Development of Nanoscale Technologies for Applications in Single-Molecule Sequencing

By

© 2021

Uditha S. Athapattu

Submitted to the graduate degree program in Department of Chemistry and the Graduate Faculty of the University of Kansas in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

_____

Chair: Prof. Steven A. Soper

_____

Prof. Susan Lunte

_____

Prof. Robert C. Dunn

_____

Prof. David Benson

_____

Prof. Elizabeth Friis

Date Defended: 26th August 2021

The dissertation committee for Uditha S. Athapattu certifies that this is the approved version of the following dissertation:

Development of Nanoscale Technologies for Applications in Single-Molecule Sequencing

_____

Chair: Prof. Steven A. Soper

Date Approved: 3rd September 2021

Abstract

Recent advancements in next-generation sequencing (NGS) technologies have completely transformed the way in which quantitative genomics and transcriptomics can be done. However, there are a few limitations associated with NGS that restrict the use of these technologies in a clinical setting. For example, the NGS methods have a complicated workflow, can only read short nucleic acid sequences up to about 150 bases, and because these methods use PCR, errors can be introduced, and important modification information can be lost during the amplification process. Furthermore, most of these methods are time-consuming, costly, and require fluorescence labeling. Therefore, new strategies for nucleic acid sequencing that can provide simpler workflow, longer reads, and amplification-free formats that are rapid, accurate, and low cost are required. We propose a novel method for single-molecule nucleic acid sequencing which can address the limitations in NGS. Our strategy, coined as “exonuclease time-of-flight (XToF),” consists of a covalently immobilized exonuclease/exoribonuclease enzyme in a thermoplastic nanofluidic device. The sequentially released mononucleotides by the processive enzymatic activity will be identified via molecular-dependent time-of-flight (ToF) and current transient amplitudes between two in-plane nanopores embedded in a nanofluidic channel.

In this study, we lay the groundwork for single-molecule RNA sequencing using the XToF method. Using a microfluidic device, we demonstrated the capability to covalently attach exoribonuclease 1 (XRN1) enzyme to a thermoplastic surface. The covalent immobilization of XRN1 to a plastic solid support was achieved using EDC/NHS coupling chemistry. Moreover, the ability of immobilized XRN1 to maintain its catalytic activity to digest canonical and methylated RNA transcripts was demonstrated. The processivity and clipping rate of immobilized XRN1 were secured using single-molecule fluorescence measurements of a single RNA transcript. Moreover, the nanofluidic XToF devices were fabricated in thermoplastics via nano-imprint lithography.

Studies conducted using labeled RNA demonstrated that there was no statistically significant difference in the clipping rate of the immobilized enzyme due to nano-confinement. Currently, we are working on identifying unlabeled ribonucleotide monophosphates using ToF and current transient amplitudes. In the future, the sequential digestion of ssRNA by immobilized XRN1 will be integrated with the identification of released rNMPs to demonstrate the RNA sequencing using the XToF method.

We also report a simple method for tailoring the size of in-plane nanopores fabricated in thermoplastics for single-molecule sensing. We were able to decrease the depth and width of the in-plane nanopores from ~30 x 30 nm to ~10 x 17 nm during the thermal fusion bonding (TFB) process. The cross-sectional area of the in-plane nanopores reduced with increasing pressures during TFB. In-plane nanopore devices assembled at higher TFB pressures showed a higher current transient amplitude when utilized to detect single λ-DNA molecules via resistive pulse sensing. Furthermore, the in-plane nanopore sizes were tuned using TFB pressures to successfully detect ssRNA and single ribonucleotide adenosine monophosphates (rAMPs). However, co-ion exclusion arising due to the high surface charge of the thermoplastic surface reduced the event frequency. To address this issue, the thermoplastic surface was modified using EDC/NHS and ethanolamine. This simple surface modification significantly increased the event frequency from ~1 s^{-1} to >136 s^{-1} for a 100 nM concentrated ssRNA.

To my ever-loving parents, my wife, my sister and all my friends and family for your unconditional love, support, and encouragement throughout this journey.

In honor of my father:

Upali Ananda Athapattu

In memory of my late mother:

Piyaseeli Athapattu

Acknowledgments

First and foremost, I would like to express my sincere gratitude to my advisor, Professor Steven Allan Soper, for his continuous support and guidance throughout my Ph.D. career. I am genuinely grateful for his valuable advice and confidence, which motivated me to carry out the research work successfully. During my Ph.D. research, Dr. Soper encouraged me to develop critical thinking, data analysis, and interpretation skills. I got to work in a range of fields obtaining valuable experiences in areas such as chemistry, engineering, and biology. As a result of his mentorship, I am a better scientist, and I am truly thankful to Dr. Soper for giving me this opportunity. I would also like to thank my dissertation committee members. Prof. Susan Lunte, Prof. Robert Dunn, Prof. David Benson, and Prof. Elizabeth Friis for their valuable comments, advices, time, and support throughout my research career. I am honored to have you all on my dissertation committee.

I would like to thank all my colleagues in the Soper research group, past and present. A special thank you goes to Dr. Maggie Witek, Dr. Matt Hupert, Dr. Kumuditha Weerakoon-Rathnayake, Dr. Charuni Amarasekara, Chathurika Rathnayake, Chad Vietz, Dr. Thilanga Pahattuge, Dr. Anishkumar Manoharan, Dr. Swarnagowri Vaidyanathan and Lindsey Roe. I would also like to thank our collaborators, Dr. Sunggook Park, Dr. Junseo Choi, and Mr. Aaron Nagel, for their support in carrying out the research projects. It was a pleasure working with all of you and I wish the best to each of you in all your future endeavors.

A special heartfelt thank you goes to my family for their endless love, support, and encouragement throughout my life. I am sincerely grateful to my parents, Upali Athapattu and Piyaseeli Athapattu, for their unconditional love and support and all the sacrifices they made to make better lives for my sister and me. Even though my mother is not here with us today, I am

glad to make her dream come true. I would also like to thank my loving sister, Shashini Athapattu, for all the love and support and for taking care of my parents while I am away. You are the best sister anyone could ever have.

I would like to thank all my friends for your support and encouragement in many ways. A special thank you goes to Pubudu, Indeewara, Kumuditha, Loshini, Thilanga, Achala, Akalanka, Amila, Sajith, Manjula, Piyanka, Dhanushka, Chamalee, Gihan, Dimuthu, Sahan, Prabhavie, Sachini, Nilan, Shiwanthi, Chathurika, Sanjaya, Shakya, Chanaka, Harshani, Nishantha, Thilina, Sumudu, Muditha, Nethni, Nalin, Chamani, Kasun, Nishama, Suresh, Meenu, Dishan, Chamiranga, Devinda, Milani, Miyuru and Indu. Thank you for all your help, and I am truly blessed to have you all as my dearest friends. Finally, my heartiest gratitude and thanks go to my lovely wife, Sayuri, for her continuous support and encouragement throughout this journey. I am forever grateful for all the sacrifices you made for me, and I would not have succeeded in finishing my studies on time without your help, love, and support. You were by my side in every situation and thank you for bearing with me all this time. I love you and thank you for everything.

May triple gem bless you all!

Table of Contents

List of Figures

List of Tables

Abbreviations and Acronyms

NGS	Next generation sequencing
SMS	Single-molecule sequencing
SBS	Sequencing by synthesis
tSMS	true single-molecule sequencing
SMRT	Single-molecule real-time sequencing
FRET	Fourier resonance energy transfer
αHL	α-hemolysin
MspA	*Mycobacterium smegmatis* porin A
PNPase	Polynucleotide phosphorylase
ssRNA	Single stranded RNA
dNTP	Deoxynucleotide triphosphate
ddNTP	Dideoxynucleotide triphosphate
rNDP	Ribonucleotide diphosphate
XRN1	Exoribonuclease 1
rNMP	Ribonucleotide monophosphate
EDC	3-(3-dimethylaminopropyl) carbodiimide
NHS	N-hydroxysuccinimide
λ-Exo	Lambda exonuclease
λ-DNA	Lambda DNA
dsDNA	double stranded DNA
nt	Nucleotide

nt s^{-1}	Nucleotides per second
PC	Polycarbonate
PMMA	Poly (methyl methacrylate)
COC	Cyclic olefin copolymer
IMER	Immobilized microfluidic enzymatic reactor
INER	Immobilized nanofluidic enzymatic reactor
IVT	*In vitro* transcription
FLuc	Firefly Luciferase gene, 1,766 nucleotide ssRNA
DMD	Duchenne Muscular Dystrophy gene, 11,163 nucleotide ssRNA
Cas9	CRISPR Associated Protein 9
m6A	N^6- methyladenosine
m5C	5-methylcytosine
Δnt_{ave}	the average length of FLuc RNA remaining after the XRN1 reaction
dNMP	deoxynucleotide monophosphates
EOF	Electroosmotic flow
EDA	Ethanoldiamine
PUA	polyurethane acrylate
ToF	Time-of-flight
XToF	Exonuclease time-of-flight

Chapter 1. Single-Molecule Sequencing (SMS) Technologies

1.1 Introduction

Significant technological advancements in Next Generation Sequencing (NGS) platforms have empowered an explosion of genomic discovery and applications across an extensive range of scientific disciplines over the last ten years. Based on recent market research estimates, the global NGS market will continue to grow through 2025, generating an impressive $20+ billion (USD) with a compound annual growth rate (CAGR) of nearly 20% owing to the great demand. One area of significant emphasis is cancer genomics and the use of NGS as both a powerful and high-throughput discovery tool for interrogating various tumor samples across thousands of patient samples, as exemplified by The Cancer Genome Atlas (TCGA) and the International Cancer Genome Consortium (ICGC). In total, more than 10,000 genomes and exomes across more than 50 different tumor types have been sequenced, resulting in a vast collection of valuable data to process and thus, assemble a road map for critically understanding and attacking the complex landscape of these cancer types at the genomic and transcriptomic level.[1-6]

The substantial growth in the NGS market has primarily been driven by the short-read sequencing-by-synthesis (SBS) paradigm, which continues to dominate the commercial landscape. Although short-read NGS has evolved and sequencing costs have dramatically dropped per base pair sequenced, challenges still prevail: [1] Tedious, and time-consuming sample library preparation is required, which demands technical skill and precision for successful library production. Although automated sample purification and liquid-handling is an option, this requires additional costs as part of the workflow; [2] various kits, reagents, and accessory instruments from multiple commercial vendors are needed for to secure results; [3] numerous quality-control steps within the library preparation step(s) are necessary to verify library performance, such as library size selection, adding additional time and cost to the workflow while also losing sample; [4] at least one step involving PCR amplification is required that can mask

important epigenetic or epitranscriptome information combined with the inability to adequately cover complex genomic regions or regions with high GC content; [5] single or paired-end sequencing reads are typically <300 bp making *de novo* sequencing problematic; [6] cost to purchase an instrument significantly varies from $50K up to ~$1M. Thus, there continues to be a critical need and great opportunity for new technologies that can potentially disrupt the NGS market by further improving the entire sample-to-sequence workflow while providing superior performance capability at reduced capital and reagent costs. In addition, the ability to require less mass input (NGS typically requires ~30 ng of DNA) would accommodate the revolution in new biomarkers, such as liquid biopsy markers that can produce mass inputs in the low picogram range.

This disruption is materializing with the evolution of Single-Molecule Sequencing (SMS), which has become both a viable and compelling alternative within the NGS market since its commercial availability in 2010.[7-8] Moreover, SMS has also spawned a tremendous surge in research and development activities within both the academic and commercial sectors to improve upon existing commercial SMS technologies or develop new, innovative SMS technologies.[9-12] This incredible interest is driven by the considerable potential of SMS compared to short-read NGS, which includes: [1] The use of the native nucleic acid molecules directly retaining crucial epigenetic or epitranscriptome information while eliminating front-end chemical treatment strategies, such as bisulfite conversion; [2] significantly simplifying the entire library preparation steps maximizing cost savings while decreasing time-to-sequence; [3] elimination of a PCR step(s) within the workflow allowing much better sequence interrogation of highly complex genomic regions or regions with high GC content; and [4] providing superior read length by several orders of magnitude. Additionally, nanopore-based SMS further offers lower cost advantages by using electrical readout, thus eliminating the need for fluorescent-labeling and simplifying the

instrument hardware by avoiding sophisticated and expensive optical detection formats.[13] While there are notable challenges with nanopore-based SMS that are well documented,[11, 14-15] such as the poor single base calling accuracy (~80%) and the high input mass requirements (~1 μg) requiring an in many cases an amplification step, technology advancements can improve overall sequencing performance and accuracy.[11, 14, 16-19]

In this chapter we will summarize the current sequencing technologies, emphasizing SMS technologies. First, we will briefly discuss conventional and next-generation bulk sequencing. Next, we will discuss in detail advancements in SMS technologies. Finally, we will briefly discuss the potential application of solid-phase enzymatic reactors and nanoscale electrophoresis for SMS.

1.2 Conventional bulk sequencing

1.2.1 Sanger sequencing

Sanger sequencing, considered the gold standard for DNA sequencing, was first introduced by Frederick Sanger in 1977 and is known as the first-generation sequencing technology.[20] Sanger methods involve a mixed-mode process of synthesizing a complementary DNA template using deoxynucleotides (dNTPs) and termination of synthesis using dideoxynucleotides (ddNTPs).[20-23] ddNTPs lack a 3'-OH group, which is required to form a bond with a 5' phosphate group of another dNTP, preventing further elongation of the DNA strand.[24] Initial Sanger sequencing approaches were split into four reactions. Each reaction contained a ssDNA template, DNA polymerase, four unmodified dNTPs, and a single type of chain-terminating ddNTP, which was labeled.[20] By balancing between dNTP and ddNTP concentrations, a set of nested DNA fragments could be obtained. After primer hybridization and polymerase activity, a mix of varying length DNA fragments is obtained, which are terminated in a ddNTP bearing a reporter. Because different

ddNTPs were added to the four reactions, all combinations of terminations at the 3' end can be produced. The resulting 3' terminated DNA fragments are heat denatured and separated by size using high-resolution gel electrophoresis (see Figure 1.1). Radioactive labeling or, more commonly, fluorescent labeling enables the visualization of the DNA fragments in the gel, which migrate based on their size, enabling the identification of the sequence one-base at a time.[25]

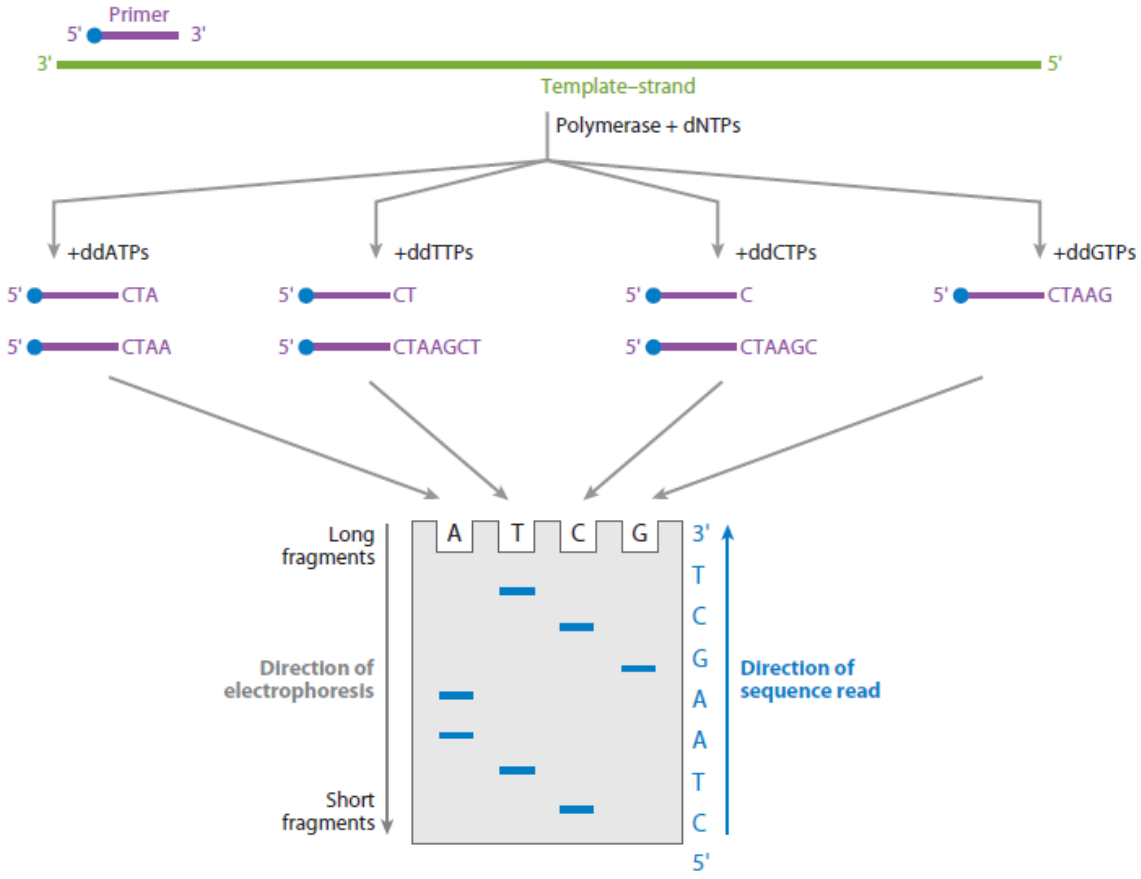


Figure 1.1. The procedure for Sanger sequencing. The template DNA is primed with a primer, and the labeled ddNTPs are added, followed by gel electrophoresis to identify the sequence. (Adapted from reference [37]).

Several improvements in fluorescent dyes, fluorescence detection, and capillary array electrophoresis have made enhancements to the original Sanger sequencing method leading to more automation of the workflow.[25-28] One of the significant advancements includes tagging each of the four different ddNTPs with a specific fluorescent dye. When the chain terminated fragments pass through a laser impinging on the gel positioned in the DNA sequencer, the fluorophore is

excited. It produces a particular fluorescence color based on the terminating ddNTP. Labeling the ddNTPs with spectrally distinct dyes enables a single reaction compared to four different reactions and the use of a single lane during the gel fractionation.[25]

Furthermore, the development of "shotgun sequencing" addressed the read length limitation of first-generation Sanger sequencers where only DNA fragments of length slightly less than 1 kB were sequenced.[29-30] In this method, overlapping DNA fragments were cloned and sequenced separately and finally assembled into a long continuous sequence *in silico*. Additionally, advancements in recombinant DNA technologies and the development of polymerase chain reaction (PCR) provided the high DNA concentrations needed for DNA sequencing[31-33] further advancing the automation of the sequencing process in turn leading to an increase in parallelization, high throughput, and reproducibility.[34]

Even though recent advancements in sequencing technologies have not been related to Sanger sequencing, it is still widely used in many applications, where high throughput is not required, such as validation of genetic variants, verification of plasmid constructs and PCR products, and securing high-quality reads of 300 – 900 bases.[21, 35]

1.2.2 Next-generation sequencing (NGS)

In the mid-2000s with the release of the first high-throughput sequencing technology based on NGS, the cost of sequencing the human genome dropped ~50,000-fold compared to Sanger sequencing.[36-37] Since then, advances in NGS have enabled the use of sequencing in the clinic.[36, 38]

The first step in NGS is DNA library preparation using the target DNA template that needs to be sequenced.[37-38] Briefly, the template DNA is fragmented, and to each end of the fragment, a

synthetic DNA adapter is covalently added using a DNA ligase. These DNA adapters are universal sequences unique to each NGS platform and amplify the DNA fragments using polymerases during the sequencing process. The fragments can be amplified in beads or solid glass surfaces covalently functionalized with complementary sequences to the adapter sequence. Next, amplification of the template DNA fragments is carried out to increase the signal-to-noise during sequencing.[37] A typical NGS sequencing reaction involves introducing nucleotides, adding and detecting the appropriate nucleotide to the sequence, and a wash step that removes the florescent tag/blocking groups for the subsequent attachment of the next nucleotide. NGS is also known as massively parallel sequencing because the sequencing is done on hundreds to thousands to billions of reactions to generate a large number of sequencing data per instrument. In this review, we will discuss two main types of NGS platforms that are being widely used.

1.2.2.1 NGS with reversible dye terminators

In 2007, Solexa Inc. developed an NGS system that involved the use of reversible dye terminators for enzymatic sequencing of amplified DNA fragments. This technology was later acquired by Illumina, Inc.[39] In this method, a DNA library is generated by fragmenting the DNA template, enzymatic trimming, adenylation of the fragment ends, and specific adapter ligation (see Figure 1.2.A).[39] The flow cell of the Illumina system consists of eight microfluidic glass channels that are derivatized with a complementary strand of the library adapter sequences. Next, the library DNA fragments are amplified *in situ* using bridge amplification to create sequencing clusters (see Figure 1.2.B). The fragment ends are released chemically and primed with a complementary synthetic DNA primer to provide free 3'-OH groups for subsequent stepwise extension during sequencing.

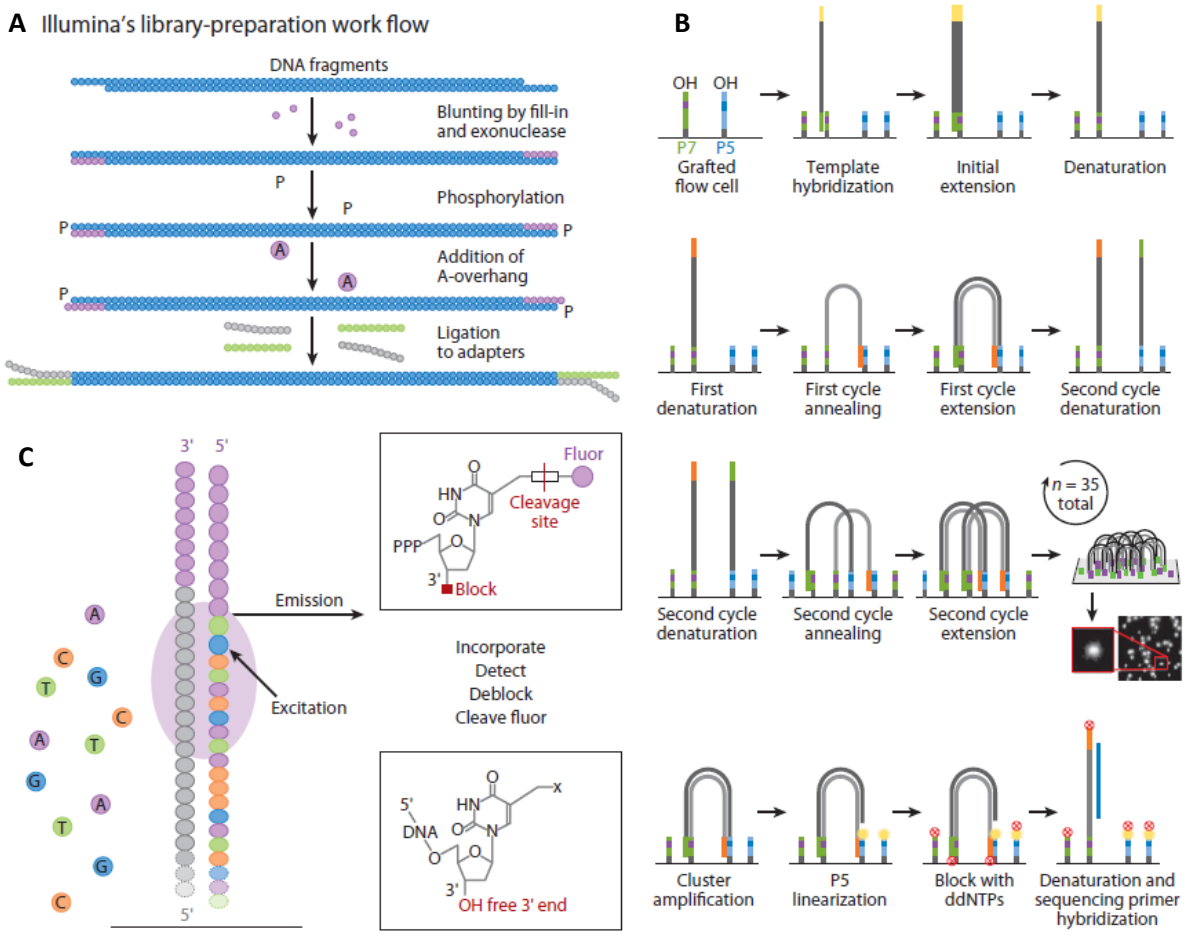


Figure 1.2. Reversible dye terminator sequencing. **(A)** Library construction process in Illumina NGS platform. **(B)** Cluster generation process by bridge amplification. **(C)** Sequencing using reversible dye terminators. (Adapted from reference [37]).

In Illumina reversible dye terminator sequencing, all of the nucleotides are labeled with a dye specific to their identity, and the 3' end is modified with a blocking agent to prevent further incorporation of a nucleotide (*i.e.*, only a single nucleotide is added in each enzymatic step with the subsequent fluorescent read of the added nucleotide).[39-40] The labeled nucleotides are introduced into the flow cell, and the polymerase adds a complementary nucleotide to the target DNA fragment. The unincorporated nucleotides are washed away, and the DNA clusters on the flow cell are imaged to identify the fluorescent signal, which basically corresponds to a single nucleotide in the sequence of the fragment. Finally, the fluorescent tag is chemically cleaved, and

the 3'-OH deblocked paving the way to a subsequent nucleotide incorporation event (see Figure 1.2.C). These sequencing steps are repeated up to 150 times. To read from the opposite end of the fragment cluster, the synthesized strands are removed by denaturation, and the clusters are regenerated by bridge amplification. Then, the opposite end of the fragments is released by a chemical cleavage agent specific to the labile group on the reverse adapter. The fragments are then primed with the reverse primer, and the sequencing is carried out similar to the first read.

Errors in the Illumina sequencing system are mainly substitution errors and are ~0.5%.[37] Read length of the Illumina Hiseq instrument is 150-bp paired-end reads, which is an improvement from the initial 25-bp single-end read, increasing the throughput per instrument.

1.2.2.2 NGS by monitoring pH changes

In 2010 a completely different approach to NGS was introduced by Ion Torrent, which Life Technologies Corp. later acquired.[41] In this NGS platform, hydrogen ions, a byproduct of nucleotide incorporation, are detected and quantified as a change in pH using an innovative Si-based detector. In this method, the library preparation includes fragmentation of the input DNA, enzymatic end polishing, and adapter ligation. Library amplification is achieved using emulsion PCR, where small beads that contain complementary sequences to the library adapter sequence, template DNA, DNA polymerase, and PCR reagents are mixed in equimolar quantities. An emulsion is formed by mixing, and the beads and template DNA fragments are encapsulated in a 1:1 ratio in oil micelles along with PCR reagents. Thermal cycling of the emulsion is carried out, facilitating thousands of individual PCR reactions in-parallel inside oil micelles. After amplification, “emulsion breaking” is carried out to separate the beads that were amplified successfully. Then, a priming sequence is annealed to amplified target DNA fragments on the beads, and the beads are kept in the wells of the Ion Chip, which is the Si device that detects the pH changes within

each well. The Ion chip consists of a microfluidic channel for reagent introduction, and the lower surface of the chip is directly connected to the hydrogen ion detector (see Figure 1.3.A). The hydrogen ion detector converts the pH to the number of nucleotides incorporated in each reaction cycle. Each nucleotide type is introduced one at a time systematically instead of all types of nucleotides at once due to the absence of specific labels for each nucleotide type (see Figure 1.3.B).

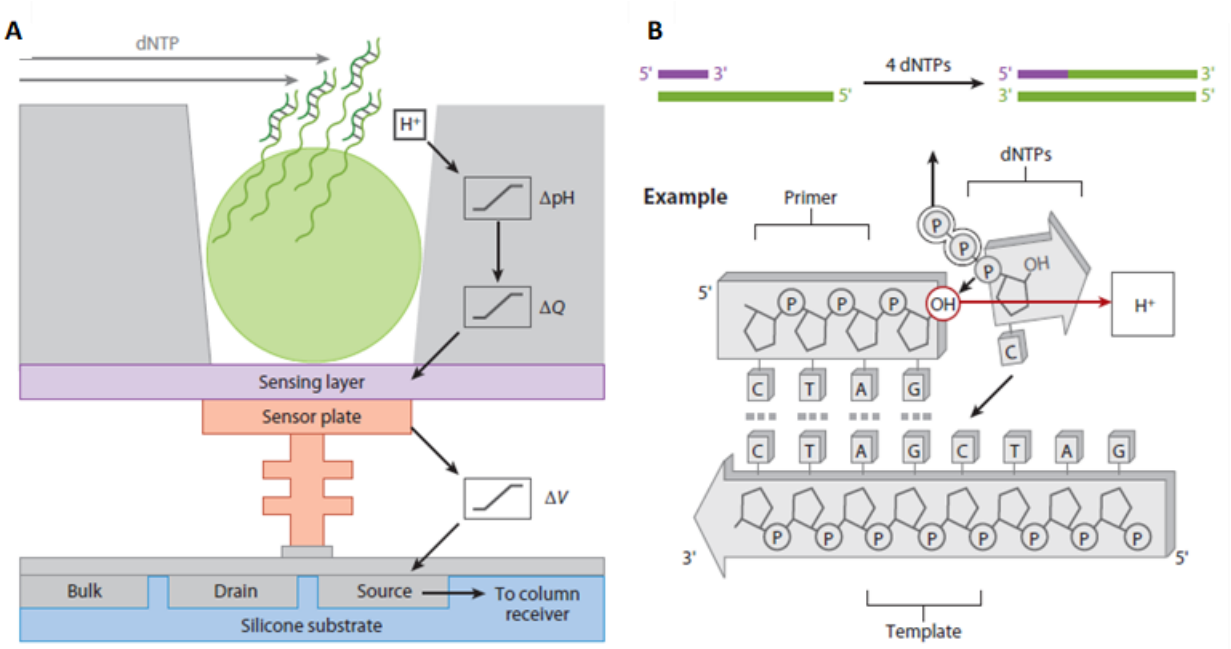


Figure 1.3. Ion torrent sequencing platform. **(A)** Schematic of the Ion chip used in Ion torrent sequencing system. **(B)** Sequence determination using the change in pH. (Adapted from reference [37]).

Since native nucleotides are used, noise arising from fluorescence or blocking groups is not observed in the Ion Torrent sequencing system. The errors in the Ion Torrent system mainly occur from insertions and deletions, which are predominant in homopolymers. Another source of error is phasing, which arises due to differences in nucleotide incorporation in different fragments.[37]

1.2.2.3 RNA sequencing using NGS technology

Since the advent of NGS, there has been a tremendous advancement in RNA sequencing as well.[42-44] In next-generation RNA sequencing, commonly known as RNA-Seq, a population of RNA

molecules are first converted to a cDNA fragment library using reverse transcriptase (RT) with adaptors attached to one or both ends. After library amplification, the cDNA sequences are read either using single-end or paired-end sequencing. In principle, any NGS platform can be used for RNA-seq.

1.3 Single-molecule sequencing (SMS)

Single-molecule sequencing (SMS) is a non-sanger-based method that is considered as the "next-next generation" or "third-generation" sequencing technology. SMS provides answers to some of the most vexing problems with NGS, such as library preparation, DNA template amplification, and sample mass requirements. During sample manipulation and amplification during NGS library preparation, bias and artifacts to the sequence are introduced, which has a detrimental impact on quantitative applications. Especially in RNA-Seq, during the cDNA synthesis, important nucleotide modification information is lost. Additionally, using SMS approaches such as nanopore sequencing, longer reads can be achieved. In this section, we discuss several methods for SMS in detail.

1.3.1 Sequencing-by-synthesis (SBS)

Sequencing-by-synthesis (SBS) is the most commonly used method for nucleic acid sequencing in NGS and third-generation sequencing technologies. The SBS principle involves determining the individual nucleotides incorporated into the nucleic acid strand during synthesis by a polymerase enzyme. In most of these techniques, identifying the particular nucleotide incorporated is determined by measuring the fluorescence of labeled nucleotides. Below, we discuss the most commonly used SBS-SMS methods in detail.

1.3.1.1 True single-molecule sequencing (tSMS)

The first commercially available single-molecule system, HeliScope from Helicos BioSciences, originated from the tSMS method initiated by the work of Braslavsky *et al.*[45-48] In this method, the template DNA is first fragmented and denatured to obtain single-stranded DNA (ssDNA). Then, these ssDNA fragments are polyadenylated and the terminal adenosine of each fragment is fluorescently labeled with a Cy5 fluorescent tag to create a poly(dA)-tailed template library (see Figure 1.4). The polyadenylated DNA is then introduced into a flow cell where millions of poly(dT)-oligonucleotides are immobilized on the cover plate. The polyadenylated target DNA pairs with the poly(dT) oligonucleotides on the cover plate and a CCD camera is used to determine the position of the paired DNA. After imaging, the fluorescent tag on the terminal adenosine is cleaved off. To determine the sequence, labeled nucleotides are systematically added cyclically in the presence of a DNA polymerase.[46-47] After each reaction cycle, the excess unincorporated nucleotides are washed away. Next, the incorporated nucleotides are excited using a laser to determine a successful attachment of the nucleotide. Then, the fluorescent tag is cleaved off the nucleotide for the attachment of the next nucleotide.

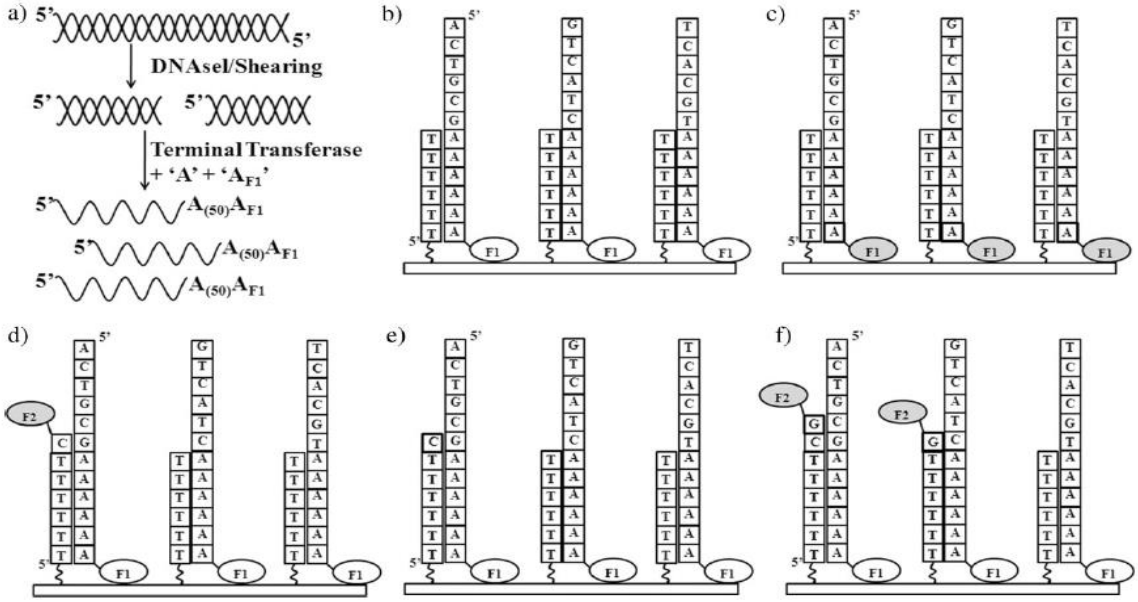


Figure 1.4. True single-molecule sequencing (tSMS) technique. (a) fragmentation, denaturation, 3' polyadenylation, labeling and blocking of the template DNA. (b) hybridization of template DNA onto poly(dT) oligonucleotide bound surface. (c) Imaging of the surface to identify sequencing by synthesis sites. (d) Incubation with one type of fluorescently labeled nucleotide with polymerase for incorporation and washing away the excess molecules and imaging. (e) Cleaving of the fluorescent tag. (f) Introduction of the next type of nucleotide. F1 – Cy3 and F2 – Cy5. (Adapted from reference [48]).

tSMS has been used for the sequence determination of an individual human genome,[49] to quantify a yeast transcriptome,[50] and to resequence the M13 virus genome.[48] When the M13 virus genome was sequenced using tSMS, a method called "two-pass" sequencing was used to reduce the error rate. During the library preparation, asynchronous fragmentation results in dephasing, which leads to misincorporations. This was avoided by monitoring each template molecule discretely, preventing the need for synchronization between different template molecules.[48] Moreover, using the tSMS method, homopolymers can be read accurately compared to NGS methods, where errors in the read are caused when a homopolymer segment is encountered.[51]

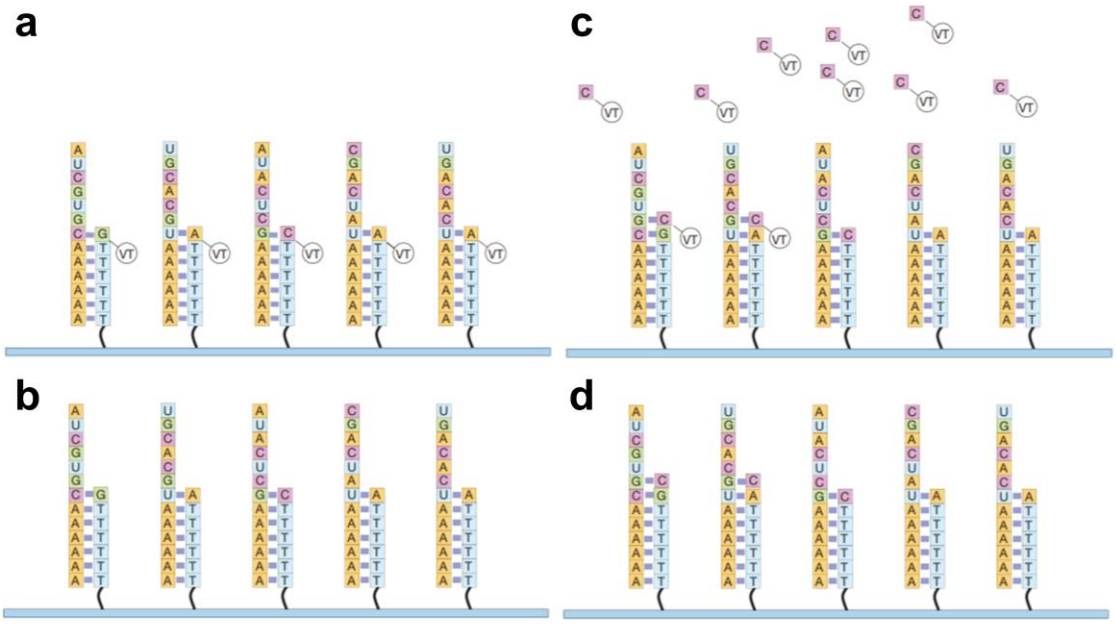


Figure 1.5. Direct RNA sequencing (DRS) method. (a) 3' blocked, polyadenylated RNA capture by the poly(dT) oligonucleotides on the surface. A "fill" using excess dTTP followed by a "lock" step using fluorescently labeled VT-A, -C and -G is are performed. (b) Dye-nucleotide linker is cleaved. (c) Incubation with polymerase and only one type of VT nucleotide. The incorporated nucleotides are located by imaging. (d) Cleavage of the dye allowing the next nucleotide incorporation. (Adapted from reference [52]).

Furthermore, the tSMS method was used by the Helicos Biosciences as the basis to develop an RNA sequencing method called direct RNA sequencing (DRS).[52-53] Like the tSMS, the flow cell used for the DRS method is comprised of a cover plate with an ultraclean glass surface containing millions of poly(dT) oligonucleotides attached at their 5' end to the glass surface.[53] These poly(dT) oligonucleotides are used to capture 3' poly-A containing RNA molecules for priming and initiation of the sequencing. For the RNA sequencing using DRS, the RNA molecules need to be polyadenylated at their 3' end, which is "blocked" at the end by a terminal 3'deoxy nucleotide. For polyadenylation and blocking, rATP and 3'dATP are incorporated, respectively, using either *Escherichia coli* or yeast poly-A polymerases.[53] For RNA with naturally occurring poly-A tails, only the blocking step is needed before hybridization to the flow cell cover plate. Prior to sequencing, each primer-template duplex is filled using excess dTTP and locked in position with A-, C-, and G-Virtual Terminator (VT) nucleotide (see Figure 1.5). These steps prevent misalignments and

make sure that the sequencing starts in the template rather than in the polyadenylated tail.[52] VT nucleotides contain a fluorescent dye and a chemically cleavable group that can prevent the incorporation of another nucleotide.[53-54] Similar to the tSMS method, after washing away the excess unincorporated nucleotides, the attached fluorescently labeled molecules are irradiated with a laser at an angle that causes total internal reflection at the surface, which creates an evanescent field that excites only the molecules close to the surface reducing background fluorescence for single-molecule detection. The single-molecule fluorescence is detected using a CCD camera. After image acquisition, the VT and the fluorescent dye are chemically cleaved, generating a suitable site for the subsequent incorporation event. The sequence-by-synthesis cycle is repeated by adding VT nucleotides followed by rinsing, imaging, and cleavage to generate a large set of images from which the sequence is identified.[52-54]

The DRS technology allows genome-wide analysis to be performed.[50, 55-57] Because native RNA is used, the reverse transcription is avoided resulting in minimum distortion of RNA templates.[52-53] Each DRS run contains ~50 independent channels and produces between 800,000 and 12,000,000 aligned reads that are 25-55 nucleotides in length per channel. Each channel requires ~300 picograms of polyadenylated RNA, and the error rates are in the range of 4%.[53]

1.3.1.2 Single-molecule real-time sequencing (SMRT)

Single-molecule real-time sequencing (SMRT) is a proprietary method developed by Pacific Biosciences.[58-59] This method consists of a SMRT chip which is made of a 100 nm thick aluminum layer on a SiO_2 substrate.[58, 60] Each SMRT chip contains thousands of zero-mode waveguides (ZMW), which are 50 – 70 nm diameter cavities.[58, 61] Inside each ZMW cavity, a single DNA polymerase molecule is immobilized (see Figure 1.6.A).[58, 62] To determine the sequence, the real-time synthesis of a DNA strand from the template DNA is visualized. Each ZMW contains

nucleotides labeled with a fluorophore giving out a specific color based on the molecule's identity.[59]

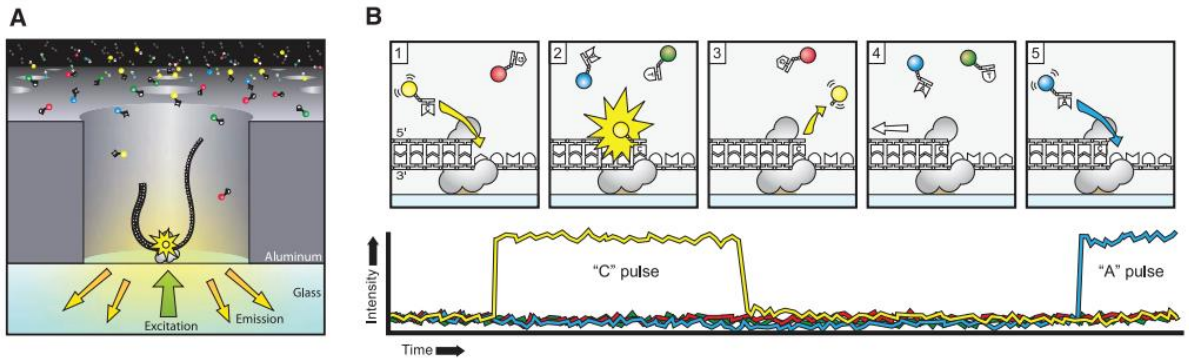


Figure 1.6. Single-molecule real-time sequencing (SMRT) technique. **(A)** Schematic of the ZMW with an immobilizes single molecule of DNA polymerase bound to the DNA template. **(B)** Schematic representation of the sequencing procedure.1 – a labeled nucleotide is incorporated to the template DNA at the polymerase active site. 2 – The fluorescent signal of the corresponding nucleotide is detected. 3 – The phospholinked dye molecule is cleaved due to polymerase activity and drifted away. 4 – the polymerase enzyme moves onto the next nucleotide. 5 – the next nucleotide is incorporated to the template giving a fluorescence signal. (Adapted from reference [59]).

In the SMRT method, the nucleotides are labeled at the phosphate bond rather than at the base like in most other sequence-by-synthesis methods.[63-64] As shown in Figure 1.6.B, when a labeled nucleotide is incorporated, the fluorophore of the nucleotide is lit up due to the laser-mediated illumination of the small detection volume 20 zL (20 x 10^{-21} L).[51] The small volume localizes the excited light into dimensions that cannot be obtainable by techniques such as confocal or total internal reflection microscopy.[65] A holographic optical element, a confocal pinhole array, and a dispersive optical element are used to detect the light emitted. During the incorporation, the nucleotides are held in the detection volume for a longer time than the free-flowing nucleotides in the ZMW. The free-flowing nucleotides diffuse out of the detection volume in microseconds which is not enough for detection. Therefore, ZMW enables observation of the fluorescence of incorporated nucleotides maintaining a high signal-to-noise ratio.[62] During the incorporation of a labeled nucleotide, due to the DNA polymerase activity, the fluorophore attached to the phosphate bond is cleaved. The phosphate-chain-fluorophore complex is quickly

diffused away from the detection volume, making sure the background signal is kept at a minimum level and facilitating the identification of the next incorporated nucleotide. This method enables the nucleotides to be incorporated at a speed of 10 nucleotides per second, promoting the synthesis of a DNA strand that is thousands of nucleotides long in minutes.[66]

1.3.1.3 FRET-based sequencing

A SBS approach using the fluorescence resonance energy transfer (FRET) method also was proposed.[51, 66-67] In this method, a DNA polymerase with a donor fluorophore is used in combination with nucleotides containing acceptor fluorophores specific for each nucleotide. Each time a nucleotide is incorporated, a unique FRET signal is generated due to the close proximity of the donor and the acceptor fluorophores. The continuous monitoring of the FRET signals is used to determine the sequence. The pyrophosphate group which contains the acceptor fluorophore is cleaved, and the FRET signal is quenched prior to incorporation of the subsequent nucleotide.

1.3.1.4 Pyrosequencing

Pyrosequencing is a sequence-by-synthesis technique that monitors the sequence in real-time by detecting pyrophosphate release during DNA synthesis.[68-71] This method uses a cascade of enzyme reactions resulting in the generation of light that is directly proportional to the number of nucleotides incorporated (see Figure 1.7). The reaction starts with nucleic acid polymerization resulting in an inorganic pyrophosphate group PPi being released whenever a nucleotide is incorporated.[72] In standard pyrosequencing, the Klenow fragment of DNA polymerase 1 is used for the DNA polymerization reaction.[73] The released PPi is then converted to ATP-by-ATP sulfurylase.[74] In the presence of luciferase from American firefly *photinus pyralis,* the ATP that was generated is used to oxidize luciferin to generate light.[75] Since the initially added nucleotide

is known, generation of light is used to deduce the sequence. The pyrosequencing technique can be used for both DNA and RNA sequencing.[72] From the initial polymerization to the release of light, the reaction takes around ~4 s. One pmol of DNA releases 6 x 10^{11} of ATP molecules, resulting in more than 6 x 10^{9} photons at 560 nm.[71] This light is usually detected by a photomultiplier tube, photodiode, or a CCD camera.

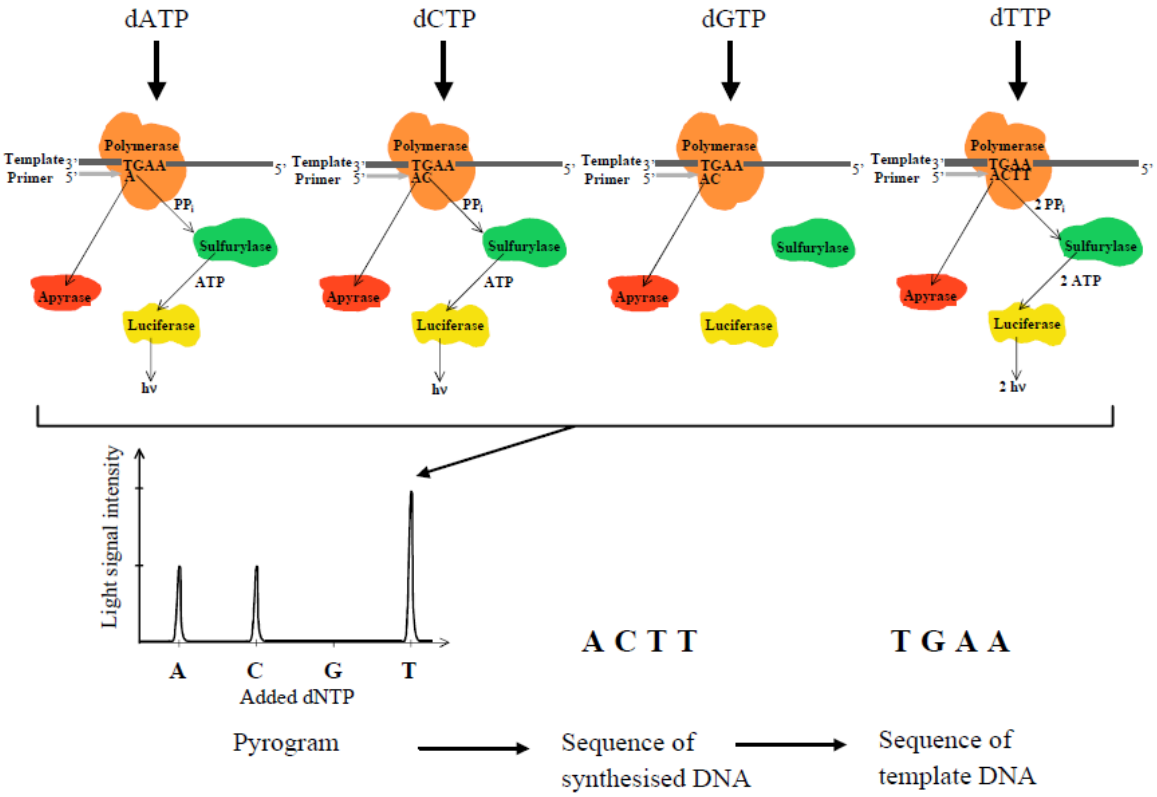


Figure 1.7. Pyrosequencing enzyme system. When the added dNTP is incorporated by the Klenow polymerase, pyrophosphate (PPi) is released. The released PPi is converted to ATP by ATP luciferase which is used to the production of light by luciferase enzyme. (Adapted from reference [72]).

There are two main types of pyrosequencing, solid phase and solution phase pyrosequencing. In solid-phase pyrosequencing, immobilized DNA is used in combination with DNA polymerase I, ATP sulfurylase, and luciferase.[71] After the addition of each nucleotide, a washing step is performed to remove the excess substrate nucleotides. In solution-phase pyrosequencing, in addition to the enzymes mentioned above, an additional enzyme apyrase is added to degrade

the excess nucleotides.[76] The addition of apyrase eliminates the need for the solid support and the washing step needed to remove excess nucleotides enabling the sequencing to be done in a single tube. The pyrosequencing technique has been combined with single-molecule emulsion PCR and developed into SMS technology.[67, 77]

1.3.1.5 Microfluidic devices for sequencing-by-synthesis (μSBS)

Kartalov *et al.* reported a SBS method that uses a polydimethylsiloxane (PDMS) microfluidic device.[78] In this method, a primed DNA template is exposed to a mixture containing a known nucleotide, its fluorescently tagged analog, and DNA polymerase. If the labeled nucleotide is complementary to the nucleotide in the DNA template, the DNA polymerase incorporates the nucleotide and extends the sequence. The excess nucleotides are washed away, and the fluorescence is detected to reveal the DNA sequence. The PDMS channels are first biotinylated, and streptavidin is deposited, followed by immobilization of the biotinylated DNA. Immobilization of the DNA prevents the loss of DNA during feeds and washes. An average read length of 3 bp was demonstrated as a proof-of-concept, and a fully integrated microfluidic system containing active plumbing and parallelism was also shown.[78]

1.3.2 Nanopore sequencing

Even though single-molecule SBS approaches have significantly reduced the cost and avoided error-prone enzymatic DNA amplification processes, they still require fluorescence labeling and expensive imaging and data handling equipment for nucleic acid sequence determination.[51, 66-67] These problems are addressed in nanopore sequencing, where electrical readout is carried out to determine the sequence in a "reagent-free" environment.[15, 79] In nanopore sequencing, a single-stranded nucleic acid is driven through a biological or a solid-state nanopore electrophoretically to obtain the sequence. When the nucleic acids translocate through the nanopore, the current will

be characteristically modulated based on the size of each nucleotide. Moreover, nanopore sequencing approaches do not require DNA amplification by polymerases or labeled nucleotide analogs. Furthermore, nanopore sequencing can deliver longer reads and identify modified nucleotides in both DNA and RNA molecules.[79] Nanopore sequencing is conducted in both biological and solid-state nanopores.[79-83]

1.3.2.1 Biological nanopores

Nanopore sequencing using biological nanopores has been mainly conducted using α-hemolysin (αHL) or MspA protein nanopores.[84] These protein nanopores are found naturally in cell membranes and serve as transportation channels for small molecules or ions to and from the cells. MspA, which is an octameric channel protein, has a funnel-like geometry that narrows down to a ~1.2 nm diameter, with a <0.6 nm long aperture (see Figure 1.8.A).[85-87] The diameter of the sensing region of the MspA nanopore is approximately the same size as αHL, but the length is about an order of magnitude smaller than αHL.[88] One of the problems in nanopore sequencing is the thickness of the pore being larger than that of a single nucleotide leading to current blockage amplitude generation by several nucleotides resident within the pore at a given time.[84] The sensing region of αHL is ~5 nm in length and it was shown that ~12 nucleotides can occupy the nanopore and contribute to the generated current blockage amplitude.[89] This affects the accurate identification of individual bases. In MspA pores, even though the sensing region is <0.6 nm, molecular dynamics studies have shown that nucleotides within ~1.2 nm around the pore, which is ~5 nucleotides, can contribute to the generated current blockage amplitude.[19, 90]

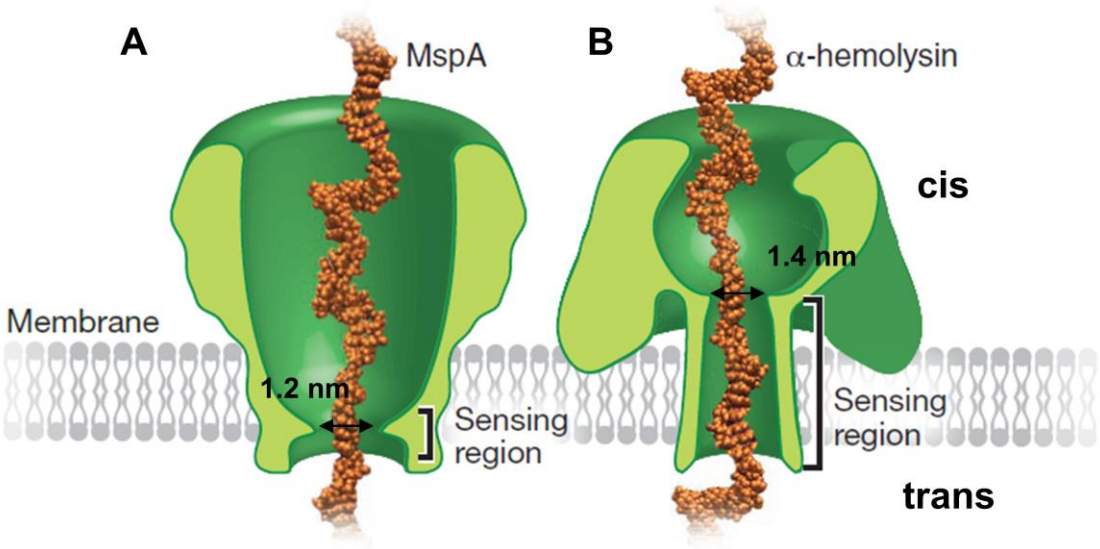


Figure 1.8. Schematic representation of **(A)** MspA and **(B)** αHL protein nanopores. (Reproduced from reference [84]).

αHL is a 293-amino acid protein secreted by *Staphylococcus aureus* as a monomeric protein that forms a mushroom-shaped heptameric pore in lipid bilayers separating a donor chamber (cis) and acceptor chamber (trans).[91] As shown in Figure 1.8.B, on the cis side, the αHL nanopore contains a vestibule, and on the trans side, there is a transmembrane β-barrel. The length of the channel through the αHL pore is ~10 nm. The diameter of the vestibule on the cis side is ~4.9 nm, and the opening is ~2.9 nm. The diameter of the β-barrel is ~2 nm on average. The two domains are separated by a constriction of ~1.4 nm.[91] Currently, protein nanopore sequencing is conducted using two primary methods: (1) Exo-sequencing; or (2) strand sequencing.

1.3.2.1.1 Exo-sequencing

For the exo-sequencing approach, the nucleic acid polymer is cleaved into its constituent mononucleotides using an exonuclease or an exoribonuclease for DNA and RNA, respectively.[80, 92-93] The generated mononucleotides generated by the exonuclease activity are then

electrophoretically driven through the nanopore to identify each nucleotide uniquely by the current transient amplitude.[93]

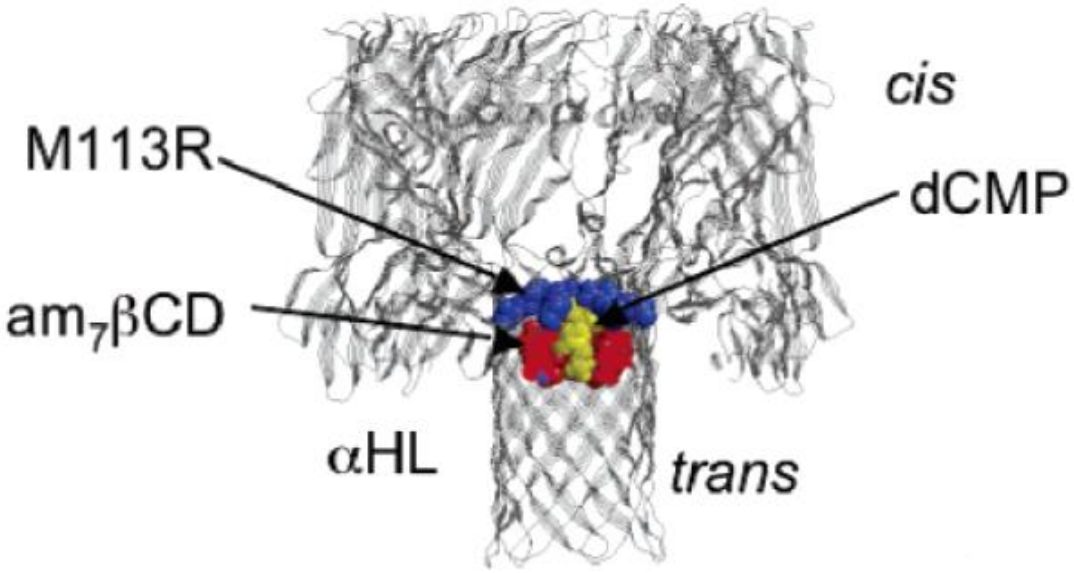


Figure 1.9. Heptameric αHL pore where the Met-113 is replaced with Arg (blue). $am_7\beta CD$ (red) is positioned at van der Waals distances from the Arg side chains. The phosphate groups of dCMP (yellow) ineracts with protonated amines of $am_7\beta CD$ and the cytosine interacts with the guanidinium groups of the Arg side chains. (Reproduced from reference [92]).

Exo-sequencing using the αHL pore was demonstrated in 2006 by Astier *et al.*[92] The identification of ribonucleoside 5'-monophosphates (rNMPs) and dNMPs were carried out in a mutant αHL pore, $(M113R)_7$, engineered with a molecular adapter heptakis-(6-deoxy-6-amino)-β-cyclodextrin ($am_7\beta CD$), which is a β-cyclodextrin with seven primary hydroxyl groups replaced by amino groups (see Figure 1.9). The $(M113R)_7$ αHL pore has a ring of seven arginines near the ~1.4 nm constriction in the β-barrel.[94] The $(M113R)_7$ αHL pore did not show current blockage amplitudes for rNMP or dNMP in the absence of the $am_7\beta CD$ ring.[92] By introducing of the $am_7\beta CD$ ring, the size of the constriction was reduced to a size comparable to the rNMP/dNMP molecule. Therefore, the mononucleotides could be identified with the engineered $(M113R)_7.am_7\beta CD$ αHL pore with an accuracy ranging from 93-98% based on their current transient amplitudes alone.[92] The dwell time measurements, which represents the time the nucleotide is required to translocate through the pore, were too similar for each nucleotide to be distinctly identified.

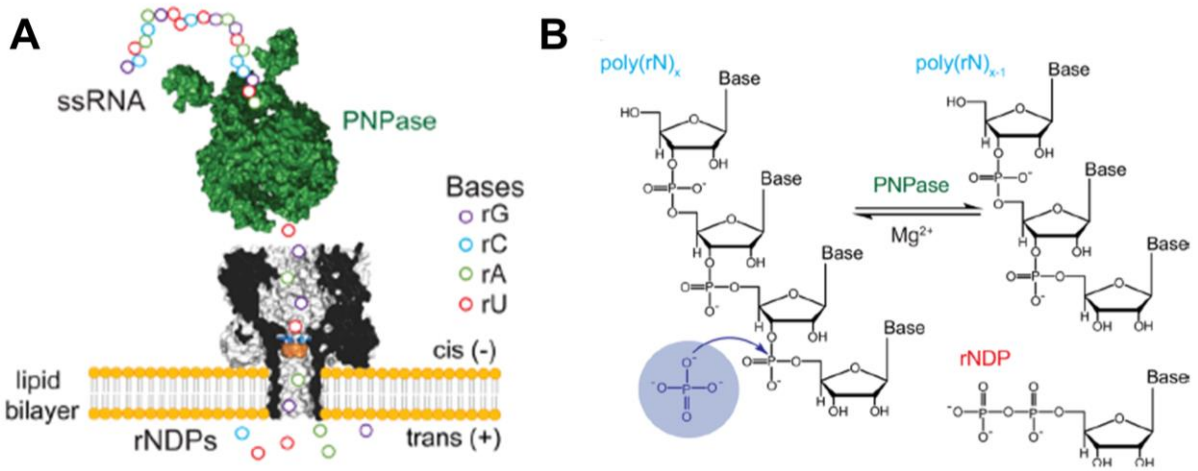


Figure 1.10. Exo-sequencing using PNPase and αHL nanopore. **(A)** Schematic representation of a ssRNA sequentially digested by a PNPase enzyme to produce rNDPs. The rNDPs are translocated through the αHL nanopore for identification using current monitoring. **(B)** The processive reaction of PNPase to produce rNDPs in the presence of inorganic phosphate and Mg^{2+}. (Reproduced from reference [93]).

In another report, Ayub *et al.* demonstrated the use of mutant M113R-RL2 of αHL pore fitted non-covalently with an am_7βCD ring to deduce the sequence of single-stranded homopolymers and heteropolymers.[93] In this report, the engineered αHL pore was used to distinguish all four canonical rNMPs and ribonucleotide diphosphates (rNDPs) successfully. Individual nucleotides interact with the Arg 113 residue and the phosphate groups in the cyclodextrin to successfully identify the mononucleotides. The use of permanently attached cyclodextrin for the identification of rNDPs was also demonstrated. The additional charge on the rNDPs due to the extra phosphate group assisted the capture of the released rNDPs by the nanopore. For continuous sequencing, a permanent attachment of the cyclodextrin ring is desirable. This was achieved by covalently attaching *β*CD within αHL pore in stable orientations using mutations at position 113 and bifunctional linker.[95]

The exo-sequencing capability of the M113R-RL2 αHL pore fitted with am_7*β*CD was demonstrated using polynucleotide phosphorylase (PNPase) (see Figure 1.10.A). The PNPase enzyme cleaves ssRNA molecules from 3'-5' to produce rNDPs in the presence of Mg^{2+} and

phosphate (P_i) (see Figure 1.10.B).[96-97] The pH was kept asymmetrically at pH = 7.0 in the cis compartment and pH = 6.0 in the trans compartment to enable optimum enzyme activity and nucleotide discrimination. The sequence identification of three homopolymers that were 30 nucleotides long each and a heteropolymer with 30 nucleotides containing all the rNMPs were used. The RNA oligomers were introduced into the cis compartment in the absence of P_i and Mg^{2+} to bind with the PNPase enzyme and the subsequent capture of the enzyme by the protein nanopore. The enzyme was prevented from entry into the nanopore due to its size. After that, the enzyme was activated by introducing P_i and Mg^{2+}, and the resulting rNDPs were observed as binding events. The homopolymers could be identified with an overall resolution of 4.4 ±0.8%, and the rNDPs generated from the heteropolymer were distinguished with an overall resolution of 4.1 ±0.8%.[93]

In exo sequencing approach, problems arise such as misordering of the mononucleotides once cleaved from the intact nucleic acid strand and low injection efficiency of the released mononucleotides into the pore. Reiner *et al.* reported that many nucleotides could escape to the bulk before detection by the nanopore in exo sequencing approach.[98] The capture efficiency could be increased by increasing the voltage which will increase the electric field outside the pore but result in an increase in translocation rate of the nucleotides through the pore leading to a decrease in identification accuracy. Due to the low probability of capturing a base by the nanopore which is ~74% to accurately read the sequence the number of nucleic acid templates that need to be read increases exponentially. The capture efficiency can be increased by orienting the exonuclease enzyme so that the cleaved mononucleotide is closer to the nanopore.[98]

1.3.2.1.2 Strand sequencing

In the strand sequencing approach, single-stranded DNA or RNA molecules are ratcheted through the nanopore to identify ionic current transitions with an increased signal-to-noise ratio.[99-100] Due to the high translocation rate of nucleic acid strands through the nanopores, the individual nucleotides travel through the sensing region in ~1-10 μs making it difficult to achieve signal-to-noise ratios required for the identification of picoampere current blockage signals. To achieve the required signal-to-noise ratio the nucleotides must be resident in the pore for >100 μs.[84] Several reports have demonstrated the use of motor proteins for the slow translocation of nucleic acids through the nanopores.[99-101] In 2007, Astier *et al.* demonstrated the use of viral RNA translocase P4 ATPase in combination with αHL pore in strand sequencing of a 5'-$C_{25}A_{25}$-3' oligonucleotide (see Figure 1.11).[102] P4 ATPase, which is a viral packing motor, combines ATP hydrolysis to move along ssRNA.[103] The ssRNA-P4 complex is generated by binding the RNA to a primary binding site on the surface of the P4 enzyme.[104] The P4 binding slowed down the RNA translocation through the pore.[102]

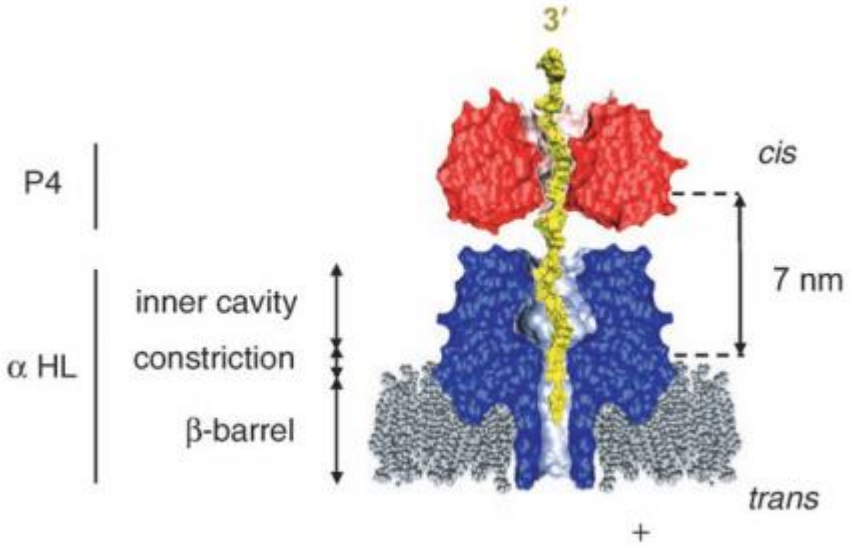


Figure 1.11. Interaction of viral RNA translocase P4 (red) with αHL pore (blue). Oligonucleotide 5'-C25A25-3' (yellow) is translocated through the P4 and αHL. (Adapted from reference [102]).

Several reports have demonstrated the strand sequencing approach using MspA nanopores in combination with phi29 DNA polymerase (DNAP).[19, 105-107] Phi29 DNAP was used to unzip and move double-stranded DNA through the MspA pore. The template DNA strand was bound to the phi29 DNAP at a frayed 5'-3' junction where the polymerase function could not be carried out.[106] Due to the electrical bias, the 5' end of the template DNA was drawn to the MspA pore, and the phi29 DNAP enzyme sat at the entrance without entering into the pore due to steric hindrance (see Figure 1.12). Due to electrostatic forces, the DNA template was pulled through the pore, and the phi29 DNAP acted as a wedge and unwound the double-stranded DNA one nucleotide at a time resulting in discrete current transient amplitudes that were ms in dwell times.[106]

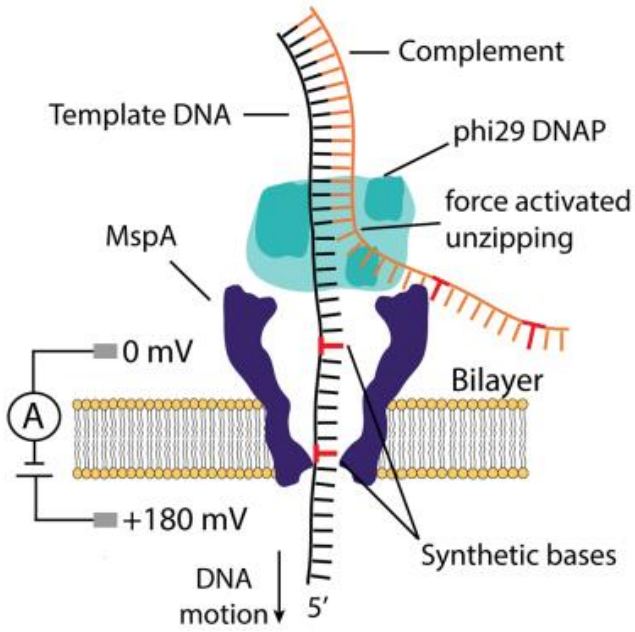


Figure 1.12. Schematic representation of the MspA/phi29 enzyme system. The voltage draws the DNA into the pore with the phi29 polymerase which rests in the rim. The electric filed pulls down the DNA, which results in dissociation of the dsDNA in single nucleotide steps. (Reproduced from reference [106]).

Cherf *et al.* reported the use of phi29 DNA polymerase for forward and reverse ratcheting of DNA templates through the αHL pore.[16] DNA strands could be translocated through the pore with median translocation speeds of 2.5 – 4.0 nucleotides per second without the need for active voltage control. Phi29 DNA polymerase remains bound to the αHL pore ~10000 times longer than

other type of polymerases such as T7 DNA polymerase.[108] Adding only a 23 nt long DNA construct annealed to a synthetic 70 nt DNA template resulted in current blockage amplitudes (22.5 pA) with ~4 ms event duration, whereas the introduction of phi29 polymerase increased event duration to several seconds with current blockage amplitudes that were ~23-24 pA.[16] Around 500 molecules with a rate of ~130 molecules per hour were processed using phi29 polymerase bound single αHL pore with the probability of an insertion and deletion error ranging from 10% to 24.5%.

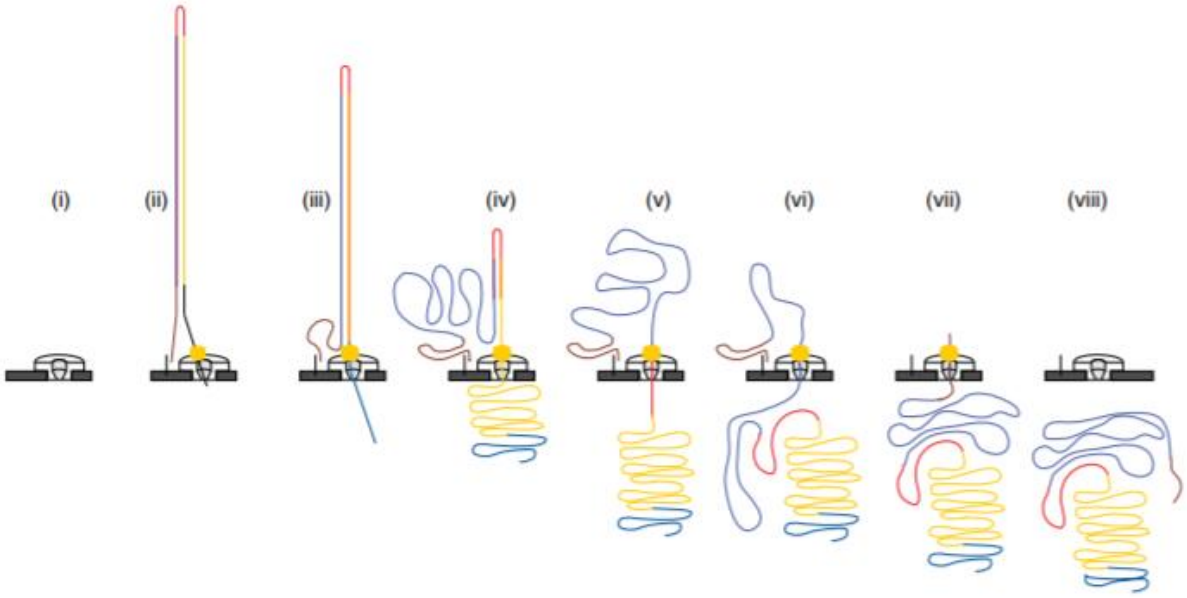


Figure 1.13. 2D read of a dsDNA strand using MinION nanopore sequencer. (i) Open pore. (ii) The template strand with hairpin (red) and adaptors (brown and blue) at the template strand ends. (iii-iv) Template strand (gold) reading. (v-vii) complement strand reading. (viii) return to open pore status. (Adapted from reference [109]).

With the advancements in the strand sequencing approach, in 2014, Oxford Nanopore Technologies (ONT) introduced the MinION nanopore sequencer.[13, 84] MinION is a portable device that weighs around ~100 g and controlled by laptop computer. The flow cell of MinION contains 2048 nanopores embedded in individual stable membranes. The nanopores can be controlled in groups of 512 using an application-specific integrated circuit (ASIC).[13] The template samples are prepared by adding a hairpin structure to one end and adapters to the other end of the genomic

DNA or the cDNA template (see Figure 1.13). The adapters help in strand capture in the cis compartment and the processive enzyme unwinds the DNA strand from the 5' end facilitating the capture of the strand by the nanopore. Due to the hairpin structures both strands can be read (2D read), which improves the base call accuracy of the read. MinION facilitates the read of long continuous strands in the range of 6000 – 48000 bases.[109] The 2D read accuracy of MinION is reported to be ~92%.[13, 84] With the ability to combine multiple reads of the same template and advancements in algorithms for nanopore sequencing will contribute to the further improvements in the read accuracy.[84]

1.3.2.2 Solid-state nanopores

The first solid-state nanopore work was demonstrated in 2001 by Li *et al.*,[110] and since this report, there has been an increasing interest in the use of solid-state nanopores for nucleic acid sequencing.[16, 81-83] Although biological nanopores have been successfully used for nucleic acid sequencing, there are several disadvantages arising mainly due to their limited stability in extreme conditions of salt, pH, temperature, and mechanical stress. Solid-state nanopores have the potential to address the limitations mentioned above associated with biological pores due to their superior mechanical, thermal, and chemical stabilities.[111-112] Additionally, due to the tunability of the pore size, the use of solid-state nanopores has the potential to increase the spatial resolution, thus leading to an increase in the read accuracy.[11, 113] Furthermore, solid-state nanopores can be integrated into other microscale sample processing units for automation leading to higher throughput.

Solid-state nanopores have been mainly fabricated in inorganic thin-film substrates using techniques such as charged particle beams,[110, 114-116] electrical breakdown,[117] exposure with a defocused beam of electrons,[114] ions,[118] thermal heating,[119] focused ion beam (FIB) milling,[120] and

deposition of materials. To achieve accurate discrimination of the nucleotides, the thickness of the nanopores must be reduced.[81] In protein nanopores, the length of the pore is fixed and is much larger than the thickness of a single base, whereas, in solid-state nanopores, ultrathin layers such as thinned SiN, atomically thin layers of graphene, molybdenum sulfide (MoS_2), and boron nitride can be used to increase the spatial resolution.[11] Several groups have also reported the use of atomic layer deposition, reactive ion etching, helium ion beam etching, and photothermal processes induced by a laser to achieve layer thicknesses <5 nm.[121-125]

Generally, ionic current blockage measurements have been the method used to identify nucleic acid sequences using nanopores. In 2013, Venta *et al.* demonstrated the use of SiN nanopores with 0.8 - 2 nm diameter and 5 – 8 nm thickness to resolve 30 base long ssDNA homopolymers by differences in ionic current signals.[122] In 2015, Feng *et al.* used an ultrathin MoS_2 nanopore with a 2.8 nm diameter to resolve between all four 30 bases long homopolymers and all four mononucleotides.[126] In this method, a high viscosity room temperature ionic liquid (RTIL) 1-butyl-3-methylimidazolium hexafluorophosphate ($BmimPF_6$), which is biomolecule friendly solvent with good ionic conductivity,[127] was used to slow down the translocation speed of single nucleotides, increasing the temporal resolution while maintaining a signal-to-noise ratio of 10. Welch's t-test conducted using ionic current drops of 500 - 3000 events showed that the four mononucleotides could be discriminated with a *p*-value <0.0001.[126]

In addition to ionic current blockage detection, several other reports have demonstrated innovative methods for solid-state nanopore sequencing. In one such report, Soni *et al.* reported a novel solid-state nanopore DNA sequencing method using optical detection.[83] In this method, two single-molecule detection methods were combined, optical detection using total internal reflection fluorescence (TIRF) and nanopore assisted DNA unzipping. First, template DNA was fluorescently labeled with specific markers using the Design DNA polymer (DDP) labeling

procedure. In this procedure, each nucleotide in the original DNA template was substituted with a 3-16 nucleotide long DNA fragment. Then, fluorescently labeled oligonucleotides (molecular beacons) with the complementary sequence to the converted DNA were hybridized. When the DNA was electrophoretically driven through a nanopore <2 nm in diameter, which was fabricated in a thin SiN layer using an electron beam, the molecular beacons were sequentially peeled off, giving fluorescence, which could be associated to the original nucleotide in the template DNA.[83] A combination of 488 nm and 640 nm diode lasers were used for illumination. Images were acquired at a speed corresponding to ~5 ms per image.

In another innovative approach, Lee *et al.* introduced nanoelectrode-gated detection of individual molecules for rapid DNA sequencing.[128] The principle of identifying single molecules using nanoelectrodes involves the preparation of nanoelectrodes with a sub-nanometer scaled gap and monitoring the perturbation in the tunneling current signal when a molecule translocates through this gap.[128] In 2008, Tsutsui *et al.* demonstrated the preparation of a nanogap electrode that had a gap between 0.5 – 10 nm using a mechanically controllable break-junction (MCBJ) method.[129] Based on these findings, Tsutsui *et al.* successfully identified single nucleotides by tunneling current using a 1 nm gapped gold electrode.[130] In 2012, Ohshiro *et al.* successfully identified the four canonical dNMPs and rNMPs using the tunneling current as the mononucleotides passed through the nanoelectrodes that possessed a gap of 0.8 nm.[131]

1.3.4 Exonucleolytic degradation

The fact that single fluorescent molecules can be identified accurately in short times has allowed for different approaches for single-molecule sequencing.[132] A single RNA or DNA molecule with fluorescently labeled mononucleotides attached to a solid support can be sequentially digested by an enzyme to produce mononucleotides.[133-135] These released and

labeled mononucleotides can then be detected in the order of their release from the biopolymer using fluorescence detection. High-speed sequence analysis can be achieved depending on the rate of enzymatic degradation.

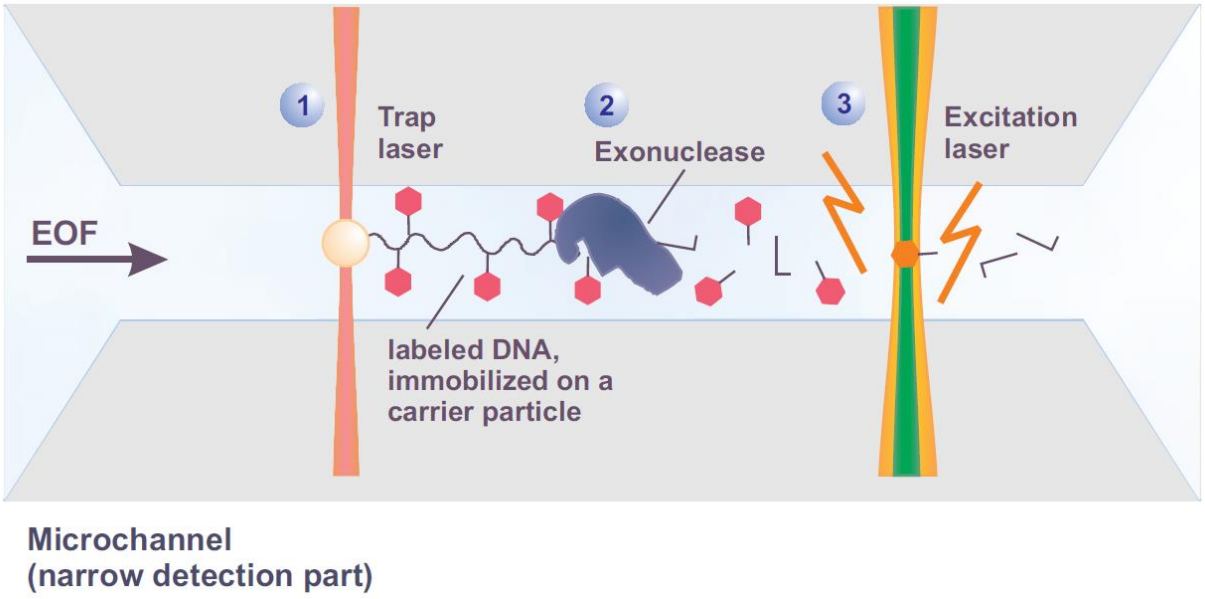


Figure 1.14. The schematic representation of the exonucleolytic sequencing. 1- A bead with immobilizes labeled DNA is held in position by a trap laser. 2 – The labeled DNA is sequentially digested by an exonuclease enzyme. 3 – The released nucleotides are electrophoresed through the detection laser and the fluorescence characteristics such as emission wavelength and fluorescence lifetime are measured. (Adapted from reference [133]).

The successful identification of single mononucleotides by fluorescence to identify the sequence requires replacing the natural nucleotides with fluorescently labeled nucleotides.[133] This can be achieved via DNA amplification using fluorescently labeled nucleotide analogs during PCR or primer extension. Additionally, the DNA needs to be 5' biotinylated in order to immobilize onto a bead that is primed with streptavidin. The concentration of the biotinylated DNA to bead ratio can be controlled to achieve a single DNA molecule immobilized on a single bead. Beads with an immobilized enzyme can be separated from the rest by measuring fluorescence. Using an "optical tweezer," which is an IR trap laser (λ = 1064 nm), a bead with immobilized DNA can be drawn and fixed into a capillary or microchannel (see Figure 1.14). Next, sequential DNA digestion is initiated by introducing a 3'-5' exonuclease enzyme. The sequentially released mononucleotides

are then transported through a laser focus by electro-osmotic flow to excite the fluorophores. The emitted fluorescence characteristics, such as the wavelength and fluorescence lifetime, are measured to identify each nucleotide uniquely.[133] Usually, the four different nucleotides are labeled with four different fluorophores that show discrete emission spectra or different lifetimes, or a combination thereof. It is essential to identify all the nucleotides that are released by the exonuclease; therefore, the entire microfluidic detection area is illuminated by the laser.[133]

1.3.5 Motion-based sequencing

Greenleaf *et al.* reported a single-molecule motion-based sequencing method that relied on resolving the motion of a processive *Escherichia coli* RNA polymerase (RNAP).[12] In this method, a pair of optical traps float two polystyrene beads, where one is attached to an RNAP enzyme and the other to the end of a DNA template. The transcriptional motion of the RNAP enzyme along the DNA template causes changes in the DNA tether joining two beads resulting in shifts that can be registered with angstrom-level accuracy, which facilitates single-base pair resolution.[136] When the DNA polymerase assay is carried out in a buffer with one of the four nucleotides in very low concentrations, the RNAP will be tempted to pause at each position that requires the incorporation of the low concentrated nucleotide.[12] The sequencing of DNA is carried out four times with each dNTP added in low concentrations in turn, and the sequence deducted directly from the order of the pauses. To obtain the sequence, the pause patterns from the four runs need to align within one base pair.[12, 137]

1.3.6 Scanning tunneling spectroscopy-based sequencing

Xu *et al.* reported the base-specific electronic signatures of DNA bases deposited on an Au (111) surface using ultra-high vacuum scanning tunneling microscopy (STM).[138] In 2008, Shapir *et al.* used STM to resolve the electronic structure of single DNA molecules.[139] In 2009, the first step

towards DNA sequencing using STM was demonstrated by Tanaka *et al.*[140] In this method, single DNA molecules were deposited on a copper surface using the oblique pulse-injection method. It was shown that guanine bases manifest a discrete electronic state that can be distinguished from other nucleic acid bases. The "electronic fingerprint" of the guanine bases was identified in the DNA molecule by comparing it with an M13mp18 single-stranded phage DNA with a known sequence. Meng *et al.* reported the theoretically 100% base identifications using electronic characteristics among the four DNA bases.[141]

1.4 Solid-phase enzymatic bio-reactors and their applications in SMS

Solid-phase enzymatic reactors, such as immobilized microfluidic/nanofluidic enzymatic reactors (IMERs/INERs) are a sub-class of solid-phase bioreactors that recently have gained interest as tools in areas such as chemical and biological engineering, nanotechnology, and single-molecule enzymology.[142-145] In IMERs and INERs, enzymes are immobilized on solid surfaces for various biotechnology applications. The attachment of enzymes on these solid-phase enzymatic reactors is plausible using an abundance of available attachment chemistries to surfaces such as silica, glass, and thermoplastic polymers.[146-149] There are numerous advantages of solid-phase enzymatic reactors compared to their liquid-based counterparts, such as enhanced stability, increased rate of catalytic activity, reduced interference in terms of downstream analysis, prevention of aggregation and autodigestion, and reusability of the enzymes.[150-155]

1.4.1 Enzyme Immobilization strategies

Immobilization of enzymes involves attaching the enzyme to a specific surface. The immobilization of enzymes usually causes a change in their conformation, leading to a change in catalytic activity.[156] In most cases, these conformational alterations lead to a decrease in enzyme activity but also can lead to an increase in stability against extreme conditions of temperature,

pH, and organic solvents.[157-159] There are several methods for enzyme immobilization, such as physical adsorption, covalent attachment, cross-linking, or encapsulation (see Figure 1.15).[160]

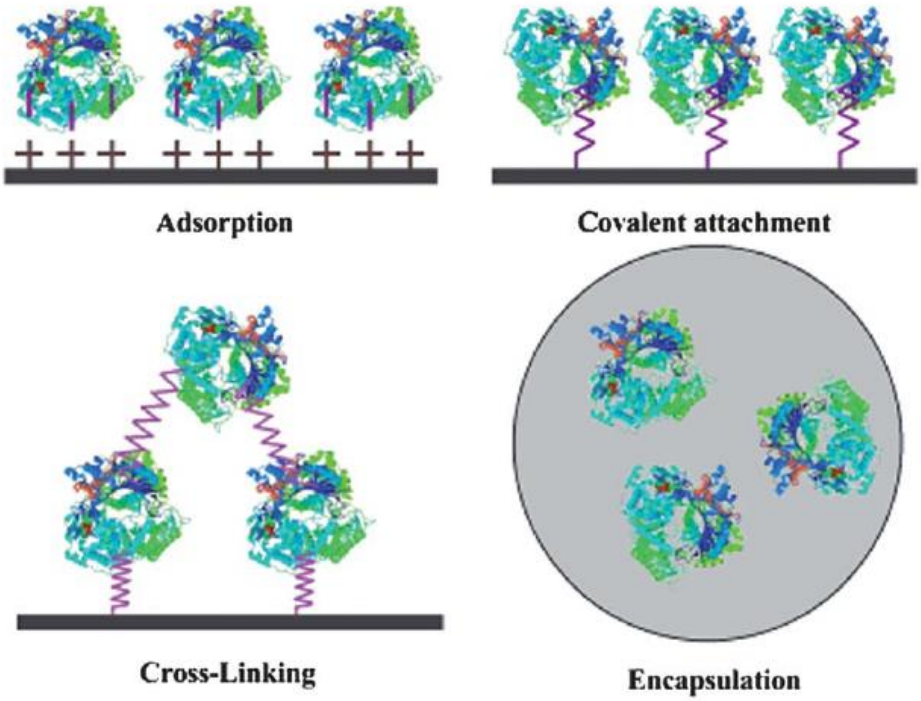


Figure 1.15. Commonly used enzyme immobilization methods. (Adapted from reference [160]).

Physical adsorption of an enzyme on an organic or inorganic carrier surface is one of the simplest methods for enzyme immobilization.[161] In this method, the enzyme is physically adsorbed on a polymer bead or membrane surface by inserting the solid support in an enzyme solution and incubating to allow the adsorption onto the surface.[160] Another approach involves the drying of the enzymes on the surface of an electrode and rinsing away the unadsorbed molecules.[162-163] Physical adsorption requires only mild conditions and no special reagents. Additionally, the activity of the enzyme is well preserved in this method. However, disadvantages of this method include the leaching of the enzyme due to the weak forces between the surfaces, such as van der Waals forces, ionic and hydrophobic interactions.[161, 164] In addition, the physical adsorption of the enzyme onto the surface can cause either masking the active site and/or denaturation of the enzyme in both cases making the immobilized enzyme inactive.

Covalent bonding of enzymes to a solid surface results in the most reliable attachment. In this technique, functional groups on the enzyme and the surface are conjugated, leading to decreased enzyme movement.[160] Due to the covalent bond formation, enzyme leaching is prevented, and the enzyme's stability is increased. However, due to conformational changes and restriction of movement, in some cases, the enzyme activity is decreased.[165-166] A number of functional groups in the enzyme can be utilized for covalent attachment, including amino, carboxyl, alcohol, thiol, and phenol groups. Several covalent attachment chemistries are used for conjugation, such as 1-Ethyl-3-(3-dimethyl aminopropyl) carbodiimide (EDC), N-hydroxysuccinimide (NHS), imine, epoxide, azide click chemistry, and photo-immobilization.[167-169]

Another widely used approach for enzyme immobilization is cross-linking by intermolecular reactions using bifunctional reagents.[160] In this approach, either chemical or physical methods are used for cross-linking. Example bifunctional reagents used for covalent linkages include dimethyl suberimidate, dimethyl adipimidate, succinimidyl trans-4-(maleimidylmethyl), cyclohexane-1-carboxylate (SMCC), and formaldehyde.[160] However, even though immobilization by cross-linking provides a very strongly bonded enzyme, due to the aggregation enzyme activity may be lost.[170] Another disadvantage is the use of harsh chemicals leading to enzyme denaturation.

Encapsulation of an enzyme into semipermeable polymer matrices is another immobilization method.[171-173] Due to the absence of covalent bonds between the substrate and the enzyme, the activity and the stability of the enzyme can be improved compared to covalent attachment of the enzyme to a solid support. Additionally, leaching of the enzyme is also prevented. Even though substrates and products with low molecular weight are allowed to pass through the membrane, the high molecular mass macromolecules are prevented from accessing the enzyme.[174] Various support materials for encapsulation have been used, including sol-gels, polymers, sol-gel/polymers, and inorganic material.[160]

1.4.2 Solid-phase enzymatic reactors for SMS

Solid-phase enzymatic reactors have mainly been used for basic biomedical research, such as solid-phase studies for leukemia treatment, immobilized urease chambers for the study of artificial kidneys, and carbonate dehydratase utilized artificial lungs.[175-177] Another research area of interest is the application of solid-phase enzymatic reactors for the SMS of nucleic acids. The most notable use of a solid-phase enzymatic reactor in SMS is the Pacific biosciences, SMRT sequencing technology (discussed in detail in section 1.3.1.2).[59] First, a single DNA polymerase enzyme molecule is immobilized at the bottom of each ZMW fabricated in a Si substrate. Then, the sequence of the template DNA is determined using a SBS approach, where the incorporation of a labeled nucleotide is identified using fluorescence.

In 2014, Oliver-Calixte *et al.* reported a thermoplastic IMER consisting of covalently attached λ-exonuclease (λ-Exo) for the sequential digestion of a λ-DNA molecule.[178] λ-Exo is an exonuclease that sequentially cleaves dsDNA from its 5' end to the 3' end and releases dNMPs. λ-Exo was covalently attached to a polymethylmethacrylate (PMMA) thermoplastic microfluidic device using EDC/NHS coupling chemistry for the attachment of primary amine-containing biological entities, such as the $–NH_2$ group of lysine. The immobilized λ-Exo manifested an increased apparent processivity of >40 kb compared to the free solution processivity of 3 kb and a clipping rate of 1100 ±100 nucleotides s^{-1}.

The use of a thermoplastic as the solid support has several advantages compared to other substrate materials such as glass or silicon. Thermoplastics are low-cost materials with favorable biocompatibility, good optical properties, and solvent and acid/ base resistant properties.[179] Moreover, there are many simple, cost-effective strategies for fabrication and surface modification of thermoplastics compared to glass and Si.

Recently, Athapattu *et al.* successfully immobilized the exoribonuclease-1 (XRN1) to a microfluidic device consisting of micropillars (discussed in detail in Chapter 2).[180] XRN1 is a processive exoribonuclease that cleaves monophosphorylated ssRNA molecules from the 5' to 3' end producing rNMPs. The immobilized enzyme, covalently attached using EDC/NHS coupling chemistry to a polymer surface, showed increased digestion efficiency compared to its solution-phase counterpart. Moreover, the immobilized XRN1 enzyme showed a significant increase in both clipping rate and processivity, 26 ± 5 nucleotides s^{-1} and 10.5 kb, respectively, when surface-immobilized compared to the free solution enzyme. The ability of the immobilized XRN1 enzyme to sequentially cleave through ssRNA sequences that contain modified nucleotides was also demonstrated. In addition, XRN1 can cleave through secondary structure in the ssRNA molecule.

These recent findings provide the groundwork for a novel single-molecule exosequencing approach that determines the identity of the sequentially cleaved dNMPs and rNMPs from an intact nucleic acid biopolymer.[178, 180] In conventional nanopore sequencing, the nucleotides are identified using current transient amplitudes alone, resulting in lower identification accuracy. However, the nanofluidic solid-phase enzymatic reactors can be integrated into other microfluidic/nanofluidic systems to increase the identification accuracy by coupling the current transient amplitudes with nanoscale electrophoresis.[181-182]

1.4.3 Identification of mononucleotides using nanofluidic devices

The use of nanofluidic devices for the identification of biomolecules has gained attention recently due to unique nanoscale phenomena, such as electric double layer (EDL) overlap and solute/wall interactions due to high surface-to-volume ratio leading to unique electrophoresis modalities.[183-184] The electrophoretic mobility of molecules in nanofluidic devices depends on factors such as ion valence, zeta potential, mobility of ions, and EDL thickness. Additionally, due to the increased

surface-to-volume ratio, there are potential solute/wall interactions, such as hydrophobic, electrostatic, and van der Waals interactions, which can influence the electrokinetic transport of biomolecules through nanofluidic channels.[185] Moreover, when the surface roughness is similar to the critical dimension of the nanochannel, field-dependent mobilities are observed due to intermittent motion of solutes.

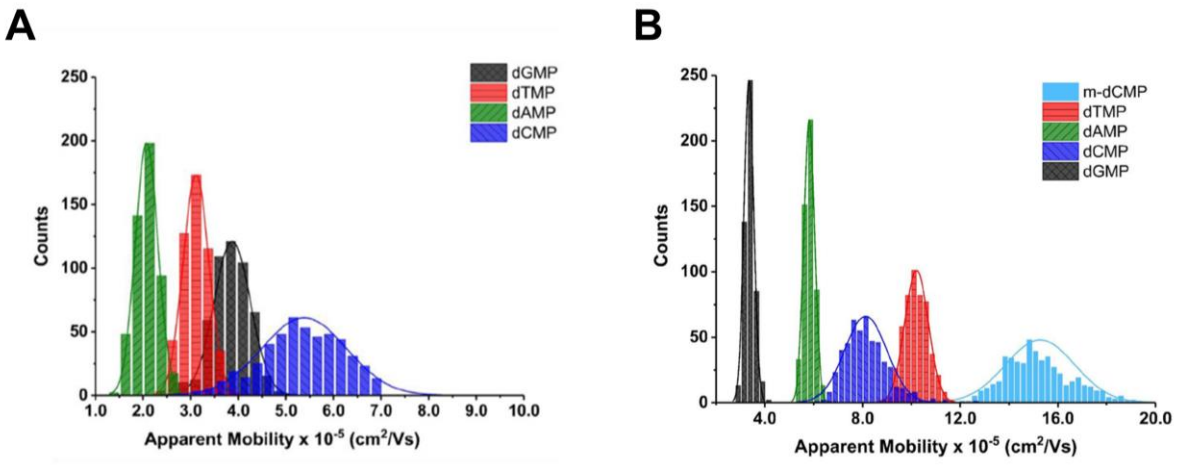


Figure 1.16. Identification of dye labeled dNMPs using nanoscale electrophoresis. Histograms of apparent mobilities of dNMPs **(A)** at pH 8.3. The resolution ranged from 0.73 – 2.13. **(B)** at pH 10.3. The resolution was ~4.84. (Reproduced from reference [181]).

O'Neil *et al.* reported identifying ATTO 532 dye-labeled dNMPs through their electrophoretic mobility in a PMMA (substrate)/COC (cover plate) thermoplastic nanochannel that was 110 nm in width and depth.[181] The sensitivity of the electrokinetic transport of dNMPs to the pH and the ionic strength of the carrier electrolyte was demonstrated. The resolution between the canonical dNMPs ranged from 0.73 – 2.13 at pH 8.3 but showed a resolution greater than 4 when the pH was increased to 10.3 (see Figure 1.16). Due to the increased thickness of the EDL at low ionic strengths, a decrease in identification accuracy and an increase in standard deviation of a histogram of single molecule electrophoretic mobilities was observed. The identification accuracy of the dNMPs based on their apparent mobilities was >99%.[181]

Recently, Amarasekara *et al.* demonstrated the identification of rNMPs in thermoplastic nanofluidic devices fabricated in PMMA/COC and COC/COC nanofluidic devices.[182] The canonical rNMPs could be separated at pH = 7.9 with a resolution >4 in both PMMA/COC (1.94 – 8.88) and COC/COC (1.4 – 7.8) devices. Some rNMPs showed field-dependent electrophoretic mobilities. Identification accuracies >99.9% were achieved using both PMMA/COC and COC/COC devices.[3]

Nanoscale electrophoresis combined with current transient amplitudes generated while traveling through in-plane nanopores could address the base call accuracy limitations associated with current nanopore sequencing technologies, which use only current transient amplitudes of the nucleotides to distinguish the bases. The identification of mononucleotides using their molecular-dependent electrophoretic mobility is coined as the time-of-flight (ToF) identification method.[180-182] ToF is particularly attractive for the separation of a variety of mononucleotides due to the ability to alter electrophoresis conditions such as substrate material type, electric field strength, pH, and ionic strength of the electrolyte to optimize the identification accuracy. Additionally, because thermoplastic substrates are used to fabricate the nanofluidic devices, it is possible to scale up the fabrication using nano-injection compression molding, facilitating the use of the production of these nanofluidic devices in clinical settings.[3, 10, 130]

1.5 References

1. Jennings, J. L.; Hudson, T. J., International Cancer Genome Consortium (ICGC). AACR: 2016.

2. Chang, K.; Creighton, C.; Davis, C.; Donehower, L., The cancer genome atlas pan-cancer analysis project. *Nat Genet* **2013,** *45* (10), 1113-1120.

3. Hoadley, K. A.; Yau, C.; Hinoue, T.; Wolf, D. M.; Lazar, A. J.; Drill, E.; Shen, R.; Taylor, A. M.; Cherniack, A. D.; Thorsson, V., Cell-of-origin patterns dominate the molecular classification of 10,000 tumors from 33 types of cancer. *Cell* **2018,** *173* (2), 291-304. e6.

4. Nik-Zainal, S.; Davies, H.; Staaf, J.; Ramakrishna, M.; Glodzik, D.; Zou, X.; Martincorena, I.; Alexandrov, L. B.; Martin, S.; Wedge, D. C., Landscape of somatic mutations in 560 breast cancer whole-genome sequences. *Nature* **2016,** *534* (7605), 47-54.

5. Ramsay, A. J.; Martinez-Trillos, A.; Jares, P.; Rodríguez, D.; Kwarciak, A.; Quesada, V., Next-generation sequencing reveals the secrets of the chronic lymphocytic leukemia genome. *Clinical and Translational Oncology* **2013,** *15* (1), 3-8.

6. Tirode, F.; Surdez, D.; Ma, X.; Parker, M.; Le Deley, M. C.; Bahrami, A.; Zhang, Z.; Lapouble, E.; Grossetête-Lalami, S.; Rusch, M., Genomic landscape of Ewing sarcoma defines an aggressive subtype with co-association of STAG2 and TP53 mutations. *Cancer discovery* **2014,** *4* (11), 1342-1353.

7. Milos, P. M., Emergence of single-molecule sequencing and potential for molecular diagnostic applications. *Expert review of molecular diagnostics* **2009,** *9* (7), 659-666.

8. Milos, P. M., Helicos single molecule sequencing: Unique capabilities and importance for molecular diagnostics. *Genome Biology* **2010,** *11* (1), 1-1.

9. Fuller, C. W.; Kumar, S.; Porel, M.; Chien, M.; Bibillo, A.; Stranges, P. B.; Dorwart, M.; Tao, C.; Li, Z.; Guo, W., Real-time single-molecule electronic DNA sequencing by synthesis using polymer-tagged nucleotides on a nanopore array. *Proceedings of the National Academy of Sciences* **2016,** *113* (19), 5233-5238.

10. Kumar, S.; Tao, C.; Chien, M.; Hellner, B.; Balijepalli, A.; Robertson, J. W.; Li, Z.; Russo, J. J.; Reiner, J. E.; Kasianowicz, J. J., PEG-Labeled Nucleotides and Nanopore Detection for Single Molecule DNASequencing by Synthesis. *Scientific reports* **2012,** *2* (1), 1-8.

11. Lindsay, S., The promises and challenges of solid-state sequencing. *Nature nanotechnology* **2016,** *11* (2), 109-111.

12. Greenleaf, W. J.; Block, S. M., Single-molecule, motion-based DNA sequencing using RNA polymerase. *Science* **2006,** *313* (5788), 801-801.

13. Jain, M.; Olsen, H. E.; Paten, B.; Akeson, M., The Oxford Nanopore MinION: delivery of nanopore sequencing to the genomics community. *Genome biology* **2016,** *17* (1), 1-11.

14. Movileanu, L., Interrogating single proteins through nanopores: challenges and opportunities. *Trends in Biotechnology* **2009,** *27* (6), 333-341.

15. Branton, D.; Deamer, D. W.; Marziali, A.; Bayley, H.; Benner, S. A.; Butler, T.; Di Ventra, M.; Garaj, S.; Hibbs, A.; Huang, X., The potential and challenges of nanopore sequencing. *Nanoscience and technology: A collection of reviews from Nature Journals* **2010**, 261-268.

16. Cherf, G. M.; Lieberman, K. R.; Rashid, H.; Lam, C. E.; Karplus, K.; Akeson, M., Automated forward and reverse ratcheting of DNA in a nanopore at 5-Å precision. *Nature biotechnology* **2012,** *30* (4), 344-348.

17. Larkin, J.; Foquet, M.; Korlach, J.; Wanunu, M., 207 Nanopore immobilization of DNA polymerase enhances single-molecule sequencing. *Journal of Biomolecular Structure and Dynamics* **2013,** *31* (sup1), 134-135.

18. Maitra, R. D.; Kim, J.; Dunbar, W. B., Recent advances in nanopore sequencing. *Electrophoresis* **2012,** *33* (23), 3418-3428.

19. Manrao, E. A.; Derrington, I. M.; Laszlo, A. H.; Langford, K. W.; Hopper, M. K.; Gillgren, N.; Pavlenok, M.; Niederweis, M.; Gundlach, J. H., Reading DNA at single-nucleotide resolution with a mutant MspA nanopore and phi29 DNA polymerase. *Nature biotechnology* **2012,** *30* (4), 349-353.

20. Sanger, F.; Nicklen, S.; Coulson, A. R., DNA sequencing with chain-terminating inhibitors. *Proceedings of the national academy of sciences* **1977,** *74* (12), 5463-5467.

21. Stranneheim, H.; Lundeberg, J., Stepping stones in DNA sequencing. *Biotechnology journal* **2012,** *7* (9), 1063-1073.

22. Heather, J. M.; Chain, B., The sequence of sequencers: The history of sequencing DNA. *Genomics* **2016,** *107* (1), 1-8.

23. Valencia, C. A.; Pervaiz, M. A.; Husami, A.; Qian, Y.; Zhang, K., Sanger Sequencing Principles, History, and Landmarks. In *Next Generation Sequencing Technologies in Medical Genetics*, Springer: 2013; pp 3-11.

24. Chidgeavadze, Z.; Beabealashvilli, R. S.; Atrazhev, A.; Kukhanova, M.; Azhayev, A.; Krayevsky, A., 2', 3'-Dideoxy-3'aminonucleoside 5'-triphosphates are the terminators of DNA synthesis catalyzed by DNA polymerases. *Nucleic acids research* **1984,** *12* (3), 1671.

25. Smith, L. M.; Sanders, J. Z.; Kaiser, R. J.; Hughes, P.; Dodd, C.; Connell, C. R.; Heiner, C.; Kent, S. B.; Hood, L. E., Fluorescence detection in automated DNA sequence analysis. *Nature* **1986,** *321* (6071), 674-679.

26. Prober, J. M.; Trainor, G. L.; Dam, R. J.; Hobbs, F. W.; Robertson, C. W.; Zagursky, R. J.; Cocuzza, A. J.; Jensen, M. A.; Baumeister, K., A system for rapid DNA sequencing with fluorescent chain-terminating dideoxynucleotides. *Science* **1987,** *238* (4825), 336-341.

27. Cohen, A.; Najarian, D.; Paulus, A.; Guttman, A.; Smith, J. A.; Karger, B., Rapid separation and purification of oligonucleotides by high-performance capillary gel electrophoresis. *Proceedings of the National Academy of Sciences* **1988,** *85* (24), 9660-9663.

28. Luckey, J. A.; Drossman, H.; Kostichka, A. J.; Mead, D. A.; D'Cunha, J.; Norris, T. B.; Smith, L. M., High speed DNA sequencing by capillary electrophoresis. *Nucleic acids research* **1990,** *18* (15), 4417-4421.

29. Staden, R., A strategy of DNA sequencing employing computer programs. *Nucleic acids research* **1979,** *6* (7), 2601-2610.

30. Anderson, S., Shotgun DNA sequencing using cloned DNase I-generated fragments. *Nucleic acids research* **1981,** *9* (13), 3015-3027.

31. Saiki, R. K.; Scharf, S.; Faloona, F.; Mullis, K. B.; Horn, G. T.; Erlich, H. A.; Arnheim, N., Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. *Science* **1985,** *230* (4732), 1350-1354.

32. Saiki, R. K.; Gelfand, D. H.; Stoffel, S.; Scharf, S. J.; Higuchi, R.; Horn, G. T.; Mullis, K. B.; Erlich, H. A., Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science* **1988,** *239* (4839), 487-491.

33. Jackson, D. A.; Symons, R. H.; Berg, P., Biochemical method for inserting new genetic information into DNA of Simian Virus 40: circular SV40 DNA molecules containing lambda phage genes and the galactose operon of Escherichia coli. *Proceedings of the National Academy of Sciences* **1972,** *69* (10), 2904-2909.

34. Karger, B. L.; Guttman, A., DNA sequencing by CE. *Electrophoresis* **2009,** *30* (S1), S196-S202.

35. Slatko, B. E.; Gardner, A. F.; Ausubel, F. M., Overview of next-generation sequencing technologies. *Current protocols in molecular biology* **2018,** *122* (1), e59.

36. Wetterstrand, K. A., DNA sequencing costs: data from the NHGRI Genome sequencing program (GSP). 2013. *URL http://www. genome. gov/sequencingcosts* **2016**.

37. Mardis, E. R., Next-generation sequencing platforms. *Annual review of analytical chemistry* **2013,** *6*, 287-303.

38. Goodwin, S.; McPherson, J. D.; McCombie, W. R., Coming of age: ten years of next-generation sequencing technologies. *Nature Reviews Genetics* **2016,** *17* (6), 333-351.

39. Bentley, D. R.; Balasubramanian, S.; Swerdlow, H. P.; Smith, G. P.; Milton, J.; Brown, C. G.; Hall, K. P.; Evers, D. J.; Barnes, C. L.; Bignell, H. R., Accurate whole human genome sequencing using reversible terminator chemistry. *nature* **2008,** *456* (7218), 53-59.

40. Guo, J.; Xu, N.; Li, Z.; Zhang, S.; Wu, J.; Kim, D. H.; Marma, M. S.; Meng, Q.; Cao, H.; Li, X., Four-color DNA sequencing with 3′-O-modified nucleotide reversible terminators and chemically cleavable fluorescent dideoxynucleotides. *Proceedings of the National Academy of Sciences* **2008,** *105* (27), 9145-9150.

41. Rothberg, J. M.; Hinz, W.; Rearick, T. M.; Schultz, J.; Mileski, W.; Davey, M.; Leamon, J. H.; Johnson, K.; Milgrew, M. J.; Edwards, M., An integrated semiconductor device enabling non-optical genome sequencing. *Nature* **2011,** *475* (7356), 348-352.

42. Hrdlickova, R.; Toloue, M.; Tian, B., RNA-Seq methods for transcriptome analysis. *Wiley Interdisciplinary Reviews: RNA* **2017,** *8* (1), e1364.

43. Nagalakshmi, U.; Waern, K.; Snyder, M., RNA-Seq: a method for comprehensive transcriptome analysis. *Current protocols in molecular biology* **2010,** *89* (1), 4.11. 1-4.11. 13.

44. Mortazavi, A.; Williams, B. A.; McCue, K.; Schaeffer, L.; Wold, B., Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nature methods* **2008,** *5* (7), 621-628.

45. Braslavsky, I.; Hebert, B.; Kartalov, E.; Quake, S. R., Sequence information can be obtained from single DNA molecules. *Proceedings of the National Academy of Sciences* **2003,** *100* (7), 3960-3964.

46. Jenkins, S.; Kahvejian, A., True single molecule sequencing (tSMS) TM by synthesis. *The handbook of plant mutation screening: mining of natural and induced alleles* **2010**, 289-306.

47. Kahvejian, A.; Kellett, S., Making single-molecule sequencing a reality. *American Laboratory* **2008,** *40* (20), 48-53.

48. Harris, T. D.; Buzby, P. R.; Babcock, H.; Beer, E.; Bowers, J.; Braslavsky, I.; Causey, M.; Colonell, J.; DiMeo, J.; Efcavitch, J. W., Single-molecule DNA sequencing of a viral genome. *Science* **2008,** *320* (5872), 106-109.

49. Pushkarev, D.; Neff, N. F.; Quake, S. R., Single-molecule sequencing of an individual human genome. *Nature biotechnology* **2009,** *27* (9), 847-850.

50. Lipson, D.; Raz, T.; Kieu, A.; Jones, D. R.; Giladi, E.; Thayer, E.; Thompson, J. F.; Letovsky, S.; Milos, P.; Causey, M., Quantification of the yeast transcriptome by single-molecule sequencing. *Nature biotechnology* **2009,** *27* (7), 652-658.

51. Gupta, P. K., Single-molecule DNA sequencing technologies for future genomics research. *Trends in biotechnology* **2008,** *26* (11), 602-611.

52. Ozsolak, F.; Platt, A. R.; Jones, D. R.; Reifenberger, J. G.; Sass, L. E.; McInerney, P.; Thompson, J. F.; Bowers, J.; Jarosz, M.; Milos, P. M., Direct RNA sequencing. *Nature* **2009,** *461* (7265), 814-818.

53. Ozsolak, F.; Milos, P. M., Single-molecule direct RNA sequencing without cDNA synthesis. *Wiley Interdisciplinary Reviews: RNA* **2011,** *2* (4), 565-570.

54. Bowers, J.; Mitchell, J.; Beer, E.; Buzby, P. R.; Causey, M.; Efcavitch, J. W.; Jarosz, M.; Krzymanska-Olejnik, E.; Kung, L.; Lipson, D., Virtual terminator nucleotides for next-generation DNA sequencing. *Nature methods* **2009,** *6* (8), 593-595.

55. Ozsolak, F.; Goren, A.; Gymrek, M.; Guttman, M.; Regev, A.; Bernstein, B. E.; Milos, P. M., Digital transcriptome profiling from attomole-level RNA samples. *Genome research* **2010,** *20* (4), 519-525.

56. Ozsolak, F.; Ting, D. T.; Wittner, B. S.; Brannigan, B. W.; Paul, S.; Bardeesy, N.; Ramaswamy, S.; Milos, P. M.; Haber, D. A., Amplification-free digital gene expression profiling from minute cell quantities. *nature methods* **2010,** *7* (8), 619-621.

57. Kapranov, P.; Ozsolak, F.; Kim, S. W.; Foissac, S.; Lipson, D.; Hart, C.; Roels, S.; Borel, C.; Antonarakis, S. E.; Monaghan, A. P., New class of gene-termini-associated human RNAs suggests a novel RNA copying mechanism. *Nature* **2010,** *466* (7306), 642-646.

58. Flusberg, B. A.; Webster, D. R.; Lee, J. H.; Travers, K. J.; Olivares, E. C.; Clark, T. A.; Korlach, J.; Turner, S. W., Direct detection of DNA methylation during single-molecule, real-time sequencing. *Nature methods* **2010,** *7* (6), 461-465.

59. Eid, J.; Fehr, A.; Gray, J.; Luong, K.; Lyle, J.; Otto, G.; Peluso, P.; Rank, D.; Baybayan, P.; Bettman, B., Real-time DNA sequencing from single polymerase molecules. *Science* **2009,** *323* (5910), 133-138.

60. Foquet, M.; Samiee, K. T.; Kong, X.; Chauduri, B. P.; Lundquist, P. M.; Turner, S. W.; Freudenthal, J.; Roitman, D. B., Improved fabrication of zero-mode waveguides for single-molecule detection. *Journal of Applied Physics* **2008,** *103* (3), 034301.

61. Levene, M. J.; Korlach, J.; Turner, S. W.; Foquet, M.; Craighead, H. G.; Webb, W. W., Zero-mode waveguides for single-molecule analysis at high concentrations. *science* **2003,** *299* (5607), 682-686.

62. Korlach, J.; Marks, P. J.; Cicero, R. L.; Gray, J. J.; Murphy, D. L.; Roitman, D. B.; Pham, T. T.; Otto, G. A.; Foquet, M.; Turner, S. W., Selective aluminum passivation for targeted immobilization of single DNA polymerase molecules in zero-mode waveguide nanostructures. *Proceedings of the National Academy of Sciences* **2008,** *105* (4), 1176-1181.

63. Vilfan, I. D.; Tsai, Y.-C.; Clark, T. A.; Wegener, J.; Dai, Q.; Yi, C.; Pan, T.; Turner, S. W.; Korlach, J., Analysis of RNA base modification and structural rearrangement by single-molecule real-time detection of reverse transcription. *Journal of nanobiotechnology* **2013,** *11* (1), 1-11.

64. Sood, A.; Kumar, S.; Nampalli, S.; Nelson, J. R.; Macklin, J.; Fuller, C. W., Terminal phosphate-labeled nucleotides with improved substrate properties for homogeneous nucleic acid assays. *Journal of the American Chemical Society* **2005,** *127* (8), 2394-2395.

65. Lundquist, P. M.; Zhong, C. F.; Zhao, P.; Tomaney, A. B.; Peluso, P. S.; Dixon, J.; Bettman, B.; Lacroix, Y.; Kwo, D. P.; McCullough, E., Parallel confocal detection of single molecules in real time. *Optics letters* **2008,** *33* (9), 1026-1028.

66. Xu, M.; Fujita, D.; Hanagata, N., Perspectives and challenges of emerging single-molecule DNA sequencing technologies. *Small* **2009,** *5* (23), 2638-2649.

67. Treffer, R.; Deckert, V., Recent advances in single-molecule sequencing. *Current opinion in biotechnology* **2010,** *21* (1), 4-11.

68. Nyrén, P.; Lundin, A., Enzymatic method for continuous monitoring of inorganic pyrophosphate synthesis. *Analytical biochemistry* **1985,** *151* (2), 504-509.

69. Hyman, E. D., A new method of sequencing DNA. *Analytical biochemistry* **1988,** *174* (2), 423-436.

70. Ronaghi, M.; Karamohamed, S.; Pettersson, B.; Uhlén, M.; Nyrén, P., Real-time DNA sequencing using detection of pyrophosphate release. *Analytical biochemistry* **1996,** *242* (1), 84-89.

71. Ronaghi, M., Pyrosequencing sheds light on DNA sequencing. *Genome research* **2001,** *11* (1), 3-11.

72. Ahmadian, A.; Ehn, M.; Hober, S., Pyrosequencing: history, biochemistry and future. *Clinica chimica acta* **2006,** *363* (1-2), 83-94.

73. Klenow, H.; Overgaard-Hansen, K.; Patkar, S. A., Proteolytic cleavage of native DNA polymerase into two different catalytic fragments: influence of assay conditions on the change of exonuclease activity and polymerase activity accompanying cleavage. *European journal of biochemistry* **1971,** *22* (3), 371-381.

74. Segel, I. H.; Renosto, F.; Seubert, P. A., Sulfate-activating enzymes. *Methods in enzymology* **1987,** *143*, 334-349.

75. Hosseinkhani, S., Molecular enigma of multicolor bioluminescence of firefly luciferase. *Cellular and Molecular Life Sciences* **2011,** *68* (7), 1167-1182.

76. Komoszynski, M.; Wojtczak, A., Apyrases (ATP diphosphohydrolases, EC 3.6. 1.5): function and relationship to ATPases. *Biochimica et Biophysica Acta (BBA)-Molecular Cell Research* **1996,** *1310* (2), 233-241.

77. Margulies, M.; Egholm, M.; Altman, W. E.; Attiya, S.; Bader, J. S.; Bemben, L. A.; Berka, J.; Braverman, M. S.; Chen, Y.-J.; Chen, Z., Genome sequencing in microfabricated high-density picolitre reactors. *Nature* **2005,** *437* (7057), 376-380.

78. Kartalov, E. P.; Quake, S. R., Microfluidic device reads up to four consecutive base pairs in DNA sequencing-by-synthesis. *Nucleic acids research* **2004,** *32* (9), 2873-2879.

79. Bayley, H., Nanopore sequencing: from imagination to reality. *Clinical chemistry* **2015,** *61* (1), 25-31.

80. Bayley, H., Sequencing single molecules of DNA. *Current opinion in chemical biology* **2006,** *10* (6), 628-637.

81. Goto, Y.; Akahori, R.; Yanagi, I.; Takeda, K.-i., Solid-state nanopores towards single-molecule DNA sequencing. *Journal of human genetics* **2020,** *65* (1), 69-77.

82. Taniguchi, M., Selective multidetection using nanopores. *Analytical chemistry* **2015,** *87* (1), 188-199.

83. Soni, G. V.; Meller, A., Progress toward ultrafast DNA sequencing using solid-state nanopores. *Clinical chemistry* **2007,** *53* (11), 1996-2001.

84. Deamer, D.; Akeson, M.; Branton, D., Three decades of nanopore sequencing. *Nature biotechnology* **2016,** *34* (5), 518-524.

85. Derrington, I. M.; Butler, T. Z.; Collins, M. D.; Manrao, E.; Pavlenok, M.; Niederweis, M.; Gundlach, J. H., Nanopore DNA sequencing with MspA. *Proceedings of the National Academy of Sciences* **2010,** *107* (37), 16060-16065.

86. Niederweis, M.; Ehrt, S.; Heinz, C.; KloÈcker, U.; Karosi, S.; Swiderek, K. M.; Riley, L. W.; Benz, R., Cloning of the mspA gene encoding a porin from Mycobacterium smegmatis. *Molecular microbiology* **1999,** *33* (5), 933-945.

87. Trias, J.; Benz, R., Permeability of the cell wall of Mycobacterium smegmatis. *Molecular microbiology* **1994,** *14* (2), 283-290.

88. Butler, T. Z.; Pavlenok, M.; Derrington, I. M.; Niederweis, M.; Gundlach, J. H., Single-molecule DNA detection with an engineered MspA protein nanopore. *Proceedings of the National Academy of Sciences* **2008,** *105* (52), 20647-20652.

89. Meller, A.; Nivon, L.; Branton, D., Voltage-driven DNA translocations through a nanopore. *Physical Review Letters* **2001,** *86* (15), 3435.

90. Bhattacharya, S.; Derrington, I. M.; Pavlenok, M.; Niederweis, M.; Gundlach, J. H.; Aksimentiev, A., Molecular dynamics study of MspA arginine mutants predicts slow DNA translocations and ion current blockades indicative of DNA sequence. *ACS nano* **2012,** *6* (8), 6960-6968.

91. Song, L.; Hobaugh, M. R.; Shustak, C.; Cheley, S.; Bayley, H.; Gouaux, J. E., Structure of staphylococcal α-hemolysin, a heptameric transmembrane pore. *Science* **1996,** *274* (5294), 1859-1865.

92. Astier, Y.; Braha, O.; Bayley, H., Toward single molecule DNA sequencing: direct identification of ribonucleoside and deoxyribonucleoside 5 '-monophosphates by using an engineered protein nanopore equipped with a molecular adapter. *Journal of the American Chemical Society* **2006,** *128* (5), 1705-1710.

93. Ayub, M.; Hardwick, S. W.; Luisi, B. F.; Bayley, H., Nanopore-based identification of individual nucleotides for direct RNA sequencing. *Nano letters* **2013,** *13* (12), 6144-6150.

94. Cheley, S.; Gu, L.-Q.; Bayley, H., Stochastic sensing of nanomolar inositol 1, 4, 5-trisphosphate with an engineered pore. *Chemistry & biology* **2002,** *9* (7), 829-838.

95. Wu, H.-C.; Astier, Y.; Maglia, G.; Mikhailova, E.; Bayley, H., Protein nanopores with covalently attached molecular adapters. *Journal of the American Chemical Society* **2007,** *129* (51), 16142-16148.

96. Hardwick, S. W.; Gubbey, T.; Hug, I.; Jenal, U.; Luisi, B. F., Crystal structure of Caulobacter crescentus polynucleotide phosphorylase reveals a mechanism of RNA substrate channelling and RNA degradosome assembly. *Open biology* **2012,** *2* (4), 120028.

97. Hardwick, S. W.; Chan, V. S.; Broadhurst, R. W.; Luisi, B. F., An RNA degradosome assembly in Caulobacter crescentus. *Nucleic acids research* **2011,** *39* (4), 1449-1459.

98. Reiner, J. E.; Balijepalli, A.; Robertson, J. W.; Drown, B. S.; Burden, D. L.; Kasianowicz, J. J., The effects of diffusion on an exonuclease/nanopore-based DNA sequencing engine. *The Journal of chemical physics* **2012,** *137* (21), 214903.

99. Cockroft, S. L.; Chu, J.; Amorin, M.; Ghadiri, M. R., A single-molecule nanopore device detects DNA polymerase activity with single-nucleotide resolution. *Journal of the American Chemical Society* **2008,** *130* (3), 818-820.

100. Olasagasti, F.; Lieberman, K. R.; Benner, S.; Cherf, G. M.; Dahl, J. M.; Deamer, D. W.; Akeson, M., Replication of individual DNA molecules under electronic control using a protein nanopore. *Nature nanotechnology* **2010,** *5* (11), 798-806.

101. Byrd, A. K.; Raney, K. D., Helicases and DNA Motor Proteins. In *Nanopore Sequencing: An Introduction*, World Scientific: 2019; pp 59-74.

102. Astier, Y.; Kainov, D. E.; Bayley, H.; Tuma, R.; Howorka, S., Stochastic Detection of Motor Protein–RNA Complexes by Single-Channel Current Recording. *Chemphyschem* **2007,** *8* (15), 2189-2194.

103. Mancini, E. J.; Kainov, D. E.; Grimes, J. M.; Tuma, R.; Bamford, D. H.; Stuart, D. I., Atomic snapshots of an RNA packaging motor reveal conformational changes linking ATP hydrolysis to RNA translocation. *Cell* **2004,** *118* (6), 743-755.

104. Lísal, J.; Lam, T. T.; Kainov, D. E.; Emmett, M. R.; Marshall, A. G.; Tuma, R., Functional visualization of viral molecular motor by hydrogen-deuterium exchange reveals transient states. *Nature structural & molecular biology* **2005,** *12* (5), 460-466.

105. Laszlo, A. H.; Derrington, I. M.; Ross, B. C.; Brinkerhoff, H.; Adey, A.; Nova, I. C.; Craig, J. M.; Langford, K. W.; Samson, J. M.; Daza, R., Decoding long nanopore sequencing reads of natural DNA. *Nature biotechnology* **2014,** *32* (8), 829-833.

106. Craig, J. M.; Laszlo, A. H.; Derrington, I. M.; Ross, B. C.; Brinkerhoff, H.; Nova, I. C.; Doering, K.; Tickman, B. I.; Svet, M. T.; Gundlach, J. H., Direct detection of unnatural DNA nucleotides dNaM and d5SICS using the MspA nanopore. *PLoS One* **2015,** *10* (11), e0143253.

107. Laszlo, A. H.; Derrington, I. M.; Brinkerhoff, H.; Langford, K. W.; Nova, I. C.; Samson, J. M.; Bartlett, J. J.; Pavlenok, M.; Gundlach, J. H., Detection and mapping of 5-methylcytosine and 5-hydroxymethylcytosine with nanopore MspA. *Proceedings of the National Academy of Sciences* **2013,** *110* (47), 18904-18909.

108. Lieberman, K. R.; Cherf, G. M.; Doody, M. J.; Olasagasti, F.; Kolodji, Y.; Akeson, M., Processive replication of single DNA molecules in a nanopore catalyzed by phi29 DNA polymerase. *Journal of the American Chemical Society* **2010,** *132* (50), 17961-17972.

109. Jain, M.; Fiddes, I. T.; Miga, K. H.; Olsen, H. E.; Paten, B.; Akeson, M., Improved data analysis for the MinION nanopore sequencer. *Nature methods* **2015,** *12* (4), 351-356.

110. Li, J.; Stein, D.; McMullan, C.; Branton, D.; Aziz, M. J.; Golovchenko, J. A., Ion-beam sculpting at nanometre length scales. *Nature* **2001,** *412* (6843), 166-169.

111. Dekker, C., Solid-state nanopores. *Nanoscience And Technology: A Collection of Reviews from Nature Journals* **2010**, 60-66.

112. ávan Loo, N., DNA nanopore translocation in glutamate solutions. *Nanoscale* **2015,** *7* (32), 13605-13609.

113. Jain, M.; Koren, S.; Miga, K. H.; Quick, J.; Rand, A. C.; Sasani, T. A.; Tyson, J. R.; Beggs, A. D.; Dilthey, A. T.; Fiddes, I. T., Nanopore sequencing and assembly of a human genome with ultra-long reads. *Nature biotechnology* **2018,** *36* (4), 338-345.

114. Storm, A.; Chen, J.; Ling, X.; Zandbergen, H.; Dekker, C., Fabrication of solid-state nanopores with single-nanometre precision. *Nature materials* **2003,** *2* (8), 537-540.

115. Gierak, J.; Madouri, A.; Biance, A.; Bourhis, E.; Patriarche, G.; Ulysse, C.; Lucot, D.; Lafosse, X.; Auvray, L.; Bruchhaus, L., Sub-5 nm FIB direct patterning of nanodevices. *Microelectronic engineering* **2007,** *84* (5-8), 779-783.

116. Yang, J.; Ferranti, D. C.; Stern, L. A.; Sanford, C. A.; Huang, J.; Ren, Z.; Qin, L.-C.; Hall, A. R., Rapid and precise scanning helium ion microscope milling of solid-state nanopores for biomolecule detection. *Nanotechnology* **2011,** *22* (28), 285310.

117. Kwok, H.; Briggs, K.; Tabard-Cossa, V., Nanopore fabrication by controlled dielectric breakdown. *PloS one* **2014,** *9* (3), e92880.

118. Stein, D. M.; McMullan, C. J.; Li, J.; Golovchenko, J. A., Feedback-controlled ion beam sculpting apparatus. *Review of Scientific Instruments* **2004,** *75* (4), 900-905.

119. Asghar, W.; Ilyas, A.; Billo, J. A.; Iqbal, S. M., Shrinking of solid-state nanopores by direct thermal heating. *Nanoscale research letters* **2011,** *6* (1), 1-6.

120. Schiedt, B.; Auvray, L.; Bacri, L.; Biance, A.-L.; Madouri, A.; Bourhis, E.; Patriarche, G.; Pelta, J.; Jede, R.; Gierak, J., Direct FIB fabrication and integration of “single nanopore devices” for the manipulation of macromolecules. *MRS Online Proceedings Library (OPL)* **2009,** *1191*.

121. Larkin, J.; Henley, R.; Bell, D. C.; Cohen-Karni, T.; Rosenstein, J. K.; Wanunu, M., Slow DNA transport through nanopores in hafnium oxide membranes. *ACS nano* **2013,** *7* (11), 10121-10128.

122. Venta, K.; Shemer, G.; Puster, M.; Rodriguez-Manzo, J. A.; Balan, A.; Rosenstein, J. K.; Shepard, K.; Drndic, M., Differentiation of short, single-stranded DNA homopolymers in solid-state nanopores. *ACS nano* **2013,** *7* (5), 4629-4636.

123. Carlsen, A. T.; Zahid, O. K.; Ruzicka, J.; Taylor, E. W.; Hall, A. R., Interpreting the conductance blockades of DNA translocations through solid-state nanopores. *Acs Nano* **2014,** *8* (5), 4754-4760.

124. Lee, M.-H.; Kumar, A.; Park, K.-B.; Cho, S.-Y.; Kim, H.-M.; Lim, M.-C.; Kim, Y.-R.; Kim, K.-B., A low-noise solid-state nanopore platform based on a highly insulating substrate. *Scientific reports* **2014,** *4* (1), 1-7.

125. Yanagi, I.; Ishida, T.; Fujisaki, K.; Takeda, K.-i., Fabrication of 3-nm-thick Si 3 N 4 membranes for solid-state nanopores using the poly-Si sacrificial layer process. *Scientific reports* **2015,** *5* (1), 1-13.

126. Feng, J.; Liu, K.; Bulushev, R. D.; Khlybov, S.; Dumcenco, D.; Kis, A.; Radenovic, A., Identification of single nucleotides in MoS 2 nanopores. *Nature nanotechnology* **2015,** *10* (12), 1070-1076.

127. Carda–Broch, S.; Berthod, A.; Armstrong, D., Solvent properties of the 1-butyl-3-methylimidazolium hexafluorophosphate ionic liquid. *Analytical and bioanalytical chemistry* **2003,** *375* (2), 191-199.

128. Lee, J. W. In *Nanoelectrode-gated detection of individual molecules with potential for rapid DNA sequencing*, Solid State Phenomena, Trans Tech Publ: 2007; pp 1379-1386.

129. Tsutsui, M.; Taniguchi, M.; Kawai, T., Fabrication of 0.5 nm electrode gaps using self-breaking technique. *Applied Physics Letters* **2008,** *93* (16), 163115.

130. Tsutsui, M.; Taniguchi, M.; Yokota, K.; Kawai, T., Identifying single nucleotides by tunnelling current. *Nature nanotechnology* **2010,** *5* (4), 286-290.

131. Ohshiro, T.; Matsubara, K.; Tsutsui, M.; Furuhashi, M.; Taniguchi, M.; Kawai, T., Single-molecule electrical random resequencing of DNA and RNA. *Scientific reports* **2012,** *2* (1), 1-7.

132. Rigler, R.; Mets, Ü.; Widengren, J.; Kask, P., Fluorescence correlation spectroscopy with high count rate and low background: analysis of translational diffusion. *European Biophysics Journal* **1993,** *22* (3), 169-175.

133. Dörre, K.; Brakmann, S.; Brinkmeier, M.; Han, K. T.; Riebeseel, K.; Schwille, P.; Stephan, J.; Wetzel, T.; Lapczyna, M.; Stuke, M., Techniques for single molecule sequencing. *Bioimaging* **1997,** *5* (3), 139-152.

134. Nguyen, D. C.; Keller, R. A.; Jett, J. H.; Martin, J. C., Detection of single molecules of phycoerythrin in hydrodynamically focused flows by laser-induced fluorescence. *Analytical chemistry* **1987,** *59* (17), 2158-2161.

135. Ambrose, W. P.; Goodwin, P. M.; Jett, J. H.; Johnson, M. E.; Martin, J. C.; Marrone, B. L.; Schecker, J. A.; Wilkerson, C. W.; Keller, R. A.; Haces, A., Application of single molecule detection to DNA sequencing and sizing. *Berichte der Bunsengesellschaft für physikalische Chemie* **1993,** *97* (12), 1535-1541.

136. Abbondanzieri, E. A.; Greenleaf, W. J.; Shaevitz, J. W.; Landick, R.; Block, S. M., Direct observation of base-pair stepping by RNA polymerase. *Nature* **2005,** *438* (7067), 460-465.

137. Herbert, K. M.; La Porta, A.; Wong, B. J.; Mooney, R. A.; Neuman, K. C.; Landick, R.; Block, S. M., Sequence-resolved detection of pausing by single RNA polymerase molecules. *Cell* **2006,** *125* (6), 1083-1094.

138. Xu, M.; Endres, R. G.; Arakawa, Y., The electronic properties of DNA bases. *Small* **2007,** *3* (9), 1539-1543.

139. Shapir, E.; Cohen, H.; Calzolari, A.; Cavazzoni, C.; Ryndyk, D. A.; Cuniberti, G.; Kotlyar, A.; Di Felice, R.; Porath, D., Electronic structure of single DNA molecules resolved by transverse scanning tunnelling spectroscopy. *Nature Materials* **2008,** *7* (1), 68-74.

140. Tanaka, H.; Kawai, T., Partial sequencing of a single DNA molecule with a scanning tunnelling microscope. *Nature nanotechnology* **2009,** *4* (8), 518-522.

141. Meng, S.; Maragakis, P.; Papaloukas, C.; Kaxiras, E., DNA nucleoside interaction and identification with carbon nanotubes. *nano letters* **2007,** *7* (1), 45-50.

142. Löffelholz, C.; Husemann, U.; Greller, G.; Meusel, W.; Kauling, J.; Ay, P.; Kraume, M.; Eibl, R.; Eibl, D., Bioengineering parameters for single-use bioreactors: Overview and evaluation of suitable methods. *Chemie Ingenieur Technik* **2013,** *85* (1-2), 40-56.

143. Liu, W.; Wang, L.; Jiang, R., Specific enzyme immobilization approaches and their application with nanomaterials. *Topics in Catalysis* **2012,** *55* (16-18), 1146-1156.

144. Ansari, S. A.; Husain, Q., Potential applications of enzymes immobilized on/in nano materials: A review. *Biotechnology advances* **2012,** *30* (3), 512-523.

145. Urban, P. L.; Goodall, D. M.; Bruce, N. C., Enzymatic microreactors in chemical analysis and kinetic studies. *Biotechnology Advances* **2006,** *24* (1), 42-57.

146. Wong, L. S.; Khan, F.; Micklefield, J., Selective covalent protein immobilization: strategies and applications. *Chemical reviews* **2009,** *109* (9), 4025-4053.

147. Hanefeld, U.; Cao, L.; Magner, E., Enzyme immobilisation: fundamentals and application. *Chemical Society Reviews* **2013,** *42* (15), 6211-6212.

148. Ducker, R. E.; Montague, M. T.; Leggett, G. J., A comparative investigation of methods for protein immobilization on self-assembled monolayers using glutaraldehyde, carbodiimide, and anhydride reagents. *Biointerphases* **2008,** *3* (3), 59-65.

149. Křenková, J.; Foret, F., Immobilized microfluidic enzymatic reactors. *Electrophoresis* **2004,** *25* (21-22), 3550-3563.

150. Laurell, T.; Drott, J.; Rosengren, L.; Lindström, K., Enhanced enzyme activity in silicon integrated enzyme reactors utilizing porous silicon as the coupling matrix. *Sensors and Actuators B: Chemical* **1996,** *31* (3), 161-166.

151. Letant, S. E.; Hart, B. R.; Kane, S. R.; Hadi, M. Z.; Shields, S. J.; Reynolds, J. G., Enzyme immobilization on porous silicon surfaces. *Advanced materials* **2004,** *16* (8), 689-693.

152. Manjon, A.; Obon, J.; Casanova, P.; Fernández, V.; Ilborra, J., Increased activity of glucose dehydrogenase co-immobilized with a redox mediator in a bioreactor with electrochemical NAD+ regeneration. *Biotechnology letters* **2002,** *24* (15), 1227-1232.

153. Matosevic, S.; Szita, N.; Baganz, F., Fundamentals and applications of immobilized microfluidic enzymatic reactors. *Journal of Chemical Technology & Biotechnology* **2011,** *86* (3), 325-334.

154. Davidson, Y. Y.; Soper, S. A.; Margolis, S.; Sander, L. C., Immobilization of the restriction enzymes HaeIII and HindIII on porous silica particles via a glutaraldehyde linkage for the micro-digestion of dsDNA with analysis by capillary electrophoresis. *Journal of separation science* **2001,** *24* (1), 10-16.

155. Seong, G. H.; Zhan, W.; Crooks, R. M., Fabrication of microchambers defined by photopolymerized hydrogels and weirs within microfluidic systems: application to DNA hybridization. *Analytical chemistry* **2002,** *74* (14), 3372-3377.

156. Hanefeld, U.; Gardossi, L.; Magner, E., Understanding enzyme immobilisation. *Chemical Society Reviews* **2009,** *38* (2), 453-468.

157. Rodrigues, R. C.; Ortiz, C.; Berenguer-Murcia, Á.; Torres, R.; Fernández-Lafuente, R., Modifying enzyme activity and selectivity by immobilization. *Chemical Society Reviews* **2013,** *42* (15), 6290-6307.

158. Polizzi, K. M.; Bommarius, A. S.; Broering, J. M.; Chaparro-Riggers, J. F., Stability of biocatalysts. *Current opinion in chemical biology* **2007,** *11* (2), 220-225.

159. Mateo, C.; Palomo, J. M.; Fuentes, M.; Betancor, L.; Grazu, V.; López-Gallego, F.; Pessela, B. C.; Hidalgo, A.; Fernández-Lorente, G.; Fernández-Lafuente, R., Glyoxyl agarose: a fully inert and hydrophilic support for immobilization and high stabilization of proteins. *Enzyme and Microbial Technology* **2006,** *39* (2), 274-280.

160. Iqbal, J.; Iqbal, S.; Müller, C. E., Advances in immobilized enzyme microbioreactors in capillary electrophoresis. *Analyst* **2013,** *138* (11), 3104-3116.

161. Brockman, H. L.; Law, J. H.; Kézdy, F. J., Catalysis by adsorbed enzymes: the hydrolysis of tripropionin by pancreatic lipase adsorbed to siliconized glass beads. *Journal of Biological Chemistry* **1973,** *248* (14), 4965-4970.

162. Joshi, K. A.; Prouza, M.; Kum, M.; Wang, J.; Tang, J.; Haddon, R.; Chen, W.; Mulchandani, A., V-type nerve agent detection using a carbon nanotube-based amperometric enzyme electrode. *Analytical Chemistry* **2006,** *78* (1), 331-336.

163. Joshi, K. A.; Tang, J.; Haddon, R.; Wang, J.; Chen, W.; Mulchandani, A., A disposable biosensor for organophosphorus nerve agents based on carbon nanotubes modified thick film strip electrode. *Electroanalysis: An International Journal Devoted to Fundamental and Practical Aspects of Electroanalysis* **2005,** *17* (1), 54-58.

164. Kumakura, M.; Kaetsu, I., Immobilization of cellulase using porous polymer matrix. *Journal of applied polymer science* **1984,** *29* (9), 2713-2718.

165. Levin, Y.; Pecht, M.; Goldstein, L.; Katchalski, E., A water-insoluble polyanionic derivative of trypsin. I. Preparation and properties. *Biochemistry* **1964,** *3* (12), 1905-1913.

166. Ogle, J. D.; Glassmeyer, C. K., Properties of an insoluble form of trypsin. *Biochemistry* **1971,** *10* (5), 786-792.

167. Pinto, P. C.; Costa, S. P.; Costa, A. D.; Passos, M. L.; Lima, J. L.; Saraiva, M. L. M., Trypsin activity in imidazolium based ionic liquids: evaluation of free and immobilized enzyme. *Journal of Molecular Liquids* **2012,** *171*, 16-22.

168. Tran, D. N.; Balkus Jr, K. J., Perspective of recent progress in immobilization of enzymes. *Acs Catalysis* **2011,** *1* (8), 956-968.

169. Datta, S.; Christena, L. R.; Rajaram, Y. R. S., Enzyme immobilization: an overview on techniques and support materials. *3 Biotech* **2013,** *3* (1), 1-9.

170. Birnbaum, S.; Nilsson, S., Protein-based capillary affinity gel electrophoresis for the separation of optical isomers. *Analytical Chemistry* **1992,** *64* (22), 2872-2874.

171. Chiang, C.-J.; Hsiau, L.-T.; Lee, W.-C., Immobilization of cell-associated enzymes by entrapment in polymethacrylamide beads. *Biotechnology techniques* **1997,** *11* (2), 121-125.

172. Subramanian, A.; Kennel, S. J.; Oden, P. I.; Jacobson, K. B.; Woodward, J.; Doktycz, M. J., Comparison of techniques for enzyme immobilization on silicon supports. *Enzyme and Microbial Technology* **1999,** *24* (1-2), 26-34.

173. Klotzbach, T. L.; Watt, M.; Ansari, Y.; Minteer, S. D., Improving the microenvironment for enzyme immobilization at electrodes by hydrophobically modifying chitosan and Nafion® polymers. *Journal of Membrane Science* **2008,** *311* (1-2), 81-88.

174. Brown, H. D.; Patel, A. B.; Chattopadhyay, S. K., Enzyme entrapment within hydrophobic and hydrophilic matrices. *Journal of biomedical materials research* **1968,** *2* (2), 231-235.

175. Kimmel, J.; Arazawa, D.; Ye, S.-H.; Shankarraman, V.; Wagner, W.; Federspiel, W., Carbonic anhydrase immobilized on hollow fiber membranes using glutaraldehyde activated

chitosan for artificial lung applications. *Journal of Materials Science: Materials in Medicine* **2013,** *24* (11), 2611-2621.

176. Mangaldas, K. S.; Rajput, Y. S.; Sharma, R., Urease immobilization on arylamine glass beads and its characterization. *Journal of plant biochemistry and biotechnology* **2010,** *19* (1), 73-77.

177. Teodor, E.; Litescu, S.-C.; Lazar, V.; Somoghi, R., Hydrogel-magnetic nanoparticles with immobilized L-asparaginase for biomedical applications. *Journal of Materials Science: Materials in Medicine* **2009,** *20* (6), 1307-1314.

178. Oliver-Calixte, N. J.; Uba, F. I.; Battle, K. N.; Weerakoon-Ratnayake, K. M.; Soper, S. A., Immobilization of lambda exonuclease onto polymer micropillar arrays for the solid-phase digestion of dsDNAs. *Analytical chemistry* **2014,** *86* (9), 4447-4454.

179. Chantiwas, R.; Park, S.; Soper, S. A.; Kim, B. C.; Takayama, S.; Sunkara, V.; Hwang, H.; Cho, Y.-K., Flexible fabrication and applications of polymer nanochannels and nanoslits. *Chemical Society Reviews* **2011,** *40* (7), 3677-3702.

180. Athapattu, U. S.; Amarasekara, C. A.; Immel, J. R.; Bloom, S.; Barany, F.; Nagel, A. C.; Soper, S. A., Solid-phase XRN1 reactions for RNA cleavage: application in single-molecule sequencing. *Nucleic acids research* **2021,** *49* (7), e41-e41.

181. O'Neil, C.; Amarasekara, C. A.; Weerakoon-Ratnayake, K. M.; Gross, B.; Jia, Z.; Singh, V.; Park, S.; Soper, S. A., Electrokinetic transport properties of deoxynucleotide monophosphates (dNMPs) through thermoplastic nanochannels. *Analytica chimica acta* **2018,** *1027*, 67-75.

182. Amarasekara, C. A.; Rathnayaka, C.; Athapattu, U. S.; Zhang, L.; Choi, J.; Park, S.; Nagel, A. C.; Soper, S. A., Electrokinetic identification of ribonucleotide monophosphates (rNMPs) using thermoplastic nanochannels. *Journal of Chromatography A* **2021,** *1638*, 461892.

183. Pennathur, S.; Santiago, J. G., Electrokinetic transport in nanochannels. 1. Theory. *Analytical Chemistry* **2005,** *77* (21), 6772-6781.

184. Pennathur, S.; Santiago, J. G., Electrokinetic transport in nanochannels. 2. Experiments. *Analytical chemistry* **2005,** *77* (21), 6782-6789.

185. Baldessari, F.; Santiago, J. G., Electrophoresis in nanochannels: brief review and speculation. *Journal of Nanobiotechnology* **2006,** *4* (1), 1-6.

Chapter 2. Solid-phase XRN1 Reactions for RNA Disassembly: Application in Single-Molecule RNA Sequencing

This chapter is based on the following manuscript,

- **Athapattu, U. S.**; Amarasekara, C. A.; Immel, J. R.; Bloom, S.; Barany, F.; Nagel, A. C.; Soper, S. A., Solid-phase XRN1 reactions for RNA cleavage: application in single-molecule sequencing*. Nucleic Acids Research* **2021***, 49,* 1-15*.*

2.1 Introduction

With the development of next generation sequencing (NGS), the field of transcriptomics has seen tremendous advancements creating opportunities for improved diagnostics, prognostics, and treatment of diseases such as cancers and infectious diseases.[1, 2] RNA sequencing enables measurement of single nucleotide variants (SNVs), insertions and deletions, detection of different transcript isoforms, splice variants, and chimeric gene fusions.[1] There is also an increasing interest in the study of post-transcriptional modifications of RNA and their potential role in modulating processes associated with cancer and other diseases.[3-6] Although NGS has been a useful technique for identifying specific post-transcriptional modifications, several technical challenges remain.[7, 8] Almost all current NGS techniques require library preparation prior to sequencing. During library preparation, the RNA molecules are fragmented and converted to cDNAs using reverse transcription and amplified using PCR, followed by a purification step.[7] Not only does PCR introduce biases and other artifacts that would affect the identification and quantification of transcripts, but also by using these pre-sequencing steps important RNA modification information can be lost.[9, 10]

Single-molecule nanopore sequencing has been viewed as an attractive alternative to NGS that can address many of the aforementioned issues associated with NGS.[11, 12] Of the many potential advantages of single-molecule nanopore sequencing, the most exciting are the simple and inexpensive sample preparation steps, which do not in some cases require amplification using PCR, and in most cases provide longer reads compared to NGS. Unlike NGS, nanopore sequencing does not require fluorescent labelling as the sequencing is done using DNAs and RNAs in their native state, significantly reducing the sequencing cost and time.[11] In addition, the lack of the need for amplification can preserve the post-transcriptionally modified ribonucleotides by not only eliminating PCR, but the reverse transcription step as well.

Nanopore sequencing is currently performed using two approaches, strand sequencing[13, 14] or exosequencing.[15] Although both methods have been used to sequence DNA, lesser considerations have been given to nanopore RNA sequencing. However, several reports do discuss RNA sequencing using both nanopore approaches.[10, 16-18] In one report, an engineered alpha hemolysin nanopore containing amino-cyclodextrin adapters were used together with an exoribonuclease enzyme, polynucleotide phosphorylase (PNPase), which cleaves single stranded RNA (ssRNA) in the 3' → 5' direction to produce ribonucleotide diphosphates (rNDPs). It was shown that the four canonical rNDPs could be discriminated using this exosequencing method with the additional charge on the rNDPs assisting in the capture of the cleaved rNDP by the nanopore.[18] In another approach, strand sequencing of RNA was demonstrated using immobilized RNA, where the four canonical bases (adenosine, uridine, guanine, cytosine) and modified bases (I, m6A, m5C) were successfully distinguished.[17] Exosequencing, where the biopolymer is cleaved into its constituent nucleotides in a sequential manner (either 5' → 3' or 3' → 5' direction) before passing through the nanopore is advantageous compared to strand sequencing because only one nucleotide is resident within the pore at any time.[11] Thus, the resultant current transient signal resulting from a single nucleotide resident within the pore gives a distinguishable signal.[11]

In this chapter, we lay the groundwork for an exosequencing technique for RNA using solid-state in-plane nanopores fabricated in thermoplastics, with exoribonuclease-1 (XRN1) immobilized onto a solid support.[19-22] XRN1 is a processive exoribonuclease that cleaves ssRNA in the 5' → 3' direction releasing ribonucleotide monophosphates (rNMPs). XRN1 plays a critical role in RNA turnover and participates in nonsense-mediated decay, gene silencing, rRNA maturation, and degradation of mRNAs within eukaryotic cells.[23-25] According to crystallographic data, the size of XRN1 is 15 nm x 15 nm x 15.49 nm at angles $\alpha = \beta = \gamma = 90^{\circ}$.[23] The narrow

entrance to the active site of XRN1, which is ~9 Å, only allows the entry of 5' monophosphorylated ssRNA and also helps in removing secondary structures as it cleaves through the substrate.[23-25] The 5' monophosphorylated ssRNA is required to be at least 4 nucleotides in length for efficient capture by the active site and the divalent cation Mg^{2+} acts as a cofactor to carry out its function as an exoribonuclease.[25] The clipping rate of XRN1 *in vivo* has been reported to be 38 – 55 nt s^{-1}.[26] Recently, Langeberg *et al.* measured the clipping rate of XRN1 *in vitro* and it was found to be 17.3 ±0.6 nt s^{-1} at 37ºC and pH = 7.9.[27] However, to the best of our knowledge, there has been no published report on XRN1's clipping rate and processivity when the enzyme is immobilized to a solid support. Furthermore, the presence of 45 lysine residues on XRN1 provides an abundance of potential attachment sites for covalent attachment onto a solid support bearing carboxylic acid groups using 3-(3-dimethylaminopropyl) carbodiimide/*N*-hydroxysuccinimide (EDC/NHS) coupling chemistry.

To understand the immobilization of XRN1 and its ability to cleave ssRNA into its constituent mononucleotides, we immobilized XRN1 onto pillars poised within a microfluidic device. Microfluidic devices, where enzymes are immobilized for biological reactions, are known as immobilized microfluidic enzymatic reactors, IMERs.[28, 29] There are several advantages of IMERs compared to solution phase bioreactors, such as enhanced enzymatic activity and stability, prevention of aggregation and auto-digestion, and reduced interference in downstream analysis.[30, 31] Previously, our group demonstrated that lambda exonuclease (λ-Exo), which cleaves double stranded DNA (dsDNA) to produce mononucleotides, can be covalently attached to a solid surface.[20] Immobilized λ-Exo demonstrated an average clipping rate of 1,100 ±100 nucleotides per second (nt s^{-1}), and a significantly higher processivity (~40,000 base pairs) compared to the free solution enzyme.

Of the many substrates that are available (silicon, glass, polymers) for the fabrication of both microfluidic and nanofluidic devices, thermoplastics offer many advantages due to their favorable biocompatibility, good optical properties, ease of surface modification, and the number of well-established fabrication technologies to produce devices.[32] The most commonly used thermoplastics for microfluidics are polycarbonate (PC), poly(methyl methacrylate) (PMMA), and cyclic olefin copolymer (COC).[20, 32-34] In this chapter, we used PMMA as the substrate due to its favorable properties, such as good UV/vis transparency, low autofluorescence, and good solvent and acid/base resistance as well as its ability to be UV/O_3 activated to generate surface confined carboxylic acid groups that can be used to attach biologics containing primary amine groups.[35, 36]

In this chapter, we report an IMER containing XRN1 as the immobilized enzyme for the sequential digestion of 5' monophosphorylated ssRNA for potential applications in single-molecule RNA exosequencing. XRN1 was immobilized onto a UV/O_3 activated PMMA device containing micropillars. Attachment consisted of using EDC/NHS coupling chemistry. AFM analysis showed that XRN1 only attached to the PMMA surface where it had been UV/O_3 activated and in the presence of EDC/NHS with little or no nonspecific binding. Fluorescence studies, UPLC/MS measurements, and electrophoresis data provided information on the digestion of both modified and unmodified 5' monophosphorylated RNA by both free solution and immobilized XRN1. Real-time digestion of dye labelled RNA by free solution and immobilized XRN1 was observed using fluorometry and fluorescence microscopy, respectively, allowing deduction of the processivity and clipping rate of both free solution and immobilized XRN1.

2.2 Materials and Methods

2.2.1 Device Fabrication and Assembly

The microfluidic devices used in this work were fabricated in poly (methyl methacrylate), PMMA (Plaskolite), using hot embossing (see Figure 2.1). First, a brass mold master containing the required microstructures were micromilled into a brass plate using high precision micromachining (Kern, MMP Feinwerktechnik, Murnau-Westried, Germany). The microstructures on the brass mold master were then replicated into PMMA by hot embossing (155°C, 950 psi, 150 s) using a Precision Press model P3H-15-PLX (Wabash MPI, USA). After embossing, devices were diced using a bandsaw and reservoirs were drilled at each end of the microchannel. The devices were checked using non-contact profilometry (VK-X250, Keyence, Japan) to assess if the dimensions of the devices were consistent with the brass mold's dimensions.

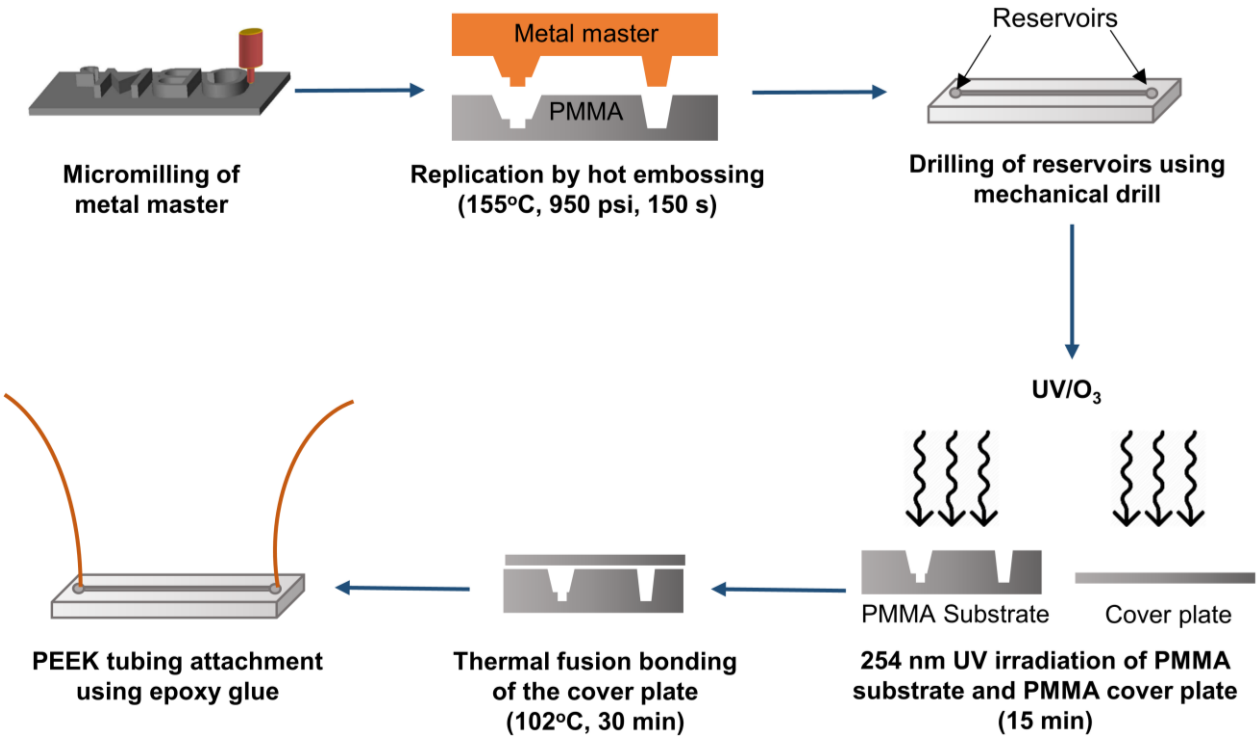


Figure 2.1. Schematic representation of fabrication and assembly of microfluidic devices.

The microchannel of the pillared IMER was 24 mm long and 1.4 mm wide and contained 3,600 micropillars each of which were 100 μm in diameter and 60 μm in height. The pillar-to-pillar spacing of the device was 35 μm. The volume and the surface area of the device were 2.9 μL and 1.17 cm^2, respectively. To allow for the real time monitoring of immobilized XRN1 digestion using fluorescence microscopy, a single channel microfluidic device was fabricated that consisted of a single flow channel (100 μm wide and 30 μm deep) made from PMMA and was hot embossed as noted above. Inlet and outlet reservoirs were drilled at each end of the microfluidic channel.

Before assembly, the microfluidic devices containing the embossed microstructures and cover plates (PMMA, 250 μm) were cleaned using isopropyl alcohol, 10% micro-90, and ddH2O followed by compressed air application to remove any debris from the microchannel. The cleaned devices were kept in an oven at 65°C for drying until further use.

Prior to attachment of the cover plate to the embossed substrate, the devices and cover plates were UV/O_3 irradiated at 254 nm (16 mW/cm^2) for 15 min, which created surface confined carboxylic acid groups to aid in the bonding as well as serving as a functional scaffold for the covalent attachment of the enzyme to the PMMA surface. After UV/O_3 treatment, the cover plate was placed on top of the microfluidic device and both were clamped together between two borosilicate glass plates (McMaster, Atlanta, GA, USA) and then, inserted into a convection oven (ThermoFisher, USA), which was set at 102°C with the assembly remaining in this oven for 30 min. After thermal fusion bonding of the cover plate to the substrate, PEEK tubing (0.007–0.020″ i.d., 1/32″ o.d., Idex Health and Science) was attached to the reservoirs using epoxy glue to facilitate reagent/sample introduction into the microchannel.

2.2.2 Enzyme Immobilization

After fabrication and assembly of the microfluidic devices, XRN1 (NEB, Ipswich, MA, USA) was covalently immobilized onto microfluidic device surfaces using EDC/NHS coupling chemistry, which is used for attachment of primary amine containing biological entities, such as XRN1, to UV/O_3 modified thermoplastic surfaces.[36, 37] For the reaction, 200 mM EDC and 50 mM NHS in 0.1 M 2-[morpholino]ethanesulfonic acid (MES) buffer at pH 4.8 was introduced into the microfluidic devices and incubated at room temperature for 15 min. This enabled the formation of succinimidyl ester groups, which are labile to nucleophilic attack.[38] When XRN1 was introduced into the microfluidic device, which displaced the EDC/NHS reagents, the amine groups on the enzyme reacted with the reactive succinimidyl ester groups forming an amide bond between the surface carboxyl groups and a primary amine resident within XRN1. After introduction of XRN1, devices were kept at room temperature for 2 h prior to washing away unbound enzyme with PBS and subsequently storing at 4°C until required for use.

Before introducing XRN1 to the EDC/NHS treated microfluidic devices, the storage buffer of the enzyme (20 mM Tris-HCl, 500 mM NaCl, 2 mM DTT, 0.1 mM EDTA, 50% Glycerol, 0.1% Triton X-100) was replaced with PBS. This was done because Tris-HCl in the storage buffer could interfere with the covalent attachment of the enzyme to the carboxylated plastic surface as Tris contains a primary amine group, which can compete with the enzyme for amide bond formation with the succinimidyl ester groups. Buffer exchange was carried out using 7 MWCO Zeba Spin columns (ThermoFisher Scientific, USA) according to the manufacturer's protocol.

2.2.3 AFM Analysis

To determine successful covalent attachment of XRN1 to activated PMMA surfaces, AFM analysis (Nanoscope IIIA, Brucker, MA, USA) was conducted. The tip used for imaging was a

DLC-300 tip with a frequency of 300 kHz and a tip radius <15 nm. Tapping mode was used with a scanning frequency of 1 Hz so that possible damage done by the tapping force applied by the tip to the immobilized enzyme was minimal. PMMA surfaces (1 cm x 1 cm) were irradiated with UV/O_3 light followed by the addition of a 40 nM XRN1 solution in the presence and absence (negative control) of EDC/NHS coupling reagents. The PMMA surfaces were kept at 4°C overnight and were rinsed with reaction buffer and distilled water and gently air dried prior to AFM imaging.

2.2.4 Protein Quantification

Pierce™ 660 nm protein quantification assay (Thermo Fisher Scientific, Waltham, MA, USA) was used to determine the amount of XRN1 covalently attached to the microfluidic channel containing micropillars. Absorbance of XRN1 solutions at 660 nm were measured (UV-VIS 1200 spectrophotometer, Shimadzu, Kyoto, Japan) before and after introduction into the IMER with attached enzyme.

2.2.5 Digestion of Monophosphorylated RNA

The model RNAs (60 nucleotides, nt) for XRN1 digestion studies were obtained from Integrated DNA Technologies, Inc. (Skokie, IL, USA). Following synthetic preparation, the 60 nt RNAs were purified using RP-HPLC and purity checked by mass spectrometry, which yielded a purity of 85-90%. The impurities were suspected to consist of truncated 60 nt RNAs lacking the 5'-monophosphorylated end and thus, would not serve as a viable substrate for XRN1 digestion.

Digestion of 5' monophosphorylated 60 nt RNA was investigated using 2.32 pmol of XRN1 enzyme in both free solution and the immobilized state. In the free solution reaction, EDTA was added to stop the reaction after the desired time. The experimental control for the free solution

reaction consisted of adding XRN1 to an RNA solution in the absence of the cofactor Mg^{2+}. In the immobilized state, the effective reaction times were achieved by hydrodynamically pumping RNA solutions through the IMERs with a suitable flow rate using a syringe pump (Harvard Apparatus 22, Harvard Apparatus, Holliston, MA, USA). The negative control for the immobilized XRN1 reaction consisted of introducing RNA solutions to the IMERs, which did not contain immobilized XRN1. The pH was set at 7.9 for both solution-phase and solid-phase reactions. For all the XRN1 digestion experiments, ≥3 trials were conducted for each data point and the average value with the standard deviation is reported.

2.2.6 Fluorescence Measurements of 5' Monophosphorylated RNA

Following the XRN1 digestion, the remaining RNA molecules were labeled with SYTO 82 (Life Technologies, Eugene, OR, USA) to assess the extent of digestion. SYTO type dyes show a quantum efficiency of ~0.4 when bound to RNA and a low quantum efficiency (0.01) in the presence of mononucleotides and the buffer alone.[39] The fluorescence emission spectra of labeled RNA solutions were measured from 490 nm to 700 nm using a Fluorolog-3 fluorimeter (Horiba Jobin Yvon, Kyoto, Japan) with 480 nm excitation. The data was analyzed using Datamax 2.0 software.

2.2.7 Denaturing Microchip Gel Electrophoresis

The sizes of each ssRNA both before and after digestion by XRN1 were measured using denaturing microchip gel electrophoresis (Agilent Tapestation 2200 instrument: Agilent Technologies, Santa Clara, CA, USA). For the experiments reported herein, we used the high sensitivity RNA Screentape gel, which is a non-rigid plastic device that contains 16 lanes each of which are 25 mm in length, 2 mm in width and 1 mm in height. The denaturing gel (50 – 75%

DMSO) consisted of 3% *N*-acryloylamido ethoxyethanol (AAEE). Gel electrophoresis data were analyzed using the Tapestation data analysis software.

2.2.8 Analysis of Digestion Products by Ultra-high-performance Liquid Chromatography (UPLC)/Mass Spectrometry (MS)

To determine the identity of XRN1 reaction products an analysis of the reaction mixture following the digestion of both unmethylated and methylated 60 nt RNA substrates with XRN1 was conducted using UPLC (Waters Acquity) coupled to a mass spectrometer (Advion Expression^{s} CMS Mass Spectrometry – electrospray ionization – system). Before carrying out UPLC/MS analysis of the XRN1 digestion products, separation conditions were optimized using an rNMP mixture containing the canonical rNMPs and the modified rNMPs as well (m5C and m6A). The concentrations used for these experiments were based on their expected abundance within the synthetic RNAs.

Unmodified and modified 60 nt RNA strands were reacted with XRN1 and after the reaction was complete, XRN1 was removed using an Amicon Ultra 3K size exclusion column (Millipore Sigma, St. Louis, MO, USA) according to the manufacturer's recommended protocol. UPLC/MS analyses of the digestion products were then performed using a Waters Acquity UPLC coupled to an Advion Expression^{s} CMS MS (electrospray ionization). UPLC used a Waters XBridge BEH C18 (2.5 µm, 4.6 x 150 mm) column and 100% (0.1% Formic acid/H_2O) mobile phase with a 1.00 mL/min flow rate.

2.2.9 *In vitro* Transcription (IVT) of Full-length Firefly Luciferase (FLuc) and Human Duchenne Muscular Dystrophy Gene (DMD) RNA Transcripts

For real time RNA digestion studies, 1,766 nt (FLuc) and 11,163 nt (DMD), homogeneous ssRNA molecules were synthesized using IVT. Plasmid Clone ID HsCD00082587 harboring the full-length dystrophin gene (DMD) open-reading frame was purchased from The PlasmID Repository within the DNA Resource Core at Harvard Medical School. Briefly, a working stock of Escherichia coli DH5α was transferred to 50 mL of Luria Broth (LB) containing spectinomycin (100 µg/mL) as the selective antibiotic marker and incubated overnight at 37°C in a rotary shaker. Plasmid DNA was isolated using mini-prep spin columns (Qiagen, MD, USA) following the manufacturer's protocol, including the RNAse treatment step. Following isolation, purified plasmid DNA was quantified using a Biophotometer D30 (Eppendorf, NY, USA) to a final stock concentration of 100 ng/µL. Linearized FLUC DNA plasmid control template was obtained from New England Biolabs (NEB, Ipswich, MA, USA).

Full-length dystrophin cDNA was amplified from 50 pg plasmid pENTR223.1 DNA using a gradient PCR protocol in combination with the following primer pair for long-range, high-fidelity PCR using Hot Start LongAmp Mastermix (NEB, Ipswich, MA, USA): DMDR, 5' – ATGCTTTGGTGGGAAGAAGTAGAG - 3'; DMDF/T7, 5' - TGA GAC ACG GGC CAG AGC TGC CAG GAA ACA GCT ATG ACC ATG TAA TAC GAC TCA CTA TAG – 3'. DMDF/T7 contains a flanking T7 promoter sequence that is subsequently incorporated during PCR. The optimized PCR conditions were as follows: Initial denature at 94oC for 1 min followed by 28 cycles at 94oC for 15 s, 62°C for 10 s, 65°C for 50 s/kb. No final extension step was implemented to avoid addition of 5'/3' dA overhangs following PCR that could potentially affect in vitro transcription reactions. The resulting PCR products were verified by standard gel electrophoresis on a 1% TBE agarose gel Bullseye Smartglow RNA/DNA stain (Midwest Scientific, Inc.; see Figure 2.2.A for results). All

lanes generated a discrete amplicon of approximately 11 kb and were subsequently pooled together for cleanup using a Monarch DNA/PCR Purification Kit (NEB, Ipswich, MA, USA) to generate DNA template that was used for in vitro transcription.

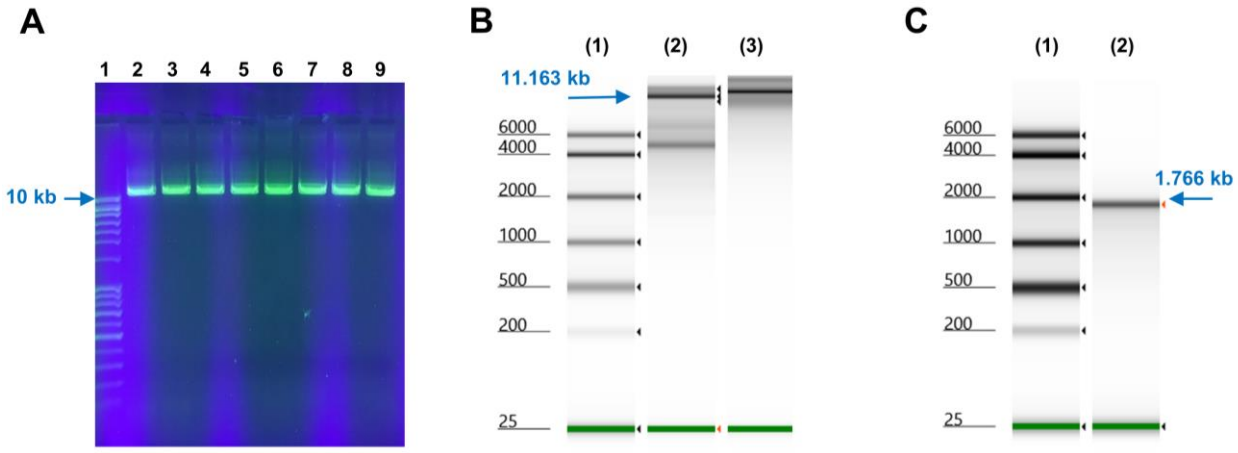


Figure 2.2. Gradient PCR amplification and IVT of full-length human dystrophin RNA (DMD) and FLUC RNA. **(A)** Agarose gel analysis of PCR product at the following gradient temperature profile: Lane 1 – DNA ladder; Lane 2 – 53.9°C; Lane 3 – 54.8°C; Lane 4 – 56.1°C; Lane 5 – 57.4°C; Lane 6 – 58.6°C; Lane 7 – 59.9°C; Lane 8 – 61.2°C; Lane 9 – 62.2°C. **(B)** High sensitivity RNA Tapestation analysis of DMD IVT products. (1) High sensitivity RNA ladder; (2) Purified IVT product. Blue arrow depicts the RNA band of interest, which was subsequently excised from an agarose gel; (3) Purified DMD RNA after gel excision, monophosphorylation, and purification. **(C)** High sensitivity RNA Tapestation analysis of FLUC IVT product. (1) High sensitivity RNA ladder; (2) Purified FLUC RNA after monophosphorylation and purification.

All IVT reactions were performed in a 20 µL final reaction volume using the HiScribe T7 High Yield RNA Synthesis kit (NEB, Ipswich, MA, USA) following the manufacturer's recommended protocol. Following IVT, 10 µL of DNase cocktail containing 10 U of DNAse I (NEB, Ipswich, MA, USA), 3 µl 10X DNase I buffer, and 2 µL nuclease-free water was added to each IVT reaction to a final reaction volume of 30 µL and incubated for 20 min to degrade DNA template. Reactions were terminated by adding 5 µL of 50 mM EDTA solution, mixed, and briefly placed on ice. Synthesized RNA product was purified using a Monarch RNA Cleanup kit (NEB, Ipswich, MA, USA) following the manufacturer's recommended protocol and eluted using nuclease-free water pre-warmed to 50°C to enhance recovery. Purified RNA product was quantified using a Biophotometer and diluted to a final stock concentration of 100 ng/µL in molecular-grade nuclease-free water (Midwest Scientific, Inc.).

Initial evaluation of full-length RNA product was determined using a standard 1% non-denaturing TBE gel pre-stained with Bullseye Smartglow RNA/DNA stain (Midwest Scientific, Inc.). First, 10 µL of RNA product was mixed with 10 µL of 2X RNA loading dye (NEB, Ipswich, MA, USA) and heat denatured for 3 min at 72°C. Following denaturation, RNA samples were immediately loaded and resolved on a non-denaturing gel. IVT consistently resulted in an ~11 kb RNA transcript in addition to several additional co-synthesized RNA products. Several optimization strategies were employed to mitigate RNA co-synthesis such as time, temperature, and nucleotide concentration that still resulted in very similar banding patterns when compared to the recommended protocol. Because it was ideal for this work to use a homogeneous RNA species, we excised the corresponding 11 kb RNA band of interest from the non-denaturing agarose gel and used an RNA gel extraction and purification kit (Zymo Research, CA, USA). Results from the gel extraction consistently showed successful isolation and purification of non-degraded, full-length RNA product (see Figure 2.2.B and C).

2.2.10 5’ Monophosphorylation of RNA

The IVT RNA products that were synthesized according to the procedure outlined in the previous section were triphosphorylated at the 5’ end. Up to 500 ng of purified IVT RNA product was treated with RNA 5´ Pyrophosphohydrolase, RppH (NEB, Ipswich, MA, USA) to remove pyrophosphate from the 5´ end of the triphosphorylated RNA to generate 5´ monophosphate RNA following the manufacturer’s recommended protocol (see Figures S3B and S3C). Following incubation, RNA samples were pooled and purified using the Monarch RNA Cleanup kit (NEB, Ipswich, MA, USA) following the manufacturer’s recommended protocol and eluted in 50 µL nuclease-free water pre-warmed to 50°C to enhance recovery. Purified RNA product was quantified using a Biophotometer D30 and diluted to a final working stock concentration of 25 ng/ µL in molecular-grade nuclease-free water (Midwest Scientific, Inc.).

2.2.11 RNA Labeling and Real Time Digestion Measurements of RNA

For real time fluorescence studies, FLUC RNA and DMD RNA were labelled either with SYTO 82 (541/560 nm; Life Technologies, Eugene, OR, USA) or RiboGreen (480/520 nm; Life Technologies, Eugene, OR, USA). SYTO 82 has an extinction coefficient >50,000 cm^{-1} M^{-1} and binds to both DNA and RNA exhibiting a quantum yield of 0.4, which is an approximate 40-fold enhancement compared to the unbound dye.[39] RiboGreen, on the other hand, is specific to RNA and shows a fluorescence enhancement of 1,000 upon binding to RNA.[40]

For labelling RNA with either of these dyes, a 5-fold molar excess of dye compared to the total number of nucleotides present in the strand was used. RNA solutions were heated at 72°C for 3 min and flash cooled in ice prior to adding the 5-fold molar excess of staining dye. The RNA-dye solutions were kept at room temperature for 30 min and the excess dye was removed from the solution using 7 MWCO size exclusion spin columns. Before using the stained RNA for experiments, 1X buffer was added with or without 10 mM $MgCl_2$ (also included 100 mM NaCl, 50 mM Tris-HCl, and 1 mM DTT) according to the experimental need.

2.2.12 Clipping Rate and Processivity of XRN1

For determination of the clipping rate and processivity of XRN1 in free solution, a method described by Han *et al.* was used with slight modifications.[41] First, 0.0875 pmol of RiboGreen dye (Life technologies, Eugene, OR, USA) labelled FLuc RNA was incubated with 0.35 pmol of XRN1 in the presence of XRN1 buffer (100 mM NaCl, 50 mM Tris-HCl, 1 mM DTT) without Mg^{2+} to enable complexation of FLuc RNA to XRN1 without clipping. Next, 8.75 pmol of a competitor RNA, in this case a 60 nt RNA saturated with SYTO 82 dye (Life Technologies, Eugene, OR, USA) in XRN1 buffer with Mg^{2+} (final Mg^{2+} concentration of 20 mM) was added to FLuc RNA complexed to XRN1, and the fluorescence intensity of the mixture was measured at 30 s intervals for 30 min

with 470 nm excitation and 500 nm emission using the Fluorolog-3 fluorimeter. Data were analyzed using Datamax 2.0 software.

To deduce the clipping rate and processivity of immobilized XRN1, a method described by Oliver-Calixte *et al.* was used.[20] XRN1 was immobilized to the cover plate of the single channel microfluidic device (no pillars) and SYTO 82 labelled DMD RNA was introduced into the device in enzyme buffer without Mg^{2+} to facilitate complexation of DMD RNA with immobilized XRN1. Complexed DMD RNA to XRN1 was determined by monitoring the fluorescence of single RNA molecules to make sure that they were not randomly diffusing (see SI for more details). Once a single DMD RNA molecule was located that was complexed to the immobilized XRN1, enzyme buffer containing Mg^{2+} cofactor was introduced into the device to initiate digestion and the fluorescence intensity (532 nm excitation, 0.01 W) of the DMD RNA-XRN1 complex was monitored continuously. For these experiments, an epifluorescence microscope was used, which consisted of a NIKON TE 2000 microscope fitted with a 100X/1.4 NA oil-immersion objective and an Andor iXon3 EMCCD camera.[21] All images were acquired using Metamorph advanced v7.5.6.0 software (10 fps acquisition rate). Acquired images were analyzed using Image J software.

2.2.13 Statistical Analysis

All reported data sets were compared by a two-sided t-test using R Studio v1.0.153 and R v3.5.1 software.

2.3 Results

2.3.1 XRN1 Immobilization and Quantification

XRN1 contains a total of 45 lysine residues, most of which reside opposite to the active site. These lysine residues act as attachment sites available for conjugation to –COOH functional groups on the UV/O_3 activated PMMA surface. To confirm successful covalent attachment of XRN1 onto UV/O_3 activated PMMA surfaces using EDC/NHS coupling chemistry (see Figure 2.3.A), an AFM analysis was carried out to determine the presence of morphological features indicative of covalently immobilized XRN1. Sheet PMMA surfaces were exposed to UV/O_3 light and a 40 nM XRN1 solution was introduced without and with EDC/NHS reagents (Figure 2.3.B and C, respectively). As shown in Figure 2.3.B and 2.3.C, surface features could be seen in EDC/NHS treated PMMA surfaces with the absence of such features in the case of no EDC/NHS reagents.

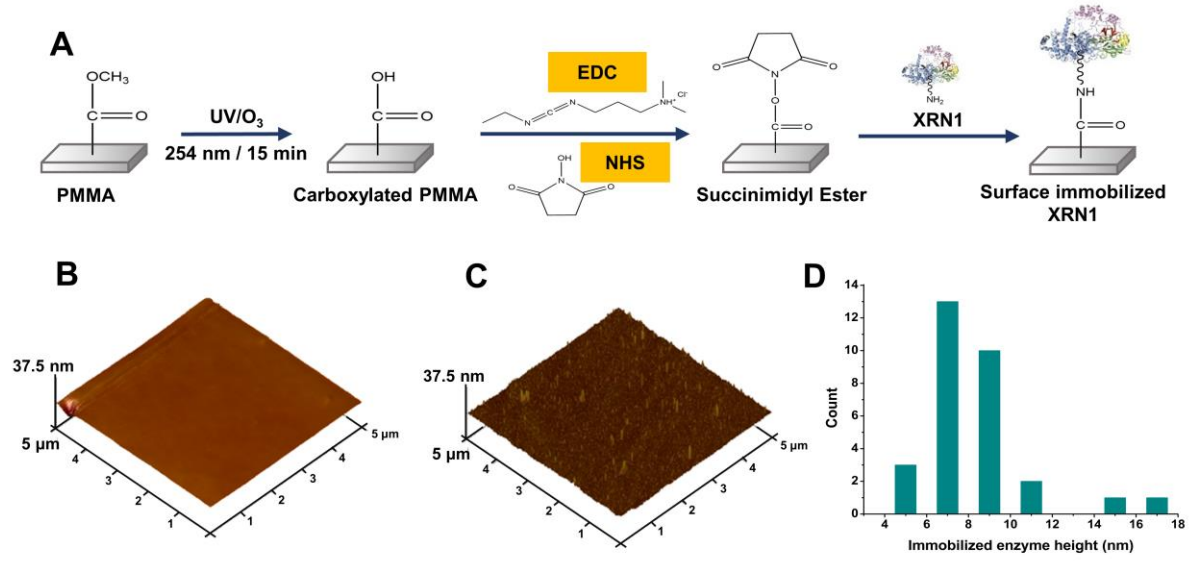


Figure 2.3. Covalent attachment of XRN1 onto UV/O_3 activated PMMA. **(A)** Schematic representation of the process of covalent attachment of XRN1 onto PMMA surface by EDC/NHS coupling reaction. 5 µm x 5 µm AFM image of PMMA surface after UV/O_3 activation, and incubation with 40 nM XRN1 enzyme **(B)** without **(C)** with EDC/NHS coupling reagents. **(D)** Height distribution of surface features present on (C). The average height of a surface structure is 8.4 ±0.5 nm.

To determine the heights of the surface features present in EDC/NHS treated PMMA surfaces, surface structures in Figure 2.3.C were measured (see Figure 2.3.D). According to the height distribution, the average height was determined to be 8.4 ±0.5 nm, which was near the size of this molecule in terms of its crystal structure, which is ~15 nm.[23] The slight disparity in size could be due to the compression of the enzyme by the tapping force applied by the AFM tip and/or size reduction of the enzyme due to dehydration.[42, 43]

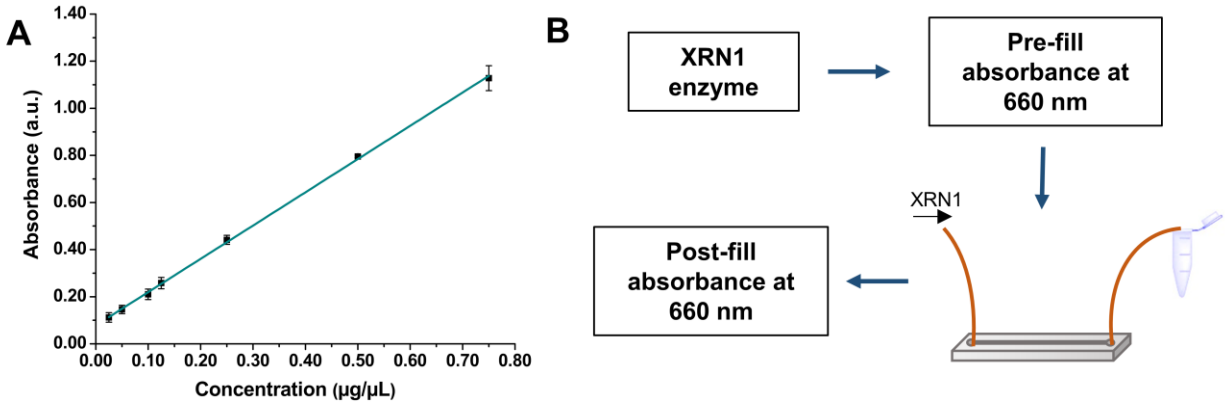


Figure 2.4. Quantification of immobilized enzyme on IMERs. **(A)** Calibration plot of Pierce 660 nm protein quantification assay (R2 = 0.9995). **(B)** Schematic representation of experimental procedure.

We used a protein quantification assay to determine the amount of XRN1 covalently attached inside the micropillared IMER (surface area 1.17 cm^2). This assay uses a dye-metal based total protein quantification methodology and the complexation of the dye-metal complex onto a protein shifts the absorbance maximum to 660 nm from 450 nm.[44] The calibration plot (R^2 = 0.9995) for the assay plotted using BSA protein standards provided by the manufacturer is shown in Figure 2.4.A. For these experiments, three different input concentrations of XRN1 were used (183 nM, 305 nM, 426 nM) based on the amount of XRN1 needed for a theoretical monolayer coverage of the IMER (1.1 x 10^{11} molecules), and the lowest XRN1 enzyme concentration that can be measured using the protein quantification assay. The IMERs were assembled and prepared for enzyme immobilization as previously described. Absorbance of each XRN1 solution was

measured at 660 nm before introducing 20 µL into the IMERs for enzyme covalent surface attachment. The eluant was collected and the absorbance was measured again at 660 nm. A schematic representation of the experimental procedure is shown in Figure 2.4.B. The extent of nonspecific adsorption was assessed by introducing XRN1 solutions to IMERs that had not been treated with EDC/NHS coupling reagents following UV/O_3 activation of the polymer.

Table 2.1. Number of pmol of enzyme nonspecifically adsorbed.

Concentration (nM)	Average pre-fill absorbance (n = 3)	Average post-fill absorbance (n = 3)	Adsorbed XRN1 amount (pmol)
183	0.1323 ±0.0007	0.1313 ±0.0012	0.0142 ±0.0086
305	0.1692 ±0.0004	0.1689 ±0.0002	0.0270 ±0.0168
426	0.2061 ±0.0002	0.2040 ±0.0003	0.1672 ±0.0247

In the absence of the coupling reagents, XRN1 would only adsorb to the surface non-specifically and the number of moles was calculated to be <2% of the input number of moles for each XRN1 concentration (see Table 2.1). When the IMERs were treated with EDC/NHS coupling reagents for 15 min prior to flowing XRN1 solutions through the devices, the number of moles of enzyme covalently attached increased with increasing input XRN1 concentration and ranged from 2.32 - 4.07 pmol of XRN1 (see Table 2.2). The total number of moles immobilized from the total input number of moles ranged from 53.4% - 39.4% while the surface density of immobilized enzyme ranged from 1.98 pmol/cm^2 to 3.48 pmol/cm^2.

Table 2.2. Number of pmol of enzyme lost to immobilization for each input concentration of XRN1.

Concentration (nM)	Average pre-fill absorbance (n = 3)	Average post-fill absorbance (n = 3)	Immobilized XRN1 amount (pmol)
183	0.1323 ±0.0007	0.1036 ±0.0083	2.32 ±0.67
305	0.1692 ±0.0004	0.1325 ±0.0048	2.97 ±0.38
426	0.2061 ±0.0002	0.1558 ±0.0082	4.07 ±0.67

2.3.2 Digestion Studies of 5' Monophosphorylated RNA

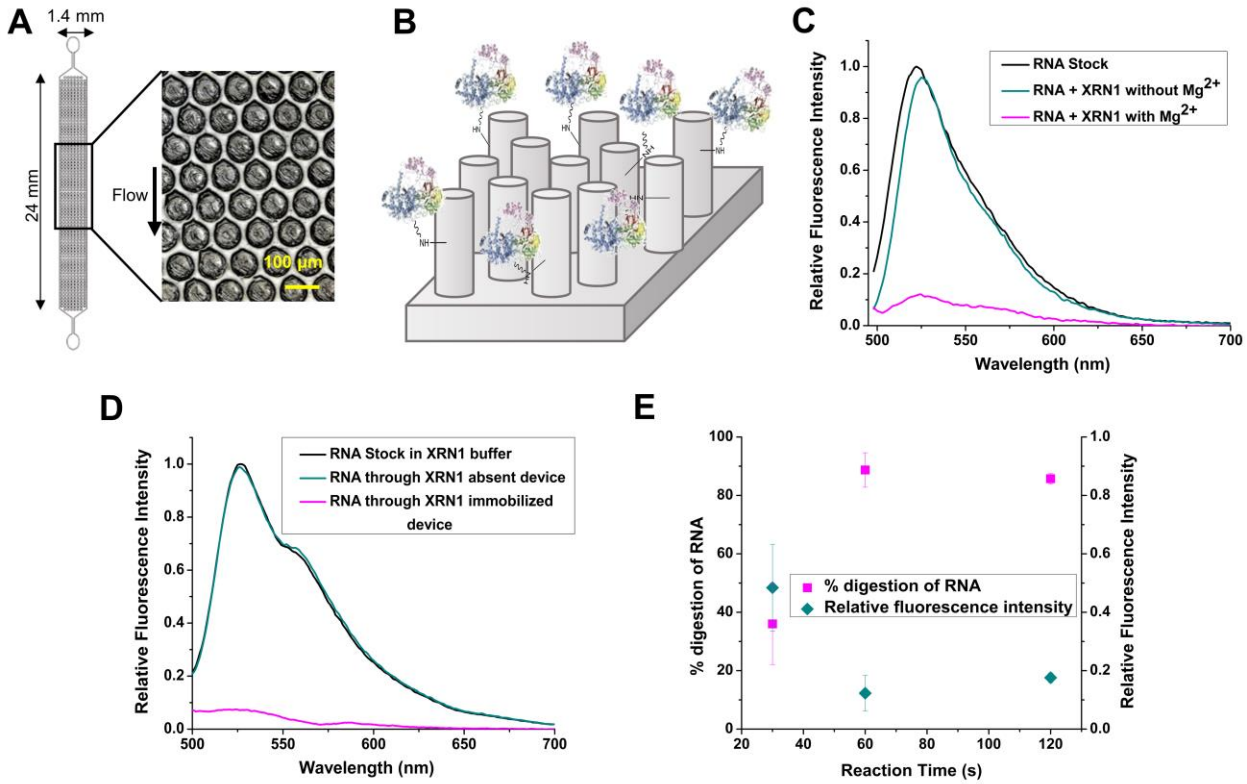


Figure 2.5. Solid-phase digestion reactions of XRN1. **(A)** Top-down view of the pillared IMER channel. **(B)** Schematic representation of the covalently attached enzyme on the micro-pillars of the device. Fluorescence emission spectra of SYTO RNASelect Green labeled monophosphorylated RNA solutions digested by XRN1 in **(C)** free solution and **(D)** Immobilized state. The reaction time was 60 s and 2.32 pmol of enzyme was used in both free solution and immobilized digestion. SYTO RNASelect Green was added after digestion and fluorescence emission spectra were taken from 495 nm to 700 nm with 480 nm excitation. **(E)** Percentage digestion and relative fluorescence intensity of digested RNA with varied reaction time and constant surface enzyme density. The XRN1 reactions were all performed at room temperature. The error bars represent standard deviations in the measurements (n ≥ 3).

To demonstrate the ability of XRN1 to digest monophosphorylated RNA, 10.6 pmol of a 60 nt RNA substrate was reacted with 2.32 pmol of XRN1 both in solution- and the solid-phase. For solid-phase XRN1 experiments, we used a micropillared IMER consisting of 3,600 micropillars (see Figure 2.5.A). A schematic representation of the enzyme-immobilized IMER is shown in Figure 2.5.B. Fluorescence emission spectra shown in Figure 2.5.C depict the digestion of RNA by free solution XRN1 in the presence and in the absence of the cofactor, Mg^{2+}. As can be seen from the spectrum depicted in dark cyan in Figure 2.5.C, in the absence of Mg^{+2} with XRN1

present there was no change in the fluorescence spectrum of the RNA labeled with SYTO RNASelect Green compared to the RNA solution with no XRN1, indicating the 60 nt RNA remained intact after 60 s. When the Mg^{2+} cofactor was introduced into the reaction mixture, the 60 nt RNA was digested as shown by the loss of fluorescence due to cleavage of the RNA. Peak area analysis of the spectra yielded a digestion efficiency of 78.3 ±4.4% (n = 3; T = 25°C) after 60 s for XRN1 solution-phase reactions.

Figure 2.5.D shows fluorescence spectra of SYTO RNASelect Green-labeled 60 nt RNA reacted with XRN1 when immobilized within the IMER. The negative control for this experiment consisted of flowing 10.6 pmol of the 60 nt RNA substrate through the IMER that did not contain immobilized XRN1. The negative control revealed that there was no loss of intact RNA molecules as evident by the emission spectrum appearing similar to the RNA stock solution. When the 60 nt RNA was flowed through the IMERs containing immobilized XRN1, the amount of intact RNA decreased as apparent from the loss of fluorescence seen in the magenta trace in Figure 2.5.D. Peak area analysis of the spectra indicated that 87.6 ±2.8% (n = 4; T = 25°C) of the 60 nt RNA was digested by the immobilized XRN1 enzyme. However, these numbers should be qualified by the fact that the RP-HPLC purified 60 nt models contained RNA fragments that were not 5'-monophosphorylated making them a non-viable substrate for XRN1.

To assess the effect of surface enzyme density on the activity of immobilized XRN1, the 60 nt RNA substrate was introduced into the XRN1 immobilized IMERs for 60 s with different XRN1 concentrations used for the immobilization reaction. The digestion percentages were above 80% for all surface enzyme densities used (Table 2.3) and showed no significant statistical difference at the 95% confidence level as determined by a *t*-test.

Table 2.3. Percent digestion as a function of enzyme load.

Varied Enzyme Concentration			
pmol of enzyme	**2.32 (n = 3)**	**2.97 (n = 3)**	**4.07 (n = 3)**
% RNA digestion	88.7 ±5.9	83.8 ±6.9	82.9 ±8.7

Next, to evaluate the effect of reaction time on the percent digestion of the 60 nt RNA substrate, surface enzyme density and the concentration of RNA solutions were kept constant, while the reaction times were varied by changing the flowrate of the RNA substrate through the IMERs. When a reaction time of 30 s was used, the digestion percentage was ~35% and when the reaction time was increased to 60 s and 120 s, the digestion percentages were >80% (see Figure 2.5.E) and showed no significant difference at the 95% confidence level (p = 0.5284).

2.3.3 Ability of XRN1 to Digest through Methylated RNA Sequences

A **m6A methylated RNA:**

5'-monophos- rArUrC rU/**iN6Me-rA**/rU rGrUrU rCrUrC rCrArC rCrArA rUrUrG rUrArC rUrArU rCrArC rUrCrU rUrCrG rCrUrU rCrUrC rUrUrC rCrUrU rUrCrG rUrUrA rUrArA rArGrU

B **m5C methylated RNA:**

5'-monophos- rArUrC rUrArU rGrUrU /**i5M-rC-TOM**/rUrC rCrArC rCrArA rUrUrG rUrArC rUrArU rCrArC rUrCrU rUrCrG rCrUrU rCrUrC rUrUrC rCrUrU rUrCrG rUrUrA rUrArA rArGrU

C **Unmethylated RNA:**

5'-monophos- rArUrC rU rA rU rGrUrU rCrUrC rCrArC rCrArA rUrUrG rUrArC rUrArU rCrArC rUrCrU rUrCrG rCrUrU rCrUrC rUrUrC rCrUrU rUrCrG rUrUrA rUrArA rArGrU

Figure 2.6. RNA sequences and digestion of methylated RNA. **(A)** Sequence of m6A methylated RNA. **(B)** Sequence of m5C methylated RNA. **(C)** Sequence of unmethylated control RNA.

We next investigated the ability of both solution-phase and solid-phase XRN1 to digest through sequences that contained methylated bases using two 60 nt RNA sequences. The sequences of RNA 60mers are shown in Figure 2.6.A – 2.6.C. Two of the 60mers contained one of the two most common RNA modifications found in eukaryotic cells. One RNA sequence contained an N^6-methyladenosine (m6A) residue (see Figure 2.7.A) at the 5th nucleotide position within the 60mer,

while the second RNA sequence contained a 5-methylcytosine (m5C) residue (see Figure 2.7.B) at the 10th nucleotide position from the 5' end. A third unmethylated RNA sequence was used as the control sequence.

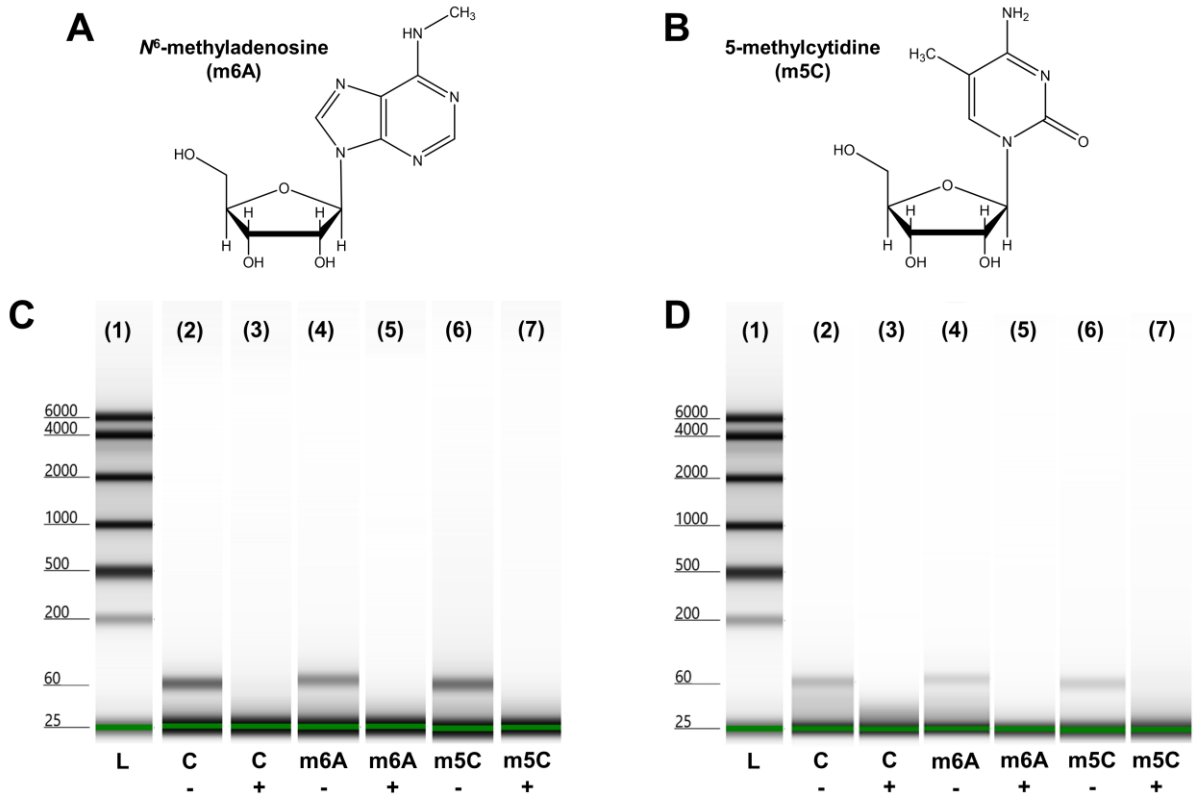


Figure 2.7. Digestion of methylated RNA sequences. Chemical structures of **(A)** m6A and **(B)** m5C. Digestion of methylated RNA sequences by **(C)** solution phase and **(D)** immobilized XRN1. (1) Ladder (L). Negative control for (2) unmethylated (c -) (4) m6A-methylated (m6A -) and (6) m5C-methylated (m5C -) RNA. Digestion results for (3) unmethylated (c +) (5) m6A-methylated (m6A +) and (7) m5C-methylated (m5C +) RNA by XRN1.

Each RNA sequence was reacted with immobilized XRN1 in the IMER for 60 s. RNA in the absence of XRN1 were used as the negative control. After the reactions were complete, denaturing microchip gel electrophoresis was conducted to determine the length of the remaining RNAs. If XRN1 was unable to digest through the methylated nucleotides, RNAs with a length of 51 nucleotides should remain for the m5C RNA and 56 nucleotides for the m6A RNA. As can be seen from Figures 2.7.C and 2.7.D, peaks corresponding to 60, 51 or 56 nucleotides were not observed after 60 s of reaction for both solution and IMER reactions demonstrating the ability of

XRN1 to digest through RNA structures containing m6A and m5C residues. We note that the signal intensity difference in the negative controls between solution-phase and the solid-phase XRN1 reactions is likely due to some sample loss while collecting the sample eluant from the IMERs.

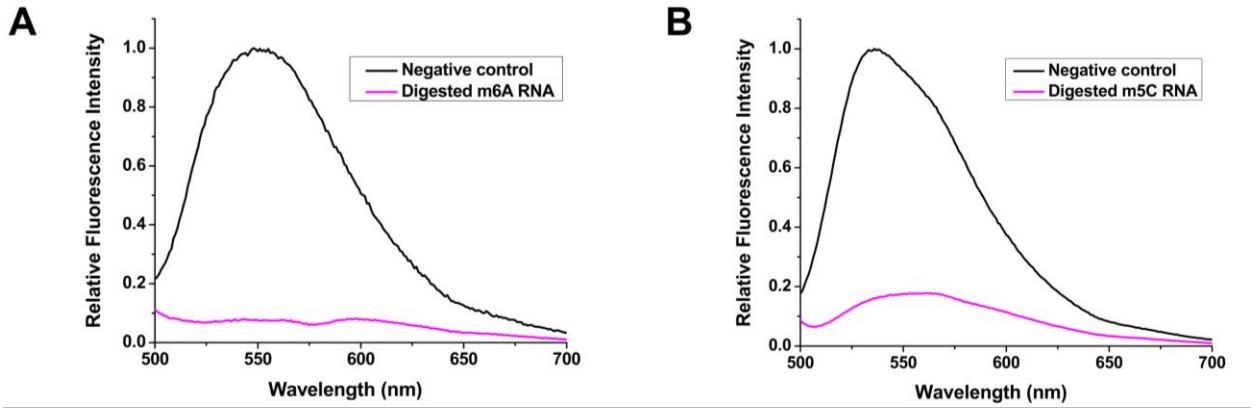


Figure 2.8. Fluorescence emission spectra of (A) m6A methylated RNA and, (B) m5C methylated RNA. After 60 s reaction time 87.0 ±4.2% (n = 4; T = 25°C) and 77.3 ±6.0% (n = 3; T = 25°C) of m6A and m5C RNA was digested, respectively by XRN1.

To determine the extent of digestion of methylated RNAs by the immobilized XRN1, the IMER digested RNA solutions were stained with SYTO 82 post-digestion and the fluorescence emission spectra were taken from 490 nm to 700 nm with 480 nm excitation. Peak area analysis of these spectra revealed that after 60 s of reaction, 87.0 ±4.2% (n = 4; T = 25°C) of m6A methylated RNA was digested (see Figure 2.8.A), while after the same amount of time, 77.3 ±6.0% (n = 3; T = 25°C) of m5C methylated RNA was digested by the immobilized XRN1 (see Figure 2.8.B). The digestion of m5C RNA seemed to be somewhat slower compared to the m6A RNA for the surface immobilized XRN1 ($P = 0.0243$). If XRN1 digestion was terminated at the methylation sites, the

fluorescence intensity would be closer to that of the negative control due to the fact that the oligomers remaining (56 nt and 51 nt) were close in size to the starting RNA 60mer.

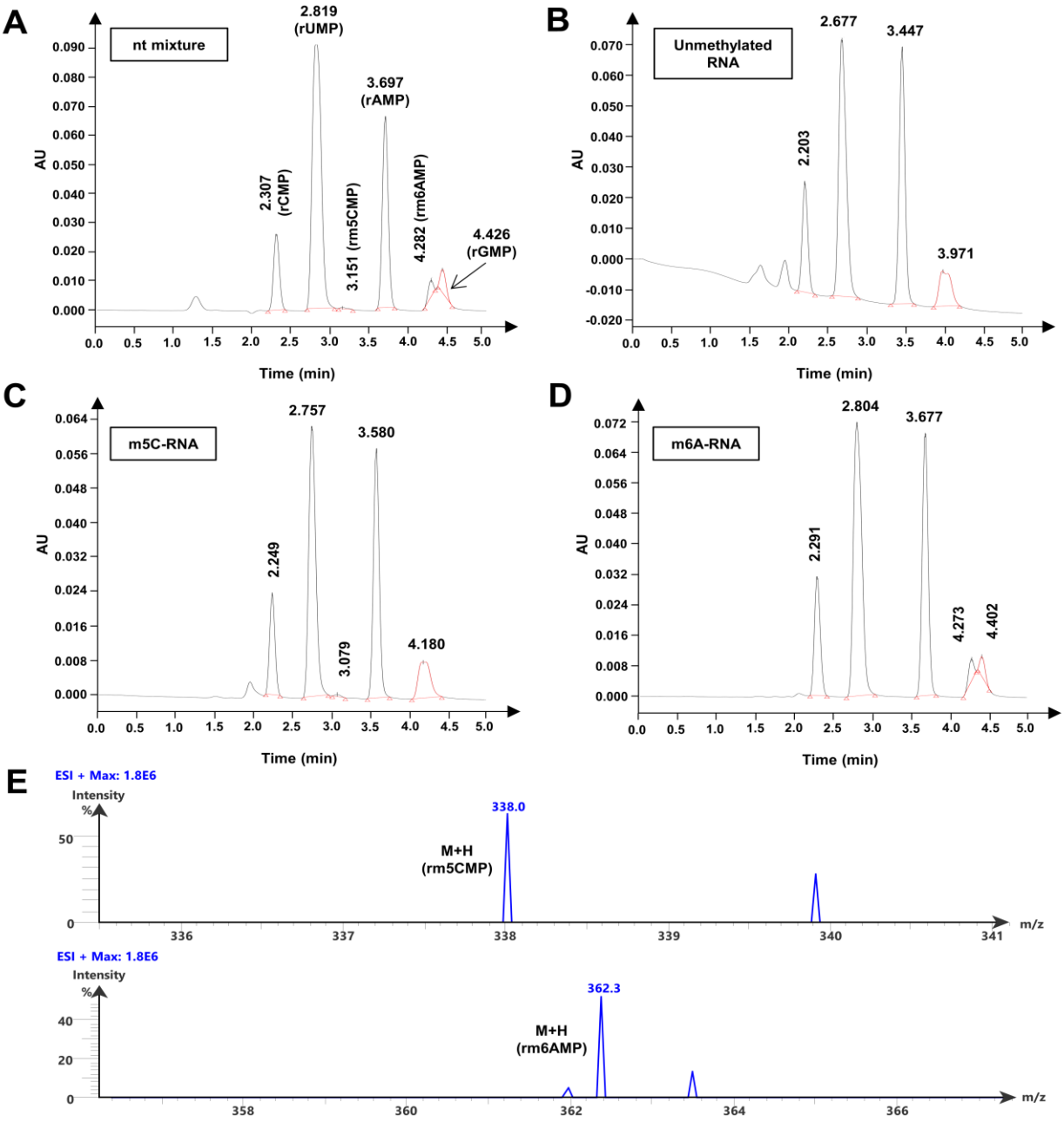


Figure 2.9. UPLC/MS analysis of digestion products from XRN1 reactions. Chromatograms (UV detection at 254 nm) of: **(A)** mixture of rNMPs; **(B)** unmethylated 60 nt synthetic RNA; **(C)** m5C methylated synthetic RNA; and **(D)** m6A methylated synthetic RNA. The UPLC/MS analysis was run after reaction with XRN1. **(E)** [M+H] peaks for m5C and m6A modified synthetic RNA oligomers obtained after digestion by XRN1.

We also subjected both unmethylated and methylated RNA strands digested by XRN1 to UPLC/MS analysis. As a reference, we ran an rNMP standard mix containing each rNMP in the expected concentration if the input 60mer RNA was fully digested by XRN1 (see Figure 2.9.A). If XRN1 was unable to digest through the m6A methylated RNA, only the first four nucleotides would be cleaved by XRN1 resulting in an RNA product of 56 nt. Therefore, the four nucleotides, rAMP, rUMP, rCMP, and rGMP, would appear in the chromatogram in a 1:2:1:0 intensity ratio, respectively. Moreover, m5C would generate a 2:5:1:1 (A:U:C:G) ratio if the digestion was terminated at the methylation site. These intensity ratios were not observed. The peaks for each ribonucleotide for XRN1 60mer RNA reactions were in the expected intensity ratio to a fully digested 60mer RNA (see Figures 2.9.B – 2.9.D).

Furthermore, the UPLC/MS results indicated that the digestion products of XRN1 were indeed 5’ rNMPs and the methylations in the resultant rNMPs were preserved (see Figure 2.9.E). We investigated the mass spectra of both the rNMP mixture and m6A methylated RNA to determine the composition of the overlapped UPLC peaks at 4.2 and 4.4 min, which could have arisen from rGMP and m6-rAMP or 8-oxo-guanosine monophosphate. Guanosine is the most susceptible nucleotide to oxidation with an oxidation product 8-oxo-guanosine monophosphate, which has a molecular weight of 379.2 g/mol.[45] The [M+H] mass spectrum for either the ribonucleotide mixture or the m6A-RNA did not contain a peak at 380.2, which indicated that 8-oxo-guanosine was not found.

2.3.4 Clipping Rate and Processivity of XRN1

The clipping rate and processivity of XRN1 are important parameters in understanding the enzyme activity both in free solution and the immobilized state for a number of applications. Thus,

we assessed these properties of both the free solution and immobilized XRN1 using IVT RNA substrates.

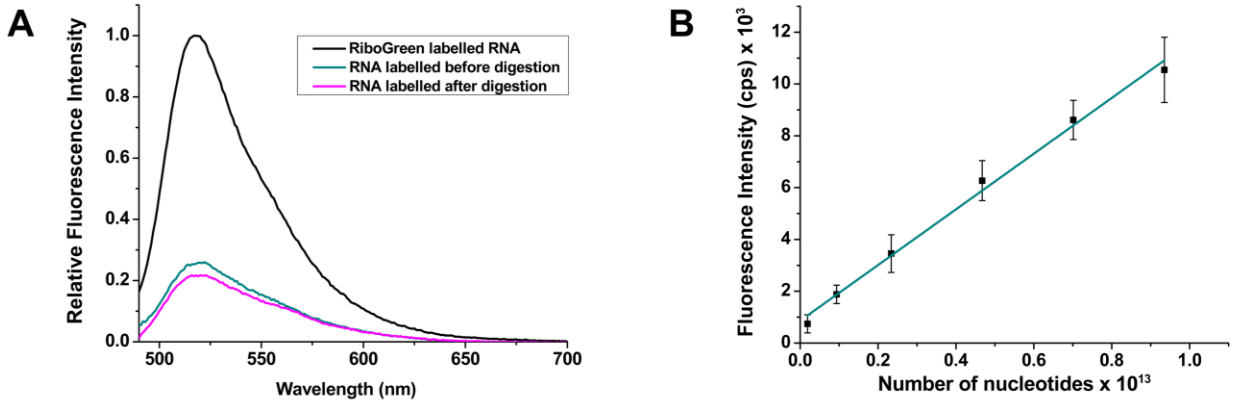


Figure 2.10. RiboGreen labelling of FLuc RNA molecules. **(A)** Fluorescence emission spectra of pre-labelled digestion and post-labelled digestion of RiboGreen labelled RNA. **(B)** Fluorescence intensity vs. number of nucleotides for RiboGreen labelled RNA (R^2 = 0.9927).

To determine the clipping rate and processivity of free solution XRN1, we used FLuc RNA labelled to saturation with RiboGreen dye as the substrate and a 60 nt RNA as the competitor that was labeled with SYTO 82. Experiments carried out using RiboGreen labelled and unlabeled FLuc RNA showed that there was no statistical difference in digestion rates for the labeled vs. unlabeled substrates at the 95% confidence interval (p = 0.5196; see Figure 2.10.A). To prevent the released RiboGreen dye molecules from attaching to the competitor 60mer RNA generating a fluorescence background, the competitor RNA molecules were labelled to saturation with SYTO 82, which did not produce a fluorescence signal using the RiboGreen filter set due to spectral dissimilarities between these dyes (see Figure 2.11.A). If the processivity of XRN1 is below 1,766 nucleotides, the FLuc RNA will detach from the enzyme and the re-engagement of the partially digested FLuc RNA to XRN1 will be prevented by the smaller competitor 60mer RNA, which was in a 100-fold molar excess compared to the FLuc RNA. Therefore, because undigested FLuc RNA will show fluorescence in the RiboGreen spectral range specifically, partial cleavage will give a

constant fluorescence signal after the reaction was terminated due to the remaining residual FLuc RNA.

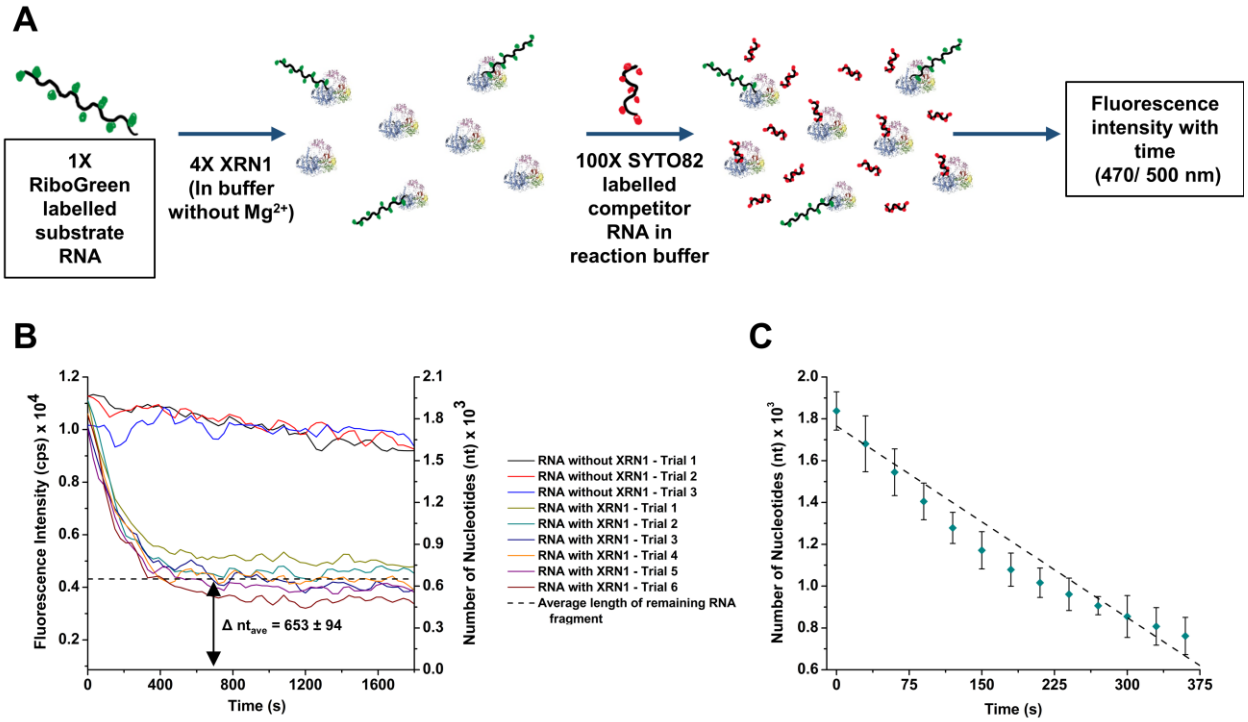


Figure 2.11. Solution phase clipping rate and processivity of XRN1. **(A)** Schematic representation of the reaction procedure. **(B)** Fluorescence intensity of RiboGreen labelled FLuc RNA with time. According to the average length of FLuc RNA fragment remaining after the reaction (Δ nt_{ave}), the processivity of XRN1 in solution phase is 1113 ±132 nucleotides. **(C)** Clipping rate calculated using the fluorescence decay portion from 5A. According to the slope of the graph (R^2 = 0.99121), the average clipping rate of XRN1 in solution is 3.06 ±0.11 nt s^{-1} at 25˚C.

As shown in Figure 2.11.B, the fluorescence intensity of the solution decreased with time and came to a constant value at ~330 s. The background was measured in the presence of SYTO 82 labeled 60 nt RNA and RiboGreen dye in the absence of FLuc RNA. The background signal was subtracted from the fluorescence emission spectra shown in Figure 2.11.B. Using a calibration plot between the fluorescence intensity and the number of nucleotides (see Figure 2.10.B), the constant fluorescence value was converted to the number of nucleotides, which was found to be 653 ±94 nt (Δnt_{ave}). The Δnt_{ave} represented the average length of FLuc RNA remaining after the XRN1 reaction, which indicated that the processivity of XRN1 in free solution was 1,113 ±132 nt.

Using the fluorescence decay portion of Figure 2.11.B, the clipping rate of XRN1 was 3.06 ±0.11 nt s^{-1} (n = 6, R^2 = 0.99332) at 25°C (see Figure 2.11.C).

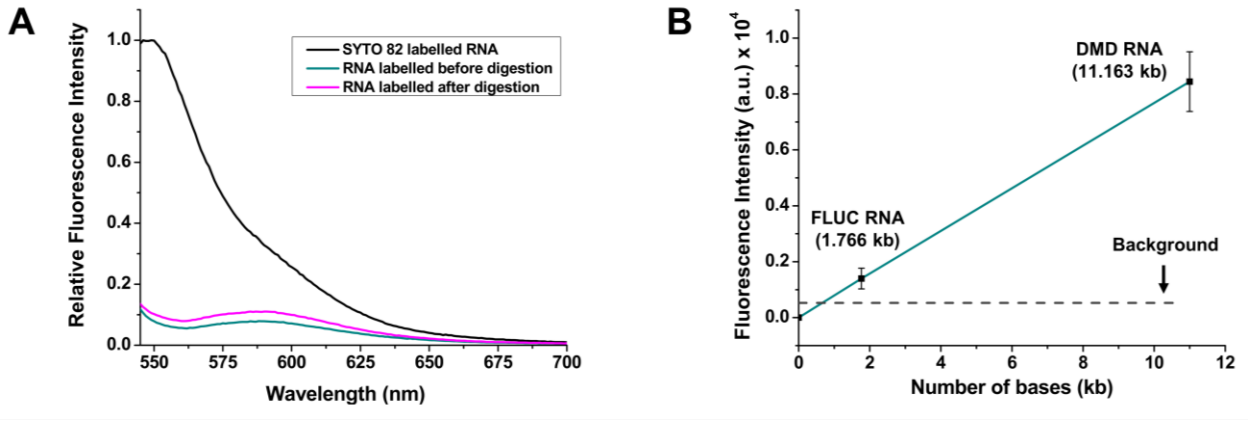


Figure 2.12. SYTO 82 labelling of RNA molecules. **(A)** Fluorescence emission spectra of pre-labelled digestion and post-labelled digestion of SYTO 82 labelled RNA. **(B)** RNA calibration plot for identification of the lowest detectable RNA fragment length using fluorescence microscope (R^2 = 0.99996).

For determining the clipping rate and processivity of immobilized XRN1, we monitored the fluorescence of a single DMD RNA (11.1 kb) molecule labeled with SYTO 82 and associated to a single immobilized XRN1 molecule using a high-sensitivity fluorescence microscope equipped with an EMCCD camera. In this case, we used a longer RNA strand compared to FLuc to produce a brighter fluorescence signal from a single RNA molecule. Analysis of the fluorescence emission obtained for DMD RNA labeled with SYTO 82 both pre-digestion and post-digestion did not show a statistical difference at the 95% confidence interval (p = 0.1573) indicating that the labeling had no influence on the activity of XRN1 (see Figure 2.12.A).

XRN1 was immobilized onto the cover plate of a single-channel microfluidic device made from PMMA. Then, single DMD RNA molecules were flowed hydrodynamically through the microchannel and when the fluorescence generated from a single RNA molecule was found to remain stationary, it was assumed to be associated to the XRN1 immobilized enzyme (see Figure

2.13.A). Unassociated RNA molecules moved in and out of the field-of-view of the microscope when the flow was stopped (see Figure 2.13.B).

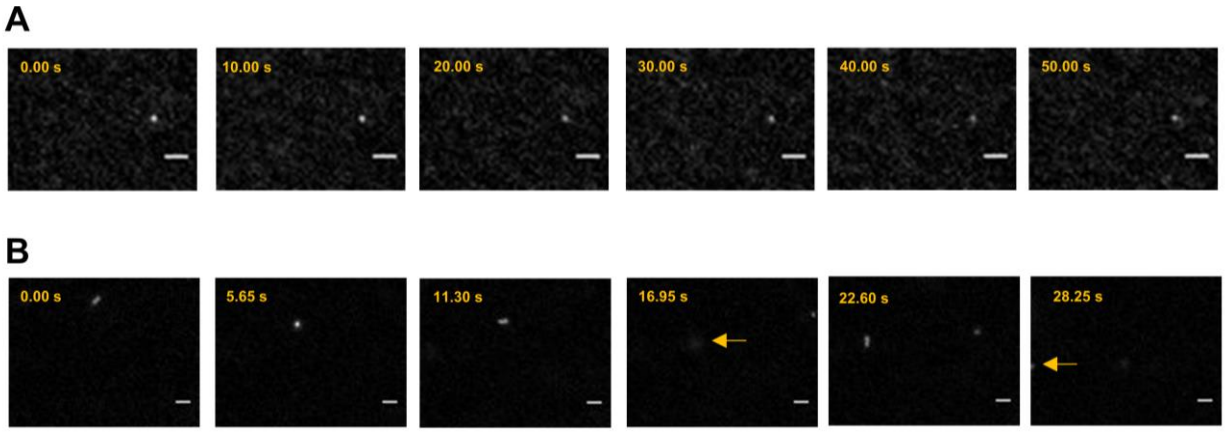


Figure 2.13. SYTO 82 labelled DMD RNA in the single channel microfluidic device. **(A)** Labelled RNA-immobilized XRN1 complex on the cover plate of the microfluidic device. Due to complexation with immobilized XRN1, the RNA molecules remained stationary with time. **(B)** Free flowing labelled DMD RNA that is moving in and out of the imaging plane and eventually moving out of the field of view with time. The yellow arrow shows the position of the out-of-plane RNA molecule. The scale bar denotes 2 µm.

The enzyme cofactor (Mg^{2+}) was then flowed into the microchannel and the fluorescence was monitored in real time to determine the processivity and clipping rate. When buffer containing Mg^{2+} was introduced into the microfluidic device, the fluorescence intensity of the stationary RNA molecules decreased with time (see Figure 2.14.A). To confirm that the fluorescence loss was due to clipping of RNA by XRN1 and not to photobleaching, control experiments were carried out in which the stationary RNA molecules were exposed to the excitation light and the fluorescence intensity was measured as a function of time. As can be seen in Figure 2.14.B, there was no significant reduction of fluorescence intensity of RNA molecules when the Mg^{2+} cofactor was absent. This is also in agreement with the literature, which reported that SYTO 82 exhibits low levels of photobleaching.[39] To determine the shortest detectable SYTO 82 labeled RNA fragment, a calibration plot was constructed using DMD RNA and FLuc RNA with SYTO 82 labeling (Figure 2.12.B). According to the calibration plot, the smallest RNA fragment that could be detected was 664 nt.

During the fluorescence imaging experiment, a single RNA molecule associated with the immobilized XRN1 did not move out of the microscope's field-of-view, indicating that the RNA molecule was associated to the immobilized XRN1 and when the fluorescence spot being imaged disappeared, the remaining fragment of RNA, if present, was below 664 nt. As can be seen from Figure 2.14.A, for a single RNA molecule, the signal completely disappeared. This suggested that the apparent processivity of immobilized XRN1 should be greater than or equal to 10,499 nt. We note that the processivity of XRN1 in the immobilized state is given as an apparent processivity due to the indirect nature of obtaining the data.[20]

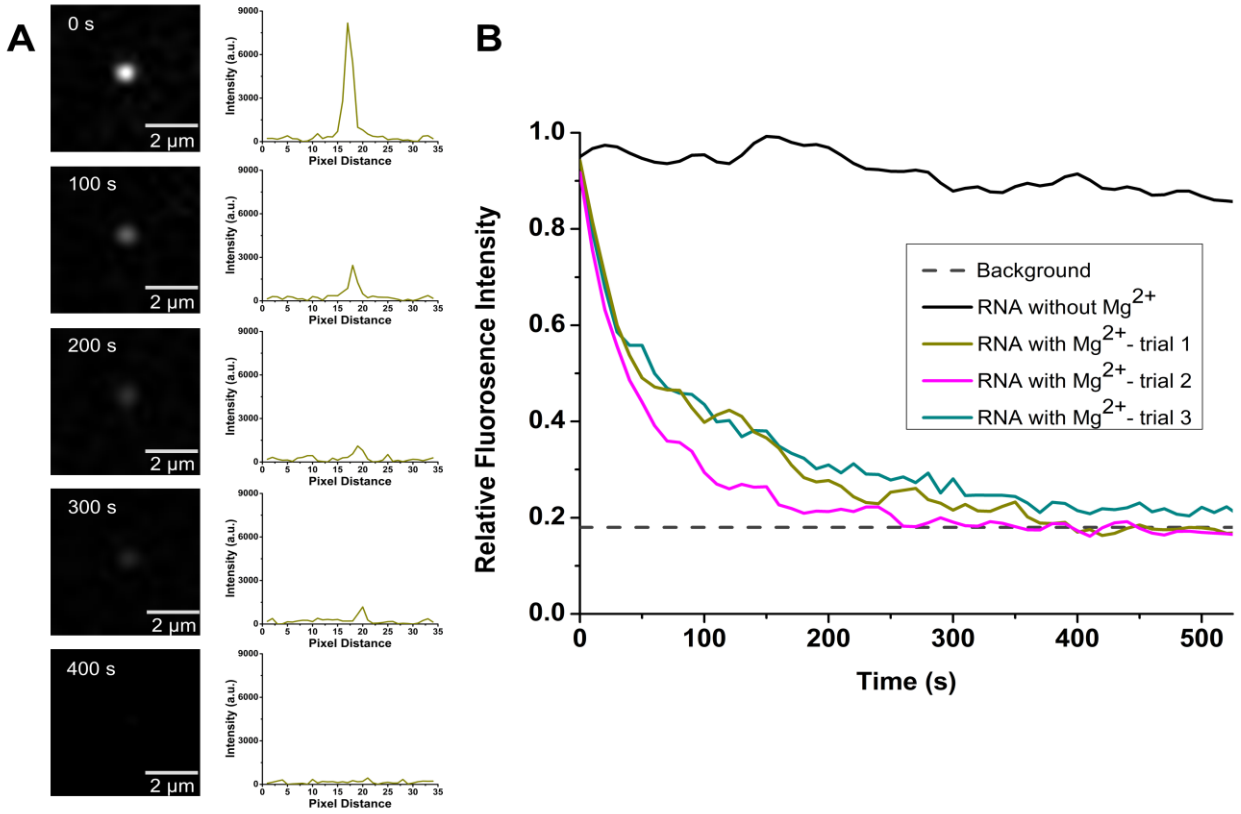


Figure 2.14. Digestion of SYTO 82 labelled DMD RNA by immobilized XRN1. **(A)** Fluorescence still images and corresponding intensity plot profiles of labelled DMD RNA-immobilized XRN1 complex acquired at different times, after introduction of Mg^{2+} to initiate digestion. **(B)** Relative fluorescence intensity of RNA-enzyme complexes with time. The black spectrum depicts the intensity of the complex in the absence of the cofactor Mg^{2+}. The dark cyan, dark yellow and magenta spectra illustrate the fluorescence intensity of the complexes when Mg^{2+} is introduced. The average fluorescence intensity becomes indistinguishable from the background intensity around 400 s.

The average clipping rate of XRN1 in the immobilized state was deduced using the decay portion of the fluorescence intensity shown in Figure 2.14.B**.** The total observable length of DMD RNA was calculated by subtracting the smallest detectable length from the length of DMD RNA. The clipping rate was then calculated by dividing the total observable length by the time the relative fluorescence intensity reached background (400 s). This yielded an average clipping rate of 26 ±5 nt s^{-1} (n = 3) for immobilized XRN1 at 25°C. We note that the clipping rate and the processivity of both free solution and immobilized XRN1 may not be optimal values as these experiments were conducted at room temperature (25°C) instead of the enzyme's optimum temperature of 37°C.

2.3.5 Monophosphorylation of RNA Transcripts

To demonstrate the ability to remove 5' cap structures found in mRNA, we treated CleanCap Fluc RNA (TriLink Biotechnologies, San Diego, CA, USA) and IVT 62mer, both of which contained a cap1 structure at their 5' end with mRNA decapping enzyme (MDE) prior to XRN1 digestion. The IVT 62mer was capped using "one-step capping and 2'-O-methylation protocol" that generates a cap1 structure at the 5' end of the 62mer RNA following the manufacturer's recommended protocol (see Figure 2.15.A; NEB, Ipswich, MA, USA). The 62mer RNA had two extra guanosine groups at the 5' end and the rest of the sequence was the same as the 60mer RNA (see Figure 2.6). The capped RNAs were then treated with MDE to remove the 5' cap1 structure (see Figure 2.15.B). MDE belongs to the Nudix family of pyrophosphohydrolases, which reacts with polyphosphate groups to make monophosphorylated RNA.[46-51] When MDE reacts with cap1 RNA, a m7G-pp- (m7GDP) group is removed and an intact RNA that is 2'-O-methylated (Nm) at the first nucleotide from the 5' end is generated (see Figure 2.15.B). MDE is reported as being highly efficient for removing both cap0 and cap1 structures.[52]

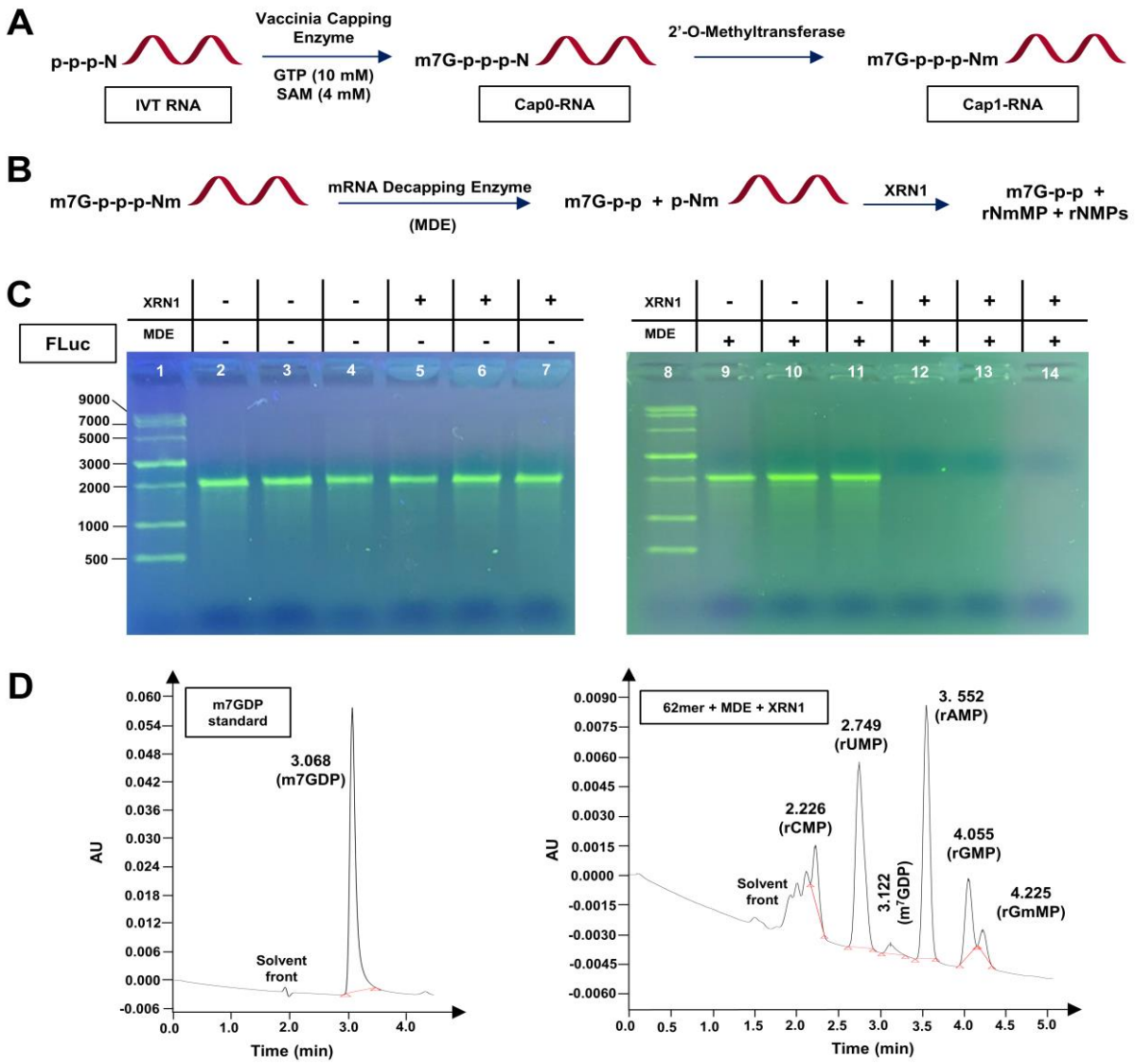


Figure 2.15. Decapping of 5' capped RNA. **(A)** Insertion of the 5' cap1 structure to the IVT RNA using "one-step capping and 2'-O-methylation protocol". **(B)** Cap1 removal using mRNA decapping enzyme (MDE) prior to XRN1 digestion. **(C)** Agarose gel analysis of capped and decapped FLuc RNA reactions with XRN1: lane **1,8** – RNA ladder; Lane **2-4** – Capped RNA without MDE and XRN1; Lane **5-7** – Caped RNA with XRN1; Lane **9-11** – Decapped RNA; Lane **12-14** – Decapped RNA with XRN1 **(D)** UPLC chromatograms of m7GDP standard solution and decapped 62mer digestion products. UV detection at 254 nm.

Next, the decapped RNA samples were purified as previously described and reacted with XRN1. After digestion, the RNA samples were subjected to agarose gel electrophoresis (see Figure 2.15.C). As can be seen from the gel images, the capped RNA remained intact when reacted with XRN1. When the RNA was decapped prior to XRN1 digestion, loss of full-length

intact RNA indicated that XRN1 was able to digest the RNA. Additionally, this experiment also demonstrated the ability of XRN1 to digest through the 2'-O-methylated RNA sequences as it was able to digest through the resulting 2'-O-methylated RNA sequences from the decapping reaction, which contained a 2'-O-methyladenosine (rAmMP) and 2'-O-methylguanine (rGmMP) for FLuc and 62mer, respectively.

For further confirmation, we analyzed the reaction products from the decapped 62mer RNA and XRN1 reaction using ultra-high-performance liquid chromatography (UPLC). UPLC was carried out using a Waters XBridge BEH C18 (2.5 µm, 4.6 x 150 mm) column and 100% (0.1% Formic acid/H_2O) mobile phase with a 1.00 mL/min flow rate. The reaction products contained the four canonical rNMPs, m7GDP and, rGmMP which confirmed successful decapping of the cap1-62mer RNA and subsequent digestion of the 2'-O-methylated RNA by XRN1 (Figure 2.15.D).

2.3.6 RNA Secondary Structures

All RNA substrates used in this report consisted of secondary structures at room temperature (25^0C; see Figure 2.16). The Minimum Free Energy (MFE) secondary structures were obtained using RNAfold webserver developed by the Institute for Theoretical Chemistry at the University of Vienna [53]. Due to the size restrictions of the server, only the first 10,000 nucleotides of the DMD RNA were input into the analyzer software. The MFEs of the most stable RNA secondary structures at room temperature were -1.99 kcal/mol for the 60 nt RNA, -574.76 kcal/mol for FLUC RNA, and -3581.78 kcal/mol for DMD RNA (10,000 nucleotides).

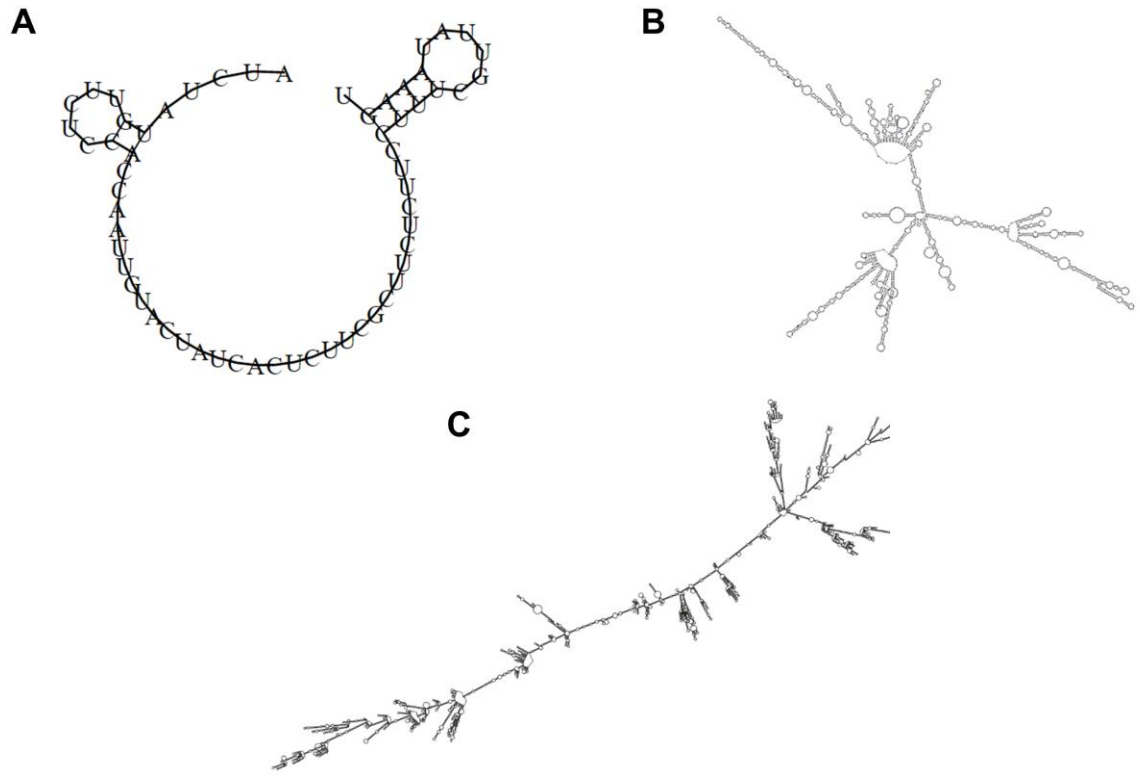


Figure 2.16. Minimum free energy (MFE) secondary structures of the **(A)** 60 b RNA, **(B)** FLUC RNA and **(C)** DMD RNA at room temperature. The MFE secondary structures were obtained using RNAfold webserver developed by Institute for Theoretical Chemistry, University of Vienna (see Reference [53]).

If XRN1 was unable to digest through secondary structures, there would be partially digested RNA molecules remaining. As an example, the 60 nt RNA first stem-loop structure is encountered around the 5th nucleotide from the 5' end, which would leave an RNA fragment of ~55 nucleotides in length if the activity of XRN1 was stalled due to secondary structures. The fact that the 60 nt RNA digested to its end confirms the ability of XRN1 to digest through secondary structures.

2.4 Discussion

RNA sequencing has become extremely important due to the recent COVID-19 pandemic. Reports using Illumina-based NGS have determined that this coronavirus, called SARS-CoV-2, has an approximate 30 kb ssRNA genome with a sequence different from that of the 2002 SARS coronavirus (~79% sequence homology) and the MERS virus (~50% sequence homology).[54-56]

Due to the evolutionary rate of RNA viruses (~10^{-4} nt substitutions per year), RNA sequencing will continue to be an important tool for not only detection, but vaccine discovery and determining resistance as well. Thus, new strategies for RNA sequencing that can provide simpler workflow, longer reads, and amplification-free formats would be particularly attractive. We have previously reported a single-molecule DNA sequencing strategy that fits the aforementioned operational criteria.[19-22] Briefly, it uses a processive exonuclease tethered to a solid-support with the cleaved nucleotides shuttled electrokinetically through a nanometer channel that measures the molecular-dependent electrophoretic mobility as an identifier. However, we are now envisioning this single-molecule sequencing strategy for RNA sequencing. As a foundation for that transition, we investigated the ability to use a solid-phase exoribonuclease reaction to sequentially generate ribonucleotides for identification using a label-free approach with high base identification accuracy via mobility matching and can identify modified ribonucleotides as well due to the lack of need for a PCR step in the workflow.

There are two categories of ribonucleases, endoribonucleases and exoribonucleases. Endoribonucleases cleave RNA internally whereas exoribonucleases cleave RNA sequentially from either the 3' or 5' end.[57] Exoribonucleases are further categorized into two types, distributive enzymes in which the RNA substrate is separated from the enzyme after each catalytic event and, processive enzymes where the RNA substrate is held by the enzyme until all of the nucleotides are cleaved from the intact substrate or the enzyme decomposes.[58] For identification of rNMPs or rNDPs using an exosequencing technique, sequential clipping of the nucleic acid substrate with high fidelity is critical.[13, 15, 18, 59]

The processivity is important in exosequencing as well because it is a factor that determines read length. Of the many exoribonucleases available only a few are processive.[60] Among these are the 5'→3' exoribonucleases (XRNs): XRN1 and XRN2 (Rat1) and 3'→5' exoribonucleases

RNase II, RNase R, and PNPase.[23-25, 60, 61] The 3'→5' exoribonucleases act on either 3'-OH or 3'-phosphate to produce rNMPs or rNDPs.[62-64] Although the 3'→5' exoribonucleases eliminate the need for prior sample treatment such as 5'-m7G decapping and dephosphorylation[65, 66], both RNase II and RNase R leave a residual oligonucleotide that is 3-5 nucleotides in length.[63] In the case of RNase II and PNPase, the enzyme activity is stalled by stable secondary structures in the RNA substrate.[64]

XRNs digest 5' monophosphorylated RNA[23-25, 61, 65, 66] to produce rNMPs as has been shown for XRN.[65] It has been reported that the activity of XRN2 is stalled when encountering secondary structure in the RNA substrate.[61] XRN2 can cleave through RNA strands with stem-loop structures only in the presence of the Rai1 protein.[61] Unlike XRN2, XRN1 can digest through RNA secondary structures due to the size of its active site and the mechanism of its action as demonstrated in several reports.[23-25] However, the narrow active site of XRN1 (~9 Å) does prevent access of double stranded RNA structures.[23-25] The 5' monophosphorylated ssRNA substrate is pulled through the narrow gap by a Brownian ratchet mechanism and this together with the steric barrier at the entrance causes duplex unwinding.[23-25] In addition to being processive, the ability to digest through secondary structures is a major advantage of using XRN1 in an exosequencing method, as it will eliminate the need for prior sample preparation to remove secondary structures in the RNA substrate.

As noted above, the model RNA substrates used in this chapter possessed significant secondary structures (see Figure 2.16). For example, the 60 nt RNA model was exhaustively digested by both the solid-phase and solution-phase XRN1 reactions (see Figure 2.7), in spite of the fact that this 60 nt RNA model possessed a stable secondary structure starting at the 5th nucleotide from the 5' end. If XRN1 was unable to cleave through this secondary structure, a 56-

nucleotide fragment would have appeared in the gel trace shown in Figure 2.7 following the reaction and this was not the case.

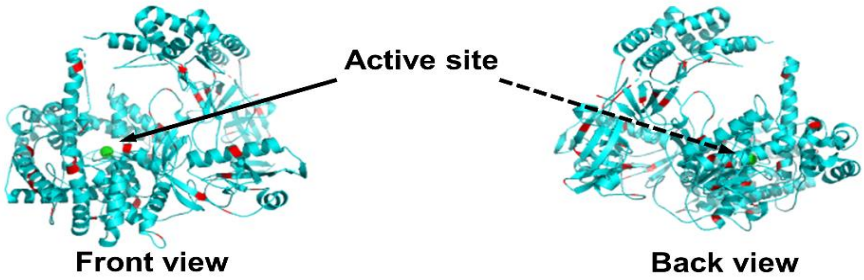


Figure 2.17. Front and back view of XRN1 with lysine groups highlighted in red. The lysine residues on the surface of the enzyme indicate potential attachment sites to PMMA surface. Structure of XRN1 was obtained from RCSB protein data bank and modified using PYMOL v2.1.1 software.

Moreover, the presence of 45 lysine groups on the surface of XRN1 as potential attachment sites onto a solid support possessing –COOH groups makes XRN1 an ideal exoribonuclease for exosequencing requiring a solid-phase reaction. For example, the UV/O_3 activation of many plastics, such as PMMA, creates surface-confined carboxylic acid groups that can be coupled to primary amines found on lysine groups using EDC/NHS coupling chemistry (see Figure 2.3.A). Figure 2.17 shows the primary structure of XRN1 with the lysine groups highlighted in red. Of the many sites available for covalent attachment to the carboxylated surface, only ~4 lysine residues reside in close proximity to the active site while most of the reactive lysine residues are located on the opposite side of the active site, which allows the efficient capture of RNA substrates by the immobilized enzyme without masking of the active site by the surface.

As previously reported, the -COOH group density of UV/O_3 activated polymers depend on the polymer type and the UV dose.[37] For enzyme immobilization onto PMMA, a lower surface density of -COOH groups is desirable to avoid multisite attachment of the enzyme to the surface, which could inadvertently lead to an inaccessible active site or denaturation of the enzyme.[67, 68] Therefore, immobilization of XRN1 that minimizes multi-site attachment is critical for efficient

accessibility to the active site and sequential digestion of an RNA substrate. The conditions used here for attachment of XRN1 onto PMMA (UV dose of 16.0 mW/cm^2 for 15 min) generates a -COOH group density of 32 pmol/cm^2 as determined by a TBO assay.[20] This is ~10 times higher than the highest surface concentration used here (see Table 2.3), which may result in multi-site attachment. However, the TBO assay, which was used to quantify the amount of -COOH groups on plastics, tends to label -COOH groups underneath the surface of the PMMA substrate that are not accessible by the enzyme.[20]

Immobilization of enzymes can, however, cause conformational changes altering performance.[69] Although in most cases these conformational changes lead to adverse effects in terms of enzyme activity, sometimes these conformational changes can lead to enhanced enzyme activity, stability, and specificity.[31] Immobilization can also increase the rigidity of the enzyme, which can result in higher stability.[70, 71] Previously we observed that immobilized λ-Exo, which is comprised of 3 subunits, demonstrated a processivity much higher than solution-phase reactions most likely arising from enzyme stabilization.[20] Proper immobilization of the enzyme to the solid support will stabilize the enzyme and maintain its structure in turn increasing the enzyme processivity and even the clipping rate.[31, 70] It is also reported in some cases that increasing the rigidity of an enzyme can reduce allosteric inhibition leading to higher enzyme activity.[72]

As noted from our data shown in Figures 4 and 5, the clipping rates of XRN1 were 3.06 ±0.11 nt s^{-1} and 26 ±5 nt s^{-1} at 25^{o}C for solution-phase and solid-phase reactions, respectively. Langeberg *et al.* recently reported the free solution clipping rate of XRN1 was 17.3 ±0.6 nt s^{-1}.[27] This value was ~6-fold higher compared to the value we obtained and is most likely due to differences in enzyme reaction temperatures (37^{o}C vs. 25^{o}C). We should also note that solution-phase (Figure 2.11) and the solid-phase (Figure 2.14) clipping rates reported in this manuscript were secured using bulk and single-molecule measurements and different templates of RNA,

which could have contributed to the observed differences. At any rate, because the immobilized XRN1 enzyme can function in a fashion similar to the solution enzymatic reaction supports our supposition that XRN1 can potentially be used for single-molecule exosequencing requiring an immobilized enzyme.

More than 170 post-transcriptional RNA modifications have been identified to date.[73] Most of these modifications occur in abundant non-coding RNAs including ribosomal RNA (rRNA) and transfer RNA (tRNA) and play an important role in structural folding and function. There are 6 nucleotides with base modifications that can influence the metabolism and function of messenger RNA (mRNA): m6A, m5C, inosine (I), pseudouridine (ψ), N^1-methyladenosine (m1A), and 5-hydroxylmethylcytidine (5-hmC). Additional functions of RNA modifications include tRNA stability, cellular stress response (m5C), and microRNA stability (2'-O-methlyation). The "epitranscriptome", which is the term used to describe RNA modifications throughout the transcriptome[74], has historically been difficult to study due to a lack of tools for deciphering the presence of the entire inventory of RNA modifications. However, technological advancements such as NGS have led to a greater understanding of the epitranscriptome and its role in normal biology and disease.[75] Recently, Koh *et al.* demonstrated the use of XRN1 to map exact locations of m6A modifications within an RNA molecule.[76] This technique, coined m6A-Crosslinking-Exonuclease-Sequencing (m6ACE-seq), involves anti-m6A antibodies photo-crosslinked onto m6A and 2'-O-methylated m6A (m6Am) modifications contained within an RNA molecule. The antibody crosslinking halts the digestion of RNA by XRN1 at the exact position of m6A/m6Am thereby indicating the position of these methylations.

For the RNA nucleotide modifications we tested, we found that XRN1 could cleave through the m5C and m6A modified nucleotides. This is in agreement with the literature, which has reported that XRN1 is involved in degradation of m6A and m5C containing RNA species.[77-81] Both

m6A and m5C were chosen as RNA models for this initial work based on literature precedence indicating the high abundance of these modifications throughout the transcriptome, which has been corroborated using next-generation RNA-sequencing.[82-84]

Our conclusion of XRN1's ability to cleave through m5C and m6A nucleotides was supported by three lines of independent evidence: (1) Denaturing gel electrophoresis, which showed the disappearance of bands near the intact 60mer RNA model (see Figure 2.7). (2) Near complete loss of fluorescence specific to intact RNA molecules (see Figure 2.8). (3) UPLC/MS data obtained that indicated the molar ratios of the ribonucleotides was consistent with complete digestion with the structures of the methylated ribonucleotides preserved after digestion (see Figure 2.9).

It has also been reported that XRN1 plays a major role in rapid tRNA decay (RTD), which digests fully modified mature tRNA species *in vivo.*[77, 85-87] In one study, Whipple *et. al.* demonstrated the ability of XRN1 to digest through fully modified mature tRNAs using 3' cytosine-3',5'-bisphosphate (pCp) labelled $tRNA^{Ser(CGA)}$ variants.[87] The wild type $tRNA^{Ser(CGA)}$ contained a 2'-O-methylated guanosine and uridine at the 18th and 44th ribonucleotide positions from 5' end, respectively. When the tRNA variants were reacted with XRN1 the full length $tRNA^{Ser(CGA)}$ was degraded to completion to give pCp as the end product. Additionally, the digestion of $tRNA^{Ser(CGA)}$ and $tRNA^{Val(AAC)}$ provided information on the ability of XRN1 to digest through other modified nucleotides such as N^1-methylguanosine (m1G), ψ, dihydrouridine (D), m1A, 2,2,7-trimethylguanosine (m2,2,7G), and i6A.[77, 87] Thus, our single-molecule exosequencing strategy may possess the ability to identify these and other epitranscriptomic modifications without the need for antibody crosslinking or PCR amplification.

All eukaryotic mRNAs contain a cap structure at their 5' end to enhance their stability by inhibiting the digestion ability of 5'-3' exoribonucleases.[48, 88-90] The most common cap structure found in mRNA includes the 5'-5' triphosphate linked 7-methylguanosine (m7G) cap, known as cap0, which is sometimes further modified by 2'-O-methylation at the 1st nucleotide, cap1, and 2nd nucleotide, cap2.[51, 88, 90] Although less abundant, other cap structures such as nicotinamide adenine dinucleotide (NAD), m2,2,7G, and pyrophosphate groups are also found in mRNA.[91-93] In our proposed exosequencing method, prior to XRN1 digestion the mRNA will need to be decapped to produce viable XRN1 substrates. We have successfully shown that RppH can remove the pyrophosphate group of IVT RNA products (see Figure 2.2.C). RppH can also be used for removing the cap0 structure of mRNA.[93, 94] In addition, Frindert *et al.* has demonstrated that *Bacillus subtilis* RppH (BsRppH), a homolog of RppH, can also be used to successfully remove the NAD cap structure *in vitro* resulting in 5' monophosphorylated RNA for subsequent digestion by 5'-3' exoribonucleases.[91] Moreover, *Schizosaccharomyces pombe* Edc1-fused Dcp1–Dcp2 decapping enzyme, also known as "mRNA decapping enzyme (MDE)", can be used to decap mRNA *in vitro* to remove both cap0 and cap1 structures.[52] We successfully used MDE for decapping RNA prior to XRN1 digestion (see Figure 2.15). The subsequent digestion of the decapped RNA by XRN1 further confirmed the ability of XRN1 to digest through 2'-O-methylated RNA as it was able to digest through the resulting 2'-O-methylated RNA following the decapping reaction (see Figure 2.15.C). UPLC analysis of the digestion products of the decapped 62mer RNA further confirmed successful decapping and digestion of 2'-O-methylated RNA (see Figure 2.15.D). In addition to removing cap0 and cap1 structures, MDE can also be used for decapping of m2,2,7G cap and also convert 5' pyrophosphate ends to 5' monophosphorylated ends.[92, 95] While the decapping of mRNA adds additional pre-processing steps, the decapping is simple using either RppH or MDE and requires only an incubation step followed by a quick RNA clean-up step for direct XRN1 RNA digestion.

In addition to using the solid-phase XRN1 reaction for single-molecule sequencing, solid-phase reactions can be used for IMER-based applications, which provide many advantages compared to solution-phase reactors, such as low sample consumption, high-throughput processing, and prevention of autodigestion of the enzyme.[30, 31, 96] But, they also have limitations arising from poor enzyme kinetics due to mass transfer limitations.[97, 98] For a successful enzymatic reaction to occur, the substrate must diffuse to the immobilized enzyme. The use of a pillared IMER reduces this problem as it increases the surface-to-volume ratio compared to an open channel IMER as well as reducing diffusional distances.[98] The IMER used in this chapter consisted of a microchannel with 3,600 micropillars that were 100 µm in diameter and 60 µm in height with the inter-pillar spacing being 35 µm. The use of micropillars in this case increased the surface-to-volume ratio by 73% compared to an open channel IMER with the same dimensions. The higher digestion efficiency achieved in this report for the immobilized enzyme (87.6 ±2.8%) compared to the free solution digestion (78.3 ±4.4%; $p = 0.0187$) was possibly due to the higher surface area-to-volume ratio coupled with several other factors such as increased stability, and prevention of enzyme aggregation.[31, 99-101] The micropillared IMER also reduced diffusional distances increasing the number of interactions between the solution-borne RNA substrates and immobilized XRN1 enzymes. Furthermore, because single enzyme molecules are attached to the substrate most likely through a lysine residue that makes the active site accessible (see Figure 2.17), the RNA substrate is able to diffuse into the active site, whereas in free solution aggregation of enzyme molecules can make the active site inaccessible.[31, 102]

2.5 Conclusion

In this chapter, we have demonstrated for the first time the covalent attachment of XRN1 onto a solid-support for potential applications in single-molecule RNA exosequencing. The immobilization of XRN1 was carried out using established EDC/NHS coupling chemistry to a

plastic support that was activated using UV/O_3 light to create surfaces containing -COOH groups. The immobilized XRN1 exhibited a higher digestion efficiency compared to free solution XRN1. The ability of XRN1 to digest through methylated sequences, demonstrated in this chapter for both free solution and the immobilized state using m6A and m5C methylated RNA sequences, is particularly advantageous for sequencing of RNA, which could eliminate the need for antibodies and bisulfite treatment used in current NGS sequencing methods.[76, 103] Furthermore, we have reported the clipping rate and the processivity of both free solution and immobilized XRN1. Immobilized XRN1 demonstrated an average clipping rate that is 9-fold higher compared to the solution phase clipping rate and an apparent processivity of >10.5 kb. However, the clipping rates were secured at room temperature and not at 37˚C, which is the manufacturer's suggested conditions. Thus, the clipping rates reported herein would be expected to increase when the XRN1 reactions are performed at the optimal temperature. In addition, the reported processivity is a lower limit for immobilized XRN1 because the number of nucleotides associated with the model (DMD) was completely digested. Therefore, it may be possible to read through the entire SARS-CoV-2 genome (~30 kb) using our proposed exosequencing approach but will depend on the upper limit of XRN1's processivity. These findings offer important information for considering the use of XRN1 in exosequencing techniques that use nanopores.

2.6 References

1. Byron, S. A.; Van Keuren-Jensen, K. R.; Engelthaler, D. M.; Carpten, J. D.; Craig, D. W., Translating RNA sequencing into clinical diagnostics: opportunities and challenges. *Nature Reviews Genetics* **2016,** *17* (5), 257.

2. Ozsolak, F.; Milos, P. M., RNA sequencing: advances, challenges and opportunities. *Nature reviews genetics* **2011,** *12* (2), 87.

3. Chandola, U.; Das, R.; Panda, B., Role of the N6-methyladenosine RNA mark in gene regulation and its implications on development and disease. *Briefings in functional genomics* **2014,** *14* (3), 169-179.

4. Kanwal, R.; Gupta, S., Epigenetic modifications in cancer. *Clinical genetics* **2012,** *81* (4), 303-311.

5. Motorin, Y.; Lyko, F.; Helm, M., 5-methylcytosine in RNA: detection, enzymatic formation and biological functions. *Nucleic acids research* **2009,** *38* (5), 1415-1430.

6. Yue, Y.; Liu, J.; He, C., RNA N6-methyladenosine methylation in post-transcriptional gene expression regulation. *Genes & development* **2015,** *29* (13), 1343-1355.

7. Corney, D. C., RNA-seq using next generation sequencing. *Materials and Methods* **2013,** *3*, 203.

8. Schaefer, M.; Pollex, T.; Hanna, K.; Lyko, F., RNA cytosine methylation analysis by bisulfite sequencing. *Nucleic acids research* **2008,** *37* (2), e12-e12.

9. Ozsolak, F.; Platt, A. R.; Jones, D. R.; Reifenberger, J. G.; Sass, L. E.; McInerney, P.; Thompson, J. F.; Bowers, J.; Jarosz, M.; Milos, P. M., Direct RNA sequencing. *Nature* **2009,** *461* (7265), 814.

10. Garalde, D. R.; Snell, E. A.; Jachimowicz, D.; Sipos, B.; Lloyd, J. H.; Bruce, M.; Pantic, N.; Admassu, T.; James, P.; Warland, A., Highly parallel direct RNA sequencing on an array of nanopores. *Nature methods* **2018,** *15* (3), 201.

11. Branton, D.; Deamer, D. W.; Marziali, A.; Bayley, H.; Benner, S. A.; Butler, T.; Di Ventra, M.; Garaj, S.; Hibbs, A.; Huang, X., The potential and challenges of nanopore sequencing. In *Nanoscience And Technology: A Collection of Reviews from Nature Journals*, World Scientific: 2010; pp 261-268.

12. Deamer, D. W.; Branton, D., Characterization of nucleic acids by nanopore analysis. *Accounts of chemical research* **2002,** *35* (10), 817-825.

13. Cockroft, S. L.; Chu, J.; Amorin, M.; Ghadiri, M. R., A single-molecule nanopore device detects DNA polymerase activity with single-nucleotide resolution. *Journal of the American Chemical Society* **2008,** *130* (3), 818-820.

14. Lieberman, K. R.; Cherf, G. M.; Doody, M. J.; Olasagasti, F.; Kolodji, Y.; Akeson, M., Processive replication of single DNA molecules in a nanopore catalyzed by phi29 DNA polymerase. *Journal of the American Chemical Society* **2010,** *132* (50), 17961-17972.

15. Astier, Y.; Braha, O.; Bayley, H., Toward single molecule DNA sequencing: direct identification of ribonucleoside and deoxyribonucleoside 5 '-monophosphates by using an engineered protein nanopore equipped with a molecular adapter. *Journal of the American Chemical Society* **2006,** *128* (5), 1705-1710.

16. Smith, A. M.; Abu-Shumays, R.; Akeson, M.; Bernick, D. L., Capture, unfolding, and detection of individual tRNA molecules using a nanopore device. *Frontiers in bioengineering and biotechnology* **2015,** *3*, 91.

17. Ayub, M.; Bayley, H., Individual RNA base recognition in immobilized oligonucleotides using a protein nanopore. *Nano letters* **2012,** *12* (11), 5637-5643.

18. Ayub, M.; Hardwick, S. W.; Luisi, B. F.; Bayley, H., Nanopore-based identification of individual nucleotides for direct RNA sequencing. *Nano letters* **2013,** *13* (12), 6144-6150.

19. Novak, B. R.; Moldovan, D.; Nikitopoulos, D. E.; Soper, S. A., Distinguishing Single DNA Nucleotides Based on Their Times of Flight Through Nanoslits: A Molecular Dynamics Simulation Study. *Journal of Physical Chemistry B* **2013,** *117* (12), 3271-3279.

20. Oliver-Calixte, N. J.; Uba, F. I.; Battle, K. N.; Weerakoon-Ratnayake, K. M.; Soper, S. A., Immobilization of lambda exonuclease onto polymer micropillar arrays for the solid-phase digestion of dsDNAs. *Analytical chemistry* **2014,** *86* (9), 4447-4454.

21. O'Neil, C.; Amarasekara, C. A.; Weerakoon-Ratnayake, K. M.; Gross, B.; Jia, Z.; Singh, V.; Park, S.; Soper, S. A., Electrokinetic transport properties of deoxynucleotide monophosphates (dNMPs) through thermoplastic nanochannels. *Analytica chimica acta* **2018,** *1027*, 67-75.

22. Xia, K.; Novak, B. R.; Weerakoon-Ratnayake, K. M.; Soper, S. A.; Nikitopoulos, D. E.; Moldovan, D., Electrophoretic Transport of Single DNA Nucleotides through Nanoslits: A Molecular Dynamics Simulation Study. *Journal of Physical Chemistry B* **2015,** *119*, 11443-11448.

23. Jinek, M.; Coyle, S. M.; Doudna, J. A., Coupled 5′ nucleotide recognition and processivity in Xrn1-mediated mRNA decay. *Molecular cell* **2011,** *41* (5), 600-608.

24. Jones, C. I.; Zabolotskaya, M. V.; Newbury, S. F., The 5′→ 3′ exoribonuclease XRN1/Pacman and its functions in cellular processes and development. *Wiley Interdisciplinary Reviews: RNA* **2012,** *3* (4), 455-468.

25. Nagarajan, V. K.; Jones, C. I.; Newbury, S. F.; Green, P. J., XRN 5′→ 3′ exoribonucleases: structure, mechanisms and functions. *Biochimica et Biophysica Acta (BBA)-Gene Regulatory Mechanisms* **2013,** *1829* (6-7), 590-603.

26. Hoek, T. A.; Khuperkar, D.; Lindeboom, R. G.; Sonneveld, S.; Verhagen, B. M.; Boersma, S.; Vermeulen, M.; Tanenbaum, M. E., Single-molecule imaging uncovers rules governing nonsense-mediated mRNA decay. *Molecular cell* **2019,** *75* (2), 324-339. e11.

27. Langeberg, C. J.; Welch, W. R.; McGuire, J. V.; Ashby, A.; Jackson, A. D.; Chapman, E. G., Biochemical Characterization of Yeast Xrn1. *Biochemistry* **2020,** *59* (15), 1493-1507.

28. Křenková, J.; Foret, F., Immobilized microfluidic enzymatic reactors. *Electrophoresis* **2004,** *25* (21-22), 3550-3563.

29. Matosevic, S.; Szita, N.; Baganz, F., Fundamentals and applications of immobilized microfluidic enzymatic reactors. *Journal of Chemical Technology & Biotechnology* **2011,** *86* (3), 325-334.

30. Homaei, A. A.; Sariri, R.; Vianello, F.; Stevanato, R., Enzyme immobilization: an update. *Journal of chemical biology* **2013,** *6* (4), 185-205.

31. Rodrigues, R. C.; Ortiz, C.; Berenguer-Murcia, Á.; Torres, R.; Fernández-Lafuente, R., Modifying enzyme activity and selectivity by immobilization. *Chemical Society Reviews* **2013,** *42* (15), 6290-6307.

32. Chantiwas, R.; Park, S.; Soper, S. A.; Kim, B. C.; Takayama, S.; Sunkara, V.; Hwang, H.; Cho, Y.-K., Flexible fabrication and applications of polymer nanochannels and nanoslits. *Chemical Society Reviews* **2011,** *40* (7), 3677-3702.

33. Mecomber, J. S.; Murthy, R. S.; Rajam, S.; Singh, P. N.; Gudmundsdottir, A. D.; Limbach, P. A., Photochemical functionalization of polymer surfaces for microfabricated devices. *Langmuir* **2008,** *24* (7), 3645-3653.

34. Qu, H.; Wang, H.; Huang, Y.; Zhong, W.; Lu, H.; Kong, J.; Yang, P.; Liu, B., Stable microstructured network for protein patterning on a plastic microfluidic channel: strategy and characterization of on-chip enzyme microreactors. *Analytical chemistry* **2004,** *76* (21), 6426-6433.

35. Dominick, W. D.; Berhane, B. T.; Mecomber, J. S.; Limbach, P. A., Covalent immobilization of proteases and nucleases to poly (methylmethacrylate). *Analytical and bioanalytical chemistry* **2003,** *376* (3), 349-354.

36. Wei, S.; Vaidya, B.; Patel, A. B.; Soper, S. A.; McCarley, R. L., Photochemically patterned poly (methyl methacrylate) surfaces used in the fabrication of microanalytical devices. *The Journal of Physical Chemistry B* **2005,** *109* (35), 16988-16996.

37. Jackson, J. M.; Witek, M. A.; Hupert, M. L.; Brady, C.; Pullagurla, S.; Kamande, J.; Aufforth, R. D.; Tignanelli, C. J.; Torphy, R. J.; Yeh, J. J., UV activation of polymeric high aspect ratio microstructures: ramifications in antibody surface loading for circulating tumor cell selection. *Lab on a Chip* **2014,** *14* (1), 106-117.

38. Lim, C. Y.; Owens, N. A.; Wampler, R. D.; Ying, Y.; Granger, J. H.; Porter, M. D.; Takahashi, M.; Shimazu, K., Succinimidyl ester surface chemistry: implications of the competition between aminolysis and hydrolysis on covalent protein immobilization. *Langmuir* **2014,** *30* (43), 12868-12878.

39. Wlodkowic, D.; Skommer, J.; Darzynkiewicz, Z., SYTO probes in the cytometry of tumor cell death. *Cytometry Part A: The Journal of the International Society for Analytical Cytology* **2008,** *73* (6), 496-507.

40. Jones, L. J.; Yue, S. T.; Cheung, C.-Y.; Singer, V. L., RNA quantitation by fluorescence-based solution assay: RiboGreen reagent characterization. *Analytical biochemistry* **1998,** *265* (2), 368-374.

41. Han, E. S.; Cooper, D. L.; Persky, N. S.; Sutera Jr, V. A.; Whitaker, R. D.; Montello, M. L.; Lovett, S. T., RecJ exonuclease: substrates, products and interaction with SSB. *Nucleic acids research* **2006,** *34* (4), 1084-1091.

42. Lubarsky, G.; Davidson, M.; Bradley, R., Hydration–dehydration of adsorbed protein films studied by AFM and QCM-D. *Biosensors and Bioelectronics* **2007,** *22* (7), 1275-1281.

43. Nnebe, I. M.; Schneider, J. W., A tapping-mode AFM study of the compression of grafted poly (ethylene glycol) chains. *Macromolecules* **2006,** *39* (10), 3616-3621.

44. Antharavally, B. S.; Mallia, K. A.; Rangaraj, P.; Haney, P.; Bell, P. A., Quantitation of proteins using a dye–metal-based colorimetric protein assay. *Analytical biochemistry* **2009,** *385* (2), 342-345.

45. Kino, K.; Hirao-Suzuki, M.; Morikawa, M.; Sakaga, A.; Miyazawa, H., Generation, repair and replication of guanine oxidation products. *Genes and Environment* **2017,** *39* (1), 21.

46. Decker, C. J.; Parker, R., mRNA decay enzymes: decappers conserved between yeast and mammals. *Proceedings of the National Academy of Sciences* **2002,** *99* (20), 12512-12514.

47. Deshmukh, M. V.; Jones, B. N.; Quang-Dang, D.-U.; Flinders, J.; Floor, S. N.; Kim, C.; Jemielity, J.; Kalek, M.; Darzynkiewicz, E.; Gross, J. D., mRNA decapping is promoted by an RNA-binding channel in Dcp2. *Molecular cell* **2008,** *29* (3), 324-336.

48. Grudzien-Nogalska, E.; Kiledjian, M., New insights into decapping enzymes and selective mRNA decay. *Wiley Interdisciplinary Reviews: RNA* **2017,** *8* (1), e1379.

49. Li, Y.; Kiledjian, M., Regulation of mRNA decapping. *Wiley Interdisciplinary Reviews: RNA* **2010,** *1* (2), 253-265.

50. Mildvan, A.; Xia, Z.; Azurmendi, H.; Saraswat, V.; Legler, P.; Massiah, M.; Gabelli, S.; Bianchet, M.; Kang, L.-W.; Amzel, L., Structures and mechanisms of Nudix hydrolases. *Archives of biochemistry and biophysics* **2005,** *433* (1), 129-143.

51. Wulf, M. G.; Buswell, J.; Chan, S.-H.; Dai, N.; Marks, K.; Martin, E. R.; Tzertzinis, G.; Whipple, J. M.; Corrêa, I. R.; Schildkraut, I., The yeast scavenger decapping enzyme DcpS and its application for in vitro RNA recapping. *Scientific reports* **2019,** *9* (1), 1-9.

52. Paquette, D. R.; Mugridge, J. S.; Weinberg, D. E.; Gross, J. D., Application of a Schizosaccharomyces pombe Edc1-fused Dcp1–Dcp2 decapping enzyme for transcription start site mapping. *RNA* **2018,** *24* (2), 251-257.

53. Lorenz, R.; Bernhart, S. H.; Zu Siederdissen, C. H.; Tafer, H.; Flamm, C.; Stadler, P. F.; Hofacker, I. L., ViennaRNA Package 2.0. *Algorithms for molecular biology* **2011,** *6* (1), 26.

54. Sah, R.; Rodriguez-Morales, A. J.; Jha, R.; Chu, D. K. W.; Gu, H.; Peiris, M.; Bastol, A.; Lal, B. K.; Ojha, H. C.; Rabaan, A. A.; Zambrano, L. I.; Costello, A.; Morita, K.; Pandey, B. D.; Poon, L. L. M., Complete Genome Sequence of a 2019 Novel Coronavirus (SARS-CoV-2) Strain Isolated in Nepal. *Genome Sequences* **2019,** *9* (11), 1-3.

55. Kim, J.-M.; Chung, Y.-S.; Jo, H. J.; Lee, N.-J.; Kim, M. S.; Woo, S. H.; Park, S.; Kim, J. W.; Kim, H. M.; Han, M.-G., Identification of Coronavirus Isolated from a Patient in Korea with COVID-19. *Osong Public Health and Research Perspectives* **20,** *11* (1), 3-7.

56. Lu, R.; Zhao, X.; Li, J.; Niu, P.; Yang, B.; Wu, H.; Wang, W.; Song, H.; Huang, B.; Zhu, N.; Bi, Y.; Ma, X.; Zhan, F.; Wang, L.; Hu, T.; Zhou, H.; Hu, Z.; Zhou, W.; Zhao, L.; Chen, J.; Meng, Y.; Wang, J.; Lin, Y.; Yuan, J.; Xie, Z.; Ma, J.; Liu, W. J.; Wang, D.; Xu, W.; Holmes, E. C.; Gao, G. F.; Wu, G.; Chen, W.; Shi, W.; Tan, W., Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding. *The Lancet* **2020,** *395*, 565-574.

57. Andrade, J. M.; Pobre, V.; Silva, I. J.; Domingues, S.; Arraiano, C. M., The role of 3′–5′ exoribonucleases in RNA degradation. *Progress in molecular biology and translational science* **2009,** *85*, 187-229.

58. Thomas, K. R.; Olivera, B. M., Processivity of DNA exonucleases. *Journal of Biological Chemistry* **1978,** *253* (2), 424-429.

59. Clarke, J.; Wu, H.-C.; Jayasinghe, L.; Patel, A.; Reid, S.; Bayley, H., Continuous base identification for single-molecule nanopore DNA sequencing. *Nature nanotechnology* **2009,** *4* (4), 265-270.

60. Zuo, Y.; Deutscher, M. P., Exoribonuclease superfamilies: structural analysis and phylogenetic distribution. *Nucleic acids research* **2001,** *29* (5), 1017-1026.

61. Xiang, S.; Cooper-Morgan, A.; Jiao, X.; Kiledjian, M.; Manley, J. L.; Tong, L., Structure and function of the 5′→ 3′ exoribonuclease Rat1 and its activating partner Rai1. *Nature* **2009,** *458* (7239), 784-788.

62. Amblar, M.; Barbas, A.; Fialho, A. M.; Arraiano, C. M., Characterization of the functional domains of Escherichia coli RNase II. *Journal of molecular biology* **2006,** *360* (5), 921-933.

63. Cheng, Z.-F.; Deutscher, M. P., Purification and characterization of the Escherichia coli exoribonuclease RNase R Comparison with RNase II. *Journal of Biological Chemistry* **2002,** *277* (24), 21624-21629.

64. Spickler, C.; Mackie, G. A., Action of RNase II and polynucleotide phosphorylase against RNAs containing stem-loops of defined structure. *Journal of bacteriology* **2000,** *182* (9), 2422-2427.

65. Stevens, A., Purification and characterization of a Saccharomyces cerevisiae exoribonuclease which yields 5'-mononucleotides by a 5'leads to 3'mode of hydrolysis. *Journal of Biological Chemistry* **1980,** *255* (7), 3080-3085.

66. Stevens, A., 5′-Exoribonuclease 1: Xrn1. In *Methods in enzymology*, Elsevier: 2001; Vol. 342, pp 251-259.

67. Ye, P.; Wan, R.-B.; Wang, X.-P., Quantitative enzyme immobilization: Control of the carboxyl group density on support surface. *Journal of Molecular Catalysis B: Enzymatic* **2009,** *61* (3-4), 296-302.

68. Zabriskie, D.; Ollis, D. F.; Burger, M. M., Activity and specificity of covalently immobolized wheat germ agglutinin toward cell surfaces. *Biotechnology and Bioengineering* **1973,** *15* (5), 981-992.

69. Hanefeld, U.; Gardossi, L.; Magner, E., Understanding enzyme immobilisation. *Chemical Society Reviews* **2009,** *38* (2), 453-468.

70. Polizzi, K. M.; Bommarius, A. S.; Broering, J. M.; Chaparro-Riggers, J. F., Stability of biocatalysts. *Current opinion in chemical biology* **2007,** *11* (2), 220-225.

71. Mateo, C.; Palomo, J. M.; Fuentes, M.; Betancor, L.; Grazu, V.; López-Gallego, F.; Pessela, B. C.; Hidalgo, A.; Fernández-Lorente, G.; Fernández-Lafuente, R., Glyoxyl agarose: a fully inert and hydrophilic support for immobilization and high stabilization of proteins. *Enzyme and Microbial Technology* **2006,** *39* (2), 274-280.

72. Fernandez-Lafuente, R.; Rosell, C. M.; Guisan, J. M., The use of stabilised penicillin acylase derivatives improves the design of kinetically controlled synthesis. *Journal of Molecular Catalysis A: Chemical* **1995,** *101* (1), 91-97.

73. Frye, M.; Harada, B. T.; Behm, M.; He, C., RNA modifications modulate gene expression during development. *Science* **2018,** *361* (6409), 1346-1349.

74. Saletore, Y.; Meyer, K.; Korlach, J.; Vilfan, I. D.; Jaffrey, S.; Mason, C. E., The birth of the Epitranscriptome: deciphering the function of RNA modifications. *Genome biology* **2012,** *13* (10), 175.

75. Xuan, J.-J.; Sun, W.-J.; Lin, P.-H.; Zhou, K.-R.; Liu, S.; Zheng, L.-L.; Qu, L.-H.; Yang, a. J.-H., RMBase v2.0: deciphering the map of RNA modifications from epitranscriptome sequencing data. *Nucleic Acids Research* **2017,** *46*, D327-D334.

76. Koh, C. W.; Goh, Y. T.; Goh, W. S., Atlas of quantitative single-base-resolution N 6-methyl-adenine methylomes. *Nature communications* **2019,** *10* (1), 1-15.

77. Chernyakov, I.; Whipple, J. M.; Kotelawala, L.; Grayhack, E. J.; Phizicky, E. M., Degradation of several hypomodified mature tRNA species in Saccharomyces cerevisiae is mediated by Met22 and the 5′–3′ exonucleases Rat1 and Xrn1. *Genes & development* **2008,** *22* (10), 1369-1380.

78. Coker, H.; Wei, G.; Brockdorff, N., m6A modification of non-coding RNA and the control of mammalian gene expression. *Biochimica et Biophysica Acta (BBA)-Gene Regulatory Mechanisms* **2019,** *1862* (3), 310-318.

79. Kretschmer, J.; Rao, H.; Hackert, P.; Sloan, K. E.; Höbartner, C.; Bohnsack, M. T., The m6A reader protein YTHDC2 interacts with the small ribosomal subunit and the 5′–3′ exoribonuclease XRN1. *Rna* **2018,** *24* (10), 1339-1350.

80. Lee, M.; Kim, B.; Kim, V. N., Emerging roles of RNA modification: m6A and U-tail. *Cell* **2014,** *158* (5), 980-987.

81. Wojtas, M. N.; Pandey, R. R.; Mendel, M.; Homolka, D.; Sachidanandam, R.; Pillai, R. S., Regulation of m6A transcripts by the 3'→ 5' RNA helicase YTHDC2 is essential for a successful meiotic program in the mammalian germline. *Molecular cell* **2017,** *68* (2), 374-387. e12.

82. Dominissini, D.; Moshitch-Moshkovitz, S.; Schwartz, S.; Salmon-Divon, M.; Ungar, L.; Osenberg, S.; Cesarkas, K.; Jacob-Hirsch, J.; Amariglio, N.; Kupiec, M., Topology of the human and mouse m 6 A RNA methylomes revealed by m 6 A-seq. *Nature* **2012,** *485* (7397), 201-206.

83. Squires, J. E.; Patel, H. R.; Nousch, M.; Sibbritt, T.; Humphreys, D. T.; Parker, B. J.; Suter, C. M.; Preiss, T., Widespread occurrence of 5-methylcytosine in human coding and non-coding RNA. *Nucleic acids research* **2012,** *40* (11), 5023-5033.

84. Meyer, K. D.; Saletore, Y.; Zumbo, P.; Elemento, O.; Mason, C. E.; Jaffrey, S. R., Comprehensive analysis of mRNA methylation reveals enrichment in 3′ UTRs and near stop codons. *Cell* **2012,** *149* (7), 1635-1646.

85. Dewe, J. M.; Whipple, J. M.; Chernyakov, I.; Jaramillo, L. N.; Phizicky, E. M., The yeast rapid tRNA decay pathway competes with elongation factor 1A for substrate tRNAs and acts on tRNAs lacking one or more of several modifications. *Rna* **2012,** *18* (10), 1886-1896.

86. Watanabe, K.; Miyagawa, R.; Tomikawa, C.; Mizuno, R.; Takahashi, A.; Hori, H.; Ijiri, K., Degradation of initiator tRNA Met by Xrn1/2 via its accumulation in the nucleus of heat-treated HeLa cells. *Nucleic acids research* **2013,** *41* (8), 4671-4685.

87. Whipple, J. M.; Lane, E. A.; Chernyakov, I.; D'Silva, S.; Phizicky, E. M., The yeast rapid tRNA decay pathway primarily monitors the structural integrity of the acceptor and T-stems of mature tRNA. *Genes & Development* **2011,** *25* (11), 1173-1184.

88. Furuichi, Y.; LaFiandra, A.; Shatkin, A. J., 5′-Terminal structure and mRNA stability. *Nature* **1977,** *266* (5599), 235-239.

89. Hsu, C. L.; Stevens, A., Yeast cells lacking 5'--> 3'exoribonuclease 1 contain mRNA species that are poly (A) deficient and partially lack the 5'cap structure. *Molecular and cellular biology* **1993,** *13* (8), 4826-4835.

90. Ramanathan, A.; Robb, G. B.; Chan, S.-H., mRNA capping: biological functions and applications. *Nucleic acids research* **2016,** *44* (16), 7511-7526.

91. Frindert, J.; Zhang, Y.; Nübel, G.; Kahloon, M.; Kolmar, L.; Hotz-Wagenblatt, A.; Burhenne, J.; Haefeli, W. E.; Jäschke, A., Identification, biosynthesis, and decapping of NAD-capped RNAs in B. subtilis. *Cell reports* **2018,** *24* (7), 1890-1901. e8.

92. Cohen, L. S.; Mikhli, C.; Jiao, X.; Kiledjian, M.; Kunkel, G.; Davis, R. E., Dcp2 Decaps m2, 2, 7GpppN-capped RNAs, and its activity is sequence and context dependent. *Molecular and cellular biology* **2005,** *25* (20), 8779-8791.

93. Deana, A.; Celesnik, H.; Belasco, J. G., The bacterial enzyme RppH triggers messenger RNA degradation by 5′ pyrophosphate removal. *Nature* **2008,** *451* (7176), 355-358.

94. Song, M.-G.; Bail, S.; Kiledjian, M., Multiple Nudix family proteins possess mRNA decapping activity. *Rna* **2013,** *19* (3), 390-399.

95. Cohen, L. S.; Mikhli, C.; Friedman, C.; JANKOWSKA-ANYSZKA, M.; Stepinski, J.; Darzynkiewicz, E.; Davis, R. E., Nematode m7GpppG and m32, 2, 7GpppG decapping: Activities in Ascaris embryos and characterization of C. elegans scavenger DcpS. *Rna* **2004,** *10* (10), 1609-1624.

96. Duan, J.; Liang, Z.; Yang, C.; Zhang, J.; Zhang, L.; Zhang, W.; Zhang, Y., Rapid protein identification using monolithic enzymatic microreactor and LC-ESI-MS/MS. *Proteomics* **2006,** *6* (2), 412-419.

97. Palm, A. K.; Novotny, M. V., Analytical characterization of a facile porous polymer monolithic trypsin microreactor enabling peptide mass mapping using mass spectrometry. *Rapid communications in mass spectrometry* **2004,** *18* (12), 1374-1382.

98. Lee, J.; Soper, S. A.; Murray, K. K., Development of an efficient on-chip digestion system for protein analysis using MALDI-TOF MS. *Analyst* **2009,** *134* (12), 2426-2433.

99. Datta, S.; Christena, L. R.; Rajaram, Y. R. S., Enzyme immobilization: an overview on techniques and support materials. *3 Biotech* **2013,** *3* (1), 1-9.

100. Cao, L., *Carrier-bound immobilized enzymes: principles, application and design*. John Wiley & Sons: 2006.

101. Mohamad, N. R.; Marzuki, N. H. C.; Buang, N. A.; Huyop, F.; Wahab, R. A., An overview of technologies for immobilization of enzymes and surface analysis techniques for immobilized enzymes. *Biotechnology & Biotechnological Equipment* **2015,** *29* (2), 205-220.

102. Ding, S.; Cargill, A. A.; Medintz, I. L.; Claussen, J. C., Increasing the activity of immobilized enzymes with nanoparticle conjugation. *Current Opinion in Biotechnology* **2015,** *34*, 242-250.

103. Motorin, Y.; Helm, M., RNA nucleotide methylation. *Wiley Interdisciplinary Reviews: RNA* **2011,** *2* (5), 611-631.

Chapter 3. Tailoring Thermoplastic In-Plane Nanopore Size by Thermal Fusion Bonding for the Analysis of Single Molecules

This chapter is based on the following manuscript,

"Tailoring thermoplastic in-plane nanopore size by thermal fusion bonding for single-molecule sensing".

Reprinted with permission from – {**Athapattu, U. S.**; Rathnayaka, C.; Vaidyanathan, S.; Gamage, S. S. T.; Choi, J.; Riahipour, R.; Manoharan, A.; Hall, A. R.; Park, S.; Soper, S. A., Tailoring Thermoplastic In-Plane Nanopore Size by Thermal Fusion Bonding for the Analysis of Single Molecules. *ACS Sensors* **2021**, *6* (8), 3133-3143.} Copyright {**2021**} American Chemical Society.

3.1 Introduction

Pores with nanometer dimensions are typically fabricated in a thin membrane separating two fluid chambers.[1] When an electrical bias is applied across the membrane, the resulting electric field can transport charged molecules through the pore producing signals in the trans-membrane ionic current that can be used to sense single molecules via resistive pulse sensing (RPS). Among single-molecule sensors, nanopores have garnered significant interest because they allow the detection of single molecules without requiring fluorescence labeling of the target and the need for sophisticated optical equipment for transduction.[2-5] These and other attractive properties have led to the development of many nanopore-based applications including detection of DNA-protein interactions,[6-9] measurement of molecular forces,[10-11] and nucleic acid sequencing.[12-13] Although biological pores, such as α-hemolysin[14] and MspA,[15] have proven to be useful sensors, several disadvantages remain primarily due to their fixed size and limited stability under extreme conditions of salt, pH, temperature, and mechanical stress. As an alternative, solid-state nanopores[3, 16] have captured attention to address challenges associated with biological pores. Moreover, solid-state nanopores can be integrated with other micro- and nanofluidic components to form lab-on-a-chip systems.

Most solid-state nanopores have been fabricated on inorganic thin-film membranes.[3, 17-19] Several approaches have been demonstrated to produce small nanopores in these substrates through charged particle beams[17-18, 20-21] or electrical breakdown[22] and to control the size of the pores *ex post facto* via exposure with a defocused beam of electrons,[17] ions,[23] direct thermal heating,[24] or focused ion beam (FIB) deposition of materials such as gold.[25] Even though these methods have proven successful in the fabrication of small-diameter pores, they are generally not conducive to production at a scale and cost that will ultimately enable them to be translated for clinical applications that require disposable devices as is necessary for *in vitro* diagnostics.

Solid-state nanopores have also been fabricated in planar substrates ("in-plane" nanopores) embedded within a fluidic network rather than suspended on a thin membrane.[26] FIB has been used to fabricate monolithic channels with micro- and nanoscale components including in-plane nanopores.[26-28] In-plane pores can also be produced in series so that other measurement modalities can be realized. For example, Kondylis *et al*. used glass nanopore devices with 2, 4, and 8 pores (width: 60 nm, depth: 70 nm) in series for real-time, resistive pulse analysis of virus capsids. They showed that the standard deviation of the pulse amplitude distributions of individual molecules decreased with increasing number of pores in series leading to increased measurement precision[29] while the electrophoretic mobility of virus particles have been determined using two in-plane pores.[30-32]

Thermoplastics provide the means for high-throughput manufacturing at low production costs even at the nanoscale due to a plethora of fabrication technologies available, such as nano-imprint lithography (NIL) and injection molding.[33-34] Additionally, due to the diverse physiochemical properties of different thermoplastics, the appropriate material can be selected according to measurement requirements.[35-36] However, it has been difficult to achieve sub-20 nm structures using thermoplastics due to challenges associated with bonding a cover plate to the nanofluidic network, which can result in deformation of the patterned nanostructures. The cover plate bonding process in thermoplastic devices typically involves the use of thermal fusion bonding (TFB), which bonds a thin cover plate to the nano-patterned substrate under a controlled pressure at temperatures near the glass transition temperature, T_g, of the substrate or cover plate. The TFB process involves motion of polymer chains between the cover plate and the substrate, which inevitably alters the dimensions of the nanostructures in the enclosed nanofluidic devices from those in the imprinted substrate. We have shown that high process yield rates of thermoplastic nanofluidic devices with minimal deformation of nanostructures can be realized using a hybrid

bonding process in which a lower T_g cover plate is thermally fusion bonded to a higher T_g substrate.[37]

The ability to control the size of in-plane nanopores imprinted from the same mold will allow for reduction in the development and production costs by obviating the need for FIB-milled Si masters to accommodate a particular application, for example reducing in-plane nanopore size to sense molecules of various sizes. Several reports have demonstrated reduction of nanostructure dimensions after initial fabrication by applying pressure to the patterned polymer substrate at an elevated temperature.[38] For example, Choi *et al.* reduced the size of micropores in a perforated SU-8 membrane produced by NIL from 3 μm to 300 nm.[39] The same group utilized polymer reflow, which consists of applying an elevated temperature without any external pressure, to reduce nanopore size from 12 nm to 6 nm.[34] In another report, Chou *et al.* described the use of a method called pressed self-perfection by liquefaction (P-SPEL), where the transiently molten thermoplastic nanostructures were pressed using a blank Si plate to achieve sub-20 nm structures.[38] However, these methods have not been demonstrated to reduce in-plane nanopore size for nanofluidic devices. Moreover, these methods employed an additional process step beyond the original production pipeline, which increases production time and cost.

The transport properties of biomolecules through nanopores depend on the specific surface interactions of analytes with the nanopore's surface.[40] Several reports discuss the functionalization of thermoplastic pore surfaces in order to achieve desired interactions with analyte of interest. For an example, Martin *et al.* reported a method to alter the surface properties of track-etched nanopores in polycarbonate with gold by electrodeless plating[41] followed by chemisorption of thiols.[42-43] In polyimide (PI) and polyethylene terephthalate (PET) the carboxyl groups created during track-etching process of nanopores have been chemically functionalized by covalent attachment through alkylation (using alkyl bromide and KF catalyst)[44-46], methylation

(treating with diazomethane)[47] or amidation (reaction with amine)[48-49]. Surface modification of various elastomers and thermoplastics have also been investigated to stabilize and alter the electroosmotic flow (EOF). Modification methods such as covalent coatings, dynamic coatings, and UV grafting were employed to modify the surfaces for separations.[50-52] Johnson *et al*. used a pulsed UV excimer laser to modify the surface charge of side walls of microchannels hot-imprinted on poly(methyl methacrylate) (PMMA) substrate.[53] In another approach, Henry *et al*. chemically modified laser ablated poly(ethylene terephthalate glycol) (PETG) microchannels to control the EOF by yielding amines.[54] previously our group reported surface modification of PMMA nanochannels and nanoslits to generate both negatively charged and positively charged surfaces.[55] A negative charge on the thermoplastic surface was generated via oxygen plasma treatment which forms carboxyl acids groups on the thermoplastic surface. These surface-confined carboxyl groups were subsequently converted in to positively charged amine groups by covalently attaching ethylenediamine (EDA) to the surface. Ali et al. used a similar amidation approach for PI nanopores by using EDC/NHS coupling chemistry followed by the covalent coupling of EDA.[56]

In this chapter, we demonstrate a post fabrication method to tailor the dimensions of in-plane nanopores enclosed in nanofluidic devices using TFB, a process step needed to produce enclosed nanofluidic devices. The thermoplastic dual in-plane nanopore devices were fabricated on either a PMMA (poly (methyl methacrylate)) or COP (cyclic olefin polymer) substrate and the O_2 plasma activated imprinted substrates and cover plates (made from COC; cyclic olefin copolymer) were subjected to different bonding pressures to vary the size of the in-plane nanopores. The change in depth and width of the nanopores with bonding pressure was measured by AFM and SEM, respectively. COMSOL simulations and experimental conductance measurements further demonstrated the pore closing behavior of the nanopores with higher

bonding pressures. The devices bonded at different pressures were used to analyze λ-DNA and showed improved signal-to-noise ratio (SNR) with lower nanopore size. Furthermore, nanopore devices with different pore sizes fabricated by changing the bonding pressure during TFB was used to analyze different types of molecules, such as ssRNA and rNMPs.

Moreover, to reduce the ion concentration polarization effects which leads to sub optimal event frequency, a simple surface modification step was carried out using Ethanolamine. The number of events observed with the O_2 plasma modified PMMA devices was less than anticipated, which could have been due to concentration polarization effects at the nanopore resulting from the surface charge being negative as well as the RNA/rNMP molecules. The high carboxyl acid group density generated during the O_2 plasma treatment prior to TFB creates a high surface charge which leads to the exclusion of co-ions passing through the nanopore. To obtain a neutral surface, ethanolamine can be used in a similar approach reported by Uba et al.[55] Ethanolamine was covalently attached to the O_2 plasma activated thermoplastic surface via EDC/NHS coupling chemistry. Surface characterization carried out showed the successful attachment of ethanolamine, which lead to a ~10- and ~9-fold decrease in surface charge and EOF, respectively, while maintaining the surface hydrophilicity. Dual in-plane nanopore devices modified with ethanolamine showed a significant increase in translocation events compared to O_2 plasma treated devices. Work reported herein provide the groundwork for integrating the dual in-plane nanopores in a novel exo-sequencing technique for the direct identification of rNMPs released sequentially from an intact RNA molecule following solid-phase enzymatic reaction for single-molecule RNA sequencing.

3.2 Material and Methods

3.2.1 Reagents and Materials

Chemicals and other materials were obtained from the following sources and used without further purification: S1813 photoresist (MicroChemicals, Germany); MF319 developer (MicroChemicals); potassium hydroxide (KOH) pellets (Fisher Scientific, Waltham, MA); isopropyl alcohol (IPA; Sigma-Aldrich, St. Louis, MO); hydrofluoric acid (HF, Sigma-Aldrich); Tri(propylene glycol) diacrylate (TPGDA, Sigma-Aldrich); Trimethylolpropane triacrylate (TMPTA, Sigma-Aldrich); 2,2-Dimethoxy-2-phenylacetophenone (photoinitiator, Sigma-Aldrich); NOA72 (Norland Products, Neuchâtel, Switzerland); Si wafers coated with 100 nm of Si_3N_4 (P/B, resistivity 5-10 Ωcm, orientation of (100), and 525 ±25 μm thickness, WaferPro, Santa Clara, CA); polyethylene terephthalate (PET) with 250 μm thickness (Goodfellow, Coraopolis, PA). PMMA was received from ePlastics (San Diego, CA). COC (Type 8007) was purchased from TOPAS Advanced Polymers (Florence, KY). COP sheets were obtained from STRATEC SE (Birkenfeld, Germany). UV curable polyurethane resin was purchased from Chansang Co. Adenosine 5'-monophosphate disodium salt was obtained from Sigma-Aldrich. Molecular biology grade water was secured from Thermo Fisher (Waltham, MA). SYTO82 dye was purchased from Life Technologies (Eugene, OR, USA).

3.2.2 Device Fabrication and Assembly

A Si master mold was prepared by using a combination of photolithography, wet-chemical etching, and focused ion beam (FIB) milling. Si wafers with a 100 nm thick silicon nitride (Si_3N_4) layer on each side were used for fabricating a Si master mold. Microchannels were fabricated using a combination of photolithography and wet-chemical etching (see Figure 3.1.A). To accomplish this, a 1.3 μm thick S1813 photoresist layer was first spin-coated at 4,000 rpm for 60 s on a Si wafer

and then baked at 115°C for 60 s. Photolithography was performed using a designed photomask in a UV exposure station (Quintel) in a class 100 cleanroom. UV exposure was conducted at 130-140 mJ/cm^2 with post-exposure baking at 95°C for 60 s. Then, the wafer was developed with a MF319 developer for 90 s, followed by washing with deionized water. The exposed Si_3N_4 layer was etched to open a window using an ICP-DRIE system (Plasmalab System 100, Oxford Instruments, Abingdon, UK). Subsequently, the wafer was transferred to a 40 wt% KOH solution with IPA (5 % v/v) at 70°C. The KOH solution was prepared by dissolving KOH pellets in deionized water. After 25 min etching to form 10 µm deep microchannels, the wafer was removed from the etchant, rinsed in water, and dried with N_2 gas. Prior to FIB milling, the Si_3N_4 layer was completely removed using a dilute HF solution. The nanochannel flight tube combined with in-plane nanopores was fabricated using FIB milling (Quanta 3D Dual Beam system, FEI, Hillsboro, OR). The milling was performed at a beam voltage and current of 30 kV and 10 pA, respectively, in a bitmap mode (see Figure 3.1.B).

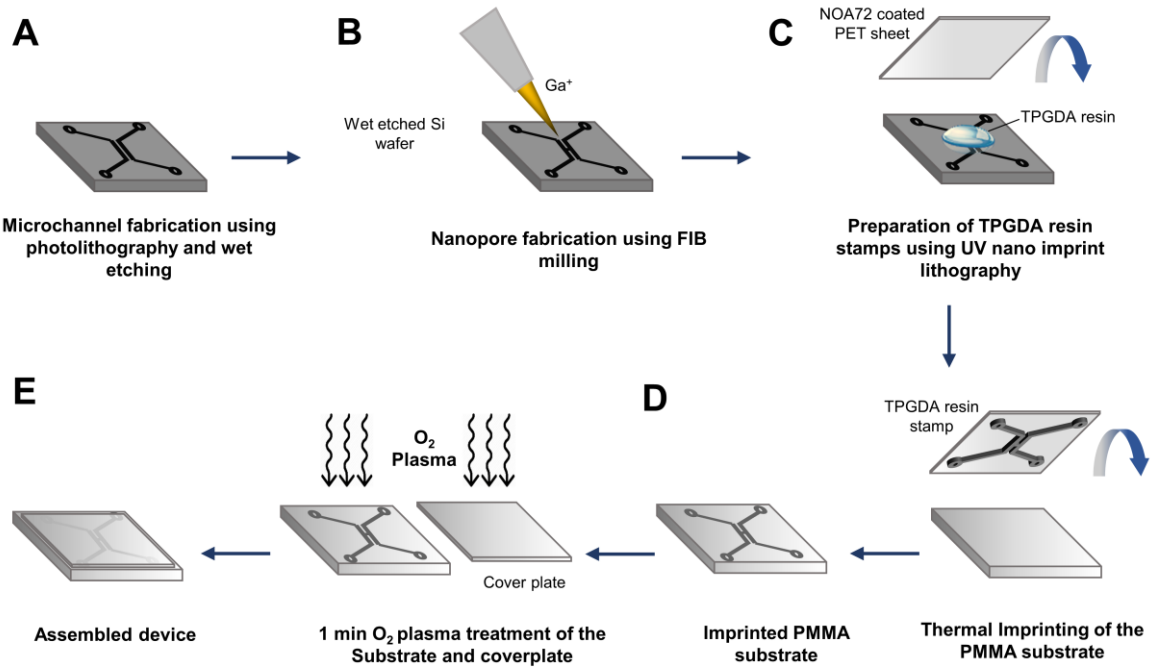


Figure 3.1. Schematic representation of the dual in-plane nanopore device fabrication. **(A)** Fabrication of microchannels using photolithography and wet etching. **(B)** FIB milling of nanopores. **(C)** TPGDA resin stamp fabrication. **(D)** Nano-imprint lithography for replication. **(E)** Cover plate assembly.

The Si master mold was used to produce a resin stamp by using a UV resin solution (70 wt% TPGDA, 28 wt% TMPTA, and 2 wt% photoinitiator). Drops of the UV-resin were dispensed against the Si master mold (see Figure 3.1.C). A flexible PET sheet coated with an adhesive layer (NOA72) was then slightly pressed against the liquid drop and used as a backbone substrate. Residual resin solution and air bubbles were gently squeezed out. During the curing process, the sample was exposed to flash-type UV light (250-400 nm) for 20 s at an intensity of ~1.8 W/cm^2 by using a nanoimprinter (Eitre6, Obducat, Lund, Sweden). After UV-curing, the molded UV-resin/PET substrate was demolded from the Si master.

Nanopores were imprinted into a plastic substrate using NIL as shown in Figure 3.1.D (Nanonex 2500, Monmouth Junction, NJ).[57] The optimized imprinting conditions for PMMA nanofluidic devices were 145°C, 300 psi, and 5 min and 130°C, 300 psi, and 5 mins for COP devices. Imprinted nanofluidic devices were then characterized using scanning electron microscope, SEM, and atomic force microscopy, AFM.

The nanopore device was sealed after thermal imprinting using a COC 8007 cover plate (see Figure 3.1.E). Thermal fusion bonding with NIL was used for sealing nanopore devices. Bonding of PMMA/COC devices were done following the method described by Uba *et al*. [37] with 1 min O_2 plasma treatment for both the substrate and cover plate at 50 W.

3.2.3 Atomic Force Microscopy (AFM)

To determine the depth of the nanopores with increasing pressure an AFM (SPM HT-9700, Shimadzu Corporation, Kyoto, Japan) analysis carried out. The tip used for imaging was a Super Sharp Silicon tip (Nanosensors, Switzerland) with a tip radius <2 nm, half cone angle of 10^o, aspect ratio 4:1 at 200 nm from the tip apex, and frequency of 300 kHz. A dynamic scanning

mode was used for imaging with a scanning frequency of 0.5 Hz. The acquired images were analyzed using SPM Manager v4.76.1 software (Shimadzu Corporation).

3.2.4 Scanning Electron Microscopy (SEM)

SEMs of the nanopore devices were acquired using a FEI VERSA 3D Dualbeam field emission/low vacuum SEM. A 2 nm thin conductive Iridium layer was sputter coated onto the thermoplastic devices using an EMS 150ES sputter coater before SEM imaging. All images were acquired using 5.0 kV accelerating voltage and 8.7 mm working distance. The SEM images of the Si mold masters were acquired using a Quanta^{TM} 3D DualBeam^{TM} FEI FIB-SEM and were analyzed using the instrument's software and Image J.

3.2.5 COMSOL Simulations

Simulations were performed in COMSOL v5.5 for the dual nanopores devices. The length of both nanopores was kept at 30 nm, but the width was varied from 10-50 nm to calculate the corresponding conductance. The electrolyte used was 1 M KCl and a DC field of 1 V at 293 K was applied in the electrostatics module to calculate the electric potential, the current density, and the corresponding conductance across the nanopores.

3.2.6 Biomolecule Translocation through Dual In-Plane Nanopores

Translocation experiments were performed for λ-DNA in PMMA dual in-plane nanopore devices bonded at 110,170, and 200 psi. Briefly, the devices were allowed to equilibrate with 50% v/v methanol/water for 5 mins and then after completely removing the methanol/water solution, 1X TBE carrier buffer was introduced for 15 mins. Finally, 100 nM of λ-DNA in 1 M KCl seeded into 1X TBE was injected into the device. The devices were placed in the Faraday cage and Pt electrodes were immersed in the reservoirs of the device. A potential of 1 V was applied, and the

data was acquired using the Axopatch Digidata 1440B, and analyzed using the Clampfit 11.1 software. The peak amplitude data (n≥120) was collected and plotted. The Wilcoxon p-test was used to calculate the statistical difference of peak amplitudes used for each TFB pressure.

For RNA and rNMP translocation nanopore devices were primed as previously described and 100 nM of RNA (60 bases/EGFP mRNA)/rAMP in 1X NEBuffer 3 was introduced to one of the reservoirs of the device's microchannels. For RNA/rAMP translocation the applied potential was increased by serially connecting a 1.5 V battery to the Axopatch circuit which increased the applied potential to 2.5 V. Potential was applied using Ag/AgCl electrodes and all data were collected with a sampling frequency of 250 kHz, a head stage configuration of β = 0.1, a gain of α = 1, and a low pass filter frequency of 10 kHz. The nanofluidic devices were kept inside a custom-built stainless-steel Faraday cage while recording current transient data. Data were collected for period of 10 mins and Clampfit 11.1 software was used for data acquisition and analysis.

3.2.7 Surface Modification with Ethanolamine

After fabrication and assembly of the dual in-plane nanopore devices, to suppress the surface charge and EOF, O_2 plasma modified surface was modified with ethanolamine. The ethanolamine was covalently attached to the surface confined carboxylic acid groups using the established EDC/NHS coupling chemistry which is used for the attachment of primary amine containing molecules to carboxylated surfaces via the formation of an amide bond.[58-59]

3.2.8 Sessile Water Contact Angle

The hydrophilicity/hydrophobicity of native, O_2 plasma modified, and ethanolamine modified PMMA surfaces were determined by water contact-angle measurements using a VCA Optima

instrument (AST Products). PMMA sheets (1.5 mm thick) were cut in to 1.5 cm × 1.5 cm and 2.0 µL of nanopure water (pH 7.5) was dispensed onto the substrate followed by capturing images and analyzing using the software provided by the manufacturer. The measurements reported were the mean± one standard deviation of five replicates at separate positions on the substrate.

3.2.9 Attenuated Total Reflectance Fourier Transform Infrared Spectroscopy (ATR-FTIR)

ATR-FTIR measurements were performed on UV/O_3 activated and ethanolamine modified PMMA plates. UV/O_3 activation was carried out instead of the O_2 plasma activation since the O_2 plasma treatment only modifies the first few monolayers which would not provide sufficient signal for viable observations. ATR-FTIR spectra were acquired from 400-4000 cm^{-1} using an ALPHA FTIR spectrometer and a Platinum ATR module (Bruker Optics). Six replicates were performed, and spectra were analyzed using Essential FTIR analysis software.

3.2.10 Surface Charge of Ethanolamine Modified Surfaces

Surface charge of ethanolamine modified PMMA-COC nanochannel device (107 µm long, 110 nm deep, and 110 nm wide) was measured by following the method described in Uba *et al.* Direct current conductance plots were utilized to assess the surface charge of ethanolamine modified nanochannel device. The surface of nanochannel device was modified with ethanolamine as described earlier and washed with ultrapure water prior to use. Nanochannels were filled with aid of capillary pulling from the inlet reservoir and vacuum suction from the outlet reservoir to confirm complete filling and avoiding air-bubble trapping inside the nanochannel. The pre-rinsed devices were filled with the KCl solutions and Ag/AgCl electrodes were placed into the access reservoirs and allowed to equilibrate for 4 min under a bias voltage. The conductance values were estimated by fitting the slope of the ionic current as a function of applied voltage, which was stepped from 1000 mV to -1000 mV with 100 mV step size and 10 s holding time for each data point. All

measurements were acquired with a low noise Axopatch 200B amplifier with a pClamp10 software and Digidata 1440B set at 10 kHz sampling frequency. The measurements were performed five times with repeated unloading and loading. The average conductance was plotted against the electrolyte concentration in a log-log plot and used in determination of surface charge.

3.2.11 Electroosmotic Flow (EOF)

The EOF of ethanolamine modified nanofluidic device was assessed using the current monitoring method.[55] A PMMA-COC device possessing a single nanochannel (107 µm long, 110 nm deep, and 110 nm wide) connecting two opposite access microchannels was fabricated as described in Amarasekara *et al.*[60] The entire chip was modified with ethanolamine as described earlier and washed with nuclease free water. After that, the device was filled with 0.1 M KCl solution and allowed to equilibrate for 5 min under a 500 mV DC bias. Next, one reservoir was replaced with 0.05 M KCl solution. Ag/AgCl electrodes were placed into the reservoirs across the channels under a 500 mV DC bias. pClamp10 software and Digidata 1440B low noise digitizer set at 10 kHz sampling frequency were used for data acquisition.

3.3 Results and Discussion

3.3.1 Device Fabrication and Assembly

Nanofluidic devices were fabricated in a thermoplastic using a method we have reported, which consists of making microstructures and nanostructures in Si masters followed by producing resin stamps via UV-NIL and production of the final device using thermal NIL into the appropriate plastic.[33, 61]

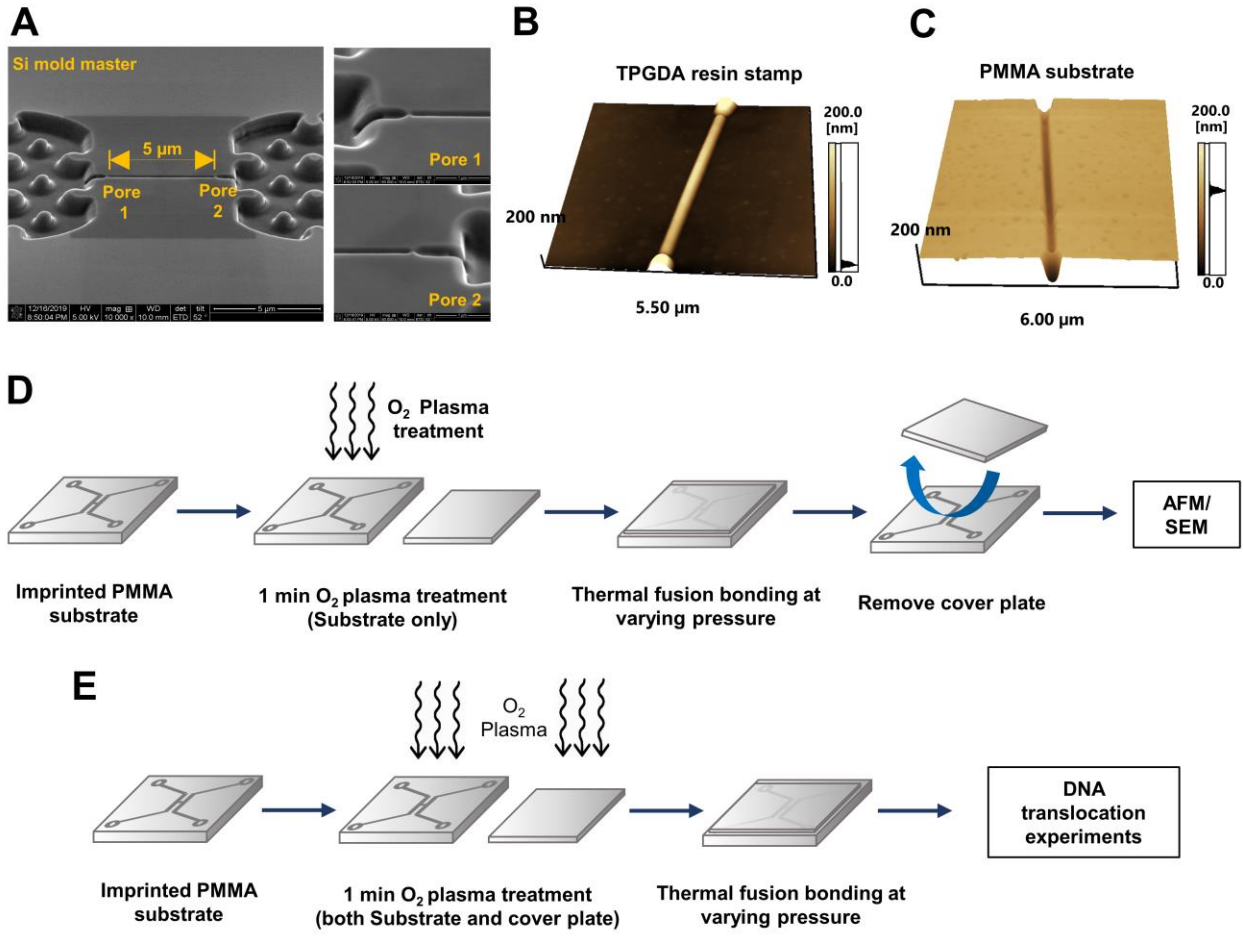


Figure 3.2. Dual in-plane nanopore device. **(A)** SEM image of the Si mold master. The two in-plane nanopores are 5 µm apart from each other. AFM scans of the **(B)** TPGDA resin stamp and **(C)** imprinted PMMA substrate. Tapping mode AFM scans were acquired at 0.5 Hz scanning frequency using a high aspect ratio tip with a radius <2 nm. **(D)** Schematic representation of experimental procedure for determining depth and width of dual in-plane nanopores. **(E)** Schematic representation of device assembly for translocation studies.

The Si master containing the in-plane nanopores and fluidic network was fabricated using a combination of photolithography and FIB milling. The in-plane nanopores were positioned at either end of a nanochannel, which was 5 µm in length and 50 nm x 50 nm in width and depth (see Figure 3.2.A). The nanochannel containing the in-plane nanopores was connected on both ends to access microchannels fabricated via photolithography and wet etching of Si. The structures on the Si master were transferred to a UV curable Tri(propylene glycol) diacrylate (TPGDA) resin situated on a NOA72 coated thin poly(ethyl terephthalate), PET, sheet (250 µm thick) serving as

a back plate by exposing to UV light for 3 min (see Figure 3.2.B).[62] Next, the TPGDA resin stamp was used to pattern the PMMA substrate at 145°C and 300 psi for 5 min using thermal NIL (see Figure 3.2.C). The average height of the in-plane nanopores on the resin stamp were 30.3 ±2.0 nm (n = 4) and the depth of the nanopores in the imprinted substrate were 29.6 ±1.7 nm (n = 3).

Following fabrication, the ability to control the size of the in-plane nanopores via TFB was examined by subjecting O_2 plasma activated (50 W, 1 min) NIL imprinted devices to different bonding pressures at 70°C for 15 min and measuring the depth and width of the in-plane nanopores using AFM and SEM. For this purpose, we only treated the imprinted PMMA substrate with O_2 plasma and not the COC cover plate prior to TFB to reduce the bonding strength so that the cover plate and substrate could be pulled apart without damaging the underlying structures (see Figure 3.2.D) for metrology. Previously we have reported the bond strength of PMMA/COC devices to be 0.086 ±0.014 mJ/cm^2 using the crack opening method.[60] In these experiments the bond strength between the O_2 plasma treated substrate and untreated cover plate was 0 mJ/cm^2 (*i.e.*, no crack was measured) meaning that the cover plate could be removed without material removal or deformation of the nanostructures in the substrate. A TFB temperature of 70°C was used as it was close to the T_g of the COC 8007 cover plate. For single-molecule translocation studies and RPS, the dual in-plane nanopore devices were assembled by O_2 plasma treatment of both the substrate and cover plate before TFB at varying pressures (see Figure 3.2.E) so that the bond strength was sufficient to sustain fluidic/translocation experiments.

The PMMA substrate was exposed to O_2 plasma to form O-containing groups.[35, 59, 63-65] In TFB, these O-containing species are involved in strong bond formation between the substrate and cover plate.[35] In addition, polymer chain scissioning can result in photo-fragments that are more thermally mobile due to their lower molecular weight. This leads to a lowering of the T_g of the polymer on the surface, making it easier for the polymer chains to fuse into the mating

substrate.[59] Depending on the type of polymer and the O_2 plasma power used as well as exposure time, polymer chain scission can occur up to several molecular layers into the bulk of the polymer. Therefore, during TFB of the cover plate to the substrate, the O_2 plasma activated polymer surfaces can rearrange leading to a change in nanostructure dimensions.

3.3.2 Nanopore Size Analysis

The depth of the in-plane nanopores were measured by dynamic mode AFM (Shimadzu Corporation) at a 0.5 Hz scanning rate (see Figure 3.3.A). For AFM, a high aspect-ratio tip with a radius <2 nm and a half cone angle of ~10^o was used. To measure the width of the in-plane nanopores following TFB, SEM was performed (see Figure 3.3.B). The depth of the in-plane nanopores reduced from 22.3 ±1.4 nm (110 psi, n = 6) to 10.2 ±1.5 nm (200 psi, n = 4) with increasing bonding pressure used for TFB (see Figure 3.3.C). The relative width of the in-plane nanopores decreased initially with bonding pressure to 0.47 ±0.04 (n = 4) at 130 psi compared to the width of the imprinted device but showed no statistically different widths at higher pressures (130-200 psi, p >0.05, see Figure 3.3.D). However, the overall cross-sectional area of the nanopores decreased with increasing pressure demonstrating the pore closing behavior at high TFB pressure (see Figure 3.3.E). Our in-plane nanopore in the imprinted PMMA is a U-shaped constriction with a flat bottom attributed to the na-ture of FIB milling process.

The results of the pore depth (Figure 3.3.C) and width (Figure 3.3.D) versus the bonding pressure provided a hint on the pore closing behavior. The decrease in both the pore depth and width in the low bonding pressure range can be attributed to the fusion of polymer chains between the two mating polymers, primarily in the thin layer of polymer subjected to O_2 plasma or UV/O_3

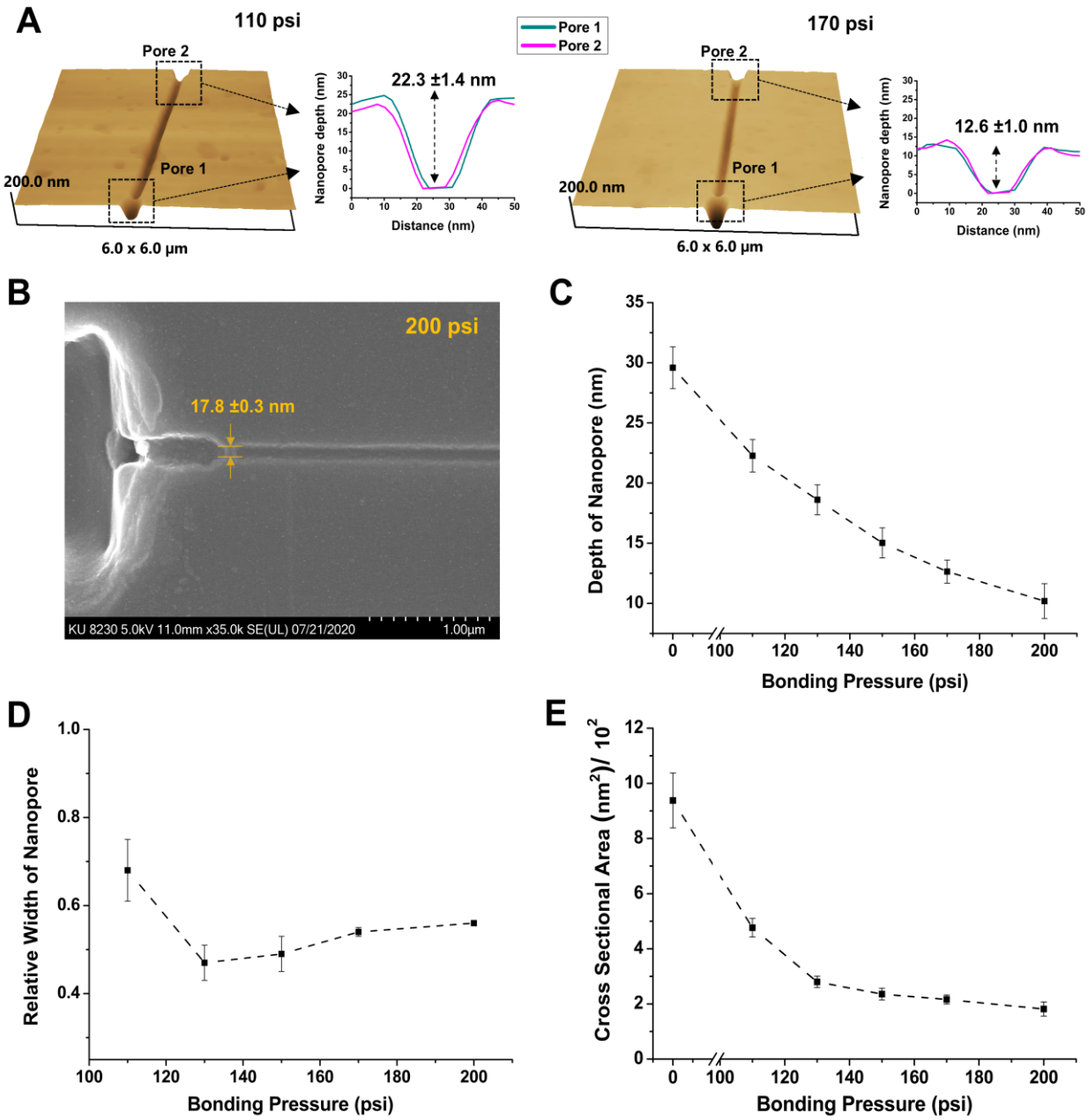


Figure 3.3. Nanopore depth and width with varying thermal fusion bonding pressure. **(A)** AFM scans of PMMA devices at 110 psi and 170 psi bonding pressures. **(B)** SEM image of PMMA device at 200 psi bonding pressure. A 2 nm thin conductive Iridium layer was sputter coated onto the PMMA device using an EMS 150ES sputter coater before SEM Imaging. **(C)** Change in the depth of the in-plane nanopores with bonding pressure. **(D)** Relative width of the in-plane nanopores after bonding at different pressures relative to the width of the nanopore before bonding (0 psi). There was no statistical difference in relative width from 130-200 psi at the 95% confidence interval ($p > 0.05$). **(E)** Cross sectional area of the in-plane nanopore with thermal fusion bonding pressure.

prior to bonding, and to the squeeze flow toward the nanopores. As the bonding pressure increases, the lateral squeeze flow will preferentially flow along the border of nanopores as was observed during the squeeze flow into hydrophilic nanostructures,[66-67] and thus further reduce the pore depth while limiting the decrease in the pore width. Consequently, the nanopore bonded at higher pressures became a shallower V shape. At the same time, the decrease in the thickness of the surface modified thermoplastic layer will further limit the squeeze flow because of increases in the T_g for polymers in the vicinity of a hydrophilic surface due to the alignment of polymer chains along the surface.[68] This may account for the slight increase in the pore width at higher bonding pressures. Further increases in the bonding pressure beyond 200 psi ultimately led to collapse of the nanopores as evidenced by the cessation of the open pore current.

3.3.3 COMSOL Simulations and Conductance Measurements of Devices

We carried out COMSOL simulations for the dual in-plane nanopore sensor to estimate the change in the conductance with change in pore size. The length of the nanopore was maintained at 30 nm, but the width changed from 10 nm to 50 nm (see Figure 3.4.A). A DC bias of 1 V was applied across the ends of the device and 1 M KCl was used as the electrolyte. From Figure 3.4.B, it can be seen that the majority of the potential drop occurred across the two in-plane nanopores (0.03 V in each pore) and the 5 µm long nanochannel. A current density graph was plotted that showed a sharp increase in current density at the nanopores due to the smaller dimensions (see Figure 3.4.C). The current density was integrated over the area of the nanopore to calculate the current. The conductance of the pores was then calculated using Ohm's law and was plotted to estimate the conductance at different pore sizes. As shown in Figure 3.5.A, a linear increase in conductance was seen as the size of the nanopore increased. The conductance values obtained from simulations and depths and widths from AFM and SEM, respectively, were used to estimate the size of the nanopore with the conductance values obtained experimentally.

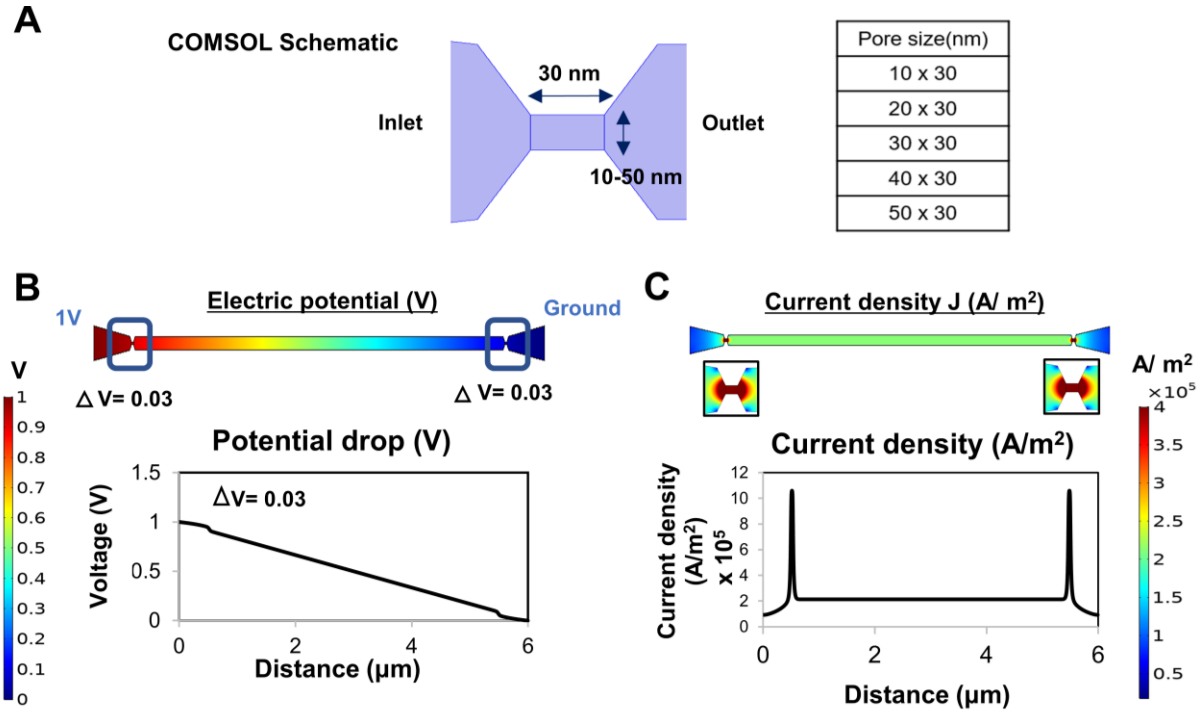


Figure 3.4. (A) The 2D design of the dual in-plane nanopores used for COMSOL simulations. The pore and intervening 5 μm long nanochannel were assumed to be cylindrical. In order to understand if a change of in-plane nanopore size (diameter) would cause an increase in conductance, the pore diameter was varied from 10-50 nm while the length was kept constant at 30 nm. **(B)** The electric potential data from COMSOL simulations shows that the majority of the potential drop appears across the two nanopores and the nanochannel implying that the overall conductance is contributed by the two nanopores and the nanochannel. **(C)** The current density was plotted from which the current and the subsequent conductance was calculated (I/V).

Experimental conductance measurements were performed using the dual in-plane nanopore devices made in PMMA and COP bonded at different pressures using 1 M KCl as the electrolyte. The conductance was calculated from the slope of the curve for all bonding pressures and is shown in Figure 3.5.B. The average conductance at 110 psi for PMMA and COP devices were 83 ±29 nS (RSD 34%) and 128 ±89 nS (RSD 69%), respectively. However, at 130 psi the average conductance of the PMMA and COP dual in-plane nanopore devices was 23 nS ±6 nS (RSD 26%) and 21 ±16 nS (RSD 76%), respectively.

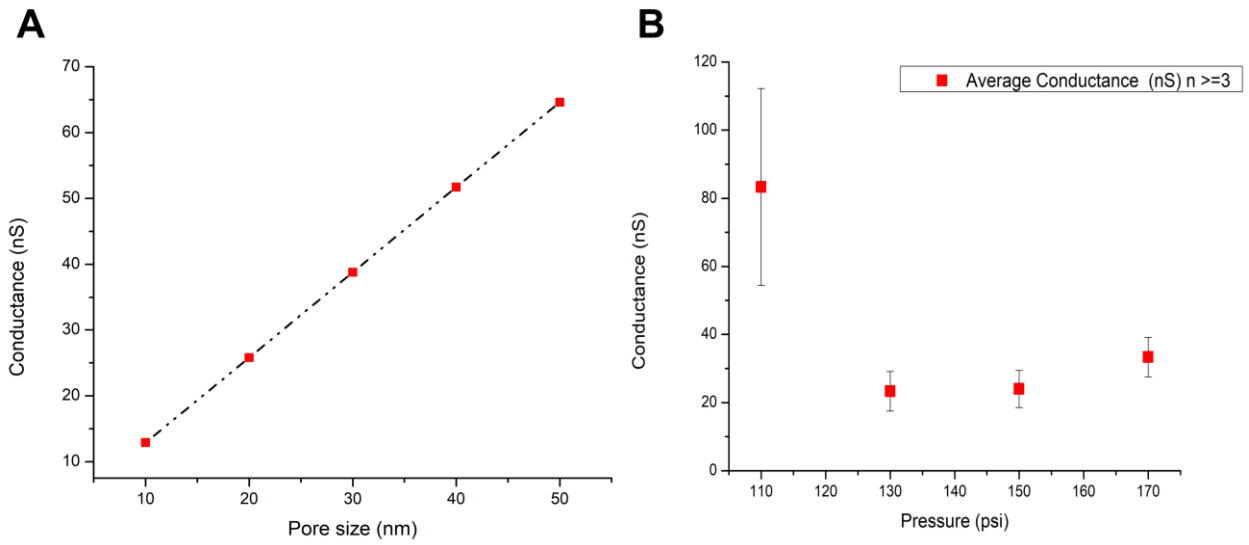


Figure 3.5. (A) Conductance (nS) calculated from COMSOL for varying pore size in 1 M KCl. There is a linear increase in conductance with increasing pore width. **(B)** Variation of measured conductance through the dual in plane nanopore PMMA and COP devices at different bonding conditions using an electrolyte of 1M KCl ($n \geq 3$). There was a decrease in conductance with increase in bonding pressure, but with no statistical differences at pressures above 130 psi ($p>0.05$). The conductance results agree with the pore size determined using AFM and SEM correlated to the results from COMSOL. The y-axis scales of graphs for figures A and B are adjusted according to their corresponding x-axis and hence the range might be different.

These values correlate well with the conductance (~23 nS) obtained from COMSOL for an 18 nm pore, which is the size of the pore obtained from AFM and SEM when using 130 psi TFB pressure. The conductance values showed a slight increase from 24 ±5 nS (RSD 22%) to 33 ±6 nS (RSD 18%) for PMMA devices bonded at 150 and 170 psi, respectively, but there was no statistical difference in the conductance values at bonding pressures >130 psi ($p >0.05$). Similarly, in the case of COP devices the conductance values decreased at 150 psi to 9 ±7 nS (RSD 77%) but showed a statistically insignificant increase at 170 psi to 12 ±6 nS (RSD 48%). This small increase in conductance can be correlated to the results from SEM and COMSOL that showed <5 nm increase in pore width. Overall, both PMMA and COP devices followed a similar trend of decrease in conductance at 130 and 150 psi followed by a slight increase at 170 psi showing the reproducibility of our approach with different substrate materials.

3.3.4 λ-DNA Translocation through the Dual In-plane Nanopores

Assembled dual in-plane nanopore devices could be used to electrokinetically drive charged single molecules through the nanopores for single-molecule sensing by monitoring the electrical current as a function of time. Figure 3.6 shows the use of a dual in-plane nanopore device as a sensing platform for the detection of unlabeled 48.5-Kbp λ-DNA that has a contour length of 16.5 µm.

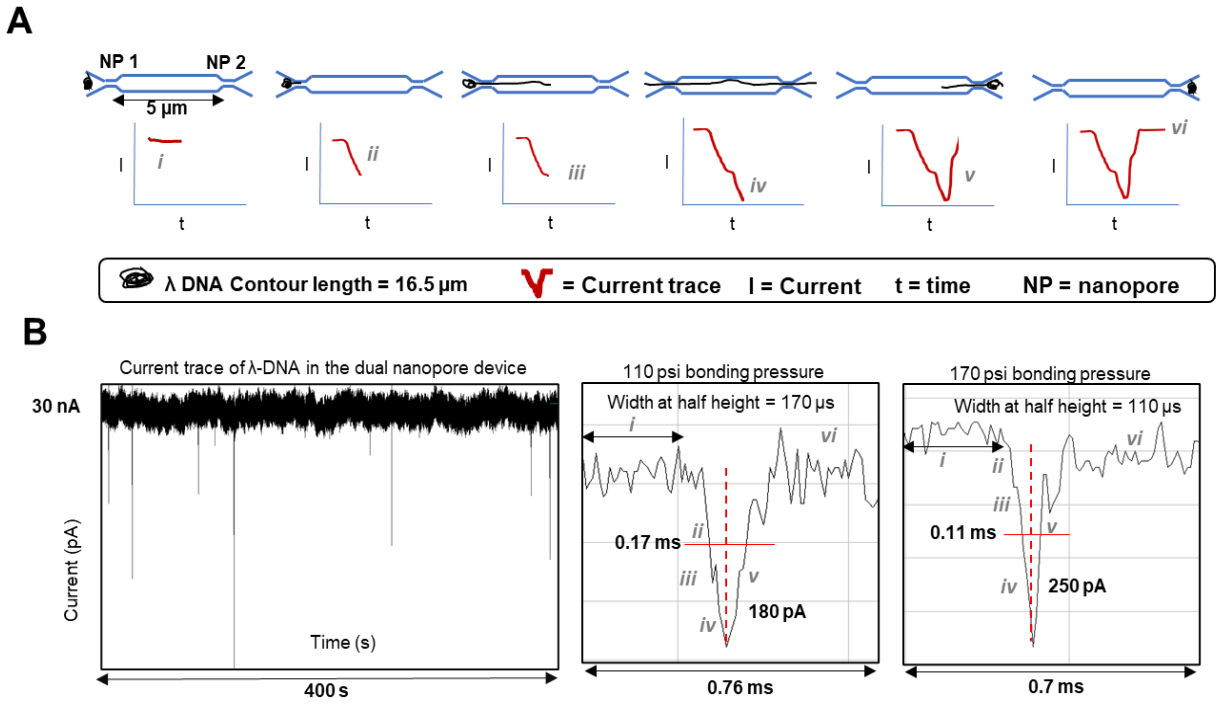


Figure 3.6. λ-DNA translocation through the dual in-plane PMMA nanopores and the ramifications of the size of the nanopore on peak amplitude. **(A)** Schematic of the λ-DNA translocation through the in plane dual nanopore device that gives rise to a negative peak as the DNA enters the first pore. Since the contour length of the DNA is longer than both the pores, there is a second subsequent peak when the DNA co-resides in both the pores. The DNA then leaves the pores very quickly which makes the current return to the baseline. **(B)** A detected current transient trace typically observed in a time interval of 400 s as a result of λ-DNA translocation and magnified images of individual peak shapes at various translocation stages of the DNA through the dual nanopore at 110 and 170 psi pressure, respectively.

When the electrokinetically driven molecule entered the first nanopore, there is a partial current blockage creating a transient increase in the electrical resistance, which is manifested as a negative peak (current drop) in the measured trace as shown in the schematic of Figure 3.6.A.

Because the contour length of λ-DNA is longer than the distance between the two in-plane nanopores (5 µm), the initial drop of current is accompanied by a subsequent drop in current when the DNA co-resides in the first and second in-plane nanopores. Further, when the DNA exits the first nanopore, it resides only in the second nanopore causing a subsequent small shoulder in the current trace and eventually returning to baseline. A current trace showing multiple current transient signals over a time interval of 400 s for λ-DNA is shown in Figure 3.6.B demonstrating the characteristic shape of the peaks at 110 psi and 170 psi. The various stages of the translocation are marked in both traces to show how the shape of the current transients agree with the aforementioned descriptions. Although the average amplitude increased 2-fold with every increase in bonding pressure, the shape of the peaks remained similar to that shown in Figure 3.6.B indicating that the size of the nanopore did not alter the translocation dynamics rather changed only the SNR in the resistive pulse.

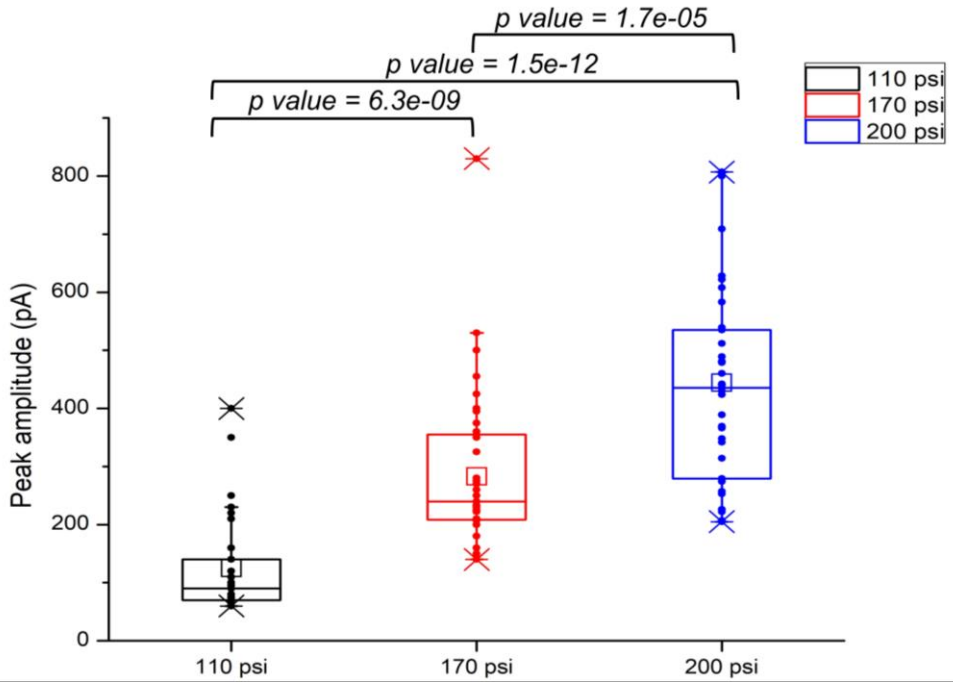


Figure 3.7. Distribution of peak amplitudes of λ-DNA at 110, 170 and 200 psi bonding pressures. The average peak amplitude increases with the increasing bonding pressure. p values calculated between each bonding pressure condition (Wilcoxon signed rank test) show statistically significant difference at 95% confidence interval ($p < 0.05$).

Peak height measurements (n ≥120) of the current transient amplitudes of λ-DNA in devices bonded at different pressures showed differences between each bonding pressure as determined by the Wilcoxon signed rank test (Figure 3.7). The peak amplitudes were collected from over >10 devices for each bonding pressure as we were only able to see ~6 events per device. The average peak amplitude of λ-DNA in devices bonded at 110 psi for 15 min was 130 pA corresponding to a pore depth of ~22 nm and a width of ~21 nm. For devices bonded at 170 psi for 15 min having a pore depth and width of ~13 nm x ~17 nm yielded a current amplitude of 280 pA and for devices bonded at 200 psi for 5 min having a depth and width of ~10 nm x ~18 nm yielded a current amplitude of 437 pA. There was a 3.5-fold increase in current amplitude when the pore dimensions decreased from 22 nm x 21 nm to 10 nm x 18 nm (depth x width). Devices bonded at 200 psi for 15 min served as the upper limit because the nanopore became unusable as the cover plate collapsed into the nanopore. However, for cases where 200 psi bonding pressure were used, the TFB time had to be reduced to 5 min to generate functional devices.

We calculated the apparent mobility (μ_{app}) of λ-DNA based on the time interval between peaks (Δt) at time points (i) and (iv) (see Figure 3.6.A). These time points were used because one end of λ-DNA is entering the first in-plane nanopore and that same end is then entering the second nanopore corresponding to a length of 5 μm (l) at the applied electric field strength (*E*); the apparent mobility was calculated from:

$$\mu_{app} = v_{app}/E \qquad v_{app} = l/\Delta t \tag{1}$$

The average apparent mobility for λ-DNA was determined to be 2.57×10^{-7} ±9.4×10^{-8} m²/Vs.

3.3.5 RNA and rAMP Translocation through O_2 Plasma Modified PMMA Dual In-Plane Nanopore Devices

To further demonstrate the use of the dual in-plane nanopores for identification of various sized molecules, we carried out translocation of 60 nt ssRNA (R_g ~6 nm) molecules using the dual in-plane nanopore devices bonded at 170 psi which resulted in a pore depth and width of 12.6 ±0.9 nm x 17.1 ±0.2 nm. As shown in Figure 3.8.A the assembled PMMA dual in-plane nanopore devices were primed with 50% v/v methanol for 5 mins followed by 15 min with 1X NEBuffer 3. Before introducing RNA/rAMP solution the current was monitored to establish the baseline (see Figure 3.8.B). Then, the buffer in one reservoir was replaced with 100 nM ssRNA solution and a potential (1 – 2.5 V) was applied.

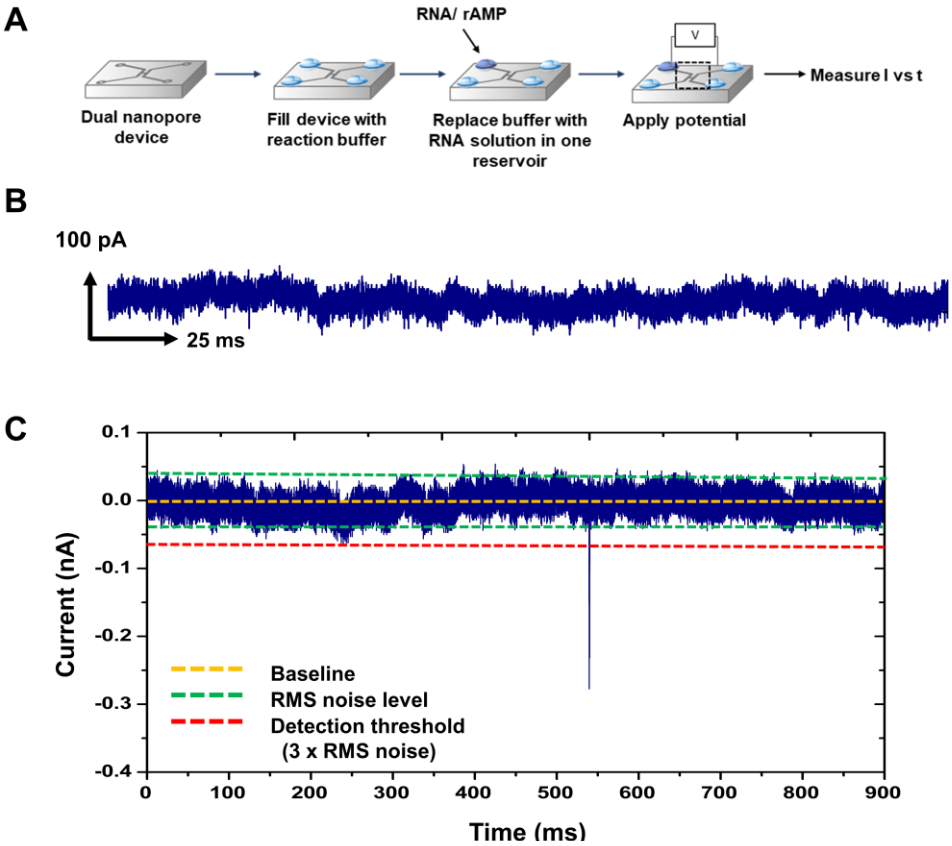


Figure 3.8. Biomolecule translocation through dual in-plane nanopores under applied electric field. **(A)** Schematic representation of the reaction procedure. **(B)** 25 ms trace of the open pore (baseline) current. **(C)** 900 ms current trace obtained for 60 bases RNA in a O_2 plasma treated nanopore device. The open pore current is subtracted from the trace.

For the 100 nM ssRNA solution, we only observed very few translocation events even after increasing the concentration to 1 µM and the driving voltage to -2.5 V (see Figure 3.8.C). Due to the low event frequency a large number of devices had to be used to collect reasonable number of events to secure meaningful statistics as was done for the λDNA data shown in Figure 3.6. The low event frequency per device could have been due to co-ion exclusion due to the high surface charge of the O_2 plasma modified PMMA devices, which is reported as -40 mC/cm^2.[55] Due to the high negative charge on the nanopore surface and partial electric double layer overlap within the nanopore, the entry of negatively charged co-ions such as ssRNAs is impeded. In addition, the opposing EOF also serves to reduce the transport. To address this issue, we investigated ways to reduce the surface charge of O_2-plasma activated PMMA/COC devices.

Previously our group reported the use of EDA to reverse the polarity of the surface charge of PMMA thermoplastic devices[55]. The EDA molecules were covalently attached to the O_2 plasma treated surfaces using EDC/NHS coupling chemistry. We investigated if the same method could be used to obtain a neutral surface by attaching ethanolamine to the O_2 plasma activated PMMA surface. Ethanolamine contains a primary amine which can be used for the covalent attachment. Due to the high pKa of the terminal hydroxyl group of ethanolamine, after immobilization the OH will be protonated at the reaction pH (pH 7.9) providing a neutral yet hydrophilic surface. The hydrophilic nature will be important for the continuous filling of the nanofluidic device by spontaneous capillary action.

3.3.6 Surface Modification with Ethanolamine and Characterization

Surface modification with ethanolamine was initially tested on planar COC and PMMA substrates (1.5 cm × 1.5 cm). As shown in Figure 3.9, PMMA and COC substrates were exposed to O_2 plasma at 50 W for 1 min to functionalize the surface with carboxylic acid groups. The plasma

treated thermoplastics were then soaked in 5 mL buffered solution (0.1 M MES, pH 4.7) containing 100 mg EDC, 10 mg NHS, and ethanolamine (final concentration – 50 mM) for 30 min at room temperature. After incubation, samples were washed with ultrapure water and dried in air. The same procedure was adopted for the ethanolamine modification of assembled nanofluidic devices. In this case, the assembled devices were immediately filled with the EDC-NHS-ethanolamine/MES solution and allowed to incubate for 30 min and rinsed with ultrapure water prior to use in experiments.

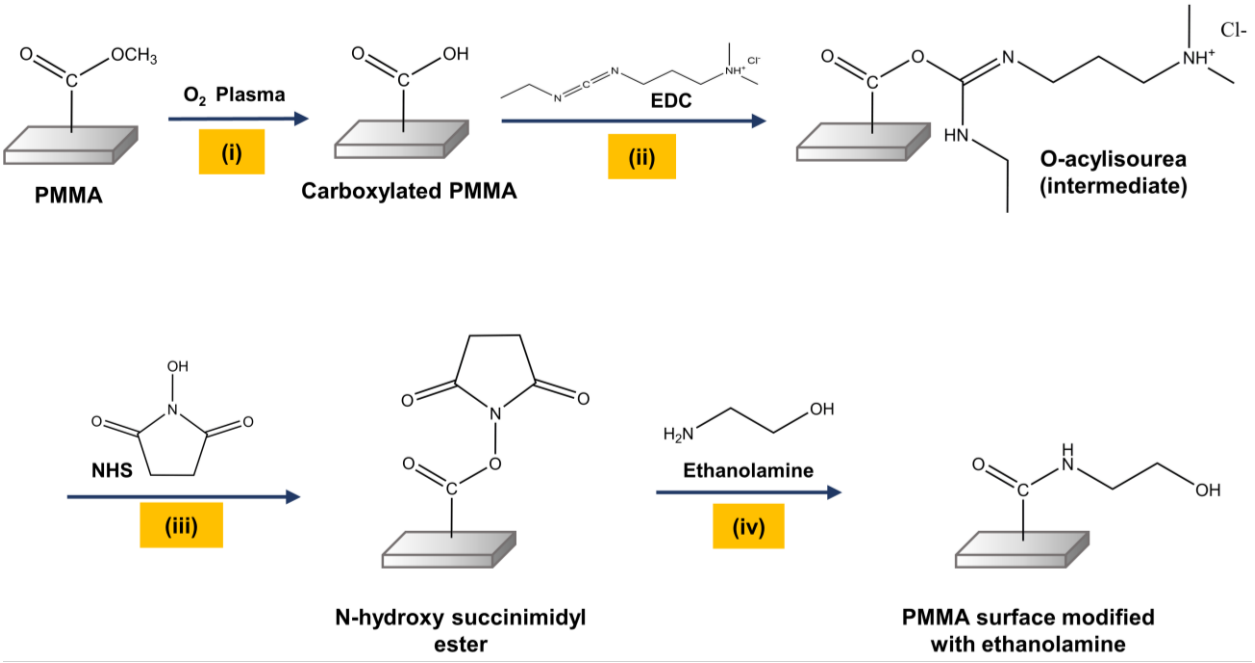


Figure 3.9. Schematic representation of the covalent attachment of ethanolamine to the O_2 plasma activated PMMA surface using EDC coupling chemistry. **(i)** Generation of surface carboxyl groups by O_2 plasma activation, **(ii)** O-acylisourea intermediate by reaction with EDC, **(iii)** formation of N-hydroxy succinimidyl ester by chemical reaction with NHS, and **(iv)** covalent attachment of ethanolamine to the surface via amide bond formation.

3.3.6.1 Water Contact Angle

During the surface modification steps, water contact angle measurements were assessed on the native, O_2 plasma modified, native- ethanolamine treated without EDC/NHS and ethanolamine modified PMMA as a simple monitor of the variations generated on the surface due to the modification methods (see Figure 3.10). The contact angle dropped from 79.23 ±1.78° for pristine

PMMA to 42.20 ±1.02° after O_2 plasma treatment, indicating the generation of surface carboxyl groups. The values have a better agreement with values reported in literature.[60] After amidation of carboxyl groups with ethanolamine via EDC/NHS coupling chemistry, sessile drop water contact angle (53.40 ±2.5°) slightly increased.

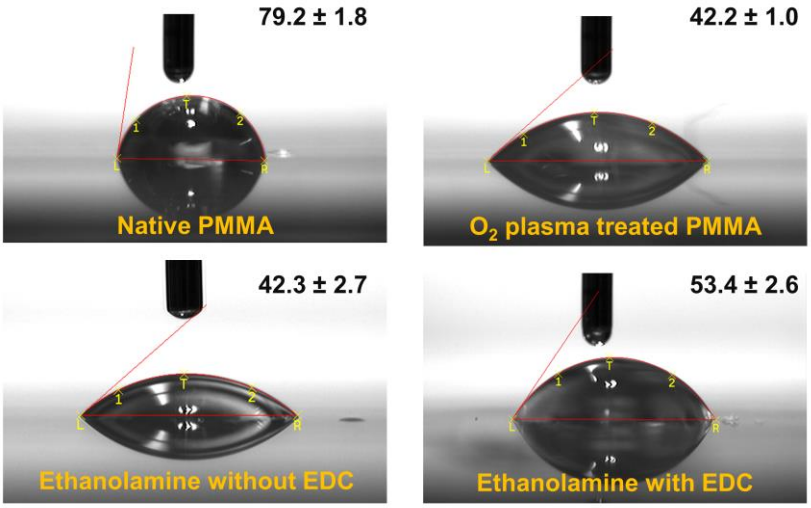


Figure 3.10. Sessile water contact angle of native PMMA, O_2 plasma treated PMMA and ETA modified PMMA surfaces with and without EDC/NHS treatment. 2 µL of 18 MΩ.cm water was dispensed using a syringe and water contact angle was measured (n ≥ 5) using a VCA optima instrument (AST products).

When the native PMMA is exposed to O_2 plasma due to the formation of surface carboxylic acid groups, the surface becomes hydrophilic. As evident from the water contact angle the ethanolamine treated surfaces remained hydrophilic compared to the native PMMA due to the presence of a terminal hydroxyl group. The hydrophilic nature of the ethanolamine treated surface is helpful in consistent filling of the nanofluidic device without generating air bubbles. The water contact angle of native PMMA treated with ethanolamine in the absence of EDC/NHS was 42.34 ±2.73°. The water contact angle of O_2 plasma treated PMMA was not changed after treatment ($p= 0.7742$, $n=5$, contact angle data for O_2 plasma modified vs. native- ethanolamine treated with ethanolamine in the absence of EDC/NHS) with ethanolamine only and this confirms that there is no nonspecific adsorption of ethanolamine on carboxyl group modified PMMA. Ethanolamine forms a covalent attachment to the PMMA surface only through EDC/NHS coupling chemistry.

4.3.6.2 ATR-FTIR Analysis of Ethanolamine Modified PMMA and COP

To examine the molecular nature of the modified and unmodified thermoplastic surfaces with ethanolamine, ATR-FTIR experiments were performed. As noted, ATR-FTIR has penetration depths of 0.5 – 2 µm into the bulk material. Therefore, thermoplastics were UV/O_3 activated for 15 min at 22 mW cm^{-2} power prior to ethanolamine modification.[69] The surfaces activated with UV/O_3 observed sufficient ATR-FTIR signal (Figure 3.11) as activation occurs into depth of the thermoplastics whereas, O_2 plasma activation occurs only within first few monolayers.

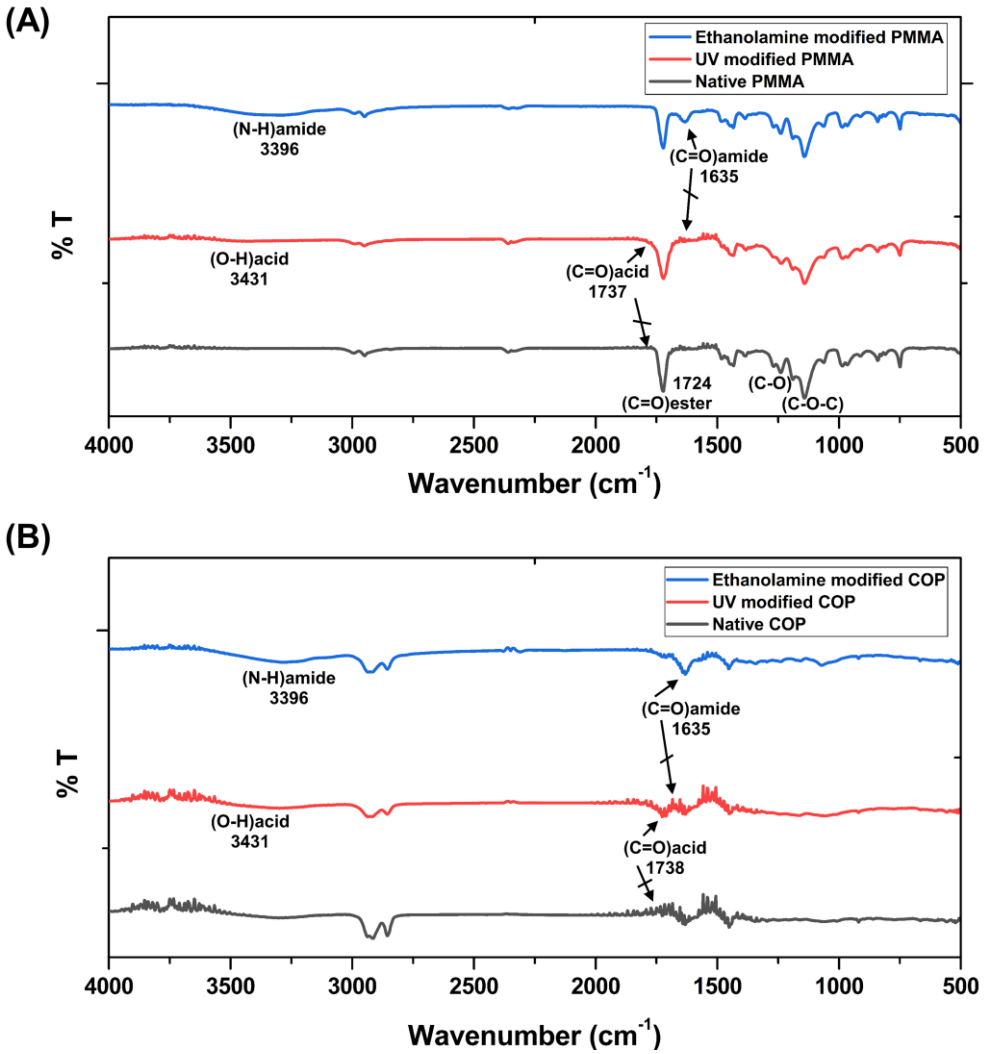


Figure 3.11. ATR-FTIR spectra of native, UV activated, and ETA modified **(A)** PMMA and **(B)** COP. ATR-FTIR spectra were acquired from 400-4000 cm^{-1} using an ALPHA FTIR spectrometer and a Platinum ATR module (Bruker Optics). Spectra (n = 6) were analyzed using Essential FTIR analysis software.

An ATR-FTIR spectrum of native PMMA with the characteristic peaks between 4000 and 500 cm^{-1} is shown in Figure 3.11.A. The most prominent band was v(C=O) at 1724 cm^{-1} assigned to the methacrylate ester stretch. The peaks at 1270, 1240 cm^{-1} and 1195, 1150 cm^{-1} could be assigned to v(C-O) and v(C-O-C) stretching of an ester. After UV/O_3 activation, there was the appearance of a band at 3441 cm^{-1} and 1737 cm^{-1}, which could be assigned to the v(O-H) and v(C=O) of a carboxylic acid group. There is a better agreement with the FTIR spectra of native and plasma modified PMMA reported in literature.[52, 55] Upon treatment with ethanolamine led to the appearance of bands at 3396 cm^{-1} and 1635 cm^{-1} corresponding to the v(N-H) stretch of a primary amine and v(C=O) of an amide. ATR-FTIR spectrum of native COP ranging from 4000 cm^{-1} - 1250 cm^{-1} is shown in Figure 3.11.B. The peaks ranging from 2900-2800 cm^{-1} and a peak at 1450 cm^{-1} could be assigned to -CH stretching and -CH bending. This spectrum showed good agreement with previously reported results.[70] After exposing to UV/O_3, a new peak appeared at 1738 cm^{-1} which is corresponded to carbonyl group (C=O) of carboxyl acid moiety. The peak showed at 3431 cm^{-1} was assigned to the hydroxyl group (-OH) of carboxyl acid. Two new bands appeared at 1635 cm^{-1} and 3396 cm^{-1} were assigned to carbonyl group of amides and -NH stretching of amine respectively.

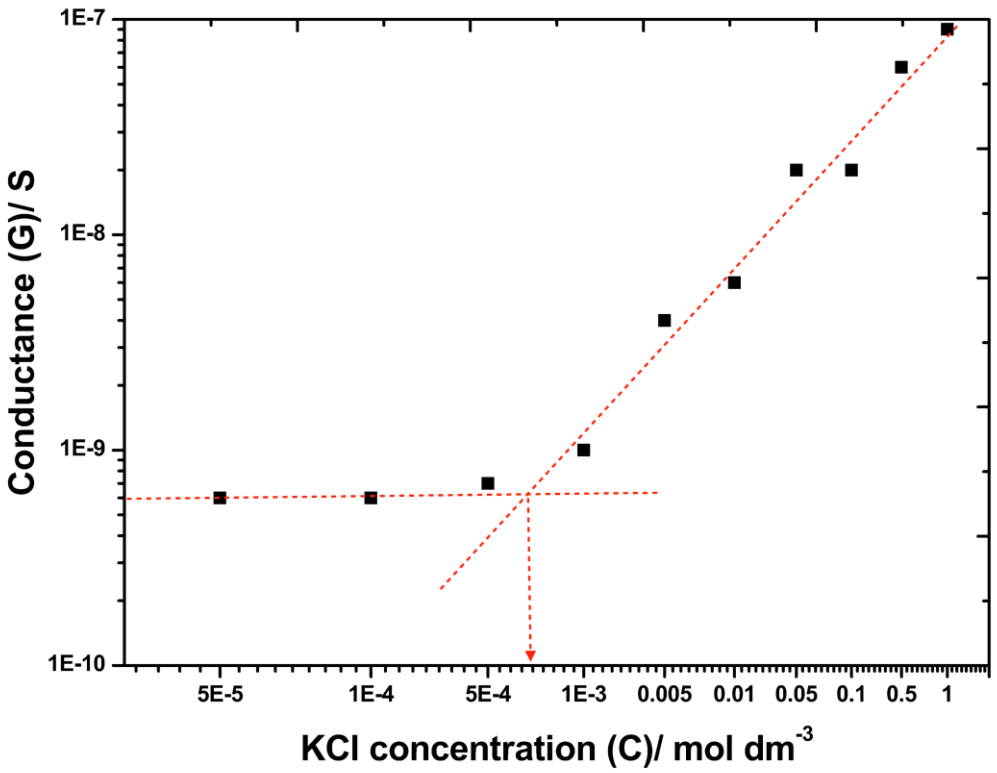
3.3.6.3 Surface Charge Density and EOF of Ethanolamine Modified PMMA Devices

Electrical conductance measurements across ethanolamine modified nanochannel device filled with different KCl concentrations have been used to work out the surface charge density. The average conductance was plotted against the electrolyte concentration in a log-log plot and the surface charge (σ_s) density was deduced by fitting the conductance plot according to;[55]

$$G_T = 10^3(\mu_{K^+} + \mu_{Cl^-})cN_A e.\frac{nwh}{L} + 2\mu_{opp}\sigma_s n\frac{(w+h)}{L} \quad (2)$$

where G_T is the total measured conductance in the nanochannel, μ_{K^+} and μ_{Cl^-} are ion mobilities of K^+ and Cl^- ions, respectively (μ_{K^+}= 7.619 × 10-8 $m^2V^{-1}s^{-1}$ and μ_{Cl^-}=7.912 × 10^{-8} m^2 V^{-1} s^{-1}), c is the electrolyte concentration in mol L^{-1}, N_A is Avogadro's number, e is the electron charge, n is the number of nanochannels in the device, w, h, and L are width, height and length of the nanochannel respectively, and $\mu_{opp} \approx \mu_{K^+}$ for the deprotonated carboxyl surface (Figure 3.12.A).

A



B

	EOF (cm^2/ V s)	Surface charge (mC/ m^2)	Zeta potential (mV)
ETA modified PMMA	3.63 (±0.10) x 10^{-5}	-3.8	-5.2
O_2 plasma modified PMMA	4.10 (±0.10) x 10^{-4}	-40.1	-59

Figure 3.12. (A) Conductance plots obtained from ETA modified PMMA devices consisting an array of four nanochannels (each 100 nm wide, 100 nm deep, and 107 µm long). Each data point represents a mean of five measurements with a scatter in the data within 5-8 % of the average value. The calculated effective surface charge density from the graph was -3.8 mC/m^2. **(B)** Measured EOF values as well as surface charge density zeta potential for ETA modified PMMA nanochannel devices investigated at pH 7.8. (The EOF and zeta potential for O_2 plasma modified PMMA were reported in reference [57]. The surface charge density for plasma modified nanochannel device was reported reference [55]).

It is reported that carboxylic acid groups are generated on PMMA and COC upon UV/O_3 activation[59] or O_2 plasma treatment.[71-72] After ethanolamine modification of O_2 plasma treated devices, the surface is dominated by hydroxyl groups. When ethanolamine modified surfaces are in contact with an electrolyte solution at pH 7.8, ~99.9% of the hydroxyl groups (pK_a ~16.0) would be protonated and ~99.9% of the carboxyl groups (pK_a = 4.66) would be deprotonated.[73] Therefore, these deprotonated carboxylic acid moieties are responsible for generating a surface charge density. The transition concentration, c_t used to calculate σ_s was 0.78 mM for ethanolamine modified surface. For ethanolamine modified PMMA nanochannels, we obtained $|\sigma_s|$ ~ 3.8 mC/m^2, which was ~10 fold less than 40.5 mC/m^{-2} reported by Uba *et al.* for O_2 plasma modified nanochannels.[55]

The EOF of ethanolamine modified nanochannel device was measured using the current monitoring method.[74] The EOF can be described by $\mu_{eof} = \upsilon_{eof}/E$, where υ_{eof} is electroosmotic flow velocity and E is the field strength. As noted above, plasma activated PMMA surfaces are negatively charged due to deprotonation of carboxylic acid groups and ethanolamine modified PMMA surfaces are uncharged at pH 7.8. The EOF for ethanolamine modified PMMA nanochannel device was 3.63 $\times$ 10^{-5} cm^2/Vs (see Figure 3.12.B) which was ~9-fold less than 4.1×10^{-4} cm^2/Vs reported by Amarasekara *et al.* for O_2 plasma modified PMMA nanochannel device.[60] The zeta potential, ζ was computed using equation (3). At low electric double layer thicknesses (λ_D ≈0.8 nm for 1X NEBuffer 3) μ_{eof} can be represented by Helmholtz-Smoluchowski equation;

$$\mu_{eof} = \frac{\varepsilon_0 \varepsilon_r \zeta}{\eta_0} \quad (3)$$

where ε_0, ε_r are the permittivity of vacuum, and the relative permittivity of the buffer (80.1), respectively, ζ is the zeta potential and η_0 is the bulk solvent viscosity (8.9 $\times$ 10^{-4} Pa/s).[37] The

computed zeta potential for ethanolamine modified PMMA device was -0.45 mV and it was 131-fold less compared to O_2 plasma modified device.[60] The surface charge density, EOF and the zeta potential values further supports evidences for successful modification of PMMA with ethanolamine.

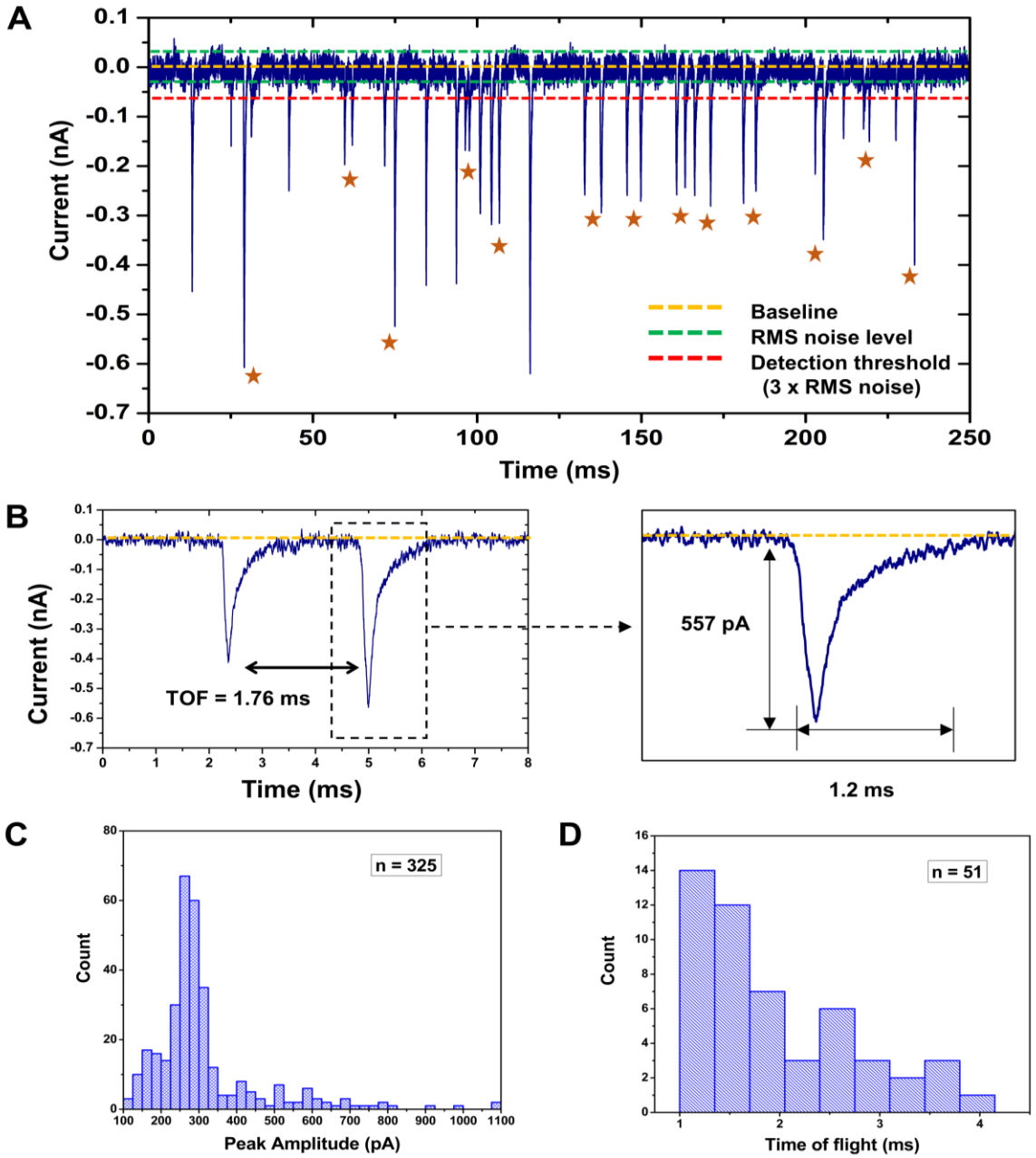


Figure 3.13. Translocation of 60 nt RNA through dual in-plane nanopore devices bonded at 170 psi. **(A)** 250 ms of the current transient amplitude signal obtained for 100 nM solution of 60 bases long RNA. The stars represent paired peaks. **(B)** An example peak pair obtained from the peak pair selection criteria. **(C)** Histogram of the current transient amplitudes for the 60 bases long RNA. **(D)** ToF distribution of 60 nt RNA.

3.3.7 RNA/rAMP Translocation through Ethanolamine-Modified PMMA Dual In-Plane Nanopore Devices

After the ethanolamine modification, RNA and rAMP translocation was repeated as previously described. First, a 100 nM solution 60 b RNA in NEBuffer 3 without Mg^{2+} was introduced to the nanopore device bonded at 170 psi and a potential of 2.5 V was applied across the nanopores. As shown in Figure 3.13.A, after the ethanolamine treatment the event frequency increased significantly compared to the O_2 plasma treated nanopore devices.

To identify peak pairs, which corresponded to a single-molecule travelling through both nanopores sequentially, we followed a selection criteria.[75] The first criterion was that the peak amplitude should be >3X the RMS noise of the open pore current. The RMS noise of the open pore current of 170-psi bonded devices was found to be 19.6 pA and therefore, only peaks with amplitudes >58.5 pA were considered as true 60 b RNA translocation events. The second criterion was the minimum time-of-flight, where the travel time between the two pores (time-of-flight, ToF) for the single-molecule should be greater than the dwell time (peak width) of each peak comprising the peak pair. The average dwell time of the peaks obtained for 60 b RNA was determined to be 0.92 ±0.38 ms. The third criterion was that the maximum ToF needed to be within 1.5 times the theoretical ToF. The theoretical ToF was calculated using the electrophoretic mobility values reported for fluorescently labelled oligonucleotides in PMMA/COC devices by Amarasekara *et al*.[60, 76] The maximum ToF was therefore determined to be 4.1 ms. An example of a peak pair obtained for 60 b RNA is shown in Figure 3.13.B. The peak amplitudes ranged from 0.10 – 0.8 nA (n ≥ 325) with an average peak amplitude of 311.75 ±137.49 pA (see Figure 3.13.C) and the ToF between the pores for 60 b RNA was 2.09 ±0.97 ms (n = 51; see Figure 3.13.D). In the 250 ms current trace shown in Figure 3.13.A, only 76.5% (26/34) of the peaks corresponded to peak pairs.

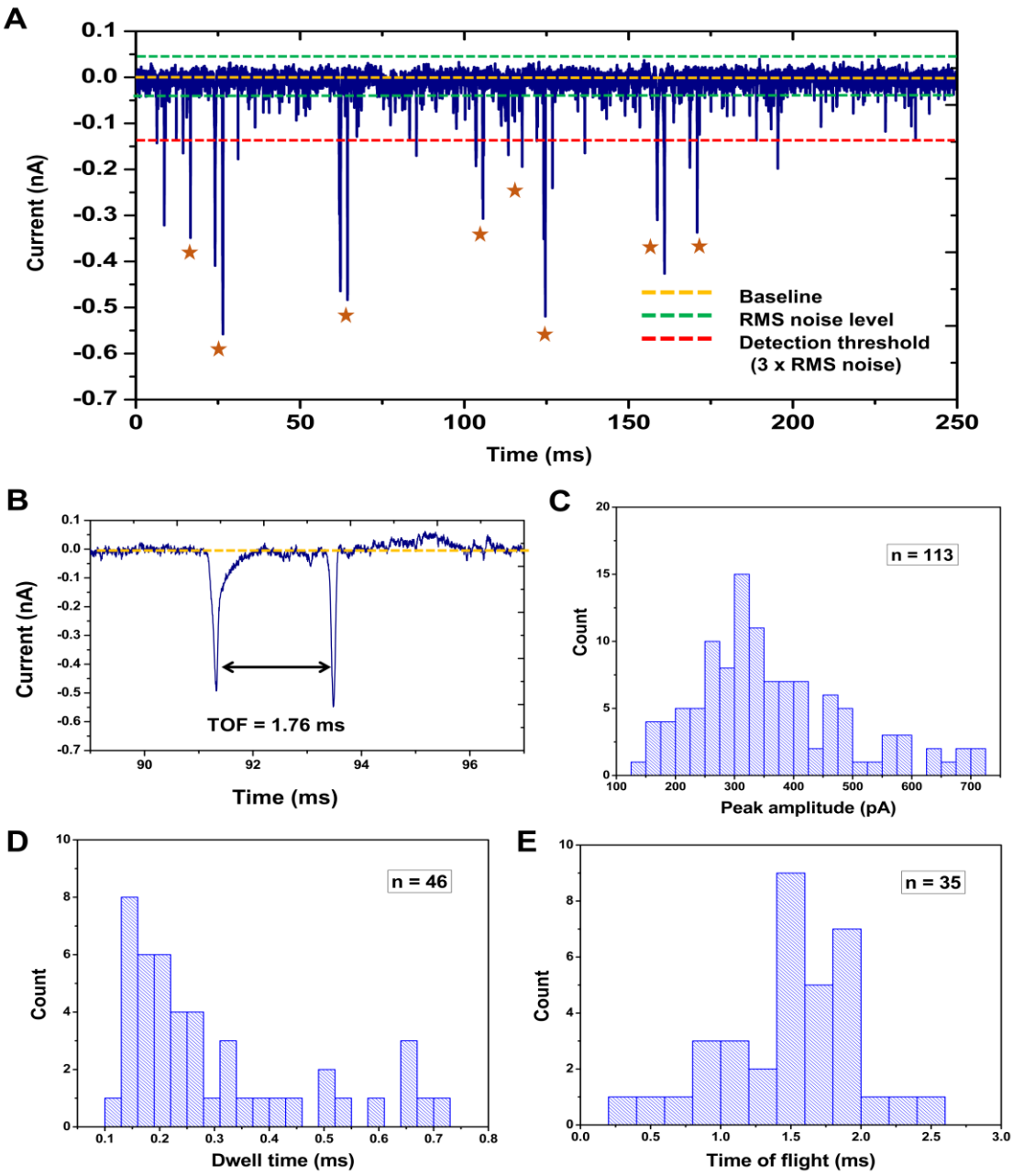


Figure 3.14. Translocation of EGFP mRNA (996 nt) through dual in-plane nanopore devices bonded at 110 psi. **(A)** 250 ms of the current transient amplitude signal obtained for 100 nM solution of EGFP mRNA. The stars represent paired peaks. **(B)** An example peak pair obtained from the peak pair selection criteria. **(C)** Histogram of the current transient amplitudes for the EGFP mRNA. **(D)** Histogram of the dwell times obtained for EGFP mRNA. **(E)** Histogram of the ToF values obtained for the EGFP mRNA.

Next, translocation of Enhanced Green Fluorescent Protein mRNA (EGFP mRNA; 996 nt) was conducted using dual in-plane nanopore devices bonded at 110 psi (nanopore depth – 22.3 nm) to accommodate the larger size of the molecule (R_g – ~21 nm). The open pore current of dual

nanopore devices bonded at 10 psi was ~48.5 nA with an RMS noise of 44.2 pA. Peak pairs were determined using the criteria described above (see Figure 3.14.A). Peaks with an amplitude greater 132.6 pA were considered as translocation events. An example peak pair is shown in Figure 3.14.B. The peak amplitude for EGFP RNA ranged from 0.15 – 0.75 nA with an average of 361.88 ±131.05 pA (n = 113; Figure 3.14.C). The dwell time and time of flight between the pores were 0.303 ±0.176 ms (n = 46; Figure 3.14.D) and 1.54 ±0.44 ms (n = 35; Figure 3.14.E).

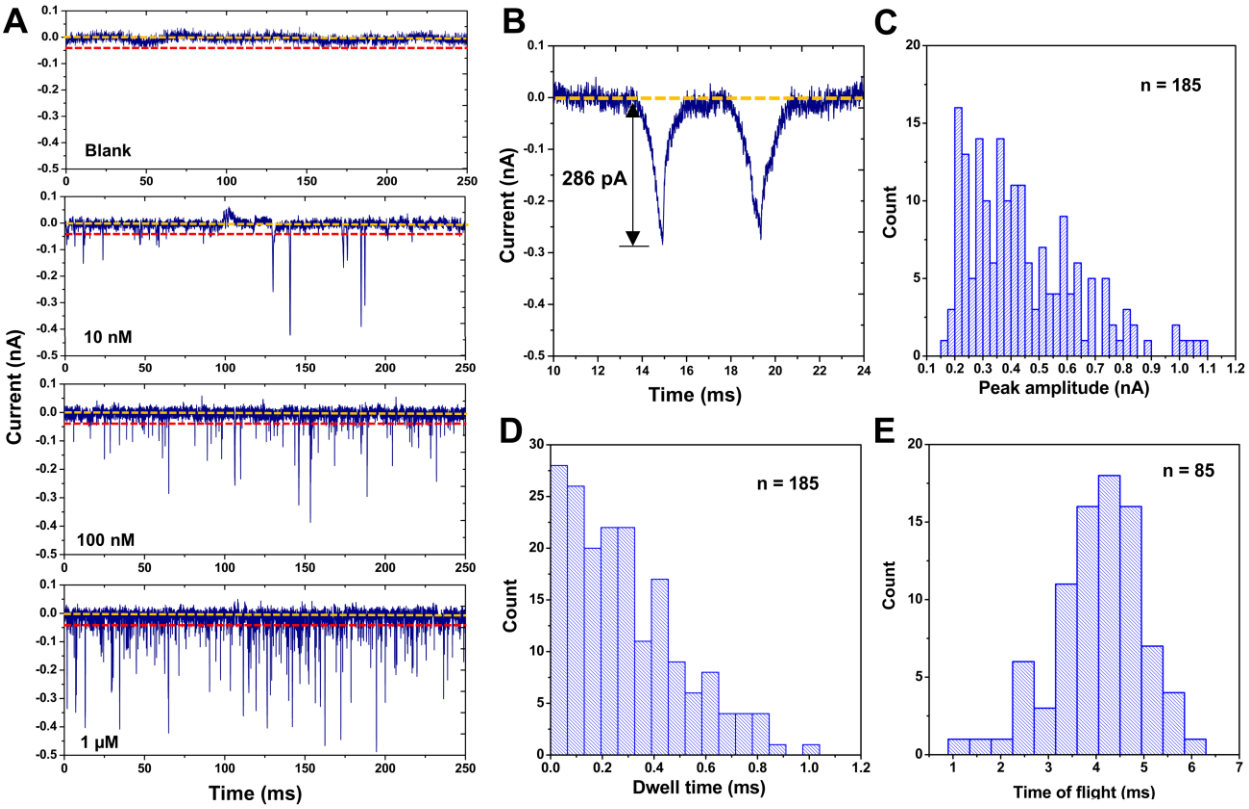


Figure 3.15. Translocation of rAMPs through dual in-plane nanopore devices assembled at 200 psi bonding pressure. **(A)** A 250 ms current transient trace of signal amplitudes obtained for a blank, 10 nM, 100 nM and 1 µM solutions of rAMP using dual in-plane nanopore devices. An increase in event frequency was observed with increasing concentration (R^2 = 0.9757). **(B)** An example peak pair selected using the peak pair criteria. **(C)** Peak amplitude distribution of rAMP events. **(D)** The dwell time distribution for rAMP events. **(E)** ToF distribution for rAMP.

Furthermore, we carried out experiments to detect and identify single rAMP molecules using the dual-in-plane nanopore devices bonded at 200 psi for 5 mins, which was used to create a smaller pore to accommodate the smaller size of the rAMP molecule. Figure 3.15.A shows the

current traces for blank, 10 nM, 100 nM and 1 µM solutions of rAMP. With increasing concentration, a linear increase in event frequency could be observed (R^2 = 0.9757). An example peak pair obtained for rAMP translocation between the two pores in series identified using the aforementioned peak pair selection criteria is shown in Figure 3.15.B. Peaks with amplitude greater than 43.2 pA (open pore current -~16.8 nA; RMS noise – 14.4 pA) were considered as peaks. The average current blockage amplitude for rAMPs was 425.89 ±175.89 pA (n = 185; Figure 3.15.C). The average dwell time and ToF between the two in-plane nanopores were 0.31 ±0.26 (n = 185; Figure 3.15.D) and 4.14 ±0.97 ms (n = 85; Figure 3.15.E), respectively. The maximum ToF (5.7 ms) was calculated using the mobility values we recently reported for ATTO-532 labelled rAMP.[57] The percentage of peak pairs was 82%, 63% and 64% for 10 nM, 100 nM and 1 µM rAMP solutions, respectively. For measurements performed with dual in-plane nanopore devices bonded at 110 psi, no current transients were observed (data not shown).

Unpaired peaks found in this work may be attributed to the relatively large pore size and the irregular shape (shallow U-shape) of the in-plane nanopores, which resulted in a large variation in the peak amplitude, as seen in Figures 3.13 – 3.15. Thus, some peaks may not produce sufficient amplitudes to be selected as paired events and account for the unpaired peaks. Also, there may be cases where entry of a molecule into the first nanopore before the previous molecule left the second nanopore may lead to unpaired events, which can be seen in the decreased percentage of the paired events with increasing concentration. In previous work by Langecker *et al*.[75] where double-stranded (ds)-DNA was detected by stacked nanopores with diameters of 23 and 28 nm, 94% of detected peaks could unambiguously be assigned to the translocating dsDNA.

3.4 Conclusion

In this chapter we have shown the ability to tailor the size of an in-plane nanopore by TFB using different pressures during an assembly step already required for thermoplastic nanofluidic chip production. We noticed that the original size and shape of the in-plane nanopore changes with different bonding pressures as measured using AFM and SEM. Further, COMSOL simulations with varying pore widths gave the approximate conductance observed in the PMMA and COP dual in-plane nanopore devices connected using a 5 μm long nanochannel. Further, λ-DNA was electrokinetically driven by a DC voltage through the dual nanopore devices bonded at 110, 170 and 200 psi. The transient current amplitudes were seen to increase with higher TFB pressures. Therefore, thermoplastic nanofluidic devices allow for tuning nanostructure size to accommodate a given application by simply altering device assembly conditions without requiring direct FIB milling of a new Si master. This was further demonstrated using EGFP mRNA, 60 bases long RNA, and rAMPs using devices bonded at 110, 170 and 200 psi bonding pressure, respectively. We should note as well that we could produce >1000 resin stamps from a single Si master and from each resin stamp, >20 nanofluidic devices could be produced with a success rate of producing functional devices following TFB of the COC cover plate to the PMMA or COP substrate >90%.[33, 37, 61] In addition, for high scale production, we mount the resin stamp produced from Si master onto a stainless steel metal block serving as a prototype model to injection mold more than 1000 devices from each stamp.

The dual in-plane nanopores with their associated fluidic network and the fact that they were fabricated in thermoplastics in a single step using NIL makes it feasible to integrate this sensing technology into other nanofluidic components for the label-free identification of biomolecules. For example, we are currently developing a chip-based single-molecule exo-sequencing method, termed exonuclease time-of-flight (XToF).[57, 77-79] This method involves a solid-phase enzymatic

reactor coupled to a nano-flight tube that contains dual in-plane nanopores to measure free nucleotides' ToFs. Recently, we reported the use of solid-phase XRN1 reactions to sequentially produce rNMPs (5' → 3' direction).[80] Previous work from our group also demonstrated the identification of labelled rNMPs via their molecular-dependent electrophoretic mobility (*i.e*., ToF) in thermoplastic nanochannels; we were able to achieve ToF identification accuracies >99%.[57]

3.5 References

1. Branton, D.; Deamer, D.; Marziali, A.; Bayley, H.; Benner, S.; Butler, T.; Di Ventra, M.; Garaj, S.; Hibbs, A.; Huang, X., Xinsheng Sean. *L., Mastrangelo, CH, Meller, A., Oliver, JS, Pershin, YV, Ramsey, JM, Riehn, R., Soni, GV, Tabard-Cossa, V., Wanunu, M., Wiggin, M., Schloss, JA: The potential and challenges of nanopore sequencing. Nat. Biotechnol* **2008,** *26* (10), 1146-1153.

2. Wanunu, M.; Sutin, J.; McNally, B.; Chow, A.; Meller, A., DNA translocation governed by interactions with solid-state nanopores. *Biophysical journal* **2008,** *95* (10), 4716-4725.

3. Dekker, C., Solid-state nanopores. *Nature nanotechnology* **2007,** *2* (4), 209-215.

4. Venkatesan, B. M.; Bashir, R., Nanopore sensors for nucleic acid analysis. *Nature nanotechnology* **2011,** *6* (10), 615-624.

5. Smeets, R. M.; Keyser, U. F.; Krapf, D.; Wu, M.-Y.; Dekker, N. H.; Dekker, C., Salt dependence of ion transport and DNA translocation through solid-state nanopores. *Nano letters* **2006,** *6* (1), 89-95.

6. Carlsen, A. T.; Zahid, O. K.; Ruzicka, J. A.; Taylor, E. W.; Hall, A. R., Selective detection and quantification of modified DNA with solid-state nanopores. *Nano letters* **2014,** *14* (10), 5488-5492.

7. Kowalczyk, S. W.; Hall, A. R.; Dekker, C., Detection of local protein structures along DNA using solid-state nanopores. *Nano letters* **2010,** *10* (1), 324-328.

8. Plesa, C.; Ruitenberg, J. W.; Witteveen, M. J.; Dekker, C., Detection of individual proteins bound along DNA using solid-state nanopores. *Nano letters* **2015,** *15* (5), 3153-3158.

9. Smeets, R.; Kowalczyk, S. W.; Hall, A.; Dekker, N.; Dekker, C., Translocation of RecA-coated double-stranded DNA through solid-state nanopores. *Nano letters* **2009,** *9* (9), 3089-3095.

10. Hornblower, B.; Coombs, A.; Whitaker, R. D.; Kolomeisky, A.; Picone, S. J.; Meller, A.; Akeson, M., Single-molecule analysis of DNA-protein complexes using nanopores. *Nature Methods* **2007,** *4* (4), 315-317.

11. Tabard-Cossa, V.; Wiggin, M.; Trivedi, D.; Jetha, N. N.; Dwyer, J. R.; Marziali, A., Single-molecule bonds characterized by solid-state nanopore force spectroscopy. *ACS nano* **2009,** *3* (10), 3009-3014.

12. Cherf, G. M.; Lieberman, K. R.; Rashid, H.; Lam, C. E.; Karplus, K.; Akeson, M., Automated forward and reverse ratcheting of DNA in a nanopore at 5-Å precision. *Nature biotechnology* **2012,** *30* (4), 344-348.

13. Ricroch, A. E., Assessment of GE food safety using '-omics' techniques and long-term animal feeding studies. *New biotechnology* **2013,** *30* (4), 349-354.

14. Song, L.; Hobaugh, M. R.; Shustak, C.; Cheley, S.; Bayley, H.; Gouaux, J. E., Structure of staphylococcal α-hemolysin, a heptameric transmembrane pore. *Science* **1996,** *274* (5294), 1859-1865.

15. Faller, M.; Niederweis, M.; Schulz, G. E., The structure of a mycobacterial outer-membrane channel. *Science* **2004,** *303* (5661), 1189-1192.

16. ávan Loo, N., DNA nanopore translocation in glutamate solutions. *Nanoscale* **2015,** *7* (32), 13605-13609.

17. Storm, A.; Chen, J.; Ling, X.; Zandbergen, H.; Dekker, C., Fabrication of solid-state nanopores with single-nanometre precision. *Nature materials* **2003,** *2* (8), 537-540.

18. Li, J.; Stein, D.; McMullan, C.; Branton, D.; Aziz, M. J.; Golovchenko, J. A., Ion-beam sculpting at nanometre length scales. *Nature* **2001,** *412* (6843), 166-169.

19. Wanunu, M.; Meller, A., Chemically modified solid-state nanopores. *Nano letters* **2007,** *7* (6), 1580-1585.

20. Gierak, J.; Madouri, A.; Biance, A.; Bourhis, E.; Patriarche, G.; Ulysse, C.; Lucot, D.; Lafosse, X.; Auvray, L.; Bruchhaus, L., Sub-5 nm FIB direct patterning of nanodevices. *Microelectronic engineering* **2007,** *84* (5-8), 779-783.

21. Yang, J.; Ferranti, D. C.; Stern, L. A.; Sanford, C. A.; Huang, J.; Ren, Z.; Qin, L.-C.; Hall, A. R., Rapid and precise scanning helium ion microscope milling of solid-state nanopores for biomolecule detection. *Nanotechnology* **2011,** *22* (28), 285310.

22. Kwok, H.; Briggs, K.; Tabard-Cossa, V., Nanopore fabrication by controlled dielectric breakdown. *PloS one* **2014,** *9* (3), e92880.

23. Stein, D. M.; McMullan, C. J.; Li, J.; Golovchenko, J. A., Feedback-controlled ion beam sculpting apparatus. *Review of Scientific Instruments* **2004,** *75* (4), 900-905.

24. Asghar, W.; Ilyas, A.; Billo, J. A.; Iqbal, S. M., Shrinking of solid-state nanopores by direct thermal heating. *Nanoscale research letters* **2011,** *6* (1), 372.

25. Schiedt, B.; Auvray, L.; Bacri, L.; Oukhaled, G.; Madouri, A.; Bourhis, E.; Patriarche, G.; Pelta, J.; Jede, R.; Gierak, J., Direct FIB fabrication and integration of “single nanopore devices” for the manipulation of macromolecules. *Microelectronic engineering* **2010,** *87* (5-8), 1300-1303.

26. Harms, Z. D.; Haywood, D. G.; Kneller, A. R.; Selzer, L.; Zlotnick, A.; Jacobson, S. C., Single-particle electrophoresis in nanochannels. *Analytical chemistry* **2015,** *87* (1), 699-705.

27. Henley, W. H.; Dennis, P. J.; Ramsey, J. M., Fabrication of microfluidic devices containing patterned microwell arrays. *Analytical chemistry* **2012,** *84* (3), 1776-1780.

28. Menard, L. D.; Ramsey, J. M., Fabrication of sub-5 nm nanochannels in insulating substrates using focused ion beam milling. *Nano letters* **2011,** *11* (2), 512-517.

29. Kondylis, P.; Zhou, J.; Harms, Z. D.; Kneller, A. R.; Lee, L. S.; Zlotnick, A.; Jacobson, S. C., Nanofluidic Devices with 8 Pores in Series for Real-Time, Resistive-Pulse Analysis of Hepatitis B Virus Capsid Assembly. *Analytical Chemistry* **2017,** *89*, 4855-4862.

30. Harms, Z. D.; Mogensen, K. B.; Nunes, P. S.; Zhou, K.; Hildenbrand, B. W.; Mitra, I.; Tan, Z.; Zlotnick, A.; Kutter, J. P.; Jacobson, S. C., Nanofluidic Devices with Two Pores in Series for Resistive-Pulse Sensing of Single Virus Capsids. *Analytical Chemistry* **2011,** *83*, 9573-9578.

31. Harms, Z. D.; Selzer, L.; Zlotnick, A.; Jacobson, S. C., Monitoring Assembly of Virus Capsids with Nanofluidic Devices. *ACS Nano* **2015,** *9*, 9087-9096.

32. Zhou, K.; Li, L.; Tan, Z.; Zlotnick, A.; Jacobson, S. C., Characterization of Hepatitis B Virus Capsids by Resistive-Pulse Sensing. *Journal of the American Chemical Society* **2011,** *133*, 1618-1621.

33. Chantiwas, R.; Park, S.; Soper, S. A.; Kim, B. C.; Takayama, S.; Sunkara, V.; Hwang, H.; Cho, Y. K., Flexible fabrication and applications of polymer nanochannels and nanoslits. *Chemical Society Reviews* **2011,** *40* (7), 3677-3702.

34. Choi, J.; Lee, C. C.; Park, S., Scalable fabrication of sub-10 nm polymer nanopores for DNA analysis. *Microsystems & Nanoengineering* **2019,** *5* (1), 12.

35. O'Neil, C. E.; Taylor, S.; Ratnayake, K.; Pullagurla, S.; Singh, V.; Soper, S. A., Characterization of activated cyclic olefin copolymer: effects of ethylene/norbornene content on the physiochemical properties. *Analyst* **2016,** *141* (24), 6521-6532.

36. Jia, Z.; Choi, J.; Park, S., Surface Charge Density-Dependent DNA Capture through Polymer Planar Nanopores. *ACS applied materials & interfaces* **2018,** *10* (47), 40927-40937.

37. Uba, F. I.; Hu, B.; Weerakoon-Ratnayake, K.; Oliver-Calixte, N.; Soper, S. A., High process yield rates of thermoplastic nanofluidic devices using a hybrid thermal assembly technique. *Lab on a Chip* **2015,** *15* (4), 1038-1049.

38. Wang, Y.; Liang, X.; Liang, Y.; Chou, S. Y., Sub-10-nm wide trench, line, and hole fabrication using pressed self-perfection. *Nano letters* **2008,** *8* (7), 1986-1990.

39. Choi, J.; Farshchian, B.; Kim, J.; Park, S., Fabrication of perforated micro/nanopore membranes via a combination of nanoimprint lithography and pressed self-perfection process for size reduction. *Journal of Nanoscience and Nanotechnology* **2013,** *13* (6), 4129-4133.

40. Siwy, Z.; Trofin, L.; Kohli, P.; Baker, L. A.; Trautmann, C.; Martin, C. R., Protein biosensors based on biofunctionalized conical gold nanotubes. *Journal of the American Chemical Society* **2005,** *127* (14), 5000-5001.

41. Martin, C. R.; Nishizawa, M.; Jirage, K.; Kang, M., Investigations of the transport properties of gold nanotubule membranes. *The Journal of Physical Chemistry B* **2001,** *105* (10), 1925-1934.

42. Siwy, Z.; Heins, E.; Harrell, C. C.; Kohli, P.; Martin, C. R., Conical-nanotube ion-current rectifiers: the role of surface charge. *Journal of the American Chemical Society* **2004,** *126* (35), 10850-10851.

43. Harrell, C. C.; Kohli, P.; Siwy, Z.; Martin, C. R., DNA− nanotube artificial ion channels. *Journal of the American Chemical Society* **2004,** *126* (48), 15646-15647.

44. Maekawa, Y.; Suzuki, Y.; Maeyama, K.; Yonezawa, N.; Yoshida, M., Visualization of chemical modification of pore internal surfaces using fluorescence microscopy. *Chemistry letters* **2004,** *33* (2), 150-151.

45. Maekawa, Y.; Suzuki, Y.; Maeyama, K.; Yonezawa, N.; Yoshida, M., Chemical modification of the internal surfaces of cylindrical pores of submicrometer size in poly (ethylene terephthalate). *Langmuir* **2006,** *22* (6), 2832-2837.

46. Li, J.; Maekawa, Y.; Yamaki, T.; Yoshida, M., Chemical modification of a poly (ethylene terephthalate) surface by the selective alkylation of acid salts. *Macromolecular Chemistry and Physics* **2002,** *203* (17), 2470-2474.

47. Pasternak, C.; Alder, G.; Apel, P. Y.; Bashford, C.; Edmonds, D.; Korchev, Y.; Lev, A.; Lowe, G.; Milovanovich, M.; Pitt, C., Nuclear track-etched filters as model pores for biological membranes. *Radiation measurements* **1995,** *1* (25), 675-683.

48. Marchand-Brynaert, J.; Deldime, M.; Dupont, I.; Dewez, J.-L.; Schneider, Y.-J., Surface functionalization of poly (ethylene terephthalate) film and membrane by controlled wet chemistry: chemical characterization of carboxylated surfaces. *Journal of colloid and interface science* **1995,** *173* (1), 236-244.

49. Papra, A.; Hicke, H. G.; Paul, D., Synthesis of peptides onto the surface of poly (ethylene terephthalate) particle track membranes. *Journal of applied polymer science* **1999,** *74* (7), 1669-1674.

50. Makamba, H.; Kim, J. H.; Lim, K.; Park, N.; Hahn, J. H., Surface modification of poly (dimethylsiloxane) microchannels. *Electrophoresis* **2003,** *24* (21), 3607-3619.

51. Liu, Y.; Fanguy, J. C.; Bledsoe, J. M.; Henry, C. S., Dynamic coating using polyelectrolyte multilayers for chemical control of electroosmotic flow in capillary electrophoresis microchips. *Analytical chemistry* **2000,** *72* (24), 5939-5944.

52. Henry, A. C.; Tutt, T. J.; Galloway, M.; Davidson, Y. Y.; McWhorter, C. S.; Soper, S. A.; McCarley, R. L., Surface modification of poly (methyl methacrylate) used in the fabrication of microanalytical devices. *Analytical chemistry* **2000,** *72* (21), 5331-5337.

53. Johnson, T. J.; Ross, D.; Gaitan, M.; Locascio, L. E., Laser modification of preformed polymer microchannels: application to reduce band broadening around turns subject to electrokinetic flow. *Analytical Chemistry* **2001,** *73* (15), 3656-3661.

54. Henry, A. C.; Waddell, E. A.; Shreiner, R.; Locascio, L. E., Control of electroosmotic flow in laser-ablated and chemically modified hot imprinted poly (ethylene terephthalate glycol) microchannels. *Electrophoresis* **2002,** *23* (5), 791-798.

55. Uba, F. I.; Pullagurla, S. R.; Sirasunthorn, N.; Wu, J.; Park, S.; Chantiwas, R.; Cho, Y.-K.; Shin, H.; Soper, S. A., Surface charge, electroosmotic flow and DNA extension in chemically modified thermoplastic nanoslits and nanochannels. *Analyst* **2015,** *140* (1), 113-126.

56. Ali, M.; Schiedt, B.; Healy, K.; Neumann, R.; Ensinger, W., Modifying the surface charge of single track-etched conical nanopores in polyimide. *Nanotechnology* **2008,** *19* (8), 085713.

57. Amarasekara, C. A.; Rathnayaka, C.; Athapattu, U. S.; Zhang, L.; Choi, J.; Park, S.; Nagel, A. C.; Soper, S. A., Electrokinetic identification of ribonucleotide monophosphates (rNMPs) using thermoplastic nanochannels. *Journal of Chromatography A* **2021,** *1638*, 461892.

58. Wei, S.; Vaidya, B.; Patel, A. B.; Soper, S. A.; McCarley, R. L., Photochemically patterned poly (methyl methacrylate) surfaces used in the fabrication of microanalytical devices. *The Journal of Physical Chemistry B* **2005,** *109* (35), 16988-16996.

59. Jackson, J. M.; Witek, M. A.; Hupert, M. L.; Brady, C.; Pullagurla, S.; Kamande, J.; Aufforth, R. D.; Tignanelli, C. J.; Torphy, R. J.; Yeh, J. J., UV activation of polymeric high aspect ratio microstructures: ramifications in antibody surface loading for circulating tumor cell selection. *Lab on a Chip* **2014,** *14* (1), 106-117.

60. Amarasekara, C. A.; Athapattu, U. S.; Rathnayaka, C.; Choi, J.; Park, S.; Soper, S. A., Open-tubular nanoelectrochromatography (OT-NEC): gel-free separation of single stranded DNAs (ssDNAs) in thermoplastic nanochannels. *Electrophoresis* **2020,** *41* (18-19), 1627-1640.

61. Wu, J. H.; Chantiwas, R.; Amirsadeghi, A.; Soper, S. A.; Park, S., Complete plastic nanofluidic devices for DNA analysis via direct imprinting with polymer stamps. *Lab on a Chip* **2011,** *11* (17), 2984-2989.

62. Jia, Z.; Choi, J.; Park, S., Selection of UV-resins for nanostructured molds for thermal-NIL. *Nanotechnology* **2018,** *29* (36), 365302.

63. Miller, A.; Lawton, E.; Balwit, J., Effect of chemical structure of vinyl polymers on crosslinking and degradation by ionizing radiation. *JPoSc* **1954,** *14* (77), 503-504.

64. Mahmood Raouf, R.; Abdul Wahab, Z.; Azowa Ibrahim, N.; Abidin Talib, Z.; Chieng, B. W., Transparent blend of poly (methylmethacrylate)/cellulose acetate butyrate for the protection from ultraviolet. *Polymers* **2016,** *8* (4), 128.

65. Dole, M., *The Radiation Chemistry of Macromolecules: Volume II*. Elsevier: 2013; Vol. 2.

66. Jeong, J. H.; Choi, Y. S.; Shin, Y. J.; Lee, J. J.; Park, K. T.; Lee, E. S.; Lee, S. R., Flow behavior at the embossing stage of nanoimprint lithography. *Fibers and Polymers* **2002,** *3* (3), 113-119.

67. Heyderman, L. J.; Schift, H.; David, C.; Gobrecht, J.; Schweizer, T., Flow behaviour of thin polymer films used for hot embossing lithography. *Microelectronic Engineering* **2000,** *54* (3-4), 229-245.

68. Fryer, D. S.; Peters, R. D.; Kim, E. J.; Tomaszewski, J. E.; de Pablo, J. J.; Nealey, P. F.; White, C. C.; Wu, W.-l., Dependence of the Glass Transition Temperature of Polymer Films on Interfacial Energy and Thickness. *Macromolecules* **2001,** *34* (16), 5627-5634.

69. Mirabella, F. M., *Internal reflection spectroscopy: theory and applications*. CRC Press: 1992; Vol. 15.

70. McCann, R.; Hughes, C.; Bagga, K.; Stalcup, A.; Vázquez, M.; Brabazon, D., Pulsed laser deposition of plasmonic nanostructured gold on flexible transparent polymers at atmospheric pressure. *Journal of Physics D: Applied Physics* **2017,** *50* (24), 245303.

71. Vesel, A.; Mozetic, M., Surface modification and ageing of PMMA polymer by oxygen plasma treatment. *Vacuum* **2012,** *86* (6), 634-637.

72. Roy, S.; Yue, C.; Lam, Y.; Wang, Z.; Hu, H., Surface analysis, hydrophilic enhancement, ageing behavior and flow in plasma modified cyclic olefin copolymer (COC)-based microfluidic devices. *Sensors and Actuators B: Chemical* **2010,** *150* (2), 537-549.

73. Meisenberg, G.; Simmons, W. H., *Principles of medical biochemistry*. Mosby Elsevier: 2006.

74. Huang, X.; Gordon, M. J.; Zare, R. N., Current-monitoring method for measuring the electroosmotic flow rate in capillary zone electrophoresis. *Analytical chemistry* **1988,** *60* (17), 1837-1838.

75. Langecker, M.; Pedone, D.; Simmel, F. C.; Rant, U., Electrophoretic time-of-flight measurements of single DNA molecules with two stacked nanopores. *Nano letters* **2011,** *11* (11), 5002-5007.

76. Stellwagen, N. C.; Gelfi, C.; Righetti, P. G., The free solution mobility of DNA. *Biopolymers: Original Research on Biomolecules* **1997,** *42* (6), 687-703.

77. Novak, B. R.; Moldovan, D.; Nikitopoulos, D. E.; Soper, S. A., Distinguishing Single DNA Nucleotides Based on Their Times of Flight Through Nanoslits: A Molecular Dynamics Simulation Study. *Journal of Physical Chemistry B* **2013,** *117* (12), 3271-3279.

78. Oliver-Calixte, N. J.; Uba, F. I.; Battle, K. N.; Weerakoon-Ratnayake, K. M.; Soper, S. A., Immobilization of lambda exonuclease onto polymer micropillar arrays for the solid-phase digestion of dsDNAs. *Analytical chemistry* **2014,** *86* (9), 4447-4454.

79. O'Neil, C.; Amarasekara, C. A.; Weerakoon-Ratnayake, K. M.; Gross, B.; Jia, Z.; Singh, V.; Park, S.; Soper, S. A., Electrokinetic transport properties of deoxynucleotide monophosphates (dNMPs) through thermoplastic nanochannels. *Analytica chimica acta* **2018,** *1027*, 67-75.

80. Athapattu, U. S.; Amarasekara, C. A.; Immel, J. R.; Bloom, S.; Barany, F.; Nagel, A. C.; Soper, S. A., Solid-phase XRN1 reactions for RNA cleavage: application in single-molecule sequencing. *Nucleic Acids Research* **2021,** *49*, 1-15.

Chapter 4. Exonuclease time-of-flight (XToF) device for single-molecule RNA sequencing

4.1 Introduction

Regulation of synthesis and maturation of RNAs are closely controlled processes that ultimately drive biological processes by shaping complex gene expression networks.[1] An in-depth understanding of the mechanism and principles of these complex processes is essential for interpreting the functional elements of the genome, revealing the molecular constituents of cells and tissues, and to better understand complex diseases such as cancer. RNA sequencing helps in expanding our understanding of these processes by providing an increasingly fuller knowledge of both the quantitative and qualitative aspects of RNA biology in both eukaryotes and prokaryotes.[2] Moreover, RNA sequencing provides ways for transcription site mapping, strand-specific measurements, gene fusion detection, small RNA characterization, and detection of splicing events, as well as the role of RNA in disease pathogenesis. Understanding the role of RNA (miRNAs, mRNAs, long non-coding RNAs, transfer RNAs, and ribosomal RNA) in disease pathogenesis can also provide important diagnostic markers to assist in managing a plethora of diseases. As a first step in understanding the biology of RNA, sequencing has been designated as an important tool.

The technological advances realized by Next Generation Sequencing (NGS) have been instrumental in generating fundamental discoveries in the field of transcriptomics and RNA biology in general, especially in terms of understanding the complex nature of the transcriptome and their role in various human disease states.[3-9] The success of NGS to accumulate such data has been driven by the extraordinary figures-of-merit associated with NGS compared to Sanger sequencing in terms of throughput (8 days to sequence an entire human genome per machine) and cost ($1,000 per genome) without sacrificing base call accuracy.

While NGS sequencing-by-synthesis platforms have streamlined the acquisition of sequencing data, there are still some challenges that need to be addressed so that full implementation of nucleic acid sequencing (both DNA and RNA) can be realized in a clinical environment, including: (1) Relaxing the extensive protocols required for library preparation; (2) reducing the heavy capital outlay associated with NGS machines; (3) eliminating the need for PCR. During PCR amplification, important nucleotide modifications can be lost, especially in RNA, where modifications are much more numerous than in DNA; (4) improving upon the short reads (~200 bp); and (5) further reducing the time of sequencing. Thus, newer technologies that can address these challenges will assist in making sequencing more distributive to enable clinical information to be secured at the point of use, for example in physicians' offices.

Single-molecule sequencing (SMS), which was first proposed by the Keller group in the late 1980's,[10] has many attributes that can address the challenges discussed above, such as simplifying or eliminating the need for library preparation, obviating the necessity for PCR or even reverse transcription required for RNA sequencing, and provide long reads that can be on the order of 10's of kbps.[11-33] Some examples of SMS platforms commercially available include Helicos,[34] Genia,[35-36] Pacific Biosystems,[37-38] and Oxford Nanopore.[39-44] Genia and Oxford Nanopore technologies are particularly interesting because they use electrical readout techniques, significantly simplifying the hardware requirements of the platform as well as eliminating the need for fluorescent reporters.

SMS based on nanopores have been compelling due to their relative simplicity in operation. But challenges do exist:[45-46] (1) The translocation times through the pore are fairly short (1-20 μs per nucleotide) – this issue has been somewhat mitigated by using the ratcheting action of a polymerase.[47-49] (2) The readout resolution requires a pore thickness equal to or less than the single base spacing of DNA molecules, ~0.34 nm. Because the thickness of both synthetic and

α-hemolysin pores is much larger (5-15 nm) than 0.34 nm, multiple bases simultaneously reside within the pore (5-6 nucleotides). The evolution of the MspA pore has somewhat reduced this issue.[50-51] But, even if nanopores could be fabricated with the prerequisite thickness, the effective electric field read region extends approximately 1 pore diameter unit on either side of the pore.[52] In addition, thinner pores result in shorter translocation times, which can be an issues due to the limited bandwidth of the readout electronics as noted above.[18] (3) The call accuracy of nanopore sequencers (single read accuracy of 92% for MinION)[53] is significantly less than conventional NGS platforms (read accuracy >99%).[54]

Readout resolution limitations can be mitigated if nucleotides are physically separated from each other while maintaining their original order following clipping from the nucleic acid, which can be accomplished through the use of an exonuclease enzyme.[55] This was first demonstrated using a processive exonuclease, which sequentially cleaves individual deoxynucleotide monophosphates (dNMPs) from an intact DNA fragment and directing these bases through an α-hemolysin nanopore fitted with a cyclodextrin collar.[56] Furthermore, the exo sequencing of RNA was demonstrated by Ayub *et al.* using polynucleotide phosphorylase (PNPase) enzyme.[57] PNPase is a 3' to 5' end processive exoribonuclease enzyme that sequentially produces ribonucleotide diphosphates (rNDPs) from an intact RNA in the presence of Mg^{2+} and inorganic phosphate (P_i). The resulting rNDPs could be identified using the current transient amplitudes generated while translocating through an α-hemolysin nanopore. Unfortunately with the exo sequencing approach, issues arise due to misordering of the mononucleotides once cleaved from the intact nucleic acid strand and also, injection efficiency of the released mononucleotides into the pore.[58]

We propose a highly innovative, low-cost SMS sequencing method, which employs a novel sequencing-by-subtraction strategy (Exo-Seq) that uses a nanofluidic device consisting of an

exoribonuclease immobilized onto a support pillar for the sequential disassembly of unamplified RNA molecules into their constituent ribonucleotide monophosphates (rNMPs). This novel RNA exo-sequencing method, coined as the exonuclease time-of-flight (XToF), consists of a nanoscale enzymatic reactor that is ~500 nm in diameter at the center of the nanofluidic XToF device with covalently attached XRN1 enzyme. Previously, we reported the successful covalent attachment of the XRN1 enzyme onto a UV/O_3 activated thermoplastic surface using EDC/NHS coupling chemistry.[59] The covalently attached XRN1 enzyme was able to retain its catalytic activity upon immobilization and demonstrated a higher clipping rate (26 ±5 nt s^{-1}) and processivity (>10.5 Kb) at pH 7.9 and 25^0C compared to its solution-based counterpart. Moreover, we showed the ability of XRN1 to digest through RNA secondary structure and also, modified nucleotides demonstrating the potential ability to use our proposed method for the identification of epitranscriptomically-modified nucleotides. It is envisioned that our technology will require very low amounts of RNA (<1 ng) while also providing unprecedented single-read base-call accuracy (>99%) compared to other SMS platforms by integrating a multi-parameter single-molecule identification approach. In addition, XToF uses a chip consisting of micro- and nanofluidic channels fabricated in plastics via high production modes, such as nano-imprint lithography (NIL) or injection compression molding that will make the consumable aspect of the instrument low-cost in spite of the nanostructures it possesses.

4.2 Materials and methods

4.2.1 Reagents and materials

Reagents and materials were obtained from the following sources: polyethylene terephthalate (PET) with 250 µm thickness Goodfellow (Coraopolis, PA); PMMA from ePlastics (San Diego, CA); COC (8007) from TOPAS Advanced Polymers (Florence, KY); UV curable polyurethane

acrylate (PUA) resin from Chansang Co.; Molecular biology grade water from Thermo Fisher (Waltham, MA); SYTO82 dye from Life Technologies (Eugene, OR, USA); Cas9 RNA from TriLink Biotechnologies (San Diego, CA, USA); XRN1 enzyme and NE buffer 3 from New England Biolabs (Ipswich, MA, USA); MF319 developer (MicroChemicals); potassium hydroxide (KOH) pellets (Fisher Scientific, Waltham, MA); isopropyl alcohol (IPA; Sigma-Aldrich, St. Louis, MO); hydrofluoric acid (HF, Sigma-Aldrich); NOA72 (Norland Products, Neuchâtel, Switzerland); Si wafers (P/B, resistivity 5-10 Ωcm, orientation of (100), and 525 ± 25 μm thickness, WaferPro, Santa Clara, CA); 5" x 5" x 0.90" Antireflective Chrome Soda Lime photomask ordered from HTA photomask (San Jose, CA, USA).

4.2.2 XToF device fabrication and assembly

XToF nanofluidic devices were fabricated and assembled as previously reported.[60-61] Briefly, a-Si mold master was fabricated using photolithography, wet etching, and focused ion beam (FIB) milling. The structures on the Si mold master were transferred to the thermoplastic substrate via a combination of UV – nano-imprint lithography (UV-NIL) and thermal NIL.

4.2.3 COMSOL simulations

COMSOL simulations were performed using COMSOL Multiphysics v5.5 software prior to fabrication and experiments to confirm the feasibility of the XToF device design. For all experiments 0.1 M NaCl was used as the electrolyte to be consistent with the experimental conditions.

4.2.4 Atomic force microscopy (AFM)

To determine the depth of the nanochannels at each fabrication step, AFM (SPM HT-9700, Shimadzu Corporation, Kyoto, Japan) was used. Tapping mode with 0.5 Hz scanning rate using

an AFM tip with a tip radius <2 nm, half cone angle of 10^{o}, and a frequency of 300 kHz was used. All acquired images were analyzed using SPM manager v4.7.6.1 software.

4.2.5 Scanning electron microscopy (SEM)

SEM images of the resin stamp and the thermally imprinted XToF devices were taken using a Hitachi FlexSEM 1000 II (Hitachi Ltd., Tokyo, Japan). Secondary electron images of the resin stamp and imprinted devices were acquired using a 5.00 kV accelerating voltage and 6.00 mm working distance. All non-conductive samples were coated with a 5 nm thin Au layer prior to imaging using a Denton Desk II sputter coater. SEM images of the Si mold master were acquired using a Quanta 3D DualBeam FEI FIB-SEM. All SEM images were analyzed using instrument software and ImageJ.

4.2.6 RNA labeling and translocation through nanochannels

The DMD and Cas9 RNA were labeled to saturation with SYTO 82 as previously described in Chapter 2 (see section 2.2.11).

4.2.7 Selective immobilization of Cy3 labeled oligonucleotides/XRN1 enzyme within the nanoscale bioreactor

Selective immobilization of Cy3-labeled oligonucleotides around the bioreactor was carried out using a “one-step immobilization” reaction. For the one-step immobilization reaction, 50 nM of 3’-Cy3-labeled oligonucleotide that was 20 nt long and was introduced to one of the enzyme input/ output reservoirs along with 20 mg/mL EDC and PBS buffer (pH 7.4). A 5 V potential was applied to electrokinetically drive the dye solution around the bioreactor.

For selective immobilization of XRN1 enzyme, first a solution of 20 mg/mL EDC and 2 mg/mL NHS in 0.1 M MES buffer (pH 4.9) was introduced to the XToF device. After a 15 min incubation, the EDC/NHS solution was completely removed and XRN1 enzyme was introduced into the enzyme input/output reservoirs and driven electrokinetically by applying a voltage.

4.2.8 XRN1 enzyme kinetics due to nano-confinement

After covalent XRN1 immobilization around the bioreactor, SYTO 82 labeled Cas9 RNA was introduced to the RNA input channel and driven electrokinetically towards the nanoscale enzymatic reactor in NE buffer 3 without Mg^{2+}. After successful complexation of a single RNA molecule to an immobilized XRN1 enzyme at the nanoscale bioreactor, Mg^{2+} containing NE buffer 3 was introduced to activate the enzyme. The fluorescence intensity of the immobilized enzyme/labeled RNA complex was measured with respect to time to obtain enzyme kinetic data.

4.3 Results and discussion

4.3.1 Exonuclease time-of-flight (XToF) device: Mechanism of action

The novel exo-sequencing method for direct RNA sequencing is a nanofluidic chip-based method. The nanofluidic device is shown in the Figure 4.1.A, which is fabricated in thermoplastics, consists of a nanoscale enzymatic reactor that is ~500 nm in diameter at the middle of the device (see Figure 4.1.B). For this nanoscale enzymatic reactor, the XRN1 enzyme is covalently attached using previously described EDC/NHS coupling chemistry of an activated thermoplastic.[59, 62]

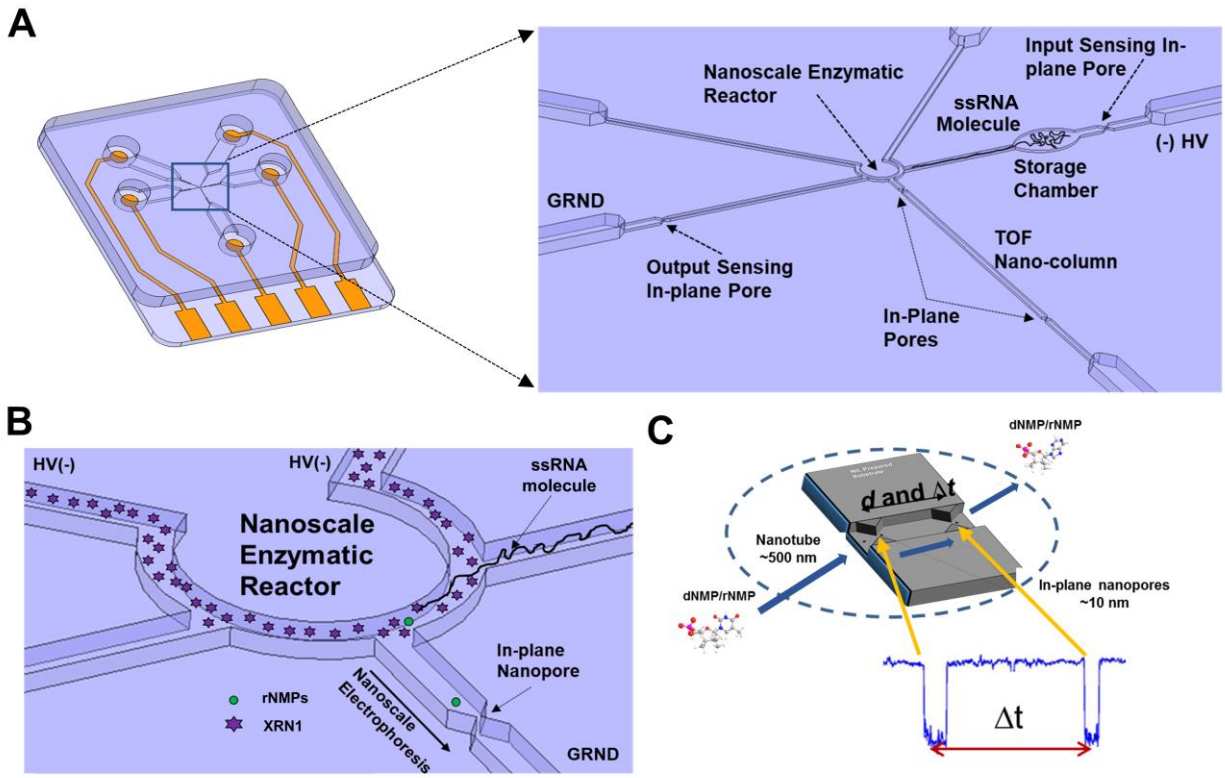


Figure 4.1. Exonuclease time-of-flight (XToF) device. **(A)** Schematic of the XToF device. **(B)** Nanoscale enzymatic reactor with immobilized XRN1 enzyme. **(C)** Identification of rNMPs using their molecular dependent time-of-flight between two in-plane nanopores. The time-of-flight is directly related to the electrophoretic mobility of the rNMP.

After immobilization of the enzyme, single-stranded RNA (ssRNA) molecules are electrokinetically introduced into the input channel and are electrokinetically driven towards the immobilized nanoscale enzymatic reactor (INER) in NE buffer 3 without the Mg^{2+} cofactor. The entropic trap/storage chamber at the entrance of the input nanochannel (see Figure 4.1.A) helps in making sure that only a single molecule of RNA is entering into the nanochannel at a given time.[63] The entropically trapped RNA molecules can be ejected from the trap by increasing the applied voltage. Once a single RNA molecule reaches the nanoscale enzymatic reactor, the 5' end of the RNA molecule enters into an active site of the immobilized enzyme and form a complex. Due to the absence of the cofactor, the RNA only complexes with the enzyme without digestion.[59] After the successful complexation of the RNA molecule, the potential applied between the input and output channels is removed, and a potential is applied across the enzyme input/output

channels and the nano-flight tube to introduce the Mg^{2+} cofactor into the nanoscale enzymatic reactor. Due to enzyme activation by the cofactor, it will initiate digestion of the complexed RNA molecule into its constituent rNMPs. The released rNMPs are then driven electrokinetically into and through the nano-flight tube. The nano-flight tube contains two in-plane nanopores, and when rNMPs travel through these nanopores, a current transient amplitude will be generated. The rNMPs will then be identified using nanoscale electrophoresis based on their travel time (time-of-flight) between two nanopores in series (one at the beginning of the flight tube and one at the end of the same flight tube) and the current transient amplitudes measured (see Figure 4.1.C).[64]

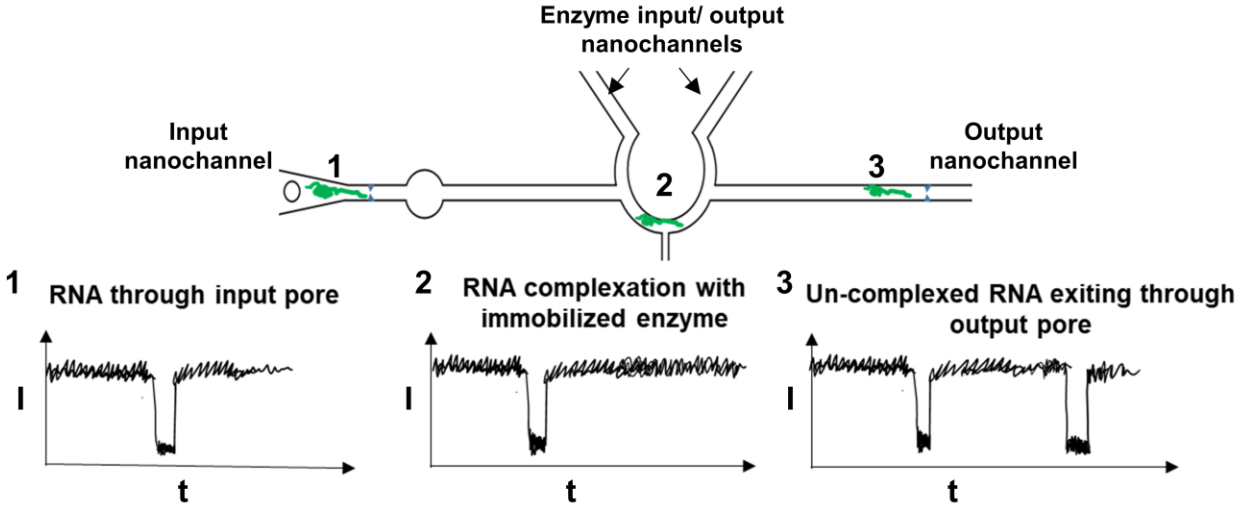


Figure 4.2. In-plane nanopores at the entrance of input and output nanochannels for the identification of entering RNA.

The XToF device will also contain two in-plane nanopores at the entrance of the input and the output nanochannels, as shown in Figure 4.2. Since the final version of XToF will use an unlabeled RNA molecule, we need a method to identify whether a single RNA molecule entered the XToF sensor, and whether it got complexed with an immobilized enzyme or exited through the output nanochannel without successful complexation. As shown in Figure 4.2, the input/output in-plane nanopores help to transduce the entry of a single RNA molecule into the nanochannel by generating a current transient amplitude. The entered RNA molecule will electrokinetically translocate towards the INER, and if it complexes with an immobilized enzyme, only a single peak

event would be observed (Figure 4.2). When a successful complexation is observed, the digestion of the complexed RNA molecule can be initiated by supplying the Mg^{2+} cofactor. If the complexation was unsuccessful, the exiting RNA molecule would result in a peak pair due to the translocation through input and output in-plane nanopores. In Chapter 3 we showed the identification of a 60 nt and EGFP (996 nt) RNA using dual in-plane nanopores fabricated in PMMA (see section 3.3.7). So, the transduction of single RNA molecules using resistive pulse sensing is feasible.

4.3.2 XToF device: design optimization

4.3.2.1 Microchannel/ micro-funnel design

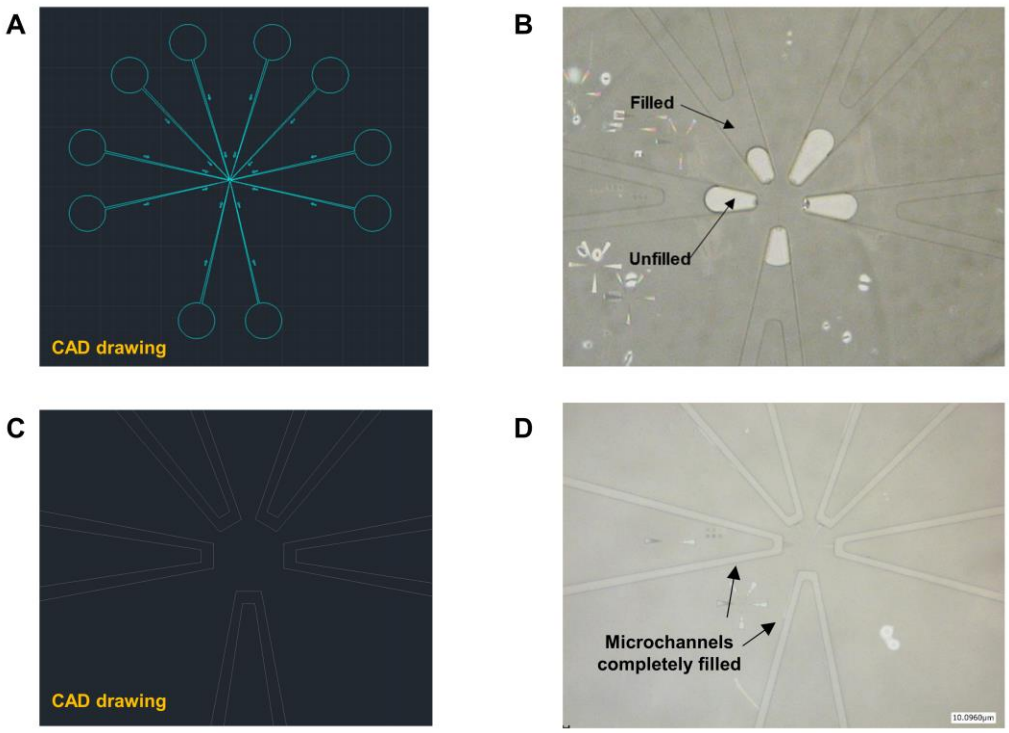


Figure 4.3. Access microchannels in the XToF device. **(A)** AutoCAD drawing of the ten access microchannels. **(B)** 20X optical image of the partially filled microchannels with buffer. **(C)** AutoCAD drawing of the modified microchannels. The shape of microchannel/ micro-funnel interface was changed to a square shape. **(D)** 10X optical image of the microchannels completely filled with buffer after the modification.

As shown in Figure 4.3.A, the XToF device contains ten access microchannels that are connected to five micro-funnels. In the first iteration of the XToF device, the microchannels did not fill

completely, and air bubbles were trapped at the microchannel/micro-funnel interface (see Figure 4.3.B). To overcome this problem, the microchannel shape at the microchannel/funnel interface was changed, as shown in Figure 4.3.C. This prevented air bubble trapping, and the microchannels could be filled continuously (see Figure 4.3.D).

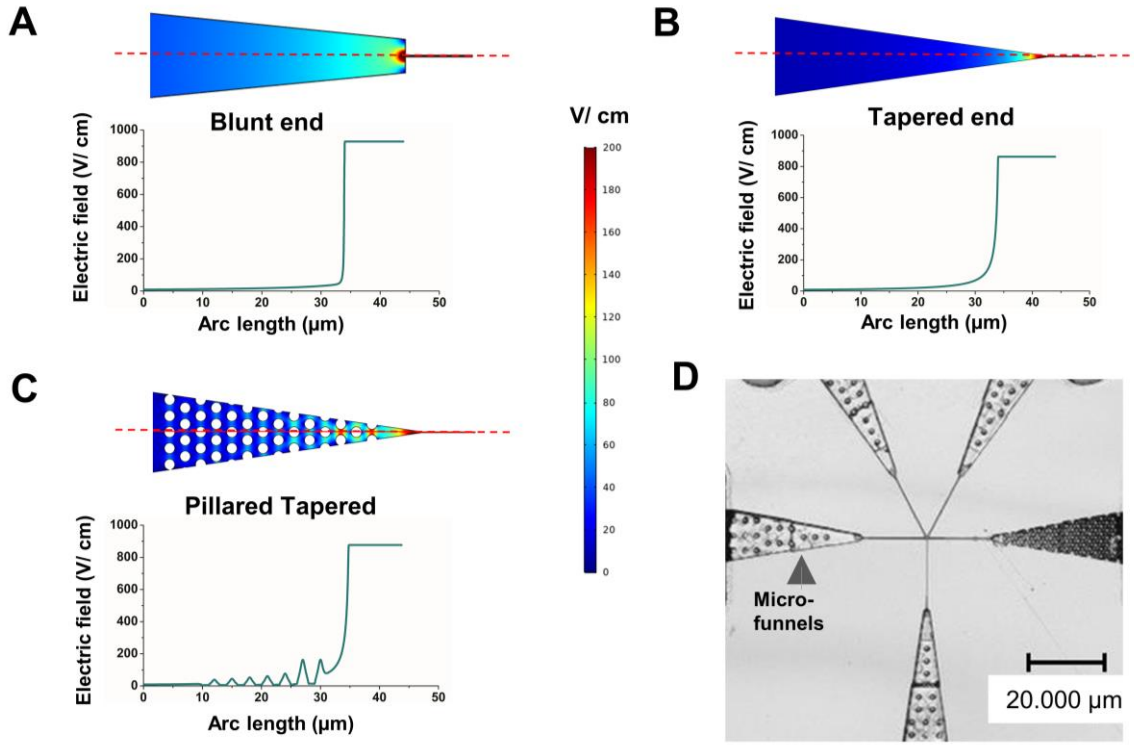


Figure 4.4. Micro-funnel geometry. COMSOL simulation results for electric field distribution in **(A)** blunt end, **(B)** tapered end, and **(C)** pillared tapered end funnel geometries. (COMSOL simulations courtesy of Dr. Swarnagowri Vaidyanathan) **(D)** Pillared tapered funnel at the RNA input side. The nanopillars were 450 nm in diameter.

The ten access microchannels of the XToF device are connected to five micro-funnels. We carried out COMOL simulations to determine the effect of funnel geometry on the electric field distribution (see Figure 4.4.A – 4.4.C). The tapered and pillared geometry of the micro-funnels gave the maximum extension of the electric field into the adjoining microchannel, which will increase sampling efficiency and reduce the mass input requirements for the XToF sensor.[63] The spreading of the electric field helps in better loading efficiency of RNA molecules from the access microchannel. Previous data reported by our group showed a higher capture efficiency of lambda

DNA molecules when a tapered funnel was used compared to other funnel geometries such as blunt end geometry.[63] Additionally, the electric field distribution helps to prevent dielectrophoretic trapping of charged molecules at the micro/nano interface by reducing the electric field gradient. Moreover, the input channel micro-funnel contains nanopillars that are 450 nm in size (see Figure 4.4.D), which further contributes to the extension of the electric field into the adjoining microchannel. The nanopillars also contribute to the efficient entry of RNA molecules into the nanochannel by reducing the entropic effects through partial unwinding of the RNA molecules.

4.3.2.2 Nanochannel design

The enzyme input/output nanochannels and the RNA input/output nanochannels were selected to be 250 nm in depth and width, whereas the nano-flight tube is 50 nm in depth and width (see Figure 4.5.A). Experiments conducted using nanochannels with 100 nm x 100 nm (depth x width) showed that the entry of a DMD RNA molecule (R_g = 70 nm) was prevented (see Figure 4.5.B) due to steric hindrance. Unlike DNA molecules, ssRNA molecules have stable secondary structures (see section 2.3.6 in Chapter 2), which can prevent unwinding of the RNA molecules at the micro-funnel/nanochannel interface and thus, result in poor loading efficiency. For RNA molecules to enter into smaller dimensional nanochannels, the RNA needs to be denatured first using a denaturing agent, such as urea.[65] The high urea concentration needed for denaturation can potentially cause problems with enzyme activity due to possible denaturation. Therefore, larger nanochannel dimensions were chosen to facilitate the entry of the RNA molecules to the nanochannels.

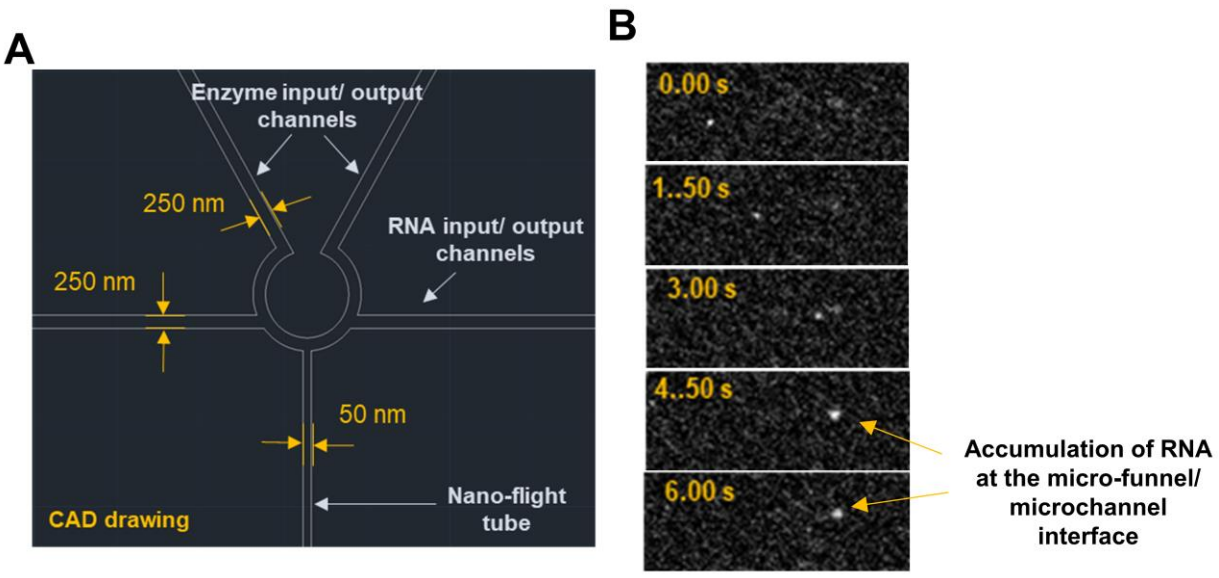


Figure 4.5. Nanochannels of the XToF device. **(A)** AutoCAD drawing of the nanochannels. Enzyme input/output channels are used for enzyme immobilization. RNA input/ output channels are used for RNA introduction to the nanoscale enzymatic reactor. The released rNMPs will be identified in nano-flight tube. **(B)** Accumulation of DMD RNA at the micro-funnel/ nanochannel interface when a 100 nm x 100 nm (W x D) nanochannels are used.

The V-shape of the enzyme input/output channels helps to assure that the enzyme is only attached around the bioreactor. Relatively smaller dimensions were chosen for nano-flight tubes, as only rNMPs will be translocating through the nano flight tube and for the ease of fabricating smaller size in-plane nanopores. In addition, any chromatographic effects that may occur during the electrophoresis can be enhanced by reducing the size of the nano-flight tube generating a higher phase ratio (*i.e.*, increased solute/wall interactions).

4.3.3 XToF device: fabrication and assembly

The fabrication and assembly of the XToF device consisted of four steps: (1) Fabrication of the Si mold master; (2) fabrication of a UV curable PUA resin stamp; (3) replication onto the thermoplastic substrate using NIL; and (4) assembly of the device by attaching a cover plate to the patterned substrate using thermal fusion bonding.

4.3.3.1 Si mold master fabrication

The Si mold master was fabricated as reported previously using photolithography and wet etching for fabricating microchannels followed by FIB milling for nanochannel fabrication (see Figure 4.6.A).

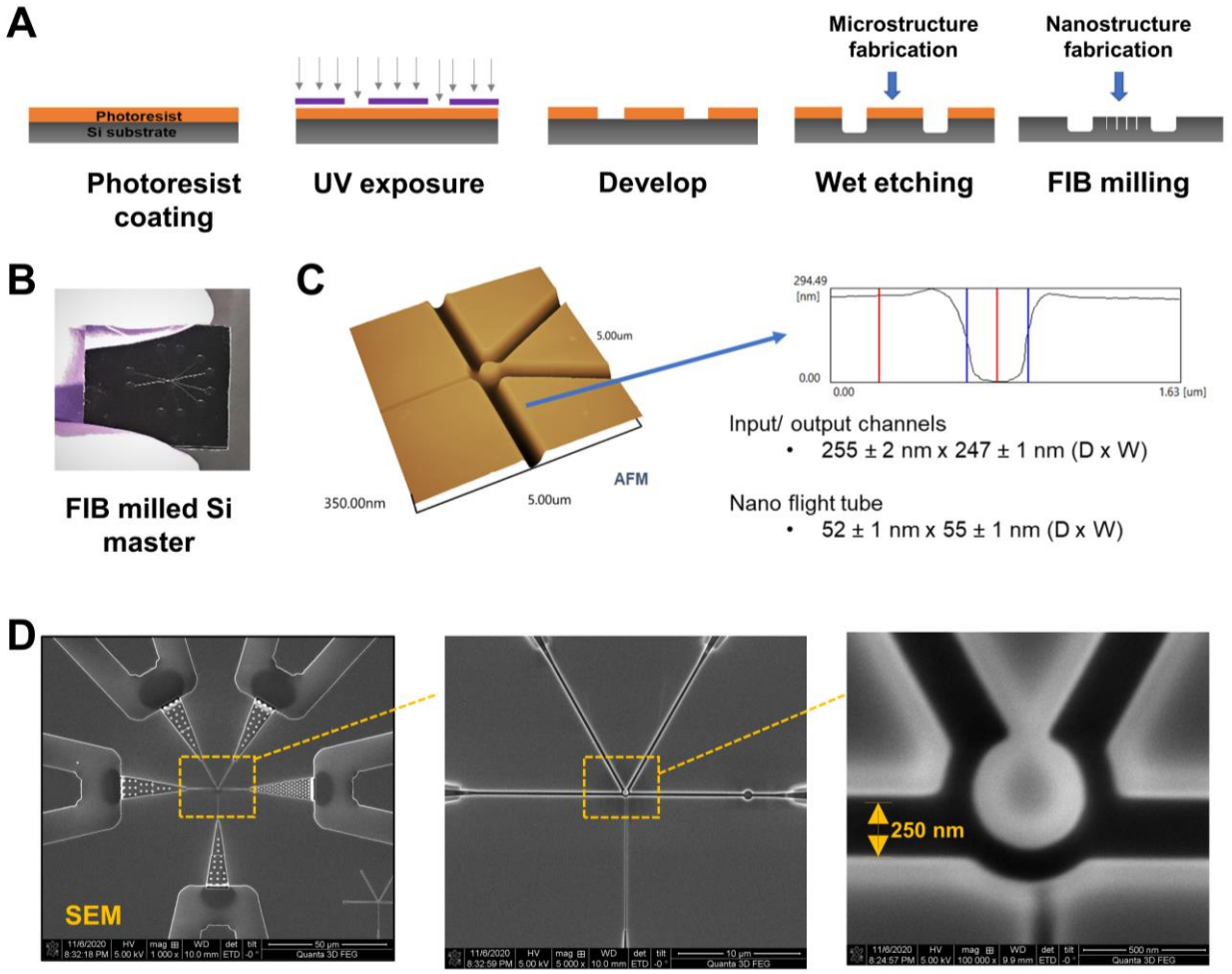


Figure 4.6. Si mold master fabrication. **(A)** Schematic representation of the Si mold fabrication procedure. **(B)** The final Si mold master after FIB milling. **(C)** AFM image of the nanochannels of Si mold master. **(D)** SEM images of the Si mold master.

First, to fabricate the microstructures photolithography and wet etching were done. For photolithography, a 4" Si wafer coated with a 100 nm layer of SiO_2 was used (p-type; resistivity 5-10 Ωcm; orientation of (100); thickness 525 ±25 μm). The SiO_2 wafer was then spin-coated with a 5 μm thick AZ9260 positive photoresist at 4000 rpm for 60 s using a spin-coater. After spin

coating, the Si wafer was baked at 100^{o}C for 10 min. Then, the Si wafer was exposed to UV light (365 nm; 20 mW/ cm^{2}) for 24 s through a photomask that contained the structures of the microchannels (5” x 5” x 0.90”; anti-reflective Chrome Soda Lime). After exposure, the exposed areas were developed using MIF200 developer solution for ~120 s. After development, the Si wafer was checked using an optical microscope to make sure that over-development was not done. Then, the SiO_2 wafer was cut into individual Si mold masters, which was determined using guidelines.

After photolithography to transfer the microstructures into the underlying Si, wet etching using KOH was carried out. Before immersing the Si wafer in the KOH solution, the SiO_2 layer was removed from the exposed areas by immersing in 10:1 buffered oxide etchant (BOE) for 3 minutes. KOH can only remove Si from the surface and does not react with SiO_2. After removing the SiO_2 layer, the Si masters were immersed in a 40% KOH/5% IPA bath heated to 70^{o}C to facilitate the wet etching of the microchannels. The wet etching was carried out until the desired depth of 6 µm was achieved. After the wet etching, the Si masters were washed with acetone to remove the photoresist and used for subsequent steps.

After the fabrication of the microchannels, to fabricate the micro-funnels and the nanochannels, FIB milling was used. FIB milling of the micro-funnels was carried out using a beam voltage of 30 kV and a current of 3 nA in a bitmap mode. For the nanochannels, a beam voltage and current of 30 kV and 50 pA were used, respectiely. The final FIB milled Si master is shown in Figure 4.6.B. Metrology studies using AFM (see Figure 4.6.C), and SEM (see Figure 4.6.D) were conducted to determine the dimensions of the nanochannels in the XToF Si mold master. The input/output nanochannels were 255 ±2 nm x 247 ±1 nm (depth x width), and the nano-flight tube was 52 ±1 nm x 55 ±1 nm (depth x width).

4.3.3.2 PUA resin stamp fabrication

To transfer the structures on the Si mold master to the thermoplastic substrate, a UV resin stamp was used. To prepare the PUA resin stamp, a PET sheet coated with NOA72 was prepared. PET sheet act as the support plate for the PUA resin stamp, and NOA 72 acts as an adhesive between the support plate and the resin. Briefly, the PET sheets were O_2 plasma-activated for 1 min at 50 W, and a solution of NOA72 mixed with acetone in a 1:2 ratio was spin-coated on the PET sheets. The coated PET sheets were then exposed to 365 nm UV light for 300 s using a UV flood source.

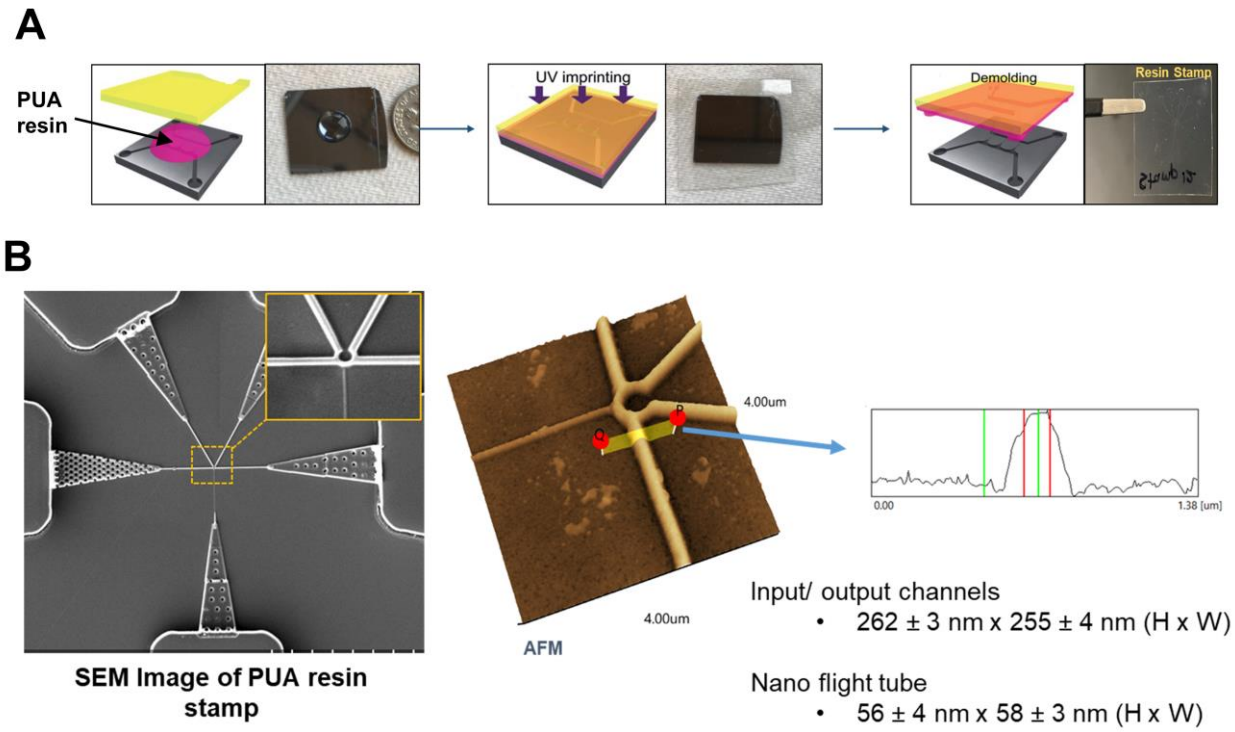


Figure 4.7. UV-curable PUA resin stamp. **(A)** PUA resin stamp fabrication procedure. (Reproduced from reference [61]). **(B)** SEM and AFM images of the PUA resin stamp. The resin stamp contains negatively toned structures compared to the Si mold master.

Next, the Si mold master was exposed to UV/O_3 radiation and 50 µL of PUA resin was placed on top of the plate (see Figure 4.7.A). Then, a PET piece prepared earlier was placed on top of the resin stamp and squeezed gently. Next, the resin was exposed to 365 nm UV light for 3 min for curing. Finally, the resin stamp was demolded from the Si mold master. An SEM and an AFM

image of the PUA resin stamp are shown in Figure 4.7.B. The resin stamp contains negatively toned structures compared to the Si mold master. The size of the input/output nanochannels on the resin stamp were 262 ±3 nm x 255 ±4 nm (height x width), and the nano-flight tube was 56 ±4 nm x 58 ±3 nm (height x width). Prior to using the resin stamps for thermal imprinting, the excess PET was cut and rinsed with IPA to remove uncured PUA resin followed by gentle air-drying.

4.3.3.3 Nano-imprint lithography (NIL)

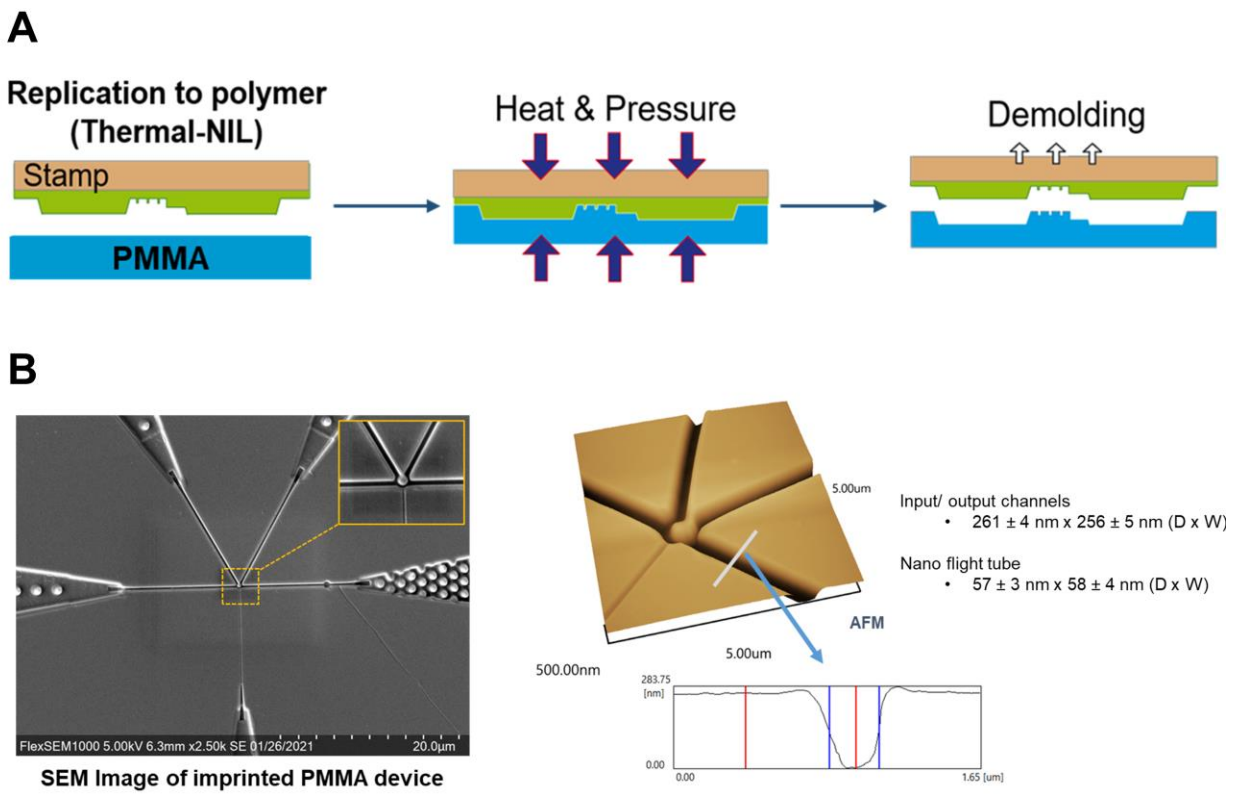


Figure 4.8. Replication using NIL. **(A)** Thermal NIL procedure using PUA resin stamps. **(B)** SEM and AFM images of the imprinted PMMA device.

The structures on the PUA resin stamp fabricated in the previous step were transferred to the thermoplastic substrate via thermal NIL. A Nanonex 2500 was used for the NIL, which uses an air cushion press (ACP) method (see Figure 4.8.A). In ACP, an inert gas such as N_2 is used to press the UV resin stamp and the thermoplastic substrate against each other at an elevated temperature, which was selected to be slightly above the T_g of the substrate. The ACP method

has several advantages compared to NIL methods that use solid parallel plates, such as reduction of damage to the mold, elimination of hot spots during the imprinting, absence of effects related to imperfections in the solid plates, immune to dust and surface variations on the backside of the mold, and faster thermal imprinting.[66-67]

For the imprinting of PMMA (T_g = 105°C), a temperature of 135°C and a pressure of 300 psi were applied for 5 min. Figure 4.8.B shows SEM and AFM images of the imprinted PMMA device. The input/output nanochannels were 261 ±4 nm x 256 ±5 nm (depth x width), and the nano-flight tube was 57 ±3 nm x 58 ±4 nm (depth x width), which indicated good replication fidelity.

4.3.3.4 Thermal fusion bonding of cover plate to imprinted substrate

After imprinting and to obtain an enclosed fluidic network, a COC 8007 (T_g – 78 °C) cover plate was attached to the substrate. Prior to thermal fusion bonding, both the imprinted substrate and cover plate were exposed to O_2 plasma for 1 min at 50 W. As shown in Figure 4.9.A, after O_2 plasma activation, the cover plate was placed on top of the substrate and heated to 72°C at 110 psi for 15 min to facilitate bonding. A final bonded device is shown in Figure 4.9.B.

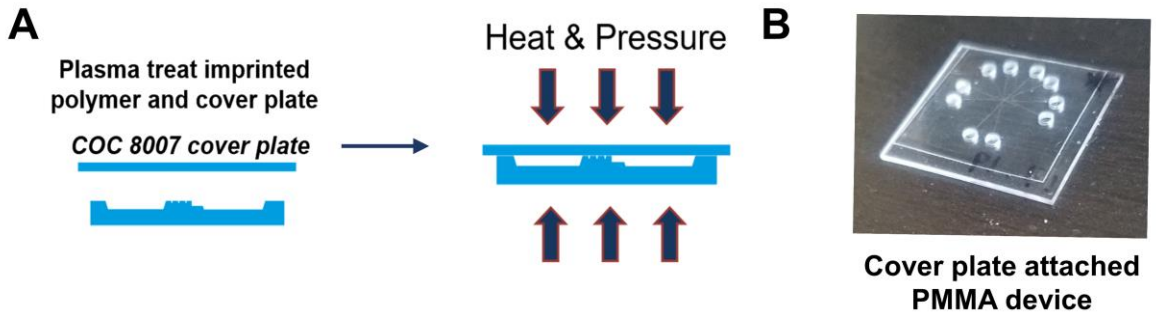


Figure 4.9. XToF device assembly. **(A)** Procedure for COC 8007 cover plate attachment. Both the cover plate and the substrate were exposed to O_2 plasma for 1 min before thermal fusion bonding. **(B)** Cover plate attached PMMA XToF device.

After thermal fusion bonding, we carried out a fluorescence dye test and current measurements to ensure that the nanochannels were connected. The assembled XToF devices were filled with 20 mM Rhodamine B dye for the dye test and imaged using a fluorescence microscope to make sure the devices were filled (see Figure 4.10.A). The current vs. voltage (I-V) measurements were carried out after priming with 50% methanol/water and refilling with NE buffer 3. The I-V curves were obtained using an Axopatch current amplifier instrument and analyzed using Clampfit 11.1 software (see Figure 4.10.B). Current rectification at negative applied voltages could be seen due to the broken symmetry of the nanochannels.[68]

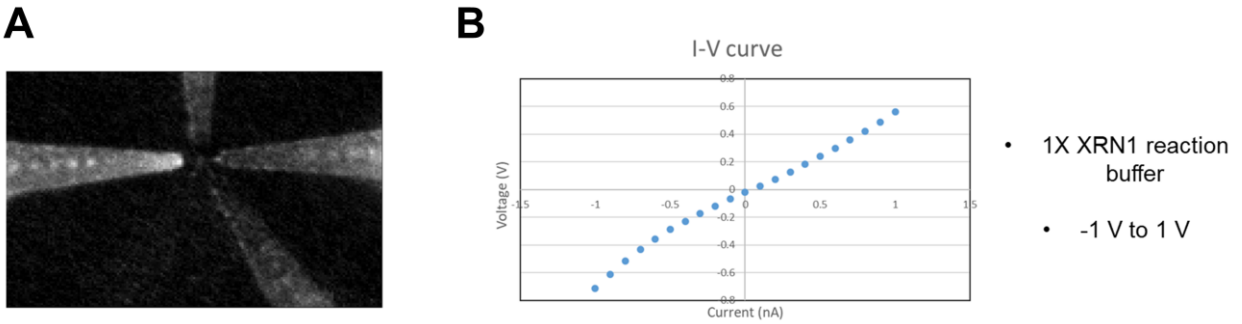


Figure 4.10. (A) Rhodamine B dye test. The microchannels and the nanochannels were filled with 20 nM Rhodamine B dye to make sure the channels are filled. **(B)** I-V curve obtained for XToF device filled with NE buffer 3 to ensure the nanochannels are properly connected.

4.3.4 Fluorescently Labeled RNA translocation through input/output nanochannels

After XToF device fabrication and assembly, fluorescently labeled RNA was used to demonstrate that RNA could be electrokinetically translocated through the input/output nanochannels without the RNA molecules adsorbing onto the nanochannel wall. A 3 nM DMD RNA solution labeled to saturation with SYTO82 dye, as reported earlier in Chapter 2, was introduced into the input reservoir of the device and in the presence of NE buffer 3 (pH 7.9; ionic strength 145 mM) without Mg^{2+} cofactor as shown in Figure 4.11.

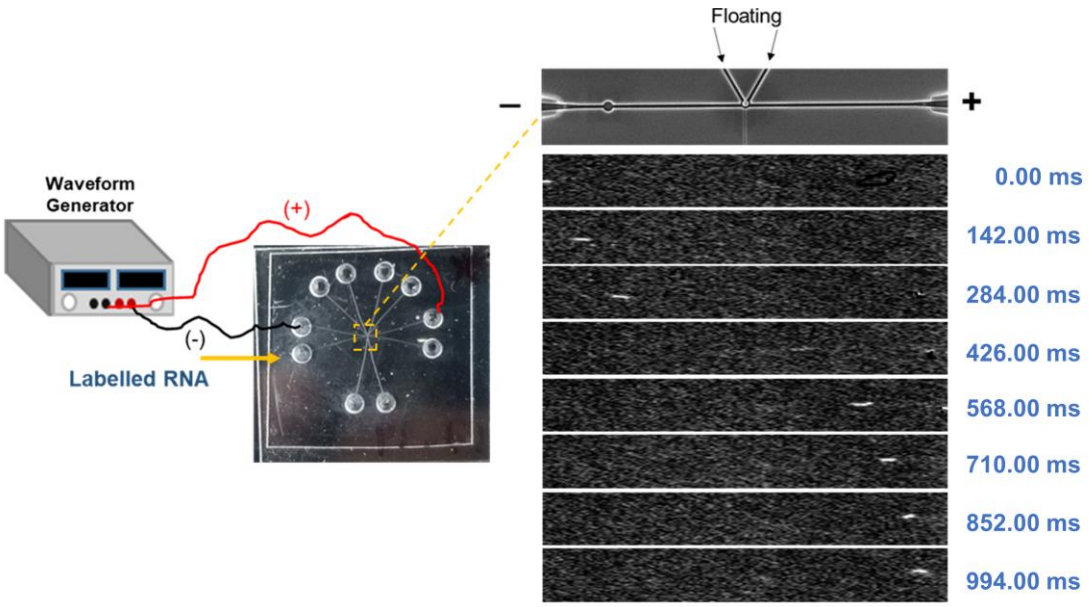


Figure 4.11. SYTO 82 labeled DMD RNA translocation through input/output nanochannels. A driving voltage of 2 V was used.

After introducing labeled RNA, a 2 V potential was applied between the input and the output channels using a waveform generator. As shown in Figure 4.11, labeled DMD RNA molecules translocated electrokinetically through the input and output nanochannels, and we were unable to observe non-moving RNA that could have arisen from non-specific adsorption of RNA molecules to the nanochannel wall. All images were acquired with an exposure time of 20 ms and 0.01 W excitation laser power (532 nm).

4.3.5 Selective immobilization of XRN1 within the bioreactor

After ensuring that RNA could be electrokinetically translocated through plastic nanochannels without non-specific adsorption artifacts, we next sought to site-specifically attach XRN1 enzyme to the nanoscale enzymatic reactor, which consisted of the nanochannel surrounding the 500 nm pillar. Before immobilizing the enzyme, we used a test molecule, which consisted of an oligonucleotide labeled with a Cy3 fluorescent reporter at its 3' end and a primary amine group at

the 5’ end that permitted coupling to the surface –COOH groups of the activated plastic via EDC/NHS coupling chemistry.

For the Cy3 oligonucleotide attachment, “One-step immobilization” was used (see Figure 4.12.A). For the one-step immobilization reaction, the Cy3 labeled oligonucleotide was introduced into the O_2 plasma-activated device in the presence of 20 mg/mL EDC and PBS buffer (pH 7.4). The EDC reacts with the surface-confined carboxylic acid groups to make a labile intermediate prone to nucleophilic attack. The primary amine on the 5’ end of the oligonucleotide reacts with this intermediate to form an amide bond.

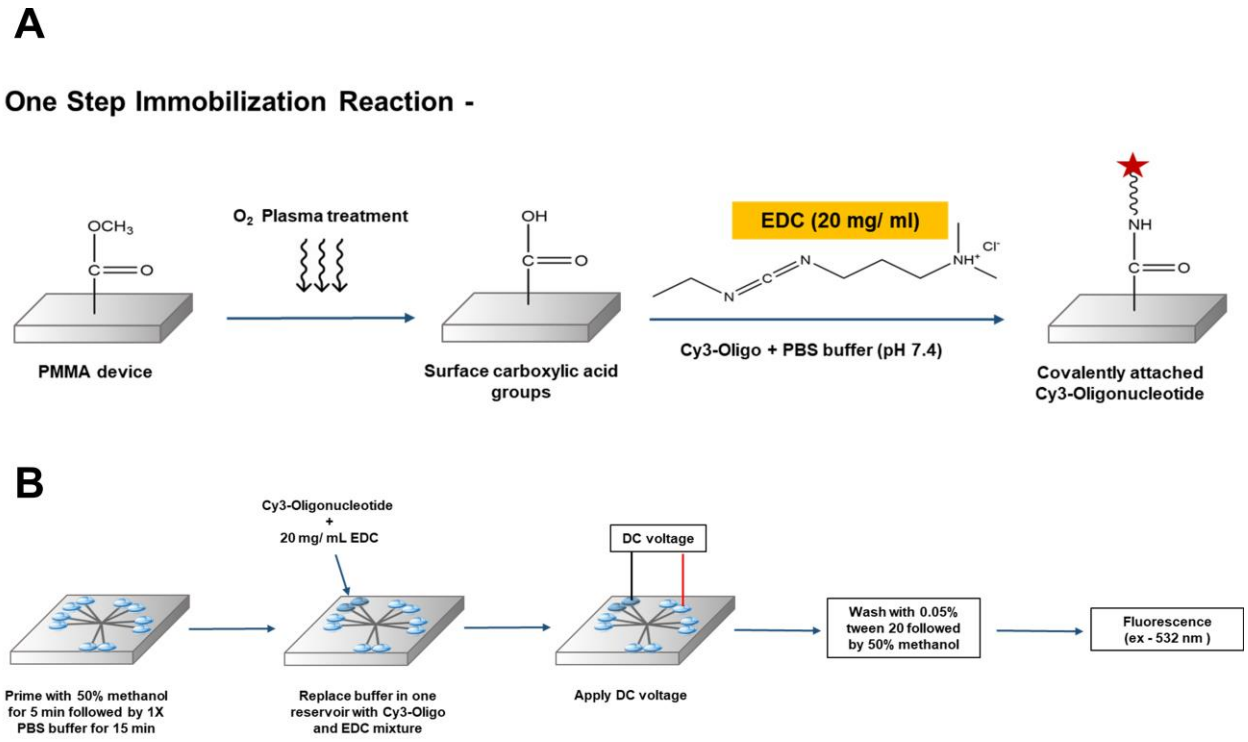
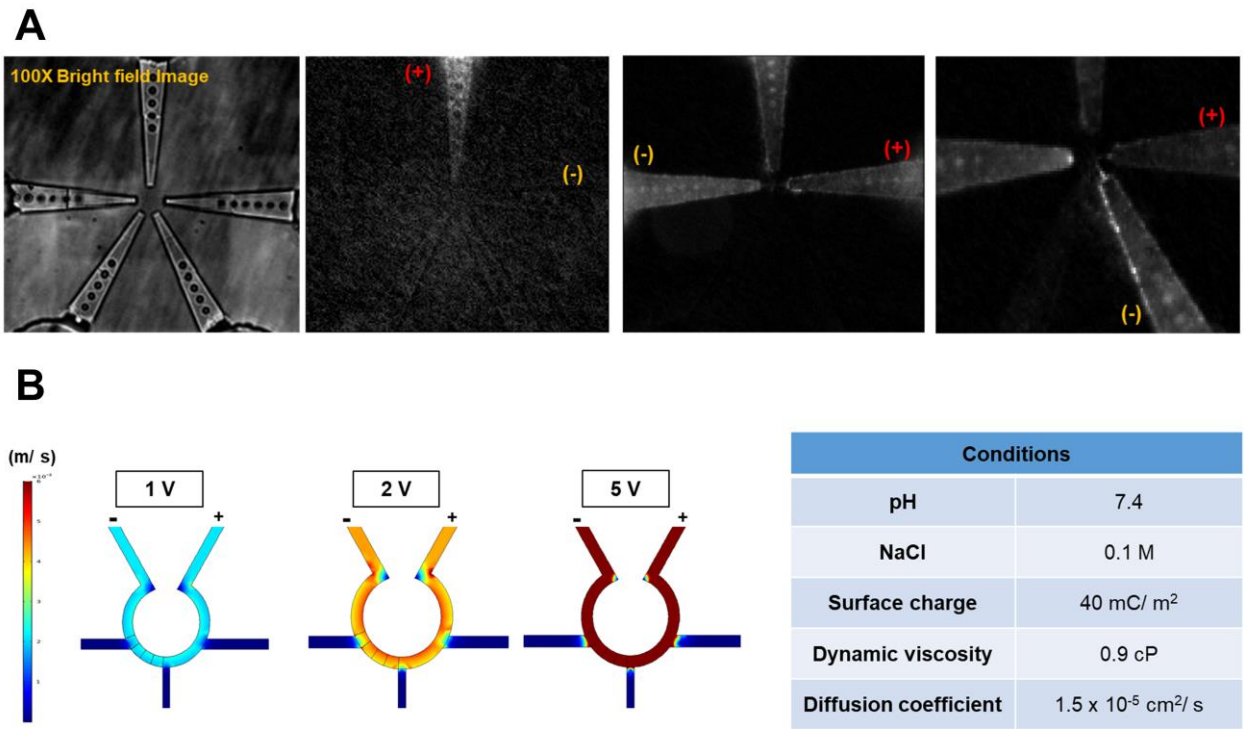


Figure 4.12. Selective immobilization of Cy3-labeled oligonucleotides within the nanoscale enzymatic reactor. **(A)** Schematic representation of the one-step immobilization reaction. **(B)** Selective immobilization procedure.

Before the immobilization reaction and to make sure that the flow of solution could be controlled electrokinetically, we carried out a control experiment using 20 nM Rhodamine B as a tracer. Rhodamine B dye solution was introduced into one microchannel as shown in Figure 4.13.A. Then, a 5 V potential was applied across the nanochannels to drive the dye from one

channel to another. As can be seen from the fluorescence image, the direction of the flow could be controlled by adjusting the applied potential. Moreover, we carried out COMSOL simulations to further demonstrate the ability to electrokinetically control the solution to selectively immobilize biomolecules around the bioreactor. As shown in Figure 4.13.B, at increasing voltage, the velocity of the solution is increased. At all voltages tested, minimum diffusion of the solution to the adjoining nanochannels was seen.



Conditions	
pH	7.4
NaCl	0.1 M
Surface charge	40 mC/ m^2
Dynamic viscosity	0.9 cP
Diffusion coefficient	1.5 x 10^{-5} cm^2/ s

Figure 4.13. Electrokinetic control of solution transport in the nanofluidic network. **(A)** Electrokinetic movement of Rhodamine B dye from one nanochannel to another. **(B)** COMSOL simulations carried out to ensure the solution does not enter other nanochannels. Simulations were carried out for 1 – 5 V driving voltages. (COMSOL simulations courtesy of Dr. Swarnagowri Vaidyanathan).

As shown in Figure 4.12.B, the assembled devices were primed with 50% methanol for 5 min followed by the introduction of 1X PBS buffer. After 15 min, the PBS buffer in one of the enzyme input/output reservoirs was replaced with a solution that contains 50 nM Cy3 labeled oligonucleotide and 20 mg/mL EDC in PBS buffer. After the introduction, a DC voltage was applied across the enzyme input/ output reservoirs to electrokinetically drive the dye molecules around

the bioreactor. After 15 min, the dye solution was completely removed using a vacuum and washed with 0.05% tween 20 followed by 50% methanol to remove non-specifically bound molecules. After washing, the devices were imaged using an epifluorescence microscopy system (ex – 532 nm).

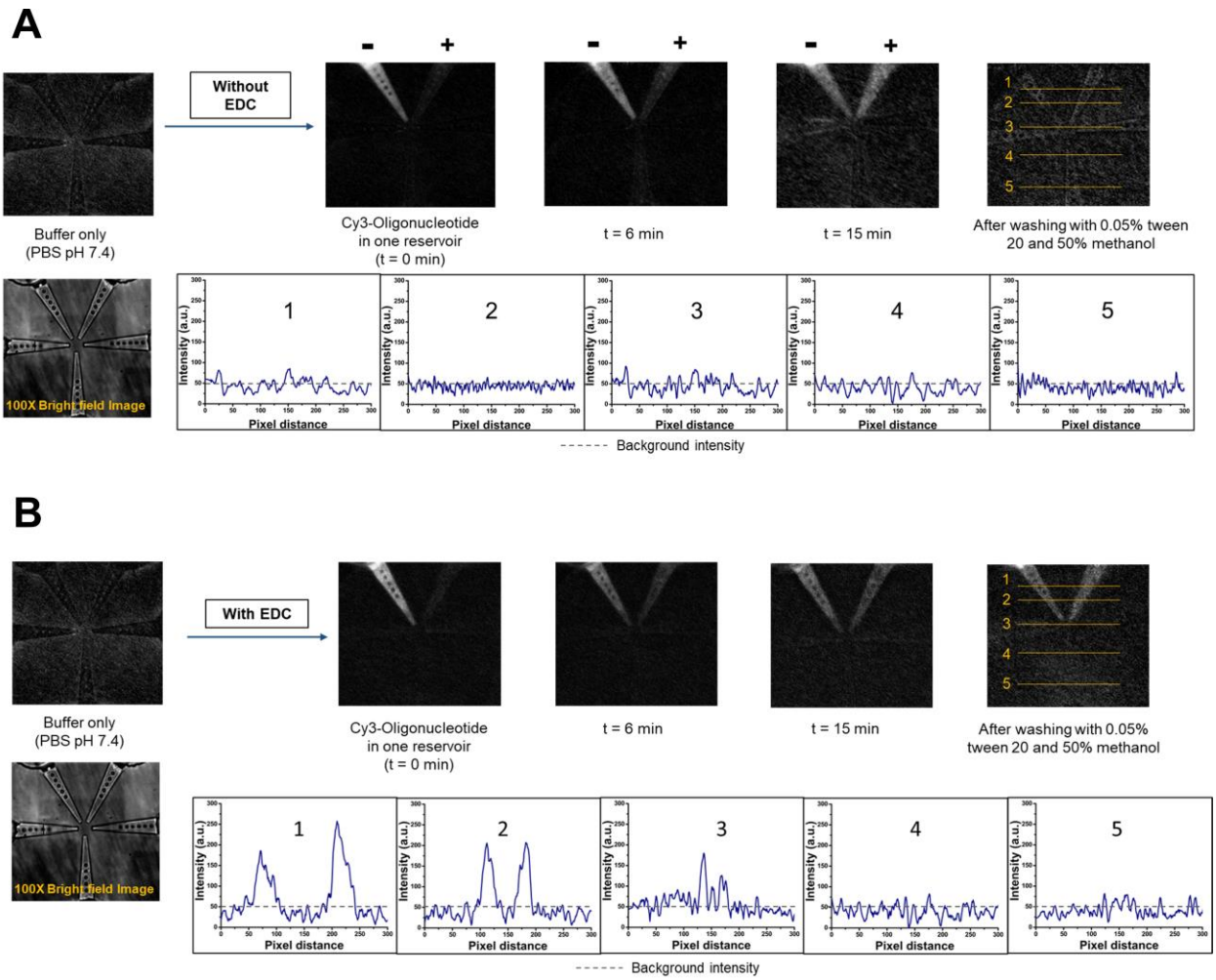


Figure 4.14. (A) Non-specific adsorption of Cy3 labeled oligonucleotides. **(B)** Covalent attachment of Cy3 labeled oligonucleotides using the one-step immobilization reaction.

To make sure that non-specific adsorption of the Cy3 labeled oligonucleotide was negligible, we used a 50 nM solution of Cy3 oligonucleotide that was introduced into the XToF device without EDC (see Figure 4.14.A). A driving potential of 5 V was applied across the enzyme input/output nanochannels and fluorescence images of the XToF device were collected at different times. After

15 min, the oligonucleotide solution was removed, and the device was washed with 0.05% tween 20 and 50% methanol. The fluorescence intensity of the device at different positions was measured as shown in Figure 4.14.A. The fluorescence intensity at each position was similar to the background intensity, which meant that non-specific adsorption was negligible.

Next, the Cy3-labeled oligonucleotide was introduced into the bioreactor, but this time in the presence of 20 mg/mL EDC followed by 5 V potential to drive the solution around the bioreactor (see Figure 4.14.B). After 15 min, the solution was removed and washed with 0.05% tween 20. Next, the fluorescence intensity was measured at different positions of the device. As can be seen from Figure 4.14.B, the fluorescence intensity was above background in the access micro-funnels (position 1 and 2) and around the bioreactor (position 3). This meant that by electrokinetically controlling the flow of the solution, biomolecules could be selectively immobilized around the bioreactor. We understand that diffusion occurring during the reaction and carry some of the immobilization reagents into the channels that were floated during the reaction. However, any diffusional effects into channels where the reaction is not slated to occur can be mitigated by using electrokinetic forces (EOF and/or electrophoretic) to reduce this diffusional artifact by applying voltages in the input/output channels and the flight tube.

4.3.6 XRN1 enzyme kinetics due to nano-confinement

After the Cy3-labeled oligonucleotide control experiments, the XRN1 enzyme was selectively immobilized around the nanoscale enzymatic reactor as was done for the case of the oligonucletoides. However, in this case, the enzyme was not labeled with a fluorescent reporter in order to maintain its activity for subsequent RNA digestion experiments. For XRN1 immobilization, 20 mg/ mL EDC and 2 mg/ mL NHS solution in 0.1 M MES buffer (pH 4.9) were introduced into the enzyme input/output nanochannels of the XToF device. This was achieved by

filling the solution thru the enzyme input/output channels and applying a 5 V potential for 15 min across the enzyme input/output channels. After 15 min, the EDC/NHS solution was completely removed, and a 0.35 mM XRN1 enzyme solution was introduced through the bioreactor using the enzyme input/output channels. Again, a potential of 5 V was applied across the enzyme input/output nanochannels for 30 min to facilitate selective immobilization of the XRN1 enzyme around the INER. Then, the enzyme solution was completely removed, and the device was washed with NE buffer 3 without Mg^{2+}.

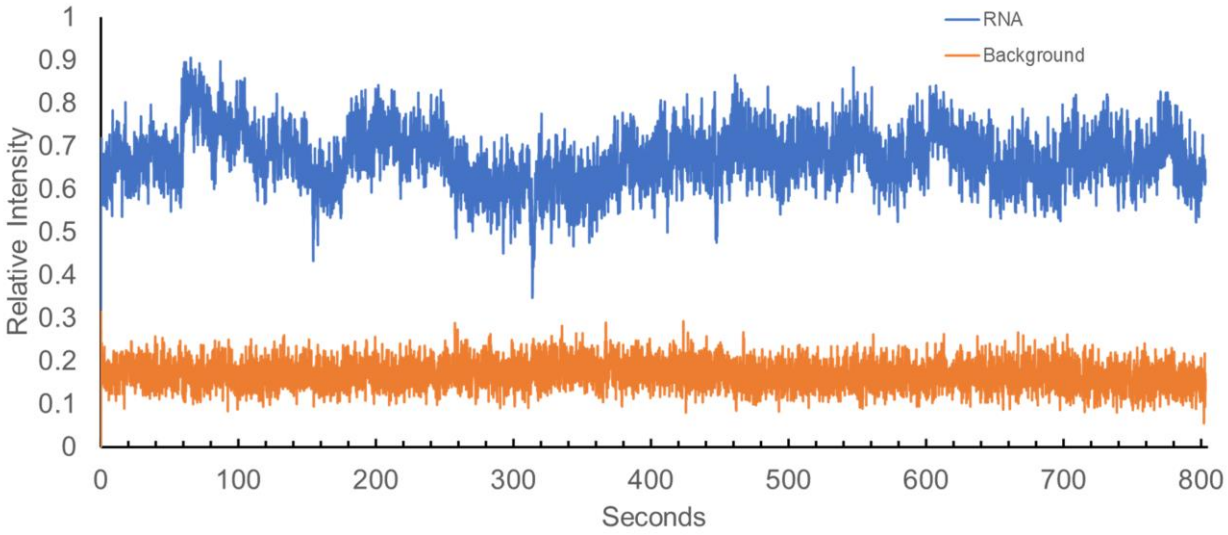


Figure 4.15. Fluorescence intensity of the SYTO 82 labeled Cas9 RNA complexed to an immobilized XRN1 at the nanoscale enzymatic reactor in NE buffer 3 without Mg^{2+} cofactor. (Data courtesy of Chad M. Vietz).

After enzyme immobilization, a 3 nM solution of SYTO 82 labeled Cas9 RNA in XRN1 enzyme buffer without Mg^{2+} was introduced into the RNA input nanochannel of the XToF device. The labeled Cas9 RNA solution contained 4% β-mercaptoethanol (β-ME) to minimize photobleaching effects.[62] The labeled RNA was electrokinetically driven towards the enzyme immobilized nanoscale bioreactor at a 2 V driving voltage. The translocating Cas9 RNA molecules were fluorescently imaged and seen to cease their translocation due to complexation with the immobilized XRN1 at the INER. Due to the absence of the cofactor Mg^{2+}, the RNA would only complex with the enzyme without digestion and this was visually seen as no loss in fluorescence

from the SYTO 82 reporter was observed indicative of an intact RNA molecule. To ensure the Cas9 RNA molecules were complexed to the immobilized enzyme, the potential was reversed to see if the RNA molecules would reverse their direction and go back into the input nanochannel. Even with the increased and reversed potentials, the RNA molecules were stationary due to complexation with the surface immobilized XRN1.

Before introducing Mg^{2+} cofactor to activate the enzyme, we measured the fluorescence intensity of an XRN1/Cas9 RNA complex with time to ensure there was no photobleaching. As shown in Figure 4.15, the fluorescence intensity was reasonably stable over an 800 s time interval, which indicated that photobleaching was negligible. The fluctuations in the intensity around 100 – 300 s were due to defocusing of the microscope while collecting data.

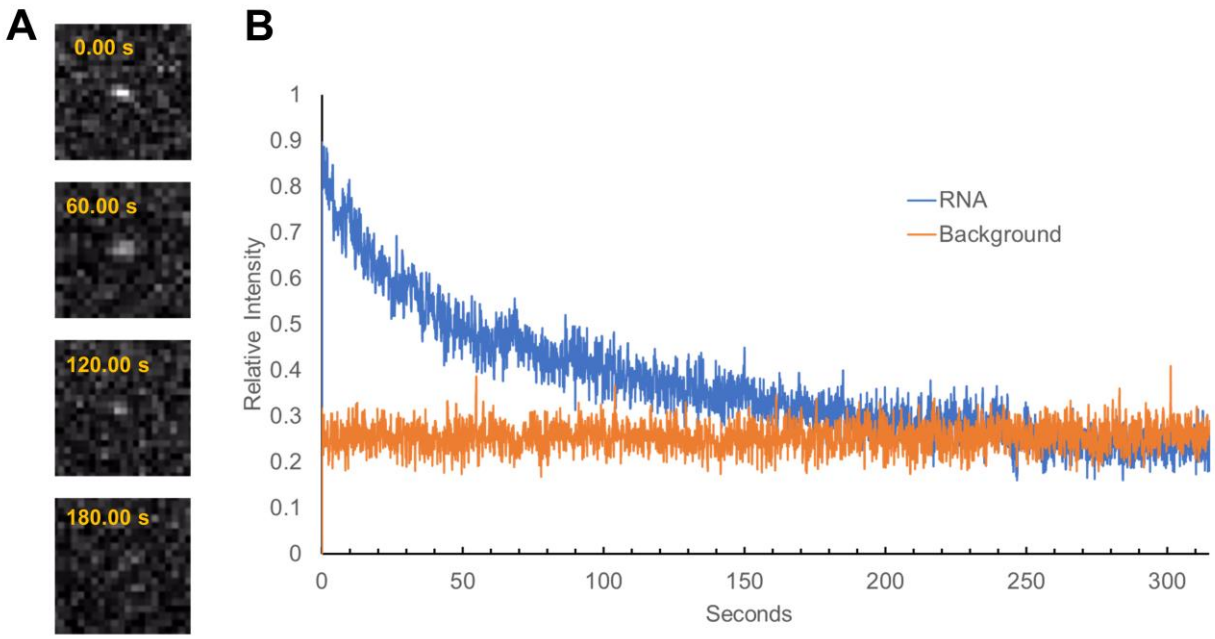


Figure 4.16. Labeled Cas9 RNA digestion by XRN1 enzyme immobilized at the nanoscale enzymatic reactor. **(A)** Fluorescence still images of Cas9 RNA/ immobilized XRN1 enzyme with time after the introduction of the Mg^{2+} cofactor. **(B)** Fluorescence intensity decay with time. (Data courtesy of Chad M. Vietz).

Next, NE buffer 3 containing Mg^{2+} cofactor was introduced into the XToF device with an immobilized XRN1-labeled RNA complex. The buffer was introduced into the enzyme input/output

channels and was electrokinetically driven towards the bioreactor by applying a 2 V potential. When Mg^{2+} was introduced, the XRN1 enzyme was activated, and the fluorescence intensity of the Cas9 RNA decreased with time (see Figure 4.16.A). The images were collected at a 20 ms exposure time at 21 frames per second with 0.01 W excitation laser power (532 nm). The fluorescence intensity was plotted against time to deduce the clipping rate of the nano-confined XRN1 (see Figure 4.16.B). As described earlier in section 2.3.4 in Chapter 2, the lowest RNA fragment that could be measured using our epifluorescence single-molecule microscope was 664 nucleotides. Therefore, the clipping rate of the nano-confined XRN1 was calculated to be 23 ±3 nt s^{-1} (n = 3). As previously reported in Chapter 2, the clipping rate of XRN1 immobilized to a micropillar (*i.e*., IMER) was 26 ±5 nt s^{-1} (n = 3). The volume of the nanoscale reactor was calculated as ~1.3 x 10^{-10} µL and the volume of the IMER used in Chapter 2 was 2.9 µL. Sometimes the activity of the enzyme could be affected due to nanoconfinement of the enzyme.[69-70] Wang et al. reported that the activity of glucose oxidase (GOx) in "free state" is two times faster than when nanoconfined to a nanochannel that was 80 nm in depth and 200 nm in width.[69] The reduction in activity could be due to the partial inactivation of the enzyme and the diffusion restricted accessibility of the substrate to the enzyme due to nanoconfinement.[70] However, with immobilized XRN1 there was no statistically significant difference in clipping rate in the INER compared to the IMER (P >0.05). This could be due to the larger size of the nanochannels (250 nm in depth and width) compared to the size of the enzyme (~15 nm). If the dimensions of the channels were reduced a significant change in the activity could be expected.

4.4 Future directions

The next step of the project is to identify the rNMPs generated by the sequential digestion of a ssRNA by the XRN1 enzyme immobilized within the INER. Recently, we reported the identification of ATTO-532 labeled rNMPs in thermoplastic nanochannels that are 110 nm x 110 nm in width and depth.[71] We were able to identify the canonical rNMPs in thermoplastic devices fabricated in PMMA/COC (see Figure 4.17.A) and COC/COC devices (see Figure 4.17.B), which is the ideal pH for the XRN1 enzyme. Identification accuracies >99.9% were achieved for both PMMA/COC and COC/COC devices with resolutions >4. This basically translates into the read accuracy of the rNMPs during the sequencing run.

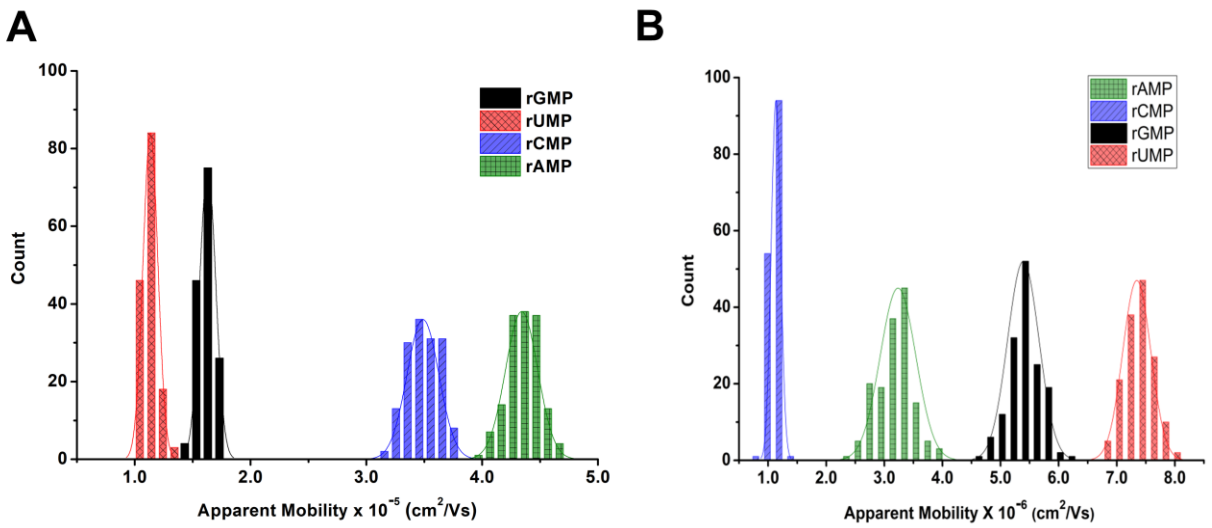


Figure 4.17. Histograms of the apparent mobilities of rNMPs in **(A)** 110 nm x 110 nm PMMA/ COC nanochannel devices at 280 V/ cm and **(B)** 110 nm x 110 nm COC/ COC nanochannel devices at 980 V/ cm. The carrier electrolyte was NE buffer 3 at pH 7.9. (Reproduced from reference [71]).

In the final version of the XToF device, we would not be using labeled RNA of rNMPs for identification because rNMP labeling with a fluorescent dye is not practical.[72-74] To obviate the need for fluorescent labels, we will seek to use current transients generated during resistive pulse sensing of unlabeled molecules to deduce the ToF. In Chapter 3, we described the use of dual in-plane nanopores for the identification of both intact ssRNA and rNMPs. Currently, we are

working on fabricating an XToF device that contains dual in-plane nanopores with both the RNA input/output nanochannel (see Figure 4.2) and the nano-flight tube for identification of the entering ssRNA and sequentially released rNMPs, respectively.

Work is also underway in our laboratory to optimize the nanoscale electrophoresis conditions for the identification of the canonical rNMPs and modified rNMPs (m-rAMP, m-rCMP) using a modified dual in-plane nanopore device, which contains nanopores that are ≥5 μm apart from each other. Modifications were done to the dual in-plane nanopore device described in Chapter 3 in order to increase the sampling efficiency as well. The width of the access microchannel was decreased to 20 μm from 130 μm at the microchannel/ micro-funnel interface (see Figure 4.18.A). The depth of the microchannels was also reduced to 3 μm from 10 μm. The micro-funnel geometry was changed to a tapered pillared shape, as shown in Figure 4.18.B.

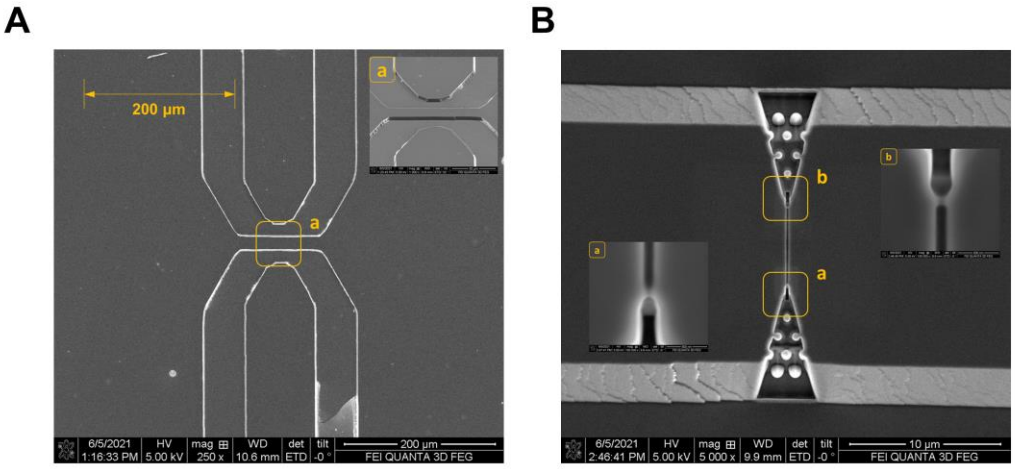


Figure 4.18. Modified dual in-plane nanopore device. **(A)** Structure of the access microchannels. The width of the channel at the position "a" was reduced to 20 μm from 130 μm. **(B)** The tapered pillared input funnels and the dual in-plane nanopores.

Finally, the sequential digestion will be integrated with the identification of the rNMPs to facilitate single-molecule sequencing of RNA. Unlabeled RNA will be introduced into the XToF device with XRN1 immobilized at the INER, and the released and unlabeled rNMPs will be

identified via their molecular dependent time-of-flight between two in-plane nanopores and current transient amplitudes.

4.6 References

1. Marguerat, S.; Bähler, J., RNA-seq: from technology to biology. *Cellular and molecular life sciences* **2010,** *67* (4), 569-579.

2. Ozsolak, F.; Milos, P. M., RNA sequencing: advances, challenges and opportunities. *Nature reviews genetics* **2011,** *12* (2), 87-98.

3. Bonifati, V., Will New Genetic Techniques Like Exome Sequencing and Others Obviate the Need for Clinical Expertise? Yes. *Movement Disorders Clinical Practice* **2017,** *4* (1), 36-38.

4. Broaddus, R. R.; Litzenburgen, B.; Shaw, K.; Lee, J.; Zhang, J. X.; Kopetz, S., Clinical Value of Larger Multi-Gene Next-Generation Sequencing Panels. *Modern Pathology* **2017,** *30*, 454A-454A.

5. Bunnell, A. E.; Garby, C. A.; Pearson, E. J.; Walker, S. A.; Panos, L. E.; Blum, J. L., The Clinical Utility of Next Generation Sequencing Results in a Community-Based Hereditary Cancer Risk Program. *Journal of Genetic Counseling* **2017,** *26* (1), 105-112.

6. Bustos, B. D. U.; Estal, R. M.; Simo, G. P.; Jimenez, I. D.; Munoz, B. E.; Serna, M. R.; de Miquel, V. A.; Ros, M. L.; Sanchez, R. B.; Enguidanos, E. N.; Suela, S. P.; Estrada, R. B., Towards Personalized Medicine in Melanoma: Implementation of a Clinical Next-Generation Sequencing Panel. *Scientific Reports* **2017,** *7*.

7. Frank, K.; Aster, J. C.; Lindsley, C., Two-Year Experience of Performing a Next-Generation-Sequencing Based Panel Test in an Academic Medical Center and Its Clinical Impact. *Blood* **2016,** *128* (22).

8. Sethi, K. D.; Lang, A. E., Will New Genetic Techniques Like Exome Sequencing Obviate the Need for Clinical Expertise? No. *Movement Disorders Clinical Practice* **2017,** *4* (1), 39-41.

9. Editorial, Knocking on the Clinical Door. *Nature Biotechnology* **2012,** *30*, 1009.

10. Davis, L. M.; Fairfield, E. R.; Harger, C. A.; Jett, J. H.; Hahn, J. H.; Keller, R. A.; Krakowski, L. A.; Martin, J. C.; Marrone, B. L.; Ratliff, R. L.; Seitzinger, N. K.; Shera, E. B.; Soper, S. A., High-Speed DNA Sequencing - an Approach Based Upon Fluorescence Detection of Single Molecules. *Abstr. Pap. Am. Chem. Soc.* **1990,** *200*, 75-ANYL.

11. Gupta, P. K., Single-molecule DNA sequencing technologies for future genomics research. *Trends in Biotechnology* **2008,** *26* (11), 602-611.

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

13. Milos, P. M., Emergence of single-molecule sequencing and potential for molecular diagnostic applications. *Expert Review of Molecular Diagnostics* **2009,** *9* (7), 659-666.

14. Pushkarev, D.; Neff, N. F.; Quake, S. R., Single-molecule sequencing of an individual human genome. *Nature Biotechnology* **2009,** *27* (9), 847-U101.

15. Turner, S., Applying single molecule real time DNA sequencing. *Cancer Research* **2009,** *69*.

16. Braslavsky, I., Single Molecule DNA Sequencing: from Demonstration to Application. *Biophysical Journal* **2010,** *98* (3), 208A-209A.

17. Flusberg, B. A.; Webster, D. R.; Lee, J. H.; Travers, K. J.; Olivares, E. C.; Clark, T. A.; Korlach, J.; Turner, S. W., Direct detection of DNA methylation during single-molecule, real-time sequencing. *Nature Methods* **2010,** *7* (6), 461-U72.

18. Asghar, W.; Ilyas, A.; Billo, J. A.; Iqbal, S. M., Shrinking of solid-state nanopores by direct thermal heating. *Nanoscale research letters* **2011,** *6* (1), 372.

19. Thompson, J. F., Single Molecule Sequencing Technology: the Applications and Implications for Biological Research. *In Vitro Cellular & Developmental Biology-Animal* **2011,** *47*, S14-S14.

20. Thompson, J. F.; Milos, P. M., The properties and applications of single-molecule DNA sequencing. *Genome Biology* **2011,** *12* (2).

21. Avent, N. D., Refining Noninvasive Prenatal Diagnosis with Single-Molecule Next-Generation Sequencing. *Clinical Chemistry* **2012,** *58* (4), 657-658.

22. van den Oever, J. M. E.; Balkassmi, S.; Verweij, E. J.; van Iterson, M.; van Scheltema, P. N. A.; Oepkes, D.; van Lith, J. M. M.; Hoffer, M. J. V.; den Dunnen, J. T.; Bakker, E.; Boon, E. M. J., Single Molecule Sequencing of Free DNA from Maternal Plasma for Noninvasive Trisomy 21 Detection. *Clinical Chemistry* **2012,** *58* (4), 699-706.

23. Davis, B. M.; Chao, M. C.; Waldor, M. K., Entering the era of bacterial epigenomics with single molecule real time DNA sequencing. *Current Opinion in Microbiology* **2013,** *16* (2), 192-198.

24. Shin, S. C.; Ahn, D. H.; Kim, S. J.; Lee, H.; Oh, T. J.; Lee, J. E.; Park, H., Advantages of Single-Molecule Real-Time Sequencing in High-GC Content Genomes. *Plos One* **2013,** *8* (7).

25. Patel, A.; Schwab, R.; Liu, Y. T.; Bafna, V., Amplification and thrifty single-molecule sequencing of recurrent somatic structural variations. *Genome Research* **2014,** *24* (2), 318-328.

26. Sakai, H.; Naito, K.; Ogiso-Tanaka, E.; Takahashi, Y.; Iseki, K.; Muto, C.; Satou, K.; Teruya, K.; Shiroma, A.; Shimoji, M.; Hirano, T.; Itoh, T.; Kaga, A.; Tomooka, N., The power of single molecule real-time sequencing technology in the de novo assembly of a eukaryotic genome. *Scientific Reports* **2015,** *5*.

27. Steinig, E. J.; Andersson, P.; Harris, S. R.; Sarovich, D. S.; Manoharan, A.; Coupland, P.; Holden, M. T. G.; Parkhill, J.; Bentley, S. D.; Robinson, D. A.; Tong, S. Y. C., Single-molecule sequencing reveals the molecular basis of multidrug-resistance in ST772 methicillin-resistant Staphylococcus aureus. *Bmc Genomics* **2015,** *16*.

28. Yang, Y.; Sebra, R.; Pullman, B. S.; Qiao, W. Q.; Peter, I.; Desnick, R. J.; Geyer, C. R.; DeCoteau, J. F.; Scott, S. A., Quantitative and multiplexed DNA methylation analysis using long-read single-molecule real-time bisulfite sequencing (SMRT-BS). *Bmc Genomics* **2015,** *16*.

29. Betz-Stablein, B. D.; Topfer, A.; Littlejohn, M.; Yuen, L.; Colledge, D.; Sozzi, V.; Angus, P.; Thompson, A.; Revill, P.; Beerenwinkel, N.; Warner, N.; Luciani, F., Single-Molecule Sequencing Reveals Complex Genome Variation of Hepatitis B Virus during 15 Years of Chronic Infection following Liver Transplantation. *Journal of Virology* **2016,** *90* (16), 7171-7183.

30. Gao, Y.; Deng, L. W.; Yan, Q.; Gao, Y. Q.; Wu, Z. D.; Cai, J. S.; Ji, D. R.; Li, G. L.; Wu, P.; Jin, H.; Zhao, L. Y.; Liu, S.; Ge, L. J.; Deem, M. W.; He, J. K., Single molecule targeted sequencing for cancer gene mutation detection. *Scientific Reports* **2016,** *6*.

31. Hargreaves, K. R.; Thanki, A. M.; Jose, B. R.; Oggioni, M. R.; Clokie, M. R. J., Use of single molecule sequencing for comparative genomics of an environmental and a clinical isolate of Clostridium difficile ribotype 078. *Bmc Genomics* **2016,** *17*.

32. Kujawa, S.; Sethuraman, A.; Eng, K.; Baybayan, P.; Heyrman, L.; Del Favero, J., Highly sensitive and cost-effective detection of somatic cancer variants using single molecule, real-time sequencing. *Cancer Research* **2016,** *76*.

33. Ozsolak, F., Attomole-level Genomics with Single-molecule Direct DNA, cDNA and RNA Sequencing Technologies. *Current Issues in Molecular Biology* **2016,** *18*, 43-48.

34. Milos, P. M., Helicos single molecule sequencing: unique capabilities and importance for molecular diagnostics. *Genome Biology* **2010,** *11*.

35. Fuller, C. W.; Kumar, S.; Porel, M.; Chien, M. C.; Bibillo, A.; Stranges, P. B.; Dorwart, M.; Tao, C. J.; Li, Z. M.; Guo, W. J.; Shi, S. D.; Korenblum, D.; Trans, A.; Aguirre, A.; Liu, E.; Harada, E. T.; Pollard, J.; Bhat, A.; Cech, C.; Yang, A.; Arnold, C.; Palla, M.; Hovis, J.; Chen, R.; Morozova, I.; Kalachikov, S.; Russo, J. J.; Kasianowicz, J. J.; Davis, R.; Roever, S.; Church, G. M.; Ju, J. Y., Real-time single-molecule electronic DNA sequencing by synthesis using polymer-tagged nucleotides on a nanopore array. *Proceedings of the National Academy of Sciences of the United States of America* **2016,** *113* (19), 5233-5238.

36. Stranges, P. B.; Palla, M.; Kalachikov, S.; Nivala, J.; Dorwart, M.; Trans, A.; Kumar, S.; Porel, M.; Chien, M. C.; Tao, C. J.; Morozova, I.; Li, Z. M.; Shi, S. D.; Aberra, A.; Arnold, C.; Yang, A.; Aguirre, A.; Harada, E. T.; Korenblum, D.; Pollard, J.; Bhat, A.; Gremyachinskiy, D.; Bibillo, A.;

Chen, R.; Davis, R.; Russo, J. J.; Fuller, C. W.; Roever, S.; Ju, J. Y.; Church, G. M., Design and characterization of a nanopore-coupled polymerase for single-molecule DNA sequencing by synthesis on an electrode array. *Proceedings of the National Academy of Sciences of the United States of America* **2016,** *113* (44), E6749-E6756.

37. McCarthy, A., Third Generation DNA Sequencing: Pacific Biosciences' Single Molecule Real Time Technology. *Chemistry & Biology* **2010,** *17* (7), 675-676.

38. Grohme, M. A.; Soler, R. F.; Wink, M.; Frohme, M., Microsatellite marker discovery using single molecule real-time circular consensus sequencing on the Pacific Biosciences RS. *Biotechniques* **2013,** *55* (5), 255-258.

39. Eisenstein, M., Oxford Nanopore announcement sets sequencing sector abuzz. *Nature Biotechnology* **2012,** *30* (4), 295-296.

40. Goodwin, S.; Gurtowski, J.; Ethe-Sayers, S.; Deshpande, P.; Schatz, M. C.; McCombie, W. R., Oxford Nanopore sequencing, hybrid error correction, and de novo assembly of a eukaryotic genome. *Genome Research* **2015,** *25* (11), 1750-1756.

41. Hargreaves, A. D.; Mulley, J. F., Assessing the utility of the Oxford Nanopore MinION for snake venom gland cDNA sequencing. *Peerj* **2015,** *3*.

42. Jain, M.; Olsen, H. E.; Paten, B.; Akeson, M., The Oxford Nanopore MinION: delivery of nanopore sequencing to the genomics community (vol 17, 239, 2016). *Genome Biology* **2016,** *17*.

43. Jain, M.; Olsen, H. E.; Paten, B.; Akeson, M., The Oxford Nanopore MinION: delivery of nanopore sequencing to the genomics community. *Genome Biology* **2016,** *17*.

44. Oikonomopoulos, S.; Wang, Y. C.; Djambazian, H.; Badescu, D.; Ragoussis, J., Benchmarking of the Oxford Nanopore MinION sequencing for quantitative and qualitative assessment of cDNA populations. *Scientific Reports* **2016,** *6*.

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

46. Zwolak, M.; Di Ventra, M., Colloquium: Physical approaches to DNA sequencing and detection. *Reviews of Modern Physics* **2008,** *80* (1), 141-165.

47. Greenleaf, W. J.; Block, S. M., Single-molecule, motion-based DNA sequencing using RNA polymerase. *Science (New York, N.Y.)* **2006,** *313* (5788), 801-801.

48. Larkin, J.; Foquet, M.; Korlach, J.; Wanunu, M., Nanopore immobilization of DNA polymerase enhances single-molecule sequencing. *Journal of Biomolecular Structure & Dynamics* **2013,** *31*, 134-135.

49. Cherf, G. M.; Lieberman, K. R.; Rashid, H.; Lam, C. E.; Karplus, K.; Akeson, M., Automated forward and reverse ratcheting of DNA in a nanopore at 5-Å precision. *Nature Biotechnology* **2012,** *30*, 344-348.

50. Brinkerhoff, H., Improving DNA Sequencing with Nanopore MSPA. *Biophysical Journal* **2016,** *110* (3), 531A-531A.

51. Manara, R. M. A.; Wallace, E. J.; Khalid, S., DNA sequencing with MspA: Molecular Dynamics simulations reveal free-energy differences between sequencing and non-sequencing mutants. *Scientific Reports* **2015,** *5*.

52. Liu, H.; Qian, S.; Bau, H. H., The effect of translocating cylindrical particles on the ionic current through a nanopore. *Biophysical Journal* **2007,** *92*, 1164-1177.

53. Jain, M.; Olsen, H. E.; Paten, B.; Akeson, M., The Oxford Nanopore MinION: delivery of nanopore sequencing to the genomics community. *Genome biology* **2016,** *17* (1), 1-11.

54. Pirooznia, M.; Kramer, M.; Parla, J.; Goes, F. S.; Potash, J. B.; McCombie, W. R.; Zandi, P. P., Validation and assessment of variant calling pipelines for next-generation sequencing. *Human genomics* **2014,** *8* (1), 1-10.

55. Davis, L. M., Fairfield, F.R., Hammond, M.L., Harger, C.A., Jett, J.H., Keller, R.A., Hahn, J.H., Krakowski, L.A., Marrone, B., Martin, J.C., Nutter, H.L., Ratliff, R.R., Shera, E.B., Simpson, D.J., Soper, S.A., Wilkerson, C.W. *Rapid DNA Sequencing Based on Single Molecule Detection*; 1992.

56. Wu, H.-C.; Astier, Y.; Maglia, G.; Mikhailova, E.; Bayley, H., Protein nanopores with covalently attached molecular adapters. *Journal of the American Chemical Society* **2007,** *129*, 16142-16148.

57. Ayub, M.; Hardwick, S. W.; Luisi, B. F.; Bayley, H., Nanopore-based identification of individual nucleotides for direct RNA sequencing. *Nano letters* **2013,** *13* (12), 6144-6150.

58. Reiner, J. E.; Balijepalli, A.; Robertson, J. W. F.; Drown, B. S.; Burden, D. L.; Kasianowicz, J. J., The effects of diffusion on an exonuclease/nanopore-based DNA sequencing engine. *The Journal of Chemical Physics* **2012,** *137*, 214903.

59. Athapattu, U. S.; Amarasekara, C. A.; Immel, J. R.; Bloom, S.; Barany, F.; Nagel, A. C.; Soper, S. A., Solid-phase XRN1 reactions for RNA cleavage: application in single-molecule sequencing. *Nucleic acids research* **2021,** *49* (7), e41-e41.

60. Chantiwas, R.; Park, S.; Soper, S. A.; Kim, B. C.; Takayama, S.; Sunkara, V.; Hwang, H.; Cho, Y. K., Flexible fabrication and applications of polymer nanochannels and nanoslits. *Chemical Society Reviews* **2011,** *40* (7), 3677-3702.

61. Wu, J. H.; Chantiwas, R.; Amirsadeghi, A.; Soper, S. A.; Park, S., Complete plastic nanofluidic devices for DNA analysis via direct imprinting with polymer stamps. *Lab on a Chip* **2011,** *11* (17), 2984-2989.

62. Oliver-Calixte, N. J.; Uba, F. I.; Battle, K. N.; Weerakoon-Ratnayake, K. M.; Soper, S. A., Immobilization of lambda exonuclease onto polymer micropillar arrays for the solid-phase digestion of dsDNAs. *Analytical chemistry* **2014,** *86* (9), 4447-4454.

63. Vaidyanathan, S.; Weerakoon-Ratnayake, K. M.; Uba, F. I.; Hu, B.; Kaufman, D.; Choi, J.; Park, S.; Soper, S. A., Thermoplastic nanofluidic devices for identifying abasic sites in single DNA molecules. *Lab on a Chip* **2021,** *21* (8), 1579-1589.

64. O'Neil, C.; Amarasekara, C. A.; Weerakoon-Ratnayake, K. M.; Gross, B.; Jia, Z.; Singh, V.; Park, S.; Soper, S. A., Electrokinetic transport properties of deoxynucleotide monophosphates (dNMPs) through thermoplastic nanochannels. *Analytica chimica acta* **2018,** *1027*, 67-75.

65. Japrung, D.; Henricus, M.; Li, Q.; Maglia, G.; Bayley, H., Urea facilitates the translocation of single-stranded DNA and RNA through the α-hemolysin nanopore. *Biophysical journal* **2010,** *98* (9), 1856-1863.

66. Cabrini, S.; Kawata, S., *Nanofabrication handbook*. CRC press: 2012.

67. Gao, H.; Tan, H.; Zhang, W.; Morton, K.; Chou, S. Y., Air cushion press for excellent uniformity, high yield, and fast nanoimprint across a 100 mm field. *Nano Letters* **2006,** *6* (11), 2438-2441.

68. Siwy, Z. S., Ion-current rectification in nanopores and nanotubes with broken symmetry. *Advanced Functional Materials* **2006,** *16* (6), 735-746.

69. Wang, C.; Ye, D.-K.; Wang, Y.-Y.; Lu, T.; Xia, X.-H., Insights into the “free state” enzyme reaction kinetics in nanoconfinement. *Lab on a Chip* **2013,** *13* (8), 1546-1553.

70. Sun, W.; Vallooran, J. J.; Zabara, A.; Mezzenga, R., Controlling enzymatic activity and kinetics in swollen mesophases by physical nano-confinement. *Nanoscale* **2014,** *6* (12), 6853-6859.

71. Amarasekara, C. A.; Rathnayaka, C.; Athapattu, U. S.; Zhang, L.; Choi, J.; Park, S.; Nagel, A. C.; Soper, S. A., Electrokinetic identification of ribonucleotide monophosphates (rNMPs) using thermoplastic nanochannels. *Journal of Chromatography A* **2021,** *1638*, 461892.

72. Goodwin, P. M.; Cai, H.; Jett, J. H.; Ishaug-Riley, S. L.; Machara, N. P.; Semin, D. J.; Orden, A. V.; Keller, R. A., Application of single molecule detection to DNA sequencing. *Nucleosides and Nucleotides* **1997,** *16* (5-6), 543-550.

73. Ambrose, W. P.; Goodwin, P. M.; Jett, J. H.; Johnson, M. E.; Martin, J. C.; Marrone, B. L.; Schecker, J. A.; Wilkerson, C. W.; Keller, R. A.; Haces, A., Application of single molecule detection to DNA sequencing and sizing. *Berichte der Bunsengesellschaft für physikalische Chemie* **1993,** *97* (12), 1535-1541.

74. Werner, J. H.; Cai, H.; Jett, J. H.; Reha-Krantz, L.; Keller, R. A.; Goodwin, P. M., Progress towards single-molecule DNA sequencing: a one color demonstration. *Journal of biotechnology* **2003,** *102* (1), 1-14.

Chapter 5. Conclusions and Future Directions

5.1 Conclusions

There is growing interest in the study of RNA modifications and their potential role(s) in many physiological abnormalities, such as modulating many processes associated with cancer development and progression.[1-4] While RNA sequencing (RNA-seq) using short-read next-generation sequencing (NGS) has proven to be a valuable tool for identifying specific RNA modifications, there remain technical challenges.[5-6] These technical challenges include: [1] a harsh bisulfite conversion generating RNA fragments followed by a required purification step; [2] an intermediate cDNA synthesis protocol to incorporate necessary 3' adaptors followed by a required purification step; [3] a PCR protocol (to incorporate 5' library barcodes) to create a heterogeneous pool of sample library templates for NGS sequencing that may require optimization while promoting potential sampling bias and still another required purification step; and [4] a quality control step to verify library enrichment with correct base-pair size distributions prior to sequencing. Importantly, using any of these pre-analytical steps, important RNA modification information can be lost.[5, 7-8]

In this dissertation, we have proposed a method for the single-molecule exosequencing of RNA molecules, coined as the Exonuclease Time-of-Flight (XToF) method, which has the potential to overcome the limitations in current sequencing technologies.[9-11] This is a transformative single-molecule sequencing technology specifically aimed at addressing each of these challenges by employing a sequencing-by-subtraction approach to detect single rNMPs (modified and unmodified) generated sequentially from full-length native RNA species. The workflow is intended to be simple, requiring only purified RNA as input without the need for bisulfite conversion, eliminating the need for cDNA synthesis and PCR, and incorporates an internal quality control RNA to verify system performance. Compared to existing single-molecule sequencing strategies, such as the Oxford Nanopore's MinION, XToF provides the following

unique attributes: [1] Uses multiple single-molecule identifiers (time-of-flight, ToF; current transient amplitudes) to identify nucleotides with unprecedented accuracy - >95%; [2] can process input samples, RNA, <1 ng negating the need for a PCR amplification step to sequence rare targets; [3] because XToF is a chip-based technology, it can be easily integrated to microscale sample processing units to provide fully automated processing – bona fide sample-to-sequencing platform; and [4] XToF chip is made in plastic to allow for high-scale production at low-cost using injection molding, even though the platform has nanometer structures.

Due to the recent COVID-19 pandemic, RNA sequencing has become increasingly important. SARS-CoV-2, which is the virus responsible for producing the COVID-19 disease, consists of a 30 kb RNA genome, and due to the evolutionary rate of RNA viruses, a rapid, amplification-free, and highly accurate RNA sequencing with long reads would be useful for the identification of variants to assist in the development of new vaccines targeted for these variants. These variants include the UK (B.1.1.7), South African (501Y.V2), and Brazilian (B.1.1.28) lineages.[12-14] For example, the B.1.1.7 UK variant consists of 17 mutations, of which 8 appear in the spike protein. The mutation rate in viruses is nearly 10^6 times that of the host and provides high viral adaptation to make the variant drug-resistant.[15] As we reported in this dissertation (please see Chapter 2), solid-phase XRN1 reactions demonstrated a processivity >10.5 kb, which provided a lower limit for the processivity as the RNA template (DMD RNA; 11.1 kb) used for the determination was entirely digested by the immobilized enzyme.[9] For producing the solid-phase enzymatic reaction, the solid-support is a thermoplastic that can be UV/O_3 activated with the enzyme attached through its lysine residue(s) using standard EDC/NHS coupling chemistry.[9] Therefore, our XToF technology may have the potential to read the complete SARs-CoV-2 genome in a single sequencing run but will depend on determining the upper limit on the processivity of XRN1 using longer RNA templates.

In the case of existing platforms for RNA sequencing, for example, the Oxford Nanopore MINION approach and that of Illumina (iSeq100™), amplicon-based sequencing of the SARS-CoV-2 genome was undertaken from clinical samples with a comparison of the sequencing metrics reported.[16] The results from this comparison study included the following: (i) Read length = 1,074 MINION, 140 iSeq100™; (ii) base error rate = 8.3% MINION, 0.2% iSeq100™; (iii) in both cases, reverse transcription followed by PCR was carried out; and (iv) total raw reads = 45,625 MINION, 3,990,760 iSeq100™. While sequencing experiments have not been performed at this time using XToF, it is instructive to compare the potential sequencing metrics for XToF: (i) read length = depends on processivity of the XRN1 enzyme, but our current data indicates this processivity to be >10 kb;[9] (ii) base error rate ≥99% based on ToF data secured for the rNMPs;[9] (iii) no reverse transcription or PCR required; and (iv) total number of raw reads will depend on the processivity of the XRN1 enzyme. As can be seen, XToF will offer some desirable sequencing metrics. Because it is a chip-based format, additional sample preparation units can be interfaced to XToF to automate the workflow fully. For example, we have recently shown that a microfluidic decorated with aptamers directed against the receptor-binding domain of the spike protein can select SARS-CoV-2 particles directly from clinical samples, such as saliva, with high specificity and recovery. This will allow for the enrichment of target sequencing material to improve data quality even under low viral load conditions and eliminate interferences arising from host nucleic acids.

5.2 Future Directions

We are now envisioning to take advantage of our experiences in generating XToF devices for solid-phase RNA digestion[9], nanoscale electrophoresis[17-19], and the isolation of liquid biopsy markers[20-21] to facilitate the development of unique technologies for analyzing and sequencing

single molecules of DNA and proteins. The XToF device can be programmed for a specific input (DNA, RNA, or protein) using the appropriate enzyme.

5.2.1 Converting XToF into a DNA Sequencer

We will investigate the ability to immobilize Exo 1 to the nanoscale enzymatic reactor in the XToF device to process ssDNA molecules in the 3' → 5' direction to generate single dNMPs. While we have reported the immobilization of λ-Exonuclease (λ-Exo) to pillars in an IMER to digest dsDNAs[11], the use of Exo 1 and its ability to process ssDNA can provide new applications, such as single-molecule sequencing using XToF not accommodated by λ-Exo.

Exo 1 is a processive enzyme that uses Mg^{2+} cofactor and cleaves ssDNA into single dNMPs.[22-23] As with most ssDNA exonucleases, Exo 1 maintains processivity by holding ssDNA in an electropositive cleft that is too narrow (<2 nm) to accommodate dsDNA.[24] The cleft can electrostatically bind to ssDNA even when the active site is inactive (no Mg^{2+}) but cannot maintain processivity through secondary structures, which do not fit within the binding cleft.[24-25] As such, in vivo Exo 1 is complexed with a single-stranded binding (SSB) protein (SSB; $K_d = 10^{-7}$ M).[26] SSB is a tetrameric protein that binds both ssDNA and Exo 1, sliding along the ssDNA substrate and unwinding secondary structures to maintain Exo 1 processivity as well as enhance clipping rates.[25] In solution, SSB-Exo 1 can clip ssDNA at ~120 nucleotides/s[25] and has a processivity ≥625 nucleotides.[22]

However, longer DNAs have not been tested to determine processivity precisely.[27-28] We will immobilize Exo 1 to the nanoscale enzymatic reactor in the XToF to determine its processivity and clipping rate when immobilized. We will first covalently link SSB onto the reaction pillar that has been UV/O_3 activated with EDC/NHS coupling chemistry as we did for XRN1. Then, Exo 1

will self-assemble onto the surface through the immobilized SSB (see Figure 5.1). This strategy will properly orient the Exo 1 active site towards solution and bind to ssDNA.

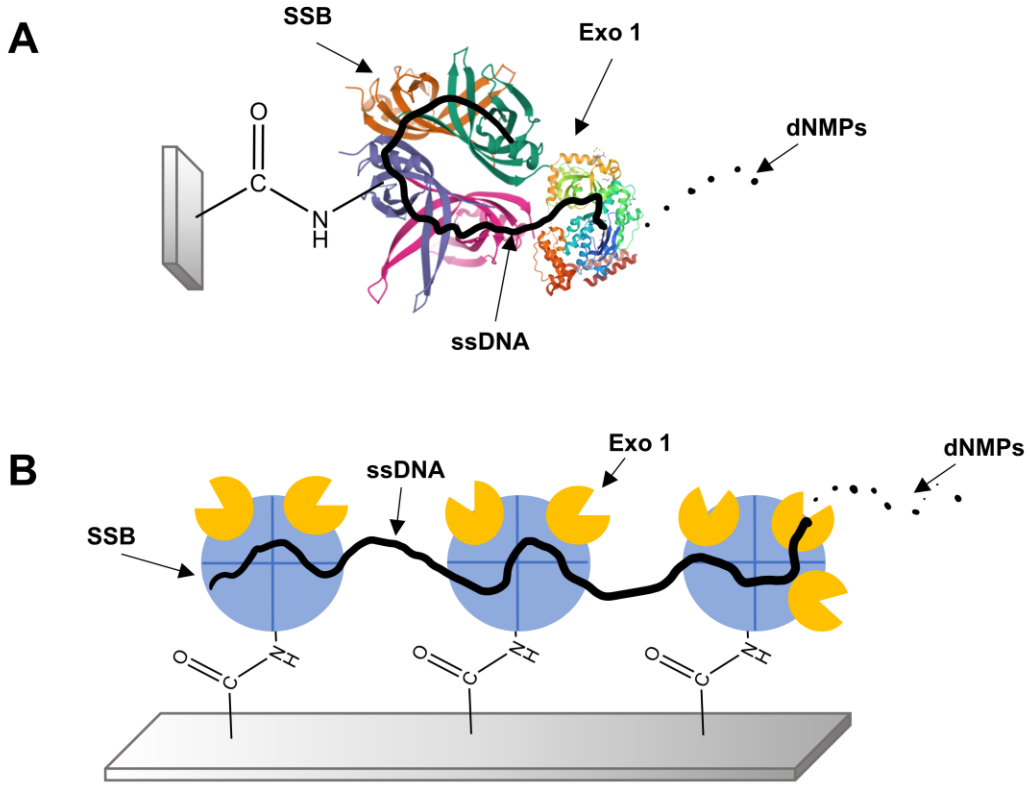


Figure 5.1. (A) Schematic representation of covalently immobilized single-stranded binding protein (SSB) to a UV/O_3 irradiated plastic support followed by assembly with Exo 1. **(B)** Cartoon showing Exo 1 assembly to SSB. SSB can bind up to 65 nt of ssDNA and 4 Exo 1 enzymes. Nanoelectrophoresis is used to interrogate reaction products (dNMPs) generated by the Exo 1 cleavage reaction.

We will measure kinetic clipping rates[29] of Exo 1/SSB in solution and solid-phase using the same methods we have reported in Chapter 2 (See Figure 2.11 and Figure 2.14). A labeled ssDNA (SYTO 82) substrate (7.249 kb, M13mp18 ssDNA substrate linearized, heat-denatured, and flash cooled) will be primed with Exo 1 and SSB without Mg^{2+} cofactor. When clipping is initiated with Mg^{2+} addition, we will simultaneously add a short (<100 nt) competitor labeled ssDNA (SYTO 60) in a high concentration that saturates excess Exo 1. The reaction will proceed, and we will monitor the fluorescence of the labeled substrate ssDNA with time to monitor the digestion rate and processivity of the ssDNA by analyzing fluorescence decay with time.

Measurements of solid-phase kinetics and processivity will employ SSB and Exo 1 immobilized to the XToF pillar. For the INERs, we will immobilize the SSB to the sub-micron pillar as we did for XRN1, followed by incubation with Exo 1 to form the complex. Exo 1 surface density and stability under electric field operation will be determined using an enzyme that is fluorescently labeled and monitored using our single-molecule fluorescence microscope. This will also allow us to determine the possible leaching of Exo 1 from the surface. Enzyme clipping rate and processivity will be determined by monitoring the real-time capture of stained ssDNA (SYTO 82) as well as observing the rate of ssDNA digestion after the introduction of Mg^{2+} by monitoring loss of fluorescence.

5.2.2 Sequencing of Single Protein Molecules using XToF

We will investigate the use of XToF for digesting single protein molecules into their constituent peptides. In these experiments, we will use trypsin as our model due to our success in using IMERs with trypsin[30-32] (see Figure 5.2). Trypsin is a serine protease that catalytically hydrolyzes proteins[33] by predominantly cleaving at the carboxyl side of lysine and arginine residues except when bound to a C-terminal proline[34].

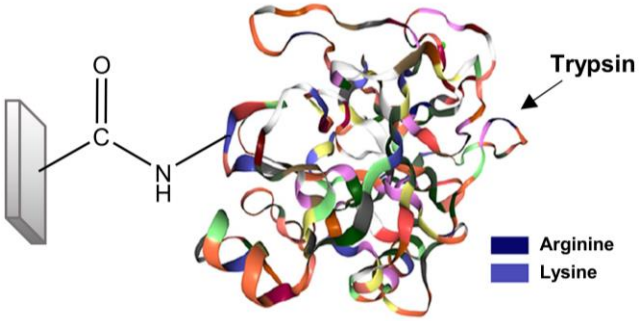


Figure 5.2. A schematic of the covalent immobilization of trypsin to a UV/O_3 activated polymer pillar. In this instance, a lysine residue is coupled to surface -COOH groups using EDC/NHS chemistry.

We have demonstrated the utility of IMERs coupled to MS for analysis of proteins using peptide mass fingerprinting with good sequence coverage as indicated by the favorable Mowse scores

for reaction times ranging from 4.8 to 48 s. In this case, the output of the IMERs was deposited onto a MALDI-MS plate with each spot containing ~6.6 pmol of peptide (3.9×10^{12} molecules).

We will use the fluidic architecture of XToF, and surface immobilize trypsin to the sub-micrometer pillar that has been UV/O_3 activated to create surface-confined carboxylic acid groups and then directly immobilizes trypsin using EDC/NHS coupling chemistry as we did for XRN1. The enzyme and reagents will be introduced into the XToF using the strategy delineated in Chapter 4 for enzyme immobilization around the bioreactor (see section 4.3.5). To assure immobilization of the enzyme to the pillar, we can covalently label trypsin with a fluorescent reporter prior to surface immobilization and monitor using fluorescence microscopy.

Following enzyme immobilization, we will electrokinetically introduce model proteins, such as cytochrome c, into the XToF device and interrogate the extent of reaction. As opposed to XRN1 and Exo 1, trypsin is not processive. Therefore, a batch-type reaction is used in which the substrate is introduced into the INER electrokinetically with the electric field terminated once the protein has entered the INER. We can detect the input of a single protein into the XToF using the input sensing in-plane pore that is placed before the input channel (see Chapter 4) to allow proper control of the driving electric field to load the protein.

We will couple the protein digestion in the XToF bioreactor to nanoscale electrophoresis in a 1D format with a nanocolumn (50 – 100 nm width and depth; >5 μm in length) that has been UV/O_3 activated to induce an EOF to allow driving the peptides from anode to cathode. Because we know the sequence of the peptides generated from the digestion of cytochrome c using trypsin (~18 peptide fragments produced), we can build histograms of the ToF for each peptide fragment to determine its molecular-dependent ToF. We will also test other proteins, such as BSA or myoglobin, which can be digested using an IMER with trypsin.[31] The reaction products can be

monitored using fluorescence by selective labeling of amino acid residues[35] or in a label-free format using the dual in-plane nanopore devices reported in Chapter 3.

We will use a model protein, in this case, Epithelial Cell Adhesion molecule, EpCAM[36-40], to test the utility of the XToF device for protein processing. EpCAM is an epithelial transmembrane glycoprotein that serves as an intercellular adhesion molecule and is overexpressed in a number of cancer cells, including OVCAR3. Upon trypsin digestion, it can generate ~20 peptides with molecular weights that range from 560 – 3441 Daltons. Single EpCAM protein molecules will be electrokinetically loaded into the chip possessing a trypsin anchored INER and nanoscale electrophoresis used to map the peptides created from the INER. Single protein molecules are loaded into the chip by applying a voltage across input and output nanochannels of the XToF. Following loading and sensing of the input into the bioreactor in XToF, the voltage is turned off, and the protein is allowed to digest in batch mode.

After the necessary reaction time, the electric field is applied to sweep the peptides into the nano-flight tube with two in-plane nanopores where the ToFs are deduced. From the number of signals generated and their ToF, we can generate a fingerprint corresponding to EpCAM. We will repeat this analysis for >1,000 events to generate the necessary statistics. For the EV-isolated proteins, single-molecule peptide fingerprints will be compared to the proteins cataloged in the proteomic analysis secured by Godwin *et al.*[41] From the database generated for the OVCAR3 EV membrane proteome, we can inspect the primary sequence of the proteins contained in the membrane of the EVs and with the use of trypsin, understand the number of peptides produced as a first-line identifier of the protein. Unique ToFs for each peptide will further provide identification of the protein.

5.2.3 Multidimensional Nanoscale Electrophoresis for the Identification of Products Generated in XToF Device

When sequencing DNA/RNA/proteins, large numbers of single molecules must be identified using their ToF. For example, for DNA, there are >5 different epigenetic modifications, and coupled to the 4 canonical bases, 9 components must be analyzed. In RNA, there are >100 various modifications, and for proteins, the number of distinct peptides proteolytically generated from a single protein molecule can be >20. As we have reported previously,[17], the separation selectivity for the 70mer and 50mer is 0.3 (based on the difference in their apparent mobilities and the average mobility of the pair). If we require a resolution of 6.0 to achieve an identification efficiency >99.75%, the total peak capacity (P_T) is 20. Therefore, we will seek to increase peak capacity using 2D nanoscale electrophoresis.

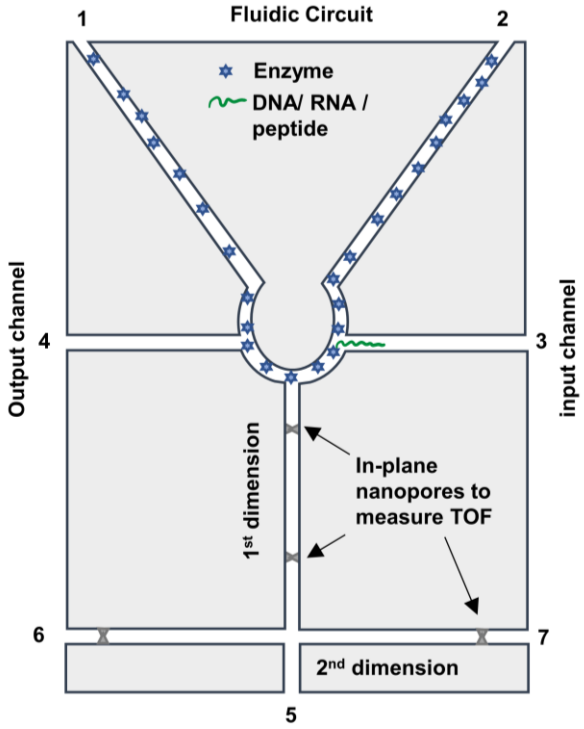


Figure 5.3. Schematic of single molecule 2D nanoscale electrophoresis unit to accept single molecules from the INERs. For the electrophoresis run, a drive voltage is applied between (1, 2) and (5) to introduce peptides into the 1st dimension column. The ToF is measured using the in-plane pores configured on both ends of this column. Once a current transient is detected at the second pore, the drive voltage is switched between (6) and (7) and the ToF measured between the last in-plane pore in the 1st dimension and the pore situated in the 2nd dimension column.

When performing multidimensional electrophoresis, the total peak capacity (P_T) can be approximated from the product of the individual peak capacities of each dimension.[42] For a 2D separation with $P_1 = 20$ and $P_2 = 20$, $P_T = 400$. Previously, we have developed 2D electrophoresis platforms using plastic microchips;[42-44] a peak capacity of 897 was achieved. For this study, we will use a 2D electrophoresis nanofluidic chip with an architecture shown in Figure 5.3, in which the 1st dimension channel will be >10 μm in length, and the 2nd dimension channel will be <5 μm in length. The nanofluidic channels in both dimensions will be 50-100 nm in width and depth. As shown in Figure 5.3, products can be fed into the nanoscale electrophoresis column from the INERs if required.

We will use our single-molecule fluorescence-tracking microscope to follow the transport of single molecules through the columns. To allow tracking, the molecules will be labeled with a fluorescent reporter, such as ATTO-532, which we have used to track the motion of dNMPs and rNMPs.[19] We understand the dye will impose a perturbation on the apparent mobility of the molecule. However, for a series of different molecules, the perturbation will be constant, and thus, differences in mobilities will be due to the molecular structure of the target molecule. For these experiments, we will test the canonical rNMPs and modified rNMPs. For the peptides, we will use bradykinin (1188 g/mol; pI = 12.51), C-peptide fragment 3-33 (3020 g/mol; pI = 2.89), C-type natriuretic peptide (2197 g/mol; pI = 9.05), fibronectin adhesion-promoting peptide (1023 g/mol; pI = 12.50), β-endorphin (3465 g/mol; pI = 10.5), and met-enkephalin (573 g/mol; pI = 9.46) as models.

We will use COC as the substrate material for the nanofluidic chip because it displays an extensive range of surface charge densities with different UV/O_3 or O_2 plasma dosing levels and the significant differences in the sessile water contact angles for non-exposed and exposed COC. For example, with no UV/O_3 dosing, the sessile water contact angle of COC 6013 is 98^{o,} and after

10 min dosing (16 mW/cm^2), the contact angle is ~37^o.[45] Therefore, for minimal dosing, the column can be considered hydrophobic, and the ToF is primarily determined by hydrophobic/hydrophilic interactions between the solute and column wall. A column receiving high dosing will contain high levels of anionic surface charges, and as such, the mobility is determined primarily by ion-exchange interactions. As a note, we can convert the anionic surface charges induced by deprotonating the surface carboxyl groups to a cationic surface by appending through EDC/NHS coupling chemistry ethylene diamine to the surface.[46]

For the 2D nano-electrophoresis, we can run each dimension at different pH values by using a different running buffer. Under high dosing conditions, a large EOF will be generated, inducing motion dominated by the EOF. For low dosing and a small EOF, we can use a running buffer with a pH greater than the pI of the components we are analyzing, so the mobility is dominated by the electrophoretic mobility of the peptides. As we have done for the dNMPs,[19] we will covalently label the peptides with a fluorescent reporter to free carboxyl or amine groups, such as ATTO-532 (anionic), or a neutrally charged reporter, such as a bodipy derivative,[47] and track the electrokinetic motion using our single-molecule fluorescence imaging microscope. We will investigate important figures-of-merit for the 2D nano-electrophoresis, such as the resolution for a series of molecules (series here is defined as the rNMPs or peptides) for each column type, and the actual peak capacity, which depends on the orthogonality of each electrophoretic dimension.[42] Because the potential peak capacity of 2D electrophoresis is the product of the peak capacities of the constituent dimensions, large peak capacities can be obtained even if the constituent dimensions produce only modest peak capacities.[48] However, 2D electrophoresis generates the theoretically available peak capacity only if the constituent dimensions are entirely orthogonal. A high degree of mobility correlation between the dimensions can reduce 2D electrophoresis to what is, in fact, 1D electrophoresis with peaks distributed along a diagonal plot of mobilities

between the constituent dimensions.[49] Therefore, for each molecule series tested, we will build a scatter plot of migration times for each dimension to determine the dimensional orthogonality.

5.3 References

1. Chandola, U.; Das, R.; Panda, B., Role of the N6-methyladenosine RNA mark in gene regulation and its implications on development and disease. *Briefings in functional genomics* **2015,** *14* (3), 169-179.

2. Kanwal, R.; Gupta, S., Epigenetic modifications in cancer. *Clinical genetics* **2012,** *81* (4), 303-311.

3. Motorin, Y.; Lyko, F.; Helm, M., 5-methylcytosine in RNA: detection, enzymatic formation and biological functions. *Nucleic acids research* **2010,** *38* (5), 1415-1430.

4. Yue, Y.; Liu, J.; He, C., RNA N6-methyladenosine methylation in post-transcriptional gene expression regulation. *Genes & development* **2015,** *29* (13), 1343-1355.

5. Corney, D. C., RNA-Seq using next generation sequencing. *Mater Methods* **2013,** *3* (203).

6. Schaefer, M.; Pollex, T.; Hanna, K.; Lyko, F., RNA cytosine methylation analysis by bisulfite sequencing. *Nucleic acids research* **2009,** *37* (2), e12-e12.

7. Ozsolak, F.; Platt, A. R.; Jones, D. R.; Reifenberger, J. G.; Sass, L. E.; McInerney, P.; Thompson, J. F.; Bowers, J.; Jarosz, M.; Milos, P. M., Direct RNA sequencing. *Nature* **2009,** *461* (7265), 814-818.

8. Garalde, D. R.; Snell, E. A.; Jachimowicz, D.; Sipos, B.; Lloyd, J. H.; Bruce, M.; Pantic, N.; Admassu, T.; James, P.; Warland, A., Highly parallel direct RNA sequencing on an array of nanopores. *Nature methods* **2018,** *15* (3), 201.

9. Athapattu, U. S.; Amarasekara, C. A.; Immel, J. R.; Bloom, S.; Barany, F.; Nagel, A. C.; Soper, S. A., Solid-phase XRN1 reactions for RNA cleavage: application in single-molecule sequencing. *Nucleic acids research* **2021,** *49* (7), e41-e41.

10. Novak, B. R.; Moldovan, D.; Nikitopoulos, D. E.; Soper, S. A., Distinguishing single DNA nucleotides based on their times of flight through nanoslits: a molecular dynamics simulation study. *The Journal of Physical Chemistry B* **2013,** *117* (12), 3271-3279.

11. Oliver-Calixte, N. J.; Uba, F. I.; Battle, K. N.; Weerakoon-Ratnayake, K. M.; Soper, S. A., Immobilization of lambda exonuclease onto polymer micropillar arrays for the solid-phase digestion of dsDNAs. *Analytical chemistry* **2014,** *86* (9), 4447-4454.

12. Voloch, C. M.; Jr, R. d. S. F.; Almeida, L. G. P. d.; Cardoso, C. C.; Brustolini, O. J.; Gerber, A. L.; Guimarães, A. P. d. C.; Mariani, D.; Costa, R. M. d.; Jr, O. C. F.; Cavalcanti, A. C.; Frauches,

T. S.; Mello, C. M. B. d.; Galliez, R. M.; Faffe, D. S.; Castiñeiras, T. M. P. P.; Tanuri, A.; Vasconcelos, A. T. R. d., A novel Brazilian SARS-CoV-2 lineage. *medRxiv* **2020**.

13. Tegally, H.; Wilkinson, E.; Giovanetti, M.; Iranzadeh, A.; Fonseca, V.; Giandhari, J.; Doolabh, D.; Pillay, S.; San, E. J.; Msomi, N.; Mlisana, K.; Gottberg, A. v.; Walaza, S.; Allam, M.; Ismail, A.; Mohale, T.; Glass, A. J.; Engelbrecht, S.; Zyl, G. V.; Preiser, W.; Petruccione, F.; Sigal, A.; Hardie, D.; Marais, G.; Hsiao, M.; Korsman, S.; Davies, M.-A.; Tyers, L.; Mudau, I.; York, D.; Maslo, C.; Goedhals, D.; Abrahams, S.; Laguda-Akingba, O.; Alisoltani-Dehkordi, A.; Godzik, A.; Wibmer, C. K.; Sewell, B. T.; Lourenço, J.; Carlos, L.; Alcantara, J.; Pond, S. L. K.; Weaver, S.; Martin, D.; Lessells, R. J.; Bhiman, J. N.; Williamson, C.; Oliveira, T. d., Emergence and rapid spread of a new severe acute respiratory syndrome-related coronavirus 2 (SARS-CoV-2) lineage with multiple spike mutations in South Africa. *medRxiv* **2020**.

14. Volz, E.; Mishra, s.; Chand, M.; Barrett, J. C.; Johnson, R.; Geidelberg, L.; Hinsley, W. R.; Laydon, D. J.; Dabrera, G.; O'Toole, A.; Amato, R.; Ragonnet-Cronin, M.; Harrison, I.; Jackson, B.; Ariani, C. V.; Boyd, O.; Loman, N.; McCrone, J. T.; Goncalves, S.; Jorgensen, D.; Myers, R.; Hill, V.; David K, J.; Gaythorpe, K.; Groves, N.; Sillitoe, J.; Kwiatkawski, D. P.; Flaxman, S.; Ratmann, O.; Bhatt, S.; Hopkins, S.; Gandy, A.; Rambaut, A.; Gerguson, N. M., Transmission of SARS-CoV-2 Lineage B.1.1.7 in England: Insight fron a linking epidemiological and genetic data. *medRxiv* **2021**.

15. Pachetti, M.; Marini, B.; Benedetti, F.; Giudici, F.; Mauro, E.; Storici, P.; Masciovecchio, C.; Angeletti, S.; Ciccozzi, M.; Gallo, R. C.; Zella, D.; Ippodrino, R., Emerging SARS-CoV-2 mutation hot spots include a novel RNA-dependent-RNA polymerase variant. *Journal of Translational Medicine* **2020,** *18*, 179.

16. Hourdel, V.; Kwasiborski, A.; Balière, C.; Matheus, S.; Batéjat, C. F.; Manuguerra, J.-C.; Vanhomwegen, J.; Caro, V., Rapid Genomic Characterization of SARS-CoV-2 by Direct Amplicon-Based Sequencing Through Comparison of MinION and Illumina iSeq100TM System. *Frontiers in Microbiology* **2020,** *11*, 1-9.

17. Amarasekara, C. A.; Athapattu, U. S.; Rathnayaka, C.; Choi, J.; Park, S.; Soper, S. A., Open-tubular nanoelectrochromatography (OT-NEC): gel-free separation of single stranded DNAs (ssDNAs) in thermoplastic nanochannels. *Electrophoresis* **2020,** *41* (18-19), 1627-1640.

18. Amarasekara, C. A.; Rathnayaka, C.; Athapattu, U. S.; Zhang, L.; Choi, J.; Park, S.; Nagel, A. C.; Soper, S. A., Electrokinetic identification of ribonucleotide monophosphates (rNMPs) using thermoplastic nanochannels. *Journal of Chromatography A* **2021,** *1638*, 461892.

19. O'Neil, C.; Amarasekara, C. A.; Weerakoon-Ratnayake, K. M.; Gross, B.; Jia, Z.; Singh, V.; Park, S.; Soper, S. A., Electrokinetic transport properties of deoxynucleotide monophosphates (dNMPs) through thermoplastic nanochannels. *Analytica chimica acta* **2018,** *1027*, 67-75.

20. Witek, M. A.; Aufforth, R. D.; Wang, H.; Kamande, J. W.; Jackson, J. M.; Pullagurla, S. R.; Hupert, M. L.; Usary, J.; Wysham, W. Z.; Hilliard, D., Discrete microfluidics for the isolation of circulating tumor cell subpopulations targeting fibroblast activation protein alpha and epithelial cell adhesion molecule. *NPJ precision oncology* **2017,** *1* (1), 1-11.

21. Jackson, J. M.; Witek, M. A.; Soper, S. A., Sinusoidal microchannels with high aspect ratios for CTC selection and analysis. *Circulating Tumor Cells: Isolation and Analysis* **2016**, 85-126.

22. Brody, R.; Doherty, K.; Zimmerman, P., Processivity and kinetics of the reaction of exonuclease I from Escherichia coli with polydeoxyribonucleotides. *Journal of Biological Chemistry* **1986,** *261* (16), 7136-7143.

23. Thomas, K. R.; Olivera, B. M., Processivity of DNA exonucleases. *Journal of Biological Chemistry* **1978,** *253* (2), 424-429.

24. Korada, S. K. C.; Johns, T. D.; Smith, C. E.; Jones, N. D.; McCabe, K. A.; Bell, C. E., Crystal structures of Escherichia coli exonuclease I in complex with single-stranded DNA provide insights into the mechanism of processive digestion. *Nucleic acids research* **2013,** *41* (11), 5887-5897.

25. Lu, D.; Myers, A. R.; George, N. P.; Keck, J. L., Mechanism of Exonuclease I stimulation by the single-stranded DNA-binding protein. *Nucleic acids research* **2011,** *39* (15), 6536-6545.

26. Lu, D.; Keck, J. L., Structural basis of Escherichia coli single-stranded DNA-binding protein stimulation of exonuclease I. *Proceedings of the National Academy of Sciences* **2008,** *105* (27), 9169-9174.

27. Clarke, J.; Wu, H.-C.; Jayasinghe, L.; Patel, A.; Reid, S.; Bayley, H., Continuous base identification for single-molecule nanopore DNA sequencing. *Nature nanotechnology* **2009,** *4* (4), 265-270.

28. Maitra, R. D.; Dunbar, W. B., Determining Optimal Voltage Inputs For Exonuclease I Experimentation from Single-Abasic DNA Strands Captured in a Nanopore. *Biophysical Journal* **2012,** *102* (3), 204a-205a.

29. Han, E. S.; Cooper, D. L.; Persky, N. S.; Sutera Jr, V. A.; Whitaker, R. D.; Montello, M. L.; Lovett, S. T., RecJ exonuclease: substrates, products and interaction with SSB. *Nucleic acids research* **2006,** *34* (4), 1084-1091.

30. Lee, J.; Soper, S. A.; Murray, K. K., A solid-phase bioreactor with continuous sample deposition for matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. *Rapid Communications in Mass Spectrometry* **2011,** *25* (6), 693-699.

31. Lee, J.; Musyimi, H. K.; Soper, S. A.; Murray, K. K., Development of an automated digestion and droplet deposition microfluidic chip for MALDI-TOF MS. *Journal of the American Society for Mass Spectrometry* **2011,** *19* (7), 964-972.

32. Lee, J.; Soper, S. A.; Murray, K. K., Microfluidic chips for mass spectrometry-based proteomics. *Journal of mass spectrometry* **2009,** *44* (5), 579-593.

33. Rawlings, N. D.; Barrett, A. J., [2] Families of serine peptidases. *Methods in enzymology* **1994,** *244*, 19-61.

34. Rodriguez, J.; Gupta, N.; Smith, R. D.; Pevzner, P. A., Does trypsin cut before proline? *Journal of proteome research* **2008,** *7* (01), 300-305.

35. Swaminathan, J.; Boulgakov, A. A.; Hernandez, E. T.; Bardo, A. M.; Bachman, J. L.; Marotta, J.; Johnson, A. M.; Anslyn, E. V.; Marcotte, E. M., Highly parallel single-molecule identification of proteins in zeptomole-scale mixtures. *Nature biotechnology* **2018,** *36* (11), 1076-1082.

36. Alaiya, A. A.; Franzén, B.; Auer, G.; Linder, S., Cancer proteomics: from identification of novel markers to creation of artifical learning models for tumor classification. *ELECTROPHORESIS: An International Journal* **2000,** *21* (6), 1210-1217.

37. Crisan, D.; Ruark, D. S.; Decker, D. A.; Drevon, A. M.; Dicarlo, R. G., Detection of circulating epithelial cells after surgery for benign breast disease. *Molecular diagnosis* **2000,** *5* (1), 33-38.

38. Meng, S.; Tripathy, D.; Shete, S.; Ashfaq, R.; Haley, B.; Perkins, S.; Beitsch, P.; Khan, A.; Euhus, D.; Osborne, C., HER-2 gene amplification can be acquired as breast cancer progresses. *Proceedings of the National Academy of Sciences* **2004,** *101* (25), 9393-9398.

39. Simon, B.; Podolsky, D. K.; Moldenhauer, G.; Isselbacher, K. J.; Gattoni-Celli, S.; Brand, S. J., Epithelial glycoprotein is a member of a family of epithelial cell surface antigens homologous to nidogen, a matrix adhesion protein. *Proceedings of the National Academy of Sciences* **1990,** *87* (7), 2755-2759.

40. Szala, S.; Froehlich, M.; Scollon, M.; Kasai, Y.; Steplewski, Z.; Koprowski, H.; Linnenbach, A. J., Molecular cloning of cDNA for the carcinoma-associated antigen GA733-2. *Proceedings of the National Academy of Sciences* **1990,** *87* (9), 3542-3546.

41. Liang, B.; Peng, P.; Chen, S.; Li, L.; Zhang, M.; Cao, D.; Yang, J.; Li, H.; Gui, T.; Li, X., Characterization and proteomic analysis of ovarian cancer-derived exosomes. *Journal of proteomics* **2013,** *80*, 171-182.

42. Shadpour, H.; Soper, S. A., Two-dimensional electrophoretic separation of proteins using poly (methyl methacrylate) microchips. *Analytical chemistry* **2006,** *78* (11), 3519-3527.

43. Osiri, J. K.; Shadpour, H.; Park, S.; Snowden, B. C.; Chen, Z. Y.; Soper, S. A., Generating high peak capacity 2-D maps of complex proteomes using PMMA microchip electrophoresis. *Electrophoresis* **2008,** *29* (24), 4984-4992.

44. Sweedler, J.; Shear, J.; Fishman, H.; Zare, R. N.; Scheller, R., Fluorescence detection in capillary zone electrophoresis using a charge-coupled device with time-delayed integration. *Analytical chemistry* **1991,** *63* (5), 496-502.

45. O'Neil, C. E.; Taylor, S.; Ratnayake, K.; Pullagurla, S.; Singh, V.; Soper, S. A., Characterization of activated cyclic olefin copolymer: effects of ethylene/norbornene content on the physiochemical properties. *Analyst* **2016,** *141* (24), 6521-6532.

46. Uba, F. I.; Pullagurla, S. R.; Sirasunthorn, N.; Wu, J.; Park, S.; Chantiwas, R.; Cho, Y.-K.; Shin, H.; Soper, S. A., Surface charge, electroosmotic flow and DNA extension in chemically modified thermoplastic nanoslits and nanochannels. *Analyst* **2015,** *140* (1), 113-126.

47. Loudet, A.; Burgess, K., BODIPY dyes and their derivatives: syntheses and spectroscopic properties. *Chemical reviews* **2007,** *107* (11), 4891-4932.

48. Cortes, H. J., *Multidimensional chromatography: techniques and applications*. CRC Press: 2020.

49. Venkatramani, C.; Xu, J.; Phillips, J. B., Separation orthogonality in temperature-programmed comprehensive two-dimensional gas chromatography. *Analytical chemistry* **1996,** *68* (9), 1486-1492.