CHARACTERIZATION NUCLEIC ACIDS UNWINDING AND EXPLORING ITS APPLICATION IN MIRNA DETECTION IN THE NANOPORE

A Dissertation

Presented to

The Faculty of the Graduate School

At the University of Missouri

In Partial Fulfillment

Of the Requirements for the Degree

Doctor of Philosophy

By

YONG WANG

Dr. Li-Qun (Andrew) Gu, Dissertation Supervisor

MAY 2012

The undersigned, appointed by the dean of the Graduate School, have examined the dissertation entitled

Characterization of Nucleic Acids Unwinding and Exploring Its Application in miRNA Detection in the Nanopore

Presented by Yong Wang

A candidate for the degree of

Doctor of Philosophy

And hereby certify that, in their opinion, it is worthy of acceptance.

Li-Qun (Andrew) Gu, Ph.D

Shi-Jie Chen, Ph.D

Kevin D. Gillis, D. Sc.

Fuhong Hsieh, Ph.D

Tzyh-Chang Hwang, Ph.D

ACKNOWLEDGEMENTS

First of all, I would like to sincerely thank my supervisor Dr. Li-Qun (Andrew) Gu for his guidance and support during my Ph.D study. I am so grateful for his mentorship, insightful scientific views and comments on the research, as well as his consideration, caring for my experiments and daily life here as a student. Without Dr. Gu's heart of giving, support and loving, I would never have been able to complete the study, and develop myself toward a scientist.

I am also thankful to my committee members: Dr. Kevin Gillis, Dr. Tzyh-Chang Hwang, Dr. Shi-Jie Chen, and Dr. Fuhung Hsieh. I am so thankful to have you as my committee members, as well as your suggestions, discussions and comments, which greatly helped my research.

I also would like to thank all the members of Dr. Gu's group, Qiulin Tan, Kai Tian, Ji Wook Shim, Insoon Kang, Brandon Ritzo for their kind assistance. Without your help and support during my experiments, I would not have been able to finish the research and achieve all the results.

Finally, I would to thank all the brothers and sisters in my church, thank you all for your love, help, support and prayers. As well as the encouragement and support from my parents, wife and friends.

And I know that my Ph.D. study was under the watchful eye and with the great help of the Lord. I thank God for all the gifts, blessings given to me, and may all the glory to be yours.

TABLE OF CONTENTS

LIST OF FIGURES

Figure **Page**

LIST OF TABLES

Table **Page**

CHARACTERIZATION NUCLEIC ACIDS UNWINDING AND EXPLORING ITS APPLICATION IN MIRNA DETECTION IN THE NANOPORE

Yong Wang

Dr. Li-Qun (Andrew) Gu, Dissertation Supervisor

ABSTRACT

MicroRNAs (miRNAs) are a class of short (~14-27 nucleotides) noncoding RNAs that regulate gene expression at the post-transcriptional level. As powerful gene regulators, miRNA binding induces either translational repression or cleavage of target mRNAs. MiRNAs play important roles in development, cell differentiation, and regulation of cell cycle, apoptosis and signaling pathways. Aberrant expression of miRNAs has been found in all types of tumors. Different cancer types have distinct miRNA expression profiles. Reverse transcription real-time polymerase chain reaction (qRT-PCR) and microarrays have been developed for detecting microRNA; however, these methods need labeling and amplification, as they also suffer from cross-hybridization, low selectivity and lack of valid internal controls.

The development of nanopore sensors for microRNA detection is a new effort. One of the superior properties of nanopore is that the ion current in a nanometer-scaled pore structure is very sensitive to the presence, location and conformation of single target molecules occupying the ion pathway. This

sensitivity allows elucidating single molecule kinetics from characteristic changes in the pore conductance, and further, quantifying the target from the occurrence of single molecule signature events.

Here we show that a nanopore sensor based on the alpha-hemolysin protein can selectively detect microRNAs at the single molecular level in the plasma samples of lung cancer patients without labeling and amplification of microRNAs. First, we uncovered a signature current pattern that can be used to electrically track the double-stranded DNA (dsDNA) unzipping process. With the signature signals, we can also distinguish the release of dsDNA without unzipping. Second, based on the electrical signatures we identified, we have designed a nanopore-based microRNA sensor that uses a programmable oligonucleotide probe to generate a signature electrical signal for the direct and label-free detection of target microRNA in a fluctuating background, such as plasma RNA extracts from clinical samples.

This sensor can quantify picomolar levels of cancer-associated microRNAs and can distinguish single-nucleotide differences between microRNA family members. This nanopore method can be a useful tool for quantitative studies of microRNAs, which are important for non-invasive screening and the early diagnosis of diseases such as cancer.

CHAPTER 1

INTRODUCTION

1.1 Nanopores

A nanopore is a small hole, with the diameter from several nanometers (nm) to hundreds of nanometers. It could be a pore-forming protein, ion channels or as a hole in synthetic materials such as glass, silicon, silicon nitride or graphene. The nanopores are used as a single-molecule detector for broad biotechnological applications (Howorka and Siwy, 2009), such as DNA sequencing (Rhee and Burns, 2006; Branton et al., 2008), sensing of nucleic acids, exploring DNA-protein, protein-protein interactions, studying enzyme functions, and for detecting small peptides and terrorist agents (Liu et al., 2010).

1.1.1 Solid state nanopores

Solid state nanopores are a hole in synthetic membranes, generally made in silicon membranes, and one of the most common being silicon nitride (Dekker, 2007; Siwy and Howorka, 2010; Keyser, 2011), and others as grapheme (Garaj et al., 2010; Siwy and Davenport, 2010; Tada et al., 2011), glass (Ding et al., 2009; Gu and Shim, 2010) or glass slides (Fertig et al., 2001).

Solid-state nanopores can be manufactured with several techniques including ion-beam sculpting (Li et al., 2001), transmission electron microscope (TEM) technique or electron beams (Storm et al., 2003).

1.1.2 Biological nanopores

Biological nanopores are protein nanopores, formed by pore-forming proteins inserted into a lipid bilayer membrane or other polymers.

1.1.2.1 Alpha-hemolysin (α-HL)

Hemolysins are toxins produced by bacteria that cause lysis of red blood cells. Alpha-Hemolysin (α-HL) secreted by *Staphylococcus aureus* forms a homo-heptameric beta-barrel in biological membranes (Protein Data Bank:7AHL) (Song et al., 1996).

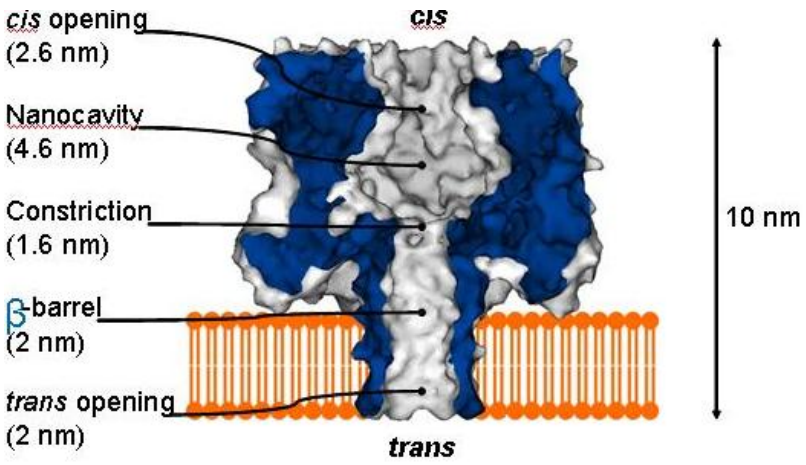

Figure 1.1: The structure of Alpha-hemolysin.

Several properties of α-HL make this membrane pore/channel suitable for various biotechnological applications: this self-assembled channel is stable over a wide range of pH and temperature. The channel stays open and can be

inserted into various biological or synthetic lipid bilayers. The inserting proceeds spontaneously and does not require specific ionic conditions. Furthermore, the precise mushroom structure (Figure 1-1) makes this pore unique for bio-detection: the big cap has a 2.6 nm opening, 4.6 nm nanocavity. In the center of the pore, it has a constriction site of 1.6 nm, which allows single-stranded DNA (ss DNA, diameter ~1.2 nm) pass through very fast, but not double-stranded DNA (dsDNA, diameter ~ 2.2 nm). The 2 nm β-barrel stem spans the lipid bilayer. The nanopore is a molecular-scale pore structure that is able to detect the position and conformation of a single molecule in the pore with great sensitivity (Bayley and Jayasinghe, 2004).

1.1.2.2 Mycobacterium smegmatis porin A (MspA)

MspA porin (MspA) is a membrane β-barrel protein produced by *Mycobacteria*, which allows nutrients to enter the bacterium.

The protein forms a tightly interconnected octamer with eightfold rotation symmetry that contains a central channel. The cylindrical geometry structure of the MspA channel was established by electron microscopy, and later the atomic structure was established (Faller et al., 2004). Dr. Gundlach and his colleagues at the university of Washington are interested in MspA nanopore sequencing, as well as bio-detection to probe nano-scale physics of proteins and polymers. They found that the wild-type MspA channel prevents ssDNA translocation, but the mutated channel was proven to be suitable for DNA sequencing (Derrington et al., 2010; Manrao et al., 2011) as well as single-molecule DNA detection (Butler

et al., 2008). The geometry dimensions of the MspA protein pore are shown in Figure 1-2.

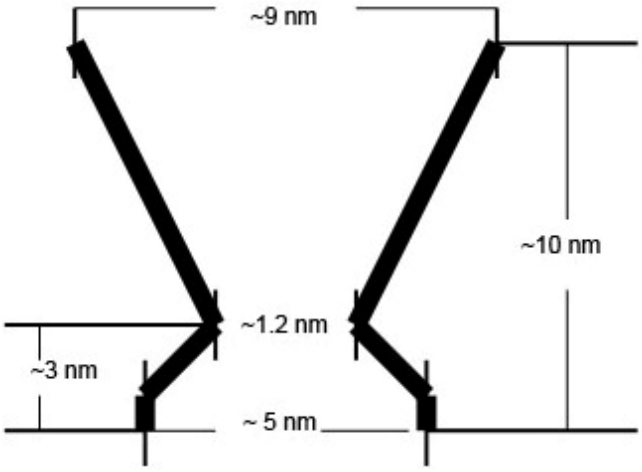

Fig 1-2: The simple geometry dimensions of MspA.

1.1.2.3 Bacterial virus phi29 motor protein

Motor proteins are powered by the hydrolysis of ATP. They are a class of molecular motors that can move along the surface of a suitable substrate. For example, myosin moves along microfilaments.

The Bacterial virus phi29 motor protein has been extensively studied since 1987 by Dr. Guo and his colleagues at the University of Kentucky (Guo et al., 1987). They incorporated the phi29 motor channel into the lipid bilayer membrane which allows the dsDNA to pass through the pore. This engineered pore can be used for single molecule sensing, dsDNA sequencing (Wendell et al., 2009; Jing et al., 2010), as well as delivery of therapeutics (Shu et al., 2011). The dimensions of the phi29 motor protein are shown in Figure 1-3.

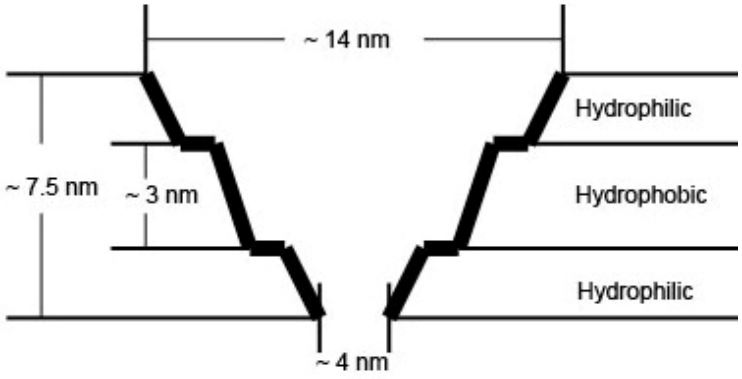

Fig 1-3: The simple geometry dimensions of Bacterial virus phi29 motor protein porin.

1.1.2.4 β-Barrel Membrane Proteins

Dr. Movileanu and his colleagues from Syracuse University are working on redesigning pores from bacterial outer membrane proteins and pore-forming toxins, which contain a robust β-barrel structure. They explored the ferric hydroxamate uptake component A (FhuA), a monomeric 22-stranded β-barrel protein from the outer membrane of *Escherichia coli*, and after engineering, the pore showed greater conductance compare to the wild-type pore, which could be applied for stochastic single-molecule sensing of proteins and nucleic acids (Mohammad et al., 2011).

1.1.2.5 Outer Membrane Carboxylate Channels (OCC)

Gram-negative bacteria contain ion channels within the outer membrane (OM), which acquires water-soluble, small compounds that are required for cell growth and normal function. The majority of molecules are taken up by the

members of the OprD outer membrane protein family. Dr. Bert van den Berg and his colleagues found that OprD channels require a carboxyl group in the substrate for efficient transport. They renamed the family Occ, for outer membrane carboxylate channels, and crystal structures of Occ family members show large variations in pore sizes. They demonstrated that Occ channels show high specificity for antibiotics uptake, as well as substrate specific (Eren et al., 2012).

1.1.2.6 Anthrax toxin

Anthrax is a disease, which is caused by *Bacillus anthracis*, a spore-forming, rod-shaped bacterium. The lethality of the disease is caused by the tripartite protein toxin, called anthrax toxin. Anthrax toxin is composed of three proteins: 1) protective antigen (PA), 2) edema factor (EF), and 3) lethal factor (LF). Like the voltage-gated channels (KcV) studied in the artificial membranes (Shim et al., 2007; Tan et al., 2010), the PA pore was found to form an ion-selective channel (Blaustein and Finkelstein, 1990; Blaustein et al., 1990).

Studies also found that PA formed pore conductance can be eliminated at positive voltage after EF (red) binding (Figure 1-4).

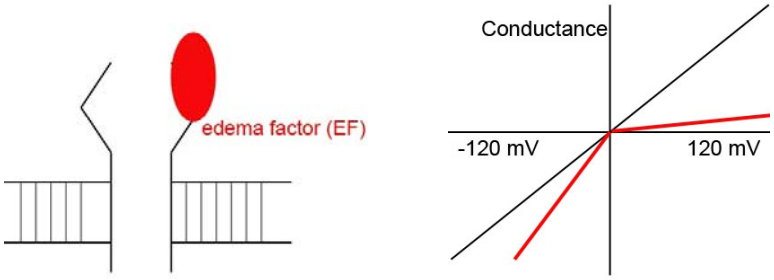

Fig 1-4: The protective antigen (PA) channel in the lipid bilayer and the conductance-voltage relationships before and after edema factor (EF) binding. The picture was modified from poster 3710, B571, biophysical meeting, 2012.

1.1.3 Hybrid nanopores

The biological nanopores offer an atomically precise structure and are available for genetic engineering, but solid state nanopores offer a range of sizes, durability, and structure control, and are also better suited for integration into devices. However, each system has its limitations: aHL relies on delicate lipid bilayers (or polymers) for mechanical support, while solid state nanopores are biological in-compatible and are difficult to modify, modification. Furthermore, the fabrication of solid-state nanopores with precise dimensions, especially in the nanometer range, still remains challenging.

Composing hybrid nanopores is a newly evolved research filed since 2007. Researchers are working on functionalizing solid state nanopores with coating biological compatible polymers, or inserting biological-forming pores into the solid state nanopores, which make the hybrid nanopores available for different applications (Venkatesan and Bashir, 2011).

1.1.3.1 Alpha-hemolysin solid state nanopore

Scientists with expertise on biological pore and solid state pore worked together and made the first hybrid pore. Dr. Dekker and his colleagues from Delft University of Technology collaborated with Dr. Bayley and his colleagues from the University of Oxford and they showed that a single, pre-assembled aHL protein pore can be inserted into a silicon nitride membrane with small holes, with a diameter of 2.4–3.6 nm (Hall et al., 2010). By attaching a dsDNA to the aHL, the protein pore can be threaded into the silicon nitride pores. And more exciting, this hybrid pore retains its function and they demonstrated the translocation of ssDNA through this hybrid pore. A schematic view of the hybrid pore is shown in Figure 1-5. The αHL pore was directed into the solid state pore by the attached dsDNA. Before αHL pore insertion, the pore current is at level a, then the dsDNA slides into the solid state pore decreasing the current to level b. Finally, a stable current level c was generated by the αHL pore inserted in the solid state pore. At all three current levels, DNA translocation events were observed.

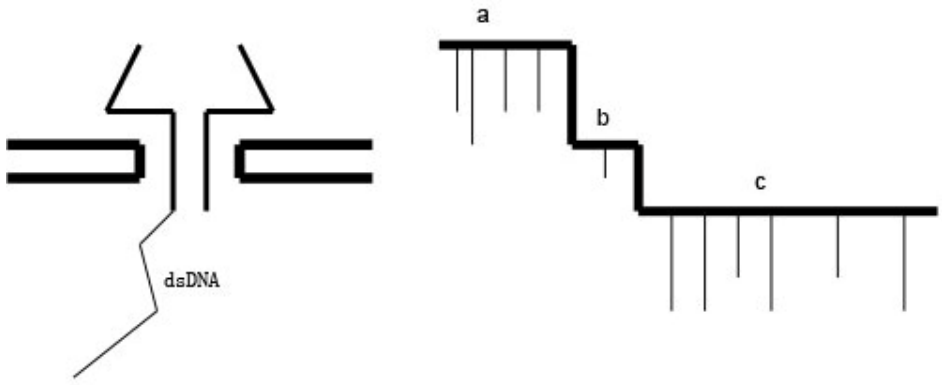

Figure 1-5: Schematic view of the alpha-hemolysin inserted into the solid state nanopore, and the DNA translocation events during the hybrid pore formation.

1.1.3.2 DNA Origami Nanopores

Inspired by the results of Hall et al. (2010) on aHL protein pore combined with the solid state nanopores, Drs. Keyser, Liedl and their colleagues from the University of Cambridge and Ludwig-Maximilians-Universitat Munchen developed a novel approach for the formation of hybrid nanopores. They demonstrated that the DNA origami structures can be inserted into the solid state nanopores to form the functional hybrid nanopores. DNA origami can repeatedly inserted into and ejected from solid state nanopores with diameters around 15 nm. These hybrid nanopores are capable of single molecule sensing (Bell et al., 2012).

1.1.3.3 Functionalized Solid Nanopore

In order to overcome the limitations of the solid state nanopore, many research groups are working on coating specific recognition sequences and receptors to the nanopore, or chemical functionalization of solid nanopores.

Dr. Bashir and his colleagues demonstrated that the SiO_2 nanopores can be functionalized with hairpin DNAs. They found different translocation time distributions between perfectly complementary ssDNA and ssDNA in which there was a single base mismatch, showing the potential of this approach to detect single-nucleotide polymorphisms (SNPs) (Iqbal et al., 2007).

Dr. Dekker and his colleagues constructed a biomimetic complex by covalently tethering either Nup98 or Nup153 (phenylalanine-glycine (FG) nucleoporins) to a solid-state nanopore, and these functionalized nanopores in

SiN can be used for the study of nucleocytoplasmic transport at the single-molecule level (Kowalczyk et al., 2011).

Dr. Mayer and his colleagues have shown that the sensitive detection and discrimination of proteins can be achieved by altering the surface chemistry in a nanopore. Inspired by the lipid-coated olfactory sensilla of insect antennae, SiN nanopores can be coated with a lipid bilayer and used to identify streptavidin proteins and antibodies (Yusko et al., 2011).

The incorporation of receptor or ligands into the bilayer allows the solid state nanopore to have chemical specificity, slows the translocation of target molecules, prevents pores from permanent blocks and decreass non-specific binding, thereby resolving many issues inherent to solid-state nanopores. A lipid-bilayer-coated nanopore architecture has been developed in both SiN (Yusko et al., 2011) and Al_2O_3 (Venkatesan et al., 2011) nanopores.

Another trend of solid state nanopore modification is metal coating of the solid state nanopore. Some groups are working on gold coated pores for using in single molecule sensing (Siwy et al., 2005; Wei et al., 2010a, 2010b). Furthermore, Dr. Martin and his colleagues found that these pores can be chemically modified by assembling a monolayer of alkane-thiols (SAM) on the gold surface of the pore interior (Sexton et al., 2007, 2010). For example, studies have shown that how the pore conductance is changing during the self-assembly of $HS\text{-}(CH_2)_{15}\text{-}(OCH_2CH_2)_3\text{-}OH$ molecules (Tinazli et al., 2005), or nitrilotriacetic acid which are used to chelate Ni^{2+} (NTAx, x = 1, 2, or 3). The structures are shown in Figure 1-6. The monolayer can be further coated through the S-S

connections and the Ni^{2+} loaded NTA receptor can specifically bind His6-tagged proteins (Sigal et al., 1996; Tinazli et al., 2005).

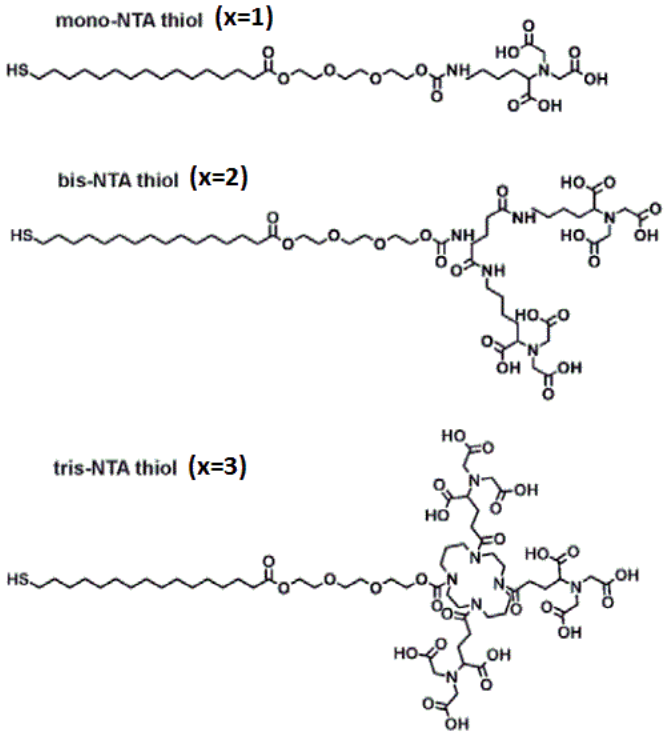

Figure 1-6: Structures of nitrilotriacetic acids.

1.2 Single Channel Recording Techniques

1.2.1 Patch clamp techniques

The Patch Clamp technique was developed by Erwin Neher and Bert Sakmann who received the Nobel Prize in 1991 for their discoveries concerning the function of single ion channels in cells (Neher et al., 1978; Neher, 1988). A patch-clamp microelectrode/pipette is placed next to a cell/neuron. In a patch

clamp experiment, gentle suction is applied through the microelectrode to draw a piece of the cell membrane (patch) into the microelectrode tip which forms a high resistance seal with the pipptte. This configuration is the "cell-attached" mode. At this time, when the micropipette is quickly withdrawn from the cell membrane, the patch of membrane still remains attached to the micropipette, and exposing the intracellular surface of the membrane to the external solutions. This forms the "inside-out patch" mode. This is helpful when one wishes to manipulate the environment at the intracellular surface of ion channels, such as the channels that are activated by intracellular ligand.

After the cell-attached mode is formed, when strong suction is applied as a pulse, the small patch of membrane in the electrode tip can be disrupted, leaving the electrode sealed to the rest of the cell and the solution in the microelectrode is connected with the cytoplasm solution, which forms the "whole-cell" mode. While the "perforated patch" technique tries to make small holes on the patch (when cell-attached mode formed) with pore-forming agents so that large molecules such as proteins can be kept inside the cell, but ions can pass through the holes freely.

After the whole-cell mode is formed, the micropipette can be withdrawn from the cell, allowing a patch of membrane to bleed out from the cell membrane, which can reform as a membrane on the tip. And the original outside of the cell membrane is facing outward. This forms the "outside-out patch" mode. This mode is very useful when the channel is activated from the extracellular side.

This mode allows the researcher to perfuse the same membrane patch with different solutions with different ligands.

Scientists Bert Sakmann and Erwin Neher edited the book "Single-Channel Recording". This book gives both theoretical background and practical advice regarding both experimental setup and data analysis. Whole-cell, single channels, and perforated patch are discussed. The practicalities of suitable patch glass, noise abatement, and capacitance cancelation are also covered.

1.2.2 Planar lipid bilayer recording

Planar lipid bilayers are used for functional studies of ion channel proteins using electrophysiological patch clamp techniques.

An important advance in biomembrane methodology was established by Wescott and his colleagues in 1962, when they reported a method for forming planar lipid bilayer membranes separating two aqueous phases (Mueller et al., 1962). A dispersion of artificial cell membranes (phospholipid or other surface-active lipid) in a nonpolar solvent (such as decane) is spread beneath an aqueous phase across an aperture several millimeters/micrometers in diameter drilled through a partition of some nonpolar material such as polyethylene (PE), polytetrafluoroethylene (PTFE), or polychlorotrifluoroethylene (PCTFE). The bilayer forms spontaneously. And this planar lipid bilayer can be used for biological studies such as Incorporation of rhodopsin (Montal and Korenbrot, 1973). A critique review of the formation of planar lipid bilayers and their electrical properties can be found in White et al. (1976).

Later, the capacitance of planar bilayer lipid membrane was studied with a patch clamp amplifier (Toyama et al., 1991), and single-channel recording studies were perform on the membrane reconstituted into a planar lipid bilayer (Kawahara et al., 1986). Voltage-gated ion channels in planar lipid bilayer membranes were also extensively studied (for review see (Latorre and Alvarez, 1981) as well the peptide/protein-lipid bilayer interactions (Jost and Griffith, 1980; Kagan et al., 1990),

In the planar lipid bilayer recording, the lipid bilayer composition is important for different ion channels to function normally. Studies have found that the function of membrane proteins varies with the thickness of the bilayer (Mitra et al., 2004; Yuan et al., 2004). The property of lipid headgroup is also important for the normal function of channel proteins. Studies have found that an anionic phospholipid is required for the KcsA potassium channel to remain its function in the artificial lipid bilayer (Valiyaveetil et al., 2002; Williamson et al., 2002, 2003).

In conclusion, the artificial planar lipid bilayer provides an alternative platform that allows the electrophysiological studies of ion channels.

1.3 References

Bayley, H., and L. Jayasinghe. 2004. Functional engineered channels and pores - (Review). *Molecular Membrane Biology*. 21:209-220.

Bell, N.A., C.R. Engst, M. Ablay, G. Divitini, C. Ducati, T. Liedl, and U.F. Keyser. 2012. DNA origami nanopores. *Nano Lett*. 12:512-517.

Blaustein, R.O., and A. Finkelstein. 1990. Voltage-dependent block of anthrax toxin channels in planar phospholipid bilayer membranes by symmetric tetraalkylammonium ions. Effects on macroscopic conductance. *The Journal of General Physiology*. 96:905-919.

Blaustein, R.O., E.J. Lea, and A. Finkelstein. 1990. Voltage-dependent block of anthrax toxin channels in planar phospholipid bilayer membranes by symmetric tetraalkylammonium ions. Single-channel analysis. *The Journal of General Physiology*. 96:921-942.

Branton, D., D.W. Deamer, A. Marziali, H. Bayley, S.A. Benner, T. Butler, M. Di Ventra, S. Garaj, A. Hibbs, X. Huang, S.B. Jovanovich, P.S. Krstic, S. Lindsay, X.S. Ling, C.H. Mastrangelo, A. Meller, J.S. Oliver, Y.V. Pershin, J.M. Ramsey, R. Riehn, G.V. Soni, V. Tabard-Cossa, M. Wanunu, M. Wiggin, and J.A. Schloss. 2008. The potential and challenges of nanopore sequencing. *Nature Biotechnology*. 26:1146-1153.

Butler, T.Z., M. Pavlenok, I.M. Derrington, M. Niederweis, and J.H. Gundlach. 2008. Single-molecule DNA detection with an engineered MspA protein nanopore. *Proc Natl Acad Sci U S A*. 105:20647-20652.

Dekker, C. 2007. Solid-state nanopores. *Nature Nanotechnology*. 2:209-215.

Derrington, I.M., T.Z. Butler, M.D. Collins, E. Manrao, M. Pavlenok, M. Niederweis, and J.H. Gundlach. 2010. Nanopore DNA sequencing with MspA. *Proc Natl Acad Sci U S A*. 107:16060-16065.

Ding, S., C. Gao, and L.Q. Gu. 2009. Capturing single molecules of immunoglobulin and ricin with an aptamer-encoded glass nanopore. *Analytical Chemistry*. 81:6649-6655.

Eren, E., J. Vijayaraghavan, J. Liu, B.R. Cheneke, D.S. Touw, B.W. Lepore, M. Indic, L. Movileanu, and B. van den Berg. 2012. Substrate Specificity within a Family of Outer Membrane Carboxylate Channels. *PLoS biology*. 10:e1001242.

Faller, M., M. Niederweis, and G.E. Schulz. 2004. The structure of a mycobacterial outer-membrane channel. *Science*. 303:1189-1192.

Fertig, N., C. Meyer, R.H. Blick, C. Trautmann, and J.C. Behrends. 2001. Microstructured glass chip for ion-channel electrophysiology. *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*. 64:409011-409014.

Garaj, S., W. Hubbard, A. Reina, J. Kong, D. Branton, and J.A. Golovchenko. 2010. Graphene as a subnanometre trans-electrode membrane. *Nature*. 467:190-193.

Gu, L.Q., and J.W. Shim. 2010. Single molecule sensing by nanopores and nanopore devices. *Analyst*. 135:441-451.

Guo, P.X., S. Erickson, and D. Anderson. 1987. A small viral RNA is required for in vitro packaging of bacteriophage phi 29 DNA. *Science*. 236:690-694.

Hall, A.R., A. Scott, D. Rotem, K.K. Mehta, H. Bayley, and C. Dekker. 2010. Hybrid pore formation by directed insertion of +Ý-haemolysin into solid-state nanopores. *Nature Nanotechnology*. 5:874-877.

Howorka, S., and Z. Siwy. 2009. Nanopore analytics: Sensing of single molecules. *In* Chemical Society Reviews. Vol. 38. 2360-2384.

Iqbal, S.M., D. Akin, and R. Bashir. 2007. Solid-state nanopore channels with DNA selectivity. *Nat Nanotechnol*. 2:243-248.

Jing, P., F. Haque, D. Shu, C. Montemagno, and P. Guo. 2010. One-way traffic of a viral motor channel for double-stranded DNA translocation. *Nano Lett*. 10:3620-3627.

Jost, P.C., and O.H. Griffith. 1980. Lipid-lipid and lipid-protein interactions in membranes. *Pharmacology, Biochemistry, and Behavior*. 13 Suppl 1:155-165.

Kagan, B.L., M.E. Selsted, T. Ganz, and R.I. Lehrer. 1990. Antimicrobial defensin peptides form voltage-dependent ion-permeable channels in planar lipid bilayer membranes. *Proc Natl Acad Sci U S A*. 87:210-214.

Kawahara, S., Y. Kirino, S. Nagao, and Y. Nozawa. 1986. Ion permeability of ciliary membrane vesicles isolated from Tetrahymena. Single-channel recording study on the membrane reconstituted into a planar lipid bilayer. *Journal of Biochemistry*. 100:1569-1573.

Keyser, U.F. 2011. Controlling molecular transport through nanopores. *Journal of the Royal Society Interface*. 8:1369-1378.

Kowalczyk, S.W., L. Kapinos, T.R. Blosser, T. Magalhaes, P. van Nies, R.Y. Lim, and C. Dekker. 2011. Single-molecule transport across an individual biomimetic nuclear pore complex. *Nat Nanotechnol*. 6:433-438.

Latorre, R., and O. Alvarez. 1981. Voltage-dependent channels in planar lipid bilayer membranes. *Physiological Reviews*. 61:77-150.

Li, J., D. Stein, C. McMullan, D. Branton, M.J. Aziz, and J.A. Golovchenko. 2001. Ion-beam sculpting at nanometre length scales. *Nature*. 412:166-169.

Liu, A., Q. Zhao, and X. Guan. 2010. Stochastic nanopore sensors for the detection of terrorist agents: Current status and challenges. *Analytica Chimica Acta*. 675:106-115.

Manrao, E.A., I.M. Derrington, M. Pavlenok, M. Niederweis, and J.H. Gundlach. 2011. Nucleotide discrimination with DNA immobilized in the MspA nanopore. *PloS one*. 6:e25723.

Mitra, K., I. Ubarretxena-Belandia, T. Taguchi, G. Warren, and D.M. Engelman. 2004. Modulation of the bilayer thickness of exocytic pathway membranes by membrane proteins rather than cholesterol. *Proc Natl Acad Sci U S A*. 101:4083-4088.

Mohammad, M.M., K.R. Howard, and L. Movileanu. 2011. Redesign of a plugged beta-barrel membrane protein. *J Biol Chem*. 286:8000-8013.

Montal, M., and J.I. Korenbrot. 1973. Incorporation of rhodopsin proteolipid into bilayer membranes. *Nature*. 246:219-221.

Mueller, P., D.O. Rudin, H.T. Tien, and W.C. Wescott. 1962. Reconstitution of cell membrane structure in vitro and its transformation into an excitable system. *Nature*. 194:979-980.

Neher, E. 1988. The use of the patch clamp technique to study second messenger-mediated cellular events. *Neuroscience*. 26:727-734.

Neher, E., B. Sakmann, and J.H. Steinbach. 1978. The extracellular patch clamp: a method for resolving currents through individual open channels in biological membranes. *Pflugers Archiv : European Journal of Physiology*. 375:219-228.

Rhee, M., and M.A. Burns. 2006. Nanopore sequencing technology: research trends and applications. *Trends in Biotechnology*. 24:580-586.

Sexton, L.T., L.P. Horne, S.A. Sherrill, G.W. Bishop, L.A. Baker, and C.R. Martin. 2007. Resistive-pulse studies of proteins and protein/antibody complexes using a conical nanotube sensor. *J Am Chem Soc*. 129:13144-13152.

Sexton, L.T., H. Mukaibo, P. Katira, H. Hess, S.A. Sherrill, L.P. Horne, and C.R. Martin. 2010. An adsorption-based model for pulse duration in resistive-pulse protein sensing. *J Am Chem Soc*. 132:6755-6763.

Shim, J.W., M. Yang, and L.Q. Gu. 2007. In vitro synthesis, tetramerization and single channel characterization of virus-encoded potassium channel Kcv. *FEBS letters*. 581:1027-1034.

Shu, D., Y. Shu, F. Haque, S. Abdelmawla, and P. Guo. 2011. Thermodynamically stable RNA three-way junction for constructing multifunctional nanoparticles for delivery of therapeutics. *Nat Nanotechnol*. 6:658-667.

Sigal, G.B., C. Bamdad, A. Barberis, J. Strominger, and G.M. Whitesides. 1996. A self-assembled monolayer for the binding and study of histidine-tagged proteins by surface plasmon resonance. *Anal Chem*. 68:490-497.

Siwy, Z., L. Trofin, P. Kohli, L.A. Baker, C. Trautmann, and C.R. Martin. 2005. Protein biosensors based on biofunctionalized conical gold nanotubes. *J Am Chem Soc*. 127:5000-5001.

Siwy, Z.S., and M. Davenport. 2010. Nanopores: Graphene opens up to DNA. *Nature Nanotechnology*. 5:697-698.

Siwy, Z.S., and S. Howorka. 2010. Engineered voltage-responsive nanopores. *Chemical Society Reviews*. 39:1115-1132.

Song, L.Z., M.R. Hobaugh, C. Shustak, S. Cheley, H. Bayley, and J.E. Gouaux. 1996. Structure of staphylococcal alpha-hemolysin, a heptameric transmembrane pore. *Science*. 274:1859-1866.

Storm, A.J., J.H. Chen, X.S. Ling, H.W. Zandbergen, and C. Dekker. 2003. Fabrication of solid-state nanopores with single-nanometre precision. *Nature Materials*. 2:537-540.

Tada, K., J. Haruyama, H.X. Yang, M. Chshiev, T. Matsui, and H. Fukuyama. 2011. Graphene magnet realized by hydrogenated graphene nanopore arrays. *Applied Physics Letters*. 99.

Tan, Q., J.W. Shim, and L.Q. Gu. 2010. Separation of heteromeric potassium channel Kcv towards probing subunit composition-regulated ion permeation and gating. *FEBS Letters*. 584:1602-1608.

Tinazli, A., J. Tang, R. Valiokas, S. Picuric, S. Lata, J. Piehler, B. Liedberg, and R. Tampe. 2005. High-affinity chelator thiols for switchable and oriented immobilization of histidine-tagged proteins: a generic platform for protein chip technologies. *Chemistry*. 11:5249-5259.

Toyama, S., A. Nakamura, and F. Toda. 1991. Measurement of voltage dependence of capacitance of planar bilayer lipid membrane with a patch clamp amplifier. *Biophys J*. 59:939-944.

Valiyaveetil, F.I., Y. Zhou, and R. MacKinnon. 2002. Lipids in the structure, folding, and function of the KcsA K+ channel. *Biochemistry*. 41:10771-10777.

Venkatesan, B.M., and R. Bashir. 2011. Nanopore sensors for nucleic acid analysis. *Nat Nanotechnol*. 6:615-624.

Venkatesan, B.M., J. Polans, J. Comer, S. Sridhar, D. Wendell, A. Aksimentiev, and R. Bashir. 2011. Lipid bilayer coated Al(2)O(3) nanopore sensors: towards a hybrid biological solid-state nanopore. *Biomedical Microdevices*. 13:671-682.

Wei, R., D. Pedone, A. Zurner, M. Doblinger, and U. Rant. 2010a. Fabrication of metallized nanopores in silicon nitride membranes for single-molecule sensing. *Small*. 6:1406-1414.

Wei, R., D. Pedone, A. Zurner, M. Doblinger, and U. Rant. 2010b. Nanopores: Fabrication of Metallized Nanopores in Silicon Nitride Membranes for Single-Molecule Sensing (Small 13/2010). *Small*. 6.

Wendell, D., P. Jing, J. Geng, V. Subramaniam, T.J. Lee, C. Montemagno, and P. Guo. 2009. Translocation of double-stranded DNA through membrane-adapted phi29 motor protein nanopores. *Nat Nanotechnol*. 4:765-772.

White, S.H., D.C. Petersen, S. Simon, and M. Yafuso. 1976. Formation of planar bilayer membranes from lipid monolayers. A critique. *Biophys J*. 16:481-489.

Williamson, I.M., S.J. Alvis, J.M. East, and A.G. Lee. 2002. Interactions of phospholipids with the potassium channel KcsA. *Biophys J*. 83:2026-2038.

Williamson, I.M., S.J. Alvis, J.M. East, and A.G. Lee. 2003. The potassium channel KcsA and its interaction with the lipid bilayer. *Cellular and Molecular Life Sciences : CMLS*. 60:1581-1590.

Yuan, C., R.J. O'Connell, P.L. Feinberg-Zadek, L.J. Johnston, and S.N. Treistman. 2004. Bilayer thickness modulates the conductance of the BK channel in model membranes. *Biophys J*. 86:3620-3633.

Yusko, E.C., J.M. Johnson, S. Majd, P. Prangkio, R.C. Rollings, J. Li, J. Yang, and M. Mayer. 2011. Controlling protein translocation through nanopores with bio-inspired fluid walls. *Nat Nanotechnol*. 6:253-260.

CHAPTER 2

CHARACTERIZATION OF DSDNA UNWINDING IN THE NANOPORE

2.1 Introduction

The nanopore is a receptive single-molecule detector for broad biotechnological applications (Maglia et al., 2010; Bayley and Cremer, 2001; Bayley and Jayasinghe, 2004; Gu and Shim, 2010; Howorka and Siwy, 2009; Ma and Cockroft, 2010; Movileanu, 2009). The ion current through a nanopore is sensitive to the target molecules that occupy the pore lumen, therefore from characteristic change in the nanopore current, different molecular states can be electrically identified and the kinetic pathway in a single molecular reaction can be elucidated. The nanopore is being developed as a rapid, label-free and low-cost technology for DNA sequencing (Bayley, 2006; Branton et al., 2008; Kasianowicz et al., 1996). Toward this goal, single-molecule nucleic acids and their interaction with nanopore has been extensively characterized (Bockelmann and Viasnoff, 2008; Dudko et al., 2007; Kasianowicz et al., 1996; Mathe et al., 2006; McNally et al., 2008; Muzard et al., 2010; Sauer-Budge et al., 2003; Sutherland et al., 2004; Vercoutere et al., 2001, 2003; Zhao et al., 2008; Ashkenasy et al., 2005; Mitchell and Howorka, 2008; Stoddart et al., 2009; Howorka et al., 2001b), which includes many topics such as how the nanopore conductance is sensitively changed with the sequence of a single-stranded DNA

or RNA in the pore and how a double-stranded DNA is unzipped driven by voltage across the pore.

Understanding molecular behaviors of nucleic acids in the nanopore not only gives insight into their important biophysical mechanisms, but is also beneficial for biosensor development (Howorka et al., 2001a, 2001b; Howorka and Bayley, 2002). We recently proposed a robust nanopore method of differentiating and quantifying cancer-associated microRNAs (miRNAs) in human blood samples, an approach with the potential in non-invasive and cost-effective cancer detection (Wang et al., 2010). For clinical detection, the circulating miRNAs were hybridized with their specific DNA probes. The hybrid was trapped in the α-hemolysin (αHL) protein pore, followed by unzipping driven by the transmembrane voltage. The key to this sensor was a s*ignature current pattern* that acts as an unzipping marker. By recognizing signature signals, single miRNA molecules in blood can be identified. Therefore, it is a priority to characterize the unzipping signature, which determines both selectivity and sensitivity of the sensor.

Unzipping of various double-stranded DNAs (dsDNAs) has been detected using the nanopore. When a short (3-9 bps) blunt-ended hairpin was unzipped in the αHL pore, it generated a “shoulder-spike” current block. By measuring the block duration, single mismatches in the hairpin stem can be discriminated (Vercoutere et al., 2001, 2003). When a dsDNA with a overhang (hairpin (Sutherland et al., 2004; Mathe et al., 2006; Dudko et al., 2007; Bockelmann et al., 2008; McNally et al., 2008; Zhao et al., 2008; Muzard et al., 2010) or blunt

end (Sauer-Budge et al., 2003) at the other terminal) carrying an overhang was trapped in the pore, the transmembrane voltage can pull the overhang that occupies the β-barrel, driving the unzipping of the double-stranded domain. By programming the voltage profile and analyzing the long block duration, the unzipping kinetics has been understood (Sutherland et al., 2004; Mathe et al., 2004, 2006; Dudko et al., 2007; Bockelmann et al., 2008; McNally et al., 2008; Zhao et al., 2008; Muzard et al., 2010; Sauer-Budge et al., 2003). The dsDNA with one strand covalently tethered to the lumen of nanopore can also be unzipped. In this configuration, the electrical force pulled the complementary strand, tearing it off from the tethered strand (Howorka et al., 2001b). Although the nucleic acids unzipping, as an important biological process, has been widely investigated using the nanopore single molecule approach, the signature current pattern we have identified for miRNA detections has not been reported previously.

Given the importance in biosensor construction, we are motivated to elucidate the molecular mechanisms behind the unzipping signatures we have identified. In this report, we designed a series of dsDNA carrying overhangs of different lengths. These DNAs were found to generate multiple characteristic current patterns, each being attributed to a specific single DNA configuration in the nanopore. Most notably, the characteristic multi-current long blocks clearly revealed the sequential steps in an unzipping procedure. The unzipping signatures enabled us to track when the unzipping occurred after the DNA is trapped in the pore, and what is the motion pathway of the unzipped single-stranded DNA (ssDNA). The signature signals also allowed us to identify whether

a trapped dsDNA is unzipped or returns to the *cis* solution without unzipping, whether the dsDNA is trapped in the "overhang-in" (Figure 2-2 a,b,c) or "blunt end-in" (Figure 2-2 d) orientation, as well as how the unzipping efficiency and DNA trapping directionality is controlled by the length and sequence of overhangs.

2.2 Materials and Methods

The electrophysiology setup and methods for nanopore experiments have been detailed elsewhere (Shim and Gu, 2008; Shim et al., 2009). Briefly, the recording apparatus was composed of two chambers (*cis* and *trans*) that were partitioned with a Teflon film. The planar lipid bilayer of 1,2-diphytanoyl-sn-glycerophosphatidylcholine (Avanti Polar Lipids) was formed spanning a 100-150 nm hole in the center of the partition. Both *cis* and *trans* chambers were filled with symmetrical 1 M KCl buffered with 10 mM Tris and titrated to pH 8.0. All solutions are filtered before use. Single α-hemolysin proteins were inserted into the bilayer from the *cis* side to form molecular pores. DNA oligonucleotides including 3'-biotinlated DNA (Table 2-1) were synthesized and electrophoresis-purified by Integrated DNA Technologies, IA. Before testing, the mixtures of ssDNAs were heated to 90°C for 5 minutes, then gradually cooled down to room temperature and stored at 4°C.

In the single channel recording, the *cis* solution was grounded and the voltage was applied from the *trans* solution, so that a positive voltage would drive the translocation of a negatively charged DNA through the pore from *cis* to *trans*.

Single-channel currents were recorded with an Axopatch 200A patch-clamp amplifier (Molecular Device Inc., formerly Axon Inc.), filtered with a built-in 4-pole low-pass Bessel Filter at 5 kHz, and acquired with Clampex 9.0 software (Molecular Device Inc.) through a Digidata 1332 A/D converter (Molecular Device Inc.) at a sampling rate of 20 kHz•s^{-1}. The data were analyzed using Clampfit 9.0 (Molecular Device Inc.), Excel (MicroSoft) and SigmaPlot (SPSS) software. Because the ssDNA translocation events (~100us μs (Meller et al., 2000, 2001) can be well distinguished from the signature blocks for trapping a dsDNA (10^0-10^3 ms), we used 1 ms as the boundary. Blocks shorter than 1 ms were counted as the passage of ssDNAs in the pore, while blocks longer than 1 ms as the events of dsDNA. The poly(dC) was chosen as the DNA overhangs, because earlier studies have shown that poly(dC) has the shortest translocation time (Meller et al., 2000; Chen and Li, 2007) with the least possibility to form coiled structures (Chen and Li, 2007; Muthukumar, 1999) compared to other oligonucleotides. According to previous studies (Meller et al., 2000, 2001; Shim and Gu, 2008), there are two translocation times that can be measured, the peak time and the exponentially decaying time constant, from the duration histogram of short-lived blocks. The peak time marks the most probable translocation time and the decaying constant reflects the temporal dispersion of the translocation. In this report, we simply measured the translocation duration by averaging all short events. The duration of signature blocks (long-lived blocks) was obtained by fitting the dwell-time histogram to an exponential distribution or by averaging all the long events in cases there were not enough events for accurate fitting. Data

were given as the mean ± SD, based on at least four separate experiments. The electrophysiology experiments were conducted at 22 ± 2°C.

2.3 Results and Discussion

2.3.1 Characteristic current patterns formed by the dsDNA with an overhang.

We first employed DNA C30 as the model to identify characteristic current patterns. As shown in Table 2-1, the double-stranded domain of C30 contains 22 base pairs and adopts the sequence of the miRNA *miR-155*, a potential biomarker in lung cancer (Rabinowits et al., 2009; Rosell et al., 2009a). The overhang of C30 comprises 30 deoxycytidines, which acts as a signal tag attached to the 3'-terminal of the anti-sense stand.

C30 presented to the *cis* side of the αHL pore generated both short- and long-lived current blocks (Figure 2-1). Monitored at +100 mV in 1 M KCl, the spike-like short blocks lasted for 200±10 μs on average and almost fully reduced the pore conductance (Figure 2-1a). As both the duration and conductance are similar to that of blocks formed by either strand of C30 alone, these short blocks should be attributed to the rapid passage of un-hybridized ssDNAs in the pore (Meller et al., 2000, 2001).

Table 2-1. Sequences of dsDNAs utilized in this research.

dsDNA	Sequence
C0	5'-TTAATGCTAATCGTGATAGGGG-3'[a] 3'-AATTACGATTAGCACTATCCCC-5'
C8	5'-TTAATGCTAATCGTGATAGGGG-3' 3'-$(C)_8$AATTACGATTAGCACTATCCCC-5'
C12	5'-TTAATGCTAATCGTGATAGGGG-3' 3'-$(C)_{12}$AATTACGATTAGCACTATCCCC-5'
C20	5'-TTAATGCTAATCGTGATAGGGG-3' 3'-$(C)_{20}$AATTACGATTAGCACTATCCCC-5'
C30	5'-TTAATGCTAATCGTGATAGGGG-3' 3'-$(C)_{30}$AATTACGATTAGCACTATCCCC-5'
HP-C30	5'-TTAATGCTAATCGTGATAGGGGCGAGACAACGCTCTCTCGTTGTCTCG-3'[b] 3'-$(C)_{30}$AATTACGATTAGCACTATCCCC-5'
SA-C30	5'-TTAATGCTAATCGTGATAGGGGAAAAAA-3'-biotin-steptavidin 3'-$(C)_{30}$AATTACGATTAGCACTATCCCC-5'

[a]: The sequence of the sense strand originated from that of microRNA *miR155*.

[b]: The sequence of 3'-hairpin domain came from a previous report (Mathe et al., 2004) for a better comparison of the results.

We focused on the long block characterization. The long blocks persisted for 430±16ms on average and should be formed by the dsDNA C30 as they did not appear when either strand of C30 alone was present. The long blocks can be classified into four types based on the current patterns (Figure 2-1b through e). Pattern I was characterized by three sequential current levels, Level 1 → Level 2 → Level 1' (Figure 2-1b). Level 1 lasted for most of the block duration and almost fully reduced the conductance to 110±6 pS. At the end of Level 1, the current was discretely promoted to Level 2 at 490±21 pS for 810±83 μs; then dropped again to level 1' with similar conductance to level 1. The current stayed at level 1' for 180±20 μs before resuming to the full open state. The Pattern II (Figure 2-1c)

was identical to Pattern I blocks in Level 1 and Level 2 profiles. However, there was no Level 1' in Pattern II. Instead, the current directly recovered to the full open state from Level 2, forming an "end-shoulder" pattern. In addition to multi-conductance blocks, we also identified two types of single-conductance blocks, Pattern III (Figure 2-1d) and IV (Figure 2-1e). The Pattern III blocks showed identical conductance to Level 1 identified in Pattern I and II, whereas the Pattern IV blocks reduced the current to the unique Level 3 at 200±36 pS, which is significantly higher than the Level 1 conductance of all other current patterns (Figure 2-1b, c and d).

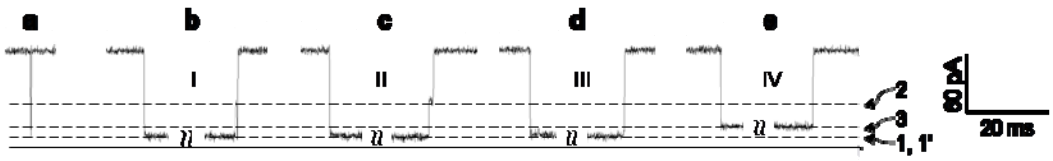

Figure 2-1: Multiple characteristic current patterns in the αHL pore generated by dsDNA with an overhang. The DNA in this detection was C30 (Table 1). **a.** Spike-like short blocks; **b** through **e**. Long blocks classified as Pattern I (d), II (c), III (d) and IV (e).

2.3.2 Specific molecular procedures and configurations revealed by current patterns.

We proposed that Pattern I is a signature of the following molecular procedure: as shown in Figure 2-2a, Level 1 is formed by a C30 captured in the pore with its overhang trapped in the β-barrel where the voltage pulls the overhang, inducing unzipping of the double-stranded domain of C30. Upon unzipping, the long strand carrying the overhang is thread through the pore from the *trans* mouth, whereas the short one temporarily halts in the wider nanocavity

of the pore, increasing the current to Level 2. When the short strand finally traverses the β-barrel, it reduces the pore current to Level 1'.

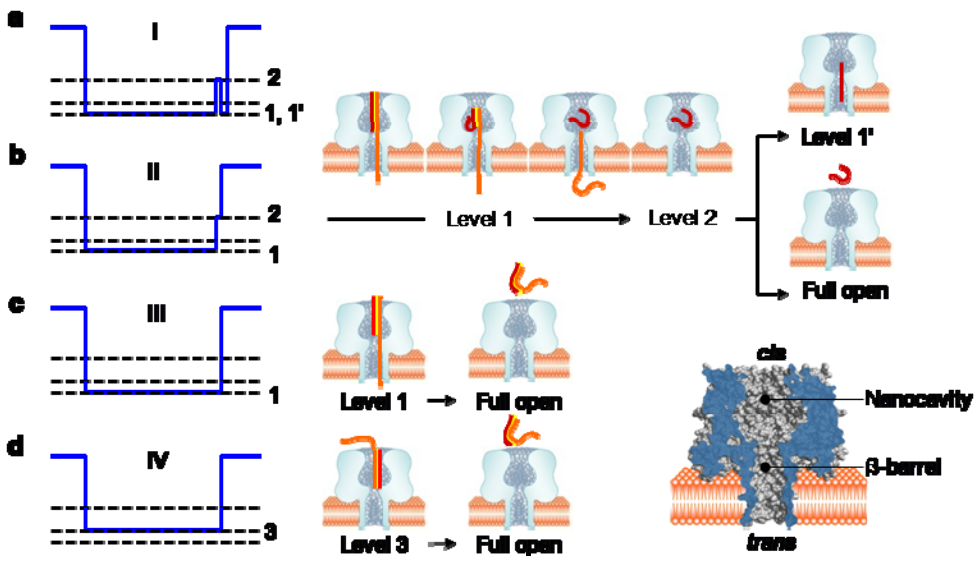

Figure 2-2. Models showing DNA configurations in the nanopore revealed by signature current patterns. Left, current patterns idealized from recordings in Figure 2-1; right, molecular configurations and processes. **a.** Pattern I, unzipping of dsDNA and translocation of the short strand after unzipping; **b.** Pattern II, unzipping of dsDNA and release of the short strand back to the *cis* solution; **c.** Pattern III, return of dsDNA to the *cis* solution without unzipping; and **d.** dsDNA trapped in the opposite orientation and released to the *cis* solution without unzipping.

The occupancy of β-barrel is the key to discriminate this molecular procedure. The low conductance of Level 1 supported the configuration that a DNA complex such as hairpin (Mathe et al., 2004) and G-quadruplex (Shim and Gu, 2008) is trapped in the pore with the overhang occupying the β-barrel. In contrast, the partially blocked Level 2 is consistent with a configuration that an oligonucleotide resides in the nanocavity, rather than in the β-barrel (Maglia et al., 2008; Shim et al., 2009). The similar low conductance of Level 1' and Level 1 suggested that Level 1' is also associated with a ssDNA in the β-barrel.

Table 2-2: Voltage-dependent durations of Level 1, Level 2 and Level 1' in the unzipping signature current pattern (Pattern I).

Voltage (mV)	T_{L1} (ms)	T_{L2} (ms)	$T_{L1'}$ (ms)
+100	432±16	0.81±0.08	0.18±0.06
+150	27±5	0.89±0.12	0.13±0.05
+180	13±2	1.1±0.1	0.12±0.02

The model for Pattern I was further evidenced by the voltage-dependent duration of each current level, τ_{L1}, τ_{L2} and $\tau_{L1'}$ (Table 2-2). τ_{L1} is comparable to the time scales for DNA unzipping in the nanopore in earlier studies. For example, it needed ~440 ms to unzip a 50 bps dsDNA at +140 mV (Sauer-Budge et al., 2003; Shim et al., 2009), and ~40 ms for a 10 bps DNA hairpin at +90 mV (Mathe et al., 2004). τ_1 can be greatly shortened by 35 fold from 430 ms to 13 ms as the voltage increases from +100 mV to +180 mV, in agreement with a voltage-enhanced unzipping process (Meller et al., 2001). Compared with τ_{L1}, τ_{L2} slightly varies between 0.81-1.1 ms in the same voltage range. The independence of τ_{L2} to voltage is consistent with a ssDNA in the nanocavity because the voltage drop in this domain is small (10% of the applied voltage) (Howorka and Bayley, 2002). This is also consistent with our earlier finding that the unfolding of a G-quadruplex DNA aptamer encapsulated in the nanocavity is independent to the voltage (Shim and Gu, 2008). $\tau_{L1'}$ becomes voltage-dependent again. It can be shortened from 180±20 μs to 120±20 μs by increasing the voltage from +100 mV to +180 mV, supporting the voltage-dependent translocation of a single-stranded oligonucleotide (Meller et al., 2001).

Because Level 2 and Level 1’ mark different configurations of the unzipped short strand of C30, it would be possible to regulate the properties of both current levels by altering the short strand structure. This hypothesis was verified through the following tests. When employing a modified C30, HP-C30 with a hairpin at the 3’-end of short strand (Table 2-1), we observed a novel type of three-level current pattern (Figure 2-3a). Its Level 1 and Level 2 were similar to C30’s Pattern I blocks in both conductance and duration (Figure 2-2a), consistent with the unzipping of HP-C30 and transference of the unzipped short strand from the nanocavity to the β-barrel. However, the duration of Level 1’ was drastically prolonged by 80 folds to 15±1.9 ms, compared to that of C30’s Pattern I blocks (Figure1-1b). The prolonged Level 1’ is in agreement with the unzipping of hairpin prior to threading in the β-barrel. When using another modified C30, SA-C30 attached with a streptavidin at the 3’-end of the short strand (Table 2-1), we also observed a new multi-level current pattern (Figure 2-3b). Again, both the fully-blocked Level 1 and partially-blocked level 2 were similar to those in C30’s Pattern I blocks. However, the current was fixed at Level 1’ until it was forced to recover by a negative voltage. The “permanent” Level 1’ can be interpreted by that although the short strand of SA-C30 moves into the β-barrel after unzipping, its translocation is prevented by the attached large streptavidin. Overall, both detections with hairpin- and streptavidin-modified validated the unzipping model (Figure 2-2a) proposed for Pattern I blocks (Figure 2-1b).

As the Pattern I blocks have elucidated, it became straightforward to understand the molecular procedures in other current patterns. The absence of Level 1' in Pattern II ("end-shoulder" profile) can be explained by the short strand DNA confined in the nanocavity directly returns to the *cis* solution (Figure 2-2b), rather than traversing the β-barrel. Therefore, both Pattern I and II are *signature current patterns* that reveal the entire unzipping procedure including the movement of unzipped ssDNAs. Since Pattern III blocks feature a single current level, and its conductance is identical to Level 1 of both Pattern I and II blocks, we propose that this type of blocks is formed by the release of trapped C30 from the *cis* entry without unzipping (Figure 2-2c). The Pattern IV blocks with much higher Level 3 conductance is due to C30 in the opposite orientation with the blunt end in the nanocavity. In this orientation, the β-barrel was unoccupied, giving higher conductance (Figure 2-2c). Similar configurations and resulting partial blocks have been reported for short blunt-ended hairpins in the nanocavity (Vercoutere et al., 2003), short dsDNA covalently tethered to the nanocavity (Howorka et al., 2001b) and the G-quadruplex aptamer encapsulated in the nanocavity (Shim and Gu, 2008; Shim et al., 2009). Because no additional current change was observed following Level 3, we expected that C30 in this trapping orientation does not undergo unzipping (Mathe et al., 2005b; Sauer-Budge et al., 2003).

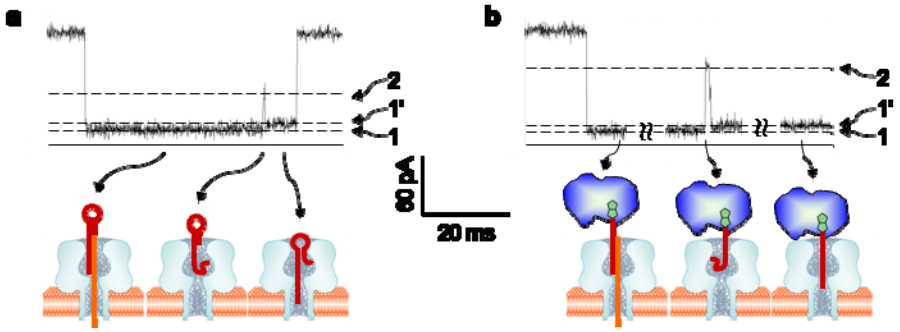

Figure 2-3: Current patterns and configurations for unzipping of HP-C30 and SA-C30 with engineered short strands. The sequences of the two DNAs were given in Table 1. **a,** HP-C30, C30 with a hairpin extended to the short strand 3'-end; and **b,** SA-C30, C30 with a streptavidin attached to the biotinylated 3'-end of the short strand.

2.3.3 Characterization and regulation of DNA unzipping

Discrimination of signature current patterns provides a quantitative method for characterizing, and further regulating the unzipping of nucleic acids by factors such as the overhang length, voltage, and trapping directionality. To test this hypothesis, we designed a series of dsDNAs that possess a common double-stranded complex attached with different overhangs. As shown in Table 2-1, these DNAs included C0, C8, C12, C20 and C30 (studied above). Their overhangs comprised 0, 8, 12, 20 and 30 deoxycytidines, respectively at the 3'-terminal of the anti-sense strand. Figure 2-4a illustrated one configuration of these DNAs in the pore in which their overhangs occupy the β-barrel to different extents.

C8, C12 and C20 were found to generate the same Pattern I, II, III and IV current blocks as C30 shown in Figure 2-1. C0 was different. We did not observe the Pattern I and II unzipping blocks for this DNA. Figure 2-4b analyzed the Level 1 and Level 3 conductance of these current patterns. The Level 1 conductance for DNAs carrying an overhang (C8 through C30) was almost identical, ranging between 99-110 pS. This character indicated that their overhangs, long or short, are trapped in the β-barrel. Notably the overhang of C8 (~2.6 nm) is shorter than the distance between the *cis* opening and the top of the β-barrel (~4 nm). Thus, in order for the overhang to enter the β-barrel, the double-stranded domain in the DNA must be trapped in the nanocavity, rather than anchored at the *cis* opening (Sauer-Budge et al., 2003). The Level 3 conductance for these DNAs slightly varied between 200-260 pS, and was significantly higher than the Level 1 conductance, suggesting that these DNAs can also be trapped in the opposite orientation with the blunt end entering the pore (Figure 2-2d). C0 only generated two single-level current patterns at 280±11 pS and 330±10 pS. Both conductance values were similar to that of Level 3 conductance for DNAs with an overhang, thus, they are probably associated with two trapping orientations, each with a blunt end trapped in the nanocavity.

Identification of signature current patterns allows for characterizing the fraction of each type of blocks. The calculation of fractions of different types of blocks is described below:

The fractions of Pattern I and II ($F_{I,II}$, unzipping), Pattern III (F_{III}) and Pattern IV (F_{IV}) blocks, were measured from the ratio of the numbers of Level 1 (N_{L1}) and Level 3 (N_{L3}) blocks,

$$a = N_{L1} / N_{L3} \quad \text{(S1)}$$

and the ratio of the numbers of unzipping (N_{unzip}) and no-unzipping ($N_{no\text{-}unzip}$) blocks,

$$b = N_{unzip} / N_{no-unzip} \quad \text{(S2)}$$

a and *b* can be expressed using $F_{I,II}$, F_{III} and F_{IV},

$$a = (F_{I,II} + F_{III}) / F_{IV} \quad \text{(S3)}$$

$$b = F_{I,II} / (F_{III} + F_{IV}) \quad \text{(S4)}$$

From Eq. S3, Eq. S4 and the condition $F_{I,II} + F_{III} + F_{IV} = 1$, we obtained

$$F_{IV} = 1/(1+a)$$

$$F_{III} = 1/(1+b) - 1/(1+a) \quad \text{(S5)}$$

$$F_{I,II} = b/(1+b)$$

Eq. S5 was used to calculate $F_{I,II}$, F_{III} and F_{IV} in Figure 2-5. Values of *a* and *b* for the DNAs with different overhangs were given in Table 2-3.

Table 2-3. Measured *a* and *b* for the DNAs with different overhangs.

	C0	C8	C12	C20	C30
$a = N_{L1} / N_{L2}$	1.16	2.4±0.3	3.2±0.4	5.7±0.5	7.6±0.6
$b = N_{unzip} / N_{non-unzip}$	0	0.10±0.02	0.44±0.05	1.2±0.1	1.7±0.2

As shown in Figure 2-4c, the total fraction of Pattern I and II unzipping blocks ($F_{I,II}$) progressively increased from 0 for C0 to 0.63 for C30, whereas the fraction of Pattern III (F_{III}) continuously decreased from 0.54 for C0 to 0.25 for C30 as the overhang lengthened. The $F_{I,II}$ and F_{III} results indicated an overhang length-determined unzipping process. A longer overhang carries more charges, thus it can be pulled more intensively than a shorter overhang under the same electrical field in the pore (once the β-barrel is "full", then change in field is the same, however, longer length represents entropic barrier to exit without unzipping). As a result, DNA with a longer overhang should demonstrate higher occurrence of unzipping ($F_{I,II}$) and, in turn, a lower occurrence of diffusive escape without unzipping (F_{III}). Figure 2-4c also shows that a contiguous decrease in the fraction of Pattern IV blocks (F_{IV}) with the overhang length. This character suggested that elongating overhang reduces the trapping occurrence in the "blunt end-in" orientation, and in turn increases the occurrence in the "overhang-in" orientation. The orientation preference can be explained as that the single-stranded overhang is narrower and more flexible than the double-stranded blunt end, thus being more favorable to the nanopore with a lower entrance resistance (Hille, 2001). C0 is a dsDNA with both blunt ends. According to Figure 2-4c, C0's F_{III} (0.55) and F_{IV} (0.45) are similar, indicating that both blunt ends of C0 enter the pore with equal probability and there is no preference in trapping directionality. C0's $F_{I,II}$ was 0, meaning that this DNA can not be unzipped in either trapping orientation. We noticed that the blunt-ended short DNA hairpins can be unzipped in previous studies (Vercoutere et al., 2001). The different observations between

the two studies may originate from the fact that C0 contains more base pairs (22 bps) than the short hairpins (3-9 bps) used in the other study (Vercoutere et al., 2001); thus, C0 should feature a longer unzipping time. When trapped in the nanocavity, the much weaker electrical field in this pore domain (10% of applied voltage) will cause a higher probability for C0 to diffusively escape before unzipping.

Characteristic current patterns have indicated two pathways for the dsDNA trapped in the nanopore, unzipping or returning to the *cis* solution (Figure 2-2). If P_{unzip} and P_{return} represent the unzipping and returning probabilities, their total should be 1. Since $F_{I,II}$ is the fraction of total unzipping events and F_{III} is the fraction of returning events, P_{unzip} can be expressed as

$$P_{unzip} = \frac{F_{I,II}}{F_{I,II} + F_{III}} \quad (1)$$

Figure 2-4d shows P_{unzip} values calculated from the fraction data in Figure 2-4c. We found that P_{unzip} contiguously increased as the overhang elongated. The trend of P_{unzip} can be fitted with the Boltzmann's distribution,

$$P_{unzip} = \frac{e^{-z(n-n_{0.5})FV/RT}}{1+e^{-z(n-n_{0.5})FV/RT}} \quad (2)$$

In this equation V is the voltage drop over the β-barrel, which is approximated as the applied transmembrane voltage (+100 mV). z is the charge valence in a nucleotide, n is the number of deoxycytidines in an overhang, thus zn represents the total charges carried by the overhang. $n_{0.5}$ is the number of deoxycytidines at which the unzipping and returning probabilities are equal, i.e. $P_{unzip} = P_{return} = 0.5$

when $n = n_{0.5}$. The P_{unzip} values between C0 and C20 was fitted using Eq. 2. z was fitted to be -0.077, comparable to -0.1 in a previous study (Sauer-Budge et al., 2003). The low charge valence is consistent with the high salt concentration (1 M KCl) in recording solutions that effectively screen the negative charge on DNA phosphates. $n_{0.5}$ was fitted to be 15.6, meaning that 15.6 deoxycytidines are needed in the overhang to drive the unzipping with a probability of P_{unzip} = 0.5 at +100 mV. Under this condition, the driving force $zn_{0.5}V/l_{\beta}$ for unzipping the double-stranded domain (22 base pairs) is 3.2 pN, where l_{β} is the length of β-barrel (~6 nm (Song et al., 1996)). This force level is lower than theoretically predicted ~12 pN as the unzipping threshold force in the solution (Cocco et al., 2001), suggesting it needs less force to unzip a DNA confined in the nanopore. Figure 2-4d also showed that the measured P_{unzip} for C30 (0.71) was far lower than the fitted value (0.94) using DNAs with shorter overhangs. This is because the overhang of C30 ($d(C)_{30}$, ~9.6 nm) is sufficiently long enough to occupy the entire β-barrel (~6 nm). Consequently, the driving force reaches its maximum and the unzipping probability stops increasing.

The voltage is another important unzipping regulator. We have measured that the ratio of total unzipping and non-unzipping blocks, $F_{I,II}/F_{III,IV}$, for C30 increases from 1.7 to 3.7 as the voltage leveled up from +100 mV to +180 mV (Figure 2-5a). This voltage-dependent block ratio indicated that the voltage enhances the unzipping occurrence. We have known that the trapped DNA is either unzipped or returns to the *cis* solution. If k_{unzip} and k_{return} are rate constants for the two reactions, their ratio can be calculated as the ratio of fractions between unzipping

and returning events, $k_{unzip} / k_{return} = F_{I,II} / F_{III}$. The rate constants are also related to the long block duration *τ*, $\tau = 1/(k_{unzip} + k_{return})$. This is because in a kinetic pathway, the lifetime in any single state is the inverted sum of transition rates that lead away from the state. Overall, k_{unzip} and k_{return} can be calculated by

$$k_{unzip} = F_{I,II} /[\tau(F_{I,II} + F_{III})] \text{ and } k_{return} = F_{III} /[\tau(F_{I,II} + F_{III})] \qquad (3)$$

Figure 2-5 showed calculated k_{unzip} using C30 as the target. k_{unzip} was 1.5 s^{-1} at +100 mV, 25 s^{-1} at +150 mV and 61 s^{-1} at +180 mV, indicating the voltage-enhanced unzipping rate constant for this. The voltage-dependent k_{unzip} follows the Woodhull's equation for a binary reaction (two-states),

$$k_{unzip}(V) = k_{unzip}(0)e^{-znVF/RT} \qquad (4)$$

We used Eq. 4 to fit the k_{unzip} data. The fitted *zn* was -1.21. Supposing the entire β-barrel (6 nm) can accommodate 8-12 nucleotides, i.e. *n*=8-12, the charge valence *z* should be 0.1-0.15. This *z* value is in agreement with the -0.077 calculated from the overhang length-dependent unzipping probability (Figure 2-5b). By extending the fitting curve to 0 mV, we obtained the unzipping rate constant in the absence of driving force, $k_{unzip}(0)$=0.014 s^{-1}.

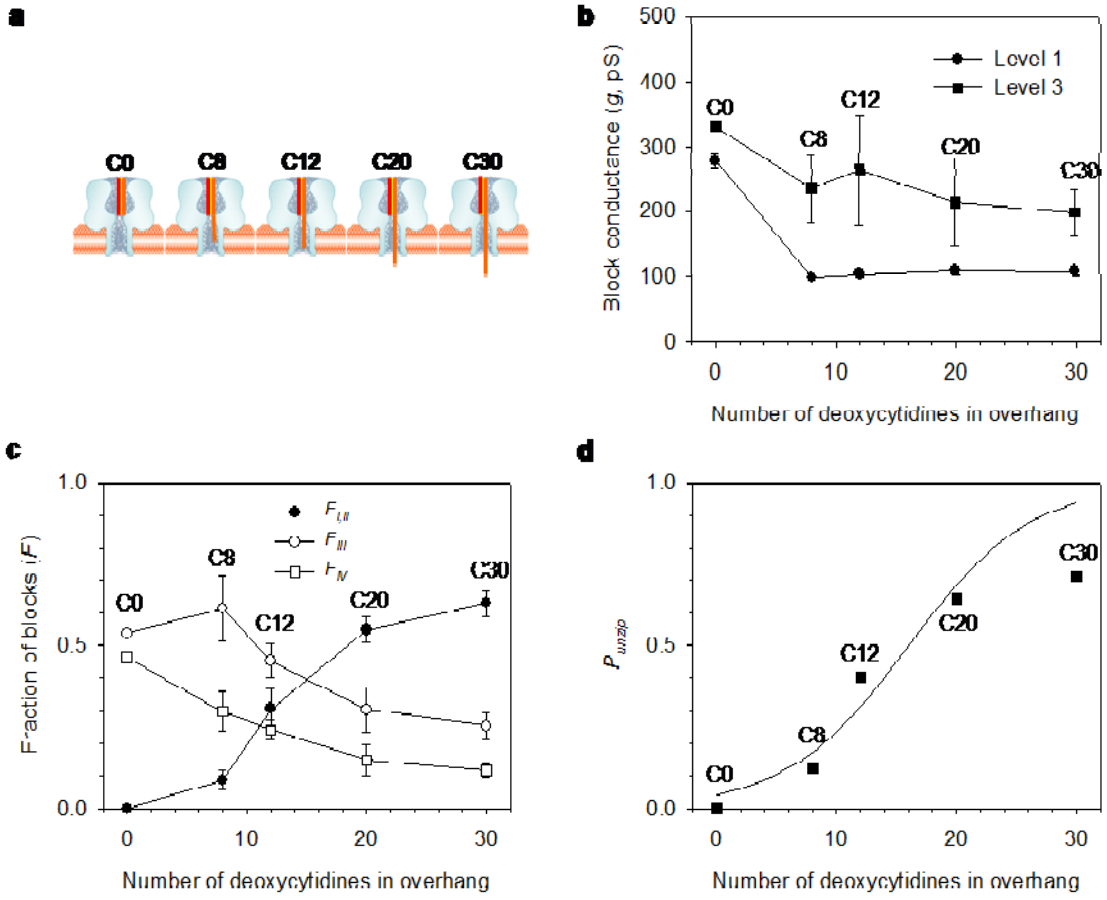

Figure 2-4. Overhang length-regulated DNA unzipping properties. **a.** Models showing the configurations of DNAs in the nanopore for unzipping. DNAs were C0, C8, C12, C20 and C30, which contain 0, 8, 12, 20 and 30 deoxycytidine in the overhang attached to the 3'-end of the anti-sense strand (Table 1). **b.** Overhang length-dependent conductance of Level 1 occurred in the Pattern I, II and III blocks, and Level 3 in Pattern IV blocks. **c.** Fractions of different types of blocks generated by these DNAs. $F_{1,II}$, the total fraction of Pattern I and II blocks (unzipping); F_{1II}, Pattern III blocks; and F_{1V}, Pattern IV blocks. **d.** Unzipping probability P_{unzip} calculated from Eq.1. The data between C0 and C20 were fitted using Eq. 2.

The nanopore as a single-molecule force spectroscope has been utilized to explore the unzipping reactions of nucleic acids (Dudko et al., 2010; Nakane et al., 2004). In previous studies, the long blocks of duration 10^1-10^2 ms observed in the presence of dsDNA were considered as signals for the unzipping reaction. The block duration, or the unzipping time, was an effective parameter for unzipping characterization. By measuring the voltage- and temperature-dependence of the block duration, the unzipping kinetics can be detected (Sutherland et al., 2004; Mathe et al., 2004, 2006; Dudko et al., 2007; Bockelmann and viasnoff, 2008; McNally et al., 2008; Zhao et al., 2008; Muzard et al., 2010; Sauer-Budge et al., 2003). In this report, we uncovered that these dsDNA-associated long blocks can be further classified. By using dsDNAs carrying an overhang as the model, we distinguished a series of characteristic current patterns, from single-conductance to multi-conductance blocks, each marking a specific molecular procedure or a configuration. Discrimination of these current patterns provides a tool to comprehensively explore molecular mechanisms of DNA in the nanopore, allowing for precise determination of kinetic pathways of these processes including unzipping.

Firstly, we clarified that not all the dsDNA-generated long blocks are due to the unzipping reaction. Only current patterns featuring multiple sequential current levels, like Pattern I and II blocks in Figure 2-1, represent the occurrence of unzipping, whereas single-conductance current patterns including Pattern III and IV blocks reveal the release of DNA from the pore without unzipping. This mechanism is different from the hairpin unzipping which generates single-

conductance long blocks. Once unzipped, the hairpin is transformed into an extended ssDNA. Thus, throughout the unzipping procedure, the β-barrel retains occupied and no distinguishable current variation can be detected. Secondly, the multi-conductance current patterns (Pattern I and II blocks) are markers of the entire unzipping procedure. The initial Level 1→Level 2 transition in these blocks demonstrates the time point at which the unzipping occurs; the following Level 2 illustrates the molecular state in which the unzipped short strand temporarily halts in the nanocavity; and the final current jump from Level 2 to Level 1' or to the full open level visualizes the motion pathway of the unzipped ssDNA, which leaves the pore from either the *trans* or *cis* opening. Thus, these sequentially-occurred characteristic currents form a signature that can be used for recognizing the unzipping occurrence. Thirdly, the overhang attached to the dsDNA has been found to have at least two functions. By changing the nucleotide number of the overhang, we identified that the overhang is a determinant of the dsDNA trapping orientations. Earlier studies have identified the trapping directionality of ssDNA (3'- and 5'-threading) (Mathe et al., 2005a), and trapping preference of polyT, polyC and polyA (Purnell et al., 2008) based on the block conductance. The DNA with both blunt ends (without overhang) can enter the pore in both orientations with equal probability, whereas the overhang attached to the dsDNA becomes more favorable to the nanopore compared with the blunt end of DNAs. This trapping preference can be enhanced as the overhang lengthens. Thus, the overhang serves as a "trapping inducer" to increase the trapping probability. As the overhang is trapped in the β-barrel, the voltage will drive it to induce the

unzipping, and the unzipping probability will increase with the number of nucleotides in the overhang. Therefore the overhang also functions as an "unzipping enhancer".

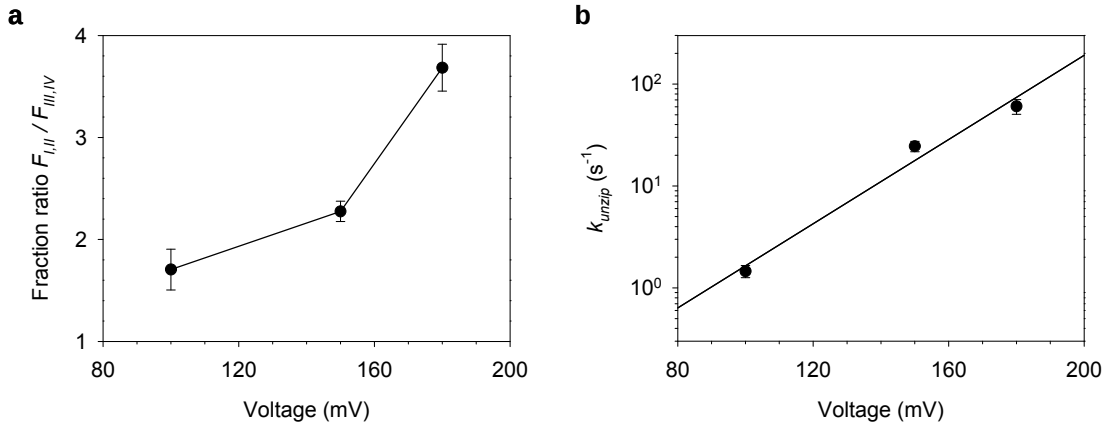

Figure 2-5. Voltage-dependent unzipping of C30. **a.** $F_{I,II}/F_{III,IV}$-V curve. $F_{I,II}/F_{III,IV}$ is the ratio of fractions $F_{I,II}$ and $F_{III,IV}$, representing the ratio of total unzipping and non-unzipping blocks; **b.** Rate constant k_{unzip}. k_{unzip} was calculated using Eq. 3 and data was fitted using Eq. 4.

The ability to discriminate different molecular configurations in the nanopore is important to the construction of a biosensor with both selectivity and sensitivity. Using the "electrical marker" to separate the target signals from interfering signals will greatly enhance the detecting selectivity. In single molecule detection, the sensitivity is determined by the frequency of signature signal occurrence. We have shown in Figure 2-4 and 2-5 that the overhang length and trapping orientation can be used to control the unzipping occurrence. The overhang sequence is also a determinant of the unzipping efficiency. We have compared the ratio of total unzipping and non-unzipping block numbers at +100 mV, $F_{I,II}/F_{III,IV}$, for dsDNA carrying overhangs of poly(dA), poly(dT) and poly(dC). $F_{I,II}/F_{III,IV}$ is 0.5 for 30 deoxythymidines, and 0.8 for 30 deoxyadenosines. By

comparing, $F_{I,II}/F_{III,IV}$ significantly increases to 1.7 for 30 deoxycytidines (Table 2-4). All these findings suggest that it is possible to engineer the overhang in order to achieve a high frequency of unzipping occurrence, thereby enhancing the sensitivity of single molecule detection. The engineered overhang can also be combined with other effective approaches to increase the sensitivity, such as increasing voltage, constructing desired charge profile lining the pore lumen (Maglia et al., 2008), and detecting in a salt concentration gradient (Wanunu et al., 2010). We are working in this direction, and recently we have applied the unzipping signatures in clinical detection of circulating microRNAs (Wang et al., 2010), small regulating RNA molecules (Carthew and Sontheimer, 2009) that are recognized as potential biomarkers of cancers (Rosell et al., 2009b).

Table 2-4: Voltage-dependent ratios of occurrences between unzipping and no-unzipping events.

Voltage (mV)	$N_{unzip}/N_{non\text{-}unzip}$		
	3'-$(dC)_{30}$ overhang	3'-$(dA)_{30}$ overhang	3'-$(dT)_{30}$ overhang
+100	1.7±0.3	0.80±0.12	0.50±0.11
+150	2.0±0.2	n.a.	n.a.
+180	3.7±0.4	n.a.	n.a.

2.4 Conclusions

In this report, we uncovered a signature current pattern that can be used to electrically track the entire unzipping process, from the time course of unzipping to the motion pathway of unzipped single-stranded DNA. With the signature signals, we can also distinguish the release of DNA without unzipping as well as DNA trapping directionality. Quantitative analysis of signature signals shows that the overhang of a DNA is more favorable to the nanopore compared to the blunt end. Therefore, the overhang is not only an unzipping driver, but also a controller of DNA trapping orientation, allowing for regulating the unzipping with programmable probability. Most notably the unzipping signature can act as a single-molecule marker for highly selective and sensitive biosensing.

2.5 References

Ashkenasy,N., J.Sanchez-Quesada, H.Bayley, and M.R.Ghadiri. 2005. Recognizing a single base in an individual DNA strand: A step toward DNA sequencing in nanopores. *Angewandte Chemie-International Edition* 44:1401-1404.

Bayley,H. 2006. Sequencing single molecules of DNA. *Current Opinion in Chemical Biology* 10:628-637.

Bayley,H. and P.S.Cremer. 2001. Stochastic sensors inspired by biology. *Nature* 413:226-230.

Bayley,H. and L.Jayasinghe. 2004. Functional engineered channels and pores - (Review). *Molecular Membrane Biology* 21:209-220.

Bockelmann,U. and V.Viasnoff. 2008. Theoretical study of sequence-dependent nanopore unzipping of DNA. *Biophysical Journal* 94:2716-2724.

Branton,D., D.W.Deamer, A.Marziali, H.Bayley, S.A.Benner, T.Butler, M.Di Ventra, S.Garaj, A.Hibbs, X.Huang, S.B.Jovanovich, P.S.Krstic, S.Lindsay, X.S.Ling, C.H.Mastrangelo, A.Meller, J.S.Oliver, Y.V.Pershin, J.M.Ramsey, R.Riehn, G.V.Soni, V.Tabard-Cossa, M.Wanunu, M.Wiggin, and J.A.Schloss. 2008. The potential and challenges of nanopore sequencing. *Nature Biotechnology* 26:1146-1153.

Carthew,R.W. and E.J.Sontheimer. 2009. Origins and Mechanisms of miRNAs and siRNAs. *Cell* 136:642-655.

Chen,P. and C.M.Li. 2007. Nanopore unstacking of single-stranded DNA helices. *Small* 3:1204-1208.

Cocco,S., R.Monasson, and J.F.Marko. 2001. Force and kinetic barriers to unzipping of the DNA double helix. *Proc Natl Acad Sci U S A* 98:8608-8613.

Dudko, O. K., Mathe, J., and Meller, A. Nanopore Force Spectroscopy Tools for Analyzing Single Biomolecular Complexes. 475[C], 565-589. 2010.
Ref Type: Serial (Book,Monograph)

Dudko,O.K., J.Math, A.Szabo, A.Meller, and G.Hummer. 2007. Extracting kinetics from single-molecule force spectroscopy: Nanopore unzipping of DNA hairpins. *Biophysical Journal* 92:4188-4195.

Gu,L.Q. and J.W.Shim. 2010. Single molecule sensing by nanopores and nanopore devices. *Analyst* 135:441-451.

Hille,B. 2001. Ion Channels of Excitable Membranes. Third ed. Sinauer Associates, Sunderland, MA.

Howorka,S. and H.Bayley. 2002. Probing distance and electrical potential within a protein pore with tethered DNA. *Biophys J* 83:3202-3210.

Howorka,S., S.Cheley, and H.Bayley. 2001a. Sequence-specific detection of individual DNA strands using engineered nanopores. *Nat Biotechnol* 19:636-639.

Howorka,S., L.Movileanu, O.Braha, and H.Bayley. 2001b. Kinetics of duplex formation for individual DNA strands within a single protein nanopore. *Proc Natl Acad Sci U S A* 98:12996-13001.

Howorka,S. and Z.Siwy. 2009. Nanopore analytics: Sensing of single molecules. *Chemical Society Reviews* 38:2360-2384.

Kasianowicz,J.J., E.Brandin, D.Branton, and D.W.Deamer. 1996. Characterization of individual polynucleotide molecules using a membrane channel. *Proc Natl Acad Sci U S A* 93:13770-13773.

Ma,L. and S.L.Cockroft. 2010. Biological nanopores for single-molecule biophysics. *ChemBioChem* 11:25-34.

Maglia, G., Heron, A. J., Stoddart, D., Japrung, D., and Bayley, H. Analysis of Single Nucleic Acid Molecules with Protein Nanopores. 475[C], 591-623. 2010. Methods in Enzymology.
Ref Type: Serial (Book,Monograph)

Maglia,G., M.R.Restrepo, E.Mikhailova, and H.Bayley. 2008. Enhanced translocation of single DNA molecules through +¦-hemolysin nanopores by manipulation of internal charge. *Proc Natl Acad Sci U S A* 105:19720-19725.

Mathe,J., A.Arinstein, Y.Rabin, and A.Meller. 2006. Equilibrium and irreversible unzipping of DNA in a nanopore. *Europhysics Letters* 73:128-134.

Mathe,J., H.Visram, V.Viasnoff, Y.Rabin, and A.Meller. 2004. Nanopore unzipping of individual DNA hairpin molecules. *Biophys J* 87:3205-3212.

Mathe,J., A.Aksimentiev, D.R.Nelson, K.Schulten, and A.Meller. 2005a. Orientation discrimination of single-stranded DNA inside the alpha-hemolysin membrane channel. *Proc Natl Acad Sci U S A* 102:12377-12382.

Mathe,J., A.Aksimentiev, D.R.Nelson, K.Schulten, and A.Meller. 2005b. Orientation discrimination of single-stranded DNA inside the alpha-

hemolysin membrane channel. *Proc Natl Acad Sci U S A* 102:12377-12382.

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.

Meller,A., L.Nivon, E.Brandin, J.Golovchenko, and D.Branton. 2000. Rapid nanopore discrimination between single polynucleotide molecules. *Proc Natl Acad Sci U S A* 97:1079-1084.

Meller,A., L.Nivon, and D.Branton. 2001. Voltage-driven DNA translocations through a nanopore. *Phys Rev Lett* 86:3435-3438.

Mitchell,N. and S.Howorka. 2008. Chemical tags facilitate the sensing of individual DNA strands with nanopores. *Angew Chem Int Ed* 47:5565-5568.

Movileanu,L. 2009. Interrogating single proteins through nanopores: challenges and opportunities. *Trends Biotechnol* 27:333-341.

Muthukumar,M. 1999. Polymer translocation through a hole. *Journal of Chemical Physics* 111:10371-10374.

Muzard,J., M.Martinho, J.Mathe, U.Bockelmann, and V.Viasnoff. 2010. DNA translocation and unzipping through a nanopore: Some geometrical effects. *Biophysical Journal* 98:2170-2178.

Nakane,J., M.Wiggin, and A.Marziali. 2004. A nanosensor for transmembrane capture and identification of single nucleic acid molecules. *Biophys J* 87:615-621.

Purnell,R.F., K.K.Mehta, and J.J.Schmidt. 2008. Nucleotide identification and orientation discrimination of DNA homopolymers immobilized in a protein nanopore. *Nano Lett* 8:3029-3034.

Rabinowits,G., C.GerÃ§el-Taylor, J.M.Day, D.D.Taylor, and G.H.Kloecker. 2009. Exosomal microRNA: A diagnostic marker for lung cancer. *Clin Lung Cancer* 10:42-46.

Rosell,R., J.Wei, and M.Taron. 2009a. Circulating MicroRNA Signatures of Tumor-Derived Exosomes for Early Diagnosis of Non-Small-Cell Lung Cancer. *Clin Lung Cancer* 10:8-9.

Rosell,R., J.Wei, and M.Taron. 2009b. Circulating MicroRNA Signatures of Tumor-Derived Exosomes for Early Diagnosis of Non-Small-Cell Lung Cancer. *Clin Lung Cancer* 10:8-9.

Sauer-Budge,A.F., J.A.Nyamwanda, D.K.Lubensky, and D.Branton. 2003. Unzipping kinetics of double-stranded DNA in a nanopore. *Phys Rev Lett* 90.

Shim,J.W. and L.Q.Gu. 2008. Encapsulating a single G-quadruplex aptamer in a protein nanocavity. *J Phys Chem B* 112:8354-8360.

Shim,J.W., Q.Tan, and L.Q.Gu. 2009. Single-molecule detection of folding and unfolding of a single G-quadruplex aptamer in a nanopore nanocavity. *Nucleic Acids Res* 37:972-982.

Song,L.Z., M.R.Hobaugh, C.Shustak, S.Cheley, H.Bayley, and J.E.Gouaux. 1996. Structure of staphylococcal alpha-hemolysin, a heptameric transmembrane pore. *Science* 274:1859-1866.

Stoddart,D., A.J.Heron, E.Mikhailova, G.Maglia, and H.Bayley. 2009. Single-nucleotide discrimination in immobilized DNA oligonucleotides with a biological nanopore. *Proc Natl Acad Sci U S A* 106:7702-7707.

Sutherland,T.C., M.J.Dinsmore, H.B.Kraatz, and J.S.Lee. 2004. An analysis of mismatched duplex DNA unzipping through a bacterial nanopore. *Biochemistry and Cell Biology* 82:407-412.

Vercoutere,W., S.Winters-Hilt, H.Olsen, D.Deamer, D.Haussler, and M.Akeson. 2001. Rapid discrimination among individual DNA hairpin molecules at single-nucleotide resolution using an ion channel. *Nature Biotechnology* 19:248-252.

Vercoutere,W.A., S.Winters-Hilt, V.S.DeGuzman, D.Deamer, S.E.Ridino, J.T.Rodgers, H.E.Olsen, A.Marziali, and M.Akeson. 2003. Discrimination among individual Watson-Crick base pairs at the termini of single DNA hairpin molecules. *Nucleic Acids Res* 31:1311-1318.

Wang,Y., D.Zheng, Q.Tan, M.Wang, and L.Q.Gu. 2010. Nanopore-facilitated single molecule detection of circulating microRNAs in lung cancer patients. *Submitted*.

Wanunu,M., W.Morrison, Y.Rabin, A.Y.Grosberg, and A.Meller. 2010. Electrostatic focusing of unlabelled DNA into nanoscale pores using a salt gradient. *Nature Nanotechnology* 5:160-165.

Zhao,Q., J.Comer, V.Dimitrov, S.Yemenicioglu, A.Aksimentiev, and G.Timp. 2008. Stretching and unzipping nucleic acid hairpins using a synthetic nanopore. *Nucleic Acids Research* 36:1532-1541.

CHAPTER 3

NANOPORE-BASED DETECTION OF CIRCULATING MICRORNAS IN LUNG CANCER PATIENTS

3.1 Introduction

MicroRNAs are important in development and cell differentiation and in the regulation of the cell cycle, apoptosis and signaling pathways (Carthew and Sontheimer, 2009; Inui et al., 2010). They regulate gene expression at the post-transcriptional level and can either repress translation or cleave target messenger RNAs (Carthew and Sontheimer, 2009). Aberrant expression of microRNAs has been found in all types of tumours, including lung cancer (Garzon et al., 2009; Ortholan et al., 2009). Different types of cancers exhibit distinct microRNAs (Calin and Croce, 2006) that are released in a stable form from the primary tumour into the bloodstream (Mitchell and Howorka, 2008). Recent studies have shown that circulating microRNAs are enveloped inside exosomal vesicles and are transferable to and functional in recipient cells (Kosaka et al., 2010; Rabinowits et al., 2009). Therefore, the detection of tumour-specific circulating microRNAs is useful for the early diagnosis, staging and monitoring of cancer (Rosell et al., 2009).

Reverse transcription real-time polymerase chain reaction (qRT-PCR) assays and microarrays have been developed for the detection of microRNA. However, these methods suffer from cross-hybridization, low selectivity and a lack of valid internal controls (Chen et al., 2005; Li and Ruan, 2009) because the shortness of

the microRNA sequences makes it difficult to design probes and primers. Other techniques based on colourimetry, bioluminescence, enzyme turnover and electrochemistry have been proposed. Nanoparticles, molecular beacons, deep sequencing (Hunt et al., 2009; Yendamuri and Kratzke, 2011) and single-molecule fluorescence (Neely et al., 2006) have also been applied to microRNA detection with enhanced sensitivity and/or selectivity, but these methods either need improvements in versatility or require expensive instruments.

The nanopore is a molecular-scale pore structure that confers great sensitivity to the position and conformation of a single molecule that binds to the pore lumen (Bayley and Jayasinghe, 2004). From the characteristic change in the nanopore conductance, one can electrically elucidate single-molecule kinetic pathways and quantify the target. Various nanopore sensors are being developed with broad biotechnological applications (Bayley and Jayasinghe, 2004; Gu and Shim, 2010; Howorka and Siwy, 2009; Ma and Cockroft, 2010; Movileanu, 2009)), including the next generation of DNA sequencing technology (Bayley, 2006; Branton et al., 2008). The development of nanopore-based microRNA detectors is a novel effort in this rapidly evolving field (Wanunu et al., 2010a). In this report, we propose a robust nanopore sensor that enables the sensitive, selective and direct quantification of cancer-associated microRNAs in the blood by discriminating single nucleotide differences in microRNA family members.

3.2 Materials and Methods

Oligonucleotides, including microRNAs and DNA probes, were synthesized and electrophoresis-purified by Integrated DNA Technologies. Before testing, the mixture of each microRNA and the probe were heated to 90°C for 5 min, gradually cooled to room temperature and stored at 4°C. RNase-free water was used to prepare the RNA samples.

Nanopore electric recording: The electrophysiology setup and methods for nanopore experiments have been detailed elsewhere (Shim et al., 2009). Briefly, the recording apparatus was composed of two chambers (*cis* and *trans*) that were partitioned with a Teflon film. The planar lipid bilayer of 1,2-diphytanoyl-sn-glycerophosphatidylcholine (Avanti Polar Lipids) was formed spanning a 100-150 um hole in the center of the partition. Both *cis* and *trans* chambers were filled with symmetrical 1 M KCl buffered with 10 mM Tris and titrated to pH 8.0. All solutions were filtered before use. Single α-hemolysin proteins were inserted into the bilayer from the *cis* side to form molecular pores. DNA and RNA oligonucleotides (Table 3-1) were synthesized and electrophoresis-purified by Integrated DNA Technologies, IA. Before testing, the mixtures of ssDNA probes and ssRNAs were heated to 90°C for 5 minutes, then gradually cooled down to room temperature and stored at 4°C.

Total RNA extraction from plasma: Peripheral blood samples were obtained from the University of Missouri Ellis Fischel Cancer Center with Institutional Review Board approval. Whole blood with the EDTA preservative was centrifuged at 1,600g for 10 min at room temperature and the plasma was

transferred to new tubes. Total RNAs containing microRNAs were extracted from 350 ml of plasma using the mirVana PARIS Kit (Ambion) according to the manufacturer's protocol. The final elution volume was 100 ml.

Table 3-1: Sequences of studied microRNAs and their probes.

mir-155	5'-UUAAUGCUAAUCGUGAUAGGGG-3'
P_{nt}	5'-CCCCTATCACGATTAGCATTAA-3'
$P_{5'\text{-}C30}$	5'-C_{30}-CCCCTATCACGATTAGCATTAA-3'
$P_{3'\text{-}C30}$	5'-CCCCTATCACGATTAGCATTAA-C_{30}-3'
P_{155}	5'-C_{30}-CCCCTATCACGATTAGCATTAA-C_{30}-3'
Let-7a	5'-UGAGGUAGUAGGUUGU**A**U**A**GUU-3'
P_a	5'-C_{30}-AACTATACAACCTACTACCTCA-C_{30}-3'
Let-7b	5'-UGAGGUAGUAGGUUGU**G**U**G**GUU-3'
P_b	5'-C_{30}-AACCACACAACCTACTACCTCA-C_{30}-3'
Let-7c	5'-UGAGGUAGUAGGUUGU**A**U**G**GUU-3'
P_c	5'-C_{30}-AACCATACAACCTACTACCTCA-C_{30}-3'
miR-39	5'-UCACCGGGUGUAAAUCAGCUUG-3'
P_{39}	5'-C_{30}-CAAGCTGATTTACACCCGGTGA-C_{30}-3'

Normalization of the nanopore and qRT-PCR data using spiked-in C. Elegans microRNA miR-39 as control: We introduced spiked-in synthetic microRNA as control to validate the nanopore sensor's capability of detecting microRNA in human samples. The spiked-in RNA oligonucleotide in the detection matches the sequence of C. Elegans miR-39, a microRNA that is absent in the human genome. About 3.5 ml of 1 nM synthetic miR-39 solution was introduced to each 350 ml plasma sample after the addition of 2× denaturing solution (mirVana PARIS Kit) to plasma; thus, the miR-39 concentration in plasma was 10 pM. The denaturing solution prevents RNAs from degradation by inhibiting endogenous

plasma RNases. For each sample, both miR-155 and spiked-in miR-39 were measured using the nanopore sensor and SYBR Green-based qRT-PCR. In nanopore detection, the probes for miR-155 and miR-39 were P155 and P39. We first measured the signature event frequencies, f155 and f39, of hybrids miR-155·P155 and miR-39·P39, respectively. The variability of f39 reflects the difference in miR-39 concentrations among samples after RNA extraction. Therefore, the ratio of the two frequencies, f_{155}/f_{39}, should eliminate this variability. Finally, the mean f_{155}/f_{39} ratio of six normal samples was used as the standard to calculate each sample's relative miR-155 level.

3.3 Results and Discussion

3.3.1 Generation of microRNA signatures in the nanopore.

We used the α-hemolysin protein pore as the sensor element. The translocation of single-stranded oligonucleotides through this 2-nm pore has been studied extensively (Kasianowicz et al., 1996; Akeson et al., 1999; Meller et al., 2000, 2001; Mitchell and Howorka, 2008). However, it is difficult to distinguish the translocation of different microRNAs because the sequences of all microRNAs are short and similar in length (~18-22 nucleotides). One way to overcome this challenge is the use of a signature that can discriminate the target microRNA in the mixture. We identified such a microRNA signature signal in the nanopore using an oligonucleotide probe (Nakane et al., 2004; Vercoutere et al., 2001; Mathe et al., 2004; Sauer-Budge et al., 2003).

The probe structure is shown in Figure 3-1a. The capture domain of the probe was used to bind the target microRNA in the solution. Each end (3' and 5') of the capture domain was extended with a poly(dC)$_{30}$ signal tag. Figure 3-1b illustrates a sequence of nanopore current events in the presence of microRNA *miR-155* (Donnem et al., 2011; Rosell et al., 2009) and its probe, P_{155}, on the *cis* side of the pore. The boxed events represent a characteristic type of multi-level block that was generated by the *miR-155*•P_{155} hybrid. This block type was not observed in the presence of *miR-155* or P_{155} alone in the *cis* solution. In a multi-level block (Figure 3-1c), Level 1 lasted for 250±58 ms (Figure 3-2a), which is almost equal to the entire event duration, and greatly reduced the nanopore current, with a relative residual conductance (g/g_0) of 0.15. Level 1 was followed by a discrete current increase to Level 2, which persisted for 410±20 μs, with a g/g_0 of 0.42. Finally, the current discretely dropped to Level 3 and remained there briefly for 270±30 μs before returning to the full open level. Similar to Level 1, Level 3 almost fully reduced the pore current, with a g/g_0 of 0.08. The right panel in Figure 3-1c depicts the molecular configurations that corresponded to the multi-level block. The amplitude and duration of Level 1 were consistent with a configuration in which the *miR-155*•P_{155} complex was trapped in the nanopore at the 2.6-nm *cis* opening, with either the 3' or 5' signal tag of P_{155} occupying the 1.6- to 2.0-nm β-barrel. Driven by the transmembrane voltage, the signal tag in the β-barrel induced the dissociation of *miR-155*•P_{155}. The dissociation time (the duration of Level 1) was comparable to previously reported time scales for DNA unzipping in the pore (Mathé et al., 2004; Sauer-Budge et al., 2003).

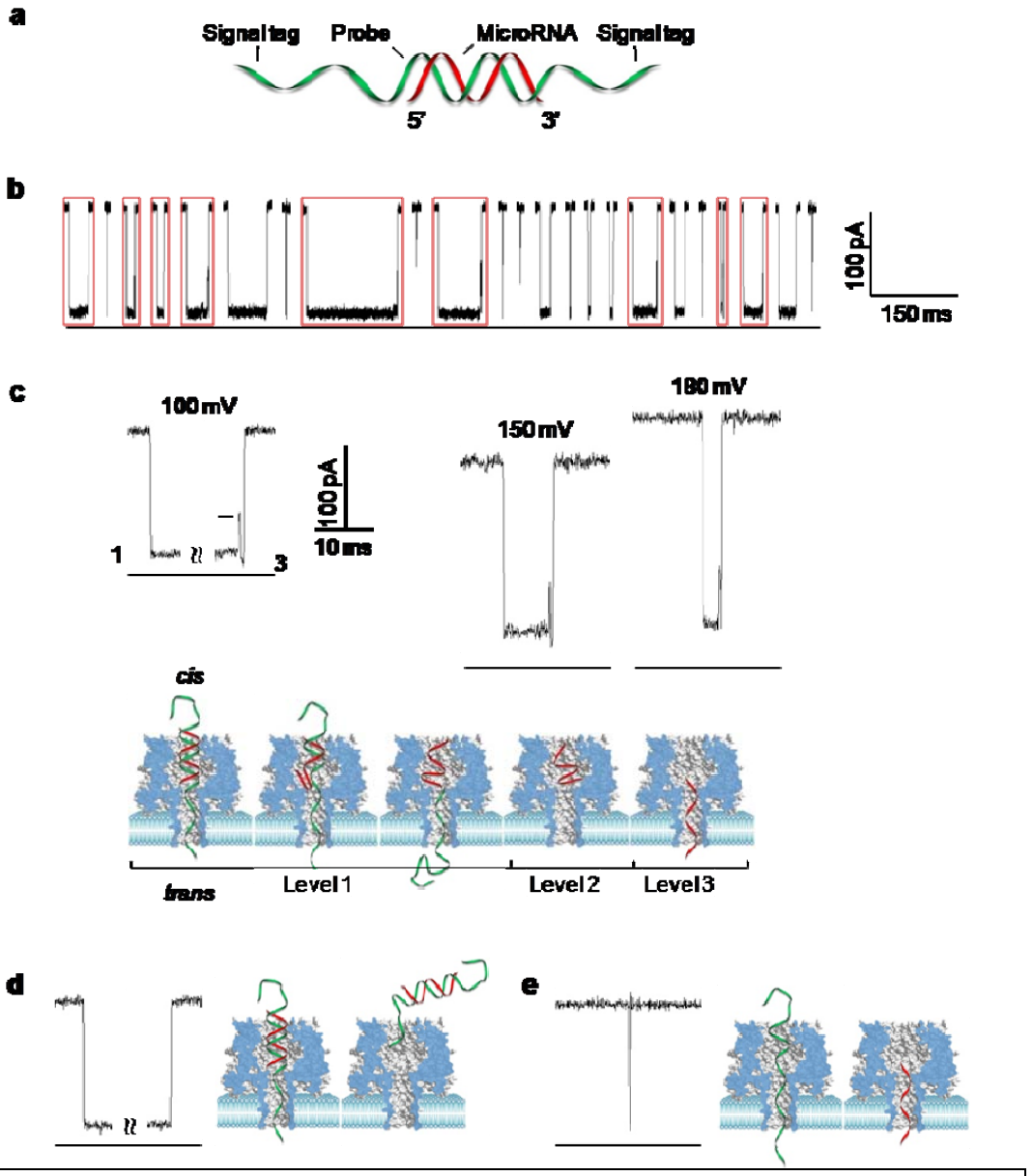

Figure 3-1: Capturing single microRNA molecules in the nanopore. **a,** Molecular diagram of a microRNA (red) bound to a probe (green) bearing signal tags on each end. **b,** A sequence of nanopore current blocks in the presence of 100 nM *miR-155* and 100 nM P_{155} in the *cis* solution. Traces were recorded at +100 mV in solutions containing 1 M KCl buffered with 10 mM Tris (pH 8.0). Red boxes represent the multi-level current pattern. **c,** Multi-level long block (from **a**) at +100 mV generated by the *miR-155*•P_{155} hybrid. Right: diagram showing the mechanism of translocation. Level 1 indicates the trapping of the microRNA•probe hybrid in the pore, the unzipping of the microRNA from the probe and the translocation of the probe through the pore; Level 2, the unzipped microRNA residing in the nanocavity of the pore; Level 3, the translocation of the unzipped microRNA through the pore. Lower: multi-level blocks at +150 mV and +180 mV. Increasing the voltage reduced the duration of Level 1 and Level 3, which supports the above mechanistic model. **d,** A single-level block (from **a**) generated by a trapped *mir-155*•P_{155} hybrid that exited the pore from the *cis* entrance without unzipping. **e,** A spike-like short block generated by the translocation of unhybridised *miR-155* or P_{155} from the *cis* solution.

After unzipping, P_{155} left the pore through the narrower (*trans*) opening, and Level 1 was terminated. Level 2 featured a large residue conductance, which corresponded to a configuration in which *mir-155* unzipped from *miR-155*•P_{155} temporarily resided in the nanocavity. This result is consistent with earlier reports that an oligonucleotide trapped in the nanocavity can generate partial blocks (Maglia et al., 2008; Shim et al., 2009). The *MiR-155* in the nanocavity finally passed through the β-barrel to yield the short-lived Level 3. The duration of Level 3 (270 μs) was consistent with the translocation duration of *mir-155* alone (220 μs) and the time scale for DNA or RNA translocations in previous studies (Kasianowicz et al., 1996; Butler et al., 2006).

The molecular mechanism described above was further evidenced by the voltage-dependent durations of Levels 1 and 3. Level 1 was shortened by 23-fold to 11 ms and Level 3 by 2-fold to 150 μs as the voltage increased from +100 mV to +180 mV (Figure 3-1c), which indicated that the voltage both enhanced the unzipping of *miR-155*•P_{155} and accelerated the translocation of *miR-155* (Meller et al., 2001; Wang and Gu, 2011). After the electrical recording, we also identified *miR-155* in the *trans* solution using RT-PCR (data not shown), which supports the hypothesis that the microRNA unzipped from a microRNA• probe hybrid can translocate to the *trans* side of the pore.

The *miR-155*•P_{155} hybrid also produced single-level long blocks that had similar conductances (g/g_0=0.15) to Level 1 of the multi-level block (Figure 3-1d). The single-level block was formed by the trapped *miR-155*•P_{155} complex that

returned to the *cis* solution without unzipping (Mathé et al., 2005; Sauer-Budge et al., 2003).

The frequency ratio of multi-level to single-level events monotonically increased with the voltage, from 0.42 at +100 mV to 1.4 at +180 mV (Table 3-2). This voltage-dependent frequency variation was consistent with the observation that higher voltage increased the unzipping probability of *miR-155*•P_{155} and decreased its escaping probability.

Table 3-2. Frequencies of mult-level long events, single-level long events and short translocation events in the presence of *miR-155* and P_{155}.

	f_{multi} (s^{-1})	f_{single} (s^{-1})	f_{multi} / f_{single}	f_{transl} (s^{-1})
+100 mV	4.2±0.6	10±1	0.42	8.5±1.8
+150 mV	45±8	49±4	0.91	50±5
+180 mV	77±13	53±2	1.4	84±13

f_{multi} : frequency of multi-level long events; f_{single}: frequency of single-level long events; f_{transl} : frequency of short translocation events. 100 nM *miR-155* and 100 nM P_{155}.

In addition to the characteristic long events, we also observed spike-like short blocks in the same recordings (Figure 3-1e). Both their duration (220±21 μs) and conductance (g/g_0=0.16) were similar to the translocation of unhybridised *miR-155* (220 μs) and the P_{155} (160 μs) present in the *cis* solution. The histogram of different block durations are shown in Figure 3-2.

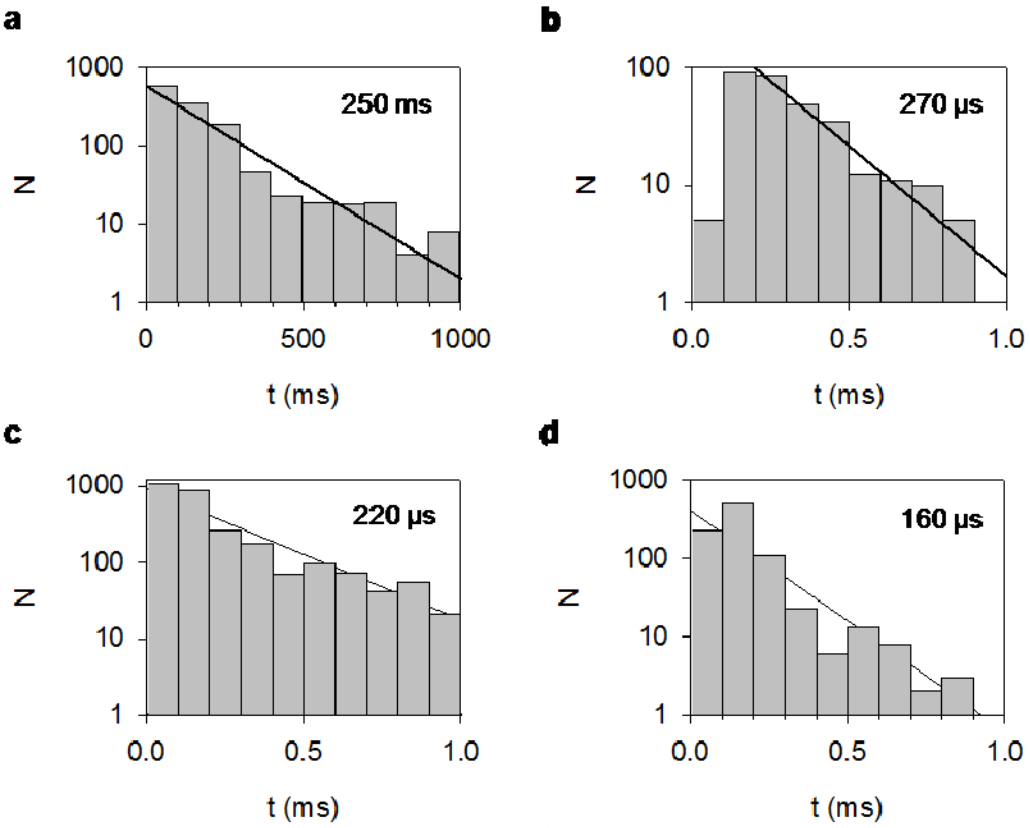

Figure 3-2. Histograms of block durations. **a.** Signature blocks generated by the *mir-155*•P_{155} hybrid. **b.** The short Level 3 state in the signature block. **c.** and **d.** Short blocks by translocation of *miR-155* (c) and P_{155} (d) alone. Data was obtained from current traces recorded in 1 M KCl at +100 mV.

The above analysis indicated two important functions performed by the signal tag of the probe: the guidance of the microRNA•probe complex entrapment in the pore and the inducement of the dissociation of the microRNA•probe complex. The configuration change during unzipping gave rise to the signature current patterns, which enabled the recognition of single target microRNA molecules. Because of the probe's specificity, the frequency of the

signature signal (f_{sig}) was independent of the presence of multiple nucleic acids such as of *Let-7a* and *of Let-7b* components. Therefore, the frequency of the signature signal (f_{sig}) could be used to quantify the target microRNA in the mixture. Overall, the signature signal ensured the high selectivity that is required for microRNA detection in plasma RNA extract.

3.3.2 Quantification of microRNA using probes with optimized sequences.

The frequency of signature events can be used to quantify microRNA by the equation $f_{sig} = k_{on}\,[miR]_0$, where $[miR]_0$ is the microRNA concentration and k_{on} is the occurrence rate constant of signature events. k_{on} is a key determinant of quantification sensitivity and can be vastly improved by the optimisation of the probe structure. Figure 3-3a shows that the probe without the signal tag (P_{nt}) gave the lowest k_{on} at $2.8\pm0.6\times10^4$ $M^{-1}s^{-1}$. k_{on} tripled to $6.8\pm1.3\times10^4$ $M^{-1}s^{-1}$ when the probe was attached to a poly$(dC)_{30}$ tag at the 5' end ($P_{5'\text{-}C30}$). However, k_{on} was vastly increased by 50-fold to $1.4\pm0.3\times10^6$ $M^{-1}s^{-1}$ when the poly$(dC)_{30}$ tag was attached to the 3' end ($P_{3'\text{-}C30}$). This orientation discrimination of single-stranded oligonucleotides in the nanopore is consistent with previous studies (Mathé et al., 2005; Butler et al., 2006; Purnell et al., 2008). As expected, the P_{155} probe that contained both 3' and 5' poly$(dC)_{30}$ achieved the highest k_{on} at $2.0\pm0.2\times10^6$ $M^{-1}s^{-1}$. In addition to the tag directionality, k_{on} is also dependent on the tag length. For example, the poly$(dC)_{30}$ tag showed much higher efficiency in the generation of signature events than a shorter tag, such as poly$(dC)_8$, and was more efficient than poly$(dA)_{30}$ and poly$(dT)_{30}$ tags (unpublished data). Using

P_{155} as the probe, we verified that the frequency of the signature event was proportional to the *miR-155* concentration range from 10 to 100 nM (Figure 3-3b). This correlation was measured at +100 mV in 1 M KCl. The frequencies in any two *miR-155* concentrations, such as 10 and 25 nM, were significantly separated (p<0.005). Wanunu et al. (2010b) have reported that a gradient of salt concentration across a synthetic nanopore vastly increased the capture rate of dsDNA. Also, the use of an engineered nanopore (Maglia et al., 2008) proved effective in increasing the event frequency for high sensitivity. In addition we have shown here that the application of high voltage also can increase the capture rate. The voltage-dependent frequency of *mir-155*•P_{155} signature events is shown in Figure 3-3c.

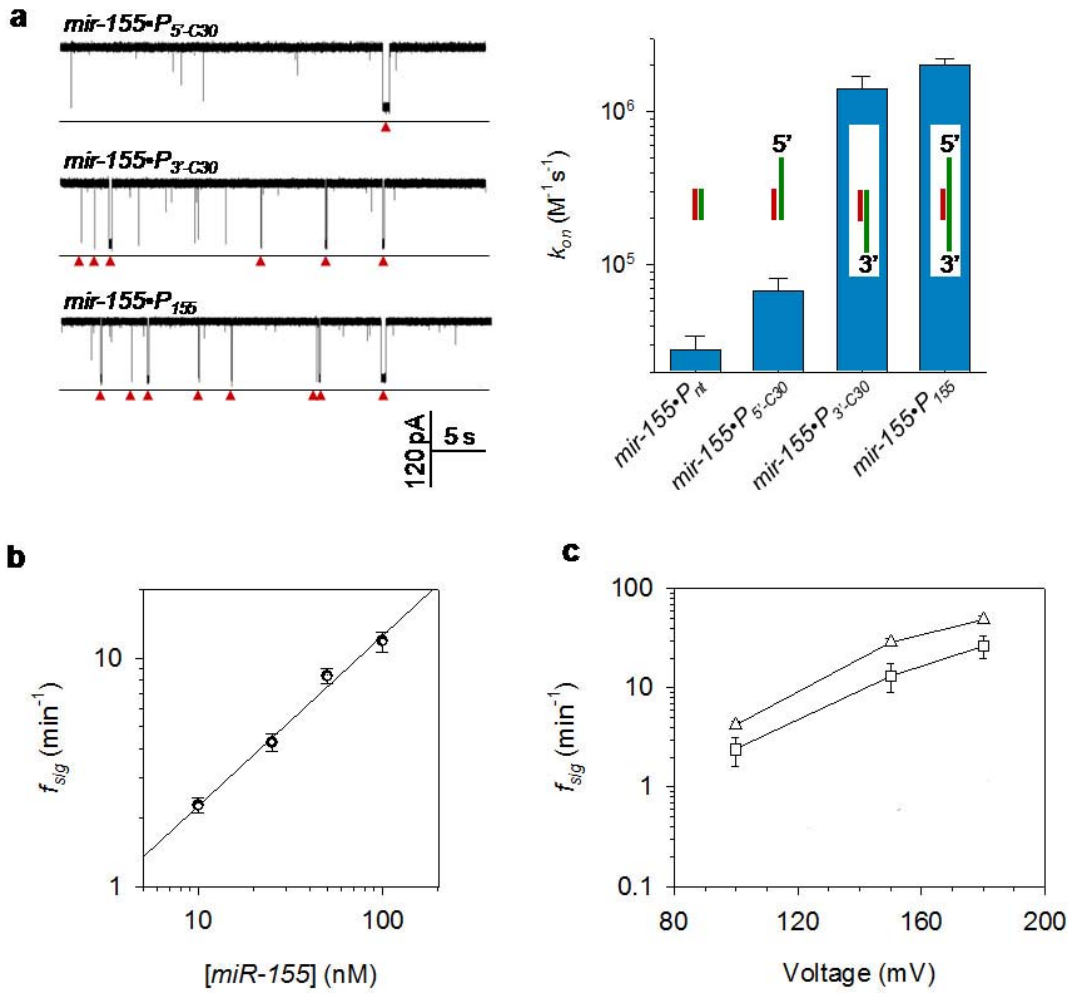

Figure 3-3. Enhancing the detection sensitivity by optimising the probe sequence. **a.** Left, current traces showing the frequency of signature events for *miR-155* hybridised to the probes $P_{5'\text{-}C30}$ (top), $P_{3'\text{-}C30}$ (middle) and P_{155} (bottom), monitored at +100 mV in 1 M KCl. Right, the occurrence rate constant of the signature events for *miR-155* detection with different probes. The results using any two probes were statistically significant (p<0.005); **b.** the [*miR-155*] - f_{155} correlation for target concentration ranges between 10 and 100 nM. The results of detections in any two *miR-155* concentrations were statistically significant (p<0.01). **c.** Voltage-dependent frequency of *mir-155*•P_{155} signature events. Data was obtained from current traces recorded in 1 M KCl with 10 (□) and 25 (□) nM *mir-155* in the presence of 100 nM P_{155}.

3.3.3 Discrimination of single nucleotide differences between microRNAs.

Among the over 800 human microRNAs that have been identified, many members of the same family possess similar sequences or single nucleotide polymorphisms (SNPs). The SNPs are associated with significant biological properties of cancers, such as susceptibility, prognosis, and the response to therapeutic agents (Cho, 2009). However, sequence-similar microRNAs or SNPs are difficult to distinguish using current PCR or hybridisation-based methods (Chen et al., 2005; Hunt et al., 2009; Li and Ruan, 2009). Because dsDNAs that contain a single nucleotide mismatch are identifiable in the nanopore based on their unzipping time (Howorka et al., 2001; Nakane et al., 2004; Sauer-Budge et al., 2003; Vercoutere et al., 2001), we decided to study whether the nanopore could discriminate single nucleotide differences in microRNA family members.

We selected the *Let-7* tumour-suppressing microRNA family (Calin and Croce, 2006; Garzon et al., 2009; Ortholan et al., 2009) as the target. *Let-7* is down-regulated in lung cancer and, therefore, is useful as a biomarker and potential therapeutic agent (Landi et al., 2010). *Let-7a* and *-7b* each contain two different nucleotides (see sequences in Table 3-1), and their probes were P_a and P_b, respectively. The hybrids *let-7a*•P_a and *let-7b*•P_b were fill-matched, but *let-7b*•P_a and *let-7a*•P_b contained 2 mismatches. Figure 3-4a shows that the unzipping time, τ_{sig}, at +120 mV decreased 3.2-fold from *let-7a*•P_a (155±28 ms) to *let-7b*•P_a (48±11 ms) (p<0.005) and 6.9-fold from *let-7b*•P_b (165±47 ms) to *let-7a*•P_b (24±2 ms) (p<0.005) (Figure 3-4a). Similarly, *Let-7a* and *-7c* each contained only one different nucleotide. Figure 3-4b shows that when using P_a

and P_c to target *Let-7a* and -*7c*, respectively, at +100 mV, τ_{sig} decreased 2.4-fold from *let-7a*•P_a (303±45 ms) to *let-7c*•P_a (124±39 ms) (p<0.005) and 2.0-fold from *let-7c*•P_c (342±49 ms) to *let-7a*•P_c (179±38 ms) (p<0.005), respectively. The duration of signature events for fully-matched microRNA•probe hybrids and that with mismatches are shown in Table 3-3.

Based on receiver operating characteristic (ROC) curves (Figure 3-4c), we measured the area under the ROC curve (AUC, Figure 3-4d), which is an indicator of the ability to discriminate nucleotide differences between microRNAs. The AUCs varied between 0.72 and 0.83 and increased with the ratio of event duration between fully-matched hybrids and mismatches. This duration ratio-dependent AUC was further verified through simulation (Figure 3-4d). The ROC analysis strongly suggested the possibility of the discrimination of SNPs in microRNAs based on the duration of signature events. This discriminatory ability was tested in a mixture of 100 nM *Let-7a* and *Let-7b* and probed with P_a. Based on the AUCs, the optimal cut-point (OCP) duration was calculated to be ~190 ms, which was the threshold duration that provided the best discriminatory ability. Block events that were longer than the OCP occurred for *Let-7a*•P_a, and events that were shorter than the OCP occurred for *Let-7c*•P_a. The numbers of events were transformed into the concentrations of *Let-7a* and *Let-7c*, which were 85 nM and 120 nM, respectively. Both concentrations varied 15-20% from the real concentration (100 nM).

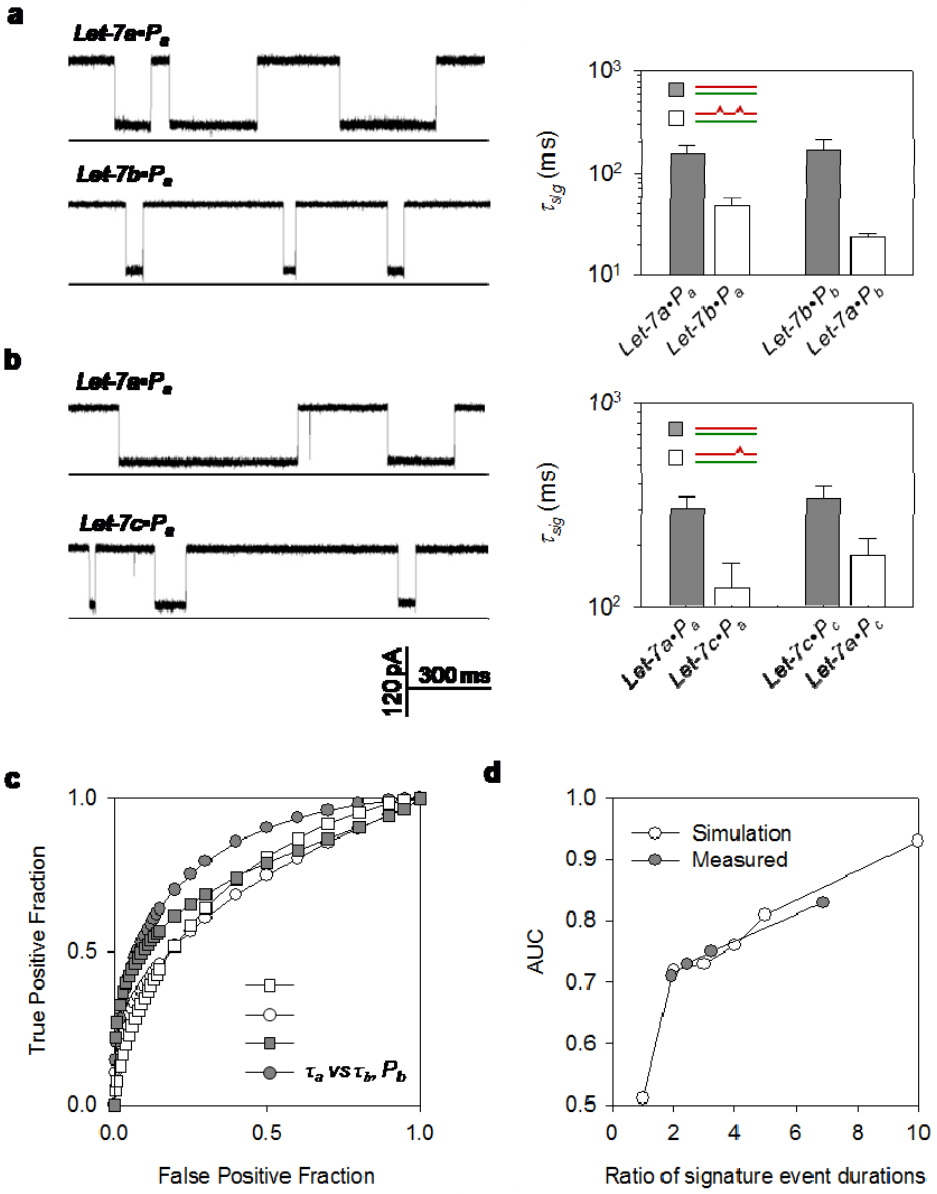

Figure 3-4. The differentiation of *let-7* microRNAs containing one or two different nucleotides. The sequences of *let-7a*, *-7b*, and *-7c* are provided in Table 3-1. **a,** The detection of *let-7a* and *let-7b* using the probe P_a or P_b at +120 mV. Left, current traces. Right, the signature event duration. **b,** The detection of *let-7a* and *-7c* using the probe P_a or P_c at +100 mV. Left, current traces. Right, the signature event duration. **c,** Receiver operating characteristic (ROC) curves for the discrimination of events for fully-matched microRNA•probe hybrids (defined as positive events) and probes containing mismatches (defined as negative events). □: *let-7a*•P_a / *let-7b*•P_a, ○: *let-7b*•P_b / *let-7a*•P_b, ■: *let-7a*•P_a / *let-7c*•P_a, and ●: *let-7c*•P_c / *let-7a*•P_c. **d,** Correlations between the areas under the ROC curves (AUCs) and the ratio of event duration for fully-matched hybrids versus mismatches. ●: AUCs measured from the ROC curves in panel **c** (Table 3-4), ○: AUCs calculated from ROC analyses based on simulated datasets (Figure 3-5 and Table 3-5). The computer-generated event duration followed an exponential distribution. The ratios of event duration for the ROC analysis were 1, 2, 3, 4, 5 and 10.

Table 3-3: Durations of signature events for fully-matched microRNA•probe hybrids and that with mismatches.

microRNA•Probe	*let-7a*•P_a	*let-7b*•P_a	*let-7b*•P_b	*let-7a*•P_b
τ_{sig} (ms) at +120 mV	155±28 (n=7)	48±11 (n=7)	165±47 (n=6)	24 ± 2 (n=5)
p-value	<0.005		<0.005	
microRNA•Probe	*let-7a*•P_a	*let-7c*•P_a	*let-7c*•P_c	*let-7a*•P_c
τ_{sig} (ms) at +100 mV	303±45 (n=6)	124±39 (n=4)	343±49 (n=6)	179±38 (n=4)
p-value	<0.005		<0.05	

Table 3-4: Areas under ROC curves (AUC) for separation of microRNAs with one nucleotide difference (*let-7a* and *let-7c*) and with two nucleotide difference (*let-7a* and *let-7b*).

	let-7a•P_a	*let-7b*•P_b
let-7b•P_a	0.75	n.a.
let-7a•P_b	n.a.	0.83

	let-7a•P_a	*let-7c*•P_c
let-7c•P_a	0.73	n.a.
let-7a•P_c	n.a.	0.71

The receiver operating characteristic (ROC) curve is a plot of the true positive rate (sensitivity) against the false positive rate (1-selectivity) for the different possible cutoff points that separate the entire duration distribution into the positive and negative components. In the microRNA detection, the events for fully matched microRNA•probe hybrids were denoted as “positive”, and that for mismatched hybrids as “negative”. The separation accuracy was measured by the area under the ROC curve (AUC). An AUC of 1 represents a perfect separation; an area of 0.5 represents no separation ability. AUC was analyzed

online using free software at the website:

http://www.rad.jhmi.edu/jeng/javarad/roc/JROCFITi.html

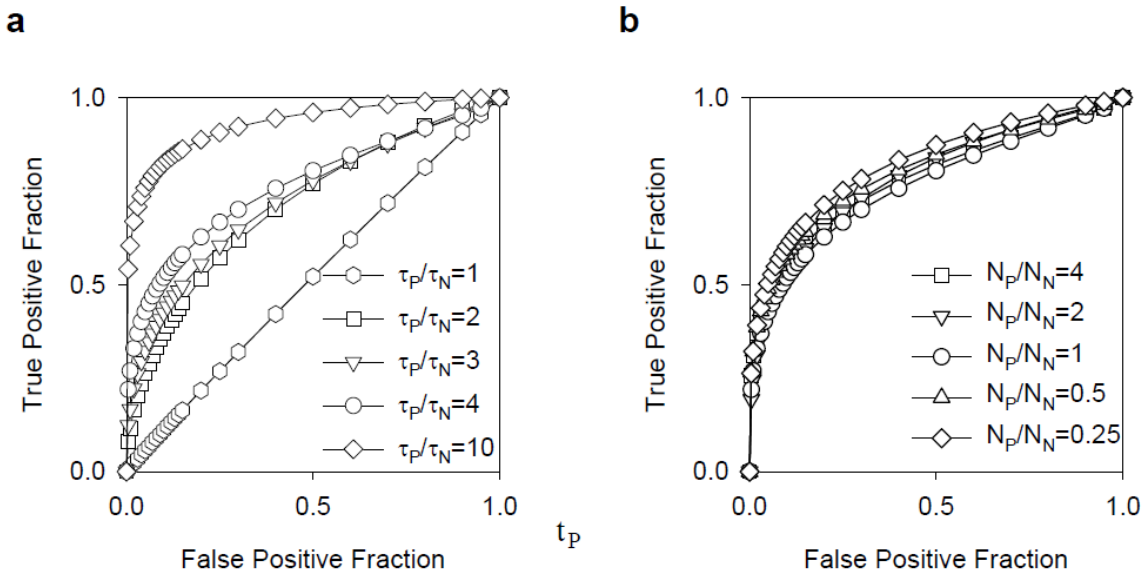

Figure 3-5: Simulation on separation of fully-match (positive) and mismatch (negative) events based on event duration. **a.** ROC curves at various duration ratios. There were 400 events of both types participating in the analysis; **b.** ROC curves at various event number ratios of the two type of events. The duration ratio $\tau_P/\tau_N = 3$.

Table 3-5: Areas under ROC curves (AUC) and optimal cutoff point (OCP) at various duration ratio and event number ratio.

τ_P/τ_N (s/s)	1/1	2/1	3/1	4/1	5/1	10/1
AUC	0.51	0.72	0.73	0.76	0.78	0.93
OCP	n.a.	1.33	1.74	1.88	1.98	2.18
N_P/N_N	200:800	200:400	200:200	200:150	200:100	200:50
AUC	0.83	0.81	0.76	0.76	0.79	0.78
OCP	1.88	1.88	1.85	1.79	1.81	1.97

Both AUC and OCP were calculated from the ROC curves shown in Figure 3-5. OCP is a cutoff duration at the maximal value of the Youden index. The Youden index is defined as {sensitivity+selectivity-1}, calculated from the ROC curve, and range between 0 and 1. A cutoff duration leading to complete separation of long and short duration distribution results in the Youden index =1, whereas complete overlap gives the Youden index = 0. The cutoff duration value that returns the maximum of the Youden index, i.e. "optimal" cutoff point (OCP) (Greiner et al., 2000 *Preventive Veterinary Medicine* **45**, 23-41) gives the most accurate separation. In Table 3-5, the τ_P / τ_N is the duration ratio of the "positive" and "negative" datasets. Each dataset contained 200 exponentially-distributed duration values. The dataset with a longer mean duration was denoted as "positive"; the shorter one as "negative". N_P / N_N: the event number ratio in the "positive" (N_P) versus the "negative" dataset. τ_P / τ_N was 5 in this simulation.

3.3.4 Detection of microRNA levels in lung cancer patients.

Lung cancer is the leading cause of cancer mortality in men and women worldwide and is responsible for approximately 1.2 million deaths each year (Silvestri et al., 2009). Because there is no valid screening procedure available, more than 70% of patients are diagnosed at advanced stages, with a 5-year survival rate of less than 15% (Silvestri et al., 2009). The development of new methods for early diagnosis is critically needed to save lives in lung cancer patients. Using microarray and qRT-PCR methods, distinct microRNA expression profiles can be revealed in normal and lung cancer tissues (Chen et al., 2005;

Hunt et al., 2009; Li and Ruan, 2009). Over 100 microRNAs are dysregulated in lung cancer (Calin and Croce, 2006; Garzon et al., 2009; Kosaka et al., 2010; Mitchell et al., 2008; Ortholan et al., 2009; Rabinowits et al., 2009; Rosell et al., 2009). High levels of *miR-155* and low levels of *let-7a-2* correlate with significantly poor prognoses and shorter survival times in lung cancer patients (Patnaik et al., 2010; Yanaihara et al., 2006).

We detected plasma *miR-155* in lung cancer patients using a prototype of the nanopore sensor. Peripheral blood samples were obtained from six lung cancer patients and six healthy volunteers after local IRB approval. Total plasma RNAs, which contained microRNAs, were extracted from 350 µl of each plasma sample using a mirVana PARIS Kit (Ambion, Austin, TX), with a final elution volume of 100 µl. The elution volumes were then divided into two 50-µl aliquots for the nanopore and RT-PCR assays (Shi and Chiang, 2005). One aliquot was mixed with the recording solution that contained the P_{155} probe. The nanopore current retained a low level of noise even in the presence of plasma samples, and distinct short and long blocks (marked with red arrows) were indentified in both the control (Figure 3-6a,c) and lung cancer groups (Figure 3-6b,d). The characteristic long blocks of *miR-155* shared the same current profile and properties with the synthetic *miR-155* RNA (Figure 3-1). In the absence of P_{155}, no long blocks were observed (Figure 3-6a,b), but short blocks were observed as a translocation of single-stranded oligonucleotides, such as free microRNAs. One reason for the identification of more short events in the presence of P_{155} (Figure 3-6a,b) than in the absence of P_{155} (Figure 3-6c,d) is that the DNA probe (P_{155})

translocation rate was 8.4 times higher than that of the microRNAs. Overall, the characteristic long blocks could be attributed to *miR-155*•P_{155} hybrids and served as signatures for the identification of single molecules of *miR-155*.

The frequencies of *miR-155* signature events (f_{155}) for all the samples in the lung cancer and control groups were measured in the presence of spiked-in synthetic *C. elegans* microRNA *miR-39*. Figure 3-6e shows that all of the f_{155} values in the lung cancer group were higher than those in the control group (p<0.001). However, Figure 3-6f indicates that the frequencies of spiked-in *miR-39* events (f_{39}) were independent of the samples (p>0.01). To evaluate the variability during sample preparation, the ratio f_{155}/f_{39} was used to normalise the assays. Figure 3-6g shows that the mean f_{155}/f_{39} in the lung cancer group (0.62±0.06) was significantly higher than the f_{155}/f_{39} in the control group (0.24±0.05) (p<0.001). The Levels of *miR-155* and spiked-in *miR-39* in human plasma samples detected by the nanopore sensor are shown in Table 3-6.

The relative levels of *miR-155* that were measured with the nanopore were further compared using the qRT-PCR method (Figure 3-6h). Collectively, the mean level of *mir-155* was increased 2.6-fold in the lung cancer group compared to the control group as measured by the nanopore sensor (p< 0.001), but a 4.3-fold increase was obtained using the qRT-PCR method (p<0.02). The latter method had a greater variability. Therefore, although both the nanopore and qRT-PCR assays indicated a significant elevation of *miR-155* in lung cancer patient samples, the nanopore method demonstrated higher accuracy with no requirement for labeling or amplification (Figure 3-6h).

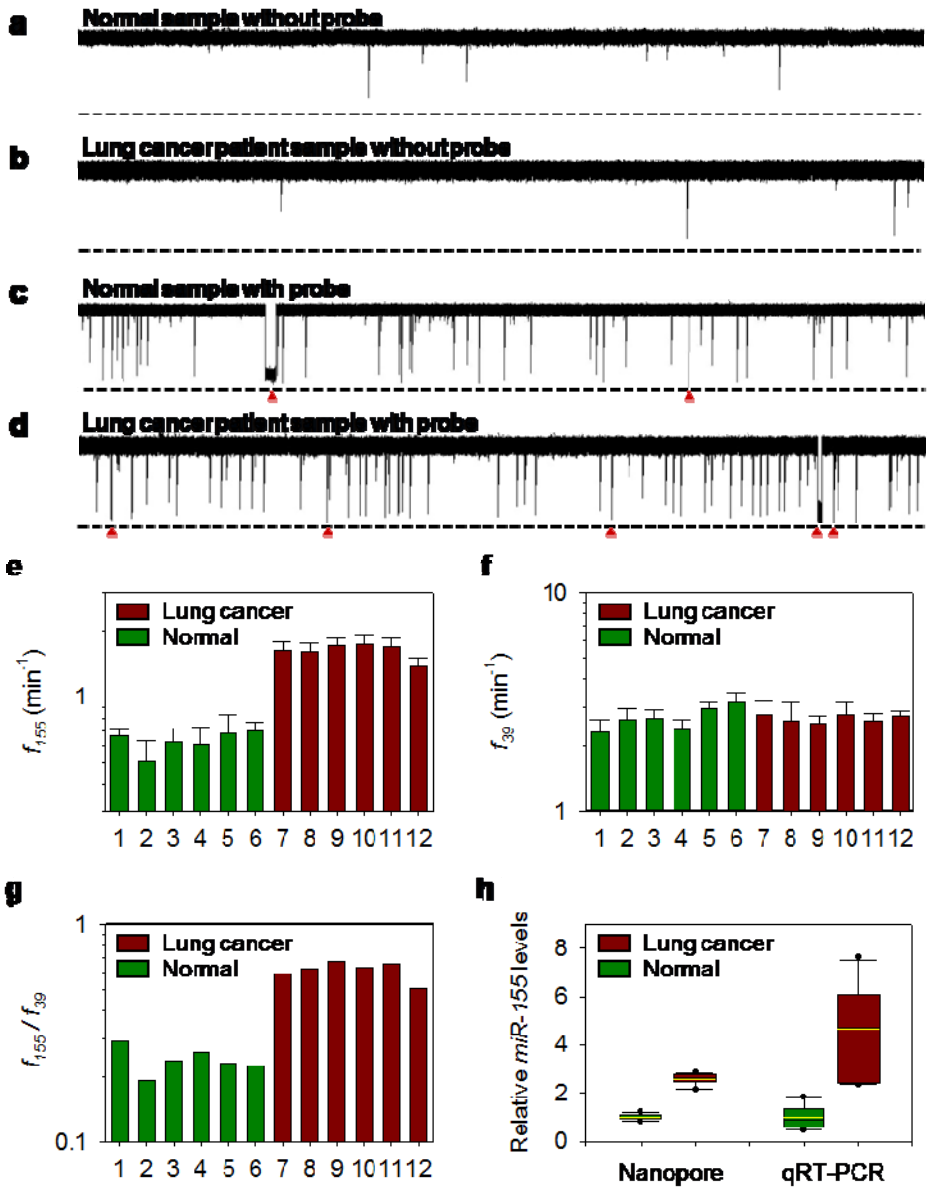

Figure 3-6: The detection of *miR-155* in the plasma of lung cancer patients. **a** through **d,** The signature events found in current traces for total plasma RNAs from healthy volunteers (**a**) and lung cancer patients (**b**) in the presence of 100 nM P_{155} probe. No signature events were observed in the absence of P_{155} (**c** and **d**). The traces were recorded in 1 M KCl at +100 mV. **e,** The frequencies of *miR-155* signature events (f_{155}) from six healthy individuals (#1 to #6) and six patients with lung cancer (#7 to #12) in the presence of spiked-in synthetic *miR-39*. **f,** The frequencies of *miR-39* signature events detected using P_{39} (sequence in Table 3-1) from all the samples that were used in **e**. Each sample was measured n times ($n \geq 4$) using independent nanopores. The data are displayed as means±SD. **g,** f_{155}/f_{39} calculated from panels **e** and **f**. **h,** Box and whiskers plots of the relative *miR-155* levels in the healthy and lung cancer groups measured with the nanopore sensor and qRT-PCR. The boxes mark the intervals between the 25th and 75th percentiles. The black lines inside the boxes denote the medians. The whiskers denote the intervals between the 5th and 95th percentiles. The filled circles indicate data points outside of the 5th and 95th percentiles.

Table 3-6: Levels of *miR-155* and spiked-in *miR-39* in human plasma samples detected by the nanopore sensor.

Sample #		f_{155} (min^{-1})	f_{39} (min^{-1})	f_{155} / f_{39}	Relative *miR-155* level [a]
Normal	1	0.67±0.04	2.29±0.34	0.292	1.22
	2	0.50±0.13	2.62±0.33	0.192	0.80
	3	0.62±0.09	2.65±0.25	0.236	0.99
	4	0.61±0.12	2.35±0.29	0.258	1.08
	5	0.68±0.15	2.95±0.20	0.230	0.97
	6	0.70±0.06	3.14±0.35	0.224	0.94
			(Mean) [a]	0.239	1
Lung cancer	7	1.63±0.15	2.75±0.44	0.593	2.49
	8	1.62±0.14	2.58±0.57	0.627	2.63
	9	1.71±0.16	2.52±0.21	0.675	2.83
	10	1.74±0.18	2.76±0.38	0.633	2.65
	11	1.69±0.16	2.57±0.25	0.658	2.76
	12	1.39±0.11	2.74±0.12	0.507	2.12
			(Mean)	0.615	2.57

[a]: Relative *miR-155* level was obtained by normalizing each sample's f_{155} / f_{39} to the mean f_{155} / f_{39} of the normal samples 1-6, which was 0.239 as highlighted in the table.

3. 4 Conclusions

In conclusion, we designed an oligonucleotide probe to generate the signature signal in a nanopore system for the direct and label-free detection of a target microRNA in a fluctuating background, such as a plasma RNA extract.

This nanopore method can be optimised by the incorporation of unnatural compounds, such as locked nucleic acids (LNA) and peptide nucleotide acids (PNA) (Kloosterman et al., 2006; Neely et al., 2006), into the probe sequence to enhance both specificity and sensitivity. During the preparation of this paper, Wanunu et al. (2010a) also reported utilising a 3-nm synthetic pore to quantify the translocation of enriched microRNAs that were hybridised to the probe. This strategy can be adapted to fabricate nanopore arrays (Kim et al., 2006) for microRNA profiling. We speculate that an ideal nanopore device in the future would integrate the protein pore in a synthetic nanopore system with the advantages of both, as reported recently (Hall et al., 2010). Overall, the nanopore method could prove a useful tool for quantitative studies of microRNAs and the discovery of disease markers. Once the microRNA markers are established, this approach will have the potential for the noninvasive screening and early diagnosis of diseases, such as cancer.

Note: This chapter was modified from a published work by: Yong Wang, Dali Zheng, Qiulin Tan, Michael X. Wang & Li-Qun Gu. Nanopore-based detection of circulating microRNAs in lung cancer patients. Nature Nanotechnology 6, 668–674, (2011).

3.5 References

Akeson,M., D.Branton, J.J.Kasianowicz, E.Brandin, and D.W.Deamer. 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.

Bayley,H. 2006. Sequencing single molecules of DNA. *Current Opinion in Chemical Biology* 10:628-637.

Bayley,H. and L.Jayasinghe. 2004. Functional engineered channels and pores - (Review). *Molecular Membrane Biology* 21:209-220.

Branton,D., D.W.Deamer, A.Marziali, H.Bayley, S.A.Benner, T.Butler, M.Di Ventra, S.Garaj, A.Hibbs, X.Huang, S.B.Jovanovich, P.S.Krstic, S.Lindsay, X.S.Ling, C.H.Mastrangelo, A.Meller, J.S.Oliver, Y.V.Pershin, J.M.Ramsey, R.Riehn, G.V.Soni, V.Tabard-Cossa, M.Wanunu, M.Wiggin, and J.A.Schloss. 2008. The potential and challenges of nanopore sequencing. *Nature Biotechnology* 26:1146-1153.

Butler,T.Z., J.H.Gundlach, and M.A.Troll. 2006. Determination of RNA orientation during translocation through a biological nanopore. *Biophys J* 90:190-199.

Calin,G.A. and C.M.Croce. 2006. MicroRNA signatures in human cancers. *Nat Rev Cancer* 6:857-866.

Carthew,R.W. and E.J.Sontheimer. 2009. Origins and Mechanisms of miRNAs and siRNAs. *Cell* 136:642-655.

Chen,C., D.A.Ridzon, A.J.Broomer, Z.Zhou, D.H.Lee, J.T.Nguyen, M.Barbisin, N.L.Xu, V.R.Mahuvakar, M.R.Andersen, K.Q.Lao, K.J.Livak, and K.J.Guegler. 2005. Real-time quantification of microRNAs by stem-loop RT-PCR. *Nucleic Acids Res* 33.

Cho,W.C. 2009. Role of miRNAs in lung cancer. *Expert Rev Mol Diagn* 9:773-776.

Donnem,T., K.Eklo, T.Berg, S.W.Sorbye, K.Lonvik, S.Al Saad, K.Al Shibli, S.Andersen, H.Stenvold, R.M.Bremnes, and L.T.Busund. 2011. Prognostic impact of MiR-155 in non-small cell lung cancer evaluated by in situ hybridization. *J Transl Med* 9:6.

Garzon,R., G.A.Calin, and C.M.Croce. 2009. MicroRNAs in cancer. *Annual Review of Medicine* 60:167-179.

Gu,L.Q. and J.W.Shim. 2010. Single molecule sensing by nanopores and nanopore devices. *Analyst* 135:441-451.

Hall,A.R., A.Scott, D.Rotem, K.K.Mehta, H.Bayley, and C.Dekker. 2010. Hybrid pore formation by directed insertion of +¦-haemolysin into solid-state nanopores. *Nat Nanotechnol* 5:874-877.

Howorka,S., S.Cheley, and H.Bayley. 2001. Sequence-specific detection of individual DNA strands using engineered nanopores. *Nature Biotechnology* 19:636-639.

Howorka,S. and Z.Siwy. 2009. Nanopore analytics: Sensing of single molecules. *Chemical Society Reviews* 38:2360-2384.

Hunt,E.A., A.M.Goulding, and S.K.Deo. 2009. Direct detection and quantification of microRNAs. *Anal Biochem* 387:1-12.

Inui,M., G.Martello, and S.Piccolo. 2010. MicroRNA control of signal transduction. *Nat Rev Mol Cell Biol* 11:252-263.

Kasianowicz,J.J., E.Brandin, D.Branton, and D.W.Deamer. 1996. Characterization of individual polynucleotide molecules using a membrane channel. *Proc Natl Acad Sci U S A* 93:13770-13773.

Kim,M.J., M.Wanunu, D.C.Bell, and A.Meller. 2006. Rapid fabrication of uniformly sized nanopores and nanopore arrays for parallel DNA analysis. *Advanced Materials* 18:3149-3153.

Kloosterman,W.P., E.Wienholds, E.de Bruijn, S.Kauppinen, and R.H.Plasterk. 2006. In situ detection of miRNAs in animal embryos using LNA-modified oligonucleotide probes. *Nat Methods* 3:27-29.

Kosaka,N., H.Iguchi, Y.Yoshioka, F.Takeshita, Y.Matsuki, and T.Ochiya. 2010. Secretory mechanisms and intercellular transfer of microRNAs in living cells. *J Biol Chem*.

Landi,M.T., Y.Zhao, M.Rotunno, J.Koshiol, H.Liu, A.W.Bergen, M.Rubagotti, A.M.Goldstein, I.Linnoila, F.M.Marincola, M.A.Tucker, P.A.Bertazzi, A.C.Pesatori, N.E.Caporaso, L.M.McShane, and E.Wang. 2010. MicroRNA expression differentiates histology and predicts survival of lung cancer. *Clin Cancer Res* 16:430-441.

Li,W. and K.Ruan. 2009. MicroRNA detection by microarray. *Anal Bioanal Chem* 394:1117-1124.

Ma,L. and S.L.Cockroft. 2010. Biological nanopores for single-molecule biophysics. *Chembiochem* 11:25-34.

Maglia,G., M.R.Restrepo, E.Mikhailova, and H.Bayley. 2008. Enhanced translocation of single DNA molecules through +¦-hemolysin nanopores by

manipulation of internal charge. *Proc Natl Acad Sci U S A* 105:19720-19725.

Mathé,J., A.Aksimentiev, D.R.Nelson, K.Schulten, and A.Meller. 2005. Orientation discrimination of single-stranded DNA inside the alpha-hemolysin membrane channel. *Proc Natl Acad Sci U S A* 102:12377-12382.

Mathé,J., H.Visram, V.Viasnoff, Y.Rabin, and A.Meller. 2004. Nanopore unzipping of individual DNA hairpin molecules. *Biophys J* 87:3205-3212.

Meller,A., L.Nivon, E.Brandin, J.Golovchenko, and D.Branton. 2000. Rapid nanopore discrimination between single polynucleotide molecules. *Proc Natl Acad Sci U S A* 97:1079-1084.

Meller,A., L.Nivon, and D.Branton. 2001. Voltage-driven DNA translocations through a nanopore. *Phys Rev Lett* 86:3435-3438.

Mitchell,N. and S.Howorka. 2008. Chemical tags facilitate the sensing of individual DNA strands with nanopores. *Angew Chem Int Ed* 47:5565-5568.

Mitchell,P.S., R.K.Parkin, E.M.Kroh, B.R.Fritz, S.K.Wyman, E.L.Pogosova-Agadjanyan, A.Peterson, J.Noteboom, K.C.O'Briant, A.Allen, D.W.Lin, N.Urban, C.W.Drescher, B.S.Knudsen, D.L.Stirewalt, R.Gentleman, R.L.Vessella, P.S.Nelson, D.B.Martin, and M.Tewari. 2008. Circulating microRNAs as stable blood-based markers for cancer detection. *Proc Natl Acad Sci U S A* 105:10513-10518.

Movileanu,L. 2009. Interrogating single proteins through nanopores: challenges and opportunities. *Trends Biotechnol* 27:333-341.

Nakane,J., M.Wiggin, and A.Marziali. 2004. A nanosensor for transmembrane capture and identification of single nucleic acid molecules. *Biophys J* 87:615-621.

Neely,L.A., S.Patel, J.Garver, M.Gallo, M.Hackett, S.McLaughlin, M.Nadel, J.Harris, S.Gullans, and J.Rooke. 2006. A single-molecule method for the quantitation of microRNA gene expression. *Nat Methods* 3:41-46.

Ortholan,C., M.P.Puissegur, M.Ilie, P.Barbry, B.Mari, and P.Hofman. 2009. MicroRNAs and lung cancer: New oncogenes and tumor suppressors, new prognostic factors and potential therapeutic targets. *Curr Med Chem* 16:1047-1061.

Patnaik,S.K., E.Kannisto, S.Knudsen, and S.Yendamuri. 2010. Evaluation of microRNA expression profiles that may predict recurrence of localized

stage I non-small cell lung cancer after surgical resection. *Cancer Res* 70:36-45.

Purnell,R.F., K.K.Mehta, and J.J.Schmidt. 2008. Nucleotide identification and orientation discrimination of DNA homopolymers immobilized in a protein nanopore. *Nano Lett* 8:3029-3034.

Rabinowits,G., C.Gerçel-Taylor, J.M.Day, D.D.Taylor, and G.H.Kloecker. 2009. Exosomal microRNA: A diagnostic marker for lung cancer. *Clin Lung Cancer* 10:42-46.

Rosell,R., J.Wei, and M.Taron. 2009. Circulating MicroRNA Signatures of Tumor-Derived Exosomes for Early Diagnosis of Non-Small-Cell Lung Cancer. *Clin Lung Cancer* 10:8-9.

Sauer-Budge,A.F., J.A.Nyamwanda, D.K.Lubensky, and D.Branton. 2003. Unzipping kinetics of double-stranded DNA in a nanopore. *Phys Rev Lett* 90.

Shi,R. and V.L.Chiang. 2005. Facile means for quantifying microRNA expression by real-time PCR. *BioTechniques* 39:519-524.

Shim,J.W., Q.Tan, and L.Q.Gu. 2009. Single-molecule detection of folding and unfolding of a single G-quadruplex aptamer in a nanopore nanocavity. *Nucleic Acids Res* 37:972-982.

Silvestri,G.A., A.J.Alberg, and J.Ravenel. 2009. The changing epidemiology of lung cancer with a focus on screening. *BMJ* 339.

Vercoutere,W., S.Winters-Hilt, H.Olsen, D.Deamer, D.Haussler, and M.Akeson. 2001. Rapid discrimination among individual DNA hairpin molecules at single-nucleotide resolution using an ion channel. *Nature Biotechnology* 19:248-252.

Wang,Y. and L.Q.Gu. 2011. Single-molecule signatures for characterization and regulation of nucleic acids unzipping in a nanopore. *Submitted*.

Wanunu,M., T.Dadosh, V.Ray, J.Jin, L.McReynolds, and M.Drndi-ç. 2010a. Rapid electronic detection of probe-specific microRNAs using thin nanopore sensors. *Nat Nanotechnol* 5:807-814.

Wanunu,M., W.Morrison, Y.Rabin, A.Y.Grosberg, and A.Meller. 2010b. Electrostatic focusing of unlabelled DNA into nanoscale pores using a salt gradient. *Nat Nanotechnol* 5:160-165.

Yanaihara,N., N.Caplen, E.Bowman, M.Seike, K.Kumamoto, M.Yi, R.M.Stephens, A.Okamoto, J.Yokota, T.Tanaka, G.A.Calin, C.G.Liu, C.M.Croce, and

C.C.Harris. 2006. Unique microRNA molecular profiles in lung cancer diagnosis and prognosis. *Cancer Cell* 9:189-198.

Yendamuri,S. and R.Kratzke. 2011. MicroRNA biomarkers in lung cancer: MiRacle or quagMiRe? *Transl Res* 157:209-215.

CHAPTER 4

FUTURE DIRECTIONS

4.1 Introduction

MicroRNAs are a class of short (~18-22 nucleotides) non-coding RNAs that are important in development and cell differentiation and in the regulation of the cell cycle, apoptosis and signalling pathways (Carthew and Sontheimer, 2009; Inui et al., 2010). By either repressing translation or cleaving the target messenger RNA (Diederichs and Haber, 2007; Bartel, 2009), these microRNAs regulate approximately 30% of human gene expression (Bartel, 2004, 2009; Diederichs and Haber, 2007) at the post-transcriptional level. Aberrant expression of microRNAs has been found in all types of tumours (Garzon et al., 2009; Ortholan et al., 2009), and different types of cancers exhibit distinct microRNAs profiles. MicroRNAs can be released from the primary tumour into the bloodstream in a stable form. Circulating microRNAs are enveloped inside exosomal vesicles and are transferable to and functional in recipient cells (Rabinowits et al., 2009; Rosell et al., 2009; Kosaka et al., 2010). Therefore, the detection of tumour-specific circulating microRNAs is useful for the early diagnosis, staging and monitoring of cancer (Mitchell et al., 2008; Rabinowits et al., 2009; Rosell et al., 2009; Kosaka et al., 2010).

The nanopore is a molecular-scale pore structure that is able to detect the position and conformation of a single molecule in the pore with great sensitivity[12]. From the characteristic change in the nanopore conductance, one can electrically

elucidate single-molecule kinetic pathways and quantify the target. Various nanopore sensors are being developed with broad biotechnological applications (Bayley and Jayasinghe, 2004; Howorka and Siwy, 2009; Movileanu, 2009; Gu and Shim, 2010; Ma and Cockroft, 2010; Olasagasti et al., 2010), including the next generation of DNA sequencing technology (Bayley, 2006; Branton et al., 2008). The development of nanopore-based microRNA detectors is a novel effort in this rapidly evolving field, and Wanunu et al. (2010a) first reported utilising a 3-nm synthetic pore to quantify the translocation of enriched microRNAs that were hybridised to the probe.

Single-nucleotide polymorphisms (SNPs) among microRNA have big effects on cancer (Song and Chen, 2011). The SNPs are associated with significant biological properties of cancers, such as susceptibility, prognosis and response to therapeutic agents. The difference of the unzipping time between microRNA family members with SNA was identified in a nanopore sensor (Wang et al., 2011).

4.2: Future Directions

Utilizing a biological nanopore, clinically-relevant miRNA-155 was detected in blood samples and miRNAs with only one or two nucleotides differences among let-7 family that can be differentiated by the unzipping time of the probe-microRNA hybrid (Wang et al., 2011). Nanopore detection needs no labeling and amplification of the target MicroRNA. The difference of the unzipping time between microRNA family members with SNA can be differentiated, but needs

further improvement. This differences are apparent at a voltage of 100 mV, but higher voltages could confer higher sensitivity in a nanopore sensor. In clinical samples, the trapping frequency of the target probe-microRNA hybrid is 0.5-2/min, so a long recording time is needed to generate enough event numbers for data analysis. Also, there may be some other dsRNAs in the RNA samples extracted from blood that may affect the results (Wang et al., 2011).

Therefore, improvements could occur in three directions: 1) improve the sensitivity of the nanopore sensor; 2) improve the resolution of the nanopore sensor to differentiate single nucleotide difference among nucleic acids, especially at high voltages and 3) eliminate the background RNAs in the clinic sample.

4.3 To Improve the Sensitivity of the Nanopore Sensor

4.3.1 Engineered nanopore

Studies have found that altering the charge profile lining the pore lumen can significantly increase the trapping rate (sensitivity) of ssDNA in the nanopore (Maglia et al., 2008). We will test two mutations, M113R and D128K, which add positively charged amino acid residues at the constriction site and the end of the stem.

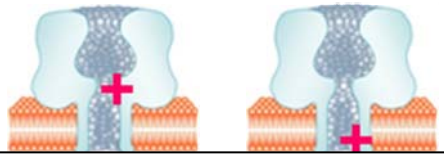

Figure 4-1: Two mutations M113R (left) and D128K (right).

We will test dsDNA using these two mutations and expect to see the trapping rates increase.

4.3.2 Salt gradient

A previous study has shown that a salt gradient across the pore can generate an enhanced electrical field (EEF) (Figure 4-2), which increases the capture rate of dsDNA at high throughput using a solid nanopore (Wanunu et al., 2010b). However, it is not clear if the salt gradient can increase the capture rate of DNA in the biological α-HL pore, or if it can increase the capture rate of other biomolecules, such as single-stranded nucleotides, peptides, or DNA-protein complexes.

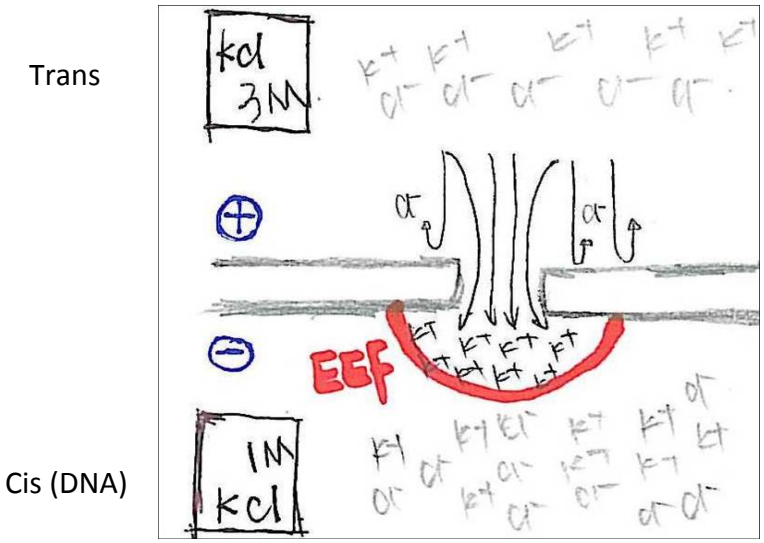

Figure 4-2: The principle of the salt gradient generates the enhanced electric filed (EEF).

Therefore, we will test the capture rate of nucleic acids, peptides and DNA-protein complexes under an enhanced electric field in the nanopore system. And

we expect to see the similar results that the molecules trapping rate will be increased under the EEF. The principle underlying the EEF: under symmetric salt conditions, the application of a potential across the membrane produces an electric field profile away from the pore mouth. This electric field near the pore attracts the negatively charged DNA into the pore. Under asymmetric salt conditions, the applied voltage results in cation selectivity, K^+ ions are continuously pumped from trans side to cis side (Figure 4-2) by both the chemical and electrical potential gradients. The pore mouth at cis side is effectively polarized and the magnitude of the electric field increases. This effect enhances the capture rate of the DNA present at cis side (Wanunu et al., 2010b).

4.3.3 Lower pH value

Studies found that a single nucleobase can be recognized at alkaline pH (Franceschini et al., 2012), and at the lower pH values, the ssDNA trapping rates was increased when compared to the neutral pH (de Zoysa et al., 2011). We will test the dsDNA in the nanopore sensor at a pH range from 6.0, 5.5, 5, 4.5 and 4 and we expect to see the dsDNA trapping rate increases under these pH values too. At lower pH values, more positively charged H^+ ions may lining inside the pore which can attract the negatively charged DNA, this may also increase the translocation time (Franceschini et al., 2012).

4.4: To Improve the Resolution of the Nanopore Sensor to Differentiate Single Nucleotide Differences Among Nucleic Acids.

4.4.1 Shear dissociation

4.4.1.1 Introduction

Studies using an atomic force microscope (AFM) have shown that the dsDNA could be unzipped through two geometries: tear and shear (Albrecht et al., 2003; Kufer et al., 2008). The unzipping forces were more than 3 times higher for shear dissociation compared to tear unzipping (Smit et al., 1988; Strunz et al., 1999). Later in a differential measurement format, the bond-rupture probabilities of the same base pairs is more than 15 times lower in shear geometry than in tear geometry (Albrecht et al., 2003). We have demonstrated the tear unzipping signature in the nanopore (Wang et al., 2011). We hypothesize that shear dissociation time is significantly longer than tear unzipping time, and a single-base pair mismatches would significantly decrease the shear dissociation time.

4.4.1.2 Preliminary data and experiment design

We have designed a 22-base pairs (bp) dsDNA with a biotin-strepdavidin link which can be unzipped by the voltage (Figure 4-3 upper panel). The dsDNA complex was trapped in the nanopore first (Figure 4-3A), and the dsDNA went through tear unzipping (Figure 4-3B) until the dsDNA unzipping was blocked when the big head (strepdavidin) was blocked by the pore opening (Figure 4-3C). Then the dsDNA is thought to transverse the pore following shear dissociation. After the dissociation process is finished, the biotin-strepdavidin-ssDNA complex stayed in the nanocavity temporarily (Figure 4-3D) and finally stretched into the

stem by the applied voltage (Figure 4-3E). In the middle panel of Figure 4-3, a single trace recorded at 180 mV follows the steps described above.

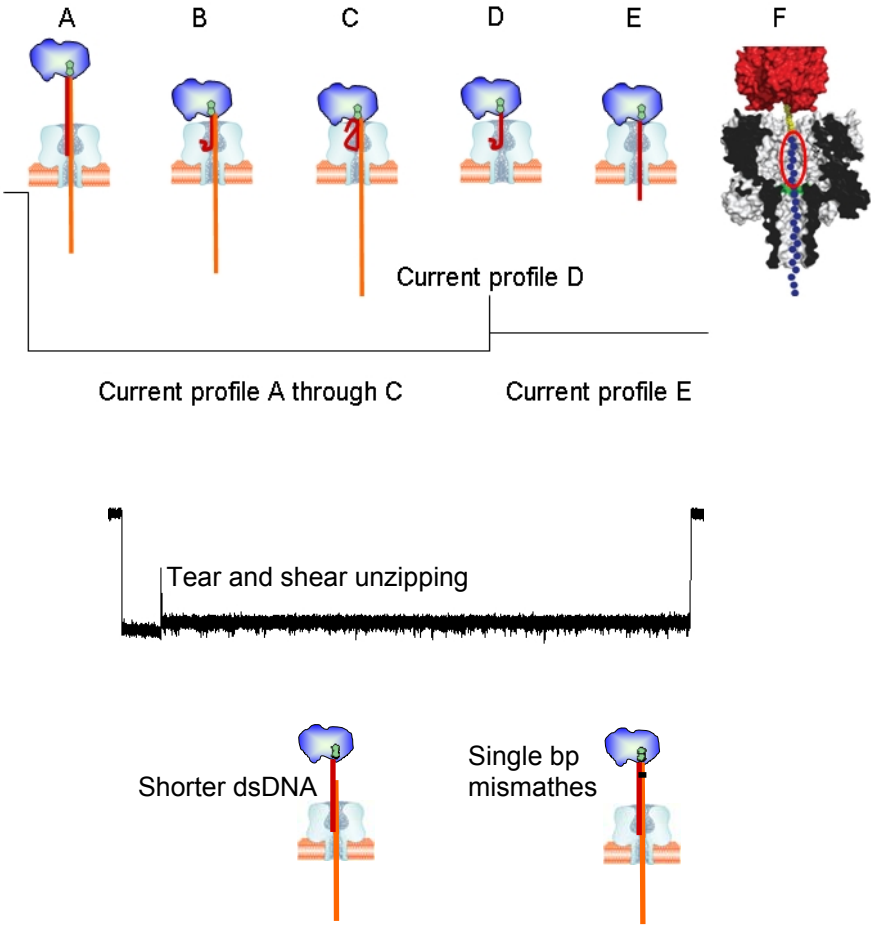

Figure 4-3: Experiment design to demonstrate shear dissociation in the nanopore. Upper panel: A: biotin (green)-strepdavidin (blue)-dsDNA complex trapped in the nanopore; B: dsDNA during tear unzipping; C: the big strepdavidin reach the pore and stop tear unzipping, shear dissociation starts; D: shear dissociation finished and biotin-strepdavidin-ssDNA complex trapped in the nanocavity temporarily; E: the ssDNA stretched into the stem by the applied voltage; F; the scheme of biotin (yellow)-strepdavidin (red)-ssDNA (blue) complex in the pore which indicates that it is 6 bases from biotion to the constriction site. Middle panel: A single trace recorded at 180 mV. Bottom panel: future experiments with a shorter dsDNA (left panel) and single-base pair mismatches in the shear dissociation geometry.

There are 6 bases from the biotin to the constriction site of the pore (Stoddart et al., 2009) (Figure 4-3F, upper panel), which corresponds to 9-bp in length for dsDNA (Figure 4-3C, upper panel). So for our 22-bp dsDNA, it will have 13-bp tear unzipping and 9-bp shear dissociation. We found that the at 180 mV, the tear and shear dissociation time for the 22-bp dsDNA is 67.3 ms by exponential fit (N=547 events), which is more than 40 times longer than the same dsDNA without the biotin-strepdavidin which contains only tear unzipping, the unzipping time is about 1.45 ms by exponential fit (N=536 events). This difference is consistent with other measurements. In AFM studies, the unzipping forces were more than 3 times higher for shear dissociation compared to tear unzipping (Smit et al., 1988; Strunz et al., 1999). Later, in a differential measurement format, the bond-rupture probabilities of the same base pairs is more than 15 times lower in shear geometry than in tear geometry (Albrecht et al., 2003). So the long dissociation time in current profile A through C (Figure 4-3 upper panel) should include the dissociation in shear geometry.

To confirm the long unzipping time in current profile A through C (Figure 4-3 upper panel) indeed includes the dissociation in shear geometry. We will carry several control experiments: 1) get recordings at different voltages and a voltage-dependence of the tear and shear dissociation time should appear, i.e. higher voltage will give a shorter dissociation time, which includs the shear dissociation time; 2) design a shorter dsDNA (Figure 4-3, bottom left panel), and we expect to see a decrease in the unzipping time. After these control experiments are confirmed, we will design a single-base pair mismatch in the shear dissociation

geometry. We expect to see a significant decrease in the dissociation time compared to what we have seen in the tear unzipping geometry (Wang et al., 2011).

4.4.1.3 Discussion

Currently, the best resolution obtained in shear and tear unzipping has been around 10-bp by AFM (Strunz et al., 1999; Krautbauer et al., 2003). If we can confirm the shear dissociation in our nanopore experiment, which has more than 40 fold difference in unzipping time in shear and tear unzipping geometries for only 9-bp, it will be possible to increase the resolution in the nanopore experiment, which could serve as a new platform for the unzipping studies at a more refined level. Furthermore, we hope to show that shear dissociation will improve the ability to discriminate SNA in the microRNA families. This study could provide a useful way to detect the SNPs and their function among microRNAs.

It may be problematic to confirm the shear dissociation signature event. The dsDNA with the big strepdavidin could possibly decrease the unzipping probability and increase the chance of diffusing back, i.e. the applied voltage would be less effective, and the highest voltage we can apply in the nanopore system is 200 mV. But we still expect to see that some of the dsDNA can be unzipped at high voltages (150 mV and 180 mV). But at low voltage we may not be able to see many signature events and we may need to use a higher concentration and record for a longer time to obtain a sufficient number of

unzipping events. For single-base pair mismatches, we expect to see a significant difference in the unzipping time, but this is still uncertain.

4.4.2 Engineered oligonucleotide for slowing-down unzipping

4.4.2.1 Introduction

Studies have found that less-charged complexes (dye-intercalated DNA) displayed strikingly longer translocation times when compared to the DNA alone (Wanunu et al., 2009). We propose that a neutral polyethylene glycol (PEG) molecule engineered into the nucleic acid will slow down the DNA unzipping in the nanopore, and a single-base pair mismatch would significantly decrease the unzipping time. We have preliminary data that show the dwell time for e dsDNA with a polyethylene glycol spacer (PEG, isp18 from Integrated DNA Technologies, Coralville, IA) is much longer (around 10 times) than the dsDNA without PEG segment (Figure 4-4A, and its current traces in Figure 4-4B). Because PEG is neutral, the applied voltage has less effect on the dsDNA with the PEG when compared to the dsDNA without the PEG segment.

The results demonstrate that dsDNA with PEG can slow down the unzipping process. More interestingly, when we used a longer PEG segment (2 isp18 spacer, Figure 4-4C), we identified a signature for the unzipping process (Figure 4-4D), and we can clearly see two current levels (two peaks in Figure 4-4E) during the unzipping process. We propose that these two current levels were generated by the dsDNA unzipping and re-zipping at the constriction site (Figure 4-4D lower panel, F). At the unzip scheme (Figure 4-4F, left), the DNA is

occupying the constriction site which generated a lower current blockage (41.1 pA). At the re-zip scheme (Figure 4-4F, right), the unzipped dsDNA is re-zipped and the PEG is occupying the constriction site which generated a higher current conductance (46.7 pA). This proposal is reasonable because the long neutral PEG segment (~3.4 nm) is occupying the most part of the β-barrel stem. So the unzipping process is slowed and stepped back and forth even when a high voltage of 180 mV was applied. This proposal is reasonable because previous studies have found that a PEG and DNA can be distinguished by different reductions in the α-HL channel conductance caused by the presence of these polymers (Sanchez-Quesada et al., 2004; Cockroft et al., 2008).

PEG isp18 spacer length calculation: the isp18 spacer contains six C-C bonds (135pm) and 12 C-O bonds (141pm), so the total length in linear form would be around 2.5nm, but each bond has a tilt around a 120 degree connection with the neighbor bond, so the exact length would be around 1.7nm. The length of two isp18 spacers would be around 3.4nm.

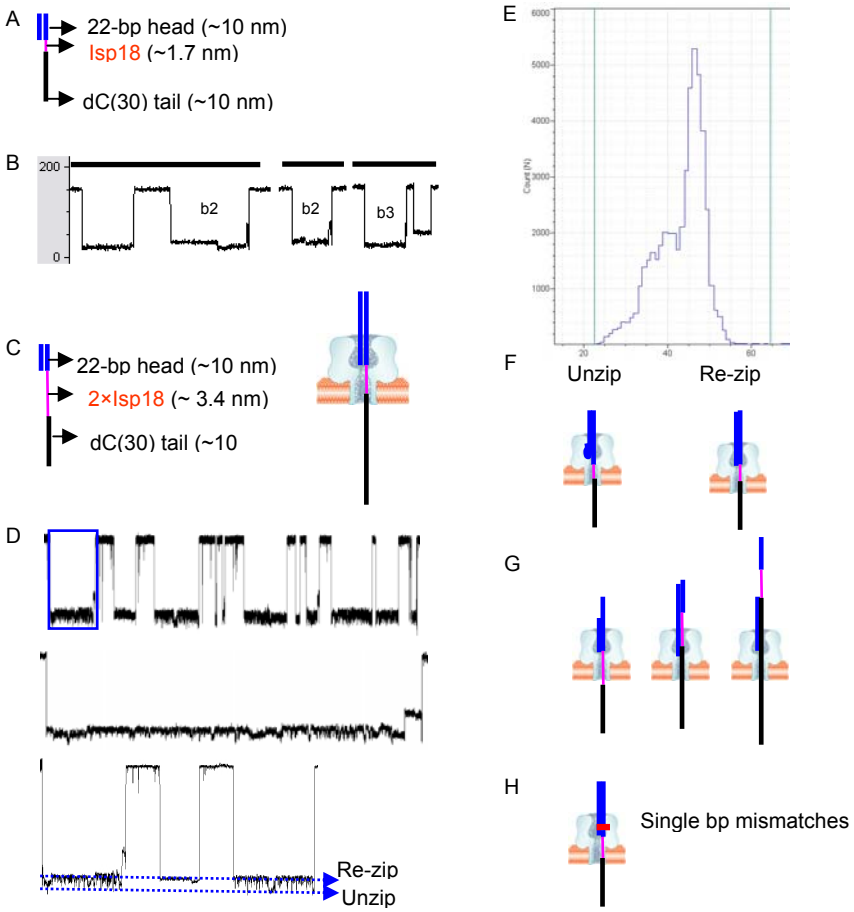

Figure 4-4: The dsDNA with PEG segment unzipping in the nanopore. A: A dsDNA with a short PEG segment (1.7 nm, left panel) and its scheme (right panel) when it was trapped in the nanopore; B: Current traces recorded at 120 mV using the dsDNA in A; C: the same dsDNA with a longer PEG segment (3.4 nm, left panel) and its scheme (right panel) when it was trapped in the nanopore; D: Current traces recorded at 180 mV using the dsDNA in C. Middle panel: enlargement of the current trace in the blue box (upper panel); Lower panel: the traces after filtering, we can clearly see two current levels; E: Histogram of current level during unzipping, we can see two peaks (41.1 pA and 46.7 pA) which correspond to the two current levels in Fig D; F: Scheme for the dsDNA in the nanopore which generate the “unzip” (DNA occupying the constriction site) and “re-zip” (PEG occupying the constriction site) current levels; G: Future negative control experiments: 1) a shorter dsDNA (left panel); 2) a longer dsDNA (middle panel); 3) a dsDNA under the PEG segment (right panel). H: Future experiments for designing a single base-pair mismatches at the end of the dsDNA, which we expect to discriminate the SNA by the dwell time and current levels during the unzipping process.

4.4.2.2 Preliminary data and experiment design

We will carry out several negative control experiments to confirm if this signature current pattern in Figure 4-4D is true: 1) a shorter dsDNA (Figure 4-4G, left panel); 2) a longer dsDNA (Figure 4-4G, middle panel); 3) a dsDNA under the PEG segment (Figure 4-4G, right panel). We expect similar results from 1) and 2) to Figure 4-4 D. The result from 3) should be different than 1) and 2), which will be similar to the dsDNA without PEG segment. Then we will design a single-base pair mismatched at the end of the dsDNA (Figure 4-4H), which we expect to see a significant decrease in the unzipping time, which has been shown in tear unzipping (Wang et al., 2011). We also expect to see a difference between the "unzip" current level (Figure 4-4D) from the full-matched dsDNA and mismatched dsDNA. This expectation is reasonable because 1) the unzipping process is significantly slowed down so the unzipped single base passes through the constriction site slowly; 2) Many studies have shown single base resolution from the current profile in the nanopore (Ashkenasy et al., 2005; Clarke et al., 2009; Purnell and Schmidt, 2009). So we expect different type of bases interact with the constriction site that will generate distinguishable current levels.

4.4.2.3 Discussion

This work will provide a new geometry for DNA unzipping in the nanopore. Besides the tear (fast unzipping) and shear dissociation (all rupturing at once), we want to demonstrate a new geometry by engineering a small molecule into the nucleic acid: slow base pair by base pair unzipping, and then unzipping. If we

can discriminate SNA during this slow unzipping process, it will be useful for studying the SNPs among microRNAs, and could serve as an alternative method. If we can discriminate SNA from the current level because the movement of the single base at the constriction site of the nanopore was slowed down, we can try to identify the current levels for four types of bases: A, T, C and G, which could provide useful information for DNA sequencing and provide some new insights for studying the DNA sequencing in the nanopore.

We have identified the unzipping signature with the PEG molecules and these neutral PEG molecules can slow down the unzipping process. The concern is that we do not know how many base pairs were unzipped in each unzip and re-zip process (Figure 4-4D, F). It could be different each time, for example, it may unzip 3-bp and re-zip, then unzip 4-bp and re-zip. So, in the single-base pair mismatches experiment, we may not be able to see the difference between the "unzip" current level (Figure 4-4D) from the full-matched dsDNA and mismatched dsDNA. That is the possible reason we noticed the sub-current level between the "unzip" and "re-zip" current level (Figure 4-4D, bottom panel), but we still can see a clear "unzip" level which suggests that there is a majority percent that a favored number of base pairs were unzipped at a fixed voltage. Even with this risk, but we expect to see a significant difference in the dwell time for the unzipping process (Figure 4-4D) from the full-matched dsDNA and mismatched dsDNA like we have seen in tear unzipping (Wang et al., 2011).

4.5 To Eliminate the Background RNAs in the Clinic Sample

4.5.1 Introduction

Ribonuclease If (RNase If) is a single strand specific RNA endonuclease which will cleave at all RNA dinucleotide bonds leaving a 5´ hydroxyl and 2´, 3´ cyclic monophosphate. RNase If cleaves single-stranded RNA to mono-, di- and trinucleotides (Meador Iii et al., 1990). It has a strong preference for single-stranded RNA over double-stranded RNA. But RNase If will not degrade DNA. So RNase If would be the perfect enzyme for digesting all the background RNAs after the DNA probe binds to its target microRNA.

Therefore Ribonuclease If (RNase If) will cut all ssRNAs after the specific DNA probe binding to its RNA target and only microRNA-DNA hybrids will remaim after RNase If digestion. Currently, RNA samples from animals and humans are available from many companies (Epicentre, Zyagen, Biocompare, etc). We also want to find a method to detect the microRNA at the tissue level. A further complication is that the total RNA samples from blood and tissues in clinical studies contain a lot of other RNAs including microRNAs, mRNAs and tRNAs, which could possibly have secondary structure (Lee et al., 2002). We tested a rat liver total RNA sample (1 mg/mL, Zyagen, CA). And the sample was treated with RNase-free Dnase to remove DNA. From the single-channel nanopore recordings, we can see many the non-specific signature events (longer than 1ms, red) (Figure 4-5), and the frequency is around 15/min. This frequency is too high and we can not use the same method as we used previously (Wang et al., 2011).

We propose these non-specific signature events were generated by the RNAs with a complex secondary structure. Many studies have shown that ssRNA can form hairpin, loop and other secondary structures which can generate long blockages in the nanopore because these RNAs cannot be translocated through the pore without overcoming the secondary structures (Bundschuh and Gerland, 2005; Butler et al., 2006; Japrung et al., 2010). Other than the long events, we also noticed some spikes with different conductances which were generated by linear ssRNA molecules (Figure 4-5) which are similar to what we have seen from the synthetic ssRNA molecules (Wang et al., 2011).

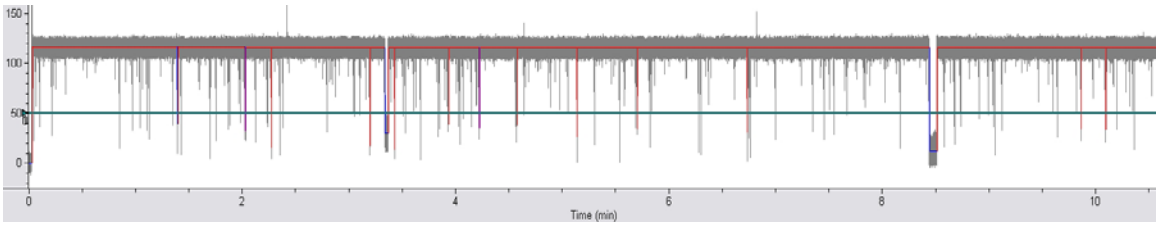

Figure 4-5: The current trace recorded at 100 mV from the total Rat Liver RNA sample (15 ug/mL). We can see lots of the non-specific signature events (> 1ms, red events in the trace), and the frequency is around 15/min. This frequency is too high and we can not use the same method as we used previously (Wang et al., 2011).

4.5.2 Preliminary data and experiment design

Based on the known function of RNase If, we will develop a protocol to denature the rat liver RNA sample with secondary structures into a linear form (Zyagen, CA) (Figure 4-6A) at 90 degrees. Then the DNA probe (P122) will be added to bind specifically to mir-122a (a liver-specific miRNA (Wanunu et al., 2010a)) (Figure 4-6B), then the sample will be incubated with the RNase If at 37 degrees. We would expect only RNA-DNA hybridization remains and all other

non-hybridized ssRNA will be cleaved into mono-, di- and trinucleotide by RNase If. In this way, we will be able to detect the target micRNA (mir-122a here) in a complex RNA background in the tissue level.

The expected nanopore current recordings are illustrated in Figure 4-6C. In the initial control experiment we expect to see only short current blockades due to mono-, di- and trinucleotide digested after RNase If (Figure 4-6C, upper panel). In the experiment with protection with the DNA probe followed by digestion by RNase If, we expect the DNA-RNA hybrids that will generate the unzipping signature event as we have shown previously (Figure 4-6C, lower panel). We expect to also see the short current blockades caused by mono-, di- and trinucleotide which are the products of background RNAs digested by the RNase If.

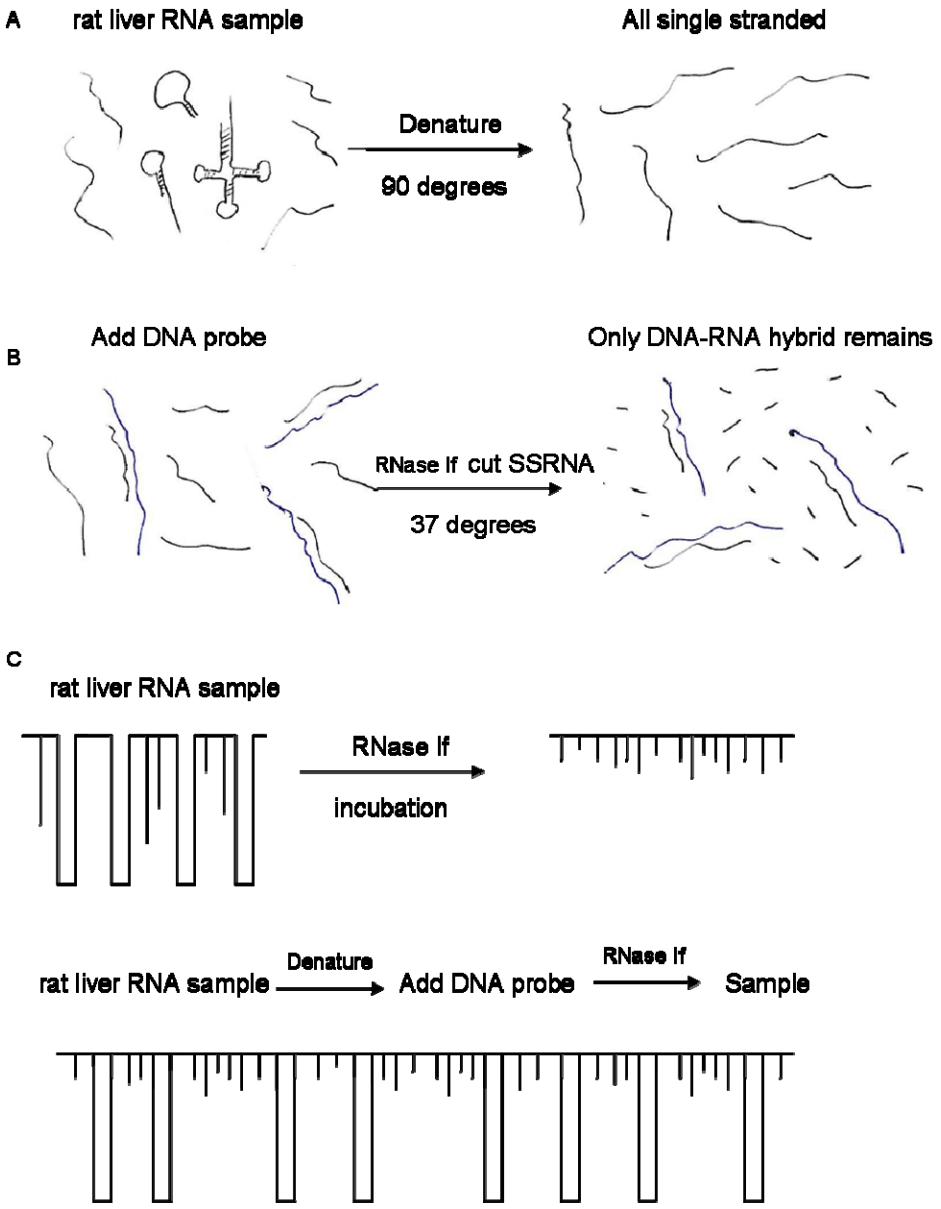

Figure 4-6: Protocol of MicroRNA detection at tissue level. A: rat liver RNA sample includes RNAs have complex secondary structure, which can be denatured at high temperature into ssRNA; B: A DNA probe can be blended with the single stranded RNAs, which will form the DNA-RNA hybrids. After cleavage with the RNase If, only DNA-RNA hybrids remain, all other ssRNA will be digested into mono-, di- and trinucleotide, which can only generate short current blockades (upper panel in C); C: Upper panel: a control test to show that rat liver RNA sample after RNase If digestion will generate short current blockades; Lower panel: the sample with the DNA-RNA hybrids will generate the signature event as we have shown previously (Wang et al., 2011).

4.5.3 Discussion

RNA samples from animals and humans are commercially available. If we can confirm this approach, it will be great to study different microRNAs from different tissues. We can apply this approach to compare the specific microRNA level between the healthy tissue and disease-related tissue. This approach is potentially useful for quantitative microRNA detection, the discovery of disease markers and non-invasive early diagnosis of cancer. The rat liver total RNA sample should not contain any DNAs (ssDNA, dsDNA), so it is very important to confirm the control experiment shown in Figure 4-6C, upper panel. If we still see the non-specific signature events (longer than 1ms, red) (Figure 4-5) after RNase If digestion, this suggests that the sample contains dsDNA or some dsRNA which can not be digested by the RNase If, because RNase If will not degrade DNA, and it has a strong preference for single-stranded RNA over double-stranded RNA. If this control experiment does not give expected results, we may try to digest for a longer time or contact different companies to obtain the workable sample.

4.6 Perspectives

An alternative method we can try is the locked nucleic acid (LNA) technique, where the ribose ring is “locked” by a methylene bridge connecting the 2’-O atom and the 4’-C atom. LNA nucleosides contain the common nucleobases (T, C, G, A, U and mC) and are able to form base pairs according to standard Watson-Crick base pairing rules. LNA makes the pairing with a complementary nucleotide strand more rapid and increases the stability of the resulting duplex which can improve the mismatch discrimination (You et al., 2006). A novel LNA-based, sequence-specific ‘Zorro-LNA’ was developed recently (Zaghloul et al., 2011). LNA has been consider as a promising technique for cancer treatment, and it has been used for in situ detection of microRNA (Kloosterman et al., 2006; Havelda, 2010) and MicroRNA profiling, as well as microRNA-based cancer diagnostics and therapeutics. We also could apply this LNA technique ito our study to see its effect in the DNA unzipping and SNA discrimination ability as well as microRNA detection in the nanopore experiment.

The key component of the nanopore sensor is the probe, the sequence of which is programmable and can be optimized to achieve high sensitivity and selectivity. We expect that the probe composition will not be limited to the four nucleotides. The incorporation of unnatural compounds such as locked nucleic acids and peptide nucleotide acids into the probe sequences may enhance selectivity because of the strengthened hybridization between probe and target. The probe can also be engineered with a specific barcode through chemical modification (Mitchell and Howorka, 2008; Singer et al., 2010) so that multiple

microRNAs can be simultaneously detected using distinct probes. The sensor can detect microRNAs, various nucleic acid fragments, genetic alterations and pathogenic DNAs/RNAs. In the future, nanopore sensor can be devised on new membrane platforms such as a droplet–interface bilayer (Bayley et al., 2008) as well as stable biological-solid hybrid nanopore arrays (Hall et al., 2010) for high-throughput microRNA detection. Overall, the nanopore method can be a useful tool for quantitative studies of microRNAs and the discovery of disease markers, which are important for non-invasive screening and the early diagnosis of diseases such as cancer.

4.7 References

Albrecht, C., K. Blank, M. Lalic-Mlthaler, S. Hirler, T. Mai, I. Gilbert, S. Schiffmann, T. Bayer, H. Clausen-Schaumann, and H.E. Gaub. 2003. DNA: A programmable force sensor. *Science*. 301:367-370.

Ashkenasy, N., J. Sanchez-Quesada, H. Bayley, and M.R. Ghadiri. 2005. Recognizing a single base in an individual DNA strand: A step toward DNA sequencing in nanopores. *Angewandte Chemie - International Edition*. 44:1401-1404.

Bartel, D.P. 2004. MicroRNAs: Genomics, Biogenesis, Mechanism, and Function. *Cell*. 116:281-297.

Bartel, D.P. 2009. MicroRNAs: Target Recognition and Regulatory Functions. *Cell*. 136:215-233.

Bayley, H. 2006. Sequencing single molecules of DNA. *Curr Opin Chem Biol*. 10:628-637.

Bayley, H., B. Cronin, A. Heron, M.A. Holden, W.L. Hwang, R. Syeda, J. Thompson, and M. Wallace. 2008. Droplet interface bilayers. *Molecular BioSystems*. 4:1191-1208.

Bayley, H., and L. Jayasinghe. 2004. Functional engineered channels and pores - (Review). *Molecular Membrane Biology*. 21:209-220.

Branton, D., D.W. Deamer, A. Marziali, H. Bayley, S.A. Benner, T. Butler, M. Di Ventra, S. Garaj, A. Hibbs, X. Huang, S.B. Jovanovich, P.S. Krstic, S. Lindsay, X.S. Ling, C.H. Mastrangelo, A. Meller, J.S. Oliver, Y.V. Pershin, J.M. Ramsey, R. Riehn, G.V. Soni, V. Tabard-Cossa, M. Wanunu, M. Wiggin, and J.A. Schloss. 2008. The potential and challenges of nanopore sequencing. *Nature Biotechnology*. 26:1146-1153.

Bundschuh, R., and U. Gerland. 2005. Coupled dynamics of RNA folding and nanopore translocation. *Physical Review Letters*. 95:1-4.

Butler, T.Z., J.H. Gundlach, and M.A. Troll. 2006. Determination of RNA orientation during translocation through a biological nanopore. *Biophysical Journal*. 90:190-199.

Carthew, R.W., and E.J. Sontheimer. 2009. Origins and Mechanisms of miRNAs and siRNAs. *Cell*. 136:642-655.

Clarke, J., H.C. Wu, L. Jayasinghe, A. Patel, S. Reid, and H. Bayley. 2009. Continuous base identification for single-molecule nanopore DNA sequencing. *Nature Nanotechnology*. 4:265-270.

Cockroft, S.L., J. Chu, M. Amorin, and M.R. Ghadiri. 2008. A single-molecule nanopore device detects DNA polymerase activity with single-nucleotide resolution. *Journal of the American Chemical Society*. 130:818-820.

de Zoysa, R.S., D.M. Krishantha, Q. Zhao, J. Gupta, and X. Guan. 2011. Translocation of single-stranded DNA through the alpha-hemolysin protein nanopore in acidic solutions. *Electrophoresis*. 32:3034-3041.

Diederichs, S., and D.A. Haber. 2007. Dual role for argonautes in microRNA processing and posttranscriptional regulation of microRNA expression. *Cell*. 131:1097-1108.

Franceschini, L., E. Mikhailova, H. Bayley, and G. Maglia. 2012. Nucleobase recognition at alkaline pH and apparent pKa of single DNA bases immobilised within a biological nanopore. *Chem Commun (Camb)*. 48:1520-1522.

Garzon, R., G.A. Calin, and C.M. Croce. 2009. MicroRNAs in cancer. Vol. 60. 167-179.

Gu, L.Q., and J.W. Shim. 2010. Single molecule sensing by nanopores and nanopore devices. *Analyst*. 135:441-451.

Hall, A.R., A. Scott, D. Rotem, K.K. Mehta, H. Bayley, and C. Dekker. 2010. Hybrid pore formation by directed insertion of +Ý-haemolysin into solid-state nanopores. *Nature Nanotechnology*. 5:874-877.

Havelda, Z. 2010. In situ detection of miRNAs using LNA probes. *Methods Mol.Biol*. 592:127-136.

Howorka, S., and Z. Siwy. 2009. Nanopore analytics: Sensing of single molecules. *In* Chemical Society Reviews. Vol. 38. 2360-2384.

Inui, M., G. Martello, and S. Piccolo. 2010. MicroRNA control of signal transduction. *Nat Rev Mol.Cell Biol*. 11:252-263.

Japrung, D., M. Henricus, Q. Li, G. Maglia, and H. Bayley. 2010. Urea facilitates the translocation of single-stranded DNA and RNA through the +Ý -hemolysin nanopore. *Biophysical Journal*. 98:1856-1863.

Kloosterman, W.P., E. Wienholds, E. de Bruijn, S. Kauppinen, and R.H.A. Plasterk. 2006. In situ detection of miRNAs in animal embryos using LNA-modified oligonucleotide probes. *Nature Methods*. 3:27-29.

Kosaka, N., H. Iguchi, Y. Yoshioka, F. Takeshita, Y. Matsuki, and T. Ochiya. 2010. Secretory mechanisms and intercellular transfer of microRNAs in living cells. *J Biol Chem*.

Krautbauer, R., M. Rief, and H.E. Gaub. 2003. Unzipping DNA oligomers. *Nano Letters*. 3:493-496.

Kufer, S.K., E.M. Puchner, H. Gumpp, T. Liedl, and H.E. Gaub. 2008. Single-molecule cut-and-paste surface assembly. *Science*. 319:594-596.

Lee, Y., K. Jeon, J.T. Lee, S. Kim, and V.N. Kim. 2002. MicroRNA maturation: stepwise processing and subcellular localization. *EMBO J*. 21:4663-4670.

Ma, L., and S.L. Cockroft. 2010. Biological nanopores for single-molecule biophysics. *ChemBioChem*. 11:25-34.

Maglia, G., M.R. Restrepo, E. Mikhailova, and H. Bayley. 2008. Enhanced translocation of single DNA molecules through +Ý-hemolysin nanopores by manipulation of internal charge. *Proceedings of the National Academy of Sciences of the United States of America*. 105:19720-19725.

Meador Iii, J., B. Cannon, V.J. Cannistraro, and D. Kennell. 1990. Purification and characterization of Escherichia coli RNase I. Comparisons with RNase M. *European Journal of Biochemistry*. 187:549-553.

Mitchell, N., and S. Howorka. 2008. Chemical tags facilitate the sensing of individual DNA strands with nanopores. *Angewandte Chemie - International Edition*. 47:5565-5568.

Mitchell, P.S., R.K. Parkin, E.M. Kroh, B.R. Fritz, S.K. Wyman, E.L. Pogosova-Agadjanyan, A. Peterson, J. Noteboom, K.C. O'Briant, A. Allen, D.W. Lin, N. Urban, C.W. Drescher, B.S. Knudsen, D.L. Stirewalt, R. Gentleman, R.L. Vessella, P.S. Nelson, D.B. Martin, and M. Tewari. 2008. Circulating microRNAs as stable blood-based markers for cancer detection. *Proceedings of the National Academy of Sciences of the United States of America*. 105:10513-10518.

Movileanu, L. 2009. Interrogating single proteins through nanopores: challenges and opportunities. *Trends in Biotechnology*. 27:333-341.

Olasagasti, F., K.R. Lieberman, S. Benner, G.M. Cherf, J.M. Dahl, D.W. Deamer, and M. Akeson. 2010. Replication of individual DNA molecules under electronic control using a protein nanopore. *Nature Nanotechnology*. 5:798-806.

Ortholan, C., M.P. Puissegur, M. Ilie, P. Barbry, B. Mari, and P. Hofman. 2009. MicroRNAs and lung cancer: New oncogenes and tumor suppressors, new prognostic factors and potential therapeutic targets. *Current Medicinal Chemistry*. 16:1047-1061.

Purnell, R.F., and J.J. Schmidt. 2009. Discrimination of single base substitutions in a DNA strand immobilized in a biological nanopore. *ACS Nano*. 3:2533-2538.

Rabinowits, G., C. GerA el-Taylor, J.M. Day, D.D. Taylor, and G.H. Kloecker. 2009. Exosomal microRNA: A diagnostic marker for lung cancer. *Clinical Lung Cancer*. 10:42-46.

Rosell, R., J. Wei, and M. Taron. 2009. Circulating MicroRNA Signatures of Tumor-Derived Exosomes for Early Diagnosis of Non-Small-Cell Lung Cancer. *Clinical Lung Cancer*. 10:8-9.

Sanchez-Quesada, J., A. Saghatelian, S. Cheley, H. Bayley, and M.R. Ghadiri. 2004. Single DNA rotaxanes of a transmembrane pore protein. *Angewandte Chemie - International Edition*. 43:3063-3067.

Singer, A., M. Wanunu, W. Morrison, H. Kuhn, M. Frank-Kamenetskii, and A. Meller. 2010. Nanopore based sequence specific detection of duplex DNA for genomic profiling. *Nano Letters*. 10:738-742.

Smit, M., E. van der Kooij-Meijs, L.P. Woudt, L.M. Havekes, and R.R. Frants. 1988. Exact localization of the familial dysbetalipoproteinemia associated HpaI restriction site in the promoter region of the APOC1 gene. *Biochem Biophys Res Commun*. 152:1282-1288.

Song, F.J., and K.X. Chen. 2011. Single-nucleotide polymorphisms among microRNA: Big effects on cancer. *Chinese Journal of Cancer*. 30:381-391.

Stoddart, D., A.J. Heron, E. Mikhailova, G. Maglia, and H. Bayley. 2009. Single-nucleotide discrimination in immobilized DNA oligonucleotides with a biological nanopore. *Proceedings of the National Academy of Sciences of the United States of America*. 106:7702-7707.

Strunz, T., K. Oroszlan, R. Schfer, and H.J. Gntherodt. 1999. Dynamic force spectroscopy of single DNA molecules. *Proceedings of the National Academy of Sciences of the United States of America*. 96:11277-11282.

Wang, Y., D. Zheng, Q. Tan, M.X. Wang, and L.Q. Gu. 2011. Nanopore-based detection of circulating microRNAs in lung cancer patients. *Nature Nanotechnology*. 6:668-674.

Wanunu, M., T. Dadosh, V. Ray, J. Jin, L. McReynolds, and M. Drndi. 2010a. Rapid electronic detection of probe-specific microRNAs using thin nanopore sensors. *Nature Nanotechnology*. 5:807-814.

Wanunu, M., W. Morrison, Y. Rabin, A.Y. Grosberg, and A. Meller. 2010b. Electrostatic focusing of unlabelled DNA into nanoscale pores using a salt gradient. *Nature Nanotechnology*. 5:160-165.

Wanunu, M., J. Sutin, and A. Meller. 2009. DNA profiling using solid-state nanopores: Detection of DNA-binding molecules. *Nano Letters*. 9:3498-3502.

You, Y., B.G. Moreira, M.A. Behlke, and R. Owczarzy. 2006. Design of LNA probes that improve mismatch discrimination. *Nucleic Acids Research*. 34.

Zaghloul, E.M., A.S. Madsen, P.M.D. Moreno, I.I. Oprea, S. El-Andaloussi, B. Bestas, P. Gupta, E.B. Pedersen, K.E. Lundin, J. Wengel, and C.I.E. Smith. 2011. Optimizing anti-gene oligonucleotide 'Zorro-LNA' for improved strand invasion into duplex DNA. *Nucleic Acids Research*. 39:1142-1154.

VITA

Yong Wang was born in DangYang, HuBei province, People's Republic of China in September, 1980. He obtained his bachelor degree in Applied Chemistry from HuBei Agriculture College and Master of Science degree in Zoology from NanKai University. He came to United State at 2006 and obtained his second Master of Science degree in Biomedical Sciences in 2009 from University of Missouri-Columbia. He achieved Doctor of Philosophy degree in Biological Engineering in 2012 from University of Missouri-Columbia.