Forcing DNA and RNA through Artificial Nanopores

Proefschrift

ter verkrijging van de graad van doctor
aan de Technische Universiteit Delft;
op gezag van de Rector Magnificus prof. ir. K.C.A.M. Luyben,
voorzitter van het College voor Promoties,
in het openbaar te verdedigen op donderdag 21 oktober 2010 om 15.00 uur
door

Michiel VAN DEN HOUT

doctorandus in de natuurkunde
geboren te Leiderdorp.

Dit proefschrift is goedgekeurd door de promotor:

Prof. dr. N.H. Dekker

Samenstelling van de promotiecommissie:

Rector Magnificus	Voorzitter
Prof. dr. N.H. Dekker	Technische Universiteit Delft, promotor
Prof. dr. U. Gerland	München, Germany
Prof. dr. G.T. Barkema	Universiteit Leiden
Prof. dr. C. Dekker	Technische Universiteit Delft
Prof. dr. H.W. Zandbergen	Technische Universiteit Delft
Dr. U.F. Keyser	Cambridge University, United Kingdom
Dr. D. Stein	Brown University, Verenigde Staten







Keywords: RNA, DNA, nanopores, optical tweezers, single molecule, nanotechnology, electrophoresis

Published by: Michiel van den Hout
Cover design: Kasper van den Hout
Printed by: Off Page

The production of this thesis was financially supported by Delft University of Technology and J. van den Hout Beheer B.V.

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

Copyright © 2010 by Michiel van den Hout

Casimir PhD series, Delft-Leiden, 2010-26
ISBN: 978-90-8593-085-3

Whatever you do will be insignificant,
but it is very important that you do it

— **Mahatma Gandhi**

Preface

This is *it*. And this *it* can mean a lot of things. First, *it* refers to this preface, the only part of this book that most of its owners will actually read. Next, to you, the reader, *it* is likely to represent the part "where he writes something about me. *It* better be nice…". Well.., it just might, you know. For me, *it* refers to this entire book, as it represents the culmination of more than four and a half years of my working life. *It* is also the opportunity to express my gratitude to those who participated in making *it* possible. But before that, I wish to briefly recall some recent history, as I believe it may be of some interest. I received my MSc degree in physics in 2001, after a year of research on carbon nanotubes under the very wise supervision of Serge Lemay and Cees Dekker. In the last months of my time there, Cees embarked on an entirely new scientific quest as he started his own *Molecular Biophysics* group. I had the privilege to be among the very first students to graduate as *master of science* in this new group. When I finally rejoined the group, after an intermezzo of four years of "real" work at ASM Europe, the group had flourished immensely and I found myself overwhelmed by the many excellent researchers of such a wide variety of backgrounds. Now, some nine years after my first Delft graduation, I feel again privileged, as I will be the first PhD student to graduate in what has evolved into an entirely new department: the department of Bionanoscience! Perhaps, if ever I might be tempted to return to academia in –say- five years, I expect nothing less than to see the birth of a completely new faculty of Bionanoscience, and a Delft University of Bionanoscience by 2025..?

But now, let's get on with *it*! I have many people to whom I wish to express my gratitude. First and foremost, my thanks go to Nynke, my promoter and supervisor, who made this all possible. Thank you for all your support, Nynke. You have taught me some valuable things, which I will benefit from for the rest of my life. First, you have taught me the importance of being accurate and precise. As my personality is such that I often prefer *not* to be very serious and accurate, this was sometimes a bit hard for me. Next, you have taught me the importance of obtaining results and the systematic approach that this typically requires. As both of these were not exactly abundantly present during the first couple of years, I am very happy that in the end everything worked out fine and this thesis actually contains a reasonable amount of pages with (to some people) even some quite interesting results. I (and with me, this thesis) have also benefited strongly from your excellent writing skills and scientific insight. Finally, I wish to express my gratitude to have been witness to the rise of the Nynke Dekker lab, which will no doubt (remain to) be very successful indeed. As the first student to be subject to your *ius promovendi* I wish you all the best in the future, particularly with the exciting new scientific directions you have recently started to explore.

Next, I want to thank Cees, who has been a source of great inspiration to me. Cees, it was you who made me enthusiastic about carbon nanotubes, some ten years ago. Five years later, it was again through you that I started on this project and I have always very much appreciated your close involvement. I have deep respect for your incredible energy, your quick mind, and your friendly and open attitude. Especially this last thing has made the group (and now the department) such a great place to work. You have the ability to gather a very interesting breed of people around you: intelligent, talented, but also very communicative and willing to have fun in other areas than science such as sports, music or just going out and having a drink. Most importantly, I have always very much appreciated your enthusiasm, which has helped me through some of the more difficult periods during my PhD.

Then there are of course many other people who were directly involved in my research. Diego, you were the first person I had the pleasure of working with when I arrived. I realized I had forgotten everything about physics in my four years at ASM and you were a great help in getting me back on track. I enjoyed your impatient and often chaotic approach, and enthusiasm. Ulrich, I regret we had less interaction, but you were definitely always available for questions. In building (and using) my version of your setup, I have continuously been amazed at the level of complexity and subtlety in its design. You are brilliant! Stijn, you too, I trust you´ll do very well in your current research. Serge Lemay, unfortunately we haven't had much interaction, but I happen to have experienced during my MSc project that you are the perfect teacher, and the best scientific speaker I know. After talking to you I always feel I have learned something new. Igor, thanks for all the help with the RNA and in the bio-lab: you, Susanne and Andrea ("*oi bellooooo, come andiomo?")* made life there a lot more bearable. It was both daunting (and I must admit: encouraging) to see the amount incredible bad luck you faced (and managed to overcome!) before your P2 experiment finally worked. And thanks for all the parties! Wiecher, thanks for skilfully exterminating the RNA ghost, it hasn't been spotted since. Gary, my good sir knight, you were my saviour when nothing worked: after your short initial stay the experiments finally started rolling. Thanks for your inexhaustible optimism. *The* Adam, thanks for everything: the Tuesday beers, the Friday squash, the many discussions about our experiments, the fact that your luck with the experiments was usually at least as bad as mine, the dinners together, for inventing the word *cinequotefortzone* (with Gary), and for being always willing to help. You must be the kindest person I know, you're a great scientist and I'm sure you'll be an awesome professor, I am proud to have you as a friend and honoured that you are my *paranimf*. Ralph, I have always very much enjoyed your company, and the pleasant human angle in most of our conversations. Meng Yue, thanks for all your help with the pores, and good luck with your company. Susanne, so many thanks for all your help with the experiments! Thanks for being always willing to help and putting in the extra hours. Serge, you too, thanks for being such a great help in the bio-lab. Aart-Jan, I can't

believe you pulled it off, and managed to make that most elusive of biological constructs, the RNA hairpin (even though it was only once..). Great work, hopefully you and others will be able to use it for some cool experiments. Stefan, thanks for being a great room-mate, and for our nice discussions about nanopores. Dimitri, and Jaap, thanks for your help on my setup. Xander, good luck with the intended further complications to my setup, I thought it was already difficult enough, but I guess I was wrong. And last, but certainly not least, I want to thank Onno, Vincent and Sven (and Alfred). Onno, your coming was (as J.R.R. Tolkien wrote) "..like the falling of small stones that starts an avalanche..": your work set off the first RNA-nanopore publication. Thanks for your drive and enthusiasm. Vincent, always happy, thanks for all your efforts! And Sven, thanks for some of the latest data in this *boekje*.

Jelle, jij krijgt een aparte plaats: bedankt voor al je inzet en hulp (ik hoefde maar te vragen, en jij had het al gemaakt voor ik uitgesproken was) en de vele uren dat we geouwehoerd hebben. Maar bovenal heb jij mijn eeuwige dank voor het mooiste kado dat ik ooit gekregen heb en waar ik nog mijn hele leven lang plezier van zal hebben (en waar ik al zovelen mee heb lastiggevallen in de avonduurtjes): de vleugel.

Then there are of course many people who, though less involved, still each in their own way contributed somehow to this book. First, my fellow PhD-maties: I enjoyed our occasional PhD dinners, it was good to be able to complain to each other about our supervisors and the rest of the lab. Frank, and Daniel, my former room-mates, thanks for making these incredibly annoying beat-box sounds together: I now know the value and beauty of silence… Thijn, it was interesting to see someone juggling apples, minks and magnetic beads simultaneously, while still being able to produce the biggest thesis of the group (although you beat me only by a measly two pages). Koenie, I still sometimes miss your sniggering laugh, and the fact that you were the only other person in the lab who also knew all the *Urbanus* songs. Martin, you´re a smart and eloquent person and not afraid to speak your mind, please consider going into politics, it would be good for our country. Marcel, I loved playing you in your PhD movie, you must be a very happy person. Have fun in Eindhoven. Oppi, so long, and thanks for all the fun, I reckon we'll stay in touch! Sjoerd-Jan, I have enjoyed seeing you grow into the fun person you are now (or more likely it was me, growing used to you). I'll miss your shouts of "Meekeeyou!" across the hallway, but I rest in the fact that there will always be other people's names to be shouted. Such as "Hol!": Felix, it's too bad we haven't spent more time together. But it's good to know that there is always a safe haven (although now unfortunately in Delft in stead of Rotterdam), whenever I happen to wake up in a closed-off train in the middle of the night. Rifka, I wish you all the best with combining motherhood with a PhD. Marijn, after having built such a cool setup, I hope you can actually measure some cool things with it. And lastly, Charl, good luck with your challenging new project.

And then a big hand for all you post-docs out there, who made the group such a lively place: Derek, Fernando, Irene, Dirk, Aurelién (I´m terribly sorry, but your *outrageous* French accent has been topped by David), Freek, Christine, Jaan, Peter, David, Juan (I'll just count you as post-doc), Greg, Matthew, Tim and Gautam. Thank you all for your enthusiasm in science, for all the chats and fußball matches we played. Bernard, thanks for the fun times and the great conversations we've had. Jan, thanks for always willing to help and think along with everyone. The squash-group, Edgar, Pradyumna and Francesco, I've enjoyed our many squash matches and conversations. Iwijn, thanks for your Flemish influence on the group (which has been –in perfect agreement with the Flemish attitude- extremely modest, yet pleasant) and your humour and enthusiasm.

Of all the Bachelor and Master students that visited the group in the past five years, I'd like to thank a few in particular: Douwe-Jan, Stijn, Johan, thanks for the many fussball matches and borrels. Anne, you are a wonderful person, always fun to be around with. I can't thank you enough for your brilliant insight to introduce me to your sister. Andrea (and parents), thanks for making that meeting possible in the most marvellous of settings, Triëste.
Thanks also to Elsemarieke, Anouk, Liset, Emmylou en Leonie, for all the moral support. To all the new people in the BN department, I am sorry I haven´t had the time to get to know you all. I hope you will enjoy your time and make this new department into the centre of excellence it is intended to become!

And finally, I want to finish by thanking those who are closest to me, who each in their own way helped me with this project: my friends, and family. Vrienden, zonder jullie was mijn leven ondraaglijk saai geworden. Kasper, heel erg bedankt voor de geweldig omslag. Lies, Dick, jullie zijn de beste ouders die ik me kan wensen, dank voor jullie onvoorwaardelijke, nooit aflatende steun, voor de veiligheid en vrijheid die jullie me altijd hebben gegeven. En tenslotte, Roos, dank je voor al je vrolijkheid, je liefde en je ondernemendheid. Zolang ik met jou ben, schijnt elke dag de zon.

Utrecht, September 2010
Michiel van den Hout

Contents

1. Introduction **13**
- 1.1 The ultimate question of life, the universe, and very tiny things... 14
- 1.2 DNA 15
- 1.3 RNA 16
- 1.4 (Bio)-Nanoscience 19
  - 1.4.1 Optical tweezers 21
  - 1.4.2 Nanopores 22
  - 1.4.3 Optical tweezers combined with nanopores 23
- 1.5 Outline of this thesis 25
- Bibliography 27

2. An RNA toolbox for single-molecule force spectroscopy studies **31**
- 2.1 Introduction 32
- 2.2 Materials and methods 35
- 2.3 RNA construct assembly for s.m. force spectroscopy studies 38
  - 2.3.1 The central section synthesis for ssRNA constructs 40
  - 2.3.2 Labeling of the terminal sections of the ssRNA construct 43
  - 2.3.3 Joining central and labeled terminal sections of the ssRNA 49
  - 2.3.4 Double-stranded RNA constructs 50
  - 2.3.5 Isolation and purification of RNA molecules and constructs 54
- 2.4 Conclusions 54
- 2.5 Acknowledgments 55
- Bibliography 55
- Appendix 2 64

3. End-joining long nucleic acid polymers **65**
- 3.1 Introduction 66
- 3.2 Materials and methods 67
- 3.3 Results 70
  - 3.3.1 Formation of DNA-STV-DNA hetero-dimers 70
  - 3.3.2 Applications of the method 72
- 3.4 Discussion 76
- 3.5 Conclusions 78
- 3.6 Acknowledgments 78
- Bibliography 78
- Appendix 3 83

4. Distinguishing single-stranded and double-stranded nucleic acid molecules using solid-state nanopores **93**
4.1 Introduction 94
4.2 Materials and methods 95
4.3 Results 97
4.4 Discussion 102
4.5 Acknowledgments 105
Bibliography 106
Appendix 4 109

5. Direct force measurements of dsRNA in solid-state nanopores **115**
5.1 Introduction 116
5.2 Materials and methods 117
5.3 Results and discussion 119
5.4 Conclusions 126
5.5 Acknowledgments 127
Bibliography 127
Appendix 5 130

6. Controlling nanopores size, shape and stability **133**
6.1 Introduction 134
6.2 Materials and methods 135
6.3 Results and discussion 136
6.4 Conclusions 143
6.5 Acknowledgments 143
Bibliography 144
Appendix 6 146

7. Distinguishable populations report on interactions of single DNA molecules with solid-state nanopores **147**
7.1 Introduction 148
7.2 Materials and methods 150
7.3 Results 152
7.4 Discussion 158
7.5 Conclusions 162
7.6 Acknowledgments 163
Bibliography 163
Appendix 7 165

8. Forcing RNA Homopolymers through Solid-State Nanopores **169**
8.1 Introduction 170
8.2 Materials and methods 173

8.3 Results 174
8.3.1 RNA homopolymer translocations in pores < 8 nm 175
8.3.2 Measuring forces on poly(A) in a nanopore 182
8.4 Towards RNA unfolding through a nanopore 187
8.4.1 Molecular constructs 188
8.4.2 Experimental challenges 191
8.5 Conclusions 194
8.6 Acknowledgments 195
Bibliography 195

Summary **199**

Samenvatting **203**

Curriculum Vitae **207**

List of publications **209**

List of figures and tables **211**

Chapter 1

Introduction

Abstract

In this dissertation we describe experiments on RNA and DNA molecules using nanopores, and optical tweezers. This first chapter provides some background information on these concepts. Starting with the discovery of DNA and a description of DNA molecular structure, we next introduce RNA and describe some of its many biological functions. Having provided the biological context for our experiments, we move on to discuss some recent experimental techniques to study these biological molecules at the level of single molecules. We focus on two techniques in particular, that allow us to manipulate and study single molecules: optical tweezers and nanopores, both of which are used for the experiments described in this thesis. Optical tweezers allow us to grab on to single biological molecules, exert forces on them, and measure these forces. Next, we introduce nanopores: nanometer-sized holes in a very thin membrane. We discuss how both biological and synthetic nanopores have contributed to the understanding of polymer transport, and how they could be exploited in the future for applications such as genomic screening and sequencing or the determination of RNA secondary structure. We finish this chapter with an overview of the contents of this thesis.

"The most incomprehensible thing about the world is that it is at all comprehensible."
— **Albert Einstein**

1.1 The ultimate question of life, the universe, and very tiny things…

For many thousands of years mankind has wondered about the answer to "the ultimate question of life, the universe and everything", as science fiction writer Douglas Adams so eloquently put it in his book "The hitchhiker's guide to the galaxy"[1]. Although the ultimate answer to this question will likely remain forever elusive, the past few hundred years have seen incredible progress towards the understanding of what life, the universe and everything "is", and what laws govern its functioning. We have managed to explore, study and even manipulate the physical world to incredible extent. For example, we have sent probes into outer space, the farthest of which is removed well over 10^{10} km from the earth (satellite Voyager-1). At the other end of the scale, using the very sharp tip of a device such as a Scanning Tunneling Microscope (STM), we can manipulate single atoms of about 10^{-10} m (see Section **1.4**), as if they were tiny bricks of LEGO™. To appreciate how small this is, imagine zooming in a million times on this particular experiment: the ~5 mm STM tip would now be approximately the size of the Matterhorn mountain (Swiss Alps), the top of which would be used as a "sharp" tip to manipulate individual atoms, now the size of a single grain of salt! Throughout this thesis we will be working at a similar size scale, of about 1 to 10 nanometer, approximately the size of biological molecules (1 nanometer = 1nm = 10^{-9} m).

We will start this first chapter by introducing two of the most important molecules that form the basis of all life as we know it: DNA and RNA. We will briefly discuss their discovery, biological function, and some of their structural features. Next, we will introduce the new field of bio-nanotechnology, which allows us to study these molecules by visualizing and even manipulating *individual* molecules, focusing in particular on two powerful single-molecule techniques: optical tweezers and nanopores. Both have been used for the experiments described in this thesis.

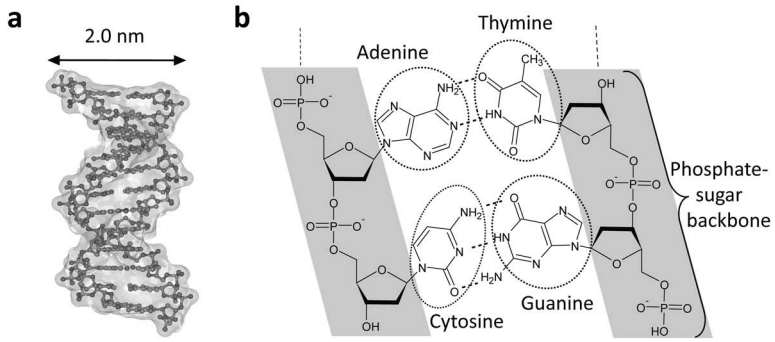


Figure 1.1. *Chemical structure of DNA*. **a**. Picture of a small part of a DNA molecule. Two molecular chains are linked together in a double-helical shape. The chain consists of a sugar-phosphate backbone on the outside of the molecule, with hydrophobic bases attached to it facing inward to the centre. Hydrogen bonds between the bases of both chains stabilize the structure. **b**. Chemical structure of two base-pairs in DNA, showing all four bases: adenine always pairs with thymine, and cytosine always with guanine. The grey shaded area depicts the sugar-phosphate backbone.

1.2 DNA

The discovery of DNA

It is now widely known that DNA (deoxyribose nucleic acid) is the carrier of our genetic information. DNA was first discovered in 1869 by Friedrich Miescher, who was trying to "..track down the basic prerequisites of cellular life.."[2]. At the time, he was studying the biochemical composition of leukocytes (a type of white blood cell), which he could obtain in high purity from –of all things– pus (from surgical bandages). During his experiments he stumbled upon a strange precipitate, which he called *nuclein* as it originated from the cell nuclei. Although he speculated that the material might play a role in the transmission of hereditary traits, it took another 75 years before the suggestion was made by Avery and coworkers[3] that DNA itself was the carrier of hereditary information. Shortly after that (in 1963), Rosalind Franklin and Raymond Goslin performed X-ray scattering measurements on crystallized DNA, in order to determine the structure of this molecule[4]. Watson and Crick then proposed the famous double-helix structure for DNA (see Figure **1.1a**) to explain these X-ray data[5]. A simultaneously published separate X-ray study by Maurice Wilkins and co-workers also supported this structure[6].

DNA structure

In Figure **1.1** we show the structure of DNA in some detail. The double-helix (Figure **1.1a**) contains two long polymers, or chains, the backbone of which consists of a regular alteration of sugar and phosphate groups (see Figure **1.1b**). To each sugar is attached a nitrogenous base, which can be of four different types. Two of the bases - adenine and guanine - are so-called purines (containing two aromatic rings), and the other two - thymine and cytosine - are pyrimidines (containing one aromatic ring). The monomer unit, consisting of a phosphate, sugar and base, is known as a nucleotide. The two chains are held together by hydrogen bonds between the bases in a pair-like fashion, as shown schematically in Figure **1.1b**. A single base from one chain is hydrogen-bonded (dashed lines in Figure **1.1b**) to a single base from the other, such that adenine always faces thymine, and guanine always faces cytosine. Because of this unique base-pairing, both chains are effectively each other's complement, and the sequence of one chain dictates that of the other. The sequence of the bases is known as "the genetic code", as it forms the template from which proteins are eventually produced.

The central dogma of molecular biology

The discovery of DNA's function and structure was a tremendous step ahead in the field of molecular biology, the science aimed at understanding biology on the molecular level. It quickly led to the insight of how hereditary information was encoded and how this information was transferred into the molecules of life. By the 1970s, the general consensus was that DNA contained the cell's hereditary information, which could be transferred into either RNA or other DNA. The RNA in turn could be translated into a protein (or other RNA). This principle is now generally referred to as the "central dogma" of molecular biology by Francis Crick[7]. In this picture, RNA (more commonly known as messenger RNA, or mRNA) is a temporary copy of a small section of the DNA, its only function being to facilitate protein production. In more recent years, however, this picture has changed dramatically, and RNA is now recognized to perform many other functions in the cell as we will see in the next section.

1.3 RNA

RNA is a hugely versatile molecule, involved in a large range of cellular regulation processes, of which we will now discuss a few. For example, during protein synthesis from mRNA, individual protein building blocks (amino acids) are chained together by a very large molecular complex, the ribosome. It turns out that the

ribosome itself is constituted for approximately 70% out of RNA. Also, the amino acids are put into place by special RNA molecules, called transfer-RNA (tRNA). Furthermore, even before the RNA is transferred to the ribosome, it is modified by a molecular complex called the spliceosome, which consists primarily of – again – RNA. Then, there are countless other RNA molecules involved in the regulation of gene transcription, the process in which RNA molecules are produced from DNA. For example, the discovery of RNA interference[8] has prompted a new wave of research into RNA: here, small double-stranded RNA (dsRNA) molecules were found to be responsible for inhibiting the transcription of certain targeted genes. Finally, when the human genome was sequenced in 2001[9,10], it was found that only ~1.4% of our DNA is actually transcribed into proteins. This means that over 98% of our DNA apparently serves some other purpose (if any). It is now known that at least 27% (and possibly even up to 45%) of our DNA is actually transcribed into RNA, performing (as yet) unknown roles in the cell[11].

RNA structure

There are three major chemical differences between DNA and RNA, which give rise to RNA's large variability: First, the base thymine (Figure **1.2a**, top) does not occur in RNA, but is replaced by uracil (Figure **1.2a**, bottom), which, like thymine, readily base-pairs with adenine. Second, although RNA is also a polymer consisting of nucleotides, the sugar in the sugar-phosphate backbone is slightly different: it contains an additional hydroxyl (-OH) group (Figure **1.2a**, bottom, dotted circle). This small difference has large consequences for the chemical stability of RNA: the hydroxyl group in RNA is more reactive than the C-H bond occurring in DNA. By reacting with the phosphate backbone it can actually cleave the polymer in two[12], thus making RNA intrinsically much less stable than DNA. Finally, RNA typically occurs in single-stranded form (ssRNA), in other words, it is a single chain of nucleotides. Since its individual bases can still locally form base pairs like in DNA, this gives rise to a large multitude of structural variety in RNA molecules (see Figure **1.2b**): the base sequence of each RNA molecule determines its local folding, and thus its structure. This so-called secondary structure is essential to RNA function: for example, many dsRNA segments of specific lengths are known to be binding domains for other RNA or protein molecules. In addition, higher order folding (known as tertiary structure) in which secondary structure elements are connected in a complicated 3-dimensional arrangement are also possible and again strongly determine their function in the cell. Although also helical, the local dsRNA segments are slightly different in structure than dsDNA because of the presence of the RNA's hydroxyl group (see also Chapters **4** and **5**). The diversity of RNA structure is even further increased by the fact that the basepairing in RNA is slightly less precise than in DNA (eg. basepairing between uracil and guanine is also

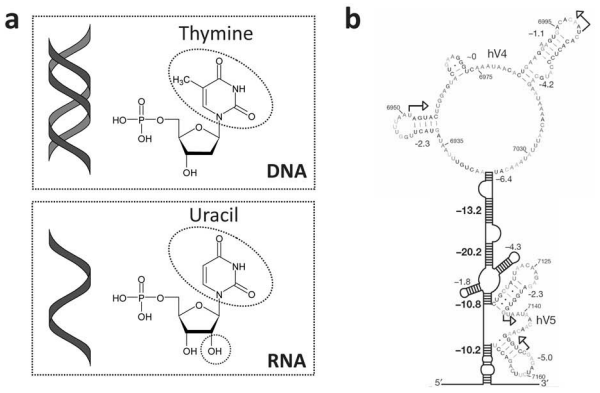


Figure 1.2. *Chemical structure of RNA*. **a**. The three major differences between DNA and RNA: 1) The base thymine (top) occurs only in DNA and is replaced with uracil (bottom) in RNA. 2) The sugar ring in RNA molecules contains a hydroxyl group (small dotted circle, bottom), which the sugar ring in DNA does not have. 3) DNA occurs primarily as a double-stranded molecule, whereas RNA is typically single-stranded. The base sequence along the RNA can lead to complex structures. **b**. Example of a section of the HIV1-RNA genome. The single stranded RNA molecule locally forms short regions of (base-paired) dsRNA, depending on the sequence of the molecule. Base-paired areas are shown as short ladder-like structures. On top a section is highlighted with the actual bases spelled out. Calculated free energies of the duplex regions are indicated in numbers (in kcal/mol). Image adjusted from ref [44].

possible).

Although essential for our understanding of their function, the three-dimensional structure of most RNA molecules is still unknown[13]. One of the reasons for this is that existing techniques to determine the three-dimensional structure of RNA molecules[14] are very time-consuming. For example, X-ray crystallography[15] (from which the three-dimensional structure of DNA was deduced) relies on the fact that under certain conditions a biological molecule can crystallize, thus forming a lattice of regularly spaced molecules. Using the diffraction pattern of X-rays at different angles through such a crystal, it is possible to reconstruct the shape of the individual molecules. However, most RNA molecules do not readily crystallize into regular lattices, in which case X-ray diffraction cannot be applied. Other techniques such as nuclear magnetic resonance[16] (NMR), cryo-electron microscopy[17] (cryo-EM), small angle x-ray scattering[18] (SAXS) are less cumbersome, but suffer from lower resolution and typically require very large amounts of RNA material. Many biochemical techniques exist as well[14], but these typically rely on cutting the molecules in small pieces and determining the structure of the individual pieces (which might have refolded, and so do not necessarily correspond to the structure in the complete molecules). Finally, there are many computational approaches to predict RNA structure[19], usually based on calculating the lowest free energy

configuration of the molecule. However, for many large (>100 bases) RNA sequences such computational methods do not predict a unique structure, but many alternative folds can have free energies within several k_BT of the lowest energy structure. This means that thermal fluctuations can favour any of the predicted structures. Computationally predicted structures are therefore not always reliable. In most cases, a large number of techniques are combined for the ultimate determination of an RNA molecule's three-dimensional structure.

In this thesis, we discuss a novel technique for the potential determination of RNA secondary structure, in which an RNA molecule is pulled at one end through a narrow hole. By choosing the size of the hole such that (base-paired) regions of dsRNA cannot pass through, and therefore must unfold, it might be possible to determine the location and size of such dsRNA regions. The next section will introduce a number of techniques that could make this experimentally possible.

1.4 (Bio)-Nanoscience

In the past decades, a large number of new techniques have been developed that allow the study and manipulation of matter at the molecular scale. Such technology is commonly referred to as nanotechnology, for the size scale of molecules is the nanometer, one billionth of a meter (10^{-9} m). In 1959, Nobel laureate Richard Feynman already speculated upon the possibility of controlling matter at this scale in his famous lecture "There's Plenty of Room at the Bottom" at the annual meeting of the American Physical Society. He wondered ".. whether – in the great future – we can arrange the atoms the way we want, the very atoms, all the way down!". Three decades later, this vision was actually realized by Don Eigler and Erhard Schweizer: using a low-temperature version of the recently developed (by Gerd Binnig and Heinrich Rohrer) Scanning Tunnelling Microscope they "wrote" the letters I B M with individual atoms[20], see Figure **1.3a**. This device consists of a very sharp tip, which can extremely accurately measure the distance to a nearby (conducting) surface. Eigler used such a tip to push individual Xenon atoms over an atomically flat nickel surface. This major breakthrough in the manipulation of matter at the molecular scale is now seen as one of the landmark experiments defining the field of nanoscience.

Control at the molecular scale is particularly useful for the study of biological molecules, as the structure and the functioning of many biological molecules are still unknown. In this exciting new area of research, aptly named bio-nanoscience, the classical sciences biology, chemistry, and physics converge, as biological processes on the molecular scale require an understanding of all these fields. We will mention a number of recent advances made in this area. For example, new optical imaging techniques such as STED microscopy[21] (stimulated emission

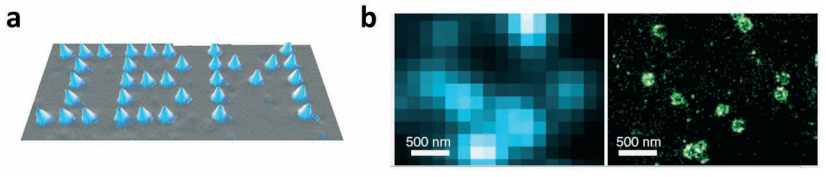


Figure 1.3. *Examples of advances in nanotechnology*. **a**. The letters I, B, and M, spelled out by individual Xenon atoms that were moved into place and imaged using the tip of a Scanning Tunnelling Microscope (STM). This device consist of an atomically sharp tip, which can extremely accurately measure the distance between sample and tip by measuring a (quantum) tunnelling current. Image reproduced from ref [20]. **b.** Recent improvements in optical microscopy now allow the very accurate localization of individual molecules. In this example, a regular optical microscopy image (left image) is compared to an image of the same area acquired with stochastic optical reconstruction microscopy (STORM, right image). Not only can the individual molecular complexes (clathrin-coated pits, in this particular example) be localized to much higher precision, their shape can even be determined very accurately. Image reproduced from ref [52].

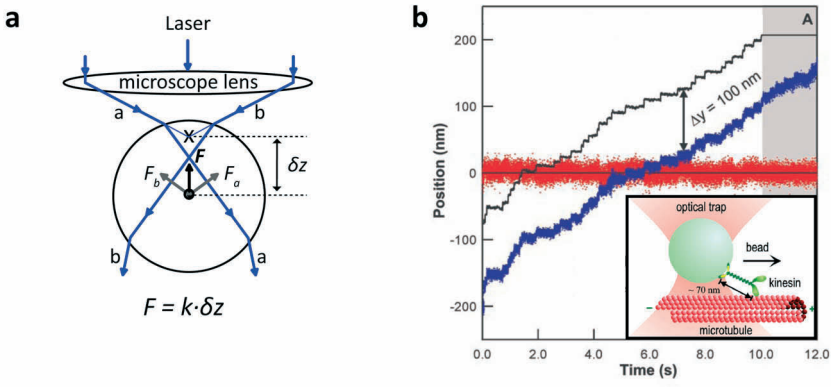


Figure 1.4. *Working principle, and example of optical tweezers.* **a**. A micron-sized bead next to a strongly focused laser beam will experience a force ***F***, resulting from refraction of the light in the surface of the bead. The force is directed towards the centre of the laser focus (denoted as x), and is proportional to the deviation of the bead from the centre, $\boldsymbol{F}=k\cdot\delta z$, where k is a constant, depending among other things on the laser power. The bead is thus "trapped" inside a harmonic potential well. **b**. Such an optical trap was used to measure the single steps of a *kinesin* molecule, a molecular "motor" protein. In cells, this protein transports large vesicles through the cell, by moving along scaffolding-like molecular rods (microtubules) that span the interior of the cell. A single molecule of *kinesin* was attached to an optically bead, and its movement was measured accurately as it traversed the microtubule. In addition, a force feedback was applied, moving the sample stage to compensate for the movement of the bead. The graph shows two sloped lines, of which the top trace corresponds to the stage movement, and the bottom trace to the bead position. Clearly, both traces indicate that the molecule moves in small steps. The constant trace at around 0 nm is the difference between the two sloped traces. Image reproduced from ref [27].

depletion), PALM[22] (Photo-activated Localization Microscopy) and STORM[23] (Stochastic Optical Reconstruction Microscopy) allow the localization of single biomolecules with a resolution of several tens of nanometres, see Figure **1.3b**. Scanning probe microscopy (in which samples are imaged and manipulated by scanning a sharp tip over its surface) is now routinely used for visualizing, and manipulating single molecules or small biological samples. In particular, the recent development of single-molecule force spectroscopy techniques (where molecules are manipulated individually) has opened up new ways to elucidate the kinetics and dynamics of biological molecules, which are typically averaged out in traditional ensemble measurements. For example, the kinetics of molecular "motors" such as *kinesin* could be elucidated to very high precision by studying its movement with nanometre precision (see Optical Tweezers, below). In the next two subsections we will highlight two recent, very powerful single-molecule techniques: optical tweezers, and nanopores.

1.4.1 Optical tweezers

In 1986 Arthur Ashkin showed that small dielectric particles could be manipulated using a strongly focused laser[24]. With these so-called optical tweezers[25] particles can be held and moved around in three dimensions. The principle of operation of optical tweezers is based on refraction: objects with a refractive index slightly different from their surrounding medium will refract any incident light. Since light has momentum, such a change in direction must result in a net force. Close to the laser focus this net force is directed towards the focus, thus "trapping" the object into the focus (see Figure **1.4a**). By moving the optical trap (i.e. the laser focus) the molecules can be moved around. External forces applied to the object can also be measured using this instrument, as the force pulling the object towards the focus is (for small distances δz) proportional to the distance from the laser focus, analogous to a Hookean spring: $F = k_{trap}\,\delta z$. Here, k_{trap} is the trapping constant, characterizing the stiffness of the optical trap. By determining the distance δz of the particle to the centre of the trap, the external force on the particle can thus be measured directly.

Optical tweezers have been used throughout the 1990s and afterwards for the study of biological systems. In most experiments biological molecules are biochemically attached to a microscopic bead, which in turn is trapped in a laser focus. Via the bead, forces can thus be applied to the biological molecule. For example, the movement of the single molecular motor *kinesin* was determined by Steve Block and coworkers in 1990[26]. This protein has two "legs", and is known to "walk" over large fibre-like structures in the cell called microtubules. By attaching a kinesin molecule to a microscopic bead, the force exerted by the molecule as it moved along a microtubule could be determined, and even individual steps could be resolved[27] (see Figure **1.4b**). Optical tweezers have also been employed to study

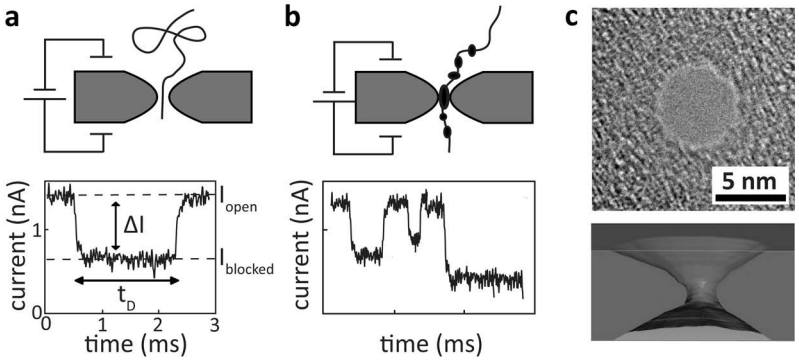


Figure 1.5. *Nanopore detection principle.* **a**. A single nanopore (top) separates two compartments filled with electrolyte. An applied bias voltage over the nanopore will cause a small ionic current to flow (below, I_{open}). When a molecule (eg. DNA or RNA) traverses the pore it will temporarily block the current ($I_{blocked}$), because its volume excludes the ions from passing the nanopore. **b**. Using this principle, it should be possible to determine local structure along a molecule, as the total current through the pore is dominated by the current flowing through the nanopore centre. For example, proteins bound locally to a DNA molecule (top) could potentially be detected from the different levels of current blockade, caused by their presence inside the pore centre (bottom). **c**. Example TEM image (top) of an artificial nanopore. The nanopore is fabricated into a 20 nm thick Si_3N_4 membrane. Below we show a computer rendering of the 3-dimensional shape of a nanopore in a 60 nm membrane, reconstructed from TEM tomography data (image reproduced from ref [38]).

RNA, by pulling on the RNA molecule from both ends and thus unfolding its secondary structure[28-30].

1.4.2 Nanopores

Another recent technique for the study of individual molecules is the use of nanopores, nanometer-sized holes in a thin membrane. Single molecules can be forced to translocate through such a nanopore, which can be detected and can be used to provide information about the molecule's structure. The first application of this idea to the study of biological molecules was done by Kasianowicz, and coworkers[31], who used a protein nano channel α-Hemolysin (α-HL, a toxin originating from a bacterium called Staphylococcus Aureus), embedded inside a lipid membrane, to detect the passage of individual molecules of RNA. This allowed them to characterize some of the structural features of these molecules. The detection principle used for this (and most subsequent nanopore experiments) is very simple: A voltage difference is applied over a nanopore, causing an ionic current to flow through the nanopore. The electrical field, strongest close to the

nanopore, will pull nearby charged molecules into the nanopore. As the molecule enters and traverses the nanopore, it will temporarily block part of the volume inside the pore, causing a change in the current (see Figure **1.5a**). This current change depends on the size and charge distribution of the molecule inside the nanopore[32], and can thus potentially be used to characterize its local structure (see Figure **1.5b**).

Although the α-HL nanopore is still widely used for such translocation experiments, nanopores can also be made artificially in solid-state membranes[33-35]. This has the advantage that the nanopore can be shaped to an arbitrary size, allowing for example the translocation of larger molecules. A popular technique to fabricate such solid-state nanopores is the use of a transmission electron microscope (TEM), which can drill a hole in a very thin solid-state membrane, usually consisting of SiN or SiO_2 of several tens of nm thick. With this technique[35,36], the size and shape[37,38] (see also Chapter **6**) of the nanopores can be controlled with sub-nanometer precision (see Figure **1.5c**).

Such artificial pores have already been used to translocate large molecules such as dsDNA[33,39], proteins[40,41] or protein coated-DNA[42], and also recently – presented in Chapter **4** of this thesis – RNA molecules. It is even possible to fabricate nanopores of such precise dimensions that double-stranded DNA cannot enter the nanopore, but single-stranded molecules can (i.e. nanopores with diameters between about 1 to 2 nm). Such nanopores have already been used to unfold the double-stranded part of partially ss-dsDNA molecules[43], by first trapping the single-stranded part into the nanopore and then denaturing the dsDNA by the electrical force acting on the single-stranded part of the molecule. This offers the possibility to study RNA secondary structure, as this also contains double-stranded regions.

However, in all experiments described above the forces used to pull the molecule through the nanopore are applied electrically. This has several disadvantages: First, the exact force exerted on the molecule is not known. Second, in most cases, the translocation speed is too high to resolve any fine local details along the molecule. Finally, the manner of translocation is not controlled and molecules may get stuck during translocation in a folded state, thus permanently clogging the nanopore. Such disadvantages could potentially be overcome by pulling the molecule through the nanopore using a local force probe, such as optical tweezers.

1.4.3 Optical tweezers combined with nanopores

The feasibility of such a combination of optical tweezers and nanopores was demonstrated in 2006 by Keyser and coworkers[44-46]. They used optical tweezers to arrest single molecules of dsDNA inside nanopores of about 6 to 10 nm in diameter, and measure the forces pulling the molecule into the nanopore (see Figure **1.6a**). In

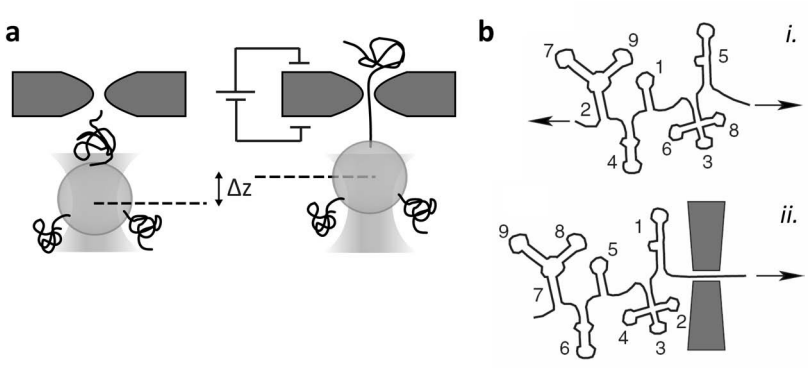


Figure 1.6. *Combination of optical tweezers and nanopores*. **a**. An optically trapped bead coated with biological molecules can be moved close to a nanopore. A strong electrical field over the pore may capture the molecule into the pore, reeling it through the pore until the optical restoring force on the bead equals the total force pulling the molecule in the pore. This force can be measured accurately by monitoring the displacement of the bead. **b**. Optical tweezers can be used to determine RNA secondary structure, by pulling on the molecule from both ends. The dsRNA stems along the molecule will then sequentially unfold based on their relative stability, which provides information about their length. However, in this case, the unfolding sequence does not necessarily correspond to their position along the molecule (see i. above, where the numbers correspond to the sequence in which the duplex regions unfold) and reconstructing the full secondary structure of the molecule can thus be cumbersome[28-30]. A potentially better approach is to pull the molecule from one end through a very small nanopore (ii.), such that the duplex regions do not fit into the pore and must unfold to pass through. In this case, the stems will have to unfold one by one, in the sequence corresponding to their position along the molecule. This method would therefore determine both the position and length of each duplex region along the molecule directly. Image reproduced from ref [48].

this experiment, molecules are first attached to an optically trapped bead, which is held close to the nanopore. An electrical field is then used to capture a molecule into the nanopore, which will pull the bead out of the centre of the optical trap, until the combined electrophoretic and electro-osmotic forces on the molecule balance the optical restoring force on the bead. As the optical tweezers also provide a direct readout of the force acting on the bead, the total force on a molecule inside a nanopore can be determined directly[44,47].

This experimental configuration could also be used to study local structure along a molecule with very high precision, as the molecule can be moved at an arbitrarily low speed through the pore mechanically (as opposed to the case of free translocation). Furthermore, by using sufficiently small nanopores the mechanical constriction of the nanopore can even be used to disrupt such local structure, for example by dislodging proteins bound to DNA. Alternatively, the force applied by molecular motors moving along a dsDNA or dsRNA molecule could be monitored by

allowing the molecular motor to move along the dsDNA until it reaches the nanopore. Provided the pore is small enough, the motor cannot pass through and will thus pull the dsDNA through the nanopore, exerting a force on the bead to which the dsDNA is attached. Such a measurement could be advantageous compared to alternative force spectroscopy methods, as it requires no tethering of the molecular motor to a surface.

In this thesis, we use such a combination of optical tweezers and nanopores to measure the forces on dsRNA (Chapter **5**) and ssRNA (Chapter **8**) as they are held inside a nanopore. These experiments are aimed towards using this setup for the study of RNA secondary structure by the sequential unfolding of dsRNA regions along the molecule[48], as schematically depicted in Figure **1.6b**. By pulling only from one end, only the double-stranded region directly inside the nanopore will experience the unfolding force, and will thus unfold before any other secondary structure is unfolded. Such sequential unfolding could potentially allow a more accurate determination of secondary structure than by unfolding from both ends of the molecule[28-30], as in the latter case the sequence of unfolding is determined by the relative stability of the various secondary structure elements along the molecule (see Figure **1.6b**).

1.5 Outline of this thesis

This dissertation describes experiments on RNA and double-stranded DNA using nanofabricated solid-state nanopores and optical tweezers. The work presented here was motivated by the ultimate goal to combine these two techniques to study the secondary structure of RNA molecules, by unfolding such RNA molecules through very small nanopores. The thesis can roughly be divided into two parts: the first part provides an overview of biochemical methods and procedures available to manipulate and synthesize RNA molecules suitable for our experiments (Chapters **2** and **3**). In the second part, we present several separate studies involving nanopores designed with the aim to study RNA and DNA structure (Chapters **4** to **8**). Below, we briefly discuss the contents of each chapter.

In Chapter **2**, we review the biochemical methods and procedures currently available for the modification of RNA molecules, to make them suitable for single-molecule experiments. For example, for many single-molecule experiments the molecules must be modified with molecular “handles” to allow their attachment to a force probe, such as a magnetic or optically trapped microbead. Many biochemical methods are available to modify DNA molecules at will, but for RNA the range of such biochemical tools is much more limited. We also discuss some of the practical difficulties specific to RNA, such as degradation and the synthesis of very long molecules exceeding 10 kilobases (kb).

In Chapter **3**, we present a new procedure that allows the cross-linking of long DNA or RNA molecules, eliminating some of the problems with existing biochemical methods, such as the low yields typically achieved once the molecules reach lengths above 10 kb. We demonstrate that we can efficiently end-join long (20 kb) DNA molecules, and also DNA to dsRNA or ssRNA or other combinations, by using the protein streptavidin to link these molecules.

We then move on to the nanopore experiments. In Chapter **4**, we present the first translocation measurements of long homopolymeric ssRNA and dsRNA molecules through artificial nanopores of approximately 10 nm in diameter, together with similar measurements on dsDNA. We compare the characteristic blockades caused by these molecules and show that from these blockades we can distinguish the single-stranded molecules from the double-stranded molecules in real time.

Next, in Chapter **5** we employ the combination of optical tweezers to measure forces on dsRNA and dsDNA over a wide range of nanopore size ranging from 3.5 to 35 nm. We find that in any given nanopore the forces on these molecules are approximately equal, but that the magnitude of the forces depends on the diameter of the nanopore (which was reported previously on dsDNA and RecA-coated dsDNA[47,49]). We also demonstrate that the force on the molecules is independent of the distance of the optical trap to the nanopore surface down to a distance of 0.5 μm, allowing for force measurements quite close to the nanopore.

Having established that we understand the physical processes at play in pores with diameters significantly larger than those of the molecules under study, we have studied what happens to the passage of DNA and RNA in very small nanopores of about 2 nm in diameter. We discovered that some of our small nanopores would grow in size over hours, which we report in Chapter **6**. There, we demonstrate that this behaviour originates from the TEM beam size employed to fabricate the pores: using a large TEM beam gives rise to hourglass shaped nanopores, which are likely to grow in size. Stable nanopores can be fabricated by using a beam size equal to the nanopore size.

We then go on to present dsDNA translocation results on small nanopores ranging from 2 to 7 nm in diameter (Chapter **7**). We can distinguish several groups of events with different conductance blockades, which we attribute to normal translocations and to instances in which the molecules interact strongly with the nanopore (not necessarily translocating the pore), respectively. Interestingly, our results differ from previous results by Wanunu, and coworkers[50], who did not observe such different groups of events. We show that this discrepancy may be related to the shape of our nanopores (which is set by the TEM beam employed during fabrication).

Finally, we present in Chapter **8** ongoing preliminary work towards our goal to use these nanopores to study the secondary structure of RNA molecules. We demonstrate for the first time that we can force single-stranded polymers of RNA into nanopores down to 2 nm in diameter. Second, we present very preliminary force measurements on poly(A) ssRNA with our optical tweezers. We conclude this final chapter with a discussion of the remaining challenges to unfold RNA molecules through a nanopore using optical tweezers.

Bibliography

1. Adams, D. (1979) The Hitchhiker's Guide to the Galaxy. *Pan Books*.
2. Dahm, R. (2008) Discovering DNA: Friedrich Miescher and the early years of nucleic acid research. *Hum Genet*. **122**: 565-581.
3. Avery, O.T., C. M. Macleod, and A.M. McCarty (1944) Studies on the chemical nature of the substance inducing transformation of pneumonococcal types. Induction of transformation by a deoxyribo-nucleic acid fraction isolated from pnuemococcus type III. *J. Exp. Med.* **79**: 137-157.
4. Franklin, R. and R.G. Gosling (1953) Molecular Configuration in Sodium Thymonucleate. *Nature.* **171**: 740-741.
5. Watson, J.D. and F.H. Crick (1953) A Structure for Deoxyribose Nucleic Acid. *Nature*. **171**: 964-967.
6. Wilkins, M.H.F., A.R. Stokes, and H.R. Wilson (1953) Molecular Structure of Deoxypentose Nucleic Acids. *Nature*. **171**: 738-740.
7. Crick, F.H. (1972) The Central Dogma of Molecular Biology. *Nature*. **227**: 561-563.
8. Fire, A., et al. (1998) Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. *Nature*. **391**: 806-811.
9. Consortium, I.H.G.S. (2001) Initial sequencing and analysis of the human genome. *Nature*. **409**: 860-921.
10. Venter, J.C., et al. (2001) The Sequence of the Human Genome. *Science*. **291**: 1304-1351.
11. Hüttenhofer, A., P. Schattner, and N. Polacek (2005) Non-coding RNAs: hope or hype? *Trends in Genetics*. **21**: 289-297.
12. Soukop, G.A. and R.R. Breaker (1999) Relationship between Internucleotide Linkage Geometry and the Stability of RNA. *RNA*. **5**: 1308-1325.
13. Popenda, M. et al. (2010) RNA FRABASE 2.0: an advanced web-accessible database with the capacity to search the three-dimensional fragments within RNA structures. *BMC Bioinformatics.* **11**: 231.1-231.11.
14. Felden, B. (2007) RNA structure: experimental analysis. *Curr. Opin. Struct. Biol.* **10**: 286-291.

15. Blundell, T. L., H. Jhoti and C. Abell (2002) High-throughput crystallography for lead discovery in drug design. *Nature Rev. Drug Disc*. **1**: 45-54.
16. Scott, L.G., and M. Hennig (2008) RNA structure determination by NMR. *Methods Mol. Biol.* **452**: 29-61.
17. Frank, J. (2002) Single-particle imaging of macromolecules by cryo-electron microscopy. *Annu. Rev. Biophys. Biomol. Struct*. **31**:303–319.
18. Lipfert, J. and S. Doniach (2007) Small angle X-ray scattering from RNA, proteins, and protein complexes. *Annu. Rev. Biophys. Biomol. Struct.* **36**: 307-327.
19. Shapiro, B.A., et al. (2007) Bridging the gap in RNA structure prediction. *Curr. Opin. Struct. Biol*. **17**: 157-165.
20. Eigler, D.M. and E.K. Schweizer (1990) Positioning single atoms with a scanning tunnelling microscope. *Nature*. **344**: 524-526.
21. Hell, S.W. (2003) Toward fluorescence nanoscopy. *Nat Biotech*. **21**: 1347-1355.
22. Betzig, E., et al. (2006) Imaging Intracellular Fluorescent Proteins at Nanometer Resolution. *Science*. **313**: 1642-1645.
23. Rust, M.J., M. Bates, and X. Zhuang (2006) Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM). *Nat Meth*. **3**: 793-796.
24. Ashkin, A., et al. (1986) Observation of a single-beam gradient force optical trap for dielectric particles. *Optics Letters*. **11**: 288-290.
25. Neuman, K.C. and S.M. Block (2004) Optical trapping. *Rev. Sci. Instr.* **75**: 2787-2809.
26. Block, S.M., L.S.B. Goldstein, and B.J. Schnapp (1990) Bead movement by single kinesin molecules studied with optical tweezers. *Nature*. **348**: 348-352.
27. Lang, M.J., et al. (2002) An Automated Two-Dimensional Optical Force Clamp for Single Molecule Studies. *Biophys. J*. **83**: 491-501.
28. Liphardt, J., et al. (2001) Reversible unfolding of single RNA molecules by mechanical force. *Science*. **292**: 733-737.
29. Onoa, B., et al. (2003) Identifying kinetic barriers to mechanical unfolding of the T-thermophila ribozyme. *Science*. **299**: 1892-1895.
30. Green, L., et al. (2008) Characterization of the mechanical unfolding of RNA pseudoknots. *J. Mol. Biol.* **375**: 511-528
31. Kasianowicz, J.J., et al. (1996) Characterization of individual polynucleotide molecules using a membrane channel. *Proc. Nat. Ac. Sc. USA*. **93**: 13770-13773.
32. Smeets, R.M.M., et al. (2006) Salt dependence of ion transport and DNA translocation through solid-state nanopores. *Nano Lett.* **6**: 89-95.
33. Li, J., et al. (2001) Ion-beam sculpting at nanometre length scales. *Nature*. **412**: 166-169.
34. Li, J.L., et al. (2003) DNA molecules and configurations in a solid-state nanopore microscope. *Nat. Mater.* **2**: 611-615.

35. Storm, A.J., et al. (2003) Fabrication of solid-state nanopores with single-nanometre precision. *Nat. Mater*. **2**: p. 537-541.
36. Wu, M.Y., et al. (2005) Formation of nanopores in a SiN/SiO_2 membrane with an electron beam. *Appl. Phys. Lett*. **87**: 113106.1-113106.3.
37. van den Hout, M., et al. (2010) Controlling nanopore size, shape, and stability. *Nanotechnology*. **21**: 115304.1-115304.5.
38. Wu, M.-Y., et al. (2008) Control of Shape and Material Composition of Solid-State Nanopores. *Nano Lett.* **9**: 479-484.
39. Storm, A.J., et al. (2005) Translocation of double-strand DNA through a silicon oxide nanopore. *Phys. Rev. E*. **71**: 051903.01-051903.10.
40. Talaga, D. and J. Li (2009) "Single-molecule protein unfolding in solid state nanopores". *J. Am. Chem. Soc*. **131**: 9287–9297.
41. Firnkes, M., et al. (2010) Electrically Facilitated Translocations of Proteins through Silicon Nitride Nanopores: Conjoint and Competitive Action of Diffusion, Electrophoresis, and Electroosmosis. *Nano Lett*. **10**: 2162-2167.
42. Smeets, R.M.M., et al. (2009) Translocation of RecA-Coated Double-Stranded DNA through Solid-State Nanopores. *Nano Lett*. **9**: 3089-3095.
43. McNally, B., M. Wanunu, and A. Meller (2008) Electromechanical Unzipping of Individual DNA Molecules Using Synthetic Sub-2 nm Pores. *Nano Lett.* **8**: 3418-3422.
44. Keyser, U.F., et al. (2006) Direct force measurements on DNA in a solid-state nanopore. *Nat. Phys*. **2**: 473-477.
45. Keyser, U.F., et al. (2006) Optical tweezers for force measurements on DNA in nanopores. *Rev. Sc. Instr.* **77**: 105105.1-.9.
46. Keyser, U.F., et al. (2009) Inserting and Manipulating DNA in a Nanopore with Optical Tweezers. *Micro and Nano Technologies in Bioanalysis: Methods and Protocols*: 95-112.
47. van Dorp, S., et al. (2009) Origin of the electrophoretic force on DNA in solid-state nanopores. *Nat. Phys.* **5**: 347-351.
48. Gerland, U., R. Bundschuh, and T. Hwa (2004) Translocation of structured polynucleotides through nanopores. *Phys. Biol*. **1**: 19-26.
49. Hall, A.R., et al. (2009) Electrophoretic Force on a Protein-Coated DNA Molecule in a Solid-State Nanopore. *Nano Lett.* **9**: 4441-4445.
50. Wanunu, M., et al. (2008) DNA Translocations Governed by Interactions with Solid-State Nanopores. *Biophys. J.* **95**: 4716-4725.
51. Watts, J.M., et al. (2009) Architecture and secondary structure of an entire HIV-1 RNA genome. *Nature*. **460**: 711-716.
52. Huang, B. et al. (2008) Three-dimensional super-resolution imaging by Stochastic Optical Reconstruction Microscopy. *Science*. **319**: 810-813.

Chapter 2

An RNA Toolbox for Single-Molecule Force Spectroscopy Studies

Abstract

Precise, controllable single-molecule force spectroscopy studies of RNA and RNA-dependent processes have recently shed new light on the dynamics and pathways of RNA folding and RNA-enzyme interactions. A crucial component of this research is the design and assembly of an appropriate RNA construct. Such a construct is typically subject to several criteria. First, single-molecule force spectroscopy techniques often require an RNA construct that is longer than the RNA molecules used for bulk biochemical studies. Next, the incorporation of modified nucleotides into the RNA construct is required for its surface immobilization. In addition, RNA constructs for single-molecule studies are commonly assembled from different single-stranded RNA molecules, demanding good control of hybridization or ligation. Finally, precautions to prevent RNase- and divalent cation-dependent RNA digestion must be taken. The rather limited selection of molecular biology tools adapted to the manipulation of RNA molecules, as well as the sensitivity of RNA to degradation, make RNA construct preparation a challenging task. We briefly illustrate the types of single-molecule force spectroscopy experiments that can be performed on RNA, and then present an overview of the toolkit of molecular biology techniques that is at one's disposal to assemble such RNA constructs. Within this context, we evaluate the RNA construct assembly protocols in terms of their effectiveness in producing long and stable RNA constructs.

This chapter was published as: Igor D. Vilfan, Wiecher Kamping, Michiel van den Hout, Andrea Candelli, Susanne Hage, and Nynke H. Dekker, *Nucleic Acids Research.* **35**: 6625–6639 (2007)

2.1 Introduction

RNA molecules can be divided into several classes that are responsible for a diverse set of metabolic and regulatory functions *in vivo*. As early as the 1940s, it was shown that RNA molecules are the carriers of the genetic information in Tobacco Mosaic Virus[1]. Since then, other viruses with so-called RNA genomes have been reported in the literature[2-4]. In 1956, "DNA-like RNA" was discovered[5], which serves as an intermediate between DNA and proteins and was later renamed messenger RNA. Subsequent identification and characterization showed that RNA molecules can perform a diverse set of biological reactions such as peptidyl transferase reaction, degradation and ligation of nucleic acids, and recognition of metabolites[6-17]. Furthermore, a diverse class of short RNA molecules involved in the regulation of gene expression has recently emerged[18-22].

A host of bulk biophysical techniques have been successfully applied to elucidate the thermodynamics and kinetics of these RNA-dependent processes. More recently, the biological importance of RNA has stimulated a wealth of single-molecule research. The single-molecule approach offers several advantages over more traditional ensemble studies. For example, single-molecule techniques enable sensitive micromanipulation of individual biomolecules (e.g. extension by force, supercoiling)[23-25], which can be used to lay bare the states a molecule can adopt, or to precisely control its initial state. Also, by observing a single molecule or a single event, it is possible to measure individual fluctuations in thermodynamic or kinetic parameters, thus eliminating the averaging inherent in ensemble studies[26]. Furthermore, the amounts of sample required for single-molecule studies are usually far smaller than for ensemble studies[27,28]. Finally, single-molecule detection can be used to investigate biological systems *in vitro* as well as in the living cell[29-31]. In this manner, single-molecule studies of RNA have provided new insights into the polymer physics of RNA[32,33], RNA folding[34-43] and enzyme-RNA interactions[44-49]. For example, single-molecule force spectroscopy has been performed to measure the persistence lengths of both homopolymeric single-stranded RNA[33] and random sequence double-stranded RNA (Figure **2.1a**)[32]. Elegant experiments based on optical tweezers have been carried out by Liphardt and co-workers[38] to measure the folding thermodynamics and kinetics of an RNA hairpin (Figure **2.1b** and **2.1c**). A similar experimental scheme has been used to investigate step-wise unwinding of an RNA hairpin by an RNA helicase (Figure **2.1d**)[45].

In order to study RNA and RNA-dependent processes with single-molecule force spectroscopic techniques, specialized RNA constructs have to be designed that differ from the bulk RNA substrates in several respects: (i) single-molecule force-spectroscopy studies can require long RNAs of lengths greater than ~1 kb[32,33,38,39,50]; (ii) the incorporation of modified nucleotides is often necessary at one or both ends of the RNA construct to immobilize the RNA to a surface or to suspend it between two surfaces, respectively; and (iii) the RNA construct is often composed of different sections that need to be assembled by hybridization or

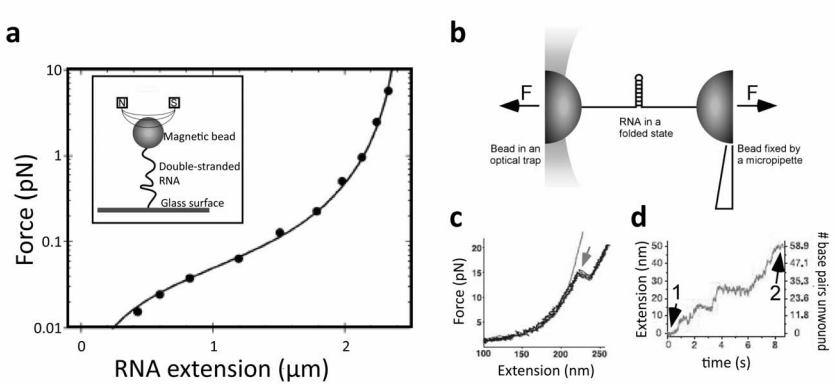


Figure 2.1. *Examples of single-molecule studies of RNA and RNA-dependent processes.* **a.** Measurement of the persistence length of double-stranded RNA (dsRNA) with magnetic tweezers[32]. A force-extension curve for dsRNA is obtained on an 8.3 kbp dsRNA. The experimental points were fitted to the worm-like chain model (smooth line). The fit yielded a persistence length of dsRNA of 63.0 ± 2.5 nm. Inset: The experimental configuration used to measure the persistence length of dsRNA[32]. The dsRNA is suspended between the magnetic bead and glass surface. At a given applied magnetic force, the distance between the glass surface and the magnetic bead is measured, yielding the force-extension curve. **b.** An experimental configuration for single-molecule study of RNA hairpin unfolding by force or through the activity of an RNA helicase[38,45]. The RNA hairpin is suspended between two beads. One bead is optically trapped, while the other bead is held by a micropipette that can be moved with a nanometer precision to apply a force. **c.** An example of a force-extension curve obtained during the RNA hairpin unfolding by externally applied force[38]. The force on the RNA construct is increased by moving the micropipette in the direction indicated. The RNA extension is obtained from measurement of the distance between the centers of the two beads, corrected for the bead diameter. The RNA hairpin unfolding occurs at the point indicated by the red arrow. **d.** An increase in RNA extension at constant force in the presence of RNA helicases[45]. The RNA construct extension is the result of RNA helicase binding to the bottom of the RNA hairpin (position 1) and subsequent unwinding of the RNA hairpin to produce a completely denatured RNA molecule (position 2). Between the binding of RNA helicase and complete denaturation of the RNA hairpin, several plateaus in the construct extension can be observed, suggesting pauses in helicase activity.

ligation. These specific requirements can present several difficulties for the assembly of an RNA construct. For example, the maximum length of RNA that can be chemically synthesized is often limited to ~100 nucleotides. The low yield of longer RNA oligomers renders non-enzymatic RNA synthesis unfeasible for such applications. Enzymatic synthesis of RNA by viral polymerases provides an alternative, but the presence of termination signals and the limited processivity of RNA polymerase nonetheless substantially reduce the yield of RNAs longer than

~10 kb. In addition, the yield of many enzymatic reactions necessary for the assembly of RNA constructs (e.g. RNA ligation) decreases considerably as the RNA length increases beyond ~1 kb[51]. A controlled joining of different sections of RNA construct presents additional challenges due to lack of RNA restriction enzymes[50]. Finally, due to the chemical nature of RNA and the prevalence of RNA-degrading enzymes special precautions need to be taken when working with RNA. High pH (> 9), RNase contamination and/or the presence of divalent cations can decrease the yield of active RNA constructs and thus decrease the probability to observe the biophysical processes of interest. While RNases can be inhibited through the use of RNase inhibitors, the presence of divalent cations often cannot be avoided, since their presence is necessary for a variety of RNA-dependent processes (e.g. RNA folding, RNA polymerization). The limited selection of molecular biology tools for RNA construct synthesis and RNA sensitivity to degradation by RNases and divalent cations make the design of suitable RNA constructs a challenging task.

Here, we provide an overview of the toolkit available for RNA construct assembly, in particular for single-molecule force spectroscopy techniques that require RNA molecules longer than those usually handled in bulk studies or in other single-molecule techniques such as FRET (the assembly protocols for other single-molecule techniques have been reviewed elsewhere[52,53]). We focus on RNA construct design for magnetic and optical tweezers experiments, and illustrate techniques by examples from both the literature and our own results. This article is divided into several chapters. The first section gives an introduction to the basic design of the RNA construct for single-molecule studies, and defines the basic steps of the RNA construct assembly. Examples of RNA constructs from the literature are given to illustrate the type of experiments that can be performed. This is followed by sections that discuss RNA synthesis, labeling, and joining of different sections of a single-stranded RNA construct, which we have evaluated in our laboratory. A separate section is devoted to the assembly of dsRNA constructs, which includes a demonstration of the synthesis of dsRNA molecules longer than 15 kilobases and an evaluation of the ligation requirements for dsRNA molecules. To complete the overview of protocols, we discuss the methods for RNA construct isolation and purification. We conclude by describing the persisting challenges in the field of RNA construct assembly.

Throughout this chapter, we refer to the RNA under consideration either as an *RNA molecule*, or an *RNA construct*. This is to distinguish building blocks (i.e. *the RNA molecules*) from the final product used in the experimental set-up (i.e. *the RNA construct*). Also, as mentioned above, RNA length is an important factor in most of the steps of RNA construct assembly, as it usually affects the yield of the enzymatic reactions. If an exact length of a specific RNA is not given, it is described with qualitative terms *long* or *short*. This is to indicate when the RNA length starts to affect the reaction yield of a particular step in the RNA construct assembly. The approximate length separating the *long* and *short* RNA is discussed for each step of the assembly process.

2.2 Materials and methods

3'-labeling of DNA and RNA oligomers with terminal deoxynucleotidyl transferase

The dry pellets of DNA oligomer with a sequence 5'-GTACCGGCTGTCTGGTATGTATGAG-3', and RNA oligomer with a sequence 5'-UCCGAUGGCUGGCUGCG-3' (Biolegio, The Netherlands) were resuspended in 1x TE (10 mM Tris-HCl, 1 mM EDTA, pH = 8.0) to a final concentration of 500 mM. For every 3'-end labeling reaction, 10 µl reaction mixtures were prepared that contained 100 pmol of either DNA or RNA oligomer, 10 units of *Tdt* (Fermentas, Canada), 1x *Tdt* reaction buffer (Fermentas, Canada) (200 mM potassium cacodylate, 25 mM Tris, 1 mM $CoCl_2$ and 0.01% (v/v) Triton X-100, pH 7.2), and varying concentrations of dUTP (Roche, The Netherlands) or digoxigenin-labeled dUTP (Roche, The Netherlands). The reaction mixtures were incubated at 37 ºC for 1h, and subsequently denatured by the addition of an equal volume of 9 M urea. The denatured samples were loaded on a urea-denaturing PAGE (20% polyacrylamide, 8 M urea, 1x TBE). The 3'-labeling efficiency was evaluated by staining the gel with SYBR® Gold (Invitrogen Molecular Probes, USA) and imaging the bands under UV light.

Synthesis and hybridization of complementary single-stranded RNA molecules

Two 4.1 kilo base pair (kbp) and a 0.5 kbp PCR fragments were generated using pBADb10 (a recombinant pBAD vector) encompassing the sequences from base pair 230 to base pair 4373, and base pair 230 to base pair 681, respectively. The PCR primers (Biolegio, The Netherlands) were designed to incorporate a T7 RNA polymerase promoter sequence at one end of each PCR fragment (Appendix **2**, Figure **A2.1**, Table **A2.1**). In the 4.1 kbp PCR fragments, the T7 RNA polymerase promoter was inserted at position 230 or 4373 for the synthesis of the sense or antisense 4.1 kb single-stranded RNA (4.1 kb ssRNA), respectively (Appendix **2**, Figure **A2.1**). The T7 RNA promoter sequence was inserted at the position 681 in the 0.5 kbp PCR fragment to yield a 0.5 kb ssRNA complementary to the 3'-end of the sense 4.1 kb ssRNA (Appendix **2**, Figure **A2.1**). (The positions of the promoter sequence insertion refer to the base pair in the original pBADb10 vector.) Each 50 µl PCR reaction contained 0.01 U Phusion™ High-Fidelity DNA Polymerase (Finnzymes, Finland), 1x Phusion™ HF (Finnzymes, Finland), 0.2 µM of each primer, 200 µM of each dNTP (Promega, USA), 3% DMSO, and 5 ng of the pBADb10 vector. The PCR temperature cycling was programmed as suggested by the manufacturer with an annealing temperature of 55 ºC. PCR fragments were purified with a

NucleoSpin® Extract II clean-up kit (Macherey-Nagel, Germany). The lengths of the PCR fragments were confirmed on a 0.75% agarose gel.

In vitro transcription of the three ssRNA molecules was performed with T7 RiboMAX™ Large Scale RNA Production System (Promega, USA) according to the manufacturer's protocol. The *in vitro* transcription reactions were purified with RNeasy MiniElute Cleanup kit (Qiagen, USA). The aliquots of purified ssRNA were denatured with RNA Sample Loading Buffer (Sigma-Aldrich, USA) and loaded on a 0.75% agarose gel to confirm the length of the *in vitro* transcripts. The concentrations of ssRNA transcripts were determined spectroscopically by measuring the absorbance at 260 nm using an extinction coefficient of 25 l $cm^{-1}g^{-1}$.

To hybridize the complementary 4 kb ssRNA molecules, the sense and the antisense 4.1 kb ssRNA transcripts were mixed in a 1:1 molar ratio at a total ssRNA concentration of 2 µg per hybridization reaction. The same quantity of the sense 4.1 kb ssRNA was hybridized with 0.5 kb ssRNA at 1:1 and 1:4 molar ratios. The final ssRNA concentrations depended on the hybridization volume, which is noted below for each of the hybridization conditions. Hybridization reactions under aqueous conditions were carried out in 0.5x SSC (Promega, USA) supplemented with 1x TE (Sigma-Aldrich, USA) (SSCTE buffer) with a final volume of 150 µl. Hybridization of the sense 4.1 kb ssRNA and 0.5 kb ssRNA was also performed in 150 µl 1x TE. Hybridization reactions were also performed in a formamide-based buffer used by Liphardt and co-workers[38]. Briefly, 1 µg of sense 4.1 kb ssRNA and 1 µg of antisense 4.1 kb ssRNA were resuspended in 17.5 µl 1x TE and mixed with 80 µl of the formamide buffer (prepared by mixing 100% formamide, 1 M EDTA (pH=8.0), 1 M Pipes (pH=6.3), and 5 M NaCl in a 400:1:20:40 volume ratio, respectively)[38]. Two different temperature programs were tested for hybridization. First, a "gradual-cool" temperature program was adopted from Dekker and co-workers[50] (denaturation for 1h at 65 ºC followed by a hybridization step with 1.5 ºC/5min cooling rate with the final temperature of 25 ºC). Second, a "step-cool" temperature program was adopted from Liphardt and co-workers[38] (denaturation at 85 ºC for 10 min, followed by 1 h incubation at 62 ºC, 1.5 h incubation at 52 ºC, and finished with a cooling to 10 ^{o}C within 10 min. Hybridization reactions were purified with the RNeasy MiniElute Cleanup kit (Qiagen, USA), and analyzed on a 0.75% agarose gel.

Synthesis of double-stranded RNA constructs

Long random sequences of single-stranded RNA (ssRNA) were synthesized using Polynucleotide phosphorylase (*Pnpase*) from *Bacillus stearothermophilus* (Sigma-Aldrich, USA). 4 µg of *Pnpase* was incubated in 100 mM Tris-HCl buffer (pH=9), supplemented with 5 mM $MgCl_2$, 0.4 mM EDTA, 0.04% w/v BSA and 20 mM nucleotide diphosphates (NDP), distributed as follows: CDP, UDP, ADP, and GDP at 9 mM, 9 mM, 1 mM and 1 mM, concentrations, respectively (Dr. Roman Tuma

personal communication). The total reaction volume of 200 µl was incubated at 37 ºC for 15 minutes. The reaction mixture was purified with RNeasy MiniElute Cleanup kit (Qiagen, USA), and the ssRNA eluted with 20 µl 1x TE. To obtain double-stranded RNA (dsRNA), ssRNA was subsequently incubated with 5 µg of Φ6 P2 RNA-dependent RNA polymerase (Finnzymes, Finland) in the presence of 50 mM Tris-HCl (pH=8.9), 80 mM NH_4OAc, 2 mM DTT, 0.1 mM EDTA, 5 mM $MnCl_2$, and 2.5 mM of each NTP. The final reaction volume was 200 µl. The reaction was incubated at room temperature for 1h and subsequently purified with RNeasy MiniElute Cleanup kit (Qiagen, USA) and loaded on a 0.75% agarose gel. 15-20 kbp dsRNA was then gel-extracted using a TAKARA Recochip (Takara Bio, Japan). AFM images of the gel-extracted dsRNA on a Ni-coated mica surface were taken as previously[46] described.

Intramolecular ligation of 5'-overhangs in double-stranded RNA molecules

The sense and antisense 4.1 kb ssRNA were synthesized as described above. In addition, an antisense 4.1 kb ssRNA with an eight base extension at its 5'-end (5'-GGGCUACC-3') was transcribed from a PCR fragment as described previously by Dekker and co-workers[50]. Briefly, a forward primer used for the PCR fragment of the antisense 4.1 kb ssRNA was modified to include a 5'-GCTACC-3' sequence between the T7 promoter sequence and the sequence complementary to the pBADb10 vector (Appendix **2**, Table **A2.1**). The reverse primer was the same as the one used to obtain the PCR fragment for the antisense 4.1 kb ssRNA. The 5'-extended antisense 4.1 kb ssRNA was alternatively transcribed *in vitro* in the absence or in the presence of GMP as described by Ke and co-workers[54] to obtain 5'-triphosphate and 5'-monophosphate transcripts of 4.1 kb ssRNA, respectively. The sense 4.1 kb ssRNA was hybridized to the antisense 4.1 kb ssRNA, and to 5'-extended antisense 4.1 kb ssRNA (alternatively synthesized with or without GMP) in three separate hybridization reactions. The SSCTE buffer and the "gradual-cool" temperature program were used. The hybridization reactions were purified with RNeasy MiniElute Cleanup kit (Qiagen, USA), and analyzed on a 0.75% agarose gel. The intramolecular ligation of the hybrids was carried out in a 50 µl volume containing 2 µg of hybridized dsRNA, 4 U of RNA Ligase 2 (RNL2) (New England BioLabs, MA, USA), 1x RNL2 New England BioLabs buffer, and 20 U Superase•In™ (Ambion, TX, USA). The ligation reactions were incubated at 37 ºC for 1 h and purified with RNeasy MiniElute Cleanup kit (Qiagen, USA). An aliquot of each reaction was denatured with RNA Sample Loading Buffer (Sigma-Aldrich, USA) and loaded on 0.75% agarose gel.

2.3 RNA construct assembly for single-molecule force spectroscopy studies

As illustrated by the examples in Figure **2.1**, single-molecule force spectroscopy can be used to sensitively monitor and manipulate RNA molecules and RNA-dependent processes. We now describe the general principles underlying the assembly of suitable RNA constructs. For force spectroscopy, molecules are typically tethered on both ends to allow the application of force. Therefore, an RNA construct can be divided into a central section and two terminal sections (Figure **2.2a**). The central section contains the RNA sequence that is to be investigated experimentally in terms of its mechanical properties or its interactions with other biomolecules such as proteins or small ligands. The two terminal sections, commonly referred to as '*handles*', are necessary to suspend the RNA construct between the two attachment points (i.e. a glass surface and a magnetic bead in the magnetic tweezers, a glass surface and a small gold particle for the analysis of tethered particle motion (TPM), or two beads in the optical tweezers) (Figure **2.2b**). The attachment of the RNA construct to the glass surface or beads is facilitated by formation of bonds between molecules immobilized on the surface and their ligands incorporated in the terminal sections. Each terminal section has to incorporate a unique ligand to allow for controlled anchoring of the RNA construct to the two surfaces (Figure **2.2b**). The terminal section can be labeled with a single or multiple surface-binding molecules (Figure **2.2c**, **2.2d**)[38,50]. Multiple surface-binding molecules are often needed to increase the stability of anchoring of the RNA construct at high forces. In special cases, single-molecule applications do not require labeled terminal sections on both ends of the central section. For example, force spectroscopy measurements in which forces are applied via the intermediaries of fluid flow or electric fields, as in optical tweezers measurements combined with flow or with nanopores, can be performed with a labeled terminal section attached to only one end of the central section[55,56]. A second example is the case in which the second extremity of the nucleic acid is tethered via the intermediary of surface-immobilized molecular motors (e.g. polymerases, helicases, transferases), in which case a labeled terminal section at only one end of the construct suffices[44,57].

Different protocols have been reported in the literature to assemble RNA constructs[32,33,38,39,45]. These protocols can be divided into several steps that are the focus of this review: (i) synthesis of the central section of the RNA construct; (ii) synthesis of labeled terminal sections of the RNA construct; (iii) joining of the central and terminal sections of the RNA construct; and (iv) isolation and purification of the final RNA construct. While the above four steps describing the assembly of RNA constructs may seem straightforward, many single-molecule studies impose a complication, namely the RNA construct length. Here, we present each step in RNA construct assembly and point out limitations arising from this

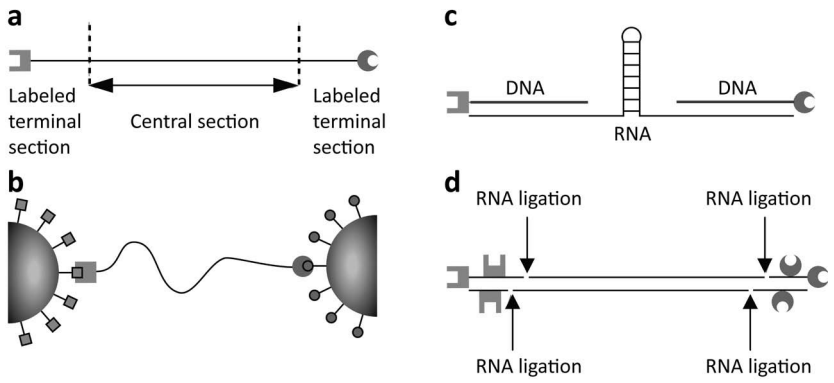


Figure 2.2. *The basic design of the RNA construct for single-molecule experiments.* **a.** The RNA construct is divided into three sections. The central section denoted in black contains the RNA sequence or structure to be studied. The two labeled terminal sections of the RNA construct contain one or multiple nucleotides modified with ligands that can bind to coated surfaces. This enables a controlled mode of RNA construct binding between two surfaces. The scheme shows a single nucleotide modification per terminal section. Modification ligands are shown as red forks and green crescents at the ends of the respective terminal sections. **b**. The RNA construct is suspended between surfaces in the single-molecule experiment. The nucleotide modifications in the terminal sections serve as ligands that bind to surface-immobilized molecules. The surfaces are shown as grey spheres. The surface-immobilized molecules are shown either as red squares or green circles, which specifically bind to the red forks and green crescents, respectively. **c**. A schematic of the RNA construct by Liphardt and co-workers that was successfully used to study the kinetics of RNA folding[38]. The central section contains an RNA hairpin. The terminal sections consist of a DNA:RNA hybrid with single modified nucleotides at the terminal ends of DNA strands. **d**. Schematic of an RNA construct used to measure the persistence length of a double-stranded RNA (dsRNA)[32]. The central section contains dsRNA. The terminal sections have been ligated with RNA Ligase 2 to the central section, and contain multiple modified nucleotides.

particular constraint. For simplicity, we first discuss the complete assembly of RNA constructs for studies of single-stranded RNA (ssRNA), and subsequently generalize this to include the assembly of double-stranded RNA (dsRNA) constructs for single-molecule experiments.

2.3.1 The central section synthesis for single-stranded RNA constructs

The central section of the ssRNA construct may be several kilobases (kb) in length, in part because the central section is usually extended at both ends to provide for subsequent attachment of the labeled terminal sections (Figure **2.2c**). Two main experimental approaches have been adopted for the synthesis of such long central sections; *in vitro* transcription with RNA polymerase[32,38,39], and polymerization of nucleotide diphosphates (NDP) using Polynucleotide phosphorylase (*Pnpase*)[33]. *In vitro* transcription has been used in the synthesis of the central sections shorter than ~15 kb, and in the synthesis of central sections that contain specific sequences (e.g. naturally-occurring RNA sequences, RNA structural motifs)[38,39,45]. In contrast, *Pnpase* has been used to synthesize central sections longer than 15 kb, but could be adopted only for random or homopolymeric RNA sequences[33] (Dr. Roman Tuma personal communication). Alternatively, the central section that contains only the naturally occurring RNA sequence and does not require any sequence extensions for subsequent attachment of the labeled terminal sections can also be isolated directly from its biological source[34].

DNA templates for *in vitro* transcription of the central section of the single-stranded RNA construct

In vitro transcription is most frequently used to synthesize the central section of the ssRNA construct. In this case, the first step is the design of a DNA fragment that serves as a template for *in vitro* run-off transcription. The optimal strategy for obtaining this DNA template depends on the type of RNA sequence desired. Cloning of a DNA sequence into a plasmid is commonly used to study short, naturally-occurring RNA sequences with a specific biological function (e.g. ribozyme, t-RNA) or to produce a central section of a ssRNA construct that folds into a specific RNA structural motif (e.g. a hairpin)[38,39,45]. In contrast, a transcription termination signal-free section of a plasmid sequence is usually sufficient to generate kb-long central sections that can be used to study polymer properties of RNA (e.g. persistence length)[32]. These two basic approaches require different protocols that we now discuss.

Most of the single-molecule studies of naturally-occurring RNA sequences and structural motifs take advantage of the protocol developed by Liphardt and co-workers[38] to obtain DNA templates for *in vitro* transcription. First, two synthetic DNA oligomers are hybridized to form a primary DNA fragment containing the sequence of interest in the center and two different restriction overhangs at the respective ends. The primary DNA fragment is then cloned into a plasmid that has been doubly digested to yield overhangs complementary to the overhangs of the

primary DNA fragment. Finally, an RNA-polymerase promoter sequence for *in vitro* transcription (Table **2.1**) is inserted upstream of the cloned sequence by PCR to yield a secondary DNA fragment. The choice of primers in this PCR step determines the extent to which the final RNA sequence extends up- and down-stream of the cloned primary DNA fragment. These extensions at the ends of the central section enable subsequent joining of the labeled terminal sections of the ssRNA construct (please, see below for the discussion on labeling of ssRNA constructs)[38,39]. This approach has been limited to relatively short, naturally occurring sequences and RNA structural motifs (< 120 base pairs) due to the length limitations of the chemical synthesis of DNA oligomers. To circumvent this problem, Mangeol and co-workers have chosen to clone cDNA sequences obtained from the RNA to be studied[39]. Once the cDNA sequence was cloned into a plasmid, the secondary DNA fragment was obtained as in the protocol described above. This approach was successfully applied to study the unfolding and refolding of 173 and 735 nucleotide long fragments of 23S *E. Coli* ribosomal RNA[39].

A simplified DNA template synthesis scheme can be utilized for single-molecule studies that do not require a specific naturally-occurring RNA sequence or structural motif. In order to produce these central sections by *in vitro* transcription, it is common to amplify a section of a plasmid with PCR using a pair of primers, one of which contains an RNA-polymerase promoter sequence at its 5'-end[32,39,50]. The choice of plasmid for this type of DNA template synthesis is crucial if the length of RNA required for single-molecule studies exceeds kilobases[32,33,39], since the probability of encountering transcription termination signals that cause abortive RNA transcripts starts to increase significantly. Prediction of the number of transcription termination signals is challenging, since the signals themselves can be quite diverse and indeed vary for different RNA polymerases. For example, it has been shown that certain sequences and structural motifs on the DNA template and in the nascent RNA can affect the activity of bacteriophage T7 RNA polymerase[58-62]. Due to the complexity of transcription termination signals, it is nearly impossible to

Type of RNA Polymerase	Promoter sequence
SP6 RNA Polymerase	5'-ATTTAGGTGACACTATA**G**AAGNG-3'
T3 RNA Polymerase	5'-AATTAACCCTCACTAAA**G**GGAGA-3'
T7 RNA Polymerase	5'-TAATACGACTCACTATA**G**G-3'

Table 2.1. *Promoter sequences of commonly used viral RNA polymerases.* The bolded base indicates the first nucleotide to be transcribed by a given viral RNA polymerase. The bases downstream from the TATA box are often not reported as part of the promoter sequence. However, these bases often influence the promoter strength. The reported bases are known to enhance the promoter strength.

predict termination sites for even the little-regulated viral RNA polymerases commonly used for *in vitro* transcription. The problem of termination signals is further compounded when designing DNA fragments for dsRNA constructs (please, see the section below on dsRNA constructs), since in this case both DNA strands have to be free of termination signal. Hence, rather than searching for the termination signals within a given DNA sequence, it is typically more efficient to use plasmids that have a demonstrated lack of termination signals for a given RNA polymerase[32,38,39,50].

In vitro transcription of the central section of the single-stranded RNA construct

In vitro transcription of RNA from the DNA template is routinely carried out using viral RNA polymerases (e.g. T7 or SP6 RNA polymerase). Viral RNA polymerases have a number of advantages: (i) The ability to transcribe in the absence of transcription factors; (ii) Selective transcription initiation from well-defined promoter sequences; and (iii) High polymerization rates[63]. However, there are also several known issues that need to be considered when *in vitro* transcription of single-stranded RNA molecules is performed. For instance, viral RNA polymerases tend to initiate transcription from within the promoter sequence (Table **2.1**). This results in an extension of the RNA transcript at its 5'-end by an RNA sequence of a defined length and base composition[64,65]. Furthermore, viral RNA polymerases are known to add a random number of nucleotides to the 3'-end of the RNA transcripts[65-68]. The exact sequence composition and length of the 3'-end extension are not known, although the relative base composition of the 3'-extension can be influenced by the relative concentrations of the free nucleotides during the *in vitro* transcription (data not shown).

Several protocols have been developed in bulk studies to overcome this limitation and obtain well-defined 3'-ends. First, the random extension at the 3'-end has been corrected with site-specific RNase H cleavage, stimulated by the hybridization of a short oligomer to the cleavage site[69-74]. Also, *cis*- or *trans*-acting ribozyme sequences have been attached to the 3'-end of the RNA transcripts to obtain homogenous 3'-ends after *in vitro* transcription[75-77]. However, these site-specific RNase H cleavage and ribozyme sequences have only been used successfully with high specificity for editing the 3'-ends of RNA transcripts shorter than 1 kb[69,75-77]. In the case of RNase H cleavage, the presence of RNA secondary structure can prevent specific binding of short oligomers to the cleavage site[69]. In addition, the folding of the ribozyme sequence into its active structure may be less efficient when attached to an RNA sequence longer than 1 kb[77]. Therefore, to prevent the random 3'-end extension of RNA transcripts longer than 1 kb, a more successful strategy may be the use of a DNA template obtained with a reverse PCR primer that contains 2'-methoxyl modifications at its last two 5' nucleotides. In this

case random nucleotide addition should be prohibited, if T7 RNA polymerase is used[78].

In addition to the terminal extensions of the *in vitro* transcripts, abortive transcripts are a common problem associated with *in vitro* transcription. In the absence of the transcription termination sites on DNA template and nascent RNA, most of the commercially available viral RNA polymerases readily transcribe DNA templates of up to ~15 kb in length. The abortive transcripts are produced when transcription termination sites are present on the DNA template and/or nascent RNA. Furthermore, limited RNA polymerase processivity can also cause the occurrence of the abortive transcripts. Thus, the onset of abortive transcription depends on the type of RNA polymerase used to transcribe RNA as well as on the particular plasmid from which the DNA template is derived.

Polymerization of long random and homopolymeric central sections of single-stranded RNA constructs

An alternative synthesis scheme can be used to obtain ssRNA molecules of ~20-30 kb in length, provided a known RNA sequence is not required. This scheme relies on the use of the *Pnpase* enzyme for the synthesis of the central section of the ssRNA construct[33]. Although *Pnpase* functions *in vivo* as an RNA exonuclease[79,80], it has been used *in vitro* to polymerize nucleotide diphosphates (NDP) into long RNA polymers in a primer-independent fashion[33,81,82]. When the *Pnpase* is provided with only one type of NDP, the enzyme will generate a homopolymer. Alternatively, an incubation of the enzyme with all four NDPs will generate random RNA sequences. While the advantage of *Pnpase* over RNA polymerases is its ability to synthesize RNA molecules longer than 10 kb, *Pnpase*-assisted RNA synthesis results in a wide distribution of RNA molecule lengths. The resulting absence of well-defined lengths can present difficulties in characterizing RNA molecules in the single-molecule experiment[33].

2.3.2 Labeling of the terminal sections of the single-stranded RNA construct

Having described the techniques that can be used to synthesize the central section of the ssRNA construct for use in single-molecule force spectroscopy experiments, we next address the introduction of labeled terminal sections for tethering of such RNA molecules to the glass or bead surfaces (Figure **2.2a** and **2.2b**). The anchoring of each terminal section of the ssRNA construct is achieved by the formation of a strong bond between the surface-immobilized molecule and its ligand in the terminal section. The most commonly used bonds are those between digoxigenin

and its monoclonal antibody (i.e. antidigoxigenin) or between biotin and streptavidin. The surface is usually coated with either antidigoxigenin or streptavidin, whereas the corresponding terminal sections of the ssRNA constructs are labeled with digoxigenin or biotin, respectively. As an alternative, fluorescein-labeled RNA molecules have been recently used for immobilization on the antifluorescein-coated surface[34]. Even though Lambert and coworkers[34] applied the fluorescein-labeled RNA in the context of a TPM experiment in the absence of force, it has been used in DNA single-molecule force spectroscopy experiments[83,84].

Labels in the terminal sections of the ssRNA construct are introduced by incorporating digoxigenin- or biotin-modified nucleotides into its terminal sections. The terminal section of ssRNA construct are principally obtained in three ways: (i) chemical modifications of the 3'- and 5'-end of the central section can be utilized for the attachment of a variety of labels (e.g. biotin)[34,85-87], (ii) by direct enzymatic extension of the 3'- and 5'-ends of the ssRNA construct central section itself with modified nucleotides, or (iii) by the hybridization of a labeled RNA or DNA complementary to the ends of the central sections (Figure **2c**).

Chemical labeling of the central sections of single-stranded RNA construct

A selective 3'- or 5'-end chemical labeling of the ssRNA central section has been the least utilized way of labeling RNA constructs in single-molecule experiments. However, this method holds several advantages over the more common enzymatic labeling and hybridization of complementary labeled strands discussed below, namely high efficiency even at ssRNA lengths longer than 1 kb, and insensitivity to the sequence at the end of the central section.

Chemical labeling consists of two steps[34,85-87]. First, the 3'- or 5'-end of the ssRNA central section is chemically activated. In the subsequent step, a chemically activated label is attached to the activated end of the ssRNA central section. The most common chemical activation of 3'-end of ssRNA is periodate oxidation resulting in a *cis*-diol ribose modification at the 3'-end[85,86]. The *cis*-diol ribose modified 3'-end then readily reacts with hydrazide-activated labels such as biotinamidocaproyl hydrazide[85,86]. The 5'-end labeling of the ssRNA construct can be carried out with initial chemical activation of the 5'-end in the presence of carbohydrazide and EDC (1-ethyl 3-(3-diethylaminopropyl) carbodiimide), followed by a reaction with isothiocyanate derivative of the label of interest[34,87]. These chemical labeling protocols have been successfully applied to 3'- and 5'- labeling of 1.6 kb long 16S rRNA from *E. Coli* and synthetic poly(U) ssRNA molecules of up to 5.5 kb[34].

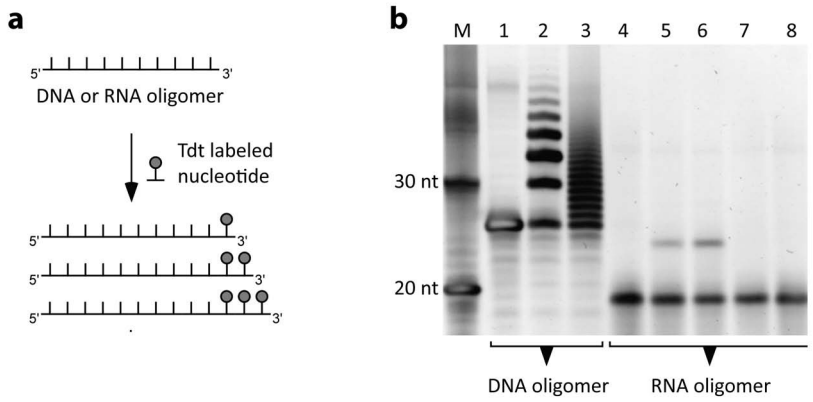


Figure 2.3. *3'-end labeling of DNA and RNA oligomers using Terminal deoxynucleotidyl transferase (Tdt).* **a**. DNA or RNA oligomers are incubated with *Tdt* in the presence of labeled deoxynucleotides, which are added to the 3'-end of the DNA or RNA oligomers by the *Tdt*. Usually, more than one labeled deoxynucleotide is added to the DNA oligomers. Previously, the addition of multiple labeled deoxynucleotide was reported for the RNA oligomers[75]. **b**. An agarose gel of *Tdt* labeling reactions in the presence of DNA or RNA oligomers and varying concentrations of dTTP, dUTP or digoxigenin-labeled dUTP (dig-dUTP). DNA or RNA oligomers were incubated with *Tdt* and respective deoxynucleotide in the 1 × *Tdt* reaction buffer for 1h at 37 ^{o}C (Materials and Methods). The reactions were stopped with the addition of EDTA and denatured with urea prior to loading them on a denaturing 20% PAGE. Lanes: M, single-stranded DNA marker. The 20 and 30 nucleotide long oligomers are indicated; 1, DNA oligomer incubated in the absence of deoxynucleotides; 2, DNA oligomer labeled with 50 µM dig-dUTP; 3, DNA oligomer extended at its 3'-end with 50 µM dTTP; 4, RNA oligomer incubated in the absence of deoxynucleotides; 5, RNA oligomer labeled with 50 µM dig-dUTP; 6, RNA oligomer labeled with 250 µM dig-dUTP; 7, RNA oligomer incubated with 50 µM of dTTP; 8, RNA oligomer incubated with 10 mM dTTP.

Enzymatic labeling of the ends of the central section of the single-stranded RNA construct

Enzymatic extension of the central section of the RNA construct can be carried out by several different enzymes (e.g. Poly(A) polymerase (*Pap*), Terminal deoxynucleotidyl transferase (*Tdt*), and T4 RNA Ligase 1 (RNL1)). Most of the labeling efficiencies cited in the literature have been measured for RNA oligomers[88-90]. It is expected that the yield of labeling ssRNA for a typical single-molecule experiment is lower than for an RNA oligomer. However, even a relatively low efficiency of enzymatic labeling of the 3'- and 5'-ends of single stranded RNA

molecules does not necessarily pose an insurmountable problem in single-molecule research, provided the unlabeled ssRNA constructs do not anchor and thus do not interfere with single-molecule experiments.

Direct labeling is possible for both the 3'- and 5'-ends of RNA molecules. The 3'-end of the central section can be extended with a single modified nucleotide by using *Pap* in the presence of a 3'-deoxy version of the modified nucleotide which, following its incorporation, prevents any further addition of modified nucleotides at the new 3'-end[89]. This reaction typically has an efficiency of approximately 50% for RNA oligomers (data not shown)[89]. Similarly, the enzyme *Tdt* has also been reported to be able to attach labeled deoxynucleotides at the 3'-end of RNA oligomers with at least 50% efficiency (Figure **2.3a**)[90]. Our experiments indicate however, that the *Tdt* labeling efficiency of RNA oligomers is consistently below 10% (Figure **2.3b**). While *Tdt* efficiently adds multiple modified or unmodified deoxynucleotides at the 3'-end of DNA oligomers (Lanes 2, and 3 in Figure **2.3b**, respectively), it exhibits poor modification yield for RNA oligomers even at higher concentrations of deoxynucleotides (Lanes 4 to 8 in Figure **2.3b**). It should be noted, that a comparison of Lanes 5 and 6 to Lanes 7 and 8 reveals an intriguing higher modification yield when modified deoxynucleotides are used with RNA oligomers (Figure **2.3b**). The 5'-end of ssRNA molecules can be labeled using T4 RNA ligase 1[88]. The labeling efficiency of this method is approximately 50% for RNA oligomers[88].

Hybridization of labeled complementary ssRNA and ssDNA to the ends of the central section of single-stranded RNA construct

An alternative to the enzymatic labeling of the central section of the RNA construct is the hybridization of labeled complementary ssRNA or ssDNA to the ends of the central section (Figure **2.2c**). While the enzymatic labeling of the central section is the faster alternative, it can suffer from low yield due to the length of the central section RNA. Such low yield can be problematic since an incomplete enzymatic modification of ssRNA construct termini may result in a low yield of tethered ssRNA constructs in TPM or tweezers measurements, where one end of the construct binds to the bead and the opposite end is left unlabeled and free to attach to a surface-immobilized protein. In this type of an experimental set-up, the yield of tethered constructs would be decreased by unlabeled construct competition for the binding to the surface-immobilized motors[44].

Labeling via hybridization consists of two steps: (i) synthesis of complementary ssRNA or ssDNA with modified nucleotides or deoxynucleotides, respectively, and (ii) hybridization of the complementary strands to the ends of the central section of the RNA construct. The labeled ssRNA or ssDNA molecules must be shorter than the central section itself (Figure **2.2c**). The labeled ssRNA and ssDNA experimentally used have been 500 to 2000 base pairs in length. This

provides the stability required in tweezer experiments in the presence of applied forces and/or low loading rates[39,91]. Labeling via hybridization offers a greater degree of flexibility since the labeled molecules can contain single or multiple modified nucleotides. Synthesis of the complementary ssRNA transcript with a single modified nucleotide is carried out enzymatically as described above for the central section of the ssRNA construct. The complementary ssRNA labeled with multiple modified nucleotides is obtained by *in vitro* transcription in the presence of modified nucleotides[32,50]. Similarly to the complementary ssRNA, ssDNA can be labeled with a single or multiple modified deoxynucleotides. The labeling of the ssDNA with a single modified nucleotide at the 3'-end can be obtained by labeling of the corresponding dsDNA with *Tdt* prior to denaturation[92]. Alternatively, 3'-end labeling of ssDNA can be achieved by performing a replacement reaction on a corresponding dsDNA using T4 DNA polymerase[38]. Labeling of ssDNA with a single modified deoxynucleotide at the 5'-end can be obtained by labeling of the corresponding dsDNA with T4 RNA ligase 1 (*RNL1*) prior to denaturation[88,90]. More commonly however, 5'-labeled PCR primers can be used to obtain a 5'-labeled dsDNA, which is denatured in the subsequent step to obtain a 5'-labeled complementary ssDNA[38]. To obtain ssDNA with multiple modified nucleotides, a PCR is performed in the presence of modified deoxynucleotides to obtain dsDNA[93]. The ssDNA required is produced by denaturing the dsDNA during the subsequent hybridization step (see below).

Hybridization of the labeled complementary ssRNA or ssDNA to the ends of the central section of the RNA construct is subsequently performed. The basic protocol for the hybridization of complementary ssRNA or ssDNA to the central section of the RNA construct starts with a denaturation step followed by the actual hybridization. The denaturation of labeled ssRNA and dsDNA requires temperatures above 60 ºC and can be promoted by the presence of denaturing agents such as formamide[94]. The subsequent hybridization is usually achieved by slow cooling from the denaturing temperature to the ambient temperature, which promotes the formation of the thermodynamically stable state of RNA:RNA or DNA:RNA hybrids between the ends of the central section and labeled complementary single-stranded nucleic acids. The higher thermodynamic stability of DNA:RNA compared to DNA:DNA allows for the formation of DNA:RNA hybrid when labeled dsDNA is used in formamide-based buffers[94-97]. At the high temperatures required for the denaturation, care must be taken to avoid degradation of RNA. The presence of chelating agents (e.g. EDTA, sodium citrate) is essential to prevent RNA hydrolysis by trace amounts of divalent ions[98-101]. RNA degradation by RNases is typically avoided through the use of commercially available RNase inhibitors, but this does not apply in the case of hybridization, as most of the commercially available RNase inhibitors are inactivated at temperatures above 60 ºC. In this case, care must be taken to use only RNase-free materials.

Several different RNA hybridization protocols have been applied in single-molecule research[32,38,39,50]. They differ in the composition of the denaturing buffer

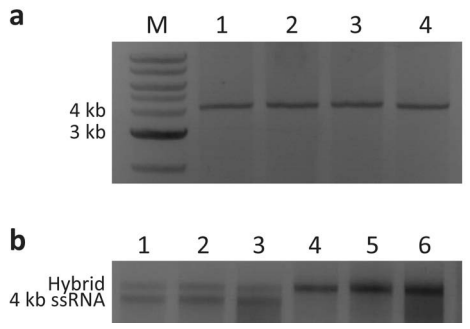


Figure 2.4. *Comparison of hybridization of RNA molecules under various sets of experimental conditions.* **a**. Hybridization of the sense and antisense 4.1 kb single-stranded RNA molecules (4.1 kb ssRNA) in aqueous or formamide denaturing buffer at a 1:1 molar ratio. Lanes: M, double-stranded DNA marker; 1, hybridization of the sense and antisense 4.1. kb ssRNA performed in SSCTE buffer. Hybridization reaction was performed using 'gradual-cool' temperature program; 2, hybridization performed in SSCTE buffer. Hybridization reaction carried out according to "step-cool" temperature protocol; 3, hybridization performed in FEPS buffer. Hybridization reaction was done using the 'gradual-cool' temperature program; 4, RNA hybridization in FEPS buffer using the "step-cool" temperature protocol. **b**. Hybridization of a 0.5 kb ssRNA to the 5'-end of the sense 4.1 kb ssRNA. All the hybridizations were performed using the 'gradual-cool' temperature program. The sense 4.1 kb ssRNA and the hybrid of 0.5 kb ssRNA and the sense 4.1 kb ssRNA are indicated. Lanes: 1-3, hybridization with a 1:1 strand molar ratio in TE buffer (Lane 1), SSCTE buffer (Lane 2) or FEPS buffer (Lane 3); 4-6, hybridization with a 4:1 strand molar ratio with an excess of the 0.5 kb ssRNA in TE buffer (Lane 4), SSCTE buffer (Lane 5) or FEPS buffer (Lane 6).

as well as in the cooling rate used for the hybridization step. A comparison of the denaturing buffer and the cooling rate effects on hybridization of complementary 4.1 kb long ssRNA molecules is shown in Figure **2.4a**. It can be concluded that the complementary RNA molecules hybridize equally well in aqueous-based and formamide-based buffers, as judged by the identical electrophoretic mobilities of hybridization products (Figure **2.4a**). The two temperature programs reported in the literature[38,50] also did not notably affect the final result of hybridization as judged by agarose gel electrophoresis (Lane 1 vs. Lane 2, and Lane 3 vs. Lane 4 in Figure **2.4a**). Conversely, the molar ratio of complementary RNA strands did influence the efficiency of the reaction. Figure **2.4b** shows that the efficiency of hybridization reaction of a 0.5 kb ssRNA to a section of 4.1 kb ssRNA is higher when the short strand is in a molar excess regardless of the denaturation buffer used in the reaction.

A ssRNA construct obtained with the hybridization protocol discussed above has been used successfully in the force-extension experiments employing terminal sections with a single modified nucleotide each[38,39]. This is not surprising as the

high rupture forces of streptavidin-biotin, digoxigenin-monoclonal antibody pairs enable the ssRNA construct with a single modified nucleotide to resist forces up to 100 pN[102-106].

2.3.3 Joining central and labeled terminal sections of the single-stranded RNA construct by ligation

Ligation of the single-stranded central and labeled terminal sections with RNA Ligase 1 (RNL1) can represent an alternative to labeling of the ends of the central section, which often exhibit low efficiency for long ssRNA molecules. In addition, ligation can substitute for hybridization of complementary ssRNA or ssDNA to the ends of the central section, if completely single-stranded RNA construct is required (e.g. measurement of the persistence length of ssRNA). An overview of possible ligation reactions performed by RNL1, which catalyze the formation of a phosphodiester bond between the 3'-OH and 5'-monophosphate termini of nucleic acids, is given in Table **2.2**. The 3'- and 5'-ends of two ssRNA molecules can be ligated by an intermolecular reaction using RNL1[107,108]. The efficiency of this reaction decreases with an increase in ssRNA length and hence often requires high concentrations of the substrate molecules[107,108]. Fortunately, the rate of ligation of two ssRNA molecules can be increased, if a DNA splint is used to bring the 3'- and 5'-end of the two ssRNA molecules into close proximity prior to ligation with RNL1 to form an RNA loop structure[51,107,109-112]. An alternative way to increase the rate of ligation is Y-ligation (Table **2.2**)[107]. Briefly, the 3'- and 5'-end of two ssRNA molecules are hybridized to obtain a short stem structure with free 3'- and 5'-ends. In the Y-ligation step, the free 3'- and 5'-ends are joined by RNL1 to form an RNA loop structure[113]. The length of the resulting RNA loop should not exceed 15 nucleotides[113].

When ligating RNA molecules synthesized by *in vitro* transcription, it should be kept in mind that while the viral RNA polymerases create a 3'-OH at the end of RNA transcripts, they yield a triphosphate at the 5'-end of the RNA transcript. This 5'-triphosphate can be removed after *in vitro* transcription by use of Calf Intestinal Phosphatase (*CIP*) and then the desired monophosphate moiety is added by a kinase[32,50]. Alternatively, the *in vitro* transcription reaction can be supplemented with an excess of nucleotide monophosphate (NMP) corresponding to the first transcribed nucleotide[54]. For example, GMP was added in a twenty time molar excess compared to GTP in a transcription reaction with T7 RNA polymerase[45]. The 20:1 molar excess of GMP to GTP ensures a 95% yield of transcripts with a monophosphate at their 5'-end, assuming equal binding affinity and reaction kinetics of the two nucleotides[54].

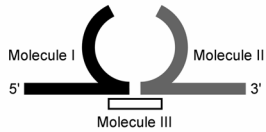
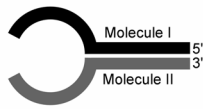
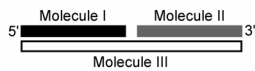
Ligation type	Ligation structure	Ligation conditions
Single-stranded RNAs ligation with RNL1		Molecule I: RNA Molecule II: RNA or DNA
Splint ligation with RNL1		Molecule I: RNA or DNA Molecule II: RNA Molecule III: DNA Resulting hairpin length < 8nt
Y-ligation with RNL1		Molecule I: RNA Molecule II: RNA or DNA Resulting hairpin length < 15nt
Nick ligation with RNL2		Molecule I: RNA Molecule II: RNA or DNA Molecule III: RNA or DNA

Table 2.2. *An overview of RNA ligation schemes using RNA ligase 1 (RNL1) or RNA ligase 2 (RNL2).* The directionality of the ligated nucleic acid strands are denoted by the non-ligated end of each strand. 'Ligation type' column includes a common name for a particular ligation reaction and RNA ligase type to be used for a particular ligation. RNA Ligase 1, and RNA Ligase 2 are abbreviated with RNL1 and RNL2, respectively. The 'Ligation structure' column provides a schematic representation of the RNA structure to be ligated. The 'Ligation conditions' column lists the combinations of nucleic acids resulting in an efficient ligation[107,109].

2.3.4 Double-stranded RNA constructs

So far, the assembly of ssRNA constructs has been discussed due to its importance in single-molecule studies of naturally occurring ssRNA sequences and RNA structural motifs. However, the interest in dsRNA has steadily grown in recent years, prompted by a growing appreciation of dsRNA as an important molecule in the cellular regulatory processes[18-22]. As in the case of ssRNA constructs several different protocols have been applied to obtain dsRNA constructs for single-molecule studies. Dekker and co-workers have obtained approximately 8 kbp long dsRNA by *in vitro* transcribing the antisense and sense RNA strands from the same DNA fragment with a subsequent hybridization to form a dsRNA under the solution conditions discussed above[32,50]. In chapter **5** we describe an even longer dsRNA construct of 12.7 kbp, also obtained through hybridization of two transcribed

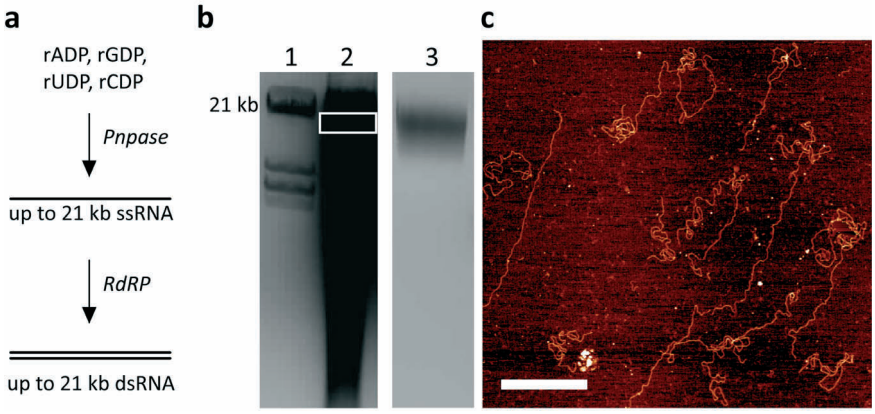


Figure 2.5. *Synthesis of dsRNA with a combination of Polynucleotide phosphorylase (Pnpase) and RNA-dependent RNA polymerase (RdRP) yields longer dsRNA molecules than standard in vitro transcription.* **a**. A reaction scheme for dsRNA synthesis. ssRNA molecules are first synthesized enzymatically with *Pnpase* in the presence of four nucleotide diphosphates (rADP, rGDP, rUDP, and rCDP). In the subsequent step, ssRNA molecules are used as template strands for *RdRP* replication reaction. **b**. An agarose gel of the products of dsRNA synthesis with *PnPase* and *RdRp*. Lanes: 1, double-stranded DNA marker obtained with digestion of 49kb λ DNA. The 21 kbp double stranded DNA band is indicated on the left; 2, dsRNA products of the *RdRp* replication reaction. The length of synthesized dsRNA varies from several hundred base pairs to molecules longer than 20 kbp. This is due to *de novo* initiation of *Pnpase* polymerization that yields a broad length distribution of ssRNA molecules in the first step of the reaction scheme in Figure **2.5a**; 3, the gel-extracted dsRNA. The range of RNA lengths extracted is indicated by the white rectangle in Lane 2. **c**. A representative AFM image of the gel-extracted dsRNA from Lane 3 in Figure **2.5b**. The gel-extracted dsRNA was adsorbed on a Ni-functionalized mica surface and imaged as described in the Materials and Methods. We measured an average dsRNA length of 5 µm, which corresponds to the length of approximately ~20 kbp dsRNA. The scale bar represents 1 µm.

complementary strands. The length of these dsRNA molecules is limited by the same factors affecting ssRNA *in vitro* transcription. To circumvent the limits of *in vitro* transcription, we have recently used RNA-dependent RNA polymerase (*RdRP*) to obtain a double-stranded RNA molecule longer than 13 kbp (see Chapter **5**) by using ssRNA molecule obtained with *Pnpase* as a template (Figure **2.5a**). Using this scheme, the products of the *Pnpase* polymerization reaction were used to obtain dsRNA molecules (Lane 2 in Figure **2.5b**). The dsRNA molecules between 15 to 25 kbp in length were then gel-extracted (Lane 3 in Figure **2.5b**) and imaged with AFM to verify their lengths (Figure **2.5c**). The combined use of *Pnpase* and *RdRP* for the synthesis of dsRNA constructs longer than 15 kbp (Figure **2.5a**) does not enable a

control over the final sequence of dsRNA since it can be used only to synthesize homopolymeric or random dsRNA sequences. This limits the use of these dsRNA construct to studies that require either a homopolymeric RNA sequence or permit the use of an unknown random RNA sequence.

Terminal sections of the double-stranded RNA constructs

By analogy with the methods used to assemble DNA constructs for single-molecule manipulation[114], the addition of terminal sections to a dsRNA construct has been achieved by ligation of one double-stranded terminal section to each end of the double stranded central section of the RNA construct. This is most effectively accomplished using RNA Ligase 2 (RNL2) (Figure **2.2d**)[50], which has a higher affinity for nicked dsRNA substrates compared to RNL1[115]. The double-stranded terminal sections were labeled with multiple modified nucleotides on both strands (Figure **2.2d**)[50]. One end in each terminal section was engineered to contain an overhang that was complementary to the one of the engineered overhangs on the double stranded central section (Figure **2.2d**)[50]. After hybridization of the complementary overhangs, the junctions were sealed with RNL2 (Figure **2.2d**). Figure **2.6a** shows an agarose gel of unligated control and products of a ligation of a double-stranded central section and two double-stranded terminal sections through complementary overhangs using RNL2. The additional two bands with lower electrophoretic mobility in Lane 3 represent the two products of the ligation reaction, namely the 4.1 kbp dsRNA with 0.5 kbp dsRNA ligated to one or both of its ends (Figure **2.6a**). Comparison of Lane 2 with Lane 3 in Figure **2.6a** suggests that the ligation reaction did not go to completion, as unligated product was still observed.

There are several factors that can decrease the yield of ligation of the terminal and central sections of the dsRNA construct, including: (i) low rate of ligation due to a low concentration of the overhangs; (ii) folding of the complementary overhangs back on themselves to form hairpins thus producing an intramolecular nick which can be sealed by RNL2 ; (iii) the presence of random extensions at the 3'-end of the terminal and central sections that result from the viral RNA polymerase transcription. These extensions would yield imperfect junctions, that have been shown to be poor substrates for RNL2[112]; and (iv) partial digestion of the complementary overhangs during the ligation reaction would result in an unstable junction or in a gapped junction, that has been shown to be a poor substrate for RNL2[112]. In this particular case, the incomplete ligation reaction is not due to the low concentration of overhangs, as the yield of the ligation reaction did not increase with longer reaction times (data not shown). Furthermore, the intramolecular hairpin ligation did not affect the yield of intermolecular ligation during the ligation time as the yield of intramolecular ligation of a 4.1 kbp dsRNA with a 5'-overhang at one end was negligible. This was suggested by the low intensity of 8.2 kb long ssRNA (Lane 3 in Figure **2.6b**) that would result from such

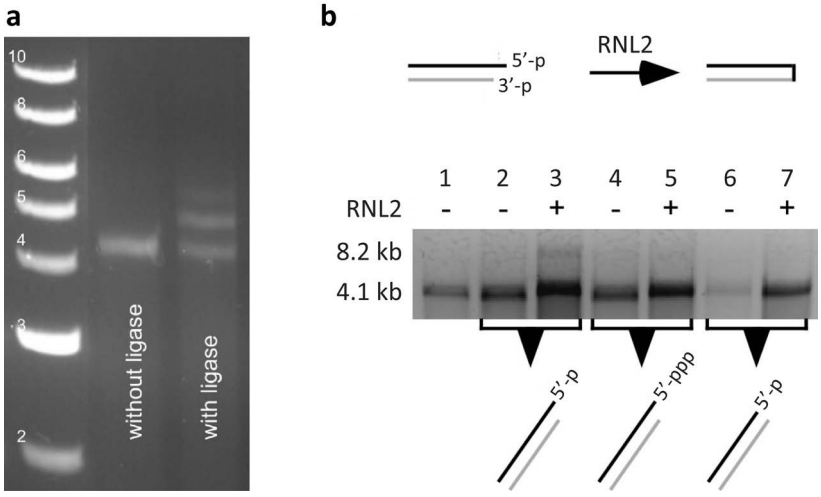


Figure 2.6. *Ligation of double-stranded RNA molecules with complementary 5′-overhangs using RNA Ligase 2 (RNL2).* **a**. Intermolecular ligation of a 0.5 kbp dsRNA molecule to each end of a 4.1 kbp dsRNA with RNL2 through complementary 5′-overhangs (Figure 2D) *(gel image from[50])*. Lanes: 1, double-stranded DNA marker. The corresponding dsDNA lengths are noted next to each band; 2, 4.1 kbp dsRNA with 5′-overhang at both ends incubated in the RNL2 ligation buffer; 3, ligation of 0.5 kbp dsRNA to the 3′- and 5′-ends of 4.1 kbp dsRNA with RNL2. **b**. The efficiency of intramolecular ligation with RNL2 was tested with a 4.1 kbp dsRNA with a monophosphate at its 5′-overhang (activated dsRNA). Two control molecules 4.1 kbp dsRNA with a triphosphate at its 5′-overhang (deactivated dsRNA), and the 4.1 kbp dsRNA with a blunt end (blunt dsRNA) were also incubated with RNL2. Lanes: 1, 4.1 kb ssRNA; 2, activated dsRNA; 3, activated dsRNA ligated with RNL2; 4, deactivated dsRNA; 5, deactivated dsRNA ligated with RNL2; 6, blunt dsRNA; 7, blunt dsRNA ligated with RNL2. The unligated dsRNA molecules run as 4.1 kb ssRNA after denaturation whereas intramolecularly ligated dsRNA molecules run as 8.2 kb ssRNA after dentauration. The 4.1 kb and 8.2 kb ssRNA mobilities are indicated in the gel image.

an intramolecular ligation. It is thus possible that the random 3′-end extensions and/or overhang digestion decreased the yield of the dsRNA ligation reaction. If the digestion of the overhangs results in a low yield of the intermolecular ligation reaction, an increase in the RNL2 reaction rate may prove effective. This can be accomplished by using RNA:DNA hybrids as the double-stranded terminal sections, as these have been shown to present a more efficient substrate for RNL2 activity than RNA:RNA junctions (Table **2.2**)[115].

2.3.5 Isolation and purification of RNA molecules and constructs

As we have shown above, the assembly of RNA constructs for single-molecule measurements is a process that consists of several steps, typically requiring purification after completion of each step. In addition, it can be advantageous to isolate the final RNA construct from a complex mixture of side products of RNA construct assembly. The isolation and purification protocols for RNA molecules and constructs reported in the literature are almost exclusively based on agarose gel-extraction and ethanol precipitation, respectively[116]. The stability of RNA molecules and constructs during all these purification and isolation procedures should always be given consideration. Usually, the addition of RNase inhibitors for inactivation of RNases (e.g. RNasin) and the addition of EDTA for chelation of the divalent cations is sufficient to prevent appreciable RNA degradation.

Agarose gel extraction is the most common way of RNA construct isolation. Once a gel slice containing the desired RNA molecule or construct has been selectively cut out of the gel, the RNA can be extracted by centrifuging a crushed gel slice at high speed, which sediments the gel particles while RNA remains in the supernatant[117]. Alternatively, the polymerized agarose can be enzymatically digested to release the RNA construct into the solution[118]. Several precautions must be noted with gel separation and extraction. First, the yield of agarose gel-extraction strongly depends on the length of the RNA construct. In addition to the length of the RNA molecule, we find that the isolation of dsRNA exhibits higher yields and lower degrees of RNA degradation compared to the ssRNA molecules, presumably due to higher stability of dsRNA. Certain homopolymeric sequences of ssRNA can not be detected on agarose gels due to their poor staining with standard nucleic acid dyes. Agarose gel electrophoresis results in poor length separation for solutions with wide and continuous RNA length distributions as in the case of *Pnpase* polymerization products. In cases where gel extraction results in a low yield, HPLC has been shown to be a plausible alternative (*Dr. Roman Tuma personal communication*).

2.4 Conclusions

RNA is involved in many of the most important processes in biology. The ability to make stable and active RNA constructs for single-molecule experiments enables one to complement RNA bulk studies by taking the advantage of the unique property of single-molecule techniques to measure the properties of the biological processes in the absence of ensemble averaging. We have both discussed and analyzed a number of protocols and molecular biology tools for the assembly of RNA construct that have been successfully used to study RNA and RNA-related

processes on a single-molecule level. Many relevant RNA properties and RNA-dependent processes have thus been investigated on a single-molecule level[32,33,35-49], however many still remain relatively unexplored at the single-molecule level (e.g. protein translation). Currently, the scope of many RNA single-molecule studies is limited by the limited ability to synthesize single- or double-stranded RNA molecules of known sequence that are longer than ~15 kb. The progress in the molecular biology of RNA should thus focus on further increasing the maximum length of the attainable RNA of a known sequence. Recently a subsequent use of *in vitro* transcription and replication with *RdRP* has been used by Bamford and co-workers to synthesize long dsRNAs. This protocol successfully reduces the effect of termination signals on the maximum length of attainable dsRNA by requiring only one of the two RNA strands to be transcribed *in vitro*. Consequently, ssRNA and dsRNA of known sequences have an identical length limit, determined by the presence of the termination signals on the transcribed strand and DNA fragment used for transcription. Nevertheless, some single-molecule experiments may require even longer RNA molecules of known sequence (e.g. optical tweezers combined with nanopores). Furthermore, a higher efficiency of labeling with the modified nucleotides and higher RNA stability under RNA construct assembly conditions (i.e. hybridization temperatures above 60 ºC) should turn out beneficial for high-throughput single-molecule studies. Several commercial RNase inhibitors have been optimized to provide for RNA stability at temperatures above 60 ºC (e.g. RNasin), however they are guaranteed their efficiency only for a limited reaction time. In addition to RNases and divalent cations, RNA construct stability can also be affected by the interaction with glass and bead surfaces, which has not been systematically investigated.

2.5 Acknowledgments

We thank Jeroen Abels, Michiel Bruinink, and Peter Veenhuizen for developing preliminary experiments on RNA in our laboratory, and Roman Tuma and Dennis Bamford for a critical reading of the manuscript. This research was supported by FOM and NWO.

Bibliography

1. Pennazio, S. and P. Roggero (2000) Tobacco mosaic virus RNA as genetic determinant: Genesis of a discovery. *Riv. Biol.-Biol. Forum*. **93**: 431-455.
2. Domingo, E., Escarmis, C., Sevilla, N., Moya, A., Elena, S.F., Quer, J., Novella, I.S. and J.J. Holland (1996) Basic concepts in RNA virus evolution. *Faseb J.* **10**: 859-864.

3. Domingo, E. and J.J. Holland (1997) RNA virus mutations and fitness for survival. *Annu. Rev. Microbiol.* **51**: 151-178.
4. Holland, J.J., Delatorre, J.C. and D.A. Steinhauer (1992) Rna Virus Populations As Quasi-Species. *Curr. Top. Microbiol. Immunol.* **176**: 1-20.
5. Volkin, E. and L. Astrachan (1956) Phosphorus Incorporation In Escherichia-Coli Ribonucleic Acid After Infection With Bacteriophage-T2. *Virology*. **2**: 149-161.
6. Stahley, M.R. and S.A. Strobel (2006) RNA splicing: group I intron crystal structures reveal the basis of splice site selection and metal ion catalysis. *Curr. Opin. Struct. Biol.* **16**: 319-326.
7. Weinger, J.S., Parnell, K.M., Dorner, S., Green, R. and S.A. Strobel (2004) Substrate-assisted catalysis of peptide bond formation by the ribosome. *Nat. Struct. Mol. Biol.* **11**: 1101-1106.
8. Moore, M.J. and P.A. Sharp (1993) Evidence For 2 Active-Sites In The Spliceosome Provided By Stereochemistry Of Premessenger Rna Splicing. *Nature*. **365**: 364-368.
9. Kazantsev, A.V. and N.R. Pace (2006) Bacterial RNase P: a new view of an ancient enzyme. *Nat. Rev. Microbiol.* **4**: 729-740.
10. Beebe, J.A. and C.A. Fierke (1994) A Kinetic Mechanism For Cleavage Of Precursor Trna(Asp) Catalyzed By The Rna Component Of Bacillus-Subtilis Ribonuclease-P. *Biochemistry*. **33**: 10294-10304.
11. Steitz, T.A. and J.A. Steitz (1993) A General 2-Metal-Ion Mechanism For Catalytic Rna. *Proc. Natl. Acad. Sci. U. S. A.* **90**: 6498-6502.
12. Smith, D. and N.R. Pace (1993) Multiple Magnesium-Ions In The Ribonuclease-P Reaction-Mechanism. *Biochemistry*. **32**: 5273-5281.
13. Bergman, N.H., Johnston, W.K. and D.P. Bartel (2000) Kinetic framework for ligation by an efficient RNA ligase ribozyme. *Biochemistry*. **39**: 3115-3123.
14. Birikh, K.R., Heaton, P.A. and F. Eckstein (1997) The structure, function and application of the hammerhead ribozyme. *Eur J Biochem*. **245**: 1-16.
15. Brimacombe, R. and W. Stiege (1985) Structure And Function Of Ribosomal-Rna. *Biochem. J.* **229**: 1-17.
16. Serganov, A., Polonskaia, A., Phan, A.T.n., Breaker, R.R. and D.J. Patel (2006) Structural basis for gene regulation by a thiamine pyrophosphate-sensing riboswitch. *Nature*. **441**: 1167.
17. Winkler, W.C., Nahvi, A., Roth, A., Collins, J.A. and R.R. Breaker (2004) Control of gene expression by a natural metabolite-responsive ribozyme. *Nature*. **428**: 281.
18. Ahlquist, P. (2002) RNA-dependent RNA polymerases, viruses, and RNA silencing. *Science*. **296**: 1270-1273.
19. Fire, A., Xu, S., Montgomery, M.K., Kostas, S.A., Driver, S.E. and C.C. Mello (1998) Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. *Nature*. **391**: 806.

20. Haley, B. and P.D. Zamore (2004) Kinetic analysis of the RNAi enzyme complex. *Nat. Struct. Mol. Biol.* **11**: 599-606.
21. Hannon, J. (2002) RNA interference. *Nature*. **418**: 244-251.
22. McManus, M.T. and P.A. Sharp (2002) Gene silencing in mammals by small interfering RNAs. *Nat. Rev. Genet.* **3**: 737-747.
23. Charvin, G., Strick, T.R., Bensimon, D. and V. Croquette (2005) Tracking topoisomerase activity at the single-molecule level. *Annu. Rev. Biophys. Biomolec. Struct.* **34**: 201-219.
24. Williams, M.C., Rouzina, I. and V.A. Bloomfield (2002) Thermodynamics of DNA interactions from single molecule stretching experiments. *Accounts Chem. Res.* **35**: 159-166.
25. Williams, M.C., Wenner, J.R., Rouzina, I. and V.A. Bloomfield (2001) Entropy and heat capacity of DNA melting from temperature dependence of single molecule stretching. *Biophys. J.* **80**: 1932-1939.
26. Tamarat, P., Maali, A., Lounis, B. and M. Orrit (2000) Ten years of single-molecule spectroscopy. *J. Phys. Chem. A*. **104**: 1-16.
27. Hong, F. and D.D. Root (2006) Downscaling functional bioassays by single-molecule techniques. *Drug Discov. Today*. **11**: 640-645.
28. Skinner, G.M. and K. Visscher (2004) Single-molecule techniques for drug discovery. *Assay Drug Dev. Technol.* **2**: 397-405.
29. Mashanov, G.I. and J.E. Molloy (2007) Automatic detection of single fluorophores in live cells. *Biophys. J.* **92**: 2199-2211.
30. Sako, Y. and T. Uyemura (2002) Total internal reflection fluorescence microscopy for single-molecule imaging in living cells. *Cell Struct. Funct.* **27**: 357-365.
31. Sako, Y. and T. Yanagida (2003) Single-molecule visualization in cell biology. *Nat. Cell Biol.* SS1-SS5.
32. Abels, J.A., Moreno-Herrero, F., van der Heijden, T., Veenhuizen, P.T.M., Bruinink, M.M., Dekker, C. and N.H. Dekker (2005) Single-molecule measurements of the persistence length of double-stranded RNA. *Biophys. J.* **88**: 2737-2744.
33. Seol, Y., Skinner, G.M. and K. Visscher (2004) Elastic properties of a single-stranded charged homopolymeric ribonucleotide. *Phys. Rev. Lett.* **93**: 118102.1-118102.4
34. Lambert, M.N., Vocker, E., Blumberg, S., Redemann, S., Gajraj, A., Meiners, J.C. and N.G. Walter (2006) Mg2+-induced compaction of single RNA molecules monitored by tethered particle microscopy. *Biophys. J.* **90**: 3672-3685.
35. Bokinsky, G., Rueda, D., Misra, V.K., Rhodes, M.M., Gordus, A., Babcock, H.P., Walter, N.G. and X.W. Zhuang (2003) Single-molecule transition-state analysis of RNA folding. *Proc. Natl. Acad. Sci. U. S. A.* **100**: 9302-9307.
36. Bokinsky, G. and X.W. Zhuang (2005) Single-molecule RNA folding. *Accounts Chem. Res.* **38**: 566-573.

37. Li, P.T.X., Bustamante, C. and I. Tinoco (2006) Unusual mechanical stability of a minimal RNA kissing complex. *Proc. Natl. Acad. Sci. U. S. A.* **103**: 15847-15852.
38. Liphardt, J., Onoa, B., Smith, S.B., Tinoco, I. and C. Bustamante (2001) Reversible unfolding of single RNA molecules by mechanical force. *Science*. **292**: 733-737.
39. Mangeol, P., Cote, D., Bizebard, T., Legrand, O. and U. Bockelmann (2006) Probing DNA and RNA single molecules with a double optical tweezer. *Eur. Phys. J. E*. **19**: 311-317.
40. Onoa, B. and I. Tinoco (2004) RNA folding and unfolding. *Curr. Opin. Struct. Biol.* **14**: 374-379.
41. Pan, T. and T. Sosnick (2006) RNA folding during transcription. *Annu. Rev. Biophys. Biomolec. Struct.* **35**: 161-175.
42. Tinoco, I., Li, P.T.X. and C. Bustamante (2006) Determination of thermodynamics and kinetics of RNA reactions by force. *Q. Rev. Biophys.* **39**: 325-360.
43. Zhuang, X.W. and M. Rief (2003) Single-molecule folding. *Curr. Opin. Struct. Biol.* **13**: 88-97.
44. Davenport, R.J., Wuite, G.J.L., Landick, R. and C. Bustamante (2000) Single-molecule study of transcriptional pausing and arrest by E. coli RNA polymerase. *Science*. **287**: 2497-2500.
45. Dumont, S., Cheng, W., Serebrov, V., Beran, R.K., Tinoco, I., Pyle, A.M. and C. Bustamante (2006) RNA translocation and unwinding mechanism of HCVNS3 helicase and its coordination by ATP. *Nature*. **439**: 105-108.
46. Henn, A., Medalia, O., Shi, H.P., Steinberg, M., Franceschi, F. and I. Sagi (2001) Visualization of unwinding activity of duplex RNA by DbpA, a DEAD box helicase, at single-molecule resolution by atomic force microscopy. *Proc. Natl. Acad. Sci. U. S. A.* **98**: 5007-5012.
47. Rueda, D., Bokinsky, G., Rhodes, M.M., Rust, M.J., Zhuang, X.W. and N.G. Walter (2004) Single-molecule enzymology of RNA: Essential functional groups impact catalysis from a distance. *Proc. Natl. Acad. Sci. U. S. A.* **101**: 10066-10071.
48. Yin, H., Wang, M.D., Svoboda, K., Landick, R., Block, S.M. and J. Gelles (1995) Transcription against an applied force. *Science*. **270**: 1653-1657.
49. Zhuang, X., Bartley, L.E., Babcock, H.P., Russell, R., Ha, T., Herschlag, D. and S. Chu (2000) A Single-Molecule Study of RNA Catalysis and Folding. *Science*. **288**: 2048-2051.
50. Dekker, N.H., Abels, J.A., Veenhuizen, P.T.M., Bruinink, M.M. and C. Dekker (2004) Joining of long double-stranded RNA molecules through controlled overhangs. *Nucleic Acids Res.* **32**: e140.
51. Stark, M.R., Pleiss, J.A., Deras, M., Scaringe, S.A. and S.D. Rader (2006) An RNA ligase-mediated method for the efficient creation of large, synthetic RNAs. *RNA.* **12**: 2014-2019.

52. Walter, N.G. (2001) Structural dynamics of catalytic RNA highlighted by fluorescence resonance energy transfer. *Methods*. **25**: 19-30.
53. Qin, P.Z.F. and A.M. Pyle (1999) Site-specific labeling of RNA with fluorophores and other structural probes. *Methods*. **18**: 60-70.
54. Ke, A. and J.A. Doudna (2004) Crystallization of RNA and RNA-protein complexes. *Methods*. **34**: 408-414.
55. Amitani, I., Baskin, R.J. and S.C. Kowalczykowski (2006) Visualization of Rad54, a chromatin remodeling protein, translocating on single DNA molecules. *Molecular Cell*. **23**: 143-148.
56. Keyser, U.F., Koeleman, B.N., Van Dorp, S., Krapf, D., Smeets, R.M.M., Lemay, S.G., Dekker, N.H. and C. Dekker (2006) Direct force measurements on DNA in a solid-state nanopore. *Nat. Phys.* **2**: 473-477.
57. Yin, H., Landick, R. and J. Gelles (1994) Tethered Particle Motion Method For Studying Transcript Elongation By A Single Rna-Polymerase Molecule. *Biophys. J.* **67**: 2468-2478.
58. Hartvig, L. and J. Christiansen (1996) Intrinsic termination of T7 RNA polymerase mediated by either RNA or DNA. *Embo J.* **15**: 4767-4774.
59. He, B., Kukarin, A., Temiakov, D., Chin-Bow, S.T., Lyakhov, D.L., Rong, M.Q., Durbin, R.K. and W.T. McAllister (1998) Characterization of an unusual, sequence-specific termination signal for T7 RNA polymerase. *J. Biol. Chem.* **273**: 18802-18811.
60. Jeng, S.T., Gardner, J.F. and R.I. Gumport (1990) Transcription termination by Bacteriophage T7 RNA polymerase at Rho-independent terminators. *J. Biol. Chem.* **265**: 3823-3830.
61. Lyakhov, D.L., He, B., Zhang, X., Studier, F.W., Dunn, J.J. and W.T. McAllister (1998) Pausing and termination by bacteriophage T7 RNA polymerase. *J. Mol. Biol.* **280**: 201-213.
62. Macdonald, L.E., Durbin, R.K., Dunn, J.J. and W.T. McAllister (1994) Characterization of 2 types of termination signal for Bacteriophage T7 RNA polymerase. *J. Mol. Biol.* **238**: 145-158.
63. Logel, J., Dill, D. and S. Leonard (1992) Synthesis of cRNA probes from PCR-generated DNA. *Biotechniques* **13**: 604-610.
64. Melton, D.A., Krieg, P.A., Rebagliati, M.R., Maniatis, T., Zinn, K. and M.R. Green (1984) Efficient in vitro synthesis of biologically active RNA and RNA hybridization probes from plasmids containing a bacteriophage-SP6 promoter. *Nucleic Acids Res.* **12**: 7035-7056.
65. Milligan, J.F., Groebe, D.R., Witherell, G.W. and O.C. Uhlenbeck (1987) Oligoribonucleotide synthesis using T7 RNA polymerase and synthetic DNA templates. *Nucleic Acids Res.* **15**: 8783-8798.
66. Draper, D.E., White, S.A. and J.M. Kean (1988) Preparation of specific ribosomal-RNA fragments. *Method Enzymol.* **164**: 221-237.
67. Kholod, N., Vassilenko, K., Shlyapnikov, M., Ksenzenko, V. and L. Kisselev (1998) Preparation of active tRNA gene transcripts devoid of 3 '-extended products and dimers. *Nucleic Acids Res.* **26**: 2500-2501.

68. Milligan, J.F. and O.C. Uhlenbeck (1989) Determination of RNA-protein contacts using thiophosphate substitutions. *Biochemistry*. **28**: 2849-2855.
69. Donis-Keller, H. (1979) Site specific enzymatic cleavage of RNA. *Nucleic Acids Res.* **7**: 179-192.
70. Lapham, J. and D.M. Crothers (1996) RNase H cleavage for processing of in vitro transcribed RNA for NMR studies and RNA ligation. *RNA*. **2**: 289-296.
71. Li, J. and R.M. Wartell (1998) RNase H1 can catalyze RNA/DNA hybrid formation and cleavage with stable hairpin or duplex DNA oligomers. *Biochemistry*. **37**: 5154-5161.
72. Lima, W.F. and S.T. Crooke (1997) Cleavage of single strand RNA adjacent to RNA-DNA duplex regions by Escherichia coli RNase H1. *J. Biol. Chem.* **272**: 27513-27516.
73. Yu, Y.T. and J.A. Steitz (1997) A new strategy for introducing photoactivatable 4-thiouridine ((4S)U) into specific positions in a long RNA molecule. *RNA*. **3**: 807-810.
74. Zamaratski, E., Pradeepkumar, P.I. and J. Chattopadhyaya (2001) A critical survey of the structure-function of the antisense oligo/RNA heteroduplex as substrate for RNase H. *J. Biochem. Biophys. Methods*. **48**: 189-208.
75. FerreDamare, A.R. and J.A. Doudna (1996) Use of cis- and trans-ribozymes to remove 5' and 3' heterogeneities from milligrams of in vitro transcribed RNA. *Nucleic Acids Res.* **24**: 977-978.
76. Walker, S.C., Avis, J.M. and G.L. Conn (2003) General plasmids for producing RNA in vitro transcripts with homogeneous ends. *Nucleic Acids Res.* **31**: e82.
77. Wichlacz, A., Legiewicz, M. and J. Ciesiolka (2004) Generating in vitro transcripts with homogenous 3 ' ends using trans-acting antigenomic delta ribozyme. *Nucleic Acids Res.* **32**: e39.
78. Kao, C., Zheng, M. and S. Rudisser (1999) A simple and efficient method to reduce nontemplated nucleotide addition at the 3 ' terminus of RNAs transcribed by T7 RNA polymerase. *RNA.* **5**: 1268-1272.
79. Grunbergmanago, M. and S. Ochoa (1955) Enzymatic synthesis and breakdown of polynucleotides - Polynucleotide phosphorylase. *J. Am. Chem. Soc.* **77**: 3165-3166.
80. Grunbergmanago, M., Ortiz, P.J. and S. Ochoa (1955) Enzymatic synthesis of nucleic acid-like polynucleotides. *Science*. **122**: 907-910.
81. Adami, R.C. and K.G. Rice (1999) Metabolic stability of glutaraldehyde cross-linked peptide DNA condensates. *Journal of Pharmaceutical Sciences*. **88**: 739-746.
82. Vanzi, F., Vladimirov, S., Knudsen, C.R., Goldman, Y.E. and B.S. Cooperman (2003) Protein synthesis by single ribosomes. *RNA.* **9**: 1174-1179.
83. Zlatanova, J., Lindsay, S.M., and S.H. Leuba (2000) Single molecule force spectroscopy in biology using the atomic force microscope. *Prog. Biophys. Mol. Biol.* **74**: 37-61.

84. Gore, J., Bryant, Z., Nollmann, M., Le, M.U., Cozzarelli, N.R. and C. Bustamante (2006) DNA overwinds when stretched. *Nature*. **442**: 836-839.
85. Hansske, F. and F. Cramer (1979) Modification of the 3' Terminus of tRNA by Periodate Oxidation and Subsequent Reaction with Hydrazides. *Method Enzymol.* **59**: 172-181.
86. Willkomm, D.K. and R.K. Hartmann (2005) In Hartmann, R. K., Bindereif, A., Schön, A. and E. Westhof (eds.), *Handbook of RNA Biochemistry*. Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim. **1**: 86-94.
87. Bakin, A.V., Borisova, O.F., Shatsky, I.N. and A.A. Bogdanov (1991) Spatial-Organization Of Template Polynucleotides On The Ribosome Determined By Fluorescence Methods. *J. Mol. Biol.* **221**: 441-453.
88. Kinoshita, Y., Nishigaki, K. and Y. Husimi (1997) Fluorescence-, isotope- or biotin-labeling of the 5'-end of single-stranded DNA/RNA using T4 RNA ligase. *Nucleic Acids Res.* **25**: 3747-3748.
89. Martin, G. and W. Keller (1998) Tailing and 3 '-end labeling of RNA with yeast poly(A) polymerase and various nucleotides. *RNA.* **4**: 226-230.
90. Rosemeyer, V., Laubrock, A. and R. Seibl (1995) Nonradioactive 3'-end labeling of RNA molecules of different lengths by Terminal deoxynucleotidyltransferase. *Anal. Biochem.* **224**: 446-449.
91. Green, N.H., Williams, P.M., Wahab, O., Davies, M.C., Roberts, C.J., Tendler, S.J.B. and S. Allen (2004) Single-molecule investigations of RNA dissociation. *Biophys. J.* **86**: 3811-3821.
92. Itakura, K., Rossi, J.J. and R.B. Wallace (1984) Synthesis and use of synthetic oligonucleotides. *Annu. Rev. Biochem.* **53**: 323-356.
93. Koster, D.A., Croquette, V., Dekker, C., Shuman, S. and N.H. Dekker (2005) Friction and torque govern the relaxation of DNA supercoils by eukaryotic topoisomerase IB. *Nature*. **434**: 671-674.
94. Casey, J. and N. Davidson (1977) Rates of formation and thermal stabilities of RNA-DNA and DNA-DNA duplexes at high concentrations of formamide. *Nucleic Acids Res.* **4**: 1539-1552.
95. Bodkin, D.K. and D.L. Knudson (1985) Assessment Of Sequence Relatedness Of Double-Stranded-Rna Genes By Rna-Rna Blot Hybridization. *J. Virol. Methods*. **10**: 45-52.
96. Bolton, E.T. and B.J. McCarthy (1962) A General Method For Isolation Of Rna Complementary To Dna. *Proc. Natl. Acad. Sci. U. S. A.* **48**: 1390.
97. McConaug.Bl, Laird, C.D. and B.J. McCarthy (1969) Nucleic Acid Reassociation In Formamide. *Biochemistry*. **8**: 3289-3295.
98. Butzow, J.J. and G.L. Eichhorn (1975) Different susceptibility of DNA and RNA to cleavage by metal ions. *Nature*. **254**: 358-359.
99. Eichhorn, G.L., Clark, P. and E. Tarian (1969) Interaction of metal ions with polynucleotides and related compounds. 13. Effect of metal ions on enzymatic degradation of ribonucleic acid by bovine pancreatic ribonuclease and of deoxyribonucleic acid by bovine pancreatic deoxyribonuclease I. *J. Biol. Chem.* **244**: 937-942.

100. Kaga, E., Nakagomi, O. and S. Uesugi (1992) Thermal degradation of RNA-RNA hybrids during hybridization in solution. *Mol. Cell. Probes*. **6**: 261-264.
101. Tenhunen, J. (1989) Hydrolysis of single-stranded RNA in aqueous solutions - Effect on quantitative hybridizations. *Mol. Cell. Probes*. **3**: 391-396.
102. Dammer, U., Hegner, M., Anselmetti, D., Wagner, P., Dreier, M., Huber, W. and H.J. Guntherodt (1996) Specific antigen/antibody interactions measured by force microscopy. *Biophys. J.* **70**: 2437-2441.
103. Lee, G.U., Kidwell, D.A. and R.J. Colton (1994) Sensing discrete streptavidin biotin interactions with atomic force microscopy. *Langmuir*. **10**: 354-357.
104. Moy, V.T., Florin, E.L. and H.E. Gaub (1994) Intermolecular forces and energies between ligands and receptors. *Science*. **266**: 257-259.
105. Pincet, F. and J. Husson (2005) The Solution to the Streptavidin-biotin paradox: the influence of history on the strength of single molecular bonds. *Biophys. J.* **89**: 4374-4381.
106. Merkel, R., Nassoy, P., Leung, A. Ritchie, K., and E. Evans (1999) Energy landscapes of receptor-ligand bonds explored with dynamic force spectroscopy. *Nature*. **397**: 50-53.
107. Persson, T., Willkomm, D.K. and R.K. Hartmann (2005) In Hartmann, R. K., Bindereif, A., Schön, A. and E. Westhof (eds.), *Handbook of RNA Biochemistry*. Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim. **1**: 53-74.
108. Sugino, A., Goodman, H.M., Heyneker, H.L., Shine, J., Boyer, H.W. and N.R. Cozzarelli (1977) Interaction of Bacteriophage T4 RNA and DNA ligases in joining of duplex DNA at base-paired ends. *J. Biol. Chem.* **252**: 3987-3994.
109. Bain, J.D. and C. Switzer (1992) Regioselective Ligation Of Oligoribonucleotides Using Dna Splints. *Nucleic Acids Res.* **20**: 4372-4372.
110. Kurschat, W.C., Muller, J., Wombacher, R. and M. Helm (2005) Optimizing splinted ligation of highly structured small RNAs. *RNA.* **11**: 1909-1914.
111. Moore, M.J. and C.C. Query (2000), *Rna-Ligand Interactions Pt A*. **317**: 109-123.
112. Nandakumar, J., Ho, C.K., Lima, C.D. and S. Shuman (2004) RNA substrate specificity and structure-guided mutational analysis of bacteriophage T4 RNA ligase 2. *J. Biol. Chem.* **279**: 31337-31347.
113. Nishigaki, K., Taguchi, K., Kinoshita, Y., Aita, T. and Y. Husimi (1998) Y-ligation: An efficient method for ligating single-stranded DNAs and RNAs with T4 RNA ligase. *Mol. Divers.* **4**: 187-190.
114. Strick, T.R., Allemand, J.F., Bensimon, D., Bensimon, A. and V. Croquette (1996) The elasticity of a single supercoiled DNA molecule. *Science*. **271**: 1835-1837.
115. Bullard, D.R. and R.P. Bowater (2006) Direct comparison of nick-joining activity of the nucleic acid ligases from bacteriophage T4. *Biochem. J.* **398**: 135-144.
116. D.J. Shapiro (1981) Quantitative ethanol precipitation of nanogram quantities of DNA and RNA. *Anal. Biochem.* **110**: 229-231.

117. Wu, W. and M.J. Welsh (1995) A method for purification of DNA species by high-speed centrifugation of agarose gel slices. *Anal. Biochem.* **229**: 350-352.
118. W. Yaphe (1957) The use of agarase from Pseudomonas atlantica in the identification of agar in marine algae (Rhodophyceae). *Can. J. Microbiol.* **3**: 987-993.

Appendix 2

The following figure and table are provided as supporting information.

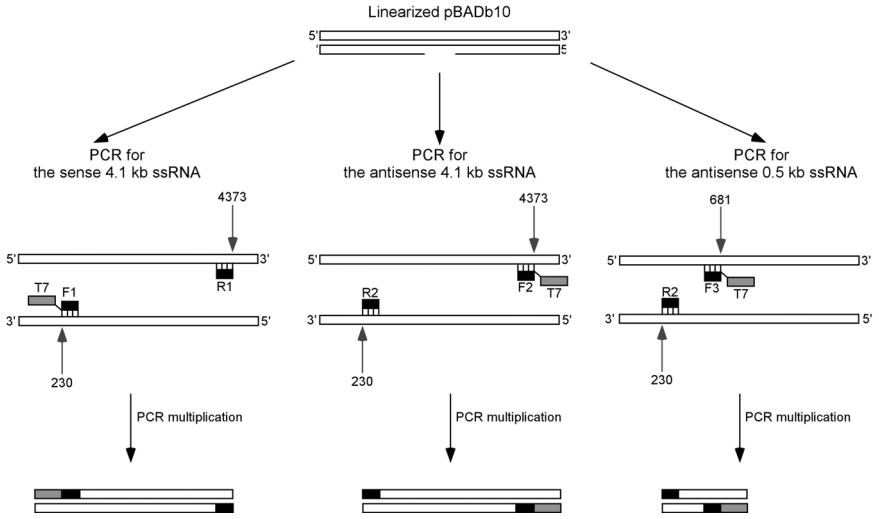


Figure A2.1. *Schematic showing the synthesis procedure for the three single-stranded RNA strands (4.1 kb sense, 4.1 kb antisense and 0.5 kb antisense).*

PCR fragment	Primers	
Sense 4.1 kb ssRNA	forward:	5'-**TAATACGACTCACTATAGG**AAGATTAGCGGATCCTACCTGAC-3'
	reverse:	5'-GGTTAACCTCAACTTCCATTTCC-3'
Antisense 4.1 kb ssRNA	forward:	5'-AAGATTAGCGGATCCTACCTGAC-3'
	reverse:	5'-**TAATACGACTCACTATAGG**GGTTAACCTCAACTTCCATTTCC-3'
0.5 kb ssRNA	forward:	5'-AAGATTAGCGGATCCTACCTGAC-3'
	reverse:	5'-**TAATACGACTCACTATAGG**AACGGCTTGATATCCACTTTACG-3'
Ext. antisense 4.1 kb ssRNA	forward:	5'-AAGATTAGCGGATCCTACCTGAC-3'
	reverse:	5'-**TAATACGACTCACTATAGG***GCTACC*GGTTAACCTCAACTTCCATTTCC-3'

Table A2.1. *Forward and reverse primers used for the synthesis of single-stranded RNA strands (4.1 kb sense, 4.1 kb antisense, 0.5 kb antisense and extended antisense 4.1 kb). The T7 promotor sequence is indicated in bold.*

Chapter 3

End-Joining Long Nucleic Acid Polymers

Abstract

Many experiments involving nucleic acids require the hybridization and ligation of multiple DNA or RNA molecules to form a compound molecule. When one of the constituents is single-stranded, however, the efficiency of ligation can be very low and requires significant individually tailored optimization. Also, when the molecules involved are very long (>10 kilobases (kb)), the reaction efficiency typically reduces dramatically. Here, we present a simple procedure to efficiently and specifically end-join two different nucleic acids using the well-known biotin-streptavidin linkage. We introduce a two-step approach, in which we initially bind only one molecule to streptavidin. The second molecule is added only after complete removal of the unbound streptavidin. This primarily forms hetero-dimers and nearly completely suppresses formation of unwanted homo-dimers. We demonstrate that the joining efficiency is 50 ± 25% and is insensitive to molecule length (up to at least 20 kb). Furthermore, our method eliminates the requirement for specific complementary overhangs and can therefore be applied to both DNA and RNA. Demonstrated examples of the method include the efficient end-joining of DNA to single-stranded and double-stranded RNA, and the joining of two double-stranded RNA molecules. End-joining of long nucleic acids using this procedure may find applications in bio-nanotechnology and in single-molecule experiments.

This chapter was published as: Michiel van den Hout, Susanne Hage, Cees Dekker and Nynke H. Dekker, *Nucleic Acids Research.* **36**: e104 (2008)

3.1 Introduction

Many experiments involving the manipulation of nucleic acids require the synthesis of complicated molecular constructs from different constituting molecules of DNA or RNA (see Chapter **2**). In single-molecule experiments, for example, constructs typically contain molecular handles[1-3] or specific sequences such as DNA hairpins[4-5], DNA-RNA hybrids[6-7], or promoter sequences[8]. For certain applications the molecules involved are very long, either because the experimental setup requires a certain minimum length[9-11] or because the studied phenomena take place over long distances[12]. An example of an experiment requiring very long molecules involves the translocation of single DNA molecules through solid-state nanopores[13-15]. Here, molecules are passed through a nanometer-sized hole in a thin membrane by an applied electrical field and can be detected by a commensurate change in the ionic current through the pore. As this process is typically very fast[13], individual molecules have to be very long (>> 1kb) in order to be resolved. Furthermore, increased sophistication of these experiments, for instance via integration with optical tweezers[9,16], will permit the interrogation of more complex lengthy molecular constructs.

In such cases, the synthesis of molecular constructs can pose a serious challenge to the experimentalist, because standard biochemical reactions can become much less efficient for long molecules[17-18]. The standard way to construct these molecules is to create complementary overhangs, which hybridize together and can then be ligated to form a stable construct. Although this typically works well for relatively short molecules, the yield of these reactions can drop off dramatically when the molecules involved reach lengths over 10 kilobases. Similar techniques using a complementary 'splint' molecule to join single-stranded molecules of DNA or RNA have been shown to be useful for joining short oligonucleotides[19-20], but they require the careful selection of the optimal splint molecule and have not demonstrably led to high yields for molecules longer than several hundred bases. To resolve some of these difficulties, we here present a straightforward procedure to efficiently end-join two arbitrary molecules in a manner that is independent of the molecule length (up to at least 20 kb).

Our method relies on the use of streptavidin (STV) to facilitate efficient end-coupling between two arbitrary nucleic acids **A** and **B**, where both **A** and **B** can be either DNA or RNA and of arbitrary length. In this scheme the STV acts as a linker between the molecule ends, which can be biotinylated chemically or enzymatically (e.g. using terminal transferase, Klenow polymerase, or poly(A) polymerase) using standard available techniques[21-26]. Although a considerable amount of previous work using STV to crosslink DNA molecules has been performed, this has primarily focused on the coupling of oligonucleotides or relatively short DNA molecules[27-31] up to 2 kilobases[32]. Also, such research typically had the aim of either forming multimers of DNA-STV complexes or long strings of *bis*-biotinylated molecules

(biotin on both ends) for the purpose of scaffolding in nanoengineering[30-31,33] or as useful aids in detection assays[29,34]. Methods that specifically aim to optimize the formation of a unique end-product, such as a hetero-dimer **A+B**, are lacking.

STV is an excellent choice as a coupling agent, because it binds biotin with an extremely high binding affinity ($K_d \sim 10^{-15}$ M) and has multiple biotin binding sites[35-36]. Indeed, wild-type STV contains four binding sites for biotin, so it could in principle join up to four biotin-labeled molecules. By simply adding STV to a collection of different biotinylated molecules, one would therefore expect to form a variety of dimers, trimers and tetramers. To specifically promote the formation of the hetero-dimer **A+B**, we here introduce a two step approach, in which the STV is first bound only to molecule **A**. Molecule **B** is then added only after the remaining unbound STV is removed, so that **B** must bind to the STV attached to molecule **A**. This not only avoids the formation of homo-dimers **A+A** and **B+B**, but, as we show below, also significantly suppresses the formation of trimers and tetramers.

3.2 Materials and methods

Synthesis of DNA and RNA

DNA molecules were synthesized by PCR derived from phage lambda DNA (Promega). For all DNA molecules used to bind STV, only one of the primers was 5'-biotinylated, so that only one end of each molecule can bind STV. The following forward and reverse primers (Biolegio, Leiden, the Netherlands) were used for the fragments:
2.2 kbp (2250 bp) DNA: 5'-CCAGAAAAATGCATTCCGTG-3' (forward) and 5'-biotin-GGATATTAATACTGAAACTGAGATCAAGC-3' (reverse).
3.5 kbp (3524 bp) DNA: 5'-AAAAGAATTCAGCCTCAGCTGACCAGCCAGAAAACGACC-3' (forward) and 5'-biotin- CTTGTTGGGCTTGTTTAATCCAGTAACTGC-3' (reverse).
4.2 kbp (4215 bp) DNA: 5'-TGATATTGCCAAAACAGAGCTG (forward) and 5'-GGAAAGGGCCCGTAAAGTGATAATGATTATCATC-3' (reverse). The underlined sequence corresponds to a restriction site for *ApaI*.
9.6 kbp (9571 bp) DNA: 5'-AAAAAAGCTTGCGAGAATTTTTAGCCCAAGC-3' (forward) and 5'-biotin-GGATATTAATACTGAAACTGAGATCAAGC (reverse).
12.7 kbp (12668 bp) DNA: 5'-biotin-GAGGCCGGGTTATTCTTGTTCTCTGG (forward) and 5'-GGAAAGGGCCCGTAAAGTGATAATGATTATCATC-3' (reverse).
20 kbp (19945 bp) DNA: 5'-biotin-GACGCAGGGGACCTGCAG (forward) and 5'-AAAAGGTCTCTTCATGCGTTCAGTCTTAAAAGCAATT (reverse).

The 4.2 kbp and 12.7 kbp DNA molecules were used in the ligation reaction in Figure **3.2b**. The reverse primer is identical for both molecules and includes an *ApaI* site (underlined). The molecules were digested overnight with an excess of *ApaI* at 25 ºC. After digestion, both molecules contain the same overhang for ligation. The ligation reactions were performed in equimolar concentrations in a 20 µl volume

with T4 DNA ligase (New England Biolabs) in T4 ligase buffer. All PCR fragments and digested molecules were purified with the Nucleospin extract II kit (Machery-Nagel, Düren, Germany) and concentrations were determined using the commercial instrument Nanodrop (Isogen, IJsselstein, The Netherlands).

The 4.2 kb single-stranded RNA (ssRNA) and double-stranded RNA (dsRNA) molecules used in Figure **3.3a** were synthesized via run-off transcription of DNA PCR fragments derived from pBADb10 (a recombinant pBAD vector) using T7 RNA polymerase. The PCR primers (Biolegio, The Netherlands) for the template DNA were designed to incorporate a T7 RNA promoter sequence at one end of the PCR fragment. The template DNA for the ssRNA molecule contained the following forward and reverse primers (the T7 promoter is underlined):
5'-TAATACGACTCACTATAGGAAGATTAGCGGATCCTACCTGAC-3' (forward) and
5'- CGCAGCCAGCCATCGGAACCGGGGTTAACCTCAACTTCC-3' (reverse).
A small biotinylated oligonucleotide, 5'-bio-CGCAGCCAGCCAUCGGAACC-3', was hybridized to the 5'-end of the 4.2 kb ssRNA molecule to facilitate the binding to STV.

For the dsRNA molecule, the complementary strand of the same DNA template was also transcribed by introducing the T7 promoter in the reverse primer. In this case, the PCR primers used for the DNA template were:
5'-AAGATTAGCGGATCCTACCTGAC-3' (forward) and
5'-TAATACGACTCACTATAGGGGTTAACCTCAACTTCCATTTCC-3' (reverse).

Run-off transcription of this DNA template yields an ssRNA molecule that is complementary to and slightly shorter than the 4.2 kb ssRNA molecule described above. After hybridization of these two RNA molecules, the double-stranded product includes a 20 nt 3' overhang, to which the same biotinylated oligonucleotide was hybridized as used above. All hybridizations were performed following a procedure described previously[37].

DNA-STV and RNA-STV binding

All incubations were performed in buffer **1** (0.25x TBE at pH 8.3, 2 mM $MgCl_2$). Unless stated otherwise, the first incubation step was always performed at 37 ºC for a duration of 30 min, and the second incubation step was performed at 4 ºC for 30 min up to several hours.

Removal of STV

We used two different methods to remove unbound STV following the first incubation step: column purification using the Microcon YM-100 (Millipore), and gel purification. The column purification was carried out by loading the sample onto the column and adding ddH_2O up to a total volume of 200 µl. After

approximately 1 min of centrifugation at 10.000 rpm, 200 µl ddH_2O was added, and the sample was again centrifuged for approximately 1-3 min at 10.000 rpm. The sample was then eluted in 20 µl ddH_2O and used for further experiments. Removal of the free STV by gel purification was done by loading the sample onto an agarose gel, excising the **A+STV** fragment and cutting the gel slice into small pieces, which were then centrifuged for 15 min at 14.000 rpm[38]. The supernatant containing the A+STV fragments was afterwards collected and used in the second incubation step. Although this technique for gel extraction has a relatively low recovery for long molecules, it was chosen because the nucleic acid-STV complex does not survive the buffers used in most commercial gel extraction kits. Other techniques either require high temperatures and/or ethanol precipitation, which destabilize the nucleic acid-STV complex. We note that extraction of the nucleic acid-STV complex using electro-elution might also be a viable alternative, with possibly higher recovery.

Gel analysis

Samples containing molecules up to ~10 kb in length were all analyzed by gel electrophoresis in non-denaturing 0.8% agarose gels in TAE buffer at 5 V/cm. DNA samples of longer molecules were analyzed by pulsed field electrophoresis (CHEF Mapper, BioRAD Technologies) on 1.0% agarose gels in 0.5x TBE buffer using a program optimized for separating DNA molecules between 10 and 80 kbp. All agarose gels were stained with ethidium bromide (EtBr).

AFM images

The gel-extracted heterodimers of 2.2 kbp + 3.5 kbp DNA were imaged with tapping-mode AFM (National Instruments, Nanoscope IV) in air by depositing 5 µl of 1 ng/ml of DNA solution in the presence of 10 mM $NiCl_2$ on freshly cleaved mica. After 2 minutes, the sample was rinsed for 10 seconds with ddH_2O water, blown dry with N_2, and directly imaged in the AFM. The AFM image in Figure **3.1c** was analyzed using the freeware WsXM[39].

Determination of binding efficiency

The efficiency of the joining reactions is determined as follows: the intensities of the bands corresponding to the biotinylated molecules **A+STV**, **B+STV**, **A+B, A+A** and **B+B** were determined using the program LabImage 1D (Kapelan GmbH, Germany) and each was divided by the respective molecular weight to obtain a relative molar amount N_{A+STV}, N_{B+STV}, N_{A+B}, N_{A+A}, and N_{B+B}. From this calculation, we

determined which of the reactants **A** or **B** was limiting (*i.e.* which reactant was present in the lowest amount before the reaction). The efficiency of the reaction was then calculated as follows: efficiency (**A+B**) = N_{A+B} / (N_{A+STV} + N_{A+A} + N_{A+B}) in the case of limiting reactant **A**, and efficiency (**A+B**) = N_{A+B} / (N_{B+STV} + N_{B+B} + N_{A+B}) in the case of limiting reactant **B**. We do not take into account the **A** and **B** molecules that did not bind STV, because these represent molecules that were not biotinylated and therefore do not participate in the reaction. Depending on the application, the final product itself may require further purification, which, depending on the protocol employed, could additionally affect the overall efficiency (eg. the gel purification method described above has a typical recovery of 50-80%[38]).

3.3 Results

3.3.1 Formation of DNA-STV-DNA hetero-dimers

In our experiments, we aim to optimize the specific formation of the dimer **A+B** starting from unlinked molecules **A** and **B**. To achieve this, we perform a two-step reaction, shown schematically in Figure **3.1a**. In the first step, molecule **A** is incubated in buffer **1** for 30 min together with a large excess of STV (>30x molar excess). Due to the excess STV, all biotinylated molecules will rapidly be bound to single STV molecules. Because these **A+STV** complexes cannot bind to each other (they can only bind to those **A** molecules in which the biotin remains unbound), this has an additional advantage that it almost completely suppresses the formation of homo-dimers **A+A**. After this first step, the free STV is removed by gel purification and the second nucleic acid molecule **B** is added in approximately a 1:1 molar ratio to molecule **A**. This is incubated at 4 ºC for a duration of 30 minutes up to several hours. Provided that all free STV has been successfully removed after the first incubation, the **B** molecules exclusively bind to the STV-tagged **A** molecules, resulting in the formation of hetero-dimers **A+B**.

The result of this protocol with two mono-biotinylated DNA molecules of lengths 3.5 kilo base pair (kbp) (molecule **A**) and 2.2 kbp (molecule **B**) is shown in Figure **3.1b**. For reference, Lane M contains a DNA marker. In lane 2 we show the result after following the protocol *without* STV, which as expected shows clear bands corresponding to 2.2 kbp DNA and 3.5 kbp DNA. In lane 1, molecules **A** and **B** were incubated according to the protocol described above. The clear band at approximately 5.7 kbp (circled) corresponds to the dimer **A+B**.

An AFM image of the resulting **A+B** dimer deposited on mica is shown in Figure **3.1c**: here, we can clearly distinguish two DNA molecules joined at the ends via a slightly higher intermediary. The intermediary is likely the STV, as STV is significantly larger (4-6 nm diameter) than the DNA (2.2 nm diameter) and should therefore appear higher in the AFM[40] (note that the apparent height of both the

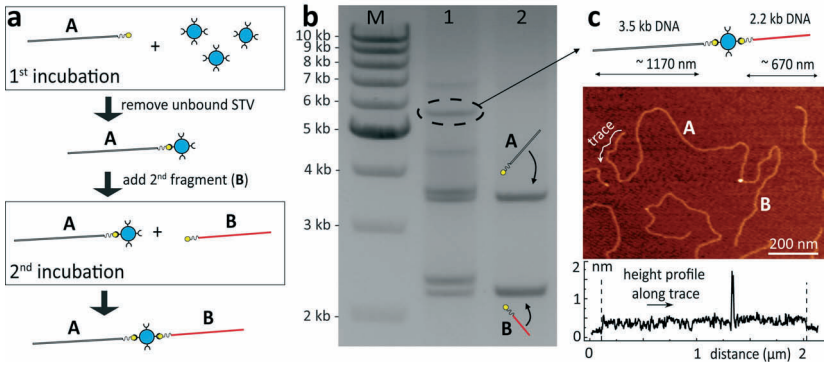


Figure 3.1: *Protocol and formation of DNA-STV-DNA dimers.* **a**. Schematic showing the 2-step incubation to optimize the efficiency of the dimer **A+B**, where **A** and **B** can be any nucleic acid with a biotin at one end. First, we bind molecule **A** to STV by 30 min incubation with a 30x molar excess of STV. We then remove all the unbound STV using gel purification and incubate the STV-bound molecule **A** together with molecule **B** at 4 ºC in a 1:1 ratio for 30 min up to 15 hrs depending on the length of the molecules. **b**. 0.8% agarose gel image showing the result of this protocol with 3.5 kbp DNA (**A**) and 2.2 kbp DNA (**B**) molecules (lane 1), and the control experiment without STV (lane 2). There is a very clear band at 5.7 kbp (lane 1), corresponding to the dimer **A+B**. The efficiency of this reaction is approximately 60%. **c**. AFM image of the **A+B** hetero-dimer, and a height profile traced along the molecule. The STV can be seen as a clear peak in the height profile, and joins two molecules of about 1200 +/- 50 nm and 750 +/- 50 nm in length, corresponding well to the expected lengths for molecule **A** and **B**.

DNA and the STV in the AFM is lower, because the AFM tip deforms the DNA and STV as is commonly observed in AFM imaging). The height profile of a trace along the contour of the construct is taken to confirm the lengths: we find the distances along the DNA to the STV are about 1200 ± 50 nm and 750 ± 50 nm. This matches well with the expected length values for molecule **A** (3524 bp x 0.34 nm/bp = 1192 nm) and molecule **B** (2250 bp x 0.34 nm/bp = 765 nm).

We now investigate the bands in Figure **3.1b**, lane 1, that do not correspond to the desired end-product dimer **A+B**. These bands migrate at approximately the same speed as the unbound molecules **A** and **B** in lane 2. However, closer inspection reveals a slight shift in the band corresponding to molecule **A**. Because this shift only appears in the presence of STV, it can be attributed to molecules **A** that are bound to STV, and is observable due to the fact that the bound STV reduces the electrophoretic mobility of DNA molecules up to at least 3.5 kbp in length (Appendix **A3**). In the context of the second step of our reaction, the appearance of this band corresponding to **A+STV** may nonetheless appear

somewhat surprising: for if STV is bound, why does dimer formation not follow and remove this band? To address this question, we have investigated whether the presence of this band could be decreased by longer incubation times or by varying the relative concentrations of **A** and **B**; however, no observable effect was found *(*Appendix **A3***)*. Consequently, we speculate that it is an imperfection in the STV that causes this band to remain: it is consistent with the existence of a fraction of STV molecules that can only bind single, but not multiple, biotins.

Using the intensities of the bands at approximately 3.7 kbp (**A+STV**) and 5.7 kbp (**A+B**), we can estimate the efficiency of formation of the dimer **A+B** and find that to be approximately 60% (see Methods). Note that in order to achieve such a high efficiency, it is necessary to minimize the formation of homo-dimers **A+A** and **B+B** (Appendix **A3**). We have employed two technical steps to do so: first, the gel extraction following the first incubation ensures that *all* free STV is removed. Unfortunately, this is not the case for many other purification methods, and small amounts of free STV can remain and will lead to the formation of **B+B** homo-dimers (Appendix **A3**). Second, the low temperature in the second step prevents dissociation of the STV from the **A+STV** complex, which would otherwise also allow formation of homo-dimers (Appendix **A3***)*.

3.3.2 Applications of the method

End-joining of very long molecules

Having demonstrated that we can specifically form the hetero-dimers **A+B** with high efficiency, we now demonstrate the utility of this method in three applications: the end-joining of two very long DNA molecules, the end-joining of DNA to RNA, and the end-joining of RNA to RNA. In Figure **3.2a**, lane 1, we demonstrate that we can efficiently end-join two very long molecules of DNA of 20 kbp (**A**) and 9.6 kbp (**B**). This is shown by the appearance of a band that migrates more slowly than the 23 kbp marker in lane M2. This likely corresponds to the dimer **A+B**, which should migrate at approximately the same speed as 30 kbp DNA. From the intensities of the bands, we deduce an efficiency of formation of the **A+B** dimer of approximately 50%, similar to the case of the much shorter molecules in Figure **3.1**. We therefore conclude that the efficiency does not significantly decrease when joining substantially longer molecules.

This is in sharp contrast to ligation, where the efficiency can decrease dramatically for such long molecules. This is illustrated in Figure **3.2b**: Lane 1 shows the self-ligation of a 4.2 kbp DNA molecule, resulting in an 8.4 kbp DNA molecule. This 4.2 kbp DNA molecule was digested with a restriction enzyme at a site in one of the PCR primers, leaving a 4 base pair self-complementary overhang (*Materials and Methods*). After overnight incubation of this molecule with T4 DNA ligase, almost all molecules are ligated (lane 1). However, when the same reaction was

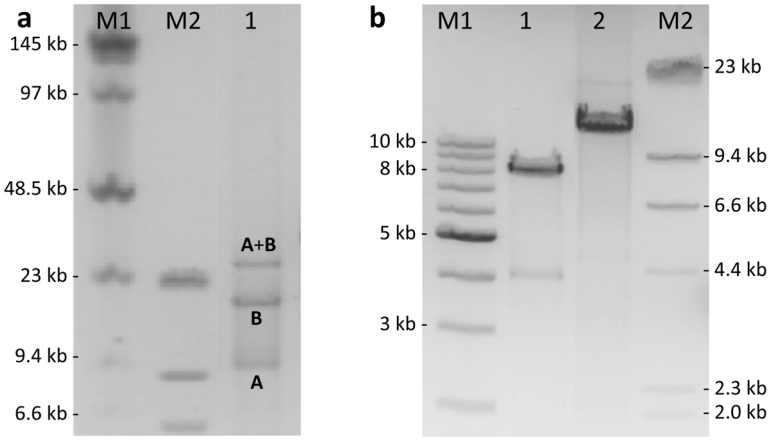


Figure 3.2: *End-joining very long DNA.* **a**. Pulsed field electrophoresis gel image illustrating the efficient end-joining of two very long DNA molecules. Lane 1 contains the product after joining a 20 kbp (**A**) DNA molecule to a 9.6 kbp (**B**) DNA molecule using our 2 step protocol. Using markers M1 and M2 as a reference, we can see bands corresponding to the molecules **A**, **B**, and an additional band corresponding to DNA longer than 23 kbp, which is likely the **A+B** dimer of ~ 30 kbp DNA. The efficiency of dimer formation (~50%) is similar to that of the shorter molecules in Figure **3.1**. **b**. Pulsed field electrophoresis image illustrating that the ligation efficiency to join two molecules is strongly reduced for longer molecules: In lane 1 we show the ligation of a 4.2 kbp long DNA molecule to itself (the molecules have a small self-complementary overhang), which clearly has a very high efficiency of about 90%, judging from the relative intensity of the band at about 8.5 kbp DNA. However, the ligation of a 12.7 kbp molecule with exactly the same overhang is much less efficient: a band at approximately 25 kbp is not observed.

carried out with a 12.7 kbp molecule containing exactly the same self-complementary overhang, the efficiency of the ligated product was dramatically reduced, to almost zero (lane 2). Note that in these two ligation reactions the molar concentrations of both molecules were kept the same for the long and the short molecules. Efforts to improve the efficiency of the ligation reaction by performing the same reaction in the presence of poly-ethylene glycol[41], did not have noticeable effects (results not shown). Further optimization may yield a higher efficiency, but this may be a time-consuming task. Conveniently, our biotin-streptavidin linkage procedure offers a way to join such molecules without the need for further optimization.

End-joining of DNA to RNA

A second advantage of using STV as a linker is its versatility: for instance, it allows us to very straightforwardly join RNA molecules to DNA molecules. Normally, this would require the synthesis of complementary overhangs on the DNA and RNA molecules, and possibly the subsequent ligation of the RNA-DNA hybrid. This can be challenging, because secondary structure in the single-stranded DNA parts and RNA can prevent efficient hybridization. In addition, DNA-RNA hybrid ligation reactions can be very inefficient[42-43]. In contrast, our biotin-streptavidin linkage procedure only requires that both molecules are biotinylated at one end.

In Figure **3.3a** we show two examples in which we bind DNA molecules to single-stranded and double-stranded RNA molecules, respectively. The DNA molecules are the 3.5 kbp and 2.2 kbp mono-biotinylated PCR products used above, and the RNA molecules were synthesized by *in vitro* transcription from a 4.2 kbp long piece of DNA. A short 5'-biotinylated RNA oligonucleotide was hybridized to one of the ends of both RNA molecules to provide the biotin-group for STV binding. The result of our protocol in the absence of STV with the 3.5 kbp DNA piece (**A**) and the single-stranded 4.2 kb RNA (**B**) molecule is shown in lane 1. The ssRNA migrates approximately at the same speed as 1.5 kbp DNA and clearly no dimer formation occurs. The result of our protocol in the presence of STV is shown in lane 2: an additional band appears at approximately 5 kbp DNA, which is where the ssRNA-DNA dimer is expected. The intensity of this band compared to the others implies an efficiency of ~40% for the formation of **A+B** dimers, demonstrating that the efficiency of DNA-ssRNA dimers is quite similar to that obtained for the DNA-DNA dimers above. We also demonstrate the use of our protocol in binding a 2.2 kbp DNA (**A**) molecule to a 4.2 kbp double-stranded RNA (**B**) molecule. Lane 4 shows the result after following the protocol *without* STV: the double-stranded RNA migrates at the same speed at 4.2 kbp DNA, and no dimer formation occurs. In lane 3, on the other hand, we observe an additional band at about 6.4 kbp, corresponding to the hetero-dimer **A+B**. From the intensities of the bands we calculate an efficiency of approximately 75% for the formation of **A+B** dimers, and indeed almost all of the RNA is bound to DNA. Achieving a similar efficiency in binding long RNAs to DNA using enzymatic ligation would be very challenging because the efficiency of these reactions may vary depending on the molecules involved.

End-joining of RNA to RNA

As a final application, we demonstrate that our method can also be applied to efficiently end-join two RNA molecules (Figure **3.3b**). Here, we used our protocol to end-join two 4.2 kbp double-stranded RNA molecules (**A**) to create dimers **A+A**. In

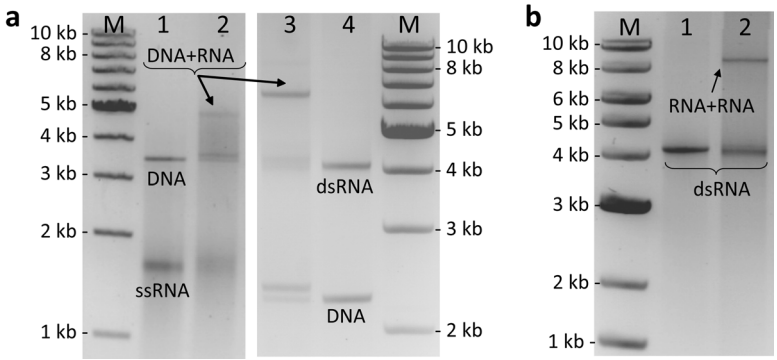


Figure 3.3: *End-joining DNA to RNA and RNA to RNA.* **a**. Joining DNA to RNA. Lane 1 contains the result following our two step protocol *without* addition of STV, using a 3.5 kbp DNA (**A**) molecule and a 4.2 kb ssRNA molecule (**B**). The ssRNA is seen to migrate at the approximately at the same speed as 1.5 kbp DNA and –as expected- we see only bands corresponding to molecules **A** and **B**. In lane 2, where the protocol was followed in the presence of STV, a band appears at approximately 5 kbp DNA, corresponding to the **A+B** dimer. In lanes 3 and 4 we illustrate the same procedure with two different molecules: a 2.2 kbp DNA molecule (**A**) and a 4.2 dsRNA molecule (**B**). Again, in the control experiment without STV no dimers are formed (lane 4). Conversely, in lane 3, after following our regular protocol in the presence of STV, an additional band is observed at approximately 6.5 kbp. As the bare dsRNA molecule migrates at approximately 4.2 kbp DNA, the band at 6.5 kbp most likely corresponds to the **A+B** dimer. The efficiencies in both cases are similar to or even better than DNA-DNA binding shown in Figure **3.1**. **b**. Joining RNA to RNA. The same 4.2 kbp dsRNA molecule used above was used with our protocol to create an 8.4 kbp dsRNA molecule. Lane 1 contains the result following our two-step protocol *without* addition of STV, using the 4.2 kbp RNA molecule as both **A** and **B**: no dimers are formed. However, in the presence of STV (lane 2) an additional band is formed at approximately 8.5 kbp DNA corresponding to the dsRNA-dsRNA dimer.

lane 2, two distinct bands can be seen, one migrating at approximately 4.2 DNA (corresponding to molecule **A**) and one migrating at approximately 8.5 kbp DNA, which must correspond to the dimer **A+A**. We find that approximately 20% of the final product consists of dimers **A+A**. As before, we have also performed a control in the absence of STV (Figure **3.3b**, lane 1), and here no dimer formation occurs, as expected.

3.4 Discussion

Having demonstrated several applications of our method, we now analyze in more detail a number of aspects that influence the overall efficiency of dimer formation. Specifically, we will address the importance of temperature and the role of STV quality in dimer formation, and we explain why our protocol disfavors the formation of multimers.

Influence of temperature on homo-dimer formation

We have mentioned above that a low temperature during the second incubation step is important in preventing homo-dimer formation (see also Appendix **A3**). The reason for this is that the rate of dissociation of STV is much reduced at a lower temperature, so that the STV dissociates less frequently from the **A+STV** complex and renders formation of **A+A** dimers (and also **B+B** dimers) less likely. Assuming that the dissociation rate of STV at 37 ºC is approximately $k_{off} = 10^{-4}\ s^{-1}$ (estimated from ref[44]), we expect that after 1 hr about 30% of the STV has dissociated once and re-associated again to form a new complex (either a DNA-STV-DNA dimer or a single DNA-STV complex). At 4 ºC, however, the rate of dissociation is approximately three orders of magnitude lower[44], meaning that only 0.03% of the STV molecules would have had time to dissociate once. The formation of unwanted dimers such as **A+A** and **B+B** due to STV dissociation is therefore dramatically suppressed. One could argue that the lower temperature may also require longer incubation times to compensate for slower diffusion. A longer incubation time would in turn allow for more dissociation to occur, which could cancel the benefit gained from the lower dissociation rate. However, the rate of diffusion is reduced by only about 10% by lowering temperature from 37 ºC to 4 ºC, which is hardly significant compared to the three orders of magnitude gained in the stability of the **A+STV** complex.

STV quality and the efficiency of dimer formation

We have found that the number of molecules that binds STV but does not form dimers (in other words, **A+STV** and **B+STV** complexes) is unaffected by either the incubation time or the relative concentration of molecules **A** and **B** (Appendix **3**). Our speculation is that the STV molecules in these single DNA-STV complexes do not have multiple binding sites available for biotin, but only a single one. Presumably, partial degradation of the STV molecules may be the cause of this. We have indeed observed degradation over time in individual stocks of STV: in all

cases, the efficiency of forming single DNA-STV complexes was not diminished (judged from the observable electrophoretic shift due to STV binding), but the relative amount of (homo- or hetero-) dimers would dramatically reduce over the course of several weeks to months, to the point where dimer formation was no longer observed. Switching to a fresh STV stock would always immediately result again in high efficiencies of dimer formation. This observation also rules out any possible influence of other factors, such as the buffers or DNA molecules used. The STV used for the experiments presented here (purchased from Roche Applied Science) showed the best results: stocks were found to be stable for at least 2 weeks up to several months when stored in ddH_2O at 4 ºC.

Formation of multimers

We have observed almost no formation of DNA trimers and tetramers in our experiments, even though wild-type STV has four binding sites for biotin. This is in accordance with earlier reports[29,32], where it was also found that direct incubation of biotinylated DNA with STV would not result in multimer formation, except at certain specific ratios of STV:DNA. Indeed, in some cases we have also observed the formation of trimers and tetramers when both molecules were incubated directly with STV (in a single incubation rather than our preferred two step approach). Formation of trimers and tetramers becomes more likely when the STV concentration is less than twice that of the DNA molecules, although it probably remains much slower than dimer formation because of steric hindrance[29]. Indeed, Monte-Carlo simulations[45] imply that the probability of forming trimers and tetramers under optimized molar concentrations is still orders of magnitude lower than the probability of forming dimers. It is likely that our two-step protocol similarly disfavors the formation of trimers and tetramers. The likelihood of multimer formation in the first step of our protocol is reduced significantly, due to the very large excess of STV used. This makes it much more likely that all the biotin-groups on the molecules **A** are rapidly saturated by binding to single STV molecules rather than forming dimers. In the second step, when molecule **B** is added, it can only bind to **A+STV** complexes. Due to steric constraints, the probability that two molecules **B** bind to the same **A+STV** complex is lower than for two molecules **B** to bind two different **A+STV** complexes. Also, the concentrations of **A** and **B** are typically matched so that all **B** molecules can bind to **A+STV** molecules. Dimer formation is thus much more likely then formation of trimers and tetramers.

3.5 Conclusions

We have presented a simple procedure to specifically end-join two biotinylated nucleic acid molecules **A** and **B** using STV as an intermediary. The method avoids the time-consuming optimization usually required for ligation reactions when long molecules are involved or when joining DNA to RNA or RNA to RNA. The method presented here appears to be insensitive to the length of the molecules (up to at least 20 kbp long DNA) or to whether the molecules are DNA or RNA. The motivation to use STV to end-join two nucleic acids was based on the fact that the STV-biotin linkage is the strongest non-covalent intermolecular bond known, but also because it is relatively straightforward to enzymatically biotinylate DNA and RNA molecules. We note that this procedure could in principle make use of any other pair of strongly binding antigens, such as digoxygenin and anti-digoxygenin or fluorescein and anti-fluorescein[21,24-26]. The use of such antigen-pairs could also be useful in typical single molecule applications, such as magnetic or optical tweezers, where also the other ends of the molecules **A** and **B** will be attached to either a bead or a surface[46-48]. Conveniently, such an approach will automatically filter out the **A+B** product (the **A+STV** and **B+STV** alone will not tether between both surfaces), which will increase the overall yield of the protocol even more. The binding protocol presented here can be performed easily and rapidly in a standard biochemical buffer, and no special chemicals and subsequent purifications are required. It is therefore an ideal choice to create long nucleic acid constructs that may find many applications in bio-nanotechnology and single-molecule experiments.

3.6 Acknowledgments

We thank Martin van den Heuvel and Gary Skinner for useful discussions and Serge Donkers for help with the final experiments. This work was financially supported by the Netherlands Organization for Scientific Research (Vidi grant to N.H.D).

Bibliography

1. Marko, J.F. and E.D. Siggia (1995) Stretching DNA. *Macromolecules*. **28**: 8759-8770.

2. Seol, Y., Skinner, G.M., Visscher, K., Buhot, A. and A. Halperin (2007) Stretching of homopolymeric RNA reveals single-stranded helices and base-stacking. *Phys. Rev. Lett.* **98**: 158103.1-158103.4.
3. Seol, Y., Skinner, G.M. and K. Visscher (2004) Elastic properties of a single-stranded charged homopolymeric ribonucleotide. *Phys. Rev. Lett.* **93**: 118102.1-118102.4.
4. Woodside, M.T., Behnke-Parks, W.M., Larizadeh, K., Travers, K., Herschlag, D. and S.M. Block (2006) Nanomechanical measurements of the sequence-dependent folding landscapes of single nucleic acid hairpins. *Proc. Nat. Acc. Sci. U.S.A.* **103**: 6190-6195.
5. Woodside, M.T., Anthony, P.C., Behnke-Parks, W.M., Larizadeh, K., Herschlag, D. and S.M. Block (2006) Direct measurement of the full, sequence-dependent folding landscape of a nucleic acid. *Science*. **314**: 1001-1004.
6. Liphardt, J., Onoa, B., Smith, S.B., Tinoco, I. and C. Bustamante (2001) Reversible unfolding of single RNA molecules by mechanical force. *Science*. **292**: 733-737.
7. Onoa, B., Dumont, S., Liphardt, J., Smith, S.B., Tinoco, I. and C. Bustamante (2003) Identifying kinetic barriers to mechanical unfolding of the T-thermophila ribozyme. *Science*. **299**: 1892-1895.
8. Skinner, G.M., Baumann, C.G., Quinn, D.M., Molloy, J.E. and J.G. Hoggett (2004) Promoter binding, initiation, and elongation by bacteriophage T7 RNA polymerase - A single-molecule view of the transcription cycle. *J. Biol. Chem.* **279**: 3239-3244.
9. Keyser, U.F., Koeleman, B.N., Van Dorp, S., Krapf, D., Smeets, R.M.M., Lemay, S.G., Dekker, N.H. and C. Dekker (2006) Direct force measurements on DNA in a solid-state nanopore. *Nat. Phys.* **2**: 473-477.
10. Storm, A.J., Chen, J.H., Zandbergen, H.W., C. Dekker (2005) Translocations of double-strand DNA through a silicon oxide nanopore. *Phys. Rev. E*. **71**: 051903.1–051903.10
11. Li, J.L., McMullan, C., Stein, D., Branton, D. and J. Golovckenko (2001) Solid state nanopores for single molecule detection. *Biophys. J.* **80**: 339A.
12. Bianco, P.R., Brewer, L.R., Corzett, M., Balhorn, R., Yeh, Y., Kowalczykowski, S.C. and R.J. Baskin (2001) Processive translocation and DNA unwinding by individual RecBCD enzyme molecules. *Nature*. **409**: 374-378.
13. Dekker, C. (2007) Solid-state nanopores. *Nat. Nanotechnol.*, **2**, 209-215.
14. Storm, A.J., Storm, C., Chen, J.H., Zandbergen, H., Joanny, J.F. and C. Dekker (2005) Fast DNA translocation through a solid-state nanopore. *Nano Lett.* **5**: 1193-1197.
15. Li, J.L., Gershow, M., Stein, D., Brandin, E. and J.A. Golovchenko (2003) DNA molecules and configurations in a solid-state nanopore microscope. *Nat. Mat.* **2**: 611-615.

16. Keyser, U.F., van der Does, J., Dekker, C. and N.H. Dekker (2006) Optical tweezers for force measurements on DNA in nanopores. *Rev. Sci. Instr.* **77**: 105105.1-105105.9.
17. Pheiffer, B.H. and S.B. Zimmerman (1983) Polymer-Stimulated Ligation - Enhanced Blunt-End or Cohesive-End Ligation of DNA or Deoxyribooligonucleotides by T4-DNA Ligase in Polymer Solutions. *Nucleic. Acids Res.* **11**: 7853-7871.
18. Rusche, J.R. and P. Howardflanders (1985) Hexamine Cobalt Chloride Promotes Intermolecular Ligation of Blunt End DNA Fragments by T4 DNA-Ligase. *Nucleic Acids Res.* **13**: 1997-2008.
19. Kurschat, W.C., Muller, J., Wombacher, R. and M. Helm (2005) Optimizing splinted ligation of highly structured small RNAs. *RNA*. **11**: 1909-1914.
20. Stark, M.R., Pleiss, J.A., Deras, M., Scaringe, S.A. and S.D. Rader (2006) An RNA ligase-method for the efficient creation of large, synthetic RNAs. *RNA*. **12**: 2014-2019.
21. Schmitz, G.G., Walter, T., Seibl, R. and C. Kessler (1991) Nonradioactive Labeling of Oligonucleotides Invitro with the Hapten Digoxigenin by Tailing with Terminal Transferase. *Anal. Biochem.* **192**: 222-231.
22. Yehudai-Resheff, S. and G. Schuster (2000) Characterization of the E.coli poly(A) polymerase: nucleotide specificity, RNA-binding affinities and RNA structure dependence. *Nucleic Acids Res.* **28**: 1139-1144.
23. Feinberg, A.P. and B. Vogelstein (1983) A Technique for Radiolabeling DNA Restriction Endonuclease Fragments to High Specific Activity. *Anal. Biochem.* **132**: 6-13.
24. Huang, Z. and J.W. Szostak (1996) A simple method for 3'-labeling of RNA. *Nucleic Acids Res*. **24**: 4360-4361.
25. Proudnikov, D. and A. Mirzabekov (1996) Chemical methods of DNA and RNA fluorescent labeling. *Nucleic Acids Res.* **24**: 4535-4542.
26. Lingner, J. and W. Keller (1993) 3'-End Labeling of Rna with Recombinant Yeast Poly(A) Polymerase. *Nucleic Acids Res.* **21**: 2917-2920.
27. Niemeyer, C.M., Adler, M., Gao, S. and L.F. Chi (2001) Nanostructured DNA-protein aggregates consisting of covalent oligonucleotide-streptavidin conjugates. *Bioconj. Chem.* **12**: 364-371.
28. Niemeyer, C.M., Ceyhan, B. and D. Blohm (1999) Functionalization of covalent DNA-streptavidin conjugates by means of biotinylated modulator components. *Bioconj. Chem.* **10**: 708-719.
29. Niemeyer, C.M., Adler, M., Pignataro, B., Lenhert, S., Gao, S., Chi, L.F., Fuchs, H. and D. Blohm (1999) Self-assembly of DNA-streptavidin nanostructures and their use as reagents in immuno-PCR. *Nucleic Acids Res.* **27**: 4553-4561.
30. Niemeyer, C.M., Burger, W. and J. Peplies (1998) Covalent DNA - Streptavidin conjugates as building blocks for novel biometallic nanostructures. *Angew. Chem. Int. Ed.* **37**: 2265-2268.

31. Niemeyer, C.M., Sano, T., Smith, C.L. and C.R. Cantor (1994) Oligonucleotide-Directed Self-Assembly of Proteins - Semisynthetic DNA Streptavidin Hybrid Molecules as Connectors for the Generation of Macroscopic Arrays and the Construction of Supramolecular Bioconjugates. *Nucleic Acids Res.* **22**: 5530-5539.
32. Tian, Y., He, Y., Ribbe, A.E. and C.D. Mao (2006) Preparation of branched structures with long DNA duplex arms. *Org. Biomol. Chem.* **4**: 3404-3405.
33. Niemeyer, C.M., Boldt, L., Ceyhan, B. and D. Blohm (1999) DNA-directed immobilization: Efficient, reversible, and site-selective surface binding of proteins by means of covalent DNA-streptavidin conjugates. *Anal. Biochem.* **268**: 54-63.
34. Niemeyer, C.M., Wacker, R. and M. Adler (2001) Hapten-functionalized DNA-streptavidin nanocircles as supramolecular reagents in a competitive immuno-PCR assay. *Angew. Chem. Int. Ed.* **40**: 3169-3172.
35. Green, N.M. (1990) Avidin and Streptavidin. *Methods Enzym*. **184**: 51-67.
36. Jones, M.L. and G.P. Kurzban (1995) Noncooperativity of Biotin Binding to Tetrameric Streptavidin. *Biochem.* **34**: 11750-11756.
37. Vilfan, I.D., Kamping, W., van den Hout, M., Candelli, A., Hage, S. and N.H. Dekker (2007) An RNA toolbox for single-molecule force spectroscopy studies. *Nucleic Acids Res.* **35**: 6625-6639.
38. Wu, W. and M.J. Welsh (1995) A Method for Purification of DNA Species by High-Speed Centrifugation of Agarose-Gel Slices. *Anal. Biochem.* **229**: 350-352.
39. Horcas, I., Fernandez, R., Gomez-Rodriguez, J.M., Colchero, J., Gomez-Herrero, J. and A.M. Baro (2007) WSXM: A software for scanning probe microscopy and a tool for nanotechnology. *Rev. Sci. Instrum.* **78**: 013705.1-013705.8.
40. Pignataro, B., Chi, L., Gao, S., Anczykowski, B., Niemeyer, C., Adler, M. and H. Fuchs (2002) Dynamic scanning force microscopy study of self-assembled DNA-protein nanostructures. *Appl. Phys. A*. **74**: 447-452.
41. Zimmerman, S.B. and A.P. Minton (1993) Macromolecular Crowding - Biochemical, Biophysical, and Physiological Consequences. *Ann. Rev. Biophys. Biomol. Struct*. **22**: 27-65.
42. Sugino, A., Goodman, H.M., Heyneker, H.L., Shine, J., Boyer, H.W. and N.R. Cozzarelli (1977) Interaction of Bacteriophage-T4 Rna and DNA Ligases in Joining of Duplex DNA at Base-Paired Ends. *Jour. Biol. Chem.* **252**: 3987-3994.
43. Persson, T.W., D.K., and Hartmann, R.K. (2005) In Hartmann, R. K., Bindereif, A., Schön, A. and E. Westhof (ed.), *Handbook of RNA Biochem.* Wiley-VCH GmbH & Co. **1**: 53-74.
44. Chilkoti, A. and P.S. Stayton (1995) Molecular-Origins of the Slow Streptavidin-Biotin Dissociation Kinetics. *Jour. Am. Chem. Soc.* **117**: 10622-10628.

45. Richter, J., Adler, M. and C.M. Niemeyer (2003) Monte Carlo simulation of the assembly of bis-biotinylated DNA and streptavidin. *Chem. Phys. Chem.* **4**: 79-83.
46. Smith, S.B., Finzi, L. and C. Bustamante (1992) Direct Mechanical Measurements of the Elasticity of Single DNA-Molecules by Using Magnetic Beads. *Science*. **258**: 1122-1126.
47. Lionnet, T., Allemand, J-F., Revyakin, A., Strick, T.R., Saleh, O.A., Bensimon, D., and V. Croquette (2008) In Selvin, P. R., and T. Ha (ed.), "Single Molecule Techniques". **1**: 347-369.
48. Gore, J., Bryant, Z., Nollmann, M., Le, M.U., Cozzarelli, N.R. and C. Bustamante (2006) DNA overwinds when stretched. *Nature*. **442**: 836-839.

Appendix 3

Here we present detailed information on the origin of the formation of homo-dimers (**A+A** and **B+B**), the electrophoretic shift caused by the binding of STV to molecules A and B, and the influence of the STV concentration during and after the first incubation step. We then investigate the influence of incubation time and concentration of molecule **B** on the overall efficiency of dimer formation. Finally, we present detailed experimental information on the figures presented in both the main text and those presented here.

Origin of homo-dimer formation: incubation temperature and incomplete STV removal

In the main text we used gel purification to remove the excess unbound STV after the first incubation, combined with a lower temperature during the second incubation. Here we demonstrate why these measures are necessary to prevent the formation of homo-dimers (**A+A** and **B+B**).

In Figure **A3.1a**, lane 1, we demonstrate what happens when the second incubation step is performed at 37 ºC instead of 4 ºC and the purification is performed using column purification rather than gel purification. We used the same DNA molecules as in Figure **3.1b** from the main text: 3.5 kbp DNA (**A**) and 2.2 kbp DNA (**B**). Apart from the bands that were already discussed in Figure **3.1b** of the main text, we now find two additional bands at 4.4 kbp and 7.0 kbp, corresponding to the homo-dimers **B+B** and **A+A**, respectively. We now demonstrate that gel purification can prevent formation of the **B+B** homo-dimers, and a lower temperature during the second incubation can prevent the formation of the **A+A** homo-dimers.

The **B+B** homo-dimer forms primarily due to the fact that not all STV is removed by the Microcon column purification (which we also show experimentally in Figure **A3.2**). If some STV remains after the purification, **B+STV** complexes will form as soon as the **B** molecules are added and so allow for the formation of **B+B** homo-dimers (in addition to **A+B** hetero-dimers). To demonstrate that this is indeed what happens, we carry out the same protocol in Figure **A3.1b**, but this time again using gel purification to ensure complete removal of the unbound STV. The temperature during the second incubation, however, is kept at 37 ºC instead of 4 ºC. Molecule **A** (3.5 kbp DNA) was incubated with excess STV and then loaded on an agarose gel in lanes 2 and 3. The band containing only the fraction of **A** bound to STV was extracted from the gel and used in the second incubation. As a control, the same amount of **A** but without STV was loaded in lane 4 and gel extracted in the same manner. The gel-extracted samples from lanes 2 and 4 were subsequently

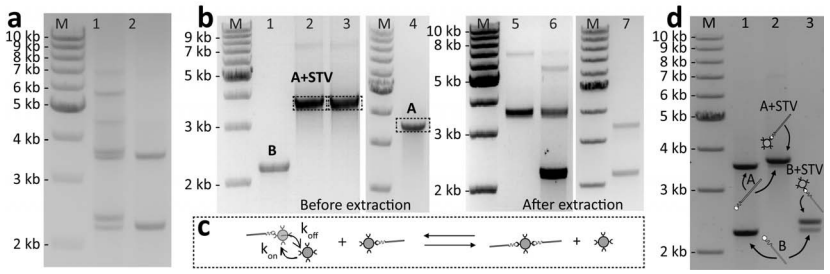


Figure A3.1. *Influence of temperature and STV purification.* **a**. 0.8% agarose gel image of the result of our protocol at 37 ºC instead of 4 ºC and using column purification instead of gel purification. The same **A** (3.5 kbp DNA) and **B** molecules (2.2 kbp DNA) were used as in Figure **3.1b** in the main text. Apart from the bands discussed earlier, we now find two additional bands, corresponding to the homo-dimers **A+A** (at 7.0 kbp) and **B+B** (at 4.4 kbp). **b**. 0.8% agarose gel image of the result of our protocol at 37 ºC, but using gel purification to remove the unbound STV after the first incubation. The **B+B** homo-dimer no longer appears, whereas the **A+A** homo-dimer is still present. This indicates that the **B+B** homo-dimer formation primarily forms because of residual STV after the removal. By using gel purification, all free STV is separated from the sample, and the homo-dimer **B+B** can no longer form. The **A+A** band, however, can still form because of dissociation of STV from the **A+STV** complex, as depicted in: **c**. This dissociation is more pronounced because of the higher temperature used compared to the normal protocol at 4 ºC. Combining both low temperature with gel purification to remove the unbound STV is therefore essential to prevent homo-dimer formation. **d**. Electrophoretic shift due to STV: lane 1 contains **A** and **B** in the absence of STV and lanes 2 and 3 contain the molecules **A** and **B**, respectively, incubated with an excess of STV. The two upper bands correspond to molecules **A** and **B** bound to STV molecules. The relative intensity of the upper compared to the lower band is a measure of how many of the molecules are biotinylated (and can bind STV).

incubated with molecule **B** (shown for reference in lane 1) at 37 ºC. The resulting products can be seen in lanes 6 and 7, respectively. In lane 7, containing the control in the absence of STV**,** no dimer formation occurs. In lane 6, we see the band corresponding to the hetero-dimer **A+B** (at 5.7 kbp), and indeed, the band corresponding to the homo-dimer **B+B** has disappeared: the gel purification completely removes the unbound STV and formation of the homo-dimer **B+B** is suppressed.

In Figure **A3.1b**, lane 6, a small band is still visible at 7 kbp, corresponding to the hetero-dimers **A+A**. Interestingly, this band also appears in lane 5, where we loaded the gel-extracted **A+STV** complex (from lane 3) on gel after incubation at 37 ºC for 45 minutes, but without adding molecule **B**. Apparently, even when all free STV is removed and molecule **B** is omitted, the **A+A** homo-dimer still forms. This implies that the **A+STV** complex is itself unstable and can dissociate into a free

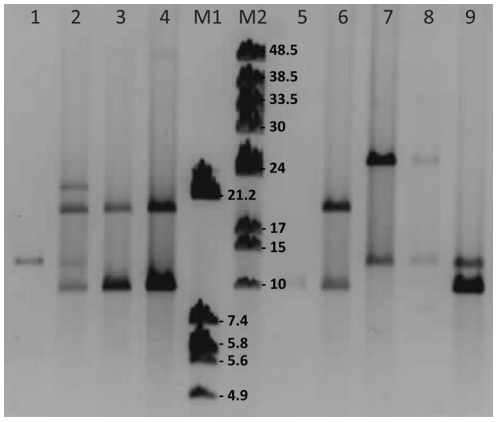
molecule **A** and free STV; when this occurs, it is possible that the free molecule **A** binds another **A+STV** complex, forming the homo-dimer **A+A**. This mechanism is schematically shown in Figure **A3.1c**. The formation of the homo-dimer **A+A** is then the result of the continuous dissociation and association of STV from the biotinylated DNA. Indeed, as we argued in the main text, the dissociation rate of the STV-biotin complex is strongly temperature dependent and lowering the temperature from 37 ºC to 4 ºC can reduce the rate of formation of this homo-dimer by several orders of magnitude. Indeed, as was shown in Figure **3.1b** of the main text, at 4 ºC formation of both the **A+A** and **B+B** homo-dimers is completely suppressed.

Small electrophoretic shift due to STV binding

In Figure **A3.1a**, lane 1, the bands at approximately 3.5 kbp (**A**) and 2.2 kbp (**B**) and are each split into two separate bands: a lower band, migrating at exactly the same speed as the molecules in the absence of STV (lane 2), and an upper band, migrating slightly slower. Since the latter only appear in the presence of STV, these upper bands must correspond to molecules **A** and **B** that are bound to a single STV molecule. We observe that this gel shift, observable due to the fact that the bound STV reduces the electrophoretic mobility of DNA molecules up to at least 3.5 kbp in length, can also conveniently function as a rapid assay to determine the amount of biotinylated product following a labeling reaction (under the assumption that all biotinylated molecules bind STV). To demonstrate this we show in Figure **A3.1d** the molecules **A** (lane 2) and **B** (lane 3) after a single incubation of 30 min with an excess of STV. Comparing the intensities of the upper and lower bands in lanes 2 and 3 it can thus be concluded that >90% of the **A** molecules and >65% of the **B** molecules are biotinylated. For reference, lane 1 contains the molecules **A** and **B** in the absence of STV. In this particular example, the difference in the biotinylated fraction of molecules **A** and **B** originates from different purities of the PCR primers used to synthesize molecules **A** and **B**.

Influence of STV concentration during and after the 1st incubation

We have investigated how the STV concentration during and after the 1st incubation influences the formation of **A+B** dimers. In Figure **A3.2** we show a pulsed field electrophoresis image of an EtBr stained 1% agarose gel in 0.5x TBE. Here, we end-joined two molecules of DNA, **A** (12.7 kbp long) and **B** (9.6 kbp long), using our two-step protocol. For convenience, column purification (Microcon YM-100, Millipore) was used to remove the STV after the first incubation (leading to the formation of **B+B** homo-dimers). The table below the gel image explains in detail the contents of each lane.



Lane	1st incubation	purification	2nd incubation	remark
1	12.7 kbp + 100x STV	YM-100	n.a.	excess STV
2	12.7 kbp + 100x STV	YM-100	9.6 kbp	normal protocol
3	STV only	YM-100	9.6 kbp	incomplete removal of STV
4	-	-	9.6 kbp + 1x STV	equimolar STV in step 2
M1	-	-	-	
M2	-	-	-	
5	9.6 kbp + 100x STV	YM-100	n.a.	excess STV
6	9.6 kbp + 100x STV	YM-100	9.6 kbp	protocol for homo-dimers
7	12.7 kbp + 100x STV	YM-100	12.7 kbp	protocol for homo-dimers
8	12.7 kbp + 100x STV	YM-100	-	dissociation of STV
9	12.7 kbp	YM-100	9.6 kbp	no STV

Figure A3.2 and Table A3.2. *Influence of concentration and incomplete removal of STV.* Pulsed field electrophoresis agarose gel image showing the influence of STV during and after the 1st incubation. See table **A3.2** for details on the contents of each lane. Lanes 1 and 5 show that an excess of STV prevents formation of homo-dimers in the 1st incubation step. In contrast, lane 4 was loaded with a 2:1 ratio of DNA:STV, clearly leading to dimer formation. Lane 8 shows the result of following our normal protocol, but not adding the **B** molecule. The dissociation of STV from the **A+STV** complex causes dimer formation after overnight incubation. The incomplete removal of STV by the YM-100 purification is illustrated in lane 3: only STV in buffer was purified over the YM-100 column, after which molecule **B** was added and loaded in lane 3 after 45 min of incubation. The presence of dimers indicates that not all STV was removed. The regular 2-step protocol is demonstrated in lane 2 for hetero-dimer formation (12.7 kbp+9.6 kbp), and in lanes 6 and 7 for homo-dimers (9.6 kbp+9.6 kbp and 12.7 kbp+12.7 kbp respectively). No dimers are formed when the protocol is performed in the absence of STV, see lane 9.

Lane 2 contains the product after binding 12.7 kbp DNA to 9.6 kbp DNA using the two-step protocol described in the main text. Using markers M1 (lamda EcoR1 marker, Promega) and M2 (lambda monocut marker, Promega) as a reference, we

can see that bands form at approximately 19 kbp (**B+B** dimer) and 22 kbp (**A+B** dimer). A band at 25 kbp (**A+A** dimer) is faintly visible, which can be attributed to the dissociation of the **A+STV** complex. As we argued before, the band corresponding to dimer **B+B** forms because the YM-100 purification does not completely remove all free STV. This is demonstrated directly in lane 3: here we incubated a sample containing only STV in the first step (no DNA present), purified with the YM-100 column, and added molecule **B** in the second incubation. We clearly see a band corresponding to dimer **B+B**, implying there must still have been some STV present after the YM-100 purification. In comparison, lane 9 contains the product in the absence of STV in the first incubation, and clearly no dimers form.

The figure also demonstrates that an excess of STV is required to prevent formation of homo-dimers in the first incubation: Lanes 1 and 5 were loaded with molecules **A** and **B**, respectively, which were both incubated with a 100-fold molar excess of STV for 30 min, purified over a Microcon YM-100 column, and then immediately loaded on gel. Clearly, formation of dimers is not observed. In contrast, lanes 4 and 8 show the results of incubation of the 9.6 kbp and 12.7 kbp DNA molecules respectively with a 1:1 ratio of STV to DNA. In both cases homo-dimer formation occurs.

Homo-dimers also form after following the regular protocol without adding **B** in the second incubation (lane 8). In this case, the **A+STV** construct is kept at 37 ºC overnight, allowing for the dissociation of STV and thus causing dimer formation. We finally demonstrate that the formed homo-dimers are really the dimers **A+A** and **B+B** by using our 2-step protocol to end-join molecules **A** and **B** to themselves, see lanes 6 and 7. Here, we incubated molecules 9.6 kbp (lane 6) and 12.7 kbp (lane 7) with themselves using the two-step protocol, forming only homo-dimers **A+A** and **B+B**.

Optimization of dimer formation

As discussed in the main text, we have found that there is typically a significant amount (up to about 50%) of molecules **A** and **B** in the final product that do not form dimers but still bind STV. In an effort to reduce this fraction and optimize dimer formation, we have investigated two factors: the influence of the duration of the second incubation step and the relative concentrations of the molecules used.

In Figure **A3.3**, we varied the duration of the second incubation step from 10 minutes to 24 hours. We used the same 2.2 kbp DNA (**A**) and 3.5 kbp DNA (**B**) molecules as used in the main text. In this experiment we used column purification to remove the STV and the incubations were performed at 37 ºC, and samples were taken out after the times indicated in the figure and immediately stored at -20 ºC. After 24 hours, all samples were simultaneously loaded on a 0.75% agarose gel. As is clear from Figure **A3.2**, no noticeable change in the relative amount of dimers (both homo-dimers and hetero-dimers) is visible between 10 minutes incubation

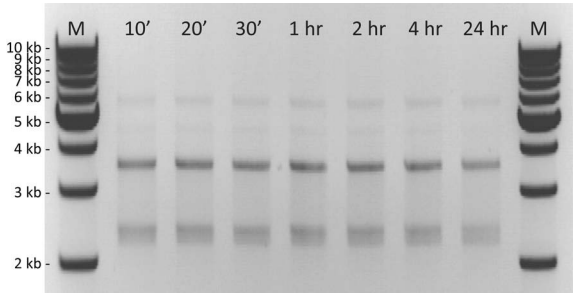


Figure A3.3. *Influence of duration of the second incubation step.* Experiment in which the duration of the second incubation step was varied from 30 min to 24 hr. Molecule **A** (3.5 kbp DNA) was incubated at 37 ºC with a 30x molar excess of STV for 30 min and loaded on a gel. After gel extraction of the **A+STV** complex, molecule **B** (2.2 kbp DNA) was added and the sample was incubated at 37 ºC. Small samples were taken out after the times indicated above the lanes and immediately stored at -20 ºC. After all samples were taken, the molecules were simultaneously loaded on the gel. No significant changes are observed in the relative intensity of the bands at 3.5 kbp DNA, indicating that the equilibrium situation is already reached after 30 min.

and 24 hours incubation. We conclude that a longer incubation does not demonstrably reduce the amount of **A+STV** and **B+STV** molecules.

Furthermore, we have investigated the effect of DNA concentration on the maximum yield of homo- and hetero-dimers. In Figure **A3.4** we varied the concentration of the second molecule **B** in the second incubation. Here, we used the 3.5 kbp DNA molecule (**A**) with a 100-fold excess of STV at 4 ºC in the first incubation. STV was removed by gel purification, after which the sample was divided over 5 different samples and incubated with 5 different amounts of 2.2 kbp DNA (**B**). After the gel extraction the concentration of the first molecule could not be determined precisely, but it can be estimated from the intensity of the lanes on the agarose gel. We estimate this to be about 10-20 ng, because the lanes in marker M contain 25 ng each (except for the lane at 5 kbp, which contains 75 ng). The molar ratios of **B** to **A** stated in the figure result from this estimation. The figure shows that the fraction of molecules **A** that does not form dimers does not significantly change even when we add approximately a 24-fold excess of the second molecule **B**. From the fact that neither a longer incubation time nor a higher concentration of molecule **B** improves the final yield of homo- or hetero-dimers we conclude that the maximum dimer formation is probably limited by the STV itself, as discussed also in the main text.

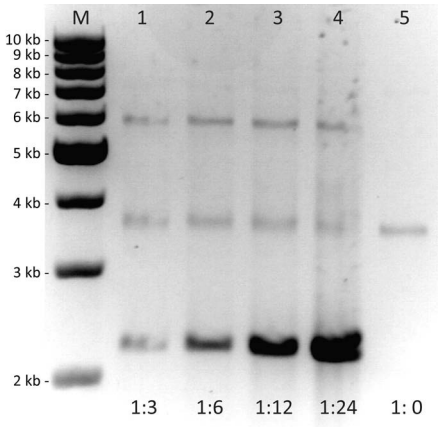


Figure A3.4. *Influence of relative concentration on dimer formation.* Molecule **A** (3.5 kbp) was incubated with different amounts of molecule **B** (2.2 kbp) following our 2-step protocol at 4ºC; gel extraction was used to remove all STV after the first incubation. The approximate relative molar concentrations of **A**:**B** are indicated below each lane. The intensity of the band at 3.5 kbp (**A** molecules that do not form dimers) does not decrease significantly even when molecule **B** is added in 24-fold excess.

Detailed experimental information on figures 3.1-3.3 and A3.1-A3.4

The 3.5 kbp and 2.2 kbp DNA fragments used in Figure **3.1**, lane 2, were incubated as follows: 470 fmol of 3.5 kbp DNA was incubated at 37 ºC for 30 min with 20 pmol STV in buffer **1** (0.25x TBE at pH 8.3, 2 mM $MgCl_2$) in a total volume of 20 µl. Subsequently, the STV was removed using gel purification and 65 fmol 2.2 kbp DNA was added. The buffer and volume were adjusted to 30 µl of buffer **1**, and this was incubated for 45 min at 4 ºC, after which the entire sample was loaded on gel. For lanes 3 and 4 we incubated 40 fmol 3.5 kbp DNA with 4 pmol STV and 65 fmol 2.2 kbp DNA with 6.5 pmol STV, respectively, at 37 ºC for 30 min.

The sample loaded in lane 1 in Figure **3.2a** was prepared by incubating 250 fmol of 20 kbp DNA (**A**) with 25 pmol STV in buffer **1** in a total volume of 40 µl at 37 ºC for 30 min. The STV was then removed using a Microcon YM-100 column, after which 250 fmol of 10 kbp DNA (**B**) was added and buffer and volume were adjusted to 100 µl of buffer **1**. After overnight incubation at 4 ºC, 10 µl of the sample was loaded on gel.

The sample loaded in lane 2 in Figure **3.3a** was prepared by incubating 70 fmol of 3.5 kbp DNA (A) with 7 pmol STV in buffer **1** for 30 min at 37 ºC. After removal of the STV using column purification, the 70 fmol of 4.2 kb ssRNA was added, the buffer was adjusted to buffer **1**, and the sample was incubated for 30 min at 4 ºC. For lane 1 the same procedure was followed except that no STV was added. In lane 3 we loaded a sample prepared as follows: 100 fmol of 2.2 kbp DNA (**A**) was incubated with 10 pmol STV in buffer **1** for 30 min at 37 ºC. After removal of STV using column extraction 65 fmol of 4.2 kbp dsRNA was added and this was incubated for another 30 min at 4 ºC in buffer **1**.
For the samples loaded in Figure **3.3b**, lanes 1 and 2, we incubated a total amount of 75 fmol of 4.2 kbp dsRNA in buffer 1 for 30 min at 37 ºC. In lane 2, 7.5 pmol of STV was added before the incubation. After column purification, again 75 fmol of 4.2 kbp dsRNA was added to both samples and the samples were incubation for another 30 min at 4 ºC in buffer **1**.

The 3.5 kbp and 2.2 kbp DNA fragments used in Figure **A3.1a**, lane 2, were incubated as follows: 100 fmol of 3.5 kbp DNA was incubated at 37 ºC for 30 min with 30 pmol STV in buffer **2** (0.5x TBE, pH=8.3 and 4 mM $MgCl_2$) in a total volume of 20 µl. Subsequently, the STV was removed using a Microcon YM-100 column and 100 fmol 2.2 kbp DNA was added. The buffer and volume were adjusted to 30 µl of buffer **2**, and this was incubated for 45 min at 37 ºC, after which the entire sample was loaded on gel. The samples loaded in lanes 2 and 3 in Figure **A3.1b** were prepared by incubating 200 fmol of 3.5 kbp DNA (**A**) with 60 pmol of STV in a total volume of 20 µl. Incubation was done at 37 ºC for 45 min. Lane 4 was loaded with the same amount of molecule **A** in buffer **2**, but no STV was added. Lane 1 was loaded with 100 fmol of 2.2 kbp DNA. The bands in lanes 2, 3 and 4 were gel extracted and one of the samples was incubated with 140 fmol of 2.2 kbp DNA (**B**). Buffer and volume were adjusted to 30 µl of buffer **2** for all three samples. After incubation for 45 min at 37 ºC, the entire samples were again loaded on gel in lanes 5 to 7. In Figure **A3.1d** we used approximately 75 fmol of 2.2 kbp DNA and 50 fmol of 3.5 kbp DNA. Both molecules were incubated with a 100x excess of STV in buffer **1** at 37 ºC for 30 min and loaded in lanes 2 (3.5 kbp) and 3 (2.2 kbp). In lane 1 the same amount of molecules was loaded in the absence of STV.

Detailed information on the samples loaded in Figure **A3.2** is given in Table **A3.1**.

In Figure **A3.3** we used a single sample prepared as follows: 1 pmol of 2.2 kbp DNA (**A**) was incubated with 30 pmol of STV in buffer **2** for 30 min at 37 ºC. STV was then removed using column purification, and 650 fmol of 3.5 kbp DNA (**B**) was added. The buffer and volume was adjusted to 100 µl of buffer 2 and the sample was incubated at 37 ºC. Samples of 7 µl were taken out after the times indicated in the figure and stored at -20 ºC. After the last sample was taken out all samples were loaded simultaneously on an agarose gel.

In Figure **A3.4** we used a single sample prepared as follows: 360 fmol of 3.5 kbp DNA (**A**) was incubated with 20 pmol of STV for 30 min at 4 ºC in buffer **2**. STV was removed using gel purification, and the sample was divided in 5 samples containing equal amounts of the recovered **A+STV** fragment. To different samples we added 20 fmol, 40 fmol, 80 fmol and 160 fmol of 2.2 kbp DNA (**B**), respectively and the samples were incubated at 4 ºC for another 45 min.

Chapter 4

Distinguishing Single- and Double-stranded Nucleic Acid Molecules Using Solid-State Nanopores

Abstract

Solid-state nanopores offer a promising method for rapidly probing the structural properties of biopolymers such as DNA and RNA. We have for the first time translocated RNA molecules through solid-state nanopores, comparing the signatures of translocating double-stranded RNA molecules and of single-stranded homopolymers poly(A), poly(U), poly(C). Based on their differential blockade currents, we can rapidly discriminate between both single- and double-stranded nucleic-acid molecules, as well as separate purine-based homopolymers from pyrimidine-based homopolymers. Molecule identification is facilitated through the application of high voltages (~600 mV), which contribute to the entropic stretching of these highly flexible molecules. This striking sensitivity to relatively small differences in the underlying polymer structure greatly improves the prospects for using nanopore-based devices for DNA or RNA mapping.

This chapter was published as: Gary M. Skinner*, Michiel van den Hout*, Onno Broekmans, Cees Dekker, and Nynke H. Dekker, *Nano Letters.* **9**: 2953-2960 (2009) (* = equal contribution)

4.1 Introduction

Due to their central importance within the machinery of life, nucleic acid structures such as the DNA double-helix, RNA hairpins, RNA pseudoknots, etc, have been intensively studied. Given their nanometer-scale sensitivity and ability to probe dynamics in real time, single-molecule techniques have recently provided a new tool for the investigation of these elastic molecules[1,2]. For example, force spectroscopy techniques such as optical tweezers have been used to study the unfolding of single RNA structural motifs[3,4]. A novel single-molecule tool that allows for the detection of local molecular structure has emerged in the form of nanopores[5,6,7]. As molecules translocate through a nanopore carrying an ionic current, a portion of the current is blocked, resulting in a characteristic blockade of the pore that depends upon the molecule that is translocating[8,9]. In principle, these pores could be used to rapidly identify and map local structures along biopolymers.

Naturally-occurring nanopores have been employed in many translocation studies, e.g. the α-Hemolysin (α-HL) protein[10] which forms a continuous channel across a lipid bilayer[11]. These α-HL nanopores have been used to probe DNA translocations[8,12,13], including forced unzipping of the double-helix[14]. They have also been used to discriminate the differing structural nature of RNA homopolymers[15]. However, this discrimination relies on the unique translocation dynamics of each molecule and strongly depends upon the interaction of the RNA molecules directly with the pore surfaces. In addition, the use of a protein-based pore limits the kinds of analysis to only systems at close to physiological conditions of temperature, pH, denaturants etc. Furthermore, the small size of the α-HL pore (diameter ~1.3 nm)[10], also limits bio-molecular analysis to molecules of this order of size, e.g. dsDNA cannot pass through α-HL[8].

For these reasons, efforts have been directed towards the development of artificial nanopores fabricated using existing semiconductor technologies[16,17]. Such solid-state nanopores have been used to measure local force upon DNA arrested within a pore[18,19,20], nucleic-acid unzipping[21] and protein/DNA complexes[9]. Here, we have translocated single-stranded (ss) RNA homopolymers poly(A)[22] poly(C)[23], poly(U)[24] and double-stranded (ds) A-RNA. We have measured the pore conductance blocked by these molecules while sweeping the applied bias from 100 mV to 600 mV. This conductance blockage vs. applied bias relationship is distinctly nonlinear for both classes of molecules, with the double-stranded molecules blocking more conductance as a function of voltage, while the single-stranded molecules block less. This differing response to voltage allows one to readily identify molecule type.

4.2 Material and methods

Nanopore experiments

We employed solid-state nanopores for the investigation of single- and double-stranded nucleic acid polymers. Nanopores of 8-14 nm in diameter were drilled into freestanding 20 nm thick Silicon Nitride membranes, using a 300 keV transmission electron microscope[17]. Figure **4.1a** shows an example TEM image of one of our pores. Prior to drilling, the surface of the chip containing the pore was coated in a layer of PDMS, extending to within a few tens of microns from the edge of the free-standing SiN membrane[33]. Immediately after electron-beam drilling, the pore was submerged within a solution of 20 nm filtered 50:50 ethanol/dH2O until use. We found that these two simple steps yielded a far greater number of useable pores, and also significantly improved the signal-to-noise ratio of our pores. Our solid-state nanopores exhibit good low-noise (~5-15 pA I_{RMS} at 100 mV, at 1 kHz low-pass filter).

The nanopore-containing chip was mounted within a custom-built Teflon flowcell containing two reservoirs for the introduction of ionic solutions containing the various molecules under study. A schematic of the setup is shown in Figure **4.1b**. The flowcell has ports that allow for the independent introduction of the various desired solutions into either chamber of the flowcell. Additional ports allow for the connection of the nanopore device to an Axon Instruments Axopatch 200B amplifier. Pore resistance was determined by measuring the pore current as a function of applied bias voltage and applying a linear fit to the slope of the resulting IV plot (Figure **4.1c**). Nucleic-acid molecules for study (Figure **4.1d**; see also below) were introduced in 10 mM Tris-HCl (pH 8.0), 1 mM EDTA, 1 M KCl, onto one side of the nanopore chip, and a voltage bias was applied across the silicon nitride membrane. The voltage and current were sampled at 150 kHz or higher bandwidth, and subsequently low-pass filtered at 35 kHz up to 70 kHz (Appendix **4**, Figure **A4.2**). For each molecule, traces of translocation events were collected at many different applied bias voltages, ranging from 100 mV up to 600 mV. After low-pass filtering the traces translocation events were identified based on preset selection criteria as described earlier[25] (Appendix **4**, Figure **A4.1**).

For each dataset, the filter frequency was chosen such that peak amplitudes in the current were not significantly affected by the filter. This was done by comparing the extreme values of the filtered detected events with the unfiltered raw data (Appendix **4**, Figure **A4.2**). The current traces of each event together with a small number of baseline (i.e. open pore current) points around it were each normalized by subtracting the average baseline value and then added together into a single long trace. Examples of such traces and histograms of these traces are shown in Figure **4.2**. To ensure that most of the detected events correspond to actual translocations, a simple check is performed: prior to each measurement we ensure that no events are observed upon applying a negative bias voltage, i.e. no

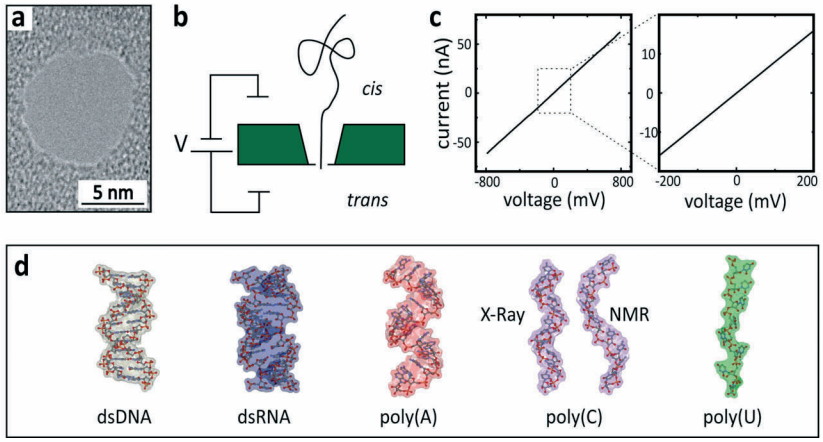


Figure 4.1. *Experimental setup and molecules used for translocation experiments.* **a.** A representative transmission electron micrograph of the nanopores used in these experiments. **b.** Schematic of the experiment: green represents the nanopore-containing chip. A pair of electrodes is inserted into each reservoir for bias application and current measurement. **c.** Typical I-V curves for one of our solid-state nanopores. The right panel is a zoom in on the low voltage regime of +/- 200 mV. **d.** Atomic models of each of the nucleic-acid polymers used in this study, with surface enclosing molecular volume applied. In the case of poly(C) there are two proposed structural models, a right-handed helix determined by X-ray studies[23] and a left-handed helix, suggested from NMR studies[29].

molecules translocate from *trans* to *cis*. Then, right after an actual measurement where translocations from *cis* to *trans* have been observed, the same check is performed: a small number of translocations are observed, which then cease after several seconds. This is a clear indication that molecules actually traversed the pore, and molecules that recently translocated can be recaptured into the pore[26]. However, such events quickly cease as the majority of the translocated molecules will have diffused away from the pore.

Synthesis of molecules

Single-stranded homopolymers of poly(A), poly(C) and poly(U) were synthesized by incubating 20 mM of ADP, CDP and UDP, respectively, with polynucleotide phosphorylase (PnP) (Sigma Aldrich) at 37 ºC for 20 minutes in a buffer containing 100 mM Tris-HCl pH 9.0, 5 mM MgCl2, 0.4 mM EDTA and 4% BSA; for dsRNA a random-sequence polydisperse length heteropolymer was synthesized using PnP and equimolar rNDPs, this ssRNA molecule was then converted to dsRNA by use of

the RNA dependent RNA polymerase from phage Φ6[27]; from gel electrophoresis, we estimated the length distributions to be 10 to 20 kilo bases (kb) for poly(A) and poly(U), 3-10 kb for poly(C), and 10-30 kilo base pair (kbp) for the ds A-RNA molecules. We used a 12.7 kbp ds B-DNA fragment produced by PCR amplification from a λ-phage DNA template (Promega)[28].

Molecular modelling and volume computation

Molecular models of the known structures of each of the molecules are shown in Figure **4.1d**. Both ds B-DNA and ds A-RNA are double-stranded and have a helical diameter of ~2.0 and ~2.3 nm respectively (as we deduced from the crystal structures (pdb ID# 2BNA and 3CIY (minus protein), respectively). For the single-stranded RNA molecules, poly(A) is known to form a helical structure with a diameter of 2.1 nm[22]. In the case of poly(C) there are two alternative models for the structure, a right-handed helix of 1.3 nm diameter as determined from X-ray studies[23] or a left-handed 1.7 nm diameter helix determined through the use of H1 NMR[29] (both shown in Figure **4.1d**); poly(U) instead adopts a random coil in solution, with an effective diameter of ~1.3 nm as determined from our molecular modeling (Figure **4.1d**). For the fully stretched homopolymers, all-atom models of 15-20 bases of each fully stretched linear molecule ($x/L_0 \approx 1$) were constructed using Xenoview software[30]. These structures correspond to how these polymers would appear when extended under very high forces (~50 pN). The helical homopolymer structures were built using Accelrys DS Visualizer (Accelrys, Inc.) taking the structural data previously reported for poly(A)[22] and the two poly(C) structures mentioned above [23,29]. For the ds B-DNA we used pdb ID# 2BNA[31], and 3CIY[32] for the dsRNA, after deletion of the bound protein, which does not distort the ds A-RNA helix. Each model was imported into the software package VEGA ZZ (Università degli Studi di Milano) to calculate its molecular volume; we define this to be the solvent inaccessible volume using a probe radius of 2.0 Å. The resulting pdb files for each of the homopolymers are available upon request.

4.3 Results

Examples of translocation events for each of the molecules at different applied bias values are shown in Figure **4.2a-e**, left panels. Each event is characterized by the amplitude of the current blockade, ΔI, and by the length of time the molecule spends in the pore. ΔI is defined as the open pore current (I_{open}) minus the current in the blocked state ($I_{blocked}$), resulting in positive values for all experiments presented here. Events were identified (see Materials and methods, and Appendix **4**) and used to construct histograms of the current values during the blockade

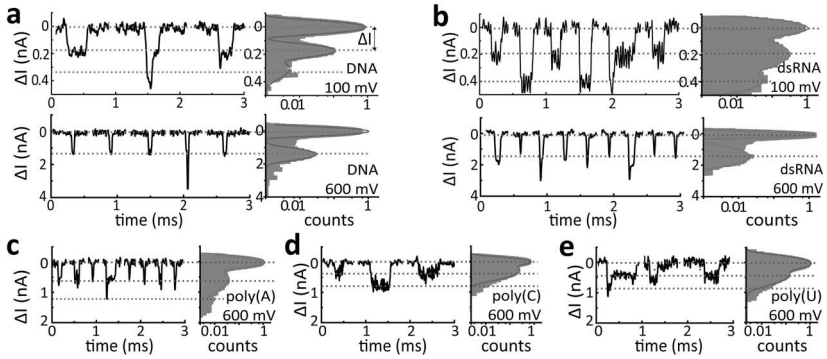


Figure 4.2. *Examples of translocation events.* Here we show sample events as well as complete histograms for all molecules studied. **a.** dsDNA; **b.** dsRNA **c.** poly(A); **d.** poly(C); **e.** poly(U). The double-stranded molecules are shown at both low (100 mV) and high (600 mV) bias. All other molecules are shown at 600 mV. In each case, the left panel shows sample events, while the righthand panel shows a typical histogram of the nanopore current, normalized to the baseline current when the pore is empty. The peak current blockade for each molecule at a given voltage was determined from the mean of the Gaussian fitted to the first peak. The resulting ΔI (= $I_{open} - I_{blocked}$) levels at a given voltage are indicated by red dashed lines.

events (Figure **4.2a-e**, right panels).

In this study, ds B-DNA was used as a calibration tool (Figure **4.2a**), to which the properties of the other molecules under study are compared. The mean current blockade, ΔI, was determined from a Gaussian function fitted to the first peak in the current histogram (in all cases, the peak number is relative to that of the open pore current). At low voltage, we measure a ΔI corresponding to a mean conductance change, ΔG (= ΔI/V) of 1.6 ± 0.5 nS at 100 mV, consistent with previous measurements of the same parameter for ds B-DNA[19,25,34-36]. Additional peaks in the current histogram are also observed at integer multiples of this ΔI: as shown in previous work, the first peak corresponds to DNA passing through the pore in a straight fashion, while further peaks at integer multiples of this ΔI correspond to molecules passing through in differing folded states (2 or more double-helical strands of the DNA occupying the pore at a time)[25]. From analysis of the translocation event rates (Appendix **4**, Figure **A4.3**), we have determined that the probability of two molecules traversing the pore simultaneously is extremely low, even at the highest applied voltages applied in our experiments; thus the peaks in the histograms correspond to single molecules translocating the nanopore. The red dotted lines spaced along the y-axis of the translocation traces (Figure **4.2**) correspond to integer multiples of the fitted values of ΔI, and these correspond well to the levels of current observed within each trace, as expected.

When RNA molecules are translocated through our nanopores at low voltage, a similar behavior is observed. In the case of ds A-RNA, translocation events can be identified as before, and the current histogram reveals multiple distinct peaks, as in the case of dsDNA (Figure **4.2b**, histogram of data at 100 mV). As the subsequent current blockades are again integer multiples of each other, these events likely correspond to molecules traversing the pore in different folded states. Similarly, histograms were made from the events detected when single stranded homopolymers were translocated at low bias, showing again a clearly distinguishable peak relative to the baseline (see Appendix **4**, Figure **A4.5**).

We next repeated these measurements at higher voltages to determine the effect of voltage on molecule translocation. In practice, we found that effects such as pore instability and clogging limited the practically achievable maximum voltage to 600 mV. When the measurements were repeated at these higher applied biases, events could still be detected, albeit with shorter translocation times, and the current histograms predominantly revealed a single peak for all molecules studied (Figure **4.2a-e**, right panels). We constructed such histograms for each molecule tested throughout the entire voltage range of 100 mV to 600 mV at 50 mV increments, and we repeated these measurements in several nanopores. We then looked at the voltage dependence of the conductance change, $\Delta G = \Delta I/V$, for the first peak in each histogram, for each molecule studied (Figure **4.3a**). ΔG was selected as the parameter of interest, as it allows normalized comparison of blockades between different experiments on pores of similar size. The translocation times are not discussed in great detail, here, because of the large variation in length of the RNA molecules. In all cases, the translocation time decreased as the voltage was increased, as expected (Appendix **4**, Figure **A4.4**). In certain cases the translocation times were unexpectedly long, which might be attributed to interactions with the nanopore walls, but these did not affect the conductance blockade: no relationship between ΔG and the translocation times was observed.

Double-stranded nucleic acids exhibit an increasing current blockade at increasing bias voltage

We will first discuss the response of ΔG to voltage for the double-stranded nucleic acids. In the case of dsDNA (Figure **4.3a**, black circles), we observe that the current blockade increases as the applied bias voltage increases from 100 to 600 mV. The measured values for ΔG at 600 mV is, dsDNA = 2.25 ± 0.09 nS, (versus ≈ 1.6 ± 0.5 nS at 100 mV). The errors in our values of ΔG reflect the maximum and minimum values from multiple experiments (see Figure **4.3a** caption, and Figure **A4.6**, showing all individual measurement points). The ΔG for the ds A-RNA also increases with increasing bias, but the effect is somewhat less pronounced than for ds B-DNA (Figure **4.3a**, blue circles). These results suggest that the molecules are actually

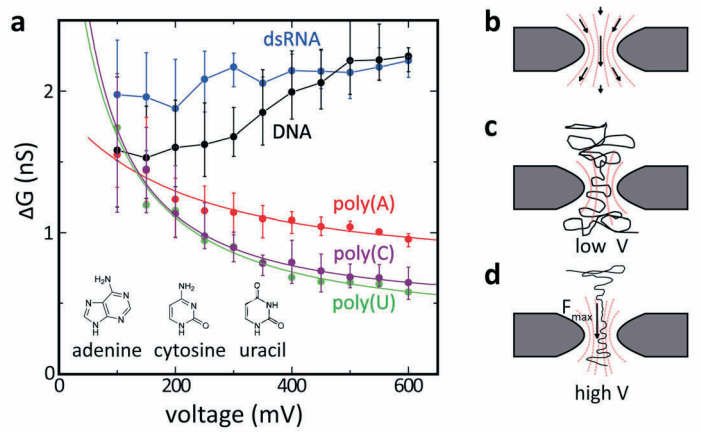


Figure 4.3. *Relationship between the applied bias and the amplitude of the conductance blockade.* **a.** Plot of the change in pore conductance, ΔG as a function of applied bias across the nanopore (ΔG= ΔI/V). At each voltage, a single experiment with a given nanopore contributed many hundreds of events, for which ΔI, and hence ΔG, were determined by fitting (see Figure **4.2**). Points shown on this plot are the mean values from N independent experiments at a given voltage, where N = 2 for dsRNA (Blue), N = 3 for dsDNA (Black), N = 7 for poly(A) (Red), N = 2 for poly(C) (Violet) and N = 1 for poly (U) (Green). Errors bars represent the maximum and minimum values from the N experiments. For poly (U), the error bars are left out as only one measurement was performed (N = 1). The solid red, green and magenta lines are fits to a model assuming the polymers are entropically stretched inside the pore (see main text). A version of this figure with all separate measurement points is available in Appendix **4**, Figure **A4.6**.
b. The force exerted upon a charged particle that experiences the electric field within the pore is proportional to the field strength at that location. Black arrows represent the magnitude of this force. **c.** At low V, these stretching forces are lower, and the highly flexible ($L_p \approx 1$ nm[38,39]) homopolymers have low relative extension, thus occupying more of the pore volume. **d.** At high V, the stretching force is higher, and so the molecule has a higher relative extension, occupying less of the pore volume, resulting in a lower measured ΔG. The stretching force is a combination of the electrical force and drag forces on the polymer, and varies along the nanopore, reaching a peak somewhere above the pore.

occupying a larger effective volume in or close to the nanopore at higher voltage, increasingly contributing to the pore resistance as the applied bias increases. Possible explanations for this unexpected result are discussed below.

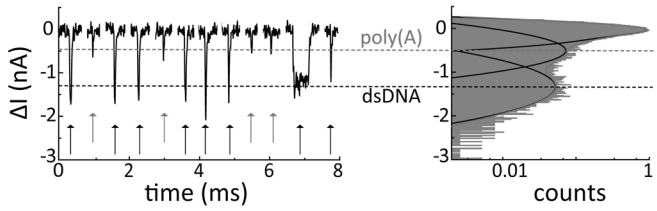


Figure 4.4. *Experiment to test the feasibility to discriminate double and single stranded nucleic acids*. Translocation experiment, with equimolar concentrations of 12.7 kbp DNA and poly(A). Data was taken at 600 mV in a 10 nm nanopore and was digitally filtered at 35 kHz. A sample of events is shown on the left, and the resulting current blockade histogram is seen on the right. Two well-separated peaks of ΔI are seen, each corresponding to either dsDNA or poly(A). Relating these peak values to the events themselves, it then becomes possible to individually identify the molecule involved from only a single translocation event, as indicated by the red and black arrows.

Single-stranded homopolymers exhibit a decreasing current blockade at increasing applied bias voltage

In contrast to the behaviour observed for the double-stranded polymers, as the applied voltage was increased, the values for ΔG showed a very pronounced decrease for the single-stranded RNA homopolymers poly(A), poly(C) and poly(U) (Figure **4.3a**, red, magenta, and green circles, respectively). At low applied bias voltage (100 mV) the ΔG values for all three homopolymers were indistinguishable, both from each other and from the double-stranded molecules (all yielding ~1.6 ± 0.5 nS). But as the voltage increases up to 600 mV, the ΔG were reduced to 0.95 ± 0.04 nS (poly(A)), 0.65 ± 0.11 nS (poly(C)), and 0.58 nS (poly(U)). Interestingly, at these higher applied bias limits, the purine poly(A) and the pyrimidines poly(U) and poly(C) can also be well separated (Figure **4.3a**).

Individual double and single-stranded translocation events can be distinguished at high applied bias voltage.

We have carried out a mixing experiment to see whether the application of high voltages allows for rapid discrimination between single- and double-stranded nucleic acids, using the large difference in conductance change between double- and single-stranded molecules. High voltages should also facilitate such separation, as in this regime all current histograms consist of only single peaks. As a proof-of-principle, approximately equimolar amounts of 12.7 kbp DNA and poly(A)

molecules were mixed and translocated through the nanopore as before. A series of events from this mixing experiment are shown at left in Figure **4.4**, displaying differing values of the current blockade. The histogram of these blockade currents demonstrates two well-separated peaks, presumably corresponding to the translocation of either poly(A) or dsDNA, respectively (Figure **4.4**). Values for ΔG for each molecule determined from this mixing experiment were dsDNA = 2.23 nS and poly(A) = 0.81 nS. This difference in the conductance change is consistent with the values measured for each molecule in isolation, and it is clear that the two different molecule populations can be distinguished. On the basis of the average ΔG computed, the suggested attribution of individual events to the translocation of dsDNA versus poly(A) molecules is indicated in Figure **4.4** by black and red arrows, respectively.

4.4 Discussion

We have successfully translocated both double-stranded as well as various types of single-stranded RNA molecules through solid-state nanopores. The homopolymeric nature of the single-stranded RNA molecules eliminated any potential contributions of internal basepairing to the signal of the molecules traversing the nanopore, under normal conditions without the need for denaturarants. As the homopolymers do have a widely dispersed length, we have focused on the magnitude of the conductance blockade rather than the pore dwell times, as the latter information is obscured by the molecular length variations. As shown, translocation of single- versus double-stranded molecules across a wide range of applied bias voltages yields surprisingly different molecular signatures.

In the case of the double-stranded molecules, we observe a slightly increasing ΔG as a function of the applied bias voltage. If the molecules occupied the same fraction of pore volume at every applied bias, we would expect ΔG to remain constant, based on a geometrical argument relating ΔG to amount of ions excluded by the molecule passing through the pore[34]. Since the persistence length L_p for dsDNA and dsRNA are ~50 and ~60 nm respectively, these molecules are expected to be fully taut inside the nanopore. Therefore, any further changes in the measured ΔG must be due to either some form of distortion of the molecular structure within the pore, or changes in access resistance due to the presence of the DNA residing just outside the pore. A similar effect has been observed for long molecules of DNA in small pores (3-8 nm)[36], where it was argued that the access resistance is increased for longer molecules due to their larger radius of gyration and subsequent enhanced interactions of the DNA with the pore membrane. It is possible that these DNA-membrane interactions are further enhanced at higher voltages, thus explaining the increased conductance blockade that we measure. As the applied bias is increased and the molecule is pulled more strongly towards the pore, perhaps it is driven to a greater level of interaction with the pore membrane.

A similar argument would apply for the dsRNA, but not for the homopolymers, as these molecules have much smaller radii of gyration ($R_{g,\ poly(A)}$ ~ 50-100 nm, $R_{g,\ DNA}$ ~ 500 nm) and this effect would therefore be much less pronounced.

The ss homopolymers display a very different dependence of the ΔG on voltage, namely a monotonic decrease, and we now discuss possible explanations for this. It is known that poly(A)[22] and poly(C)[37] both adopt a helical conformation in solution (Figure **4.1d**). Since these helices can be unwound at moderate forces of the order 10-30 pN[38], it is possible that the decrease in ΔG we observe for the homopolymers is, at least partly due to such force-induced denaturation of the base-stacking in these molecules. However, we observed the same decrease in ΔG for poly(U) (Figure **4.3a**), which is known to adopt a random-coil structure in solution at room temperature[24,39]. On this basis, we consider it unlikely that such a helix-coil transition would provide the dominant contribution to the observed decrease in ΔG with voltage.

An alternative explanation for this decreasing ΔG relies on the different degrees of flexibility between single- and double-stranded molecules. Since the single-stranded homopolymers are much more flexible than the double-stranded molecules (homopolymer $L_p \approx$ 1nm, dsDNA $L_p \approx$ 50 nm)[38,39], these molecules will be significantly coiled by the entropic forces on the length scale of the nanopore. We therefore considered that the resulting decreasing ΔG value might be due to stretching of these molecules inside the pore resulting from the electric field and opposing drag forces (Figure **4.3b**); at higher voltages V, the average stretching force F also increases. Based on this argument, at low forces (Figure **4.3c**) the molecule has a lower relative extension, $x/L_0 < 1$ (with L_0 the contour length of the molecule), thus more polymer material would occupy the volume within the nanopore, resulting in a higher ΔG value. At higher forces (Figure **4.3d**) the molecule will be stretched out more within the nanopore, and the total volume blocking the nanopore will be correspondingly reduced, resulting in a lower ΔG. This argument does not apply to the double stranded molecules, as their much larger persistence length ensures that such entropic coiling within the pore is negligible.

To test whether this effect could account for the experimental observations, we have fitted the ΔG values for the homopolymers with a simple model based on the well-accepted worm-like chain description of polymer stretching[40] (Figure **4.3a**, solid lines). The worm-like chain model relates the stretching force F on a polymer to its fractional extension, x/L_0. Our model assumes that the conductance change ΔG is inversely related to the relative extension x/L_0, such that low x/L_0 results in high ΔG, and vice-versa: $\Delta G = \frac{b}{c + x/L_0}$, where b and c fit parameters. Next, we assume that the stretching force F is proportional to the applied bias voltage[18], F = $f \cdot V$, where f is a conversion constant in pN/mV. Making these substitutions into the worm-like chain model (Appendix **4**) results in the following relationship between the applied voltage and ΔG:

$$V(\Delta G) = a\left\{\frac{1}{4(1-\frac{b}{\Delta G}+c)^2} - \frac{1}{4} + \frac{b}{\Delta G} - c\right\},$$

where a, b and c are free fitting parameters. The red, magenta and green solid lines in Figure **4.3a** are fits of this model to our data, and can be seen to describe trends observed in the data well. Note that the worm-like chain description assumes that a constant force acts on the polymer ends, which may not be strictly true in our case: the stretching force arises from a complicated combination of electrical forces and viscous drag forces, and is expected to vary in magnitude along the polymer, peaking at some point above the nanopore (Figure **4.3d**) where the molecule may not yet sense the influence of the electrical field. For simplicity, however, we have assumed that the integral of this stretching force over the entire region of the nanopore simply leads to an *average* extension x/L_0 of the molecule inside the pore.

From the resulting fit parameters, we can extract values for the conductance blockade ΔG_{max} at zero extension ($x/L_0 = 0$), ΔG_{min} at full extension ($x/L_0 = 1$), and the conversion constant f relating the applied voltage to an effective stretching force. The values we find for ΔG_{max} are 1.86 ± 0.48 nS for poly(A), 5.53 ± 4.25 nS for poly(C), and 4.22 ± 1.77 nS for poly(U). In this case, the polymer experiences no force and can be approximated as a random coil of approximate volume $V=(4/3)\pi R_g^3$, where R_g is the radius of gyration. One can estimate that the relative volume comprised by the polymer itself inside this volume is on the order of 1 to 5% for the homopolymers used here (using different persistence lengths for the homopolymers from Seol, et al[38,39]). Thus, approximately 1 to 5% of the pore volume (and thus of the ionic current) will be blocked by the polymer. Given that the total conductance for a 10 nm nanopore is about 75-80 nS, the values of ΔG_{max} are in good agreement with this, assuming that the conductance blockade is indeed proportional to the volume occupied by the polymer inside the nanopore.

Second, we can compare the plateau values ΔG_{min} derived from our fit to the expected blockade for single-stranded molecules passing through the pore in a fully stretched manner, based on simple volumetric comparison of the molecules. Again, if the conductance blockade is indeed proportional to the volumes occupied by these molecules inside the nanopore, the ΔG_{min} values from our fit would reflect the volume occupied by the fully stretched homopolymers. We therefore determined the solvent inaccessible volumes of equally long sections of the fully stretched homopolymers and compared these to that of DNA from our computer models (Materials and methods). We find the following volume ratios $V_{poly(A)}:V_{DNA}$ = 0.34 , $V_{poly(C)}:V_{DNA}$ = 0.26 , and $V_{poly(U)}:V_{DNA}$ = 0.31. These are to be compared to the ratios $\Delta G_{min}/\Delta G_{dsDNA}$ that we deduce from the fits, where ΔG_{dsDNA} (= 1.6 ± 0.5 nS) equals the value of ΔG at 100 mV, selected as the ΔG for DNA at higher voltages of 600 mV is unexpectedly high. These ratios equal 0.41 ± 0.12 for poly(A), 0.25 ± 0.18 for poly(C) and 0.18 ± 0.13 for poly(U), in reasonable agreement with the expected values. However, we note that van Dorp et al. have observed that even in the static

situation where DNA is arrested in the nanopore using optical tweezers, a full explanation of ΔG is challenging[41] due to the unknown effects of the molecule above the nanopore on the access resistance of the nanopore. More precise quantification would require further investigation of this issue.

Finally, we can extract the value of f from the fit. Optical tweezer measurements on stalled dsDNA inside a nanopore yielded a force of ~ 24 pN/100 mV or f = 0.24 pN/mV in pores of similar size as used in this study. Taking into account the approximately two to three-fold lower charge density for a single-stranded homopolymer at almost full extension, we would expect f ~ 0.1 pN/mV in our case. However, the value for f that we find from our fit is significantly lower, ~ 0.008 ± 0.004 pN/mV. Potential reasons for this discrepancy include the fact that in our experiments the molecules are moving, which may yield quite different forces acting on the polymer. It is also possible that the charge screening and subsequent drag forces of the counter ions may be significantly different for single-stranded polymers than for DNA, which might give rise to a lower force. Future experiments with optical tweezers probing the stretching force on the homopolymers could shed light on these issues (see Chapter **8**).

An exciting prospect is that the distinctive ΔG values observed for each molecule at high applied bias imply that one could use solid-state nanopores to discriminate between single and double-stranded nucleic acid molecules based on current blockade; and it may even be possible to use this phenomenon for discriminating purine and pyramidine bases (Figure **4.3a**). In our proof-of-principle "mixing" experiment (Figure **4.4**) at a voltage of 600 mV, dsDNA exhibits a larger current blockade than poly(A) molecules at the same voltage. Potentially, by counting the numbers of translocation events of each kind, it could become possible to measure the absolute concentrations of each of double and single-stranded nucleic acid molecules within a mixture, in a single measurement. This type of application amounts to molecular "Coulter Counting"[42] and could have wide utility in measuring absolute concentrations of diverse biopolymer mixtures in a single measurement.

4.5 Acknowledgments

We thank S. W. Kowalczyk, A. R. Hall, M.Y. Wu, and V. Svechnikov for fabrication of the nanopores, S. Hage, S. Donkers and for assistance with molecule preparation, J. van der Does and J. Beekman for assistance with instrument construction. We especially thank I. Vilfan for his expertise in handling and preparation of our RNA molecules. We also wish to thank S. Lemay and K. Visscher for their valuable discussions. NHD acknowledges financial support from the TU Delft and from the European Science Foundation via an EURYI award.

Bibliography

1. Bustamante, C., Z. Bryant, and S. B. Smith (2003) Ten years of tension: single-molecule DNA mechanics. *Nature.* **421**: 423-427.
2. Larson, M. H., et al. (2008) Applied force reveals mechanistic and energetic details of transcription termination. *Cell*. **132**: 971-982.
3. Liphardt, J., et al. (2001) Reversible unfolding of single RNA molecules by mechanical force. *Science*. **292**: 733-737.
4. Onoa, B., et al. (2003) Identifying kinetic barriers to mechanical unfolding of the T-thermophila ribozyme. *Science*. **299**: 1892-1895.
5. Dekker, C. (2007) Solid-state Nanopores. *Nat. Nanotech*. **2**: 209-215.
6. Healy, K., B. Schiedt, and A. P. Morrison (2007) Nanopore-based single-molecule DNA analysis. *Nanomedicine*. **2**: 875-897.
7. Deamer, D. and D. Branton. (2002) Characterization of nucleic acids by nanopore analysis. *Acc. Chem. Res.* **35**: 817-825.
8. Kasianowicz, J.J., et al. (1996) Characterization of individual polynucleotide molecules using a membrane channel. *Proc. Natl. Ac. Sci. USA*. **93**: 13770-13773.
9. Smeets, R. M., et al. (2009) Translocation of RecA-Coated Double-Stranded DNA through Solid-State Nanopores. *Nano Lett.* **9**: 3089-3095.
10. Song, L., et al. (1996) Structure of staphylococcal alpha-hemolysin, a heptameric transmembrane pore. *Science*. **274**: 1859-1865.
11. Kang, X.F., et al. (2007) A Storable Encapsulated Bilayer Chip Containing a Single Protein Nanopore. J. Am. Chem. Soc. **129**: 4701-4705.
12. Meller, A., L. Nivon, and D. Branton (2001) Voltage-driven DNA translocations through a nanopore. *Phys. Rev. Lett.* **86**: 3435.
13. Khulbe, P. K., M. Mansuripur, and R. Gruener (2005) DNA translocation through alpha-hemolysin nanopores with potential application to macromolecular data storage. *J Appl Phys*. **97**: 104317-104321.
14. Mathe, J., et al. (2004) Nanopore unzipping of individual DNA hairpin molecules. *Biophys J*. **87**. 3205-3212.
15. Akeson, M., et al. (1999) Microsecond time-scale discrimination among polycytidylic acid, polyadenylic acid, and polyuridylic acid as homopolymers or as segments within single RNA molecules. *Biophys. J.* **77**: 3227-3233.
16. Li, J., et al. (2003) DNA molecules and configurations in a solid-state nanopore microscope *Nat. Mater*. **2**: 611-615.
17. Storm, A. J., et al. (2003) Fabrication of solid-state nanopores with single-nanometre precision. *Nat. Mater.* **2**: 537-40.
18. Keyser, U. F., et al. (2006) Direct force measurements on DNA in a solid-state nanopore. *Nat. Phys.* **2**: 473-477.
19. Storm, A. J., et al. (2005) Fast DNA translocation through a solid-state nanopore. *Nano Lett.* **5**: 1193-1197.

20. Trepagnier, E.H., et al. (2007) Controlling DNA capture and propagation through artificial nanopores. *Nano Lett.* **7**: 2824-2830.
21. McNally, B., M. Wanunu, and A. Meller (2008) Electromechanical Unzipping of Individual DNA Molecules Using Synthetic Sub-2 nm Pores. *Nano Lett.* **8**: 3418-3422.
22. Saenger, W., J. Riecke, and D. Suck (1975) A structural model for the polyadenylic acid single helix. *J. Mol. Biol.* **93**: 529-534.
23. Arnott, S., R. Chandrasekaran, and A.G. Leslie (1976) Structure of the single-stranded polyribonucleotide polycytidylic acid. *J. Mol. Biol.* **106**: 735-748.
24. Saenger, W. (1984) Principles of Nucleic Acid Structure. Springer-Verlag: New York. p 556.
25. Storm, A. J., et al. (2005) Translocation of double-strand DNA through a silicon oxide nanopore. *Phys. Rev. E*. **71**: 051903.1-051903.5.
26. Gershow, M., and J. A. Golovchenko (2007) Recapturing and trapping single molecules with a solid-state nanopore. *Nat Nanotech.* **2**: 775-779.
27. Vilfan, I.D., et al. (2007) An RNA toolbox for single-molecule force spectroscopy studies. *Nucleic Acids Res.* **35**: 6625-6639.
28. van den Hout, M., et al. (2008) End-joining long nucleic acid polymers. *Nucleic Acids Res.* **36**: e104.
29. Broido, M. S., and D. R. Kearns (1982) Proton NMR evidence for a left-handed helical structure of poly(ribocytidylic acid) in neutral solution. *J. Am. Chem. Soc.* **104**: 5207-5216.
30. Shenogin, S., and R. Ozisik (2005) Deformation of Glassy Polycarbonate and Polystyrene: The Influence of Chemical Structure and Local Environment. *Polymer*. **46**, 4397-4404.
31. Drew, H. R., S. Samson, and R. E. Dickerson (1982) Structure of a B-DNA dodecamer at 16 K. *Proc. Natl. Ac. Sci. USA*. **79**: 4040-4044.
32. Liu, L., et al. (2008) Structural basis of toll-like receptor 3 signalling with double-stranded RNA. *Science*. **320**: 379-381.
33. Tabard-Cossa, V., et al. (2007) Noise analysis and reduction in solid-state nanopores. *Nanotechnology*. **18**: 305505.1-5.
34. Smeets, R. M. M., et al. (2006) Salt dependence of ion transport and DNA translocation through solid-state nanopores. *Nano Lett.* **6**: 89-95.
35. Fologea, D., et al. (2005) Detecting single stranded DNA with a solid-state nanopore *Nano Lett.* **5**: 1905-1909.
36. Wanunu, M., et al. (2008) DNA translocations governed by interactions with solid-state nanopores. *Biophys. J.* **95**: 4716-4725.
37. Arnott, S., R. Chandrasekaran, and A. G. Leslie (1976) Structure of the single-stranded polyribonucleotide polycytidylic acid. *J. Mol. Biol.* **106**: 735-48.
38. Seol, Y, et al. (2007) Stretching of homopolymeric RNA reveals single-stranded helices and base-stacking. *Phys. Rev. Let.* **98**: 158103.1-158103.4.

39. Seol, Y., G. M. Skinner, and K. Visscher (2004) Elastic properties of a single-stranded charged homopolymeric ribonucleotide. *Phys. Rev. Let.* **93**: 118102.1-118102.4.
40. Marko, J. F., and Siggia, E. D. (1995) Stretching DNA. *Macromolecules*. **28**: 8759-8770.
41. van Dorp, S, et al. (2009) Origin of the electrophoretic force on DNA in solid-state nanopores. *Nat. Phys. **5**: 347-451.*
42. COULTER, W. H. (1953) U. S. Patent No. 2656508, issued Oct. 20. 1953.

Appendix 4

Determination of current blockade

The histograms presented in Figures **4.2-4.4** were composed from current traces sampled at a frequency of 150 kHz or 200 kHz. Generally, the traces were digitally low-pass filtered at a frequency determined by the criteria discussed below (see **Filtering**), prior to event detection. For high voltage data (V_{bias} > 500 mV), low-pass filtering was frequently not necessary. Events were then selected according to the following algorithm:

1. The filtered current trace is parsed into segments of 5000 data points I_x, and for each segment the average <I> and standard deviation σ are determined (Figure **A4.1**). Within each segment, events are identified whenever an excursion of amplitude 4.5σ from the average <I> is detected.
2. The duration of each event is determined by finding the positions x_{start} and x_{end}, at which the current values are within 0.5σ from the average <I>. The duration of the event is defined as $(x_{end} - x_{start})*\Delta t$ (where Δt is the minimum time step, $\Delta t = 1/f_{sample}$). Events shorter than a minimum duration t_{min} and longer than t_{max} are discarded (typical values: t_{min} = 40 µs, t_{max} = 10 ms). For each remaining event, the data corresponding to the indices $[x_{start} - n, x_{end} + n]$ is selected (Figure **A4.1**, bottom panel) and added to the final output trace. A check is performed that overlapping points for neighbouring events are removed.
3. The output traces for each segment are then merged into a final trace, from which a histogram of the current values is deduced (see Figures **4.2**, **4.4**, and **A4.4**). Gaussian functions are then fitted to each peak and the current blockade is determined from the difference in the positions of the peaks.

Filtering

The frequency at which the original current trace was low-pass filtered prior to event detection varied slightly for different datasets. Its value was selected depending on the average noise level of the current in the nanopore, as well as the signal-to-noise ratio of the individual events. For lower voltages, the filter frequency was chosen such that peaks could clearly be distinguished from the average noise level, all the while verifying that they were not significantly distorted by the filter.

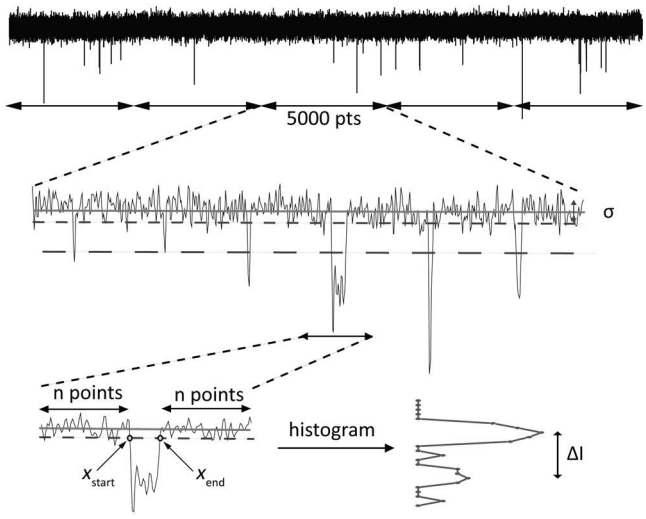


Figure A4.1. *Explanation of the event detection algorithm*. See text for a full description.

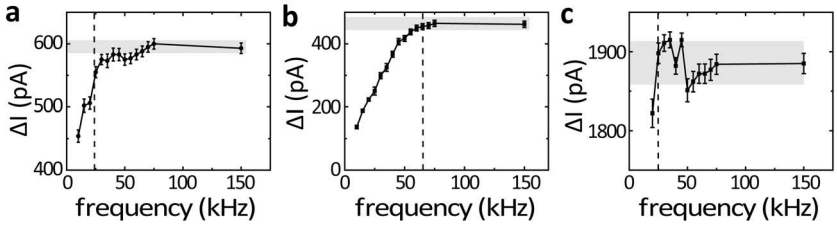


Figure A4.2. *Influence of filtering the raw data before event detection*. ΔI as determined from the histograms is plotted at different filter frequencies for poly(A) (**a**), poly(C) (**b**) and DNA (**c**) at 600 mV. In this manner, the minimum filter frequency at which the value ΔI is not distorted by the filter can be determined for each dataset.

To prevent possible filtering distortion, we analyzed the current blockade deduced for datasets low-pass filtered at different frequencies. An example of the resulting current blockade as a function of filter frequency is shown in Figure **A4.2a** for poly(A) at 600 mV. One can observe that the values level off at filter frequencies

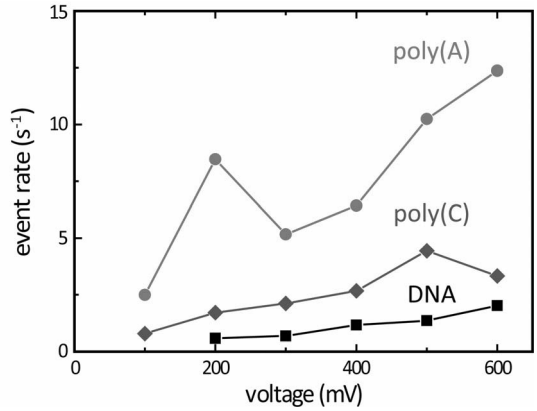


Figure A4.3. *Typical average event rate for representative molecules*. In our experiments, the molecule concentrations were chosen such that the event rate was always low enough that it is very unlikely that two molecules simultaneously traverse the pore.

of 30 kHz and higher. At lower frequencies, the computed current blockades are distorted and therefore cannot be trusted. In Figures **A4.2b,c**, a similar analysis is performed for poly(C) and DNA at 600 mV, respectively. For this poly(C) dataset, the minimum low-pass filter frequency that can be applied equals 55 kHz. This higher value is likely a consequence of the shorter mean lengths obtained in poly(C) synthesis (see Materials and methods), which result in a concomitant reduction of the translocation times.

In a few cases the minimum filter frequency could not easily be determined via this method, because at higher filter frequencies no clear peaks could be determined from the histograms. In these cases, a different method for avoiding event distortion as a result of filtering was employed. Specifically, the criterion for the minimum duration t_{min} for each event was adapted, so that only events with translocation times longer than $2*(1/f_{filter})$ were selected. Events satisfying this condition are also not significantly distorted by the influence of the filter.

Event rate

To ensure that all peaks in the current histograms correspond to contributions from single molecules only, all measurements were performed with low molecule concentrations (1-10 ng/µl). In Figure **A4.3**, we plot the typical event rates as a function of voltage for poly(A), poly(C), and dsDNA: the maximum rate is approximately 10 events/s. As the average translocation time is on the order of 100 µs, there is typically a 0.1% chance of finding a molecule in the pore. The likelihood

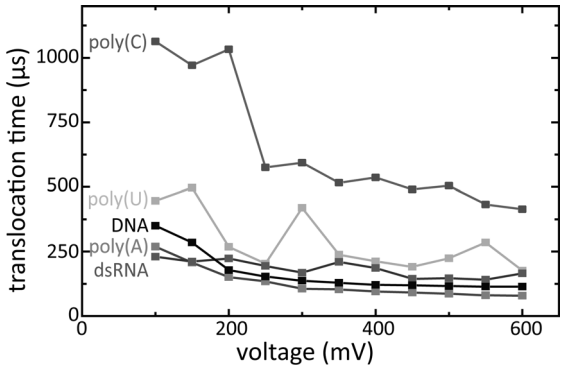


Figure A4.4. *Examples of the average translocation time for each of the molecules versus the applied bias voltage*. In general, the translocation time decreases at higher voltages. The long translocation time for the poly(C) data is in this particular dataset probably due to interactions with the nanopore surface.

of two molecules simultaneously traversing the pore is therefore negligible, which implies that the observed peaks in the histograms reflect the contributions from single molecules.

Translocation Time

The translocation times are plotted as a function of the applied bias voltage for all molecule types sampled (Figure **A4.4**). From this representative dataset, it can be seen that the translocation time roughly scales as 1/V, as expected. The relatively long translocation times of the poly(U) and poly(C) homopolymers were rather unexpected, and may reflect interactions with the nanopore. However, for other datasets we have observed significantly shorter translocation times for poly(C), indicating that the level of interaction with the nanopore surface may vary between different experiments. Our datasets do not, however, demonstrate any correlation between the translocation time and the conductance blockade per events.

Modelling

In the main text, we have proposed a model based on entropic stretching to account for the observed decrease in ΔG as a function of voltage for the homopolymers. In this model, we assume that the conductance blockade ΔG is

inversely proportional to the amount of polymer material inside the nanopore, and thus inversely proportional to the relative extension x/L_0. We therefore express the conductance blockade as:

$$\Delta G = \frac{b}{\frac{x}{L_0} + c}, \text{ or conversely: } \frac{x}{L_0} = \frac{b}{\Delta G} - c$$

where *b* and *c* are fit parameters. Second, we assume that the stretching force is proportional to the applied bias voltage: $F = f \cdot V$, where f is a conversion factor in pN/mV. For any flexible polymer, the force–extension relationship is well described by the worm-like chain model[40]:

$$F(\frac{x}{L_0}) = \frac{k_B T}{L_P}\left\{\frac{1}{4(1-\frac{x}{L_0})^2} - \frac{1}{4} + \frac{x}{L_0}\right\}$$

Using the expressions above, this can be rewritten as:

$$V(\Delta G) = a\left\{\frac{1}{4(1-\frac{b}{\Delta G}+c)^2} - \frac{1}{4} + \frac{b}{\Delta G} - c\right\}, \text{ with } a = \frac{k_B T}{L_P f} \text{ a free parameter}$$

This expression can be fitted to the homopolymeric data in Figure **4.3a**. From the fit, we can extract values for the fit parameters $\Delta G_{\min} = \frac{b}{c+1}$, the conductance blockade at full extension; $\Delta G_{\max} = \frac{b}{c}$, the conductance blockade at zero extension (or V = 0); and $f = \frac{k_B T}{L_p a}$, the conversion factor in pN/mV. These fit parameters have the following values:

	ΔG_{min}	ΔG_{max}	f
Poly(A)	0.65 ± 0.17 nS	1.86 ± 0.48 nS	0.0076 ± 0.0029 mV/pN
Poly(C)	0.39 ± 0.29 nS	5.53 ± 4.25 nS	0.0062 ± 0.0014 mV/pN
Poly(U)	0.29 ± 0.20 nS	4.22 ± 1.77 nS	0.0089 ± 0.0038 mV/pN

These values are discussed in the main text.

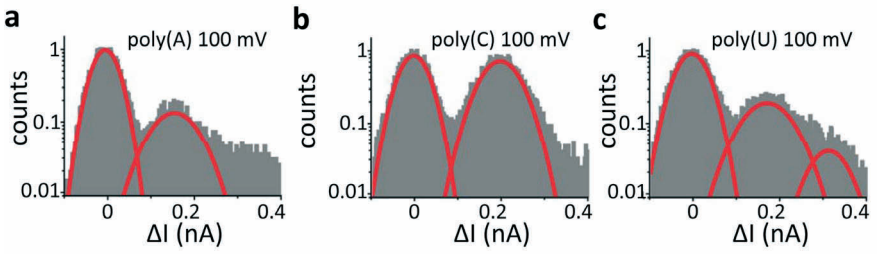


Figure A4.5. *Typical histograms of the observed current blockade for the homopolymers poly(A)* **(a)**, *poly(C)* **(b)** *and poly(U)* **(c)** *at 100 mV bias voltage*. The peak at ΔI=0 nA corresponds to the baseline current, while the other peaks in the current blockade are due the presence of molecules in the nanopore. In subfigures **(a)** and **(c),** a peak corresponding to two strands in the pore can be observed.

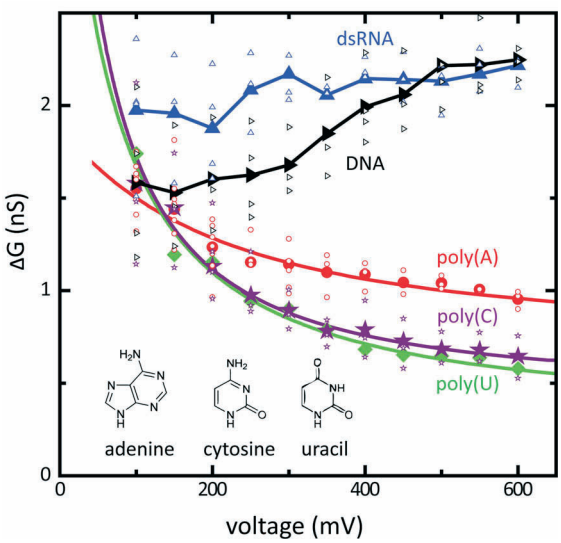


Figure A4.6. *Extended version of Figure **4.3**.* This plot contains the same information as Figure **4.3** in the main text, but explicitly shows all measured data points to indicate their full spread: open symbols represent individual measurements, whereas the closed symbols correspond to the average values. The double-stranded molecules showed more variation than the single-stranded homopolymers. The mixing experiment from Figure **4.4** is not included in this graph.

Chapter 5

Direct Force Measurements on Double-Stranded RNA in Solid-State Nanopores

Abstract

Solid-state nanopores can be employed to detect and study local structure along single molecules by voltage driven translocation through the nanopore. Their sensitivity and versatility can be augmented by combining them with a direct force probe, e.g. optical tweezers. Such a tool could potentially be used to directly probe RNA secondary structure through the sequential unfolding of duplex regions. In this chapter, we demonstrate the first application of such a system to the study of RNA by directly measuring the net force on individual double-stranded RNA (dsRNA) molecules. We have probed the force on dsRNA over a large range of nanopore sizes from 35 nm down to 3.5 nm and find that it decreases as the pore size is increased, in accordance with numerical calculations. Furthermore, we find that the force is independent of the distance between the optical trap and the nanopore surface, permitting force measurement on quite short molecules. By comparison with dsDNA molecules trapped in the same nanopores, we find that the force on dsRNA is on the order of, but slightly lower than that on dsDNA. With these measurements we expand the possibilities of the nanopore-optical tweezers to the study of RNA molecules, with potential applications to the detection of RNA-bound proteins, the determination of RNA secondary structure, and the processing of RNA by molecular motors.

This chapter was published as: Michiel van den Hout, Igor D. Vilfan, Susanne Hage, and Nynke H. Dekker, *Nano Letters.* **10**: 701-707 (2010)

5.1 Introduction

Recently, a new tool for single-molecule studies has emerged in the form of nanopores. Here, charged molecules of interest immersed in an electrolyte are translocated through a narrow pore of several nm in cross-section by the application of a strong electrical field over the pore. When the molecule is inside the nanopore, it blocks a certain portion of the ionic current. Using this approach, nanopores can not only be employed for molecular counting, but also for the analysis of molecular structure or nucleic acid-protein interactions. For example, they can be used to detect single proteins in solution[1], or to monitor protein binding on dsDNA[2-3]. More recently, it was shown that solid-state nanopores can also be employed for the study of RNA: homopolymeric ssRNA was shown to be easily distinguishable from dsDNA and dsRNA as it traverses 8-10 nm pores, which was attributed to stretching of the flexible ssRNA on the length scale of the pore[4]. Both biological and solid-state nanopores have also been proposed as candidate systems for genomic sequencing of dsDNA[5-6], and indeed, it has recently been shown in the biological α-Hemolysin nanopore that different bases could be distinguished, albeit under the condition that they are translocated through the nanopore as isolated nucleotides[7].

The versatility of solid-state nanopores in particular can be enhanced through their integration with optical tweezers, permitting the controlled positioning of molecules with respect to the pore and the direct measurement of the forces acting on them[8-9]. In such an experiment, to date demonstrated for dsDNA[8], molecules are attached to an optically trapped bead and moved close to the nanopore where the electrical field can 'capture' the molecule into the pore. In contrast to free translocation experiments, here the optically trapped bead prevents the molecule's subsequent translocation to the other side of the pore. Furthermore, the trapped bead's position directly monitors the force applied to the molecule in the nanopore. Such a force measurement is performed in parallel with the ionic current recording, providing a simultaneous readout of the molecule's spatial configuration in the nanopore and the force to which it is subjected.

Since both the force and the nanopore current report on the local charge density and size of the molecules inside the nanopore[10-11], their combined readout can serve as an indicator for the presence of proteins or other local structure along nucleic acids. In addition, forces could be selectively applied to proteins bound to nucleic acids simply by choosing a nanopore size smaller than the nucleic acid-protein complex. The use of a local force probe holds particular interest for the study of RNA as it provides a method to sequentially unravel and probe its secondary structure[12-13]. However, for this to be accomplished, several challenges must be overcome. First, while translocations of RNA molecules through biological nanopores[14-15] and, more recently, solid-state nanopores[4] have been demonstrated, the use of the combined nanopore-optical tweezers system has not yet been applied to RNA. One major difficulty here is that trapping a single

molecule inside the nanopore requires the optically trapped bead to be moved into close proximity to the nanopore, so that the molecule can enter it with reasonable probability. Upon capture, however, the bead must be prevented from being pulled against the nanopore membrane to avoid aspecific attachment. In practice, this imposes a criterion on the minimum length of the molecule under study: its radius of gyration (R_g) should be at least on the order of hundreds of nanometers, as this is the typical distance the bead is displaced from its equilibrium position in the optical trap to balance the force on the molecule. Specifically for RNA this is a stringent requirement, as many RNA molecules are single-stranded and thus very flexible, resulting in R_g ~ 10-100 nm even when they comprise many kilobases (kb). Even for dsRNA[16], which is less flexible[17-18], this criterion is challenging, as the biochemical synthesis of such molecules rapidly decreases in efficiency for transcript lengths greater than 5 kilo base pair (kbp). In addition, one must consider the potential influence of the resulting short contour lengths on the measured experimental signals, i.e. the force and nanopore current.

Here, we demonstrate the first study of RNA in the combined nanopore-optical tweezers apparatus. We have successfully synthesized a long (12.7 kbp) dsRNA molecule that can readily be captured inside our nanopores, and we present the forces measured on these dsRNA molecules, comparing these to the forces on dsDNA molecules of the same contour length. We show that the ionic current blockades occasioned by single dsRNA and dsDNA molecules upon capture in the nanopore are in excellent agreement with results from translocation studies[4]. Furthermore, we demonstrate that the measured forces are independent of the distance of the bead to the nanopore, down to distances as close as 250 nm from the nanopore, which would allow such measurements to be performed with molecules of even shorter contour lengths. Finally, we show that our results agree qualitatively with numerical calculations of the forces on these molecules over a large range of nanopore sizes from 35 nm down to 3.5 nm in diameter. In particular, the exploration of this small nanopore size range opens the way for the study of local structure of smaller molecules such as single-stranded RNA, for which secondary structure determination using established techniques such as NMR and X-Ray scattering remains challenging[19].

5.2 Materials and methods

Experimental setup

Our combined nanopore-optical tweezers setup has been previously described[9,20], with several minor differences: the trapping laser used is a 4W Coherent Compass 1064-4000M (Coherent, Santa Clara, US), and the piezo stage employed is a P561.3CD (Physik Instrumente, Karlsruhe, Germany).

Sample preparation

Nanopores were fabricated in 20 nm SiN membranes by a highly focused TEM beam as previously described[21] and stored in a solution containing 50% ethanol and 50% ddH_2O directly after drilling. Prior to use, the SiN membrane was thermally oxidized at 200 ºC for 10-30 min before use in order to reduce aspecific attachment of beads or nucleic acids to the surface. Subsequently, the sample was treated with O_2 plasma at room temperature for 30 seconds, mounted into the flow cell, and immediately flushed with electrolyte. All experiments were performed in 1M KCl, 10 mM Tris-HCl (pH=8.0), 1 mM EDTA and 0.01% Triton-X100 (Roche, Switzerland). These high ionic strength conditions were chosen for a number of reasons: good signal-to-noise in the nanopore current, reduced sticking of the molecules to the nanopore, optimal capturing efficiency over the entire range of nanopore sizes studied, as well as straightforward comparison with published data[4,8-9]. Conversely, the influence of low ionic conditions on the forces experienced in nanopores is not yet well understood[11]. The addition of Triton-X100 was found to further reduce sticking of the beads to the membrane.

All experiments were performed with streptavidin-coated polystyrene microbeads of 1.87 ± 0.03 µm in diameter (Spherotech, IL, USA). The 12.7 kbp dsDNA fragment was synthesized via PCR amplification as previously described[22]. A single biotin molecule was incorporated at the 5´ end of the forward PCR primer, allowing for the specific attachment of one end of the molecule to our beads. The 12.7 kbp dsRNA was synthesized via in vitro transcription of the two complementary strands of a different dsDNA construct that was obtained by reverse transcription of the equine arteritis virus (EAV) genome and subsequent PCR amplification thereof[23]. The strand corresponding to the original EAV genomic RNA was fully transcribed, whereas the complementary strand was transcribed into two adjoining segments of 530 bases and 12161 bases. Transcription of the 500-nt segment was performed in the presence of a 1:1 mixture of rUTP and bio-rUTP, causing the segment to have a number (up to 60, estimated) biotin-labeled bases, which allows for the specific attachment of the molecules to our beads. The three segments were then hybridized in a formamide-based buffer previously described by Liphardt and coworkers[24]. The hybridization mixture was incubated at 55 ºC for 3h, and subsequently cooled down to room temperature with a cooling rate of 1.5 ºC / 5 min. The hybridization mixture was immediately dialyzed against 2 L of 0.5 x SCC at 4 ºC[25]. Dialysis was ended after 2h, and the resulting RNA hybrid was purified by RNeasy (Promega), eluted with 1x TE and stored at -20 ºC.

Prior to each experiment, the nucleic acid molecules were incubated with the beads for at least 30 min in a solution containing 200 mM KCl and TE (10 mM Tris-HCl, 1 mM EDTA, pH = 8.0). For the experiments involving RNA, the beads were additionally rinsed, first twice with 0.1M NaOH and then twice with TE prior to the incubation with the RNA molecules, to reduce potential RNAse contamination. Concentrations of the molecules on the beads were chosen such that during the

experiments single molecules are easily captured in the nanopore (within tens of seconds), but capture of multiple molecules is avoided as much as possible. We found that for nanopores of 5nm and smaller this required molecule concentrations about 100x higher than those for large nanopores (>20 nm).

Trap calibration

Prior to each experiment, the trap stiffness of each individual bead was determined by a Lorentzian fit to the power spectrum of the bead position recorded over tens of seconds[26] at a distance of more than 20 μm from the nanopore surface. For beads that were incubated with very high molecule concentrations the apparent trap stiffness was found to be significantly (10-30%) lower than for uncoated beads of the same size. In these cases, we have used the average trap stiffness determined from the uncoated beads (as the presence of the DNA should not influence the trap stiffness[32]), and taken the difference between the two values as the error in the trap stiffness.

Modelling

Numerical calculation of the forces on dsRNA and dsDNA was done using code from van Dorp and coworkers[11]. The dsDNA and dsRNA are approximated as cylinders with uniformly distributed surface charges. The surface charge density of dsDNA, which has a linear charge density λ_{bare} of -0.96 nC/m and a diameter D of 2.0 nm, equals -0.15 C/m^2. The surface charge density of dsRNA, which has linear charge density λ_{bare} = -1.23 nC/m and D = 2.6 nm diameter, likewise equals -0.15 C/m^2. Values for the linear charge densities and diameters are obtained from the published crystal structures of these molecules (pdb ID 2BNA[27] for dsDNA and ID 3CIY for the dsRNA, after the bound protein, which does not distort the double-stranded helix of the A-RNA, was deleted[28]).

5.3 Results and discussion

We have measured forces on dsRNA and dsDNA molecules in nanopores integrated with optical tweezers. At the start of an experiment, an optically trapped bead with one or more molecules attached to its surface is held in the *cis* side in close vicinity of the nanopore (Figure **5.1a**, left). A voltage bias is then applied over the nanopore (using two electrodes on both sides of the nanopore), leading to a strong electrical field in the vicinity of the nanopore, and an ionic current passing through it. Together with the ionic current, we track the position of the bead by monitoring

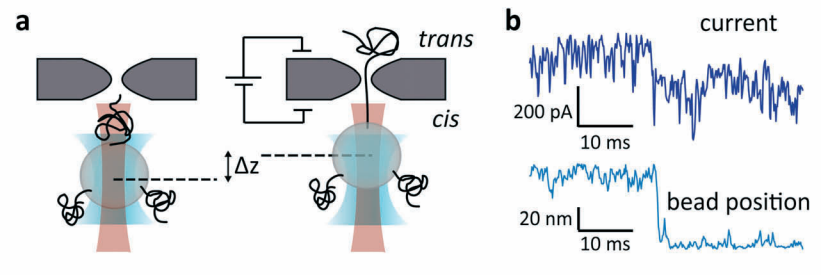


Figure 5.1. *Principle of the experiments*. **a**. Schematic of the experiment. A dsRNA molecule, attached to an optically trapped bead, is held close to the nanopore. The bead position is monitored by a separate detection laser (red). Application of a voltage bias across the nanopore results in an electrical field that pulls the charged molecule into the pore. The optical restoring force on the bead, however, prevents complete translocation of the molecule, and consequently the molecule is held taut inside the nanopore. **b**. Example of such a 'capture' event. When the molecule enters the pore, it blocks a portion of the ionic current through the nanopore (dark blue curve). This current drop is accompanied by a simultaneous change in the position of the trapped bead (light blue curve). The bead position stabilizes to a new equilibrium position at which the optical force balances the total electro-osmotic force on the molecule.

the optical reflection signal of a separate detection laser incident on the bead. When the bead is held in close proximity to the nanopore, the electrical field (which typically extends only tens to hundreds of nanometers into the solution) will pull the negatively charged nucleic acid molecule through the nanopore towards the *trans* side until the force is balanced by the optical force on the bead (Figure **5.1a**, right). The details of this force balance were described in detail recently[11,29] and will be discussed below in the context of our measurements. Upon capture into the nanopore, the ionic current through the nanopore is partially blocked by the presence of the molecule, which together with the change in bead position serves as a clear indication that a molecule is captured.

A typical capture event for a single dsRNA molecule in a 6 nm pore is shown in Figure **5.1b**. A short while after the application of a bias voltage across the pore (+100 mV in this experiment), the current signal is observed to suddenly decrease in a step-wise fashion (Figure **5.1b**, dark blue line). Simultaneously the position signal changes, commensurate with movement of the bead towards the nanopore (Figure **5.1b**, light blue line). From the change in the current signal, we compute the change in nanopore conductance ΔG ($= \Delta I/V_{bias}$) which for this event equals 2.07 ± 0.35 nS. This value is in good agreement with the ΔG observed for untethered dsRNA molecules translocating through similarly-sized nanopores[4], for which an average value of 2.0 ± 0.4 nS was found at this voltage. The change in the position signal of the bead (in V) can be related to an actual displacement Δz (Appendix **5**) of

approximately 64 ± 5 nm, which corresponds to an effective force of 8.6 ± 2.3 pN given the typical trap stiffness of 135 ± 25 pN/µm (Materials and Methods).

By sequentially trapping dsRNA and dsDNA in the same nanopore, we can directly compare both the conductance changes caused by the presence of these molecules in the nanopore and the forces exerted on them under the influence of the electric field. In Figure **5.2a** we show the changes in conductance and bead position occasioned by dsRNA (blue trace) compared to that of dsDNA (black trace), sequentially trapped in a 22 nm nanopore. Focusing first on the conductance changes, it is clear from Figure **5.2a** that the change in conductance is very similar for the two molecules, which is not unexpected given their similar dimensions (Figure **5.2d**). Histograms summarizing all our measurements of the conductance change occasioned by trapped molecules of dsRNA and dsDNA, where these changes were determined either upon capture into or extraction out of the pore, are shown in Figure **5.2b**. For the dsRNA we find a mean value of ΔG_{dsRNA} = -1.96 ± 0.70 nS (mean and standard deviation determined from a maximum likelihood Gaussian fit to the data). This value is in excellent agreement with translocation results of dsRNA of similar length through nanopores of the same diameter and thickness, where a mean value of ΔG_{dsRNA} = -2.0 ± 0.4 nS was found at 100 mV bias voltage[4]. Similarly, for the conductance blockade from dsDNA we find an average value of ΔG_{dsDNA} = -1.97 ± 0.81 nS, which is again in good agreement with results obtained from translocation studies involving the same dsDNA fragment. There, a mean value of ΔG_{DNA} = -1.6 ± 0.5 nS was found at 100 mV bias voltage[4]. The value for ΔG_{dsDNA} is somewhat higher than the value previously reported in nanopore-optical tweezers trapping experiments[8], which is readily explained by the fact that these previous experiments were performed with nanopores formed in thicker membranes, resulting in a smaller change in conductance.

We now compare the forces experienced by dsRNA and dsDNA molecules as they are tethered in the nanopore. The electrical field in the pore vicinity causes a force on the molecules towards the *trans* side, as the phosphodiester backbone of nucleic acids is negatively charged in aqueous solution. At the same time, positively charged counterions screening the nucleic acid are forced by the electrical field to flow in the opposite direction, leading to an electro-osmotic drag force from the fluid on the molecule in the direction of the *cis* side[11]. The resulting net force (always directed towards the *trans* side) is balanced by the optical restoring force on the bead, which we determine directly by measuring the bead position. Comparing the structures of the molecules (Figure **5.2d**), it is not a priori clear how the force on dsRNA should compare to that on dsDNA. Certainly, the purely electrical force on dsRNA towards the *trans* side should be higher, reflecting the 30% increase in its linear charge density compared to dsDNA. However, this increase is partially cancelled by a larger electro-osmotic drag force in the opposite direction, as the surface area that experiences fluid drag from the counterions is also increased (the dsRNA helical diameter being ~30% bigger than that of dsDNA). Since the magnitude of this latter effect also depends on the pore size as discussed

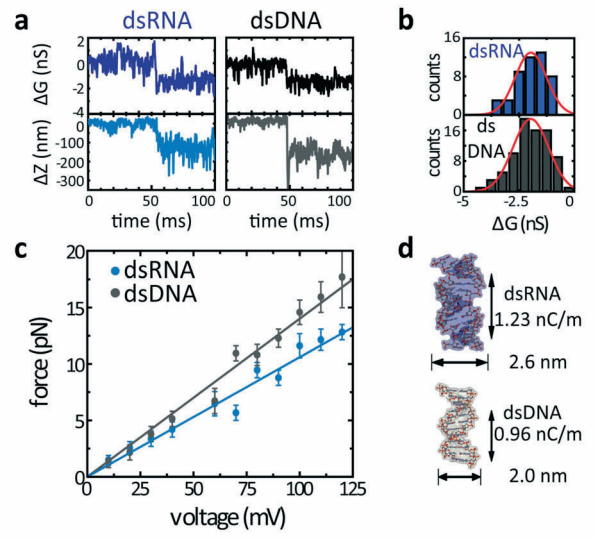


Figure 5.2. *Capturing dsRNA and dsDNA molecules*. **a**. Comparison of the raw data in a 22 nm pore at 100 mV bias voltage illustrating captures of a dsRNA (dark- and light-blue) and dsDNA molecule (black and grey). The conductance signal (top, dark traces) and bead position signal (bottom, light traces) are shown for both types of molecules. **b**. Histograms of the conductance change ΔG (= $\Delta I/V_{bias}$) for dsRNA and dsDNA, extracted from the current signal on capture or extraction. The values for ΔG compare well to earlier results in which molecules were freely translocated through similarly-sized pores[4]. **c**. Force versus voltage relationship for dsRNA (light blue) and dsDNA (grey) molecules successively captured in the same 22 nm pore. The solid lines correspond to linear fits to the data. As the bias voltage is varied, the corresponding change in bead position Δ*z* is measured, from which the force can be calculated. Here, the initial bead position is ~0.75 μm from the pore surface, and error bars stem from the uncertainties in the determination of the trapping strength and the bead position (Appendix **5**). **d**. Images of the dsRNA and dsDNA helices as determined via x-ray crystallography, indicating the molecular diameters and charge densities[27,28].

below, the difference between the net force on dsRNA and dsDNA is not readily estimated.

By varying the bias voltage across the nanopore we can directly determine force curves for dsRNA (Figure **5.2c**, blue) and for our control molecule, dsDNA (Figure **5.2c**, grey). We find that the measured force on the bead depends linearly on the voltage for all molecules probed, allowing us to determine the force per voltage f on the molecule from the slope of a linear fit to the forces at different bias voltages. For the measurements shown in Figure **5.2c**, conducted in the same 22 nm pore as used in Figure **5.2a-b**, the linear fit through the points yields a force per voltage of f_{dsRNA} = 0.11 ± 0.02 pN/mV and f_{DNA} = 0.14 ± 0.03 pN/mV, where the

errors reflect the uncertainties in the determination of both the trap stiffness and the position of the beads. Therefore, within error, the forces dsRNA and dsDNA are roughly equal in this nanopore.

Prior to investigating this in greater detail, we will first address an important question related to performing these force measurements with short molecules and at very close distances to the nanopore: are the measured forces strongly dependent on the molecule length and the distance separating the bead from the nanopore membrane? The maximum distance achievable whilst maintaining a molecule inside the pore, is limited to the contour length of the molecules (a condition achieved at high voltages, where the molecule is fully stretched). For our 12.7 kbp dsRNA and dsDNA molecules this distance amounts to about 3.2 µm and 4.2 µm, respectively, at least four-fold shorter than in the case of the 48 kbp (~16 µm contour length) bacteriophage-λ dsDNA molecules utilized in previous measurements using this experimental setup[8,11,30]. To check first if the force depends on the length of the molecule, we have compared the forces on our shorter dsDNA molecules to those on λ-DNA in the same nanopores (at equal distances to the nanopore membrane) and found no noticeable difference (data not shown). Next, to see if the proximity of the nanopore membrane influences the optical trap stiffness and hence the values of the measured forces, we have investigated the forces over distances ranging from 250 nm to 3 µm from the nanopore membrane. Here, a molecule was initially captured in the pore, and subsequently the distance between the bead and the nanopore was varied while maintaining the molecule in the pore through the application of a small voltage bias (30 mV). Force curves such as in Figure **5.2c** were then recorded at different positions z, where z is the distance between the bead surface and the nanopore membrane (Figure **5.3a**). The resulting dependence of the force per voltage versus z for dsRNA molecules (blue) and dsDNA molecules (black) in a nanopore of 32 ± 4 nm diameter is plotted in Figure **5.3b**. Each data point corresponds to the average force per voltage at a given position z, $\langle f(z) \rangle$, determined from measurements in this nanopore repeated on different single molecules. For both dsRNA and dsDNA, we fit the data using a linear relation (solid lines) and determine accompanying slopes $\partial f(z)/\partial z$ of -12.7% / µm (dsRNA) and 1.2% / µm (dsDNA). Thus, for this particular pore, the average force changes on the order of ten percent per µm of distance to the pore membrane. To provide a complete picture including all our measurements on dsDNA and dsRNA in different nanopores, we took the following approach: for each molecule for which force curves (Figure **5.2b**) could be recorded at minimally 3 different z positions, we calculated the slope $\partial f(z)/\partial z$, normalized it to the average and expressed it into % change per µm, and binned it in a histogram (Figure **5.3c**). A maximum likelihood Gaussian fit to the values gives a mean value of 2 ± 17% (the error is the standard deviation). We therefore conclude that within experimental error there is no dependence of the force on the distance z to the membrane, even at distances as close as 250 nm from the surface. This result ensures that it is possible to measure on even shorter trapped molecules, such as

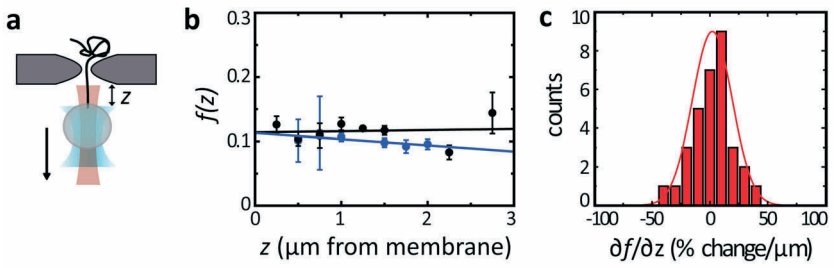


Figure 5.3. *Measured forces are independent of the separation between optical trap and nanopore*. **a**. Schematic. The bead position z is defined as the distance between the surface of the bead and that of the nanopore membrane. After capture the distance z is varied (signified by the black arrow) and at every position z the force per voltage $f(z)$ is calculated from force curves (Figure **5.2c**). **b**. Average force per voltage determined for dsRNA (blue) and dsDNA (black) as a function of bead position z in a 32 ± 4 nm pore. The dependence of the force per voltage on the bead position is determined by the slope of a linear fit to the data. **c**. Histogram of the slopes $\partial f(z)/\partial z$ expressed in % change/µm for all dsRNA and dsDNA molecules for which f was determined at more than two positions. The mean of a maximum likelihood Gaussian fit to the histogram is 2 ± 17%/µm, where the error is the standard deviation.

very short duplexes or even single-stranded RNA, without having to correct the values of the measured forces.

Above, we have discussed the force measured on dsRNA in a 22 nm nanopore, and found this to be roughly equal to the force on dsDNA in the same pore. However, for a thorough comparison, the forces must be probed in nanopores of different transverse sizes. This is because the nanopore size influences the net electro-osmotic force on the captured molecule[29,31] and hence the net force measured in our experiments. In a recent publication by Van Dorp et al.[11], this effect was demonstrated at low ionic strength (<50 mM KCl), where the total force on dsDNA was found to decrease as the size of the nanopore was increased. This was attributed to a hydrodynamic coupling of the dsDNA counterions to the nanopore surface: close to this surface, the non-slip condition on the water molecules causes a transfer of part of the counterions´ momentum to the molecule and to the nanopore surface. Therefore, the nanopore geometry influences the magnitude of the drag force exerted by these counterions on the nucleic acid molecule inside the pore. Specifically, for larger pores this coupling is less pronounced, and the drag force of the counterions is mostly directed against the nucleic acid molecule, leading to a lower net force measured in the optical tweezers with increased pore size.

To get a complete picture of the forces on dsRNA molecules, we have therefore investigated the dependence of this force on the size of the nanopore,

additionally probing dsDNA as a control. Measurements of the average force per voltage <f> on both dsRNA (blue) and dsDNA (black) over a tenfold variation in nanopore sizes from 3.5 to 35 nm in diameter are shown in Figure **5.4**. Each point in the graph is the mean force per voltage <f>, averaged over molecules on different beads probed at different distances z from the nanopore membrane (the latter is justified by the independence of f on z, see Figure **5.3**). We observe that as the nanopore size increases, the average force decreases monotonically for dsRNA, as it does for dsDNA. The solid lines in Figure **5.4** correspond to numerically calculated values using the model from van Dorp and co-workers: these can be seen to capture the trends in the experimental data well. The fact that the dsRNA follows the same trend as the dsDNA also serves as an additional validation of the theoretical assumptions underlying these calculations. However, exact quantitative agreement between model and data is not expected, given that the mean-field theory underlying this model becomes less reliable at high ionic strength. Particularly for experiments conducted in very small nanopores, we start to see clear deviations of the model from our data, where the measured forces lie below the predicted values. This is not surprising, as for such small nanopores finite ion sizes will start to become important, which is not accounted for in the model.

To directly compare the forces on dsRNA and dsDNA we compute the ratio f_{dsRNA}/f_{dsDNA} for all measurements in which both dsRNA and dsDNA were probed in the same nanopore (Figure **5.4**, inset). For this ratio, numerical simulations (indicated by the dashed line) predict a value just above unity over the entire range of pore sizes. On average, we indeed find the force on dsRNA molecules to be very similar to that on dsDNA, with the ratio between the two values falling slightly below unity (data points in Figure **5.4**, inset; f_{dsRNA}/f_{dsDNA} = 0.95 +/- 0.18, where the error indicates the standard deviation). This experimentally established similarity between the forces on the two molecules could be convenient for future experiments involving RNA. For instance, this would allow the use of hybrid DNA-RNA constructs in which the DNA could serve as a stiff double-stranded handle for a short RNA.

The slight discrepancy between the experimental data and the numerical model may reflect an inherent oversimplification of assumptions in the model compared to the experimental situation. For instance, the model approximates the molecules as cylinders with a uniformly surface charge. The highly localized nature of the charge groups on both dsRNA and dsDNA, however, may lead to the localization of at least a fraction of the counterions. As the dsRNA helix is significantly different in structure from that of dsDNA (the RNA bases being strongly tilted in the helix), the contribution of this effect to the net force may be quite different for the two molecules. For the same reason, the electro-osmotic drag force from the counterions may well be more significant for the dsRNA helix than for dsDNA, leading to a larger reduction of the purely electrical force acting on the phosphodiester backbone. As such structural differences are not captured in the model, it would tend to overestimate the force on dsRNA. The lower values for

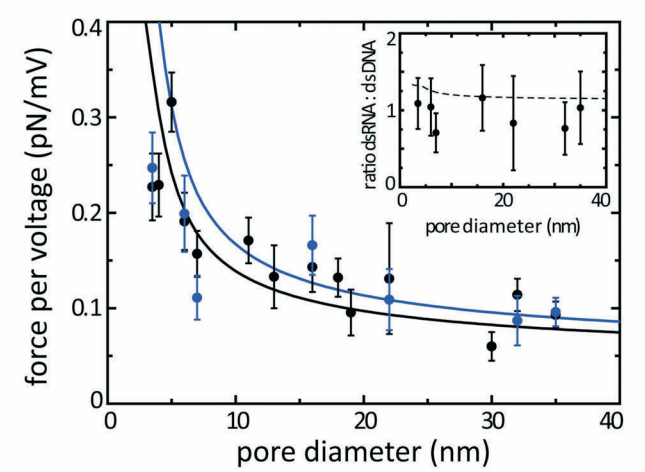


Figure 5.4. *Average force per voltage <f> on dsRNA (blue) and dsDNA (black) versus nanopore diameter*. Points indicate the experimentally measured data, while the solid lines are numerical calculations[11]. The inset shows the ratio $\langle f_{dsRNA} \rangle / \langle f_{dsDNA} \rangle$, where the black dotted line is the ratio expected from the numerical calculations. In all cases the error bars derive from the uncertainty in the determination of the trap strength (typically around 15-20%, see Materials and Methods).

the forces on dsRNA measured in the experiments can be viewed in this light.

5.4 Conclusions

With the measurements presented here we have expanded the force measurements in solid-state nanopores to include individual dsRNA molecules. The dsRNA molecules are quite long given their required synthesis via in vitro transcription, but are fourfold shorter than dsDNA molecules previously studied. We demonstrate that our versatile instrument integrating optical tweezers and nanopores can reliably measure forces for these shorter molecules, even in very small nanopores down to 3.5 nm in diameter. For dsRNA we find the force per applied voltage to be on the order of, but slightly lower than, the force per applied voltage on dsDNA, convenient for experiments involving possible hybrid dsDNA-RNA constructs. Furthermore, by demonstrating that force measurements are still reliable at very close distances to the nanopore membrane, these experiments open up the way for the use of even shorter molecular constructs, the biochemical synthesis of which is typically much easier. As the interest in RNA continues to grow

this technique could provide valuable information on the structure of RNA molecules and their interactions with proteins.

5.5 Acknowledgments

We thank S. W. Kowalczyk, A.R. Hall, M.Y. Wu, and V. Svechnikov for assistance in nanopore fabrication, A.J. te Velthuis and S. Donkers for assistance with RNA preparation, and U.F. Keyser, J. van der Does, and J. Beekman for advice on instrument construction. We also wish to thank S. van Dorp and S. Lemay for valuable discussions. NHD acknowledges financial support from the TU Delft, the Netherlands Organisation for Scientific Research (NWO), Nanoned, and from the European Science Foundation via an EURYI award.

Bibliography

1. Fologea, D., et al. (2007) Electrical characterization of protein molecules by a solid-state nanopore. *Appl. Phys. Lett*. **91**: 053901.1-053901.3.
2. Smeets, R.M.M., et al. (2009) Translocation of RecA-Coated Double-Stranded DNA through Solid-State Nanopores. *Nano Lett.* **9**: 3089-3095.
3. Wiggin, M., et al. (2008) Nonexponential Kinetics of DNA Escape from α-Hemolysin Nanopores. *Biophys. J*. **95**: 5317-5323.
4. Skinner, G.M., M. van den Hout, et al. (2009) Distinguishing Single- and Double-Stranded Nucleic Acid Molecules Using Solid-State Nanopores. *Nano Lett*. **9**: 2953-2960.
5. Soni, G.V. and A. Meller (2007) Progress toward ultrafast DNA Sequencing using solid-state nanopores. *Clin. Chem*. **53**: 1996-2001.
6. Branton, D., et al. (2008) The potential and challenges of nanopore sequencing. *Nat. Biotech*. **26**: 1146-1153.
7. Clarke, J., et al. (2009) Continuous base identification for single-molecule nanopore DNA sequencing. *Nat. Nano.* **4**: 265-270.
8. Keyser, U.F., et al. (2006) Direct force measurements on DNA in a solid-state nanopore. *Nat. Phys*. **2**: 473-477.
9. Keyser, U.F., et al. (2006) Optical tweezers for force measurements on DNA in nanopores. *Rev. Sci. Instr*. **77**: 105105.1-105105.9.
10. Smeets, R.M.M., et al. (2006) Salt dependence of ion transport and DNA translocation through solid-state nanopores. *Nano Lett*. **6**: 89-95.
11. van Dorp, S., et al., (2009) Origin of the electrophoretic force on DNA in solid-state nanopores. *Nat. Phys*. **5**: 347-351.
12. Gerland, U., R. Bundschuh, and T. Hwa (2004) Translocation of structured polynucleotides through nanopores. *Phys. Biol*. **1**: 19-26.

13. Vocks, H., D. Panja, and G.T. Barkema (2009) Amplitude and frequency spectra of thermal fluctuations of a translocating RNA molecule. *J. Phys.: Condens. Matter*. **21**: 375105.1-375105.4.
14. Akeson, M., et al. (1999) Microsecond time-scale discrimination among polycytidylic acid, polyadenylic acid, and polyuridylic acid as homopolymers or as segments within single RNA molecules. *Biophys. J*. **77**: 3227-3233.
15. Kasianowicz, J.J., et al. (1996) Characterization of individual polynucleotide molecules using a membrane channel. *Proc. Nat. Ac. Sci. USA*. **93**: 13770-13773.
16. Fire, A., et al. (1998) Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature*. **391**: 806-811.
17. Abels, J.A., et al. (2005) Single-molecule measurements of the persistence length of double-stranded RNA. *Biophys. J*. **88**: 2737-2744.
18. Hagerman, P.J. (1997) Flexibility of RNA. *Ann. Rev. Biophys. Biomol. Struct*. **26**: 139-156.
19. Ke, A. and J.A. Doudna (2004) Crystallization of RNA and RNA-protein Complexes. *Methods.* **34**: 408-414.
20. Keyser, U.F., et al. (2009) Inserting and Manipulating DNA in a Nanopore with Optical Tweezers. *Micro and Nano Technologies in Bioanalysis: Methods and Protocols*. 95-112.
21. Keyser, U.F., et al. (2005) Nanopore Tomography of a Laser Focus. *Nano Lett*. **5**: 2253-2256.
22. van den Hout, M., et al. (2008) End-joining long nucleic acid polymers. *Nucleic Acids Res*. **36**: e104.
23. van Dinten, L.C., et al. (1997) An infectious arterivirus cDNA clone: Identification of a replicase point mutation that abolishes discontinuous mRNA transcription. *Proc. Nat. Ac. Sci. USA*. **94**: 991-996.
24. Liphardt, J., et al. (2001) Reversible unfolding of single RNA molecules by mechanical force. *Science*. **292**: 733-737.
25. Friedrich, R. and G. Feix (1972) RNA-RNA hybridization in aqueous solutions containing formamide. *Anal. Biochem*. **50**: 467-476.
26. Berg-Sorensen, K. and H. Flyvbjerg (2004) Power spectrum analysis for optical tweezers. *Rev. Sci. Instr.* **75**: 594-612.
27. Young, M.A., G. Ravishanker, and D.L. Beveridge (1997) A 5-nanosecond molecular dynamics trajectory for B-DNA: analysis of structure, motions, and solvation. *Biophys. J*. **73**: 2313-2336.
28. Liu, L., et al. (2008) Structural Basis of Toll-Like Receptor 3 Signaling with Double-Stranded RNA. *Science*. **320**: 379-381.
29. Keyser, U.F., S. van Dorp, and S.G. Lemay (2010) Tether Forces in DNA Electrophoresis. *Chem. Soc. Rev.* **39**: 939-947.
30. Hall, A.R., et al. (2009) Electrophoretic Force on a Protein-Coated DNA Molecule in a Solid-State Nanopore. *Nano Lett.* **9**:4441-4445.

31. Ghosal, S. (2007) Effect of salt concentration on the electrophoretic speed of a polyelectrolyte through a nanopore. *Phys. Rev. Lett.* **98**: 238104.1-238104.4.
32. C. Gutsche, et al. (2006) The flow resistance of single DNA-grafted colloids as measured by optical tweezers. *Microfluid. NanoFluid.* **2**: 381-386

Appendix 5

A5.1 Position detection

Here we describe the bead position detection and its translation into force measurement. Additional details regarding the experimental configuration can be found elsewhere[8-9,20]. Position detection of the bead in the optical tweezers is facilitated by monitoring the reflection signal of a separate detection laser incident on the bead. The reflected light is directed onto a quadrant photo detector (QPD), which following amplification converts the total incident light into a voltage signal. The resulting signal is related to the position *z* of the bead with respect to the nanopore membrane (where *z* is defined in Figure **A5.1a**). Therefore, for each bead that is captured in the optical trap a calibration curve is recorded: using a piezo-controlled stage, the bead is moved slowly away from the membrane, while the QPD signal is recorded. As the piezo position is recorded simultaneously with nanometer precision, a calibration curve between the piezo position and the QPD signal is established. An example of such a curve is given in Figure **A5.1b**. The detailed shape of this curve is unique for each bead, and is determined by interference of the detection laser light between the membrane surface and bead surface. Since a single value of the QPD signal can correspond to multiple positions in *z*, camera detection is additionally employed to establish the absolute distance between the bead and the nanopore membrane[9].

Immediately after a calibration curve is recorded, the bead is moved in close proximity of the nanopore membrane. Using the just recorded calibration curve and the position information from the piezo controller and camera detection this initial position can be determined very accurately. A molecule attached to the bead can then be captured into the nanopore by application of a bias voltage over the nanopore. Upon capture, the bead will move, and the QPD signal will change accordingly, following the same path as the calibration curve. After capture, the voltage is varied, while still recording the QPD signal. In Figure **A5.1c** the voltage is stepwise increased from 10 mV to 110 mV, thereby increasing the net forces on the molecule inside the pore, which causes movement of the bead towards the nanopore membrane. This movement is reflected by the change in the QPD signal, which changes from a starting value of approximately 2.25 V (at V_{bias} = 10 mV) to a final value of 1.7 V (at V_{bias} = 110 mV). As the initial position of the bead prior to capture of the molecule in the nanopore is known, the subsequent bead positions at different bias voltages can be determined for each voltage by following the changing QPD signal along the calibration curve. This is indicated by the different symbols in Figures **A5.2b** and **A5.2c**. In this particular case, the bead is moved by a total distance Δz of approximately 130 nm when the bias voltage is changed from 10 mV to 110 mV. This distance can be converted to a force using the stiffness of

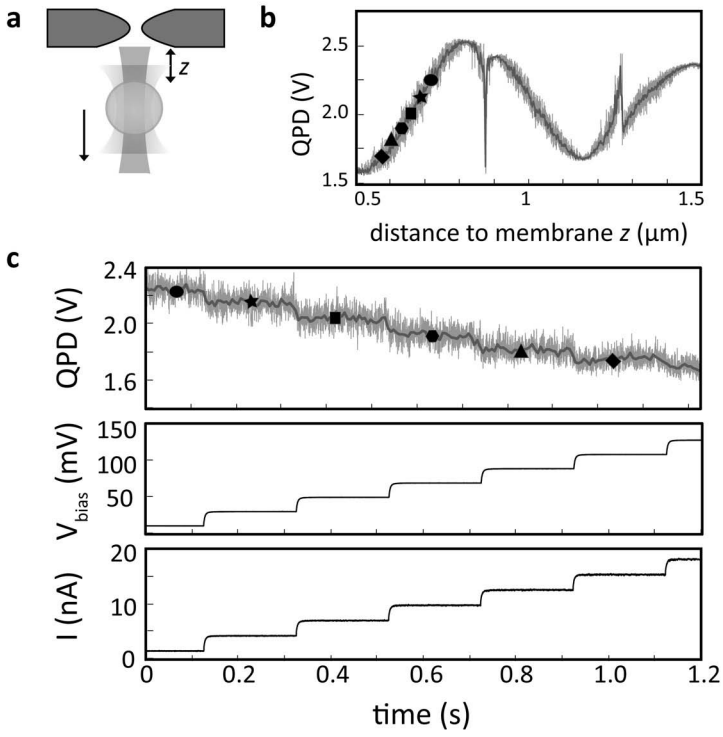


Figure A5.1. *Position detection*. **a**. Schematic defining the position *z* as the distance between the nanopore membrane surface and the bead surface. Prior to a measurement, the bead is moved away from the nanopore membrane surface while the position signal is recorded. **b**. Recorded QPD signal as a function of the position *z*. This curve is used as a calibration file relating QPD signal to position. **c**. After a molecule is captured in the pore, the voltage can be varied, leading to a commensurate change in bead position: the QPD signal (top) changes together with the bias voltage (middle), and the nanopore current (below). Using the calibration curve in **b** recorded previously, the QPD signal can be related to a position.

the optical trap k_{trap} as determined from a Lorentzian fit to the power spectrum of the bead position, using the relation $F = k_{trap} \bullet \Delta z$.

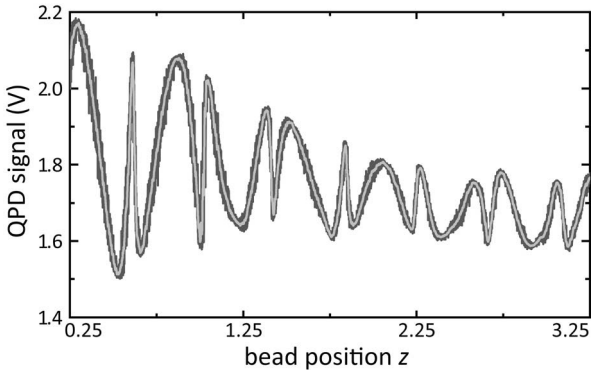


Figure A5.2: *Calibration curve over a distance of 3 µm*. The calibration curve, shown here, starts from a distance *z* of approximately 250 nm away from the nanopore membrane surface. Clearly, the sensitivity of QPD signal is higher when the bead is closer to the surface: at *z* = 0.25 µm, the QPD response is approximately fourfold larger than at *z* = 3.25 µm. This increases the signal-to-noise in determining the bead position at these small distances from the membrane, which is advantageous when working with short molecules.

A5.2 Determination of the position of the nanopore membrane

The exact position of the nanopore membrane can only be determined by physically moving the bead against the nanopore membrane, which dramatically changes the fluctuations in the QPD signal. Unfortunately, this can lead to aspecific attachment of the bead to the surface and can permanently change the conduction and current noise of the nanopore. However, we have found that the QPD noise signal increases about an order of magnitude when the bead is within approximately 100 ± 30 nm (error being the observed variation in where this crossover occurs over many different nanopores) from the membrane surface. This clear change in the QPD signal can be used to estimate the distance to the membrane with a precision of approximately 30 nm.

A5.3 Signal-to-noise in determination of position

Figure **A5.2** explains how the signal-to-noise in determining the bead position improves as the bead approaches the surface.

Chapter 6

Controlling Nanopore Size, Shape and Stability

Abstract

Solid-state nanopores are considered a promising tool for the study of biological polymers such as DNA and RNA, due largely to their flexibility in size, potential in device integration, and robustness. In this chapter, we show that the precise shape of small nanopores (~5 nm diameter in 20 nm SiN membranes) can be controlled by using transmission electron microscope (TEM) beams of different sizes. However, when some of these small nanopores are immersed in aqueous solution, their resistance is observed to decrease over time. By comparing nanopores of different shapes using (scanning) TEM both before and after immersion in aqueous solution, we demonstrate that the stability of small nanopores is related to their three-dimensional geometry, which depends on the TEM beam size employed during pore fabrication. Optimal stability is obtained using a TEM beam size of approximately the same size as the intended nanopore diameter. In addition, we show that thermal oxidation can serve as a means to independently control nanopore size following TEM fabrication. These observations provide key guidelines for the fabrication of stable solid-state nanopores on the scale of nucleic acids and small proteins.

This chapter was published as: Michiel van den Hout, Adam R. Hall, Meng-Yue Wu, Henny W. Zandbergen, Cees Dekker, and Nynke H. Dekker, *Nanotechnology*. **21**: 115304.1-115304.5 (2010)

6.1 Introduction

Nanopores are widely used for the detection and study of nucleic acids and other biopolymers[1-3]. In an electrolyte, such biomolecules can cause a characteristic change in the nanopore ionic conductance as they are driven through the pore by an electrical field. Although a variety of different biological and synthetic nanopores have been used for such studies[1-3], solid-state nanopores, which are typically fabricated in thin membranes of SiN or SiO_2, are the most obvious candidates for device integration and offer the largest flexibility in nanopore size and experimental conditions. A number of methods exist for the fabrication of solid-state nanopores[4-7], but the method of choice is to drill them using a transmission electron microscope (TEM) beam[6-7], as in this case the nanopore can be directly visualized during the drilling process. In this technique, nanopores are drilled by focusing a high-energy electron beam (typically 200-300 keV) on a thin membrane, resulting in the creation of a single pore of several nanometers in diameter. Subsequently, the nanopore can then be enlarged by continued exposure to the beam, or it can be shrunk by reducing the beam intensity[6,7], allowing for very precise specification of the pore size (which is not possible using other techniques, such as focused ion beam sculpting). Such control is increasingly important as nanopores are used to probe complex biological systems at the molecular scale, examples of which include the unfolding of DNA hairpins[8], DNA-protein interactions[9] and, potentially, the study of RNA secondary structure[10-11], and DNA sequencing[12].

Key to these applications is not only the nanopore size, but also the corresponding three-dimensional nanopore shape. For example, optimal resolution for DNA sequencing would require a very thin nanopore constriction with a local thickness on the sub-nanometer scale, so that the changes in the current blockade would reflect the presence of individual bases along the DNA. This can potentially be realized by fabricating a pore with a heavily compressed hourglass shape (i.e. a shape described by two mirrored cones with very wide bases). Recent tomography studies on solid-state nanopores have indicated that the shape of a nanopore depends on the material composition and precise conditions employed during the drilling, and can therefore be controlled to a certain extent[13-14]. Here, we demonstrate that the TEM beam size itself provides a straightforward tool to control the shape of ~5 nm diameter nanopores drilled in 20 nm SiN membranes. Unexpectedly, we also find a correlation between nanopore shape and nanopore stability: a TEM beam that has a large width compared to the intended nanopore diameter (30 nm full width at half maximum (FWHM) compared to a 5 nm diameter pore) produces nanopores with a compressed hourglass profile (Figure **6.1a**), but these pores are frequently observed to rapidly increase in size over time upon immersion in aqueous solutions; conversely, nanopores drilled with a narrow beam size are more cylindrical in shape, and their size is found to increase to a lesser extent, if at all. Our results indicate that for optimal size stability the TEM beam size

should be approximately of the same size as the intended nanopore diameter. We also demonstrate control of the nanopore size following TEM drilling, showing that the size of these SiN nanopores can be gradually reduced by thermal oxidation at modest temperatures (250 ºC), increasing their versatility even more.

6.2 Materials and methods

Nanopore fabrication and imaging

Nanopores were fabricated in 20 nm SiN membranes as previously described[15] using a Philips CM300 UT TEM, operated at an accelerating voltage of 300 kV. Pores were drilled using 3 different beam sizes (30 nm, 15 nm, and 7.5 nm), where this beam size corresponds to the full width at half maximum (FWHM) of the Gaussian intensity profile in the transverse focal plane. Scanning TEM (STEM) measurements were performed with a Cs corrected Titan Cubed Supertwin/STEM FP5600/40 instrument operated at an accelerating voltage of 300 kV and a beam size of 0.14 nm. The images were acquired with convergent angle of about 10 mrad, and a camera length of 184 mm.

Ionic current measurements

For the electrical measurements, a microfabricated chip containing a single or multiple nanopores was first rinsed with acetone and ethanol, blown dry, and exposed to O_2 plasma for 30 s. The chip was then mounted in a custom-built Teflon flowcell[16], and a solution containing 1M KCl, 10 mM Tris-HCl (pH = 8.0) and 1 mM EDTA was introduced on both sides of the nanopore. The *cis* and *trans* sides of the nanopore are connected to separate reservoirs containing Ag/AgCl electrodes that allow for the connection of the nanopore device to an Axon Instruments Axopatch 200B amplifier. Nanopore resistance was determined by measuring the pore current as a function of applied bias voltage and applying a linear fit to the slope of the resulting IV curve using custom-written Labview software.

Thermal oxidation

For the experiments on nanopore oxidation, membranes containing nanopores were thermally oxidized in a Lindberg Blue Mini/Mite TF55030C-1 tube oven, operated at 250 ºC for 1 hr with a constant O_2 flow of 10 sccm. Both before and after the oxidation, the membranes were rinsed with acetone and IPA and blown dry.

Image analysis

The areas of the nanopores were computed from TEM images using ImageJ[17]. Radial intensity profiles from the STEM images were analyzed by taking a 360º polar transformation (using the Polar Transformer plugin for ImageJ, written by F. Donnely) of the STEM image around the nanopore center and averaging line scan profiles from the center of the pore outwards. The corresponding membrane thickness was scaled such that the minimum intensity was set to 0 (corresponding to 0 membrane thickness) and the maximum intensity was set to 20 nm (corresponding to the full membrane thickness of 20 nm). A 3D image was reconstructed from these profiles by assuming cylindrical symmetry about the pore center[14]. In addition, the non-zero noise in the area at the center of the nanopore (shaded green area in Figure **6.1b**) was removed, as TEM images indicate that there is no material there.

6.3 Results and discussion

We first briefly explain how our nanopores are formed by the electron beam. As described in the Materials and Methods, we use a focused electron beam to fabricate our nanopores in 20 nm thick SiN membranes. The electrons locally ablate the material, leading to the formation of a small pore. In Figure **6.1a** we sketch the intensity profiles for two different beam sizes and the corresponding nanopore shapes that these beams are expected to form in the SiN, based on the following reasoning: for the beam with the larger waist, a large area is exposed to a high intensity, leading to appreciable material depletion in the area around the nanopore that is formed. This will lead to the formation of a nanopore with a compressed hourglass shape (Figure **6.1a**, solid black lines). Conversely, a beam with a smaller waist will expose only a small area around the nanopore to high intensity, giving rise to straighter sidewalls (Figure **6.1a**, dashed black lines) and a pore that is more cylindrical in shape. Thus, we expect that by choosing the correct TEM beam size, the nanopore shape should be controllable to some extent, which would be beneficial for many applications. For instance, a compressed hourglass shape would be particularly useful for the probing of local structure along a molecule, as the ionic resistance through the pore is dominated by the narrowest constriction; a change in the ionic conductance caused by local differences along the molecule can therefore be monitored with a resolution of the length of that constriction.

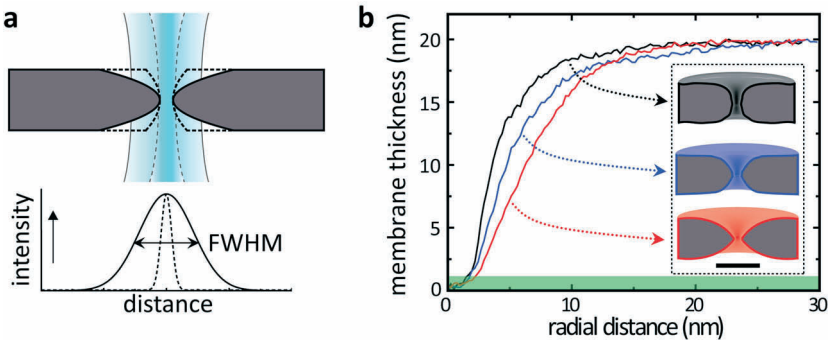


Figure 6.1. *TEM beam size influences nanopore shape*. **a**. Sketch of the expected shapes of nanopores drilled with different TEM beam sizes, whose approximate 2D intensity profiles are sketched in blue. The corresponding 1D intensity profiles of the two beams are plotted below. The solid grey line delineates a TEM beam much larger than the intended pore size, whereas the dotted grey line marks a beam of approximately the same size as the targeted nanopore. The nanopore drilled with the wider beam is expected to have a compressed hourglass shape (dark grey area with solid black perimeter), whereas the nanopore for the narrower beam is expected to be more cylindrical in shape (open area with a dashed black perimeter). **b**. Thickness profiles, derived from scanning transmission electron microscopy (STEM) images (Materials and Methods) of nanopores drilled with the three beam sizes (30 nm, red line; 15 nm, blue line; and 7.5 nm, black line). (insets) The thickness profile can be used to reconstruct a 3D-image, assuming the nanopore is point symmetric about the centre of the pore and assuming a uniform material stoichiometry. In addition, the non-zero noise in the green shaded area was removed to create the 3D-images, as the TEM images indicate there is no material there. The black scale bar for the insets is 20 nm. As the TEM beam size is increased, the nanopore shape changes from roughly cylindrical towards a compressed hourglass shape.

To examine whether the shape of our nanopores can indeed be controlled by using TEM beams of different sizes, we have used scanning transmission electron microscopy (STEM) to measure the shape of our nanopores. In this technique, a very small (0.14 nm FWHM) focused electron beam is scanned over the surface to create a 2D image. From the STEM image, a thickness profile can be derived, assuming a uniform material composition in the membrane surrounding the nanopore. A 3D shape can then be reconstructed from the STEM thickness profile by further assuming that the nanopore is point-symmetrical from its centre (Materials and Methods). Previous work has shown that the shape determined in this way corresponds well to the actual 3D structure, as determined both by TEM tomography and a number of additional analysis techniques (e.g. EELS, EFTEM)[14]. In Figure **6.1b**, we show the radial thickness profile as determined from 2D STEM images of three different nanopores of ~5 nm in diameter. Each nanopore was

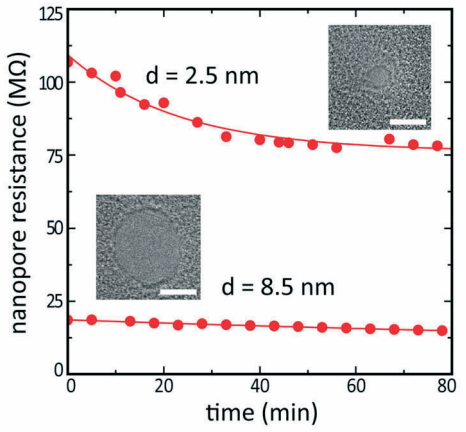


Figure 6.2. *Ionic resistance of nanopores decreases over time*. Ionic resistance (at 1M KCl, 10 mM Tris-HCl, 1 mM EDTA, pH = 8.0) as a function of time of two nanopores (2.5 nm and 8.5 nm, see insets), both drilled with a TEM beam size of size 30 nm (FWHM). The solid lines are added as guides to the eye. The scale bars in the insets correspond to 5 nm. The ionic resistance of the small nanopore is observed to decrease strongly over the course of 1 hr, whereas that of the large nanopore decreases to a lesser extent.

drilled with a different beam size: 30 nm FWHM (red line), 15 nm FWHM (blue) and 7.5 nm FWHM (black line). The insets show the reconstituted 3D shapes of these pores. Indeed, the shape of the nanopore drilled with the 30 nm beam (red) displays the highest curvature at the center. The shape of the nanopore drilled with the smaller 15 nm beam is less curved at the center, whereas the nanopore drilled with the smallest beam is shown to have nearly vertical sidewalls. This demonstrates that the nanopore shape can indeed be controlled by choosing the correct TEM beam size.

We have tested our nanopores under conditions typically employed for biological experiments, in which the nanopore is immersed in an electrolyte and a constant bias voltage is applied. Surprisingly, we found that these small nanopores are frequently observed to be unstable: the ionic resistance through the nanopore continuously decreases over time during the course of an experiment, rendering such pores less useful for quantitative analysis. Examples of this behaviour for two nanopores with initial diameters of 2.5 nm and 8.5 nm, both drilled with a TEM beam size of 30 nm FWHM, are shown in Figure **6.2**. Here, we plot the time evolution of the ionic resistance through these pores over the course of 80 min, while they where held under a constant bias voltage of 100 mV. For both pores, we observe that the resistance decreases over time. Note, however, that the effect is

more pronounced for the smaller nanopore, for which the resistance drops by 26% from 107 MΩ to 79 MΩ in 80 min, compared to a 15% drop from 18.5 MΩ to 15.8 MΩ for the larger pore. We have found that such large changes in resistance occur most frequently for small (<5 nm) nanopores, and within this size range, for those pores drilled with wide TEM beam sizes (15 nm or wider). Given the fact that the TEM beam size influences the nanopore shape, this result implies that perhaps their shape is related to the instability of these nanopores.

As nanopore stability is key to the ability to perform reproducible experiments on molecule translocation, we have set out to account for the observed drop in resistance over time by examining the influence of a number of experimental parameters. First, to verify that this behaviour could be not accounted for by evaporation of water from the electrolyte, we introduced fresh measurement buffer during the course of the experiment. The measurement of identical resistances before and after this operation indicated that evaporation could not underlie the observed behaviour. Replacement of the Ag/AgCl electrodes likewise did not affect the measured resistance, excluding possible electro-chemical depletion of the electrodes. A further potential explanation for the observed changes in nanopore resistance is that the nanopores themselves actually change in size over the course of a measurement in ionic solution. Indeed, even a small absolute change in diameter could have a strong effect on the resistance of a small nanopore. The effect on larger nanopores would be less pronounced, as the ionic resistance is inversely proportional to the square of the nanopore diameter[18]. Nonetheless, we note that such a size change over time is a somewhat surprising possibility, as the nanopores are drilled into SiN solid-state membranes, which are not expected to etch in aqueous solutions such as those used in our experiments[19].

To test whether our nanopores indeed change in size, we have compared several nanopores, drilled in a single membrane, both before and after a 1 hour immersion in measurement buffer with at an applied bias of 100 mV. Because we suspected that the size of the TEM beam used to drill the pores might be related to their stability, the sample contained 3 different pores with the same initial diameter (5 nm, see the left column in Figure **6.3b**), but each drilled with a different beam size (30 nm, 15 nm, and 7.5 nm FWHM). The resistance of this membrane containing 3 nanopores was then measured as a function of time (Figure **6.3a**). As before, the total resistance of the membrane gradually decreased, in this case from 18 MΩ to about 9 MΩ over the course of 1 hour. Images of the nanopores were taken both before (Figure **6.3b**, left column) and after (Figure **6.3b**, right column) the measurement, taking care to employ a very low beam intensity to ensure that the nanopores were not changed during the imaging process. Clearly, the nanopore drilled with the large 30 nm beam (Figure **6.3b**, top row) increases dramatically in diameter. Similarly, the nanopore drilled with the 15 nm beam also increases in size, though less strongly (Figure **6.3b**, middle row). Meanwhile, the nanopore drilled with the smallest beam studied (7.5 nm) shows very little change in diameter (Figure **6.3b**, bottom row). This striking result implies that nanopores drilled with a wide TEM beam are inherently less stable upon

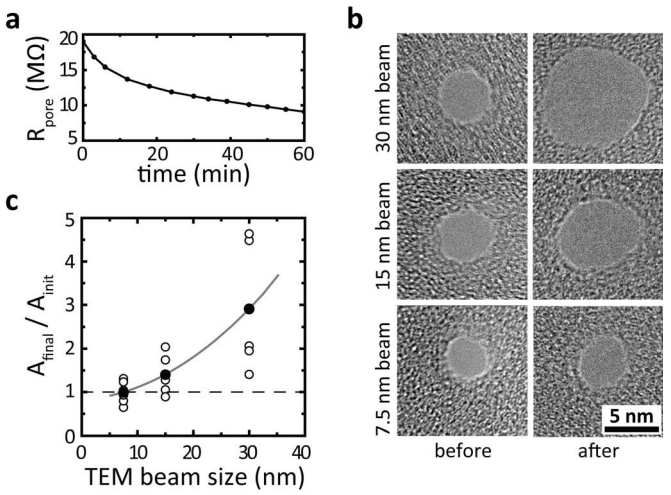


Figure 6.3. *Nanopores drilled with wide TEM beams increase in size*. **a**. Resistance versus time (at 1M KCl, 10 mM Tris-HCl, 1 mM EDTA, pH = 8.0) of a membrane containing multiple nanopores of approximately 5 nm in initial diameter, but drilled with different beam sizes. A constant bias voltage of 100 mV is applied over the nanopores during the course of the entire measurement. The resistance is observed to decrease continuously with time. **b**. TEM images of the nanopores in this membrane imaged both before (left column) and after (right column) placement in the ionic solution. The nanopore drilled with the widest beam (30 nm FWHM, top row) has become significantly larger, but this behaviour is less pronounced for the nanopores drilled with smaller beams (15 nm FWHM, middle row and 7.5 nm FWHM, bottom row). The scale bar is 5 nm. **c**. Graph of the resulting change in nanopore (cross-sectional) area, defined as the final nanopore area divided by the initial area, as a function of the TEM beam size for different repeats of this experiment. Open circles correspond to individual measurements and closed circles indicate the average values. The solid line is added as a guide to the eye. The trend clearly shows that nanopores drilled with the largest beam size expand in size following immersion in an aqueous solution for 1 hr. The nanopores drilled with the smallest beam size (approximately equal to the initial nanopore size of 5 nm) do not, on average, increase in size.

immersion in an electrolyte than nanopores drilled with a narrower beam. Also, it appears that the observed change in resistance can indeed be explained by the fact that some of the nanopores changed size.

To see whether this effect of the electron beam size on the nanopore stability was reproducible, we have repeated this experiment multiple times. In each case we used the same three different beam sizes as above to produce nanopores with initial sizes ranging from 4 to 7 nm, as determined by TEM microscopy. The results are shown in Figure **6.3c**, where we plot the ratio of the nanopore areas before and

after the experiment, A_{final}/A_{init}. The dashed line at A_{final}/A_{init} = 1 is added as a reference and corresponds to no change in the area of the nanopore. Indeed, the trend of the average values confirms our initial result: pores formed with a large diameter electron beam increase in size much more than those formed with a small beam size. The scatter in the data, especially for the large beam size, indicates that the exact size change for a given individual nanopore cannot be predicted. In some instances pores were also found to decrease in size, rather than stay constant or increase in size, and possible reasons for the observed variability will be discussed below. Yet the averages indicate a clear trend in which pore size stability decreases with the TEM beam size.

It is likely that the clear difference in shape for nanopores drilled with different TEM beam sizes is related to the observed differences in stability. Our nanopores are drilled in amorphous SiN, and it is possible that the material at the edge of the nanopores has a tendency to rearrange over time into a shape that minimizes surface tension. Sharp edges with high curvature are therefore energetically unfavorable and thus less stable. In this scenario, rearrangement is predicted to occur even in the absence of an ionic current flowing through the nanopore, as it results from energy minimization of the SiN only. Indeed, we have found that our nanopores also increase in size with time when stored in de-ionized water only, although the rate of change appears to be decreased (data not shown). Furthermore, it is also possible that the observed decrease in size for some of the nanopores drilled with the small electron beam in Figure **6.3c** results from similar rearrangement: A nanopore with a highly cylindrical shape will also have sharp edges at the cylinder ends, and rearrangement of these could produce the apparent decrease in diameter observed in the cross-sectional TEM image. Indeed, an additional experiment using even smaller beam sizes (2 nm to 5 nm) showed a decrease in area for the pores drilled with smaller beams, whereas the pores drilled with larger beam sizes again showed a size increase (Appendix **6**, Figure **A6.1**). Thus from our two observations, namely that 1) nanopores drilled with an electron beam much wider than the nanopore size have a compressed hourglass shape, and 2) such nanopores are more likely to increase in pore diameter, we can extract an important design consideration for the fabrication of nanopores: for optimal stability, a beam size of approximately the intended nanopore size is recommended.

The increasingly sophisticated control over nanopore fabrication desired for a variety of experiments can also be expanded by the ability to modify the size and surface properties of a nanopore following drilling. Here, we employ thermal oxidation as a simple tool to change the diameter of a nanopore after drilling. Upon oxidation of SiN, nitrogen atoms are displaced by oxygen atoms, as formation of SiO_2 is energetically favoured over SiN[20]. Since the packing density of SiO_2 is lower than that of SiN, this leads to a small increase in the thickness of the material, and thus to a small decrease in the nanopore diameter, assuming a uniform material stoichiometry around the nanopore (recent results indicate that, in fact, the nitrogen content decreases close to the centre of the pore[14], which would further

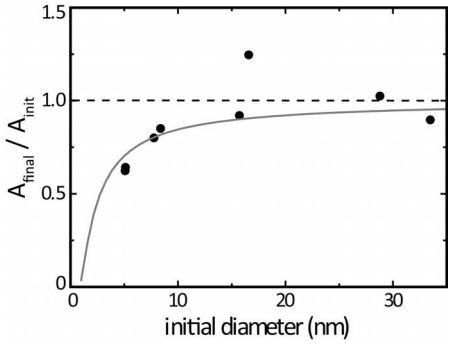


Figure 6.4. *Nanopore size can be changed by thermal oxidation*. Nanopores of different sizes were thermally oxidized at a temperature of 250 ºC for 1 hr. The ratio of the final nanopore area to the initial nanopore area is plotted as a function of the initial nanopore diameter. Almost all pores reduce in size, as expected from the fact that the oxidation of SiN should uniformly increase the membrane thickness. The solid line is a fit to the data, assuming a uniform thickness increase of the material in all directions. The fit yields a growth of 0.8 ± 0.2 nm, corresponding to an average decrease in nanopore diameter of 1.6 ± 0.4 nm.

enhance the effect of the oxidation, as pure Si is more rapidly oxidized than SiN). We have investigated whether this oxidation treatment would also improve the stability of our nanopores, but no improvement was observed. Nevertheless, apart from allowing for additional size control after fabrication, thermal oxidation can have several additional advantages. For example, changing the nanopore surface into SiO_2 allows for a large variety of surface chemistry reactions to further modify the nanopore surface[21]. Also, we have frequently observed that surface interactions with biological molecules are less prevalent for SiO_2 than for SiN[22]. Such a surface modification thus also serves to increase nanopore versatility.

Thermal oxidation of thin films of crystalline Si3N4 is reported to be very a slow process, with a (non-linear) rate averaging only several nm/min at temperatures > 900 ºC for such thin films as employed in our experiments[20]. This slow rate is useful, as it allows for very fine control over the thickness. To demonstrate that this can indeed be used to modify the size of our nanopores, we have drilled several nanopores of varying initial diameter in a single membrane (to ensure identical conditions) and thermally oxidized this membrane at 250 ºC for 1 hr (see Materials and Methods). Such a low temperature ensures that the rate is slow enough such that even the smallest nanopores do not completely close due to the treatment. The nanopores were drilled with two different beam sizes (30 nm and 7.5 nm), but this parameter did not appear to influence the results. In Figure **6.4** we plot the change in area of our nanopores, again expressed as the quantity A_{final}/A_{init}. As expected, the area of the nanopores generally decreases by the

treatment (Figure **6.4**, black points). Also, the area reduction is much more pronounced for smaller nanopores. This is not surprising, as the oxidation is expected to change the material thickness in all directions by a constant value. A 1 nm increase in thickness will therefore reduce the diameters of all pores by 2 nm. Such a change yields a much more dramatic area change for small pores than for large pores. Indeed, we can fit (solid line in Figure **6.4**) our results well with a simple relation: $A_{final} = A_{init} (r_{pore}-\delta)^2 / (r_{pore})^2$, where r_{pore} is the initial pore radius, to find an average thickness increase of $\delta = 0.8 \pm 0.2$ nm. Comparison of our observed growth rate with reported values[20] is not possible as these were measured under very different conditions (900-1100 ºC, in partial H2O atmosphere and on crystalline SiN, whereas our nanopores are drilled in amorphous SiN). The fact that over the duration of an hour the nanopores change only several nanometers in diameter demonstrates that thermal oxidation offers an additional sub-nanometer size control for nanopores, applicable following fabrication. We also note that this technique can easily be applied at a wafer scale, making it particularly useful for applications using large arrays of nanopores.

6.4 Conclusions

We have shown that the shape of small nanopores of ~5 nm in diameter is dependent on the TEM beam size used during their fabrication and that this influences their stability upon immersion in aqueous solution. A large TEM beam compared to the targeted nanopore size creates compressed hourglass shaped pores, which are found to increase in diameter over time during immersion in aqueous solutions. This effect is much reduced for nanopores drilled with a TEM beam size of approximately the same size as the nanopore diameter. In addition, we have shown that oxidation can further be used to modify the size of nanopores after drilling. Our results add to the understanding of nanopore formation and stability, and may serve as important design considerations for the sub-nanometer control of nanopore shape and size required for many of the more sophisticated applications involving solid-state nanopores.

6.5 Acknowledgments

We thank Vincent Krudde and Andrew Scott for help with electrical measurements. Financial support is acknowledged from the Netherlands Organisation for Scientific Research (NWO), Nanoned, the TU Delft, and from the European Science Foundation via an EURYI award.

Bibliography

1. Dekker, C. (2007) Solid-state nanopores. *Nat. Nanotechnol.* **2:** 209-215.
2. Healy, K. (2007) Nanopore-based single-molecule DNA analysis. *Nanomed. Nanotechnol. Biol. Med.* **2**: 459-481.
3. Rhee, M. and M.A. Burns (2006) Nanopore sequencing technology: research trends and applications. *Trends Biotechnol*. **24**: 580-586.
4. Chang, H., et al. (2006) Fabrication and characterization of solid-state nanopores using a field emission scanning electron microscope. *Appl. Phys. Lett.* **88**: 103109.1-103109.3.
5. Li, J., et al. (2001) Ion-beam sculpting at nanometre length scales. *Nature*. **412**: 166-169.
6. Storm, A.J., et al. (2003) Fabrication of solid-state nanopores with single-nanometre precision. *Nat. Mater.* **2**: 537-541.
7. Wu, M.Y., et al. (2005) Formation of nanopores in a SiN/SiO_2 membrane with an electron beam. *Appl. Phys. Lett.* **87**: 113106.1-113106.3.
8. McNally, B., M. Wanunu, and A. Meller (2008) Electromechanical Unzipping of Individual DNA Molecules Using Synthetic Sub-2 nm Pores. *Nano Let*. **8**: 3418-3422.
9. Wiggin, M., et al. (2008) Nonexponential Kinetics of DNA Escape from α-Hemolysin Nanopores. *Biophys. J.* **95**: 5317-5323.
10. Gerland, U., R. Bundschuh, and T. Hwa (2004) Translocation of structured polynucleotides through nanopores. *Phys. Biol.* **1**: 19-26.
11. Vocks, H., D. Panja, and G.T. Barkema (2009) Amplitude and frequency spectra of thermal fluctuations of a translocating RNA molecule. *J. Phys.: Condens. Matter*. **21**: 375105.1-37105.7.
12. Branton, D., et al. (2008) The potential and challenges of nanopore sequencing. *Nat. Biotech*. **26**: 1146-1153.
13. Kim, M.J., et al. (2007) Characteristics of solid-state nanometre pores fabricated using a transmission electron microscope. *Nanotechnology*. **18**: 205302.1-205302.5.
14. Wu, M.-Y., et al. (2008) Control of Shape and Material Composition of Solid-State Nanopores. *Nano Lett*. **9**: 479-484.
15. Keyser, U.F., et al. (2005) Nanopore Tomography of a Laser Focus. *Nano Lett*. **5**: 2253-2256.
16. Skinner, G.M., van den Hout, M. et al. (2009) Distinguishing Single- and Double-Stranded Nucleic Acid Molecules Using Solid-State Nanopores. *Nano Lett.* **9**: 2953-2960.
17. Rasband, W. *ImageJ*, available at http://rsb.info.nih.gov/ij; National Institute of Health, Bethesda, MD.
18. Smeets, R.M.M., et al. (2006) Salt dependence of ion transport and DNA translocation through solid-state nanopores. *Nano Lett.* **6**: 89-95.

19. Williams, K.R., K. Gupta, and M. Wasilik (2003) Etch Rates for Micromachining Processing - Part II. *J. Microelectromech. Syst.* **12**: 761-778.
20. Enomoto, T., et al. (1978) Thermal Oxidation Rate of a Si3N4 Film and its Masking Effect against Oxidation of Silicon. *Jpn. J. Appl. Phys.* **17**: 1049-1058.
21. Wanunu, M. and A. Meller (2007) Chemically modified solid-state nanopores. *Nano Lett.* **7**: 1580-1585.
22. Van den Hout, M., et al. (2010) Direct Force Measurements of Double-Stranded RNA in Solid-State Nanopores. *Nano Lett.* **10**: 701–707.

Appendix 6

Figure **A6.1** is supplied as supplement.

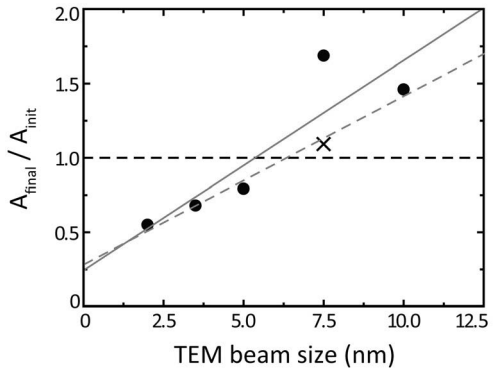


Figure A6.1: *Nanopore stability versus TEM beam size for small beam sizes*. Graph of the change in nanopore area as a function of the TEM beam size, after the pores were placed in ionic solution for 1 hr under an applied bias voltage of 100 mV. The change in area is defined as the final nanopore (cross-sectional) area divided by the initial area. In this experiment 5 pores of 5 nm (initial) diameter were drilled in a single membrane using TEM beam sizes of 2 nm, 3.5 nm, 5 nm, 7.5 nm and 10 nm FWHM. The red line is a linear fit to the data. The crossing point at zero area change (5.4 ± 0.1 nm) implies that the optimal beam size for stability is indeed similar to the pore size. The area of the pores drilled with beams smaller than the initial pore size is found to decrease. For comparison, we have also added the average value for the 7.5 nm beam size from Figure **6.3** (denoted by x).

Chapter 7

Distinguishable Populations Report on Interactions of Single DNA Molecules with Solid-State Nanopores

Abstract

Solid-state nanopores have received increasing interest over recent years for their potential for genomic screening and sequencing. Small nanopores of 2-5 nm in diameter, in particular, allow the detection of local structure along biological molecules, such as proteins bound to DNA or possibly the secondary structure of RNA molecules. In a typical experiment, individual molecules are translocated through a single nanopore, thereby causing a small deviation in the ionic conductance. Correct interpretation of these conductance changes is essential for our understanding of the process of translocation, and for further sophistication of this technique. In this chapter, we present new translocation measurements of double-stranded DNA through nanopores down to the diameter of the DNA itself (1.8 – 7 nm at their narrowest constriction). In contrast with previous findings on such small nanopores, we find that single molecules interacting with these pores can cause three distinct levels of conductance blockades. We attribute the smallest conductance blockades to molecules that briefly skim the nanopore entrance without translocating, the intermediate level of conductance blockade to regular head-to-tail translocations, and the largest conductance blockades to obstruction of the nanopore entrance by one or multiple (duplex) DNA strands. Our measurements are an important step towards understanding the conductance blockade of biomolecules in such small nanopores, which is essential for future applications involving solid-state nanopores.

This chapter was submitted as: M. van den Hout, Vincent Krudde, Xander J.A. Janssen, and Nynke H. Dekker (2010).

7.1 Introduction

The use of solid-state nanopores for the detection and study of biological molecules has expanded greatly in recent years[1-3]. For example, solid-state nanopores have now been used for the detection of locally bound proteins[4-6] or other molecules[7-9] along DNA, the unzipping of single stranded DNA hairpins[10], and the discrimination of single-stranded and double-stranded nucleic acids[11,12]. In all studies involving solid-state nanopores, charged biological molecules in an electrolyte are driven in or through the nanopore by an electrical field (Figure **7.1a**). Upon entry inside the nanopore, these molecules cause a characteristic change in the ionic conductance through the pore, which can serve as a means to detect local structure along the molecule. Although such translocation studies can also be performed with biological pores[1,13], the use of nanopores fabricated in solid-state membranes offers a number of advantages, such as the flexibility to tune their size[14,15], shape[16,17] and surface chemistry[18], and the fact that they could easily be integrated into electronic devices. For solid-state nanopores, the conductance blockade of translocating molecules is well-understood for pore sizes much larger than the molecular diameter: simple models, based on the total area or volume blocked by the polymer inside the nanopore, provide quantitative agreement with the observed blockades[19]. However, when the pore size approaches the size of the translocating molecule, new effects, such as the three-dimensional shape of the pore, its surface roughness[20], and the overlap of counterion screening layers need to be taken into account.

Here, we present translocation results of double-stranded DNA (henceforth simply referred to as DNA) through solid-state nanopores ranging from 1.8 nm to 7 nm in diameter. We characterize the DNA translocation events by their average conductance change $\Delta G = \Delta I/V$, as well as their dwell time t_D (Figure **7.1b** and Materials and Methods). ΔI is defined as the open pore current (I_{open}) minus the current in the blocked state ($I_{blocked}$), resulting in positive values for all experiments presented here. Following Wanunu *et al.*[21], we also define a relative current blockade $I_B = 1 - (\Delta I/I_{open}) = I_{blocked}/I_{open}$ as the current through the nanopore in the blocked state divided by the open pore current (Figure **7.1b**); For example, I_B equals 1 when no molecule blocks the nanopore, and I_B equals 0 in the fully blocked state.

In our results we identify three distinct event populations, each with a characteristic I_B and/or t_D. The first population consists of very short events with very small conductance blockades. Such events have also frequently been observed before in larger pores[22,23], and are attributed to DNA molecules that skim the entrance of the nanopore, but do not translocate through it. To explain how the other two populations can arise from a mono-disperse sample of DNA molecules, we have investigated how the relative weights of the populations change with pore size and DNA length. First, we show that in the smallest nanopores the majority of events result in a very large blockade, in which the conductance through the nanopore is nearly entirely blocked. When the nanopore size is increased, however,

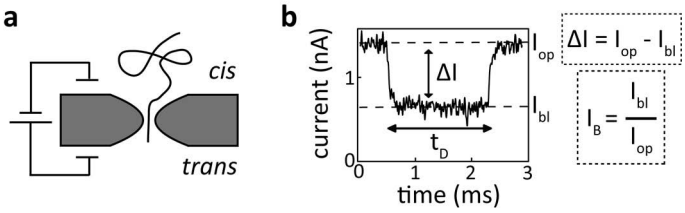


Figure 7.1. *Experimental setup*. **a**. Schematic of our setup. A nanopore immersed in electrolyte separates two compartments (*cis* and *trans*). After introduction of DNA on the *cis* side, an external bias voltage can thread DNA molecules through the nanopore. Upon traversing the nanopore, the molecule blocks a portion of the ionic current. **b**. Sample trace of the current blockade caused by a DNA molecule in a 3.2 ± 0.5 nm diameter pore at 100 mV bias voltage. For each translocation event we determine the average blockade $\Delta I = I_{op} - I_{bl}$ as the current in the open state (I_{op}) minus the current in the blocked state (I_{bl}), the dwell time t_D as the total duration of the event, and the relative current blockade $I_B = I_{bl} / I_{op} = 1 - \Delta I/I_{op}$.

the number of events with this large blockade gradually decreases, while more events appear with an intermediate blockade. We then show that in nanopores of 2.5 - 3 nm in diameter the population associated with the largest blockade increases in weight as the DNA length is increased. These observations indicate that this population is associated with events in which part of the DNA blocks the entrance of the nanopore, since such interactions would be more likely in smaller pores, and for longer DNA. A possible scenario for such an interaction would be the capture of a polymer in a transverse orientation at the nanopore entrance, after which the molecule may (through buckling or sliding one of its ends into the nanopore) or may not translocate.

We then focus on the population of events with intermediate blockades. Experimentally, the magnitude of this intermediate blockade corresponds very well to the conductance blockade observed in larger pores of about 10 nm in diameter[12] using longer DNA. Furthermore, we demonstrate that the magnitude of this conductance blockade is constant over a large range of pore sizes (2.5-12 nm). We therefore surmise that these events correspond to regular head-to-tail translocations in which the polymer enters the nanopore with one end and is "reeled" through the pore in a linear fashion. For such events the magnitude of the conductance blockade is not expected to depend strongly on the pore size, as the total volume of ions blocked by the DNA should be constant.

Interestingly, our results deviate somewhat from earlier results in small nanopores by Wanunu and coworkers[21], who also reported the existence of different event populations. In their study, however, the different populations of events were less clearly distinguishable by their conductance blockade, and were rather inferred by detailed study of the event dwell times. Also, in contrast with our

measurements, they concluded that the population with intermediate conductance blockade corresponds to molecules that do not translocate through the nanopore. We postulate that this difference is a consequence of differences in pore shape or the pore surface roughness, caused by altered nanopore fabrication or storage conditions. In support of this reasoning, we demonstrate that by tuning of the TEM beam size employed in nanopore fabrication results similar to those of Wanunu *et al*. can be obtained. Our results thus indicate that small differences in the nanopore fabrication process can lead to significantly different results for biomolecular translocation. In this light, this work contributes to the understanding of the conductance blockades caused by single molecules in very small nanopores, crucial for the further sophistication of nanopores as single molecule sensors.

7.2 Materials and methods

Sample preparation

Nanopores were fabricated in 20 nm SiN membranes by a highly focused TEM beam as previously described[24] and stored in a solution containing 50% ethanol and 50% ddH_2O directly after drilling. Nanopore size was directly determined from the TEM image prior to storage, the error bar quoted for each nanopore stems from the determination of the size from the image. The measured ionic resistance through the nanopore was also used to double-check the diameter and stability of the nanopore. In several cases, the nanopore size was also remeasured following DNA translocation experiments, to verify the nanopores did not change during the course of the experiment[16]. Only results obtained in nanopores that exhibited less than 25% change in open pore current during the experiment were used. Prior to use, the chip containing the membrane with the nanopore was rinsed with de-ionized H_2O, followed by acetone and IPA, and was then blown dry. Subsequently, the sample was treated with O_2 plasma at room temperature for 30 seconds, mounted into the flow cell, and immediately flushed with electrolyte. For each experiment, about 30 seconds of conductance traces were recorded and analyzed prior to adding DNA to the *cis* chamber, to ensure that no events are detected in that situation. All experiments were performed in 1M KCl, 10 mM Tris-HCl (pH=8.0), and 1 mM EDTA.

Analysis

The voltage and current were sampled at 200 kHz bandwidth, and subsequently low-pass filtered at 35-70 kHz. For every nanopore, several hundreds to thousands of conductance blockade events were collected at 300 mV bias voltage. The typical

requirement for the detection of an event was such that a minimum of 10 consecutive data points (corresponding to a minimum dwell time of 50 μs) must deviate by more than 4.5x the standard deviation from the average current. This criterion ensures that detected events are extremely unlikely to be caused by random fluctuations of the current. After low-pass filtering, conductance blockade events were identified and corresponding histograms were constructed as previously described[12]. For each identified event, a dwell time t_D and average conductance blockade (expressed as both $\Delta G = \Delta I/V$ and I_B, see Figure **7.1b**) were determined.

Each individual event was additionally assigned a "type", such as ***0***, ***H*** or ***L*** (see Figure **7.2**), corresponding to its conductance blockade level. To perform this assignment, we first determined the conductance blockade for each of the three levels ***0***, ***H***, and ***L*** from Gaussian fits to the current histogram comprised of all the events (see Figure **7.2b**). Next, each individual event was scanned point-by-point, to determine to which of the three levels ***0***, ***H***, or ***L*** the blockade value of each point corresponded most closely. When the number of consecutive points closest to the same level exceeded a critical value (typically ~7), the event type was assigned accordingly. Multi-level events, in which a single event reached different blockade levels consecutively (such as ***LH*** or ***HL*** or more complex versions, see Figure **7.2a**) were also frequently observed. In those cases, the same minimum number of consecutive points was required for the event to be named accordingly. For example, an ***LH*** event must have minimally 7 consecutive points that are closest to the ***L***-level, followed by minimally 7 consecutive points that are closest to the ***H***-level. In the graphs of Figure **7.4a**,**b** in which we plot the percentage of ***L***-events, such multilevel events are taken into account by adding one half of the percentage of multilevel events to the percentage of ***L***-events (and similarly to the percentage of ***H***-events, see Appendix **7**, Figure **A7.2a**). This number is additionally reflected in the size of the error bars for both event types.

The dotted squares in Figure **7.3** indicating the different event types were attributed by first determining the I_B values of the ***0***-, ***H***- and ***L***-levels. The I_B values closest to the ***0***-level were then surrounded by a blue dotted square, those closest to the ***H***-level by a red dotted square, and those closest to the ***L***-level by a green dotted square. The left and right sides of the squares were determined by the observed minimum and maximum dwell time.

Molecule synthesis

In our experiments, we have employed DNA molecules of different lengths, all synthesized from DNA templates via PCR amplification. Synthesis of the 3.5 kbp and 2.2 kbp molecules is detailed elsewhere[25]. The 500 bp molecule was synthesized by PCR amplification of pBluescript1,2,3,4 using forward primer GATCGGTACCTATGGCTGAACCGGTAGG and backward primer

AAAACTGCAGTACGATCTGCCGTCAGTC, and the 6 kbp molecule was synthesized by PCR amplification of λ-phage (Promega) using forward primer AAAAGAATTCAGCCTCAGCTGACCAGCCAGAAAACGACC and backward primer GATTTGTTCGTGACCGATATGC.

7.3 Results

In Figure **7.2a**, we show a number of typical events caused by 3.5 kbp DNA molecules in a nanopore of 2.5 ± 0.4 nm in diameter (total open pore conductance G = 9.8 ± 0.4 nS). In such a small nanopore, we expect that only a single duplex DNA can occupy the nanopore at any given time, as the diameter of duplex DNA determined via X-ray crystallography is 2.0 nm[26]. Since the DNA sample used in this experiment is mono-disperse, one might expect to see only a single blockade level, corresponding to the characteristic conductance blockade of a single molecule inside the nanopore. However, the events displayed in Figure **7.2a** clearly indicate the existence of multiple levels (here indicated by ***0***, ***H***, and ***L***). Although the majority (typically ~85%) of the events are single-levelled (Figure **7.2a**, events **I** through **III**), multi-level events (Figure **7.2a**, events **IV** and **V**) are not uncommon, indicating that an individual molecule alone can produce different conductance blockades in the nanopore.

To analyze the entirety of events, we have grouped the acquired events in this nanopore together into a corresponding conductance blockade histogram (N = 950, Figure **7.2b**, and Materials and Methods). We can now clearly identify a number of peaks in the histogram (fitted by the sum of multiple Gaussians and coloured distinctly), whose different means reflect the different blockade levels observed in the individual events in Figure **7.2a**. Examining the conductance blockade histogram from left to right, we can easily attribute the first peak (grey, centered at ΔG=0 nS) to the baseline (open pore) current, as our event selection algorithm always collects a number of baseline points accompanying individual events. In addition to this baseline conductance peak, we observe a second peak (blue labelled ***0***, centered at ΔG = 0.61 ± 0.08 nS), which corresponds to the smallest observed conductance blockade. Next, there are two peaks of similar magnitude corresponding to larger blockades (red, centered at ΔG = 1.66 ± 0.05 nS and labelled ***H***; and green, centered at ΔG = 7.32 ± 0.15 nS and labelled ***L***). The peak labelled ***L*** corresponds, in this particular example, to a blockade of almost three quarters of the conductance through the nanopore.

While the conductance histogram in Figure **7.2b** allows us to clearly identify the different possible levels of blockade caused by single DNA molecules, it has the disadvantage that the information it contains cannot be traced back to the individual events. For example, the existence of multi-level events (such as in Figure **7.2a**, events **IV** and **V**) cannot be deduced from the histogram. Thus, we also plot in Figure **7.2c** a two-dimensional histogram, showing for each event the

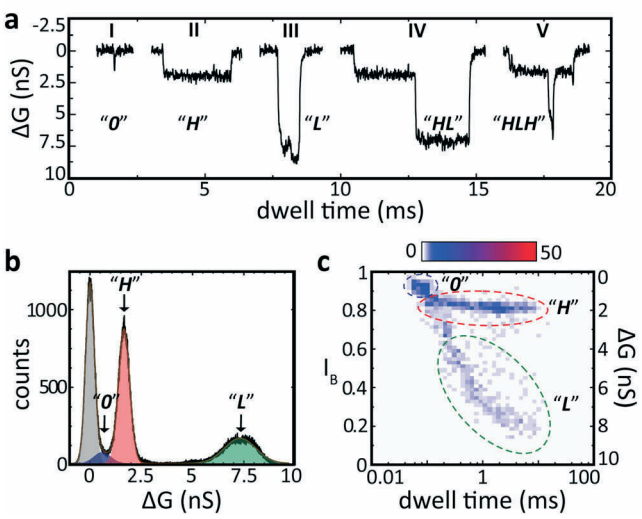


Figure 7.2. *Typical conductance blockade events and corresponding histograms from 3.5 kbp long DNA molecules, interacting with a 2.5 ± 0.4 nm diameter nanopore*. **a**. Representative conductance traces of individual events. Events **I** to **III** are single-level events: they reach a single blockade level, and return to the baseline after a while. However, there are always also a smaller number of multi-level events, such as events **IV** and **V**. The different blockade levels are labelled as ***"0"***, ***"H"***, and ***"L"***, and are discussed in the main text. **b**. Histogram of the conductance traces of all the events (black solid line). The change in conductance is plotted with respect to the open pore conductance. Each event contributes multiple counts to this histogram. A number of peaks can be distinguished, fitted by Gaussian curves and denoted by different colours. The brown solid line is the sum of these fitted Gaussians. **c**. 2-Dimensional histogram of the same set of events, showing the dwell time and relative current blockade I_B for each event. The events cluster into distinct populations (encircled by hand) which correspond to the peaks in the histogram of Figure **7.2b**.

(average) relative current blockade I_B on the vertical axis and its event dwell time t_D on the horizontal axis. This also allows us to investigate whether different conductance blockades are associated with different dwell times. Looking at the two-dimensional histogram, we can distinguish three groups of events, which can be related back to the different peaks in Figure **7.2b**: the blue peak in Figure **7.2b**, corresponding to events with a very small conductance blockade (ΔG = 0.61 ± 0.08 nS, in Figure **7.2b**), appears as a small group of events at its corresponding value of I_B = 0.94 ± 0.08 (encircled by a blue oval); the red peak in Figure **7.2b**, with ΔG = 1.66 ± 0.05 nS, corresponds to the group of events encircled by a red oval (with corresponding I_B = 0.83 ± 0.05, Figure **7.2c**); Finally, the green peak in Figure **7.2b**, with ΔG = 7.32 ± 0.15 nS, corresponds to the group of events encircled by a green oval (with corresponding I_B = 0.27 ± 0.15, Figure **7.2c**). Since for each event its

average value of I_B is plotted, multi-levelled events (as in Figure **7.2a**, events **IV** and **V**) will appear somewhere in between the ***H***- and ***L***-levels in the figure.

Using the combined information from such histograms as Figure **7.2b**,**c**, we now investigate the different groups of events to understand their origin. Starting first with the events in the ***0***-group, we find that these are associated with the shorter dwell times and smallest conductance blockade, both approaching the limits of our detection. While the number of events in this group is therefore subject to the precise parameters employed in our analysis (minimum blockade and dwell time), it is nevertheless clear that such events cannot be caused simply by random fluctuation of the current (Materials and Methods). In support hereof, no such events are ever observed when control experiments are performed in the absence of DNA (Materials and Methods). Thus, this group of events must derive from a form of interaction of the molecules in the sample with the nanopore. However, given the low magnitude of conductance blockade, it is reasonable to assume that these events do not correspond to molecules passing through the nanopore: in a 2.5 nm pore, the presence of a 2.0 nm diameter DNA molecule would likely block more than 6% of the total current (corresponding to the average $I_B = 0.94 \pm 0.08$ for these events) and would likely take longer to translocate[22,23]. We therefore propose that these events correspond to molecules that briefly skim the nanopore entrance without actually entering the nanopore itself. Additional evidence supporting this hypothesis is that such short events with only a very small blockade are also frequently observed in other translocation studies, involving larger nanopores and longer molecules[22,23].

We will now turn our attention to the ***H***- and the ***L***-events. First, we investigated the relative magnitudes of these groups of events in nanopores of different sizes and determined their respective characteristic blockade and dwell time. In Figure **7.3**, we show 2-dimensional histograms of I_B versus dwell time for nanopores with diameters varying between 1.8 and 5 nm to roughly indicate the main trends. All measurements in Figure **7.3** were performed using 3.5 kbp DNA at a bias voltage of 300 mV. As before, we can identify the three different groups of events in each size of pore (indicated by dotted squares). The dotted squares reflect the I_B range corresponding to each of the three groups as determined from Gaussian fits to the conductance blockade histogram (Materials and Methods. See also Appendix **7**, Figure **A7.4** for a plot of I_B versus pore diameter). The group of ***0***-events appears to be independent of the nanopore size, in accordance with our hypothesis that these correspond to molecules that briefly skim pore entrance. The populations of the other two groups, however, vary with pore size: We find that in the smaller nanopores in Figure **7.3a**,**b**, the ***H***- and ***L***-events are visibly not very clearly separated in these 2D histograms. Nevertheless, for the small nanopores (Figure **7.3a**-**c**), the majority of the events appears to be of type ***L***. As the pore size is increased, however, the fraction of ***L***-events decreases, until in the largest pores almost no ***L***-events are found. We also see that, for the larger nanopores (Figure

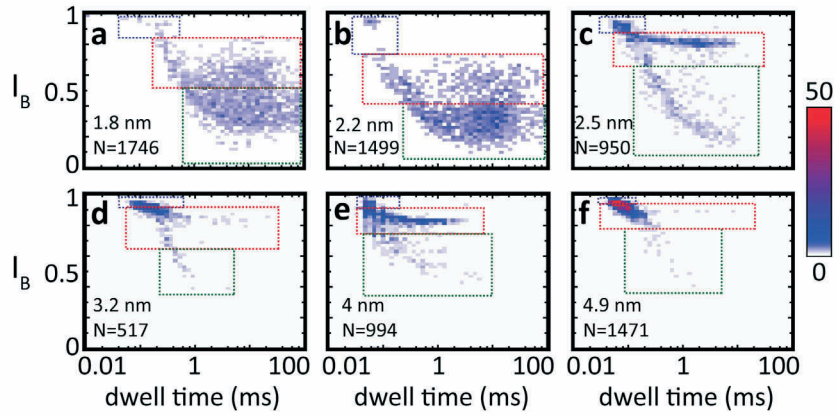


Figure 7.3. *Individual examples of 2-dimensional event histograms of dwell time versus relative current blockade I_B for nanopores of different diameters*. In all cases, events were recorded using 3.5 kbp long DNA at a bias voltage of 300 mV. The three event populations can be distinguished as before (denoted by dotted squares, see Materials and Methods), but their relative weights vary with pore size. The general trend is that ***L***-events are very frequent in the smallest pores, but as the pore size increases their frequency reduces gradually, together with a concomitant increase in ***H***-events. The variation in the conductance blockades is further quantified in Figure **7.5a** and Appendix **7**, Figure **A7.4**.

7.3e,**f**), the average event dwell time decreases (see also Appendix **7**, Figure **A7.3**), and that, while some ***L***-events can still be separately identified, the distinction between ***H***- and ***0***-events becomes less clearly visible. These last two observations can be readily explained, as in larger nanopores the translocation time of the DNA is expected to reduce due to decreased interactions with the nanopore. Also, as the total nanopore conductance scales quadratically with nanopore diameter, but the conductance blockade itself does not, the I_B of the events will approach 1 as the nanopore size is increased (recall that I_B is defined as $1 - \Delta I/I_{open}$). Thus, in particular the distinction between the ***H***-events and the ***0***-events becomes more difficult to see as the nanopore size increases. By using longer DNA and zooming in on I_B in a narrower range (e.g. ranging from I_B = 0.9 to I_B = 1), such distinction is again possible (Appendix **7**, Figure **A7.1**).

In the following, we will further quantify these trends. In Figure **7.4a**, we have plotted the percentage of ***L***-events (see Materials and Methods for a description of the determination of this percentage) relative to the total number of events as a function of pore size, for a fixed DNA length of 3.5 kbp at a bias voltage of 300 mV (for completeness, the corresponding percentages of ***H***- and ***0***-events as a function of pore size displayed in Appendix **7**, Figure **A7.2a**). Clearly, the fraction of ***L***-events increases gradually with decreasing nanopore size: whilst in the nanopores larger than 4 nm, ***L***-events are very rare, for the smallest nanopores almost all events are

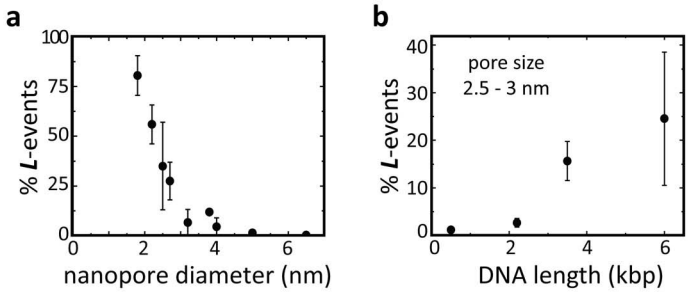


Figure 7.4. *Dependence of **L**-events on pore size and DNA length*. **a**. Fraction of ***L***-events plotted as a function of nanopore diameter for 3.5 kbp DNA at a bias voltage of 300 mV. A clear reduction of the percentage of ***L***-events is observed as the pore size increases. **b**. Fraction of ***L***-events plotted for different DNA lengths (0.5 kbp: 1214 events, 2.2 kbp: 321 events, 3.5 kbp: 2081 events, 6 kb: 2672 events) in nanopores ranging from 2.5 to 3 nm diameter. As the DNA length is increased, the relative amount of ***L***-events increases. See Materials and Methods for a definition of the percentage of ***L***-events. Both graphs also include multi-level events, reflected by the size of the error bars.

of type ***L***. Since a notable effect of reducing the nanopore diameter is presumably an increased probability that the DNA transversely approaches the nanopore, this result suggests that ***L***-events may be events in which the DNA partially obstructs the nanopore entrance. In such a scenario, an increase in ***L***-events would also be expected when the DNA length is increased, since longer DNA should also increase the likelihood for obstruction of the nanopore entrance. As can be seen from Figure **7.4b**, in which the percentage of ***L***-events in 2.5-3 nm pores is plotted for DNA lengths ranging from 500 bp up to 6 kbp in length at a bias voltage of 300 mV, an increase in the fraction of ***L***-events is indeed observed for longer DNA (for completeness, we also show the corresponding percentages of ***H***- and ***0***-events in Appendix **7**, Figure **A7.2b**). Clearly, for the longer DNA molecules the occurrence of ***L***-events is more likely than for shorter DNA molecules. Although experiments with even longer DNA lengths are in principle possible, such results were found to be very difficult to obtain, as the likelihood for permanent clogging of the nanopore was found to increase dramatically with larger DNA lengths in such small nanopores. For this reason, Figure **7.4b** does not feature DNA molecules longer than 6 kbp. Nevertheless, a clear increase in ***L***-events is visible even in this range. Together, the two observations that ***L***-events are more likely to occur when the nanopore size is reduced and when the DNA length is increased strongly suggest that these events are associated with events in which the DNA approaches the nanopore transversely, thus obstructing the entrance of the nanopore with one or multiple (duplex) strands. We will discuss possible mechanisms for this in the Discussion below.

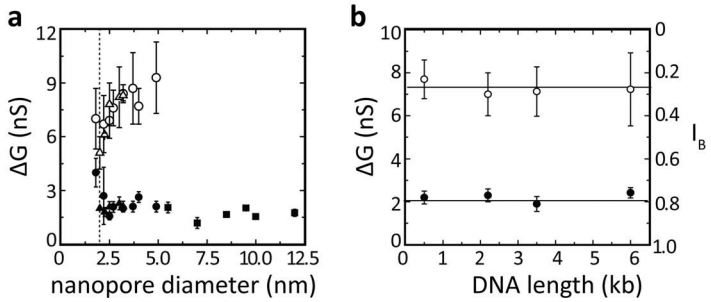


Figure 7.5. *Dependence of the blockade level of the **H**- and **L**-events on nanopore size and DNA length.* **a**. Conductance blockades of ***H***- (filled symbols) and ***L***-events (open symbols) as a function of nanopore diameter for DNA molecules of varying lengths (triangles: 500 bp, circles: 3.5 kbp, squares: 12.7 kbp). The dotted line indicates the diameter of DNA. For pores sizes approximately equal to the DNA we find a large variation in the conductance blockade of the ***H***-events, and a small increase in the conductance blockade. However, for pores larger than 2.5 nm we observe an approximate 30% variation in the conductance blockade of the ***H***-events, but find no obvious dependence on pore diameter, as the pore size is varied ~5-fold. The conductance blockade of the ***L***-events gradually increases, and the ***L***-events completely disappear in pores larger than ~5 nm. **b**. Conductance blockades of both ***H***- and ***L***-events as a function of DNA length in 2.5-3 nm diameter pores. Filled circles correspond to ***H***-events, open circles to ***L***-events. Solid lines are guides to the eye. On the right, the approximately corresponding I_B is plotted for reference (based on a 2.5 nm pore). Both types of events exhibit no dependence on DNA length for lengths varying between 500 and 6000 bp. The values and error bars for Figures **7.5a** and **7.5b** have been determined from Gaussian fits to the conductance change histogram (see Figure **7.2b**).

We now switch our attention to the ***H***-events, whose fraction gradually increases with increasing nanopore diameter (Figure **7.3**). This permits us to compare the magnitude of the conductance blockade of the ***H***-events over the pore size range probed here to results obtained previously in larger nanopores. We have reported previously that 12.7 kbp long DNA molecules translocating in a head-to-tail fashion through nanopores of approximately 10 nm in diameter caused a characteristic conductance blockade of $\Delta G = 1.68 \pm 0.15$ nS (at 300 mV bias voltage)[12]. This is strikingly close to the blockade level of the ***H***-events in the 2.5 nm pore from Figure **7.2b**, where we found $\Delta G = 1.66 \pm 0.05$ nS. Assuming that the conductance change does not depend on the nanopore diameter or DNA length, this similarity suggests that the ***H***-events may be of the same type as the translocation events observed in the larger pores involving longer DNA. Furthermore, similar ΔG values in nanopores down to 3.5 nm in diameter were also found in previous measurements, where DNA molecules tethered to an optically trapped bead were confirmed to enter the nanopores by a commensurate movement of the bead[4,27,28].

It is therefore of interest to probe how the magnitude of the conductance blockade of the ***H***-events varies over a wide range of nanopore sizes and DNA lengths. In Figure **7.5a**, we show the conductance blockade ΔG of the ***H***-events (closed symbols) and ***L***-events (open symbols) for nanopore diameters varying between 1.8 and 12 nm. The dotted line indicates the DNA helical diameter of 2.0 nm. For the smallest pore diameters < 2.5 nm, we observe an increase and a large variation in ΔG. For the pores with diameters greater than 2.5 nm, however, the conductance blockade of the ***H***-events appears to be quite independent of nanopore diameter: the conductance blockade averages around 1.9 ± 0.4 nS over the entire range (where the error is the standard deviation). In contrast, the ***L***-events show a gradual increase with increasing pore size. Also, the ***L***-events almost completely disappear for pore sizes above 5 nm (as discussed above). We note further that Figure **7.5a** includes data points for DNA molecules of a number of different lengths. There are several reasons for this. First, we found that experiments involving longer DNA (such as the 12.7 kbp DNA used previously in larger pores[12]) in nanopores smaller than 5 nm in diameter were impractical because the nanopores were very quickly permanently clogged, before enough translocation events could be collected. The smallest nanopores (< 5 nm) therefore required the use of shorter DNA. Second, the shorter DNA could not be used for the larger nanopores (> 5 nm in diameter), as in those cases the translocation time reduced below the limit of detection and the conductance blockades could no longer be properly determined. One might raise the question whether DNA of different lengths may not cause different conductance blockades in a nanopore of given diameter, thereby complicating the interpretation of Figure **7.5a**. To investigate this, we show in Figure **7.5b** the conductance blockade in nanopores of 2.5 to 3 nm in diameter for different DNA lengths ranging from 500 bp to 6 kbp. For completeness, we show both the ***H***-events (closed circles) and the ***L***-events (open circles). We find that, for both populations, the conductance blockade is a constant independent of the DNA length. Thus, given the independence of ΔG on DNA length and nanopore size (Figures **7.5b**,**a**, respectively), and the observation that the ***H***-events observed in the smaller pores exhibit conductance blockades of similar magnitude as the "regular" head-to-tail events observed in larger pores, we surmise that the ***H***-events must correspond to this type of translocations.

7.4 Discussion

When examining the interaction of DNA molecules with small solid-state nanopores, we have shown that we typically distinguish three groups of events, whose relative weight varies with pore size. In the following, we expand on the interpretation of these different groups of events.

We start with the ***H***-events, for which the most straightforward interpretation that these correspond to regular head-to-tail translocations needs to be examined

in light of the magnitude of the relative current blockade I_B measured for these events. Our experiments yield three primary results for the ***H***-events: 1) the fraction of ***H***-events increases gradually as the nanopore size is increased from ~2 nm to 5 nm (Figures **7.3**, **7.4a**); 2) for nanopores larger than 2.5 nm the magnitude of the conductance blockade of the ***H***-events is independent of both nanopore size and DNA length (Figure **7.5a**,**b**); and 3) the magnitude of the conductance blockade is similar to that of the translocations in larger pores[12,22,23] and measurements involving DNA molecules tethered to optically trapped beads[4,27,28]. Such observations would be expected from a simple model in which the conductance blockade is determined only by the exclusion of ions due to the presence of the DNA inside the nanopore. In such a model, the total volume of the DNA occupying the nanopore is the same, regardless of the nanopore diameter or DNA length (provided the membrane thickness is constant and excluding possible bending or folding of the DNA, in which case multiple duplex DNA strands could occupy the nanopore). The corresponding conductance blockade is therefore also expected to be independent of nanopore diameter and DNA length. Since this is exactly what we observe for the ***H***-events, and given that the ***H***-events gradually disappear when the pore size is reduced below 2 nm, we propose that the ***H***-events correspond to regular translocations.

Although this conclusion is compelling based on our presented results, it is nonetheless surprising that the typical I_B value for our ***H***-events ($I_B \approx 0.8$) is quite high, even for the nanopores as small as approximately 2.5 nm in diameter. This means that only ~20% of the open pore conductance is blocked. If we assume that all the ions in the volume occupied by the DNA are blocked from traversing the pore, a DNA molecule of about 2.0 nm in diameter (corresponding to its crystallographic diameter[26]) would rather be expected to block about 65% of the current through a cylindrical pore of 2.5 nm in diameter (corresponding to an I_B of 0.35). We offer two potential explanations for this discrepancy. First, the DNA is not in actual fact a closed cylinder, but a highly charged, permeable and flexible structure, attracting counterions that can possibly move within the DNA helix. In fact, recent molecular dynamics simulations[29] have demonstrated that counterions can indeed penetrate within the DNA helix, and remain mobile (albeit less so than in bulk). This implies that a portion of the ionic current may be carried along the major or minor grooves of the DNA, thus making the DNA less of a restriction for ionic conduction than a solid cylinder would be, and effectively increasing the corresponding value of I_B. Furthermore, the large amount of counterions attracted by the highly charged DNA backbone may increase the ionic conductance through the nanopore[19,20], thus also leading to a higher value of I_B during DNA translocation. Second, it is possible that the low conductance blockade observed for the ***H***-events results from the particular three-dimensional shape of the nanopores employed in these experiments. We have previously demonstrated that nanopore shape depends sensitively on the conditions employed during fabrication[16,17]. In particular, the shape depends on the size (full width at half maximum) of the focused electron beam of the TEM: a large beam compared to the nanopore size

causes an hourglass shape, whereas a small beam leads to a cylindrical shape[16]. The smallest nanopores used in this study were all drilled with a beam size of at least three times larger than the nanopore diameter, and thus these pores have an hourglass shape. For such a shape, only a very small section (in z) at the center of the SiN membrane contains the narrowest nanopore constriction. The electrical field at this location is likely higher than that in nanopores that are more cylindrical in shape, which could lead to a stronger stretching force on the DNA, making it more permeable for ions to contribute to the current. Also, the relative volume (to the total volume of the pore) excluded by the presence of the DNA is much smaller for a compressed hourglass shape than for a cylindrical nanopore, effectively lowering the relative amount of conductance blocked (thus leading to a higher value of I_B). Both explanations, namely that the DNA is an open cylinder, partially permeable to ions and that the hourglass shape of our nanopores possibly stretches the DNA at the smallest constriction, making it even more permeable for ionic conduction, may explain why the ***H***-events only block 20% of the nanopore conductance.

Another surprising fact is that, although their occurrence is greatly reduced, ***H***-events still seem to occur in even the smallest pores studied here, implying that DNA can translocate through nanopores smaller than the DNA diameter. Indeed, it has been reported previously that DNA can even penetrate nanopores as small as 0.5 nm[30], albeit at higher bias voltages. This was also supported by molecular dynamics simulations[20,31]. In our experiments, the hourglass shape of our nanopores may further facilitate the translocation of DNA through nanopores slightly smaller than the DNA helical diameter: as discussed above, the electrical field at the narrowest constriction is likely very high, allowing for possible deformation of the DNA, and subsequent translocation through the nanopore.

It is interesting to note that our results deviate from previous results published by Wanunu and coworkers[21]. In that study, the existence of different groups of events was also reported, but in contrast to the results presented here, the different populations of events were less clearly distinguishable by their conductance blockade. Rather, most events resembled our ***L***-events in the magnitude of the conductance blockade, and the existence of two populations was predominantly inferred from a detailed study of the event dwell times. From their analysis, Wanunu *et al*. concluded that the events corresponding to the smallest conductance blockade (possibly the equivalent of our ***H***-events) did not correspond to translocations, whereas we draw the exact opposite conclusion. Possibly, this striking difference could be the result of different procedures during manufacturing of the nanopore membranes or sample preparation, giving rise to different surface roughness or surface charge. Alternatively, this discrepancy with our results may originate from the shape of our nanopores: supposing that the nanopores used by Wanunu *et al*. were drilled with a significantly smaller TEM beam size than our nanopores, they could thus be expected to be more cylindrical in shape. Based on the above arguments, this would increase the conductance blockade of ***H***-events (thus lowering their value of I_B), moving them closer to the blockade level of the ***L***-

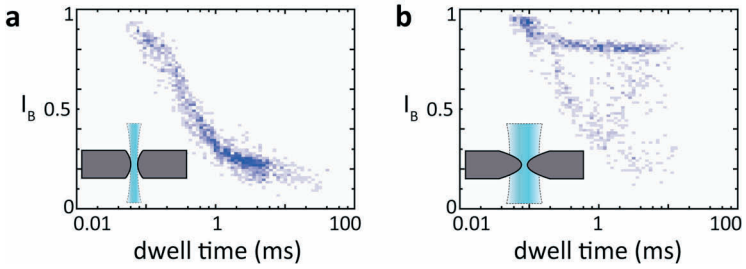


Figure 7.6. *Comparison of results obtained in nanopores drilled with different TEM beam sizes*. **a**. Results for a 2.5 ± 0.5 nm pore drilled with a TEM beam of 2 nm diameter, smaller than the beam size used for the earlier results (where a minimum beam size of 7.5 nm was employed). The smaller beam size is expected to give rise to a cylindrical shape[15], as indicated by the picture in the inset. Clearly, the conductance blockade distribution is different from the results obtained in the other nanopores: no clear distinction into ***H***-events and ***L***-events is observed. For comparison, we show in **b**. another typical 2D histogram for a 2.5 ± 0.5 nm pore drilled with a 15 nm beam, where a large population corresponding to the ***H***-events is present. In both experiments we used 3.5 kbp DNA at a bias voltage of 300 mV.

events and rendering distinction on the basis of their conductance blockade more difficult. To quantitatively investigate whether the nanopore shape influences the event distribution, we performed measurements in 2.5 to 3 nm diameter pores that were drilled with a small TEM beam size of nominally 2 nm in diameter. A typical example is shown in Figure **7.6a**, where we show a 2D histogram for a 2.5 ± 0.5 nm diameter pore. For comparison, we show in Figure **7.6b** a similar 2D histogram for one of our "regular" nanopores of 2.5 ± 0.5 nm diameter, drilled with a TEM beam of 15 nm diameter. Indeed, a clear difference is observed for the pore drilled with the smaller beam: no clearly distinguishable population of ***H***-events is observed in this case. This demonstrates that a seemingly small difference in the manufacturing conditions (TEM beam size) may indeed give rise to very different conductance blockades.

An altogether different interpretation of our data is that the ***L***-events correspond to actual translocations, and the ***H***-events correspond instead to events in which the DNA partially blocks the nanopore but does not translocate. For example, the ***H***-events could correspond to instances in which the DNA transversely blocks the nanopore, presumably leading to a much smaller conductance blockade than in the case of a translocating molecule. This interpretation would perhaps make it easier to explain the magnitude of the I_B levels of both the ***H***- and ***L***-events for small nanopores of approximately 2 to 4 nm in diameter. Yet it would be very difficult to explain why the fraction of ***L***-events decreases so dramatically with increased pore size: in nanopores with diameters exceeding 5 nm, ***L***-events are hardly detected at all (Figure **7.4a**). The observed increase in the fraction of ***L***-events for longer DNA in

Figure **7.4b** would be similarly difficult to explain. In both cases, one would rather expect the opposite dependence: translocations should be more likely when the nanopore size is increased and should be less likely when DNA length is increased, precisely the behaviour we observed for our ***H***-events.

In summary, our results therefore imply that ***H***-events correspond to regular translocations whilst ***L***-events correspond instead to a different event type in which the DNA partially obstructs the entrance of the nanopore with one or multiple (duplex) DNA strands. In the latter case, the DNA most probably approaches the nanopore in a transverse orientation, since this would be increasingly likely for smaller pores or longer DNA. Presumably, given the hourglass shape of our nanopores, the high electrical field in our small nanopores can easily trap transversely-oriented molecules at the nanopore entrance without necessarily blocking the central constriction of the nanopore. Once the molecule is trapped in this fashion, several scenarios are possible. For instance, the molecule might move away from the pore through thermal motion, thus not translocating at all. Alternatively, it might be forced through the pore by buckling or local denaturation (at 300 mV the force on the DNA can be higher than the overstretching limit of 60 pN[27]), or it might ultimately translocate in the regular head-to-tail fashion by slithering one of its ends into the pore through thermal motion. Lastly, the molecule can remain stuck to the pore, clogging it until manual reversal of the bias voltage pushes it back out. Future experiments are expected to shed light on the likelihood of these different scenarios.

7.5 Conclusions

We have presented new results of DNA molecules translocating through small solid-state nanopores with diameters ranging from 1.8 nm to 7 nm in size. In contrast to previous results involving small solid-state nanopores[21], we can clearly distinguish different levels of conductance blockade for a single species of DNA molecules. We attribute this separation capability to differences in the nanopore shape caused by small differences in the nanopore fabrication and demonstrate that we can tune the separation capability by varying the TEM beam size employed during nanopore drilling.

By varying DNA length and nanopore size, we can assign different interaction mechanisms between the DNA and the nanopore to the different conductance blockade levels. We propose that the largest conductance blockades are caused by obstruction of the nanopore entrance with one or multiple duplex DNA strands, and do not necessarily correspond to translocations. Actual head-to-tail translocations, on the other hand, appear to cause a smaller absolute blockade, even in small nanopores approaching the diameter of the DNA helix. Such an ability to distinguish different modes of translocation by the magnitude of their

conductance blockade could potentially be further exploited in the detection of local structure along DNA molecules. Finally, by demonstrating that small differences in the way nanopores are fabricated may lead to significantly different conductance blockades, our results contribute to the understanding of conductance blockades in nanopores, required for further sophistication of nanopores as single-molecule detectors.

7.6 Acknowledgments

We thank Adam R. Hall and Meng Yue Wu for help with nanopore fabrication and Aleksei Aksimentiev for useful discussions. Financial support is acknowledged from the Netherlands Organisation for Scientific Research (NWO), the Dutch Foundation for Research on Matter (FOM), and the European Science Foundation via an EURYI award.

Bibliography

1. Branton, D. et al. (2008) The potential and challenges of nanopore sequencing. *Nature Biotech*. **26**: 1146-1153.
2. Dekker, C. (2007) Solid-state nanopores. *Nature Nanotech*. **2**: 209-215.
3. Howorka, S. and Z. Siwy (2009) Nanopore analytics: sensing of single molecules. *Chem. Soc. Rev*. **38**: 2360-2384.
4. Hall, A.R., et al. (2009) Electrophoretic Force on a Protein-Coated DNA Molecule in a Solid-State Nanopore. *Nano Letters*. **9**: 4441-4445.
5. Kowalczyk, S.W., A.R. Hall, and C. Dekker (2009) Detection of Local Protein Structures along DNA Using Solid-State Nanopores. *Nano Letters*. **10**: 324-328.
6. Smeets, R.M.M., et al. (2008) Translocation of RecA-Coated Double-Stranded DNA through Solid-State Nanopores. *Nano Letters*. **9**: 3089-3095.
7. Singer, A., et al. (2010) Nanopore-Based Sequence Specific Detection of Duplex DNA for Genomic Profiling. *Nano Letters*. **10**: 738-742.
8. Tabard-Cossa, V., et al. (2009) Single-Molecule Bonds Characterized by Solid-State Nanopore Force Spectroscopy. *ACS Nano*. **3**: 3009-3014.
9. Wanunu, M., J. Sutin, and A. Meller (2009) DNA Profiling Using Solid-State Nanopores: Detection of DNA-Binding Molecules. *Nano Letters*. **9**: 3498-3502.
10. McNally, B., M. Wanunu, and A. Meller (2008). Electromechanical Unzipping of Individual DNA Molecules Using Synthetic Sub-2 nm Pores. *Nano Letters*. **8**: 3418-3422.

11. Kowalczyk, S.W., et al. (2010) Unraveling Single-Stranded DNA in a Solid-State Nanopore. *Nano Letters*. **10**: 1414-1420.
12. Skinner, G.M., van den Hout, M., et al. (2009) Distinguishing Single- and Double-Stranded Nucleic Acid Molecules Using Solid-State Nanopores. *Nano Letters*. **9**: 2953-2960.
13. Healy, K. (2007) Nanopore-based single-molecule DNA analysis. *Nanomedicine.* **2**: 459-481.
14. Storm, A.J., et al. (2003) Fabrication of solid-state nanopores with single-nanometre precision. *Nature Mat*. **2**: 537-541.
15. Wu, M.Y., et al. (2005) Formation of nanopores in a SiN/SiO2 membrane with an electron beam. *Appl. Phys. Lett*. 87**:** 113106.01-113106.05.
16. van den Hout, M., et al. (2010) Controlling nanopore size, shape, and stability. *Nanotechnology*. **21**: 115304.01-115304.05.
17. Wu, M.Y., et al. (2008) Control of Shape and Material Composition of Solid-State Nanopores. *Nano Letters*. **9**: 479-484.
18. Wanunu, M. and A. Meller (2007) Chemically modified solid-state nanopores. *Nano Letters*. **7**:1580-1585.
19. Smeets, R.M.M., et al. (2006) Salt dependence of ion transport and DNA translocation through solid-state nanopores. *Nano Letters*. **6**: 89-95.
20. Aksimentiev, A. (2010) Deciphering ionic current signatures of DNA transport through a nanopore. *Nanoscale*. **2**: 468-483.
21. Wanunu, M., et al. (2008). DNA Translocations Governed by Interections with Solid-State Nanopores. *Biophys. J*. **95**: 4716-4725.
22. Li, J.L., et al. (2003) DNA molecules and configurations in a solid-state nanopore microscope. *Nature Mat*. **2**: 611-615.
23. Storm, A.J., et al. (2005) Translocation of double-strand DNA through a silicon oxide nanopore. *Phys. Rev. E*. **71**: 051903.01-051903.10.
24. Keyser, U.F., et al. (2005) Nanopore Tomography of a Laser Focus. *Nano Letters*. **5**: 2253-2256.
25. van den Hout, M., et al. (2008) End-joining long nucleic acid polymers. *Nucl. Ac. Res*. **36**: e104.
26. Young, M.A., G. Ravishanker, and D.L. Beveridge. (1997) A 5-nanosecond molecular dynamics trajectory for B-DNA: analysis of structure, motions, and solvation. *Biophys. J.* **73**: 2313-2336.
27. Keyser, U.F., et al. (2006) Direct force measurements on DNA in a solid-state nanopore. *Nature Phys*. **2**: 473-477.
28. Van den Hout, M., et al. (2010) Direct Force Measurements of Double-Stranded RNA in Solid-State Nanopores. *Nano Letters*. **10**: 701-707.
29. Luan, B. and A. Aksimentiev (2008) Electro-osmotic screening of the DNA charge in a nanopore. *Phys. Rev. E.* **78**: 021912.01-021912.04.
30. Heng, J.B., et al. (2005) Stretching DNA using the electric field in a synthetic nanopore. *Nano Letters*. **5**: 1883-1888.
31. Comer, J., et al. (2009) Microscopic Mechanics of Hairpin DNA Translocation through Synthetic Nanopores. *Biophys. J*. **96**: 593-608.

Appendix 7

The following figures are supplied as additional information.

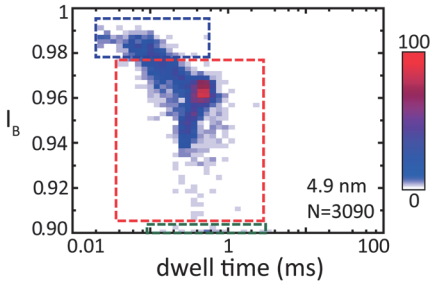


Figure A7.1. *2D-Histogram of I_B versus dwell time of the same nanopore as was used in Figure **7.3f**.* The dotted squares indicate the different event populations (***0***-events: blue, ***H***-events: red, ***L***-events: green). A longer DNA (48 kbp) and a zoomed-in vertical scale were employed, demonstrating that ***0***- and ***H***- events can be distinguished even in relatively large nanopores, such as this 4.9 nm diameter pore.

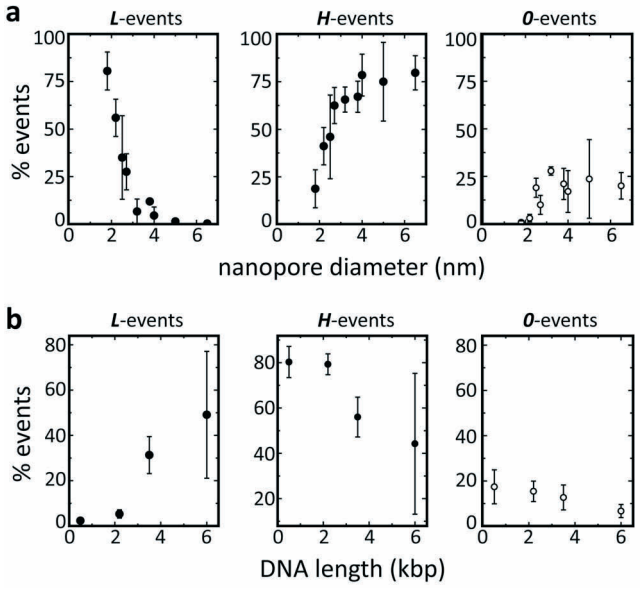


Figure A7.2. *Plots of the percentages of **L**-, **H**-, and **0**-events accompanying the graphs shown in Figure **7.4***. **a**. Percentages of ***L***-, ***H***-, and ***0***-events for different nanopore diameters ranging from 1.8 nm to 6.5 nm. **b**. Percentages of ***L***-, ***H***-, and ***0***-events for different DNA lengths ranging from 0.5 kbp to 6.0 kbp.

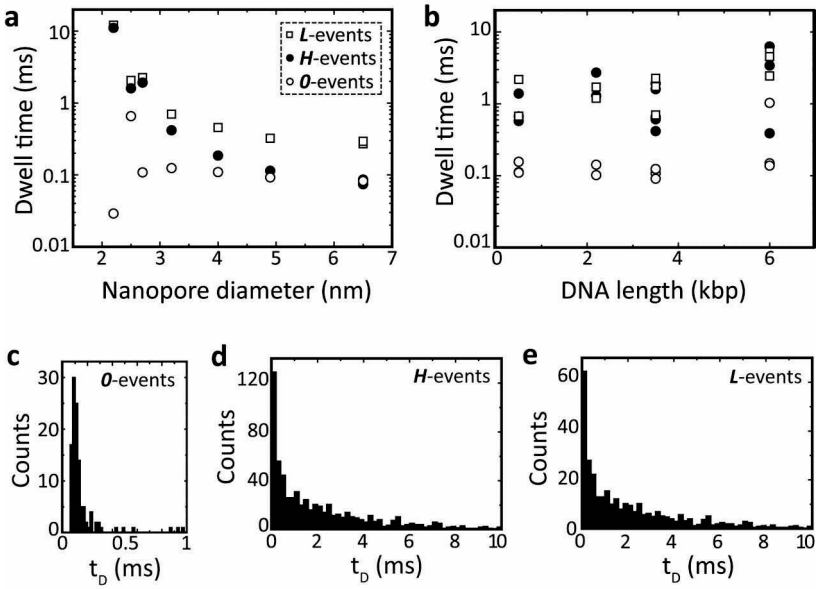


Figure A7.3. *Analysis of the dwell time (t_D) distributions for the different event types (**0**-events: open circles; **H**-events: closed circles; **L**-events: open squares).* **a.** Average dwell time versus nanopore diameter for 3.5 kbp DNA. The average dwell time is observed to decrease with increased nanopore size for both the ***H***- and ***L***- events. Typically, however, the ***L***-events are associated with the longest dwell times. **b.** Average dwell time versus DNA length in nanopores of 2.5–3 nm in diameter. The average dwell time of both ***H***- and ***L***-events increases slightly with increased DNA length. **c – e**. Dwell time distribution of ***0***-, ***H***- and ***L***-events in a 2.7 nm pore. Note the different timescale for the ***0***-events. Multilevel events are included in the ***H***- and ***L***-events as described in the Materials and Methods.

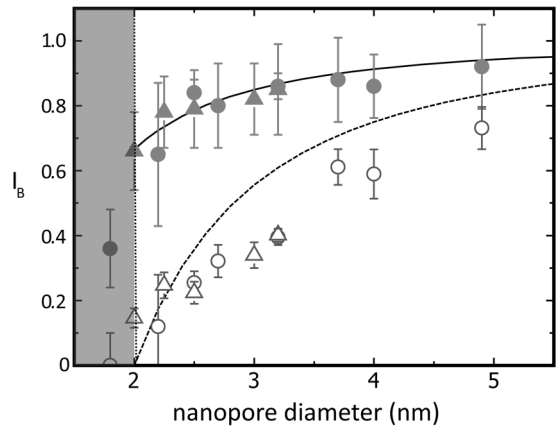


Figure A7.4. *Average I_B versus nanopore diameter for 500 bp DNA (triangles) and 3.5 kbp DNA (circles).* Closed symbols represent average I_B values for the ***H***-events, open symbols correspond to average I_B values for the ***L***-events. The I_B values (and errors) are calculated from the ΔG values (and their errors), as determined from fits to the conductance histograms using $I_B = 1 - \Delta G/G_{pore,}$, where G_{pore} is the average pore conductance. The dotted line corresponds to the value of I_B predicted under the assumptions that both the DNA and nanopore are perfect, impermeable, uncharged cylinders (d_{DNA} = 2.0 nm), and that the current blockade is fully determined by volume exclusion. The solid line corresponds to the I_B predicted under the assumptions that the pore is hourglass-shaped with a surface charge of -0.015 C/m^2, and the DNA is a permeable, charged cylinder with a linear charge density of -0.96 nC/m attracting counterions, of which 71% are assumed to be immobile[1]. The conductance blockade is described by formula (2) from Smeets *et al*[1]. The grey region denotes the region where the nanopore diameter is smaller than that of the DNA double-helical diameter of 2 nm.

[1] Smeets, et al. 2006. Salt dependence of ion transport and DNA translocation through solid-state nanopores. *Nano Letters*. 6:89-96.

Chapter 8

Forcing RNA Homopolymers through Solid-State Nanopores

Abstract

In this chapter we discuss ongoing experiments towards the detection of local structure along RNA molecules using nanopores and optical tweezers. We present preliminary translocation data on poly(U) and poly(A) through small nanopores down to about 2 nm in diameter. We find that the conductance blockade events caused by these homopolymers can be distinguished into different groups of events, very similar to our observations on dsDNA in similarly-sized pores: we again find a group of very short events with very small blockade (***0***-events), a group of events with intermediate blockade and long dwell time (***H***-events), and a group with very large blockade and long dwell times (***L***-events). As before, ***L***-events are more frequent as nanopore size is decreased, but disappear in large nanopores (> 8 nm). We also find that the magnitude of the conductance blockade of the ***H***-events is approximately constant for pores down to 4 nm in diameter. Furthermore, in pores of about 5 nm in diameter we find that the ***H***-events can cause different blockade levels, corresponding to approximately integer multiples of the lowest level, which is explained by the presence of multiple strands occupying the nanopore simultaneously. Together, our observations again support the interpretation that ***H***-events correspond to regular translocations, thus strengthening our similar conclusion for dsDNA. In the next section of this chapter, we present preliminary force measurements on poly(A) molecules arrested inside a nanopore of approximately 30 nm in diameter, and we also discuss some of the difficulties in interpreting these data. Finally, we discuss several remaining experimental considerations and challenges for the eventual detection of local structure along RNA molecules using optical tweezers and nanopores.

8.1 Introduction

Solid-state nanopores have proven to be useful tools for studying biological molecules such as dsDNA (double-stranded DNA) and RNA. Molecules of interest can be forced through the nanopore by the application of an electrical field over the pore, while their passage is monitored by a change in the ionic current. The magnitude of this ionic current during passage is related to the local size and charge density of the passing molecule[1], and can therefore possibly report on the molecule's local structure. Such detection of local structure has already been demonstrated in a number of nanopore experiments, such as the detection of locally bound proteins and ligand molecules to dsDNA[2-4], and the detection of single-stranded RNA (ssRNA) and double-stranded RNA (dsRNA) molecules (see also Chapter **4**)[5-6].

By using the pore as a mechanical constriction, however, it would be possible to probe such local structure in some more detail, as the pore would be used to disrupt the local structure along a molecule. This could be done by forcing it through a nanopore smaller than some of the molecule's constituent parts, or by dislodging other molecules that are attached to it. In this manner, the binding stability of the locally structured regions can be determined. For example, biological nanopores have already been employed to dislodge bound streptavidin molecules from dsDNA[7]. In the same way, local structure, such as duplex regions within mostly single-stranded DNA molecules have been detected and their unfolding kinetics could be studied[8-9]. In these experiments, however, the unfolding force was applied electrically, through application of a bias voltage over the nanopore. Since the exact charge distribution in those pores is not known, the magnitude of the net force on the molecules cannot be determined directly. Also, the only readout signal in those cases is the electrical current, which can primarily be used to report whether the molecule is still in the nanopore, but does not typically yield information about which parts of the molecule are being unfolded or denatured. A potentially better way of studying the local structure along molecules is by the use of a mechanical force probe to pull the molecules through a nanopore, allowing the readout of both the electrical current and the applied force. Such a force probe could potentially be used to study binding kinetics of dsDNA- or RNA-binding proteins or determine the secondary structure of complex RNA molecules directly, by sequentially unfolding duplex regions along the molecule[10] (see Figure **8.1**).

Indeed, it was recently demonstrated that such a force probe combined with nanopores was possible by combining optical tweezers with nanopores[11-12]. In this experimental configuration, molecules of interest are attached to an optically trapped bead that is moved close to a nanopore. Through the application of an electrical field over the nanopore, the molecule can then be pulled into the pore until the electrophoretic forces acting on the molecule are balanced by the optical

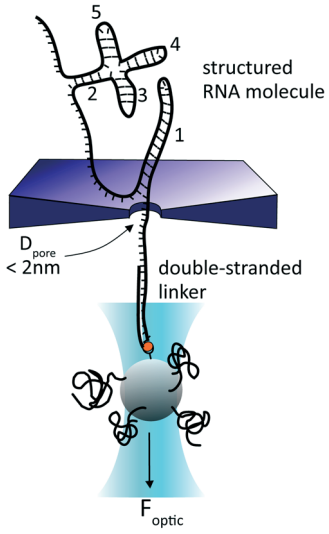


Figure 8.1. *Schematic of experimental setup to unfold an RNA molecule through a nanopore*. By attaching the RNA molecule on one end to an optically trapped bead, the molecule can be pulled through the nanopore. This setup ensures that the molecule is unfolded sequentially: region 1 must unfold first, after which the next regions unfold sequentially, independent of the relative stability of each of the subsequent duplex regions.

restoring force on the bead. Using such a setup, the forces on bare dsDNA[11,13-14], Rec-A coated dsDNA[15] and double-stranded RNA (see Chapter **5**)[16] were studied in nanopores ranging from 3.5 to 80 nm in diameter. These experiments have established a good basis for the understanding of the complex interplay of the electrophoretic and electro-osmotic forces acting on the molecule inside the nanopore. To date, however, experiments exploiting this experimental configuration to mechanically denature or unfold local structure along DNA or RNA molecules have not yet been reported. In this final chapter of this thesis we further explore our experimental progress towards this goal, in particular with respect to the study of single-stranded RNA structure. This chapter is divided into three sections: 1) translocations of RNA in very small nanopores, 2) force measurements on poly(A), and 3) an outlook of the remaining experimental challenges towards the goal of studying RNA structure using optical tweezers and nanopores.

In the first section of this chapter, we present preliminary translocation measurements on poly(A) and poly(U) molecules through nanopores of 8 nm down to 2 nm in diameter. These experiments thus approach the size range required for the potential unfolding of RNA duplex regions (< 2 nm in diameter). Similar to the results on dsDNA in small nanopores[23] (Chapter **7**), we typically find three populations of events, each with a characteristic conductance blockade, which we

again denote as ***0***-, ***H***-, and ***L***-events. As was also observed for dsDNA, we argue that these different populations correspond to attempts, in which molecules briefly skim the nanopore entrance (***0***-events), regular head-to-tail translocations (***H***-events), and translocations in the molecule enters the nanopore with many strands simultaneously, temporarily blocking the entrance of the nanopore (***L***-events). We show that at a bias voltage of 300 to 500 mV the magnitude of the conductance blockade of the ***H***-events is independent of nanopore size over the range of nanopore sizes studied. We also show that in nanopores of approximately 5 nm in diameter the ***H***-events can have several different conductance blockade levels, corresponding to approximately integer multiples of the smallest level. Such a signature was previously observed for DNA in larger nanopores[17-18], and was then attributed to the occurrence of multiple dsDNA strands in the nanopore due to folding of the molecule. Since ssRNA is more flexible, its folding could occur also in smaller pores, and this observation thus supports our interpretation that the ***H***-events correspond to regular translocations. Furthermore, we find that the ***L***–events become more abundant as the nanopore size is decreased, but vanish as the pore size is increased above 5 nm or the RNA length is decreased considerably. Finally, aiming to optimize parameters such as dwell time and signal-to-noise, we present some translocation results in 1M LiCl, instead of the 1M KCl analyte used in previous experiments. In agreement with results for dsDNA[19] we find that LiCl effectively slows down the translocation time, potentially allowing the use of shorter molecules (which may be helpful to prevent permanent clogging of the nanopore).

In the second part of this chapter, we present preliminary force measurements on poly(A) RNA molecules in a nanopore of ~30 nm in diameter, by attaching long poly(A) molecules of approximately 10-30 kilobases (kb) to an optically trapped bead and capturing the molecules inside the nanopore using an electrical field. These experiments serve as a first step to demonstrate the feasibility of capturing single-stranded nucleic acids inside a nanopore. The limited number of measurements performed so far suggest that the effective force acting on the molecule is very low: we calculate a total force of ~0.0016 pN/mV, approximately 50 times lower than that acting on the double-stranded nucleic acids, dsDNA and dsRNA. However, interpretation of the force measurements is complicated by the fact that these measurements were performed very close (< 100 nm) to the surface of the nanopore: as poly(A) molecules are very flexible on the scale of the nanopore (the persistence length of poly(A) approximately equals L_p ~ 1 nm[20]), the size of these molecules in solution is very small, forcing us to move the optically trapped bead in very close proximity to the nanopore surface in order to capture the molecule inside the nanopore. However, at such close distances, the trapping stiffness is likely to be affected, and the presence of a large number of other poly(A) molecules attached to the bead may strongly influence bead movement towards the nanopore surface. We therefore expect that the force determined from these measurements is likely an underestimation. A number of

control experiments and more force measurements must be performed before these results can be conclusive.

We finish this chapter with an outlook towards our ultimate goal to unfold an RNA molecule through a nanopore using optical tweezers. We first discuss several possible molecular constructs required to feasibly perform this experiment. Our construct of choice is a double-stranded RNA construct, which can be modified at the ends to include a single-stranded RNA molecule of interest. The double-stranded part acts as a long linker that allows the optically trapped bead to be situated at greater distances from the nanopore surface, thus eliminating the problems described above with our force measurements on poly(A). Next, we discuss a number of experimental challenges that remain in order to sequentially unfold RNA secondary structure with this setup and offer some ideas how to overcome these challenges.

8.2 Materials and methods

Sample preparation and analysis

Nanopores were fabricated and samples were prepared as described in detail in Chapter **7.2**. Nanopore sizes were determined from the TEM image directly after drilling, and double-checked by measuring the resistance. Translocation measurements were recorded at sampling frequencies ranging from 90 kHz to 200 kHz, and subsequently filtered to 35 kHz.

Homopolymer synthesis for translocation experiments

The homopolymers poly(A) and poly(U) were enzymatically synthesized using the enzyme *polynucleotide phosphorylase* (*Pnpase*), as described in Chapter **2**[21]. This protocol yields RNA molecules with varying lengths ranging from 10 to 30 kb (as estimated from gel electrophoresis measurements). In the case of poly(U), 0.5 µl of Superase·In™ (Ambion, USA) RNAse inhibitor was added to the reaction mixture which was found to slightly increase the average length (results not shown). Shorter fragments (500-2500 bases) of poly(A) were enzymatically synthesized using *poly(A) polymerase* (*Pap*), from a template A_{20} oligo-nucleotide, using the poly(A) polymerase tailing kit (Epicentre, USA), following the supplier's protocol. This reaction typically adds approximately 300-500 bases of poly(A) to the template molecule. Subsequent reactions were performed up to 5 times, yielding a final length of about 2500 bases. After synthesis, all molecules were purified and concentrated with RNeasy MiniElute Cleanup kit (Qiagen, USA) into a final volume of 30 µl of 1x TE (10 mM Tris-HCl (pH = 8.0), 1 mM EDTA) and stored at -20 ºC until

use. For the translocation experiments the molecules were diluted into final solutions of 1M KCl, 1x TE (pH=8.0) or 1M LiCl, 1x TE (pH=8.0) with final concentrations of approximately 1-10 ng / µl.

Homopolymer synthesis for optical tweezers experiments

For the optical tweezers experiments, *Pnpase*-synthesized poly(A) was subsequently labeled with biotin-ATP with the enzyme *Pap*, using the poly(A) polymerase tailing kit (Epicentre, USA). Identical reaction conditions were used as described above, except that the 1 mM ATP was replaced with 75% (by volume) of 100 µM biotin-ATP (Perkin-Elmer, USA) and 25% (by volume) of 1 mM ATP. After the labeling reaction, molecules were concentrated and purified with RNeasy MiniElute Cleanup kit (Qiagen, USA). The molecules were then attached to polystyrene beads (Spherotech, USA) of 1.87 µm in diameter by incubating 100 ng of the labeled poly(A) molecules in 25 µl of a solution containing 200 mM KCl, 10 mM Tris-HCl (pH=8.0), 1mM EDTA, containing also 0.5 µl Superase·In^{TM} (Ambion, USA) RNAse inhibitor and 0.5 µl of beads solution (1% w/v) for >30 min at 4 ^{o}C. Prior to incubation the beads were washed twice with 1M NaOH (1 min) and subsequently twice with 1x TE. For our measurements 5 µl of the RNA-functionalized bead solution was diluted into 250 µl buffer with a final concentration of 1 M KCl, 10 mM Tris-HCl (pH=8.0), 1 mM EDTA and used immediately.

Force measurements

For the force measurements, the trapping strength was calibrated as described in Chapter **5**[16].

8.3 Results

This section is subdivided into two parts. First, we will discuss preliminary results on RNA homopolymer translocations through nanopores smaller than 7 nm in diameter. Then, we will present preliminary force measurements on poly(A) in nanopores of 5 to 10 nm, and discuss some of the experimental difficulties in obtaining and interpreting these results.

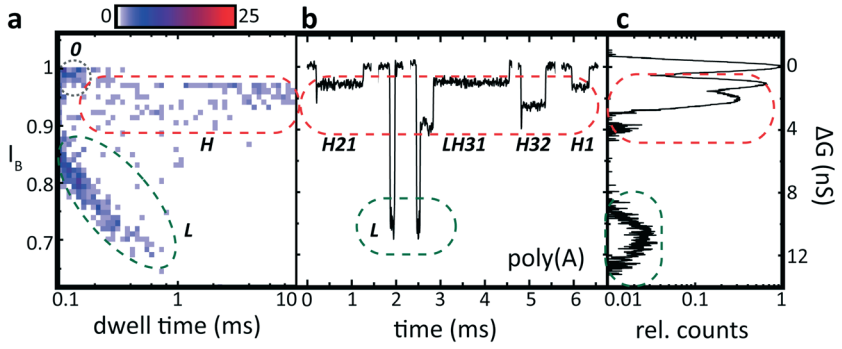


Figure 8.2. *Histograms and example event traces of poly(A) translocations in a 5 ± 2 nm pore*. **a**. Two-dimensional histogram showing I_B versus dwell time for all the 471 events in this nanopore. Analogous to our observations on DNA[23] (Chapter **7**), we distinguish several groups of events: ***0***-events, very small blockade (high I_B) with short dwell time, ***H***-events, intermediate blockade with dwell times up to ~10 ms, and ***L***-events, deep conductance blockades with typical dwell times up to ~1 ms. Dotted ovals are drawn by hand to roughly indicate the different populations. Note, that I_B is plotted in the range 0.62-1.06, meaning that even the deepest blockades only block ~40% of the total current. **b**. Example traces of individual conductance blockade events. The conductance change ΔG (in nS) is plotted versus time (in ms). A number of typical single-levelled events are shown as well as several multi-levelled events (containing more than one blockade level). Because poly(A) is much more flexible than DNA, multiple strands can easily occupy the pore simultaneously. We can distinguish these different foldings within the ***H***-events: ***H2***, for instance, means that 2 strands simultaneously occupy the nanopore. The most complicated event, labeled ***LH31***, first reaches level ***L***, then ***H3***, and then ***H1***. **c.** Conductance change histogram. The histogram clearly shows the different levels for the ***H***-events, corresponding to multiple foldings of the molecule.

8.3.1 RNA homopolymer translocations in pores < 8 nm

We have already demonstrated the feasibility of translocating RNA homopolymers through nanopores[6] ranging between 8 and 12 nm in diameter in Chapter **4**. In this section we will discuss similar measurements, in which we have extended the pore size range down to about 2 nm in diameter, the size required for the unfolding of secondary RNA structure. In Figure **8.2**, we show an example of poly(A) translocations through a nanopore of approximately 5 ± 2 nm diameter at a bias voltage of 300 mV. A two-dimensional histogram is plotted in Figure **8.2a**, showing the relative current blockade I_B (defined as $I_B = I_{blocked}/I_{open} = 1 - \Delta I/I_{open}$, see Section **7.2**) versus the dwell time for all events. Analogous to our results obtained with dsDNA, we find that events can be divided into several groups, with different conductance blockades (I_B) and dwell times. The first group, labeled ***0*** and encircled

by a black dotted oval, contains events with very short dwell times and small conductance blockade (high corresponding I_B). The second group, labeled ***H*** and encircled by a red dashed oval, consists of events with a larger conductance blockade, and longer dwell times up to about 10 ms. Finally, the third group, labeled ***L*** and encircled by a green oval, contains events with the largest conductance blockades and dwell times up to about 1 ms.

In Figure **8.2b**, we show several individual conductance blockade events: we plot the conductance change ΔG (=$\Delta I/V_{bias}$) on the vertical axis versus time on the horizontal axis. The vertical scale is chosen such that the conductance blockade ΔG aligns correctly with the corresponding I_B values from Figure **8.2a**. First, we notice that events of the ***H***- and ***L***-groups can be distinguished easily from the magnitudes of their conductance blockade (no examples of ***0***-events are shown, as the dwell time and conductance blockade of these events are too small to be easily distinguishable on this scale). Next, we notice that the ***H***-events do not all have the same conductance blockade. Rather, three different levels are found, their magnitudes corresponding to ΔG ≈ 1.1 ± 0.4 nS (which we have labeled ***H1***), ΔG ≈ 2.2 ± 0.4 nS (labeled ***H2***) and ΔG ≈ 3.5 ± 0.4 nS (labeled ***H3***). Here, the error is the estimated uncertainty from determining these values by eye from Figure **8.2b** (a better estimation is obtained by fitting Gaussians to the conductance histogram in Figure **8.2c**, see below). As this nanopore is rather large compared to a single strand of poly(A), and poly(A) is extremely flexible on the scale of the nanopore (L_p ~ 1 nm[20]), it is likely that these different blockade levels correspond to multiple strands of the molecule simultaneously occupying the nanopore. Indeed, the conductance blockade of levels ***H2*** (corresponding to two strands in the nanopore) and ***H3*** (corresponding to three strands in the nanopore) are roughly integer multiples of the blockade of level ***H1***. Similar results have also been observed previously in larger nanopores for DNA[17,22] and RNA[6]. We have labeled the individual events according to their average conductance blockade level (Figure **8.2b**). Multi-levelled events (individual events exhibiting different levels of conductance blockade) are labeled accordingly: ***H21*** means that first level ***H2*** is reached, and then level ***H1***. Interestingly, in this particular nanopore, only multi-level events were found terminating in one of the ***H***-levels, which we will further discuss below (Discussion).

Finally, we show in Figure **8.2c** the conductance blockade histogram of all the combined events in this nanopore. Again, the different blockade levels can be distinguished: from top to bottom we find first a peak corresponding to the open pore current (see also Chapter **7.2**). Next, we find three peaks corresponding to levels ***H1*** (at 1.1 ± 0.4 nS), ***H2*** (at 2.0 ± 0.7 nS) and ***H3*** (at 4.0 ± 1.7 nS). Finally, we find a peak at 10.8 ± 3.1 nS, corresponding to the ***L***-events. Here, the errors reflect the width of Gaussian fits to the peaks in the histogram.

Since our results on poly(A) show similar groups of events as observed with dsDNA[23] (Chapter **7**), we have also investigated how the magnitude of the conductance blockade for the ***H***-events depends on nanopore size. Hereto, we summarize in Figure **8.3** all our results to date on poly(A) translocations in small

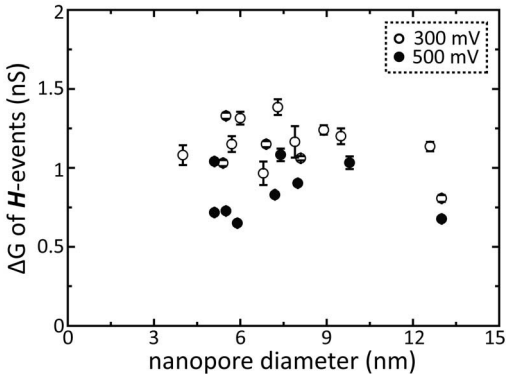


Figure 8.3. *Conductance blockade ΔG of the **H**-events from poly(A) translocations through nanopores ranging from 4 to 13 nm*. Graph shows datasets at 300 mV (open circles) and 500 mV (filled circles) bias voltage. The ΔG of the events at 500 mV are lower than those at 300 mV, in accordance with previous observations[9], due to the stronger force acting on the translocating molecules at 500 mV. No obvious dependence of ΔG on nanopore diameter is found, in agreement with results on DNA molecules in small nanopores (Chapter **7**[23]).

nanopores. The conductance blockade of the ***H***-events is plotted versus the nanopore diameter at two different bias voltages: 300 mV (open circles) and 500 mV (filled circles). Although a large spread in the ΔG values is observed at both voltages, the conductance blockade appears to be mostly independent of pore diameter. This is in accordance with our observations on dsDNA, where the ***H***-events were also found to be approximately independent of pore diameter (Chapter **7**[23]). A similar plot of ΔG versus nanopore size for the ***L***-events could not be made, as these primarily occurred in nanopores smaller than 5 nm (for which less data exists) and a well-defined conductance blockade level could in many of those cases not be determined from the conductance histograms.

The smallest nanopore in Figure **8.3** is 4 nm in diameter, still significantly larger than the diameter of the poly(A) molecule. To date, however, we were unable to get reliable results in nanopores smaller than 4 nm in diameter. This was caused in part by the lack of stability of our nanopores, which were often found to grow in size during our measurements. This effect was described in detail in Chapter **6**[24], and is attributable to the fact that a very large TEM beam was employed to fabricate these nanopores. During our measurements on poly(A), we were not yet aware how to solve this stability problem. In the few small nanopores (< 4 nm) that happened to be stable, not enough measurements could be performed to obtain reliable histograms, such as shown in Figure **8.2**. It is expected, however, that new measurements with nanopores, drilled with optimized TEM beam settings, will be relatively straightforward. Another difficulty was that the

small nanopores of less than 5 nm in diameter were found to permanently clog very quickly, before a large number of events could be collected. This may be related to the fact that the poly(A) molecules used in these experiments are very long (up to 30 kb long). In addition, gel analysis and AFM results indicate that these molecules tend to clutter together, further increasing the risk of clogging (results not shown). Experiments with much shorter poly(A) of approximately 500-2500 bases (enzymatically synthesized using consecutive runs of poly(A) polymerase, see Materials and Methods) were also tried, but here, the dwell times were too short to obtain reliable results in those cases. It is, however, interesting to mention here that in the experiments with the shorter RNA almost no ***L***-events were observed, suggesting that the occurrence of ***L***-events is also related to RNA length (although it is also conceivable that the lack of ***L***-events in those cases was due to the very short dwell times, which would distort the magnitude of their blockade). Attempts to synthesize poly(A) molecules of a length optimized for sufficiently long dwell times, but reduced likelihood for clogging (we estimate this to be between about 3000 and 8000 bases) have not been successful.

More recently, we have managed, to obtain a few preliminary results with poly(U) in nanopores of approximately 2 nm in size, using improved TEM beam settings for the fabrication of these small nanopores[24] (a TEM beam size similar to the targeted nanopore diameter, see Chapter **6**). As mentioned in Chapter **4**, poly(U) differs from poly(A) in that it is not believed to form any base-stacking, and therefore even more closely resembles a random coil than poly(A)[6]. Also, the base uracil is slightly smaller than adenine, as uracil is a pyrimidine (a base consisting only of one aromatic ring). Together, these differences result in a cross-sectional diameter of poly(U) that is significantly smaller than that of poly(A). This difference was also reflected in our measurements in larger pores[6] (Chapter **4**), where we found the average conductance blockade ΔG for poly(U) to be slightly smaller than that for poly(A). In Figure **8.4a**, we show an example of poly(U) translocations in a nanopore of 2.0 ± 0.4 nm in diameter at 500 mV bias voltage. Again, we can distinguish ***H***- (encircled by red ovals), and ***L***-events (encircled by green ovals). The variation in I_B and dwell time for the ***L***-events is quite large in this example. Also, the magnitude of I_B for the ***H***-, and ***L***-events is significantly lower than was the case in Figure **8.2a** (note the different vertical scales for both figures), which is a consequence of the smaller size of this nanopore. The magnitude of the conductance blockade of the ***H***-events (as obtained from Gaussian fits to the conductance histogram, such as **8.2c**) for this nanopore is ΔG = -1.0 ± 0.1 nS, somewhat higher than the value observed in larger pores[6] (ΔG = 0.65 ± 0.1 nS) for this molecule at 500 mV. Assuming as before (Chapter **7**) that the ***H***-events correspond to regular translocations[23], the existence of a group of ***H***-events implies that the poly(U) molecules can still easily penetrate this small nanopore, which is encouraging for future experiments on studying local structure along ssRNA molecules. This would also be expected, given the small cross-sectional diameter of a stretched poly(U) molecule[6] (<1 nm in diameter, Chapter **4**) and its highly flexible nature.

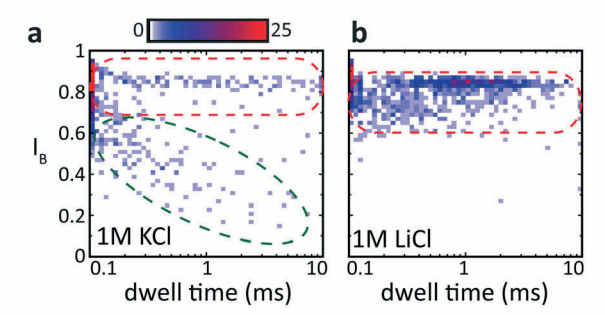


Figure 8.4. *Example 2D-histograms of I_B versus dwell time of poly(U) translocations in nanopores of ~ 2 nm in diameter.* **a**. Poly(U) translocations of a 2.0 ± 0.4 nm nanopore at 500 mV bias voltage in 1M KCl electrolyte. Again, populations of ***H***- and ***L***-events can be distinguished, indicated manually with red (***H***) and green (***L***) ovals. It is expected that poly(U) molecules with their single-stranded radius of <1.0 nm in diameter[6] can still easily translocate through the nanopore. The I_B values of the ***L***-events are significantly lower (averaging around I_B = 0.5 ± 0.3) than in Figure **8.2a**, as the nanopore diameter is much smaller (note the different range of the vertical scales in both figures). **b**. Poly(U) translocations through the same nanopore shown in **a**, but with a 1M LiCl electrolyte. The distribution of the events is significantly different: First, we observe a reduction in ***L***-events, and second, the dwell times of the ***H***-events are significantly longer.

In a further effort to optimize for dwell time and signal to noise, we have recently also explored the use of different electrolytes, such as LiCl, which was recently found to slow down dsDNA translocations by an order of magnitude (Kowalczyk, S.W. et al., *in preparation*[19]). Preliminary experiments with poly(U) and poly(A) indicate that a similar increase in dwell times is also observed for these single-stranded molecules. To illustrate this, we show in Figure **8.4b** translocation results of poly(U) in 1M LiCL, in the same nanopore as used in Figure **8.4a**. The poly(U) molecules were taken from the same stock supply as the experiment from Figure **8.4a**. Clearly, the dwell times of the ***H***-events are significantly longer for the experiment performed in 1M LiCl (Figure **8.4b**), than for that in 1M KCl (Figure **8.4a**). Interestingly, in the experiment performed in 1M LiCl, almost no ***L***-events are found. More experiments are required to find out if this is typical for measurements performed in 1M LiCl. Nevertheless, this promising example shows that the use of LiCl may be useful to extend our studies to shorter molecules, such as the short poly(A) molecules synthesized by poly(A) polymerase mentioned above.

Discussion

We have presented preliminary conductance blockade results of the RNA homopolymers poly(A) and poly(U) through nanopores down to 2 nm in diameter. In accordance with the results on dsDNA (Chapter **7**)[23], the conductance blockade events typically occur in three groups, differing in average conductance blockade and dwell time: ***0***-events (small conductance blockades, shortest dwell times), ***H***-events (intermediate conductance blockades, long dwell times) and ***L***-events (large conductance blockades, long dwell times). Here, we will briefly discuss possible physical mechanisms underlying these different levels of blockade.

We will start with the first group, the ***0***-events, corresponding to events with very short dwell times and small conductance blockades. In our measurements on dsDNA (Chapter **7**)[23] a similar group was identified, and these events were attributed to "attempts": molecules that briefly skim the nanopore entrance, but do not actually translocate through the nanopore. It is therefore reasonable to also attribute the ***0***-events in our RNA experiments to attempts, because of the very similar dwell times and conductance blockades of these events. However, it is also conceivable that these events correspond to actual translocations of very short (ie. too short to determine their ΔG properly) RNA molecules. As the synthesis procedure for our RNA molecules results in a wide range of RNA lengths, and RNA is quite easily degraded, it is likely that our samples contain reasonable amounts of such small molecules. The translocation time of these could easily be much shorter than our sampling time, giving rise to events with very short dwell time and small conductance blockade, some of which could be detected as ***0***-events. Distinction between these possibilities (attempts, or translocations of very small molecules) is currently not possible, as the occurrence of short RNA molecules cannot be prevented entirely. In any case, the short duration of these events excludes them for quantitative analysis, as the magnitude of the conductance blockade is certainly affected by the short dwell time.

Clearly, this is not the case for the ***H***-events, which have dwell times that are much longer than the minimum detectable time, and can therefore be analyzed more quantitatively. We can make several observations about ***H***-events: First, the magnitude of the conductance blockade of these events appears to be independent of nanopore diameter for nanopores down to 4 nm in diameter. These preliminary data therefore suggest a similar trend as observed for the ***H***-events of dsDNA (Chapter **7**)[23]. There, we proposed that these events must correspond to regular head-to-tail translocations, based on two observations, namely that the magnitude of the conductance blockade of these events was 1) similar to that observed in larger pores, and 2) independent of the dsDNA length used. Although a similar length dependence cannot readily be investigated for the RNA homopolymers (as synthesizing long homopolymeric RNA molecules of fixed length is not possible), the similar independence of ΔG on nanopore diameter observed for the RNA homopolymers is suggestive that these events also

correspond to regular translocations. More compelling evidence for this interpretation comes from a second observation about our ***H***-events, namely that multiple levels of conductance blockade are observed in pores of about 5 nm in diameter (***H1***, ***H2*** and ***H3***), where the conductance blockade of the higher levels is roughly an integer multiple of the smallest conductance blockade level. This was previously observed for dsDNA[17,22] and RNA[6] in large nanopores (of about 10 nm in diameter) and was attributed to folding of the molecules inside the nanopore: When two strands of a molecule occupy the pore, the conductance blockade is doubled. Such folding is typically not observed in small nanopores (<5 nm) for dsDNA molecules, because that would require very large bending energies on such length scales due to the stiffness of the dsDNA. However, the single-stranded homopolymers are very flexible, and can therefore easily translocate through the nanopore with multiple strands simultaneously, giving rise to the multiple levels of conductance blockade observed for the ***H***-events.

We have mentioned that ***L***-events were mostly found in the smaller nanopores (< 5 nm in diameter). For the larger pores (>8 nm in diameter), we have observed almost no ***L***-events. Comparing this qualitative observation again with our previous results on dsDNA (Chapter **7**[23]), we find the same trend for the dsDNA and RNA experiments: ***L***-events are hardly observed in larger nanopores, but are increasingly more frequent as the nanopore size is decreased. For the dsDNA this was seen as evidence that ***L***-events cannot correspond to regular translocations, as these are expected to be *less* likely in small nanopores and *more* likely when nanopore size is increased. The same argument applies to the ssRNA molecules, and the qualitative observation that the ***L***-events do not occur in larger pores thus implies that these cannot correspond to regular translocations. Possibly, the ***L***-events correspond instead to cases in which many entangled strands (possibly from a single molecule) enter the entrance of the nanopore simultaneously, after which the molecule can either translocate or permanent clog the nanopore (which for poly(A) was often observed). Experiments with much shorter poly(A) synthesized by poly(A) polymerase (see Materials and Methods), in which almost no ***L***-events were found, also support this interpretation: a short molecule is probably less likely to entangle with itself and thus clog the nanopore entrance.

The occurrence of multi-level events such as shown in Figure **8.2b** could possibly also shed more light on the physical mechanisms underlying the different levels of conductance blockade. For instance, the observation that for the nanopore in Figure **8.2** only multi-level events terminating in a ***H***-level were found, can be seen as supporting evidence of the above interpretations of the ***H***- and ***L***-events: If the ***L***-events correspond to obstruction of the nanopore entrance by a multitude of strands, it is likely that when the molecule is eventually forced through the nanopore, the last parts of the molecule will translocate in regular head-to-tail fashion (thus causing a ***H***-level blockade at the end of the event). However, we have also found nanopores in which most of the multi-level events ended with an ***L***-level. The multi-level events can therefore not be used to support our interpretations.

In summary, our translocations experiments with homopolymers in small nanopores indicate that these molecules can indeed translocate through nanopores as small as 2 nm in diameter, thus supporting the feasibility of studying local structure along ssRNA molecules. Furthermore, the observation of different levels of conductance blockade stresses the importance of further experiments that can potentially elucidate the underlying physical mechanisms giving rise to the different blockades. In this light, experiments in which a single molecule can be trapped and held fixed inside the nanopore using optical tweezers, could potentially also be very useful. We will discuss the first steps towards such a measurement in the next section.

8.3.2 Measuring forces on poly(A) in a nanopore

Here we will describe preliminary force measurements on poly(A) molecules in a nanopore. For these measurements we use the setup that was described in detail[16] in Chapter **5**. In short, poly(A) molecules are attached to an optically trapped bead, which is moved into the vicinity of a nanopore. An applied bias voltage over the nanopore can then capture a molecule inside the pore, causing a characteristic change in conductance, very similar to the situation described above. However, since the molecule is now attached to an optically trapped bead, it cannot completely traverse the nanopore, but instead remains inside the nanopore, pulling the bead out of its equilibrium position of the optical trap until the optical restoring force on the bead balances the electrophoretic force on the molecule (see Figure **8.5a**, and Chapter **5**)[16]. When the trapping strength k_{trap} is known, the displacement Δz of the bead can be related to a force: $F = k_{trap} \cdot \Delta z$.

In Chapter **5**, we have presented such measurements using dsDNA and dsRNA molecules of 12.7 kbp in length[16]. In order to capture these molecules in the nanopore, we found that the bead had to be moved to within ~1.5 µm from the nanopore. For experiments with poly(A), however, the molecule's much shorter persistence length requires us to move much closer to the nanopore surface, to distances of below 100 nm. At such close distances, it is probably not possible to directly relate bead displacement to force using the relationship above as the surface may affect the trapping strength and bead movement. Also, other molecules attached to the bead may obstruct bead movement towards the surface, as shown schematically in Figure **8.5b**: in our experiment, the bead is probably coated with a very large number of molecules, given the high concentration of poly(A) used during binding. As the bead is extremely close to the surface prior to capture of the molecule, it is possible that, upon capture, the bead cannot move much closer to the surface because entropic forces, resulting from the confinement of the other molecules on the bead, will "push" the bead away from the surface. Keeping these caveats in mind, we will nevertheless present some very preliminary initial results, which suggest that molecules of poly(A) can indeed be captured

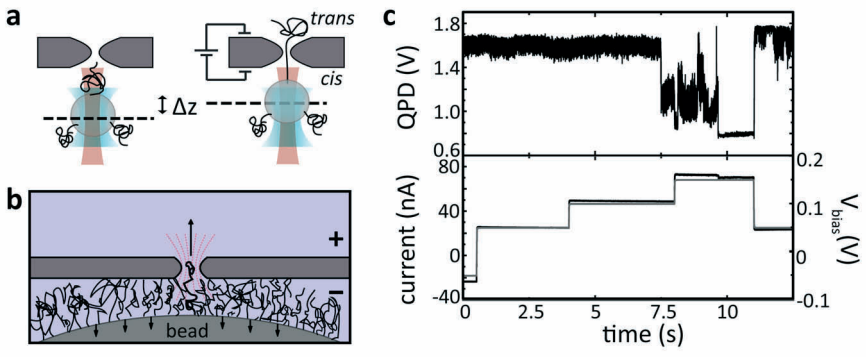


Figure 8.5. *Schematic of the experiment and example trace corresponding to possible poly(A) capture events.* **a**. Schematic overview of the experiment. An optically trapped bead, coated with charged molecules, is moved very close to the nanopore. When a bias voltage is applied, a molecule attached to the bead can be pulled in the nanopore. As it traverses the pore it will pull the bead out of its equilibrium position of the optical trap, until the electrophoretic forces on the molecule balance the optical restoring force on the bead. The change in bead position (Δz) is determined from the reflection of a separate tracking laser on the bead surface monitored by a quadrant photo detector (QPD). **b.** More realistic schematic for the case of our experiment involving a poly(A)-coated bead. The bead is very close to the nanopore surface, and is likely coated with a very large number of molecules. These may effectively push the bead (small arrows downwards) away from the surface, decreasing the bead displacement due to the force acting on the molecule inside the nanopore (arrow upwards). Electrical field lines are shown in red. **c.** Sample trace corresponding to several consecutive possible capture events of poly(A) molecules. The QPD signal is shown in the top graph, whereas in the bottom graph we show the nanopore current (black, left axis) and bias voltage (red, right axis). At t=0 the bead is positioned at about 100 nm from the nanopore surface. A positive (towards the *trans*-side) bias voltage of 50 mV is applied at t ≈ 0.5 s. Nothing happens for several seconds, then the voltage is increased, until at t ≈7.5 s the QPD signal (top graph) suddenly changes. In this nanopore the current was very noisy, and no commensurate change in current could be determined. However, at t ≈ 9.5 s, the QPD signal changes dramatically, into a signal indicating that the bead is pulled against the nanopore surface. This change is simultaneous with a significant change in the nanopore current. When the bias voltage is again reduced to 50 mV, the bead is immediately pulled back from the surface towards the optical trap.

in the nanopore using this configuration.

In Figure **8.5c**, we show sample traces of an experiment in which a bead coated with poly(A) molecules was brought to within ~100 nm of a nanopore of 30 nm in diameter. On top, we show the signal of the quadrant photo-detector (QPD), which is directly related to the position of the bead with respect to the nanopore surface (see Chapter **5**). Below, we show the current through the nanopore, and

the applied bias voltage (a positive voltage means that negative molecules are attracted towards the *trans*-side). At t ≈ 0.5 s a positive bias voltage of 50 mV is applied. For a while nothing happens, so the voltage is increased to 100 mV, until at t ≈ 7.5 s the QPD signal (top graph) suddenly changes, indicating sudden movement of the bead towards the nanopore. As this behaviour was never observed in this nanopore with uncoated beads (even when brought within ~20 nm of the surface), it is suggestive that the bead movement is due to a possible interaction of the poly(A) molecules with the nanopore: perhaps a molecule that is pulled into the pore. In this particular nanopore, the current was very noisy, and no commensurate change in current could be determined, which would be a clear indication that a molecule is indeed occupying the nanopore. However, after the voltage has been increased again to 150 mV, the QPD signal again changes dramatically at t ≈ 9.5 s, after which it appears to be quite stable. Prior to this particular event, this particular signal was determined to correspond to the situation where the bead is stuck against the nanopore surface (determined by mechanically pushing the bead against the surface, which could be done repeatedly without permanent sticking of the bead to the surface). This last change occurs simultaneously with a significant change in the nanopore current. It is likely that at this higher voltage suddenly an additional molecule (or more) is pulled into the nanopore, leading to a high enough force to press the bead against the surface. When the voltage is finally lowered back to 50 mV, the bead comes free of the surface, but does not return to its exact initial position, suggesting that several molecules may still be inside (perhaps only partially) the nanopore.

We have managed to obtain several more traces like the one shown in Figure **8.5c**, ie. situations in which the bead responded to a change in the bias voltage. Fortunately, with this nanopore the bead could be pushed against the surface and just as easily released, without any permanent attachment (sticking) of the bead to the surface. As the position of the bead could also be controlled with an accuracy of about 10 nm, we could record a calibration curve, relating QPD signal to bead displacement (see Appendix **5**, for more information), by manually moving the bead towards the surface and back, seconds prior to the capture experiment. It was thus possibly to determine the bead displacement directly in several different experiments in this nanopore. Assuming that the nanopore surface and other effects described above do not influence the trapping strength k_{trap} (which, as already mentioned, is probably incorrect), these limited data can be related to a force using the relationship $F = k_{\mathrm{trap}} \cdot \Delta z$.

In Figure **8.6**, we show the inferred force on poly(A) molecules as a function of the bias voltage, for all cases at which bead position change Δz could be determined reliably. A linear fit through zero (solid line), gives us the average force per voltage, $\mathcal{f}_{\mathrm{poly(A)}} = (1.6 \pm 0.2) \cdot 10^{-3}$ pN/mV. In the same nanopore the force per voltage on dsDNA and dsRNA were also determined to be $\mathcal{f}_{\mathrm{dsDNA}} = 0.08 \pm 0.02$ pN/mV and $\mathcal{f}_{\mathrm{dsRNA}} = 0.08 \pm 0.04$ pN/mV, respectively. The latter two values were, however, determined at distances of 0.5 - 2.5 µm from the surface. Taking the value for the force per voltage of poly(A) at face value, we would conclude that the

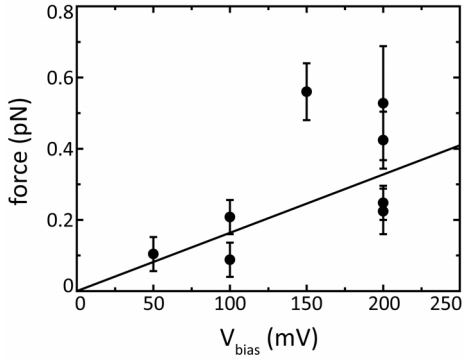


Figure 8.6. *Force on poly(A) molecules inside a nanopore versus applied voltage.* Force is calculated using $F = k_{trap} \cdot \Delta z$ from the observed bead displacement Δz in response to a change in bias voltage, in the situation in which a poly(A) molecule is inside the nanopore. The solid line is a linear fit through the data, and implies a very low force per voltage on the poly(A) of $f_{poly(A)} = 0.0016 \pm 0.0002$ pN/mV. This low value may be caused by an underestimation of the effective trapping strength k_{trap} (see text) and steric forces from the other molecules on the bead (Figure **8.5b**).

force on poly(A) is approximately 50 times lower than that on the double-stranded molecules. Interestingly, this value is in the same range as the stretching force on poly(A) inferred from our translocation experiments in Chapter **4**[6]. There, we proposed that the reduction of the conductance blockade ΔG for the homo-polymers at higher voltages could result from stretching of the molecules caused by the drag forces and the electrical field on the molecules. For poly(A), we thus inferred that the stretching force per voltage was $f_{poly(A)} = 0.008 \pm 0.004$ pN/mV, which is a factor of 5 higher than the force inferred from our optical tweezers measurement.

Discussion

Although the last observation may be interesting, we wish to point out again that, for the experiment described here, the assumption that the bead position can be related directly to force is most likely incorrect. First of all, surface charge effects (presumably very short-ranged in our case, because of our high salt conditions), and long range repulsive forces, caused by the molecules on the bead surface, may strongly affect the effective trap stiffness at the very close distances employed here. In a recent study, Schäffer, et al.[25] demonstrated that the axial trapping stiffness close to a surface can increase significantly for distances <200 nm, depending on the surface properties. Also, reflections of the trapping laser in the nanopore surface may significantly alter the trapping strength. It is therefore

expected that our trapping stiffness at the distance employed in our experiment is different from that at larger distances. Unfortunately, the close vicinity of the nanopore membrane will certainly affect the drag forces on the bead[26], which complicates trap calibration and thus makes it difficult to determine the trap stiffness directly. As a final complication, the steric effects of the other molecules on the bead (Figure **8.5b**) may strongly prevent bead movement towards the surface, and thus reduce the total force. The extent to which these combined effects influence our results is, however, very hard to estimate as our particular situation is extremely complicated. For example, the number of charged molecules attached to the bead surface is unknown, because the labeling efficiency of the poly(A) is unkown, as well as the precise number of molecules in our solutions (due to the polydispersity in RNA length). Also, the surface properties of the nanopore membrane, such as roughness and surface charge, are not known, due to the unknown influence of the TEM beam. It is therefore very difficult to estimate how much our results are influenced by these effects, but it is very likely that the overall effect is significant. The inferred very low force on poly(A) might be a consequence of these combined influences.

A relatively straightforward experiment to potentially determine how much these effects influence our data would be to measure force curves on DNA at positions far away, as well as very close (< 100 nm) to the nanopore surface. In this way, the bead response to varying voltages could be compared directly. Since the force on the DNA can be measured accurately at distances far away (> 2 μm) from the surface, the trapping stiffness very close to the surface could be determined by assuming that the actual force acting on the DNA in that situation should be the same as far away from the surface. However, it is also not obvious that this last assumption is reasonable: the influence of the bead at such close distance might well affect the access resistance of the nanopore, and thus the ion flow through the pore. This, through the effects of hydrodynamic drag, will also influence the force on the molecule. Also, the close proximity of the trapping laser is known to affect the magnitude[27] of the ion current and thus also the total of the electrophoretic and viscous drag forces.

A far better alternative, however, is to perform this experiment with the bead far removed (> 2 μm) from the nanopore surface. Here, all effects described above will not play a significant role, and forces can thus be determined directly. We have already mentioned that at such distances capturing the poly(A) molecules into the nanopore was not possible, as the radius of gyration R_g of these molecules is very small. For example, a poly(A) molecule of ~10 kb in length with a persistence length of ~1 nm has a radius of gyration of R_g ~ 40 nm. It was previously established experimentally with dsDNA that molecules can be captured into the pore from distances up about 5 x R_g from the surface[11,28]. Assuming a similar rule of thumb for single-stranded molecules, capturing should then be possible from about 200 nm distance to the nanopore surface, which is already in the problematic regime where deviations in trap stiffness are expected. Furthermore, in our experiments to date, we have not been able to obtain results that could potentially be interpreted

as captures of poly(A) molecules into the nanopore from distances larger than 150 nm from the pore membrane. In order to capture the molecules from farther away, therefore, alternative molecular constructs must be considered. We will discuss the possibilities for such a construct in the next section.

In conclusion, we have presented some very preliminary force measurements on poly(A) captured in a nanopore. Although reliable values for the force could not be extracted from these measurements, the results suggest that molecules of poly(A) can indeed be captured and held inside our nanopores using this setup. More reliable measurements require a molecular construct that allows the bead to be farther removed from the surface. We will discuss such a construct in the next section, where we will also address other outstanding challenges for the forced unfolding of RNA secondary structure through a nanopore.

8.4 Towards RNA unfolding through a nanopore

Most of the work presented in this thesis was motivated by the ultimate goal to use these nanopores to unfold the secondary structure of RNA molecules. The measurements in this last chapter are still very preliminary, but nevertheless constitute important progress towards this goal. In section **8.3.1**, we experimentally demonstrated for the first time that we can force single-stranded polymers of RNA into artificial nanopores of the size required for the unfolding of secondary structure. Second, with our optical tweezers experiments on poly(A) in section **8.3.2**, the first initial steps towards measuring forces on single-stranded RNA have been taken. There are, however, several remaining challenges before the first RNA molecule with secondary structure can be unfolded through a nanopore using optical tweezers, which we will discuss now. We already mentioned the requirement that the molecular construct must be sufficiently long that the vicinity of the nanopore membrane does not influence the measurement. Other experimental challenges of the unfolding experiment are considered next, as well as some of the limitations of our setup.

8.4.1 Molecular constructs

In order to allow the capturing of an RNA molecule of interest into the pore at sufficient distances from the nanopore membrane, we have considered several possible options. In all cases, the RNA molecule of interest will be attached to a long linker molecule, which satisfies the requirement for bead-pore separation. This means that the contour length of the linker molecule must be several micrometers long (even at low stretching forces of several pN). The most obvious candidate for such a linker molecule is dsDNA, as this can be easily synthesized to

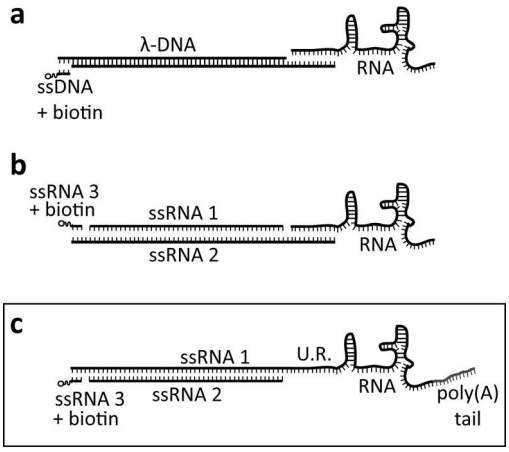


Figure 8.7. *Schematic showing several possible molecular constructs suitable for an RNA unfolding experiment.* **a**. An RNA molecule of interest (*right*) can be hybridized to a long dsDNA molecule, such as λ-DNA (*middle*). Hereto, the RNA molecule of interest can be modified slightly, such that it contains an overhang complementary to the end of one of the strands of the λ-DNA molecule. In addition, a short oligo-nucleotide (*left*) containing a biotin-label is similarly hybridized to the other end of the molecule, allowing the molecule to be attached to streptavidin-coated beads. **b**. Alternatively, the RNA molecule can be similarly hybridized to a dsRNA molecule, such as the one described in Chapter **5**[16]. Again, a short piece of RNA containing one or more biotin-labels can be hybridized to the other end of the molecule to facilitate binding to a bead. **c**. An even better option is to allow the RNA molecule of interest to be transcribed directly following transcription of one of the dsRNA linker strands, thus automatically attaching the molecule to this strand. In this way, the molecule will always be attached, independent of the hybridization efficiency of the ssRNA to the linker. The construct additionally includes a significant region of unstructured RNA (labeled U.R.) to facilitate localization of the onset of the RNA molecule of interest (see main text).

very long lengths and can also be manipulated very easily. For example, most experiments using optical tweezers in combination with nanopores were performed with λ-bacteriophage dsDNA (λ-DNA), which is approximately 16 μm long (~48 kbp). Using this molecule as a linker would seem the most straightforward option. For example, the RNA molecule of interest could potentially be hybridized to one end of a λ-DNA molecule, by creating small complementary single-stranded overhangs on both molecules (see Figure **8.7a**). In fact, λ-DNA is already equipped with a single-stranded overhang of 12 bases, which could be utilized for the hybridization of the RNA molecule, although stability would be strongly improved by creating longer overhangs. Creating such long

single-stranded overhangs on DNA is, however, not straightforward. Also, as we have described in Chapter **3**, we have found that hybridization efficiency reduces dramatically as the length of the molecules is increased. This is problematic, as once the molecules are attached to the beads, it will not be possible to distinguish correctly synthesized λ-DNA molecules (ie. with RNA on it) from λ-DNA without RNA on it. Purification of the correct fraction by gel extraction is also very difficult for these long molecules, as the difference in migration velocity will be very small due to the relative small molecular weight of the RNA molecule to the linker dsDNA. Obtaining high yields of product using dsDNA as a linker is therefore very important, and as we demonstrated in Chapter **3** this is very challenging for such long molecules.

Another possibility is the use of dsRNA as a linker (see Figure **8.7b**). One advantage here is that we have already demonstrated the synthesis of dsRNA of sufficient length for our experiments in Chapter **5**, where we used a 12.7 kbp long dsRNA molecule[16]. This construct is synthesized by in vitro transcription of complementary strands of RNA, which are subsequently hybridized. To attach an RNA molecule of interest to this construct, all that would be required is an extension of the RNA molecule of interest with a single-stranded overhang complementary to the final part of the dsRNA construct. Unlike for DNA, however, these RNA single-stranded overhangs can easily be synthesized to be quite long (hundreds of bases), as the RNA molecule is single-stranded to begin with. A long overhang strongly increases the specificity of the hybridization reaction and thus the yield of "correct" desired product (ie. constructs with the RNA of interest attached). Similarly it will increase stability of the construct (under application of mechanical force) and would eliminate the need for subsequent ligation, which is typically very inefficient[21] (Chapter **2**). Since the hybridization of multiple ssRNA molecules can be very well controlled and is known to be quite efficient[21] (see Chapter **2**), a scheme such as shown in Figure **8.7b** would be relatively straightforward.

Finally, our most promising construct is shown in Figure **8.7c**, where we present an alternative in which the RNA molecule of interest is no longer hybridized to the linker, but is incorporated directly into the linker during *in vitro* transcription: by extending the DNA template coding for the dsRNA linker with DNA that encodes for the RNA molecule of interest, *in vitro* transcription of the resulting DNA would yield a ssRNA linker strand with the RNA molecule of interest attached directly (see Figure **8.7c**). Hybridization to the complementary linker strand would thus immediately yield the desired construct with optimal efficiency. This construct also includes a section of unstructured RNA (denoted as U.R. in the figure) between the linker and the RNA molecule of interest (eg. poly(A), or otherwise a sequence of RNA which is known to have no or very little secondary structure). During an experiment, this would be helpful to determine the start of the RNA molecule of interest during an unfolding experiment: a long stretch of single-stranded RNA would first have to be pulled through the nanopore before the RNA of interest is actually unfolded. Finally, to aid capture of the molecule inside the nanopore at the

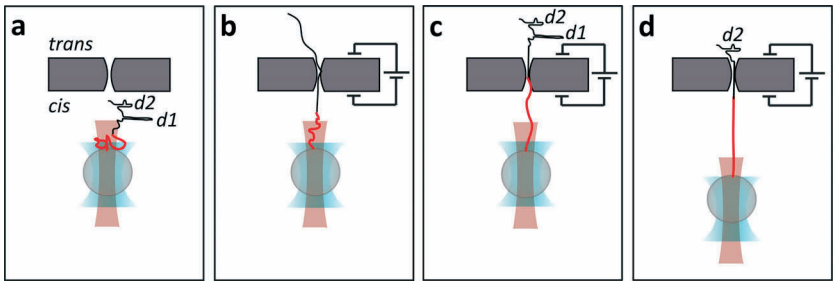


Figure 8.8. *Schematic showing the procedure for an RNA unfolding experiment.* **a**. A molecule such as the construct shown in Figure **8.7c** is attached to an optically trapped bead, which is moved close to a nanopore of ≤ 2 nm in diameter. In this particular example, the dsRNA linker is shown in red, and the attached RNA molecule of interest is shown in black. It contains two RNA duplex regions, labeled *d1* and *d2*, which must be unfolded in order to pass through the nanopore. **b**. By applying a bias voltage the single-stranded end of the RNA molecule is pulled in the nanopore towards the *trans*, and the duplex regions *d1* and *d2* are unfolded by the electrical force acting on the molecule. **c**. The bead is held in place, while the RNA molecule is allowed to refold on the *trans* side. **d**. After a while, the bead is slowly moved down, pulling the first duplex region *d1* against the entrance of the nanopore. The force initially increases very slowly, stretching first the dsRNA linker part of the construct, until finally the force is acting fully on the duplex region *d1*, which finally unfolds, suddenly lowering the force, until region *d2* reaches the nanopore entrance and can be unfolded similarly.

start of the experiment, a long stretch of poly(A) is added to the end of the molecule. This could be done by using *poly(A) polymerase* (see Materials and Methods) which would extend the molecule with 500 bases (or more through subsequent reactions) of poly(A). We thus consider the construct of Figure **8.7c** as the most suitable for our experiments and we estimate that synthesis of such a construct should be experimentally feasible.

8.4.2 Experimental challenges

Assuming that the construct shown in Figure **8.7c** can be synthesized, the next challenge is to capture such a molecule into a nanopore sufficiently small to unfold the duplex regions along the RNA. In Figure **8.8**, we show an experimental procedure how this could be done. The RNA molecule of interest (in black) in this example contains two duplex regions *d1* and *d2* (Figure **8.8a**). First the molecule, attached to the optically trapped bead is moved close enough to the nanopore to allow capturing inside the nanopore (for our 12.7 kb dsRNA molecule used in

Chapter **5**[16] a typical distance of 1.5 µm or less was sufficient)[16]. Next, a voltage is applied until the molecule is captured inside the nanopore (Figure **8.8b**). However, since the nanopore is very small, the duplex regions along the RNA molecule must first unfold to allow the molecule to pass through the nanopore. Once the molecule is pulled through the nanopore, the molecule must therefore be allowed to refold into its equilibrium structure (Figure **8.8c**)*. Depending on how complicated the RNA structure is, this could take several seconds up to minutes. Finally, the molecule can be pulled back out of the nanopore, thus again unfolding its duplex regions *d1* and *d2* (Figure **8.8d**).

As stated such, the experiment may sound straightforward, but we will now address several experimental considerations and challenges involved in performing such an experiment.

Working with small (≤ 2nm) nanopores

In this chapter, we have demonstrated that we can translocate ssRNA through nanopores of about 2 nm in diameter. This implies that capturing of such molecules while they are attached to an optically trapped bead should also be possible. Furthermore, in Chapter **5**, we have already demonstrated the feasibility capturing of dsRNA and dsDNA attached to beads into nanopores as small as 3.5 nm[16]. In those small nanopores we found that the chance of capturing a molecule into the nanopore was much lower than for larger pores. To overcome this, the number of molecules on the bead had to be increased by two orders of magnitude. For pores of 2 nm, we may therefore require even higher molecule densities on the beads. Given the low total amount of RNA material used in these experiments, however, this does not pose a problem. It is also possible that our single-stranded RNA at the end of our construct will be captured into such small nanopores much more easily than a double-stranded molecule would, eliminating the need for a higher density of molecules on our beads. It is thus expected that working with small nanopores of about 2 nm should not be problematic.

** In the experiment described here, it is likely that, once the molecule is captured, the dsRNA linker will also be forced inside the nanopore and thus be partly denatured. For this reason a section of mostly unstructured ssRNA is included in the construct which could (hopefully) be differentiated from the linker molecule. Also, upon retraction of the bead, the both strands of the linker will presumably re-hybridize, and thus cause little interference with the measurement.*

Unfolding forces

The initial unfolding of the molecule directly after capture into the nanopore should be facilitated by the electrical field forcing the molecule inside the nanopore. McNally, and coworkers have recently demonstrated[9] that DNA duplex molecules could indeed be unfolded in this manner through artificial nanopores, at modest voltages of 300 mV. It is not expected that the total force on the molecule in that situation will pull the bead out of the optical trap, as the maximum force of our optical trap is estimated to be above 80 pN (the typical force required to denature RNA secondary structure is 10 – 30 pN[29-30]). We thus expect that RNA molecules could easily be unfolded through our pores in both directions (first by the electrical field, and later by mechanical force in the reverse direction).

Experimental resolution

Given that the situation depicted in Figure **8.8c** can be reached, the molecule can then be unfolded through mechanical force. We can estimate the minimum size of secondary structure that we can detect. Excluding drift and laser instabilities, our fundamental resolution is determined by two separate effects: the trapping strength and the thermal motion of the RNA construct. The trapping strength, first of all, determines the thermal movement of the bead: increasing it lowers the thermal noise, through $\langle \delta x^2 \rangle = k_B T / k_{trap}$, where $\langle \delta x^2 \rangle$ is the variance of the bead movement, implying that a high trapping strength will improve our resolution. Increasing the trapping strength will, however, also reduce our sensitivity to detect forces (as the same force will lead to a smaller movement of the bead), but the relative large unfolding forces for RNA duplex regions (> 10 pN) should be sufficiently high to be easily detected. For an estimated maximum trapping strength of ~0.2 pN/nm the equation above gives a position accuracy of ~4.5 nm (corresponding to the standard deviation of the bead position). We expect that time averaging could probably further increase the accuracy of determining the bead position.

Next, we must consider the thermal noise of the RNA construct. This depends on the stiffness of the construct between the pore and the bead (on the *cis* side, see Figure **8.8**). Because of its short persistence length, the single-stranded part on the *cis* side will particularly cause a significant amount of thermal noise, washing out part of the information about secondary structure inside the nanopore. Recent simulations by Vocks, et al. indicate that these fluctuations will fundamentally limit the resolution to about 8 nucleotides[31], which corresponds to approximately ~5 nm (using an inter-base distance of 0.6 nm for ssRNA). They also argued that adding an electrical field pulling the molecule towards the *trans* direction (ie. in the *opposite* direction of the force from the bead) will allow the mechanical force to be increased, thus increasing the tension in the single-stranded part on the *cis* side,

which will in turn reduce the thermal noise: we estimate (from ref[31]) that a modest voltage of approximately 100 mV should already reduce the mean square displacement by a factor of ~3. Since our setup already allows us to apply such an electrical force, this would certainly be experimentally feasible.

In conclusion, we expect that our current setup should in principle allow us to detect features smaller than 5 nm in length, estimating our practical resolution to be around 3 nm. In reality, however, experimental limitations such as laser instability and, more importantly, thermal drift, due to handling of the flow cell and heating of the nanopore membrane by the trapping laser, will probably ultimately limit our resolution to about 3 to 5 nm. Putting this in a more biological context, this would allow us to detect the unfolding of approximately ~ 5 to 8 bases (using the single-stranded inter-base distance of 0.6 nm). Although this is of course not as good as single-base resolution, most biologically relevant RNA structural features such as dsRNA binding motifs are typically much longer than 10 nucleotides[32].

RNA degradation

We have previously mentioned that RNA is easily degraded, which may complicate our experiments. However, RNA degradation can typically be reduced sufficiently by working with RNAse-free solutions and materials to allow single molecule measurements[21] (see Chapter **2**). It is also not expected that the nanopore surface (composed of SiN and/or SiO_2) will particularly promote degradation of ssRNA. Therefore, we expect that RNA degradation will not pose a major challenge, although the requirements of working RNAse-free will probably make the experiment more laborious.

Effects from the trapping laser

We mentioned above that the trapping laser will heat the nanopore surface, thus raising the local temperature around the nanopore. At a laser power of about 1000 mW this can locally raise the temperature in the nanopore by 20 K[27]. Apart from introducing temporary effects, such as bending of the nanopore membrane (causing drift), the heating may affect the RNA structure, for instance by lowering unfolding barriers. This will therefore limit our ability to determine the correct unfolding forces. However, the molecular construct described above would allow us to perform these measurements at > 3 µm from the nanopore surface. The focus of the trapping laser will then be relatively far removed from the nanopore, and the heating effect is much reduced (to about 5-10 K)[27]. Also, the added value of this setup over using for example dual optical tweezers (which can determine the unfolding barriers with higher resolution) is in determining the *sequence* and

position of structured regions, for which very accurate determination of the forces is less important.

Other experimental difficulties

Last but not least, a number of experimental difficulties inherent to our setup will influence the practical feasibility of the RNA unfolding experiment. First of all, the electrical and mechanical stability of our nanopores will reduce throughput of our experiment: although we can readily produce electrically and mechanically stable nanopores, certainly not all nanopores we fabricate satisfy these requirements. Also, during the experiment instabilities in the nanopores can be induced by the trapping laser, which cannot be fully excluded, thus reducing the success rate of our experiments.

Another major difficulty in all nanopore experiments is sticking: either the molecules or the bead can aspecifically attach to the nanopore surface. In both cases, this will often terminate the experiment, because recovery to a stable nanopore after a sticking event is typically not possible. Although additives such as Tween-20 and Triton-X100 have been found to reduce sticking in some cases, they certainly do not always do so. Furthermore, in order to study native RNA structure, divalent ions such as Mg^{2+} must be present in the electrolyte because these strongly affect RNA structure[33-34]. Unfortunately, we have found that using divalent ions in our nanopore experiments will strongly enhance sticking of the molecules to the surface, making this even more challenging. Possibly, this last difficulty could, however, be overcome by using EDTA to chelate excess free ions.

Finally, leaks and air bubbles in our flow cell further limit the success rate of our experiments. Although we have continuously endeavoured to improve our procedures to reduce these effects, permanent exclusion of these experimental difficulties has not yet been achieved.

8.5 Conclusions

In this final Chapter, we have presented our progress towards the unfolding of RNA secondary structure through a nanopore using optical tweezers. Our preliminary results indicate that RNA homopolymers can be captured into nanopores as small as 2 nm, the size required to unfold duplex regions along an RNA molecule. Also, we have identified the remaining challenges to perform an actual RNA unfolding experiments, and discussed the feasibility of overcoming these experimental difficulties. Based on these assessments we conclude that it should indeed be experimentally possible to unfold a structured RNA molecule through a nanopore using optical tweezers in the future.

8.6 Acknowledgments

We thank Sven Klijnhout, Vincent Krudde and Gary M. Skinner for help with the measurements, Aart-Jan te Veldhuis for his ongoing efforts on improving the RNA construct, and Adam R. Hall, and Meng Yue Wu for help with nanopore fabrication. Financial support is acknowledged from the Netherlands Organisation for Scientific Research (NWO), the Dutch Foundation for Research on Matter (FOM), and the European Science Foundation via an EURYI award.

Bibliography

1. Smeets, R.M.M., et al. (2006) Salt dependence of ion transport and DNA translocation through solid-state nanopores. *Nano Lett.* **6**: 89-95.
2. Kowalczyk, S.W., A.R. Hall, and C. Dekker (2009) Detection of Local Protein Structures along DNA Using Solid-State Nanopores. *Nano Lett.* **10**: 324-328.
3. Kowalczyk, S.W., et al. (2010) Unraveling Single-Stranded DNA in a Solid-State Nanopore. *Nano Lett.* **10**: 1414-1420.
4. Wanunu, M., J. Sutin, and A. Meller (2009) DNA Profiling Using Solid-State Nanopores: Detection of DNA-Binding Molecules. *Nano Lett.* **9**: 3498-3502.
5. Lin, J., A. Kolomeisky, and A. Meller (2010) Helix-Coil Kinetics of Individual Polyadenylic Acid Molecules in a Protein Channel. *Phys. Rev. Lett*. **104**: 158101.1-158101.4.
6. Skinner, G.M., M. van den Hout, et al. (2009) Distinguishing Single- and Double-Stranded Nucleic Acid Molecules Using Solid-State Nanopores. *Nano Lett.* **9**: 2953-2960.
7. Tabard-Cossa, V., et al. (2009) Single-Molecule Bonds Characterized by Solid-State Nanopore Force Spectroscopy. *ACS Nano*. **3**: 3009-3014.
8. Mathe J, et al. (2004) Nanopore unzipping of individual DNA hairpin molecules. *Biophys. J.* **87**, 3205-3212.
9. McNally, B., M. Wanunu, and A. Meller (2008) Electromechanical Unzipping of Individual DNA Molecules Using Synthetic Sub-2 nm Pores. *Nano Lett.* **8**: 3418-3422.
10. Gerland, U., R. Bundschuh, and T. Hwa (2004) Translocation of structured polynucleotides through nanopores. *Phys. Biol.* **1**: 19-26.
11. Keyser, U.F., et al. (2006) Direct force measurements on DNA in a solid-state nanopore. Nat. Phys. **2**: p. 473-477.
12. Keyser, U.F., et al. (2006) Optical tweezers for force measurements on DNA in nanopores. *Rev. Sci. Instr*. **77**: 105105.

13. Keyser, U.F., S. van Dorp, and S.G Lemay (2010) Tether Forces in DNA Electrophoresis. Chem. Soc. Rev. **39**: 939 - 947.
14. van Dorp, S., et al. (2009) Origin of the electrophoretic force on DNA in solid-state nanopores. *Nat Phys.* **5**: p. 347-351.
15. Hall, A.R., et al. (2009) Electrophoretic Force on a Protein-Coated DNA Molecule in a Solid-State Nanopore. *Nano Lett.* **9**: 4441-4445.
16. Van den Hout, M., et al. (2010) Direct Force Measurements of Double-Stranded RNA in Solid-State Nanopores. *Nano Lett.* **10**: 701-707.
17. Li, J.L., et al. (2003) DNA molecules and configurations in a solid-state nanopore microscope. *Nat. Mater*. **2**: 611-615.
18. Storm, A.J., et al. (2005) Fast DNA translocation through a solid-state nanopore. *Nano Lett*. **5**: 1193-1197.
19. Kowalczyk, S.W. and C. Dekker, in preparation, 2010.
20. Seol, Y., et al. (2007) Stretching of homopolymeric RNA reveals single-stranded helices and base-stacking. *Phys. Rev. Lett.* **98**: 158103.1-158103.4.
21. Vilfan, I.D., et al. (2007) An RNA toolbox for single-molecule force spectroscopy studies. *Nucleic Acids Res.* **35**: 6625-6639.
22. Storm, A.J., et al. (2005) Translocation of double-strand DNA through a silicon oxide nanopore. *Phys. Rev. E.* **71**: 051903.01-051903.10.
23. van den Hout, M., et al. (2010) Translocations of DNA through molecular-scale solid-state nanopores. *submitted*, 2010.
24. van den Hout, M., et al. (2010) Controlling nanopore size, shape, and stability. *Nanotechnology*. **21**: 115304.1-115304.5.
25. Schäffer, E., S.F. Nørrelykke, and J. Howard (2007) Surface Forces and Drag Coefficients of Microspheres near a Plane Surface Measured with Optical Tweezers. *Langmuir*. **23**: 3654-3665.
26. Happel, J. and H. Brenner (1983) Low Reynolds Number Hydrodynamics, Martinus Nijhoff Publishers.
27. Keyser, U.F., et al. (2005) Nanopore Tomography of a Laser Focus. *Nano Lett.* **5**: 2253-2256.
28. Trepagnier, E.H., et al. (2007) Controlling DNA capture and propagation through artificial nanopores. *Nano Lett.* **7**: 2824-2830.
29. Liphardt, J., et al. (2001) Reversible unfolding of single RNA molecules by mechanical force. *Science*. **292**: 733-737.
30. Onoa, B., et al. (2003) Identifying kinetic barriers to mechanical unfolding of the T-thermophila ribozyme. *Science*. **299**: 1892-1895.
31. Vocks, H., D. Panja, and G.T. Barkema (2009) Amplitude and frequency spectra of thermal fluctuations of a translocating RNA molecule. *J. Phys.: Condens. Matter*. **21**: 375105.1-37105.7.
32. Tian, B., et al. (2004) The double-stranded-RNA-binding motif: interference and much more. *Nat Rev Mol Cell Biol*. **5**: 1013-1023.
33. Chu, V.B., et al. (2008) A repulsive field: advances in the electrostatics of the ion atmosphere. *Curr. Op. Chem. Biol*. **12**: 619-625.

34. Draper, D.E. (2004) A guide to ions and RNA structure. *RNA*. **10**: 335-343.

Summary

In this thesis we describe measurements on single molecules of DNA and RNA, using so-called nanopores. These are nanometer-sized holes, drilled in a silicon-nitride membrane using a strongly focused high-energy electron beam. By pulling a DNA or RNA molecule through such a nanopore, we can study some of its structural features. Here, we do this by mounting the nanopore in a setup such that it separates two reservoirs filled with an electrolyte. Using electrodes on either side we create an electrical field through the nanopore, causing an ionic current to flow. The passage of a single DNA or RNA molecule through the nanopore will then cause a characteristic change in the ionic current, reflecting some of the molecules' local structural properties, such as its diameter and charge density. Using this approach we present the first of such measurements on RNA and characterize the typical current changes for both RNA and DNA in nanopores of sizes that only barely allow the passage of these molecules (35 nm down to ~2 nm).

For a number of measurements in this thesis we have additionally used optical tweezers, allowing us to arrest a single molecule of double-stranded (ds) RNA (dsRNA) or dsDNA in a nanopore and to measure the force such a molecule experiences as the electrical field pulls it into the nanopore. Together, these measurements help to increase our understanding of nanopores in their role as new tools to study single molecules, and RNA in particular.

RNA molecules, like DNA, are long chains (polymers) of nucleotides, each of which can contain one of the four bases adenine (A), cytosine (C), guanine (G) or thymine (T) (which is replaced by uracil (U) in RNA). The sequence of these bases in DNA forms the so-called genetic code. In DNA molecules, two chains are packed together into the well-known double-helix structure, each chain being the others' complement: the structure is stabilized through hydrogen bonds between the bases of both strands, where the base A can only bond with T (or U in the case of RNA) and C only with G. RNA, on the other hand, typically consists of only a single chain. Because the chain itself is very flexible, it can fold up onto itself, locally forming base pairs with other parts of the chain. The final three-dimensional structure into which an RNA molecule is folded thus depends directly on the base sequence, and is therefore unique to each molecule, giving rise to a huge variety in the structure of different RNA molecules. Surprisingly, theoretical predictions of the structure of RNA molecules are still very difficult, despite great advances in computing power. Also, existing techniques to directly determine RNA structure are very cumbersome. For example, one widely used technique is X-ray diffraction, in which the diffraction of X-rays is measured as they pass through a regular crystal of RNA molecules. The challenge here is to create a useable RNA crystal, which can typically take many years of intensive research. In this thesis we present

experiments that pave the way towards an entirely new method to determine RNA structure, namely by pulling a single RNA molecule through a very narrow nanopore, forcing it to unfold its double-stranded regions in a sequential manner, thus revealing information about it structure.

We start this thesis by providing some relevant biological background, focusing in particular on RNA molecules, which can fold into many different shapes, due to their single-stranded polymeric nature (Chapter **1**). We also introduce the new field of bio-nanotechnology, with its many new methods to study biology at the level of single molecules. We then focus on how optical tweezers and in particular nanopores can be employed for the study of DNA and RNA.

As the nanopore experiments described in this thesis typically involve the preparation of suitable RNA molecules, we next address the challenges presented by the (rather limited) set of biochemical tools available to synthesize or modify RNA molecules for single molecule experiments (Chapter **2**), and illustrate the progress we have made in expanding this toolkit. We additionally present a new method to end-join very long molecules of DNA or RNA, which can significantly improve the efficiency of such a reaction over existing methods (Chapter **3**).

The remainder of this thesis describes experiments with solid-state nanopores. We describe the first translocation measurements of homopolymeric (i.e. consisting only of 1 type of base, and therefore unstructured) RNA, demonstrating that single-stranded RNA (ssRNA) can be clearly distinguished from dsRNA (and dsDNA) by the different current changes these molecules cause as they traverse the nanopore (Chapter **4**). We also observe a stretching effect of the ssRNA molecules due to the electrophoretic forces on the molecules in nanopores of approximately 10 nm in diameter.

Next, we present measurements on dsRNA and dsDNA, in which we have combined our nanopore experiments with optical tweezers (Chapter **5**). This allows us to arrest these molecules inside the nanopore and measure the forces exerted on them by the electrical field. We confirm earlier measurements on dsDNA that indicated that the force on a molecule inside the nanopore results from a combination of the purely electrical force acting on the molecule's charges, and a viscous drag force, due to counterions (shielding those charges) moving in the opposite directions[1]. As a result the net force on those molecules should depend on the nanopore size, which we indeed observe for both dsDNA and dsRNA in nanopores ranging from 3.5 to 35 nm in diameter. Furthermore, we find that the forces on dsDNA and dsRNA in our nanopores are, on average, approximately equal.

In Chapter **6**, we demonstrate that the stability of small nanopores depends on the electron (TEM) beam used for their fabrication: we discovered that small nanopores (~5 nm in diameter), fabricated using a TEM beam size significantly larger (~3x) than the nanopore diameter, slowly increase in size! In a typical experiment in which an ionic current passes through the pore, the current

gradually increases over time. Subsequent imaging after such an experiment confirmed that the diameter of these nanopores had indeed increased. We demonstrate that this instability is correlated with the shape of these nanopores: nanopores drilled with a large TEM beam adopt a compressed hourglass shape with very thin regions in the nanopore centre that have a decreased stability compared to the more cylindrically-shaped nanopores that result from the use of a smaller TEM beam during fabrication.

In the final part of this thesis we explore smaller nanopore sizes, down to a size of ~2 nm in diameter, smaller than the diameter of dsDNA and dsRNA molecules. In this size range dsRNA can no longer translocate the nanopore, thus allowing the potential unfolding of double-stranded regions along a single RNA molecule. Before investigating RNA in such small pores, however, we have first characterized the conductance blockades of dsDNA molecules in nanopores ranging from 1.8 to 7 nm (Chapter **7**). Here, we can distinguish groups of events with different conductance blockades, which we attribute to different modes of interaction of the dsDNA with the nanopore walls: a molecule can traverse the pore in a head-to-tail fashion, or it can transversely approach the pore. In this respect, our results differ from previous results by Wanunu and coworkers[2], who could not clearly distinguish such different conductance blockades. We hypothesize that this might be caused by a difference in nanopore shape (again related to the TEM beam size used during the fabrication of the nanopores).

Finally, we present in Chapter **8** preliminary measurements of homopolymeric poly(A) and poly(U) ssRNA through pores down to about 2 nm in diameter. As was the case for dsDNA, we observe similar groups of conductance blockade events, analysis of which indicates that unstructured ssRNA can still easily translocate through nanopores with diameters as small as 2 nm. This thus allows for the potential unfolding of more complicated RNA molecules through such small nanopores. We also present preliminary force measurements on poly(A) ssRNA, and discuss some of the difficulties involved in collecting and interpreting these data. Finally, we conclude this thesis with a number of considerations on the feasibility of using optical tweezers to unfold a single (structured) RNA molecule through a nanopore.

Michiel van den Hout
September 2010

Bibliography

1. van Dorp, S. et al. (2009) *Nat Phys*. **5**: 347-351.
2. Wanunu, M. et al. (2008) *Biophys. J.* **95**: 4716-4725.

Samenvatting

Dit proefschrift beschrijft metingen aan enkele moleculen van DNA en RNA, waarbij we gebruik maken van zogenaamde kunstmatige nanogaatjes. Dit zijn gaatjes met een diameter van enkele nanometers, die in een silicium-nitride membraan zijn "geboord" met een sterk gefocuseerde straal van elektronen. Door een DNA of RNA molecuul door zo'n klein gaatje te trekken, kunnen sommige structurele eigenschappen van zo'n molecuul bepaald worden. In dit proefschrift doen wij dit door een ionenstroom door het nanogaatje te laten lopen, waardoor we heel nauwkeurig kunnen meten wanneer een enkel DNA of RNA moleculen door het gaatje beweegt. Dit kan door het nanogaatje in een opstelling te monteren zodat het de enige verbinding vormt tussen twee reservoirs, gevuld met een waterige oplossing met ionen (een electroliet). Met behulp van electrodes kan dan een electrisch veld worden aangelegd, waardoor er een ionenstroom door het gaatje gaat stromen. Als een van de reservoirs nu DNA of RNA bevat, zullen deze moleculen af en toe door het gaatje bewegen, wat een karakteristieke blokkade in de ionenstroom door het gaatje veroorzaakt. Deze stroomblokkade zegt iets over de lokale structuur van het passerende molecuul, zoals bijvoorbeeld zijn ladingsverdeling en diameter. We presenteren hier de eerste van dit soort metingen met RNA moleculen en karakteriseren deze blokkade voor DNA en RNA in gaatjes die zó klein zijn, dat de moleculen er nog maar net doorheen kunnen (~ 2 nm in diameter).

Daarnaast hebben we voor een aantal experimenten in dit proefschrift een zogenaamd optisch pincet gebruikt, waarmee we met behulp van een laserstraal een enkel molecuul kunnen "vastpakken" en er krachten op kunnen uitoefenen. We gebruiken het hier om een enkel dubbelstrengs (ds) RNA (dsRNA) of dsDNA molecuul in zo'n nanogaatje tegen te houden, en zodoende de krachten te meten waarmee we met het elektrisch veld aan het molecuul trekken. Gezamenlijk dragen deze metingen bij aan het algemeen begrip van het gebruik nanogaatjes als instrumenten om biologische moleculen en RNA in het bijzonder te bestuderen.

RNA moleculen zijn, net als DNA, eigenlijk lange sliertvormige moleculen (polymeren) bestaande uit een lange reeks nucleotiden, die ieder een van vier verschillende basen adenine (A), cytosine (C), thymine (T) (voor RNA is dat uracil, U) of guanine (G) kunnen bevatten. Deze basen vormen in het DNA de zogenaamde genetische code. Maar waar ze voor DNA heel ordelijk zijn opgerold in de vorm van een dubbele helix, zijn RNA moleculen over het algemeen heel complex opvouwen. Dit verschil komt voort uit het feit dat een DNA molecuul twee afzonderlijke strengen bevat die ieder elkaars complement zijn, dat wil zeggen, de basen in beide strengen hebben een sterke voorkeur om tegen elkaar aan te "plakken" (door middel van zogenaamde waterstofbruggen). Zo heeft de base A een sterke

voorkeur om te paren met de base T (of U), en hetzelfde geldt voor C en G. Een RNA molecuul, daarentegen, bestaat uit slechts één streng. Omdat de streng echter wel alle vier de basen bevat kan hij lokaal met zichzelf gaan paren, waardoor het molecuul zich op een gecompliceerde manier kan opvouwen. Deze opvouwing is volledig afhankelijk van de base volgorde in het RNA, en verschilt dus voor ieder molecuul. Opmerkelijk genoeg kunnen we ondanks de enorme vooruitgang in computerrekenkracht nog steeds niet voor een gegeven basevolgorde voorspellen hoe het molecuul precies zal zijn opgevouwen. Ook zijn bestaande technieken om de structuur van RNA direct te meten vaak uiterst ingewikkeld. Een veelgebruikte techniek is bijvoorbeeld röntgendiffractie, waarin de verstrooïng wordt gemeten van röntgenstraling in kleine kristallen van RNA moleculen. De uitdaging hier ligt voornamelijk in het maken van een bruikbaar kristal van RNA moleculen, wat vaak jaren aan intensief onderzoek kan kosten. Dit proefschrift beschrijft experimenten die uiteindelijk een totaal nieuwe techniek mogelijk moeten maken om de structuur van een RNA molecuul te bepalen, namelijk door het door een nanogaatje te trekken, zodat iedere opvouwing in het molecuul sequentieel ontrafeld wordt.

We beginnen dit proefschrift met een algemene introductie, met de noodzakelijke biologische achtergrond informatie (Hoofdstuk **1**). We richten daarbij onze aandacht voornamelijk op DNA en RNA moleculen. Daarnaast introduceren we de bio-nanotechnologie, een spannend nieuw onderzoeksgebied, waarin biologie bestudeert wordt op het niveau van enkele moleculen. We kijken daarbij specifiek naar twee technieken, die in dit proefschrift uitgebreid aan bod komen: het optisch pincet, en kunstmatige nanogaatjes.

Aangezien we voor onze experimenten veelvuldig RNA moleculen hebben moeten synthetiseren, geven we in Hoofdstuk **2** een uitgebreid overzicht van de (tamelijk beperkte) biochemische technieken die momenteel beschikbaar zijn om RNA moleculen te manipuleren, zodanig dat ze bruikbaar zijn in enkel-molecuul experimenten. We bespreken de voortgang die op dit gebied geboekt is, maar ook de nog bestaande beperkingen en uitdagingen. Daarnaast beschrijven we in Hoofdstuk **3** een nieuwe methode om zeer lange DNA of RNA moleculen aan elkaar vast te maken, met een hogere efficientie dan de hiervoor bestaande methodes.

In de rest van dit proefschrift beschrijven we experimenten met kunstmatige nanogaatjes. Hoofdstuk **4** beschrijft de eerste translocatie metingen van homopolymeer's (dwz. bestaande uit slechts 1 soort base en daarom zonder specifieke structuur) RNA. Hier laten we zien dat zulk enkelstrengs RNA (esRNA) duidelijk van dsRNA of dsDNA te onderscheiden is, aan de hand van de verschillende stroomblokkades die deze moleculen veroorzaken wanneer ze door een nanogaatje bewegen. Uit deze metingen blijkt bovendien dat de elektroforetische kracht in het centrum van de nanogaatje de (zeer flexibele) esRNA moleculen lokaal kan uitrekken.

Vervolgens presenteren we metingen aan dsRNA en dsDNA, waarin we onze nanogaatjes met een optisch pincet combineren (Hoofdstuk **5**). Hierdoor kunnen we deze moleculen tegenhouden, terwijl ze in het nanogaatje worden getrokken door het elektrisch veld. Zodoende kunnen we direct de electroforetische kracht meten die zo'n molecuul naar binnen trekt. Onze metingen bevestigen eerdere metingen aan dsDNA waaruit bleek dat deze kracht een combinatie is van de puur elektrische kracht op de ladingen van het dsDNA, en een visceuze wrijvingskracht van tegengesteld geladen ionen die deze ladingen afschermen[1]. Dit laatste verlaagt de totale kracht op het molecuul en zorgt ervoor dat deze afhankelijk is van de diameter van het nanogaatje. Wij observeren dit effect voor zowel dsRNA als dsDNA in nanogaatjes varierend in diameters van 3.5 tot 35 nm. Bovendien wijzen onze metingen uit dat de krachten op dsDNA en op dsRNA gemiddeld genomen gelijk zijn.

In Hoofdstuk **6** laten we zien dat de stabiliteit van kleine nanogaatjes af kan hangen van de elektronenbundel die gebruikt is voor hun fabricage: wij ontdekten dat kleine nanogaatjes (~5 nm in diameter), die met een aanzienlijk (~3x) grotere elektronenbundel gemaakt zijn dan de diameter van het gaatje, in de loop der tijd langzaam groter worden. In een typisch experiment waarin een ionenstroom door het nanogaatje loopt bleek de totale stroom namelijk langzaam toe te nemen. Door achteraf het nanogaatje opnieuw te bekijken in de elektronenmicroscoop bleek de diameter van de nanogaatjes inderdaad toegenomen te zijn. We laten zien dat deze instabiliteit gecorrelleerd is aan de vorm van het nanogaatje: nanogaatjes geboord met behulp van een grote elektronenbundel krijgen een vorm van een soort ingedrukte zandloper, waardoor ze in het centrum lokaal erg weinig materiaal bevatten. Hierdoor kunnen zij gemakkelijker vervormen dan de meer cylindervormige nanogaatjes die verkregen worden wanneer een smallere TEM bundel gebruikt wordt voor de fabricage.

In het laatste gedeelte van dit proefschrift behandelen we nóg kleinere nanogaatjes, tot aan een diameter van circa 2 nm, kleiner dan de diameter van dsDNA of dsRNA moleculen. Hiermee zou het mogelijk moeten zijn om een RNA molecuul dat gedeeltelijk uit dsRNA bestaat door zo'n nanogaatje te ontvouwen (het dsRNA is dan namelijk te groot voor het nanogaatje en moet dus ontrafelen). Alvorens ons op RNA te richten, hebben wij eerst gekarakteriseerd wat de stroomblokkades zijn die dsDNA veroorzaakt in nanogaatjes met diameters van 1.8 tot aan 7 nm (Hoofdstuk **7**). Hier zien we dat het dsDNA verschillende niveaus van stroomblokkades kan veroorzaken, wat wij toeschrijven aan verschillende manieren waarop het dsDNA in contact komt met het nanogaatje: zo kan het ofwel in een kop-staart manier door het nanogaatje heen bewegen, ofwel dwars op het gaatje terecht komen. In dit opzicht verschillen onze metingen van eerdere resultaten van Wanunu[2], et al., die zulke verschillende niveaus van stroomblokkade niet zo duidelijk konden onderscheiden. Wij schrijven dit verschil toe aan een mogelijk verschil in de drie-dimensionale vorm van de nanogaatjes die we gebruikt

hebben (wat weer gerelateerd is aan de elektronenbundel gebruikt voor hun fabricage).

Tot slot presenteren we in Hoofdstuk **8** voorlopige metingen van translocaties van homopolymeer's poly(A) en poly(U) esRNA in kleine nanogaatjes tot aan circa 2 nm in diameter. Net als bij het DNA blijken deze moleculen verschillende niveaus van stroomblokkades te kunnen veroorzaken, waaruit wij concluderen dat onopgevouwen esRNA nog vrij eenvoudig door deze nanogaatjes van 2 nm in diameter kan bewegen. Dit maakt het dus mogelijk om een opgevouwen (gedeeltelijk bestaande uit dsRNA) RNA molecuul door een gaatje te kunnen ontrafelen. Verder laten we voorlopige krachtmetingen zien aan poly(A) esRNA en bespreken we een aantal van de complicaties die bij deze metingen en de interpretatie daarvan optreden. We besluiten dit proefschrift met een inschatting over de haalbaarheid om met behulp van een optisch pincet een enkel opgevouwen RNA molecuul door een nanogaatje te trekken, met als doel daarbij iets over de vorm van het molecuul te leren.

Michiel van den Hout
September 2010

Bibliografie

1. van Dorp, S. et al. (2009) *Nat. Phys.* **5**: 347-351.
2. Wanunu, M. et al. (2008) *Biophys. J*. **95**: 4716-4725.

Curriculum Vitae



28-9-1976

Michiel van den Hout

Born in Leiderdorp, the Netherlands

About the author

Michiel van den Hout was raised in Leiden, where he attended high school at the *Stedelijk Gymnasium.* He then studied physics at *Utrecht University* , graduating in *Delft* under supervision of prof. dr. Cees Dekker on "2-Dimensional imaging of electronic wave functions in carbon nanotubes". After this, he joined ASM Europe, a world-leading manufacturer of furnaces for the semi-conductor industry, where as project manager he was (among other projects) involved in the development of a new vertical furnace capable of atomic layer deposition. In 2006 he started this PhD research at *Delft University of Technology* under the supervision of prof. Nynke H. Dekker.

2006-2010	Ph.D research at Delft University of Technology *under supervision of prof. dr. Nynke H. Dekker*
2003-2005	Project Manager at ASM Europe B.V.
2001-2003	Project Engineer at ASM Europe B.V.
1994-2001	M.Sc. Physics *Utrecht University*
1999	Internship at Chemisch Weekblad *(Wolters Kluwer)*
1988-1994	Secondary School (VWO Gymnasium) *Stedelijk Gymnasium,* Leiden
1980-1988	Elementary school *Zijlwijkschool,* Leiden

List of publications

1 *Translocations of homopolymeric RNA through nanometer-sized solid-state nanopores*
Michiel van den Hout, G.M. Skinner, S. Klijnhout, V. Krudde, and N.H. Dekker
in preparation

2 *Distinguishable populations report on different interactions of single DNA molecules*
Michiel van den Hout, V. Krudde, X.R. Janssens and N.H. Dekker
under review at Biophysical Journal

3 *Controlling nanopore size, shape and stability*
Michiel van den Hout, A.R. Hall, M.Y. Wu, H.W. Zandbergen, C. Dekker and N.H. Dekker
Nanotechnology. **21:** 115304 (2010)

4 *Direct force measurements on dsRNA in solid-state nanopores*
Michiel van den Hout, I.D. Vilfan, S. Hage and N.H. Dekker
Nano Letters. **10:** 707-707 (2010)

5 *Distinguishing ss- and ds- nucleic acids molecules in solid-state nanopores*
G,M. Skinner*, Michiel van den Hout*, O.D. Broekmans, C. Dekker and N.H. Dekker (* = *equal contribution*)
Nano Letters. **21:** 115304 (2009)

6 *End-joining long nucleic acid polymers*
Michiel van den Hout, I.D. Vilfan, S. Hage and N.H. Dekker
Nucleic Acids Research. **36:** e104 (2008)

7 *An RNA toolbox for single-molecule force-spectroscopy studies*
I.D. Vilfan, W. Kamping, Michiel van den Hout, A. Candelli, S. Hage and and N.H. Dekker
Nucleic Acids Research. **35:** 6625-6639 (2007)

8 *2-dimensional imaging of electronic wave functions in carbon nanotubes*
S.G. Lemay, J.W. Janssen, Michiel van den Hout, M. Mooij, M.J. Bronikowski, P.A. Willis, P.A. Smalley, L.P. Kouwenhoven and C. Dekker
Nature. **412:** 617-620 (2001)

9

Scanning Tunneling Spectroscopy on a Carbon Nanotube Buckle
J.W. Janssen, S.G. Lemay, Michiel van den Hout, M. Mooij, L.P. Kouwenhoven, and C. Dekker
Proc. of the American Institute of Physics Conference, H. Kuzmany, S. Roth, J. Fink, and M. Mehring (eds.), Kirchberg, Austria. **591:** 293-297 (2001)

10

Elektrongolven in koolstof nanobuisjes
Michiel van den Hout en J.W. Janssen
Nederlands Tijdschrift voor Natuurkunde. **65:** 294-297 (2001)

List of figures and tables

Figure	Title	page
1.1	Chemical structure of DNA	15
1.2	Chemical structure of RNA	18
1.3	Examples of advances in nanotechnology	20
1.4	Working principle and example of optical tweezers	20
1.5	Nanopore detection principle	22
1.6	Combination of optical tweezers and nanopores	24
2.1	Examples of single-molecule studies of RNA	33
2.2	Basic design of RNA constructs for single-molecule experiments	39
2.3	3' end labeling of DNA and RNA oligomers using TDT	45
2.4	Comparison of hybridization of RNA molecules	48
2.5	Synthesis of dsRNA using pnpase and RdRP	51
2.6	Ligation of dsRNA molecules with complementary 5' overhangs	53
A2.1	Synthesis procedure for three ssRNA strands	64
3.1	Protocol and formation of DNA-STV-DNA dimers	71
3.2	End-joining very long DNA	73
3.3	End-joining DNA to RNA and RNA to RNA	75
A3.1	Influence of temperature and STV purification	84
A3.2	Influence of concentration and incomplete removal of STV	86
A3.3	Influence of duration of the second incubation step	88
A3.4	Influence of relative concentration on dimer formation	89
4.1	Experimental setup and molecules used for translocation	96
4.2	Examples of translocation events	98
4.3	Relationship between the applied bias and ΔG	100
4.4	Discrimination between ds and ss nucleic acids	101
A4.1	Explanation of the event detection algorithm	110
A4.2	Influence of filtering the raw data before event detection	110
A4.3	Typical average event rate for representatvie molecules	111
A4.4	Average translocation time versus the applied bias	112
A4.5	Histograms of ΔI for poly(A), poly(C) and poly(U) at 100 mV	114
A4.6	Extended version of Figure **4.3**	114
5.1	Principle of the experiments	120
5.2	Capturing dsRNA and dsDNA molecules	122
5.3	Force dependence on distance of bead to nanopore surface	124
5.4	Force per voltage on dsRNA and dsDNA versus pore diameter	126
A5.1	Position detection	131
A5.2	Calibration curve over a distance of 3 µm	132

Figure	Title	page
6.1	TEM beam size influences nanopore shape	137
6.2	Ionic resistance of nanopores decreases over time	138
6.3	Nanopores drilled with wide TEM beams increase in size	140
6.4	Nanopore size can be changed by thermal oxidation	142
A6.1	Nanopore stability vs TEM beam size for small beam sizes	146
7.1	Experimental setup	149
7.2	Typical ΔG events and histograms for DNA in a 2.5 nm pore	153
7.3	2D event histograms for different nanopore diameters	155
7.4	Dependence of ***L*** -events on pore size and DNA length	156
7.5	ΔG Dependence of ***H*** - and ***L*** -events on pore size and DNA length	157
7.6	Comparison of nanopores drilled with different TEM beam sizes	161
A7.1	2D-histogram with different scale and longer DNA	165
A7.2	Plots showing the % of ***L*** -, ***H*** - and ***0*** -events of Figure **7.4**	165
A7.3	Dwell time distributions for different event types	166
A7.4	Average I_B versus nanopore diameter	167
8.1	Experimental setup to unfold an RNA molecule through a pore	171
8.2	Histograms and example event traces of poly(A) in a 5 nm pore	175
8.3	ΔG of ***H*** -events of poly(A) in nanopores of 4-13 nm in diameter	177
8.4	2D event histograms of poly(U) in nanopore of ~2nm in diameter	179
8.5	Schematic and example trace of poly(A) capture event	183
8.6	Force on poly(A) in a nanopore versus applied voltage	185
8.7	Possible molecular constructs for an RNA unfolding experiment	188
8.8	Schematic of procedure for an RNA unfolding experiment	190

Table	Title	page
2.1	Promoter sequences of commonly used viral RNA polymerases	41
2.2	Overview of RNA ligation schemes using *RNL1* or *RNL2*	50
A2.1	Primers used for the synthesis of ssRNA strands	64
A3.2	Influence of concentration and incomplete removal of STV	86