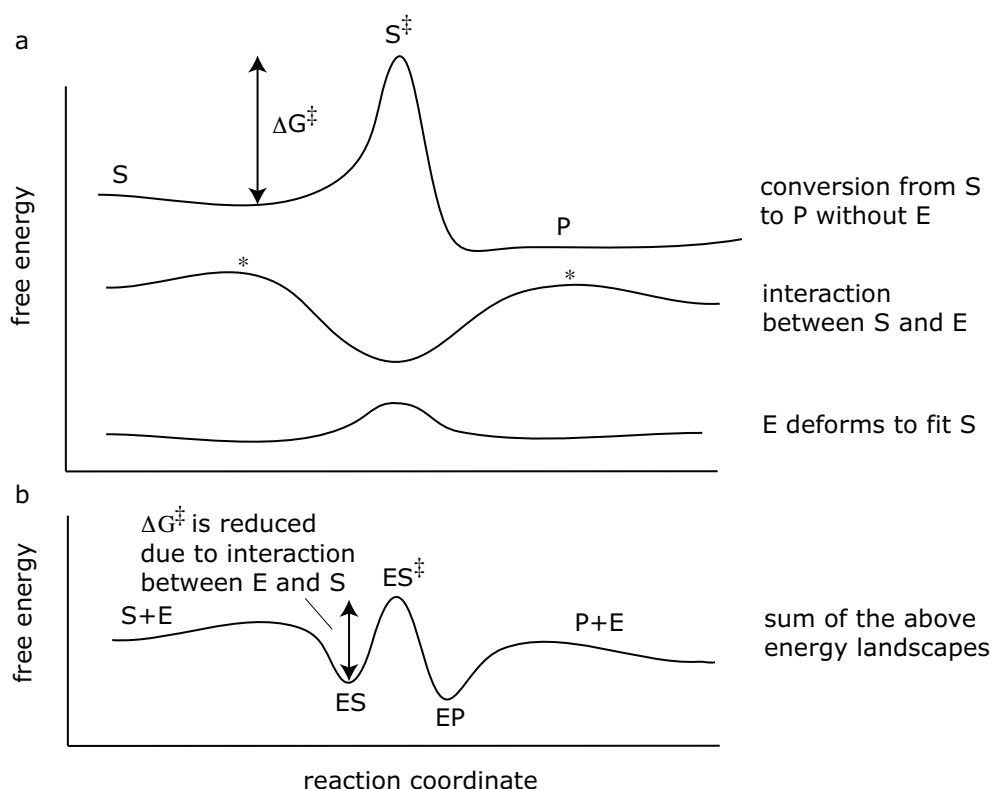


Figure 1.3: Schematic representation of the energy landscape of a chemical reaction. The text describes how enzymes lower the activation energy $\Delta G^\ddagger$ and increase the rate of the chemical reaction. Diagram adapted from [9].

1.4 The structure of double-stranded DNA

The structure of double-stranded DNA is one of the best-known structures in biology. As shown in Fig. 1.4a, each strand consists of a backbone, which is a polymer with a deoxyribose molecule linked to a phosphor group as a monomer. Connected to the deoxyribose molecule of each monomer is a base molecule that can have one of four distinct chemical identities: an adenine (A), thymine (T), cytosine (C) and guanine (G). An A can pair stably with a T through two hydrogen bonds, whereas a G can pair with a C through three hydrogen bonds. Moving along the contour of the DNA, one encounters thus a sequence of bases that carries information. As each monomer of the DNA can be linked to one of four bases, a sequence of N bases can express 4^N unique combinations of bases. The backbone is not symmetrical and is linked to different carbon atoms of the deoxyribose, specifically the 3' and the 5' carbon atoms. The terminology used to define the two ends of a linear DNA molecule is based on this numbering of the carbon atoms: there is a 5'-end and a 3'-end to the DNA, see Fig. 1.4a.

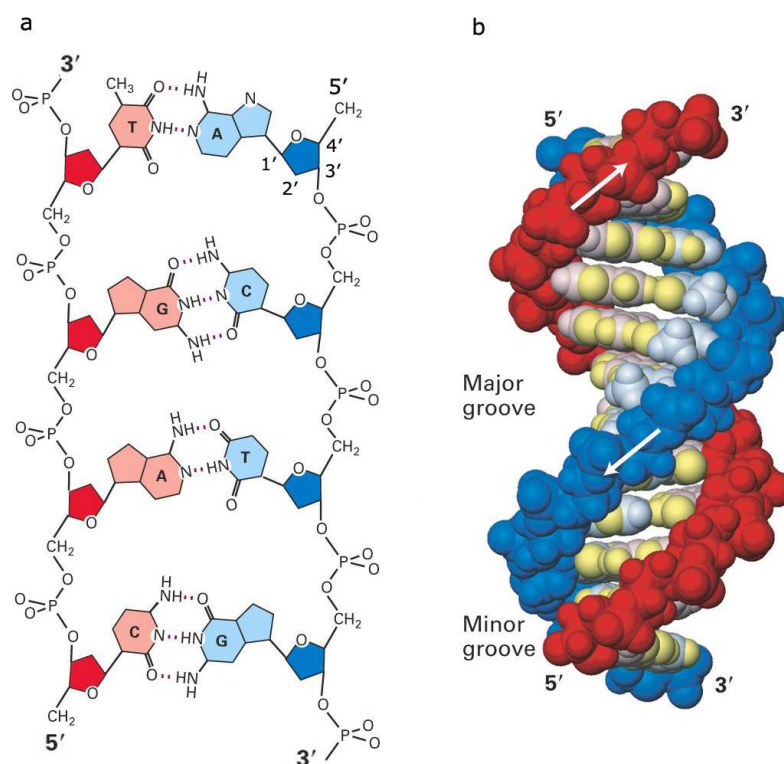


Figure 1.4: The structure of double-stranded DNA: (a) DNA consists of two polymers (strands) that are linked together by hydrogen bonds between bases. A can pair with T, while C can pair with G. (b) They are wrapped around each other to form a double helix. Image from [8].

This terminology is also used to describe enzyme motion along the contour of the DNA: an arbitrary enzyme is said to move in the $5' \rightarrow 3'$ direction, or vice versa. The chemical structures of the four bases enable the formation of hydrogen bonds with each other, which, together with "stacking" interactions between the bases, is the basis for the staircase structure of DNA, shown in Fig. 1.4b. The double helical structure was elucidated using X-ray crystallography experiments and first reported by James Watson and Francis Crick in 1953 [10]. However, much of the crystallography data originated with Rosalind Franklin. Due to the base pairing, the information stored in the DNA polymer is redundant: one strand carries the complementary piece of information of the other strand. This redundancy is exploited *e.g.* upon duplication of DNA prior to cell division or as a backup mechanism in case of DNA damage. Having DNA molecules in double-stranded form is thus of crucial importance for the viability of the cell. However, the helical structure, which is apparently the most stable structure in which to realize double strandedness also poses topological challenges to the cell, as we will now

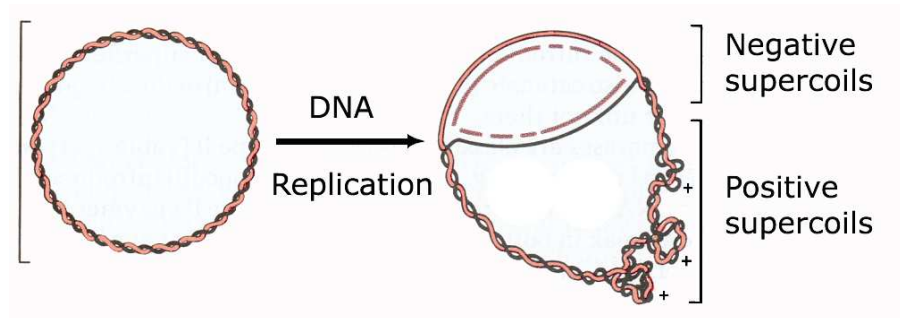


Figure 1.5: DNA replication introduces DNA supercoiling. In DNA replication, the helical DNA is locally unwound by DNA polymerases (enzymes not shown), which creates a region of negative supercoils. As DNA is often topologically constrained, meaning that the DNA's linking number (Lk) is fixed (as is naturally the case for circular DNA, shown here), negative supercoiling will be compensated by positive (overwound) supercoils.

discuss.

1.5 Cellular processes induce DNA supercoiling

Unlike a single DNA strand that contains many chemical single bonds whose rotation is not associated with significant increase in internal energy - they are "freely" rotatable - a double-stranded DNA molecule accumulates torque when twisted. Twisting of DNA is a result of several processes in the cell, such as DNA replication and DNA transcription. In DNA replication, an enzyme called DNA polymerase (DNAP) sequentially reads off the information stored in the DNA bases while polymerizing a complementary DNA copy at both strands. In this process, DNAP tracks the helix and thus rotates relative to the DNA. The size of this DNAP complex prohibits its rotation, and thus the double-stranded DNA will accumulate torque and will start to rotate about its axis. Many organisms, including humans, have their DNA in linear form. *In vivo*, however, there are several factors that resist rotation of the molecule over its entire contour length. First, the rotational drag of a very long and curved DNA molecule is quite large, as will be the topic of experimental and quantitative discussion in chapter 3. Second, many protein complexes are bound to DNA that will increase resistance to rotation. Third, the DNA is anchored at several points to the nuclear matrix, a fiber structure in the nucleus of a cell. As a result of these three factors, the torque generated by the advancing replication machinery is not fully dissipated through rotation of the entire DNA, but continuously accumulates ahead of the machinery

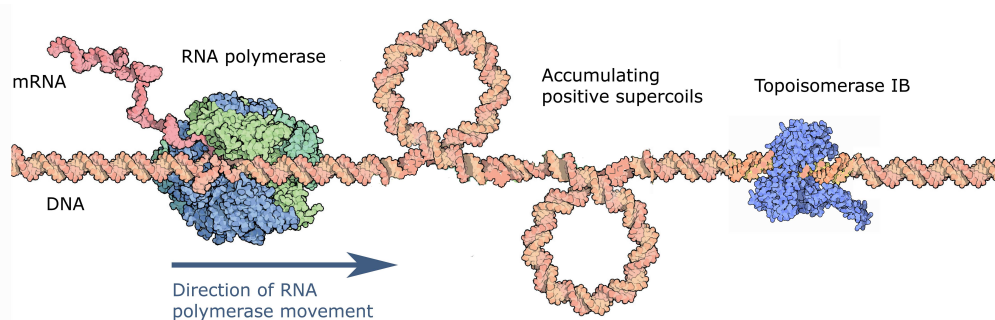


Figure 1.6: In transcription, RNA polymerase synthesizes a mRNA molecule from template DNA. As it does so, it tracks the DNA helix. However, rotational drag of the RNA polymerase complex prevents its rotation around the DNA and therefore, the DNA itself will twist. Positive supercoils (overwinding of the DNA) will accumulate ahead of the moving RNA polymerase. Topoisomerase IB is an enzyme that can remove supercoils and plays an important role in facilitating transcription. Figure adapted from RSCB protein data bank.

as replication progresses. This torque leads to the generation of positive supercoils (overwinding of the DNA double helix) ahead of DNAP, while the newly formed daughter DNA strands are negatively supercoils (underwound), as shown in Fig. 1.5.

In transcription, an enzyme called ribonucleic acid (RNA) polymerase (RNAP) synthesizes an RNA copy of the template DNA while tracking the DNA duplex. In the local vicinity of the RNAP, the DNA is denatured (breaking of the hydrogen bonds connecting the bases). This structure is called the "transcription bubble". As in the case of DNAP, RNAP and the nascent RNA (which itself may contain proteins bound to it) constitute a hydrodynamic drag that is sufficiently large to prevent their rotation around the DNA. Thus, ahead of the transcription bubble, the DNA gets positively supercoiled as transcription progresses, while compensatory negative supercoils (underwinding of the DNA) are left in its wake. This topological structure is called the "twin supercoiled" domain and consists of positive and negative supercoils [11]. Fig. 1.6 shows a piece of DNA ahead of an advancing RNAP in which positive supercoils accumulate. As transcription proceeds, more and more positive supercoils accumulate ahead of the moving RNAP. DNA replication and transcription are hampered by the overwinding of the DNA that these processes generate and it is therefore of crucial importance for a cell to dissipate this overwinding. In principle, this can occur via at least two mechanisms: by annihilation of positive and negative supercoils during transcription of a circular DNA molecule or by enzymatic regulation by topoisomerases that can change the linking number of DNA and remove DNA

supercoiling. The first mechanism, which is crucially dependent upon the dynamics of supercoil propagation in DNA molecules is the subject of single-molecule experimental work described in chapter 3. The second mechanism is the subject of experimental investigation in chapters 4 and 6.

Throughout this dissertation, we frequently describe the topology of DNA quantitatively and thus we shall introduce the nomenclature for DNA topology here. The nomenclature of the DNA holds for topologically constrained molecules, *i.e.* DNA molecules that can not freely rotate about their axis and whose ends are constrained from rotating, either because the ends are physically tethered to fixed surface, as is the case in the experiments described below, or because the ends are linked to one another, in the case of a covalently closed circular molecule. The topology of DNA can be described by three quantities: the linking number (Lk), the twist (Tw) and the writhe (Wr). The Tw of DNA is defined by the total number of times DNA of a given contour length twists around its contour axis. The Wr is defined by the total number of times the contour axis of the DNA crosses over itself. The Lk of DNA is the global topological quantity of DNA and is defined by

$$Lk = Tw + Wr \quad (1.1)$$

In topologically constrained molecules, where Lk is fixed, Tw and Wr can be interconverted. Note that canonical DNA (B-form DNA) has a natural twist of 1 turn per 10.4 base pairs and thus, in the absence of Wr, torsionally relaxed DNA of 23 kilobases (kb) has a natural Lk, called Lk_0 , that is equal to $23,000/10.4=2190$ turns.

1.6 Topoisomerases remove DNA supercoiling

Ever since the native structure of DNA was shown to be double-stranded, the potentially adverse topological implications occurring during transcription and replication have been recognized. Therefore, enzymes capable of regulating supercoils, *i.e.* enzymes that can alter the linking number of DNA *in vivo* were anticipated even before they were discovered by James C. Wang in 1971. Topoisomerases were later shown to be ubiquitous enzymes, conserved throughout all kingdoms of life [11–16]. Interestingly, not all topoisomerases have similar structures and functions in the cell. They can be divided into two main categories: those that change Lk by cutting one of the two DNA strands in the duplex (Type I and III topoisomerases etc.), and those that cut both strands of the duplex (Type II and IV topoisomerases). All topoisomerases, however, cleave their DNA sub-

strate by forming a transient phosphodiester bond between a tyrosine residue (side group) in their active site and the broken DNA end. Type I topoisomerases are also divided into type IA (attaching to the 5'-end of the DNA) or type IB (attaching to the 3'-end of the DNA). It is important to note at this point that the chemistry involved in this enzymatic reaction precludes detachment of the enzyme from the DNA when the DNA is nicked. In other words, preventing the rejoining step of the two loose ends of the DNA (religation) effectively covalently traps the topoisomerase on the DNA. This covalent trapping by definition persists over the same timescale as the prevention of religation.

We now discuss the most prominent cellular roles of topoisomerases in *Saccharomyces cerevisiae* as this yeast is the system in which our *in vivo* experiments have been carried out [12–16]. In yeast, three types of topoisomerases are expressed: a topoisomerase I (Type IB), a topoisomerase II (Type IIA) and a topoisomerase III (Type IA). The role of topoisomerase I is to remove the excess positive supercoiling that is generated ahead of the transcription bubble and ahead of the replication fork. As such, topoisomerase I acts as a "swivel" to effectively facilitate these processes. Although its predominant role *in vivo* is to remove positive supercoils, mechanistically topoisomerase I is able to relax both positive and negative supercoils, and *in vitro* it readily does both. Topoisomerase II acts prior to cell division and its presence is absolutely required for proper chromosome segregation during mitosis and thus for cell proliferation. Its function is to untangle the newly synthesized duplex daughter DNA molecules that are intertwined. As topoisomerase II can in principle remove positive as well as negative supercoils, it is reasonable to expect that topoisomerase II will support fork progression together with topoisomerase I. The extent to which topoisomerase II is involved in this particular activity is unclear, although there are indications that its role is fairly minor relative to that of topoisomerase I. Nonetheless, while in higher eukaryotes, such as humans, the presence of functional topoisomerase I is indispensable for cell growth, in yeast this is not the case. Its activity can be compensated by that of topoisomerase II.

The mechanism by which supercoils are removed has been elucidated over the years by a combination of X-ray crystallography [17–20], molecular dynamics simulations [21], supercoil relaxation assays using gel electrophoresis *e.g.* [12, 13, 22–28] and, more recently, by single-molecule experiments [29–31]. We refrain here from providing an in-depth description of the rather sophisticated mechanisms of action of Type IA and Type II topoisomerases but refer instead to [13]. Briefly, Type IA topoisomerases link to the 5-end of DNA strand covalently, but bridge the gap in the single DNA strand by holding on non-covalently to the 3'-end as well. Following cleavage, a gap in the enzyme is created by a conformational

change in the enzyme. Through this gap, the intact strand is passed, after which the cleaved DNA strand is religated. Interestingly, this process does not require ATP or any other external energy factor, though does require Mg^{2+} . Type II topoisomerases do require ATP for continued relaxation of supercoils or decatenation. In a first ATP-independent step, they bind to a crossing of two DNA duplexes, and cleave both strands of one DNA duplex. Then, a conformational change of the enzyme pulls the two broken ends of this strand, called the gated segment (G-segment), apart after which the intact duplex, called the transfer segment, or T-segment, is passed through the G-segment with a rate that is greatly accelerated by ATP hydrolysis. After the transfer of the T-segment, the DNA is religated and the enzyme is reset for another round of catalysis.

1.7 Topoisomerase IB

The structure of the 91-kDa human topoisomerase IB consists of the domains: an N-terminal domain (214 amino acids), that contains sequences that direct the enzyme to the cell nucleus [13], an adjacent core-domain (421 amino acids) that contains the amino acids that are directly involved in catalysis, except for the tyrosine residue that is involved in the chemical reaction with the DNA backbone, a linker domain (77 amino acids) that is not required for catalytic activity and a C-terminal domain (53 amino acids) that contains the active-site tyrosine723. The co-crystal structure of human topoisomerase I with a DNA duplex shows that the topoisomerase has a footprint on the DNA of approximately 14 base pairs and tightly encompasses the DNA, as shown in Fig. 1.7[17]. Many structural features of human topoisomerase IB and the vaccinia poxvirus topoisomerase IB are highly similar to another family of enzymes, called site-specific tyrosine recombinases [32]. The functional core of both families is not only virtually identical, both families also operate with a topoisomerase IB-like reaction mechanism to cleave and religate DNA: tyrosine recombinases, like topoisomerases transiently expunge a free 5'-OH group while transiently establishing a phosphotyrosine linkage to the DNA. Recombinases join different DNA duplexes covalently by ligating the 5'-OH end of one DNA duplex to the 3'-end of another DNA duplex, while topoisomerases perform the same chemistry on the same DNA duplex. Indeed, most recombinases are able to perform topoisomerase activity *in vitro*, while vaccinia topoisomerase IB has been shown to efficiently catalyze Holliday-junction resolution. As such, topoisomerase IB is an example of an enzyme that displays catalytic promiscuity. Enzyme promiscuity is considered to be an important mechanism in the evolution of novel functionality in enzymes [33]. It is an in-

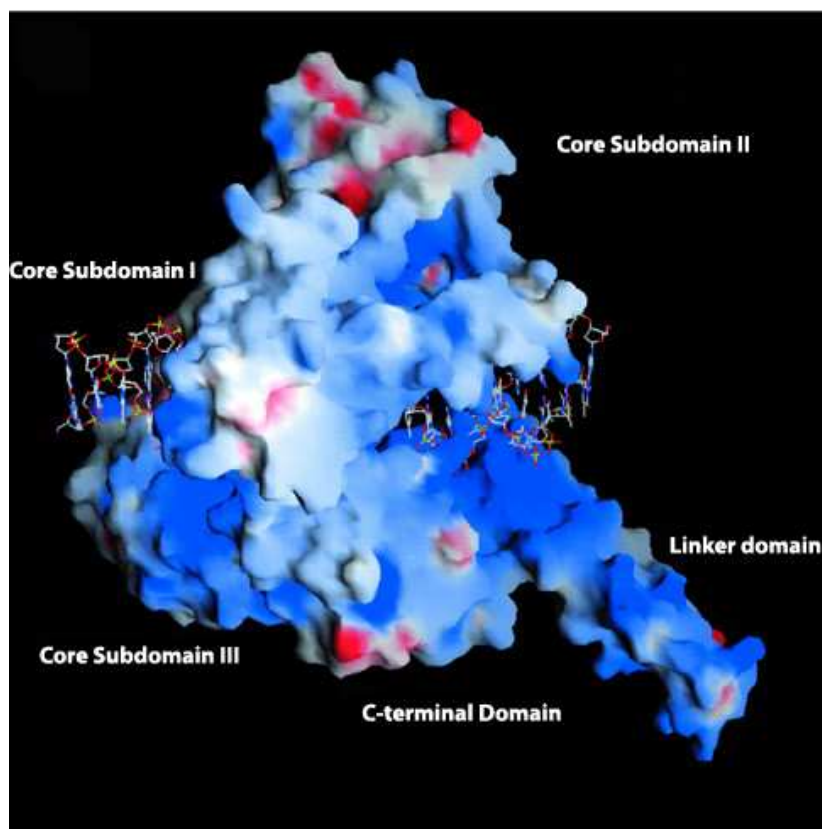


Figure 1.7: The crystal structure of human topoisomerase IB in complex with a 22-bp double-stranded DNA molecule. The topoisomerase shown here is the 70-kD amino terminally truncated form [17].

interesting speculation that a common enzyme ancestor of topoisomerase IB and site-specific tyrosine recombinase was capable of basal nicking-religation activity. From that ancestor, both the site-specific tyrosine recombinases and Type IB topoisomerases have evolved to fine-tune and narrow down their catalytic repertoire to efficiently catalyze recombination and topoisomerisation respectively [32].

The availability of the crystal structure of human topoisomerase IB has certainly contributed to an understanding of parts of the mechanism in atomic detail and thus led to a significant increase in understanding of the type IB topoisomerases in general. However, already before its appearance, the mechanism was studied and understood on a more coarse-grained level. Oligonucleotide cleavage and religation assays using vaccinia topoisomerases were used in these studies because of this enzyme's relatively large preference (compared to human topoisomerase) for site-specific cleavage of the DNA [24–27, 34, 35]. The site-specific cleavage facilitated the effective design of so-called suicide DNA constructs that

enabled detailed kinetic studies of either cleavage or religation reactions in isolation. Using the thus-obtained rate constants for cleavage and religation, plasmid relaxation assays were analyzed and a mechanism of action for supercoil relaxation was proposed. Stivers et al. [24] proposed a free rotation mechanism in which the DNA is allowed to swivel about its intact strand as long as topoisomerase is covalently linked to the DNA, thus removing supercoils. Under plasmid relaxation conditions, an average of 5 supercoils were found to be released prior to religation. However, what remains in the dark is the functioning of the swivel, *i.e.* an answer to the following questions: 1. How much does the presence of the protein surface allow for DNA rotation inside the enzyme cavity? This is an interesting question, as the tight wrap of topoisomerase that was shown by crystallography experiments was thought to hinder the rotation of DNA [17], especially given that the cross-sectional diameter of DNA increases roughly to 4 nm, instead of its canonical 2 nm, when the 5-end is misaligned with the 3'-end by 180 degrees [17, 18, 21, 22, 31, 36]. 2. What is the influence of torque on the religation and thus on the number of supercoils removed per cleavage-religation cycle? 3. On a more conceptual level, can one effectively incorporate the notion of friction between an enzyme and its substrate in the description of enzymatic activity? 4. How is the swivel affected by chemotherapeutic drugs that are known to act on human topoisomerase IB specifically and what are the consequences for our understanding of the mechanism of action of these drugs?

1.8 Overview of this dissertation

This dissertation has a large component of experimental study at the level of individual DNA molecules, but also contains the expansion of mathematical tools to analyze single-molecule experiments.

Chapter 2 provides a more technical description of the capabilities and limitations of magnetic tweezers. We give an overview of the different strategies in which one can calculate the force that the magnets exert on the magnetic bead. We also focus on the relation between temporal and spatial resolution.

Chapter 3 describes the physics of stretching a single linear DNA molecule due to an instantaneous increase in the stretching force, for both relaxed DNA molecules and supercoiled DNA molecules. These experiments are carried out in the absence of proteins and yield an upper limit for the timescales at which DNA can be uncoiled and stretched in the magnetic tweezers experiments. These experiments also answer the question whether the hydrodynamic drag of rotating and translating DNA molecules through water needs to be taken into account

in the context of these experiments. We find that equilibrium dynamics can predict all salient features we experimentally observe and thus can be used as a quantitative baseline for the study of more complex DNA dynamics that include proteins.

Chapter 4 describes the mechanism of DNA supercoil relaxation by vaccinia and human topoisomerase IB. Here we observe that the uncoiling of DNA is hindered due to the presence of the topoisomerase. Our data on various kinds of topoisomerases, combined with previous information from crystallography experiments, suggest that the tight clamp that the topoisomerase forms around the DNA during uncoiling hinders the uncoiling. We also observe that not all supercoils are removed in a single cleavage-religation cycle and find that the number of supercoils removed each enzymatic event is distributed exponentially, indicative of a stochastic process. We study the torque dependence of the number of supercoils that is removed by the enzyme and can rationalize our findings using a model that includes the hindrance in the rotation.

Chapter 5 shows how constraints in a single-molecule experiments, *e.g.* in the topoisomerase-mediated uncoiling of supercoils, can introduce large biases in the estimation of parameters used for modeling the system. We reasoned that as single-molecule measurements are often distributed quite broadly, the constraints on the measurements are broadly distributed as well. Therefore, we developed a maximum likelihood method that incorporates a bias for each single measurement separately. We show that with such "dynamic" treatment of constraints, the parameters of the measured distributions can be corrected perfectly, even when the bias initially exceeded 100%. Furthermore, we show examples of common experiments in single-molecule biophysics that may be hampered by such constraints and are thus likely to benefit from our analysis method.

Chapter 6 describes how the swivel mechanism by which human topoisomerase IB removes supercoils is affected by the camptothecin class of chemotherapeutic drugs that is in current clinical use against a variety of cancers. Surprisingly, we find that the drug significantly impedes the rotation of DNA, with a more pronounced effect for the removal of positive supercoils than for negative supercoils. Using the slow uncoiling as a signature of a single drug molecule bound to the topoisomerase-DNA complex, we measure the timescales over which the drug remains bound to be approximately 2 minutes. During this time, we observed no religation of the DNA, while in the absence of drug, we observed religation after typically 0.3 seconds. Therefore, we conclude that the drug prevents religation and thus covalently traps topoisomerase on the DNA, where it may stall the advancing replication fork, eventually killing the cell. To investigate the implication of the observed slowdown in uncoiling of positive supercoils for

a living cell, we moved to *in vivo* studies in yeast. There, we asked if positive supercoils indeed accumulate during replication and transcription. The results of these *in vivo* experiments showed that they indeed do. However, cells expressing a drug resistant topoisomerase exhibit no such accumulation of positive supercoils. With this unique combination of complementary single-molecule and *in vivo* experiments we highlight the role of positive supercoil accumulation in drug poisoning, shedding new light on the mechanism of action of the drug.

Chapter 7 describes experiments performed using atomic force microscopy (AFM) on the binding of vaccinia virus topoisomerase IB on DNA molecules. The focus of this work was to visualize the formation of filaments of topoisomerase IB. By quantitative analysis, we show that the filament formation proceeds in a highly cooperative fashion.

Chapter 8 presents ongoing work towards the single-molecule investigation of the kinetics of telomerase using a combination of single-molecule fluorescence and nanofabricated structures. This work, while clearly separate from the force spectroscopy experiments on topoisomerase, is included in this dissertation because it is an interesting experiment from both a biological and a technological point of view that I have spent a significant amount of time on.

References

- [1] Bustamante, C., Bryant, Z., and Smith, S. B. *Nature* **421**(6921), 423–7 (2003).
- [2] Strick, T., Allemand, J., Croquette, V., and Bensimon, D. *Prog Biophys Mol Biol* **74**(1-2), 115–40 (2000).
- [3] Smith, S. B., Finzi, L., and Bustamante, C. *Science* **258**(5085), 1122–6 (1992).
- [4] Yin, H., Wang, M. D., Svoboda, K., Landick, R., Block, S. M., and Gelles, J. *Science* **270**(5242), 1653–7 (1995).
- [5] Visscher, K., Schnitzer, M. J., and Block, S. M. *Nature* **400**(6740), 184–9 (1999).
- [6] Dogterom, M., Kerssemakers, J. W. J., Romet-Lemonne, G., and Janson, M. E. *Curr Opin Cell Biol* **17**(1), 67–74 (2005).
- [7] Weiss, S. *Science* **283**(5408), 1676–83 (1999).
- [8] Lodish, e. a. *Molecular Cell Biology*. W.H. Freeman, (2004).

- [9] Nelson, P. C. *Biological Physics*. W.H. Freeman and Company, New-York, (2004).
 - [10] Watson, J. D. and Crick, F. H. *Nature* **171**(4356), 737–8 (1953).
 - [11] Giaever, G. N. and Wang, J. C. *Cell* **55**(5), 849–56 (1988).
 - [12] Bjornsti, M. A. and Osheroff, N. *Methods Mol Biol* **94**, 1–8 (1999).
 - [13] Champoux, J. J. *Annu Rev Biochem* **70**, 369–413 (2001).
 - [14] Corbett, K. D. and Berger, J. M. *Annu Rev Biophys Biomol Struct* **33**, 95–118 (2004).
 - [15] Wang, J. C. *Annu Rev Biochem* **65**, 635–92 (1996).
 - [16] Wang, J. C. *Nat Rev Mol Cell Biol* **3**(6), 430–40 (2002).
 - [17] Stewart, L., Redinbo, M. R., Qiu, X., Hol, W. G., and Champoux, J. J. *Science* **279**(5356), 1534–41 (1998).
 - [18] Redinbo, M. R., Stewart, L., Kuhn, P., Champoux, J. J., and Hol, W. G. *Science* **279**(5356), 1504–13 (1998).
 - [19] Redinbo, M. R., Stewart, L., Champoux, J. J., and Hol, W. G. *J Mol Biol* **292**(3), 685–96 (1999).
 - [20] Chrencik, J. E., Staker, B. L., Burgin, A. B., Pourquier, P., Pommier, Y., Stewart, L., and Redinbo, M. R. *J Mol Biol* **339**(4), 773–84 (2004).
 - [21] Sari, L. and Andricioaei, I. *Nucleic Acids Res* **33**(20), 6621–34 (2005).
 - [22] Woo, M. H., Losasso, C., Guo, H., Pattarello, L., Benedetti, P., and Bjornsti, M. A. *Proc Natl Acad Sci U S A* **100**(24), 13767–72 (2003).
 - [23] Brown, P. O. and Cozzarelli, N. R. *Proc Natl Acad Sci U S A* **78**(2), 843–7 (1981).
 - [24] Stivers, J. T., Harris, T. K., and Mildvan, A. S. *Biochemistry* **36**(17), 5212–22 (1997).
 - [25] Stivers, J. T., Shuman, S., and Mildvan, A. S. *Biochemistry* **33**(51), 15449–58 (1994).
 - [26] Stivers, J. T., Shuman, S., and Mildvan, A. S. *Biochemistry* **33**(1), 327–39 (1994).
 - [27] Shuman, S., Golder, M., and Moss, B. *J Biol Chem* **263**(31), 16401–7 (1988).
 - [28] McClendon, A. K., Rodriguez, A. C., and Osheroff, N. *J Biol Chem* **280**(47), 39337–45 (2005).
-

-
- [29] Strick, T. R., Croquette, V., and Bensimon, D. *Nature* **404**(6780), 901–4 (2000).
 - [30] Dekker, N. H., Rybenkov, V. V., Duguet, M., Crisona, N. J., Cozzarelli, N. R., Bensimon, D., and Croquette, V. *Proc Natl Acad Sci U S A* **99**(19), 12126–31 (2002).
 - [31] Koster, D. A., Croquette, V., Dekker, C., Shuman, S., and Dekker, N. H. *Nature* **434**(7033), 671–4 (2005).
 - [32] Cheng, C., Kussie, P., Pavletich, N., and Shuman, S. *Cell* **92**(6), 841–50 (1998).
 - [33] Khersonsky, O., Roodveldt, C., and Tawfik, D. S. *Curr Opin Chem Biol* **10**(5), 498–508 (2006).
 - [34] Sekiguchi, J. and Shuman, S. *Embo J* **15**(13), 3448–57 (1996).
 - [35] Shuman, S. *J Biol Chem* **266**(17), 11372–9 (1991).
 - [36] Carey, J. F., Schultz, S. J., Sisson, L., Fazzio, T. G., and Champoux, J. J. *Proc Natl Acad Sci U S A* **100**(10), 5640–5 (2003).
-

Chapter 2

Magnetic tweezers

This chapter provides a technical description of the magnetic tweezers, a description of their temporal and spatial resolution, and technical limitations of the technique. A number of authors have in the past already provided technical descriptions of this apparatus, and we have therefore chosen to discuss only those subjects that are either, to our knowledge, not reported in the literature, or are specific to our particular experiment.

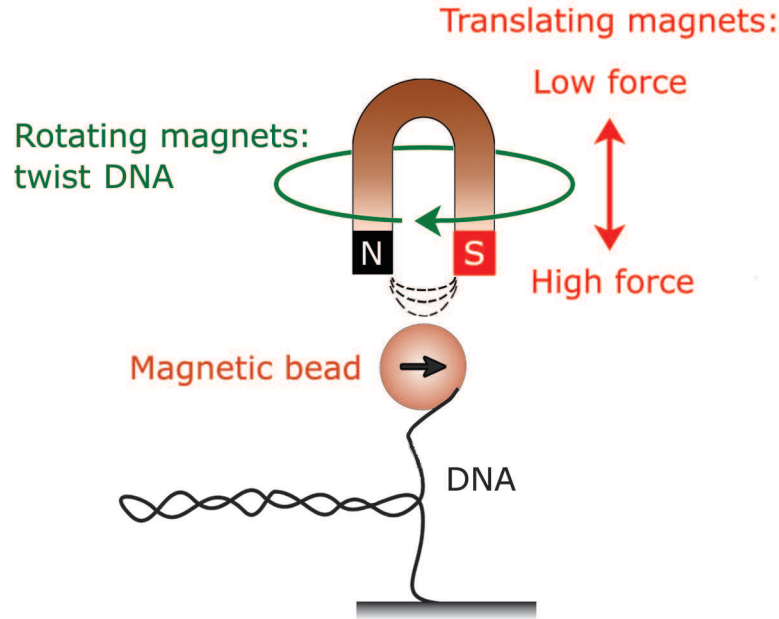


Figure 2.1: Schematic drawing of the magnetic tweezers setup. A pair of magnets exert a force on a magnetic bead, which is attached to one end of a single DNA molecule. The other end is attached to a glass surface. Increasing the distance between the magnets and the bead decreases the magnetic stretching force that is exerted on the DNA, and vice versa. Rotating the magnets and thus rotating the bead twists the DNA. Using magnetic tweezers, one can thus precisely control the topological state of a single DNA molecule.

2.1 Measuring forces in real space

The core of the magnetic tweezers is a single linear DNA molecule that is coated with multiple digoxigenin molecules at one end and with multiple biotin molecules at the other end. The digoxigenin-coated end connects in a non-covalent fashion with an anti-digoxigenin coated glass surface, while a streptavidin-coated spherical μm -size magnetic bead connects to the other end of the DNA molecule. By placing a pair of magnets (in our setup permanent magnets) above the flow cell in which the DNA molecule is immersed, we exert a magnetic force F upward. F will of course depend on the distance between the magnets and the magnetic bead. The setup is schematically drawn in Fig. 2.1. The forces that one applies also depend of the quantity of magnetic particles that the magnetic bead contains, which scales with the diameter of the bead. For a combination of a Dynal bead of $1\ \mu\text{m}$ in diameter, which is the bead size mostly used in our experiments, and neodymium magnets (a combination of neodymium, iron and boron), we exert a maximum force of order 5 pN. In theory there is no minimum force one can exert

magnetically, as the magnets can be arbitrarily far removed from the bead. In practice, however, forces of order 10 fN are considered to be the lowest workable forces. We observe the bead with an ordinary light microscope and can determine its 3-dimensional position in time (at 120 Hz for a single bead) to about 5 nm accuracy in the z-direction and about 5 nm in the $(x-y)$ -plane.

We will now consider the fluctuations of the bead under a force F , as we will measure the magnitude of the force exerted by the magnets using these fluctuations. The bead-DNA system can be thought of as an inverted pendulum, where the deviations δx from the bead's equilibrium position are driven by thermal agitation. When the bead deviates from its equilibrium position, the DNA with end-to-end distance l then generates a force $F\delta\theta$ to return it to its center position, where θ is the angular deviation of the DNA molecule measured at the anchoring point of the DNA at the glass. For small angular displacements ($\sin\theta = \theta$), the restoring force equals $\frac{F}{l}\delta x$. These dynamics are described by a spring with a spring constant $k_x = \frac{F}{l}$. As this spring is powered by thermal agitation that has $\frac{1}{2}k_B T$ available for each degree of freedom, we can thus write

$$\frac{1}{2}k_x \langle \delta x^2 \rangle = \frac{1}{2}k_B T, \quad (2.1)$$

where k_B is the Boltzmann constant and T the temperature in K. Thus, we obtain

$$F = \frac{k_B T l}{\langle \delta x^2 \rangle}. \quad (2.2)$$

This equation demonstrates that we only need to measure the variance of the bead excursions in time in one dimension in the (x, y) -plane, while simultaneously measuring l to obtain a measurement of the force F , provided that one measures longer than the characteristic frequency of the system. Such a measurement is shown in Fig.2.2, for two different forces: low force (green) and higher force (red). Qualitatively, one notices that as the stretching force increases, the variance of the excursions is reduced, as expected from equation 2.2.

2.2 Measuring forces in Fourier space

In practice, however, we do not measure the force using a measurement in real space. Instead, we analyze the fluctuations in Fourier space for a number of reasons. First, it provides a measure of the characteristic frequency of the system, called the cut-off frequency (f_c), which enables us to quantitatively assess whether we are in fact probing the expected Gaussian distribution of bead ex-

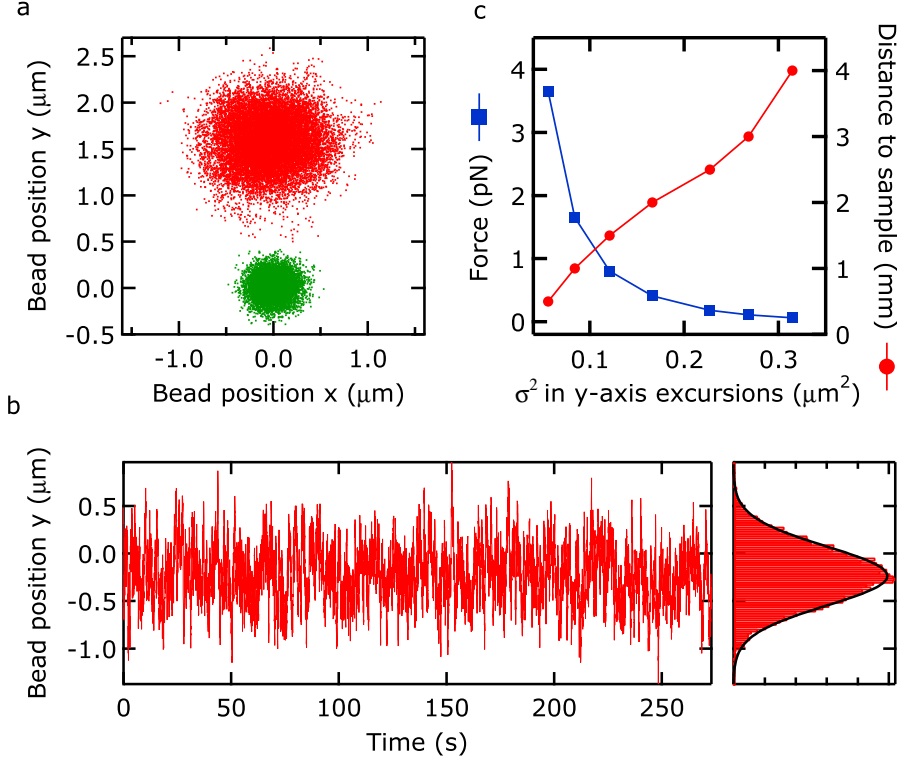


Figure 2.2: Measuring the stretching force using the bead's fluctuations in real-space. (a) Excursions of the bead center in the (x, y) -plane for low force (red) and higher force (green). (b) Bead excursion in a single dimension in time (low force measurement is shown), with a histogram on the right, including a Gaussian fit to the histogram. (c) Computed variance of the bead excursions (not obtained from a fit to the Gaussian) for a variety of magnet positions (red solid circles) as well as the calculated corresponding forces (blue solid squares).

cursions instead of a skewed distribution as a result of a measurement duration that is shorter than the correlation time of the motion. Second, we can filter low frequency drift in the bead positions, although we find that a differential measurement of two beads, one with a DNA tether and one fixed on the microscope slide is effective at removing the effects of drift from the measurement, both in the z and (x, y) -plane. Third, the camera has a finite time over which the optical image is integrated. This will lead to an overestimation of the measured force, as we show below. This effect can be countered in the Fourier domain. Finally, to obtain the correct functional form of the spectrum of bead fluctuations, we need to perform an anti-aliasing operation to correct for spectral folding of the camera.

The analysis in frequency space is shown in Fig. 2.3. Fig. 2.3c shows that the

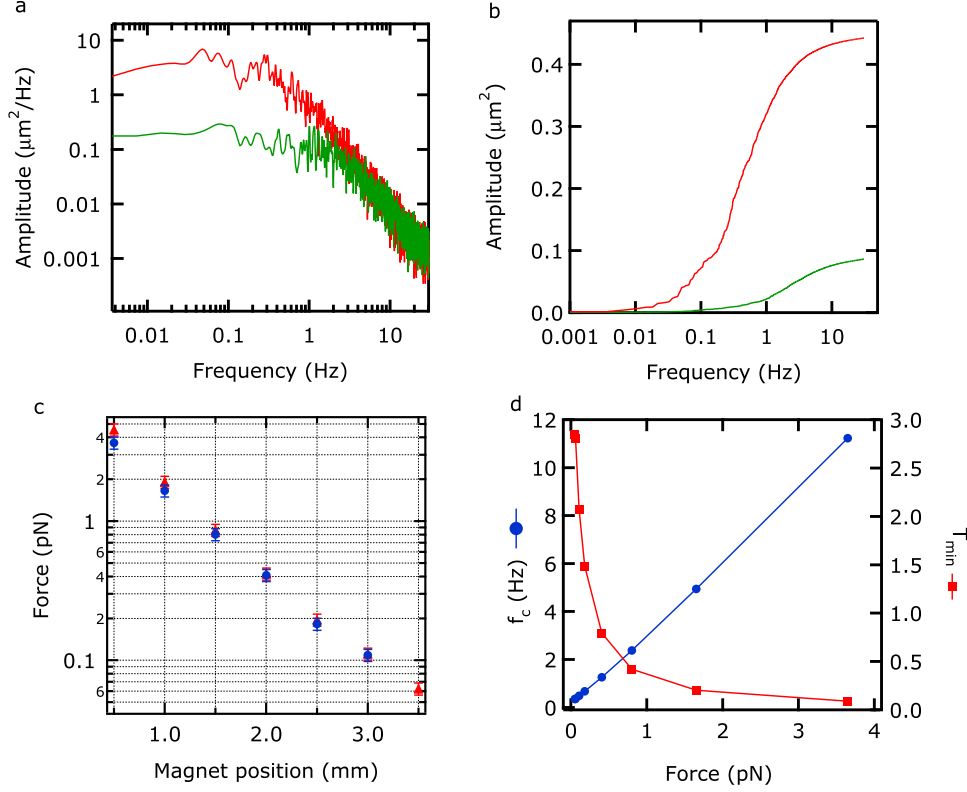


Figure 2.3: Measuring the spectrum of the bead's fluctuations. (a) Spectra of the one-dimensional fluctuations for the same two forces as shown in Fig. 2.2a. The spectrum is flat until the cut-off frequency f_c (see main text), after which it decays off with $1/f^2$. Note that f_c increases with stretching force. (b) The integral of the spectra shown in (a); the variance of the bead excursions can be obtained by integrating the spectra. Note that the variance increases with decreasing stretching force. (c) shows a comparison between the force measurement in real-space (red solid triangles) and fourier-space (blue solid circles). Note that in the low force regime (lower than 1 pN), the two methods yield similar results. Only at higher forces, where the fluctuations of the bead are faster, one does run into the problem of underestimating the variance of the bead fluctuations, leading to an overestimation in the force, according to equation 2.2. This filtering effect due to the integration time of the camera can be corrected for in fourier-space only. (d) shows the cut-off frequency, f_c , as a function of stretching force F . It also shows $1/f_c$, the time one has to minimally average the signal to sample the entire excursion distribution and obtain a statistically unbiased estimate for its center.

measurement in real-space (solid red triangles) yields values for the force that are indistinguishable from the measurements in Fourier space for forces lower than approximately 1 pN for a bead with a diameter of 1 μm (the error on the force measurement is app. 5%). However, for higher forces, the discrepancy between the two measurements becomes apparent: the real-space measurement yields force values that are higher than the measurement in Fourier space (solid blue circles). This discrepancy is caused by the integration time of the camera that cannot be effectively countered by in real space analysis compared to Fourier space analysis. As the bead fluctuations are recorded by the camera, their extreme positions are not well captured due to the averaging. Consequently, the calculation of the bead's variance in real space is underestimated, and thus its force overestimated, according to equation 2.2.

To understand the expected functional form for the spectrum of bead fluctuations, we consider the equation of motion of the bead-DNA system.

$$m \frac{d^2 x(t)}{dt^2} = -6\pi\eta R \frac{dx(t)}{dt} - k_x x(t) + f_L(t), \quad (2.3)$$

where m is the mass of the bead, $x(t)$ is bead position in time, η the viscosity of water (10^{-3} Nsm^{-2}), R the radius of the bead, k_x the effective trap stiffness of the bead-DNA pendulum in the transverse direction and f_L a Langevin force caused by water molecules transferring their momentum to the bead. The actual force in time is a quantity that stochastically fluctuates with a characteristic time scale of $\tau_{\text{collision}}$, set by the frequency at which the water molecules collide with the bead. Since water molecules have equal probability to collide with the bead from all directions, it has the following property:

$$\langle f_L(t) \rangle = 0, \quad (2.4)$$

where the brackets denote the time average. If we assume $\tau_{\text{collision}}$ to be very short, we can approximate the correlation function as a delta-function:

$$\langle f_L(t) \cdot f_L(t') \rangle = 4k_B T \Gamma \delta(t - t'), \quad (2.5)$$

where Γ is a numerical factor equal to $6\pi\eta R$ and $\delta(t - t')$ is a delta function. The power spectral density of the Langevin force is related to the autocorrelation function and is given by [1, 2]:

$$|F_L(\omega)|^2 = 4k_B T \Gamma, \quad (2.6)$$

in units of $[\text{N}^2 \cdot (\text{Hz})^{-1}]$. The square-root of Eq. 2.6:

$$F_L(\omega) = \sqrt{4k_B T \Gamma}, \quad (2.7)$$

in units of $[\text{N} \cdot (\text{Hz})^{-1/2}]$ gives the root-mean-square of the Langevin force per unit of $\sqrt{Hz}$. To calculate the force that the bead experiences due to the bombardment of the water molecules, in a given time slice Δf , one thus multiplies this quantity with $\sqrt{\Delta f}$, *i.e.*

$$\int_0^{\sqrt{\Delta f}} F_L(\omega) d\omega = \sqrt{4k_B T \Gamma \Delta f}, \quad (2.8)$$

For example, in the bandwidth of 1Hz, the root-mean-square Langevin force on a bead with $R=0.5\mu\text{m}$, $\eta=10^{-3} \text{ N} \cdot \text{s} \cdot (\text{m})^{-2}$ and at room temperature equals approximately 12 fN. We apply a Fourier transform to the equation of motion (equation 2.3) and obtain:

$$X(\omega) = \frac{F_L(\omega)}{-m\omega^2 + 6\pi i \eta R \omega + k_x}, \quad (2.9)$$

where $i = \sqrt{-1}$. We have used the following definition of the Fourier transform:

$$A(\omega) = \int_{-\infty}^{\infty} a(t) e^{-i\omega t} dt \quad (2.10)$$

and

$$a(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} A(\omega) e^{i\omega t} d\omega \quad (2.11)$$

The square of the amplitude is given by

$$X^2(\omega) = \frac{|F_L(\omega)|^2}{k_x^2} \frac{1}{1 + (\frac{\omega}{\omega_c})^2} = \frac{24\pi k_B T \eta R}{k_x^2} \frac{1}{1 + (\frac{\omega}{\omega_c})^2}, \quad (2.12)$$

where $\omega_c = \frac{k_x}{6\pi\eta R}$. Equation 2.12 is a Lorentzian, which is thus the functional form to expect from the spectral measurement, as we set out to show above. One obtains $\omega_c = 2\pi f_c$ by fitting the power spectrum of the measured time series by this equation with a single parameter and can compute the k_x directly. However, one is required to assume a bead radius R that needs to be plugged in. One can alternatively fit the power spectrum with two fitting parameters ω_c and R . The measurement of R is a nice and healthy sanity check as it better be close to the

radius of the beads that were bought. Instead of fitting the spectrum, one can integrate the power spectrum over all frequencies, which yields the total variance for all frequencies. This can be then used to calculate k_x and thus F .

2.3 Temporal and spatial resolution

We now discuss the accuracy with which one can measure the position of the bead in the z -direction. As enzymes indirectly induce changes in the height of the bead, the spatial resolution in the z -direction is directly related to the resolution with which one probes the enzymatic activity. Assuming that one measures long enough (*i.e.* longer than the cutoff frequency in the z -direction $1/f_{cutoff,z}$) to probe the full excursion probability distribution of the bead motion, one can in principle obtain an unbiased and infinitely small spatial resolution by measuring an infinite number of time points. For a measurement of N data points, the accuracy with which one measures the x, y or z -position is given by the standard error of the mean (SEM):

$$SEM = \frac{\sigma}{\sqrt{N}}, \quad (2.13)$$

where σ is an estimate of the standard deviation of the fluctuations in z and which is obtained using $\sigma = \sqrt{\frac{k_B T}{k_z}}$. In other words, measuring faster will only lead to an increase in spatial resolution when the sampling takes place longer than $T_{min,z} = 1/f_{cutoff,z}$ and thus $f_{cutoff,z}$ sets the temporal resolution to a large degree. The value for $\omega_{cutoff,z}$ can be obtained in a similar manner as for the x and y directions, but substituting k_x for k_z yielding $f_c = \frac{k_z}{12\pi^2\eta R}$, where k_z is given by $\frac{\partial F}{\partial l}$ as opposed to $\frac{F}{l}$ for the x and y directions. We have calculated this fraction using the relation between the stretching force F and the resulting end-to-end extension l of the double-stranded DNA that is widely applied in the field and is called the Worm-like chain. It is given by [3]:

$$F = \frac{k_B T}{\xi} \left(\frac{l}{l_0} - \frac{1}{4} + \frac{1}{4(1 - l/l_0)^2} + \sum_{i=2}^7 a_i \left(\frac{l}{l_0} \right)^i \right), \quad (2.14)$$

where $a_2 = -0.5164228$, $a_3 = -2.737418$, $a_4 = +16.07497$, $a_5 = -38.87607$, $a_6 = +39.49944$, $a_7 = -14.17718$, ξ is the persistence length of DNA, for which we have taken the generally accepted value of 50 nm [4] and l_0 is the contour length of the DNA, which for our experiments typically is 7 μm .

Fig 2.4a plots both $f_{cutoff,z}$ (in Hz) and its inverse $T_{min,z}$ (in s.) as a function

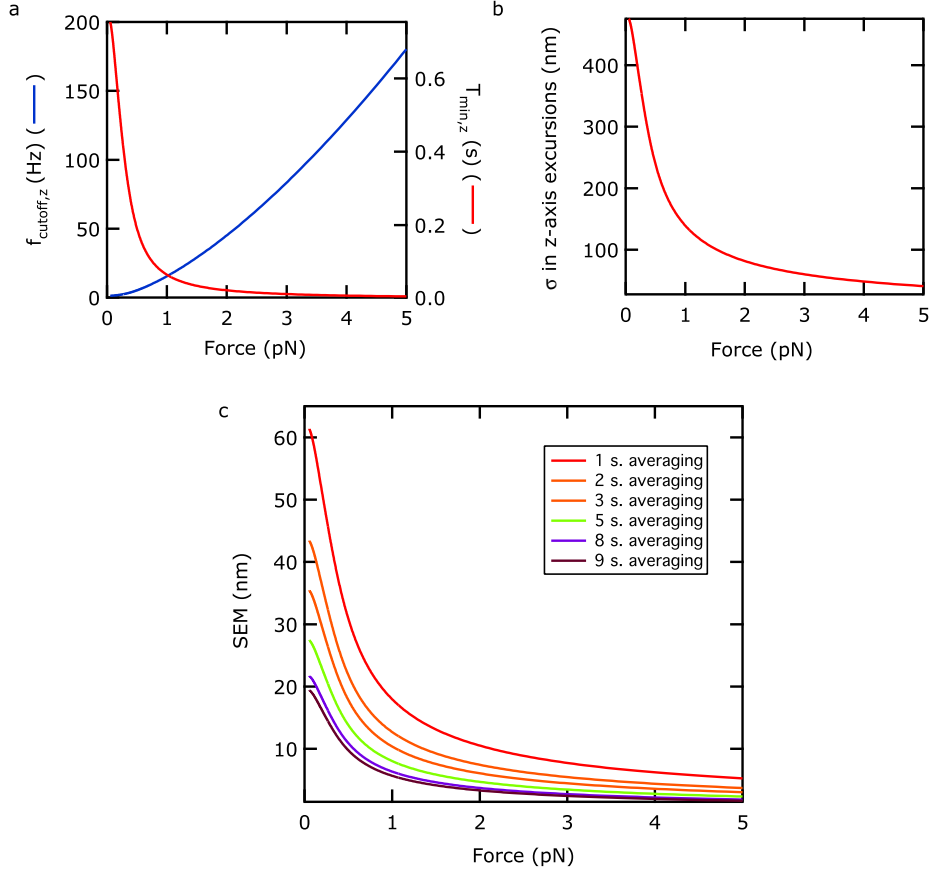


Figure 2.4: Temporal and spatial resolution of the magnetic tweezers. (a) The temporal resolution is set by the characteristic frequency of the bead excursions in the z -direction, $f_{\text{cutoff},z}$ in Hz. Its inverse, the minimum time required to sample the excursions, $T_{\text{min},z}$ is plotted as well. (b) The standard deviation of the bead excursions in the z -direction, σ decreases with increasing force as the stiffness of the DNA increases. (c) The spatial resolution of the magnetic tweezers in the z -direction, the standard error of the mean (SEM) decreases with increasing integration time and stretching force.

of the stretching force F . Fig. 2.4 thus exhibits the temporal resolution of the magnetic tweezers, *i.e.* the minimum time one should probe the excursions of the bead in the z direction. Sampling less than that time will yield a skewed average.

Given that one probes long enough, the SEM will decrease with increasing force and averaging times. This is reasonable as an increased force will decrease the σ of the fluctuations, thus decreasing the numerator in Eq. 2.13. Increasing the averaging time will increase the number of data points N , increasing the denominator. Fig. 2.4b plots σ of the bead excursions along the z axis as a

function of force. To obtain the SEM of the excursions as a function of averaging time $T_{average}$, one substitutes $N = f_{sampling} \cdot T_{average}$ in Eq. 2.13, yielding $SEM = \sqrt{\frac{k_B T}{k_z}} \frac{1}{\sqrt{f_{sampling} \cdot T_{average}}}$. Here, $f_{sampling}$ is the sampling frequency and we choose 60 Hz as this is the acquisition frequency of the camera used. The SEM as a function of stretching force is plotted in Fig. 2.4c for a number of averaging times.

A drawback of the magnetic tweezers is that the trap is very weak and the values for f_c are very low and thus the technique suffers from a temporal (and practically thus a spatial resolution) that is low in comparison to techniques with higher trap stiffnesses, such as AFMs or optical tweezers. However, an advantage is the macroscopic dimensions (order mm) over which the gradient of the magnetic field changes significantly (recall from Fig. 2.3c that, as a rule of thumb, the force change by a factor of two for each translation of the magnets by 0.5 mm). As such, with respect to the "length scale of the experiment", *i.e.* the changes in height that the bead undergoes in a typical experiment (usually maximally the contour length of λ DNA (16 μ m)), the force does not change appreciably. Measurements at constant force are thus readily achieved, in contrast to more complicated force feedback loops in an optical tweezer. Recently, the group of Steven Block at Stanford has implemented a force-clamp optical tweezer without feedback, but the workable length scale of the force clamp is extremely small (order 10 nm) in comparison to a magnetic trap (many μ m). Then again, that particular group is primarily interested in displacements that are significantly smaller than a single nanometer.

References

- [1] de Grooth, B. *Am. J. Phys.* **67**(12), 1248–1252 (1999).
- [2] Allemand, J. *Micro-manipulations de molecules d'ADN isolees*. PhD thesis, Ecole Normale Supérieure, Département de physique, (1997).
- [3] Bouchiat, C., Wang, M. D., Allemand, J., Strick, T., Block, S. M., and Croquette, V. *Biophys J* **76**, 409–13 (1999).
- [4] Bustamante, C., Bryant, Z., and Smith, S. B. *Nature* **421**(6921), 423–7 (2003). 0028-0836 Comment Journal Article Review Review, Tutorial.

Chapter 3

Fast dynamics of supercoiled DNA revealed by single-molecule experiments

The dynamics of supercoiled DNA play an important role in various cellular processes such as transcription and replication that involve DNA supercoiling. We present experiments that enhance our understanding of these dynamics by measuring the intrinsic response of single DNA molecules to sudden changes in tension or torsion. The observed dynamics can be accurately described by quasistatic models, independent of the degree of supercoiling initially present in the molecules. In particular, the dynamics are not affected by the continuous removal of the plectonemes. These results set an upper bound on the hydrodynamic drag opposing plectoneme removal, and thus provide a quantitative baseline for the dynamics of bare DNA.

3.1 Introduction

The degree of DNA supercoiling affects a number of important cellular processes such as gene expression [1], initiation of DNA replication [2], binding kinetics of sequence-specific proteins to their targets [3], and site-specific recombination [4, 5]. A strict regulation of DNA supercoiling is therefore essential for cell survival. This regulation results from a complex interplay between the occurrence of processes which generate local supercoiling of DNA, such as replication and transcription, and the action of topoisomerases, which are able to modify the global linking number (Lk) of DNA molecules via a mechanism of transient DNA strand breakage and religation (for reviews, see [6, 7]).

DNA supercoiling dynamics, *i.e.* the rate at which supercoils are created, propagated and removed on a DNA molecule, represent an important aspect of the regulation process. This can be clearly illustrated by the example of transcription-induced supercoiling. As initially proposed by Liu and Wang [8] and confirmed by later experiments *in vitro* [9] and *in vivo* [10, 11], the inability of a transcription complex of increasing molecular weight to rotate around helical DNA results in the supercoiling of DNA in its immediate vicinity: positively supercoiled domains are generated ahead of the transcription complex while negatively supercoiled domains are generated behind it. In the simple case of a single transcription complex bound to a circular DNA molecule, these supercoiled domains of opposite sign can be relaxed in one of two ways, either by the action of topoisomerases or by their mutual annihilation following their propagation along the connecting DNA segment. These two processes can have very different consequences for DNA topology: the action of topoisomerases induces a modification of Lk unless these enzymes relax positive and negative supercoils in a perfectly balanced way, while the merging of oppositely supercoiled domains does not influence the global Lk. Thus, the relative kinetics of these two processes appears as a major determinant of the degree of supercoiling of DNA in steady state. Within this context, DNA internal dynamics play an important role because they determine the rate at which oppositely supercoiled domains propagate and merge.

Interestingly, various *in vitro* experimental studies [9, 12, 13] have shown that the topological changes induced by transcription (in the presence of a controlled topoisomerase-mediated supercoil removal rates) are significantly larger than initially predicted by Liu and Wang [8], which suggests that the propagation of supercoiled domains occurs more slowly than initially expected. Indeed, Liu and Wang described the axial rotation of DNA according to a simple model in which DNA behaves as a perfectly straight "speedometer cable" of 1 nm radius. To

explain the observed experimental results, Nelson proposed that the presence of natural bends along DNA, not included in the "speedometer cable" model, may dramatically enhance the hydrodynamic drag encountered by DNA during its rotational motion [14]. However, this hypothesis was challenged by recent *in vivo* studies, in which Stupina and Wang monitored the supercoiling generated on DNA rings containing a single *tetA* transcript in cells expressing gyrase but lacking DNA topoisomerase IA. In these experiments, no excess of negative supercoiling was observed, a conclusion that remained valid even in the presence of stable bends inserted into the DNA rings [15]. This particular work therefore appears to demonstrate that, *in vivo*, the merging of oppositely supercoiled domains occurs more rapidly than the relaxation of supercoils by topoisomerases. Nonetheless, drawing quantitative conclusions about DNA internal dynamics from these experiments as well as from those cited before remains challenging, as bulk experiments, whether performed *in vivo* or *in vitro*, do not monitor dynamics directly and involve the simultaneous and interdependent action of many actors (at the very least DNA, RNA polymerases and topoisomerases).

In this paper, we describe single-molecule experiments on both torsionally relaxed and torsionally constrained molecules that provide a quantitative basis for understanding DNA dynamics. The first set of experiments provides a baseline by addressing the stretching dynamics of torsionally relaxed DNA molecules following the sudden application of an external force. These experiments are facilitated by the combination of two widely used techniques, magnetic and optical tweezers. A significant advantage of this combined experimental configuration is the possibility to control the rate of DNA stretching, as the external force used to stretch DNA can be modified by more than an order of magnitude. Our results are compared to previous studies that addressed the relaxation dynamics of tethered DNA molecules [16–19]. In addition, we present analytical results which, with only minimal approximations, accurately describe the quasistatic DNA dynamics observed. Using this basis, we next present results on the stretching of supercoiled DNA molecules which represent, to our knowledge, the first quantitative measurement of the dynamics of such molecules. We investigate the sudden application of an external force to a torsionally constrained molecule which induces a conversion of its writhe (Wr) into twist (Tw) with increasing tension [20] while maintaining a constant Lk . The analysis of these experiments shows that the frictional drag induced by this large conformational change of the DNA molecule is not large enough to alter the quasistatic character of the dynamics, and thus sets an upper bound on the drag opposing plectoneme removal. Finally, we present complementary experiments on supercoiled DNA molecules in which plectonemes are dissipated not by their conversion to Tw , but by the cre-

ation of a site-specific nick in the DNA molecule. These experiments likewise demonstrate that the process of plectoneme removal is fast compared to DNA stretching. We conclude by discussing the biological implications of the relative timescales determined.

3.2 Results

Pulling of torsionally relaxed DNA. Our integrated magneto-optical tweezers (described in the Materials and Methods) allow us to apply a nearly instantaneous force switch to a DNA molecule tethered to a magnetic bead. To do so, the extension of a DNA molecule, initially imposed by the external force created on the bead by a pair of permanent magnets, is reduced by optically trapping the bead and moving the trap position towards the surface. This leads to the initial, weakly stretched configuration illustrated in fig. 8.3a (left). Subsequently shutting off the laser trap leads to motion of the magnetic bead back to its equilibrium position under the magnetic force F_{mag} (fig. 8.3a, right). In this way, we can perform experiments in which DNA is stretched under the nearly instantaneous application of external forces up to 5 pN (fig. 8.3b). As expected, at low forces ($F_{mag}=0.90$ pN, red points in fig. 8.3b), both the rate of DNA stretching and the end-to-end distance of the DNA in steady-state are lower than at high forces ($F_{mag}=4.56$ pN, black points in fig. 8.3b). Successive runs performed at a given magnetic force are highly reproducible (fig. 6.7 of Supplementary Materials) and can therefore be averaged to reduce the effect of thermal fluctuations. The resulting traces are then analyzed using the following equation of motion obtained by balancing the forces involved (diagrammed in fig. 8.3a):

$$F_{mag} = F_{drag}(z) + F_{DNA}(z) = \zeta_{bead}(z) \frac{dz}{dt} + F_{DNA}(z), \quad (3.1)$$

where z is the distance between the surface and the nearest edge of the bead (equivalent to the DNA end-to-end distance), F_{mag} represents the external force exerted by the magnets on the bead, F_{drag} represents the hydrodynamic drag that opposes the motion of the bead, and F_{DNA} represents the force exerted by the DNA molecule on the bead. The inertia of the bead can be neglected because it contributes to a force that is several orders of magnitude smaller than the other forces involved [19]. Variations of the external force experienced by the magnetic beads during their motion can also be neglected, as this external force only varies appreciably on a ~ 1 mm length scale while the extent of bead motion is limited to only a few microns (maximally equal to the contour length

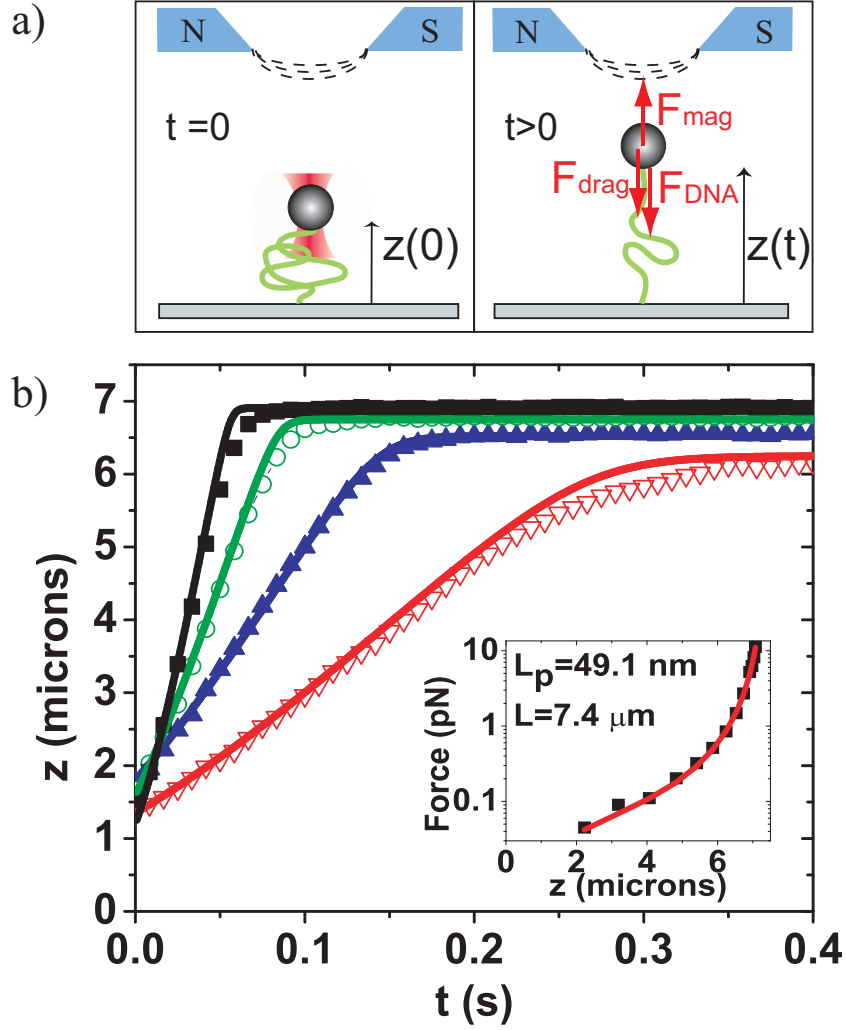


Figure 3.1: Stretching experiments with torsionally relaxed DNA molecules. (a) Principle of the experiments: an optical trap is used to reduce the end-to-end extension of a DNA molecule attached to a magnetic bead and held under a constant external force F_{mag} (left). After release of the optical trap, the end-to-end extension of the DNA molecule increases as a function of time, until it reaches its steady-state value under F_{mag} . The three forces considered in the analysis are illustrated (right). (b) Experimental traces obtained at $F_{mag} = 0.90$ pN (base-up triangles), $F_{mag} = 1.48$ pN (base-down solid triangles), $F_{mag} = 2.70$ pN (open circles) and $F_{mag} = 4.56$ pN (solid squares). Data points were averaged over eight successive runs. The solid lines are obtained by solving numerically the equation of motion assuming a quasistatic behavior. Inset: black squares: experimental force-extension curve, solid red line: fit with the interpolation formula from the Worm-Like Chain model [21].

L of the DNA molecules used). We note that use of Eq. 7.1 assumes a purely vertical motion of the bead, which is well verified experimentally, as illustrated in fig. 3.8 of the Supplementary Materials. The presence of the glass surface to which DNA molecules are anchored is known to lead to an enhancement of the hydrodynamic drag experienced by the bead compared to its value in bulk solution. This correction is taken into account according to the derivation by Bevan and Prieve for a motion perpendicular to the surface [22]:

$$\zeta_{bead}(z) = 6\pi\eta R_b \left(1 + \frac{R_b}{z} + \frac{R_b}{2R_b + 6z}\right), \quad (3.2)$$

where η is the viscosity of the solution, R_b the radius of the bead.

With the inclusion of surface corrections, all experimental traces are described accurately using a quasistatic model. This conclusion was reached by fitting the experimentally determined force-extension curve of the molecule under study (fig. 8.3b inset, black points) according to the interpolation formula derived by Bouchiat *et al.* for the Worm-Like Chain elasticity [21] with DNA contour length L and persistence length l_p as parameters (fig. 8.3b inset, solid red line), replacing $F_{DNA}(z)$ in Eq. 7.1 by this optimal fit, and solving this equation numerically. The solutions to this equation (indicated by the solid traces at different magnetic forces shown in fig. 8.3b) illustrate the excellent description of the experimental traces provided by the quasistatic model. The absence of memory effects is also supported by the fact that traces starting at different initial positions overlap well (fig. 6.10 of Supplementary Materials).

Although a numerical quasistatic solution to Eq. 7.1 is sufficient for this analysis, we can also describe the DNA dynamics analytically using only minimal approximations. Indeed, the Worm-Like Chain elasticity of DNA can be approximated by its high-force limit $F_{DNA}(z) = k_B T / (4l_p(1 - z(t)/L)^2)$ with an error smaller than 10% for relative extensions exceeding 0.3, and in this approximation a simple analytical solution for Eq. 7.1 can be derived if surface effects are neglected (approximating $\zeta_{bead}(z) = 6\pi\eta R_b$):

$$\frac{tF_{mag}}{\zeta_{bead}L} = \frac{z}{L} - \frac{1}{2}\sqrt{\epsilon} \tanh^{-1}\left(\frac{2(z/L) - 1}{\sqrt{\epsilon}}\right) + C, \quad (3.3)$$

where $\epsilon = k_B T (F_{mag} l_p)$, and C is a constant of integration. However, neglecting surface effects is unreasonable in the context of our experiments, as attested by the comparison of this analytical solution (red line in fig. 8.4) with the experimental data points (blue triangles in fig. 8.4). Nonetheless, as described in the Supplementary Materials, an expanded analytical solution can also be obtained

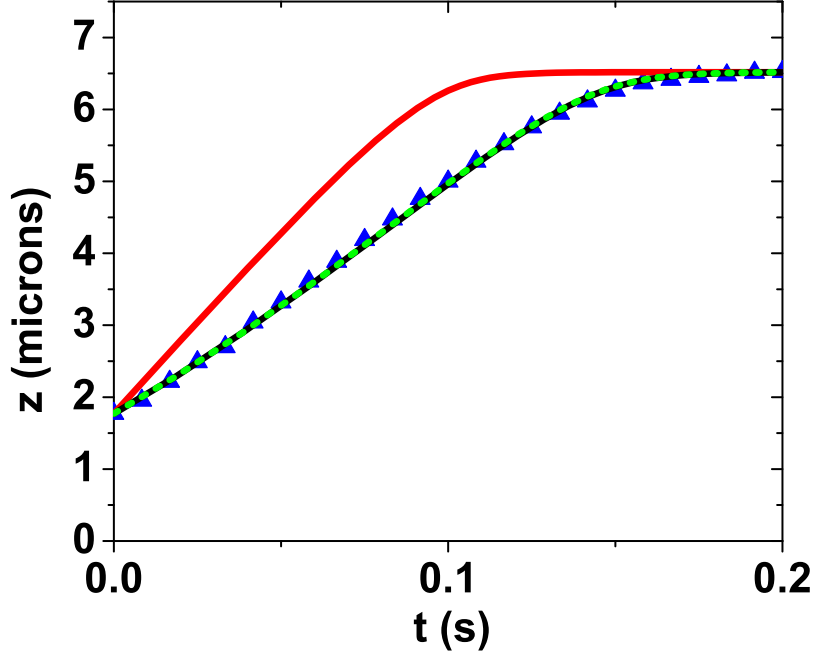


Figure 3.2: Comparison of experimental data with numerical and analytical models. The experimental data presented here (blue triangles) were collected at $F_{mag}=1.48$ pN. The traces issued from three models are represented: (i) analytical solution in the high-force approximation $F_{DNA}(z) = k_B T / (4l_p(1 - z(t)/L)^2)$, without inclusion of surface effects (solid red line); (ii) analytical solution in the high-force approximation, with inclusion of surface effects (dashed green line); (iii) complete quasistatic solution of Eq. 7.1 obtained numerically, including the complete interpolation formula from the Worm-Like Chain model [21] and surface effects (solid black line in the background). Clearly, (i) does not provide a satisfying description of the dynamics, but (ii) and (iii) are nearly indistinguishable and in excellent agreement with the experimental data.

in the high-force limit when the surface effects are modeled according to Eq. 7.2. Moreover, this analytical solution (red dashed line on fig. 8.4) coincides almost perfectly with the numerical solution including the complete Worm-Like Chain elasticity (black line in fig. 8.4).

pulling of supercoiled DNA. As we have shown, the stretching dynamics of torsionally relaxed DNA can be quantitatively understood using a simple equilibrium description, providing a convenient starting point from which to test the more complicated behaviour of supercoiled DNA. These experiments are conducted in a manner similar to those involving torsionally relaxed DNA, except that the Lk of the molecule under study is modified by a rotation of the magnets prior to stretching. The starting configuration of these experiments is thus a

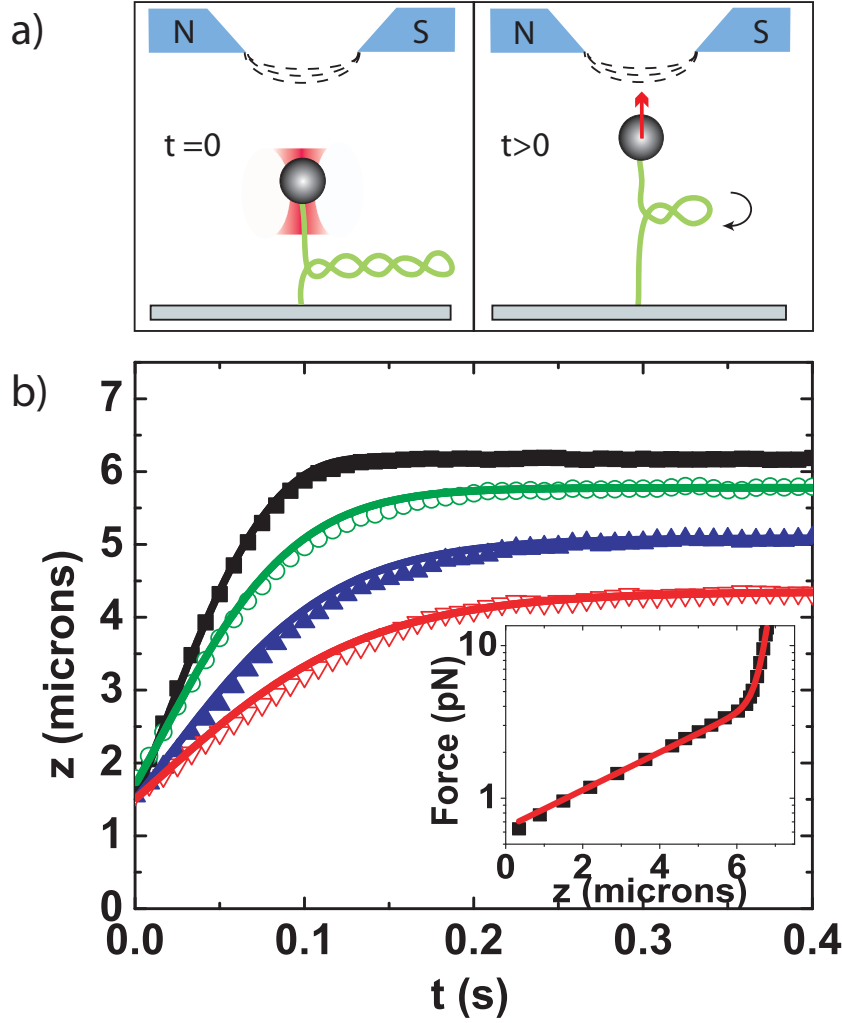


Figure 3.3: Stretching experiments with supercoiled DNA molecules. (a) These experiments are similar to those involving torsionally relaxed DNA (fig. 8.3), except that plectonemes have been created in the initial configuration by a preliminary rotation of the magnets (left). The stretching phase involves the conversion of plectonemic structures into twist (right). (b) Experimental traces (data points averaged over eight successive runs) obtained at $F_{mag} = 2.20$ pN (base-up open triangles), $F_{mag} = 2.72$ pN (base-down solid triangles), $F_{mag} = 3.36$ pN (open circles) and $F_{mag} = 4.14$ pN (solid squares). The linking number of the molecule had been modified by +100 turns prior to these pulling experiments. The solid lines are obtained by solving numerically the equation of motion assuming a quasistatic behavior. Inset: black squares: experimental equilibrium force-extension curve with $\Delta Lk = +100$; solid red line: biexponential fit.

weakly stretched DNA molecule containing plectonemic structures, as illustrated in fig. 8.5a (left). As the magnets are held fixed following the initial application of supercoils into the DNA, the Lk of the DNA molecules is fixed during the stretch-

ing phase (fig. 8.6, "force switch" arrow). However, since the partition between twist (Tw) and writhe (Wr) depends on the tension along the DNA molecule [20, 23], the final steady-state conformation of the DNA molecule is expected to include a reduced number of plectonemes compared to its initial configuration. The dynamics of this Wr to Tw conversion necessarily involves the rotation of the plectonemes about their axis (fig. 8.5a, right). The experimental traces as a function of increasing magnetic force are shown in fig. 8.5b ($F_{mag}=2.20$ pN, red points; $F_{mag}=2.72$ pN, blue points; $F_{mag}=3.36$ pN, green points; $F_{mag}=4.14$ pN, black points). As in the case of relaxed DNA, the DNA stretching rate is observed to increase with increasing applied magnetic force. We find that the dynamics of covalently-closed, supercoiled DNA under an applied magnetic force are also well-described by a quasistatic model. Numerical modeling in the absence of DNA internal dynamics proceeded as in the case of relaxed DNA, with one exception. While in the case of relaxed DNA, an interpolation formula for the molecules elasticity exists, there is no complete description of a molecules elasticity as a function of force at non-zero torque. We therefore opted to describe the experimentally-determined equilibrium force-extension behavior (fig. 8.5b inset, black points) phenomenologically, by fitting it to a double exponential relation (fig. 8.5b inset, solid red curve). The resulting best fit to the data was then substituted into $F_{DNA}(z)$ in Eq. 7.1. The solid curves shown in fig. 8.5b represent the dynamical behavior predicted according to this model, which, as in the case of torsionally relaxed DNA, indicates that the experimental data are well described by a quasistatic model.

Relaxation of supercoiled DNA. In addition to probing the dynamics of plectonemic supercoil removal by stretching, we also study its dynamics following the abrupt release of the torsional constraint. In this case, supercoiled DNA molecules are initially tethered as before and held under a constant magnetic force (fig. 8.7a, left). In the presence of the nicking enzyme N.BbvCIA, whose target site occurs only once in the 20-kb DNA molecules used in our experiments, cleavage of a single DNA strand by the nicking enzyme removes the torsional constraint imposed by the magnets. This allows the molecule to release its torsional stress by a relative rotation of the two DNA strands, changing Lk in the process and resulting in a reduction in the number of plectonemic supercoils (fig. 8.7a, right). The resulting upward motion (fig. 8.7b) shows how the resulting dynamics vary as a function of force (compare traces with $Lk=100$ turns at $F_{mag}=1.4$ pN, green points, and with $Lk=100$ turns at $F_{mag}=2.5$ pN, black points) and as a function of initial number of supercoils (compare traces with $F_{mag}=1.4$ pN and $Lk=100$ turns, green points and with $F_{mag}=1.4$ pN and $Lk=200$ turns, red points). Interestingly, for all the conditions tested, a successful description of the

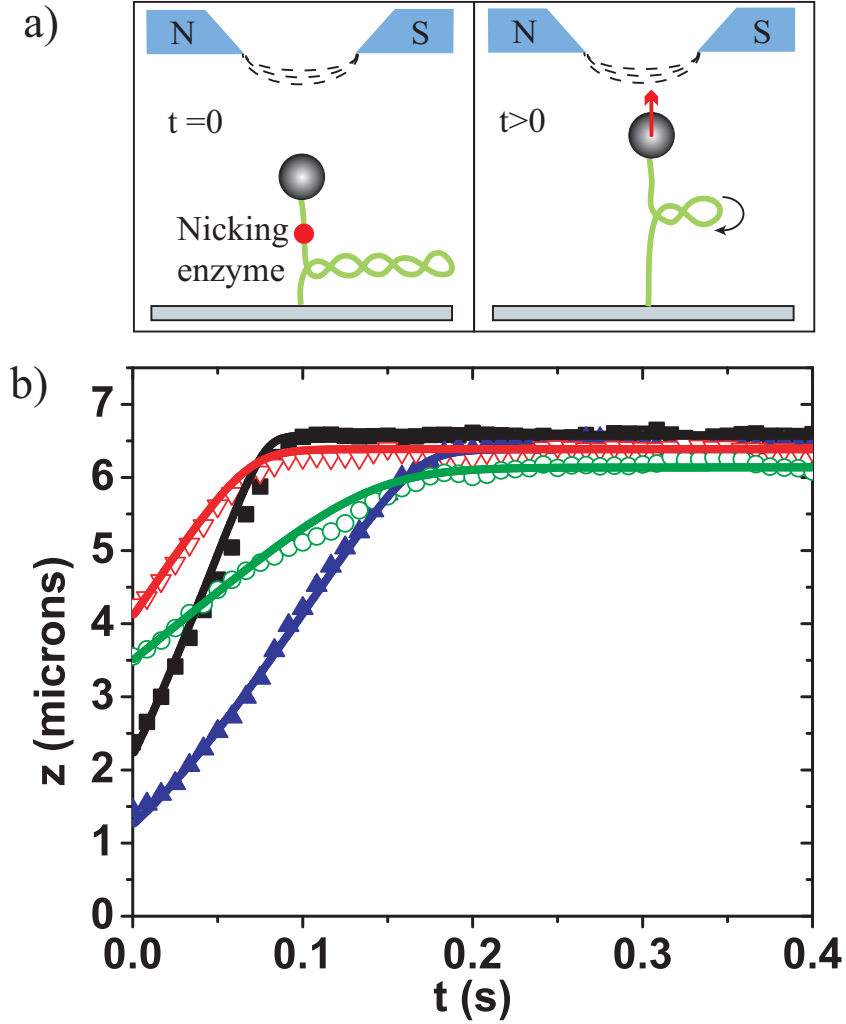


Figure 3.4: Scenarios of experiments involving supercoiled DNA. The experimentally measured DNA extension at equilibrium for different values of its linking number (0 turn represents a torsionally relaxed molecule) is illustrated for $F_{mag} = 1.7$ pN (solid squares), and $F_{mag}=0.35$ pN (solid circles). The blue squares indicate initial and final configurations in the two types of experimental situations described in this paper. Force switch experiments take place at constant linking number but involve an increase of the tension along DNA (green dotted line). On the contrary, in relaxation experiments the initial and final tensions are equal, but the initial torsional constraint is totally relaxed in the final state. This transition can *a priori* occur in two different ways (brown dashed lines): in the first scenario (1), plectoneme relaxation is much faster than DNA stretching, so that most of the motion involves torsionally relaxed DNA. In the second scenario (2), DNA stretching is faster than plectoneme relaxation so that plectonemes are progressively removed during the experiment.

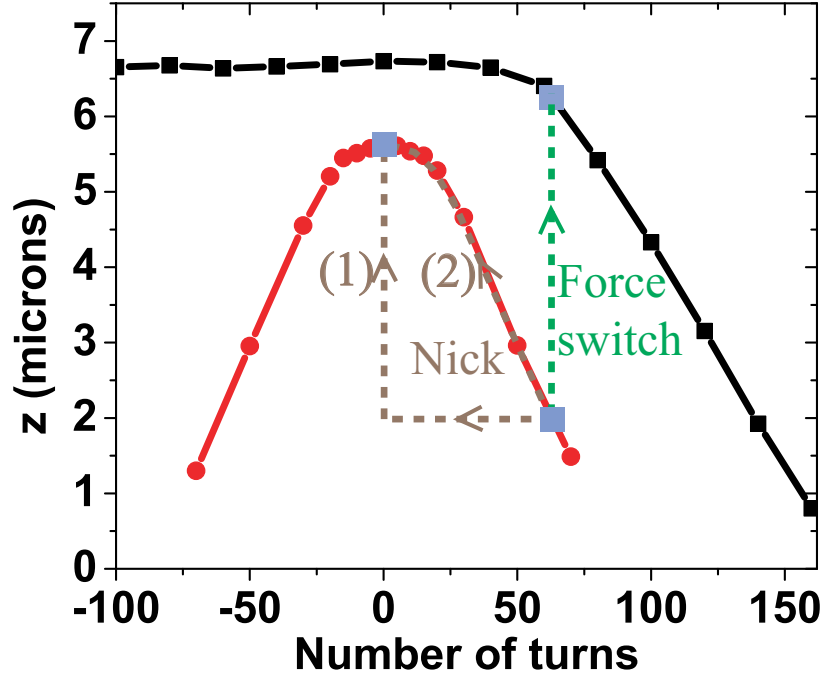


Figure 3.5: Relaxation of supercoils by a nicking enzyme. (a) Principle of the experiments: the linking number of a DNA molecule is initially modified in order to create plectonemes. Experiments take place in the presence of nicking enzymes N.BbvCIA (left). DNA extension shows a sudden increase after a nick is induced by one of these enzymes (right). (b) Individual experimental traces obtained with $F_{mag} = 2.5$ pN and $\Delta Lk = +100$ turns (solid squares), $F_{mag} = 1.4$ pN and $\Delta Lk = +100$ turns (open circles) and $F_{mag} = 1.4$ pN and $\Delta Lk = +200$ turns (solid triangles). Time $t=0$ corresponds to the action of the nicking enzyme. The solid lines are obtained by solving the equation of motion for torsionally relaxed DNA molecule, assuming quasistatic dynamics.

DNA extension trajectories in time is obtained by only taking into account the quasistatic stretching of torsionally relaxed DNA (fig. 8.7, solid lines). As described in the next section, these results indicate that plectoneme removal must occur on a timescale that is significantly faster than the upward movement of the bead due to the stretching, placing an upper bound on its rate.

3.3 Discussion

A large part of the results described here was obtained by the use of an apparatus that combines optical and magnetic tweezers. While the building of a similar experimental configuration has been reported in previous studies [24, 25], our work constitutes its first application to the study of a biologically relevant prob-

lem. This setup unites two features that are essential for studying the intrinsic dynamics of supercoiled DNA: first, the ability to twist DNA, provided by the magnetic tweezers, and second, the possibility to apply a sudden force switch. Whereas classical magnetic tweezers do not meet this second requirement, because changes in the magnetic force are typically obtained only through the slow translation of the permanent magnets, laser traps can be switched on and off within a few milliseconds using a commercial shutter. An additional advantage of these experiments is that two parameters are directly available to control the dynamics, the initial extension of DNA and the force used for DNA stretching. In comparison, relaxation experiments have only one readily accessible parameter that influences the dynamics, namely the initial extension of DNA.

The dynamics of torsionally relaxed DNA have previously been studied in two ways. In a first approach, the relaxation of hydrodynamically stretched, fluorescently labeled DNA molecules was monitored by fluorescence microscopy [26]. In a second approach, a DNA molecule attached to a bead was initially stretched by optically trapping the bead, and relaxed to an unstretched configuration following trap release [16–19]. The physics probed in the two situations is very different: whereas in the first case the relaxation dynamics of a DNA molecule is strongly out of equilibrium and governed by tension propagation along the molecule [27, 28], in the second configuration the presence of a bead and its associated hydrodynamic friction considerably slows down the motion, and quasistatic dynamics are thus expected to describe the motion of the DNA-bead complex [19]. Nonetheless, whether this is actually the case has been a long-standing debate. Early results were only adequately described by the quasistatic model upon assuming a three-fold higher value of the DNA persistence length than that commonly accepted [18]. Later, Bohbot-Raviv *et al.* developed a model including non-equilibrium effects to explain an apparent discrepancy of their data with the quasistatic model [16]. However, it was subsequently established that the apparent discrepancy was not attributable to non-equilibrium effects, but rather by a failure to correctly take into account the effect of the surface on the effective hydrodynamic drag of the bead [17, 19]. In the present work, we observed an excellent agreement between the experimental traces and the predictions of the quasistatic model. This agreement was observed over a large range of stretching velocities (up to about $100 \mu\text{m/s}$), due to our ability to tune the magnetic force that drives DNA stretching (cf Eq. 7.1 and traces from fig. 8.3), whereas classical relaxation experiments, in which dynamics reduces to $\zeta \text{d}z/\text{d}t = F_{\text{DNA}}(z)$, lack an equivalent parameter to easily influence the motion velocity.

The dynamical behavior of supercoiled DNA differs fundamentally from that

of torsionally relaxed DNA, as the elongation of a supercoiled DNA molecule requires the removal of its plectonemic structures, both when the motion is driven by a force switch (fig. reffigure3) and by the action of a nicking enzyme (fig. 8.7). Thus, it was not *a priori* clear which of the two processes - DNA stretching (governed, as demonstrated in the previous section by the magnetic force, by the hydrodynamic drag on the bead and the DNA elasticity) and plectoneme removal - would dominate the elongation kinetics. Our experiments provide an unambiguous answer to this question, as under the experimental conditions tested we never observed any influence of plectoneme removal on the elongation kinetics.

In experiments in which a sudden increase of the force exerted on a supercoiled DNA molecule is imposed, plectoneme removal is induced by DNA stretching. This causes a progressive increase of the tension along DNA and mediates the conversion of its writhe into twist. The good fit of the experimental data to a quasistatic model implies that the internal dynamics of DNA, which includes plectoneme removal, is so rapid that the DNA molecule is effectively always in an equilibrium configuration. The nature of the motion induced by the action of a nicking enzyme on supercoiled DNA is different, as in this configuration plectoneme removal can in principle occur independently of DNA stretching. Two principal scenarios are *a priori* plausible in these experiments (fig. 8.6). The first one (fig. 8.6, path (1)) assumes that plectoneme removal is much faster than DNA stretching; in this case the greater part of the observed motion takes place at $\Delta Lk \approx 0$. In the second scenario (fig. 8.6, path (2)), plectoneme removal is rate-limiting; in this case, the tension along DNA remains always close to F_{mag} , so that the rate of the DNA extension is expected to be slow (cf Eq. 7.1). Since our experiments are accurately described with the model used to describe the dynamics of torsionally relaxed DNA (fig. 8.7), the first scenario, in which the rate of plectoneme removal exceeds the rate of DNA stretching, applies. Therefore, the dynamics are governed by the rate of DNA stretching rather than the rate of plectoneme removal in both the force switch and nicking enzyme experiments, even though the manner in which plectonemes are removed differs in the two cases.

To quantitatively understand the observed separation of time scales between plectoneme removal and DNA stretching in both experiments on supercoiled DNA, it is useful to estimate the drag torque that a plectonemic region can exert during its shortening. As the starting point of this analysis, we calculate this torque in a simple model under the assumptions that the supercoiled DNA molecule contains a single branch of plectonemes (simulations clearly exclude the other extreme, in which the DNA contains as many branches as individual plectonemes) and that the shortening of this region involves its global rotation. The

rotation around a given axis of a rigid but not necessarily straight DNA fragment of diameter d and length L_{plect} at an angular velocity ω encounters a hydrodynamic drag μ_{drag} which depends crucially on the fragment distribution relative to the rotation axis and contributes to the following torque (14):

$$\tau_{drag} = \mu_{drag} \omega L_{plect} \langle r_{\perp}^2 \rangle, \quad (3.4)$$

where $\mu_{drag} = 4\pi\eta/(0.8 + \ln(X/d))$. In this equation, $r_{\perp}$ is the distance of a DNA point to the rotation axis, brackets represent the average along the whole DNA fragment, and X is a cutoff length representing the distance over which a DNA fragment is straight. Nelson proposed to equate X with the structural persistence length of DNA, equal to 130 nm [14]. Nonetheless the weak dependence of μ_{drag} on X implies that its precise numerical value does not play an important role. Using the simple model detailed in [29] to describe a DNA molecule under tension F and torsion, one can obtain an analytical expression for the the plectonemic radius $R_{plect} = \sqrt{k_B T l_p / 2F}$. One can then replace $\langle r_{\perp}^2 \rangle$ in the expression of τ_{drag} by $R_{plect}^2/2$.

Force switch experiments are then reanalyzed by including the additional force term τ_{drag}/R_{plect} contributed by the rotating plectonemes in Eq. 7.1. As before, the resulting equation of motion is solved numerically, again assuming that the configuration of DNA at a given extension is the same as at equilibrium. This analysis indicates that the rotational drag opposing plectoneme rotation is not expected to significantly affect the dynamics within the context of the model described above, since the predicted extension versus time (fig. 8.8, red trace) almost precisely overlaps with predictions in which this drag torque is ignored (fig. 8.8, black trace). Nonetheless, it can be concluded that the rotational drag opposing plectoneme rotation would have a discernable effect if it were ten times or more larger (fig. 8.8, green trace, 10-fold larger drag; blue trace, 50-fold larger drag; these traces do not refer to a specific model for plectoneme relaxation, but are drawn as references to estimate the minimal detectable drag in our experiments).

The relaxation experiments are analysed using the same framework in a more straightforward manner. A direct estimation of the time scale of plectoneme removal is obtained by equating the torque driving plectoneme relaxation, equal to $\sqrt{2k_B T l_p F_{mag}}$ in the initial configuration [29] with the drag torque τ_{drag} opposing it. One gets $\omega = 4\sqrt{2}(k_B T l_p)^{-1/2} F_{mag}^{3/2} / (\mu_{drag} L_{plect})$. Using typical experimental parameters ($F_{mag}=1.5$ pN, $L_{plect}=5$ μm at $t=0$) yields an initial value of the initial angular velocity $\omega \approx 6.104$ rad/s. This rate is expected to be enhanced as the size of the plectonemic region decreases. In this model, the total

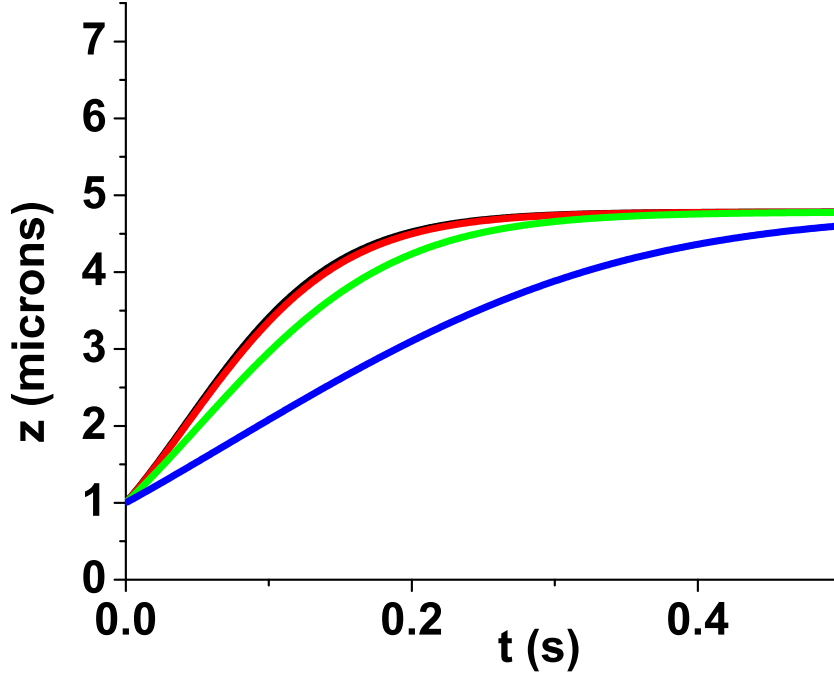


Figure 3.6: Predicted stretching behavior of supercoiled DNA following a sudden force switch, as a function of the magnitude of the hydrodynamic drag induced by plectonemes. In this example, $F_{mag}=2.5$ pN and $\Delta Lk=+100$ turns. The hydrodynamic drag opposing plectoneme removal corresponding to the simple model described in the text has been added in the equation of motion 7.1. Numerically solving this equation (red line) shows that this leads only to a nearly indistinguishable modification of the quasistatic dynamics (black line). However, traces generated by arbitrarily considering 10-fold (green line) and 50-fold (blue line) larger hydrodynamic drags are significantly different.

removal of 200 plectonemic units is accomplished within $T \approx 10$ ms, similar to the acquisition time in our experiments and much shorter than the total duration of DNA stretching under these conditions (typically 0.1 s, cf fig. 8.7). Similarly to force switch experiments, our data are therefore consistent with the simple model described by Eq. 7.4, but disagree with a rotational drag larger by ten times or more than the prediction of this simple model. These experiments therefore exclude models for plectoneme relaxation that would generate a global friction larger by a factor of ten or more than the simple model above, caused for instance by large intrinsic bends in the plectonemic region.

Thomen *et al.* have addressed the rotational drag of a double-stranded DNA molecule during its unzipping [30], and showed that the induced drag of such a relatively simple structure deviated from the naive "speedometer model" of Levinthal and Crane [31] by only a factor of ten. Our experiments complement

this work by addressing the rotational drag of plectonemic structures in a number of different cases, and place an upper bound on its value. Together, these experiments provide the main features of the dynamics of an individual DNA molecule under torsion and exclude a large contribution of DNA intrinsic bends to these dynamics. An interesting perspective of our work is offered by its potential extension to situations in which DNA interacts with proteins. For instance, stretching and nicking experiments may be performed in the presence of transcription complexes, nucleosomes or more generally DNA-binding proteins. This might considerably slow down the dynamics, as suggested by Leng and McMacken from bulk *in vitro* experiments [32]. It is likely that these experiments will have to be analyzed using the theoretical framework of Nelson's work [14], contrary to the ones presented here. Such experiments have the potential to provide a quantitative description of many aspects of supercoiled DNA dynamics that are difficult to extract from bulk experiments.

3.4 Materials and Methods

Experimental setup. All experiments were performed with 20.7 kb long DNA molecules containing a unique site for the nicking enzyme N.BbvCIA. The two ends of the molecules were ligated to 0.6 kb long biotin and digoxigenin PCR fragments, respectively. The DNA molecules were incubated with streptavidin-functionalized Micromer magnetic beads (3 μm diameter, Micromod) selected for their low magnetic content which permitted good optical trapping, and introduced in a custom-made flow cell. The lower slide of the flow cell was coated with polystyrene (1% w/v in toluene), anti-digoxigenin (50 $\mu\text{g}/\text{ml}$ in PBS) and finally polyglutamic acid (50 mg/ml in PBS); the latter step aimed at reducing non-specific interactions.

The detailed experimental configuration of the magnetic tweezers has been described previously [20]. Briefly, a pair of magnets was used to apply forces and rotations to a magnetic bead bound to a tethered DNA molecule to control the tension and the linking number the DNA molecules. The beads 3D position was determined with 10 nm accuracy from the video images with an acquisition frequency of 120 Hz. The force corresponding to a given position of the magnets was measured from the lateral fluctuations of the bead. The standard magnetic tweezers setup was expanded to include an optical trap. An infrared laser beam (1064 nm, 500 mW, Crystalaser) was attenuated to 50-100 mW and expanded through a 10X beam expander (CVI) to fill the back aperture of the objective (N.A. 1.4, 100X, Olympus). This generated a strongly focused spot in the flow

cell, which was vertically translated using the piezoelectric objective positioner (Physik Instrumente) and switched on and off using a shutter (Melles Griot) inserted along the optical path.

Force switch experiments. These experiments were carried out in 10 mM Phosphate Buffer pH 7.4, 10 mM Sodium Azide, 0.1% Tween 20 and 200 $\mu\text{g/ml}$ BSA. The magnets were first translated and rotated until the desired values for the magnetic force and the DNA linking number were reached. Then the bead under study was optically trapped and brought closer to the surface by vertical translation of the optical trap. The trap was suddenly switched off by triggering the closure of the shutter located in the optical path. Afterwards, the bead moved up until it reached again its equilibrium position under the magnetic force (fig. 8.3 and 8.5). In these experiments, the position of the optical trap was carefully adjusted to minimize the lateral motion of the bead, as the analysis was facilitated by the assumption that the bead primarily executed vertical motion away from the surface (fig. 6.10 of Supplementary Materials). Eight successive traces were taken and averaged for each experimental condition tested.

Experiments with the Nicking enzyme N.BbvCIA. These experiments were performed in a buffer containing 10 mM Tris-HCl pH 8.0, 50 mM NaCl, 10 mM MgCl_2 , 1 mM DTT, 0.1% Tween 20, 200 $\mu\text{g/ml}$ BSA, 0.5 units/ μL of the nicking enzyme N.BbvCIA and 0.04 units/ μL of T4 DNA Ligase (both enzymes from New England Biolabs). The simultaneous presence of nicking enzyme and ligase allowed us to acquire multiple traces with the same molecule, using the following procedure: first, the magnets were rotated to induce supercoils in the DNA. Second, after a time interval, single-strand cleavage by the nicking enzyme caused the DNA to relax to a torsionally unconstrained state. After a subsequent time interval necessary for ligase to repair the created nick, magnet rotation was used to induce supercoils again. Successive experiments were very reproducible in these conditions.

3.5 Acknowledgements

We thank Ralf Seidel for his large contribution to the building of the setup, Ya-Hui Chien and Susanne Hage for preparing the DNA constructs, Alexander Volgodskii for numerical simulations of the number of plectoneme branches, Cyril Claudet for helpful technical advice and Derajavan Thirumalai, Alexander Volgodskii and David Bensimon for useful discussions. NHD acknowledges financial support from FOM and NWO.

3.6 Supplementary Information

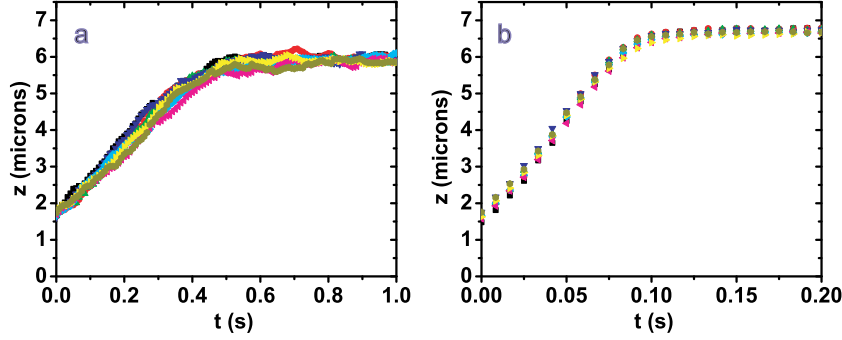


Figure 3.7: High reproducibility of force switch experiments. Eight successive runs obtained with torsionally relaxed DNA, each represented with a different color, are illustrated for $F_{mag} = 0.5$ pN (a) and 2.7 pN (b). All experiments were carried out with an acquisition frequency of 120 Hz. The excellent overlap between successive runs illustrates the high reproducibility of these experiments.

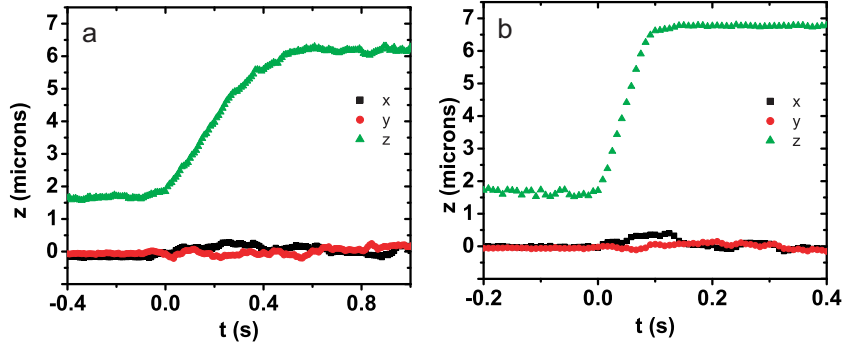


Figure 3.8: Experiment to verify the assumption that the bead moves primarily vertically. Throughout the analysis of experimental data, it is assumed that the bead moves vertically (z -direction). The data presented here as an example are typical force switch experiments with torsionally relaxed DNA performed with $F_{mag} = 0.5$ pN (a) and 2.7 pN (b). No averaging has been performed. Vertical (solid triangles) and lateral (circles) and (squares) motions of the bead before and after the release of the laser trap (at $t=0$ on the plots) are represented. In view of the small amplitude of the lateral motion, the assumption of a purely vertical motion made throughout the analysis is fully justified.

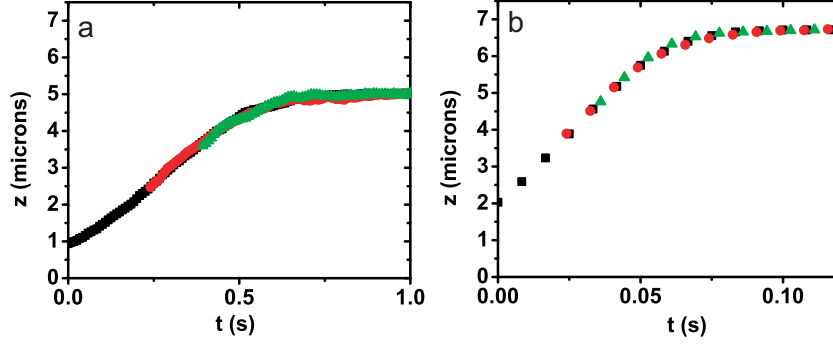


Figure 3.9: Effect of the initial conditions. The traces represented here were obtained with torsionally relaxed DNA with $F_{mag} = 0.4$ pN (a) and 3.8 pN (b). Data points were averaged over eight successive runs. The conditions of these experiments differ only by the initial position of the optical trap. The time origin was shifted to synchronize these traces. The excellent overlap favors the quasistatic model as an accurate description of the dynamics.

Analytical solution for the quasistatic dynamics of torsionally relaxed DNA. In the main text of the paper, we have derived a simple analytical solution for the quasistatic dynamics of torsionally relaxed DNA under the high-force approximation in the case where surface effects on bead dynamics were not considered. However, we have also shown that this expression does not describe our experimental data accurately (fig. 8.4, solid red line). A more complicated, but still analytic expression for $t(z)$ results if we take into account the surface-induced enhancement of the hydrodynamic friction felt by the bead described by Eq. 7.2:

$$\begin{aligned} \frac{tF_{drag}}{6\pi\eta R_b L} = & u - \frac{\ln(4(1-u)^2 - \epsilon)(2R_b^3\epsilon + 21L^2R_b\epsilon - 21L^2R_b\epsilon^2/4 + 12LR_b^2\epsilon)}{L(4R_b^2 - 9L^2\epsilon + 24LR_b + 36L^2)(4 - \epsilon)} \\ & - \frac{\tanh^{-1}(2(u-1)\epsilon^{-1/2})(72L^3\epsilon^{1/2} - 36L^3\epsilon^{3/2} + 132L^2R_b\epsilon^{1/2})}{L(4R_b^2 - 9L^2\epsilon + 24LR_b + 36L^2)(4 - \epsilon)} \\ & - \frac{33L^2R_b\epsilon^{3/2} + 60LR_b^2\epsilon^{1/2} - 3LR_b^2\epsilon^{3/2} + 9/2L^3\epsilon^{5/2} + 8R_b^3\epsilon^{1/2}}{L(4R_b^2 - 9L^2\epsilon + 24LR_b + 36L^2)(4 - \epsilon)} \\ & + \frac{\ln(R_b + 3Lu)(2/3R_b^3 + 6L^2R_b + 4LR_b^2)}{L(4R_b^2 - 9L^2\epsilon + 24LR_b + 36L^2)} + 4\frac{R_b\ln(u)}{L(4 - \epsilon)} + C, \end{aligned}$$

where $u = z(t)/L$, $\epsilon = k_B T / F_{mag} l_p$ and C is a constant of integration. As shown in fig. 8.4 of the main text, this expression (dashed green line) describes our experimental data (blue triangles) accurately.

References

- [1] Fisher, L. M. *Nature* **307**(5953), 686–7 (1984).
- [2] Funnell, B. E., Baker, T. A., and Kornberg, A. *J Biol Chem* **262**(21), 10327–34 (1987).
- [3] Gowers, D. M. and Halford, S. E. *EMBO J* **22**(6), 1410–8 (2003).
- [4] Huang, J., Schlick, T., and Vologodskii, A. *Proc Natl Acad Sci U S A* **98**(3), 968–73 (2001).
- [5] Vologodskii, A. and Cozzarelli, N. R. *Biophys J* **70**(6), 2548–56 (1996).
- [6] Champoux, J. J. *Annu Rev Biochem* **70**, 369–413 (2001).
- [7] Wang, J. C. *Annu Rev Biochem* **65**, 635–92 (1996).
- [8] Liu, L. F. and Wang, J. C. *Proc Natl Acad Sci U S A* **84**(20), 7024–7 (1987).
- [9] Tsao, Y. P., Wu, H. Y., and Liu, L. F. *Cell* **56**(1), 111–8 (1989).
- [10] Giaever, G. N. and Wang, J. C. *Cell* **55**(5), 849–56 (1988).
- [11] Wu, H. Y., Shyy, S. H., Wang, J. C., and Liu, L. F. *Cell* **53**(3), 433–40 (1988).
- [12] Droge, P. and Nordheim, A. *Nucleic Acids Res* **19**(11), 2941–6 (1991).
- [13] Wang, Z. and Droge, P. *J Mol Biol* **271**(4), 499–510 (1997).
- [14] Nelson, P. *Proc Natl Acad Sci U S A* **96**(25), 14342–7 (1999).
- [15] Stupina, V. A. and Wang, J. C. *Proc Natl Acad Sci U S A* **101**(23), 8608–13 (2004).
- [16] Bohbot-Raviv, Y., Zhao, W. Z., Feingold, M., Wiggins, C. H., and Granek, R. *Phys Rev Lett* **92**(9), 098101 (2004).
- [17] Coelho Neto, J., Dickman, R., and Mesquita, O. *Physica A* **345**(1-2), 173–184 (2005).
- [18] Feingold, M. *Physica E* **9**(3), 616–620 (2001).
- [19] Goshen, E., Zhao, W. Z., Carmon, G., Rosen, S., Granek, R., and Feingold, M. *Phys Rev E Stat Nonlin Soft Matter Phys* **71**(6 Pt 1), 061920 (2005).
- [20] Strick, T. R., Allemand, J. F., Bensimon, D., Bensimon, A., and Croquette, V. *Science* **271**(5257), 1835–7 (1996).
- [21] Bouchiat, C., Wang, M. D., Allemand, J., Strick, T., Block, S. M., and Croquette, V. *Biophys J* **76**(1 Pt 1), 409–13 (1999).
- [22] Bevan, M. and Prieve, D. *J Chem Phys* **113**(3), 1228–1236 (2000).

-
- [23] Strick, T. R., Allemand, J. F., Bensimon, D., and Croquette, V. *Biophys J* **74**(4), 2016–28 (1998).
 - [24] Claudet, C. and Bednar, J. *Appl Opt* **44**(17), 3454–7 (2005).
 - [25] Sacconi, L., Romano, G., Ballerini, R., Capitanio, M., De Pas, M., Giuntini, M., Dunlap, D., Finzi, L., and Pavone, F. *Opt Lett* **26**(17), 1359–1361 (2001).
 - [26] Perkins, T. T., Quake, S. R., Smith, D. E., and Chu, S. *Science* **264**(5160), 822–6 (1994).
 - [27] Brochard-Wyart, F. *Europhys Lett* **30**(7), 387–392 (1995).
 - [28] Manneville, S., Cluzel, P., Viovy, J., Chatenay, D., and Caron, F. *Europhys Lett* **36**(6), 413–418 (1996).
 - [29] Charvin, G., Allemand, J., Strick, T., Bensimon, D., and Croquette, V. *Contemp Phys* **45**(5), 383–403 (2004).
 - [30] Thomen, P., Bockelmann, U., and Heslot, F. *Phys Rev Lett* **88**(24), 248102 (2002).
 - [31] Levinthal, C. and Crane, H. R. *Proc Natl Acad Sci U S A* **42**(7), 436–8 (1956).
 - [32] Leng, F. and McMacken, R. *Proc Natl Acad Sci U S A* **99**(14), 9139–44 (2002).
-

Chapter 4

Friction and torque govern the relaxation of DNA supercoils by eukaryotic topoisomerase IB

Topoisomerases relieve the torsional strain in DNA that is built up during replication and transcription. They are vital for cell proliferation [1–3] and are a target for poisoning by anti-cancer drugs [4, 5]. Type IB topoisomerases (TopIB) form a protein clamp around the DNA duplex [6–8] and create a transient nick that permits removal of supercoils. Using real-time single-molecule observation, we show that TopIB releases supercoils by a swivel mechanism that involves friction between the rotating DNA and the enzyme cavity, *i.e.*, the DNA does not freely rotate. Unlike a nicking enzyme, TopIB does not release all the supercoils at once, but it typically does so in multiple steps. The number of supercoils removed per step follows an exponential distribution. The enzyme is found to be torque-sensitive, since the mean number of supercoils per step increases with the torque stored in the DNA. We propose a model for topoisomerization in which the torque drives the DNA rotation over a rugged periodic energy landscape in which the topoisomerase has a small but quantifiable probability to religate the DNA once per turn.

4.1 Introduction

Type IB topoisomerases (TopIB) alter DNA topology by cleaving and rejoining one strand of the DNA duplex [1] and are able to remove both positive and negative supercoils. *In vivo*, TopIB removes positive supercoils generated in advance of the replication fork [9]. Cleavage occurs via a transesterification reaction in which the scissile phosphodiester is attacked by a tyrosine of the enzyme, resulting in the formation of a DNA-(3'-phosphotyrosyl)-enzyme intermediate and the expulsion of a 5'-OH DNA strand. In the rejoining step, the DNA 5'-OH group attacks the covalent intermediate resulting in expulsion of the active site tyrosine and restoration of the DNA phosphodiester backbone. Based on structural [10] and kinetic [11] data, a mechanism has been proposed for TopIB, whereby supercoils are relaxed by swiveling of the DNA about the phosphodiester opposite the nick (Fig. 4.1a-c). Furthermore, the clamp-like structure of human TopIB around the DNA duplex suggested that the DNA would not be allowed to swivel unhindered and a controlled rotation model was proposed [10, 11]. The swivel mechanism of TopIB action is in stark contrast to the protein-assisted strand-passage mechanism of type IA and type II topoisomerases, whereby there is an obligate step size of 1 and 2 supercoils removed per cleavage-religation cycle, respectively [12–18].

4.2 Topoisomerase IB uncoils DNA in a step-wise fashion

Our experimental strategy entails anchoring a single continuous linear dsDNA molecule between a glass surface and a paramagnetic bead. We use a pair of magnets to apply a stretching force F and a degree of supercoiling σ by translating a pair of magnets in the vertical direction or rotating them about their axis, respectively [19]. The height of the bead from the surface is equal to the extension of the DNA molecule. Fig. 4.1d plots the DNA extension as a function of force and torque. In a typical measurement, the DNA molecule is prepared in a positively supercoiled state ($\sigma > 0$), which corresponds to a short extension. When Vaccinia TopIB, a prototypical eukaryotic type IB topoisomerase, is added, we observe discrete step-wise increases in the extension (Fig. 4.2a). Each step signifies the removal of DNA supercoils during a single cleavage-religation cycle by a single TopIB enzyme. Using the rotation curve (Fig. 4.1d), we convert the changes in DNA extension to a number of rotations removed from the DNA, which equals the change in the linking number ΔLk . The distribution of ΔLk ,

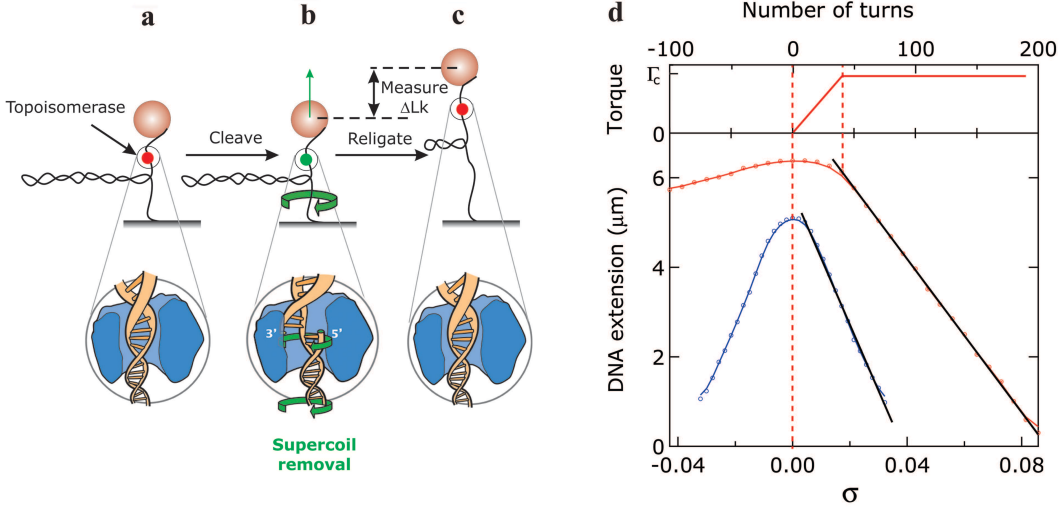


Figure 4.1: Single-molecule assay for measuring DNA supercoil removal by topoisomerase. (a) Type 1B topoisomerase is bound non-covalently to a supercoiled DNA molecule. (b) Topoisomerase establishes a covalent bond to the DNA, creating a nick that allows for rotation of the DNA about the remaining intact strand (green arrow). Consequently, supercoils are removed. (c) The DNA has been religated by the topoisomerase, terminating the release of supercoils. The height of the bead above the surface has increased, proportional to the number of supercoils removed (ΔLk). (d) The behavior of dsDNA under torsion is dependent on the stretching force (data shown for half bacteriophage λ DNA). Increasing σ at 1 pN (red curve), the extension initially remains constant and the torque builds up linearly with σ (Twist regime). After a critical buckling torque Γ_c , the torque saturates and the DNA forms plectonemic supercoils. It then contracts linearly with σ (Writhe regime) with a slope of 37nm/turn (1 pN) and 65nm/turn (0.2 pN, blue curve). At low stretching force (0.2 pN) the DNA extension is decreased irrespective of the rotation direction as the molecule is supercoiled. At a higher force (1 pN), this occurs only for positive rotations ($\sigma \geq 0$), while at negative rotations, the DNA denatures. Solid colored lines in rotation curves are splines through experimental data points, black lines are fits to a linear function.

$P(\Delta Lk)$, has a mean far greater than unity (Fig. 4.2b) and is well fitted by an exponential:

$$P(\Delta Lk) \sim \exp(-\Delta Lk / \langle \Delta Lk \rangle) \quad (4.1)$$

(red line in Fig. 4.2b). At higher stretching forces, the functional forms of the distributions are unchanged, but the corresponding values for $\langle \Delta Lk \rangle$ increase significantly (Fig. 4.2c, solid circles).

The functional forms of the distributions are also found to be insensitive to the sign of the supercoiling (data not shown). Within the experimentally accessible

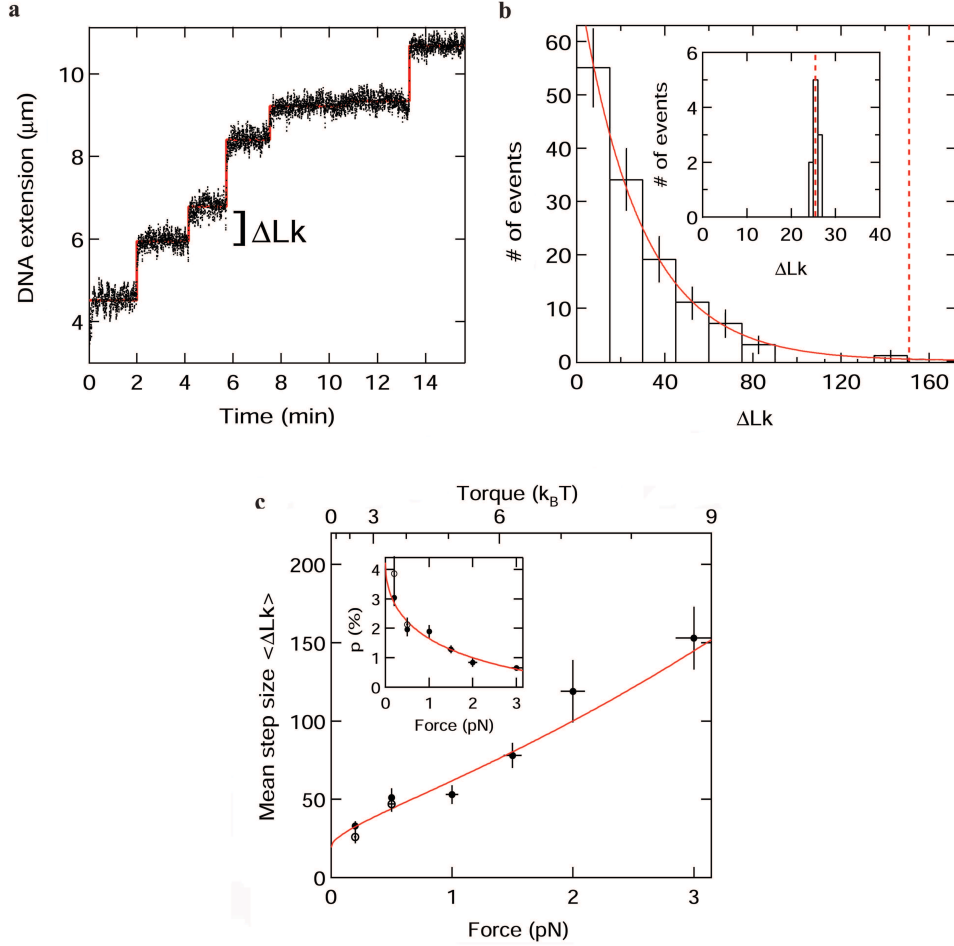


Figure 4.2: Real-time enzymatic activity and step-size distribution for TopIB acting on a single DNA molecule. (a) Each time TopIB removes supercoils from the DNA, a step is observed in the DNA extension. (b) $P(\Delta Lk)$ for TopIB (main panel) and for nicking enzyme (inset), in units of ΔLk . Vertical dashed red lines show the number of plectonemic supercoils initially incorporated into the DNA. The solid red line is a fit of Eq. 4.1 to the data. Error bars denote the square root of the number of events. (c) $\langle \Delta Lk \rangle$ as a function of applied force and torque. The solid line is the fit of Eq. 4.2 to the data. The inset shows the force-dependence of p (see text). Solid and open circles represent data taken at positive and negative supercoils, respectively.

regime at negative supercoiling (see Fig. ??d), the values for $\langle \Delta Lk \rangle$ for negative supercoils (Fig. 4.2c, open circles) are similar to those measured for positive supercoils. In contrast to Vaccinia TopIB, a control measurement of $P(\Delta Lk)$ for a nicking enzyme shows a peak that coincides with the number of supercoils initially applied to the DNA molecule (Fig. 4.2b, inset). Unlike Vaccinia TopIB, the nicking enzyme does not possess the ability to religate DNA and therefore all supercoils are released at once.

The exponential functional form of $P(\Delta Lk)$ for Vaccinia TopIB can be understood by considering the DNA as it rotates during supercoil release. After nicking of the DNA by transesterification of Vaccinia TopIB to the DNA 3'-phosphate end, the DNA rotates inside the enzyme cavity and its 5'-OH end passes the tyrosine-3'-DNA adduct once every turn. At each pass, there is a finite probability p for Vaccinia TopIB to religate the DNA and a probability $q \equiv 1 - p$ that the enzyme does not religate the DNA. In correspondence with this picture, the probability distribution for observing steps of size ΔLk is given by the discrete probability function of the geometric distribution [20] ($\Delta Lk = 1, 2, \dots$): $P(\Delta Lk) = pq^{\Delta Lk-1} = p(1 - p)^{\Delta Lk-1}$, where $\langle \Delta Lk \rangle \equiv 1/p$. The continuum limit of this distribution is equation 4.1. We find that by increasing the stretching force from 0.2 pN to 3 pN, the probability to religate per turn decreases from 3% to 0.7% (Fig. 4.2c, inset). Because a strand-passage mechanism implies that $\langle \Delta Lk \rangle$ is independent of the stretching force [15], we conclude that our data are inconsistent with such a model. Our results are however fully consistent with a swivel mechanism.

4.3 Topoisomerase hinders the uncoiling

We are able to resolve in real time the velocity of DNA extension during supercoil removal. We measure this both for Vaccinia TopIB and for endonucleolytic cleavage of one strand by nicking enzymes (Fig. 4.3). The velocity value is a measure of the rate at which the DNA swivels in the topoisomerase cavity as the supercoils are released. The data of Fig. 4.3b clearly show that Vaccinia TopIB (red diamonds) slows down the DNA rotation rate compared to the unhindered rate observed for the nicking enzymes (blue triangles).

Control experiments with two different Vaccinia TopIB mutants and with human TopIB provide further evidence that the decreased rotation rate in the TopIB reaction is caused by friction between the topoisomerase and the rotating DNA. First, we excluded the possibility that multiple topoisomerases bound to the DNA at sites other than the cleavage site were interacting with each other

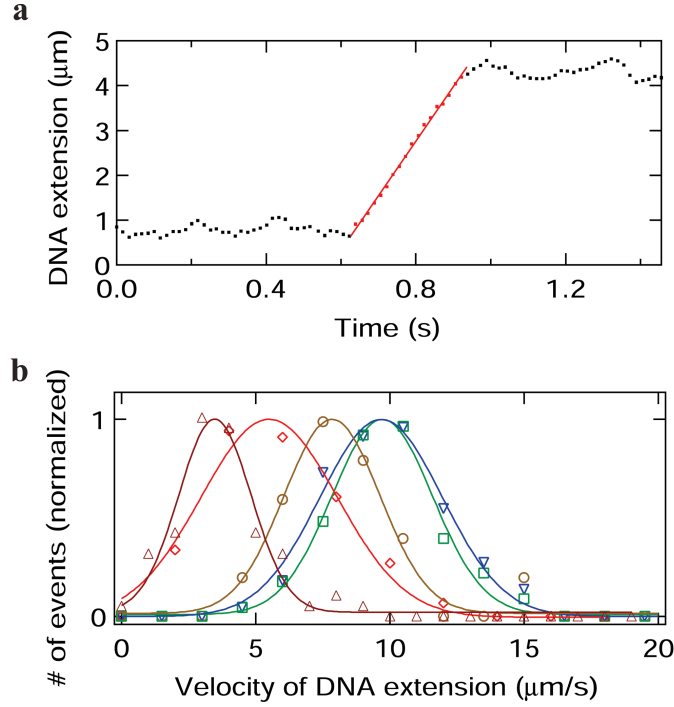


Figure 4.3: Measurement of velocity of DNA extension during supercoil release. (a) A step in DNA extension as a consequence of supercoil removal is characterized by a linear increase in DNA extension (red points) and is fitted with a linear function to obtain the rate (solid red line). (b) Velocity of DNA extension distributions taken at 0.2 pN for human TopIB (dark red triangles, $\langle v \rangle = 4.1 \pm 0.2 \mu\text{m/s}$), wild-type TopIB (red diamonds, $\langle v \rangle = 6.7 \pm 0.2 \mu\text{m/s}$), the TopIB Y70A mutant (beige circles, $\langle v \rangle = 8.9 \pm 0.6 \mu\text{m/s}$), nicking enzyme (blue triangles, $\langle v \rangle = 10.5 \pm 0.2 \mu\text{m/s}$) and for a mix of TopIB mutant Y274F and nicking enzyme (green squares, $\langle v \rangle = 10.5 \pm 0.2 \mu\text{m/s}$). Solid lines are Gaussian fits to the data. Means are numerical averages with corresponding standard errors of the means.

[21] in a manner that caused a decreased rotation rate. This was accomplished by a mixing experiment in which the nicking enzyme was reacted with the DNA in the presence of the Vaccinia TopIB active site mutant Y274F (green squares), which binds DNA noncovalently but is incapable of transesterification [22]. The rotation rate observed in the mixed reaction was indistinguishable from the rate observed with the nicking enzyme only. Second, we measured the effect of the Vaccinia Y70A mutation on the DNA rotation rate. Tyr70 is located on the concave surface of the N-terminal domain of Vaccinia TopIB that clamps over the DNA duplex in the major groove on the helical face opposite the scissile phosphodiester [23]. Due to the loss of this amino acid side chain lining the protein-DNA interface, we observe an increase in the DNA rotation rate (beige circles) compared to wild-type Vaccinia TopIB. Extending these measurements

to human TopIB (dark red triangles), we found that the DNA rotation rate is lower than for wild-type *Vaccinia* TopIB. Human TopIB encircles the DNA duplex fully [7], while *Vaccinia* TopIB has the form of a C-shaped clamp [6, 8]. Accordingly, human TopIB presumably offers less freedom for the DNA to rotate than *Vaccinia* TopIB. These data provide what is, to our knowledge, the first direct measurement of friction [24, 25] in the TopIB relaxation mechanism.

4.4 Torque drive DNA rotation over a rugged energy landscape

We propose a model that describes the effect of friction and torque on the enzymatic activity. Our model has three ingredients. First, the rotation of the DNA inside the enzyme clamp is not free, but hindered by friction, as indicated by our velocity measurements. This is modeled by a random walk over a rugged free energy landscape (Fig. 4.4) with the rotation angle θ between the 5'-OH end of the noncovalently held strand and the tyrosine-3'-DNA adduct as the reaction coordinate. The DNA rotation is not smoothly continuous, but the free energy profile and accordingly the rotation rate varies during a single rotation. This variability in rotation speed could stem from the notion that the cross-sectional size of the DNA at the nick changes dramatically, from 2 nm at $\theta = 0$ to about 4 nm at $\theta=180^\circ$. We make no assumptions about the exact shape of the energy landscape. The rate k across each of the barriers with height ΔG in this landscape follows an Arrhenius relation $k \sim \exp(-\Delta G/k_B T)$, where k_B is the Boltzmann constant and T the temperature in Kelvin. Second, the mechanically applied constant torque Γ_c drives the uncoiling. It is modeled by tilting the entire energy landscape by $-\Gamma_c \theta$. This decreases ΔG by an amount $\Gamma_c \delta \theta$, where $\delta \theta$ is the angle from the well to the transition state. The rate in the presence of a torque Γ_c thus becomes: $k' \sim \exp(-\Delta G'/k_B T) = \exp(-(\Delta G - \Gamma_c \delta \theta)/k_B T)$. The force dependence of Γ_c is given by [26], where ξ is the bending persistence length of a dsDNA molecule [27] ($\xi=53\pm 2$ nm). Third, within each 2π rotation of the DNA inside the topoisomerase cavity, there is only one position with a significant probability to religate the DNA. This is reasonable because the rotating 5'-OH end needs to be in close proximity (on the order of a few Ångströms) to the DNA-3'-phosphotyrosine adduct before a religation can occur. At this position in the energy landscape, there is a possibility to establish a covalent bond, with rate k_r . The religation probability per turn p is thus given by $p \equiv T_{res} \cdot k_r = k_r/k'$, where $T_{res} \equiv 1/k'$ is defined as the residence time in the well at the religation location. As a function of torque, and thus force, we deduce

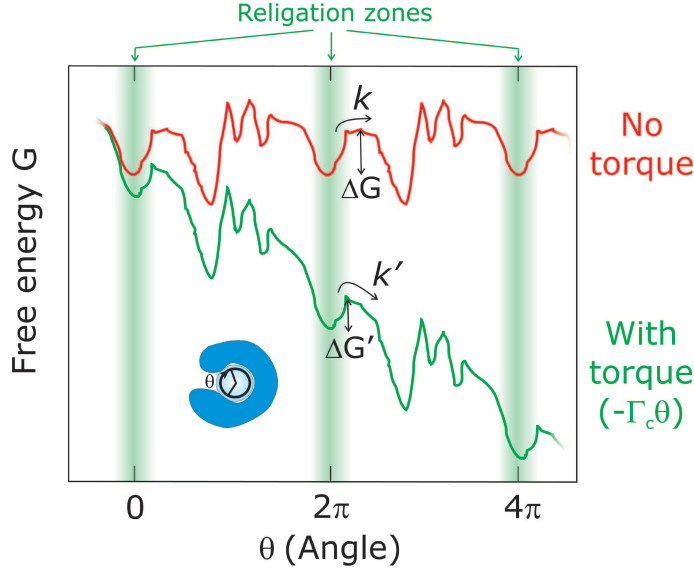


Figure 4.4: Schematic description of the model. The free energy associated with the angle of rotation between the 5'-OH end of the noncovalently held strand and the tyrosine-3'-DNA adduct in the absence (red curve) and presence (green curve) of torque in the DNA. The effect of the torque is to tilt the landscape, decreasing the barrier height ΔG to $\Delta G'$. This increases the escape rate k to k' from the well in which the rate of religation is maximal. This effectively decreases the probability per turn to religate (see text).

that $p(F) \sim \exp(-\Gamma_c \delta\theta / k_B T) = \exp(-\delta\theta \sqrt{2\xi F / k_B T})$ or,

$$\langle \Delta Lk(F) \rangle = \langle \Delta Lk \rangle_{F=0} e^{\delta\theta \sqrt{2\xi F / k_B T}}. \quad (4.2)$$

Equation 4.2 provides a good fit of the data for $\langle \Delta Lk \rangle$ vs. force (red line in Fig. 4.2c). We conclude that our model provides a good description of the single-molecule data at positive supercoils. We obtain an estimate for $\delta\theta$ of 0.23 ± 0.02 radians (13°) and an estimate for $\Delta Lk_{F=0}$ of 19.3 ± 2.3 supercoils/cycle. These numerical values could be specific to the removal of positive (rather than negative) supercoils. Bulk measurements performed by Stivers et al.¹¹ on plasmids containing 15 ± 2 supercoils yielded an average number of supercoils removed per cleavage-religation cycle of 5 ± 1.5 . This value was indirectly obtained using ensemble-averaged rate constants, whereas our experiment measures it directly. In addition, the low initial number of supercoils applied in the plasmid (roughly an order of magnitude lower than in our measurements) may have restricted the observation of large numbers of supercoils removed per cleavage-religation cycle. This could have biased the average number of supercoils released in bulk towards

lower values.

An alternative model could consider the stretching force F applied to the DNA rather than the torque as the parameter governing the religation probability. In such a model, the enzyme would have to perform work over a distance δx against F to religate the DNA. Fitting such a force-dependent model to our data yields a very large value of 20 Å for the force-sensitive step δx . However, earlier work [28] showed that although Vaccinia TopIB can religate 5'-OH DNA across a 1-nucleotide gap, *i.e.* a distance roughly 3 Å larger than a nick, the religation rate for this reaction is already decreased by a factor of 200 in comparison to the religation rate across a nick. It thus seems unrealistic that one would observe any religation events with a separation of 20 Å. Accordingly, we do not favor such a force-dependent mechanism.

4.5 Acknowledgements

We thank Peter Veenhuizen for help in constructing various DNA constructs, David Bensimon, Ulrich Keyser, Ralf Seidel, Keir Neuman, Ligeng Tian and Bastiaan Spanjaard for stimulating discussions, Chris Wiggins for advice on statistical data analysis, David Lubensky for discussions on polymer physics, Jelle van der Does for machining of flow cells, and an anonymous referee for helpful comments. We thank FOM and NWO for financial support. The authors declare that they have no competing financial interests. Correspondence and requests for materials should be addressed to N.H.D. (e-mail: nynke.dekker@mb.tn.tudelft.nl).

4.6 Materials and methods

Enzymes and buffers

Wild-type and mutant vaccinia TopIB proteins were purified as described previously [29]. Nicking enzymes N.BbvCIA, N.BbvCIB and N.BstNBI were purchased from New England Biolabs. Human TopIB (100 kDa fragment) was purchased from Topogen. The step size distributions were measured in buffer containing 10 mM Tris-HCl, pH 8.0, 50 mM NaCl, 10mM MgCl₂, 1 mM DTT, 0.1% Tween-20, 200 µg/ml BSA. TopIB concentrations varied between 0.5 and 20 nM. The rotation rate measurements were performed in identical buffer except that 2 mM MgCl₂ was used. The three nicking enzymes gave identical results, both in the step size and the rotation rate measurements.

DNA constructs Step-size measurements were performed with bacteriophage λ DNA molecules (48kb or 16µm contour length) that were coated at one extrem-

ity with multiple biotin groups and at the other with multiple digoxigenin groups. The rotation rate and the ΔLk distribution measurements at 0.2 pN were performed on half bacteriophage λ DNA molecules. In all molecules, the consensus sequence for Vaccinia TopIB cleavage (5'-CCCTT↓, where ↓ denotes the cleavage site) appears frequently (63 times for the full-length λ DNA molecule). The DNA molecules were incubated with Streptavidin-functionalized magnetic beads (1 μ m diameter, Dynal) and introduced to the flow cell.

Magnetic tweezers/flow cell The detailed experimental configuration of the magnetic tweezers, three-dimensional bead tracking, and force measurements have been described previously [19]. We measure F with 5% accuracy by continuously determining the three-dimensional bead position with 10 nm accuracy [19].

A custom-made flow cell was used, consisting of two rectangular glass microscope cover slides separated by a single layer of parafilm. The lower slide was coated with polystyrene and anti-digoxigenin. This surface was subsequently passivated with BSA.

Data analysis For the step size measurements, data traces were low-pass filtered at 2 Hz and averaged for > 2 s, depending on force. Changes in DNA extension were analysed by making use of a sliding averaging window in which steps were accepted that were larger than three standard deviations of the Brownian noise of the bead [15]. The steps can be identified as arising from a single topoisomerase enzyme because the time it takes for the enzyme to remove supercoils is much smaller than the time between successive relaxation events. The bead velocity measurements were analysed from raw data obtained at 60 Hz. Treatment of step size distributions. Values for $\langle \Delta Lk \rangle$ were obtained using a modified maximum-likelihood method that takes into account the experimental bead noise and the fact that one cannot observe steps of $\Delta Lk > n$, where n is the number of supercoils in the DNA prior to cleavage. Therefore, the final step leading to the complete removal of plectonemes from the DNA is discarded. To correct for the ensuing overrepresentation of small steps, one takes into account that each measured step ΔLk_i is not drawn from the entire ΔLk distribution but rather from the probability distribution $P_i(\Delta Lk) = (N_i \langle \Delta Lk \rangle)^{-1} \exp(-\Delta Lk_i / \langle \Delta Lk \rangle)$, where N_i is given by $N_i = \exp(-\Delta Lk_{noise} / \langle \Delta Lk \rangle) - \exp(-\Delta Lk_{constrained,i} / \langle \Delta Lk \rangle)$, ΔLk_{noise} is the smallest observable step given the bead noise and $\Delta Lk_{constrained,i}$ is the remaining number of supercoils in the DNA molecule prior to the i -th step. The corresponding likelihood function L is thus given by: $L = \prod_{i=1}^N (N_i \langle \Delta Lk \rangle)^{-1} \exp(-\Delta Lk_i / \langle \Delta Lk \rangle)$. This modified maximum likelihood method is especially useful at higher stretching forces, where failure to apply the method underestimates $\langle \Delta Lk \rangle$ by approximately 25%.

References

- [1] Champoux, J. J. *Annu Rev Biochem* **70**, 369–413 (2001).
 - [2] Corbett, K. D. and Berger, J. M. *Annu Rev Biophys Biomol Struct* **33**, 95–118 (2004).
 - [3] Wang, J. C. *Annu Rev Biochem* **65**, 635–92 (1996).
 - [4] Liu, L. F., Desai, S. D., Li, T. K., Mao, Y., Sun, M., and Sim, S. P. *Ann N Y Acad Sci* **922**, 1–10 (2000).
 - [5] Pommier, Y., Pourquier, P., Fan, Y., and Strumberg, D. *Biochim Biophys Acta* **1400**(1-3), 83–105 (1998).
 - [6] Cheng, C., Kussie, P., Pavletich, N., and Shuman, S. *Cell* **92**(6), 841–50 (1998).
 - [7] Redinbo, M. R., Stewart, L., Kuhn, P., Champoux, J. J., and Hol, W. G. *Science* **279**(5356), 1504–13 (1998).
 - [8] Sekiguchi, J. and Shuman, S. *J Biol Chem* **269**(50), 31731–4 (1994).
 - [9] Kim, R. A. and Wang, J. C. *J Mol Biol* **208**(2), 257–67 (1989).
 - [10] Stewart, L., Redinbo, M. R., Qiu, X., Hol, W. G., and Champoux, J. J. *Science* **279**(5356), 1534–41 (1998).
 - [11] Stivers, J. T., Harris, T. K., and Mildvan, A. S. *Biochemistry* **36**(17), 5212–22 (1997).
 - [12] Brown, P. O. and Cozzarelli, N. R. *Science* **206**(4422), 1081–3 (1979).
 - [13] Brown, P. O. and Cozzarelli, N. R. *Proc Natl Acad Sci U S A* **78**(2), 843–7 (1981).
 - [14] Brown, P. O., Peebles, C. L., and Cozzarelli, N. R. *Proc Natl Acad Sci U S A* **76**(12), 6110–4 (1979).
 - [15] Dekker, N. H., Rybenkov, V. V., Duguet, M., Crisona, N. J., Cozzarelli, N. R., Bensimon, D., and Croquette, V. *Proc Natl Acad Sci U S A* **99**(19), 12126–31 (2002).
 - [16] Liu, L. F., Liu, C. C., and Alberts, B. M. *Cell* **19**(3), 697–707 (1980).
 - [17] Mizuuchi, K., Fisher, L. M., O’Dea, M. H., and Gellert, M. *Proc Natl Acad Sci U S A* **77**(4), 1847–51 (1980).
 - [18] Strick, T. R., Croquette, V., and Bensimon, D. *Nature* **404**(6780), 901–4 (2000).
-

- [19] Strick, T. R., Allemand, J. F., Bensimon, D., Bensimon, A., and Croquette, V. *Science* **271**(5257), 1835–7 (1996).
 - [20] Rice, J. A. *Mathematical Statistics and Data Analysis*. Wadsworth & Brooks/Cole Advanced Books & Software, Pacific Grove, California, (1988).
 - [21] Shuman, S., Bear, D. G., and Sekiguchi, J. *Embo J* **16**(21), 6584–9 (1997).
 - [22] Sekiguchi, J. and Shuman, S. *Nucleic Acids Res* **22**(24), 5360–5 (1994).
 - [23] Sekiguchi, J. and Shuman, S. *Embo J* **15**(13), 3448–57 (1996).
 - [24] Carey, J. F., Schultz, S. J., Sisson, L., Fazzio, T. G., and Champoux, J. J. *Proc Natl Acad Sci U S A* **100**(10), 5640–5 (2003).
 - [25] Woo, M. H., Losasso, C., Guo, H., Pattarello, L., Benedetti, P., and Bjornsti, M. A. *Proc Natl Acad Sci U S A* **100**(24), 13767–72 (2003).
 - [26] Moroz, J. D. and Nelson, P. *Proc Natl Acad Sci U S A* **94**(26), 14418–22 (1997).
 - [27] Bustamante, C., Bryant, Z., and Smith, S. B. *Nature* **421**(6921), 423–7 (2003).
 - [28] Krogh, B. O. and Shuman, S. *Biochemistry* **39**(21), 6422–32 (2000).
 - [29] Shuman, S., Golder, M., and Moss, B. *J Biol Chem* **263**(31), 16401–7 (1988).
-

Chapter 5

Multiple events on single molecules: unbiased estimation in single-molecule biophysics

Most analyses of single-molecule experiments consist of binning experimental outcomes into a histogram and finding the parameters that optimize the fit of this histogram to a given data model. Here we show that such an approach can introduce biases in the estimation of the parameters and that thus great care must be taken in the estimation of model parameters from the experimental data. The bias can be particularly large when the observations themselves are not statistically independent and are subjected to global constraints, as, for example, when the iterated steps of a motor protein acting on a single molecule must not exceed the total molecule length. We have developed a maximum-likelihood analysis, respecting the experimental constraints, which allows for a robust and unbiased estimation of the parameters, even when the bias well exceeds 100%. We demonstrate the potential of the method for a number of single-molecule experiments, focusing on the removal of DNA supercoils by topoisomerase IB, and validate the method via numerical simulation of the experiment.

5.1 Introduction

Over the past few years, single-molecule techniques have started to deliver on their promise as high-resolution tools for the study of biological systems. The activity of single proteins such as kinesin, myosin, and topoisomerases [1–3], amongst others, has been monitored in real time. A hallmark of such single-molecule experiments, in contrast to bulk experiments, is their unparalleled ability to yield the functional form of the distribution of experimental outcomes, and not merely their averages [4] or other statistics [5]. Estimating the parameter values that characterize these distributions often yields the information required to construct detailed mechanical models of the system under investigation.

To obtain these parameter values from an experiment, observables are typically binned into a histogram and the histogram is fitted to the predictions of a model. An alternative method to obtain a distribution parameter is to use the maximum likelihood method, in which one calculates the value of an unknown parameter in a distribution which maximizes the likelihood of the experimentally observed data [6, 7]. The maximum likelihood method has the advantage that one does not discard information, or introduce one’s own biases, in the data through binning. Moreover, the histogram-fitting approach, at least when squared loss is used, ignores the fact that the errors induced in the construction of the histogram are themselves a function of the model and of the number of counts represented in each bin of the histogram. Another important advantage of using the maximum likelihood method, which we demonstrate below, is the possibility to build a model that is more faithful to the experimental reality. Particularly in biophysical experiments where a multitude of factors, such as finite-size or other experimental or biological constraints, unavoidably thwart the assumption that each individual observation is independent and identically distributed (referred to as the ‘i.i.d.’ assumption below), the maximum-likelihood approach facilitates building a model which is both more experimentally sound and more statistically robust.

Frequently, constraints emerge because of experimental limitations in detecting all values of experimental outcomes in a distribution: one receives from the measurement a limited range of values instead of the entire domain. In some experiments, for example the DNA translocation by the enzyme FtsK described in [8], the experimental outcomes are uncoupled from one another and are ‘i.i.d.’. However, in the scenario that experimental outcomes are coupled to each other by a global constraint, the range of values that can be detected varies with every new measurement taken. As we demonstrate below, the presence of global constraints is a factor that absolutely requires maximum likelihood analysis if

the biological parameters of a system are to be measured accurately. In principle, the analysis of bulk experiments can be hampered as much as the analysis of single-molecule experiments. However, in single-molecule measurements, one can evaluate each experimental outcome with respect to the constraints. Armed with this knowledge, one can apply the mathematical treatment outlined here and counter the bias in the data accurately.

In this article, we illustrate the problem of global constraints by showing how the measurement process in a single-molecule study of DNA supercoil relaxation by the enzyme topoisomerase IB imposes global constraints on the probability distribution from which the experimental outcomes are drawn. Subsequently, we generalize the maximum likelihood method for parameter estimation, enabling one to faithfully recover the unbiased estimate of the distribution parameter from data subject to global constraints. We also derive an expression for the standard deviation of the recovered parameter as a function of the available statistics. Numerical methods confirm our ability to recover the unbiased distribution parameter within the error estimation derived. Finally, we show that the method introduced here can play an important role in the extraction of biological parameters from several other single-molecule experiments.

5.2 Topoisomerase IB steps are subjected to a global constraint

We first illustrate the concept of global constraints using data obtained from the single-molecule analysis of topoisomerase IB [9]. Topoisomerase IB is an enzyme that removes supercoils from a dsDNA molecule by transiently introducing a nick [10, 11]. As long as the dsDNA molecule is nicked, torque present in the molecule will swivel the DNA about its intact strand. After a random number of supercoils are released, the enzyme religates the DNA, which terminates the removal of supercoils [9].

We can follow the action of the topoisomerase in real time using magnetic tweezers [12]. The experimental strategy is described elsewhere [1] and summarized in Fig. 8.3. Each time the topoisomerase removes supercoils from the DNA molecule, one observes a discrete step in the height of a μm -sized bead attached to the molecule. The height of the bead is equal to the extension of the DNA molecule and is directly related to its linking number (Lk) and to the number of supercoils present in the DNA. A small extension of the DNA corresponds to a large number of supercoils present, while a large extension corresponds to few supercoils present in the DNA. Thus, each time the topoisomerase removes su-

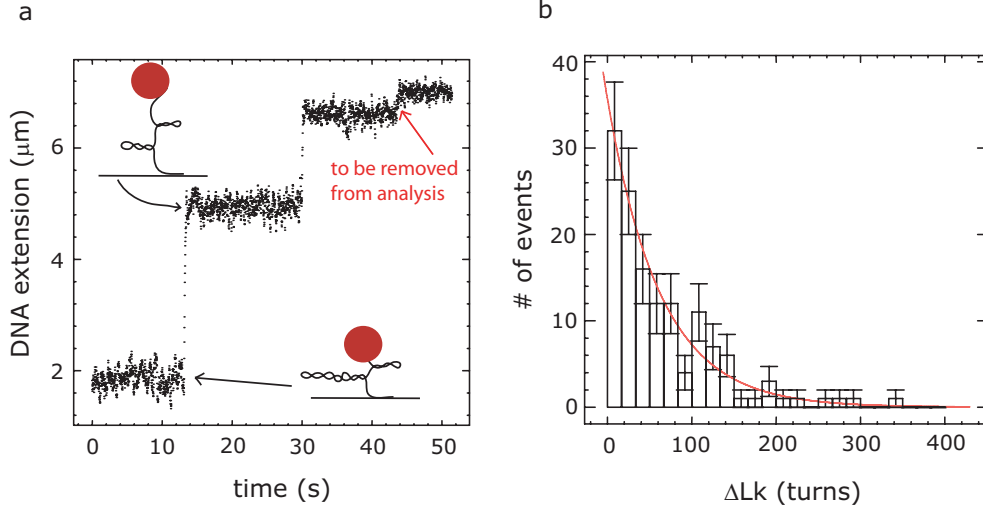


Figure 5.1: An example of a system that includes global constraints. (a) Topoisomerase IB removes DNA supercoils in steps. Each time the topoisomerase removes a number of supercoils, the DNA extension rises in a step-wise fashion. The final step which leads to the removal of the remaining supercoils in the DNA is artificially constrained and should be removed from the analysis. (b) the size of the steps (in units of change in linking number ΔLk) is distributed exponentially. In the text, this will be referred to as the 'measured' distribution. This 'measured' distribution may differ from an underlying 'true' distribution due to the presence of global constraints.

percoils from the DNA, we observe a discrete step in the DNA extension, which is proportional to ΔLk . If the probability of religation (per turn) is constant, the distribution of (ΔLk) should be an exponential (Fig. 8.3b); the average of ΔLk , which is denoted $\langle \Delta Lk \rangle$, is the parametric description of topoisomerase activity we wish to deduce from the experiment.

The setup of the experiment, in which the DNA molecule only contains a limited number of supercoils, necessarily introduces global constraints on the distribution. Consequently, at some point the topoisomerase will inevitably remove the last few supercoils that remain in the DNA (red arrow, Fig. 8.3a). This final step towards the level of zero supercoils contains only limited information in comparison to previous steps. This is because the final step is artificially constrained by the fact that no more supercoils remain in the DNA for the topoisomerase to remove. Therefore, when drawing conclusions about the working of the enzyme, one should discard this final step. For convenience, we will define sub-steps as those steps that do not extend to the level of zero supercoils. Effectively, steps so large that they become the final step are discarded, whereas steps so small that they become sub-steps are not discarded. This leads to an overrepresentation of small steps. When one simply analyzes the surviving sub-steps, one obtains

a skewed distribution with an incorrect parameter, which can hamper a proper interpretation of the system under investigation. In an actual experiment, one cannot distinguish between the "true" distribution and the distribution that is skewed as a result of the measurement. After all, all one has is the measured distribution (Fig. 8.3b), which is skewed. Fortunately, in a single-molecule measurement this can be corrected for by the method we describe below.

5.3 Maximum likelihood and domain constraints

We briefly review the concept of parameter estimation using the maximum likelihood method [7]. Let $P(s|k)$ be a properly normalized probability density function (pdf) for step size s , with parameter k . The goal of the maximum likelihood method is to obtain an estimate for the parameter of the pdf, which in this general case is k . Since the experimental outcomes are assumed to be statistically independent, the combined probability to find, in n measurements, the data $s_1, s_2, \dots, s_n$ is given by:

$$L(k) = \prod_{i=1}^n P(s_i|k), \quad (5.1)$$

where L is the likelihood function. To avoid working with very large or small numbers that can cause computational inaccuracies, and to facilitate the analysis, one often works with the logarithmic likelihood $\ln L(k) = \sum_{i=1}^n \ln P(s_i|k)$. We now introduce k_* , the value of k that maximizes L , otherwise known as the maximum likelihood value. It is the best estimate for k and we wish to calculate its value. We obtain k_* by solving

$$0 = \partial_k \ln P(k|s) \quad (5.2)$$

(see supporting information). Assuming that the shape of $\ln P(k|s)$ near k_* is a Gaussian distribution (see supporting information), we can calculate the variance σ^2 of k_* , using

$$-\sigma^{-2} \equiv \partial_k^2 \ln P(k|s) \Big|_{k=k_*}. \quad (5.3)$$

In many experimental scenarios, constraints apply to the measurable domain of P . In other words, it may not be experimentally possible to sample all possible

values of s . A proper analysis of the data taken in this experimental scenario then requires P to be renormalized by a weighting function $g(k)$:

$$g(k) = \int_{s_{min}}^{s_{max}} ds P(s|k), \quad (5.4)$$

where s_{min} is the minimum value for s that can be detected, and s_{max} is the maximum value for s that can be detected. In this simplest case, s_{min} and s_{max} are constant for each measurement i of s . However, an alternative possibility is for a global constraint to couple all observations (indexed by i) of the variable s to each other. In this second case, one requires a weighting function that varies with each measurement i of s :

$$g_i(k) = \int_{s_{min,i}}^{s_{max,i}} ds P(s|k), \quad (5.5)$$

where $s_{min,i}$ and $s_{max,i}$ are again the minimum and maximum detectable values for s , respectively, but their values are not fixed for all measurements of s . Instead, $s_{min,i}$ represents the minimum detectable value for s that is valid only for the i -th measurement of s . Similarly, $s_{max,i}$ represents the maximum detectable value for s that is valid only for the i -th measurement of s . Analogously to the 'i.i.d.' case, one can calculate the likelihood function for all of the constrained data, maximize this function and obtain k . The value for k_* we obtain in this manner is then the unbiased estimate of the parameter of the distribution.

To illustrate the method explicitly, we use an exponential function as a pdf, the appropriate model for a topoisomerase that removes supercoils from DNA with constant probability per turn of religation. The normalized pdf for $0 < s < \infty$ is then given by

$$P(s|k) = ke^{-ks}, \quad (5.6)$$

where $k = \frac{1}{\langle s \rangle}$. Here and in the following, brackets indicate averages over the experimental observations. The corresponding likelihood function is given by

$$L(s_1, s_2, \dots, s_n|k) = \prod_{i=1}^n P(s_i|k) = \prod_{i=1}^n ke^{-ks_i}. \quad (5.7)$$

In the case of 'i.i.d.' observations, we obtain

$$g(k) = \int_{s_{min}}^{s_{max}} ds ke^{-ks} = e^{-ks_{min}} - e^{-ks_{max}}. \quad (5.8)$$

However, in the case of global constraints, we obtain:

$$g_i(k) = \int_{s_{min,i}}^{s_{max,i}} ds k e^{-ks} = e^{-ks_{min,i}} - e^{-ks_{max,i}} \quad (5.9)$$

and the values for s are drawn from

$$p(s_i|k) = \frac{k e^{-ks_i}}{e^{-ks_{min,i}} - e^{-ks_{max,i}}}. \quad (5.10)$$

Having obtained a relation for P , we can calculate the corresponding likelihood. The probability of the data in the presence of global constraints is

$$L_i(k) = \prod_{i=1}^N \frac{k e^{-ks_i}}{e^{-ks_{min,i}} - e^{-ks_{max,i}}}, \quad (5.11)$$

where N is the number of experimental outcomes of s . The logarithm of L is given by

$$\ln L_i = N \ln k - k \sum_{i=1}^N s_i - \sum_{i=1}^N \ln(e^{-ks_{min,i}} - e^{-ks_{max,i}}). \quad (5.12)$$

The parameter measured in the topoisomerase IB experiment is the average change in linking number $\langle \Delta Lk \rangle$, which is equal to $\langle s \rangle$ in the terminology used above. Since $\langle \Delta Lk \rangle = 1/k$, we take the derivative of equation [12] with respect to $1/k$:

$$\partial_{\langle \Delta Lk \rangle} \equiv \partial_{1/k} = -kN + k^2 \sum_{i=1}^N s_i - k^2 \sum_{i=1}^N \frac{s_{min,i} e^{-ks_{min,i}} - s_{max,i} e^{-ks_{max,i}}}{e^{-ks_{min,i}} - e^{-ks_{max,i}}}. \quad (5.13)$$

We find the maximum in the likelihood by setting equation [13] equal to zero,

$$0 = 1/k_* - \langle s \rangle + \left\langle \frac{s_{min,i} e^{-k_* s_{min,i}} - s_{max,i} e^{-k_* s_{max,i}}}{e^{-k_* s_{min,i}} - e^{-k_* s_{max,i}}} \right\rangle, \quad (5.14)$$

where k_* is again the maximum likelihood value for k , the value that solves equation [14] (the summation signs in equation [13] have been replaced by brackets in equation [14] to denote averages). Equation [14] can be evaluated numerically to yield $\langle \Delta Lk \rangle_*$, the maximum likelihood value of $\langle \Delta Lk \rangle$. We deduce that

the variance of $\langle \Delta Lk \rangle_*$ is given by

$$\sigma_{\langle \Delta Lk \rangle}^{-2} = k_*^2 N - \frac{k_*^4}{4} \left\langle \frac{s_{min,i} - s_{max,i}^2}{\sinh^2(\frac{1}{2}k_*(s_{min,i} - s_{max,i}))} \right\rangle \quad (5.15)$$

(see also supporting information). Comparing [14] and [15] to the case in which no constraints apply, or $s_{max} = \infty$ and $s_{min} = 0$, we recover

$$\langle s \rangle = \frac{1}{k} = \langle \Delta Lk \rangle \quad (5.16)$$

and

$$\sigma_{\langle \Delta Lk \rangle} = \frac{1}{k\sqrt{N}} = \frac{\langle \Delta Lk \rangle}{\sqrt{(N)}}, \quad (5.17)$$

as expected. Equations [14] and [15] can be directly applied to the single-molecule data of DNA supercoil relaxation by TopIB (Fig. 8.3b) in order to determine the true parameter of the underlying true distribution and its associated standard deviation as a function of the number of experimental outcomes N .

5.4 Numerical simulation and the consequences of ignoring global constraints

We simulate the measurement process in a single molecule experiment to quantify the biasing effect on a "true distribution" as a result of the global constraints and the "sampling error" inherent in the finite number of observations performed. The applicability of equations [14] and [15] to the determination of the value and the standard deviation of the distribution parameter can therefore be assessed.

We start by generating an exponential distribution characterized by a parameter that we define as the "true parameter" and is denoted $\langle \Delta Lk \rangle_{true}$. We arbitrarily set it to $\langle \Delta Lk \rangle_{true} = 60$. $\langle \Delta Lk \rangle_{true}$ represents the parameter of the distribution that would be measured in the absence of constraints. We call this unbiased distribution the "true distribution". Since it is unbiased, we can think of this distribution as representing the physics governing the workings of the enzyme. We then simulate the process of removing supercoils from a DNA molecule that has a maximum of 130 supercoils present (the global constraint ΔLk_{max}^0 , Fig. 8.4a, inset). We use these values for all simulations. The number of supercoils that the topoisomerase removes each time from the DNA is randomly drawn from our "true distribution". As described above, all final steps

are subsequently discarded, while the sub-steps that remain are displayed in a histogram. This histogram reflects what we would measure experimentally and we call it the "measured distribution", characterized by a "measured parameter", which is biased and therefore denoted $\langle \Delta Lk \rangle_{biased}$. The "true distribution" is shown in blue in Fig. 8.4a, while the "measured distribution" is shown in red. As can be clearly seen from Fig. 8.4a, the two distributions are not identical. The "true distribution" obviously yields an average of 60 (in units of ΔLk). We obtain $\langle \Delta Lk \rangle_{biased}$ by fitting the "measured distribution" to an exponential in the range between zero and ΔLk_{max}^0 . In fact, the functional form is altered slightly due to the global constraints, as discussed formally in the supporting information. Note that the value for $\langle \Delta Lk \rangle_{biased}$ is thus biased due to a combination of factors: (i) the presence of global constraints, (ii) the number of experimental outcomes N , and (iii) the analysis via histogram-fitting rather than maximum likelihood. For the particular values for $\langle \Delta Lk \rangle_{true}$ and ΔLk_{max}^0 we used, $\langle \Delta Lk \rangle_{biased}$ was 46, which is an underestimate of $\sim 23\%$ in comparison to $\langle \Delta Lk \rangle_{true}$. Indeed, the measurement process has biased small steps over large steps, skewing the "measured distribution" towards lower values. We now focus more closely on the relationship between the magnitude of the constraint and the resulting degree of bias. Fig. 8.4b plots $\langle \Delta Lk \rangle_{biased}$ as a function of the severity of the global constraint ΔLk_{max}^0 . We plot $\langle \Delta Lk \rangle_{true}=60$ as a blue horizontal line in Fig 8.4b. The discrepancy between $\langle \Delta Lk \rangle_{biased}$ and $\langle \Delta Lk \rangle_{true}$ caused by the global constraints is thus reflected graphically as the distance between the red curve and the blue line; in the absence of any biasing effect, all values for $\langle \Delta Lk \rangle_{biased}$ would fall on top of the blue line. We describe three salient features of Fig 8.4b. First, as the constraint becomes less severe (ΔLk_{max}^0 increases), the magnitude of the bias decreases. Conversely, as the constraint becomes more severe (ΔLk_{max}^0 decreases), the magnitude of the bias increases. Second, the discrepancy between $\langle \Delta Lk \rangle_{biased}$ and $\langle \Delta Lk \rangle_{true}$ is very large ($>100\%$) for small values of ΔLk_{max}^0 . For example, for $\Delta Lk_{max}^0=20$, $\langle \Delta Lk \rangle_{biased}$ equals 27, which constitutes an underestimation of $\langle \Delta Lk \rangle_{true}$ by approximately 120%. Although this is an example which might not generally be observed experimentally, we include it to emphasize that our method can recover $\langle \Delta Lk \rangle_{true}$ robustly even in the case of extreme bias, as we show below. The third feature of Fig 8.4b highlights that in a regime where one naively would expect virtually no biasing effect due to the constraints, the bias is significant nevertheless. Indeed, for $\Delta Lk_{max}^0=800$, which is well over an order of magnitude larger than $\langle \Delta Lk \rangle_{true}$, one still observes that $\langle \Delta Lk \rangle_{true}$ is underestimated by approximately 7%. This surprising behavior stems from the fact that the constraints on the distribution vary from step to step and are on average smaller

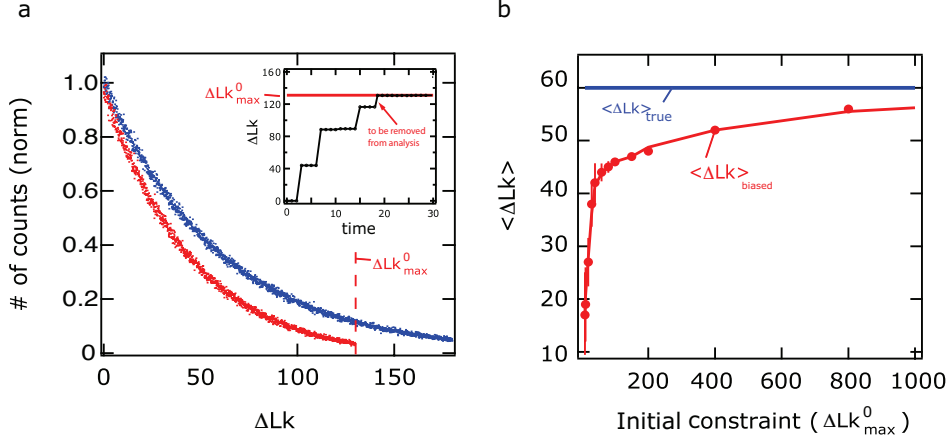


Figure 5.2: Simulated step-size distributions for the enzymatic removal of supercoils from the DNA molecule. (a) the number of supercoils that the enzyme removes each time from the DNA molecule is randomly drawn from a generated exponential distribution, called the "true distribution" (blue dots). The "true distribution" is characterized by an average of 60 (units of ΔLk). After discarding the final steps leading to the level of zero supercoils (ΔLk^0_{max} , see text), one obtains a "measured distribution" (red dots) whose parameter is underestimated ($\langle \Delta Lk \rangle = 46$). Inset: Numerical simulation of the enzymatic removal of supercoils. The size of each step is drawn from the "true distribution". As in reality, the DNA molecule simulated contains only a limited number of supercoils. The level at which no supercoils are present is depicted as a horizontal red line and acts as a constraint for the removal of supercoils by the enzyme. Since the final step towards the level of zero supercoils (red arrow) is artificially constrained, this final step is removed from the data analysis (see text). (b) the degree to which the "measured" parameter is underestimated is a function of the constraints (the initial maximum number of supercoils in the DNA, denoted ΔLk^0_{max}). As the constraints become more pronounced, the underestimation grows. In some cases, the underestimation of ΔLk due to global constraints is severe ($>100\%$). The "true" value for $\langle \Delta Lk \rangle$ is depicted as a horizontal blue line, which the measured value for $\langle \Delta Lk \rangle$ (red dots) approaches asymptotically (red line is a spline through the data points).

than ΔLk^0_{max} . We now describe how we can nonetheless obtain an accurate value for $\langle \Delta Lk \rangle_{true}$, even in cases where $\langle \Delta Lk \rangle_{true}$ is severely underestimated.

By monitoring the DNA extension, either in a real experiment or in the simulation as discussed here, we know the number of supercoils that remain in the DNA molecule before the topoisomerase removes a number of supercoils; that is to say, we know the constraints that apply to the measurement of each sub-step. The important point is that although the constraints vary for each step, they are known, and we can therefore substitute their values for $s_{max,i}$ in equation [10]. We also know the value of $s_{min,i}$, which is the minimum detectable num-

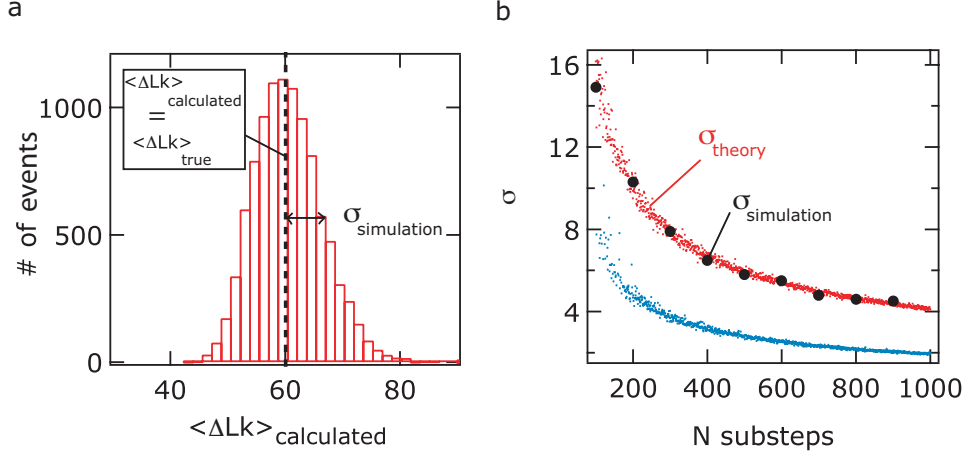


Figure 5.3: Recovery and error calculation of the true distribution parameter using the maximum likelihood method (see text). (a) the distribution of the calculated distribution parameter is generated by solving equation [14] for 10^5 times and binning the outcome of the calculation into bins. The distribution is peaked at the value that characterizes the unbiased step-size distribution ($\langle \Delta Lk \rangle_{\text{true}}$). Importantly, the method thus successfully recovers the unbiased parameter in spite of the biasing effect of global constraints. The standard deviation of the distribution, σ , is numerically calculated. (b) the theoretical standard deviation σ , obtained by solving equation [15], as a function of the number of sub-steps per exponential distribution. The theoretical standard deviation is calculated for constrained (maximum initial $\Delta Lk_{\text{max}}^0 = 130$, red points) and unconstrained (maximum initial $\Delta Lk_{\text{max}}^0 = \infty$, blue points) distributions. The theoretical error in the case of the constrained distribution is compared with the error as calculated from simulations as in Fig. 3a and is shown as black solid circles. The theoretical error calculated using equation [15] predicts the measured error very well.

ber of supercoils removed and which is determined by the noise in the height of the bead. This is beyond the scope of this work, and for all practical purposes, we give $s_{\text{min},i}$ the fixed value of zero. We now solve equation [14] and call the solution $\langle \Delta Lk \rangle_{\text{calculated}}$. In this calculation, we have used $N = 10^5$ sub-steps ΔLk . To get an idea of the reproducibility in $\langle \Delta Lk \rangle_{\text{calculated}}$, we repeat the calculation $Q = 10^5$ times and build a histogram of the solutions (Fig. 8.5a). Importantly, we calculate that the mean of the distribution of $\langle \Delta Lk \rangle_{\text{calculated}}$ is 60, which is identical to the value we have chosen as $\langle \Delta Lk \rangle_{\text{true}}$, the "true parameter" of the "true distribution". Therefore, we conclude that the analysis method accurately recovers the "true parameter", despite the biasing effect of the measurement.

In an experiment, it is not only important to recover the true parameter of the distribution but also to know its associated standard deviation as a

function of the number of experimental outcomes N . We have therefore calculated the standard deviation of the ΔLk -distribution (Fig. 8.5a) according to $\sigma_{simulation} = \sqrt{\frac{1}{Q-1} \sum_{j=0}^Q [\langle \Delta Lk \rangle_{calculated,j} - \langle \langle \Delta Lk \rangle_{calculated} \rangle]^2}$. This procedure is repeated for nine different values of N and their values are plotted as solid black circles in Fig. 8.5b. They can be compared with the theoretically predicted values for σ , denoted as σ_{theory} , calculated using equation [15]. Fig. 8.5b also plots σ_{theory} as a function of N as red and blue dots. Red dots are calculations of σ_{theory} with global constraints, whereas blue dots are calculations of σ_{theory} in the absence of constraints ($\Delta Lk_{max}^0 = \infty$). As is evident from Fig. 8.5b, the solid circles fall on top of the theoretical prediction σ_{theory} given by equation [15]. Thus, we have shown that equation [15] predicts the standard deviation associated with $\langle \Delta Lk \rangle_{true}$ accurately. From this result, we can draw an important conclusion, namely that in any given situation with global constraints, an experimenter can assess whether enough statistics have been obtained in order to determine the unbiased value of the true distribution to the desired accuracy.

5.5 Application of the method

The method outlined above deals with global constraints in the domain of the distribution of experimental outcomes. Therefore, the method should in principle be used in all experiments that involve global constraints and whose experimental outcomes are not distributed like (a series of) δ -functions. An example of outcomes distributed like a δ -function is the fixed step-size of 37 nm with which a myosin protein walks over an actin filament [3]. While experimentally it seems that one measures a Gaussian distribution of observables, the Gaussian shape in fact arises from stochastic fluctuations around a fixed true value. The function describing these processes is a δ -function, peaked at the fixed true value of the observable. Mathematically, this implies that the weight-function ($g_i(k)$, equation [9]) is always equal to one, and consequently the pdf is unaltered by constraints. Therefore, the likelihood function and the observable that maximizes it remain unaffected and one is not required to use this method.

We expect that the analysis method outlined in this manuscript could guide the proper design and analysis of experiments including assays of the processivity of helicases, polymerases and other translocation enzymes, single-molecule Förster resonance energy transfer (FRET) measurements and real-time single-molecule tracking of DNA condensation. For clarity, we describe a few of these experiments in more detail below.

5.6 Processivity measurements on limited substrate

Some substrates, such as ssRNA or dsRNA molecules, are practically hard to prepare in lengths longer than a few kb if they are to be used in single-molecule techniques [13]. If one wishes to measure the distribution of the processivity of a biomolecule that tracks along the RNA, one may find that the processivity exceeds the length of the RNA. In such a case, one is required to discard the final processive action, because it is artificially constrained by the fact that there is no more dsRNA substrate for the biomolecule to move on. The constraint is global, since the length of the RNA molecule that is available for the biomolecule shrinks as it proceeds. This dilemma is summarized in Fig. 8.6a. An example of an enzyme translocating on RNA is the RNA-dependent RNA polymerase P2 from ϕ_6 bacteriophage [14]. This polymerase can perform an RNA synthesis reaction using either dsRNA or ssRNA as a template. The processivity, which can only roughly be estimated from experiments, is on the order of 10 kilobases or more [15] and is comparable to the length of the RNA substrate. In single-molecule processivity measurements for P2 polymerase and other enzymes, we expect that our treatment would be instrumental in determining the mean processivity correctly.

5.7 Transitions in FRET efficiency

FRET efficiency is dependent on the distance between a donor dye and acceptor dye and ranges between zero (no FRET) and one (maximum FRET) [4, 16] (in practice, the range in which meaningful FRET measurements can be performed is even smaller due to lack of sensitivity close to both the no-FRET and the maximum-FRET regimes). Changes in FRET efficiency can in theory be used to quantify conformational changes in biomolecules (*e.g.* in the folding of RNA molecules or in proteins). Future experiments measuring distributions of changes in FRET efficiency could be biased due to the global constraint imposed by the limited meaningful range in FRET efficiency. For example, one could measure a series of conformational changes in an RNA molecule in which each conformational change is associated with a transition in FRET efficiency between a donor and acceptor dye attached to two parts of the RNA molecule (*e.g.* [17–19]), as schematically depicted in Fig. 8.6b. In such an experiment, one would be required to discard those transitions that extend to or exceed the limits of the FRET efficiency range. To correct for the ensuing bias towards small FRET

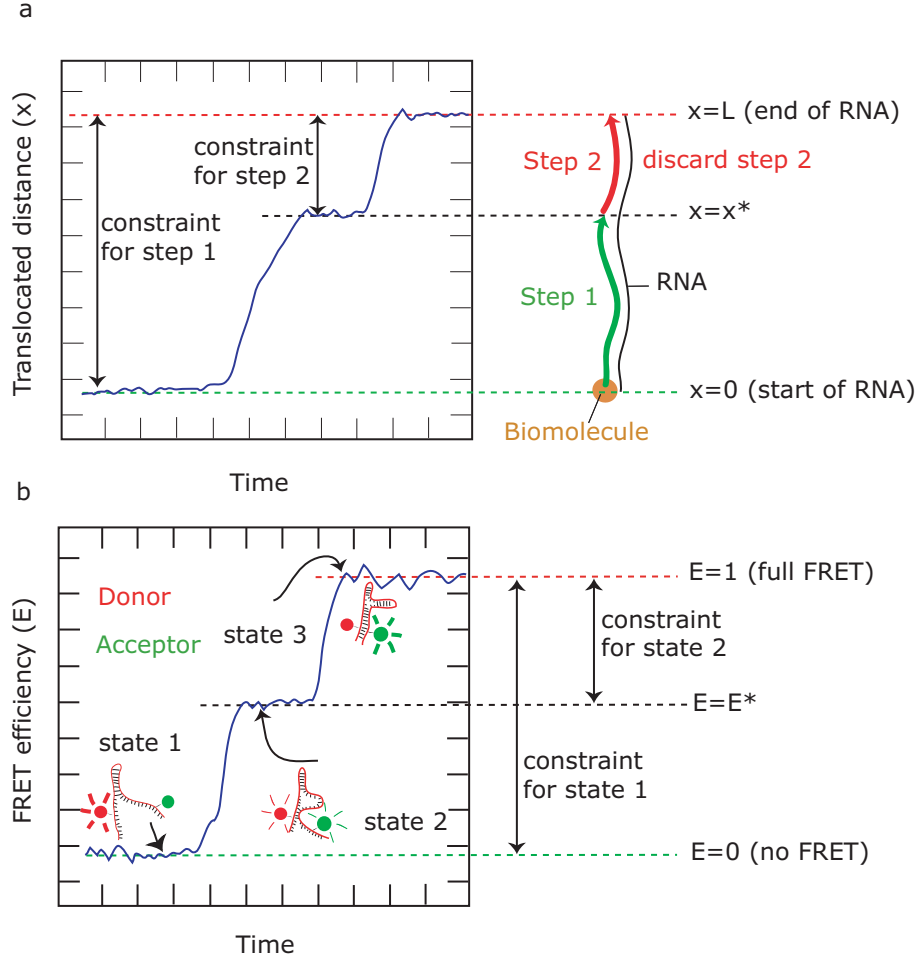


Figure 5.4: Sketch of experiments in which global constraints can bias parameter estimation. (a) processivity of a biomolecule (beige circle) along a short biopolymer such as a ssRNA or dsRNA molecule. When the biomolecule starts its procession, it has the total length of the RNA molecule (the global constraint) at its disposal ($s_{max,0} = L$). It then moves a distance x and stops. From there, it can start moving again, but the biomolecule can now only travel a length $s_{max,1} = L - x^*$, before falling off the RNA. The constraint on the distance the biomolecule can travel along the RNA is different for the first and for the second step. (b) conformational changes in *e.g.* an RNA molecule studied using FRET. The FRET efficiency is defined between 0 and 1, which is the global constraint for the experiment. *E.g.*, at state 1, the FRET efficiency $E = 0$. From this state, the FRET efficiency can only change by 1 at maximum ($s_{max,0} = 1$). However, from an arbitrary intermediate state 2 (at $E = E^*$), it can increase its E only by $s_{max,1} = \Delta E = 1 - E^*$. The constraint on the change in FRET efficiency is thus different for the first and for the second state, as described in the main text.

transitions and thus for a correct analysis of the distribution, one needs to apply the method described here.

5.8 Concluding remarks

Experimental outcomes that are non-globally constrained in that they can be assumed to be 'i.i.d.' can be relatively easily analyzed in their measured range. However, when such analysis is performed on outcomes that are coupled by global constraints, severe bias in the parameter estimation can occur. We have therefore generalized the maximum likelihood method for parameter estimation to include distributions that have global constraints. Using this method, we robustly recover the unbiased distribution parameter from biased data, independent of the severity of the bias. In addition, we have adapted the relation describing errors in the estimation for distribution parameters for the case of global constraints. This allows an experimenter to assess whether enough data points have been accumulated to predict the true parameter to the desired accuracy. Finally, we show that global constraints can occur in a variety of experiments, all of which would benefit from using this method.

5.9 Acknowledgements

We would like to thank Ulrich Keyser for stimulating discussions and Cees Dekker and Thijn van der Heijden for a critical reading of the manuscript. We also acknowledge financial support from the TU Delft, FOM and NWO.

5.10 Supplemental Materials

Proof of $\partial_k \ln P(k|\{l\}) = \partial_k \ln P(\{l\}|k)$

By Bayes' rule,

$$P(k|\{l\}) = P(\{l\}|k)P(k)/P(\{l\}). \quad (5.18)$$

However, we have no prior belief about $P(k)$ so we assume that this is constant, and $P(\{l\})$ does not depend on k .

Error in $1/k$

The calculation of the error bar for $1/k$ is closely related to the error bar in k , as presented in the text. Defining $\theta \equiv 1/k$ (referred to as $< \Delta Lk >$ in the text),

we have

$$\frac{d}{d\theta} = \frac{dk}{d\theta} \frac{d}{dk} = -k^{-2} \frac{d}{dk}. \quad (5.19)$$

Retaining primes for differentiating with respect to k , then, we find

$$\partial_\theta f = -k^2 f' \quad \partial_\theta^2 f = -k^2 (-k^2 f')' = k^4 f'' + k^2 f' (k^2)'. \quad (5.20)$$

If we wish to evaluate the second derivative at a value k_* (or, equivalently, θ_*) when $f' = 0$, we have:

$$\partial_\theta^2 f|_{\theta_*} = k_*^4 f''|_{k_*} \quad (5.21)$$

from which it follows that $\sigma(k)$ and $\sigma(\theta)$, the error bars in k and in $\theta = 1/k$, are related by k^2 :

$$\sigma(\theta)^{-2} = \partial_\theta^2 \ln P|_{k_*} = k_*^4 \partial_k^2 \ln P|_{k_*} = k_*^4 \sigma(k)^{-2} \quad \sigma(\theta) = \sigma(k)/k^2 \quad (5.22)$$

In the case $\Lambda \rightarrow \infty$, we have

$$\sigma(\theta) = \sigma(k)/k_*^2 = k_*/\sqrt{N}/k_*^2 = \theta_*/\sqrt{N}. \quad (5.23)$$

Gaussian approximation to the error

If we approximate $P(k|\{\ell\})$ as a Gaussian, we have:

$$\begin{aligned} P(k|\{\ell\}) &= \frac{e^{-\frac{1}{2}(k-k_*)^2/\sigma^2}}{\sqrt{2\pi\sigma^2}} \\ \ln P(k|\{\ell\}) &= -\frac{1}{2}(k-k_*)^2/\sigma^2 - \frac{1}{2}\ln(2\pi\sigma^2) \\ \partial_k^2 \ln P(k|\{\ell\}) &= -1/\sigma^2. \end{aligned} \quad (5.24)$$

Calculation of the Fuctional Form of the Globally Constrained Distribution

The measured distribution in Fig. 2a can be analytically calculated, and has a simple form in the statistical steady state (*i.e.* the limit of many simulation rounds, after which the initial condition $s = 0$ has effectively been forgotten). We demonstrate this here.

Consider moving from displacement x to displacement y constrained such that $x, y < z$, where z is the constraint (denoted $\Delta L k_{max}^0$ in the text). Displacements are incremented in steps s drawn from the true distribution $f(s) = e^{-s}$ (to

simplify the calculation, we will measure distances in this section in units such that $1/k = 1$). We then observe a distribution of steps t which obeys the following rules:

If the addition of the step size s to the initial position x does not exceed the constraint z , or

$$x + s < z \iff s < z - x, \quad (5.25)$$

then we register a step size $t = s$. The new position y then is given by $y = x + s$.

If the addition of the step size s to the initial position x exceeds the constraint z , or

$$x + s > z \iff s > z - x, \quad (5.26)$$

then we do not register a step size t . The new position y is set to 0.

These relationships will be expressed in terms of conditional distributions in the analysis below:

$$\begin{aligned} s < z - x &\Rightarrow p(y|x, s) = \delta(y - (x + s)), p(t|x, s) = \delta(t - s) \\ s > z - x &\Rightarrow p(y|x, s) = \delta(y), p(t|x, s) = \delta(t). \end{aligned} \quad (5.27)$$

Definitions

The conditional independence conditions give

$$P \equiv p(y, t, x, s) = p(y|x, s)p(t|x, s)p(x)p(s) \quad (5.28)$$

from which we define

$$\begin{aligned} f(s) &\equiv \int dy \, dt \, dx P \\ p^n(x) &\equiv \int dy \, dt \, ds P \\ p^{n+1}(y) &\equiv \int dt \, dx \, ds P \\ q(t) &\equiv \int dy \, dx \, ds P, \end{aligned} \quad (5.29)$$

where n is the number of steps observed. We are interested in calculating the function $q(t)$ in the limit $n \rightarrow \infty$. To do this we will first find $p^\infty(x)$ and then evaluate

$$q(t) = \int dx \, ds \, p(t|x, s)p^\infty(x)f(s). \quad (5.30)$$

Transition element $p(y|x)$

The conditional distribution $p(y|x)$ is the propagator of the distribution of possible lengths from one "roll" to the next. We can calculate this as

$$\begin{aligned}
 p(y|x) &= \int ds p(y|x, s) f(s) \\
 &= \int_0^{z-x} ds p(y|x, s) f(s) + \int_{z-x}^{\infty} ds p(y|x, s) f(s) \\
 &= \int_0^{z-x} ds \delta(y - (x + s)) f(s) + \delta(y) \int_{z-x}^{\infty} ds f(s) \\
 &= e^{-y+x} \Theta((z-x) - (y-x)) \Theta((y-x) - 0) + \delta(y) e^{-z+x} \\
 &= e^{-y+x} \Theta(z-y) \Theta(y-x) + \delta(y) e^{-z+x} \\
 &= e^{-y+x} \Theta(y-x) + \delta(y) e^{-z+x}
 \end{aligned} \tag{5.31}$$

(since $z \geq y$) from which

$$p^{n+1}(y) = \int_0^z dx p(y|x) p^n(x) = e^{-y} \int_0^y dx p^n(x) e^x + \delta(y) e^{-z} \int_0^z dx p^n(x) e^x. \tag{5.33}$$

The integration over a delta function in Eq. 5.32, and the resulting heaviside function, is a special case of the more general caveat

$$\begin{aligned}
 \int_a^b ds f(s) \delta(s - s_0) &= \begin{cases} f(s_0), & a < s_0 < b \\ 0, & \text{otherwise} \end{cases} \\
 &\equiv f(s_0) \Theta(s_0 - a) \Theta(b - s_0).
 \end{aligned} \tag{5.34}$$

Ansatz

Consider the ansatz

$$p^n(u) \equiv \pi_0^{(n)} \delta(u) + e^{-u} \sum_{j=0}^{\infty} c_j^{(n)} \frac{u^j}{j!}. \tag{5.35}$$

Clearly when $n = 0$ we have $p^{(0)}(u) = \delta(u)$, so

$$\begin{aligned}
 \pi_0^{(0)} &= 1 \\
 c_j^{(0)} &= 0 \quad \forall j
 \end{aligned} \tag{5.36}$$

and the initial distribution is within this functional form. We need only show that all later distributions have this functional form to have solved for the distribution

of lengths for all times.

$$\begin{aligned}
p^{(n+1)}(y) &= e^{-y} \int_0^y dx e^x p^{(n)}(x) + \delta(y) e^{-z} \int_0^z dx e^x p^{(n)}(x) \\
e^x p^{(n)}(x) &= \sum_{j=0}^{\infty} \frac{c_j^{(n)}}{j!} x^j + \pi_0^{(n)} \delta(x) \\
p^{(n+1)}(y) &= e^{-y} \sum_{j=0}^{\infty} \frac{c_j}{(j+1)!} x^{j+1} + e^{-y} \pi_0^n + \delta(y) \\
&\quad \left(e^{-z} \sum_{j=0}^{\infty} \frac{c_j}{j+1!} z^{j+1} + e^{-z} \pi_0^n \right). \tag{5.37}
\end{aligned}$$

So clearly the functional form is preserved compared to the ansatz (Eq.5.35).

$$\begin{aligned}
c_0^{n+1} &= \pi_0^n \\
c_{j+1}^{n+1} &= c_j^n \\
\pi_0^{n+1} &= e^{-z} \left(\pi_0^n + \sum_{k=1}^{\infty} \frac{c_{k-1}}{k!} z^k \right). \tag{5.38}
\end{aligned}$$

Statistical steady state

At steady state, we must have $\pi_0^{n+1} = \pi_0^n \equiv \pi_0$, $c_j^{n+1} = c_j^n \equiv c_j$. Consequently,

$$\begin{aligned}
c_0^{n+1} &= \pi_0^n \Rightarrow c_0 = \pi_0 \\
c_{j+1}^{n+1} &= c_j^n \Rightarrow c_j = c_0 \\
\pi_0^{n+1} &= e^{-z} \left(\pi_0^n + \sum_{k=1}^{\infty} \frac{c_{k-1}}{k!} z^k \right) \Rightarrow c_0 = e^{-z} (c_0 + c_0 e^z - c_0) \tag{5.39}
\end{aligned}$$

the last of which yields $1 = 1$. Normalization sets c_0 :

$$\begin{aligned}
1 &= \int dx c_0 (\delta(x) + 1) \Rightarrow c_0 = 1/(1+z) \\
p^\infty(x) &= \frac{\delta(x) + 1}{1+z}. \tag{5.40}
\end{aligned}$$

Calculation of $q(t)$

We can now finally calculate the distribution of observed steps $q(t)$. Note first the simplifications (recalling the procedure for definite integration over delta functions from Eqn. ??):

$$\begin{aligned}
\int_0^{z-x} ds f(s) p(t|x, s) &= f(t) \Theta((z-x) - t) \Theta(t - 0) \\
&= f(t) \Theta((z-t) - x)
\end{aligned}$$

$$\begin{aligned}\int_{z-x}^{\infty} ds f(s) p(t|x, s) &= \delta(t) \int_{z-x}^{\infty} ds f(s) \\ &= \delta(t) e^{-(z-x)}\end{aligned}\tag{5.41}$$

from which

$$\begin{aligned}q(t) &= \int_0^z dx ds p(t|x, s) p^{\infty}(x) f(s) \\ &= \int_0^z dx p^{\infty}(x) \left(\int_0^{z-x} ds f(s) p(t|x, s) + \int_{z-x}^{\infty} ds f(s) p(t|x, s) \right) \\ &= \int_0^z dx \left(\frac{\delta(x) + 1}{1+z} \right) (f(t) \Theta((z-t) - x) + \delta(t) e^{-z+x}) \\ &= f(t) \frac{1}{1+z} \int_0^{z-t} dx (\delta(x) + 1) + \delta(t) e^{-z} \frac{1}{1+z} \int_0^z dx e^x (1 + \delta(x)) \\ &= f(t) \frac{1+z-t}{1+z} + \delta(t) e^{-z} \frac{1}{1+z} (e^z - 1 + 1) \\ &= \frac{(1+z-t)}{1+z} e^{-t} + \delta(t) \frac{1}{1+z}\end{aligned}\tag{5.42}$$

which is properly normalized, since

$$\begin{aligned}\int_0^z dt (1+z-t) e^{-t} + 1 &= (1+z)(1 - e^{-z}) - (-ze^{-z} - e^{-z} + 1) + 1 \\ &= (1+z)(-e^{-z} + 1) + ze^{-z} + e^{-z} \\ &= 1+z\end{aligned}\tag{5.43}$$

Result: comparison of cumulative probability distributions

$$\begin{aligned}P(t' < t | t' > 0) &= \frac{P(t > t' > 0)}{P(t > 0)} \\ &= \frac{\int_{0+}^t dt' q(t')}{\int_{0+}^z dt' q(t')} \\ &= \left(\frac{1+z}{z} \right) \frac{(z + (t-z) e^{-rt})}{1+z} \\ &= 1 - e^{-rt} (1 - (t/z))\end{aligned}\tag{5.44}$$

In Fig 5.5, we plot the experimental result (from the simulations described in the text) against the theoretical prediction, for $z = 130$, $r = 60$. Note that for $t \ll z$, $P(t' > t) \approx 1 - e^{-rt}$, as expected.

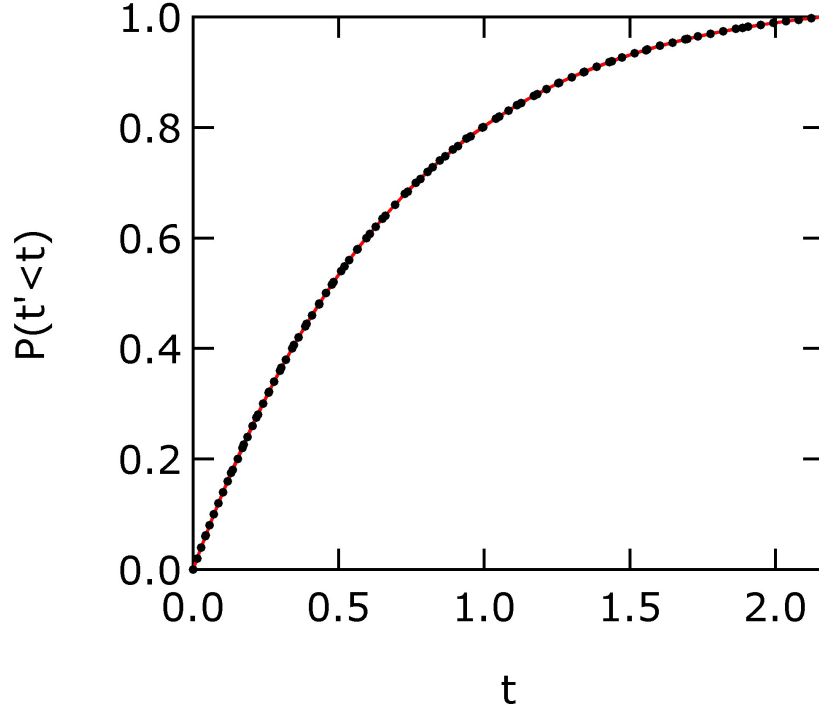


Figure 5.5: Comparison between the simulation result and the theoretical prediction. Black solid circles represent simulation data points, as described in the text. Red line represents the theoretical prediction, Eq. 5.44

Global constraints do not permit estimation of the distribution parameter by simply counting the fraction of discarded events

Consider the probability distribution for possible step sizes s :

$$p(s) = ke^{-ks} \quad (5.45)$$

A step of size s is defined by a transition between two levels, from level x to level y . From the probability distribution we can derive P_i^+ , the probability that the i th step will be terminated by the global constraint:

$$P_i^+ \equiv \int_{s=z-x_i}^{\infty} p(s)ds = e^{-k(z-x_i)}, \quad (5.46)$$

where z is the highest possible value for y , which is given by the constraint. The likelihood Λ of the observed data D , summarized in terms of $N_+/N_+ + N_- \equiv N_+/N$, the fraction of the reaction steps terminated by the global constraint, is (using i_+ and i_- to index steps terminated or not, respectively, by the global

constraint)

$$\begin{aligned}\Lambda &= P(D) = \prod_{i_+} P_{i_+}^+ \prod_{i_-} P_{i_-}^- \\ \ln \Lambda &= \sum_{i_+} -k(z - x_{i_+}) + \sum_{i_-} \ln(1 - e^{-k(z-x_i)}) \\ &= -N_+ k(z - \langle x_+ \rangle) + N_- \langle \ln(1 - e^{-k(z-x_-)}) \rangle.\end{aligned}\quad (5.47)$$

Eq. 5.47 makes it clear that we cannot estimate k by only keeping track of N_+ and N_- ; rather, we must keep track of the initial location of all the steps that did not terminate as well as of $\langle x_+ \rangle$: the average location of the reactions that did terminate.

References

- [1] Strick, T. R., Allemand, J. F., Bensimon, D., Bensimon, A., and Croquette, V. *Science* **271**(5257), 1835–7 (1996).
 - [2] Visscher, K., Schnitzer, M. J., and Block, S. M. *Nature* **400**(6740), 184–9 (1999).
 - [3] Yildiz, A., Forkey, J. N., McKinney, S. A., Ha, T., Goldman, Y. E., and Selvin, P. R. *Science* **300**(5628), 2061–5 (2003).
 - [4] Weiss, S. *Science* **283**(5408), 1676–83 (1999).
 - [5] Svoboda, K., Mitra, P. P., and Block, S. M. *Proc Natl Acad Sci U S A* **91**(25), 11782–6 (1994).
 - [6] Hald, A. *Statistical Science* **14**(2), 214–222 (1999).
 - [7] Rice, J. A. *Mathematical Statistics and Data Analysis*. Wadsworth & Brooks/Cole Advanced Books & Software, Pacific Grove, California, (1988).
 - [8] Saleh, O. A., Peral, C., Barre, F. X., and Allemand, J. F. *Embo J* **23**(12), 2430–9 (2004).
 - [9] Koster, D. A., Croquette, V., Dekker, C., Shuman, S., and Dekker, N. H. *Nature* **434**(7033), 671–4 (2005).
 - [10] Stivers, J. T., Shuman, S., and Mildvan, A. S. *Biochemistry* **33**(1), 327–39 (1994).
 - [11] Wang, J. C. *Annu Rev Biochem* **65**, 635–92 (1996).
 - [12] Strick, T. R., Croquette, V., and Bensimon, D. *Nature* **404**(6780), 901–4 (2000).
-

-
- [13] Dekker, N. H., Abels, J. A., Veenhuizen, P. T., Bruinink, M. M., and Dekker, C. *Nucleic Acids Res* **32**(18), e140 (2004).
 - [14] Bamford, D. H. *EMBO Rep* **3**(4), 317–8 (2002).
 - [15] Bamford, D. *personal communication* .
 - [16] Ha, T., Ting, A. Y., Liang, J., Caldwell, W. B., Deniz, A. A., Chemla, D. S., Schultz, P. G., and Weiss, S. *Proc Natl Acad Sci U S A* **96**(3), 893–8 (1999).
 - [17] Blanchard, S. C. e. a. *Proc Natl Acad Sci U S A* **101**, 12893–8 (2004).
 - [18] Zhuang, X., Kim, H., Pereira, M. J., Babcock, H. P., Walter, N. G., and Chu, S. *Science* **296**(5572), 1473–6 (2002).
 - [19] Zhuang, X. *Annu Rev Biophys Biomol Struct* **34**, 399–414 (2005).
-

Chapter 6

Antitumor drugs impede DNA uncoiling by Topoisomerase I

Increasing the ability of chemotherapeutic drugs to kill cancer cells is often hampered by a limited understanding of their mechanism of action. Camptothecins, such as topotecan, induce cell death by poisoning DNA topoisomerase I, an enzyme capable of removing DNA supercoils. Topotecan is thought to stabilize a covalent topoisomerase-DNA complex, rendering it an obstacle to DNA replication forks. Here we employ single-molecule nanomanipulation to monitor the dynamics of topoisomerase in the presence of topotecan. This allowed us to detect the binding and unbinding of an individual topotecan molecule in real time and quantify the drug-induced trapping of topoisomerase on DNA. Unexpectedly, our findings also show that topotecan significantly hinders topoisomerase-mediated DNA uncoiling, with a more pronounced effect on the removal of positive (overwound) versus negative supercoils. *In vivo* experiments in yeast verified the resulting prediction that positive supercoils would accumulate during transcription and replication as a consequence of camptothecin poisoning of topoisomerase I. Positive supercoils, however, were not induced by drug treatment of cells expressing a catalytically active, camptothecin-resistant topoisomerase I mutant. This unique combination of single-molecule and *in vivo* data suggests a novel cytotoxic mechanism for camptothecins, in which the accumulation of positive supercoils ahead of the replication machinery induces potentially lethal DNA lesions.

This chapter will be published in shortened format by *Nature*: Daniel A. Koster, Komaraiah Palle, Elisa S. M. Bot, Mary-Ann Bjornsti and Nynke H. Dekker

6.1 Introduction

The helical structure of double-stranded DNA allows for the faithful deciphering and transmission of genetic information, where each DNA strand serves as a template for the synthesis of a complementary polynucleotide chain [1]. The intertwining of the DNA strands further ensures DNA integrity by physically linking the individual chains in a structure stabilized by hydrogen bonding and base stacking. However, these same features pose a number of topological problems during cell cycle progression and cell division. For example, the progressive unwinding of the DNA template during DNA replication and the segregation of multiply intertwined daughter DNA molecules require changes in the linkage of DNA strands and helices. Similarly, RNA transcription can produce local unwinding of the DNA helix behind the transcription complex and local overwinding of the duplex ahead [2, 3]. DNA topoisomerases resolve these problems via a mechanism of transient DNA strand breakage and religation [4–6]. However, the stabilization of topoisomerase-DNA complexes, as a consequence of drug action, may also induce potentially lethal DNA damage. Indeed, topoisomerases constitute the cellular targets of anticancer and antibacterial drugs. The camptothecin class of chemotherapeutics targets eukaryotic DNA topoisomerase IB (Top1B) [7–10]. Several camptothecin analogues, including the pro-drug irinotecan and the water-soluble derivative, topotecan, have significant activity against adult and pediatric solid tumors and have FDA approval for the treatment of ovarian and small cell lung cancer [11–13]. A detailed understanding of topoisomerase-drug interactions is critical for optimal clinical development of these chemotherapeutics. However, the dynamic interactions underlying this poisoning and their biological ramifications remain largely unknown.

Mechanistically, Top1B removes DNA supercoils in the following manner: Top1B clamps around the DNA [14, 15], after which the active-site tyrosine serves as a nucleophile to cleave a single strand of the DNA duplex, thus forming a DNA-(3'-phosphotyrosyl)-enzyme 'covalent complex' while transiently creating a free 5'-OH DNA end. Torsional energy present in the DNA can then drive uncoiling about the intact DNA strand. The reverse ligation reaction, which necessarily terminates uncoiling and is torque-dependent, occurs after a random number of supercoils are removed and leads to restoration of the DNA backbone [16–19]. Topotecan intercalates into the nick that is generated by Top1B and prevents religation [20–22]. This reversible binding of the drug stabilizes the covalent complex, thereby trapping Top1B on the DNA [20–23]. During S-phase, these reversible ternary topotecan-Top1B-DNA complexes are converted into cytotoxic DNA lesions that result in cell death [24–26]. As processive DNA replication is

required for topotecan-induced cell lethality [24, 26], it has been proposed that these lesions occur as a result of the covalent complex colliding with an advancing replication fork [27].

Here we describe single-molecule experiments (for a review, see Bustamante *et al.* [28]) that test key predictions of the proposed mechanism of poisoning of Top1B. Simultaneously, they reveal novel dynamical interactions between topotecan and Top1B. The consequences of these interactions are observed in the context of a living cell, suggesting an alternative mechanism for cell death and demonstrating the power of combining single-molecule and *in vivo* experiments.

6.2 Drug binding hinders DNA uncoiling

To examine the consequences of camptothecin binding on the functioning of the Top1B swivel, we explored its dynamics in the presence of topotecan in the context of a single molecule and in real time. Our experimental configuration, the magnetic tweezers, is described elsewhere [29], and in short entails anchoring a single double-stranded (ds)DNA molecule between a coated glass surface and a coated paramagnetic bead [30]. An upward magnetic stretching force F on the DNA is generated by means of a pair of magnets positioned above the sample. By changing the distance from the magnets to the sample, F can be varied, and by rotating the magnets about their axis, supercoils can be mechanically injected. By continuously monitoring the height of the bead above the surface, *i.e.* the DNA extension, the degree of supercoiling σ is determined in real time.

In a typical experiment the DNA molecule is mechanically coiled, introducing multiple plectonemes that reduce the DNA extension (Fig. 8.3a). We then typically observe a plateau (*e.g.* from app. 25 to 100 s in Fig. 8.3a), which is indicative of a DNA molecule in a supercoiled and unnicked state. In the presence of human Top1B but in the absence of topotecan, the DNA is subsequently (at 100 s, shown in red) rapidly uncoiled by the enzyme. Removal of the plectonemes can occur either in a single enzymatic event or in multiple steps (supplementary materials I), as was previously reported for vaccinia virus Top1B [16]. Surprisingly, however, upon addition of topotecan, a dramatically different signature is observed (Fig. 8.3b). First, drug-mediated uncoiling occurs slowly (red data points) compared to uncoiling in the absence of drug. Slow uncoiling is observed immediately following mechanical coiling and proceeds in a fashion that appears to be continuous and which we fit by a linear relation (blue line). Second, the slow uncoiling continues for long periods of time. Third, during the slow uncoiling, we observe no apparent plateaus, suggesting that no religation takes place

(supplementary materials II). The most likely scenario to explain our observations at the molecular level is provided by X-ray crystallography studies on the covalent complex with a bound topotecan [20, 22]. The flat ring-like structure of topotecan (magenta molecule in insert of Fig. 8.3b) was shown to intercalate into a Top1B-generated nick, and is stabilized by both specific human Top1B-topotecan contacts and base stacking interactions with adjacent bases (inset, Fig. 8.3b). In this configuration, topotecan locally deforms the DNA duplex and displaces the 5'-OH, decreasing the probability for religation. Indeed, this is further consistent with the reversible stabilization of covalent Top1B-DNA complexes induced by camptothecins in biochemical assays [23, 31]. It was also speculated that as a result of topotecan's contacts with Top1B and DNA, rotational motion ought to be significantly hindered [14, 22, 32], a speculation that is fully compatible with our observations.

To quantify the reduction in uncoiling rate caused by the topotecan binding, we examine the complete distribution of DNA extension velocities during enzymatic uncoiling (Fig. 8.3c). In the presence of 5 μM TPT, we identify two populations: one topotecan-mediated and one non-topotecan-mediated. The fraction of events that are topotecan-mediated increases with topotecan concentration (data not shown), but even at the highest workable topotecan concentrations (approaching the limits of topotecan solubility), a significant fraction of the events remains non-topotecan mediated. The non-topotecan-mediated population is unambiguously identified by experiments in the absence of topotecan, shown in Fig. 8.3c (inset) and takes place on average at 4.1 $\mu\text{m s}^{-1}$. Topotecan-mediated events, however, on average take place at 0.2 $\mu\text{m s}^{-1}$, corresponding to a plectoneme relaxation rate of app. 3 Hz (supplementary materials III) which is independent of topotecan concentration (suggesting its attribution to the binding of an individual topotecan molecule, supplementary materials II). As we increase the stretching force, thereby increasing the torque that drives the DNA uncoiling, the angular velocity of uncoiling also increases, as expected (data not shown). We conclude that topotecan-mediated uncoiling by Top1B takes place at roughly a factor of 20 slower than uncoiling by Top1B alone. We note that the broadness in the topotecan-mediated uncoiling velocities cannot be attributed to measurement errors, but is a true reflection of the fact that uncoiling events that take place at different times in the experiment are not identical. As Top1B is only mildly sequence specific and bulk studies have shown that camptothecin stabilizes distinct complexes with different efficiencies [31, 33], we surmise that uncoiling events at different sequences are characterized by different uncoiling velocities.

Several control experiments were performed to exclude the possibility that topotecan binding to DNA or non-specific Top1B-topotecan interactions could

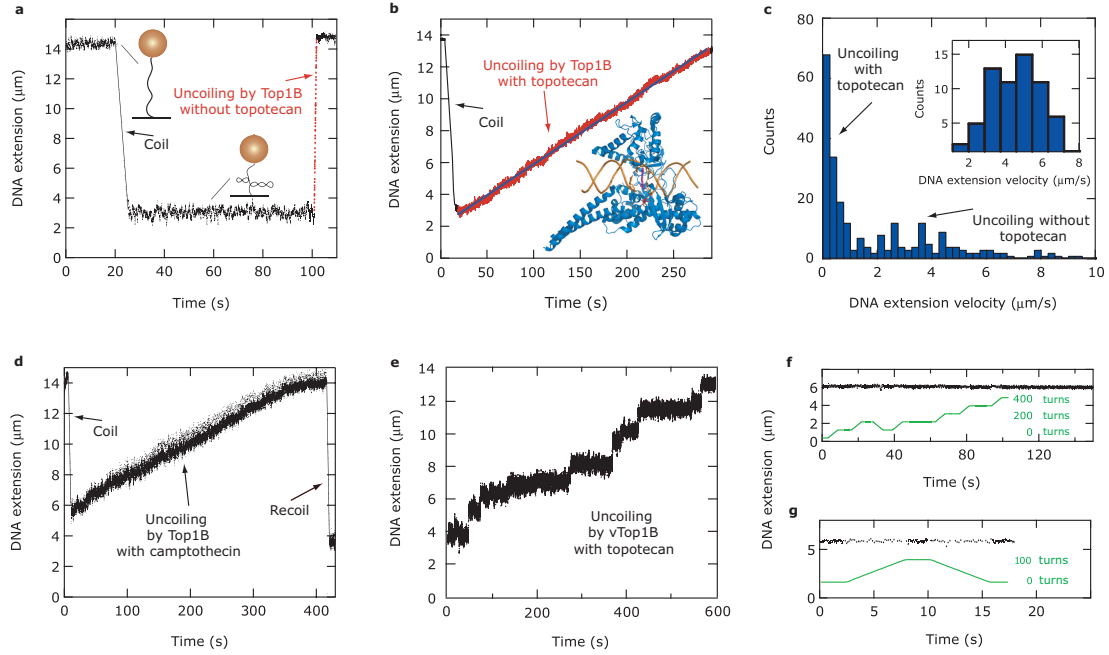


Figure 6.1: Slow DNA uncoiling is only observed with a functional human Top1B. (a) In the absence of topotecan, uncoiling (red dots, at 100 s) proceeds rapidly. (b) In the presence of topotecan, slow and continuous uncoiling is observed, apparently without religation. The uncoiling rate is constant, as expected for the constant torque regime in which measurements were performed (supplementary materials II). (c) The distribution of uncoiling velocities in the presence of 5 μM topotecan ($n=275$) shows contributions from topotecan-mediated slow uncoiling and topotecan-independent uncoiling (inset shown measurement in the absence of topotecan ($n=64$)). Topotecan-mediated uncoiling proceeds roughly 20 times more slowly than topotecan-independent uncoiling. (d) human Top1B in the presence of camptothecin yields slow uncoiling very similar to that observed with its clinical analogue topotecan. (e) in the presence of topotecan, slow continuous supercoil removal is not observed with the camptothecin-resistant vaccinia Top1B, as expected. (f) neither positive supercoils nor plectonemes can be induced by mechanical coiling of a singly nicked DNA molecule in the presence of topotecan only, or (g) in the presence of topotecan and catalytically inactive human Top1Y723F.

give rise to the slow uncoiling described above. First, we verified that the presence of topotecan alone does not alter the mechanical properties of DNA, *i.e.* the force-extension and supercoiling behaviour (supplementary materials IV). Second, we incubated human Top1B in the presence of camptothecin, the parent compound of topotecan that induces a similar mechanism of Top1B poisoning [34]. Fig. 8.3d shows a similar signature as that seen with topotecan, indicating that the effects observed were not uniquely limited to the water-soluble topotecan analog. Third, we studied uncoiling by vaccinia virus topoisomerase I, known

to be resistant to camptothecins [35], in the presence of high concentrations of topotecan. Here, we reproduced the characteristic step-wise uncoiling (Fig. 8.3e) that was observed before in the absence of topotecan [16] and that does not feature continuous slow uncoiling at concentrations similar to and higher than those used in the experiments with human Top1B. Fourth, we generated a single nick in the DNA by incubating the DNA with the sequence-specific nicking enzyme N.BbvCIA. By subsequently rotating our magnets, we attempted to introduce plectonemes in the presence of topotecan alone (Fig 8.3f), and in the presence of topotecan and the catalytically inactive human Top1Y723F (Fig 8.3g). In both experiments, no plectonemes could be introduced, in contrast to the results obtained in the presence of catalytically active human Top1B and topotecan (Fig 8.3b). These experiments show that a catalytically active human Top1B is required to yield slow uncoiling, consistent with the *in vivo* activity and co-crystal structures of camptothecin analogues [20, 22, 36]. We thus conclude that slow uncoiling represents a clear signature for a topotecan molecule bound to the Top1B-DNA complex.

6.3 Drug binding traps Top1B on the DNA

A timescale highly relevant for any consideration of the "collision" model is the lifetime of the covalent complex with a bound topotecan. Indeed, the likelihood of a replication fork colliding with the covalent complex is expected to increase with this lifetime. Therefore, we ask over what timescale topotecan remains bound and thus, over what timescale Top1B remains covalently linked to the DNA. Fig. 8.4a shows a time trace where the times at which the topotecan enters (t_{start}) and exits (t_{end}) the covalent complex are indicated. The precise assignment of t_{start} is observed by zooming in on the entry event (inset Fig. 8.4a), where a plateau that is interrupted by a fast rise in DNA extension as a result of DNA cleavage by Top1B is followed by the appearance of slow uncoiling at t_{start} , which we attribute as above to a topotecan molecule entering a Top1B-generated nick. Figs. 8.4b-e show a collection of traces meant to provide insight into how t_{start} and t_{end} are chosen. The majority of events observed (roughly 70%) are similar to the ones shown in Fig. 8.4b and Fig 8.4c. In Fig. 8.4b, around 340 s, the magnets are rotated for a substantial number of turns on two occasions (the interruption of the DNA extension signal at 334 and 338 s signifies rotation), but no plectonemes are introduced. This situation corresponds to a nicked DNA molecule as a result of the formation of the covalent complex. At t_{start} , a topotecan molecule appears present in the covalent complex, as slow uncoiling is observed. Fig 8.4c shows the

abrupt transition from slow supercoil removal to fast supercoil removal, prompting the designation of t_{end} . The bound time Δt is defined as $t_{end} - t_{start}$ and reflects the time supercoils are enzymatically removed at a constant slow rate (shaded box in Fig. 8.4a). Note that three recoiling events take place during Δt . Fig. 8.4f shows the distribution of Δt , which has a mean of about 121 ± 11 s ($n=146$) that is globally unchanged both within the practical force range of the technique (Fig. 8.4f, inset) and upon changing topotecan concentration by a factor of 10 (supplementary materials II). The latter indicates that the lifetime is not likely to be overestimated by topotecan unbinding/rebinding during the mechanical recoiling nor by the presence of occasional very short-lived plateaus (supplementary materials II). Although the measurement of $\langle \Delta t \rangle$ was unaltered upon changing the Top1B concentration by a factor of 2, Δt may nonetheless represent a lower limit for the time topotecan remains bound, since t_{end} can be caused by topotecan exiting from the covalent complex, or, in principle, by the start of topotecan-independent uncoiling by a second Top1B at another cleavage site on the DNA. We conclude that in the presence of topotecan, Top1B remains trapped on the DNA for at least 121 s.

The time that Top1B remains covalently trapped on the DNA in the presence of topotecan ought to be compared with this time in the absence of the drug, which is by definition the religation time. Fig. 8.4g depicts the strategy for measuring this quantity (supplementary materials V). When Top1B nicks the DNA (at t_{nick}) and the bead consequently travels upwards due to DNA uncoiling, it reaches above a threshold (at t_{rotate}), beyond which we prompt the magnets to spin continuously. In the absence of topotecan, DNA uncoils faster than the magnets can recoil. Consequently, plectonemes are removed and the bead keeps traveling upwards until Top1B religates the DNA and plectonemes can be once again introduced (at $t_{religated}$). A histogram of the religation time ($t_{religated} - t_{nick}$) is shown in Fig. 8.4h. It is characterized by a most probable time of app. 0.3 s, with a tail extending beyond that, as shown in Fig 8.4h, inset. A comparison of Δt with the religation time indicates that Δt is larger by roughly a factor of 100, providing the first quantitative support for a critical aspect of the "collision" model, namely the topotecan-induced increase in the lifetime of the covalent complex.

Our single-molecule techniques also permitted us to focus on the enzymatic uncoiling rate of positive *vs.* negative supercoils. Although in the absence of topotecan, no difference was detected (data not shown), we observe a clear and unexpected difference in the supercoil removal rate in the presence of topotecan. Fig. 8.5a shows that the uncoiling of positive supercoils (pink regions) is significantly slower than the uncoiling of negative rotations (blue region). This

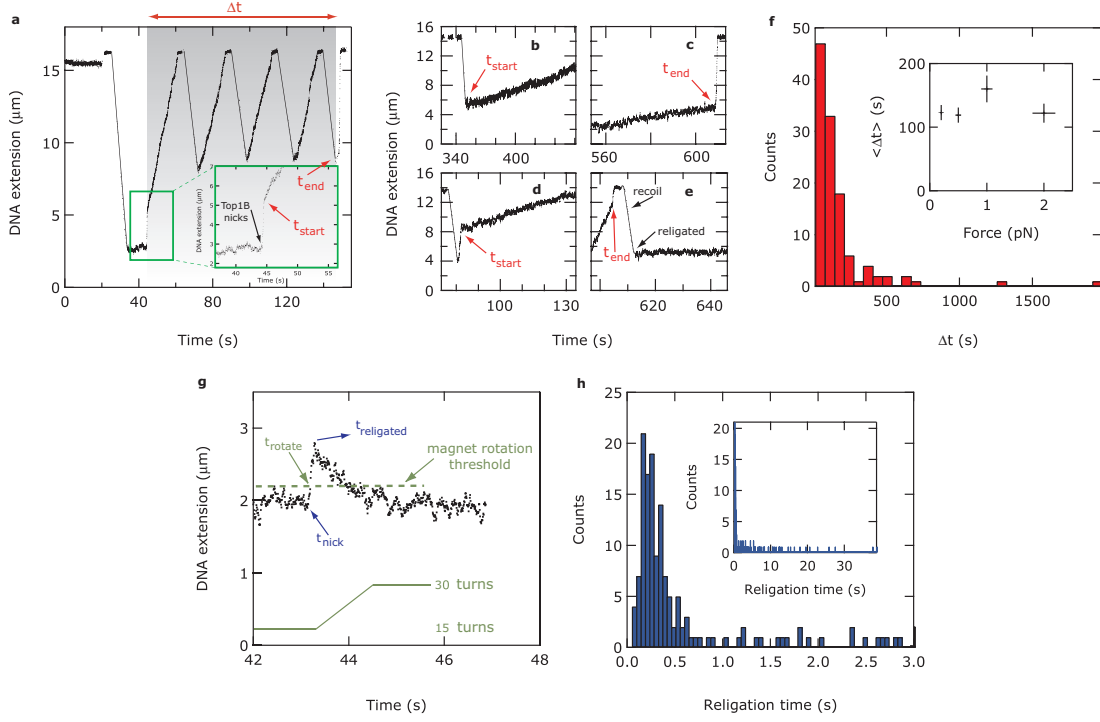


Figure 6.2: Slow supercoil removal constitutes a signature for a bound topotecan. (a) the duration Δt for a bound topotecan (grey area) is defined as the period of time during which slow and constant uncoiling is observed, *i.e.* between t_{start} and t_{end} . (b) a typical illustration (70% of events, $n=660$) of t_{start} and (c) t_{end} events. (d) t_{start} event, in which nicking appears directly prior to the onset of slow uncoiling. (e), an event suggestive of a topotecan molecule exiting at t_{end} , after which religation has taken place (plateau onset at 610 s). (f) histogram of Δt for $F=0.5\text{pN}$, with mean time of 121 ± 11 s ($n=146$), which is not significantly force dependent (inset). (g) strategy for measuring Top1B religation time in the absence of topotecan, as described in the text and in supplementary materials V. (h) distribution of Top1B religation times, with a most probably religation time of 0.3 s ($n=132$) and a tail generating an average time off 4 s., still well below the value for Δt (121 s, see Fig. 8.4f).

asymmetry in uncoiling rates is quantified in Fig. 8.5b as the differential in rates, denoted $\Delta\omega$ and defined as $|\omega_+| - |\omega_-|$, where $|\omega_+|$ is the uncoiling rate of positive supercoils and $|\omega_-|$ is the uncoiling rate of negative supercoils.

When $|\omega_-|$ is larger than $|\omega_+|$, $\Delta\omega$ is negative and its mean value is found to be -8 ± 5 Hz ($n=17$). The relative difference, defined by $\Delta\omega/|\omega_+|$, is -2.4 ± 2.2 . The microscopic interactions responsible for the asymmetrical rate of DNA uncoiling in the presence of topotecan are difficult to investigate and have not been foreseen by crystallography experiments, as these experiments provide a relatively static picture of a crystalline Top1B. However, the origins of the asymmetry may be

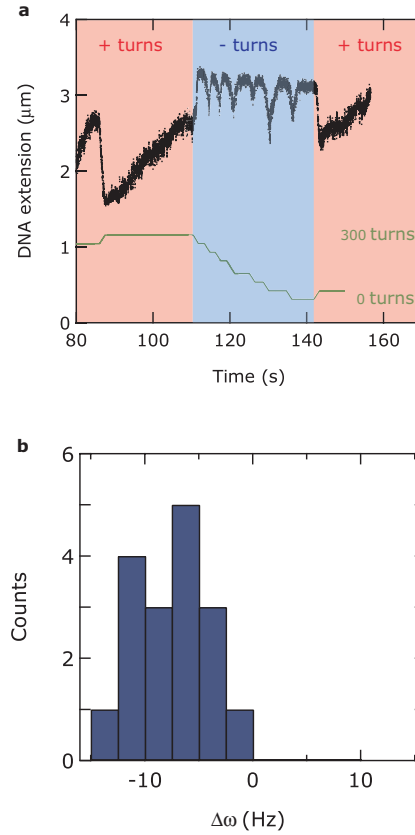


Figure 6.3: Asymmetry in topotecan-mediated DNA supercoil removal rate. (a) Traces indicating the removal rate of positive supercoils (pink areas) and the removal rate of negative supercoils (blue areas). It is observed that the removal of positive supercoils proceeds more slowly. The slow removal of positive supercoils is verified following the removal of negative supercoils in order to ascertain that the topotecan molecule remained bound in the interim. The green line represents the rotational position of the magnets (in turns), specifying the absolute number of turns that are mechanically injected into the DNA molecule. (b) a histogram of the difference in uncoiling velocity, $\Delta\omega$, between positive and negative supercoils, measured in the presence of topotecan (see main text).

described by molecular dynamics simulations that predict distinct mechanisms for positive and negative supercoil relaxation in the absence of topotecan [37].

6.4 Implications for *in vivo* transcription and replication

In order to investigate the biological ramifications of the asymmetrical hindrance of DNA uncoiling induced by topotecan, we asked if this bias was evident in drug treated yeast cells. The single-molecule observation that positive supercoils are removed more slowly than negative supercoils in the presence of topotecan led us to posit that positive supercoils would accumulate during cellular processes that induce DNA supercoiling, such as transcription and replication [2, 3].

Using the yeast *Saccharomyces cerevisiae* as a model system [2], it is possible to selectively assess the activity of a unique topoisomerase and detect any asymmetry in the topoisomerase-catalyzed relaxation of supercoils that accumulate during transcription. We note that the process of transcription produces local domains of DNA supercoiling, as the tracking of the transcription machinery between the two strands of duplex DNA will induce the accumulation of positive supercoils ahead of the complex, while an equal number of compensatory negative supercoils are left in its wake [2] (Fig. 8.6a, center). Consequently, in plasmid DNA the transcription of divergent genes will partition the DNA into two domains containing an equal number of positive and negative supercoils, respectively. In the appropriate genetic background, yeast cells will lack Top1B and Top2 activity, due to deletion of the chromosomal *TOP1* gene sequences (*top1Δ*) and the expression of a thermosensitive Top2ts mutant enzyme, which is inactive at high temperature (materials and methods). Under these conditions, transcription will fail to alter the topology of the plasmid DNA, because the compensatory positive and negative supercoils will annihilate each other as the plasmid DNA is purified. However, if these cells express bacterial type IA topoisomerase, which selectively relaxes negative supercoils, this enzymes asymmetric relaxation of transcription-generated supercoils can be detected by the accumulation of positively supercoiled DNA topoisomers [2].

Here, we adopt a similar experimental strategy by treating G1-phase arrested *top1Δ* yeast cells (Fig. 8.6b) that express low levels of plasmid-encoded human Top1B with (Fig. 8.6a, left) or without camptothecin (Fig. 8.6a, right). Also, the cells either express the thermosensitive Top2ts mutant enzyme, or wild-type Top2. Arresting the cells in G1-phase restricted the analysis to transcription-induced alterations in DNA supercoiling, while shifting the cells to the non-permissive temperature prior to drug treatment largely abolished any contribution of Top2. To avoid complications of drug transporter efflux of topotecan from yeast cells [7], camptothecin was used in these studies. Fig. 8.6c shows the distribution of

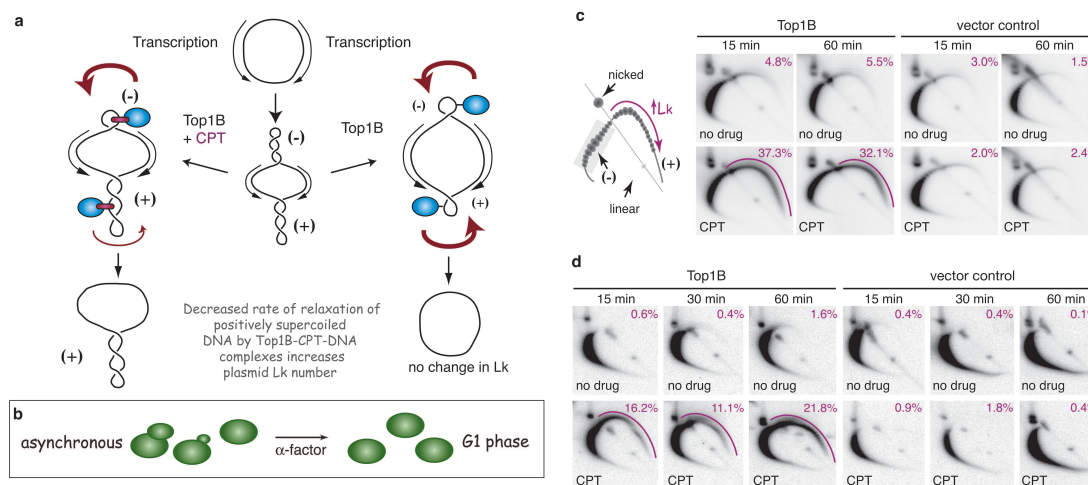


Figure 6.4: Camptothecin poisoning of Top1B induces the accumulation of positively supercoiled DNA in G1-phase yeast cells. (a) during transcription, positive supercoils are generated ahead of advancing transcription bubbles, while compensatory negative supercoils are generated in their wake. Diverging transcription units bisect a plasmid into twin domains of positive and negative supercoils. If, in the presence of drug (left-hand side), positive supercoils are removed more slowly than negative supercoils, it is expected that the plasmid should display an excess of positive supercoils. In the absence of drug (righthand side), no such asymmetry between the removal rates of positive supercoils and negative supercoils is observed and, consequently, no net change in linking number (Lk) is expected. (b) to eliminate the potential contribution of positive supercoils induced by replication forks, asynchronously growing haploid yeast cells are arrested in G1-phase of the cell cycle with the mating pheromone α -factor. (c) 2D agarose gel analysis of the Lk distribution of $2\mu m$ plasmid topoisomers isolated from G1-phase $top1\Delta$, $top2^{ts}$ cells, expressing plasmid encoded human Top1B (Top1B) or vector control, shifted to the nonpermissive temperature ($36^\circ C$), and treated with camptothecin (CPT) or no drug for 15 or 60 minutes. The activity of the thermosensitive Top2 mutant enzyme is largely abolished at $36^\circ C$. A description of plasmid DNA topoisomer resolution in 2D agarose gels is provided in supplementary material VI. To determine the percentage of topoisomers exhibiting an increase in linking number (indicated by purple numbers) a PhosphorImager was used to quantify the signal intensity of the portion of the arc to the right of the diagonal linear DNA (purple arc) relative to the amount of label detected across the entire arc. (d) $2\mu m$ topoisomer distribution in G1-phase $top1\Delta$, wild-type $TOP2$ yeast cells expressing plasmid encoded Top1B or vector control. As in (c), the accumulation of positive supercoils is only observed in Top1B expressing cells treated with camptothecin.

2 μ m plasmid DNA topoisomers isolated from G1-phase cells and resolved by 2D gel electrophoresis (supplementary materials VI). In the absence of drug (upper row in Fig. 8.6c), one observes an expected bias in the plasmid topoisomer distribution towards negative supercoils. This deficit in linking number is due to DNA wrapping around the histone core in nucleosomes. However, camptothecin treatment of Top1B-expressing cells induces a remarkable skewing of the plasmid topoisomer distribution, where one-third of the topoisomers exhibit an increased linking number or positive supercoils (purple arcs), which persists throughout G1 phase. By contrast, no alteration in linking number was induced by camptothecin treatment of cells that expressed only a vector control and were lacking Top1B; the topoisomer distribution appears as in the case of no treatment. Yeast cells expressing the catalytically inactive human Top1Y723F mutant also exhibit no change in plasmid DNA topology when treated with camptothecin (supplementary materials VII). Thus, the shift in plasmid DNA topology towards positive supercoils requires the expression of a catalytically active Top1B enzyme in the presence of camptothecin. In the absence of camptothecin, expression of Top1B does not skew the topoisomer distribution, consistent with our single-molecule experiments indicating no difference in the rate of enzyme-catalyzed uncoiling of positive vs. negative supercoils. Recent biochemical studies suggest higher levels of Top1B-DNA complexes are formed with positively *vs.* negatively supercoiled DNA and that this bias is increased in the presence of camptothecin [38]. *In vivo*, such preferential scission of positively supercoiled DNA by Top1B might be expected to skew the distribution towards more negative supercoils, due to the preferential relaxation of the local domain of positive supercoils diagrammed in Fig. 8.6a. However, such effects were not observed in our genetic background.

The same experiment was also carried out in isogenic yeast strains expressing wild-type Top2 (Fig. 8.6d). Interestingly, a similar skewing of the topoisomer distribution towards positive supercoils was also observed, albeit at a lower level. Although Top2 is essential for cell viability [39], these observations indicate that the positive supercoils induced by camptothecin poisoning of Top1B were not effectively resolved by Top2. Combined with our single-molecule observations, these *in vivo* findings support a model in which the positive supercoils that accumulate ahead of the transcription bubble are removed less effectively by Top1B in the presence of drug in comparison to the compensatory negative supercoils, which leads to a persistent overwinding of DNA in the absence of DNA replication in G1 phase.

Camptothecin toxicity is linked to on-going DNA replication in S-phase [24, 26], thus it is of particular relevance for the anticancer activity of these drugs to monitor positive supercoil accumulation during this phase of the cell

cycle. The pattern of gene transcription differs as cells transit from G1-phase to S-phase [40], although both processes occur during S-phase transit. However, unlike during transcription, twin supercoil domains are not generated during replication. Rather, positive supercoils accumulate ahead of the replication fork and/or precatenates are generated behind the fork. In a closed circular DNA template such as the $2\mu\text{m}$ plasmid, Top1B will act as a swivelase ahead of the fork to remove positive supercoils (Fig. 8.7a), while Top2 acts to decatenate the replicated DNA duplexes, allowing for the separation of the newly synthesized daughter plasmids. To study the effect of camptothecin treatment on Top1B activity during S-phase, cells arrested in G1-phase were allowed to synchronously enter S-phase by removing α -factor from the culture medium, after which the experiment was performed as described above. To ensure that we were assessing the effects of camptothecin on living cells, we first verified that drug treatment was insufficient to induce cell death, yet sufficed to induce DNA damage, which results in a slow progression through S-phase due to activation of the S-phase checkpoint [41]. Relative to untreated Top1B expressing cells, camptothecin-treated cells exhibited a slight decrease in cell viability (supplementary materials VIII). Yet, flow cytometry profiles indicate camptothecin-treated Top1B expressing cells were retained in S-phase throughout the treatment time, which is evidence of S-phase checkpoint activation [41], while untreated or *top1 Δ* control cells rapidly traversed S-phase and continued cycling asynchronously (supplementary materials VIII). During the 60 minutes that the cells were exposed to camptothecin, a number of the $2\mu\text{m}$ plasmids are expected to have completed replication, as the replication origin fires early in S-phase and the multicopy plasmid is a relatively small DNA template.

An analysis of $2\mu\text{m}$ DNA topoisomers purified from *top2_{ts}* cells in S-phase (Fig. 8.7b) yields similar results to that observed in the G1-phase experiments: the accumulation of positive supercoils occurs only in Top1B-expressing cells and only in the presence of camptothecin. In this case, about half of the plasmid topoisomers exhibited an increase in linking number. This outcome was somewhat diminished in wild-type *TOP2* yeast cells (Fig. 8.7c). Thus, as in G1-phase, alterations in Top2 activity affect the extent of positive supercoil accumulation in S-phase, induced by camptothecin-bound Top1B. The strand passage events catalyzed by Top2 would serve to disentangle any catenated daughter molecules induced by DNA replication [4–6, 42]; however, this activity is apparently less efficient in relaxing the local domain of positive supercoils generated by camptothecin-bound Top1B-DNA complexes in advance of the fork. We conclude that the camptothecin-induced accumulation of positive supercoils is not confined to a single phase of the cell cycle and may derive from the tracking of

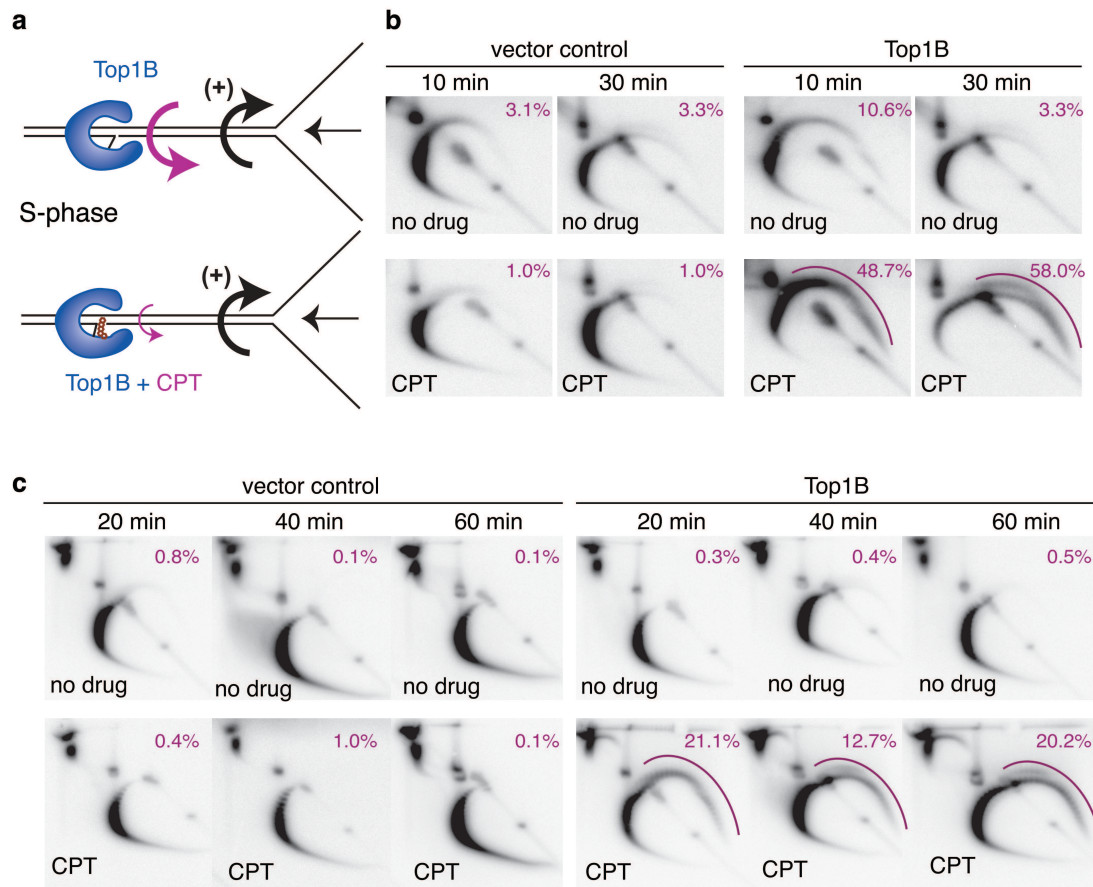


Figure 6.5: Positive supercoils also accumulate in S-phase in camptothecin-treated cells. (a) during processive replication, a replication fork (here moving leftwards) generates positive supercoils in advance of itself (DNA coiling is depicted as a black curved arrow). These can be removed by Top1B activity (magenta curved arrow). The velocity with which positive supercoils are removed by Top1B is reflected graphically by the size of the magenta arrow: in the absence of camptothecin (upper panel) positive supercoils are removed efficiently, while in the presence of the drug (lower panel), they are removed more slowly. (b) analysis of the $2\mu\text{m}$ DNA topoisomer distribution in *top1 Δ , top2^{ts}* cells expressing plasmid-encoded human Top1B or vector control, shifted to the nonpermissive temperature (36°C), and released into S-phase in the presence or absence of camptothecin for 10 or 30 minutes. (c) topoisomer distribution of the $2\mu\text{m}$ plasmid in wild-type *TOP2* yeast cells released into S-phase. Here too, the accumulation of positive supercoils is observed only when cells express Top1B and are treated with camptothecin. Purple numbers indicate the percentage of topoisomers exhibiting an increase in linking number.

a variety of complexes along the DNA duplex, including the transcription and replication machinery.

6.5 Camptothecin resistant Top1B fails to accumulate positive supercoils

Based on our data, we propose an alternative mechanism for the cytotoxic activity of camptothecins. As evidenced by our single-molecule data, we have made the first direct observation of the time scale of topotecan binding to Top1B, and we observe that the drug vastly increases the lifetime of the covalent complex. In addition, however, we observe a second consequence of topotecan binding, namely that drug binding hinders DNA uncoiling by Top1B such that the drug-stabilized Top1B complex relaxes negative supercoils more readily than positive supercoils. This was supported by our single-molecule experiments and also observed in yeast experiments, which demonstrated the accumulation of positive supercoils as a consequence of camptothecin poisoning of Top1B. Certainly, the increased lifetime of the covalent complex in the presence of drug is quantitatively compatible with the "collision" hypothesis which has traditionally been postulated to explain the cytotoxic activity of camptothecins (Fig. 8.8a, right). However, the increased lifetime of the covalent complex combined with the observation of a reduced relaxation rate of positive supercoils, provides a compelling argument for an intriguing alternate mechanism of drug-induced cell death that has not previously been considered (Fig. 8.8a, left). Instead of the covalent complex itself physically presenting an obstacle to the advancing fork, our single-molecule and *in vivo* data suggest that the local domain of positive supercoils generated ahead of the fork, which cannot be efficiently removed by the drug-bound Top1B or by wild-type levels of Top2, may itself hamper fork progression. This stalling of the replication machinery could result in fork collapse and the formation of potentially lethal DNA lesions that eventually induce cell death.

Implicit in this model is that the ability of camptothecins to selectively hinder the uncoiling of positively supercoiled DNA by Top1B would coincide with the cytotoxicity of the drug. Therefore, it follows that if the accumulation of positive supercoils is predictive of cellular response to this class of chemotherapeutics, then cells expressing a catalytically active, yet camptothecin-resistant Top1B mutant enzyme would fail to accumulate positively supercoiled plasmid DNA in response to camptothecin. To test this critical aspect of the model, the same TOP2 yeast strain used in Fig. 8.6d and 8.7c, was transformed with a plasmid expressing wild-type human Top1B or a camptothecin-resistant human

topoisomerase I mutant, Top1G365C. This mutant enzyme is catalytically active in yeast and *in vitro*, yet the single amino acid substitution of Cys for Gly365 renders yeast cells expressing this enzyme resistant to camptothecin [43]. Gly365 lies within a loop of the Lip1 domain of Top1. Crystallographic data suggest that the structure of this loop accommodates the deformation of the DNA helix upon camptothecin binding and dictates the orientation of residue Asp533, which makes direct contact with the bound drug [20]. From these studies, it follows that mutation of the flexible Gly at position 365 would also distort the structure of Lip1 to affect camptothecin binding of the Top1B-DNA complex. As the *in vivo* activity of this camptothecin-resistant mutant has previously been established in the same yeast genetic background used in these studies [43], we investigated whether the accumulation of positive supercoils was predictive of yeast cellular response to camptothecin. Fig. 8.8b shows the topoisomer distribution of 2 μ m DNA in asynchronously growing yeast cells containing wild-type Top1B (upper panel) and Top1G365C (lower panel). Interestingly, this mutant failed to induce positive supercoil accumulation in the presence of camptothecin, in full agreement with our proposed mechanism that positive supercoils are involved in Top1B-mediated and camptothecin-induced cell death. Thus, a single point mutation in an otherwise identical cellular background provided a stringent test of the specificity of Top1B-camptothecin interactions in causing the accumulation of positive supercoils observed in camptothecin-mediated cell death. Camptothecin analogues have demonstrated remarkable activity in the treatment of pediatric and adult cancer, yet the mechanism underlying their activity has been puzzling. It is known that topoisomerase I constitutes the cellular target of these agents and that they induce lethal Top1B-mediated DNA damage [7–10]. However, the specific DNA lesions that cause cell death remain poorly understood. In addition, the levels of Top1B protein or drug-stabilized Top1B-DNA complexes are not predictive of tumor cell response [34]. Indeed, the clinical development of camptothecins has been hampered by the lack of biomarkers with which to predict drug efficacy. Our observations highlight, for the first time, the role of positive supercoil accumulation in Top1B-mediated camptothecin poisoning. Importantly, our findings thereby also provide a novel conceptual and experimental framework with which to assess the efficacy of camptothecins and other Top1B-targeted agents *in vivo*. The unusual activity of camptothecins defined in our studies may also be exploited in the development of novel therapeutics that inhibit Top1B-catalyzed relaxation of negatively or positively supercoiled DNA.

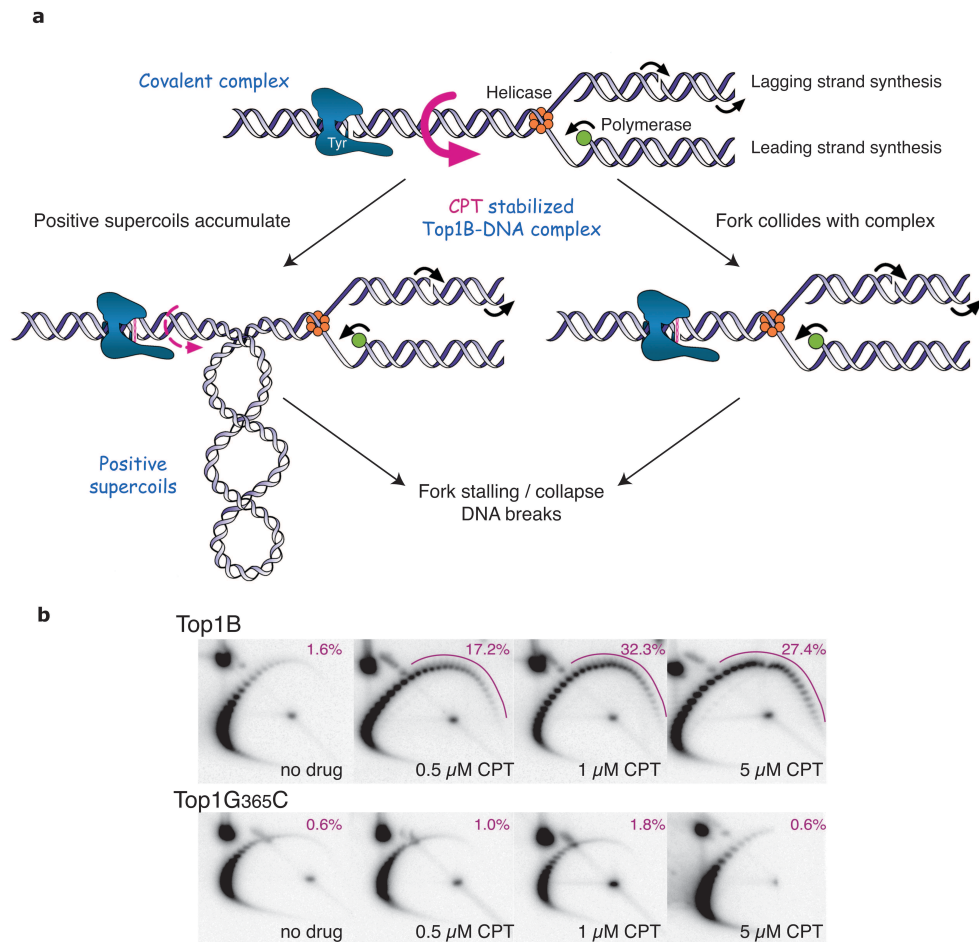


Figure 6.6: Camptothecin-induced toxicity results from Top1B-dependent accumulation of positive supercoils. (a) Schematic representation of distinct mechanisms of camptothecin-induced cell death *in vivo*. The replication fork (diagrammatically simplified to include only the replicative helicase and polymerase complex) generates positive supercoils in the DNA, which are removed (magenta curved arrow) by Top1B (covalent complex shown in blue). In the presence of camptothecin, fork stalling and fork collapse have been predicted to result from the physical collision of the advancing replication complex itself with the drug stabilized Top1B-DNA-covalent complex (lower right). Our data suggest a second scenario, in which fork progression and integrity are indirectly impaired by the unresolved positive supercoils (lower left). (b) analysis of the 2 μ m DNA topoisomer distribution in asynchronous cultures of *top1 Δ* cells expressing plasmid encoded wild-type Top1B or the catalytically active, camptothecin-resistant Top1G365C mutant enzyme and treated with the indicated concentration of camptothecin for 15 minutes. The accumulation of positive supercoils is only observed in drug-treated cells expressing the wild-type enzyme. Purple numbers indicate the percentage of topoisomers exhibiting an increase in linking number.

6.6 Methods

6.6.1 DNA constructs

Single-molecule experiments were performed with bacteriophage λ DNA (48 kilobases (kb) or 16 μ m in contour length) or half-length bacteriophage λ DNA (24 kb or 8 μ m in contour length). Measurements of the religation time were performed on shorter DNA (8 kb or 2.7 μ m contour length) to decrease the magnitude of the Brownian fluctuations of the bead and to increase the time resolution. The flow cell coating with anti-digoxigenin was performed as previously described [16, 42].

6.6.2 Enzyme and buffers

Full-length human Top1B and catalytically inactive Top1Y723F, each containing an N-terminal Flag epitope, were partially purified from galactose-induced cultures of EKY3 *top1 Δ* yeast cells as described [36]. To obtain homogenous protein preparations, Top1B fractions, bound to an anti-Flag M2 affinity gel (Sigma Chemical Co., St. Louis MO), and eluted with an excess of Flag peptide in TBS [50 mM Tris, pH 7.4, 150 mM KCl] were applied to a phosphocellulose column to remove the peptide. Homogeneous Top1B was eluted in TEEG buffer [50 mM Tris (pH 7.4), 1 mM EDTA, 1 mM EGTA, 10% glycerol] plus 1.0 M KCl and protease inhibitors, diluted with 50% glycerol and stored at -20°C. Top1B activity was assayed in plasmid DNA relaxation reactions and protein integrity assessed in immunoblots as described [36].

6.6.3 Strains and plasmids

S. cerevisiae strains EKY2 (MATa, *top1 Δ*), EKY3 (MAT?, *top1|Delta*) and JCW28 (MATa, *top1 Δ* , *top2^{ts}*) have been described [44]. For constitutive expression of Top1B or Top1G365C, the epitope tagged hTOP1 sequences were excised from YCpGAL1-ehTOP1 [36] or YCpGAL1-top1G365C [43] and cloned under the yeast TOP1 promoter in YCpScehTOP1.U or YCpScehTOP1G365C.U. The empty vector, YCpSc.U served as control.

6.6.4 *In vivo* assays and 2-D gel electrophoresis

Exponential cultures of MATa cells, transformed with YCpScehTOP1.U or vector control, were arrested in G1-phase with α -factor, then either incubated with additional α -factor and 30 μ M camptothecin or 0.25% DMSO (*v/v*) (no drug), or washed by filtration and released into S-phase with or without camptothecin. For

top2^{ts} strains, α -factor arrested cells were then shifted to 36°C for two hours with additional α -factor, then treated as above. Asynchronous cultures, incubated with 0.5, 1 or 5 μ M camptothecin for 15 minutes at 30°C, were also treated as above. To assess the distribution of 2 μ m plasmid DNA topoisomers, DNA isolated from cultures fixed with toluene/ethanol, were resolved in 2D gels electrophoresis and subjected to southern blotting as previously described [44, 45].

6.7 Acknowledgements

We thank Koen Besteman and Ulrich Keyser for useful discussions, Suzanne Hage and Ya-Hui Chien for DNA sample preparation, Robert van Waardenburg for help with PyMol, and Cees Dekker for critical reading and general support. We thank FOM, NWO, NIH (CA58755) and ALSAC for financial support.

6.8 Supplementary Materials

I

As has been shown for vaccinia virus Top1B [16], human Top1B does not always relax all DNA plectonemes present in the DNA in a single enzymatic event (Fig. 8.3a of main article). S.M. Figure 6.7 contains a collection of traces that exhibit this distributive behaviour of human Top1B.

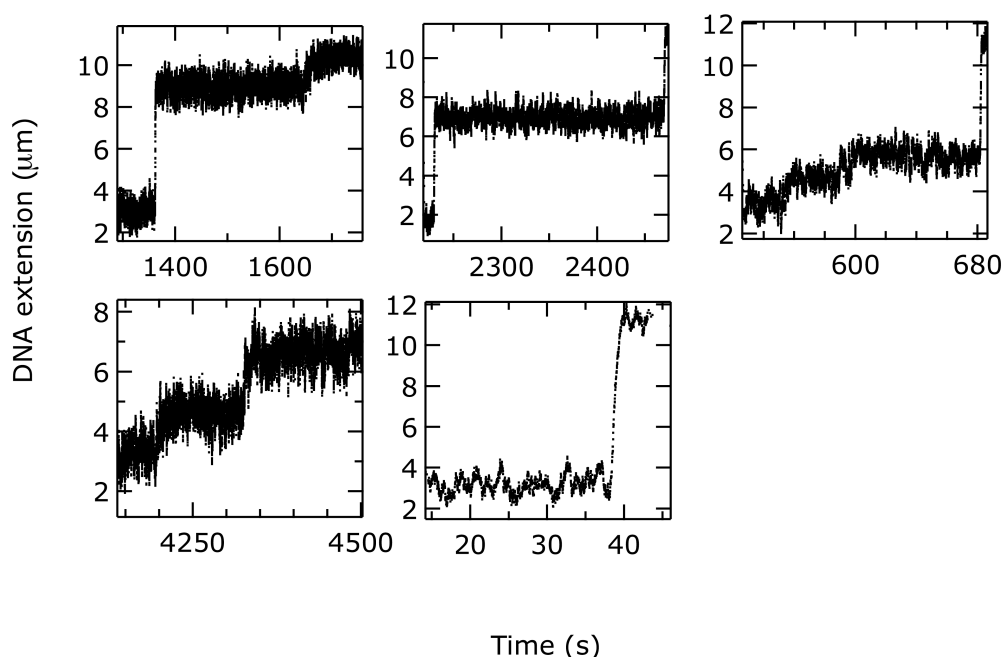


Figure 6.7: A collection of traces demonstrating that human Top1B can exhibit stepwise supercoil relaxation. The duration of the plateaus separating cleavage-religation cycles is dependent on the concentration of Top1B.

II

We show that the DNA extension velocity during topotecan-mediated Top1B uncoiling is independent of the ratio between the Top1B concentration and the topotecan concentration, as observed in S.M. Figure 6.8, where the ratio is varied by a factor of 25. We note that the uncoiling velocities shown in S.M. Figure 6.8 are not measured in a topotecan concentration-saturated regime, as a further slight decrease of the topotecan concentration leads to the disappearance of slow uncoiling within reasonable experimental timescales. The uncoiling events that are observed at such lower topotecan concentrations are fast and indistinguishable from Top1B-mediated uncoiling events in the absence of topotecan. These arguments suggest that a single topotecan molecule is responsible for the slow

uncoiling that is observed.

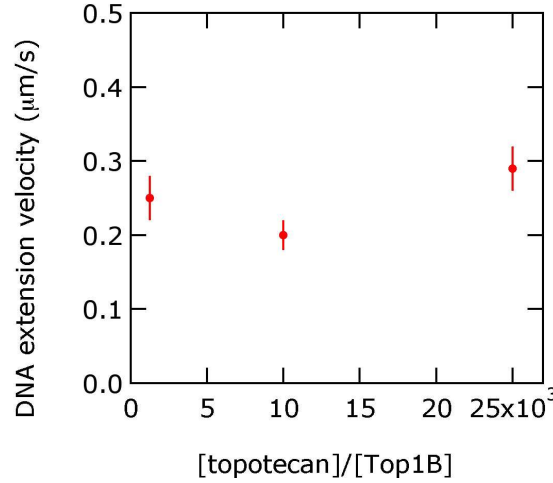


Figure 6.8: The DNA extension velocity of topotecan-mediated uncoiling is independent of the ratio of topotecan concentration, [TPT], to Top1B concentration, [Top1B].

Furthermore, we show that the time during which slow uncoiling is observed, $\langle t \rangle$ (see the main text under the heading "Drug binding traps Top1B on the DNA"), is independent of the topotecan concentration. This is important in light of a potential concern that $\langle t \rangle$ might be an overestimate if multiple topoisomerase enzymes, each associated with topotecan, induce successive slow uncoiling events separated by intervals smaller than our time resolution. If this were the case, however, one would expect $\langle t \rangle$ to depend on the topotecan concentration, with low topotecan concentrations leading to shorter $\langle t \rangle$ and high topotecan concentrations leading to longer $\langle t \rangle$. However, over the range of topotecan concentrations tested (which range from a lower limit below which no further slow relaxation events are observed to an upper limit approaching the topotecan solubility), no observable change in $\langle t \rangle$ is observed in our data (S.M. Figure 6.9). Thus, our measurement of $\langle t \rangle$ is not likely to be overestimated.

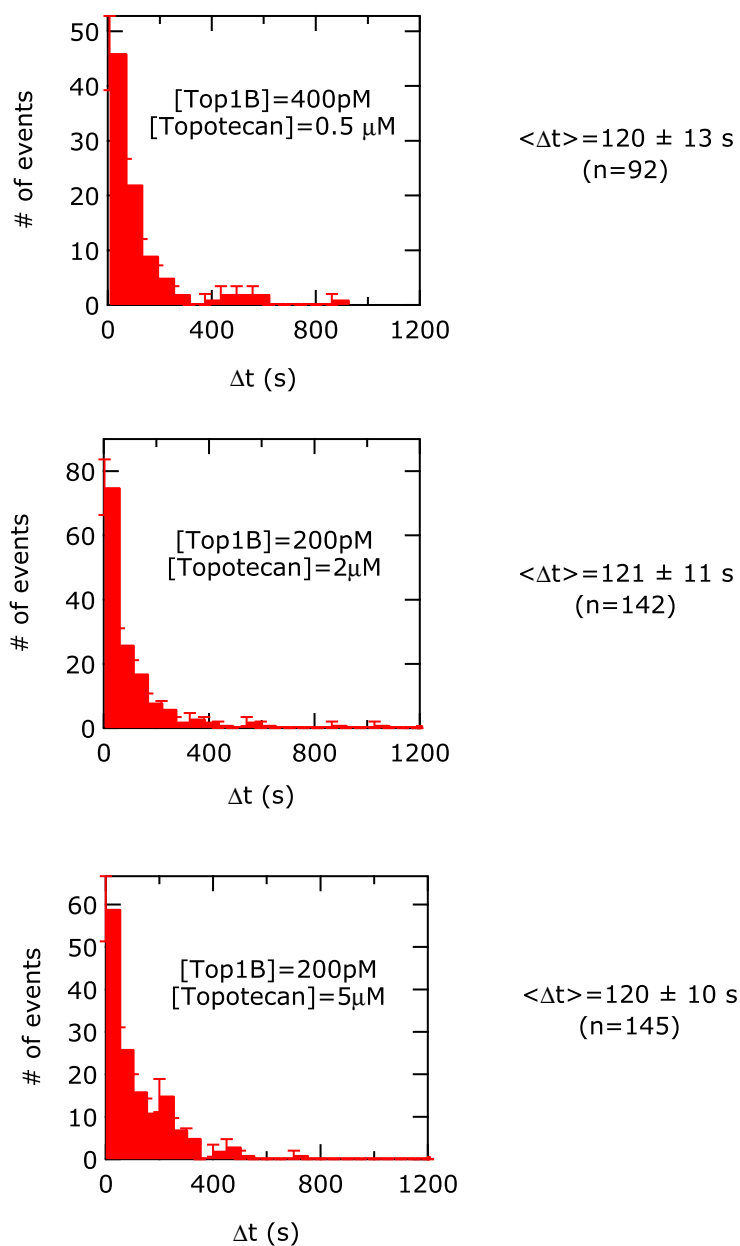


Figure 6.9: Upon varying the topotecan concentration by an order of magnitude, we do not observe a change in $\langle t \rangle$. This leads us to conclude that the measurement of $\langle t \rangle$ is not influenced by the topotecan concentration.

III

We show that topotecan-mediated supercoil removal occurs at thermodynamic equilibrium, allowing us to faithfully convert (linear) DNA extension velocities to angular velocities of DNA rotation. S.M. Figure 6.10 shows the strategy for

demonstrating this. First we measure the DNA extension as a function of magnet rotation, where each measurement point constitutes an average over 5 seconds and the waiting time between measurements is 10 s (S.M. Figure 6.10a). The values for the DNA extension obtained in this manner can be assumed to equal the equilibrium values. From these measurements, we extract the linear slope by fitting to a linear relation (red line), corresponding to the gain (or loss) in DNA extension upon removal (or addition) of plectonemes in the constant torque (writhe) regime, *e.g.* 36 nm/turn at a stretching force of 0.9 pN. We then measure the DNA extension as a function of time, while continuously rotating the magnets at a preset and known angular velocity (S.M. Figure 6.10b). We again fit the linear part of the increase in DNA extension with a linear relation and obtain a number for the DNA extension velocity at this given magnet rotation rate. We verify that the DNA molecule completes as many rotations as the magnets. From these two measurements, we can calculate the inferred angular velocity of the DNA. This procedure is repeated for various preset magnet rotation speeds and the inferred DNA rotation rate is plotted against the preset magnet rotation speed, as shown in S.M. Figure 6.10c. It is clearly seen that the two quantities are in excellent agreement with each other and that a faithful measurement of the DNA rotation rate can be made by measuring the DNA extension velocity, at least up to 3 Hz.

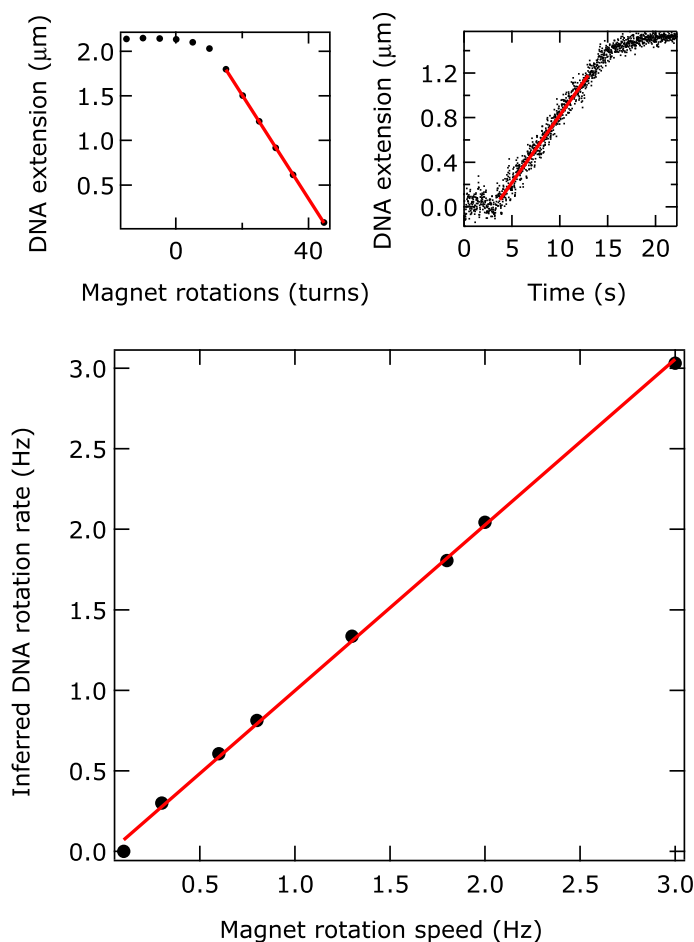


Figure 6.10: The strategy for proving that topotecan-mediated slow DNA uncoiling proceeds at thermodynamic equilibrium. S.M. Figure 6.10a shows the equilibrium measurement of the DNA extension *vs.* the number of magnet rotations. In the Writhe regime, where all additional twisting of the DNA molecule leads to the formation of plectonemic supercoils, we fit a linear relation (red line). S.M. Figure 6.10b shows the DNA extension as a function of the time during which the DNA molecule is uncoiled at a constant angular velocity. The resulting increase in DNA extension is also fitted with a linear relation (red line). S.M. Figure 8.5c shows the angular velocity of the magnets (magnet rotation speed) *vs.* the angular velocity of the DNA rotation (inferred DNA rotation rate), as calculated using the values obtained from the linear fits in S.M. Figures 6.10a and b (see the Supplementary Materials III text for a detailed description of the calculation). The regime of angular velocities probed here includes the angular velocity of topotecan-mediated uncoiling of DNA, as described in the main text.

IV

We show that the global mechanical properties of bare DNA are unaltered upon addition of 5 μM of topotecan in the same buffer used in the other experiments (S.M. Figure 6.11). The slow enzymatic uncoiling observed in the presence of topotecan can thus not be explained by a global stiffening or by a global change in the torque response of DNA.

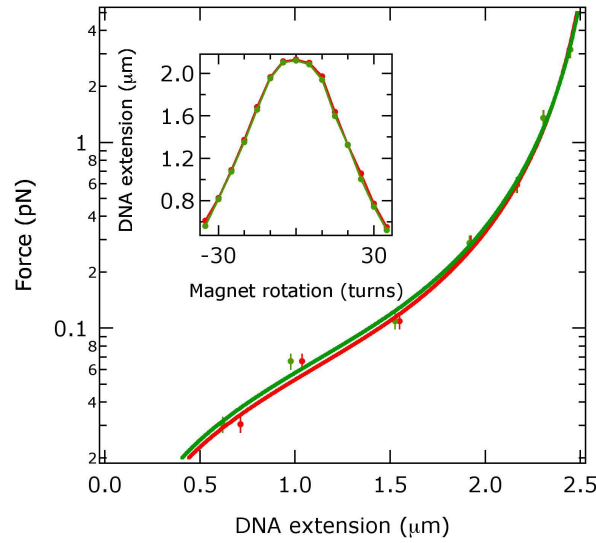


Figure 6.11: Force-extension and supercoiling (inset) behaviour of bare double-stranded DNA in the presence (red) and absence (green) of 5 μM of topotecan.

V

The main text describes the strategy for measuring the religation time, *i.e.* the time interval between cleavage of the DNA by Top1B and subsequent religation. As described in supplementary materials I, not all injected plectonemes are always removed in a single enzymatic event. Here we present a technical point concerning the implications of the latter for the measurement of the religation time.

If, following a single enzymatic event, plectonemes remain in the DNA, the religation time is measured as described in the main text and summarized in Fig 8.4g. In short, above a threshold in DNA extension (set above the experimental noise of the bead), which is crossed soon after Top1B nicks the DNA (at t_{nick}), the magnets spin continuously (red blocks in S.M. Figures reffigure5sma-c). At a certain point in time, the DNA is religated by Top1B and continuous rotation of the magnets will then lead to mechanical addition of plectonemes and consequently to a decrease in DNA extension. This point in time, designated $t_{\text{religated}}$, is used to calculate the religation time according to $t_{\text{religation}} = t_{\text{religated}} - t_{\text{nick}}$,

as described in the main text. However, a subtle complication occurs when all plectonemes are removed from the DNA in a single enzymatic event. In this case, the magnets continue to impose rotation on the bead after all plectonemes have been removed, *i.e.* the DNA is rotated while at $\sigma=0$, and the DNA extension initially remains constant, as shown in S.M. Figure Vb. At a later point in time during the continuous coiling of the DNA, the Top1B religates the DNA, and plectonemes can once again be introduced (treligated in S.M. Figure Vb). However, monitoring of the supercoiling state in the magnetic tweezers is largely limited to the observation of the plectonemic regime, which occurs beyond the buckling instability [29]. The time ΔT_{delay} it takes to reach this buckling instability from $\sigma=0$ therefore needs to be subtracted from the religation time. Thus for events in which all plectonemes are removed in a single enzymatic events, the correct religation time is $t_{religation} = t_{religated} - t_{nick} - \Delta T_{delay}$. S.M. Figure 6.12c shows the procedure for measuring ΔT_{delay} . Starting at $\sigma=0$ (at 59 s) we continuously rotate the magnets at the maximum angular velocity that the stepper motor, which rotates the magnets, allows. For our experimental configuration, this is 30 Hz. We monitor the DNA extension and observe at what time plectonemes appear. ΔT_{delay} is calculated by subtracting the time at $\sigma=0$ from the time at which plectonemes form, as shown in S.M. Figure 6.12c. The correct value of $t_{religation}$ is computed taking ΔT_{delay} into account.

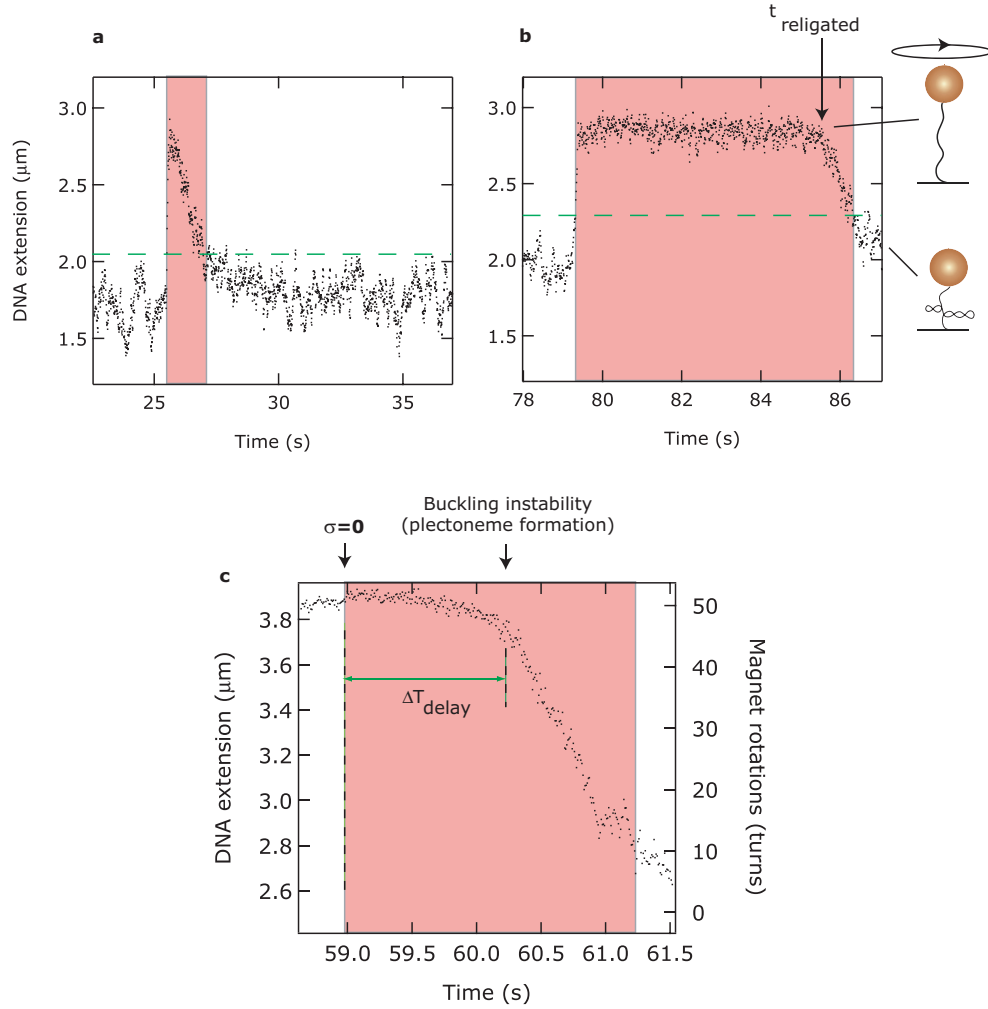


Figure 6.12: Strategy for measuring the religation time of Top1B. (a) When plectonemes remain in the DNA following a single enzymatic event, religation effectively results in an immediate decrease of the DNA extension. The religation time, $t_{religation}$, is then readily obtained by calculating $t_{religation} = t_{religated} - t_{nick}$, as described in the main text. (b) However, when all plectonemes are removed from the DNA in a single enzymatic event, a systematic correction of the religation time needs to be taken into account. The correction is the time delay between the instant of religation and the instant at which plectonemic supercoils form. (c) This time delay, ΔT_{delay} , is measured by continuously twisting a DNA molecule from $\sigma=0$ (at 59.0 s) to the onset of the formation of plectonemes (the buckling instability, at 60.2 s) and is approximately 1.2 s. The correct calculation for the religation time incorporates the correction according to $t_{religation} = t_{religated} - t_{nick} - T_{delay}$.

VI

In this part of the Supplementary Materials, we explain the principle of two-dimensional gel electrophoresis.

For a double-stranded plasmid DNA of a given size, the overall compaction of the molecule will determine its electrophoretic mobility in an agarose gel [46]. The topological state of a plasmid is described by the linking number (Lk), which is the sum of two geometrical properties, the twist (Tw) plus the writhe (Wr). Tw describes the number of times one strand wraps around the other and Wr describes the coiling of the helical axis of the DNA. Plasmid DNAs of identical sequence, but with different Lk, can be resolved in agarose gels on the basis of their Wr, where the Wr of a DNA topoisomer being considered is proportional to Lk, the difference between the Lk of the plasmid topoisomer being considered and what the Lk of the same plasmid would be if it were relaxed. Under physiological conditions, the latter may be approximated by (size of plasmid in bp) / (10.5 bp per helical turn). The Lk of an intact circular plasmid is invariant and will always be an integral value. However, the relative contribution of Tw and Wr may vary.

As an example, let us consider a negatively supercoiled 2,100 bp plasmid DNA with an Lk of 192. The same plasmid, when relaxed, would have an Lk of 200. We also assume that the entire Lk deficit is partitioned into Wr, such that $Lk = Tw + Wr$ is described by $192 = 200 + (-8)$, as shown for form I in S.M. Figure 6.13a. If this plasmid is incubated with increasing concentrations of the DNA intercalator chloroquine (indicated by grey rectangles) and in the absence of a topoisomerase, the decrease in Tw induced by chloroquine binding will induce a compensatory increase in Wr, since Lk must remain constant. As depicted in S.M. Figure 6.13a, the effect of increasing chloroquine binding (decreased Tw and increased Wr) will affect the compaction of the molecule and hence its electrophoretic mobility. In the absence of chloroquine, the DNA would be very compact ($Wr = -8$, S.M. Figure 6.13a. I) and migrate rapidly; in a gel containing low concentrations of chloroquine (such that $Wr = -4$, S.M. Figure 6.13a. II), it would migrate more slowly, and at somewhat higher concentrations of chloroquine (such that $Wr = 0$, S.M. Figure 6.13a. III), it would migrate more slowly still. However, at even higher concentrations of chloroquine, the DNA would acquire a positive Wr ($Wr = +4$, S.M. Figure 6.13a. IV), and would again migrate more rapidly, as if positively supercoiled.

These same principles apply to the electrophoretic resolution of plasmid DNA topoisomers in two dimensions. In the first dimension (indicated by the vertical arrow in S.M. Figure 6.13b), the DNAs are resolved in the presence of low concentrations of chloroquine, such that a DNA topoisomer with a slight deficit in Lk would acquire the lowest Wr and run at apex I (S.M. Figure 6.13b). DNAs of lower or higher Lk would migrate with a mobility difference proportional to absolute value of the linking number difference from the topoisomer at apex I.

The gel is then rotated 90° and run in a second dimension (indicated by the horizontal arrow in S.M. Figure 6.13b) in the presence of an increased concentration of chloroquine. Here, a plasmid of even lower Lk will now acquire the lowest W_r (apex II, S.M. Figure 6.13b). As in the first dimension, DNAs of lower or higher Lk will now migrate with a mobility difference proportional to absolute value of the linking number difference from the topoisomer at apex II. Once the gel has been run in the two dimensions with the different chloroquine concentrations as described, a topoisomer arc as diagrammed in S.M. Figure 6.13b results, where a given topoisomer spot differs from the adjacent spot by a linking number difference of one. This analysis permits a clear discrimination between plasmids with a linking number deficit, which are skewed towards the lefthand side of the arc, and plasmids with a linking number excess, which are skewed towards the righthand side of the arc. Plasmid DNA that is nicked is free to assume the most relaxed conformation and will migrate at the apex in both dimensions, as shown.

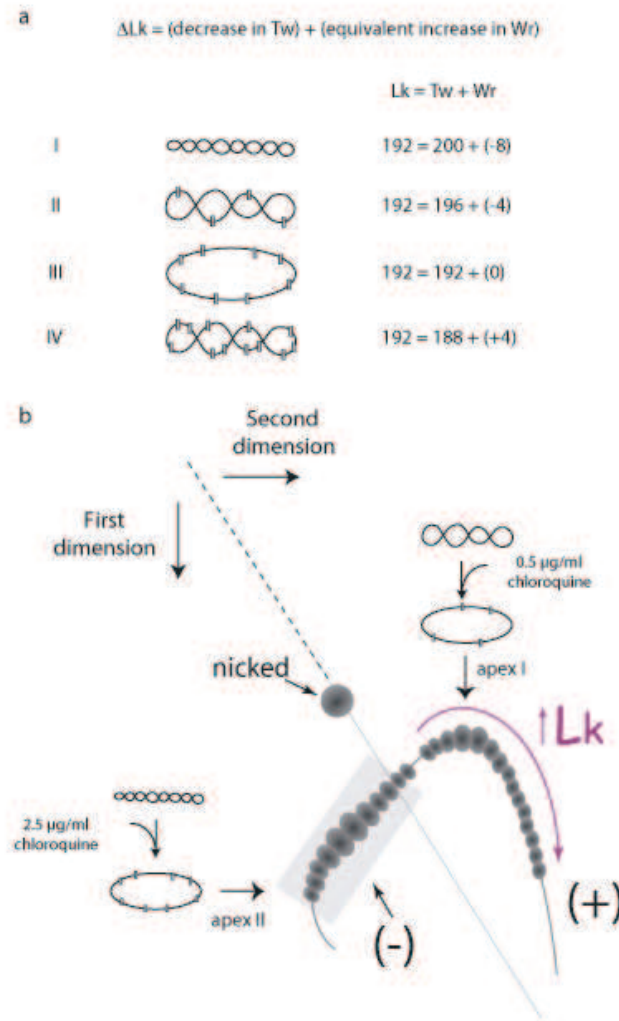


Figure 6.13: Topological isomers of DNA can be resolved using 2D agarose gel electrophoresis.

VII



Figure 6.14: Camptothecin-induced accumulation of positive supercoils requires the expression of catalytically active Top1B. Exponential cultures of *top1Δ, top2ts* cells, transformed with plasmids expressing low levels of wild-type human Top1B or the catalytically inactive human Top1Y723F mutant, were shifted to 36° C for 1 hour to inactivate Top2, then treated with camptothecin (CPT) for 30 minutes. The resulting distribution of 2 μ m plasmid DNA topoisomers was assessed in 2-D gels, as described in the main text. These data demonstrate expression of a catalytically active Top1B is necessary for camptothecin-induced accumulation of positive supercoils.

VIII

We show that camptothecin treatment of yeast cells, as they synchronously transit S- phase of the cell cycle, induces Top1B-dependent DNA damage which has modest effects on cell viability, yet nonetheless suffices to trigger the slow kinetics of S-phase transit characteristic of S-phase checkpoint activation. Yeast *top1Δ* cells, transformed with a plasmid that expresses low levels of human Top1B or an empty vector control were arrested in G1-phase with the mating pheromone -factor and then released into S- phase in the presence or absence of camptothecin (CPT).

In panel (a), samples of cells, taken at the times indicated, were serially ten-fold diluted and 5 l aliquots were spotted onto agar plates. Following incubation at 30°C, cell viability was assessed by the formation of colonies. The viability of untreated Top1B expressing cells (labelled Top1B), or untreated or camptothecin vector control cells (labelled vector control and vector control + CPT, respectively) was unchanged during the course of the experiment. Top1B expressing cells treated with camptothecin (labelled Top1B + CPT) began to exhibit a slight reduction in cell viability at 40 minutes following release into S-phase, with about a 10-fold reduction in cell viability relative to 0 time evident at 60 minutes drug treatment. As shown in panel (b), the cell cycle distribution of cells, treated as in panel a, was also assessed. Cells taken at the indicated times were fixed with

70% ethanol, stained with propidium iodide and assayed for DNA content by flow cytometry [47]. The peak height reflects the number of cells in the sample, while the distribution along the X axis is an indication of DNA content. Haploid yeast cells in G1 phase of the cell cycle accumulate with a 1N DNA content, while cells in G2/M phases of the cell cycle have a 2N DNA content. As cells transit S-phase, DNA content increases from 1N to 2N. In untreated Top1B expressing cells and in the untreated or CPT treated vector controls, the cells exhibit a similar pattern of S-phase progression, with the majority of cells in G2/M phase at 60 minutes as indicated by the presence of a 2N peak. In contrast, camptothecin treatment of Top1B expressing cells (Top1B+CPT) exhibit a slow, synchronous transit through S-phase which results from DNA damage-induced activation of the S-phase checkpoint.

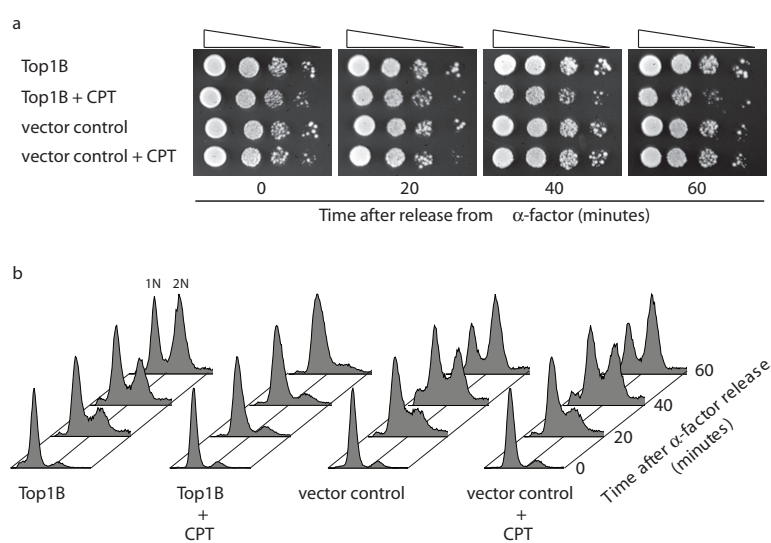


Figure 6.15: Camptothecin treatment of yeast cells in S-phase induces a Top1B-dependent decrease in cell viability (a) and a delay in cell cycle progression (b).

References

- [1] Watson, J. D. and Crick, F. H. *Nature* **171**(4356), 737–8 (1953).
- [2] Giaever, G. N. and Wang, J. C. *Cell* **55**(5), 849–56 (1988).
- [3] Liu, L. F. and Wang, J. C. *Proc Natl Acad Sci U S A* **84**(20), 7024–7 (1987).
- [4] Champoux, J. J. *Annu Rev Biochem* **70**, 369–413 (2001).
- [5] Corbett, K. D. and Berger, J. M. *Annu Rev Biophys Biomol Struct* **33**, 95–118 (2004).
- [6] Wang, J. C. *Nat Rev Mol Cell Biol* **3**(6), 430–40 (2002).
- [7] Bjornsti, M. A. *Cancer Cell* **2**(4), 267–73 (2002).
- [8] Li, T. K. and Liu, L. F. *Annu Rev Pharmacol Toxicol* **41**, 53–77 (2001).
- [9] Nitiss, J. L. *Curr Opin Investig Drugs* **3**(10), 1512–6 (2002).
- [10] Pommier, Y. *Nat Rev Cancer* **6**(10), 789–802 (2006).
- [11] Minsky, B. D. *Oncology (Williston Park)* **18**(14 Suppl 14), 49–55 (2004).
- [12] Rodriguez-Galindo, C., Radomski, K., Stewart, C. F., Furman, W., Santana, V. M., and Houghton, P. J. *Med Pediatr Oncol* **35**(4), 385–402 (2000).
- [13] Stewart, D. J. *Oncologist* **9 Suppl 6**, 33–42 (2004).
- [14] Redinbo, M. R., Stewart, L., Kuhn, P., Champoux, J. J., and Hol, W. G. *Science* **279**(5356), 1504–13 (1998).
- [15] Sekiguchi, J. and Shuman, S. *J Biol Chem* **269**(50), 31731–4 (1994).
- [16] Koster, D. A., Croquette, V., Dekker, C., Shuman, S., and Dekker, N. H. *Nature* **434**(7033), 671–4 (2005).
- [17] Stivers, J. T., Harris, T. K., and Mildvan, A. S. *Biochemistry* **36**(17), 5212–22 (1997).
- [18] Stivers, J. T., Shuman, S., and Mildvan, A. S. *Biochemistry* **33**(51), 15449–58 (1994).
- [19] Stivers, J. T., Shuman, S., and Mildvan, A. S. *Biochemistry* **33**(1), 327–39 (1994).
- [20] Chrencik, J. E., Staker, B. L., Burgin, A. B., Pourquier, P., Pommier, Y., Stewart, L., and Redinbo, M. R. *J Mol Biol* **339**(4), 773–84 (2004).
- [21] Porter, S. E. and Champoux, J. J. *Nucleic Acids Res* **17**(21), 8521–32 (1989).
- [22] Staker, B. L., Hjerrild, K., Feese, M. D., Behnke, C. A., Burgin, A. B., J., and Stewart, L. *Proc Natl Acad Sci U S A* **99**(24), 15387–92 (2002).

-
- [23] Hsiang, Y. H., Hertzberg, R., Hecht, S., and Liu, L. F. *J Biol Chem* **260**(27), 14873–8 (1985).
- [24] Holm, C., Covey, J. M., Kerrigan, D., and Pommier, Y. *Cancer Res* **49**(22), 6365–8 (1989).
- [25] Horwitz, S. B. and Horwitz, M. S. *Cancer Res* **33**(11), 2834–6 (1973).
- [26] Hsiang, Y. H., Lihou, M. G., and Liu, L. F. *Cancer Res* **49**(18), 5077–82 (1989).
- [27] Liu, L. F., Desai, S. D., Li, T. K., Mao, Y., Sun, M., and Sim, S. P. *Ann N Y Acad Sci* **922**, 1–10 (2000).
- [28] Bustamante, C., Bryant, Z., and Smith, S. B. *Nature* **421**(6921), 423–7 (2003).
- [29] Strick, T., Allemand, J., Croquette, V., and Bensimon, D. *Prog Biophys Mol Biol* **74**(1-2), 115–40 (2000).
- [30] Smith, S. B., Finzi, L., and Bustamante, C. *Science* **258**(5085), 1122–6 (1992).
- [31] Champoux, J. J. and Aronoff, R. *J Biol Chem* **264**(2), 1010–5 (1989).
- [32] Stewart, L., Ireton, G. C., and Champoux, J. J. *J Biol Chem* **274**(46), 32950–60 (1999).
- [33] Jaxel, C., Capranico, G., Kerrigan, D., Kohn, K. W., and Pommier, Y. *J Biol Chem* **266**(30), 20418–23 (1991).
- [34] Slichenmyer, W. J., Rowinsky, E. K., Donehower, R. C., and Kaufmann, S. H. *J Natl Cancer Inst* **85**(4), 271–91 (1993).
- [35] Shuman, S., Golder, M., and Moss, B. *J Biol Chem* **263**(31), 16401–7 (1988).
- [36] Woo, M. H., Vance, J. R., Marcos, A. R., Bailly, C., and Bjornsti, M. A. *J Biol Chem* **277**(6), 3813–22 (2002).
- [37] Sari, L. and Andricioaei, I. *Nucleic Acids Res* **33**(20), 6621–34 (2005).
- [38] McClendon, A. K., Rodriguez, A. C., and Osheroff, N. *J Biol Chem* **280**(47), 39337–45 (2005).
- [39] Goto, T. and Wang, J. C. *Cell* **36**(4), 1073–80 (1984).
- [40] Pramila, T., Wu, W., Miles, S., Noble, W. S., and Breeden, L. L. *Genes Dev* **20**(16), 2266–78 (2006).
- [41] Osborn, A. J., Elledge, S. J., and Zou, L. *Trends Cell Biol* **12**(11), 509–16 (2002).
-

- [42] Strick, T. R., Croquette, V., and Bensimon, D. *Nature* **404**(6780), 901–4 (2000).
 - [43] Woo, M. H., Losasso, C., Guo, H., Pattarello, L., Benedetti, P., and Bjornsti, M. A. *Proc Natl Acad Sci U S A* **100**(24), 13767–72 (2003).
 - [44] Hann, C., Evans, D. L., Fertala, J., Benedetti, P., Bjornsti, M. A., and Hall, D. J. *J Biol Chem* **273**(14), 8425–33 (1998).
 - [45] Megonigal, M. D., Fertala, J., and Bjornsti, M. A. *J Biol Chem* **272**(19), 12801–8 (1997).
 - [46] Hanai, R. and Roca, J. *DNA topoisomerase protocols: DNA topology and enzymes.*, volume 19-27. Humana Press, NJ, (1999).
 - [47] Haase, S. B. and Lew, D. J. *Methods Enzymol* **283**, 322–32 (1997).
-

Chapter 7

Atomic force microscopy shows that vaccinia topoisomerase IB generates filaments on DNA

Type IB DNA topoisomerases cleave and rejoin one strand of the DNA duplex, allowing for the removal of supercoils generated during replication and transcription. In addition, electron microscopy of cellular and viral TopIB-DNA complexes has suggested that the enzyme promotes long-range DNA-DNA crossovers and synapses. Here, we have used the atomic force microscope to visualize and quantify the interaction between vaccinia topoisomerase IB (vTopIB) and DNA. vTopIB was found to form filaments on nicked-circular DNA by *intramolecular* synapsis of two segments of a single DNA molecule. Measuring the filament length as a function of protein concentration showed that synapsis is a highly co-operative process. At high protein:DNA ratios, synapses between *distinct* DNA molecules were observed, which led to the formation of large vTopIB-induced DNA clusters. These clusters were observed in the presence of Mg^{2+} , Ca^{2+} , or Mn^{2+} , suggesting that the formation of *intermolecular* vTopIB-mediated DNA synapsis is favored by screening of the DNA charge.

7.1 Introduction

Type IB DNA topoisomerases (TopIB) remove the torsional stress that accumulates in double-stranded DNA molecules during replication and transcription. TopIB cleaves and rejoins one strand of the DNA duplex [1], allowing for the removal of positive or negative supercoils from supercoiled DNA. The topoisomerase from vaccinia virus (vTopIB) is the smallest topoisomerase known (314 amino acids), suggesting that vTopIB constitutes the minimal functional unit of a type IB enzyme [2]. vTopIB binds non-specifically to duplex DNA [3], although cleavage preferentially occurs at the target sequence 5'-(T/C)CCTT↓, where ↓ denotes the cleavage site [4, 5]. The sequence-specific cleavage of vTopIB has allowed for accurate determinations of the rate constants for cleavage and religation, hence its use as a prototype. Sedimentation analysis, gel filtration, N-terminal sequencing of the native protein, and conceptual translation of the open reading frame of the vTopIB gene indicate that vTopIB is a monomeric protein with a molecular mass of 36.7 kDa [6–8]. vTopIB consists of two domains joined through a hinge and binds circumferentially to DNA [2, 9, 10]. DNase I and exonuclease III footprinting give an upper bound to the number of base pairs covered by vTopIB when bound to DNA (25 base pairs, ~8.5 nm) [4, 11].

The binding of eukaryotic Type IB topoisomerase to DNA duplexes has been the subject of two previous studies that relied on visualization using electron microscopy [12, 13], both of which demonstrated binding of topoisomerases at intramolecular DNA crossovers. vTopIB was found to form intramolecular loop structures in which non-contiguous DNA segments were synapsed within filamentous protein stems [12]. Shuman *et al.* suggested that these loops arise through protein-protein mediated DNA synapsis [12]. The presence of filament-like structures at high topoisomerase:DNA ratios and individual intramolecular nodes at low topoisomerase:DNA ratios led to the suggestion that binding of vTopIB might be cooperative. However, the experimental strategy, which included glutaraldehyde fixation, prevented reliable quantification of the binding vTopIB to DNA. Atomic force microscopy (AFM) [14] presents several advantages over electron microscopy techniques, most notably that no fixation procedures are needed to adsorb biological molecules on a surface. Furthermore, AFM allows rapid visualization of many individual molecules and the measurements of heights and volumes. Indeed, AFM has proven to be a very suitable technique to characterize and quantify protein-protein and DNA-protein interactions [15–18], and therefore we employed this technique to study the interaction between vTopIB and DNA. AFM revealed individual vTopIB binding events, intramolecular synapsis, and filament-like structures on individual DNA molecules at vTopIB:DNA

ratios from 10:1 to 60:1. Our data is supported by a model in which DNA-bound vTopIB protein forms a single intramolecular node in the DNA. From this single intramolecular node a filament-like structure extends by the binding of new vTopIB protomers. We have used the theory published by McGhee and von Hippel to quantify the generation of filaments by vTopIB on DNA [19, 20]. By measuring the length of the filaments and the fractional saturation of the DNA as a function of the concentration of vTopIB, the cooperativity parameter ω and the association constant K^* to form a node were determined. At higher protein:DNA ratios, intermolecular synapses of DNA molecules were observed in the presence of divalent ions such as Mg^{2+} , Ca^{2+} , and Mn^{2+} .

7.2 Materials and methods

DNA constructs

In all experiments the pGEM-3Z plasmid (2743 bp) was used in nicked-circular or in linear form. The nicked-circular DNA was generated by relaxing the plasmid using the enzyme N.BstNB I (New England Biolabs). This enzyme nicks the plasmid at five positions. To generate a linear fragment, the plasmid was digested with BamHI (New England Biolabs). The nicked or linearized DNA molecules were then purified using a DNA purification kit (MoBio Laboratories), checked by gel electrophoresis, and inspected by AFM. This method resulted in nicked-circular and linear DNA molecules with mean contour lengths of 919 ± 69 nm and 928 ± 52 nm, respectively. These values agree with the expected B-form contour length of the plasmid. The cleavage consensus sequence of vaccinia topoisomerase ($5'-(\text{C/T})\text{CCTT}\downarrow$), occurred 16 times in the sequence of the DNA used.

Vaccinia topoisomerase IB

vTopIB was purified as previously described [7]. vTopIB was aliquoted and stored at -80°C in 10 mM Phosphate Buffer (PB) at pH 7.4. For each experiment a new aliquot was thawed and diluted in PB to the desired concentration.

AFM sample preparation

DNA molecules can be adsorbed on a mica surface using divalent ions (such as Mg^{2+} , Mn^{2+} , or Ca^{2+}), or using polylysine (PL) [21, 22]. These two different adsorption methods are known to yield significantly different appearances of bare DNA on mica because in the first case DNA molecules equilibrate in a two-dimensional configuration, while in the second case they are kinetically trapped onto the mica surface [23]. Fig. 8.3 shows DNA molecules adsorbed using 13 mM MgCl_2 (similar concentrations were used in experiments involving MnCl_2 or CaCl_2 , discussed below), and Fig. 8.3b shows DNA molecules adsorbed on

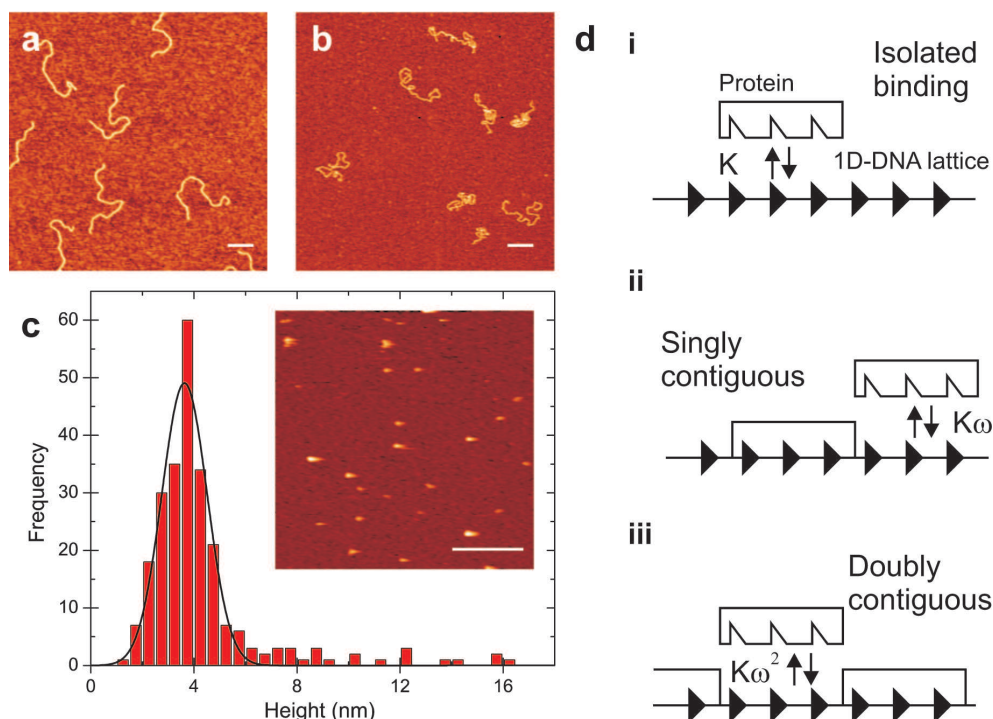


Figure 7.1: Control experiments. DNA molecules are not observed to cluster in the absence of vTopIB when adsorbed on a mica surface using MgCl_2 (a) or polylysine-coated mica (b). The appearance of the DNA molecules differs as a result of the adsorption processes involved. (c), Histogram of the height of vTopIB molecules measured in buffer in the absence of DNA. The average height of vTopIB in solution is 3.6 ± 1.8 nm, compatible with a monomeric state of the protein. An AFM image of vTopIB in buffer is shown as an inset in (c). Bar size is 200 nm. (d), Summary of cooperativity theory [19, 20]. Three different binding events can take place: isolated binding with an association constant K (d, i), singly contiguous binding with an association constant $K\omega$ (d, ii), and doubly contiguous binding with association constant $K\omega^2$ (d, iii). In our model an isolated binding event is attributed to the formation of an intramolecular node.

0.01% polylysine-treated mica. In both cases, there was no vTopIB present in the solution and the length of the DNA used was identical.

We observed individual DNA molecules homogeneously distributed over the mica surface in both cases. The polylysine-coated mica substrates were extensively rinsed with water and dried to remove any surface-mobile positive counterions [24, 25]. A control experiment was done with vTopIB imaged in buffer in the absence of DNA (Fig. 8.3c). The height distribution obtained from AFM (#events=246) displayed a peak and was fitted to a Gaussian function, yielding an average height of 3.6 ± 0.9 nm, where we have quoted the standard deviation

of the Gaussian function fit. From X-ray crystallography, the structure of the vTopIB catalytic domain shows it to be an oblong shaped globular protein with dimensions 5.6 x 2.8 x 2.6 nm [2]. Our AFM height measurements agreed reasonably well with these dimensions given that they reflected both the average height of different orientations of the protein and the small deformation exerted by the tip in buffer (estimated at 15%, given that average heights of dsDNA molecules under the same conditions measured 1.7 ± 0.3 nm, compared to a nominal value of 2 nm). Modeling of the crystal structure of vTopIB onto B-form DNA indicated that vTopIB likely covers an 18-bp DNA segment [2]. This value, $n=18$ bp, will be used in our modeling below.

vTopIB-DNA binding reactions included 0.22 nM DNA (circular or linear) and 2.5 to 40 nM vTopIB in PB buffer, supplemented, when needed, with 13 mM MgCl_2 to promote adsorption of the DNA molecules onto the mica. After mixing, the sample was incubated at 37°C for 15 min and deposited on a flat freshly-cleaved mica surface, rinsed with water, and dried under a gentle stream of nitrogen.

AFM imaging

Samples were imaged in air (except Fig. 8.3c, which was taken in buffer solution) with a commercial AFM from Nanotec Electronica operating in dynamic mode, using soft cantilevers (0.39 N/m, Olympus OMCL-RC800PSA). Typical tapping amplitudes were 15-17 nm peak-to-peak at the resonance frequency of the cantilevers, ~ 75 kHz. Image processing, including measurement of filament lengths, was done using WSxM freeware (www.nanotec.es). Standard image processing included plane subtraction and flattening.

Theoretical description of protein binding to DNA

We used the theory described by McGhee and Von Hippel [20] to quantify the generation of filaments of vTopIB on DNA. We briefly summarize the theory's main elements here. The binding of a protein that covers n basepairs on a DNA is characterized by an association constant K (Fig. 8.3d,i). The binding of a protein on one side of a previously-bound protein (while on the other side there is no bound protein) has an association constant $K\omega$. This is denoted singly-contiguous binding (Fig. 8.3d,ii). If a new protein binds between two previously bound proteins one speaks of a doubly-contiguous binding event with an association constant $K\omega^2$ (Fig. 8.3d,iii). ω is a measure for the degree of cooperativity of the binding process: $\omega = 1$ is defined as non-cooperative binding whereas $\omega > 1$ is defined as cooperative binding.

The parameters K and ω that characterize binding can be experimentally obtained from the AFM data by measuring the filament length distribution at a given DNA saturation θ , which is defined as the average fraction of DNA covered

by the protein. The filament length c is defined as the number of proteins that are bound next to each other without a gap between them. θ can be experimentally determined from the images by calculating the fraction of the total contour length of all DNA molecules present in the sample that has proteins bound to it. At a given θ , the distribution of filament lengths depends on the values of n, K , and ω [19]. The normalized probability P_c of finding a filament of length c is

$$P_c = [P(b_1|b_n)]^{c-1}[1 - P(b_1|b_n)], \quad (7.1)$$

where $P(b_1|b_n)$ is the probability of finding a protein bound right next to another one that is already bound to DNA [20]. $b_1|b_n$ denotes the protein b_1 binding to the protein b_n :

$$P(b_1|b_n) = \frac{a - (n - 2\omega + 1)\nu - R}{2\nu(\omega - 1)}, \quad (7.2)$$

where

$$R = \sqrt{[1 - (n + 1)\nu]^2 + 4\omega\nu(1 - n\nu)} \quad (7.3)$$

and $\nu = \frac{\theta}{n}$. From experimental data, P_c can be calculated and plotted as a function of the filament length c . Using a maximum-likelihood data analysis (see below), a value for ω was obtained for different DNA saturations θ . Once the cooperativity parameter ω is known, K can be deduced using Eq.7.4.

$$\frac{\nu}{L} = K(1 - n\nu) \left[\frac{(2\omega - 1)(1 - n\nu) + \nu - R}{2(\omega - 1)(1 - n\nu)} \right]^{n-1} \left[\frac{1 - (n + 1)\nu + R}{2(1 - n\nu)} \right]^2, \quad (7.4)$$

where R and ν are defined in Eq. 7.3, N is the total number of binding sites (bases of DNA) and L is the free protein concentration defined as $L = L_T - \nu N$ with L_T being the total protein concentration [20]. A fit of Eq. 7.4 to the plot of θ versus the total concentration of vTopIB (L_T) yields a value for the association constant K .

Data analysis

The length of the vTopIB filaments was measured by tracing the filaments with a segmented line. Tracing resulted in a height profile along the contour of the DNA molecule. The length of the filament was measured by attributing the beginning and the end of the filament to the positions where the height of the profile was reduced to half the maximum value. Next, the effect of tip convolution was taken into account by subtracting from the measured filament length the extra length induced by tip convolution, which was 13 nm. Tip convolution was

estimated following ref. [26], taking into account the tip radius specified by the manufacturer, 10 nm, and the DNA radius, 1 nm.

The maximum-likelihood (ML) method was used to estimate the cooperativity parameter ω . The ML-method is a powerful method to unambiguously estimate a parameter from a series of measurements [27]. The ML-method computes the distribution parameters that maximize the likelihood to observe a set of experimental outcomes given a known probability function. The likelihood function (LH) for the filament length distribution is defined as

$$LH = \prod_{i=1}^S P_{c_i} = \prod_{i=1}^S ([P(b_1|b_n)]^{c_i-1} [1 - P(b_1|b_n)]), \quad (7.5)$$

where S is the number of data points in the data set and c_i is the length of the i th filament. Eq. 7.5 can be rewritten as

$$\ln LH = S \ln[1 - P(b_1|b_n)] - S \ln P(b_1|b_n) + \sum_{i=1}^S (c_i \ln P(b_1|b_n)). \quad (7.6)$$

We then maximize this equation with respect to ω . Using the Maple software package together with Eq. 7.2 and 7.3, an analytical solution for ω can be obtained:

$$\omega = \frac{S\theta n \sum_{i=1}^S c_i + \theta S^2 - \theta n (\sum_{i=1}^S c_i)^2 - \theta S \sum_{i=1}^S c_i - n S \sum_{i=1}^S c_i + n (\sum_{i=1}^S c_i)^2}{\theta S^2}. \quad (7.7)$$

Using Eq. 7.7, ω can be calculated for different DNA saturations θ , which are experimentally obtained from the images. The uncertainty in ω corresponds to the places where $\ln LH$ is below its half maximum [28]. Because LH is not symmetric around the peak value, the corresponding uncertainty is also asymmetric. However, due to the negligible difference between positive and negative errors found, we simply quote their mean.

7.3 Results

vTopIB generates filament-like structures at low vTopIB:DNA ratios

We characterized the binding of vaccinia topoisomerase IB to linear and circular DNA at vTopIB:DNA ratios from 10:1 to 60:1. These experiments are summarized in Fig. 8.4a (linear DNA) and Fig.8.4b (circular DNA). We found

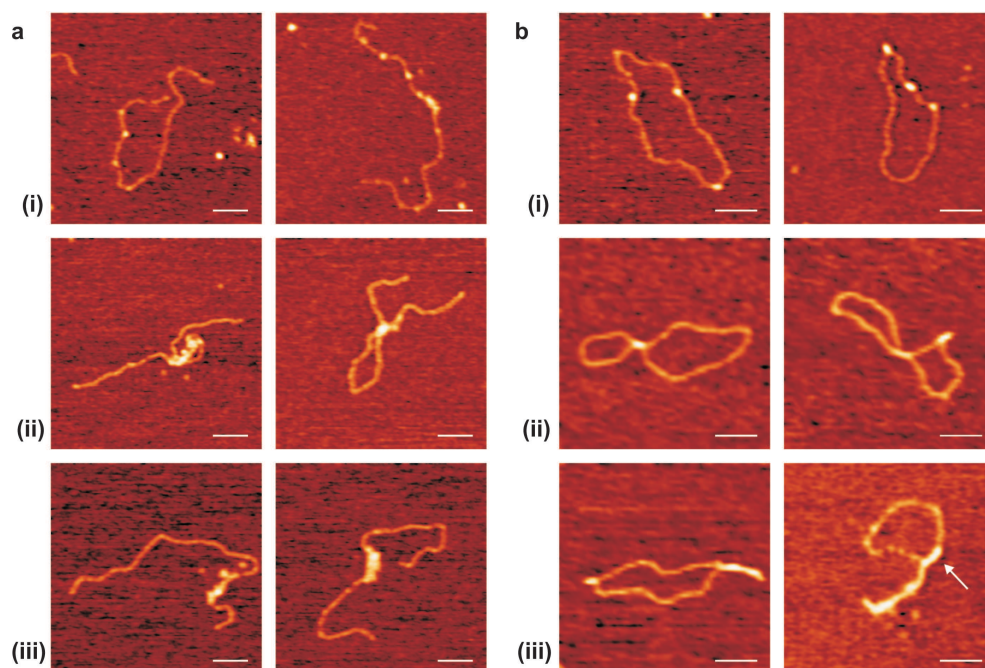


Figure 7.2: vTopIB-DNA complexes at low vTopIB:DNA values. The concentration of DNA is constant in all experiments and equal to 0.22 nM. (a), Gallery of AFM images of vTopIB bound to linear DNA. Three different types of complexes were found. a,*i*), individual vTopIB proteins bound to DNA ([vTopIB]=4nM); nodes (a,*ii*)), and filament-like structures (a,*iii*)) ([vTopIB]=6.2 nM). (b), Gallery of AFM images of vTopIB bound to nicked-circular DNA. Similar types of complexes were found. b,*i*), individual vTopIB-DNA complexes ([vTopIB]=4 nM); nodes (b,*ii*)), and filament-like structures (b,*iii*)) ([vTopIB]=13.4nM). Bar size is 100 nm.

three different types of binding events: (*i*) isolated binding, (*ii*) nodes, and (*iii*) filament-like structures. We confirmed that binding of individual vTopIB proteins to DNA did not cause wrapping of the DNA around vTopIB as evinced by the absence of shortening of the DNA contour length upon vTopIB binding. The filament-like structures shown in Fig. 8.4a,(*iii*) and Fig. 8.4b,(*iii*) are qualitatively different from each other. In the first case, the filament-like structures involved a single DNA duplex while in the second case, with circular DNA, the filament-like structures involved two segments of the circular DNA, hence two DNA duplexes. This conclusion is further corroborated by measurements of the contour length of the uncovered and the protein-covered parts of the DNA. For instance, the full contour length of the DNA molecule in Fig. 8.4b,(*iii*) is only recovered if the length of the filament is counted twice, clearly showing that two duplexes participate in that filament. A small fraction of filaments (less than 10%) involving a single DNA duplex were also found when using circular DNA

(see arrow in Fig. 8.4). Although a tendency towards type *(iii)* events was observed with increasing vTopIB:DNA ratio, type *(ii)* or *(iii)* events were also found at low vTopIB:DNA ratios. This suggests cooperative behavior in the binding of vTopIB to DNA [12], which we quantify below.

The heights of isolated DNA-bound proteins, node events, and filament-like structures were measured in air. In every experiment, we measured the height of the DNA (hDNA) as a reference height. This is especially relevant for AFM measurements conducted in air, as height measurements are then sensitive to experimental imaging conditions [29]. Values for hDNA measured in air ranged from 0.6-0.9 nm. Isolated DNA-binding events yielded a value for the height of a single DNA-bound protein of 2.1 ± 0.5 times hDNA (#events=127); node events, a value of 2.4 ± 0.5 times hDNA (#events=20); and filament-like structures, a value of 2.1 ± 0.6 times hDNA (#events=66). The height of individual DNA-free vTopIB molecules deposited on the mica surface was 1.4 ± 0.3 times hDNA (#events=63). The height distribution of isolated DNA-bound proteins consisted of a single peak in all tested experimental conditions (data not shown), consistent with the fact that vTopIB binds to DNA as a monomer [2].

Intermolecular synopsis of DNA molecules is found at high vTopIB:DNA ratios

Focusing on the protein filament-like structures, we increased the concentration of vTopIB per DNA molecule in order to see if longer protein filaments would result. Instead, this resulted in large clusters containing many DNA molecules (Fig. 8.5). Such intermolecular synapses were found on both linear and circular DNA in the presence of divalent ions such as Mg^{2+} , Mn^{2+} and Ca^{2+} (Fig. 8.5a). Qualitatively, we did not note any cation-specific effect. Hence, we focused the remainder of our study on the biologically more relevant Mg^{2+} . We quantified this phenomenon counting the number of DNA molecules involved in such clusters as a function of vTopIB concentration. Fig. 8.5b plots the analysis of intermolecular synopsis using ~ 1300 molecules, all adsorbed in the presence of MgCl_2 . The number of DNA molecules involved in a cluster increased monotonically with the vTopIB:DNA ratio (Fig. 8.5b). This trend saturated at a protein:DNA ratio of 125:1. We expect the number of DNA molecules involved in a cluster to be underestimated at such high vTopIB:DNA ratios for two reasons. First, the identification of the individual DNA molecules becomes progressively more difficult. Second, since the number of DNA molecules was constant in all experiments the clusters were found increasingly far from each other. This results in uncertainties in determining the average number of DNA molecules per cluster because very large areas must be scanned with the AFM to yield enough statistics.

We also explored the effect of divalent ions in the generation of intermolecular

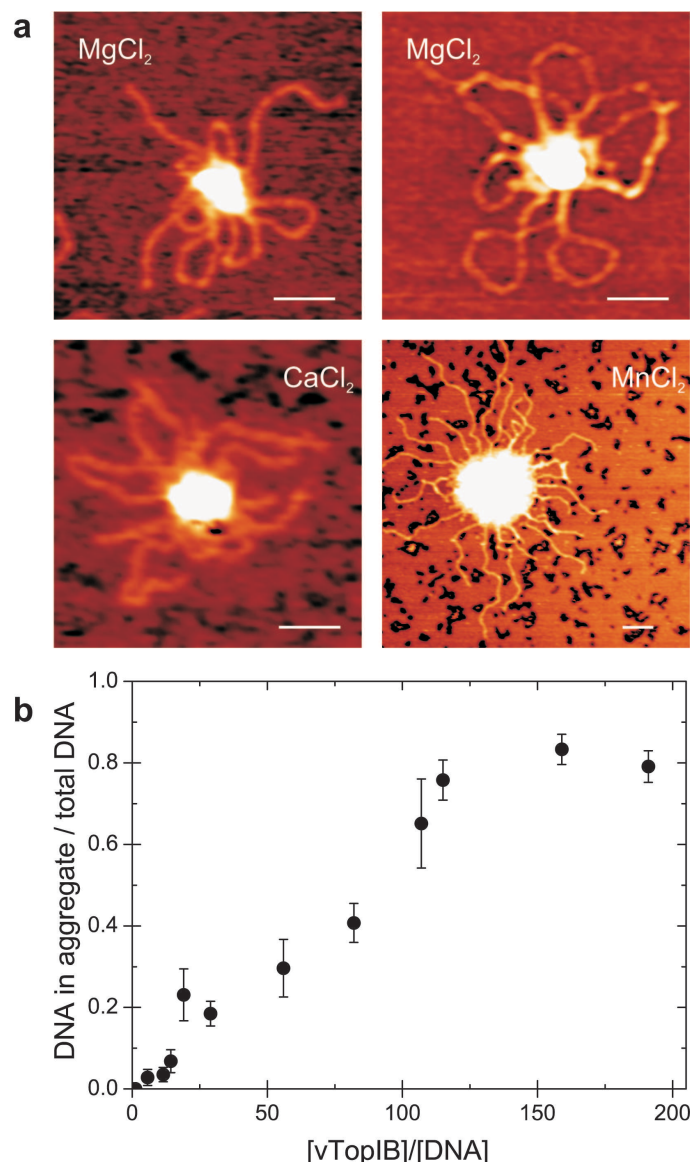


Figure 7.3: Intermolecular synopsis of DNA at high vTopIB:DNA values. The concentration of DNA is constant in all experiments and equal to 0.22 nM. (a), AFM images of linear and circular DNA molecules incubated with [vTopIB]=12.4 nM and [vTopIB]=45 nM, respectively, showed clustering in the presence of Mg^{2+} cations (top). This effect was not cation-specific, as incubation of linear DNA and 12.4 nM [vTopIB] together with Ca^{2+} or Mn^{2+} likewise resulted in the formation of aggregates. Bar size is 100 nm. (b), Intermolecular synopsis of DNA was quantified by counting the number of DNA molecules involved in an cluster at different vTopIB:DNA ratios in the presence of MgCl_2 . The fraction of DNA molecules in a cluster increased linearly with increasing vTopIB concentration and saturated at very high vTopIB concentrations. This saturation was due to an underestimation of the number of DNA molecules involved (see text for details)

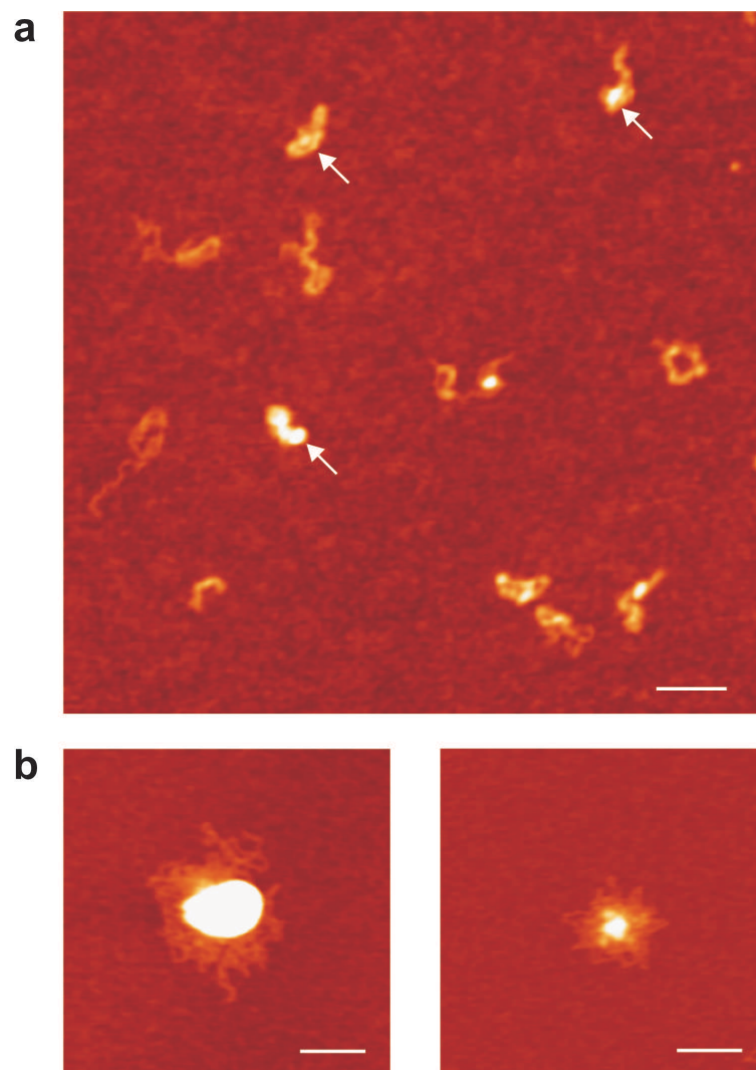


Figure 7.4: Intermolecular synapsis of DNA is affected by the presence of MgCl_2 . (a), Linear DNA molecules were deposited on polylysine-coated mica and incubated with vTopIB in the absence of MgCl_2 again resulting in the observation of vTopIB-bound DNA (white arrows). Under these conditions, however, the DNA molecules did not cluster. (b), Linear DNA molecules deposited on polylysine-coated mica showed intermolecular synapses when incubated with vTopIB and MgCl_2 . Bar size is 200 nm.

synapses. Typically, vTopIB-DNA binding reactions included 13 mM MgCl_2 in order to adhere the DNA to the mica (DNA does not adsorb on a bare mica surface). To isolate the effect of divalent ions, we instead treated the mica surface with polylysine, which is a polymer that positively charges the surface of the mica. Again we observed binding of vTopIB to DNA (Fig. 8.6a, white arrows). Surprisingly, in the absence of divalent ions DNA molecules were always isolated

from each other, even at high vTopIB concentrations (vTopIB:DNA=123:1). To conclude that the observation of isolated DNA molecules was not caused by the chemistry of mica adsorption, we verified that adding 13 mM MgCl₂ to the binding reaction and adsorbing the product on polylysine mica again resulted in the observation of DNA clusters (Fig. 8.6b). Note that the difference in the physical appearance of the molecules and clusters shown in Fig. 8.5 and in Fig. 8.6 is due to the different adsorption processes on magnesium-coated and polylysine-coated mica (see materials and methods).

vTopIB generates filaments on DNA in a cooperative fashion

We used the AFM images obtained to quantify the binding of vTopIB to DNA. The quantification of the binding of vTopIB to DNA that follows was done on circular DNA and restricted to low DNA saturation, $\theta \leq 0.25$. We used circular DNA to favor protein-protein interactions and low θ to avoid intermolecular synapsis of DNA.

We now describe the extension of vTopIB filaments on DNA using the terminology of Ref. [20], which is detailed in the materials and methods. The formation of vTopIB filaments is likely to happen in the following steps (Fig. 8.7). First, a single intramolecular node is formed (Fig. 8.7a,*i* and Fig. 8.7b,*i*). This intramolecular node may be composed of either a single vTopIB engaging two DNA duplexes or a dimer of vTopIB in which each monomer is bound to a single DNA duplex. We describe the association of a vTopIB monomer or dimer to the two DNA duplexes with the association constant K^* . Second, the filament is extended by the addition of new vTopIB protomers (Fig. 8.7, panel *ii*). The extension of the node by vTopIB has an association constant of $K^*\omega$, where ω is the cooperativity parameter for extending the filament. Values of ω larger than one indicate a cooperative process. Finally, a filament of length c is generated (Fig. 8.7a,*iii* and Fig. 8.7b,*iii*). Fig. 8.7 shows a general cartoon depicting the formation and extension of a filament by either vTopIB monomers or dimers. The theory used here is applicable to both cases (see Discussion).

In order to determine the association constant K^* and the cooperativity parameter ω of filament extension by vTopIB, we measured the length of filaments for several hundreds of DNA molecules at different vTopIB:DNA ratios. An average filament length was calculated and the DNA saturation θ was deduced. The distribution of filament lengths P_c is shown in Fig. 8.8a for three different DNA saturations ($\theta=0.07$, $\theta=0.18$, and $\theta=0.23$). For each data set, the cooperativity parameter ω was estimated using the ML method described in the Materials and Methods. The results are summarized in Table 1. Subsequently, Eq. 7.1 was plotted in Fig. 8.8a using the ω obtained for each data set. The error-weighted average cooperativity parameter was $7.7 \pm 1.1 \cdot 10^3$. This value shows the genera-

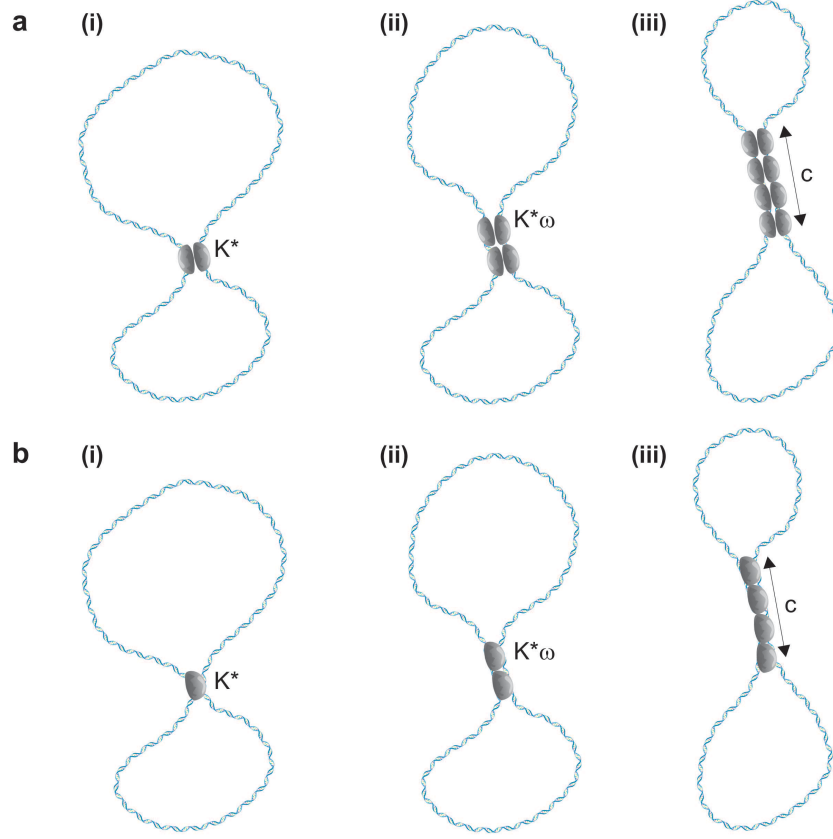


Figure 7.5: Model describing the formation of filaments of vTopIB on DNA. The constituents of the filaments may be vTopIB dimers (a) or monomers (b). The formation of filaments requires several steps. *i)*, The formation of an intramolecular node with an association constant K^* . *ii)*, The node is extended by binding of new vTopIB protomers with an association constant of $K^*\omega$. *iii)* A filament of length c is generated.

tion of filaments on DNA by vTopIB is highly cooperative.

We also estimated the binding constant K^* of the formation of a vTopIB-DNA node. Fig. 8.8b shows DNA saturations (θ) as a function of total vTopIB concentrations. This data was fitted to Eq. 7.4 using the least squares fitting method, yielding $K^* = 4.0 \pm 0.6 \cdot 10^4 \text{ M}^{-1}$ given $\omega = 7.7 \pm 1.1 \cdot 10^3$. The product $K^*\omega$ was found to equal $3.1 \pm 0.4 \cdot 10^8 \text{ M}^{-1}$.

7.4 Discussion

We quantified the generation of filaments of vTopIB on DNA using a theory of DNA binding [19, 20]. vTopIB was found to form filaments on DNA in a highly cooperative fashion. Our analysis gave a value of $7.7 \pm 1.1 \cdot 10^3$ for the cooperativ-

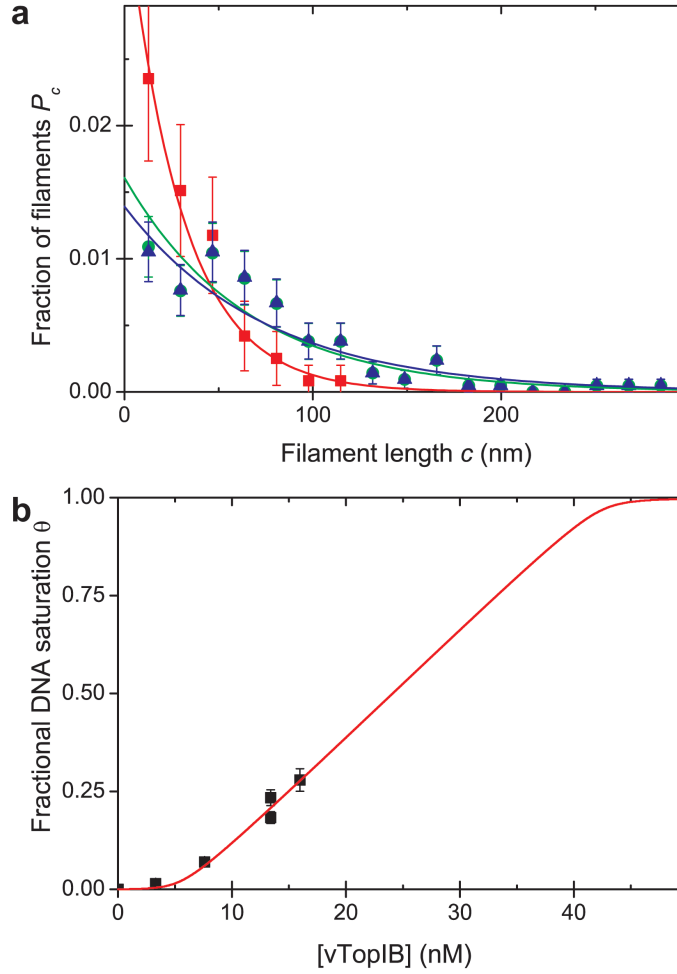


Figure 7.6: Quantitative analysis of the binding of vTopIB to DNA. (a), Filament length distribution for three different data sets. For each set ω was estimated using the Maximum-Likelihood method resulting in $\omega=5.6\pm1.4\cdot10^3$ for DNA saturation $\theta=0.07$ (red squares); $\omega=9.1\pm1.7\cdot10^3$ for DNA saturation $\theta=0.18$ (green circles); and $\omega=8.8\pm1.8\cdot10^3$ for DNA saturation $\theta=0.23$ (blue triangles). The solid lines in a are plots of Eq. 7.1 using the obtained ω . (b), Fractional DNA saturation as a function of total vTopIB concentration. The fit of Eq. 7.4 to these data using $n=18$, $N=2743$ and $\omega=7.7\pm1.1\cdot10^3$ yielded a value for $K^*=4.0\pm0.4\cdot10^4$ M $^{-1}$.

ity parameter ω , and $4.0\pm0.6\cdot10^4$ M $^{-1}$ for the association constant to generate a single intramolecular node K^* . Kowalczykowski *et al.* [19] defined sigmoidal (cooperative) filament growth by the criterion $\omega > 10n$. Our measurement of ω exceeds this threshold 40-fold. Our values are similar to those obtained for different cooperative binding proteins. For example, the gene 32 protein binds cooperatively to poly(rA) with parameters $\omega=1.2\cdot10^3$, $n=7.5$ and $K=9\cdot10^3$ M $^{-1}$ [30]; and The T4 UvsX protein forms filaments on ssDNA with parameters $\omega=1.0\cdot10^2$,

$n=4.0$ and $K=4.3 \cdot 10^4 \text{ M}^{-1}$ [31]. The values for K^* and ω quoted in this work may have a slightly different interpretation compared to other reported values. The reason for this is that in our work, the parameters K^* and ω relate to the respective formation and extension of a node (consisting of two monomers of vTopIB (Fig. 8.7a) or a single monomer (Fig. 8.7b)), whereas the parameters obtained in the binding of gene 32 protein and the T4 UvsX-protein relate to single monomers binding to their substrate.

We now discuss whether vTopIB monomers or dimers are the minimal constituents of a single intramolecular synapsis. AFM-volume analysis has been proven to be useful in distinguishing between proteins adsorbed as monomers or as dimers [32]. However, the volume analysis becomes less straightforward when proteins are small, as in the case of 36.7 kDa vTopIB, and becomes especially non-transparent when they are bound to DNA, because the proteins are of comparable height to the bare DNA. We found that, within the error, the heights of isolated-binding events and filament-like structures were identical and equal to 2.1 ± 0.5 times h_{DNA} . In the case of vTopIB binding to DNA, it was therefore impossible to unambiguously determine the number of vTopIB monomers that make up the intramolecular node, and hence we cannot rule out the possibility of a monomer of vTopIB being involved in the synapse of two DNA molecules.

The DNA aggregation found at high vTopIB:DNA ratios can be understood as an extension of the cooperative binding behavior presented above. Essentially, at similar protein:DNA ratios we found that intermolecular synapses are favored on linear DNA molecules while intramolecular synapses are favored in circular ones. This is unsurprising considering that the radius of gyration is larger for linear DNA than for circular DNA. Topological constraints of circular DNA could facilitate the propagation of a filament built up of vTopIB. Hence, interactions between two duplexes of the same molecule are more likely to occur in circular than in linear DNA. As expected, at very high protein-DNA ratios we also found clusters of circular DNA.

Interestingly, we found a correlation between DNA aggregation and the presence of the divalent ions Mg^{2+} , Ca^{2+} , or Mn^{2+} . The fact that the aggregation was not an effect only caused by one specific ion supports the notion that DNA aggregation is a more general effect. In general, divalent ions reduce the Debye length, which is the characteristic length scale of the electrostatic interaction between two charged objects in an electrolyte solution [33]. Since vTopIB does not bind divalent ions [34], their presence in the binding reaction has the effect of reducing the electrostatic repulsion between two DNA molecules and, as a consequence, in favoring interactions between two DNA-bound vTopIB molecules. Divalent ions consequently favor the formation of intermolecular synapsis.

Although our measurements provide clear evidence for cooperative extension of vTopIB filaments on DNA duplexes, the question remains what mechanism gives rise to the observed cooperativity. This question is especially intriguing given the fact that vTopIB is most likely a monomer in solution [6], in accordance with Fig. 8.3. The clustering reported here is strictly DNA mediated. We speculate that a conformational change in the vTopIB due to DNA binding may be responsible for the cooperativity observed here. Such conformational changes of proteins upon DNA binding have been previously reported [35, 36]. Specifically, dimerization of the DNA binding domains of the retinoid X receptor has been shown to be induced by DNA binding [37]. Similarly, a conformational change in vTopIB upon DNA binding may trigger multimerization of vTopIB. Such a conformational change in vTopIB may seem reasonable given its clamp-like structure and the presence of a hinge domain [2, 9].

Besides altering the number of supercoils in DNA molecules, Type IB topoisomerase have been shown to possess strand transferase activity that can result in DNA recombination [38–42]. Furthermore, Type IB topoisomerase and site-specific tyrosine recombinases have been shown to share structural features in their catalytic domains and descend from a common ancestor [2]. The vTopIB-mediated DNA synapsis shown and quantified here may be the first step in its strand transferase activity and we speculate that it may have been inherited from its ancestor. A filament structure built up of vTopIB dimers similar to those reported here could help to keep the two DNA duplexes to be joined in close proximity. In addition, our observation of vTopIB-mediated DNA condensation is consistent with the suggestion that this topoisomerase assists packaging the DNA into progeny virus particles [12], where it plays a role in the transcription of viral early genes [43].

In summary, we have presented the first quantitative AFM study of vTopIB binding to DNA. AFM images showed individual vTopIB proteins binding to DNA, single intramolecular synapsis and filament-like structures. vTopIB generated filament-like structures on individual DNA molecules at protein:DNA ratios between 10:1 and 60:1. We calculated the association constant of creating an intramolecular node by vTopIB and the cooperativity parameter from the filament-lengths distribution at different DNA saturations. From these values we conclude that the formation of filaments on DNA by vTopIB is a highly cooperative process. In addition, intermolecular synapses of DNA were observed at high vTopIB:DNA ratios. Interestingly, this interaction was promoted by divalent ions.

7.5 Acknowledgements

F.M-H is supported by a postdoctoral fellowship from La Fundacion Ramon Areces. We acknowledge Dick Korbbee for technical support in the mechanical and acoustical isolation of the AFM. C.D. and N.D. acknowledge financial support from FOM and NWO (Netherlands Organization for Scientific Research). S.S. acknowledges support from NIH grant GM46330.

References

- [1] Champoux, J. J. *Annu Rev Biochem* **70**, 369–413 (2001).
 - [2] Cheng, C., Kussie, P., Pavletich, N., and Shuman, S. *Cell* **92**(6), 841–50 (1998).
 - [3] Sekiguchi, J. and Shuman, S. *Nucleic Acids Res* **22**(24), 5360–5 (1994).
 - [4] Shuman, S. *J Biol Chem* **266**(17), 11372–9 (1991).
 - [5] Shuman, S. and Prescott, J. *J Biol Chem* **265**(29), 17826–36 (1990).
 - [6] Shaffer, R. and Traktman, P. *J Biol Chem* **262**(19), 9309–15 (1987).
 - [7] Shuman, S., Golder, M., and Moss, B. *J Biol Chem* **263**(31), 16401–7 (1988).
 - [8] Shuman, S. and Moss, B. *Proc Natl Acad Sci U S A* **84**(21), 7478–82 (1987).
 - [9] Sekiguchi, J. and Shuman, S. *J Biol Chem* **269**(50), 31731–4 (1994).
 - [10] Sharma, A., Hanai, R., and Mondragon, A. *Structure* **2**(8), 767–77 (1994).
 - [11] Cheng, C. and Shuman, S. *Biochemistry* **38**(50), 16599–612 (1999).
 - [12] Shuman, S., Bear, D. G., and Sekiguchi, J. *Embo J* **16**(21), 6584–9 (1997).
 - [13] Zechiedrich, E. L. and Osheroff, N. *Embo J* **9**(13), 4555–62 (1990).
 - [14] Binnig, G., Quate, C. F., and Gerber, C. *Phys. Rev. Lett.* **56**(9), 930–933. (1986).
 - [15] Hansma, H. G. *Annu. Rev. Phys. Chem.* **52**, 71–92 (2001).
 - [16] Moreno-Herrero, F., Herrero, P., Colchero, J., Baro, A. M., and Moreno, F. *Biochem. Biophys. Res. Commun.* **280**(1), 151–7. (2001).
 - [17] van Noort, S. J., van der Werf, K. O., Eker, A. P., Wyman, C., de Grooth, B. G., van Hulst, N. F., and Greve, J. *Biophys J* **74**(6), 2840–9. (1998).
 - [18] Viani, M. B., Pietrasanta, L. I., Thompson, J. B., Chand, A., Gebeshuber, I. C., Kindt, J. H., Richter, M., Hansma, H. G., and Hansma, P. K. *Nat. Struct. Biol.* **7**(8), 644–7. (2000).
-

- [19] Kowalczykowski, S. C., Paul, L. S., Lonberg, N., Newport, J. W., McSwiggen, J. A., and von Hippel, P. H. *Biochemistry* **25**(6), 1226–40 (1986).
 - [20] McGhee, J. and von Hippel, P. *J. Mol. Biol.* **86**(2), 469–89 (1974).
 - [21] Bussiek, M., Mucke, N., and Langowski, J. *Nucleic Acids Res* **31**(22), e137 (2003).
 - [22] Williams, R. C. *Proc Natl Acad Sci U S A* **74**(6), 2311–5 (1977).
 - [23] Rivetti, C., Guthold, M., and Bustamante, C. *J Mol Biol* **264**(5), 919–32. (1996).
 - [24] Crampton, N., Bonass, W. A., Kirkham, J., and Thomson, N. H. *Langmuir* **21**(17), 7884–91 (2005).
 - [25] Fang, Y. and Hoh, J. H. *FEBS Lett* **459**(2), 173–6 (1999).
 - [26] Bustamante, C., Vesenska, J., Tang, C. L., Rees, W., Guthold, M., and Keller, R. *Biochemistry* **31**(1), 22–6. (1992).
 - [27] Chatfield, C. *Statistics for technology*. Chapman & Hall, 3rd edition edition, (1996).
 - [28] Brock, I. *Statistical methods of data analysis*. (2000).
 - [29] Jiao, Y. and Schaffer, T. E. *Langmuir* **20**(23), 10038–45 (2004).
 - [30] Lonberg, N., Kowalczykowski, S. C., Paul, L. S., and von Hippel, P. H. *J Mol Biol* **145**(1), 123–38 (1981).
 - [31] Ando, R. A. and Morrical, S. W. *J Mol Biol* **283**(4), 785–96 (1998).
 - [32] Ratcliff, G. C. and Erie, D. A. *J Am Chem Soc* **123**(24), 5632–5 (2001).
 - [33] Israelachvili, J. *Intermolecular and Surface Forces*. Academic Press, London, (1991).
 - [34] Stivers, J. T., Shuman, S., and Mildvan, A. S. *Biochemistry* **33**(1), 327–39 (1994).
 - [35] Boskovic, J., Rivera-Calzada, A., Maman, J. D., Chacon, P., Willison, K. R., Pearl, L. H., and Llorca, O. *Embo J* **22**(21), 5875–82 (2003).
 - [36] Moreno-Herrero, F., de Jager, M., Dekker, N. H., Kanaar, R., Wyman, C., and Dekker, C. *Nature* **437**(7057), 440–3 (2005).
 - [37] Holmbeck, S. M., Dyson, H. J., and Wright, P. E. *J Mol Biol* **284**(3), 533–9 (1998).
 - [38] Christiansen, K. and Westergaard, O. *J Biol Chem* **269**(1), 721–9 (1994).
 - [39] Shuman, S. *Proc Natl Acad Sci U S A* **86**(10), 3489–93 (1989).
-

-
- [40] Shuman, S. *Proc Natl Acad Sci U S A* **88**(22), 10104–8 (1991).
 - [41] Shuman, S. *J Biol Chem* **267**(12), 8620–7 (1992).
 - [42] Shuman, S. *J Biol Chem* **267**(24), 16755–8 (1992).
 - [43] Da Fonseca, F. and Moss, B. *Proc Natl Acad Sci U S A* **100**(20), 11291–6 (2003).
-

Chapter 8

Coated nanofabricated structures for the single-molecule study of DNA polymerizing enzymes

This chapter describes the nanofabrication, coating and characterization of "zero-mode" waveguides (ZMW) for the single-molecule study of DNA polymerizing enzymes, such as telomerase and DNA polymerase. These enzymes have in common that they incorporate deoxyribose nucleic acid (DNA) building blocks, nucleotides, into a pre-existing DNA molecule. The aim is to monitor the incorporation of fluorescently-labeled nucleotides in real-time and thus gain insight into the mechanisms and kinetics of these enzymes. The experimental strategy is designed to satisfy two conflicting requirements of this experiment: 1. to monitor the incorporation at the single-nucleotide level and 2. to perform the experiment at relatively high and biologically relevant nucleotide concentrations.

8.1 Single-molecule fluorescence at high concentrations. Why?

Setting up an experimental apparatus to monitor fluorescence molecules with single-molecule sensitivity at high bulk concentration is technically challenging, but can be justified by considering typical K_M -values of enzymes, the Michaelis-Menten constant. Enzymes display Michaelis-Menten kinetics and the relation between the concentration of enzyme substrate, $[S]$ in units of mole/l, turnover velocity, v in units of mole/min, maximum turnover velocity, v_{max} , in units of mole/min and K_M , in units of mole/l is given by:

$$v = \frac{v_{max}[S]}{K_M + [S]}. \quad (8.1)$$

The value of K_M thus represents the substrate concentration at which the enzyme catalyzes the chemical reaction with $v = \frac{1}{2}v_{max}$. A survey of the K_M values of 30,000 enzymes shows that $\sim 100 \mu\text{M}$ is the expectation value, while the distribution decays off to 0 at the lower end of K_M values between roughly 1 and 10 nM [1]. For effective catalysis and processivity, one is required to work at a nucleotide concentration that is comparable to the K_M -value. When setting the goal to study these reactions at the single nucleotide level using fluorescence techniques, upper limits are set for the observation volume, *i.e.*, the volume from which the fluorescent light, emitted by the nucleotides, is collected. A K_M -value of $100 \mu\text{M}$ implies an observation volume of app. $1.6 \cdot 10^{-20}$ liter, while a K_M -value of 10 nM implies an observation volume of app. $1.6 \cdot 10^{-16}$ liter. Widely used fluorescence techniques, such as confocal, or total internal reflection microscopy are characterized by observation volumes that are too large for detecting single molecules at K_M -concentrations [2]. For example, a focussed laser beam has diffraction limited dimensions of app. 500 nm in the (x, y) -plane, and 1 μm in the z -direction [3], depending on the wavelength of the light, and thus has a typical volume of $\sim 1 \cdot 10^{-12}$ L. Total internal reflection [2] comes closer, but is still diffraction limited in the (x, y) -plane, *i.e.* app. 500 nm and evanescent, *i.e.* order 100 nm in the z -direction, constituting an effective volume of 10^{-16} L, or 0.1 fL, reasonable for the study of enzymes that have very low K_M -values. Most enzymes, including that of telomerase (K_M -value for a single nucleotide is order 1 μM [4]), can thus not be studied at the single-nucleotide level at nucleotide concentrations equal to their K_M value.

To further reduce the observation volume, we follow the lead of Levine *et al.* [5], who have previously reported the concept of "zero-mode" waveguides. They

consist of nanofabricated holes in a 100 nm-thin metal layer, which is mounted on a fused silica microscope coverslip, see Fig. 8.1. The reduction of excitation volume is a result of two different processes. First, as the aperture of the ZMW is significantly smaller than the wavelength of the light (see below), light will not propagate through the ZMW. Instead, an evanescent field is generated whose intensity decays off exponentially on a length scale that is set by the dimensions of the guide, but is typically much less than the thickness of the metal (*e.g.* ~ 20 nm for a 25 nm radius ZMW, see below). Thus, in comparison to propagating light, the observation volume is reduced in the z -direction by virtue of the length scale over which the evanescent field decays off. Second, in the (x, y) -plane, the dimensions are not set by the diffraction limited size of the laser spot, but rather by the dimensions of the ZMW, that are smaller by roughly a factor of 10 in each dimension. Typically, ZMWs thus define an effective observation volume of tens of zls ($1 \text{ zl} = 10^{-21}$, which facilitates the measurement of less than a single fluorophore at a bulk concentration of order $10 \mu\text{M}$, well in range of the bulk of K_M -values, including that of telomerase [4].

8.2 Setup of the experiment

We now describe the use of zero-mode waveguides to measure enzymatic activity of DNA polymerizing enzymes, including telomerase [6–8]. The idea is to anchor the substrate for the enzyme, a single stranded piece of DNA, at the bottom of the ZMW. Freely diffusing telomerase can then bind to the free extremity of the single strand. Using an RNA template it bears within itself (telomerase is a ribonucleoprotein) [8], it extends the single DNA strand using nucleotides that also freely diffuse in solution. Human telomerase will repeatedly synthesize a tandem repeat of 6 nucleotides, whose sequence is determined by the RNA part (human telomerase synthesizes the hexanucleotide sequence TTAGGG) [6]. As such, for each hexanucleotide repeat, there will be an incorporation of a single adenine nucleotide. Telomerase will be provided with this nucleotide in modified form, *i.e.* labeled with the fluorescent molecule tetramethylrhodamine (TMR). As described above, the dimensions of the excitation volume created inside the ZMW are such that on average there will be less than a single fluorescent nucleotide in the volume. A schematic picture of the ZMW with a tethered ssDNA and telomerase is shown in Fig. 8.1. Furthermore, the diffusion time of a single nucleotide is on the order of a μs . This can be calculated using a rough estimate for the size of TMR-dATP (radius R taken as 1 nm) diffusing over a distance x of 50 nm. The diffusion constant D for such a molecule can thus

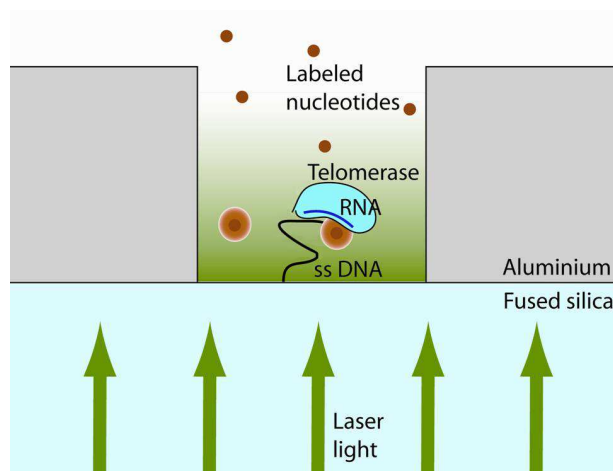


Figure 8.1: The use of zero-mode waveguides (ZMW) to monitor the activity of telomerase at the single-molecule level and in real time. ZMWs are small cavities in an aluminum film, on top of a fused silica microscope coverslip. Because of the small dimensions of the guide (see text), an evanescent field is generated inside the guide, when a laser shines through the coverslip onto the guide. The evanescent field decays exponentially with increasing height and is shown here as the gradient inside the guide. Fluorescently labeled nucleotides, which will be incorporated into the DNA are only excited in the very small excitation volume created by the waveguide. When telomerase binds a fluorescent nucleotide for incorporation, the nucleotide will presumably reside inside the excitation volume for a time period that is longer than the time that it spends to diffuse through the volume. The fluorescence intensity as a function of time is collected and is thus reflects the kinetics of incorporation of the nucleotides by telomerase.

be approximated by $D = k_B T / 6\pi\eta R \approx 2 \cdot 10^{-10} \text{ m}^2/\text{s}$. The diffusion time t is then $t = x^2 / 2D \approx 5\mu\text{s}$, in agreement with fluorescence correlation spectroscopy measurements [5]. We constantly collect the fluorescence originating from the excitation volume either using an electronmultiplication CCD (EMCCD) or an avalanche photodiode (APD). The first method has the advantage of capturing multiple ZMWs in parallel, but has the disadvantage of being slow (roughly an integration time of 2 ms is required to collect enough photon statistics to identify a single fluorescent molecule, given experimentally realistic excitation intensities and full frame, 512x512 capture without binning). The APD limits the detection of fluorescence from a single ZMW, but has the advantage of being fast enough to characterize the timescales associated with the diffusion of a single dye through the ZMW. The experiment relies on a separation between the timescale for a freely diffusion labeled nucleotide to traverse the ZMW and the timescale at which that the telomerase incorporates the nucleotide. When the latter timescale is larger

than the former (*e.g.* T7 DNA polymerase incorporates nucleotides at roughly 330 bp/s [9]), monitoring the fluorescence intensity in time will allow the monitoring of the incorporation kinetics of telomerase in real-time. We are initially interested to study the timescales at which telomerase moves from repeat to repeat, which is why we have chosen to label adenine with TMR as it is only incorporated once per hexanucleotide repeat. Furthermore, we have obtained adenines that have a TMR connected to the γ -phosphate, instead of the more common base attachment (nucleotides kindly provided by Susan Hardin from Visigen Biotechnologies Inc.). The advantage of this particular place of attachment in the context of our experiment is that upon incorporation, the diphosphate is released into solution together with the TMR, after which it rapidly diffuses out of the excitation volume. As such, the fluorescence intensity will not build up in time nor change stochastically due to blinking and/or bleaching. We envision that a signal will thus be generated that is far less prone to misinterpretation.

A particularly pressing issue is whether telomerase is capable of incorporating these γ -phosphate-labeled nucleotides, as the experiment crucially depends on it. The activity of telomerase can be measured at the ensemble level using the telomeric repeat amplification protocol (TRAP) assay, in which telomerase is first allowed to extend the single-stranded part by synthesizing multiple hexanucleotide repeats [10]. These repeats are subsequently amplified using traditional polymerase chain reaction (PCR). The reaction products are run on a gel and separated by length using gel electrophoresis and differ in length by six basepairs. Fig. 8.2 shows such a gel with TRAP assay product, showing that the products of a telomerase reaction containing unmodified nucleotides are indistinguishable from products of a telomerase reaction in which γ -phosphate-labeled adenines were used. We thus conclude that telomerase is able to successfully incorporate these modified nucleotides.

8.3 Finite-element simulations of zero-mode waveguides

To investigate how light interacts with ZMWs, we solved Maxwell's wave equations with the specific boundary conditions set by the geometry of the ZMW, *i.e.* an Aluminum layer of thickness and refractive index d_{Al} and n_{Al} respectively and with a square water-filled hole of width w_{ZMW} . The solution of the Maxwell equations yield *e.g.* the light intensity inside the ZMW (Fig. 8.3a). We also plot how the intensity of the field decays along the center z -line of the ZMW, as is shown in Fig. 8.3b. For the particular geometry of ZMW used here, we observe

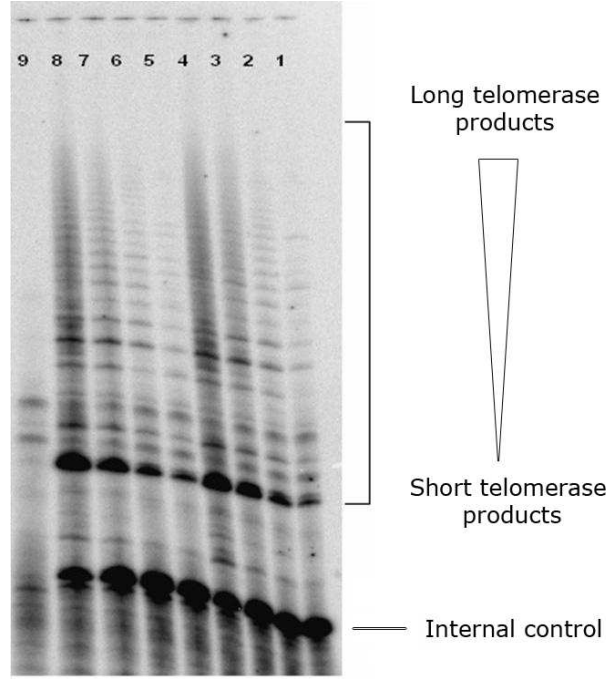


Figure 8.2: Human telomerase is capable of incorporating γ -phosphate-labeled adenines. Lanes 1-4 represent positive controls, containing TRAP assay products from a reaction with unmodified nucleotides at various concentrations. Lanes 5-8 contain TRAP products of a reaction in which γ -phosphate adenines are used. Lane 9 is the negative control as no dATP is added to the reaction mix. TRAP assay and gel characterization performed by Dr. Jue Lin in the Blackburn group at UCSF.

that the light intensity decays exponentially with a characteristic length scale of ~ 20 nm.

ZMW thus create a sufficiently small excitation volume to contain less than a single fluorescent nucleotide at high (*e.g.* ~ 10 μ M) concentration, Fig. 8.4a. The signal that is collected during the experiment is ideally limited to the fluorescence light that is generated inside the ZMW only (denoted "signal" in Fig. 8.4b). However, the bulk solution above the ZMW however contains a very large number of fluorescent molecules, that potentially contaminates the signal the detector collects (denotes "Background noise" in Fig. 8.4c).

In this context it is of importance to quantitatively assess the residual light intensity at the top of the ZMW, as this light is available to excite the bulk solution above the ZMW. We define a transmission factor TF as

$$TF = \frac{I(z = d_{Al})}{I(z = 0)}, \quad (8.2)$$

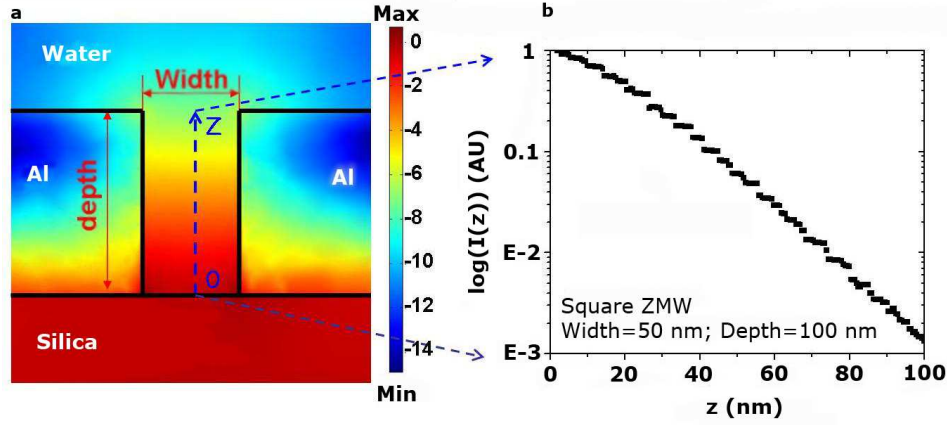


Figure 8.3: Finite-element simulations of the light intensity inside a ZMW. (a) shows the ZMW, consisting of a glass coverslip (silica), on top of which is a 100 nm-thick Aluminum layer, with a water-filled hole in it that is 50 nm in width. (b) The light intensity $I(z)$ as a function of the z -axis, where 0 represents the bottom of the ZMW. $I(z)$ decays off exponentially with z and has (for this particular geometry) a characteristic length scale of ~ 20 nm.

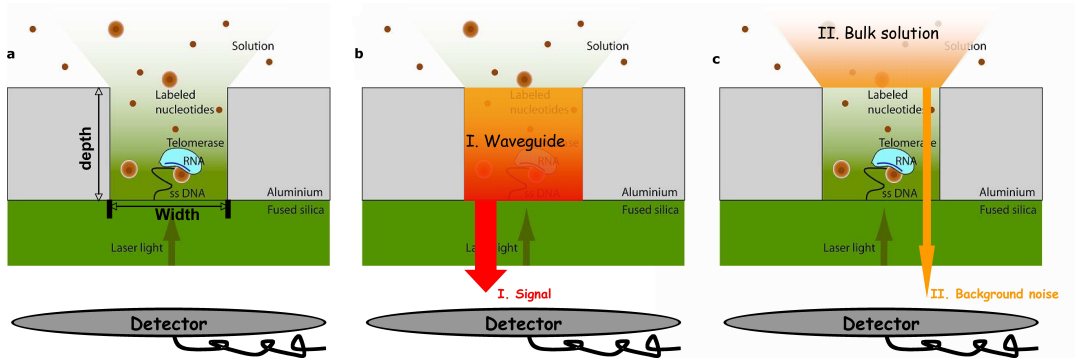


Figure 8.4: Various contributions to the observed fluorescence signal collected by the detector. The system (a) can be divided up into two regimes: the region inside the excitation volume, the fluorescence from which is called "signal" (b) and the region above the excitation volume, primarily the bulk solution, whose fluorescence contribution is considered background noise (c). In order to evaluate the contribution from the bulk solution, we simulate the light intensity inside the ZMW, as described in the text.

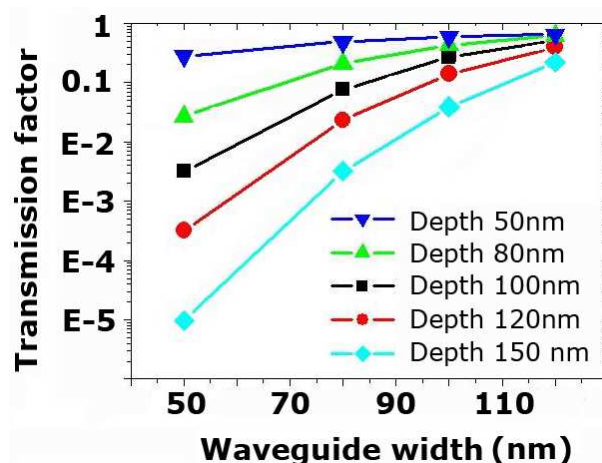


Figure 8.5: Transmission properties of ZMWs as a function of their width, plotted for a variety of depths.

where $I(z = d_{Al})$ is the light intensity at the top of the ZMW and $I(z = 0)$ is the intensity at the bottom. We can expect TF to be a function of both d_{Al} and the width of the ZMW, as the latter sets the decay length of the light intensity inside the ZMW and d_{Al} defines at what point along the z -axis the light intensity is probed. Fig. 8.5 shows how TF depends on both parameters. From the onset, we made an effort to design and fabricate ZMWs with the smallest possible width, as smaller guides obviously yield smaller excitation volumes. The issue of transmission, however, only came to light in a quantitative fashion by these simulations. Based on these simulations, we chose to increase the depth of the guides from 100 nm to 120 nm, roughly further attenuating the transmission by an order of magnitude. Further increasing the depth seems a natural and obvious choice, but nanofabrication considerations, most notably the difficulties stemming from the increased aspect ratio of the guides, made us decide on 120 nm. Furthermore, it should be noted that these curves report on the efficiency of transmitting excitation laser light through the guides, while contaminating background fluorescent light originates from light emitted by molecules in the solution above the guides. These molecules could be excited by the transmitted light, but the emitted fluorescence would have to couple back into the guides and undergo a similar attenuation due to the evanescent field. Although we have not performed simulations on this reverse coupling, given the magnitude of the transmission factor, we are confident that the relatively weak fluorescence light from the bulk molecules would not overwhelm the intensity of even a single molecule inside the guide.

8.4 Nanofabrication of zero-mode waveguides

Zero-mode waveguides were fabricated in the clean room using e-beam lithography and reactive ion (dry) etching as tools to respectively define and transfer the pattern into the metal surface [5]. Starting from a cleaned fused silica coverslip, a 120 nm layer of aluminum was thermally evaporated, after which a layer of positive tone resist was spun on the sample. The e-beam was used to define an array of dots of various sizes. The spacing between the dots was chosen as 5 μm , as is rationalized below. After development, a mixture of BCl_3 , Cl_2 and N_2 gasses was used to etch through the aluminum anisotropically. The resist was removed by a combination of UV light treatment and subsequent organic solvent after which the guides were ready for use. The etching process yielded waveguides that could be used for testing. However, the time required to fully etch through the aluminum was found to depend sensitively on the width of the guide: smaller guides etch slower due to the lower number of ionized BCl_3 molecules that hit the bottom of the narrow pit in the aluminum where the reaction ought to take place. Consequently, the relative etch rate of the pit in the lateral directions increases with decreasing width, leading to guides that are somewhat larger than defined by e-beam lithography. Also, the width of the guides varies, judging from the fluorescence intensity of freely diffusing dyes, and AFM microscopy (data not shown).

Because of these issues, we have embarked on an alternative route for the fabrication of guides, that involves the e-beam definition of pillars made of negative tone resist directly on the glass surface. In a next step, a 120 nm thick layer of metal film is evaporated, after which the resist pillars are developed. The guides are then ready for use. The fabrication process is simpler in theory, mainly because it contains fewer steps and does not contain etching. However, the accurate fabrication of thin and high pillars on glass is non-trivial and requires further development. Therefore, in this chapter, we will use waveguides that were fabricated using the etch technique.

8.5 Characterization of waveguides

We will now describe the characterization of the ZMWs using atomic force spectroscopy (AFM) and optical microscopy, with the aim of verifying their correct fabrication and expected optical properties. Fig. 8.6, upper row shows an optical microscope image in transmission mode of an array of ZMWs for 500, 200 and 100 nm in width. We have chosen to create an array spacing of 5 μm with several considerations in mind. First, the guides should not be spaced by a distance smaller

than the Rayleigh criterion, *i.e.* ~ 230 nm. After all, when two guides are separated by a distance smaller than the Rayleigh criterion, one cannot distinguish them from one another. Second, arrays of guides can exhibit enhanced transmission with a power that is a function of the waveguide spacing [11]. Clearly, enhanced transmission needs to be avoided, as the objective of using ZMWs is *not* to have transmitting light. However, above roughly $1\text{ }\mu\text{m}$ spacing, one does not expect significant contributions of this effect. Third, the stage on top of which the waveguide array is mounted on the microscope is driven by a stepper motor, that needs to be able to position the objective carefully over a single waveguide, in the case of collecting the fluorescence signal from a single guide. The guide spacing thus ought to be compatible with the absolute accuracy of the stepper motor, which is on the order of a few hundred nm after a travel distance of about $1\text{ }\mu\text{m}$, given the software optimization routine currently used. Thus, these considerations dictate a waveguide spacing of over $1\text{ }\mu\text{m}$, and we chose $5\text{ }\mu\text{m}$ for initial testing. From the top row of Fig. 8.6 it is apparent that the transmission decreases with decreasing waveguide width, as expected. The 100 nm guides do not show significant transmission, although these pictures do not quantitatively answer the question whether the transmission is sufficiently low as to not generate a significant background fluorescence signal during the experiment. Note that the bright spots visible also in the smallest guides are due to defects and/or pinholes in the metal that are primarily generated in the fabrication process. They are, however, not dense enough to disturb the measurement and therefore, we have not yet optimized the fabrication process to avoid the creation of these artifacts. The middle row in Fig. 8.6 shows AFM images of the waveguides. Note that for the widths shown here, the AFM tip readily probes the (flat) bottom fused silica of the guides and thus, we can establish that the waveguides were successfully etched through. The thickness of the metal layer was app. 120 nm , as expected and the width of the small guides was slightly larger than planned, as is expected from the particular process used to fabricate them, see above.

8.6 PEG coating of glass surfaces

To reduce the non-specific adsorption of fluorescent nucleotides and enzyme inside the waveguides and on the large metal surface outside the waveguides, we chose to coat these surfaces with polyethyleneglycol (PEG), a technique widely used to achieve these goals [12]. Fig. 8.7, top row, shows optical microscopy images of the fluorescence originating from fluorescently labeled dATP nucleotides, that have the fluorescent group (tetramethyl rhodamine, TMR) covalently attached to

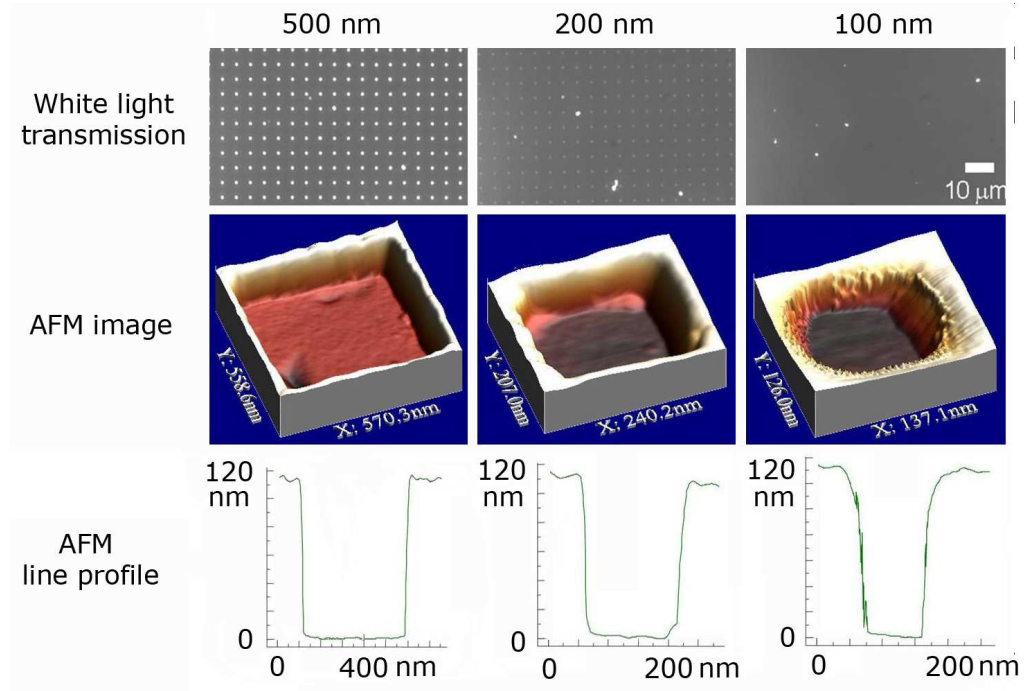


Figure 8.6: Transmission optical microscopy and AFM characterization of arrays of zero-mode waveguides of various sizes. Upper row shows optical transmission images, middle row shows AFM images and bottom row shows AFM line scans through the waveguides, see text.

the γ -phosphate. One notices that the modified nucleotides (the experimentally realistic concentration of $1 \mu\text{M}$ was chosen) readily adsorb on clean glass, even after rinsing with water, as shown by the bright fluorescence signal observed (left image). In contrast, a PEG coated glass surface shows much reduced adsorption (middle image, top row). A zoom (right image, top row) shows the adsorption of single labeled nucleotides on the surface, whose density is ~ 1 nucleotide per μm^2 . Comparing this quantity with the area of a single waveguide ($\sim 0.01 \mu\text{m}^2$) demonstrates that PEG-coating of the surface reduces the non-specific adsorption of fluorescent nucleotides to an acceptable level. The PEG surfaces are also effective at repelling proteins [12], as demonstrated by AFM measurements shown in Fig. 8.7, bottom row. The left image shows the non-specific adsorption of human telomerase enzymes (white spots) on clean glass, while the right image shows that no such non-specific adsorption occurs for PEG-coated glass. Note that glass naturally contains holes, shown as black dots on the AFM images.

Aluminum, when exposed to air, has a "native" oxide (alumina), that enables PEGylation in an identical manner compared to glass. Indeed, the results shown above for glass were reproduced for an aluminum film [11].

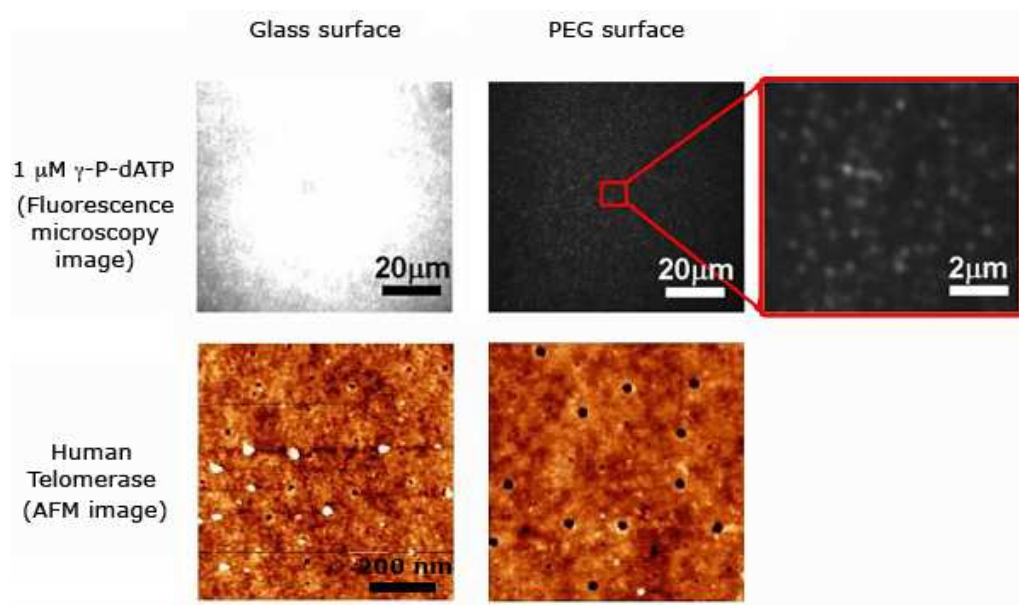


Figure 8.7: Optical microscopy and AFM images showing that PEG coating of glass surface reduces the non-specific adsorption of fluorescently labeled nucleotides and proteins to acceptable levels, see text.

8.7 Biotinylated PEG specifically binds streptavidin on a surface

The next step is to show that biotinylated PEG surfaces are capable of specifically binding streptavidin. Fig. 8.8 shows a series of AFM images of non-fluorescent streptavidin proteins as function of the percentage of biotinylated PEG used to coat the surface of glass. The left image (Fig. 8.8a) shows that no streptavidin binds to non-biotinylated PEG, as expected. However, as the fraction of biotinylated/non-biotinylated PEG increases (middle image (b) shows 5%, right image (c) shows 33%), the observed density of streptavidins on the surface increases. At 33% biotinylated PEG, we observe roughly 400 adsorbed streptavidins per 200 nm^2 , *i.e.* 2 streptavidins/ nm^2 , which is a number higher than 1 streptavidin per waveguide area. These experiments demonstrate that we can tune the surface density of specifically bound streptavidins in the range of 1 streptavidin per waveguide area.

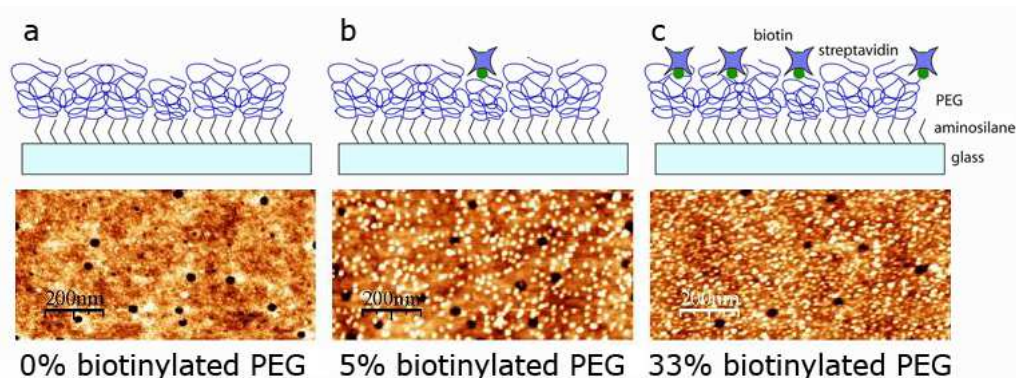


Figure 8.8: AFM images of streptavidin proteins specifically adsorbed on PEGylated surfaces that contains various (0% (a), 5% (b) and 33% (c)) concentrations of biotinylated PEG. Using varying concentrations of biotinylated PEG, we can thus tune the surface density of specifically adsorbed streptavidin.

8.8 DNA polymerase activity on PEG surface

We now turn to monitoring enzymatic activity on PEG-coated surfaces. We initially choose not to work with telomerase, but rather use DNA polymerase, as the latter enzyme is readily commercially available in high concentrations. This in contrast with telomerase, whose purification is a time-consuming endeavor that ultimately does not nearly yield the same quantity of active enzyme. To test DNA polymerase, we specifically anchored its DNA substrate, a primed DNA molecule to a PEG-coated surface, as shown in Fig. 8.9(a). The template DNA is attached by its single-stranded 5'-end to the surface through a biotin. A primed region at the other end of the DNA is created by hybridizing a primer. DNA polymerase will then start to synthesize the complementary strand, provided that all the required nucleotides are present in solution. The template DNA strand contains the following sequence: A(TCATCG)₁₀-5'-biotin and therefore a maximum of 10 fluorescent TMR-dCTP molecules can be incorporated by DNA polymerase (Fig. 8.9b schematically shows the incorporation of only 5 TMR-dCTP nucleotides). After the attachment of the DNA substrate (10 nM) to the biotinylated (5%) PEG surface, Klenow fragment DNA polymerase was added. After 90 minutes of incubation, the fluorescence signal was collected and is shown in Fig. 8.9c. For comparison, the same reaction was performed using heat-inactivated Klenow fragment, in which case it is expected that no TMR-dCTP will be incorporated (schematically shown in Fig. 8.9d). The fluorescence signal collected after this negative control experiment is shown in Fig. 8.9e. We conclude that Klenow fragment was not only active on the prepared surfaces, but was also able to

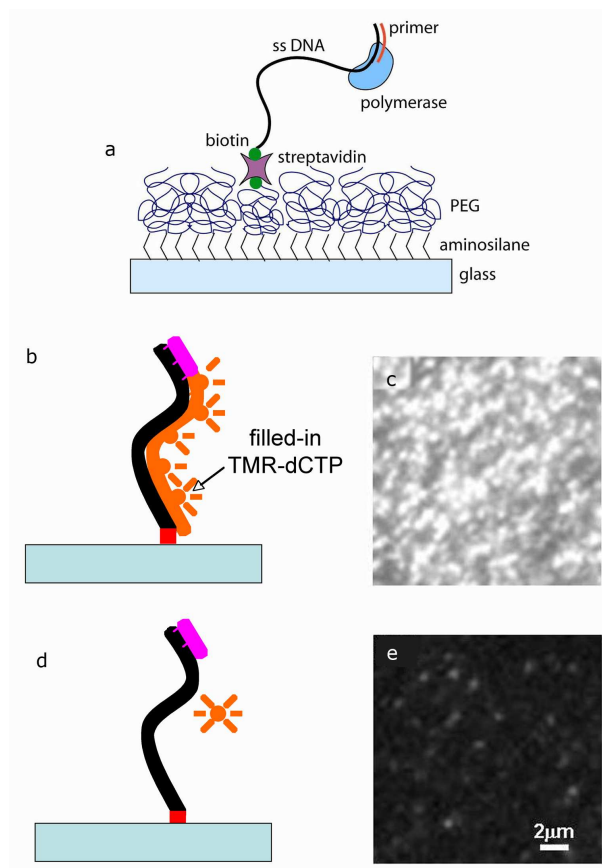


Figure 8.9: Activity of Klenow fragment DNA polymerase on a PEGylated surface. (a) shows the design of the experiment, in which a biotinylated primer DNA substrate is linked through streptavidin to a biotinylated PEG-coated glass surface. The polymerase will incorporate the nucleotides present in solution, among which TMR-labeled dCTP's, of which the polymerase can maximally incorporate 10 given the sequence of the template DNA strand, see text. Here only 5 are schematically drawn (b). After 90 minutes of incubation, the fluorescence from the sample is collected (c). As a negative control, the reaction was repeated, except for that the enzyme was heat inactivated (d). After the same amount of incubation time, the fluorescence was collected and is shown in (e).

incorporate TMR-dCTP in this geometry.

8.9 Acknowledgements

Thanks goes to Zhuangxiong Hang who performed many of the PEGylation experiments, as well as the Femlab simulations. Thanks also goes to the other member of the telomerase-team: Aurélien W. Crut. Many thanks as well to the

entire DIMES clean-room staff, most notably Marc Zuiddam and Marco van der Krogt for help with the various machines.

References

- [1] Samiee, K. T., Foquet, M., Guo, L., Cox, E. C., and Craighead, H. G. *Biophys J* **88**(3), 2145–53 (2005).
 - [2] Moerner, W. E. and Fromm, D. P. *Review Of Scientific Instruments* **74**(8), 3597–3619 (2003).
 - [3] Keyser, U. F., Krapf, D., Koeleman, B. N., Smeets, R. M., Dekker, N. H., and Dekker, C. *Nano Lett* **5**(11), 2253–6 (2005).
 - [4] Lee, M. S. and Blackburn, E. H. *Mol Cell Biol* **13**(10), 6586–99 (1993).
 - [5] Levene, M. J., Korlach, J., Turner, S. W., Foquet, M., Craighead, H. G., and Webb, W. W. *Science* **299**(5607), 682–6 (2003).
 - [6] Blackburn, E. H. *Annu Rev Biochem* **61**, 113–29 (1992).
 - [7] Greider, C. W. and Blackburn, E. H. *Cell* **43**(2 Pt 1), 405–13 (1985).
 - [8] Greider, C. W. and Blackburn, E. H. *Cell* **51**(6), 887–98 (1987).
 - [9] Tabor, S., Huber, H. E., and Richardson, C. C. *J Biol Chem* **262**(33), 16212–23 (1987).
 - [10] Kim, N. W., Piatyszek, M. A., Prowse, K. R., Harley, C. B., West, M. D., Ho, P. L., Coviello, G. M., Wright, W. E., Weinrich, S. L., and Shay, J. W. *Science* **266**(5193), 2011–5 (1994).
 - [11] Huang, Z. *Towards single-molecule studies of telomerase kinetics in zero-mode waveguides*. PhD thesis, Delft University of Technology, (2007).
 - [12] Rasnik, I., McKinney, S. A., and Ha, T. *Acc Chem Res* **38**(7), 542–8 (2005).
-

Summary

This dissertation describes experimental studies of the interaction between DNA and enzymes called topoisomerases. These enzymes are capable of removing twist from DNA, and in particular looped DNA structures called supercoils. These structures are naturally introduced into cellular DNA as a result of DNA replication, DNA transcription and recombination. Topoisomerases are vital for cell proliferation and are a target for poisoning by the camptothecin-class of anti-cancer drugs. We use a variety of techniques, which have in common that they yield dynamic information from a single molecule. Most prominently, we have used magnetic tweezers, a technique that enables both stretching and twisting of a single DNA molecule. As such, we can control and monitor the topological state of the DNA molecule with high precision and study single topoisomerase enzymes that alter this topological state.

We first test how a single DNA molecule intrinsically responds to sudden increases in force, either in the absence of supercoils, or in their presence. To achieve this, we have developed a combination of magnetic tweezers and optical tweezers. We find that the observed dynamics can be accurately described by quasistatic models, independent of the degree of supercoiling initially present in the molecules. These quasistatic models include contributions from the entropy of stretching DNA in thermodynamic equilibrium, as described by the Worm-like chain model, and a Stokes drag force opposing the motion of the magnetic bead. In particular, the dynamics are not affected by the continuous mechanical removal of the plectonemes, indicating that hydrodynamic drag of the DNA itself does not need to be taken into consideration in these experiments. Practically, our results allow us to predict the DNA extension trajectory in time upon an imposed change of the topology of the DNA. This constitutes a quantitative baseline for the dynamics of topological changes induced by enzymes.

We then study supercoil removal by topoisomerases. Type 1B topoisomerases (Top1B) form a protein clamp around the DNA duplex and create a transient

nick that permits removal of supercoils. We show that Top1B releases supercoils by a swivel mechanism that involves friction between the rotating DNA and the enzyme cavity, indicating that the DNA does not rotate freely. We compare the uncoiling dynamics by Top1B with that of a nicking enzyme, which is incapable of religating the nick it creates. One would thus expect that a nicking enzyme removes all supercoils initially present in the DNA. This is indeed what we find experimentally. In contrast to a nicking enzyme, Top1B does not release all the supercoils at once, but it typically does so in multiple steps. We find that the number of supercoils removed per step follows a single exponential distribution, which is indicative of a stochastic process for religation. Furthermore, the enzyme is found to be torque-sensitive, since the mean number of supercoils per step increases with the torque stored in the DNA. We propose a model for topoisomerization in which the torque that is present in the DNA drives the DNA rotation over a rugged periodic energy landscape in which the topoisomerase has a small but quantifiable probability to religate the DNA once per turn.

To facilitate the faithful analysis of the average number of supercoils that Top1B removes from DNA, we have developed a mathematical methodology based on maximum likelihood. The need for such a methodology arose because of a global constraint that is set by the experimental strategy. The global constraint is that Top1B cannot remove more supercoils than the DNA contains. Inevitably, Top1B will remove the last supercoils that the DNA contains, which is a number that is artificially constrained and should be ignored. Effectively, the removal of a sufficiently small number of supercoils that does not deplete all supercoils is taken into account in the calculation of the average, whereas a large number of supercoils removed is discarded from the calculation. This leads to an overrepresentation of small numbers of supercoils removed and biases the average towards lower values. The maximum-likelihood methodology allows for a robust and unbiased estimation of the average, even when the bias well exceeds 100%. We demonstrate the potential of the method for a number of single-molecule experiments, focusing on the removal of DNA supercoils by Top1B, and validate the method via numerical simulation of the experiment.

We then study the supercoil removal by Top1B in a more complex setting, *i.e.* in the presence of chemotherapeutic drugs that specifically poison human Top1B. In general, increasing the ability of chemotherapeutic drugs to kill cancer cells is often hampered by a limited understanding of their mechanism of action. The subject of our study is the camptothecin class of drugs, of which topotecan which is in clinical use against a variety of cancers, is a representative. Topotecan is thought to stabilize a covalent topoisomerase-DNA complex, rendering it an obstacle to DNA replication forks. Using magnetic tweezers, we monitor the

dynamics of topoisomerase in the presence of topotecan. This allowed us to detect the binding and unbinding of an individual topotecan molecule in real time and quantify the drug-induced trapping of topoisomerase on DNA. Unexpectedly, our findings also show that topotecan significantly hinders topoisomerase-mediated DNA uncoiling, with a more pronounced effect on the removal of positive (overwound) versus negative supercoils. Testing the biological ramifications of our single-molecule findings in yeast cells, we verified the resulting prediction that positive supercoils would accumulate during transcription and replication as a consequence of camptothecin poisoning of Top1B. Positive supercoils, however, were not induced by drug treatment of cells expressing a catalytically active, camptothecin-resistant topoisomerase I mutant. This unique combination of single-molecule and *in vivo* data suggests a novel cytotoxic mechanism for camptothecins, in which the accumulation of positive supercoils ahead of the replication machinery induces potentially lethal DNA lesions.

We also describe measurements that visualized the binding of Top1B to DNA using atomic force microscopy. The goal of this study was to inspect whether the enzyme promotes long-range DNA-DNA crossovers and synapses. Top1B was found to form filaments on nicked-circular DNA by *intramolecular* synapsis of two segments of a single DNA molecule. Measuring the filament length as a function of protein concentration showed that synapsis is a highly cooperative process. At high protein:DNA ratios, synapses between *distinct* DNA molecules were observed, which led to the formation of large vTopIB-induced DNA clusters. These clusters were observed in the presence of Mg^{2+} , Ca^{2+} , or Mn^{2+} , suggesting that the formation of *intermolecular* vTopIB-mediated DNA synapsis is favored by screening of the DNA charge.

We report on the characterization and coating of "zero-mode" waveguides for the single-molecule study of DNA polymerizing enzymes, such as telomerase and DNA polymerase. These enzymes have in common that they incorporate nucleotides into a pre-existing DNA molecule. The aim is to monitor the incorporation of fluorescently-labeled nucleotides in real-time and thus gain insight into the mechanisms and kinetics of these enzymes. The experimental strategy is designed to satisfy two conflicting requirements of this experiment: 1. to monitor the incorporation at the single-nucleotide level and 2. to perform the experiment at relatively high and biologically relevant nucleotide concentrations. "Zero-mode" waveguides thus constitute an example on how nanotechnology can be used in biophysical experiments.

In summary, we have performed biophysics experiments at the level of individual biomolecules. Using magnetic tweezers, we have provided new quantitative insights on the topoisomerase 1B swivel dynamics during DNA supercoil removal,

both in the absence and presence of chemotherapeutic drugs. Furthermore, we have performed experiments in yeast cells that verified the predictions generated by our single-molecule measurements about the behavior of an entire cell.

Daniel Koster

June 2007

Samenvatting

Dit proefschrift beschrijft experimentele studies van de interactie tussen DNA en topoisomerase enzymen. Deze enzymen zijn in staat om torsie uit DNA te verwijderen en in het bijzonder de lussen in DNA die een gevolg zijn van deze torsie. De lussen worden in cellulair DNA geïntroduceerd als gevolg van DNA replicatie, transcriptie en recombinitie. Topoisomerasen zijn in de meeste organismen essentieel voor celdeling en kunnen vergiftigd worden door kankermedicijnen van de camptothecin klasse. We gebruiken een verscheidenheid aan technieken, die met elkaar gemeen hebben dat zij informatie verschaffen over de dynamiek van een enkel molecuul. In het bijzonder hebben wij een magnetisch pincet gebruikt, een techniek die ons in staat stelt om enkele DNA moleculen zowel uit te rekken als te roteren. Zo kunnen we met een hoge mate van precisie de topologische toestand van het DNA molecuul definiëren en observeren en zijn we ook in staat om enkele enzymen te bestuderen die deze topologische toestand veranderen.

We beginnen door te toetsen hoe een enkel DNA molecuul intrinsiek reageert op plotselinge toenames in trekkracht, in af- dan wel aanwezigheid van lussen. Hiervoor hebben we een opstelling ontwikkeld waarin een magnetisch pincet en een optisch pincet verenigd zijn. De dynamica die wij meten kan goed worden beschreven door quasistatische modellen, onafhankelijk van het aantal lussen dat initieel in het DNA molecuul aanwezig was. Deze quasistatische modellen beschrijven het uitrekken van het DNA molecuul in thermodynamisch evenwicht en bevatten een entropische bijdrage beschreven door het wormachtige ketting (“worm-like chain”) model, en een hydrodynamische (Stokes) bijdrage die de beweging van het magnetische bolletje door een vloeistof tegenwerkt. In het bijzonder wordt de dynamica niet beïnvloed door de continue en mechanische verwijdering van lussen, hetgeen erop wijst dat de hydrodynamische wrijving van het DNA zelf geen noemenswaardige bijdrage levert. Praktisch gezien maken onze resultaten het mogelijk om het tijdspad van de DNA extensie als gevolg van een opgelegde verandering in de topologie van het DNA te voorspellen. Hiermee

wordt een kwantitatieve fundering gelegd voor de beschrijving van de dynamica van topologische veranderingen door de werking van enzymen.

Vervolgens bestuderen we de enzymatische verwijdering van lussen door topoisomerasen. Type 1B topoisomerase (Top1B) vormt een eiwitklem om het DNA heen en knipt tijdelijk één van de twee DNA strengen door, hetgeen het mogelijk maakt om lussen uit het DNA te verwijderen. Nadat Top1B een aantal lussen heeft verwijderd, repareert Top1B de breuk in het DNA weer, een proces dat religatie heet. We laten zien dat Top1B lussen verwijderd via een tolmechanisme waarin wrijving bestaat tussen het tollende DNA en de centrale holte in het enzym, hetgeen er op wijst dat het DNA niet vrij kan roteren. We vergelijken deze dynamica van het DNA met die van een eenvoudig enkel-strengs knipenzym, dat niet in staat is tot religatie. Het valt te verwachten dat een dergelijk enzym alle aanwezige lussen uit het DNA verwijdert, hetgeen we experimenteel ook observeren. In tegenstelling tot zo'n enkel-strengs knipenzym, verwijdert Top1B niet alle lussen ineens, maar doet Top1B dat stapsgewijs. Onze metingen wijzen erop dat het aantal lussen dat per stap ontwonden wordt een exponentiële verdeling volgt. Dit geeft aan dat de enzymatische reparatie van het DNA door Top1B een stochastisch proces is. Verder meten we dat het enzym gevoelig is voor de torsie in het DNA molecuul, aangezien het gemiddelde aantal lussen dat per enzymatische stap verwijderd wordt toeneemt met de torsie in het DNA. Deze data zijn in een model te vangen, waarin het DNA, als gevolg van de opgelegde torsie, over een hobbelig en periodiek energielandschap roteert, waarbij het topoisomerase eens per omwenteling een kleine, doch meetbare, kans heeft om het DNA te religeren.

Om een natuurgetrouwe analyse mogelijk te maken van het aantal lussen dat Top1B uit het DNA verwijdert, hebben we een wiskundige methodiek ontwikkeld die gebaseerd is op het principe van "maximum likelihood". De noodzaak voor zulk een methodiek kwam aan het licht door de aanwezigheid van randbeperkingen die intrinsiek zijn aan onze experimentele strategie. In het geval van de studie van Top1B is de intrinsieke randbeperking dat het enzym niet méér lussen uit het DNA kan verwijderen dan er in het DNA zitten. Top1B zal noodzakelijkerwijs eens de laatste lussen uit het DNA verwijderen, maar deze hoeveelheid lussen is kunstmatig gelimiteerd en dient niet meegenomen te worden in de berekening van het gemiddelde aantal verwijderde lussen. Kleine aantallen verwijderde lussen die er niet toe leiden dat alle lussen uit het DNA verwijderd worden dus wel meegenomen in de berekening van het gemiddelde, terwijl grote aantallen die er wel toe leiden dat de laatste lussen uit het DNA verdwijnen niet meegenomen worden. Dit leidt tot een overwaardering van kleine hoeveelheden verwijderde lussen, hetgeen het gemiddelde naar beneden brengt. De "maximum likelihood" methodiek stelt ons wel in staat om een robuuste berekening van het gemiddelde

te maken, zelfs in het geval dat de vertekening groter is dan 100%. We toetsen de methodiek door middel van numerieke simulaties van het experiment en laten de kracht van deze methodiek zien aan de hand van een aantal experimenten aan enkele moleculen.

Vervolgens beschrijven we de verwijdering van lussen door Top1B in een meer complexe omgeving, namelijk in de aanwezigheid van kankermedicijnen die specifiek menselijk Top1B vergiften. Het beperkte mechanistische begrip van de werking van kankermedicijnen limiteert hun effectiviteit. Het onderwerp van onze studie is de camptothecin-klasse medicijnen, met topotecan als representatief voorbeeld dat hedendaags gebruikt wordt voor de klinische behandeling van verschillende tumoren. Topotecan zou het covalente complex van Top1B en DNA stabiliseren, waardoor dit complex een obstakel voor DNA replicatievorken wordt. Met behulp van ons magnetisch pincet bekijken we de dynamica van Top1B in de aanwezigheid van topotecan. Hierdoor zijn wij in staat om “live” de binding en ontbinding van een enkel topotecan molecuul waar te nemen en om de verlengde levensduur van het Top1B aan het DNA te meten. Tegen de verwachtingen in tonen onze resultaten ook aan dat de ontwindings van het DNA significant wordt gehinderd door topotecan, terwijl dit effect op positieve (overwonden) lussen geprononceerder blijkt dan op negatieve (onderwonden) lussen. Om de biologische gevolgen van onze enkel-molecuul bevindingen in gistcellen te testen, hebben we geverifieerd dat positieve lussen zich ophopen tijdens transcriptie en replicatie als een gevolg van behandeling met topotecan. Positieve lussen hoopten zich echter niet op na topotecan behandeling in cellen die een topotecan-resistente mutant tot expressie brachten. Deze unieke combinatie van enkel-molecuul experimenten en *in vivo* studies duiden op een nieuw mechanisme voor de werking van camptothecin medicijnen, waarin positieve lussen die zich voor de replicatiemachinerie ophopen potentieel dodelijke DNA lesies induceren.

In een gerelateerde studie bekijken we ook direct de binding van Top1B op DNA met behulp van de atomaire tastmicroscop. Het doel hiervan was het inspecteren of de enzymen lange-afstands DNA-DNA kruisingen kunnen induceren, alsmede DNA synapsen. Top1B bleek filamenten te vormen op “nicked” circulair DNA door *intramoleculaire* synapsis van twee onafhankelijke segmenten van een enkel molecuul. De formatie van DNA synapsen bleek een hoogst coöperatief proces, hetgeen bleek uit een meting van filamentlengte als functie van de eiwitconcentratie. Bij grote eiwit-DNA verhoudingen werden ook DNA synapsen tussen *verschillende* DNA moleculen waargenomen, die leidden tot de vorming van grote kluwes DNA. De kluwes werden gezien in de aanwezigheid van Mg^{2+} , Ca^{2+} en Mn^{2+} , hetgeen erop wijst dat de vorming van *intermoleculaire* DNA synapsvorming bevordert wordt door afscherming van de elektrische lading van

DNA.

Tenslotte rapporteren we ook de karakterisatie en coating van optische golfgeleiders voor enkel-molecuul studies van enzymen zoals telomerase en DNA polymerase. Deze enzymen hebben met elkaar gemeen dat ze nucleotiden inbouwen in een reeds bestaand DNA molecuul. Het doel is om het inbouwen van fluorescent gemarkeerde nucleotiden “live” te bekijken en zo inzicht te verschaffen in hun mechanismes en kinetiek. Deze experimentele strategie is ontworpen om te voldoen aan twee tegenstrijdige eisen van het experiment: ten eerste om het inbouwen van nucleotiden te bekijken op het niveau van een enkel molecuul, en ten tweede om het experiment uit te voeren bij een relatief hoge en biologisch relevante nucleotide concentratie. Optische golfgeleiders zijn een voorbeeld van hoe nanotechnologie gebruikt kan worden in biofysische experimenten.

Samengevat hebben we biofysica-experimenten uitgevoerd op het niveau van individuele moleculen. Met een magnetisch pincet hebben we nieuwe kwantitatieve inzichten verschaft over de tol-dynamica van topoisomerase 1B tijdens de ontwinden van DNA lussen, zowel in af- als aanwezigheid van kankermedicijnen. Daarbij hebben we door middel van metingen aan gistcellen laten zien dat de inzichten die we hebben verkregen op het niveau van een enkel molecuul, voorspellingen kunnen genereren over het gedrag van een complete cel.

Daniel Koster
Juni 2007

Curriculum Vitae

Daniel Koster

- 18-06-1978 Born in Amsterdam, The Netherlands.
- 1991-1996 Secondary school, Vossius Gymnasium in Amsterdam.
- 1995 Dr. Bessie F. Lawrence International Summer Science Institute at the Weizmann Institute, Rehovot, Israel.
- 1995-1996 Preparatory Program for Clarinet at the Sweelinck Conservatory, Amsterdam.
- 1996-1997 Preparatory Program Tel-Aviv University, Tel-Aviv Israel.
- 1997-2003 M.Sc. Chemistry, University of Amsterdam.
- 2001-2002 Research Project at the Rowland Institute at Harvard University, Cambridge, MA, USA.
- 2003-2007 Ph.D research at Delft University of Technology under supervision of prof. dr. C. Dekker and Dr. Nynke H. Dekker.

List of Publications

1. *Antitumor drug impede DNA uncoiling by Topoisomerase I*
Daniel A. Koster, Komaraiah Palle, Elisa S. Bot, Mary-Ann Bjornsti and Nynke H. Dekker.
Nature, in press.
2. *Fast dynamics of supercoiled DNA revealed by single-molecule experiments*
Aurélien Crut, Daniel A. Koster, Chris H. Wiggins and Nynke H. Dekker.
Under review at Proceedings of the National Academy of Sciences of the USA.
3. *Multiple events on single-molecules: unbiased estimation in single-molecule biophysics*
Daniel A. Koster, Chris H. Wiggins and Nynke H. Dekker.
Proceedings of the National Academy of Sciences of the USA, Feb. 7 2006: 103(6):1750-5.
4. *Atomic force microscopy shows that vaccinia topoisomerase IB generates filaments on DNA*
Fernando Moreno-Herrero, Laurent Holtzer, Daniel A. Koster, Stewart Shuman, Cees Dekker and Nynke H. Dekker.
Nucleic Acid Res. Oct 19 2005; 33:5945-53.
5. *Friction and torque govern the relaxation of DNA supercoils by eukaryotic topoisomerase IB*
Daniel A. Koster, Vincent Croquette, Cees Dekker, Stewart Shuman and Nynke H. Dekker.
Nature, Mar 31 2005; 434:671-4.
6. *Coated nanofabricated structures for the single-molecule study of DNA polymerizing enzymes*
Daniel A. Koster*, Aurélien Crut*, Zhuangxiong Huang, Elizabeth H. Blackburn

and Nynke H. Dekker. *equal contribution. In preparation.

7. *Topotecan-mediated enzymatic relaxation of DNA supercoils at a specific sequence*

Daniel A. Koster, Fabian Czerwinski, Paola Arimondo, and Nynke H. Dekker.
In preparation.
