Investigation of the Formation of some Biologically Relevant Small Molecules Using Laser Tweezers and Capillary Electrophoresis

Dissertation Submitted to:
Department of Chemistry and Biochemistry
Kent State University, Kent, OH 44242
In Partial Fulfilment of the Requirements for Doctor of Philosophy

By

Philip M. Yangyuoru
August, 2014
A Dissertation Written by

Philip M. Yangyuoru

B.S., University of Ghana, Ghana, 2002
M.S., Illinois State University, USA, 2007
Ph.D., Kent State University, USA, 2014

Approved by

___________________________________, Chair, Doctoral Dissertation Committee
Hanbin Mao, Ph. D

___________________________________, Member, Doctoral Dissertation Committee
Songping Huang, Ph. D

___________________________________, Member, Doctoral Dissertation Committee
Michael Tubergen, Ph. D

___________________________________, Member, Doctoral Dissertation Committee
Gail C. Fraizer, Ph. D

Accepted by

___________________________________, Chair, Dept. of Chemistry & Biochemistry
Michael Tubergen, Ph. D

___________________________________, Dean, College of Arts and Sciences
James Blank, Ph.D
# TABLE OF CONTENTS

LIST OF FIGURES ........................................................................................................... V  

LIST OF TABLES ............................................................................................................. VII  

ACKNOWLEDGEMENTS ............................................................................................. VIII  

ABSTRACT ....................................................................................................................... 1  

CHAPTER 1.0 INTRODUCTION ................................................................................. 4  

1.1 Aptamers and G-Quadruplexes ............................................................................. 4  

1.2 Advantages and Challenges in Use of Aptamers ................................................. 8  

1.3 Aptamers as Sensors .......................................................................................... 9  

1.4 Bioanalytical Applications of Aptamers .............................................................. 11  

1.5 The Role G-Quadruplex Structures in Therapeutics ......................................... 13  

1.6 Measurement of Chemical Binding Affinity ($k_d$) ............................................ 14  

1.7 Mechanochemistry of Binding ......................................................................... 16  

1.8 Effects of Force and Transition distance on Mechanical Stability .................... 18  

CHAPTER 2.0 MATERIALS AND METHODS ......................................................... 21  

2.1 Molecular Biology ............................................................................................ 21  

2.1.1 Preparation of DNA Handles ...................................................................... 21  

2.1.2 Synthesis of DNA Constructs for Single-Molecule Experiments ............... 22
2.1.3 RNA Construct Preparation for Single-Molecule Experiments ........................24

2.2 Laser Tweezers Experimental Setup............................................................26

2.3 Single-Molecule Force-Ramp Experiments .................................................27

2.4 Calculation of Formation Probability of a Species ........................................30

2.5 Force-Jump Experiments ............................................................................31

2.6 Contour Length Calculations .......................................................................31

2.7 Determination of Unfolding Free Energy .....................................................31

2.8 Population Deconvolution at Nanometer Resolution (PoDNano) .................32

2.9 Capillary Electrophoresis ...........................................................................33

2.10 CD/UV Melting Experiments .....................................................................34

CHAPTER 3 Single-Molecule Measurements of the Binding between Small
Molecules and DNA Aptamers ........................................................................37

3.1 Summary......................................................................................................37

3.2 Introduction..................................................................................................39

3.3 Results ........................................................................................................41

    3.3.1 The Mechanical Stability of an Aptamer is Enhanced by Small-Molecule
        Binding ........................................................................................................41

    3.3.2 Dissociation Constant Determination at only one Concentration of a Ligand
        .....................................................................................................................45
3.3.3 Determination of Dissociation Constant by CE ........................................48

3.4 Discussion .............................................................................................................53

3.5 Conclusion ............................................................................................................56

CHAPTER 4 Mechanochemical Properties of Individual Human Telomeric RNA 
(TERRA) G-Quadruplexes .................................................................................57

4.1 Summary .............................................................................................................57

4.2 Introduction .........................................................................................................58

4.3 Results and Discussion .......................................................................................60

4.3.1 RNA G-Quadruplexes Form in the TERRA Sequence ...............................60

4.3.2 Multiple Species are Present in the TERRA ..............................................69

4.3.3 Partially Folded Structures are Intermediates to the RNA G-Quadruplexes
...............................................................................................................................72

4.3.4 Comparison between Human Telomeric DNA and RNA G-Quadruplexes
...............................................................................................................................76

4.4 Conclusion .........................................................................................................77

4.5 Supporting Information ......................................................................................78

CHAPTER 5 Determination of Sphingosine Kinase 2 Activity using Fluorescent
Sphingosine by Capillary Electrophoresis ...............................................................88

5.1 Summary .............................................................................................................88

5.2 Introduction .........................................................................................................90

5.3 Results and Discussion .......................................................................................94

5.3.1 Separation of the In Vitro SpFl (substrate) and S1PFl (product) by CE.....94
5.3.2 Phosphorylation of SpFl by SphK2 ........................................100
5.3.3 Inhibition of SphK2 ...............................................................101
5.3.4 In Vitro Inhibition of SphK1 and SphK2 .................................104
5.3.5 Inhibition of SphK(s) in Cells ...............................................106
5.4 Conclusion .............................................................................108
5.5 Experimental Section .............................................................109
5.5.1 Materials ........................................................................109
5.5.2 Capillary Electrophoresis .....................................................110
5.5.3 Sample Preparation ............................................................111
5.5.4 Sphingosine Kinase 2 In Vitro Assay .....................................111
5.5.5 Cell Culture ....................................................................112
5.5.6 Michaelis-Menten Kinetics .................................................123

CHAPTER 6 CONCLUDING REMARKS...........................................114
REFERENCES ............................................................................116

List of Figures

Figure 1. Common types of secondary structural conformation formed by aptamers ......6
Figure 2. Energy landscape of unfolding a rigid or flexible structure of a biomolecule ..19
Figure 3. Flow chart for making DNA constructs for single-molecule experiments ......23
Figure 4. Flow chart for making RNA construct and DNA/RNA hybrid handles ........25
Figure 5. Schematics of laser tweezers instrument used for single-molecule experiment .................................................................27
Figure 6. Schematic of a three-channel microfluidic reaction chamber for performing single-molecule laser tweezers experiments ..................................................28
Figure 7. Schematic of single-molecule force-ramp experiments ................................30
Figure 8. Schematic Diagram of Capillary Electrophoresis Instrument ......................34
Figure 9. Schematic of the experimental setup and a typical force-extension curve ......42
Figure 10. Unfolding force histograms of the ATP aptamer ........................................44
Figure 11. Histograms of the work done to unfold free aptamer .................................47
Figure 12. The Hess-like cycle of the aptamer-ATP binding process .........................48
Figure 13. Electropherograms for 0.5 nM aptamer with varying ATP concentrations...49
Figure 14. Electropherograms of a 0.5 nM scrambled control sequence .....................51
Figure 15. Mechanical unfolding and refolding of TERRA ......................................61
Figure 16. The PoDNano analyses on the change in contour length (ΔL) upon unfolding the structures in the TERRA ..............................................................63
Figure 17. Plots of the change in contour length (ΔL) vs force (F) based on the F-X curves ..................................................................................................................64
Figure 18. Biophysical characterization of TERRA-4G .............................................66
Figure 19. End-to-end distance measurements .........................................................68
Figure 20. Rupture force analyses of TERRA construct ...........................................71
Figure 21. Transition kinetics of the TERRA species along the 5′-3′ trajectory ..........74
Figure 22. Electropherogram of separation of and SpFl, BodipyFl and S1PF1 ............97
Figure 23. Electropherogram showing the effect of Mg^{2+} ions on separation ..........98
Figure 24. Lineweaver-Burk plot of initial rate versus SpFl substrate concentration ...99
Figure 25. Percent phosphorylation of SpFl with time in 300 µL assay buffer ...........101
Figure 26. A plot showing the SphK2 catalyzed phosphorylation of SpFl ...............104
Figure 27. Inhibitions of SphK2 and SphK1 ...............................................................105
Figure 28. CE analysis of the inhibition of SpFl phosphorylation in NIH-3T3 cells ....107
List of Tables

Table 1. $\Delta L$ Calculation According to Equation 1 in the main text ($\Delta L = L - x$) ………..86

Table 2. Rate constants ($k_1$ through $k_6$, [Average $\pm$ SD], $s^{-1}$), distance between the folded and the transition state ($x^*$, nm), unfolding force ($F_{\text{rupture}}$, $pN$), change in free energy of the 5′-3′ unfolding of GQ ($\Delta G_{\text{unfold-GQ}}$, kcal/mol) for human telomeric RNA (top) and DNA (bottom) species…………………………………………………………………………………………………………………..87

Table 3. Average elution times, and the percent relative standard deviation (%RSD) of peak heights, and peak areas, for ten sequential runs of the SpFl and S1PFl standards separated by CE-LIF using 3 different buffers …………………………………………………….95

Table 4. Average elution times and normalized peak areas and heights of SpFl and S1PFl for the in vitro assay using BodipyFl C16 as internal standard. (n = 8). SpFl gave two peaks as a result of isomer separation ……………………………………………..96
Acknowledgements

I wish to express my sincere gratitude to Kent State University and the department of Chemistry and Biochemistry for providing financial support through graduate assistantship, instruments and research facilities throughout my Ph.D. program. The department is a miniature world with diverse faculty and students from all over the globe with different cultural backgrounds. Such a warm environment both on campus and in the department makes up for the frequent piles of snow and the sometimes long and never-ending winters.

Special thanks to my research advisor, Dr. Mao for his constant provision of research ideas and guidelines to enable me carry out successful research throughout the years. Most importantly, for his scientific ideas, proactive and hands-on approach to solving research-related hurdles and being available all the time made my research experience very productive. His common question, what is the worst case scenario? Is a constant reminder to remain open minded and calm and not get too excited or discouraged as heartbreaks are part of the tools of the trade in scientific research.

Thank you to Dr. Mwongela, my first advisor in Kent State University for his valuable time and research direction for my first $2^{1/2}$ years and for getting me started with research in the Chemistry department at Kent State University. I am also very grateful to my committee members Drs. Michael Tubergen, Songping Huang and Gail C. Fraizer for their precious time, valuable comments and suggestions on my proposal and dissertation and for every support offered me. I truly appreciate it.
To all the researchers in the Mao group (both former and present). I am particularly grateful to Soma Dhakal, Zhongbo Yu, Zhe Shi and Deepak Koirala for their help, support and valuable contributions to projects we worked on together. I am also grateful to Yunxi Cui, Chiran Ghimire, Prakash Shrestha, Sangeetha Selvam, Jibin Punnoose and Shankar Mandal for being great and resourceful group members and for their contributions in many ways to enable research in the lab run smoothly.

Thank you to Dr. Mahinda Gangoda for helping me run and troubleshoot instruments used for teaching Analytical labs. To Kristen Camputaro, the stock room manager for making sure I got all the instruments, chemicals and materials to teach my labs, and a big thank you to Erin Michael, Graduate Program Administrator for taking care of all the paperwork, information and updates.

Last but most importantly, thanks to my wife for taking care of our son Earl and for putting up with my long hours in the lab, fishing or treating single-molecules data and for being a constant reminder to me that there is life outside of the lab. My sincere gratitude to all my siblings for their every support and for keeping me motivated.
Abstract

The interaction between nucleic acids and small molecule ligands is continuously generating significant interest due to their widespread biological and bioanalytical applications. We investigated the mechanical property of the binding between aptamers and small molecules. Using an ATP binding aptamer as an example, we observed that the mechanical stability of the aptamers that are bound with ATP is higher than those without a ligand. Therefore, a force-based sensor can be developed to detect small molecules using aptamers as a platform. We determined the dissociation constant, $K_d$, for aptamer-ligand interactions at the single-molecule level by applying a Hess-like cycle. Our experiments allow the $K_d$ determination from only one ligand concentration which was further validated by our capillary electrophoresis (CE) method. By using only one ligand concentration, such a method not only saves time and material, but also is less susceptible to reduced reproducibility due to run-to-run fluctuations.

G-quadruplex forming sequences which are widespread in the genome particularly in telomeres and promoter regions have been shown to be therapeutic targets. G-quadruplex structures have been extensively studied using mostly conventional methods. However, the mechanical stability, thermodynamics, kinetics properties of these interactions at the single-molecule level, remains to be fully understood. While the formation of these G-quadruplex structures is highly dynamic, they are mechanically stabilized upon ligand binding which may affect their biological functions. Small molecules which bind to nucleic acid structures may interfere with vital cellular processes.
such as transcription and protein translation during cell division by acting as energy barriers or mechanical blockage. Therefore, understanding stability of nucleic structures from a mechanical stability standpoint is critical to fully explore their therapeutic potentials.

We have investigated the human telomeric repeat containing RNA (TERRA) G-quadruplex structures. By using a TERRA sequence that hosts only one G-quadruplex at the single-molecule level in a laser-tweezers instrument, we are able to investigate intramolecular G-quadruplex without implications from high order structures due to the tendency of the TERRA sequence to self-associate. Using a mechanical unfolding and refolding method, we revealed multiple TERRA G-quadruplexes in a solution mixture. One of the structures is consistent with the parallel conformation determined by NMR/X-ray techniques. In addition, we identified a partially folded structure that serves as an intermediate to the unfolding and refolding of the TERRA G-quadruplex.

By applying a force-jump approach, the refolding kinetics of the TERRA G-quadruplex and the intermediate were measured and compared with those of human telomeric DNA (hTelo) G-quadruplex. We have found that formation of the TERRA G-quadruplex is slower than that of hTelo G-quadruplex. Together with the higher thermodynamic and mechanical stability of the TERRA with respect to the hTelo G-quadruplex, these results suggest different regulatory capacities of RNA and DNA G-quadruplexes for processes associated with human telomeres.
Lastly, a quantitative assay was developed to investigate sphingosine kinase 2 (SphK2) activity both \textit{in vitro} and with cell lysates, using capillary electrophoresis with laser-induced fluorescence (CE-LIF) and sphingosine fluorescein as the substrate. Sphingosine and sphingosine-1-phosphate are found in very minute quantities in cells and therefore require highly sensitive techniques for quantitative analysis.
Chapter 1

1.0 Introduction

Parts of the last 3 sections (1.6-1.8) of this chapter have been published in Reviews in Analytical Chemistry (DOI 10.1515/revac-2013-0004) and used with permission. Copyright © 2011–2013 by Walter de Gruyter GmbH.

1.1 Aptamers and G-Quadruplexes

Apart from storing and transmitting genetic information, nucleic acids also interact with other compounds and are important in the recognition of other biomolecules. To explore the bio-recognition property of nucleic acids, certain nucleic acids have been raised to have high selectivity and specificity for their targets. Also called aptamers, these specific nucleic acids which are usually single stranded and may vary in length from 10-100 bp, have been shown to bind not only nucleic acid but also proteins, cells and small organic molecules. Aptamers have comparable properties and binding affinities in the nanomolar to picomolar range to antibodies and serve as alternatives to antibodies in antibody-based diagnostics and research. The technique used to select functional DNA or RNA aptamers for a particular target from nucleic acid combinatorial libraries is systematic evolution of ligands by exponential enrichment (SELEX). First started in 1990s the selection and development of aptamers for targets
has progressed considerably. With nucleic acid aptamers selected for more than one hundred targets, aptamers can now be routinely generated for almost any kind of target. Recently, through a process called cell-Selex, aptamers have been raised to target certain cell membranes in their native states. A general method for engineering split DNA aptamers using three-way junction aptamers to selectively bind small target molecules has been described to show a significant potential for use in biosensing and biorecognition. There is on-going research to broaden and advance the field of aptamer design and engineering to target a wider scope of molecules and biomarkers for biomedical research, diagnostics and imaging.

Aptamers are capable of adapting different structural conformations that allow them to form different shapes such as hairpins (Figure 1A), bulge (Figure 1B), forming binding pockets for their target molecules (Figure 1). One such secondary structure formed as a result of conformational changes in aptamers is G-quadruplex structures (Figure 1C). Nucleic acid sequences with tandem G-rich repeats can fold into tetraplex structures called G-quadruplexes (GQ), made up of a stack of four Guanines interconnected by Hoogsteen base pairs and stabilized by monovalent cations. The stability of the GQ structure depends on the length of the sequence, the loops the stabilizing metal ion, and strand alignment. GQs can compete with DNA duplex formation in vivo because they are known to exhibit melting temperatures greater than 70 °C under physiological conditions and are favorably formed over the duplex under molecularly crowded conditions.
G-quadruplex-forming nucleic acid sequences are under intense investigation in recent years from different perspectives including biological, chemical and physical due to their structural stability in physiological conditions. Guanine-rich sequences are predominantly located at the ends of chromosome in the telomeres. As telomeres maintain the stability of the genome during replication, the existence of G-quadruplex structures has been demonstrated both in vitro and in cells and maybe involved in important biological processes such as cancer regulation. Small-molecule ligands can stabilize the G-quadruplex structures upon binding and prevent cell growth by inhibiting
telomerase function. The design and screening of small-molecule ligands which can selectively regulate the stability of GQs in vitro is currently being pursued.

DNA aptamers compared to their RNA counterparts are more stable towards nucleases, but they can be chemically modified to increase their stability. On the other hand, RNA aptamers are able to form more ligand binding scaffolds and are capable of being transcribed in vivo by cells whereas DNA aptamers have to be introduced externally.

Conventional methods such as fluorescence, CD, and NMR, are most commonly used to determine only the bulk average of all mixed species. However, the coexistence of the nucleic acids DNA/RNA in ssDNA, dsDNA, triplex, hairpins and G-quadruplex structures in the same solution may result in very complex mixed structures. These highly complex systems could be rather challenging for the ensemble methods to distinguish individual species, let alone follow detailed transition events between them. Single molecule experiments produce discrete signals that can be statistically analyzed to reveal the property of sub-groups in a population and the energetics of a particular reaction trajectory in a transition process. Single-molecule approaches by nature, can interrogate one molecule at a time, and hence, the most applicable techniques to deconvolute and probe individual nucleic acid species in a mixture. There are many methods and techniques used to study single molecules. Force-based techniques, such as atomic force microscope (AFM), optical tweezers, micro-needles, and magnetic tweezers, provide a unique capabilities to investigate the force generated by or applied to individual molecules. More appropriately, we designed a single-molecule
optical method using laser tweezers to carefully analyze and scrutinize nucleic acid structures in solution on a molecule-by-molecule basis. Such measurements allow for direct interrogation of the microscopic properties of individual molecules within a subpopulation. The mechanical unfolding and refolding and force-jump methods are therefore very useful techniques to probe structures, measure mechanical and thermodynamic stabilities, and follow transition kinetics for the formed structures in solution.

1.2 Advantages and Challenges in Use of Aptamers

Aptamers are similar to antibodies in their applications in diagnosis and target recognition. However, aptamers have the added advantages of being more stable and able to tolerate high pH and temperature ranges than antibodies. Being single stranded and self-folding, aptamers do not aggregate due to the lack of the hydrophobic protein cores as in antibodies. Aptamers are easier and less expensive to produce and can be readily modified\textsuperscript{37} chemically by attaching appropriate reporters and functional groups or even combined with antibodies, making it amenable to different detection methods. Production of aptamers can easily be scaled up as it does not require animal, bacteria or cell cultures which may limit their production for the detection of many different targets.\textsuperscript{37} Due to their smaller size, aptamers have improved tissue penetration,\textsuperscript{38} allowing for better cell specific targeting.
Through the process called cell-SELEX, aptamers are exponentially selected and enriched to target certain membrane difference between cell lines in their native states which has the potential to specifically target cancer cells.\textsuperscript{39,40} This is a significant development in biomarker discovery as cancer cells can be targeted without prior knowledge of the biomarker proteins expressed by the diseased cell\textsuperscript{6} as is the case with monoclonal antibodies. Apart from cancer cells, different strains of cells including viral and bacterial can be specifically targeted by aptamers with high affinity.\textsuperscript{41} The main challenge to the application of aptamers in medical diagnosis their instability in serum and in vivo conditions due to nuclease degradation. To tackle this nuclease degradation problem certain moieties in nucleic acid structure can be changed. e.g. modifying the sugars\textsuperscript{6} on nucleic or the phosphodiester\textsuperscript{42} backbone may prevent enzymatic recognition and slow down the degradation process. Delivery and low biocompatibility are also some of the problems associated with the advancement of use in aptamers in therapeutics. These can be addressed by encapsulating or conjugating aptamers to cholesterol,\textsuperscript{43} polyethylene glycol\textsuperscript{44} and gels\textsuperscript{45} to enhance delivery and biostability.

1.3 Aptamers as Sensors

Nucleic acids play important roles in molecular recognition and biological applications by storing and transmitting information, regulation of chemical activities, provide mechanical support and movement of small molecules. To recognize and transport biomolecules, these oligonucleotides serve as affinity ligands to their target molecules. Most common affinity ligands are antibodies and aptamers with aptamers
having the additional advantages of being more thermally stable, longer shelf life, easy production and modification than antibodies.\textsuperscript{37} Aptamers unlike antibodies fold into stable secondary structural conformations,\textsuperscript{46} forming binding pockets for their targets. The change in conformation is usually incorporated into signal transducers and reporters. Aptamers are used in different formats for many diagnostic applications and as biosensors. They may also serve as affinity ligands by coupling to surfaces in ELISA formats as substitutes for antibodies for the detection of target proteins with comparable results,\textsuperscript{47} or better sensitivity\textsuperscript{48} and on column surfaces as affinity chromatography for separation or purification of target molecules or capture ligands on the surface of sensors and microarrays.\textsuperscript{49} Aptamers have been successfully used for targeting whole cells\textsuperscript{50} indicating their potential for detecting specific molecules or biomarkers even on cell surfaces.\textsuperscript{51} One of the most sensitive methods of studying molecular interaction is fluorescence based detection. When fluorescently labeled, aptamers are very useful in cell-staining experiments for direct visualization of target cells in different detection formats including flow cytometry.\textsuperscript{51,52} In a biologically relevant application, an aptamer-based nanosensor was used for real-time fluorescence measurements of ATP levels in vivo, during glycolysis in yeast.\textsuperscript{53}

In the form of molecular beacons (MB), the aptamer is labeled at both ends with a fluorophore at one end and quencher at the other. In the closed state, the fluorescence is quenched due to efficient energy transfer between the fluorophore and the quencher. However, upon binding to target, results in conformational changes in the aptamer which increases the distance between fluorophore and quencher with a corresponding increase
in the fluorescence intensity. MBs are highly specific in detecting even single-nucleotide polymorphism\textsuperscript{54} or non-nucleic acid targets such as small molecule ligands\textsuperscript{55} or protein.\textsuperscript{56}

1.4 Bioanalytical Applications of Aptamers

Aptamers play important roles in bioanalytical applications. Aptamers can be immobilized onto surfaces in affinity chromatography,\textsuperscript{57,58} microfluidics or as selective capture ligands.\textsuperscript{59} A gold electrode coated with aptamers as electrochemical sensor has been successfully applied to detect cocaine in human saliva.\textsuperscript{60} Most electrochemical detection methods use aptamers immobilized onto electrode surfaces. However, homogenous methods have also been designed.\textsuperscript{61,62} An electrochemical sandwich assay system formed by two split aptamers fragments one of which was attached to a gold electrode and the other labeled with methylene blue was used to detect small-molecule targets in complex matrices and blood.\textsuperscript{63} The fragments associate into a folded structure in binding to the target which results to change in current at the electrode surface. These electronic aptamer-based sensors provide label-free alternatives to detection of small-molecule targets even in biological fluids or adulterated samples.\textsuperscript{60}

Capillary electrophoresis is also a very attractive probe for separation and detection of aptamers. Here, separations of the free aptamers from the aptamer-target complex are routine due to changes in electrophoretic mobility or change in size upon binding to the target. Most CE assays are based on laser-induced fluorescence (LIF)
detection which enhances the sensitivity and selectivity. The fact that aptamers can be easily labeled with a fluorescent molecule makes this LIF method of detection common.

Different modes of capillary electrophoresis have been useful in the study of aptamer-ligand interaction such as affinity capillary electrophoresis. However, non-equilibrium capillary electrophoresis of equilibrium mixture methods have made it possible to characterize many aptamer-ligand interactions or screen aptamers (SELEX) with fewer rounds\textsuperscript{64} and also to screen aptamers against a target protein in a cell lysate.\textsuperscript{65} Quantitative aptamer-ligand interaction assays using CE have been performed.\textsuperscript{66,67,68} Aptamer-based micellar electrokinetic chromatography (MEKC) has demonstrated the capability of differentiating slight structural changes such as enantiomers.\textsuperscript{57} A study using CE to probe the interaction between lysozyme and aptamer mediated by cations found that the binding between the aptamer and target is dependent on aptamer conformation.\textsuperscript{69}

Recently, immobilized aptamers in simple square capillaries have been demonstrated as platforms for selective cancer cells capture\textsuperscript{59} and highly efficient enrichment of rare cells.\textsuperscript{70} For effective bioanalytical application, the high affinity biorecognition and tunable properties of aptamers are explored along with their amenability to different detection methods.
1.5 The Role G-Quadruplex Structures in Therapeutics

Non-canonical nucleic acid structures such as G-quadruplexes have recently attracted significant attention due to their potential implications in the regulation of biological processes. G-quadruplexes are formed with a minimum of four guanine (G)-rich repeats in DNA and RNA sequences. They consist of a stack of planar guanine tetrants called G-quartets that are stabilized by Hoogsteen hydrogen bonds and monovalent cation coordination. G-quadruplex-forming sequences are prevalent in the human genome and particularly enriched at biologically important regions such as telomeres and in gene-promoter regions. The formation of stable G-quadruplex structures in the telomeric 3’ overhang has shown inhibitory roles on telomerase, an enzyme up-regulated in the majority of cancer cells. For normal cells, the telomeres shorten during chromosome replication due to reduced telomerase activity leading to eventual cell death. Studies on G-quadruplexes have revealed their structural integrity in physiological and molecularly crowding conditions. These stable secondary structures may cap and maintain the telomeres by preventing end-to-end fusion, thus maintain the stability of the telomere and telomerase regulation.

G-quadruplex structures have recently been directly visualized in cells using monoclonal antibodies thus confirming their formation in vivo and there is intense interests in the design of small-molecule ligands that can selectively bind and stabilize DNA G-quadruplexes (GQs) as a potential strategy for cancer therapy.
The RNA counterparts of these telomeric G-quadruplex DNA, Telomeric repeat containing RNA (TERRA) also form G-quadruplex structures which are known to be more stable than those formed by DNA (hTelo).\(^{85,86}\) The enhanced stability by TERRA GQ’s may have a regulatory role in biological processes such as polymerase function, transcription and translation. Moreover, protein expression from the 5’ untranslated regions (5’ UTR) of mRNAs have been shown to decrease in the presence of GQ forming ligands and increase upon mutating the GQ forming sequence.\(^{87}\)

1.6 Measurement of Chemical Binding Affinity (\(k_d\))

Binding between two components is an important process. For cells, a ligand and receptor binding is the first step to trigger signal transduction that involves a multitude of biochemical reactions, most of which are carried out after an enzyme is bound with a substrate. In the human biology, the binding between antibody and antigen initiates immune responses. In biosensing, binding between an analyte and a detector constitutes the recognition stage that generates signals to be amplified in subsequent steps. In all these examples the dissociation constant, \(K_d\), measures chemical binding affinity. It represents the concentration at which half of the binding sites are occupied by ligands. Binding with higher affinities implies that fewer ligands (lower concentration) are required to occupy the binding sites, which yields a smaller \(K_d\) value. The measurement of \(K_d\) is straightforward as only a series of ligand concentrations is all that is usually required.
The most widely used method for $K_d$ determination is based on the Langmuir isotherm approach.\textsuperscript{88} In this method, the fraction of ligand-bound receptors increases with increasing ligand concentration until a plateau is reached, indicating binding sites are saturated. The Langmuir model,\textsuperscript{88} is based on the assumption that all binding sites are equivalent and that binding ability of molecules is not affected by the occupancy of nearby sites.

\[
\Gamma = \Gamma_{\text{max}} \left[ \frac{K_a C}{1 + K_a C} \right]
\]

where $K_d$ is $1/K_a$, $C$ is the concentration of adsorbate (or ligand), $\Gamma$ is the amount of adsorbed ligand, and $\Gamma_{\text{max}}$ is the maximum amount of adsorbate on a surface.

Different techniques have been applied to measure $K_d$ of the interaction between a ligand and a receptor.\textsuperscript{89} Label-free approaches such as, UV-visible absorption spectroscopy is the simplest and commonly used technique.\textsuperscript{90} Isothermal titration calorimetry (ITC) is used not only to measure the $K_d$ of the binding, but also can reveal the thermodynamic parameters associated with the binding.\textsuperscript{91-92} Recently, surface plasmon resonance (SPR) has become a very popular optical technique to probe binding on a metal surface.\textsuperscript{93-95} This also provides quantitative information on the binding affinity as well as the binding kinetics. All these experiments require increasing concentrations of either a ligand or a receptor. The other challenge in these methods is their low sensitivity. Fluorescence based approaches have become useful techniques to monitor binding interactions due to their high sensitivity. Of the different fluorescence methods, fluorescence polarization anisotropy,\textsuperscript{96-97} fluorescence resonance energy
transfer (FRET),\textsuperscript{98} and laser induced fluorescence capillary electrophoresis (LIF-CE),\textsuperscript{68,99} have been extensively used. However, labeling with a fluorophore may interfere with a binding. In addition, fluorescence signal from environment presents noise to the measurement of binding. To address this issue, recently, $K_d$ has been determined by mechanical force that is required to dissemble a ligand-receptor complex.\textsuperscript{68,100-102}

Binding is however, a dynamic process in which association and dissociation proceed simultaneously. Although $K_d$ provides a measurement of binding strength, it does not reveal the binding dynamics which is better described by the rates of association ($k_{on}$) and dissociation ($k_{off}$). To obtain dynamic information of binding events, it is desirable to adopt a universal variable, similar to the $K_d$, that can compare different binding processes. Mechanical affinity has been proposed as a parameter which better represents both the dynamic and energetic aspects of such binding events.\textsuperscript{103}

1.7 Mechanochemistry of Binding

Mechanochemistry is an emerging field to study the coupling between chemical energy and mechanical energy. Mechanical force is a universal parameter that can be used to characterize a wide range of systems ranging from chemical bonds, intermolecular interactions, macromolecular structures, and many biological and biochemical processes. Change in the conformation of a macromolecule is often accompanied by a variation of the tension sustained by that molecule. Many vital biological processes can be monitored and characterized in terms of mechanical forces
including biochemical actions DNA\textsuperscript{104} or RNA polymerase,\textsuperscript{105-106} exonuclease,\textsuperscript{107} myosin,\textsuperscript{108} and kinesin,\textsuperscript{32} generate forces up to tens of picoNewtons. In ligand-receptor binding, the mechanical force that is required to destroy the interaction directly correlates the strength of the interaction between two binding partners.\textsuperscript{68, 100-102, 109} In such a case, force is applied at a constant rate to overcome the energy barrier. The dynamics of the binding process can be retrieved by measuring the unbinding rate at a constant force or by measuring rupture force as a function of force loading rate.

Despite its importance, there exists a lack of ensemble-average methods to interrogate forces in solution. Hydrodynamic forces from a flowing fluid have been used to monitor the mechanical aspect of biologically relevant processes such as adhesion and rolling of white blood cells inside blood capillaries during inflammation.\textsuperscript{110-111} However, such methods have difficulties to explore the roles of individual binding units at the molecular level. As a result, very little information has been obtained to study the coupling between chemical and mechanical energy.\textsuperscript{112}

Single molecule techniques, however, start to address this problem. Compared to bulk assays where ensemble information is obtained, single-molecule experiments produce stochastic and discrete signals that can be statistically analyzed to reveal the property of sub-groups in a population and the energetics of a particular reaction trajectory in a transition process.\textsuperscript{27-28} Among the many tools to manipulate single molecules, force-based techniques, such as atomic force microscope (AFM),\textsuperscript{29-30} optical tweezers,\textsuperscript{31-33} micro-needles,\textsuperscript{34} and magnetic tweezers,\textsuperscript{35-36} provide a unique capability to
investigate the force generated by or applied at the single molecules level. Details of each method have been recently reviewed.\textsuperscript{113-115}

1.8 Effects of Force and Transition Distance on Mechanical Stability

To unfold a structure, both the rupture force and the distance between the folded and the transition states ($x^\ddagger$) are important. The force ($F$) reduces the free energy of a system by a factor of $F \times x$, where $x$ is the unfolding distance of a biomolecule and brings down the energy barrier of a two-state folding/unfolding system by $F \times x^\ddagger$. With application of a specific force ($F_{\text{rupture}}$), the energy barrier is reduced to a level within reach of thermal energy of environment ($k_B T$, where $k_B$ is the Boltzmann constant and $T$ is absolute temperature). This leads to unfolding of the structure. Assuming $x^\ddagger$ does not change during the application of force, the magnitude of $F_{\text{rupture}}$ is inversely dependent on the transition distance with a larger distance corresponding to a smaller $F_{\text{rupture}}$ value.\textsuperscript{116} The relationship between $F_{\text{rupture}}$ and $x^\ddagger$ can be understood by the elasticity of a molecule. A soft and flexible molecule requires a long distance to be unfolded (larger $x^\ddagger$). The force ($F_{\text{rupture}}$) along this long distance is expected to be small. On the other hand, a rigid but fragile molecule only needs a little stretch in its structure (smaller $x^\ddagger$) to induce unfolding. However, the force leading to this little perturbation ($F_{\text{rupture}}$) is expected to be significantly higher than the previous case (Figure 2).
Figure 2: Energy landscape of unfolding a rigid or flexible structure of a biomolecule, showing the effect of force and transition distance. Distance to the transition state is smaller for a rigid molecule compared to a flexible one. To lower down the energy barrier ($F \times x^\dagger$) by the same amount, a higher force is required for the rigid structure.

Therefore $F_{\text{rupture}}$ alone does not fully account for the mechanical stability studies appropriately defined by $F \times x^\dagger$. Comparison of $x^\dagger$ among DNA secondary species, such as hairpins,\textsuperscript{117} DNA G-quadruplexes,\textsuperscript{101, 118-120} DNA i-motifs,\textsuperscript{121} and their intermediates,\textsuperscript{122-123} has suggested that $x^\dagger$ varies with different molecules. Within the same molecule, $x^\dagger$ is anisotropic among different unfolding trajectories.\textsuperscript{120} Also of important note is that, the magnitude of $F_{\text{rupture}}$ is dependent on the force loading rate. This dependence is rather pronounced for processes with slow unfolding rates.\textsuperscript{124-125} In such a case, $F_{\text{rupture}}$ for reversible folding and unfolding processes occurs in a timescale often not usually attainable in an experiment. By performing laser tweezers experiments
with a fixed loading rate, the $F_{\text{rupture}}$ and $\Delta x$ (difference in the distance between folded and unfolded structure along a trajectory) can be directly measured where $F_{\text{rupture}}$ provides information on the mechanical stability and $\Delta x$ is the size of the folded molecule.

This dissertation seeks to investigate the formation of biologically relevant small biomolecules. Chapter two is the experimental section which details the various techniques and methods used. In Chapter three, the application of DNA aptamers as force-based biosensors for small molecules is investigated from both mechanical stability perspective and fluorescence-based assay. Chapter four details studies on single-molecule RNA G-quadruplex structures with regards to their mechanical and thermodynamic stabilities, as well as their formation or transition kinetics. In Chapter 5, the biological roles and activity of sphingosine kinase 2 using fluorescently labeled lipid probes as substrate is investigated.
Chapter 2

2.0 Materials and Methods

All oligonucleotide sequences ordered were HPLC purified by the manufacturer (Integrated DNA Technologies (IDT), Coralville, IA) and further purified using 10% polyacrylamide gel electrophoresis (PAGE) and stored at -20°C. Samples were quantified by UV-Vis absorbance at 260 nm or by comparing band intensities on 0.8% agarose gels using Kodak Digital Camera and Software (Eastman Kodak Company, Rochester, NY). The ATP aptamer and a scrambled sequence for control capillary electrophoresis experiments were labeled with 6-carboxyfluorescein (FAM) to form 5′-FAM/TAC CTG GGG GAG TAT TGC GGA GGA AGG TT-3′ and 5′-FAM/TAG TGC GCG TGT GTG AGC AGA GTG AGA GT-3′, respectively (IDT, Coralville, IA). For the single-molecule experiments the unlabeled ATP aptamer sequence 5′-TAC CTG GGG GAG TAT TGC GGA GGA AGG TT-3′ was used.

Adenosine-5′-triphosphate disodium salt (ATP) was purchased from USB Corp, Cleveland, OH. N-[Tris-(Hydroxymethyl) Methyl] Glycine (Tricine) was purchased from Fisher Scientific. Trizma® base (minimum 99.9% titration) was purchased from Sigma. Polyvinylpyrrolidone (PVP) was obtained from Acros Organics. Nanopure water was purified by Millipore Synergy® UV with 185 nm UV lamp.

2.1 Molecular Biology

2.1.1 Preparation of DNA Handles
To make DNA construct for single molecule experiments, two double-stranded DNA (dsDNA) handles (2690 and 2028 bp) were first prepared. The 2690 bp handle was obtained from SacI followed by EagI (NEB) digestions of pEGFP plasmid (Clontech, Mountain View, CA). The 2690 fragment was gel purified and labeled with digoxigenin (Dig) at the 3’ end (SacI end) using 18 μM DigdUTP (Roche, Indianapolis, IN) and terminal transferase (Fermentas, Glen Burnie, MD). To ensure complete Dig labeling, excess Dig-dUTP was used (Dig-dUTP:2690 bp DNA = 150:1) and purified by ethanol precipitation. The Dig-labeled 2690 fragment was purified by ethanol precipitation (Figure 3).

To make the 2028 bp handle, a pBR322 plasmid (NEB) was first amplified by PCR using a 5' labeled biotinylated primer, 5'-GCA TTA GGA AGC AGC CCA GTA GTA GG-3' (IDT). The PCR product was purified by a kit (Qiagen, Germantown, MD), digested overnight with XbaI restriction enzyme (NEB) and purified by centrifugation using centrigon(R) (Figure 3).

2.1.2 Synthesis of DNA Constructs for Single-Molecule Experiments

Preparation of the DNA construct for single-molecule mechanical unfolding studies was carried out following a previous protocol.118 This is achieved by sandwiching the ATP DNA aptamer sequence between two dsDNA handles (2690 bp and 2028 bp). To make the DNA construct for the single-molecule experiments, an 83 mer DNA fragment which contains the ATP aptamer sequence, 5' ACC TGG GGG AGT ATT GCG GAG GAA GGT in the middle. This sequence is flanked by two dsDNA
fragments that end with 5' GGC CGA CGC GCT GGG CTA CGT CTT GCT GGC and 5' CGC ATC TGT GCG GTA TTT CAC ACC GT sticky ends, to which the 2028 bp and 2690 bp dsDNA handles were ligated using T4 DNA ligase (NEB) (Figure 3) and purified using a gel extraction kit (Qiagen). The purified final construct was dissolved in 1 mM Tris, and stored at -20 °C. The control ATP aptamer construct was made by replacing the DNA aptamer sequence shown above with a scrambled sequence, 5'-AGT GCG CGT GTG TGA GCA GAG TGA GAG.

**Figure 3:** Flow chart for making DNA constructs for single-molecule experiments. DNA sequence of interest is sandwiched between two dsDNA handles through T4 DNA ligation.
2.1.3 RNA Constructs Preparation for Single-Molecule Experiments

To prepare the DNA-RNA hybrid RNA (TERRA) construct, a DNA sequence complementary to the TERRA quadruplex-forming sequence (5'- TTAGGG TTAGGG TTAGGG TTAGGG TTA -3') was cloned into a pBR322 plasmid (NEB, Ipswich, MA) between EcoRI and HindIII sites and transformed into GM2163 competent cells and grown for 8 hours to make more copies. The transformed plasmid was harvested, purified using Qiagen mini prep kit and sequenced (DNA sequencing, in university of Maine). The template for in vitro RNA synthesis (MEGAscript, Invitrogen) was generated from PCR amplification of the cloned plasmid using a forward primer containing the T7 promoter. In vitro transcription was then performed using a MEGAscript T7 Kit (Ambion) according to the manufacturer’s protocol. A 6 μL of T7 template DNA (175 ng/μL) was incubated with 2 μL of each rNTPs (rATP, rCTP, rGTP, rUTP), 2 μL of enzyme mix, 1X reaction buffer and RNase free water to make a total volume of 20 μL and reacted for 5 hrs at 37 °C. This was followed by 1 μL of turboDNase (2 units/μL) for 15 min to destroy any unreacted DNA template.

Handle A (527 bp) was obtained from PCR amplification of the unmodified pBR322 plasmid (nucleotides 3836-1) with an EcoRI digestion site on the forward primer, followed by EcoRI (NEB, Ipswich, MA) digestion and end-labeling with DNA Polymerase I, Large (Klenow) Fragment (NEB) in the presence of digoxigenin (Dig)-dUTP. Handle B (1473 bp) was also obtained from PCR amplification of the unmodified pBR322 plasmid (nucleotides 31-1473) with a 5’-biotin labeled reverse primer. The
control construct was prepared similarly by replacing the TERRA sequence with the wild type sequence between the EcoRI and HindIII sites of pBR322.

The DNA-RNA hybrid constructs were prepared by annealing the RNA and two DNA handles via heating at 95 °C for 5 min and then gradually cooling down to room temperature for 4 hours (Figure 4).

*Figure 4*: Flow chart for making RNA construct and DNA/RNA hybrid handles.
2.2 Laser Tweezers Instrumental Setup

Figure 5 shows a schematic of the laser tweezers instrument.\textsuperscript{126-127} It is made up of a diode pumped solid state (DPSS) laser (1064 nm, 4 W, CW mode, BL-106C, Spectra-Physics) used as a trapping laser. \( P \) and \( S \) polarized laser light from the same laser source constitute two traps. The \( S \) polarized light is controlled by a steerable mirror (Nano-MTA, Mad City Laboratories) at a conjugate plane of the back focal plane of a focusing objective (Nikon CFI-Plan-Apochromat 60\( \times \), NA 1.2, water immersion, working distance \( \sim320 \) \( \mu \)m). The exiting \( P \) and \( S \) polarized beams are collected by an identical objective and detected separately by two position sensitive photodetectors (PSPD, DL100, Pacific Silicon Sensor).\textsuperscript{128} The force of the laser trap was calibrated by Stokes force and thermal motion measurement. Both methods yielded a similar trap stiffness of \( \sim307 \) pN/ (\( \mu \)m \( \times \) 100mW) (for 0.97 \( \mu \)m diameter polystyrene beads, Bangs Laboratory, Fishers, IN).
2.3 Single-Molecule Force-Ramp Experiments

Detailed description of the laser-tweezers instrument used for the single-molecule experiment has been reported. Anti-Digoxegenin (Anti-Dig) antibody coated polystyrene beads (diameter: 2.17 μm, Spherotech, Lake Forest, IL) were incubated with diluted DNA/RNA constructs obtained above (~0.43ng/μL) for 30 min at room temperature to attach the construct on the bead surfaces by Dig/anti-Dig linkage. The sample coated beads were then dispersed in the run buffer containing 100 mM KCl, 10 mM buffer (Tris, or Phosphate) pH 7.4 at 23 °C together with streptavidin coated beads (diameter: 1.87 μm, Bangs Laboratory) and injected into the reaction chamber. The
streptavidin-coated beads and the anti-Dig-coated beads linked with the DNA construct were injected separately into a reaction chamber through different channels as shown in Figure 6.

**Figure 6.** Schematic of a three-channel microfluidic reaction chamber for performing single-molecule laser tweezers experiments.

The two different beads were separately captured onto two laser spots. As one end of the DNA construct was modified with biotin and the other with Dig, the DNA construct could be immobilized between the two beads via streptavidin/ biotin and Dig/anti-Dig affinity interactions, respectively. The laser trapped anti-Dig coated bead bound with DNA constructs was approached by the streptavidin-coated polystyrene bead
until a tether between the two beads was observed (Figure 7A). The anti-Dig coated bead, controlled by the Nano-MTA positioner, was then moved away from the streptavidin-coated bead with a loading speed of ~5.5 pN/s. During this process, the tension inside the molecule increased until any secondary structure was mechanically unfolded (Figure 7 A, top). Rupture events indicated by sudden change in the end-to-end distance were observed in the force-extension ($F$-$X$) curves as a result of unfolding the secondary structure. The force-extension ($F$-$X$) curves were recorded at 1000 Hz using a Labview® program and the detailed data treatment was performed using programs in Matlab and Igor software. Single tethers were confirmed by the observation of a plateau at 65 pN in F-X curves\textsuperscript{31} or a single breakage event (force drops to 0 pN) when the tether was overstretched after each experiment. Only force-extension ($F$-$X$) curves from single tethers were considered for data analysis. The force-extension curve for each tether was recorded in a Labview program. Two data points flanking the rupture event in the $F$-$X$ curves (Figure 7 B) were used to calculate the change in contour length ($\Delta L$) during the rupture of a folded structure (Figure 7 C) according to the following equation,\textsuperscript{130}

\[
\frac{x}{L} = 1 - \frac{1}{2} \left( \frac{k_B T}{FP} \right)^2 + \frac{F}{S}
\]

where $x$ is the end-to-end distance, $L$ is the contour length, $T$ is absolute temperature (298 K), $F$ is the force, $k_B$ is the Boltzmann constant, $P = 51.95$ nm,\textsuperscript{127} is the persistent length, and $S = 1226$ pN,\textsuperscript{127} is the elastic strength modulus.
Figure 7: Single-molecule force-ramp experiments. A) Schematic of force ramping where a single-molecule tethered between two beads is repeatedly stretched and relaxed. B) Typical force extension curve. C) Force versus $\Delta L$ plot showing the contour length ($\Delta L$) measurement.

2.4 Calculation of Formation Probability of a Species

The formation probability of a folded species is calculated by taking the ratio of $F$-$X$ curves that contain rupture events versus the total number of tethered $F$-$X$ curves.
2.5 Force-Jump Experiments

To measure the refolding kinetics, single molecules of the construct were first mechanically unfolded. They were then quickly relaxed to 0 pN and incubated for a specific time (0, 8, 15, 30, 45, 60, 120, 180, and 300 s) prior to subsequent unfolding. Observed features are indicative of the refolding of the tethered structure during the incubation. The refolding probability was determined by the number of folded events divided by the total events.

2.6 Contour Length Calculation

Contour length (L) of the secondary structure was obtained by fitting the force-extension (F-X) curves to the Worm-Like-Chain (WLC) model: \[ \frac{X}{L} = 1 - \frac{1}{2} \left( \frac{k_B T}{F P} \right)^2 + \frac{F}{S} \]
, at a particular force (F). Each F-X curve was split into the stretching (black) and relaxing (red) traces (Figure 7 B), the change in extension (ΔX) was calculated by subtracting the two traces at the particular unfolding force. The resulting ΔX was then converted to the change in contour length (ΔL) using the (WLC) model above at a particular force (F).

2.7 Determination of Unfolding Free Energy

The Change in free unfolding energy (ΔG_{unfold}) for a secondary structure was calculated using the Jarzynski’s equality equation \[ \Delta G = -k_B T \ln \sum_{i=1}^{N} \frac{1}{N} \exp\left(-\frac{W_i}{k_B T}\right) \],
where $N$ is the number of data points or observed events and $W$ is the nonequilibrium work done to unfold the structure, which equals the hysteresis area between stretching and relaxing force extension curves.\textsuperscript{134} In the case of ligand binding, the $\Delta G_{\text{binding}}$ can be obtained from the expression:

$$\Delta G_{\text{binding}} = \Delta G_{\text{free-unfold}} - \Delta G_{\text{bound-unfold}}$$

and the bias of the $\Delta G_{\text{unfold}}$ calculated from the unfolding work histograms as described in the literature.\textsuperscript{135}

2.8 Population Deconvolution with Nanometer Resolution (PoDNano)

This is a statistical analysis method where the probability density for each transition between a folded and the unfolded species, $p$, is estimated according to the following Gaussian kernel expression,\textsuperscript{136}

$$p = \frac{1}{\sigma 2\pi} \exp\left(-\frac{(x - \Delta L)^2}{2\sigma^2}\right)$$

where $\Delta L$ and $\sigma$ are the change in contour length for the transition and its associated standard error, respectively, and $x$ is defined by the range $\Delta L \pm 3\sigma$. The $\sigma$ is the average standard error for the two regions before and after the transition event (the $\Delta L$-$F$ plots in Figure 7 C). The probability density for each step was then summed up to construct the kernel density estimation. Bootstrapping analyses with 3000 times random resampling was performed to identify individual populations in the constructed $\Delta L$ histogram. The highest peaks were identified using Igor\textsuperscript{®} program (WaveMatrics, Portland, OR) to
deconvolute the $\Delta L$ histogram into different $\Delta L$ populations and these selected peaks were used to construct the bootstrap histograms.

2.9 Capillary Electrophoresis

Capillary electrophoresis experiments were performed on a Beckman ProteomeLab™ PA800 CE system (Fullerton, CA), (diagramed in Figure 8), equipped with a LIF detector using an air-cooled 3.5 mW argon laser (Beckman Instruments) with 488 nm excitation and 520 nm emission. Separations were performed in 30 cm long fused-silica capillaries (Polymicro Technologies, Phoenix, AZ, USA) with 50 μm i.d. and 375 μm o.d. (20 cm effective length). New capillary columns were preconditioned by successively flushing with 1.0 M NaOH for 30 min and rinsed with deionized, filtered H$_2$O for 20 min at 20 psi. Five prime end fluorescein (FAM) tagged ATP aptamer (IDT) was used for the binding studies. Aliquots of ATP (0-60 μM) were added to the ATP aptamer (0.5 nM) in a 10 mM phosphate buffer (pH 7.4) with 100 mM KCl. The mixture was heated to 97 °C for 5 min followed by slow cooling to room temperature for 4 hours in a Bio Rad DNAEngine® Peltier Thermal Cycler and injected hydrodynamically into the capillary at 0.5 psi for 10 s. During the separation, an electric field of 500 V/cm was applied in a reverse polarity mode (anode at the detector) and the capillary column was maintained at 25 °C. The CE separation buffer was made of 25 mM Tris, 25 mM Tricine, and 0.5% polyvinylpyrrolidone (PVP) at pH 7.4.
2.10 CD and UV Melting Experiments

Circular Dichroism

A Chirascan spectropolarimeter (Applied Photophysics, Leatherhead, UK) was used to record circular dichroism experiments. Nucleic acid concentration of 10 μM in 10 mM Tris and 100 mM potassium chloride buffer (pH 7.4) was heated to 95 °C for 5 min followed by slow cooling to room temperature. Scans were taken at 25 °C, over a range of 220 nm to 340 nm run using a 1 mm path length quartz cuvette. Each final spectrum is the average of four scans taken with a step size of 1 nm per second, and a
bandwidth of 0.5 nm. For blank subtraction, the buffer only was run and data were zero corrected at 340 nm each spectrum.

**UV Thermal Experiments**

UV melting and annealing experiment were collected using a Cary 100 UV-Visible spectrophotometer (Agilent) by measuring absorbance at 295 nm as a function of the temperature. Oligonucleotide solutions were prepared to a final concentration of 5 μM in 10 mM lithium cacodylate buffer (pH 7.4) containing potassium chloride at specified concentration. Where necessary, lithium chloride was added to maintain an ionic strength equivalent to 100 mM KCl. The melting experiments were performed in 10 mm path length quartz cuvettes with 1 mL of buffered oligonucleotide solution. A steady stream of nitrogen was applied to prevent condensation at low temperatures and the solutions were covered with a layer of mineral oil to minimize evaporation. Temperature ramps consisted of low → high → low → high cycles with a temperature range of 10-90 °C. A ramp rate of 0.25 °C/min was maintained for all experiments with data collection every °C. Melting temperatures were determined as where fraction folded is equal to 0.5.

Temperature-independent van’t Hoff free energy changes for G-quadruplex formation (ΔG_melt) were determined using the equation lnK(T) = -ΔH/RT + ΔS/R, where K is the equilibrium constant for G-quadruplex formation, with the assumption of a two-state model between the single-stranded to the folded state. Plotting lnK(T) versus 1/T allowed us to access the values of ΔH and ΔS from the slope and intercept,
respectively, of the linear regression in the interval $0.15 < \theta < 0.85$. Free energies were calculated at 296 K using the Gibbs equation $\Delta G_{\text{melt}} = \Delta H - T \Delta S$. 
Chapter 3

Single-Molecule Measurements of the Binding between Small-Molecules and DNA Aptamers

This chapter has been published as an article in Analytical Chemistry (dx.doi.org/10.1021/ac300427d) and used with permission. Publisher: American Chemical Society. Copyright © 2012, American Chemical Society.

3.1 Summary

Aptamers that bind small molecules can serve as basic biosensing platforms. Evaluation of the binding constant between an aptamer and a small molecule helps to determine the effectiveness of the aptamer-based sensors. Binding constants are often measured by a series of experiments with varying ligand or aptamer concentrations. Such experiments are time consuming, material non-prudent, and prone to low reproducibility. Here, we use laser-tweezers to determine the dissociation constant for aptamer-ligand interactions at the single-molecule level from only one ligand concentration. Using an ATP binding aptamer as an example, we have observed that the mechanical stabilities of aptamers bound with ATP are higher than those without a ligand. Comparison of the change in free energy of unfolding ($\Delta G_{\text{unfold}}$) between these two aptamers yields a $\Delta G$ of $33 \pm 4$ kJ/mole for the binding. By applying a Hess-like cycle at room temperature, we obtained a dissociation constant ($K_d$) of $2.0 \pm 0.2$ µM, a value consistent with the $K_d$ obtained from our equilibrated capillary electrophoresis ($2.4 \pm 0.4$ µM) and close to that
determined by affinity chromatography in the literature (6 ± 3 µM). We anticipate that our laser-tweezers and CE methodologies may be used to more conveniently evaluate the binding between receptors and ligands and also serve as analytical tools for force-based biosensing.
3.2 Introduction

The interaction between nucleic acids and small molecules is continuously generating significant interest due to their widespread use in analytical and bioanalytical applications. Specific nucleic acids such as aptamers have been shown to bind not only to nucleic acids, but also to non-nucleic-acid targets such as small molecules, proteins, and cells. Aptamers are single-stranded ligand-binding nucleic acids with affinity and selectivity comparable to antibodies. Compared to antibodies, DNA aptamers are easier to obtain, more stable, and more resistant to biodegradation. As such, they are suitable for demanding conditions such as extreme pH and temperatures. These properties render them ideal candidates for sensors used in medical diagnostics and environmental monitoring.

The binding affinity between an aptamer and a ligand is a critical factor to reflect the sensitivity of aptamer-based biosensors. Conventionally, the binding is measured by techniques such as electrochemical detection, fluorescence, colorimetric or surface plasma resonance (SPR) based detection, capillary electrophoresis (CE) separation among others. At the single molecular level, AFM has been recently used to determine the ATP binding to an aptamer made of two independently split DNA strands. Single-molecule FRET has also been explored to evaluate the binding of a DNA aptamer with its binding target, vascular endothelial growth factor (VEGF).

However, majority of these methods require a series of experiments with varying ligand or aptamer concentrations to determine the dissociation constant. Apart from
being tedious and time consuming, such experiments are prone to decreased signal-to-
oise ratio due to run-to-run variations. Another disadvantage is that they require
substantial amount of materials. This becomes problematic especially in drug screening
processes where the amount of a sample is often limited. As an alternative, here, we
employ laser-tweezers instrument to determine the dissociation constant between ATP
and a DNA aptamer by using only one concentration of ATP.

Developed in the nineteen eighties, laser tweezers have been used extensively
to probe the mechanical stabilities not only for DNA and RNA, but also for
proteins at the single-molecule level. Recently, we have extended this method to
investigate the interaction between a DNA G-quadruplex and its ligand, as well as to
serve as a force-based biosensing platform. Here, we wish to use this technique to
scrutinize the interaction between aptamers and ligands. The binding of ATP to the
aptamer is expected to increase the mechanical stability of the aptamer. This increment is
converted to the change in the free energy of binding (ΔGbinding) by a Hess-like process.
The ΔGbinding is then used to obtain the dissociation constant (Kd) of the complex, which
is validated by our CE measurement. Such a force-based strategy is less prone to
background noise often seen in fluorescence-based methods. Our results demonstrate a
strong correlation between ensemble average assays (CE) and single-molecule
approaches (laser-tweezers) on the binding between aptamers and small molecules.
3.3 Results

3.3.1 The Mechanical Stability of an Aptamer is Enhanced by Small-Molecule Binding

We first investigated the mechanical stability of the ATP aptamer itself. After sandwiching the aptamer fragment between two double-stranded DNA (dsDNA) handles, the DNA construct was tethered between the two laser-trapped beads via biotin/streptavidin and digoxigenin (Dig)/anti-Dig antibody linkages (Fig. 9A). The tension in the DNA tether was gradually increased by moving one of the trapped beads away from the other until a sudden drop in force was observed in the force-extension ($F$-$X$) curves (Fig. 9 C insert). This drop in force is associated with the unfolding of a secondary structure, most likely the aptamer, in the DNA construct. The change in contour length ($\Delta L$, see Materials and Methods for calculation) was determined to be 9.2 ± 0.1 nm, a value consistent with the contour length of the 27-mer ATP aptamer.$^{118,121}$
Figure 9. (A) Schematic of the experimental setup. (B) ATP structure. (C) A typical force-extension curve. Inset shows a blowup of the unfolding event denoted by the dotted circle.

In a control with a scrambled sequence that does not assume any stable conformations, no folded structure was observed, which strongly suggests that the observed events are due to the ruptures of folded ATP aptamers. The unfolding force at which the rupture event occurred showed a most probable value of 21 pN without the
ATP ligand (Fig. 10 A). The probability of observing these rupture events from all tethered DNA molecules is 24 %, suggesting a significant fraction of the aptamers is unfolded.

To investigate the binding of a small-molecule, the aptamer construct was incubated with 2.0 µM ATP. The rupture force histogram for this treatment revealed two populations with rupture forces centered at 21 and 38 pN (Fig. 10 B), respectively. Based on the similar rupture force, we assigned the 21 pN population as the free aptamer. After the aptamer binds to ATP, it is expected that the mechanical stability of the aptamer should increase. Therefore, the 38 pN species was assigned as the ATP-bound aptamer (Fig. 10 B). To confirm this assignment, we increased the ATP concentration to 100 µM with the expectation that the population of the ATP-bound aptamer should increase. Indeed, this was observed in the unfolding force histogram in Figure 10 C. The increase in the mechanical stability of the aptamer upon ligand-binding can be exploited to develop force-based biosensing recently developed in our lab.
Figure 10. Unfolding force histograms of the ATP aptamer in a 10 mM phosphate buffer (pH 7.4) with ATP concentration ranging from (A) 0 ATP; (B) 2.0 µM ATP; and (C) 100.0 µM ATP. Curves depict Gaussian fittings.
3.3.2 Dissociation constant determination at only one concentration of a ligand

The dissociation constant \( (K_d) \) between the ATP and the aptamer can be calculated by the expression, \( \ln(K_d) = -\Delta G_{\text{binding}}/RT \), where \( \Delta G_{\text{binding}} \) is the change in free energy of binding, a state function that is independent of pathways. To obtain the \( \Delta G_{\text{binding}} \), we designed a Hess-like cycle\(^{156} \) in which the binding of the ligand to the aptamer is rerouted to the unfolding of a free aptamer in the presence of an unbound ligand. This is followed by the formation of the ATP-bound complex from the mechanically unfolded aptamer (Figures 11 & 12). The \( \Delta G_{\text{binding}} \) is then estimated from the difference in the change in free energy of unfolding \( (\Delta G_{\text{unfold}}) \) for the latter two processes. Each \( \Delta G_{\text{unfold}} \) is obtained using the Jarzynski’s equality equation\(^ {132} \) below,

\[
\Delta G_{\text{unfold}} = -k_B T \ln \sum_{i=1}^{N} \frac{1}{N} \exp \left[ -\frac{W_i}{k_B T} \right]
\]

where \( k_B \) is the Boltzmann constant, \( T \) is absolute temperature, \( N \) is the number of observations, and \( W \) is the non-equilibrium work done to unfold the structure(s), which is equivalent to the hysteresis area between stretching and relaxing \( F-X \) curves (See Figure 9 C as an example). Histograms for the work revealed that the average work done to unfold the free aptamer (Figure 11 A) is less than that for the ATP-bound aptamer (Figure 11 B), which confirmed that the mechanical stability of the aptamer is enhanced after ligand binding. The Jarzynski calculation yielded \( \Delta G_{\text{free-unfold}} \) of 29 ± 3 kJ/mol (bias estimation, 7 kJ/mol) and \( \Delta G_{\text{bound-unfold}} \) of 62 ± 3 kJ/mol (bias estimation, 8 kJ/mol). With these, \( \Delta G_{\text{binding}} \) was calculated by the following expression:
\[ \Delta G_{\text{binding}} = \Delta G_{\text{free-unfold}} - \Delta G_{\text{bound-unfold}} \]

This treatment yielded \( \Delta G_{\text{binding}} \) of -33 ± 4 kJ/mol, which led to a \( K_d \) of 2.0 ± 0.2 \( \mu \text{M} \) by the equation, \( \Delta G_{\text{binding}} = -RT \ln K_d \). The bias of these free energy estimations was calculated according the described method.\(^{135}\) The dissipated work, \( 12 \ k_B T \) and \( 15 \ k_B T \) for unfolding at 0 and 100 \( \mu \text{M} \) ATP, respectively, was well within the range (~20 \( k_B T \)) for effective bias calculations.\(^{135}\)
Figure 11. Histograms of the work done to unfold free aptamer (A) and ATP-bound aptamer (B). The work is calculated at 100 µM ATP. Solid curves depict Gaussian fittings. Dotted arrows show the work equivalent to the change in free energy of unfolding ($\Delta G_{\text{unfold}}$).
Figure 12: The Hess-like cycle of the aptamer-ATP binding process used for the calculation of the $\Delta G_{\text{binding}}$ and the determination of the dissociation constant.

3.3.3 Determination of Dissociation Constant by CE

To validate these single-molecule results, we set out to compare the $K_d$ value with an ensemble average measurement. It is known that binding affinity can be determined by capillary electrophoresis. Therefore, the ATP aptamer was labeled with 6-carboxyfluorescein at the 5’ end and analyzed using CE with laser-induced fluorescence detector (CE-LIF).
Figure 13. (A) Electropherograms for 0.5 nM aptamer in the presence of varying ATP concentrations (0-60 µM) in a 10 mM phosphate buffer (pH 7.4) containing 100 mM KCl at 25 °C. The left peak (4.2 min) represents the unfolded aptamer while the right one (4.6 min) indicates folded aptamer species (free and ATP bound). (B) Fraction of the ATP-bound aptamer with ATP concentration. The solid curve is a Langmuir isotherm fitting to obtain the dissociation constant, $K_d$. 
Depicted in Figure 13 are two populations (~4.6 and ~4.2 min retention times) in the electropherogram of the 0.5 nM ATP aptamer with increasing ATP concentrations (0 – 60 µM). Since a significant fraction of DNA is unfolded by laser-tweezers experiments (see above), we ascribe the larger population (the 4.2 min peak) as unfolded aptamer and the 4.6 min population as the folded aptamer bound with or without ATP. The difference in the elution time reflects the variation in the size and charge between folded and unfolded ATP aptamers, which is likely due to different interactions with potassium cations. This assignment was confirmed as the folded aptamer species (the 4.6-min peak with free and ATP bound aptamers) increase their populations relative to the unfolded ones when ATP concentration increases (Figure 13, 0 – 60 µM ATP traces). As a control, a scrambled sequence that does not form any structures did not show this trend (Figure 14). The fact that neither unfolded nor folded aptamers had tailing or fronting in the electropherogram suggests that the conversion between these two species is slow within the time scale of electrophoresis. In addition, the inseparable ATP-aptamer complex and the free folded aptamer helps to maintain the equilibrium between the free and the ATP-bound aptamers during the electrophoresis. These two facts directly contrast with the non-equilibrated conditions exploited by kinetic CE for measurement of dissociation constants.
Figure 14. Electropherograms of a 0.5 nM scrambled control sequence, 5′-/FAM/ T AGT GCG CGT GTG TGA GCA GAG TGA GAG T-3′, with and without ATP (0 – 60 uM) in 10 mM phosphate buffer (pH 7.4) that contains 100 mM KCl at 25 °C.

These equilibrated conditions were used to determine the dissociation constant for the ATP-aptamer complex. First, we measured the amount of folded aptamer without bound ATP from the 4.6 min peak in the electropherogram at 0 μM ATP. By comparing the folded aptamer with the unfolded aptamer that appeared at the 4.2 min peak, we
obtained a folding equilibrium constant of 0.30 ($K_{\text{fold}} = [\text{folded}]/[\text{unfolded}]$) between the folded and unfolded aptamers. This $K_{\text{fold}}$ value is very close to that estimated from the formation probability of folded species (24%) observed in the single-molecule studies ([probability of folded]/[probability of unfolded] = 24%/76% = 0.32). With this $K_{\text{fold}}$ value, we then determine the fraction of the free, folded aptamer in the 4.6 min peak by comparison with the unfolded aptamer population (the 4.2 min peak) for all other ATP concentrations (Figure 13, 0 – 60 µM ATP traces). Based on these fractions, a Langmuir isotherm$^{156}$ (Figure 13 B) was constructed to calculate the $K_d$ of the ATP-aptamer complex. The resultant $K_d$ (2.4 ± 0.4 µM) matches very well with the single-molecule $K_d$ measurement obtained above. It is also comparable with previous measurements based on affinity chromatography.$^{163}$
3.4 Discussion

3.4.1 $K_d$ Determination by Capillary Electrophoresis

Pioneered by Krylov, kinetic CE has been developed in the last ten years to obtain the dissociation constants from non-equilibrium conditions created during capillary electrophoresis. This method requires different electrophoretic mobilities between the ligand, the receptor, as well as the ligand-receptor complex. In addition, the binding or unbinding should be fast enough to allow the interaction to occur within the time scale of the CE separation. Here, due to the inseparable species of bound and free aptamers, and the slow folding and unfolding of the aptamer, our method allows the determination of the binding affinity in a manner totally different from kinetic CE approaches. This equilibrated CE method is also different from the previous CE approaches that utilize the different mobilities between the free and bound species to obtain binding affinity. Our CE method provides a $K_d$ value close to that from the affinity chromatography (6 ± 3 µM), validating our equilibrated CE approach for the affinity measurement.

3.4.2 $K_d$ Determination by Laser Tweezers

Single-molecule studies can reveal detailed mechanisms of the interaction between ligands and receptors. However, not many single-molecule approaches are capable of retrieving thermodynamic variables, such as $K_d$, which is best obtained at equilibrium established by millions of molecules. Being highly sensitive, single-
molecule techniques are better suited to investigate kinetics processes. Jarzynksi’s approach\textsuperscript{133} to retrieve thermodynamic information from non-equilibrated processes greatly expands the capability of single-molecule techniques. Our methods described above fully utilize this innovative discovery. By repeatedly unfolding and refolding individual structures with a laser-tweezers instrument, the change in free energy of unfolding can be obtained from non-equilibrium conditions. The single-molecule setting reduces the number of unfolding and refolding cycles that are required to retrieve the thermodynamic information as thermal bath from the environment becomes significant at the single molecular level. The validity of such an approach has been well established previously.\textsuperscript{133}

The fact that the binding energy, $\Delta G_{\text{binding}}$, which is used for the dissociation constant calculation, can be obtained in the presence of a ligand with a desired concentration enables our method to determine the binding affinity from only one concentration of the ligand (Figure 11). Previously, methods have been demonstrated to measure affinity from a ligand with a fixed concentration.\textsuperscript{165,166} However, all these approaches require a certain range of workable concentration to ensure accurate determination of a dissociation constant. In our case, this limitation is not necessary. In practice, we can choose a ligand concentration high enough to populate ligand-bound species, from which the $\Delta G_{\text{unfold}}$ is determined (in this research, we obtained $\Delta G_{\text{unfold}}$ from the $100 \mu$M ATP concentration). After correcting for the $\Delta G_{\text{unfold}}$ for an aptamer without a bound ligand, $\Delta G_{\text{binding}}$ is obtained (Figure 12). Therefore, our method can be conveniently used to evaluate the dissociation constant without prior knowledge.
This single-molecule approach provides a general platform to measure binding affinity of a receptor-ligand system beyond the aptamer-ligand complex described here. To investigate protein-ligand interaction, a protein can be tethered between two optically trapped particles. The mechanical stability of the protein can then be evaluated with or without a ligand to determine the dissociation constant using the Hess-like cycle (Figure 12). Recently, several proteins\textsuperscript{167,168} have been investigated for their mechanical stabilities by laser-tweezers using a similar molecular-construct strategy as described here. Compared to the fluorescence-based evaluations that are prone to background noises, the force-based detection has reduced noise.\textsuperscript{157} In addition, the positions from which a receptor is attached to handles can be well controlled to avoid the labeling impact on the binding pocket, which is common for fluorescence-based methods.

Finally, the aptamer containing single-molecule system serves as a natural platform for force-based sensing. Previously, we have described detection of SNP using this strategy.\textsuperscript{157} By exploiting the versatility of aptamers, different small molecules can be detected based on the increased mechanical stability of a ligand-bound aptamer.
3.5 Conclusion

In conclusion, we have used two complementary methods, an ensemble-average based CE and a single-molecule based mechanical unfolding, to investigate the binding of ATP to an ATP aptamer. The aptamer-ligand interaction contributes to the increased mechanical stability of the complex, which leads to the determination of the dissociation constant by a Hess-like cycle. This dissociation constant is consistent with that determined by our equilibrated CE analysis and close to that measured in literature.\textsuperscript{169} We anticipate the methods developed here are not only useful for the determination of the binding affinity of general receptor-ligand systems in a convenient manner, but also can be used to develop force-based biosensors.
Chapter 4

Mechanochemical Properties of Individual Human Telomeric RNA (TERRA) G-quadruplexes

This chapter has been published as an article in ChemBioChem DOI: 10.1002/cbic.201300350 and reprinted with permission. Publisher: John Wiley and Sons. Copyright © 2013 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim.

4.1 Summary

Recently, it was discovered that human telomeres are transcribed into telomeric repeat-containing RNA (or TERRA) that can form G-quadruplex structures, by which telomere functions might be regulated. Herein, we describe studies on single-molecules in laser tweezers, which interrogate the structural, kinetic, and mechanical properties of TERRA G-quadruplexes. We confirmed the formation of unimolecular, parallel TERRA G-quadruplexes. With force jump approaches, we also discovered partially folded structures that serve as intermediates to the TERRA G-quadruplexes. The observation of multiple rupture force populations indicates the presence of more than one set of G-quadruplex with its intermediate. The unfolding force of TERRA G-quadruplexes along the 5′-3′ trajectory is larger than the stall force of the polymerases that process nucleic acid templates in the same direction. This suggests a requirement for helicases to resolve G-quadruplexes during telomere related processes. These mechanochemical properties provide important implications for the biological role(s) of TERRA G-quadruplex structures.
4.2 Introduction

Long non-coding telomeric RNA (TERRA) that contains the repeat sequence, 5′-UUA GGG, was recently discovered in cells.\textsuperscript{170-171} Evidence suggests that TERRA is functionally important and can regulate telomerase activity, heterochromatin structures, and replication fork progress, among other potential roles.\textsuperscript{172} Like human DNA telomeric G-quadruplex (hTelo GQ)\textsuperscript{173-174}, TERRA can also form G-quadruplexes (GQs) both \textit{in vivo} and \textit{in vitro}.\textsuperscript{175} Recent \textit{in vitro} experiments have shown that the telomere-associated protein TRF2 binds to TERRA via interaction with the TERRA GQ structure.\textsuperscript{176} From a mechanochemical perspective, the analysis of structures and their dynamics can provide important insights into the potential biological roles for GQs. While hTelo GQs have been extensively characterized from the mechanical, thermodynamic, and kinetic aspects,\textsuperscript{101,120,122} there is only little such information on the TERRA GQ.\textsuperscript{177}

NMR and X-ray crystallographic studies on synthetic oligonucleotides based on the hTelo sequence have revealed a variety of GQ conformations.\textsuperscript{178-183} In contrast, the parallel ‘propeller’ GQ conformation is predominant in RNA sequences.\textsuperscript{184-185} However, high-resolution RNA GQ structures resolved so far are all intermolecular in nature.\textsuperscript{184,186} Part of the challenge in structurally characterizing intramolecular RNA GQs stems from the propensity for G-rich telomeric RNA strands to self-associate and form higher order species such as dimeric G-quadruplex units.\textsuperscript{186} By using a sequence consisting of only four telomeric RNA G-rich repeats, 5′- UUA (GGG UUA)\textsubscript{4} -3′ (TERRA-4G), the formation of dimeric GQs can be avoided at the single molecule level, enabling one to
directly probe the structure and transition kinetics of the monomeric TERRA GQ along the 5′-3′ trajectory in a laser-tweezers instrument. As polymerases or helicases process telomeres along the same direction, this method therefore provides a potential solution for the first time to investigate the interaction between RNA GQs and motor proteins from mechanochemical perspective.
4.3 Results and Discussion

4.3.1 RNA G-Quadruplexes Form in the TERRA Sequence

To mechanically unfold or refold possible RNA GQs, we incorporated the TERRA-4G sequence between two double-stranded DNA/RNA hybrid spacers, which were separately attached to two optically trapped polystyrene beads in a home-made laser tweezers instrument with 0.36 nM nucleic acid concentration (Figure 15 A). A force-ramp procedure was carried out to increase the tension in the DNA/RNA hybrid by moving the two laser traps apart in a 10 mM Tris buffer (pH 7.4) with 100 mM K\(^+\) at 23 °C. The increasing tension eventually led to unfolding events accompanied by a change in contour length (\(\Delta L\)) in the hybrid construct (see the force-extension (\(F-X\)) curves in Figure 15 B). Since unfolding of an intermolecular species would immediately drop the force to 0 pN, our observation that \(F-X\) curves are continuous after each rupture event confirmed the intramolecular nature of the unfolding transition.
Figure 15: Mechanical unfolding and refolding experiments. A) Laser tweezers set-up in which a single-stranded RNA containing a TERRA or a control fragment (sequences shown at the top) is sandwiched by two DNA/RNA hybrid handles attached to two optically trapped beads. Wells represent energy potentials for traps. B) Typical F-X curves that contain ~5.5 nm (two left sets) and ~9.3 nm (two right sets) transition events. The last set of the F-X curves shows Worm-Like-Chain fittings for unfolding (gray) and refolding (black) processes (see Methods). For clarity, each curve is offset in the x axis.
Figure 16 A summarizes 686 $\Delta L$ measurements in a histogram in which two major populations with $\Delta L$ of 9.3 ± 0.1 nm and 5.5 ± 0.3 nm are seen. To increase the confidence level, we calculated $\Delta L$ from each data point before the unfolding transition (see Figure 17 and SI). In the subsequent $\Delta L$ probability density distribution (Figure 16 B), two populations are again present. Finally, using bootstrap statistical analyses (see SI), these two populations clearly reveal their identities in Figure 16 C. It is noteworthy that these two populations represent predominant species in the TERRA-4G sequence. However, it does not rule out the presence of minor populations.
Figure 16: The PoDNano analyses on the change in contour length ($\Delta L$) upon unfolding the structures in the TERRA. A) Histogram of $\Delta L$ measured by the two points flanking the rupture event ($N=686$). B) Kernel density distribution of $\Delta L$ based on the data points obtained before the rupture event. C) The populations deconvoluted by bootstrapping of the kernel density distribution in B. Solid curves depict Gaussian fittings.
Figure 17. Plots of the change in contour length ($\Delta L$) vs force ($F$) based on the F-X curves shown in Figure 15 B. After an F-X curve was split into the stretching (gray) and relaxing (black) traces (Figure 15 B), the change in extension ($\Delta x$) was calculated by subtracting the former trace from the latter at a particular force. The resulting $\Delta x$ was then converted to the $\Delta L$ using equation S1. Histograms depicting folded (upper Gaussian populations) and unfolded (lower Gaussian populations) populations are shown to the left of the F-$\Delta L$ traces.
The percentage formation of these folded species in the 100 mM K\(^+\) buffer was determined as 41%. When a buffer with 100 mM Li\(^+\) buffer was used, the percentage of folded species dropped to 18%, which is consistent with the observation that K\(^+\) is a better stabilizer of G-tetrads than Li\(^+\). When we scrutinized the features observed in Li\(^+\), most of them were different from those in the K\(^+\) buffer (Figure 15 B) either in the transition size or the unfolding pattern, suggesting the structures formed in each of these two buffers are distinct. With a control sequence, 5’–UCU CAU GUU UGA CAG CUU AUC AUC GAU AA–3’, 9% of folded structure was observed in the K\(^+\) buffer. This 9% folding probability reflects the known propensity of RNA to adopt a variety of alternative structures. Our results strongly suggest the formation of TERRA GQs at the single molecular level in the 100 mM K\(^+\) buffer at pH 7.4. Additional evidence for GQ formation was provided from CD spectroscopic experiments on the same RNA in which a minimum at 241 nm and a maximum at 263 nm in the CD spectrum is consistent with known parallel conformations of GQs\(^{188}\) (Figure 18 A). In UV melting experiments, we observed increasing melting temperatures (\(T_m = 60\)–79 °C) for the folded GQ species (Figure 18 B) in a buffer series containing 5-100 mM K\(^+\) (Figure 18 C). The experimental procedure and measured parameters are described in the supporting information. This trend is consistent with that observed for the percentage formation of folded species (18-41%) observed by the laser tweezers experiments in the same buffer series.
Figure 18. Biophysical characterization of TERRA-4 oligonucleotide r[(GGGUUA)₃GGG]. A) CD spectra of TERRA in 100 mM KCl. B) Folded fraction curves derived from UV melting experiment of TERRA in 100 mM KCl. ΔGₘₐₜ is estimated for T =296 K. Tₘ is accurate to ± 1 °C. C) Dependence of TERRA melting temperature on potassium concentration.
We hypothesized that the population with $\Delta L = 9.3$ nm assumed a parallel GQ conformation. To test this, we calculated the contour length per nucleotide ($L_{nt}$) using the expression, $L = N \times L_{nt} = \Delta L + x$ (eqn 1), where $N$ is the number of nucleotides in the folded structure and $x$ is the end-to-end distance of the folded structure. By taking $N = 21$ nt and $x = 1.5$ nm measured from the bimolecular TERRA GQ, (PDB ID 2KBP, Figure 19), equation 1 gives $0.51 \pm 0.12$ nm for $L_{nt}$. This value is consistent with literature, validating our assignment.
Figure 19. End-to-end distance measurements of A) TERRA G-quadruplex, B) putative G-triplex-like structure, and C) putative duplex-like structure. These measurements are based on the PDB structure (ID, 2KBP). By using the equation, \( L = N \times L_{nt} = \Delta L + x \), (eqn. 1 in the main text), \( \Delta L \) for the triplex-like structure is calculated as 5.8 nm \( (L_{nt}=0.51 \text{ nm/nt}, N=15 \text{ nt}; x=1.9 \text{ nm}) \); \( \Delta L \) for the duplex-like structure is calculated as 2.1 nm \( (L_{nt}=0.51 \text{ nm/nt}, N=9 \text{ nt}; x=2.5 \text{ nm}) \).
4.3.2 Multiple Species are Present in the TERRA

We propose that the population with the smaller $\Delta L$ (5.5 nm) is a partially folded structure. This is based on the analysis that none of the known DNA or RNA GQ structures would give rise to $\Delta L < 8$ nm (Table 1) when unfolded along the 5'-3' trajectory. This is the first time a partially folded species has been directly observed in the TERRA sequence.

Using the Jarzynski theorem, we calculated the change in free energy of unfolding, $\Delta G_{\text{unfold}}$, for the 9.3 nm population as $11.3 \pm 1.0 \text{ kcal/mol}$ with a bias of $0.7$ kcal/mol at 296 K (Table 2). The $\Delta G_{\text{melt}}$ estimated from the thermal melting experiments is $16.1 \pm 0.5$ kcal/mol (Figure 18 B). The overestimation of $\Delta G_{\text{melt}}$ may originate from inaccuracies in extrapolating a $\Delta G$ value from the melting temperature (> 350 K) to the temperature set for the single-molecule experiments (296 K). Previously, we found that under molecularly crowded conditions, the parallel DNA GQ has $\Delta G_{\text{unfold}}$ of $8.7 \pm 0.4$ kcal/mol. Since it is known that crowded solutions stabilize GQs by ~2.0 kcal/mol, we estimated $\Delta G_{\text{unfold}}$ for a parallel DNA GQ to be 6.7 kcal/mol in diluted solutions, such as the 100 mM K$^+$ buffer used here. Compared to this value, $\Delta G_{\text{unfold}}$ for the RNA GQ (11.3 kcal/mol) is significantly larger, indicating that the RNA GQ is more stable than the DNA GQ. Such a result is consistent with the stabilizing presence of the 2'-OH group in RNA.

As mechanical stability of a structure can be evaluated by its rupture force, we analyzed the unfolding forces for the TERRA species. The overall rupture force profile
shows more than one population (Figure 20 A), confirming the formation of multiple species in the sequence. To clarify the rupture force for each $\Delta L$ sub-population, we performed trace-by-trace random deconvolution on the two $\Delta L$ populations shown in Figure 16. This treatment reveals that both the fully folded RNA GQ (Figure 20 B) and the partially folded populations (Figure 20 C) each have two separate rupture force populations, suggesting that at least two sets of RNA GQs are present with partially folded structures. It is possible that one set of structures accommodates a complete set of $K^+$ ions with stronger mechanical and thermodynamic stabilities (11.3 kcal/mol, Table S2) while the other has compromised ion intercalations with weaker stabilities (6.3 kcal/mol). It is interesting that the partially folded structure has a rupture force comparable to that of RNA GQ (Figure 20), which suggests that the partially folded structure is more rigid but also more fragile than the fully folded GQ. It has been recently observed that rupture force for a TERRA construct that contains 5 telomeric repeats is lower than what we have observed here. It is possible that the additional telomeric repeat may reduce the mechanical stability of the RNA quadruplex.
Figure 20: Rupture force analyses. A) Histogram of overall rupture force for the TERRA sequence (N=686). B) Rupture force distributions (20 ± 2 pN and 37 ± 2 pN) for the population with ΔL = 9.3 nm. C) Rupture force distributions (22 ± 3 pN and 38 ± 3 pN) for the population with ΔL = 5.5 nm. Solid curves depict Gaussian fittings.

To overcome an energy barrier that leads to the unfolding of GQs, both the rupture force and the distance between the folded and the transition states (x_{\dagger}) play a role.\(^\text{116,103}\) It has been shown that x_{\dagger} is rather small for hTelO GQ (< 1nm).\(^\text{120,194,195}\) For the TERRA GQs investigated here, Evans model\(^\text{196}\) yielded a similar x_{\dagger} value (0.45 nm, see Table S2). Such a small distance can be conveniently covered by motor proteins during their translocation steps. Therefore, the rupture force of GQs can be used to
estimate whether these structures can serve as an effective mechanical roadblock to the motor proteins by comparison with the stall force of the enzymes.

The rupture forces from the 5′-3′ trajectory for both hTelo and TERRA GQs (Table S2) are higher than the stall forces measured in the same direction for RNA polymerases, most of which are less than 20 pN.\textsuperscript{197-199} This suggests a need for a cell to have a means, for example helicase, to resolve GQ structure when it acts as a mechanical impediment to processes associated with the telomere along the 5′-3′ direction. Although the stall force for an RNA helicase has not yet been directly measured, it is estimated to be 30 pN for a DNA helicase.\textsuperscript{200} To overcome the mechanical stability of the TERRA GQ, we expect the stall force of an RNA helicase to exceed 35 pN.

4.3.3 Partially Folded Structures are Intermediates to the RNA G-Quadruplexes

The partially folded RNA structure ($\Delta L = 5.5$ nm) may serve as an intermediate to the RNA GQ according to a triangular (Figure 21 A) or linear three-state model (Figure 21 C). To investigate the kinetics, we performed force-jump experiments,\textsuperscript{101} in which the structure in the DNA/RNA hybrid was unfolded first, followed by a quick relaxation to 0 pN force with a time constant of 4 ± 1 ms. Refolding during incubation is then confirmed by the presence of a rupture event in subsequent force-ramping procedures. The identity of the folded species is then revealed by the $\Delta L$ using PoDNano analysis.
The refolding probabilities for the RNA GQ \((Q, \Delta L = 9.3 \text{ nm})\), partially folded \((PF, \Delta L = 5.5 \text{ nm})\), and unfolded \((U, \Delta L = 0)\) species are shown in Figure 19B. While the GQ and unfolded species display opposite behaviors over time, as expected, a transient probability maximum, which is statistically significant, appears to show at \(~ 60\) seconds for the partially folded species. The appearance of the probability maximum directly rules out that the partially folded structure is terminally formed, which should manifest a monotonic change in the refolding probability according to a first-order two-state model. This result strongly suggests the intermediate nature of the partially folded species. The probability density distribution for each species exhibits equally good fits to the equations analytically solved for the two models described in Figure 21 A and C). The fact that two sequential refolding or unfolding events were rarely observed led us to propose that the triangulation model (Figure 21 B) is more likely. The same transition model has been observed for hTelo GQ.\(^{28}\)
Figure 21: Transition kinetics of the TERRA species along the 5'-3' trajectory. A) Schematic of a triangular model. G-quadruplex (Q), partially folded (PF), and unfolded (U) species are shown with black, dark, and gray backbone strands, respectively. B) Refolding probability for each TERRA species with time. The solid curves depict the fitting according to the analytically solved kinetic equations for the triangular model. C) Schematic of proposed linear model.
Unfolding or refolding of a GQ is anisotropic from different directions.\textsuperscript{120} To be observable, an intermediate must have a substantial presence along trajectories surveyed by an ensemble-average technique. The partially folded intermediate observed here should adopt a conformation that is compatible with current 5′-3′ unfolding or refolding trajectory. An intramolecular structure containing two G-tracts\textsuperscript{184} or a G-triplex structure\textsuperscript{122} are examples of possible topologies that fulfill this requirement (Figure 19). It is noteworthy that the latter conformation has a similar $\Delta L$ (5.8 nm, Figure 19 B) as our measurement (5.5 nm, see above). In comparison, the two G-tracts structure is expected to have a $\Delta L$ of 2.1 nm (Figure 19 C). However we cannot unambiguously assign the intermediate structure at this stage. It is also noteworthy that the 5′$\leftarrow$$\rightarrow$3′ unfolding or refolding trajectory shares the same directionality as that employed by various motor proteins, such as a polymerase or a helicase, processing a nucleic acid template. As the motor protein moves along a DNA or RNA template in the 5′$\leftarrow$$\rightarrow$3′ direction, it may interact with a secondary structure, such as a GQ, that is formed either ahead of the protein or behind its trail. Such an interaction generates a tug-of-war force between the motor protein and the secondary structure. Whether a secondary structure can serve as an effective mechanical impediment to the motor protein is dependent on the relative strength between the mechanical stability of the secondary structure along the 5′-3′ trajectory and the stall force of the motor protein in the same direction.
4.3.4 Comparison between Human Telomeric DNA and RNA G-Quadruplexes

Comparison of the formation kinetics between hTelo and TERRA GQs reveals that, under the conditions of our force experiments (100 mM K\(^+\)), the formation of the latter GQ is slower (0.03 ± 0.01 s\(^{-1}\)) than the former (0.10 ± 0.03 s\(^{-1}\)). Although both rates are slower than those for the transcription or replication per nucleotide,\(^{201}\) it is still possible to form GQs during synthesis of DNA or RNA since polymerases pause frequently during these processes.\(^{199}\) If the relatively faster formation of hTelo GQ leads to stalling of an RNA polymerase during TERRA transcription, this may be followed by the formation of thermodynamically more stable TERRA GQs. Since the mechanical stability of the TERRA GQ is stronger than that of the hTelo GQ (Table S2), the former species is expected to resist unwinding by a helicase more efficiently than the latter. This could lead to a longer lifetime for the TERRA GQ and shift the population equilibrium towards TERRA GQ, which in turn gives rise to an enduring effect on the transcription machinery or other biochemical processes related to the telomere as suggested in recent studies.\(^ {176}\)
4.4 Conclusion

In summary, we have observed the formation of intramolecular GQs and partially folded structures in human TERRA. These species follow a three-state kinetics model in which the partially folded species serve as intermediates to the RNA GQs. We have also revealed that TERRA RNA GQs have stronger mechanical stability, more negative thermodynamic free energy, and slower formation kinetics than hTelo DNA GQs. The comparatively higher mechanical stability of both telomeric DNA and RNA GQs in relation to stall forces of RNA polymerases suggests that helicase activities may be required to resolve structures. Finally, the view that GQs possess flexible roles in telomere regulation is in part supported by the temporal separation in their formation kinetics and difference in mechanical stability, observed between telomeric DNA and RNA GQs.
4.5 Supporting Information

Materials and Methods

**TERRA Construct**

To prepare the DNA-RNA hybrid construct, an insert encoding the TERRA quadruplex-forming sequence (5'- TTAGGG TTAGGG TTAGGG TTAGGG TTA -3') was cloned into a pBR322 plasmid (NEB, Ipswich, MA) between EcoRI and HindIII sites. The template for *in vitro* RNA synthesis (MEGAScript, Invitrogen) was generated from PCR amplification of the cloned plasmid using a forward primer containing the T7 promoter. Handle A (527 bp) was obtained from PCR amplification of the unmodified pBR322 plasmid (nucleotides 3836-1) with an EcoRI digestion site on the forward primer, followed by EcoRI (NEB, Ipswich, MA) digestion and end-labeling with DNA Polymerase I, Large (Klenow) Fragment (NEB) in the presence of digoxigenin (Dig)-dUTP. Handle B (1473 bp) was also obtained from PCR amplification of the unmodified pBR322 plasmid (nucleotides 31-1473) with a 5'-biotin labeled reverse primer. The control construct was prepared similarly by replacing the TERRA sequence with the wild type sequence between the EcoRI and HindIII sites of pBR322.

The DNA-RNA hybrid constructs were prepared by annealing the RNA and two DNA handles via heating at 95 °C for 5 min and then gradually cooling down to room temperature for 4 h.
**Single-Molecule Experiments**

Laser tweezers experiments were performed by attaching the DNA-RNA hybrid constructs (~0.25 ng/μL) to the anti-Dig coated polystyrene beads (diameter 2.10 μm, Spherotech, Lake Forest, IL) via the Dig/anti-Dig linkage for 30 min at room temperature. These beads and the streptavidin coated beads (1.87 μm, Spherotech Inc.) were injected into a microfluidic chamber and dispersed into a buffer consisting of (10 mM Tris, 100 mM KCl, 1 mM EDTA) and (2 μL RNase) inhibitor (RNAsecure, Invitrogen) at the pH 7.4. The two different beads were separately trapped by two laser foci in a house-made laser tweezers instrument. The DNA-RNA construct was tethered between the two trapped beads by moving the trapped beads to touch each other. Each tethered construct was repeatedly stretched and relaxed between 0 and 60 pN at a loading rate of 5.5 pN/s to record the force-extension \( F-X \) curves at 1000 Hz using a Labview® program. Histograms of rupture force and \( \Delta L \) were prepared by measuring multiple \( F-X \) curves from different molecules. Variation between individual molecules is not significant.

The accuracy of the force measurement of the laser tweezers instrument was confirmed by the observation of the 65 pN plateau indicative of the denaturation of single dsDNA molecules. The accuracy of the distance measurement of the instrument was confirmed by the determination of the contour length per DNA nucleotide. To this end, we first obtained \( F-X \) curves of a DNA construct that contains a DNA hairpin with a tetrathymine loop (underscored), 5'-GC \( (T)_{10} \) GC \( TTTT \) GC \( (A)_{10} \) GC-3', which is sandwiched between two dsDNA handles as described in literature. These \( F-X \) curves...
were fitted by a sequential model\textsuperscript{152} that uses two Worm Like Chain (WLC) functions\textsuperscript{31} to represent the ssDNA hairpin and dsDNA handles, respectively. The WLC model is expressed as $x/L = 1 - \frac{1}{2} \left( \frac{k_B T}{F P} \right)^{1/2} + \frac{F}{S}$, in which $x$ is the end-to-end distance at a particular force $F$, $L$ is the contour length, $k_B$ is the Boltzmann constant, $T$ is absolute temperature, $P$ is the persistence length, and $S$ is the stretch modulus. This fitting yielded 0.44 ± 0.02 nm per DNA nucleotide, which is identical with that measured by Block and coworkers\textsuperscript{117} and within the range for those obtained by others\textsuperscript{202-205}.

\textit{Force-Jump Experiments}

The detailed force-jump procedures have been described elsewhere\textsuperscript{101}. Briefly, the TERRA construct tethered between two beads as stated above was stretched until the structure was unfolded, then relaxed to 0 pN with a time constant of 4 ± 1 ms, followed by incubation for varied time durations (0-300 s) to monitor the refolding of the structure. Rupture events in the subsequent unfolding curves that start at 10 pN by another force jump are indicative of secondary structures refolded during the incubation.

For each incubation time, the probability of a folded species was calculated by the ratio of the unfolding events observed for that species in subsequent events to the total subsequent pulling curves. Subtracting the probability of all folded species from unity yielded the probability of unfolded species.
In this statistical data analysis, the change in extension ($\Delta x$) due to the unfolding of a secondary structure sandwiched between two nucleic acid handles was obtained from each set of stretching and relaxing traces at a particular force $F$. The change in contour length ($\Delta L$) was then calculated using an expression derived from the WLC model,

$$\frac{\Delta x}{\Delta L} = 1 - \frac{1}{2} \left( \frac{k_BT}{FP} \right)^{\frac{1}{2}} + \frac{F}{S} \quad [\text{eqn S1}],$$

under the condition that the secondary structure has a negligible length compared to that of the nucleic acid handles (<5%), which is true for the TERRA construct. The values of $P$ (10.6 ± 1.8 nm) and $S$ (191 ± 61 pN) for the hybrid DNA/RNA handles were obtained from the sequential fitting on the $F$-$X$ curves (n=10, see Figure 15 B for one such fitting) using two WLC models to represent the TERRA and the RNA/DNA hybrid handles, respectively (see above). These values are consistent with those reported in literature and used in equation S1 to calculate $\Delta L$ at experimental temperature 23 °C. The $\Delta L$ values thus obtained were plotted with $F$ (the $\Delta L$-$F$ plots in Figure 17) from which $\Delta L$ for each transition was calculated by taking the difference between the two $\Delta L$ populations (Figure 17, left). These $\Delta L$ values (9.7 ± 0.8 nm, n=10) are identical within experimental error with those obtained from the sequential fittings discussed above (10.7 ± 1.1 nm, n=10), validating the new $\Delta L$ determination approach.

For each $\Delta L$, the average standard error ($\sigma$) from the regions immediately before and after the unfolding transitions was obtained. This value served as the width in a Gaussian kernel, which was used to construct a kernel density distribution for the $\Delta L$
accompanying an unfolding event\textsuperscript{119} (see Figure 16 B). To identify individual populations in the constructed $\Delta L$ histogram, bootstrapping analyses with 3,000 random resampling were performed. From each resampling, two highest peaks were identified by the Igor\textsuperscript{®} program (WaveMatrics, Portland, OR). These selected peaks were used to construct the bootstrap histograms as shown in Figure 16 C.

**Circular Dichroism**

Circular dichroism spectra were recorded on a Chirascan spectropolarimeter (Applied Photophysics, Leatherhead, UK) using a 1 mm path length quartz cuvette at an oligonucleotide concentration of 10 $\mu$M in 10 mM Tris buffer (pH 7.4) containing 100 mM potassium chloride. Sample preparation involved heating at 95 $^\circ$C for 5 min followed by slow cooling to room temperature. Scans were performed at 25 $^\circ$C over a range of 220 nm to 340 nm and each final spectrum is the average of four scans taken with a step size of 1 nm, a time per point of 1 s, and a bandwidth of 0.5 nm. A buffer only blank was subtracted from each spectrum and data were zero corrected at 340 nm.

**UV Thermal Experiments**

UV melting and annealing curves were collected using a Cary 100 UV-Visible spectrophotometer (Agilent) by measuring absorbance at 295 nm as a function of the temperature. Oligonucleotide solutions were prepared to a final concentration of 5 $\mu$M in 10 mM lithium cacodylate buffer (pH 7.4) containing potassium chloride at specified
concentration. Where necessary, lithium chloride was added to maintain an ionic strength equivalent to 100 mM KCl. The melting experiments were performed in 10 mm path length quartz cuvettes with 1 mL of buffered oligonucleotide solution. A steady stream of nitrogen was applied to prevent condensation at low temperatures and the solutions were covered with a layer of mineral oil to minimize evaporation. Temperature ramps consisted of low → high → low → high cycles with a temperature range of 10-90 °C. A ramp rate of 0.25 °C/min was maintained for all experiments with data collection every °C. Melting temperatures were determined as where fraction folded is equal to 0.5.  

Temperature-independent van’t Hoff free energy changes for G-quadruplex formation ($\Delta G_{melt}$) were determined using the equation $\ln K(T) = -\Delta H / RT + \Delta S / R$, where $K$ is the equilibrium constant for G-quadruplex formation, with the assumption of a two-state model between the single-stranded to the folded state. Plotting $\ln K(T)$ versus $1/T$ allowed us to access the values of $\Delta H$ and $\Delta S$ from the slope and intercept, respectively, of the linear regression in the interval $0.15 < \theta < 0.85$. Free energies were calculated at 296 K using the Gibbs equation $\Delta G_{melt} = \Delta H - T \Delta S$.

**Change in Free Energy of Unfolding ($\Delta G_{unfold}$)**

Change in free energy for the mechanical unfolding of the TERRA structures ($\Delta G_{unfold}$) were calculated by the Jarzynski equation for non-equilibrium systems:

$$\Delta G_{unfold} = -k_BT \ln \sum_{i=1}^{N} \frac{1}{N} \exp \left( -\frac{W_i}{k_BT} \right) \text{ [eqn S2]},$$

where $N$ is the number of observations and
$W$ is the non-equilibrium work done to unfold the structure(s), which is equivalent to the hysteresis area between unfolding and refolding $F$-$X$ curves (see Figure 13B).

**Analytical Solutions for Unfolding and Refolding Rate Constants**

For the model shown above, the rate equations for the formation of each species at time $t$ can be written as,

\[
\begin{align*}
\frac{d [U]}{dt} &= k_2 [PF]_t + k_6 [Q]_t - (k_1 + k_5) [U]_t \\
\frac{d [PF]}{dt} &= k_1 [U]_t + k_4 [Q]_t - (k_2 + k_3) [PF]_t \\
\frac{d [Q]}{dt} &= k_5 [U]_t + k_3 [PF]_t - (k_4 + k_6) [Q]_t
\end{align*}
\]

where $[U]_t$, $[PF]_t$, and $[Q]_t$ are the concentrations of the unfolded, partially folded, and G-quadruplex species at time $t$, respectively.

Using the boundary condition, $[U] = [U]_0$, $[PF] = 0$, and $[Q] = 0$ at $t = 0$, the expressions for the unfolded ($U$), the partially folded ($PF$), and the quadruplex ($Q$) species can be obtained as follows.

\[
[U]_t = [U]_0 \left[ \frac{\beta}{\gamma_1 \gamma_2} + \frac{(\alpha \gamma_1 - \gamma_2 - \beta) e^{-\gamma_1 t}}{\gamma_2 - \gamma_1} + \frac{(\gamma_2^2 - \alpha \gamma_2 + \beta) e^{-\gamma_2 t}}{\gamma_2 - \gamma_1} \right] \quad \ldots \ldots \ldots \text{[eqn S3]}
\]

\[
[PF]_t = [U]_0 \left[ \frac{\xi}{\gamma_1 \gamma_2} + \frac{(k_3 \gamma_1 - \xi) e^{-\gamma_1 t}}{\gamma_2 - \gamma_1} + \frac{(\beta - k_3 \gamma_2) e^{-\gamma_2 t}}{\gamma_2 - \gamma_1} \right] \quad \ldots \ldots \ldots \ldots \text{[eqn S4]}
\]
where,  
\[ \alpha = k_2 + k_4 + k_5 + k_6 \]
\[ \beta = k_2k_4 + k_2k_5 + k_4k_6 \]
\[ \varepsilon = k_1k_6 + k_2k_3 + k_3k_6 \]
\[ \delta = k_1k_4 + k_1k_5 + k_3k_5 \]
\[ \gamma_1, \gamma_2 = \beta + \varepsilon + \delta \]
\[ \gamma_1 + \gamma_2 = \alpha + k_1 + k_3 \]

The probability for each species at time \( t \) can be expressed as,

\[ P_{U,t} = \frac{[U]_t}{[U]_t + [PF]_t + [Q]_t} \]
\[ P_{PF,t} = \frac{[PF]_t}{[U]_t + [PF]_t + [Q]_t} \]
\[ P_{Q,t} = \frac{[Q]_t}{[U]_t + [PF]_t + [Q]_t} \]

Since \([U]_t + [PF]_t + [Q]_t = [U]_0\), equations S3, S4 and S5 become,

\[ P_{U,t} = \left[ \frac{\beta}{\gamma_1, \gamma_2} + \frac{(\alpha\gamma_1^2 - \beta) e^{-\gamma_1 t}}{\gamma_2 - \gamma_1} + \frac{(\gamma_2^2 - \alpha\gamma_2 + \beta) e^{-\gamma_2 t}}{\gamma_2 - \gamma_1} \right] \]  \[\text{[eqn S6]}\]

\[ P_{PF,t} = \left[ \frac{\varepsilon}{\gamma_1, \gamma_2} + \frac{(k_3\gamma_1 - \varepsilon) e^{-\gamma_1 t}}{\gamma_2 - \gamma_1} + \frac{(\beta - k_3\gamma_2) e^{-\gamma_2 t}}{\gamma_2 - \gamma_1} \right] \]  \[\text{[eqn S7]}\]

\[ P_{Q,t} = \left[ \frac{\delta}{\gamma_1, \gamma_2} + \frac{(\beta - \delta) e^{-\gamma_1 t}}{\gamma_2 - \gamma_1} + \frac{(\delta - k_4\gamma_2) e^{-\gamma_2 t}}{\gamma_2 - \gamma_1} \right] \]  \[\text{[eqn S8]}\]

The equations S6, S7, and S8 were used to fit the probabilities of the unfolded, the partially folded, and the G-quadruplex species, respectively, at different incubation time (Figure 21 B).

The analytical solutions for the linear three-state model shown in Figure 21 C are the same set of equations (S6 to S8) with \( k_i \) and \( k_j \) values set to be zero.
Table 1. $\Delta L$ calculated according to equation 1 in the main text ($\Delta L = L - x$). The contour length $L$ is calculated to be 9.45 nm and 12.21 nm respectively for DNA and RNA core sequences (21 nucleotides) that form G-quadruplexes. The end-to-end distances ($x$) are measured from NMR/X-ray structures of respective G-quadruplexes.

<table>
<thead>
<tr>
<th>Structure</th>
<th>Hybrid 1</th>
<th>Hybrid 1</th>
<th>Hybrid 2</th>
<th>Hybrid 2</th>
<th>Parallel</th>
<th>Basket</th>
<th>Basket</th>
<th>TERRA</th>
</tr>
</thead>
<tbody>
<tr>
<td>PDB code</td>
<td>2JSM</td>
<td>2GKU</td>
<td>2HY9</td>
<td>2JPZ</td>
<td>2JSL</td>
<td>1KF1</td>
<td>143D</td>
<td>2KF8</td>
</tr>
<tr>
<td>$x$ (nm)</td>
<td>1.52</td>
<td>1.41</td>
<td>1.1</td>
<td>1.12</td>
<td>1.5</td>
<td>1.043</td>
<td>2.08</td>
<td>2.02</td>
</tr>
<tr>
<td>[5'-3']</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\Delta L$ (nm)</td>
<td>8.56</td>
<td>8.67</td>
<td>8.98</td>
<td>8.96</td>
<td>8.58</td>
<td>9.04</td>
<td>8.00</td>
<td>8.06</td>
</tr>
</tbody>
</table>
Table 2. Rate constants \((k_1 \text{ through } k_6, [\text{Average } \pm \text{ SD}, \text{s}^{-1}])\), distance between the folded and the transition state \((s^\ddagger, \text{nm})\), unfolding force \((F_{\text{rupture}}, \text{pN})\), change in free energy of the 5'-3' unfolding of GQ \((\Delta G_{\text{unfold-GQ}}, \text{kcal/mol})\) for human telomeric RNA (top) and DNA (bottom) species.

<table>
<thead>
<tr>
<th></th>
<th>(k_1)</th>
<th>(k_2)</th>
<th>(k_3)</th>
<th>(k_4)</th>
<th>(k_5)</th>
<th>(k_6)</th>
<th>(x^\ddagger)</th>
<th>(F_{\text{rupture}})</th>
<th>(\Delta G_{\text{unfold-GQ}})</th>
</tr>
</thead>
<tbody>
<tr>
<td>TERRA</td>
<td>(7\pm3) \times 10^3</td>
<td>(9\pm1) \times 10^{-2}</td>
<td>(10\pm3) \times 10^{-2}</td>
<td>(7\pm5) \times 10^{-3}</td>
<td>(3\pm1) \times 10^{-2}</td>
<td>(6\pm2) \times 10^{-4}</td>
<td>0.45*</td>
<td>37*</td>
<td>11.3*</td>
</tr>
<tr>
<td>hTelo</td>
<td>(12\pm4) \times 10^3</td>
<td>(23\pm6) \times 10^{-3}</td>
<td>(13\pm2) \times 10^{-3}</td>
<td>(10\pm5) \times 10^{-4}</td>
<td>(10\pm3) \times 10^{-2}</td>
<td>(9\pm1) \times 10^{-3}</td>
<td>0.90*</td>
<td>21*</td>
<td>6.7*</td>
</tr>
</tbody>
</table>

Note, rate constants for hTelo are taken from reference.\(^{28}\)

*For the higher rupture force (37 pN) species in the \(\Delta L=9.3\ \text{nm}\) populations.

#For parallel species in 40% DMSO solution from reference [81].
Chapter 5

Determination of Sphingosine Kinase 2 Activity using Fluorescent Sphingosine by Capillary Electrophoresis

This chapter has been published as an article in Electrophoresis DOI 10.1002/elps.201000495 and reprinted with permission. Publisher: John Wiley and Sons. Copyright © 2011 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim.

5.1 Summary

The study of sphingosine and sphingosine-1-phosphate (S1P) is now widespread due to their immense role as intra- and extracellular messenger molecules. The balance and interplay of these ceramide metabolites is dependent on the activities of kinase and phosphatase enzymes. Sphingosine and S1P are found in very minute quantities in cells; thus, require highly sensitive techniques for quantitative analysis. In this study, we developed a quantitative assay for the determination of sphingosine kinase 2 (SphK2) activities both in vitro and with cell lysates, using CE-LIF. Sphingosine fluorescein (SpFl) was used as substrate. The $K_M$ of SphK2 for sphingosine fluorescein was $2.8 \pm 0.8 \mu M$ with a $V_{max}$ of $2490 \pm 520 \mu M/min$ and a $k_{cat}$ of $1920 \pm 402 \text{ s}^{-1}$. The inhibition of SphK2 was also investigated using four different inhibitors for which 2-(p-Hydroxyanilino)-4-(p-chlorophenyl) thiazole (2-HCT) inhibitor was the most potent for the in-vitro inhibition of SphK2 while N, N-dimethylsphingosine (DMS) did not inhibit but rather increased SphK2 activity. The fluorescence-based approach for the
determination of the enzymatic activity of SphK2 proves to be useful for quantitative determination of SphK2 activities \textit{in vitro} and in cell lysates, and could be extended to single cell analysis, or applied in drug screening.
5.2 Introduction

Over the last few years, sphingolipids have drawn a lot of attention not only as structural components of eukaryotic cell membranes but also as useful bioactive molecules.\textsuperscript{208-209} Some of the bioactive sphingolipids under intense investigation include: ceramide, ceramide-1-phosphate, sphingosine, and sphingosine-1-phosphate (S1P).\textsuperscript{208-212} Most of these bioactive molecules are known to be second messengers, and they play an important role in signal transduction. The sphingolipid metabolites, sphingosine and its precursor ceramide are associated with cell death (apoptosis), while S1P is involved in many critical cellular processes including proliferation, survival and migration, allergic responses, cytoskeletal rearrangements and cell motility, invasion, angiogenesis, vascular maturation, and trafficking of immune cells.\textsuperscript{213} S1P, a simple molecule, can play such diverse roles because it functions not only inside cells,\textsuperscript{214-215} but also as a ligand of cell surface receptors after it is secreted into the extracellular milieu.\textsuperscript{213} Gene deletion studies and reverse pharmacology have provided evidence that many of the biological effects of S1P are mediated via five specific G protein-coupled receptors, now designated S1P\textsubscript{1–5}.\textsuperscript{209}

Cellular levels of S1P are regulated by the concerted action of the enzymes responsible for its formation and degradation.\textsuperscript{209} There are two mammalian isoforms of the sphingosine kinase (SphK) enzyme i.e. sphingosine kinase 1 (SphK1) and sphingosine kinase 2 (SphK2). These two enzymes share overall homology and produce the same product, S1P; however, they display different catalytic properties, subcellular
locations, tissue distribution, and temporal expression patterns during development and possibly have unique and specific functions. SphK1 has pro-survival functions and is mainly a cytosolic protein, whereas SphK2 is a putative BH3-only protein, which inhibits cell growth and enhances apoptosis. SphK2 shuttles between the nucleus and the cytoplasm under different cell conditions and its subcellular location also depends on the cell type. SphK1 does not possess any hydrophobic transmembrane domains, whereas SphK2 has four predicted transmembrane domains. SphK1 prefers D-erythro-sphingosine as a substrate, whereas SphK2 phosphorylates a wider range of sphingoid base substrates, including phytosphingosine and dihydrosphingosine. These two enzymes generate S1P through the phosphorylation of sphingosine. However, S1P can be reversibly dephosphorylated by two known mammalian phosphatases; sphingosine-1-phosphate phosphatase 1 (SPP1) and sphingosine-1-phosphate phosphatase 2, or irreversibly degraded by a pyridoxylphosphate-dependent S1P lyase to hexadecenal and phosphoethanolamine. Ceramide, sphingosine, and S1P are interconvertible and their relative levels determine whether a cell lives or dies.

The role of S1P in various physiological processes is under intense investigation using pharmacological approaches, including treatments with S1P, S1P receptor agonists or antagonists, or SphK inhibitors. For example, S1P promotes estrogen-dependant tumorigenesis of MCF-7 human breast cancer cells and stimulates invasiveness of human glioblastoma cells. Also, endogenous SphK1 has been shown to regulate motility, growth, and chemoresistance of MCF-7 cells, whereas both SphK1 and SphK2 are involved in epidermal growth factor-mediated activation and migration of
MDA-MB-453 breast cancer cells. S1P has also been implicated in vasculogenesis and angiogenesis, atherogenesis, inflammation and immunity, asthma, modulation of vascular barrier integrity, the reproductive system, and the central nervous system. Many lines of evidence implicate S1P generation and signaling as potential targets for treating inflammatory diseases. For example, the potent immunosuppressive drug FTY720 (Fingolimod), a S1P analogue, has shown great clinical potential for the prevention of renal graft rejection and treatment of multiple sclerosis.

A better understanding of S1P signaling pathways, whether intra- or extracellular, might prove to be useful in the development of therapeutics for disorders, such as cancer, atherosclerosis, asthma, allergy, multiple sclerosis, and other autoimmune diseases. S1P is mainly synthesized within the cells but acts as an important extracellular signaling molecule. The mechanism of export of cellular S1P is not clearly understood; however, sphingosine and its derivatives are thought to readily go in and out of living cells. The question still remains whether the intracellular pool of S1P is from endogenous kinase activity or from uptake and storage of exogenous S1P. These studies have been hampered by the inability to quantitatively measure SphK activity in cells.

In a previous study, Lee et al., described a fluorescent CE-based assay for studying SphK1 and SPP1 activity in cells. This highly sensitive and fast technique demonstrated immense capability for determining SphK1 activity in small (10 cells) or
bulk numbers of cells, and a variety of biochemical studies in single cells. In addition, other competitive assays for the determination of SphK activity were discussed. Although the SphK1 isoform has been extensively studied and characterized, the regulation of expression and function of the other isoform, SphK2 remain largely unexplored. In addition, some of the inhibitors of SphK1 such as N, N-dimethylsphingosine (DMS) and phytosphingosine have been shown to activate SphK2.

The present study investigated the use of CE-LIF separation technique to measure the activity of purified SphK2 enzyme. We exploited the CE separation of the fluorescent sphingosine fluorescein (SpFl) and Sphingosine-1-phosphate fluorescein (S1PFl) to develop a method suitable for monitoring the biochemical enzymatic activities of the lipid modifying enzyme SphK2, in vitro. In addition, the inhibition of the purified SphK2 enzyme was investigated using four different inhibitors. The developed CE-based assay was also used to study the effect of the inhibitors on the activity of SphK(s) in NIH-3T3 fibroblast and in MDA-MB-231 breast cancer cells lysates. The in vitro inhibition studies showed that DMS was not an inhibitor of SphK2 but that it increased the SphK2 activity. DMS is known to be an inhibitor of SphK1. Even though, the 2-(p-Hydroxyanilino)-4-(p-chlorophenyl) thiazole (2-HCT) inhibitor showed some inhibition activity to both sphK1 and SphK2, it was however a better inhibitor for SphK2 under in vitro conditions. These studies will be useful for evaluating inhibitors of SphK1 and SphK2 which are valuable enzymes for cell and molecular biology, as well as clinical medicine.
5.3 Results and Discussion

5.3.1 Separation of the In Vitro SpFl (substrate) and S1PFl (product) by CE

We developed three CE separation methods based on Tris, HEPES, and phosphate buffer that can be used for the separation of SpFl and S1PFl standards and compared their migration time, peak height and peak area reproducibility. The Tris buffer was prepared as stated above. The phosphate buffer composed of 50 mM NaH$_2$PO$_4$/Na$_2$HPO$_4$, 20% propanol, 5.0 mM SDC and the HEPES buffer composed of 20 mM HEPES, 20% propanol, 5.0 mM SDC. The three buffers work well at pH 8.5. To test the reproducibility and robustness of the CE separation methods, experiments were repeated several times and the retention times and peak areas averaged. The HEPES buffer provided the best reproducible data for the separation of the standards (Table 3) but very poor reproducibility for the analysis of cell lysates. The Tris buffer was chosen because it was more robust for the separation of the standards (Table 4) and cell lysates. Nevertheless, an internal standard (BOPIDY-Fl C$_{16}$) was used to monitor shifts in the migration time, peak heights, and peaks areas of the fluorescent lipids, such as the substrate (SpFl) and product (S1PFl) in the in vitro and cell lysates assays. To calculate the % phosphorylation, the peak areas of SpFl and S1PFl were normalized with the peak area of the internal standard to correct for changes in peak areas due to sample losses.

In vitro study of the conversion of SpFl to S1PFl was carried out by adding 7.20 nM purified SphK2 enzyme to an assay buffer containing 7.43 nM SpFl and 2.10 nM BodipyFl C$_{16}$ as internal standard. The reaction mixture was separated using the CE Tris
buffer under reversed polarity conditions. The results are shown in (Figures 22 & 23). The SpFl and S1PFl peaks were well resolved and consistent with the CZE mechanism of separation. The more negatively charged S1PFl, migrated faster towards the positive end of the capillary (detector), eluting first, followed by BodipyFl C16 and then SpFl. The two peaks observed for SpFl and S1PFl in Figure 23 were a result of the separation of the 5’ and 6’ sphingosine labeled isomers. Isomer separation was achieved by addition of 2.5 mM Mg$^{2+}$ and with slight modification of the optimized separation buffer to 20% propanol, 5mM SDC. Mg$^{2+}$ seems to enhance near baseline separation of the split S1PFl and SpFl peaks. However, for the quantitative determination of SphK2 activity, experiments were performed using the optimized separation buffer i.e. 30% 1-propanol, 10 mM SDC which was more reproducible and robust (Figure 22).

Table 3. Average elution times, and the percent relative standard deviation (%RSD) of peak heights, and peak areas, for ten sequential runs of the SpFl and S1PFl standards separated by CE-LIF using 3 different buffers.

<table>
<thead>
<tr>
<th>Buffer</th>
<th>S1PFl</th>
<th>% RSD of SpFl</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Time</td>
<td>Height</td>
</tr>
<tr>
<td></td>
<td>(mean ± sd)</td>
<td>(%RSD)</td>
</tr>
<tr>
<td>Tris</td>
<td>5.89 ± 0.05</td>
<td>26.5</td>
</tr>
<tr>
<td>HEPES</td>
<td>4.54 ± 0.05</td>
<td>6.1</td>
</tr>
<tr>
<td>Phosphate</td>
<td>6.51 ± 0.52</td>
<td>10.4</td>
</tr>
</tbody>
</table>
Table 4: Average elution times and normalized peak areas and heights of SpFl and S1PFl for the in vitro assay using BodipyFl C_{16} as internal standard. (n = 8). SpFl gave two peaks as a result of isomer separation.

<table>
<thead>
<tr>
<th>Sample</th>
<th>Elution time/min (mean±SD)</th>
<th>Peak area (% RSD)</th>
<th>Peak Height (%RSD)</th>
<th>Separation Efficiency (%RSD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1PFl</td>
<td>3.83 ± 0.03</td>
<td>10930 ± 8.4</td>
<td>6006 ± 10</td>
<td>108831 ± 8.4</td>
</tr>
<tr>
<td></td>
<td>3.88 ± 0.03</td>
<td>12245 ± 8.7</td>
<td>4983 ± 11</td>
<td>68952 ± 22</td>
</tr>
<tr>
<td>BodipyFl C_{16}</td>
<td>4.47 ± 0.05</td>
<td>46604 ± 9.2</td>
<td>23387 ± 7.6</td>
<td>124581 ± 7.2</td>
</tr>
<tr>
<td>SpFl</td>
<td>4.87 ± 0.03</td>
<td>30197 ± 10</td>
<td>13150 ± 7.9</td>
<td>113140 ± 2.9</td>
</tr>
<tr>
<td></td>
<td>4.94 ± 0.04</td>
<td>46753 ± 10</td>
<td>18512 ± 8.1</td>
<td>116675 ± 2.5</td>
</tr>
</tbody>
</table>
**Figure 22.** Electropherogram of separation of and SpFl, BodipyFl and S1PF1 after addition of 7.20 nM SphK2 to 7.43 nM SpFl and 2.10 nM BodipyFl, 0.25 mM ATP, 0.10 µM Na₃OV₄ in 300 µL assay buffer followed by; (A) 5 min, (B) 30 min, and (C) 60 min reaction over a water bath at 37°C. Separation buffer pH, 8.50 made of 100 mM Tris, 30% 1-propanol, 5% EOTrol™ LR and 10.0 mM SDC.
Figure 23. Electropherogram showing the effect of Mg$^{2+}$ ions on separation of 7.43 nM SpFl, 2.10 nM BodipyFl and 2.00 nM S1PFl in a separation buffer of 100 mM Tris, 20% 1-propanol, 5.0 mM SDC, 5% EOTrol$^{TM}$, pH 8.50; and (A) in the absence of Mg$^{2+}$ and (B) 2.50 mM of Mg$^{2+}$.

In addition, *in vitro* studies using purified SphK2 enzyme, were carried out to determine the Michaelis-Menten kinetics constants $K_M$ and $V_{max}$ and the turn over number $k_{cat}$. The initial rates of conversion of SpFl substrate into the product S1PFl were
measured at varying substrate concentrations (Figure 24). Based on the conditions used in this assay, the $K_M$ for SphK2 was determined to be $2.8 \pm 0.8 \mu M$ with a $V_{max}$ of $2490 \pm 520 \mu M/min$, and $k_{cat}$ of $1920 \pm 402 \text{ s}^{-1}$. The value of $K_M$ observed is higher than the (0.6-1.3 $\mu M$) physiological range for sphingosine with SphK$^{250}$ probably due to label attached to the substrate. The determined $k_{cat}$ value is also higher than $k_{cat}$ value of 1265 s$^{-1}$ reported elsewhere for unlabeled sphingosine.$^{251}$

![Figure 24. Lineweaver-Burk plot of initial rate versus SpFl substrate concentration using 21.6 nM of SphK2, 0.10 $\mu M$ Na$_3$OV$_4$, 0.50 mM ATP and BodipyFl C$_{16}$ (1.75 nM) used as internal standard in 300 $\mu$L assay buffer. The experiment was run nine times and the data averaged.](image-url)
5.3.2 Phosphorylation of SpFl by SphK2

Time dependent *in vitro* study of the conversion of SpFl substrate to the S1PFl product was done using purified SphK2 enzyme. CE-LIF was used to monitor the migration time, peak height, and peak areas of the SpFl and S1PFl. To determine the rate of phosphorylation of SpFl by SphK2, 2.0 µL of stock SphK2 (14.4 nM) unless stated otherwise, was added to a vial containing 12.4 nM SpFl substrate, 0.50 mM ATP, 0.10 µM Na$_3$OV$_4$ in 300 µL total volume of the assay buffer. BodipyFl C$_{16}$ (1.75 nM) was used as internal standard. The reaction mixture was placed in a 37 °C water bath. At fixed reaction time intervals, the enzymatic reaction was stopped by taking 50.0 µL aliquots of the reaction mixture and mixing with 50.0 µL solution of 5.0 mM Tris and 0.10% triton X-100 followed by CE-LIF analysis. The ratios of S1PFl, SpFl, and BODIPY-Fl C$_{16}$ peak areas were calculated. The percentage of phosphorylation was calculated by dividing the average ratio of S1PFl and BodipyFl C$_{16}$ peak area by the sum of ratios of S1PFl, SpFl, and BODIPY-Fl C$_{16}$ internal standard peak areas. The activity of SphK2 was determined by the % of SpFl phosphorylated with time. With the conditions used, after one minute of reaction, the SpFl substrate was 50% phosphorylated and gradually increased to 90 % after one hour of reaction as shown in Figure 25. Subsequent experiments were carried out under different substrate, enzyme, and ATP concentrations and the results shown in Figure 25.
5.3.3 Inhibition of SphK2

To investigate the inhibition of the activity of SphK2, three different potential inhibitors namely; phytosphingosine, dihydrosphingosine, and DMS were each added separately to the in vitro reaction mixture just before the addition of the SpFl substrate. The results from these experiments are shown in Figure 26. From these studies
dihydrosphingosine was shown to be a better inhibitor of SphK2 than phytosphingosine. Addition of dihydrosphingosine resulted in a modest inhibition of the initial rate of phosphorylation of SpFl to (~25%) at the 1 min time point followed by a further drop to ~20% for the 5, 10, and 15 min reaction time experiments, and a slight increase in the rate of phosphorylation to ~28% after 1 hour reaction. Phytosphingosine resulted in ~28% initial rate of phosphorylation of SpFl at the 1 min time point followed by a further drop to ~25% for the 5, 10, and 15 min reaction time experiments, and a reasonable gradual increase in the rate of phosphorylation to ~50% after 1 hour. However, DMS did not inhibit SphK2, but rather activated it. As shown in Figure 26, addition of DMS to the reaction mixture resulted in higher % phosphorylation of SpFl at any given reaction time than is the case of the SphK2 alone. Our data is in agreement with some of the published studies which monitored the activity of SphK2 using these three inhibitors and D-erythro-spingsine as substrate. For example the studies by Liu et al., showed that D-erythro-dihydrosphingosine was a better substrate for SphK2 than D-erythro-spingsine.²²¹ Our studies also show that dihydrosphingosine is a better substrate for SphK2 and acts as a competitive substrate when added together with SpFl leading to a lower rate of phosphorylation of SpFl. In addition, studies by Liu and coworkers showed that phytosphingosine is a substrate of SphK2 but it is not as a good a substrate as D-erythro-spingsine. Based on our observation we can possibly say that phytosphingosine is a less competitive inhibitor; thus, only slightly inhibits the rate of phosphorylation of SpFl. Liu studies also showed that DMS is not a substrate for SphK2 but a substrate for SphK1. Work by Lee et al., shows DMS to be an inhibitor of SphK1 when SpFl is used.
as the substrate. In this study we investigated the effect of DMS on the rate of phosphorylation of SpFl using SphK2. Our studies show that DMS activates the activity of SphK2 leading to a higher rate of phosphorylation of SpFl. In the literature numerous studies have revealed the potency of DMS in inhibiting SphKs activity and in lowering S1P levels.  

DMS has been reported to be an inhibitor of both SphK1 and SphK2; however, the data for SphK2 inhibition were obtained with expressed enzyme. Recently, Vessey and coworkers reported the rapid and efficient separation of SphK1 and SphK2 from rat and mouse heart cytosol which were used to demonstrate that DMS and the immunomodulator FTY720 inhibit only native SphK1, whereas they activate SphK2.
**Figure 26.** A plot showing the SphK2 catalyzed phosphorylation of SpFl in the presence of SphK2 alone (no added inhibitors) (▲), and in the presence of inhibitors; 0.26 mM phytosphingosine (■), 0.25 mM N, N-Dimethylsphingosine (DMS) (▼) and 0.23 mM, dihydrosphingosine (●). 1µL of stock SphK2 (7.2 nM), 0.10 µM Na₃OV₄ and 0.25 mM ATP were added to 300 µL assay buffer. SpFl substrate concentration was 12.4 nM. The experiment was run nine times and the data collected were averaged.

### 5.3.4 In Vitro Inhibition of SphK1 and SphK2

The compound 2-(p-Hydroxyanilino)-4-(p-chlorophenyl) thiazole (2-HCT) is a competitive inhibitor of SphKs. As shown in Figure 27, 2-HCT was a more potent inhibitor of SphK2 than SphK1 for the studies using SpFl. Vessey and coworkers reported that 2-HCT which inhibits recombinant human SphK,\(^ {256} \) did not inhibit either
form of rat or mouse heart SphK. Thus, their results call for caution when interpreting SphKs inhibitor results and a closer look at prior studies that have used DMS to inhibit both SphK1 and SphK2 activity. Work by the Huwiler’s group indicates that 2-HCT reduces cellular SphK1 activity by down regulation of protein expression rather than directly inhibiting the catalytic activity of the enzyme. In contrast, SphK2 protein expression and cellular activity were not affected by 2-HCT. Their in vitro data showed that even though there is a weak direct inhibitory effect on both SphK1 and SphK2, % inhibition was higher in SphK2 which corroborates our observations.

**Figure 27.** Phosphorylation of SpFl (12.4 nM) in the presence of SphK2 (7.2 nM), 0.5 mM ATP; (▼) No inhibitor, (■) 10 µM (2-HCT) inhibitor, and phosphorylation of SpFl (12.4 nM) with SphK1 (7.2 nM), 0.5 mM ATP; (●) No inhibitor, (▲) 10 µM (2-HCT) inhibitor (n = 5).
5.3.5 Inhibition of SphK(s) in Cells

CE analysis of the NIH-3T3 cell lysate (Figure 28) shows an initial 70% phosphorylation of SpFl which gradually decreased to 50% in 30 minutes. The lysate result illustrates the interplay of SphK and phosphatase enzymes. The initial phosphorylation of the SpFl substrate is due to the enzyme SphK. However, the phosphatase also present in the cell acts on the S1PFl formed, converting it back to SpFl. Analysis of the media alone collected after incubation with the SpFl substrate, did not show S1PFl but only SpFl, implying that the SpFl diffuses through the cells, gets phosphorylated, and remains in the cell. The inhibition studies with the cell lysate indicate that % phosphorylation of SpFl was reduced 15% and 20% with dihydrosphingosine and DMS respectively. This observation is contrary to the results obtained with purified SphK2 where DMS rather increased the % phosphorylation by probably activating SphK2. DMS has been shown to be a predominant inhibitor of SphK1.221,248 Low levels of S1P might be an indication that SphK1 instead of SphK2 is implicated in the cell lysate studies. Similar results were observed when MDA-MB-231 cells were loaded with SpFl.
Figure 28. CE analysis of the inhibition of SpFl phosphorylation in NIH-3T3 cells incubated with 37.2 nM of SpFl substrate showing; no inhibitor (black); 19.0 µM DMS (Blue) and addition of 17.0 µM dihydrosphingosine (green).
5.4 Conclusion

The assay described here for the determination of SphK2 activity by CE-LIF offers a high sensitivity and fast analysis compared to conventional HPLC and TLC assays used for monitoring SphK activity. It can be applied in the study of purified enzymes, cell lysates, and may be extended to the study of single cells. Furthermore it can be used to provide a quick estimate of different SphK inhibitor activity. The easy uptake of fluorescent sphingosine into NIH 3T3 and MDA-MB-231 cells supports the notion that SpFl is a versatile lipid probe that may be easy to work with and may be applicable to many cell types. Given the increased interest in the cell signaling properties and biochemical measurements of the ceramide-sphingosine metabolic pathway, the developed assay may find many applications in biological and biomedical investigations.
5.5 Experimental Section

5.5.1 Materials.

Sphingosine fluorescein (SpFl) and sphingosine-1-phosphate fluorescein (S1PFl), 5’ and 6’ labeled isomers, and 2-(p-Hydroxyanilino)-4-(p-chlorophenyl) thiazole (2-HCT) were acquired from Echelon Biosciences Inc. (Salt lake City, UT). Recombinant human sphingosine kinase 2 (SphK2) was purchased from BPS Biosciences Inc. (San Diego, CA). MDA-MD-231 and NIH-3T3 cell lines were a kind gift from Dr. Basu (Kent State University, OH). Magnesium chloride, sodium chloride, methanol, tri-(hydroxymethyl) aminomethane (Tris), RPMI media and all other tissue culture reagents were purchased from Fisher Scientific (Fair Lawn, NJ). Sodium deoxycholate (SDC), 4-(2-hydroxyethyl) piperazine-1-ethanesulfonic acid (HEPES), 1-propanol, adenosine triphosphate (ATP), and sodium orthovanadate (Na$_3$VO$_4$), were purchased from Sigma-Aldrich Inc. (St. Louis, MO). N, N-dimethylsphingosine (DMS) also known as N,N-dimethyl-D-erythro-sphingosine, and phytosphingosine (also known as D-ribo-phytosphingosine, or 4-hydroxysphinganine) were purchased from Avanti Polar Lipids (Alabaster, AL). Dihydrosphingosine (also known as DL-erythro-1,3-dihydroxy-2-aminooctadecane, or sphinganine ) was purchased from Tocris Bioscience, (Ellisville, MO). EOTrol LR (low reverse) polymer solution was purchased from Target Discovery (Palo Alto, CA). Bodipy- 4,4-difluoro-5,7-dimethyl-4-bora-3a,4a-diaza-s-indacene-3-hexadecanoic acid (BODIPY-Fl C$_{16}$), were purchased from Invitrogen (Carlsbad, CA).
5.5.2 Capillary Electrophoresis

Capillary electrophoresis was performed on a Beckman ProteomeLab™ PA800 CE system (Fullerton, CA). The machine was equipped with LIF detector using an air-cooled 3.5 mW argon laser (Beckman Instruments), with excitation at 488 nm and emission at 520 nm, and was installed with a 32 Karat 8.0 software (Beckman Instruments) for instrument control and data analysis. Separations were performed on 30 cm fused-silica capillaries (Polymicro Technologies, Phoenix, AZ, USA) with 50 µm i.d., 375 µm o.d. (20 cm effective length). New capillaries were preconditioned by successively flushing with 1.0 M NaOH, for 30 min and rinsed with distilled deionized H₂O, for 20 min at 20 psi. Hydrodynamic injection was carried out at 0.5 psi for 10 s. During the separation the capillary column was thermostated at 25 °C and an electric field strength of 833 V/cm was applied in reverse polarity mode (anode at the detector). The CE separation buffer was made of 100 mM Tris, 10 mM SDC, 30% 1-propanol, 5% EOTrol™ LR and 2.5 mM MgCl₂ at pH 8.5. For some reactions 5 mM SDC and 20% 1-propanol was used. Increasing 1-propanol in the separation buffer makes it less conductive and leads to longer elution times. SDC reduces the interaction of the analyte with the capillary walls and therefore shortens the elution time and improves the resolution of the peaks. The magnesium ions enhance separation efficiency and the resolution of both the SpFl and S1PFl isomer peaks. EOTrol™ LR in the separation buffer serves as a dynamic coating for the capillary. The SpFl and S1PFl standards were dissolved in methanol to make the stock solutions from which aliquots were mixed and diluted with Triton X-100 or the appropriate buffer. Before each run, the capillary was
rinsed with 0.10 M NaOH for 5 min, distilled H₂O for 3 min, and the separation buffer for 3 min at a pressure of 20 psi. The elution order and identity of the peaks were determined by spiking the reaction mixture with SpFl or S1PFl standards and repeating the separation process.

5.5.3 Sample Preparation

The fluorescent lipids stock solutions were prepared in methanol and stored in -20 °C. Due to the lipids poor solubility and ability to adsorb to storage and analysis vials, fresh aliquots of the stock were prepared for each day’s experiment prior to analysis or loading into cells with minimal transfers and pipetting to prevent loss of sample during preparation. The enzymes were aliquoted into working volumes as soon as they were received from the vendor and stored at -80°C. Each enzyme aliquot was thawed in an ice bath prior to each in vitro assay.

5.5.4 Sphingosine Kinase 2 In Vitro Assay

Active SphK2, a human recombinant N-terminal His-tagged protein, 69.5 kDa of MW, in a baculovirus infected Sf9 cell expression system was used. The enzyme stock was 20 µg at either 0.53 mg/mL (7.63 µM) or 0.15 mg/mL (2.16 µM) and specific activity of 20 U/mg total proteins as reported by the supplier. For most of the assay performed, the concentration of SphK2 was 14.4 nM or 7.2 nM and a SpFl concentration
of 12.4 nM in assay buffer (50 mM HEPES, 150 mM NaCl, and 5.0 mM MgCl$_2$, 1.0 mM DTT at pH 7.4) in a total reaction volume of 300 µL. The assay was performed in the presence of 0.50 mM, 0.25 mM, or 0.00 mM ATP. The reaction mixture also contained 0.10 µM Na$_3$VO$_4$. In some experiments, DMS, phytosphingosine, dihydrosphingosine, and HCT were added to the reaction mixture each at a time to inhibit SphK2. The inhibition of SphK1 and SphK2 using HCT was also investigated. For the \textit{in vitro} inhibition reactions, 12.4 nM of the SpFl substrate was used and the inhibitor concentration was varied. The reaction mixture was placed in a water bath at 37 °C and 50.0 µL aliquots were taken at fixed time intervals. The enzymatic reaction was stopped by diluting the aliquots with 50 µL solutions of 5.0 mM Tris and 0.10% triton X-100 and frozen in dry ice, ready for CE analysis.

5.5.5 Cell Culture

Two cell lines, MDA-MB-231 and NIH-3T3, were grown simultaneously in a DMEM/high glucose medium (HyClone Laboratories. Inc., UT) to which 10 % fetal bovine serum and 1.0 % penicillin/streptomycin were added. The cells were maintained at 37 °C and 5.0 % CO$_2$ atmosphere. The cells were passed by gently sucking out the media and adding 2.0 mL of trypsin for 2 minutes to dislodge them. Medium was added (20.0 mL) to neutralize the trypsin and 2.0 mL of cells was added to 10 cm plates and incubated overnight. After 16 to 24 hrs, 1.0 µL of the stock SpFl solution (74.5 µM) was added into each 2.0 mL cell volume (37.2 nM effective concentrations) and incubated for
fixed time intervals. For some of the experiments, a SphK enzyme inhibitor was added into the cells in the 10 cm plate before the addition of SpFl substrate. Two such inhibitors were investigated in the cell lysate study namely, DMS and Dihydrosphingosine. For the DMS, 5.0 µL of the stock (7.63 mM) was injected into 2 mL cell volume (19.0 µM final) while 10.0 µL of Dihydrosphingosine stock (3.42 mM) was added into 2.0 mL cell volume to make a final concentration of 17.0 µM. After incubation, the media was sucked out and the cell plates rinsed two times with 1.0 mL PBS and lysed with the Tris-Triton X-100 buffer, centrifuged, and the supernatant frozen in dry ice, ready for analysis.

5.5.6 Michaelis-Menten Kinetics

Michaelis-Menten kinetics constant, $K_M$ and $V_{\text{max}}$ as well as the catalytic constant, $k_{\text{cat}}$ for SphK2 were determined from a Lineweaver-Burk plot of initial rate of phosphorylation versus SpFl substrate concentration. The substrate concentrations were 1.5, 2.0, 3.0, 3.5, 4.0, 5.0, 6.0 and 7.4 µM. The SphK2 enzyme concentration was fixed at 21.6 nM. Six determinations were made for each concentration, and the averaged data fitted to the Lineweaver-Burk equation using Origin (OriginLab Corporation, Northamton, MA) to determine the $K_M$, $V_{\text{max}}$ and $k_{\text{cat}}$. 
6.0 Concluding Remarks

The developments of methods to investigate small biomolecules could lead to discovery of many more biomarkers for diseases, drugs and cofactors for enzymes. Clinical diagnostic methods commonly rely on the use of double-antibody sandwich assays which require the target to have at least two binding sites to bind two different antibodies. This method is less suitable for the detection of small molecules due in part to the fact that small molecules generally have fewer binding sites, which, coupled with the inherent steric hindrance, makes this method less suitable for the detection of small molecule targets. Aptamers are therefore good alternatives to antibodies and are used in different platforms to design different strategies to investigate biologically relevant small molecules and biomarkers. One drawback in the use of aptamers, nevertheless is the limited availability of selected aptamers that can specifically bind a wide variety of small molecules targets. This calls for the need to select more aptamers against a variety of targets rather than reliance on the already available aptamers.

To better explore the biological roles of G-quadruplex structures and their possible therapeutic applications, it is important, first and foremost, to understand the structural stability and dynamics in the human genome. For G-quadruplex structures to effectively inhibit telomerase activity as proposed, the mechanical stability of these structures must be higher than the stall force of the enzyme, from the mechanical perspective. It is therefore relevant to evaluate the mechanical and thermodynamic stability and formation dynamics of these structures under physiological conditions.
It is, however, commonplace that most of these biologically relevant species are in such minute quantities, especially in the cell that more selective and sensitive detection methods are required. Such low abundance of biomolecules usually complicates their investigations. Due to their high sensitivity, fluorescence-based and single-molecule methods are probably the best suited techniques. To this end we utilized capillary electrophoresis with laser-induced fluorescence detection (CE-LIF) and optical laser tweezers, both of which have excellent temporal and spatial resolutions to investigate biologically relevant small molecules.
References


238 S.-S. Chae, R. L. Proia and T. Hla, Constitutive expression of the S1P1 receptor in adult tissues, Prostaglandins & Other Lipid Mediators 2004, 73, 141-150.


244 P. Haenel, P. Andreani and M. H. Graeler, Erythrocytes store and release sphingosine 1-phosphate in blood, FASEB Journal 2007, 21, 1202-1209.


248 D. A. Vessey, M. Kelley, J. Zhang, L. Li, R. Tao and J. S. Karliner, Dimethylsphingosine and FTY720 inhibit the SK1 form but activate the SK2 form of sphingosine kinase from rat heart, Journal of Biochemical and Molecular Toxicology 2007, 21, 273-279.


259 X. Zuo, Y. Xiao and K. W. Plaxco, *High specificity, electrochemical sandwich assays based on single aptamer sequences and suitable for the direct detection of small-molecule targets in blood and other complex matrices*, *Journal of the American Chemical Society* 2009, 131, 6944-6945.