UNIVERSITY OF CALIFORNIA, SAN DIEGO

FLUORESCENCE MICROSCOPY AND MICROFLUIDICS FOR CELLULAR AND ENZYMATIC ANALYSIS

A dissertation submitted in partial satisfaction of the requirements
for the degree Doctor of Philosophy

in

Bioengineering

by

Matthew Thomas Walsh

Committee in Charge:

Professor Xiaohua Huang, Chair
Professor Vineet Bafna
Professor Elizabeth Komives
Professor Shankar Subramaniam
Professor Kun Zhang

2015

The Dissertation of Matthew Thomas Walsh is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

_____

_____

_____

_____

_____

Chair

University of California, San Diego

2015

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I wish to express my immense gratitude towards Professor Xiaohua Huang for his unwavering support throughout my graduate studies. His optimism, extensive technical knowledge, and excitement for scientific discovery provided a fantastic environment for fostering my bioengineering research skills that will serve as a solid foundation for my future career. He has provided the freedom and facilities to explore various projects and strategies for solving scientific problems.

I would like to thank my other committee members, Professor Vineet Bafna, Professor Elizabeth Komives, Professor Shankar Subramaniam, Professor Kun Zhang, for their time and advice.

I owe a debt of gratitude to all the members of the Huang lab, who have all made contributions to my research in various ways. There are several members who have significantly aided in my projects that necessitate specific recognition. Dr. Alex Hsiao was instrumental in developing multi-layered PDMS microfluidics in our lab. His development of a microfluidic device for the automatic capturing of cells by hydrodynamic traps containing integrated processing chambers laid the foundation for the capture and enumeration of single cell mRNA. Dr. Zhixia Liu led the work involving the production of labeled DNA polymerases. Her efforts in the preparation of mRNA and polymerase samples, and assistance in performing single-molecule imaging experiments of DNA synthesis and image analysis were critical for this data collection. Dr. Eric Roller was responsible for developing the microscopy instrumentation and software for automated control, which were essential for the implementation of almost all of my experiments. Without his skillful assembly of this system and generous donation of time

to train me to operate it, these projects would not have been possible. Dr. Hosuk Lee provided support in conducting the single cell microfluidic experiments and also as responsible for maintaining the cell cultures and preparing them for capture in the device. Geuntak Lee and Kwang-Seuk Ko provided great collaboration and advice for surface chemistry and bioconjugation protocol development.

Additionally I wish to thank my family and friends for their encouragement throughout the degree program. Their support was essential in carrying me through the more difficult times of this process.

Charter 2, in part, is a reproduction of the material as it appears in: Walsh, M.T., Roller, E.E., Ko, K.S., Huang, X. (2015) Measurement of DNA polymerase incorporation kinetics of dye-labeled nucleotides using total internal reflection fluorescence microscopy. *Biochemistry*, **54**, 4019-4021; copyright 2015 American Chemical Society. Used with permission. The dissertation author was the primary investigator and author of this paper.

Chapter 3, in part, is a reproduction of the material as it appears in: Walsh, M.T., Hsiao, A.P., Lee, H.S., Liu, L., Huang, X. (2015) Capture and enumeration of mRNA transcripts from single cells using a microfluidic device. *Lab Chip*, **15**, 2968-2980; copyright 2015 Royal Society of Chemistry. Used with permission. The dissertation author was the primary investigator and author of this paper.

CURRICULUM VITAE

EDUCATION

University of California, San Diego – La Jolla, CA

Ph.D., Bioengineering (2015)
Dissertation Title : Fluorescence Microscopy and Microfluidics for Cellular and Enzymatic Analysis

M.S., Bioengineering (2011)
Thesis Title : A Microfluidic Device for DNA Sequencing by Denaturation

B.S., Bioengineering (2009)
Cum Laude

EXPERIENCE

University of California, San Diego – La Jolla, CA
Department of Bioengineering

- Genomics and Systems Biotechnology Lab
  Graduate Student Researcher (2009-2015)
  Research Advisor: Dr. Xiaohua Huang
  Designed and fabricated fluidic devices and automated instrumentation systems to perform multi-step biochemical reactions. Developed surface functionalization chemistries for the capture of nucleic acids followed by enzymatic processing and detection by high-sensitivity fluorescence microscopy. Applied these techniques to the capturing and enumeration of polyadenylated mRNA from single cells. Also studied DNA polymerase behavior during DNA synthesis at single molecule and ensemble levels through a variety of fluorescence methods, including single-molecule Förster resonance energy transfer.

- Microcirculation Lab
  Undergraduate Student Researcher (2008-2009)
  Research Advisor: Dr. Geert Schmid-Schönbein
  Tested for digestive enzyme activity in human plasma samples using chromogenic assays. Received IRB approval to obtain pancreatic cancer patient plasma samples to assay for trypsin, chymotrypsin, and elastase.

DexCom – San Diego, CA

Software Testing Group
Summer Intern (2008)
Testing software on the receivers of continuous glucose monitors and other products to ensure proper function. Completed documentation to comply with FDA requirements.

Quality Control
Summer Intern (2006)
Worked in cleanroom environment verifying the calibration of short-term, implantable glucose sensors prior to distribution for use.

Scripps Institute of Oceanography – La Jolla, CA
Geosciences Research Division

Norris Lab
Research Technician (2007-2008)
Assisted in the processing of deep-sea sediment core samples from the Ocean Drilling Program used to study paleobiology and oceanographic trends.

HONORS AND AWARDS

- Member, Tau Beta Pi National Engineering Honor Society
- BS/MS accelerated program
- Provost's Honors – 10 quarters
- Eagle Scout, Boy Scouts of America

TEACHING EXPERIENCE

Teaching Assistant
Responsible for leading discussion sections, holding office hours, and grading assignments and exams for both undergraduate and graduate courses.

- BENG 161A – Bioreactor Engineering (Fall 2010)
- BENG 227 – Transport Phenomena in Living Systems (Winter 2011)
- BENG 103B – Bioengineering Mass Transfer (Spring 2011)
- MDE 230 – Life Sciences and Technologies (Spring 2013)
- MDE 230 – Life Sciences and Technologies (Spring 2014)
- MDE 230 – Life Sciences and Technologies (Spring 2015)

PUBLICATIONS

Walsh, M.T., Hsaio A.P., Lee H.S., Liu Z., Huang X. (2015) Capture and Enumeration of mRNA Transcripts from Single Cells Using a Microfluidic Device. *Lab Chip*, **15**, 2968-2980.

Walsh, M.T., Roller E.E., Ko K.S., Huang X. (2015) Measurement of DNA Polymerase Incorporation Kinetics of Dye-Labeled Nucleotides Using Total Internal Reflection Fluorescence Microscopy. *Biochemistry*, **54**, 4019-4021.

PROFESSIONAL SKILLS

Fluorescence Microscopy: Single-molecule and ensemble imaging using total internal reflected fluorescence microscopy. Dual color FRET imaging.

Computational: Image processing, numerical modeling, and data analysis in MATLAB. Instrumentation automation and control using C++ in Microsoft Visual Studios.

Surface Chemistry: Glass functionalization by silianization and conjugation of polymers and biomolecules.

Fluidic Systems Design and Fabrication: Construction of fluidic devices and integration with syringe pumps, rotary valves, and pressure sources to control multistep biochemical reaction processes.

Microfabrication: Contact photolithography, metal deposition, dry etching

Molecular biology: DNA enzymology, purification, and characterization by gel electrophoresis

ABSTRACT OF THE DISSERTATION

FLUORESCENCE MICROSCOPY AND MICROFLUIDICS FOR CELLULAR AND ENZYMATIC ANALYSIS

by

Matthew Thomas Walsh

Doctor of Philosophy in Bioengineering

University of California, San Diego, 2015

Professor Xiaohua Huang, Chair

Breakthroughs in our understanding of biological systems are driven by technological advancements that enable the study of cellular and molecular processes at unprecedented levels. Recent efforts have provided a thrust in the development of

microfluidic devices that facilitate the process of genomic or transcriptomic library construction from single cells. The integration of molecular detection modalities within microfluidic devices will enable the direct analysis of cellular content through surface capture and probing of biomolecules. Similar techniques can be applied to monitoring the enzymatic reaction processes in a rapid and high-throughput manner. We have demonstrated the application of fluorescence microscopy and microfluidics to perform the automatic hydrodynamic capture of single mammalian cells followed by the immobilization and digital counting of polyadenylated mRNA molecules released from the individual cells. Fluorescence imaging was also utilized to monitor the incorporation kinetics of nucleotides by DNA polymerase within a microfluidic device using multiple dye-labeling strategies. The combination of advanced microfluidic devices and sensitive fluorescence detection techniques provides a powerful tool set for enabling sophisticated cell and molecular biology experimentation.

1 INTRODUCTION

There is a constant drive to improve our deterministic understanding of biological systems for better medical diagnoses and treatments. The greatest steps forward are made through methodological developments to enable molecular and cellular systems to be interrogated in novel ways that provide greater experimental throughput, higher detection sensitivity, and better accuracy. In this dissertation, I aim to demonstrate the development and application of state-of-the-art microfluidic processing and fluorescence imaging techniques, the automation and analytical methods necessary for precise and quantitative measurements of enzyme reaction kinetics, and digital counting of RNA molecules of single cells. These methods can be applied to a variety of cellular and molecular experimentation.

1.1 On-Chip Transcriptome Analysis

A central focus of biology is to quantitatively understand how an organism's genotype determines its phenotype. The phenotypes of the different cell types in a multi-cellular organism are defined by the selective expression of the genome to produce a unique set of RNA molecules (the transcriptome) and protein molecules (the proteome) that carries out the characteristic physiological functions of the cells. Due to the technical difficultly in whole proteome analysis, the transcriptomes are often used as surrogates to assess the physiological states of cells. Gene expression varies greatly amongst individual

cells, even within a population of the same cell type; the subtle differences may contribute significantly to both normal cellular physiology, such as stem cell differentiation, and abnormal cellular physiology such as cancer.[1-5] The ability to precisely and quantitatively measure whole transcriptomes at the single-cell level is essential to measure cellular heterogeneity and to define cell types.

Methods have been developed for direct detection and counting of DNA and RNA molecules in single cells, usually by *in situ* hybridization and single-molecule fluorescence imaging.[6-11] However, only a limited number of RNA species can be analyzed using these methods. Comprehensive transcriptome profiling of single cells is now feasible using methods based on high throughput sequencing such as RNA-seq[12-15] or direct single-molecule RNA sequencing.[16-18] But a large amount of input materials, equivalent to tens to thousands of cells is required, necessitating the amplification of the transcripts for single-cell analysis. Amplification is known to introduce bias into the relative abundance of the transcripts, in particular those with lower copy numbers, compromising the accuracy of the measurements.[15,19] The bias can be significantly mitigated by integrating the processes for cell isolation, RNA extraction, cDNA synthesis and amplification using microfluidic devices.[15] The small processing volumes in microfluidic devices help minimize sample loss and improve the efficiency of reactions by maintaining high concentrations of biomolecules.[15,20-22] These devices can be used to perform RT-qPCR quantification of RNA transcripts from single cells,though only on a limited number of RNA species.[23-26] Recently, the utilization of microemulsions have proven effective at preparing single-cell transcriptome libraries in a very high-throughput manner.[27,28] Other devices have also been developed to perform such tasks as cell

monitoring[29] or chip-based analysis.[30,31] The integration of polydeoxyribothymidine oligonucleotides immobilized on the surface of beads or glass has been demonstrated to serve as a useful mechanism for purifying and processing polyadenylated mRNA, yet has only provided a method for transcriptome library construction for subsequent analysis off-chip.[32,33] Despite the refinement and commercialization of sophisticated microfluidic devices, the current generation of microfluidic devices still lacks the capability for comprehensive direct on-chip detection and digital counting of RNA transcripts from single cells. Integrating on-chip analysis within a microfluidic device would provide the capability of performing the parallelized processing and analysis of single-cell transcriptomes in an expedited and economical manner, as compared to the current strategy of preparing samples on-chip and exporting the biomolecules off-chip for further processing and analysis.

As a step towards this end, we have developed a microfluidic device and methods for automatic hydrodynamic capture of single mammalian cells and subsequent immobilization and digital counting of polyadenylated mRNA molecules released from the individual cells. Using single molecule fluorescence imaging, we have demonstrated that polyadenylated mRNA molecules from single HeLa cells can be captured within minutes by hybridization to polydeoxyribothymidine oligonucleotides covalently attached on the glass surface in the device. The total mRNA molecule counts in the individual HeLa cells are found to vary significantly from one another. Our technology opens up the possibility of direct digital enumeration of RNA transcripts from single cells with single-molecule sensitivity using a single integrated microfluidic device.

1.2 DNA Polymerase Incorporation Kinetic Studies

DNA polymerases exhibit remarkably high fidelity in DNA replication. This great specificity for recognizing and catalyzing the incorporation of the correct incoming nucleotide complementary to the base on the template is due to the exquisite control over each stage in the multi-step chemo-mechanical process by the polymerases.[34-37] The stringent discrimination by the polymerases extends to the bias against fluorescently-labeled nucleotides,[38] which have been essential in molecular biology techniques and genomic research such as microarray assays and next-generation DNA sequencing. Hybridization-based target detection using DNA microarrays requires the labeling of nucleic acids with reactive groups or fluorescence dyes during the transcription or DNA replication process.[39] The ratio of the labeled nucleotides to native nucleotides must be carefully tuned to obtain the proper degree of labeling to produce full length products with minimal dye quenching and hindrance to hybridization.[40] Many of the current generation of high throughput sequencing technologies rely on the use of nucleotides labeled with fluorophores and the incorporation of the labeled nucleotides by engineered DNA polymerases.[41-44] Quantitative characterization of incorporation kinetics by DNA polymerase, however, is not trivial, requiring substantial effort and specialized instruments. Most traditional methods for measuring DNA polymerase kinetics rely on chemical-quench-flow methods followed by analysis using electrophoretic separation or binding and quantification of radioactively labeled products. More recently reported methods aim to simplify the analysis by combining stopped-flow with fluorescence measurement or using DNA polymerases or DNA substrates that are labeled with

fluorescent probes that are sensitive to the local protein environment or conformations.[45-48] We have developed a method for the rapid and automated measurements of incorporation kinetics of labeled nucleotides by DNA polymerases using an integrated system with microfluidics and real-time ensemble fluorescence imaging on a surface with a high density of DNA templates.

Total internal reflection fluorescence microscopy (TIRF) is used to study DNA polymerase incorporation behavior at the ensemble and single molecule level. To monitor the incorporation of nucleotides by DNA polymerase into surface-bound primed DNA templates, three fluorescence imaging strategies were employed. In the first approach, dye-labeled nucleotides are used for ensemble imaging. In the second approach, single-molecule Förster resonance energy transfer (smFRET) imaging is used to monitor the reaction of polymerases labeled with a FRET dye pairs. In the third approach, smFRET between a donor fluorophore on the polymerase and an acceptor fluorophore on the nucleotide is used to monitor the reaction.

Using the nucleotide label and ensemble imaging method, rapid and automated measurements of the incorporation kinetics of fluorescent dye-labeled nucleotides by DNA polymerases were performed without using stopped-flow and quench-flow methods. A microfluidic system is used to perform the incorporation of fluorescently-labeled nucleotides by DNA polymerase into surface-bound primed DNA templates and TIRF microscopy is used to monitor the reaction. We successfully demonstrated the method using Bst DNA polymerase and a set of coumarin-labeled nucleotides. Our method enables the rapid acquisition of polymerase kinetics for implementing and

improving DNA sequencing technologies that rely on labeled nucleotides and DNA polymerases.

By chemically attaching two appropriate fluorescent labels on Bst DNA polymerase, smFRET signals are measured during DNA synthesis to study the enzyme's conformational dynamics. The strategy of using smFRET between the enzyme and substrate enabled the monitoring of the interaction between the nucleotides and polymerase. The microfluidic system used to perform these reactions provided the fluidic and thermal control necessary to carry out these biochemical reactions in a manner that is compatible with high-sensitivity fluorescence imaging. This work was motivated by our effort in the development of a technology for single-molecule real-time DNA sequencing (READS).

1.3 References

1. Cai, L., Friedman, N. and Xie, X.S. (2006) Stochastic protein expression in individual cells at the single molecule level. *Nature*, **440**, 358-362.

2. Spiller, D.G., Wood, C.D., Rand, D.A. and White, M.R. (2010) Measurement of single-cell dynamics. *Nature*, **465**, 736-745.

3. Kalisky, T. and Quake, S.R. (2011) Single-cell genomics. *Nat Methods*, **8**, 311-314.

4. Bendall, S.C. and Nolan, G.P. (2012) From single cells to deep phenotypes in cancer. *Nat Biotechnol*, **30**, 639-647.

5. Coulon, A., Chow, C.C., Singer, R.H. and Larson, D.R. (2013) Eukaryotic transcriptional dynamics: from single molecules to cell populations. *Nat Rev Genet*, **14**, 572-584.

6. Hocine, S., Raymond, P., Zenklusen, D., Chao, J.A. and Singer, R.H. (2013) Single-molecule analysis of gene expression using two-color RNA labeling in live yeast. *Nat Methods*, **10**, 119-121.

7. Larsson, C., Grundberg, I., Soderberg, O. and Nilsson, M. (2010) In situ detection and genotyping of individual mRNA molecules. *Nat Methods*, **7**, 395-397.

8. Levsky, J.M., Shenoy, S.M., Pezo, R.C. and Singer, R.H. (2002) Single-cell gene expression profiling. *Science*, **297**, 836-840.

9. Lizardi, P.M., Huang, X., Zhu, Z., Bray-Ward, P., Thomas, D.C. and Ward, D.C. (1998) Mutation detection and single-molecule counting using isothermal rolling-circle amplification. *Nat Genet*, **19**, 225-232.

10. Zhong, X.B., Lizardi, P.M., Huang, X.H., Bray-Ward, P.L. and Ward, D.C. (2001) Visualization of oligonucleotide probes and point mutations in interphase nuclei and DNA fibers using rolling circle DNA amplification. *Proc Natl Acad Sci U S A*, **98**, 3940-3945.

11. Zhou, Y., Calciano, M., Hamann, S., Leamon, J.H., Strugnell, T., Christian, M.W. and Lizardi, P.M. (2001) In situ detection of messenger RNA using digoxigenin-labeled oligonucleotides and rolling circle amplification. *Exp Mol Pathol*, **70**, 281-288.

12. Deng, Q., Ramskold, D., Reinius, B. and Sandberg, R. (2014) Single-cell RNA-seq reveals dynamic, random monoallelic gene expression in mammalian cells. *Science*, **343**, 193-196.

13. Tang, F., Barbacioru, C., Wang, Y., Nordman, E., Lee, C., Xu, N., Wang, X., Bodeau, J., Tuch, B.B., Siddiqui, A., Lao, K. and Surani, M.A. (2009) mRNA-Seq whole-transcriptome analysis of a single cell. *Nat Methods*, **6**, 377-382.

14. Tang, F., Lao, K. and surani, M.A. (2011) Development and application of single-cell transcriptome analysis. *Nat Methods*, S6-S11.

15. Wu, A.R., Neff, N.F., Kalisky, T., Dalerba, P., Treutlein, B., Rothenberg, M.E., Mburu, F.M., Mantalas, G.L., Sim, S., Clarke, M.F. and Quake, S.R. (2014) Quantitative assessment of single-cell RNA-sequencing methods. *Nat Methods*, **11**, 41-46.

16. Ozsolak, F. and Milos, P.M. (2011) Transcriptome profiling using single-molecule direct RNA sequencing. *Methods Mol Biol*, **733**, 51-61.

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

18. Vilfan, I.D., Tsai, Y.C., Clark, T.A., Wegener, J., Dai, Q., Yi, C., Pan, T., Turner, S.W. and Korlach, J. (2013) Analysis of RNA base modification and structural rearrangement by single-molecule real-time detection of reverse transcription. *J Nanobiotechnology*, **11**, 8.

19. Bhargava, V., Head, S.R., Ordoukhanian, P., Mercola, M. and Subramaniam, S. (2014) Technical Variations in Low-Input RNA-seq Methodologies. *Sci Rep*, **4**, 3678.

20. Islam, S., Zeisel, A., Joost, S., La Manno, G., Zajac, P., Kasper, M., Lonnerberg, P. and Linnarsson, S. (2014) Quantitative single-cell RNA-seq with unique molecular identifiers. *Nat Methods*, **11**, 163-166.

21. Marcy, Y., Ishoey, T., Lasken, R.S., Stockwell, T.B., Walenz, B.P., Halpern, A.L., Beeson, K.Y., Goldberg, S.M. and Quake, S.R. (2007) Nanoliter reactors improve multiple displacement amplification of genomes from single cells. *PLoS Genet*, **3**, 1702-1708.

22. Streets, A.M., Zhang, X., Cao, C., Pang, Y., Wu, X., Xiong, L., Yang, L., Fu, Y., Zhao, L., Tang, F. and Huang, Y. (2014) Microfluidic single-cell whole-transcriptome sequencing. *Proc Natl Acad Sci U S A*, **111**, 7048-7053.

23. Dominguez, M.H., Chattopadhyay, P.K., Ma, S., Lamoreaux, L., McDavid, A., Finak, G., Gottardo, R., Koup, R.A. and Roederer, M. (2013) Highly multiplexed quantitation of gene expression on single cells. *Journal of Immunological Methods*, **391**, 133-145.

24. Powell, A.A., Talasaz, A.H., Zhang, H., Coram, M.A., Reddy, A., Deng, G., Telli, M.L., Advani, R.H., Carlson, R.W., Mollick, J.A., Sheth, S., Kurian, A.W., Ford, J.M., Stockdale, F.E., Quake, S.R., Pease, R.F., Mindrinos, M.N., Bhanot, G., Dairkee, S.H., Davis, R.W. and Jeffrey, S.S. (2012) Single cell profiling of circulating tumor cells: transcriptional heterogeneity and diversity from breast cancer cell lines. *PLoS One*, **7**, e33788.

25. Sanchez-Freire, V., Ebert, A.D., Kalisky, T., Quake, S.R. and Wu, J.C. (2012) Microfluidic single-cell real-time PCR for comparative analysis of gene expression patterns. *Nat Protoc*, **7**, 829-838.

26. White, A.K., VanInsberghe, M., Petriv, O.I., Hamidi, M., Sikorski, D., Marra, M.A., Piret, J., Aparicio, S. and Hansen, C.L. (2011) High-throughput microfluidic single-cell RT-qPCR. *P Natl Acad Sci USA*, **108**, 13999-14004.

27. Klein, A.M., Mazutis, L., Akartuna, I., Tallapragada, N., Veres, A., Li, V., Peshkin, L., Weitz, D.A. and Kirschner, M.W. (2015) Droplet barcoding for single-cell transcriptomics applied to embryonic stem cells. *Cell*, **161**, 1187-1201.

28. Macosko, E.Z., Basu, A., Satija, R., Nemesh, J., Shekhar, K., Goldman, M., Tirosh, I., Bialas, A.R., Kamitaki, N., Martersteck, E.M., Trombetta, J.J., Weitz, D.A., Sanes, J.R., Shalek, A.K., Regev, A. and McCarroll, S.A. (2015) Highly Parallel Genome-wide Expression Profiling of Individual Cells Using Nanoliter Droplets. *Cell*, **161**, 1202-1214.

29. He, L., Kniss, A., San-Miguel, A., Rouse, T., Kemp, M.L. and Lu, H. (2015) An automated programmable platform enabling multiplex dynamic stimuli delivery and cellular response monitoring for high-throughput suspension single-cell signaling studies. *Lab Chip*, **15**, 1497-1507.

30. Wen, J., Yang, X., Wang, K., Tan, W., Zhou, L., Zuo, X., Zhang, H. and Chen, Y. (2007) One-dimensional microfluidic beads array for multiple mRNAs expression detection. *Biosens Bioelectron*, **22**, 2759-2762.

31. Yu, Y., Li, B., Baker, C.A., Zhang, X. and Roper, M.G. (2012) Quantitative polymerase chain reaction using infrared heating on a microfluidic chip. *Anal Chem*, **84**, 2825-2829.

32. Marcus, J.S., Anderson, W.F. and Quake, S.R. (2006) Microfluidic single-cell mRNA isolation and analysis. *Anal Chem*, **78**, 3084-3089.

33. Bose, S., Wan, Z., Carr, A., Rizvi, A.H., Vieira, G., Pe'er, D. and Sims, P.A. (2015) Scalable microfluidics for single cell RNA printing and sequencing. *Genome Biol*, **16**, 120.

34. Kunkel, T.A. (2009) Evolving views of DNA replication (in)fidelity. *Cold Spring Harb Symp Quant Biol*, **74**, 91-101.

35. Johnson, K.A. (2010) The kinetic and chemical mechanism of high-fidelity DNA polymerases. *Biochim Biophys Acta*, **1804**, 1041-1048.

36. Xia, S. and Konigsberg, W.H. (2014) RB69 DNA polymerase structure, kinetics, and fidelity. *Biochemistry*, **53**, 2752-2767.

37. Joyce, C.M. and Benkovic, S.J. (2004) DNA polymerase fidelity: kinetics, structure, and checkpoints. *Biochemistry*, **43**, 14317-14324.

38. Giller, G., Tasara, T., Angerer, B., Muhlegger, K., Amacker, M. and Winter, H. (2003) Incorporation of reporter molecule-labeled nucleotides by DNA polymerases. I. Chemical synthesis of various reporter group-labeled 2'-deoxyribonucleoside-5'-triphosphates. *Nucleic Acids Res*, **31**, 2630-2635.

39. Schena, M., Shalon, D., Davis, R.W. and Brown, P.O. (1995) Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science*, **270**, 467-470.

40. Randolph, J.B. and Waggoner, A.S. (1997) Stability, specificity and fluorescence brightness of multiply-labeled fluorescent DNA probes. *Nucleic Acids Res*, **25**, 2923-2929.

41. Ju, J., Kim, D.H., Bi, L., Meng, Q., Bai, X., Li, Z., Li, X., Marma, M.S., Shi, S., Wu, J., Edwards, J.R., Romu, A. and Turro, N.J. (2006) Four-color DNA sequencing by synthesis using cleavable fluorescent nucleotide reversible terminators. *Proc Natl Acad Sci U S A*, **103**, 19635-19640.

42. Bentley, D.R., Balasubramanian, S., Swerdlow, H.P., Smith, G.P., Milton, J., Brown, C.G., Hall, K.P., Evers, D.J., Barnes, C.L., Bignell, H.R., Boutell, J.M., Bryant, J., Carter, R.J., Keira Cheetham, R., Cox, A.J., Ellis, D.J., Flatbush, M.R., Gormley, N.A., Humphray, S.J., Irving, L.J., Karbelashvili, M.S., Kirk, S.M., Li, H., Liu, X., Maisinger, K.S., Murray, L.J., Obradovic, B., Ost, T., Parkinson, M.L., Pratt, M.R., Rasolonjatovo, I.M., Reed, M.T., Rigatti, R., Rodighiero, C., Ross, M.T., Sabot, A., Sankar, S.V., Scally, A., Schroth, G.P., Smith, M.E., Smith, V.P., Spiridou, A., Torrance, P.E., Tzonev, S.S., Vermaas, E.H., Walter, K., Wu, X., Zhang, L., Alam, M.D., Anastasi, C., Aniebo, I.C., Bailey, D.M., Bancarz, I.R., Banerjee, S., Barbour, S.G., Baybayan, P.A., Benoit, V.A., Benson, K.F., Bevis, C., Black, P.J., Boodhun, A., Brennan, J.S., Bridgham, J.A., Brown, R.C., Brown, A.A., Buermann, D.H., Bundu, A.A., Burrows, J.C., Carter, N.P., Castillo, N., Chiara, E.C.M., Chang, S., Neil Cooley, R., Crake, N.R., Dada, O.O., Diakoumakos, K.D., Dominguez-Fernandez, B., Earnshaw, D.J., Egbujor, U.C., Elmore, D.W., Etchin, S.S., Ewan, M.R., Fedurco, M., Fraser, L.J., Fuentes Fajardo, K.V., Scott Furey, W., George, D., Gietzen, K.J., Goddard, C.P., Golda, G.S., Granieri, P.A., Green, D.E., Gustafson, D.L., Hansen, N.F., Harnish, K., Haudenschild, C.D., Heyer, N.I., Hims, M.M., Ho, J.T., Horgan, A.M., Hoschler, K., Hurwitz, S., Ivanov, D.V., Johnson, M.Q., James, T., Huw Jones, T.A., Kang, G.D., Kerelska, T.H., Kersey, A.D., Khrebtukova, I., Kindwall, A.P., Kingsbury, Z., Kokko-Gonzales, P.I., Kumar, A., Laurent, M.A., Lawley, C.T., Lee, S.E., Lee, X., Liao, A.K., Loch, J.A., Lok, M., Luo, S., Mammen, R.M., Martin, J.W., McCauley, P.G., McNitt, P., Mehta, P., Moon, K.W., Mullens, J.W., Newington, T., Ning, Z., Ling Ng, B., Novo, S.M., O'Neill, M.J., Osborne, M.A., Osnowski, A., Ostadan, O., Paraschos, L.L., Pickering, L., Pike, A.C., Chris Pinkard, D., Pliskin, D.P., Podhasky, J., Quijano, V.J., Raczy, C., Rae, V.H., Rawlings, S.R., Chiva Rodriguez, A., Roe, P.M., Rogers, J., Rogert Bacigalupo, M.C., Romanov, N., Romieu, A., Roth, R.K., Rourke, N.J., Ruediger, S.T., Rusman, E., Sanches-Kuiper, R.M., Schenker, M.R., Seoane, J.M., Shaw, R.J., Shiver, M.K., Short, S.W., Sizto, N.L., Sluis, J.P., Smith, M.A., Ernest Sohna Sohna, J., Spence, E.J., Stevens, K., Sutton, N., Szajkowski, L., Tregidgo, C.L., Turcatti, G., Vandevondele, S., Verhovsky, Y., Virk, S.M., Wakelin, S., Walcott, G.C., Wang, J., Worsley, G.J., Yan, J., Yau, L., Zuerlein, M., Mullikin, J.C., Hurles, M.E., McCooke, N.J., West, J.S., Oaks, F.L., Lundberg, P.L., Klenerman, D., Durbin, R. and Smith, A.J. (2008) Accurate whole human genome sequencing using reversible terminator chemistry. *Nature*, **456**, 53-59.

43. Eid, J., Fehr, A., Gray, J., Luong, K., Lyle, J., Otto, G., Peluso, P., Rank, D., Baybayan, P., Bettman, B., Bibillo, A., Bjornson, K., Chaudhuri, B., Christians, F., Cicero, R., Clark, S., Dalal, R., Dewinter, A., Dixon, J., Foquet, M., Gaertner, A., Hardenbol, P., Heiner, C., Hester, K., Holden, D., Kearns, G., Kong, X., Kuse, R., Lacroix, Y., Lin, S., Lundquist, P., Ma, C., Marks, P., Maxham, M., Murphy, D., Park, I., Pham, T., Phillips, M., Roy, J., Sebra, R., Shen, G., Sorenson, J., Tomaney, A., Travers, K., Trulson, M., Vieceli, J., Wegener, J., Wu, D., Yang, A., Zaccarin, D., Zhao, P., Zhong, F., Korlach, J. and Turner, S. (2009) Real-time DNA sequencing from single polymerase molecules. *Science*, **323**, 133-138.

44. Fuller, C.W., Middendorf, L.R., Benner, S.A., Church, G.M., Harris, T., Huang, X., Jovanovich, S.B., Nelson, J.R., Schloss, J.A., Schwartz, D.C. and Vezenov, D.V. (2009) The challenges of sequencing by synthesis. *Nat Biotechnol*, **27**, 1013-1023.

45. Frey, M.W., Sowers, L.C., Millar, D.P. and Benkovic, S.J. (1995) The Nucleotide Analog 2-Aminopurine as a Spectroscopic Probe of Nucleotide Incorporation by the Klenow Fragment of Escherichia-Coli Polymerase-I and Bacteriophage-T4 DNA-Polymerase. *Biochemistry*, **34**, 9185-9192.

46. Gong, P., Campagnola, G. and Peersen, O.B. (2009) A quantitative stopped-flow fluorescence assay for measuring polymerase elongation rates. *Anal Biochem*, **391**, 45-55.

47. Tsai, Y.C., Jin, Z. and Johnson, K.A. (2009) Site-specific labeling of T7 DNA polymerase with a conformationally sensitive fluorophore and its use in detecting single-nucleotide polymorphisms. *Anal Biochem*, **384**, 136-144.

48. Tsai, Y.C. and Johnson, K.A. (2006) A new paradigm for DNA polymerase specificity. *Biochemistry*, **45**, 9675-9687.

2 MEASUREMENT OF DNA POLYMERASE INCORPORATION KINETICS OF DYE-LABLED NUCLEOTIDES

2.1 Abstract

A method is reported for the rapid and automated measurements of the incorporation kinetics of fluorescent dye-labeled nucleotides by DNA polymerases that avoids using stopped-flow and quench-flow methods. Total internal reflection fluorescence microscopy (TIRF) is used to monitor the incorporation of fluorescently labeled nucleotides by DNA polymerase into surface-bound primed DNA templates and a microfluidic system is used to perform the reactions. We successfully demonstrated the method using Bst DNA polymerase and a set of coumarin-labeled nucleotides. Our method enables the rapid acquisition of polymerase kinetics for implementing and improving DNA sequencing technologies that rely on labeled nucleotides and DNA polymerases.

2.2 Introduction

The remarkable fidelity of DNA polymerases are determined by the binding and incorporation kinetics of the nucleotide substrates.[1-3] The pre-steady-state equilibrium

dissociation and incorporation rate constants are usually determined by stopped-flow methods in combination with fluorescence measurement and quench-flow methods followed by gel electrophoresis analysis or radiometric assays.[4-8] Fluorescently labeled DNA polymerases and fluorescent nucleotide analogs can also be used to characterize conformational dynamics and reaction kinetics by Förster resonance energy transfer (FRET) imaging.[9-12] We report a method for the rapid and automated measurement of the incorporation kinetics of fluorescent dye-labeled nucleotides by DNA polymerases. We use TIRF microscopy coupled to a custom-designed flow cell with fluidic and temperature control for real-time fluorescence imaging of the enzymatic reactions on the surface. The basic setup is illustrated in Figure 2.1. The shallow penetration depth of the evanescent wave created by total internal reflection (TIR) allows for the selective excitation of only the molecules very close to the surface (~160 nm in our case),[13] enabling the real-time monitoring of the incorporation of the labeled nucleotides into primed DNA templates immobilized on the surface of a glass substrate.

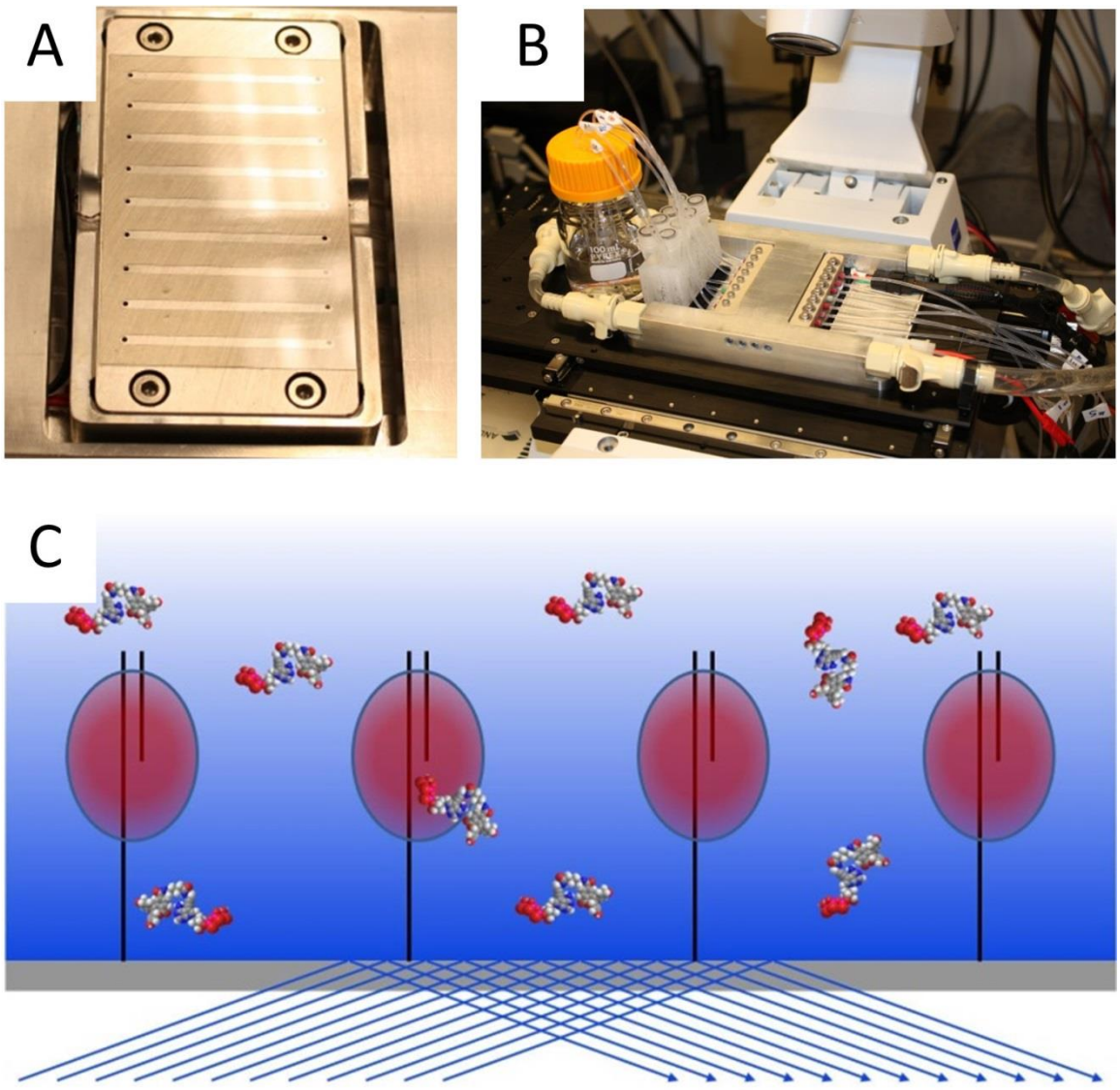


Figure 2.1. Fluidic system and TIRF microscopy for rapid measurement of incorporation kinetics of dye-labeled nucleotides by surface-bound DNA polymerases. A. Flow cell with nine fluidic channels. B. Flow cell with fluidic interface mounted to a temperature-control unit on a microscope stage. C. Schematic of nucleotide incorporation reaction on surface-bound DNA polymerases with TIR illumination.

2.3 Experimental Methods

2.3.1 Glass Coverslip Surface Functionalization

Borosilicate glass coverslips (50 mm x 75 mm x 0.170 mm, Erie Scientific) are processed[14] and derivatized with a 50-base polydeoxythymidine oligonucleotide ($poly(dT)_{50}$) through a series of chemical reactions. The cleaning, hydroxyl activation, and amino-silanization steps are performed as previously described.[14] In brief, coverslips are first sonicated in a solution of 2% Micro-90 (Cole-Parmer Instrument Co.) followed by a thorough rinse with deionized water. Organic residue is removed from the coverslips by soaking them in an RCA-1 cleaning solution (1:1:5 NH_4OH:H_2O_2:H_2O) for 60 min at 85 ºC, followed by rinsing three times with deionized water. Hydroxyl groups are next generated on the surface to prepared for silanization by treatment with a Piranha solution (3:1 H_2SO_4:H_2O_2) for 60 min at 85 ºC. After extensive rinsing with deionized water and a brief soak in ethanol, silanization is performed by placing the coverslips in a 2% solution of 3-aminopropyltriethoxysilane (Gelest) in 95:5 ethanol/water for 15 min at room temperature. The coverslips are then rinsed three times with acetone and cured at 110 ºC for 15 min in a convection oven. The amine groups on the surfaces are then conjugated to homobifunctional carboxymethyl poly(ethylene glycol) (1000 g/mol, Laysan Bio, Inc.) at 1 mM, catalyzed by 10 mM of 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide (EDC) and 5 mM of triethylamine (TEA) in dry N,N-dimethylformamide (DMF) for 2 hours by sandwiching the solution between two coverslips. Following this reaction, the coverslips are rinsed with DMF and dried by blowing with filtered argon before reacting with N-hydroxysuccinimide (NHS) activated methyl-$(PEG)_8$ (Thermo Scientific) at 25 mM with

5 mM TEA in dry DMF for 1 hour using the sandwich method. This reagent is intended to react with any remaining amine groups on the surface to help further reduce non-specific binding. The coverslips are next rinsed with methanol and blown dry with filtered argon. A 50-base poly-deoxythymidine (poly($dT)_{50}$) is then immobilized on the surface using 5 μM 5'-amine-labeled poly($dT)_{50}$ and 10 mM EDC in 100 mM 2-(N-morpholino)-ethanesulfonic acid (MES) buffer, pH 5 with 0.02% Triton X-100 for 2 hours using the sandwich method. After the reaction, coverslips are washed with Tris-EDTA buffer, pH 8 with 0.02% Triton X-100 and stored at 4 °C in this buffer until use.

2.3.2 Fluidic Device

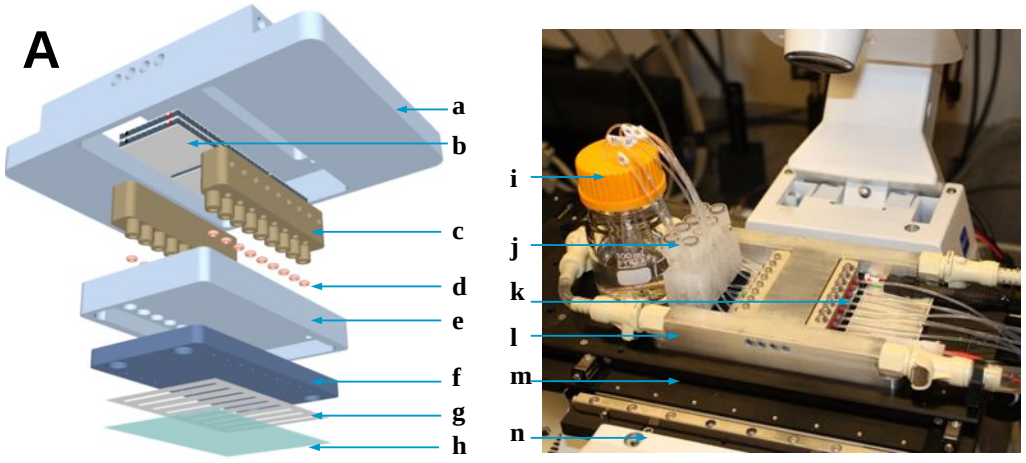
The coverslips are assembled onto a flow cell consisting of nine 40 mm x 2 mm x 100 μm channels cut from double-side adhesive silicone tape (~100 μm thick, Scapa 702) and attached to an anodized aluminum or stainless steel plate with 1 mm holes drilled for fluidic connections. The modular flow cell is mounted into an aluminum encasement block, as shown in Figure 2.1a. The encasement block serves as a temperature control unit comprised of thermoelectric modules sandwiched between the block and a liquid-cooled aluminum heatsink, which is designed to mount on a motorized XY microscope stage. A thermistor is fixed to the encasement block to provide feedback to a temperature controller. Thermoelectric modules (VT-71-1.4-1.15), thermistor and temperature controller (TC-36-25-RS232) from TE Technology were used for temperature control. Fluid control is managed by a syringe pump connected to a 9-port rotary valve. The water-cooled refrigerated circulator for cooling of the heatsink is model F25-HE from Julabo. The device is shown in Figure 2.1b mounted on the stage with fluidic connections

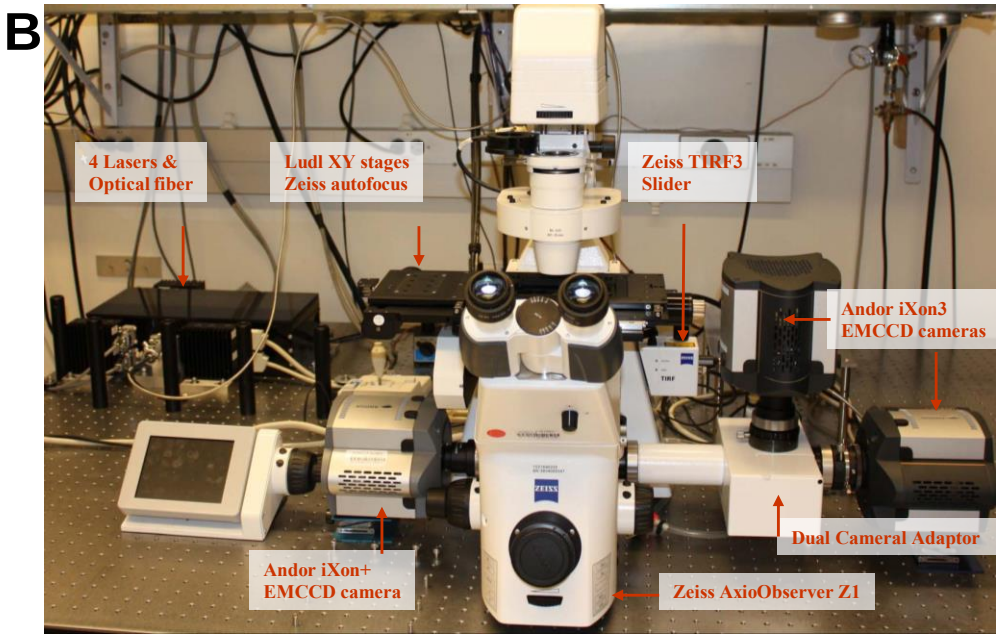
made. All components are controlled through a computer using custom software written in C++ to automate the entire process, including temperature control, delivery of solutions for multistep reactions and real-time fluorescence imaging.

2.3.3 TIRF Microscopy System

The objective-based TIRF system consisted of an inverted fluorescence microscope and a TIRF slider (AxioObserver Z1 and TIRF 3 Slider, Carl Zeiss). A 100X oil objective lens with a numerical aperture of 1.46 (Alpha Plan Apochromat, Carl Zeiss) was used for both laser excitation and fluorescence detection. Laser diodes for 405 nm, 488 nm, and 642 nm, along with a diode-pumped solid state 532 nm laser, are coupled to the slider via a polarization preserving single-mode broadband optical fiber are used for excitation. and the illumination angle is adjusted slightly beyond the TIR angle using a piezo-driven mechanism. The fluorescence from the glass surface was collected through a dichroic beamsplitter and a bandpass filter optimized the fluorophores being imaged. The images are captured on one of three cooled EMCCD cameras (1002 x 1004 pixels; 8μm x 8μm per pixel, iXon+ 885 or 512 x 512 pixels, 16 μm x16 μm, iXon3 897. Andor Technology). There are two iXon3 897 cameras have extremely high quantum efficiency for single-molecule imaging and are attached to a dual-camera adapter used to split fluorescence wavelengths for FRET imaging. A schematic of the entire system and an exploded view of the device is shown in Figure 2.2.

Figure 2.2. Fluidic and TIRF system for kinetics measurement of enzymatic reaction on the surface-bound molecules. A. Shown on the left is an exploded view of the CAD model of the flowcell mounted into the temperature control unit. The real physical assembly mounted on the microscope is shown on the right. a: heatsink; b: thermoelectric models; c: PEEK adaptor for fluidic interface; d: silicone rubber O-rings for sealing; e: aluminum encasement block; f: black hard anodized aluminum or stainless steel block as the upper plate for the flow channels; g: double-side adhesive silicone gasket with cutouts for 9 fluidic channels; h: functionalized glass coverslip; i: waste bottle; j: reagent tubes; k: Teflon tubings interfacing the flow channels and the syringe pump; l: liquid-cooled heatsink that also serves as an insert for precise mounting of the flow cell onto the microscope; m: XY motorized microscope stages (Ludl Electronics Product); n: inverted microscope. B. The objective-based TIRF setup. The system is based on a fully motorized inverted fluorescence microscope (Zeiss AxioObserver Z1). An adaptor (Zeiss RITF3 Slider) is used to introduce the laser excitation into the objective (100x oil, NA 1.46, Alpha Plan Apochromat, Zeiss). The lasers are combined and coupled into the TIRF slider by a polarization preserving single-mode broadband optical fiber (kineFLEX, Point Source).


A
a
b
c
d
e
f
g
h
i
j
k
l
m
n


B
4 Lasers &
Optical fiber
Ludl XY stages
Zeiss autofocus
Zeiss TIRF3
Slider
Andor iXon3
EMCCD cameras
Dual Cameral Adaptor
Andor iXon+
EMCCD camera
Zeiss AxioObserver Z1

2.3.4 Incorporation Reaction Kinetic Measurements

To prepare the fluidic channels for the reactions, a solution of 0.5 μM DNA template with a poly(dA) tail and 1 μM of a specific primer designed to incorporate a particular base in 2x saline sodium citrate buffer with 0.02% Triton X-100 (2x SSCt) was flowed into each separate channel. Oligo sequences are given in Table 2.1. DNA templates were simultaneously hybridized primers and poly$(dT)_{50}$ on the surface by first raising the temperature to 65°C and holding at the temperature for 5 minutes, then decreasing the temperature to 25°C at a rate of 0.1°C/second. The channels were then washed with 2x SSCt, followed by 1x Bst buffer (20 mM Tris, 50 mM KCl, 10 mM $(NH_4)_2SO_4$, 4 mM $MgSO_4$, 0.1% Triton-X 100, pH8.8) to prepare for the polymerase reaction. To ensure that most primers were bound with a polymerase prior to the incorporation reaction, Bst DNA polymerase large fragment at 1 U/μl (0.13 μM) in 1x Bst buffer solution containing 0.1 mg/ml bovine serum albumin (BSA) and 0.019 mg/ml E. coli single-stranded binding protein (SSB) was flowed into and incubated in the channels for at least 5 minutes.

Table 2.1. DNA template and primer sequences. The four bases serving as template for incorporation are underlined. Primer A contains an Alexa546 label for calibration of surface quality before performing experiments.

Oligo Name	**Sequence**
dATP Incorpration Template/Primer	5′-ATGTCCTATCCTCAGCTGTGAGCA**T**GACTGACTTCGTACATGACTGATGGTCGATAAAAAAAAAAAAAAAAAAAAAAAA-3′ 3′-CTGACTGAAGCATGTACTGACTACC-Alexa546-5′
dCTP Incorporation Template/Primer	5′-ATGTCCTATCCTCAGCTGTGAGCAT**G**ACTGACTTCGTACATGACTGATGGTCGATAAAAAAAAAAAAAAAAAAAAAAAA-3′ 3′-TGACTGAAGCATGTACTGACTACC-5′
dGTP Incorporation Template/Primer	5′-ATGTCCTATCCTCAGCTGTGAG**C**ATGACTGACTTCGTACATGACTGATGGTCGATAAAAAAAAAAAAAAAAAAAAAAAA-3′ 3′-TACTGACTGAAGCATGTACTGACTACC-5′
dUTP Incorporation Template/Primer	5′-ATGTCCTATCCTCAGCTGTGAGC**A**TGACTGACTTCGTACATGACTGATGGTCGATAAAAAAAAAAAAAAAAAAAAAAAA-3′ 3′-ACTGACTGAAGCATGTACTGACTACC-5′

The reactions were carried out with a series of nucleotide concentrations at 25°C. Each solution contained 0.2 U/µl (0.025 µM) Bst polymerase, 8.5 µg/ml SSB, 0.1 mg/ml BSA and the desired concentration of the custom-synthesized nucleotide. The nucleotides were labeled with a coumarin dye attached via an o-allyl linker, as shown in Figure 2.3. The solution was injected into the channel at a linear flow rate of 100 mm/s, and time-lapse TIRF images were immediately acquired. Excitation was provided by a 405 nm laser at 19.5 mW after the objective. The fluorescence from the glass surface was collected through a dichroic beamsplitter and a bandpass filter optimized for coumarin (FF01-485/70-25 and Di01-R405, Semrock). The images were captured on a cooled EMCCD camera (1002 x 1004 pixels; 8µm x 8µm per pixel. iXon+ 885, Andor Technology) using 200 ms exposure and 30 linear EM gain. To avoid photobleaching of the coumarin dye, each image was taken at a different field of view over the course of 2 to 250 s. Proper focusing was maintained using an autofocus system (Definite Focus, Carl Zeiss).

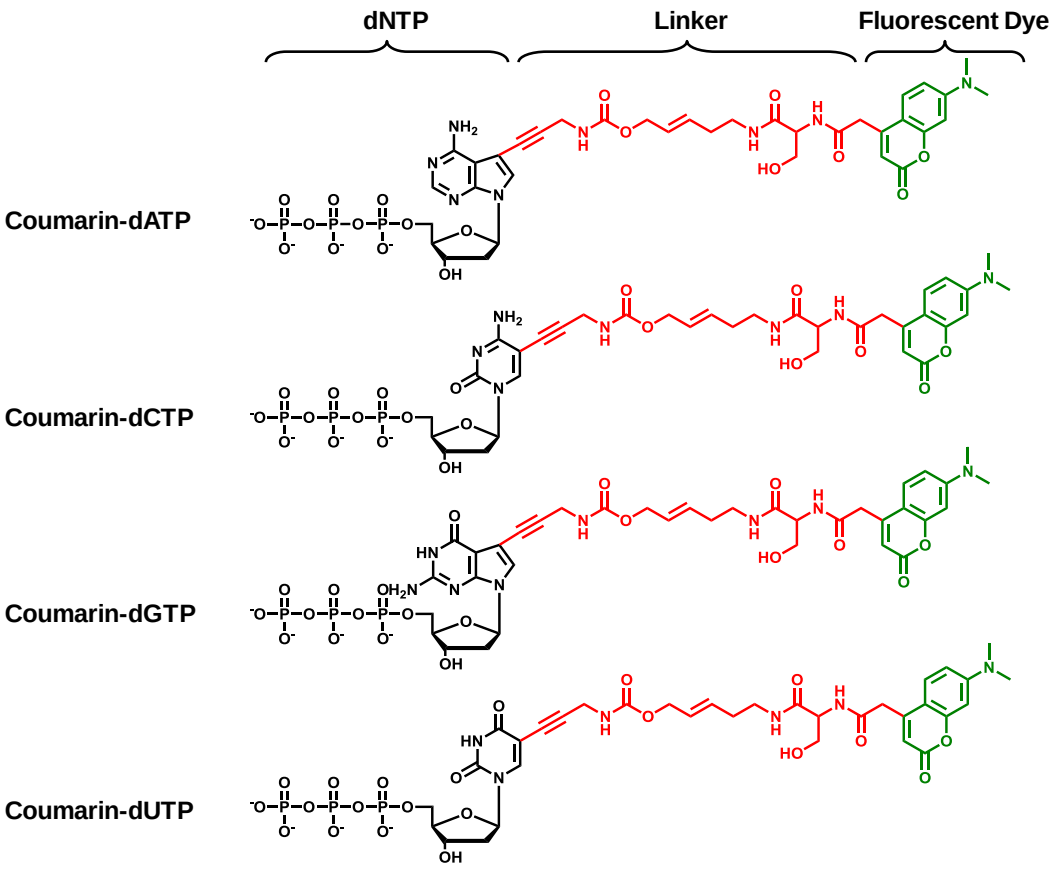


Figure 2.3. Custom-synthesized coumarin-labeled deoxyribonucleoside triphosphates. The coumarin dye is shown in green. A flexible and chemically cleavable linker is shown in red. The o-allyl linker can be chemically cleaved using Pt-based catalyst.[15] The synthesis and characterization of the coumarin-labeled deoxyribonucleotide triphosphates is described in a manuscript that has been submitted for publication (Kwang-Seuk Ko, Eric E. Roller, Matthew T. Walsh and Xiaohua Huang, "Solid-phase Synthesis of Cleavable Fluorescently Labeled Nucleotides for DNA Sequencing"). Incoporation reactions were performed in solution and the products analayzed by gel electrophoresis to show quantitative single-base extension of the primers.

2.4 Results

2.4.1 Modeling of Incorporation Kinetics

To extract the kinetic parameters, we use a Michaelis-Menten model for labeled nucleotide (represented by dNTP for simplicity) incorporation by a DNA polymerase (Pol):

$$Pol{\cdot}DNA_n + dNTP \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} Pol{\cdot}DNA_n{\cdot}dNTP \xrightarrow{k_2} Pol{\cdot}DNA_{n+1} + PP_i$$

where $Pol{\cdot}DNA_n$ is the polymerase-primer/template complex; $Pol{\cdot}DNA{\cdot}dNTP$ is complex of Pol·DNA and dNTP; and $Pol{\cdot}DNA_{n+1}$ is one-base extension product. The binding and dissociation of nucleotides are known to be much faster than the chemical reaction step, and are assumed to reach rapid equilibrium,[3] with a dissociation constant (K_d) of $K_d=([Pol\cdot DNA_n][dNTP])/[Pol\cdot DNA_n\cdot dNTP]$

The rate of product formation is then given by

$$d[Pol\cdot DNA_{n+1}]/dt=(k_2/K_d)[Pol\cdot DNA_n][dNTP] \quad (2.1)$$

Since the total polymerase species is conserved, we have

$$[Pol\cdot DNA]_{total}=[Pol\cdot DNA_n]+[Pol\cdot DNA_n\cdot dNTP]+[Pol\cdot DNA_{n+1}] \quad (2.2)$$

Using eq 2.2 and assuming large excess of nucleotide is used ([dNTP] unchanged), we can solve eq 2.1 to obtain the concentration of extension product as a function of time:

$$[Pol\cdot DNA_{n+1}]=[Pol\cdot DNA]_{total}(1-e^{-(k_2[dNTP]/(K_d+[dNTP]))t}) \quad (2.3)$$

[Pol·DNA_{n+1}] is linearly proportional to the fluorescence intensity. Intensity values at the different time points were calculated by integrating the full field of view of the time-lapse images using MATLAB (Mathworks).

2.4.2 Extracting Polymerase Kinetics Parameters from Incorporation Rates

The values were plotted as a function of time and fitted to eq 2.3 using a nonlinear least squares regression method to extract the observed reaction rate constant (k_{obs}). The data and fits are shown in Figure 2.4. The curves are smooth and the fits are excellent with adjusted R^2 values above 0.98.

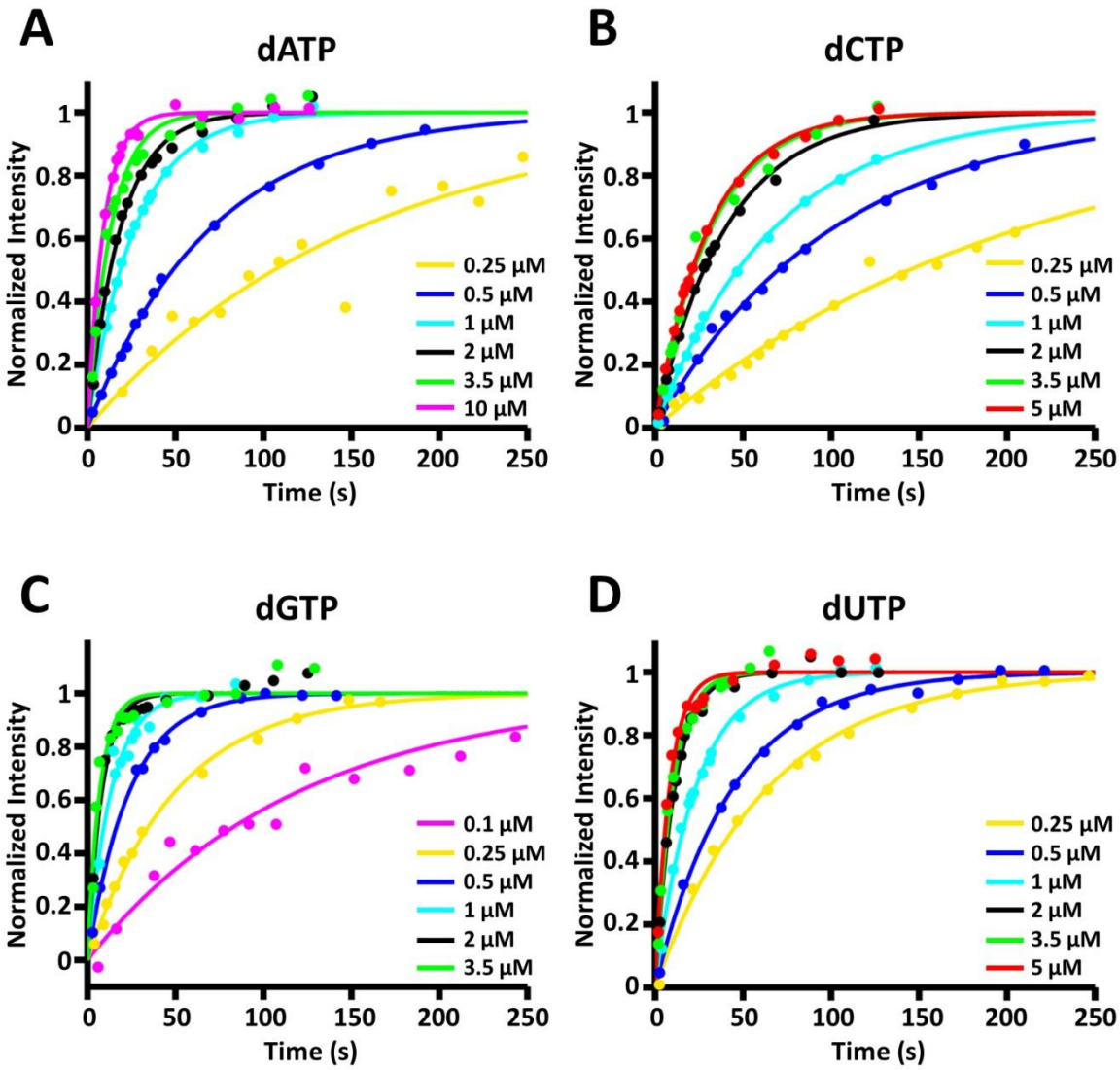


Figure 2.4. Bst DNA polymerase incorporation kinetics of coumarin-labeled dNTPs at a series of concentrations. Each data curve was fitted to an exponential in the form $a(1-e^{-k_{obs}t})+c$ and then offset and normalized.

In eq 2.3, the observed reaction rate constant is given by

$$k_{obs} = k_2[dNTP]/([K_d + [dNTP]) \quad (2.4)$$

We plotted the k_{obs} values as a function of nucleotide concentration and fitted the curves

using eq 2.4 to determine the dissociation constant (K_d) and turnover number (k_2). As shown in Figure 2.5, the data can be fitted very well with adjusted R^2 values above 0.99.

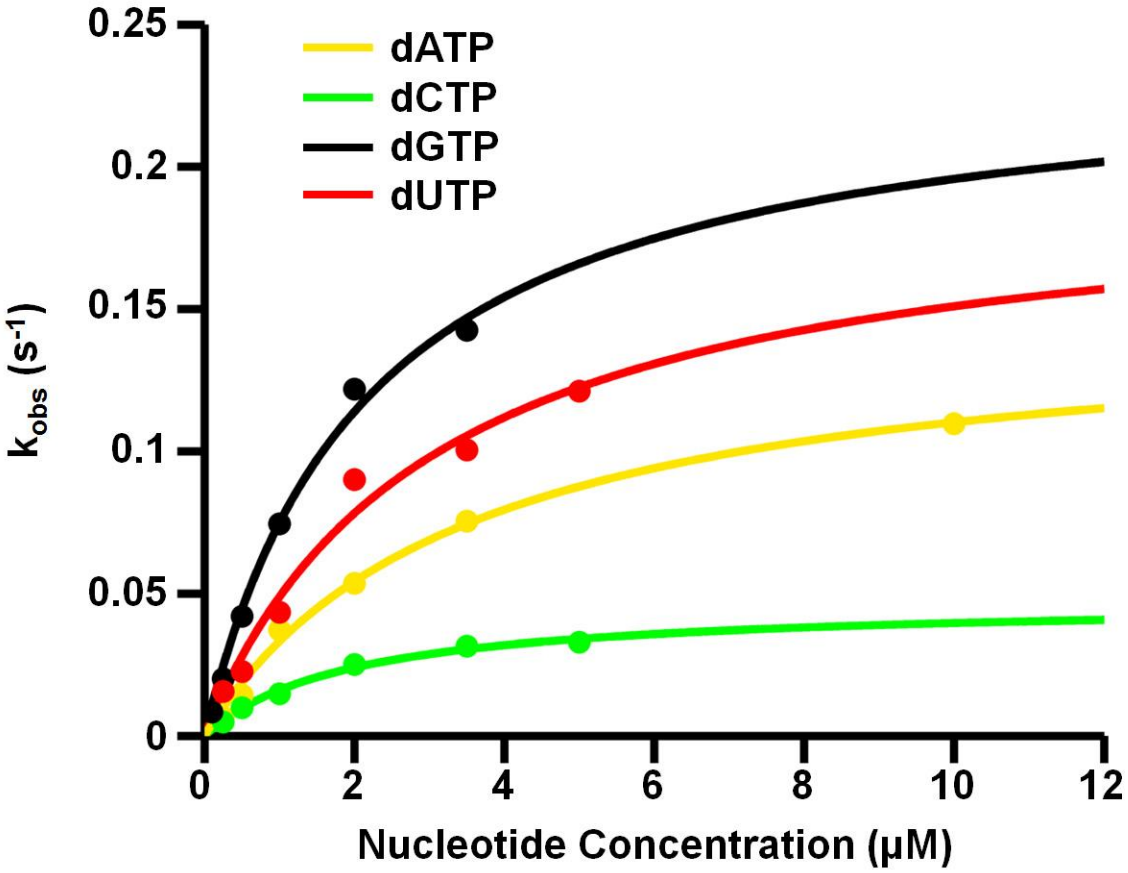


Figure 2.5. Observed reaction rate constants as a function of the concentrations of four coumarin-labeled nucleotides. The curves are fit with eq 2.4 in order to determine the kinetic parameters k_2 and K_d.

The K_d and k_2 values for the four coumarin-labeled nucleotides are listed in Table 2.2. Error values represent the 95% confidence bounds for the parameter fits. The turnover numbers measured at 25°C vary from ~0.05 s^{-1} to ~0.24 s^{-1}, which are lower than the turnover rates for natural nucleotides with other polymerases. This is expected since the reactions were carried out below the optimal temperature of Bst DNA

polymerase (~65°C) and the chemical modification could also affect the rates. Dissociation constants range from ~2 to 3.5 μM within K_M values reported for other polymerases and native nucleotides.[16] K_d equals K_M if rapid equilibrium of nucleotide binding and dissociation is assumed ($k_{-1}>>k_2$).

Table 2.2. Turnover rates and dissociation constants.

Nucleotide	**k_2 (s^{-1})**	**K_d (μM)**
coumarin-dATP	0.15±0.02	3.5±1.3
coumarin-dGTP	0.24±0.06	2.2±1.1
coumarin-dUTP	0.20±0.08	3.0±2.5
coumarin-dCTP	0.047±0.008	2.0±0.8

2.5 Discussion

Our method provides an alternative to the stopped-flow and quench-flow methods for measuring incorporation kinetics of labeled nucleotides by DNA polymerases. We use a microfluidic flowcell for the rapid initiation of the reactions and TIRF microscopy for direct real-time monitoring of the reactions on the surfaces, eliminating the needs for stopped-flow or quench-flow apparatuses and subsequent analysis by electrophoresis. Using our setup with an imaging window of 80 μm x 80 μm and a linear fluid flow velocity of 100 mm/s, a reaction can be initiated in less than 1 ms, a performance on par

or better than that of a stopped-flow and quench-flow methods. TIRF microscopy enables good signal to background kinetics measurements with up to 10 μM of fluorescently labeled nucleotides in the bulk solution. Optimization of the TIR illumination angle may allow for the use of higher concentrations of labeled nucleotides. To prevent photobleaching of the fluorescent dyes, we acquired the time-lapse images at a different field of view for each image. Since the imaged area is relatively small and very uniform, the field-to-field signal intensity variation was not considered. Due to the lack of sequence specific hybridization for the template capture, it is possible for the template to be bound at various positions on the poly$(dT)_{50}$. It is feasible that changes in hybridization location could alter the fluorescent properties of the coumarin label due to different interactions with the glass surface. We anticipate that the variations are likely not more than 25 bases and therefore differences in the proximity effect from the surface to be minor. Another consequence of our method is that there is potential for the 3' end of immobilized poly(dT)50 or template's poly(dA)25 could incorporate a base. Due to the long length of the poly$(dT)_{50}$, it is statistically unlikely that it will by hybridized such that an incorporation event can occur from its 3' when hybridized to the poly$(dA)_{25}$ tail of the template oligo. It is likely that an incorporation of a dATP onto the 3' of the poly$(dA)_{25}$ would occur. However, once a labeled nucleotide is incorporated, it is highly improbable that a second labeled nucleotide will be incorporated. Therefore this 3' poly(dA)25 will simply act as a second primer, not affecting the kinetics. This is supported by the fact that the kinetics data for dATP closely resembles the other nucleotides and fits the model well.

2.6 Future Direction

An interesting variation of our technique is to use a mixture of a native unlabeled nucleotide and its corresponding labeled nucleotide and to perform the reactions with various ratios and concentrations of the labeled and native nucleotides. We have shown that it is feasible to extract the kinetics parameters for both natural and labeled nucleotides simultaneously (Walsh and Huang, to be published). The information can then be used to predict the ratio of labeled to native nucleotides required to incorporate a desired percentage of the labeled nucleotide to implement a novel sequencing strategy called natural sequencing by synthesis (nSBS) where a small fraction of a non-terminating fluorescently labeled nucleotide is incorporated along with the corresponding natural nucleotide in the cyclic DNA sequencing by synthesis process.[17] In summary, our method enables the rapid and automated measurements of the nucleotide incorporation kinetics of DNA polymerases, which will be very useful for the mechanistic studies of DNA polymerases and for implementing and improving the current generation of DNA sequencing technologies that rely on the incorporation of labeled nucleotides by DNA polymerases.[18-21]

2.7 Acknowledgements

This chapter, in part, is a reproduction of the material as it appears in: Walsh, M.T., Roller, E.E., Ko, K.S., Huang, X. (2015) Measurement of DNA polymerase incorporation kinetics of dye-labeled nucleotides using total internal reflection

fluorescence microscopy. *Biochemistry*, **54**, 4019-4021; copyright 2015American Chemical Society. Used with permission. The dissertation author was the primary investigator and author of this paper.

2.8 References

1. Joyce, C.M. and Benkovic, S.J. (2004) DNA polymerase fidelity: kinetics, structure, and checkpoints. *Biochemistry*, **43**, 14317-14324.

2. Bertram, J.G., Oertell, K., Petruska, J. and Goodman, M.F. (2010) DNA polymerase fidelity: comparing direct competition of right and wrong dNTP substrates with steady state and pre-steady state kinetics. *Biochemistry*, **49**, 20-28.

3. Johnson, K.A. (2010) The kinetic and chemical mechanism of high-fidelity DNA polymerases. *Biochim Biophys Acta*, **1804**, 1041-1048.

4. Johnson, K.A. (1992) Transient State Kinetic Analysis of Enzyme Reaction Pathways. *The Enzymes*, **XX**, 1-61.

5. Benkovic, S.J. and Cameron, C.E. (1995) Kinetic analysis of nucleotide incorporation and misincorporation by Klenow fragment of Escherichia coli DNA polymerase I. *Methods Enzymol*, **262**, 257-269.

6. Frey, M.W., Sowers, L.C., Millar, D.P. and Benkovic, S.J. (1995) The nucleotide analog 2-aminopurine as a spectroscopic probe of nucleotide incorporation by the Klenow fragment of Escherichia coli polymerase I and bacteriophage T4 DNA polymerase. *Biochemistry*, **34**, 9185-9192.

7. Bailey, M.F., Thompson, E.H. and Millar, D.P. (2001) Probing DNA polymerase fidelity mechanisms using time-resolved fluorescence anisotropy. *Methods*, **25**, 62-77.

8. Tsai, Y.C. and Johnson, K.A. (2006) A new paradigm for DNA polymerase specificity. *Biochemistry*, **45**, 9675-9687.

9. Stengel, G., Gill, J.P., Sandin, P., Wilhelmsson, L.M., Albinsson, B., Norden, B. and Millar, D. (2007) Conformational dynamics of DNA polymerase probed with a novel fluorescent DNA base analogue. *Biochemistry*, **46**, 12289-12297.

10. Berezhna, S.Y., Gill, J.P., Lamichhane, R. and Millar, D.P. (2012) Single-molecule Forster resonance energy transfer reveals an innate fidelity checkpoint in DNA polymerase I. *J Am Chem Soc*, **134**, 11261-11268.

11. Bermek, O., Grindley, N.D. and Joyce, C.M. (2013) Prechemistry nucleotide selection checkpoints in the reaction pathway of DNA polymerase I and roles of glu710 and tyr766. *Biochemistry*, **52**, 6258-6274.

12. Maxwell, B.A., Xu, C. and Suo, Z. (2014) Conformational dynamics of a Y-family DNA polymerase during substrate binding and catalysis as revealed by interdomain Forster resonance energy transfer. *Biochemistry*, **53**, 1768-1778.

13. Burghardt, T.P. and Axelrod, D. (1981) Total internal reflection/fluorescence photobleaching recovery study of serum albumin adsorption dynamics. *Biophys J*, **33**, 455-467.

14. Barbee, K.D., Chandrangsu, M. and Huang, X. (2011) Fabrication of DNA polymer brush arrays by destructive micropatterning and rolling-circle amplification. *Macromol Biosci*, **11**, 607-617.

15. Ruparel, H., Bi, L., Li, Z., Bai, X., Kim, D.H., Turro, N.J. and Ju, J. (2005) Design and synthesis of a 3'-O-allyl photocleavable fluorescent nucleotide as a reversible terminator for DNA sequencing by synthesis. *Proc Natl Acad Sci U S A*, **102**, 5932-5937.

16. Kong, H., Kucera, R.B. and Jack, W.E. (1993) Characterization of a DNA polymerase from the hyperthermophile archaea Thermococcus litoralis. Vent DNA polymerase, steady state kinetics, thermal stability, processivity, strand displacement, and exonuclease activities. *J Biol Chem*, **268**, 1965-1975.

17. Huang, X. and Roller, E.E. (2014) United States of America patent no. US 8772473 B2.

18. Aravin, A., Gaidatzis, D., Pfeffer, S., Lagos-Quintana, M., Landgraf, P., Iovino, N., Morris, P., Brownstein, M.J., Kuramochi-Miyagawa, S., Nakano, T., Chien, M., Russo, J.J., Ju, J., Sheridan, R., Sander, C., Zavolan, M. and Tuschl, T. (2006) A novel class of small RNAs bind to MILI protein in mouse testes. *Nature*, **442**, 203-207.

19. Bentley, D.R., Balasubramanian, S., Swerdlow, H.P., Smith, G.P., Milton, J., Brown, C.G., Hall, K.P., Evers, D.J., Barnes, C.L., Bignell, H.R., Boutell, J.M., Bryant, J., Carter, R.J., Keira Cheetham, R., Cox, A.J., Ellis, D.J., Flatbush, M.R., Gormley, N.A., Humphray, S.J., Irving, L.J., Karbelashvili, M.S., Kirk, S.M., Li, H., Liu, X., Maisinger, K.S., Murray, L.J., Obradovic, B., Ost, T., Parkinson, M.L., Pratt, M.R., Rasolonjatovo, I.M., Reed, M.T., Rigatti, R., Rodighiero, C., Ross, M.T., Sabot, A., Sankar, S.V., Scally, A., Schroth, G.P., Smith, M.E.,

Smith, V.P., Spiridou, A., Torrance, P.E., Tzonev, S.S., Vermaas, E.H., Walter, K., Wu, X., Zhang, L., Alam, M.D., Anastasi, C., Aniebo, I.C., Bailey, D.M., Bancarz, I.R., Banerjee, S., Barbour, S.G., Baybayan, P.A., Benoit, V.A., Benson, K.F., Bevis, C., Black, P.J., Boodhun, A., Brennan, J.S., Bridgham, J.A., Brown, R.C., Brown, A.A., Buermann, D.H., Bundu, A.A., Burrows, J.C., Carter, N.P., Castillo, N., Chiara, E.C.M., Chang, S., Neil Cooley, R., Crake, N.R., Dada, O.O., Diakoumakos, K.D., Dominguez-Fernandez, B., Earnshaw, D.J., Egbujor, U.C., Elmore, D.W., Etchin, S.S., Ewan, M.R., Fedurco, M., Fraser, L.J., Fuentes Fajardo, K.V., Scott Furey, W., George, D., Gietzen, K.J., Goddard, C.P., Golda, G.S., Granieri, P.A., Green, D.E., Gustafson, D.L., Hansen, N.F., Harnish, K., Haudenschild, C.D., Heyer, N.I., Hims, M.M., Ho, J.T., Horgan, A.M., Hoschler, K., Hurwitz, S., Ivanov, D.V., Johnson, M.Q., James, T., Huw Jones, T.A., Kang, G.D., Kerelska, T.H., Kersey, A.D., Khrebtukova, I., Kindwall, A.P., Kingsbury, Z., Kokko-Gonzales, P.I., Kumar, A., Laurent, M.A., Lawley, C.T., Lee, S.E., Lee, X., Liao, A.K., Loch, J.A., Lok, M., Luo, S., Mammen, R.M., Martin, J.W., McCauley, P.G., McNitt, P., Mehta, P., Moon, K.W., Mullens, J.W., Newington, T., Ning, Z., Ling Ng, B., Novo, S.M., O'Neill, M.J., Osborne, M.A., Osnowski, A., Ostadan, O., Paraschos, L.L., Pickering, L., Pike, A.C., Chris Pinkard, D., Pliskin, D.P., Podhasky, J., Quijano, V.J., Raczy, C., Rae, V.H., Rawlings, S.R., Chiva Rodriguez, A., Roe, P.M., Rogers, J., Rogert Bacigalupo, M.C., Romanov, N., Romieu, A., Roth, R.K., Rourke, N.J., Ruediger, S.T., Rusman, E., Sanches-Kuiper, R.M., Schenker, M.R., Seoane, J.M., Shaw, R.J., Shiver, M.K., Short, S.W., Sizto, N.L., Sluis, J.P., Smith, M.A., Ernest Sohna Sohna, J., Spence, E.J., Stevens, K., Sutton, N., Szajkowski, L., Tregidgo, C.L., Turcatti, G., Vandevondele, S., Verhovsky, Y., Virk, S.M., Wakelin, S., Walcott, G.C., Wang, J., Worsley, G.J., Yan, J., Yau, L., Zuerlein, M., Mullikin, J.C., Hurles, M.E., McCooke, N.J., West, J.S., Oaks, F.L., Lundberg, P.L., Klenerman, D., Durbin, R. and Smith, A.J. (2008) Accurate whole human genome sequencing using reversible terminator chemistry. *Nature*, **456**, 53-59.

20. Eid, J., Fehr, A., Gray, J., Luong, K., Lyle, J., Otto, G., Peluso, P., Rank, D., Baybayan, P., Bettman, B., Bibillo, A., Bjornson, K., Chaudhuri, B., Christians, F., Cicero, R., Clark, S., Dalal, R., Dewinter, A., Dixon, J., Foquet, M., Gaertner, A., Hardenbol, P., Heiner, C., Hester, K., Holden, D., Kearns, G., Kong, X., Kuse, R., Lacroix, Y., Lin, S., Lundquist, P., Ma, C., Marks, P., Maxham, M., Murphy, D., Park, I., Pham, T., Phillips, M., Roy, J., Sebra, R., Shen, G., Sorenson, J., Tomaney, A., Travers, K., Trulson, M., Vieceli, J., Wegener, J., Wu, D., Yang, A., Zaccarin, D., Zhao, P., Zhong, F., Korlach, J. and Turner, S. (2009) Real-time DNA sequencing from single polymerase molecules. *Science*, **323**, 133-138.

21. Harris, T.D., Buzby, P.R., Babcock, H., Beer, E., Bowers, J., Braslavsky, I., Causey, M., Colonell, J., Dimeo, J., Efcavitch, J.W., Giladi, E., Gill, J., Healy, J., Jarosz, M., Lapen, D., Moulton, K., Quake, S.R., Steinmann, K., Thayer, E., Tyurina, A., Ward, R., Weiss, H. and Xie, Z. (2008) Single-molecule DNA sequencing of a viral genome. *Science*, **320**, 106-109.

3 CAPTURE AND ENUMERATOIN OF mRNA TRANSCRIPTS FROM SINGLE CELLS USING A MICROFLUIDIC DEVICE

3.1 Abstract

Accurate measurement of RNA transcripts from single cells will enable the precise classification of cell types and characterization of the heterogeneity in cell populations that play key roles in normal cellular physiology and diseases. As a step towards this end, we have developed a microfluidic device and methods for automatic hydrodynamic capture of single mammalian cells and subsequent immobilization and digital counting of polyadenylated mRNA molecules released from the individual cells. Using single-molecule fluorescence imaging, we have demonstrated that polyadenylated mRNA molecules from single HeLa cells can be captured within minutes by hybridization to polydeoxyribothymidine oligonucleotides covalently attached on the glass surface in the device. The total mRNA molecule counts in the individual HeLa cells are found to vary significantly from one another. Our technology opens up the possibility of direct digital enumeration of RNA transcripts from single cells with single-molecule sensitivity using a single integrated microfluidic device.

3.2 Introduction

Microfluidic devices provide many advantages for processing the molecular contents of cells due to the miniaturization reaction volumes to maintain physiologically relevant concentrations and the ability to process many samples in parallel. Current microfluidic technologies are used for prepare sample for off-chip analysis by RNA-seq or quantification of a limited number of transcript on-chip by RT-qPCR. These techniques have so far been incapable of performing comprehensive on-chip detection and digital counting of RNA transcripts from single cells.

We have developed a platform to enable the capture, isolation, and lysis of single cells, followed by the on-chip capture and enumeration of the total polyadenylated mRNA molecules by single-molecule fluorescence imaging. Single cells are captured along a series of hydrodynamic traps and isolated by polydimethysiloxane (PDMS) valves. The cells are then chemically lysed in individual compartments where polyadenylated (poly(A)) mRNA molecules are captured by hybridization onto polydeoxyribothymidine oligos (poly(dT)) functionalized on the glass surface. We also report a method for the selective functionalization of the PDMS and glass surfaces in the microfluidic devices, enabling the efficient capture of mRNA molecules from single cells and single-molecule fluorescence imaging. We use time-lapse single-molecule fluorescence imaging and numeric simulations to determine the kinetics of mRNA capture. It is shown that the mRNA molecules in the microfluidic chambers diffuse down and hybridize to the poly(dT) oligos on the surface very rapidly, to almost

completion within minutes. As a proof of principle, we analyze the total number of polyadenylated mRNA molecules from individual HeLa cells.

3.3 Materials and Methods

3.3.1 Microfluidic Device Design

An advanced microfluidic device consisting of 16 modules for parallel processing of 16 cells (Figure 3.1a) was designed and fabricated by Dr. Alex Hsaio. Our device is designed to integrate both single-cell capture and downstream processing in individual modules connected in serial for parallel processing of multiple cells. The cell traps of the modules are arranged in series while the processing chambers are arranged in parallel and can be isolated from one another by valves to enable parallel downstream processing. Each module contains a hydrodynamic cell trap and a compartment for processing and imaging (Figure 3.1b). The single-cell hydrodynamic traps are adapted from the work of Tan *et al*[1].[1] and the principle of operation is illustrated in Figure 3.1c. Each trap chamber has two flow paths, one through a cell trap and one through a bypass. In this work, our trap path is designed as a small channel with 8 μm x 8 μm cross section with a round 22 μm x 22 μm cup-shaped mouth. The dimensions of the paths are designed such that the cell trap path has a much lower fluidic resistance than the bypass path. Therefore, the bulk of fluid containing the cells is channeled into the single-cell trap. Any incoming cell that is larger than the narrow cross-section of the trap will be captured. Upon trapping of

the first cell, the path is immediately blocked and the fluid containing additional cells is diverted into the bypass to the next trap in series. The traps can be tailored to capture cells of specific shape, size and elasticity.

Orthogonal to the direction of cell capture channel are the compartments for cell lysis, RNA capture and further processing (Fig. 3.1b). The glass surface of this compartment is functionalized with 50-base poly(dT) oligos to capture the polyadenylated mRNA molecules by hybridization. The cell suspension is flowed into the trap channel after valve lines 2 and 3 are actuated to isolate the cell trap channel from the reagent channels and processing chambers. Upon successful single-cell trapping, the valves between the modules (V1 and V2) are then closed to compartmentalize the cells. The individual cells are then moved upward by injecting a lysis solution from the lower reagent channels into the processing chambers where the cells are lysed and the released poly(A) mRNA molecules are captured onto the glass surface by hybridization. Figure 3.2 shows a photograph of a fabricated device (Fig. 3.2a), a contrast brightfield image of one module with the area for mRNA capture outlined in red (Fig. 3.2b).

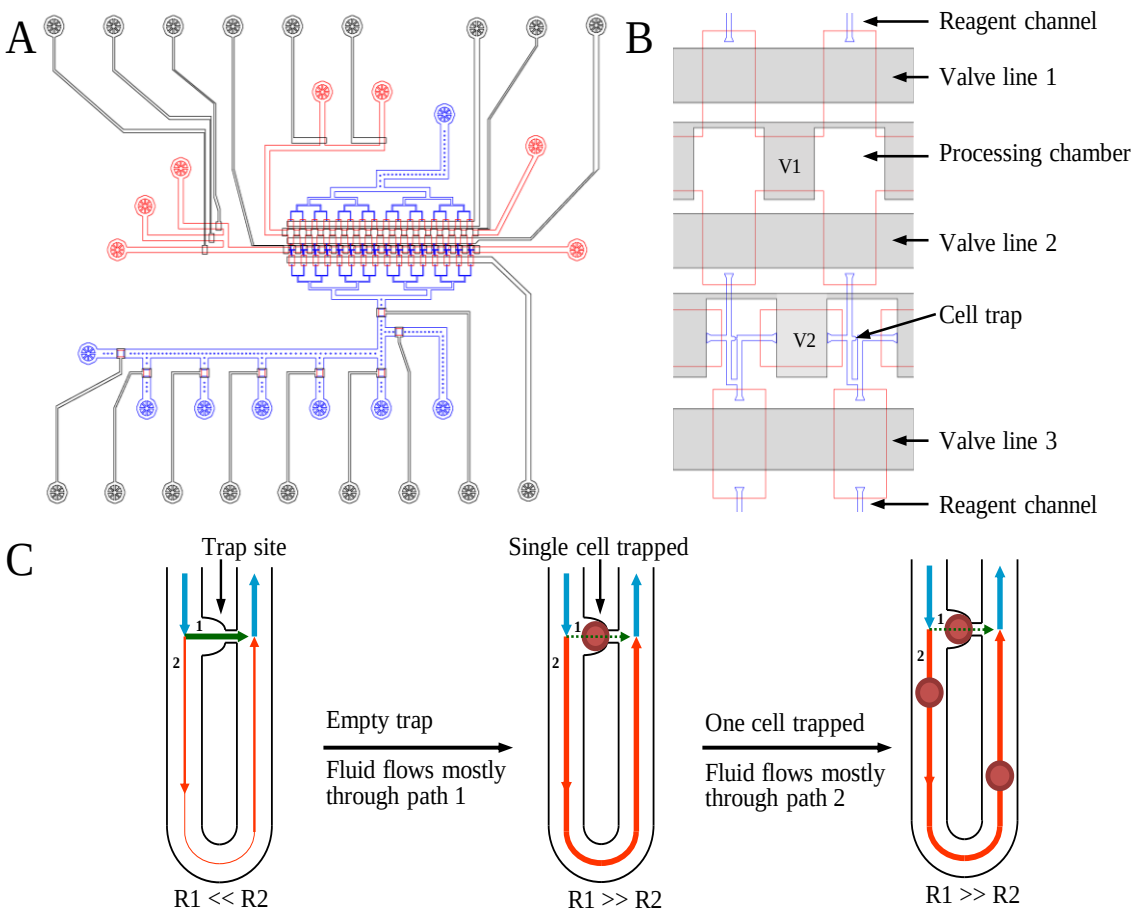


Figure 3.1. Design and operation of the microfluidic device for capturing and counting of mRNA molecules from single cells. A. 2D CAD layout of a device with 16 modules in serial. Red: fluid channels and chambers with 25 µm height; Blue: Fluidic channels/chambers with 25 µm height; Black: valves and valve control lines with 40 µm height. B. An enlarged portion of the device containing two modules. The valves (V1 and V2) and valve lines are shown in light grey. Each module consists of a cell trap and a processing chamber. C. Hydrodynamic trap design and operation. The in-flowing fluid is split into two paths, the cell trap (path 1) and bypass (path 2). The dimensions of the paths are designed such that the fluidic resistance through cell trap path (R1) is much smaller than that of the bypass (R2). In the empty trap, the fluid and cells preferentially flow into the trap path. The entry and trapping of the first single cell results in an abrupt increase in the fluidic resistance of path 1 and the fluid flow is then directed to the bypass (path 2) to the next trap in series, and so on.

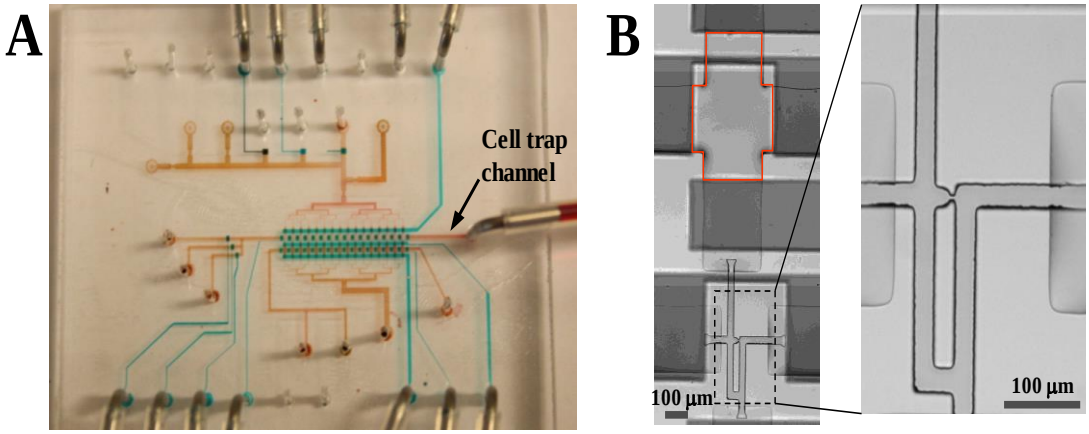


Figure 3.2. Microfluidic device for capturing and counting of mRNA molecules from single cells. A. Photograph of a functional device. The valve control lines are filled with a blue solution while the fluidic channels are filled with an orange solution. B. A contrast brightfield image of an area containing a single module. The area outlined in red is the processing chamber where the cell is lysed and the released polyadenylated mRNA molecules are captured by hybridization to the poly(dT)$_{50}$ oligonucleotides covalently functionalized on the glass surface.

3.3.2 PDMS Device Fabrication

The PDMS devices were designed and fabricated according to the procedures of Unger *et al.*[2] with some modifications as described in our previous work.[3,4] A two-step lithography process was utilized to fabricate the mold for the PDMS fluidic layer with larger rounded channels for valving and smaller channels with vertical walls for cell trapping and fluid flow. All fluidic channels were fabricated to a height of 25 µm. The width and length of the processing chamber are 250 µm. The mold for the valve control layer was patterned with a channel height of 40 µm.

The valve control layer and fluidic channel layer were prepared using Sylgard 184 (Dow Corning) with part A to part B ratio of 5:1 and 20:1, respectively. The valve control

layer was prepared by adding degassed PDMS prepolymer mixture onto the mold in a carrier. The fluidic channel layer was prepared by spin-coating of degassed PDMS mixture onto the mold to a height of about 50 μm. After the PDMS layers were cured at 65 °C for 30 min., the valve control PDMS layer was peeled off the mold and holes were punched for fluidic connections using a 0.75 mm diameter punch. The layer was aligned and laid onto the fluidic channel layer on the mold. The two layers were bonded together by heating at 65 °C for 4 hours. The bonded PDMS layers were peeled off the mold together and holes were punched for the fluidic connections.

3.3.3 Chemical Bonding of PDMS to Glass and Selective Functionalization of Surfaces

Central to our design is the surface chemistry for efficient capture and analysis of single RNA molecules. This requires surfaces that have low non-specific binding to biomolecules and reagents such as fluorescently-labeled oligos to minimize sample loss and to reduce background noises for reliable single-molecule fluorescence imaging. For this purpose, we developed a method for chemical bonding and differential functionalization of the PDMS and glass surfaces.

Prior to the assembly of the glass-PDMS device, we functionalized the PDMS surface with epoxy groups and the glass surface with primary amine groups. This enables the chemical bonding of the PDMS to the glass coverslip without using plasma treatment which usually damages the functional groups on the surface. Moreover, our strategy enables the selective passivation of the PDMS surface inside the microfluidic channels with PEG which is known to prevent non-specific binding and fouling, and the selective

covalent functionalization of only the glass surface with poly$(dT)_{50}$ oligos via a PEG linker. The selective functionalization enables the specific capture of polyadenylated RNA by hybridization onto the glass surface for processing and single-molecule imaging while the non-specific binding of RNA and other reagents to the PDMS surface is minimized.

The overall process is illustrated in Fig. 3.3. First, the bottom of the PDMS channel layer is covalently bonded onto a glass coverslip by a chemical reaction between the reactive groups functionalized on their surfaces.[5] To do so, the surface of the PDMS layer is functionalized with epoxy groups while the surface of the cover glass is activated with primary amine groups. Borosilicate glass coverslips (50 x 50 mm x 0.17 mm) functionalized with amine groups essentially as described.[6] Briefly, the glass slides were cleaned and then soaked in 2% solution of 3-aminopropyltriethoxysilane in 93% methanol plus 5% glacial acetic acid for 10 minutes, then rinsed with methanol and dried with argon gas. The surface of the PDMS was activated by oxygen plasma in a UV-ozone cleaner (Jelight) for 3 min. and then functionalized by soaking in a 2% solution of 3-glycidoxypropyl triethoxysilane in 93% methanol and 5% glacial acetic acid for 15 min. The PDMS was then rinsed with methanol and blown dry with argon. Finally, the functionalized PDMS was laid on top of the coverslip. Any trapped air bubbles between layers were removed by placing the assembly under vacuum for 15 min. and rapidly equilibrating to atmospheric pressure. The PDMS-glass assembly was then heated at 65 °C for 4 hours to facilitate the crosslinking of the epoxy groups on the PDMS to the amine groups on the coverslip.

Next, the surface of the PDMS inside the channels was passivated with polyethylene glycol (PEG) by reacting the epoxy groups with a monofunctional amine-PEG2000. The assembled device was heated at 60 °C on a hotplate and an automated syringe pump was used to deliver a solution containing 2 mM amine-PEG and 10 mM triethylamine (TEA) in anhydrous N,N-dimethylformamide (DMF) through the channels. The reaction was carried out at 60 °C for 1 hour. The solution was frequently refreshed to compensate for evaporation.

The amine groups on the glass surface of the channels was then conjugated to a bifunctional carboxyl PEG1000 by injecting a solution containing 1 mM of the carboxymethyl-PEG1000, 10 mM EDC and 5 mM TEA in anhydrous DMF into the channels and incubating for two hours at room temperature. The channels were treated then with 10 mM sulfo-NHS acetate in 100 mM sodium borate, pH 8.5 with 0.02% Triton X-100 for 30 minutes at room temperature to acetylate the residual amine groups.

Finally, the glass surface was selectively functionalized with poly$(dT)_{50}$ oligos. The valves above and below the RNA capture and processing chambers (valve lines 1 and 2 in Fig. 3.1B) were closed to isolate the channel containing the chambers and a solution containing 1 µM amine-labeled poly$(dT)_{50}$, 10 mM EDC in 100 mM (N-morpholino) ethanesulfonic acid (MES) buffer at pH 5.0 was injected into the channel and incubated for 1 hour at room temperature. The channel was then washed and stored in saline sodium citrate buffer (SSC) until use.

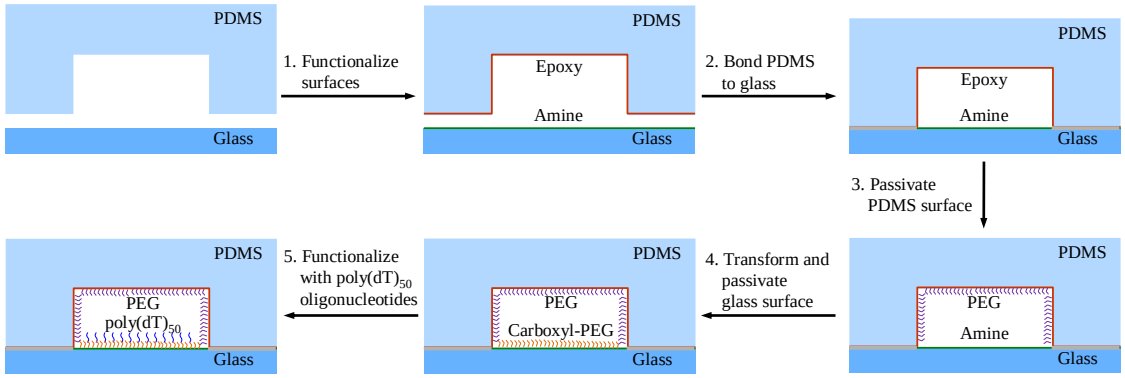


Figure 3.3. Method for chemical bonding and selective functionalization of the surfaces of the microfluidic devices. The novel strategy allows for the chemical bonding of PDMS to a glass substrate and selective passivation of the PDMS surface and functionalization of the glass surface.

3.3.4 Device Operation: Capture, Isolation, and Lysis of Single Cells

Figure 3.4 shows the hydrodynamic capture and chemical lysis of single cells in a device. An image of two modules, each with a single cell trapped is shown in Fig. 3.4a. Following cell capture, the valves between the modules were closed to isolate each trapped cell from its neighbors. The valve between the cell trap and processing chamber which has been pre-filled with air is opened and the lysis/binding buffer was injected from the bottom reagent channel. Since the path from the backside of the cell trap has a lower fluidic resistance, the cell was dislodged from the trap site and pushed towards the processing chamber (Fig. 3.4b). As more lysis/binding solution was injected, the cell in the solution was pushed into the processing chamber by completely displacing the air in the chamber. An equal volume (~1.5 nL) of lysis buffer is mixed with the solution in the

cell trap. The cell was lysed by the high concentration o f the anionic detergent, lithium dodecyl sulfate (Fig. 3.4c). Cell lysis was observed to occur within seconds upon mixing with the lysis solution.

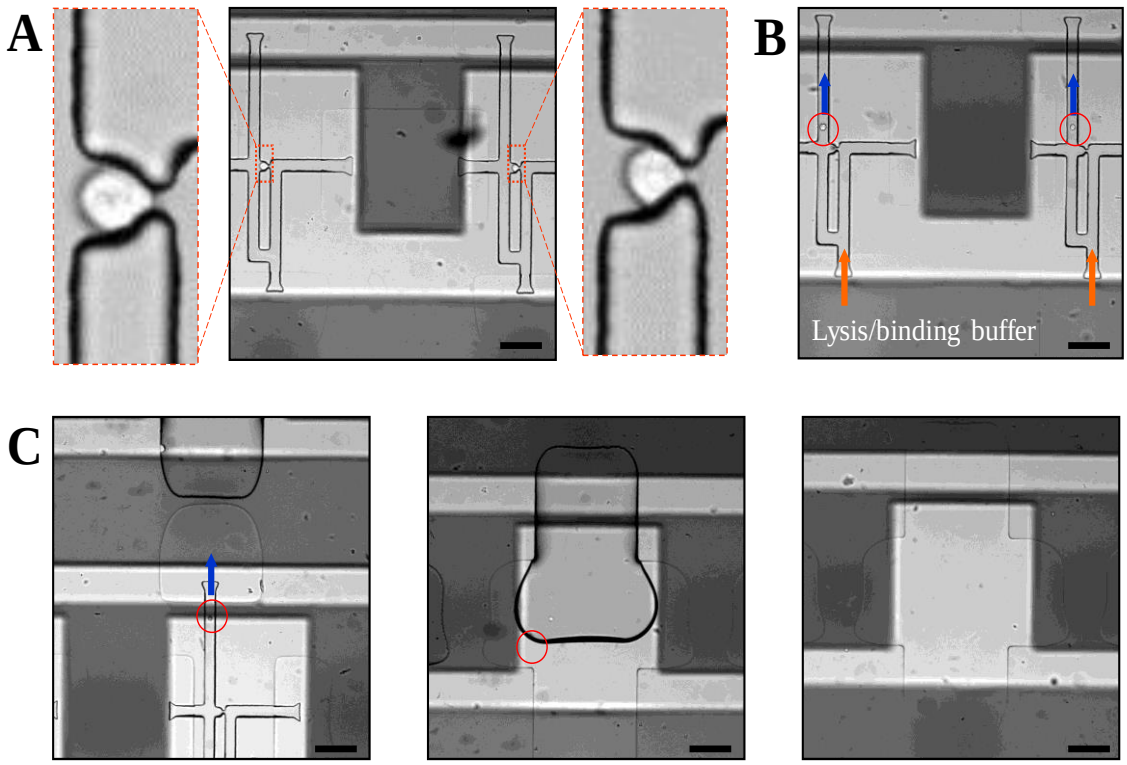


Figure 3.4. Hydrodynamic capture and chemical lysis of single cells. A. Bright-field micrograph of a portion of the device showing two modules, each with a single cell captured. The zoom-in views show that one cell (left) is larger than the other (right). B. Following cell capture, the inter-module valves were closed and a lysis buffer was injected from the bottom reagent channel (orange arrows) to move the cells (in red circles) out of the trap site towards the processing chamber (blue arrows). C. The further injection of the lysis buffer pushed the cell further toward (left panel) and into the processing chamber (middle panel) while the air was displaced from the chamber. Eventually the air was completely displaced and chamber was filled. The cell was fully lysed in the chamber and no longer observable (right panel). The scale bar is 100 µm.

3.3.5 Characterization of mRNA Capture Kinetics in a Microfluidic Device

To assess the rate and efficiency of mRNA capture in a microfluidic device two sets of experiments were devised using either synthetic mRNA containing a fluorescent label or utilizing mRNA from HeLa cell lysate. Extensive mRNA capture kinetics characterization was performed and matched to a diffusion-reaction kinetics model to compute hybridization reaction rate constants and the effects of diffusion.

The fluorescently-labeled synthetic mRNA was prepared by the following procedure. The mRNA transcription from a gene encoding a phi29 DNA polymerase on a plasmid was performed by Dr. Zhixia Liu. A 30-cycle PCR was performed to amplify the gene with a primer to include the T7 promoter and a primer designed to include a 150-base poly(dT) tail. The PCR product was used as a template to transcribe the mRNA in the presence of fluorescein-12-UTP (Enzo Life Sciences) at 2% of the total UTP concentration using the HiScribe™ T7 transcription Kit (New England Biolabs). After digest of the DNA using DNase-I, the polyadenylated fluorescein-labeled mRNA was purified using a kit (MEGAclear Kit, Ambion) and murine RNase inhibitor (Roche) was added into the product to prevent degradation. The final mRNA product is about 1700 bases long and contains a 150-base poly(A) tail (termed poly$(A)_{150}$-phi29 mRNA for short). The concentration of the product and average number of labels per molecule were quantified by UV-vis absorption.

Cells were cultured and prepared in the same manner for performing single-cell analysis in the advanced microfluidic device or for lysing to measure mRNA binding

kinetics in a simplified microfluidic device. Cell growth, maintenance and collection were performed by Dr. Hosuk Lee. HeLa cells were grown and maintained under standard culture conditions (37 °C and 5% CO_2) in DMEM supplemented with 10% fetal bovine serum and 1% penicillin/streptomycin. The cells were detached from culture flask surfaces using 0.25% trypsin/EDTA, re-suspended in DMEM and centrifuged at 1000 rpm for 6 minutes to remove the trypsin. Finally, the cells were diluted in DMEM to a concentration of 0.5 million cells per mL for the experiments.

To establish a benchmark, the pure synthetic fluorescently-labeled poly$(A)_{150}$-phi29 mRNA at 50 pM in a lysis/binding buffer (500 mM LiCl, 100 mM Tris-HCl, pH 7.5, 10 mM EDTA, 5 mM DTT, 1% lithium dodecyl sulfate) was used to determine the capture kinetics. Measurements were also performed with 50 pM mRNA in the presence of 250 pM and 1 nM of Alexa647-lableled poly$(dT)_{25}$ which can hybridize to the poly(A) tail on the mRNA to investigate the effect of the labeled poly$(dT)_{25}$ on the overall capture kinetics. To simplify the measurements, we used a device with straight channels of 2 mm width and 100 μm height fabricated on coverslips functionalized with poly$(dT)_{50}$ oligos. The mRNA solution was pulled into the microfluidic channel at a flow rate of 160 mm/s using a syringe pump and immediately time-lapse fluorescence images were acquired in both fluorescein and Alexa647 channels at the same field of view using a TIRF setup. Excitation was provided by a 488 nm laser for the fluorescein channel and a 642 nm laser for the Alex647 channel. The images were captured using 200 ms exposure and an EM gain of 20.

To simulate the condition used in our single-cell microfluidic device, we used simpler devices with 15 mm x 250 μm x 25 μm straight PDMS channels, which were

fabricated and functionalized similarly, to measure the kinetics of mRNA capture from HeLa cell lysate. To prepare the lysate, the cultured cells were counted and mixed into lysis/binding buffer solution containing 2.5 nM of Cy3-poly$(dT)_{25}$ probe at a concentration of 1 cell per 3 nL, which is the volume of the processing chamber in our single-cell microfluidic device. Capture kinetics measurements were performed in a similar manner using a 532 nm laser for excitation. To quantify the non-specific binding of the mRNA and the Cy3-labeled oligo probe to the surface, two control experiments were also performed, one with a channel that was not functionalized with poly$(dT)_{50}$ oligos, and one with a functionalized channel but only lysis/binding buffer and Cy3-poly$(dT)_{25}$ probe. The time-lapse images were analyzed and the single-molecule counts from the time-lapse images were plotted as a function of time.

3.3.6 Computational Modeling of mRNA Capture Kinetics

mRNA capture is modeled as a one-dimensional diffusion-hybridization process. The mean squared displacement ($<x^2>$) of an mRNA molecule due to Brownian diffusion is calculated using the Einstein equation ($<x^2>=2Dt$) and the Stokes-Einstein equation ($D=k_BT/6\pi\eta R_h$), where D and R_h are the diffusivity and hydrodynamic radius of the mRNA molecule respectively. R_h can be approximated by $R_h=0.5(aLp/3)^{1/2}$, where a is the distance per base of single-stranded RNA (a = 0.49 nm), L is the contour length of the mRNA (number of bases) and p is the persistence length for single-stranded RNA (p = 1.4 nm).[7,8] For simplicity, the hybridization of the poly(A) tail of the mRNA to the poly$(dT)_{50}$ oligos on the surface is treated as a pseudo first-order reaction since the poly$(dT)_{50}$ oligos are in large excess.

The entire process of mRNA capture by one dimensional diffusion and hybridization can be modeled using the diffusion equation:

$$\frac{\partial C}{\partial t} = D\frac{\partial^2 C}{\partial z^2} \qquad 3.1$$

with the boundary conditions of a total reflector at the top surface $(\partial C(z = H,t) / \partial z = 0)$, a pseudo first-order irreversible hybridization reaction at the bottom surface $(\partial C(z = 0,t) / \partial t = -kC(z = 0,t))$ and an initial condition of uniform mRNA concentration $(C(z,t = 0) = C_0)$, where C is the concentration of the mRNA in the solution, k is the rate constant for the hybridization reaction which is independent of channel height, z is the perpendicular position above the surface of the channel and H is the height of the channel.

To investigate the capture kinetics of mRNA of various lengths in chambers with different heights, the partial differential equation (eq. 3.1) was solved numerically in MATLAB using the finite difference method. A mesh of 10^7 by 10^2 nodes was used to discretize the temporal and spatial components of the one dimensional diffusion model. To do so, first, the rate constant of the pseudo first-order hybridization reaction (k) was determined using the experimental kinetics data of poly$(A)_{150}$-phi29 mRNA capture in a 100 μm high channel. By inputting the calculated RNA diffusivity and channel height, the value of k was optimized so that the overall kinetics determined numerically by solving equation 3.1 using the finite difference method matches the overall experimentally measured kinetics data. The diffusivity of the RNA molecule is calculated using the Stokes-Einstein equation. Once k is determined, the theoretical kinetics data can be numerically computed for any channel height and RNA length. This model was also

used to determine the hybridization reaction kinetics of mRNA capture from HeLa cell lysate in the presence of 2.5 nM Cy3-poly(dT)$_{25}$, as well as the synthetic mRNA in the presence of various concentrations of poly(dT)$_{25}$.

3.3.7 Capture of Polyadenylated mRNA from Single Cells

Single HeLa cells were captured, isolated, and lysed in parallel using devices with 16 modules. Valve operations and fluid delivery were performed as described in our recent work.[3,4] Several steps were performed to prepare the devices prior to the cell capture experiments. First, to reduce non-specific binding of cells, valve lines 2 and 3 were closed and the cell capture channels were passivated with a protein-free blocking buffer (PBS-T20, Thermo Fisher) by injecting the solution into the channels and incubating for >10 minutes. Second, the bottom reagent channels were filled with lysis/binding buffer solution containing 1 nM Alex647-poly(dT)$_{25}$ and 1 unit/μL of murine RNase inhibitor. Finally, the processing chambers were dried by a continuous flow of air through the channels.

Once a device was ready and mounted onto a microscope stage via an adaptor and valve lines 2 and 3 were actuated, a HeLa cell suspension was injected through the main channel at 1-10 μL per hour using an infusion syringe pump. The cells were captured one per trap automatically by hydrodynamic trapping as they flowed through the channel. The process was monitored by brightfield contrast microscopy. Then the inter-compartment valves (valves V1 and V2 in Fig. 3.1b) were closed to isolate the modules.

To process the cells, the individual cell in each module was pushed out of the trap and into the processing and mRNA capture chamber by opening valve line 2 and

injecting the lysis/binding solution from the bottom reagent channel. After the air in the chamber was completely displaced (PDMS is permeable to air), valve line 2 was actuated to close the chamber. As the cell undergoes lysis in each compartment, the polyadenylated mRNA molecules are released into the chamber, diffuse and hybridize to the poly(dT)$_{50}$ oligos functionalized on the glass surface, while simultaneously hybridizing to the Alexa647-poly(dT)$_{25}$ in solution for detection. Bright-field microscopy was used to monitor the cell trapping and lysis process using a 10x objective. After at least ten minutes of incubation of mRNA in the capture chambers, the entire surface of each mRNA capture chamber was scanned and imaged with a 10 percent overlap between fields of view.

3.3.8 Single-Molecule Fluorescence Imaging and Image Processing

Fluorescence imaging was performed using the objective-based TIRF system described in Chapter 2. A back-illuminated EMCCD camera (iXon3 897, Andor Technology) was used for image acquisition. All image processing was performed with MATLAB (MathWorks). For each chamber, a montage was constructed by positioning each image tile in its proper location. Overlapping sections were merged by performing a maximum projection calculation to select which pixel value to incorporate. The background was removed by subtracting from the image using the morphological opening algorithm, calculated using a ball structured element with a radius of eight. The images were then thresholded to select pixels with a value above 50. To remove noisy pixels, a filter was applied to calculate the average density of thresholded spots around

each pixel and a second round of thresholding was applied such that only selected pixels with at least two adjacently selected pixels were chosen. A labeling function was then employed to identify and index the separate objects in the binary image. At this point, spots with a small number of pixels were defined as a single molecule and the remaining objects were analyzed using the watershed transform to segment closely grouped molecules based on the topography in the original image. A histogram of intensity values from all spots was utilized to select a mean intensity for a single mRNA transcript. An adjusted count for the number of transcripts present was estimated by dividing all intensity values by the approximated mean intensity of a single mRNA molecule.

3.4 Results

3.4.1 Kinetics of mRNA Capture in a Microfluidic Device

The capture kinetics of the pure synthetic 1700-base fluorescein-labeled poly(A)$_{150}$-phi29 mRNA in a 100 μm high microfluidic channel in the absence of and presence of Alexa647-poly(dT)$_{25}$ at different ratios can be fit well with the diffusion-reaction model by optimizing the rate constant of the hybridization reaction. The kinetics data and fits are shown in Fig. 3.5. Table 3.1 lists the extracted time constants of the hybridization reaction, the time to reach 95% of total capture, the count of the captured mRNA and capture yields. In the absence of Alexa647-poly(dT)$_{25}$ oligos, the hybridization reaction on the surface is very rapid with a time constant of 6.1 s. As can be observed, the presence of the 25-base Alex647-poly(dT)$_{25}$, which can hybridize to the 150-base poly(A) tail of the mRNA, affects the hybridization kinetics and capture yield.

Since a maximum of 6 poly(dT)$_{25}$ oligos can hybridize to the 150-base poly(A) tail on each mRNA molecule, the ratio of poly(dT)$_{25}$ to mRNA is calculated as molar concentration ratio divided by a factor of 6. As the poly(dT)$_{25}$ to mRNA ratio increases from 0:1 to 0.83:1, and 3.3:1, the hybridization reaction time constant increases from 6.1 s to 7.5 s, and 68.3 s, respectively, and the capture yield decreases to 87% and 20%, respectively.

At a ratio of 0.83:1 which is below the level of the saturation of the poly(A) tails, the hybridization kinetics decreases only slightly. However, once the ratio is above saturation level (3.3:1 in this case), the capture kinetics decreases significantly. Intuitively, this is expected because the rate constant of the hybridization reaction between the poly(A) tail on the mRNA and the poly(dT)$_{50}$ on the surface is approximately linearly proportional to the squared root of the number of the available adenosine bases of the poly(A) tail,[9] and the fraction of available regions in the poly(A) tail for hybridization decreases drastically as the ratio is above saturation level. The capture yield or count is decreased as the ratio of poly(dT)$_{25}$ to the polyadenylated mRNA is increased. This indicates that the poly(A) tails of some mRNA molecules become saturated and inaccessible to capture by the poly(dT)$_{50}$ oligos on the surface. Interestingly, a significant fraction of the total mRNA is captured even though the labeled poly(dT)$_{25}$ is above saturation level.

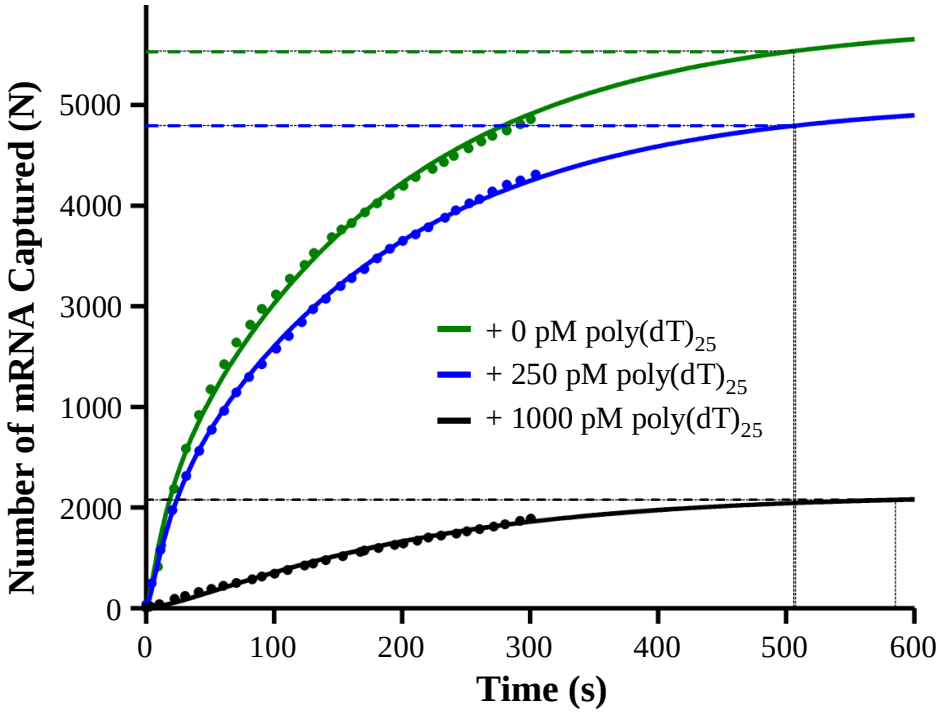


Figure 3.5. Capture kinetics of synthetic polyadenylated fluorescently-labeled mRNA (poly(A)150-phi29mRNA) in 100 µm high microfluidic channels. Shown are the kinetics data of capturing 50 pM mRNA in the presence of 0 pM, 250 pM and 1 nM of Cy3-poly(dT)$_{25}$ oligos. The data are fit to a reaction-diffusion model by optimizing the rate constant of the hybridization reaction on the surface. The dotted curves are the data and the solid lines are the model fits. The time constants of the hybridization reaction in the presence of 0, 250, and 1000 pM of Cy3-poly(T)$_{25}$ are determined to be 6.1 s, 7.5 s and 68.3 s, respectively. The horizontal dashed lines represent 95% completion of mRNA capture as approximated by the extrapolated fit. The vertical dashed lines correspond to the time at which 95% mRNA capture has occurred.

Table 3.1. Extracted parameters from the fits of capture kinetics data to the diffusion-hybridization model. The capture kinetics data were acquired using 50 pM of poly(A)150-phi29mRNA in the presence of various concentrations of Alexa647-poly(dT)$_{25}$ probe. The capture yield is calculated as percentage of the count of captured mRNA relative to the theoretical total in solution above the measurement area.

Concentration of poly(dT)$_{25}$ (pM)	Hybridization time constant (s)	Time for 95% capture (s)	Captured mRNA count	Capture yield	R^2 of fit
0	6.1	506	5535	29%	0.997
250	7.5	507	4795	25%	0.999
1000	68.3	585	1077	5.6%	0.995

3.4.2 Capture Kinetics of mRNA from HeLa Cell Lysate

To simulate the single-cell mRNA capture kinetics experiments in the microfluidic device, our strategy is to add a fluorescently-labeled poly$(dT)_{25}$ oligo into the cell lysate to enable the real-time monitoring of the process by single-molecule fluorescence imaging using TIRF microscopy. Since the concentration of the total number of mRNA from a single HeLa cell is not known, we estimated the possible range and the concentration of the labeled poly$(dT)_{25}$ to be used for our measurements.

We found that the addition of 2.5 nM of Cy3-poly$(dT)_{25}$ oligos to the HeLa cell lysate at a concentration equivalent to one cell in the volume of the processing chamber in our single-cell microfluidic device (~1 cell in 3 nL) allows for the simultaneous rapid capture and detection of the mRNA molecules.

The results from the kinetics experiments using HeLa cell lysate and a microfluidic device with 250 μm wide and 25 μm high straight channels are shown in Fig. 6. The kinetics data of mRNA capture from the HeLa cell lysate can be fit to our diffusion-hybridization model very well (R^2 = 0.988) using a 25 μm channel height and the parameters of a 1400-base RNA, which is the reported median length of the mRNA molecules in HeLa cells.[10] The time constant of the hybridization reaction was determined to be 139 s. The curve of the capture kinetics plateaus quickly around 500 s, and it takes only ~ 7 minutes to capture 95% of the total molecules that can be captured due to the use of the much lower 25 μm channel. In the absence of the Cy3-poly$(dT)_{25}$, the capture rate would have been much faster. Finally, in the chambers of the control experiments, there is only a constant and very low level of non-specific binding of the

labeled probe oligos and mRNA throughout the course of the experiments. This confirms that our surface chemistry has very low non-specific binding and is suitable for single-molecule fluorescence imaging.

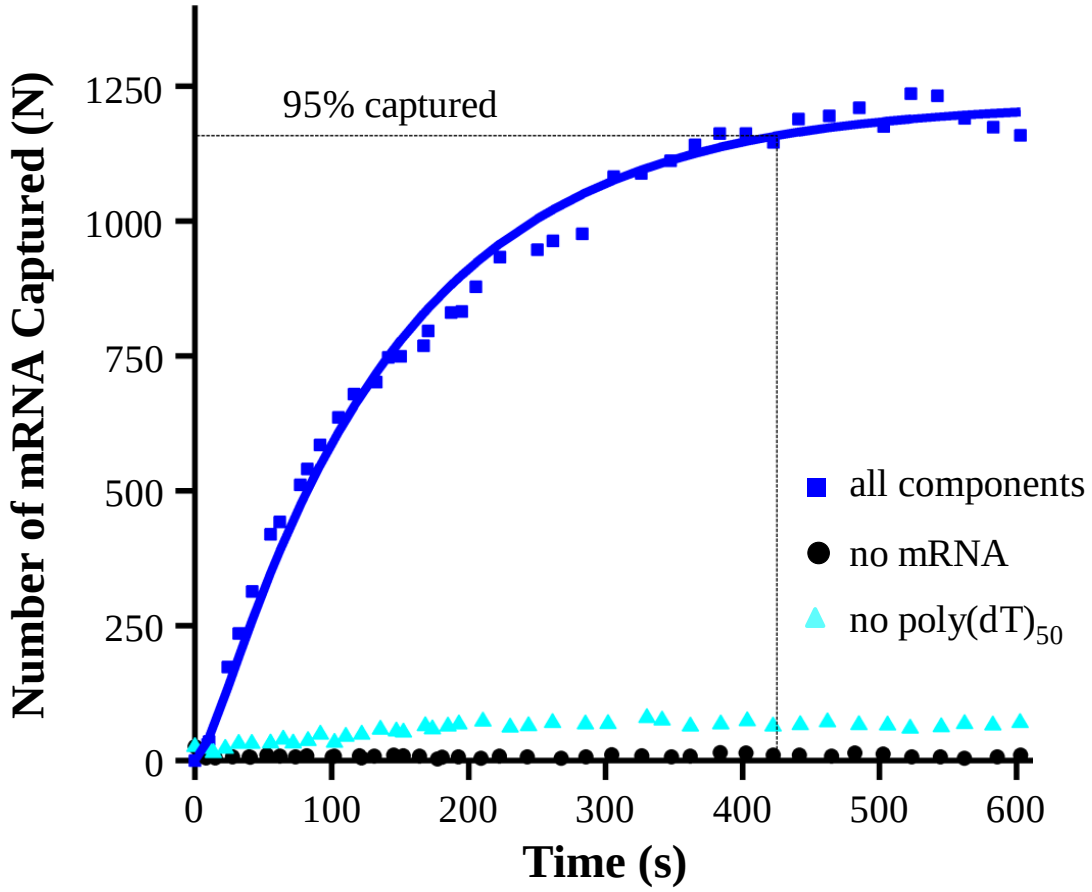


Figure 3.6. Capture kinetics of polyadenylated mRNA from HeLa cell lysate in 25 μm high PDMS microfluidic channels. The concentration of the cell lysate is about 1 cell per 3 nL of lysis/binding buffer. Cy3-poly(dT)$_{25}$ probe was added to the cell lysate to a concentration of 2.5 nM to enable real-time monitoring of the capture kinetics. The kinetics data (blue squares) along with two control experiments (black dots and cyan triangles) are shown. In one control, cell lysate was not included (black dots, no mRNA), while in another control the surface of the device was not functionalized with poly(dT)$_{50}$ oligos. The curve of the mRNA capture data was fit to the diffusion-hybridization model (solid blue line). The time constant (1/k) of the hybridization reaction is determined to be 139 s. The time required to capture 95% of total mRNA (horizontal dotted gray line) is about 7 minutes (425 s).

3.4.3 Effect of Channel Height and mRNA Size on Capture Kinetics

As can be observed in Fig. 3.5 and 3.6, the experimental mRNA capture kinetics data can be modeled very well using a one-dimensional diffusion equation with the boundary conditions of a total reflector at the top surface and a pseudo first-order hybridization reaction at the bottom surface. Fig. 3.7 shows the numerically simulated capture kinetics of RNA of four different lengths from 500 to 50,000 bases and two chamber heights of 25 μm and 100 μm using a time constant (1/k) of 6.1 s for the hybridization reaction that is determined from the experiment with the pure 1700-base poly$(A)_{150}$-phi29mRNA as described earlier. The extracted parameters and the times required for the RNA molecules diffuse across 25 μm and 100 μm high chambers calculated using the Einstein equation are listed in Table 3.2.

With the exception of the very short RNA in 25 μm high chamber, the rate of mRNA capture is limited to a higher degree by diffusion. An increase in the length of mRNA and the height of the chamber results in slower capture rate. At 100 μm chamber height, the rate of mRNA capture is significantly reduced, especially for larger transcripts. Overall, the key finding is that mRNA with length up to 50,000 bases can be captured to completion within 3 minutes in a chamber with 25 μm height.

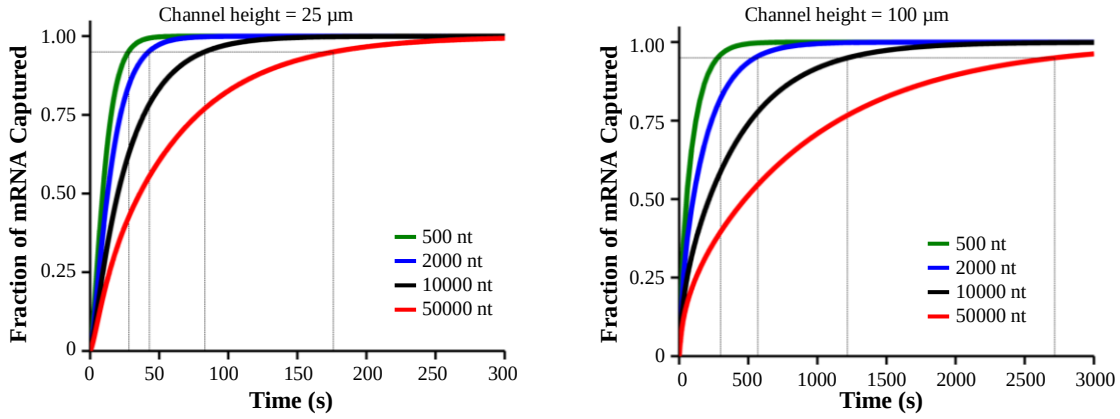


Figure 3.7. Modeling and numerical simulations of mRNA capture kinetics. The kinetics were determined by numerically solving the diffusion equation (equation 1) using finite difference method and a time constant (*1/k*) of 6.1 s for the hybridization reaction. Shown are the capture kinetics of mRNA with lengths of 500 to 50,000 bases in 25 µm (left panel) and 100 µm (right panel) high microfluidic chambers. The horizontal dotted gray line corresponds to 95% completion of total mRNA capture. The effect of mRNA length and channel height on the capture rate is distinctly visible and exaggerated for longer transcripts. In 25 µm chambers, mRNA capture is very rapid. Even for a 50,000 base long mRNA, 95% capture can be achieved in 3 minutes (176 s).

Table 3.2. Numerically simulated mRNA capture kinetics.

Channel Height	RNA Length (nt)	Time to 95% Capture (s)	Diffusion time (s)
25 µm	500	28	7.6
	2000	43	15
	10000	83	34
	50000	176	76
100 µm	500	278	122
	2000	549	245
	10000	1218	548
	50000	2718	1225

3.4.4 Capture and Enumeration of mRNA from Single Mammalian Cells

Two microfluidic devices with 16 modules were fabricated and used for the single-cell mRNA capture experiments. A total of 9 cells were analyzed in 9 modules and 8 modules without any cell captured were used as control. In these experiments, 1 nM of Alexa647-poly(dT)$_{25}$ oligos was added into the lysis/binding buffer to label the polyadenylated mRNA for single-molecule imaging and counting. The images were processed to count the total number of mRNA molecules from the individual cells. Fig. 3.8 shows an example fluorescence image and the objects recognized by our single-molecule selection algorithms overlaid on the image. As observed in Figure 3.8b, our algorithm can segment many adjacent spots into individual molecules or objects with high accuracy and the fluorescence intensities of the spots vary. The variation could be due to several factors. These include the variation in the number of Alexa647-poly(dT)$_{25}$ probe oligos that hybridize to each mRNA molecule due to the intrinsic variation in the length of the poly(A) tails of the different mRNA molecules, signal fluctuation inherent in single-molecule imaging, and possible co-localization of more than one molecule in the diffraction-limited spots (~320 nm resolution). Therefore, the raw counts represent the lower bound of the total poly(A) mRNA molecules in the single cells. To get a better estimate of the counts, an average intensity value of a single mRNA molecule is determined from the single-molecule cluster in a histogram with both the spots initially identified as isolated single molecules and the spots segmented with the watershed

function as shown in Figure 3.9. An adjusted count is determined by dividing the total intensity by the average intensity value of a molecule.

The statistics of the counts are listed in Table 3. The count varies from 30,000 to 155,000 for the adjusted counts. The average mRNA count for all 9 channels containing single-cells is 69,752 with a coefficient of variation (CV) of 66%. In contrast, the adjusted counts of objects for the 8 control modules without any cell, which are primarily due to the non-specific binding of the Alexa647-poly(dT)$_{25}$ probe to the surface and auto-fluorescing impurities on the surface, are consistently low with an average of ~2,500, less than 5% of the average count in the modules with single cells.

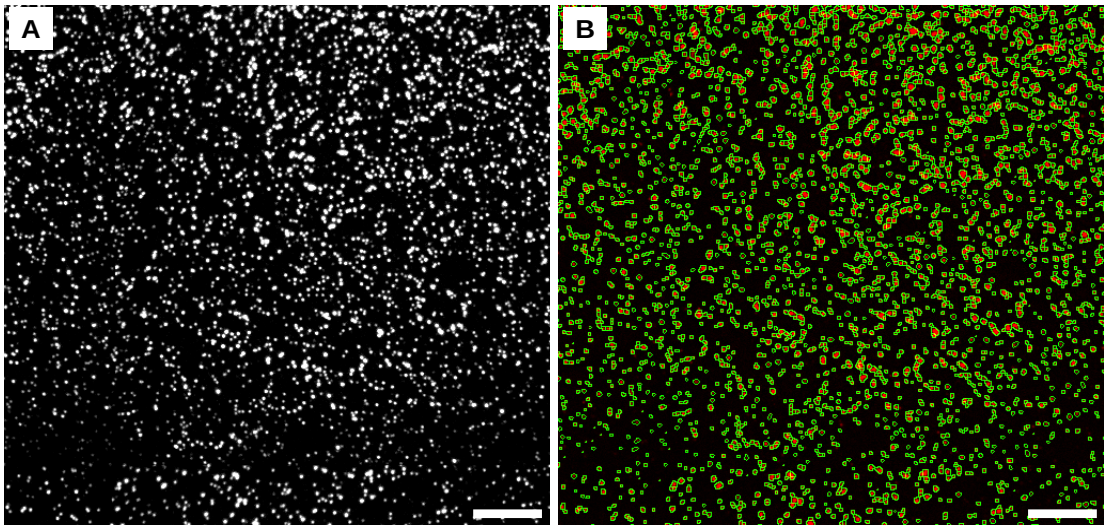


Figure 3.8. Single-molecule fluorescence imaging and counting of captured mRNA from single cells. A. Region from a montage image of an entire surface area of a chamber is shown. The background has been subtracted from the image. The region is close to where cell lysis occurred. B. The same region that has been segmented with the single-molecule selection algorithm. Fluorescence intensity is shown in red and outlines of the molecules are shown in green. Scale bar: 10 μm.

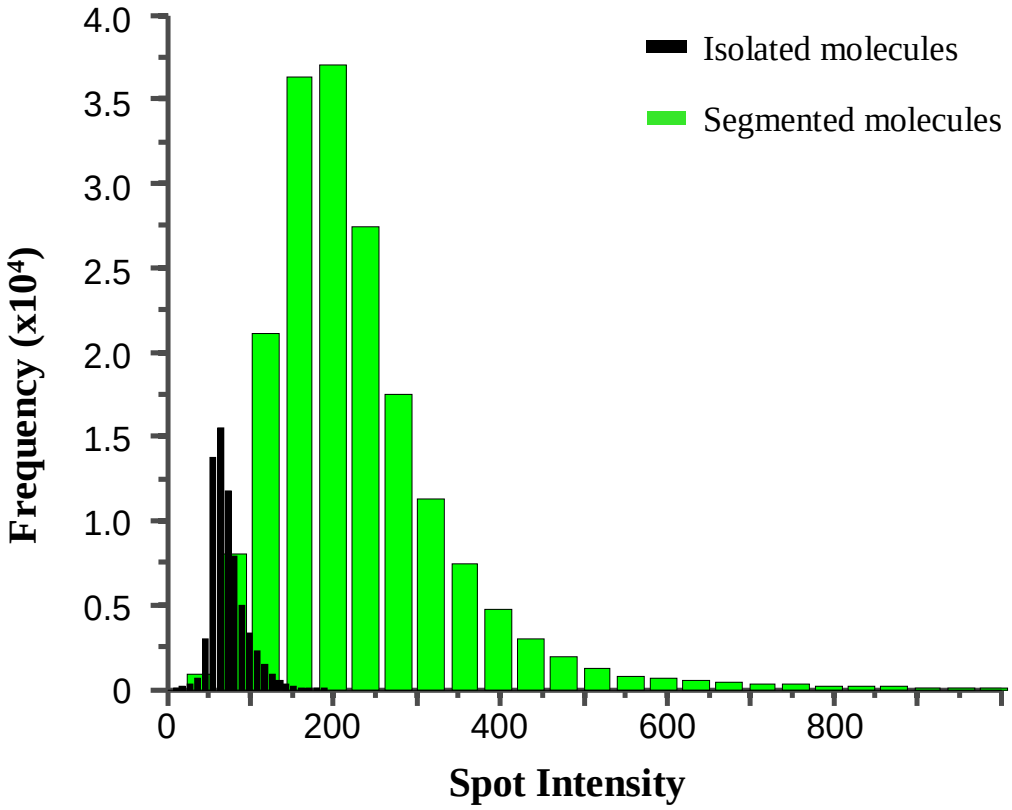


Figure 3.9. Histogram of intensity values for isolated and segmented molecules compiled from all images of mRNA capture experiments with 9 HeLa cells. The mean intensity of spots selected by the image processing algorithm as individual molecules is displayed, whose mean was used to approximate the average intensity of a single molecule. The mean intensities of the spots which were segmented using the watershed function are generally higher, indicating that each spot may contain multiple mRNA molecules.

Table 3.3. Enumeration of mRNA molecules in single cells. Listed are the counts of the total number of polyadenylated mRNA molecules from 9 single cells measured with two separate devices. Listed are both the unadjusted raw counts of discrete objects identified on the image using a segmentation algorithm and the adjusted counts in which the potential number of molecules co-localized on each individual object has been accounted for by estimating the number of molecules based on intensity measurements.

		Device 1			Device 2					
		Cell 1	Cell 2	Cell 3	Cell 4	Cell 5	Cell 6	Cell 7	Cell 8	Cell 9
RNA count	Adjusted (Raw)	40,080 (7,653)	30,406 (13,195)	56,771 (21,979)	136,709 (44,925)	40,192 (16,825)	154,734 (50,400)	37,100 (15,817)	87,256 (34,561)	44,521 (18,344)

3.5 Discussion

At present, no technology is available for direct single-molecule detection and enumeration of all mRNA transcripts from single cells. The microfluidic platform and methods described in this work represent a first attempt towards such a technology. Our strategy is to integrate single-cell capture and downstream processing into individual microfluidic modules connected in serial for scalable parallel analysis of multiple cells, and to selectively functionalize the glass surface in the processing chamber to enable the direct capture by hybridization and single-molecule imaging of polyadenylated mRNA molecules. We have shown experimentally that single cells can be captured with high efficiency using hydrodynamic traps which are tailored to capture cells of specific size and shape based on computational modeling. Critical to our design is the method we have developed for the differential passivation of the PDMS surface with PEG and the functionalization of the glass surface with poly(dT)$_{50}$ via a PEG linker to reduce non-specific binding to both surfaces and to enable the specific capture of polyadenylated mRNA molecules onto the glass for single-molecule fluorescence imaging and counting. The consistent low background from our control experiments with labeled probe but without cells underscores the suitability of our surface chemistry for single-molecule work.

Kinetics measurements of a pure 1700-base mRNA with a 150-base poly(A) tail in a 100 μm high microfluidic channel by real-time single-molecule imaging using TIRF microscopy showed that the mRNA capture fits to a one-dimensional diffusion-hybridization model. The time constant of the hybridization reaction on the surface was

determined to be 6.1 s. mRNA capture reaches almost completion within 10 minutes. The presence of fluorescently-labeled poly$(dT)_{25}$ oligo affects the capture kinetics in an interesting manner. A concentration of poly$(T)_{25}$ below a level that saturates the poly(A) tails of the mRNA molecules (83% saturation in our experiments) decreases the capture kinetics only slightly, but a concentration above saturation level decreases the capture kinetics significantly. Additionally, an increase in the concentration of labeled poly$(dT)_{25}$ oligo reduces the yield of the capture, but does not completely block mRNA capture even when a concentration above saturation is used. We postulate that this likely could be due to the random nature of the hybridization of the poly$(dT)_{25}$ oligos along the longer poly(A) tails, leaving behind some short gapped regions of poly(A) that can still promote hybridization to the longer poly$(dT)_{50}$ oligos on the surface.

For simulated single-cell mRNA capture kinetics measurements in a chamber with the same dimensions of the microfluidic device for single-cell mRNA analysis, our strategy is to add a fluorescently-labeled poly$(dT)_{25}$ oligo into the cell lysate to hybridize to the poly(A) tails of the mRNA molecules so that the capture kinetics can be monitored in real time. We were able to determine a concentration of fluorescently-labeled poly$(dT)_{25}$ that enables the successful measurement of capture kinetics of polyadenylated mRNA from cell lysate equivalent to a single HeLa cell. In the presence of 2.5 nM of Cy3-poly$(dT)_{25}$, the mRNA capture kinetics is found to reach 95% completion by 425 s with a hybridization kinetics time constant of 139 s. This indicates that polyadenylated mRNA molecules released from a single cell can be completely captured within minutes by diffusion and hybridization. Because the hybridization kinetics is found to be slower than that of poly$(A)_{150}$-phi29 mRNA in the presence of saturation level of 1 nM

poly(dT)$_{25}$ oligos (1:3.3), the poly(A) tails of the mRNA from the cells are presumably saturated to a higher degree. If the kinetics of the hybridization reaction follows a similar trend observed in the poly(A)$_{150}$-phi29 mRNA capture experiments, the capture yield would be less than 20%, potentially resulting in a total mRNA count 5 times or more lower than the expected true mRNA count in the absence of the labeled poly(dT)$_{25}$ oligos.

Our computational modeling of mRNA capture provides some useful insights into the mechanism of mRNA capture and how the length of the mRNA and the height of the microfluidic channel affect the capture kinetics. The rate constant of the hybridization reaction, which is dependent on the length of poly(A) tail but mostly independent of the length of the mRNA molecule, can be determined from experimental kinetics data using pure synthetic polyadenylated mRNA, poly(A)$_{150}$-phi29 mRNA. The rate of hybridization reaction is very fast with a time constant of about 6 s. If the chamber height is 25 μm, the capture kinetics is rate-limited to a higher degree by the hybridization reaction for shorter transcripts (500 nt), but rate-limited to a higher degree by diffusion as the length of the mRNA increases. Overall the mRNA capture kinetics is very fast and is influenced modestly by the length of the mRNA molecules (see Fig. 6 and Table 2). Even for a 50000-base long mRNA, 95% capture can be achieved in 3 minutes. However, if the chamber height is increased to 100 μm, the capture kinetics is mostly rate-limited by diffusion and is slowed down significantly, especially for longer mRNA. These findings will be invaluable for the design of microfluidic devices and experiments for single-cell mRNA capture.

More importantly, we have, for the first time, demonstrated the ability to directly capture and count polyadenylated mRNA from single cells using a single microfluidic

device. Total polyadenylated mRNA of 9 cultured HeLa cells were captured and counted. The adjusted counts vary from 30,000 to 155,000 molecules, with an average mRNA count of 69,752 and 66% CV. The large variation could be due to some intrinsic factors, including cell growth stages (we did not synchronize the cell growth) and stochastic nature of gene expression. In fact, the sizes of the individual cells are usually observed to be different. As shown in Figure 4A, the cell trapped at the left module is much larger than the one trapped at the right module. The variation could also be due to factors related to experimental measurements and image processing, for example, the variation of the number of labeled probes hybridized to transcripts with different poly(A) tail length, and limitation of the custom algorithms for single-molecule object recognition and counting.

Due to the use of labeled-poly$(T)_{25}$ oligo probe in the lysis/binding solution to allow for simultaneous capture and detection, the poly(A) tails of some mRNA molecules could potentially be saturated with poly$(dT)_{25}$ depending on the ratio of the labeled poly$(dT)_{25}$,probe to mRNA. This could significantly reduce the yield of capture of the mRNA by the poly$(dT)_{50}$ oligos onto the surface. We have observed this phenomenon in the experiments with synthetic poly$(A)_{150}$-phi29 mRNA in the presence of different ratios of poly$(dT)_{25}$ probe. The simulated single-cell mRNA capture kinetics experiment indicates that the capture yield could be below 20% of that of true complete capture in the absence of the poly$(dT)_{25}$ probe.

It has recently been reported that the median poly(A)-tail length for HeLa cells is 68 bases.[11] Because of such a relatively short median tail length, a large fraction of the mRNA poly(A)-tails could potentially be either fully saturated with the poly$(dT)_{25}$,

preventing the mRNA molecules from being captured onto the surface, or alternatively, saturated by the poly$(T)_{50}$ oligos on the surface, preventing the hybridization of the labeled probe. Both situations would result in uncounted mRNA molecules. In addition, some mRNA molecules may also bind non-specifically to the PDMS walls despite the passivation of the surface with PEG. All these factors could contribute to a significantly lower total mRNA count than the true count.

Given the uncertainties due to the limitations described above and the unavailability of a benchmark on absolute counting of total polyadenylated RNA molecules in single mammalian cells in the literature, further work will be required to improve and assess our technology in terms of absolute counting of mRNA of single cells. For example, single cells can be lysed in a lysis/binding buffer solution without a labeled poly(dT) probe oligo and the mRNA captured onto the surface. The captured mRNA molecules can then be labeled by reverse transcription using fluorescently-labeled nucleotides and subsequently detected by single-molecule imaging. To minimize the co-localization of the molecules, the total surface area of the processing chamber can be optimized and captured mRNA can also be denatured and rehybridized to the poly$(dT)_{50}$ oligos to redistribute. Both strategies may allow a more even distribution the molecules on the surface for more accurate counting. Our platform is also designed to be compatible with technologies for direct single-molecule nucleic acid detection either by hybridization using specific probes,[12-16] or direct on-chip single-molecule sequencing.[17-19] Ultimately, we want to sequence the mRNA molecule or the reverse transcribed cDNA molecules by adapting and developing methods for single-molecule RNA sequencing or DNA

sequencing. We are in the process of developing this capability in our laboratory, as is presented in the following section.

3.6 Progress Towards Total On-Chip Transcriptome Analysis

3.6.1 Implementation of Single-Molecule Sequencing

Towards the goal of achieving on-chip total mRNA analysis, a single-molecule sequencing-by-synthesis (smSBS) methodology was developed utilizing the TruSeq™ sequencing chemistry from Illumina. This regent set contains a mixture of the four nucleotides containing unique fluorescent labels and 3'-OH reversible terminators, engineered polymerase for the efficient incorporation of these nucleotides, cleavage reagent for simultaneous removal for the fluorescent moiety and 3'-OH terminator, and associated buffers. The sequencing process follows the scheme presented in Figure 3.10. A synthetic DNA template is hybridized to poly(dT) covalently bound to a coverslip surface. The complementary strand is then synthesized using a DNA polymerase. The original template is denatured and washed off before a primer is hybridized to the immobilized strand. The surface is now prepared for cycles of incorporation and cleavage reactions in order to read the sequence.

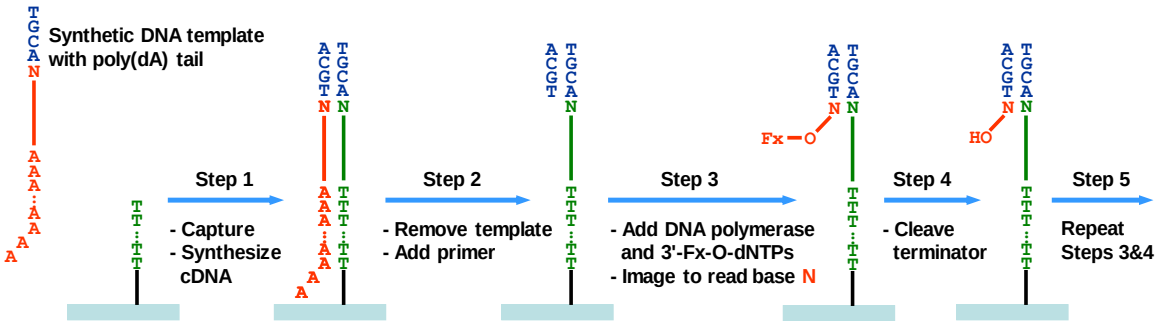


Figure 3.10. Procedure of for sequencing polyadenylated DNA templates. Templates are first covalently copied to the surface by DNA polymerase before successive rounds of fluorescent nucleotide incorporation and cleavage are performed to read the sequence of each single molecule.

In order to select a set of fluorescence emission filters to discriminate between the unique fluorophores on the bases , the absorption and emission spectra were determined. As the reagent provided as a mixture of all nucleotides, four single-base extension reactions were performed in solution with different primers designed to incorporate a specific base. Upon completion of the reaction, ethanol precipitation was performed to separate the DNA containing the incorporated nucleotide with a specific fluorophore from the bulk mixture of unincorporated nucleotides. For each reaction, a spectrophotometer was used to assess the absorbance spectrum and a spectrofluorometer was used to determine the emission spectrum. This enabled the design of a filter set that provides a distinct fluorescence signature for each base. The spectra of all four bases along with the emission detection bands and laser excitation wavelengths are displayed in Figure 3.11. The spectra of the fluorophores resemble those of Cy3, Cy3.5, Cy5, and Cy5.5; the fluorescence channels are termed accordingly. The emission filter center wavelength and bandwidth, along with the theoretical amount of photons collected using

these filters to image each nucleotide, are displayed in Table 3.4. The theoretical photon collections were calculated using a fluorescence microscopy tool developed in our lab by Dr. Eric Roller. It accounts for factors such as the collection efficiency of the objective lens, quantum efficiency of the EMCCD camera, extinction coefficient and quantum yield of the fluorophore, wavelength and power of the excitation source, and bandpass wavelength of the emission filter. For these calculations, the excitation source was a 532 nm laser for the Cy3 and Cy3.5 channels, and a 642 nm laser for the Cy5 and Cy5.5 channels, both at 10mW. A second exposure setting was used. Figure 3.12 shows the theoretical amount of photons collected for each base plotted as intensity values normalized to the channel with the highest value. This figure visually demonstrates that each nucleotide has a unique fluorescence signature and the four bases can be unambiguously identified.

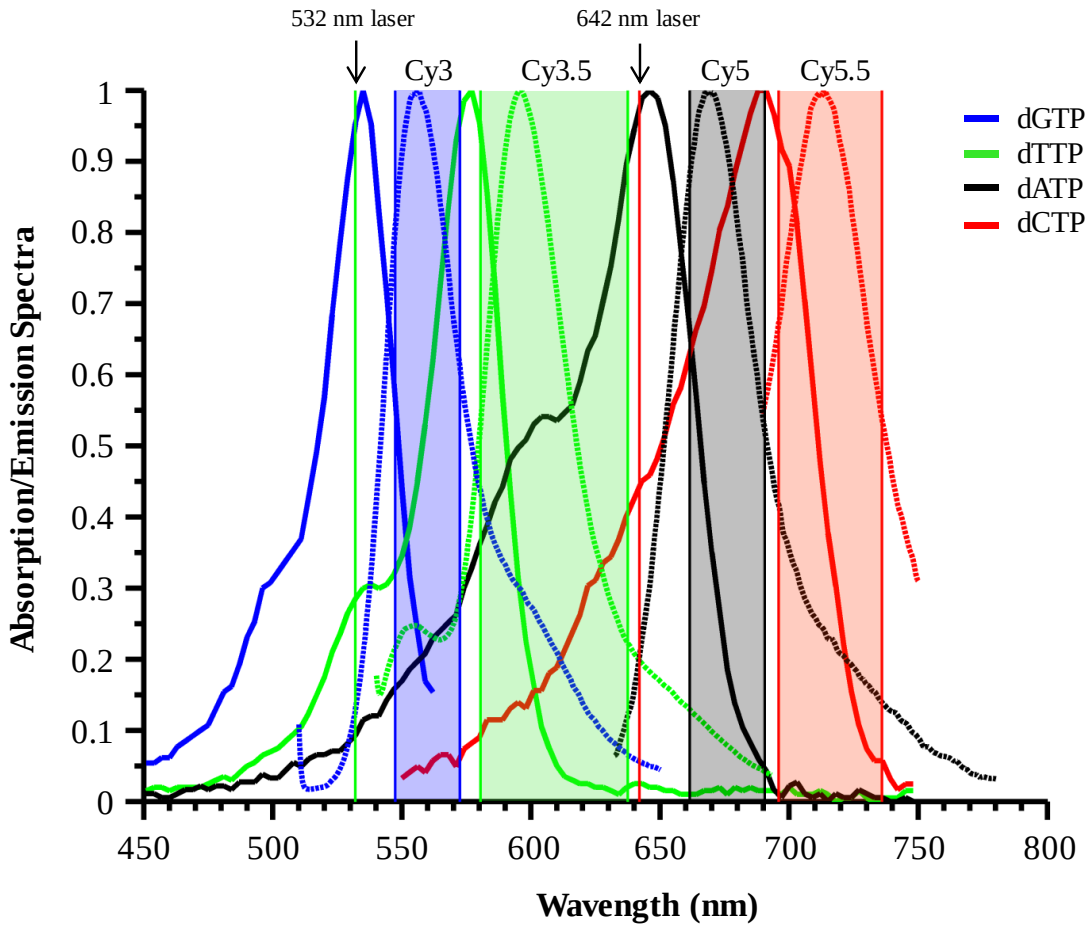


Figure 3.11. Absorption and Emission Spectra of Fluorescently-labeled Nucleotide Set with Filter Selection. Absorption spectra are shown as solid lines. Emission spectra are shown as dashed lines. The bandpass wavelengths of the emission filters, chosen to provide maximum distinction between each fluorophore, are overlaid. The laser sources used for excitation are depicted. A 532 nm laser was used as excitation for the Cy3 and Cy3.5 channels, while the 642 nm laser was used as excitation for the Cy5 and Cy5.5 channels.

Table 3.4. Fluorescence channel emission filter selection and theoretical amount of photons collected for each base. Photon collection calculations were based on the specifications of our imaging system using 10 mW excitation power and 1 second exposure time.

	Emission Filter		Theoretical Photons Collected			
	Center Wavelength (nm)	Bandwidth (nm)	dATP	dCTP	dGTP	dTTP
Cy3	560	25	0	0	1843	127
Cy3.5	609	57	11	0	965	635
Cy5	679	29	1920	31	0	5
Cy5.5	716	40	795	594	0	0

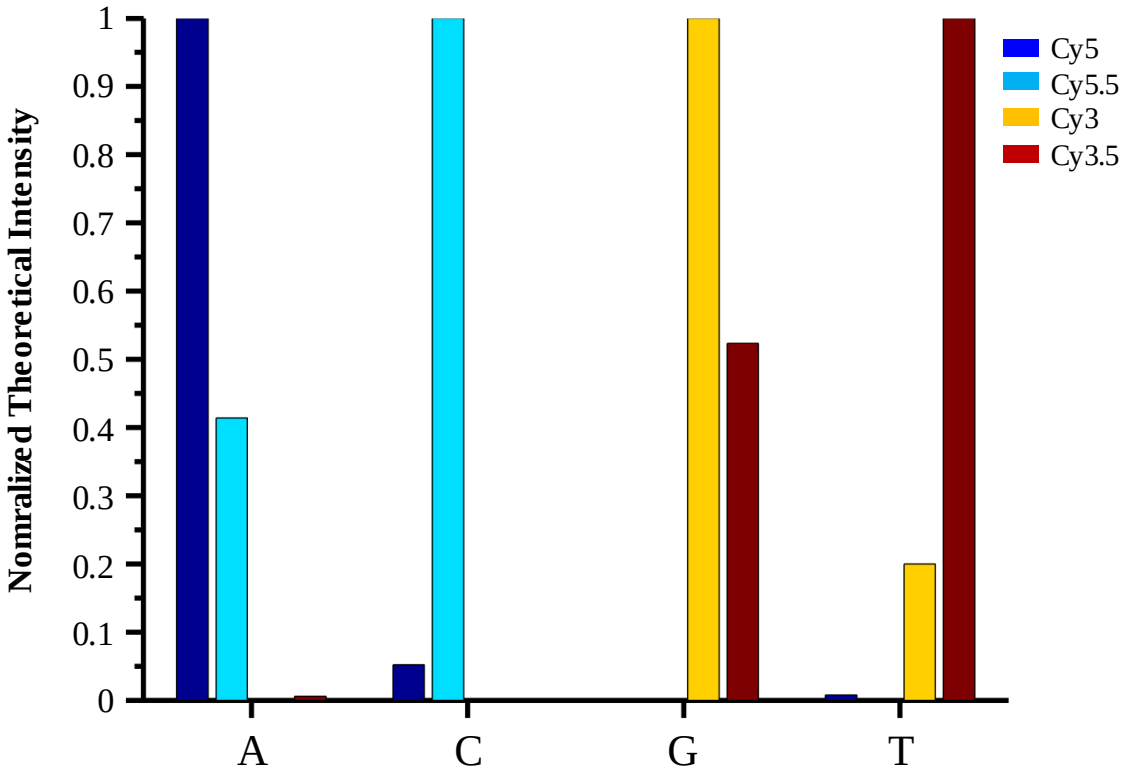


Figure 3.12. Normalized Theoretical Fluorescence Detection of four bases. The theoretical amount of photons collected from the nucleotides in each fluorescence channel is shown as intensity values normalized to the channel with the highest value. This depicts each base's distinct fluorescence signature that enables their unambiguous identification by single-molecule imaging.

3.6.2 Experimental Procedure

A synthetic template containing 55 bases of specific sequence and about 200mer poly(dA) tail was hybridized to the poly(dT) surface at a concentration of 50 pM and converted to cDNA. After denaturing the template, a 25mer Cy3-labeled primer was hybridized to the cDNA such that the subsequent four bases to be incorporated were TCGA. After imaging the primers to identify their locations, the Cy3 fluorophores were

photobleached so that they did not interfere with the imaging of the labeled nucleotides during sequencing.

Sequencing cycles of incorporation and cleavage were performed for 4 minutes at 45 °C. Imaging was performed at 25 °C. The incorporation reactions were performed using a nucleotide concentration at 25% of the standard nucleotide concentration after a brief wash with incorporation buffer. For the cleavage reaction, the solution was exchanged after 2 minutes to refresh the cleavage reagent. After each step, the channel was washed with 3 ml of wash buffer (100 mM NaCl, 10mM Tris-Cl, 1 mM EDTA and 0.02% Triton X-100, pH 8). Immediately after washing, images were taken in four different fluorescence emission filters optimized to distinguish the bases.

After manually aligning the images to account for the drift of the FOV, the images were analyzed using the primer image as a mask to serve as a selection tool for which pixels to analyze. By calculating the mean intensity within each spot in each fluorescence channel for all cycles the sequence readout is obtained. Two selected spots, which represent the correct sequence, are shown for both their integrated fluorescence intensity in each cycle and a cropped window around the spot in Figures 3.13 through 3.16.

Table 3.5. Sequences of templates and primers used for the smSBS experiments. A ~200mer poly(dA) tail is added to the smTemplate oligo to enable efficient capture and conversion into complementary strand on a poly(dT) surface. The smProbe oligo serves as a primer for sequencing and contains a fluorescent label to identify viable templates for sequencing. The bases shown underlined and in bold represent the sequence that is read through four cycles of smSBS. Because the template is converted to the complementary strand before sequencing, the smTemplate and smProbe have identical sequences.

Oligo Name	Sequence
smTemplate	5'- ATG TCC TAT CCT CAG CTG TGA GCA TGA CTG ACT TCG TAC ATG ACT GAT GG**T CGA** TAA AAA AAA AAA AAA AAA AAA AAA AA -3'
smProbe	5'-Cy3-GACTGA CTT CGT ACATGA CTG ATG G -3'

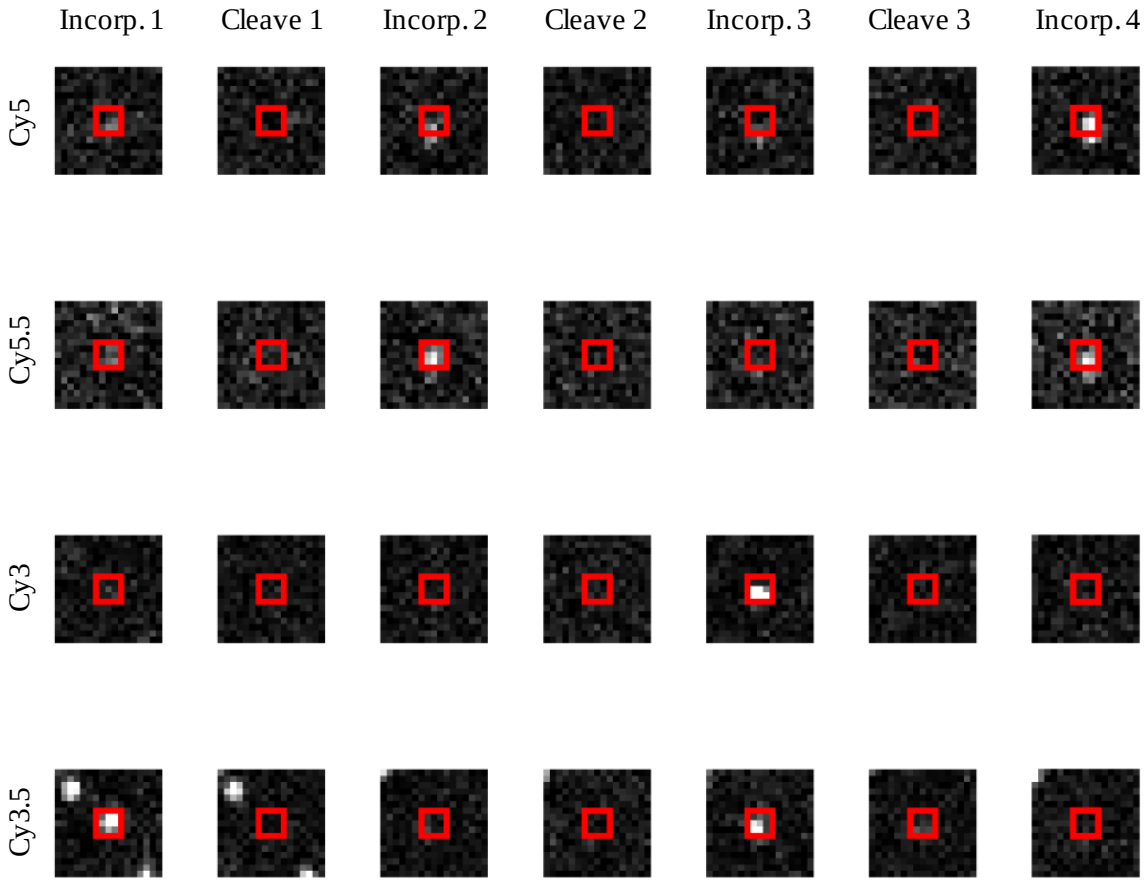


Figure 3.13. Images of a cropped region containing a sequenced template captured in four fluorescence channels over four cycles of smSBS. The 3x3 array inside the red box represents the selected pixels used for image analysis. It is evident that the fluorophores are detected with strong signal over that of background.

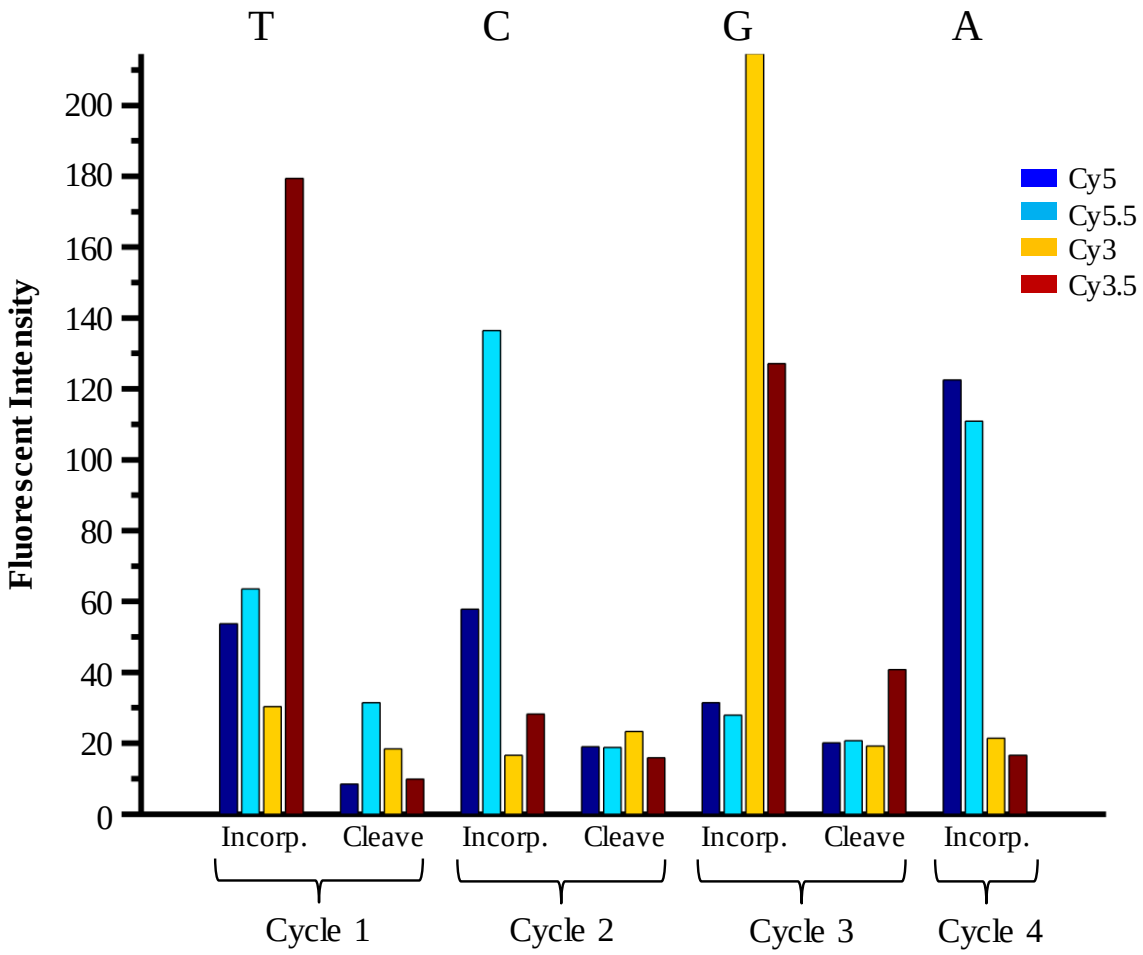


Figure 3.14. Fluorescence intensity profiles of average pixel intensity from 3 x 3 array of selected region for each channel over four cycles of incorporation and cleavage. From the plot, it is clearly discernable as to which base has been incorporated, as indicated at the top of the figure. While the intensity profiles do not precisely follow the theoretical predictions, they are unique and distinct enough to implement a base-calling algorithm.

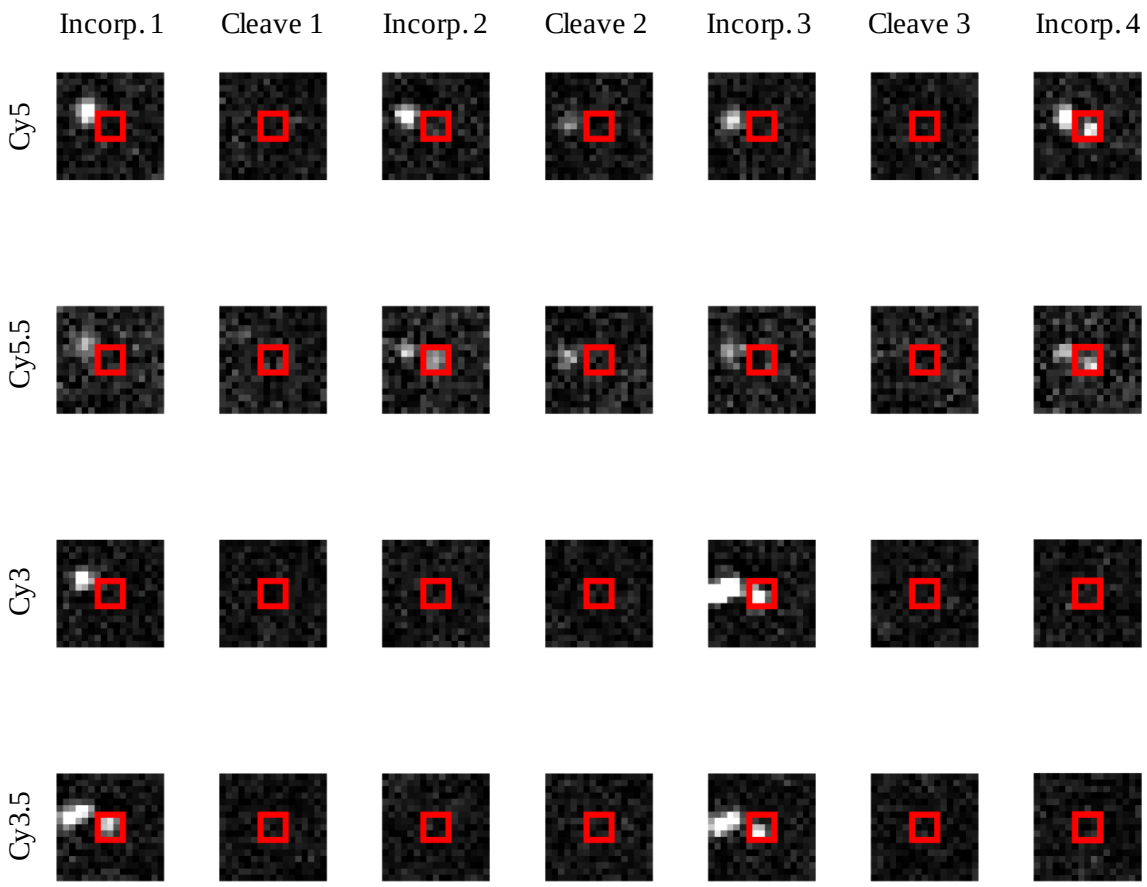


Figure 3.15. Images of a seconds cropped region containing a sequenced template captured in four fluorescence channels over four cycles of smSBS. The selected area demonstrates the ability to exclude adjacent molecules from the analysis window.

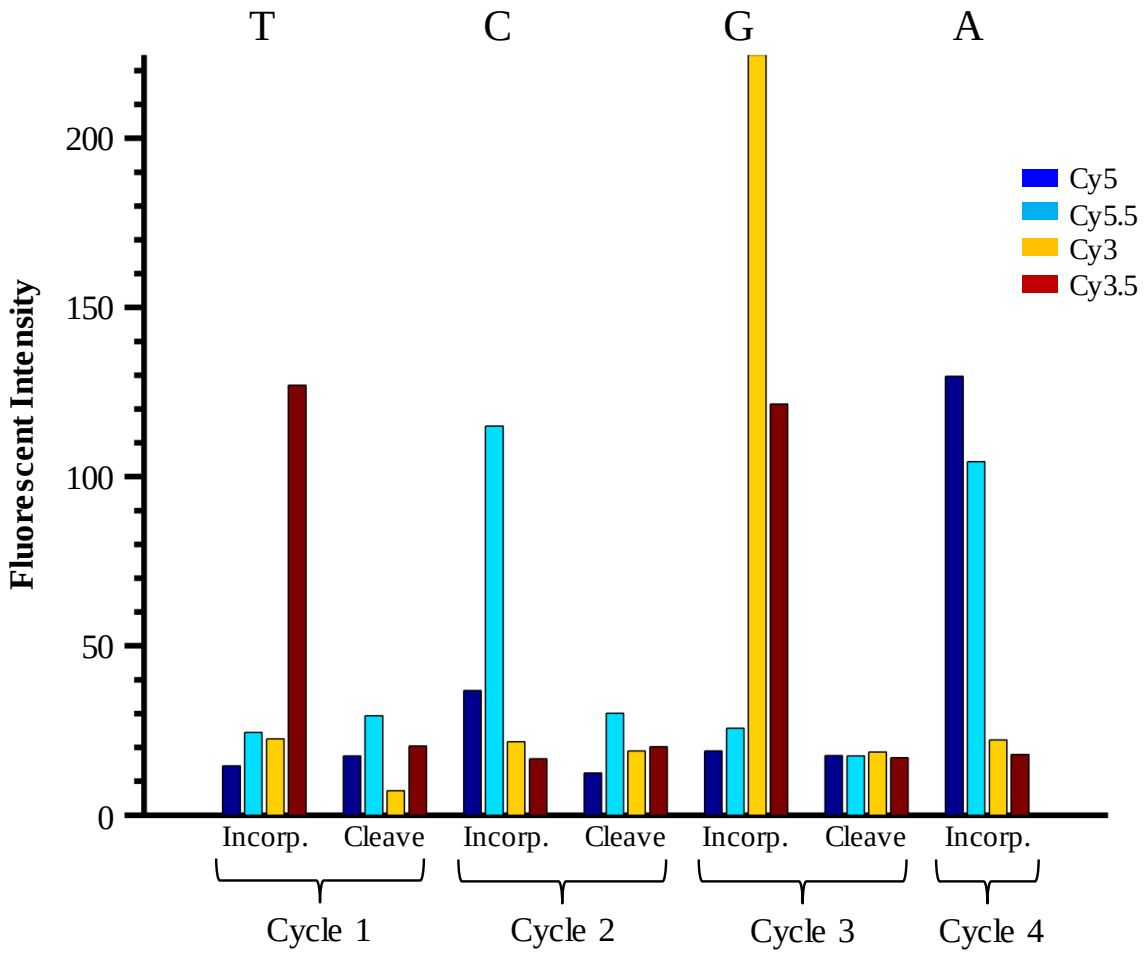


Figure 3.16. Fluorescence intensity profiles of average pixel intensity from 3 x 3 array of a second selected region for each channel over four cycles of incorporation and cleavage. The intensity values reveal that the correct sequence has been read and adjacent fluorophores do not contribute to the measurements if a properly aligned pixel selection area is used.

3.7 Conclusion

Using a 1700-base synthetic fluorescently-labeled mRNA with 150-base poly(A) tail, we have demonstrated convincingly that the mRNA can be completely captured within minutes by diffusion and hybridization to poly(dT)$_{50}$ oligos selectively

functionalized on glass the surface in a 25 μm high microfluidic chamber. Our modeling and numerical simulations of mRNA capture kinetics have also demonstrated that the hybridization reaction is very rapid with a time constant of ~6 s for mRNA with a 150-base poly(A) tail, and that short mRNA can be completely captured in less than one minute and even 50000-base long mRNA can be captured to almost completion within 3 minutes in a 25 μm high microfluidic chamber. By adding fluorescently-labeled poly$(dT)_{25}$ oligos into the HeLa cell lysate, we were able to simulate the capture kinetics of mRNA from single cells. Despite the reduced effective length or saturation of the poly(A) tails due to the hybridization of the poly$(dT)_{25}$ probe, the capture kinetics is still relatively rapid, with 95% capture in less than 7 minutes.

Using two separate integrated devices, we captured and enumerated polyadenylated mRNA molecules from 9 individual HeLa cells. The total counts of the polyadenylated mRNA from the individual cells are found to vary significantly from one cell to the other with a very large coefficient of variation. The total mRNA counts could be much lower than the true counts due to the use of the labeled poly$(dT)_{25}$ for detection and other factors. Further work will be required to improve and quantify capture efficiency and count accuracy, and potentially to enable identification of the mRNA by direct single-molecule sequencing.[17-19] Overall, our microfluidic devices and methods are designed to enable the entire workflow from single cell isolation to mRNA capture and single-molecule imaging to be performed in a single device. The work described here has demonstrated these capabilities and has set the stage for the direct digital whole transcriptome profiling of single cells using a single integrated microfluidic device and single-molecule DNA sequencing.

3.8 Acknowledgements

We thank Dr. Eric Roller for his assistance in the setup and implementation of TIRF microscopy equipment and software. We also thank Ximin Chen for her assistance in maintaining the cell culture. This work was partially funded by the NIH (grants R01HG004804, R21HG005096 and 3R21HG005096-02S1).

This chapter, in part, is a reproduction of the material as it appears in: Walsh, M.T., Hsiao, A.P., Lee, H.S., Liu, Z., Huang, X. (2015) Capture and enumeration of mRNA transcripts from single cells using a microfluidic device. *Lab Chip*, **15**, 2968-2980; copyright 2015 Royal Chemical Society. Used with permission. The dissertation author was the primary investigator and author of this paper.

3.9 References

1. Tan, W.H. and Takeuchi, S. (2007) A trap-and-release integrated microfluidic system for dynamic microarray applications. *Proc Natl Acad Sci U S A*, **104**, 1146-1151.

2. Unger, M.A., Chou, H.P., Thorsen, T., Scherer, A. and Quake, S.R. (2000) Monolithic microfabricated valves and pumps by multilayer soft lithography. *Science*, **288**, 113-116.

3. Lee, H.S., Chu, W.K., Zhang, K. and Huang, X. (2013) Microfluidic devices with permeable polymer barriers for capture and transport of biomolecules and cells. *Lab Chip*, **13**, 3389-3397.

4. Hsiao, A.P., Barbee, K.D. and Huang, X. (2010) Microfluidic Device for Capture and Isolation of Single Cells. *Proc Soc Photo Opt Instrum Eng*, **7759**, pii: 77590W_77591, doi: 77510.71117/77512.861563.

5. Nae Yoon Lee, B.H.C. (2009) Novel Poly(dimethylsiloxane) Bonding Strategy via Room Temperature "Chemical Gluing". *Langmuir*, **25**, 7.

6. Barbee, K.D., Chandrangsu, M. and Huang, X. (2011) Fabrication of DNA polymer brush arrays by destructive micropatterning and rolling-circle amplification. *Macromol Biosci*, **11**, 607-617.

7. Tinland, B., Pluen, A., Sturm, J. and Weill, G. (1997) Persistence length of single-stranded DNA. *Macromolecules*, **30**, 5763-5765.

8. Chen, H., Meisburger, S.P., Pabit, S.A., Sutton, J.L., Webb, W.W. and Pollack, L. (2012) Ionic strength-dependent persistence lengths of single-stranded RNA and DNA. *Proc Natl Acad Sci U S A*, **109**, 799-804.

9. Wetmur, J.G. and Davidson, N. (1968) Kinetics of renaturation of DNA. *J Mol Biol*, **31**, 349-370.

10. Sommer, S.S. and Cohen, J.E. (1980) The size distributions of proteins, mRNA, and nuclear RNA. *J Mol Evol*, **15**, 37-57.

11. Subtelny, A.O., Eichhorn, S.W., Chen, G.R., Sive, H. and Bartel, D.P. (2014) Poly(A)-tail profiling reveals an embryonic switch in translational control. *Nature*, **508**, 66-71.

12. Zhou, Y., Calciano, M., Hamann, S., Leamon, J.H., Strugnell, T., Christian, M.W. and Lizardi, P.M. (2001) In situ detection of messenger RNA using digoxigenin-labeled oligonucleotides and rolling circle amplification. *Exp Mol Pathol*, **70**, 281-288.

13. Levsky, J.M., Shenoy, S.M., Pezo, R.C. and Singer, R.H. (2002) Single-cell gene expression profiling. *Science*, **297**, 836-840.

14. Geiss, G.K., Bumgarner, R.E., Birditt, B., Dahl, T., Dowidar, N., Dunaway, D.L., Fell, H.P., Ferree, S., George, R.D., Grogan, T., James, J.J., Maysuria, M., Mitton, J.D., Oliveri, P., Osborn, J.L., Peng, T., Ratcliffe, A.L., Webster, P.J., Davidson, E.H., Hood, L. and Dimitrov, K. (2008) Direct multiplexed measurement of gene expression with color-coded probe pairs. *Nat Biotechnol*, **26**, 317-325.

15. Larsson, C., Grundberg, I., Soderberg, O. and Nilsson, M. (2010) In situ detection and genotyping of individual mRNA molecules. *Nat Methods*, **7**, 395-397.

16. Hocine, S., Raymond, P., Zenklusen, D., Chao, J.A. and Singer, R.H. (2013) Single-molecule analysis of gene expression using two-color RNA labeling in live yeast. *Nat Methods*, **10**, 119-121.

17. Harris, T.D., Buzby, P.R., Babcock, H., Beer, E., Bowers, J., Braslavsky, I., Causey, M., Colonell, J., Dimeo, J., Efcavitch, J.W., Giladi, E., Gill, J., Healy, J.,

Jarosz, M., Lapen, D., Moulton, K., Quake, S.R., Steinmann, K., Thayer, E., Tyurina, A., Ward, R., Weiss, H. and Xie, Z. (2008) Single-molecule DNA sequencing of a viral genome. *Science*, **320**, 106-109.

18. Ozsolak, F. and Milos, P.M. (2011) Transcriptome profiling using single-molecule direct RNA sequencing. *Methods Mol Biol*, **733**, 51-61.

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

4 TECHNOLOGIES TO ENABLE REAL-TIME DNA SEQUENCING (READS)

4.1 Abstract

This chapter describes our approaches and efforts towards the development of technologies for direct real-time monitoring of polymerase enzymatic reactions and a new sequencing technology. DNA polymerase was labeled with a pair of fluorescent labels to enable Förster resonance energy transfer (FRET) measurements. The dynamic conformational changes accompanying the nucleotide incorporation process are monitored in real-time using single-molecule FRET imaging to dissect the molecular mechanism of the DNA synthesis process. We aim to use the FRET-labeled DNA polymerase as the sensor for direct single-molecule real-time DNA sequencing, which is called READS. Described in this chapter are our initial efforts in the development of the various technical components, including engineered FRET-labeled DNA polymerases and imaging system, and the preliminary experiments to observe native and labeled nucleotide incorporation.

4.2 Introduction

The ultimate DNA sequencing technology would be capable of discerning long, continuous lengths of nucleotide sequences and epigenomic modifications in an accurate, rapid, high-throughput, and cost-effective manner. Current methods fail to address all of these requirements in one or more ways. Traditional next-generation sequencing achieves great through-put by using highly advanced instruments and reagents. While these innovations have driven down costs, the technology inherently generates short reads necessitating complex informatics processing,[1] and the ability to detect epigenomic modifications requires complex sample preparation.[2] Other technologies, such as nanopores or Pacific Biosciences sequencing, allow long reads to be generated[3,4] with some ability to directly detect epigenetic signatures.[5,6] However, these methods require sophisticated instrumentation that lacks a large degree of scalability, reducing the potential throughput of these systems. By directly labeling DNA polymerase with a FRET pair, it is feasible to observe the structural dynamics of the enzyme during nucleotide incorporation. Due to differences in the kinetic parameters between each base type,[7] including epigenetically modified bases,[5] it may be possible to extract sequencing information by observing variations in the enzyme's conformation.

Previous studies have utilized smFRET for a wide range of DNA polymerase behavioral observations. The most common method involves immobilizing primed DNA templates labeled with the donor fluorophore and monitoring the interaction with DNA polymerase labeled with the acceptor fluorophore. This strategy can provide insights into the enzymatic behavior such as incorporation rate[8] and nucleotide discrimination.[9,10] A

more sophisticated approach has been demonstrated in which both the donor and acceptor fluorophore are positioned on the fingers and palm domain of the enzyme to allow the real-time observation of conformational dynamics.[11] Confocal microscopy was used to measure the FRET efficiency of unliganded polymerase, polymerase-DNA binary complex, and polymerase-DNA-dNTP ternary complex. This method revealed the structural changes in the enzyme depending on its substrate binding state. Due to the utilization of confocal imaging of freely diffusing molecules, this polymerase study was limited to observing the equilibrium binding reactions using a dideoxy-terminated primer.

By immobilizing the DNA on the surface on a glass coverslip, the continuous monitoring on thousands of DNA polymerases during incorporation can be performed using TIRF microscopy. The main challenges presented by this strategy are to create a modified DNA polymerase with precise control of the position of the fluorescent labels and ability to optically monitor the multiple fluorescence signals from each enzyme during polymerization at high speed with high signal-to-noise. This chapter describes our work towards addressing these challenges.

4.3 Protein Engineering Using Solid-Phase Cyclic *in vitro* Translation

Current methods for protein labeling rely on reacting with fluorophores activated with N-hydroxysuccinimide (NHS) or maleimide with available amine or thiol groups, respectively, on the protein. These methods can be fairly specifically targeted if the protein contains only one cysteine for labeling using maleimide. In many applications,

multiple labels may be required for FRET experiments. In these cases, site-specific labeling would not be easy or possible. By performing cyclic *in vitro* translation using surface-bound mRNA, it is possible to precisely control protein synthesis by limiting the aminoacyl-tRNAs present in the reaction mix. This strategy is outlined in Figure 4.1. Mutagenesis is first used to convert the codons encoding the target residues for labeling to a codon that encodes for cysteine. The mutated gene is then used to prepare polyadenylated mRNA by *in vitro* translation. The mRNA molecules are immobilized onto a solid support by hybridization. Upon *in vitro* translation initiation, the ribosomes will synthesize the protein until it reaches the codon for which the specific tRNA is not present. At this point the current reaction mix is exchanged for a solution containing only the specific tRNA charged with a modified amino acid. This process is repeated for as many rounds are necessary to create the desired modified protein. The following section outlines the steps we took towards implementing this procedure. Due to difficulties in fully realizing this method, we also applied efforts towards chemically labeling DNA polymerase. We chose to use Bst DNA polymerase (Large Fragment) because it contained one native cysteine which could be mutated away without hindering enzymatic functionality. A gene was synthesized with the native cysteine removed and two native amino acids mutated to cysteine, one on the palm domain and the other on the finger domain. This enabled the *in vitro* transcription/translation, and in vivo expression of the mutant polymerases in E. coli for chemical labeling.

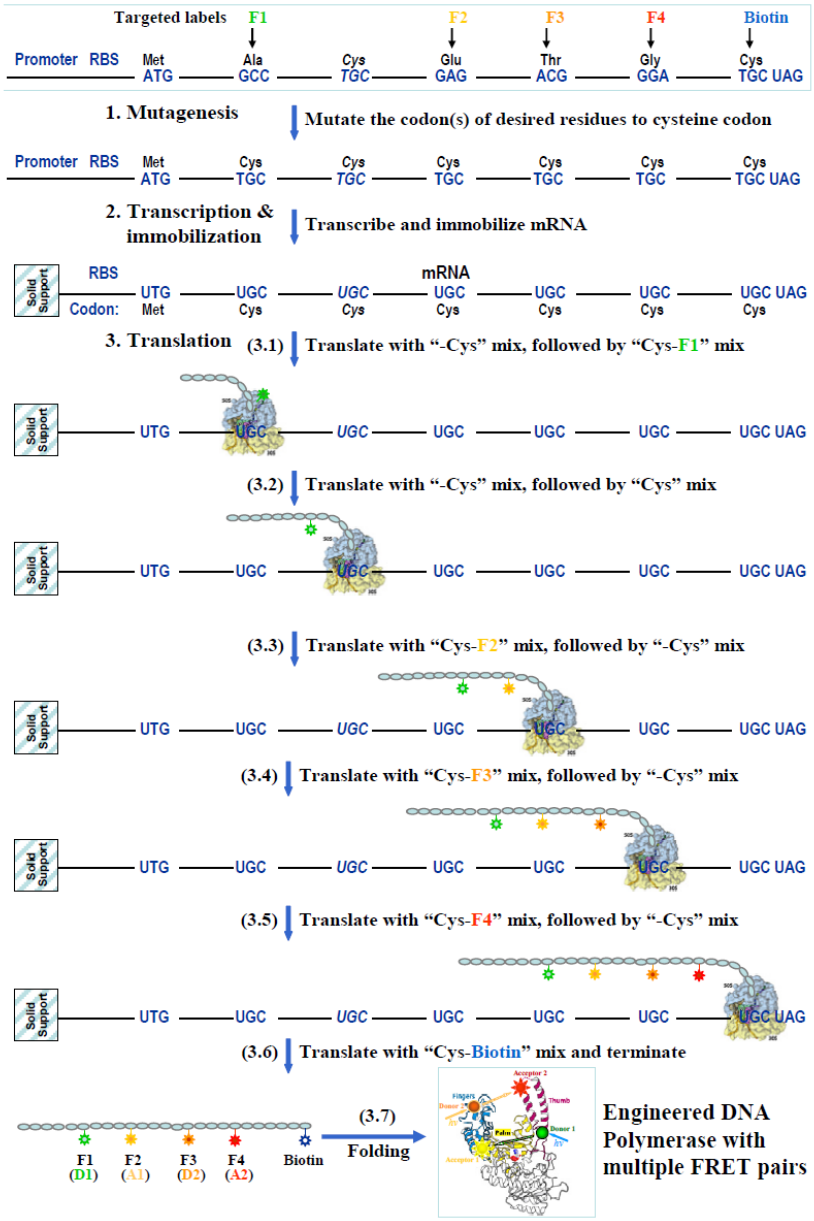


Figure 4.1. Strategy for engineering protein multiple FRET pairs by cyclic *in vitro* translation on surface-bound mRNA. The method consists of three steps. (1) Codons that encode the target residues for labeling are mutated to a codon encoding for a cysteine residue (TGC). (2) mRNA molecules for the mutated gene are generated by in vitro transcription. (3) Multiple cycles of solid-phase in vitro translation are performed with specific reaction mixes to enable the incorporation of particular labels at the desired residues.

4.3.1 Functionalization of Glass Fibers for a Solid-Phase Binding of mRNA

Pyrex fiber glass roving wool was purchased from Corning Inc. (No. 9989 glass, ~8 µm diameter). Fibers were cut into pieces approximately 5 cm long and baked overnight at 600 °C in Pyrex beakers to remove any organic residues. After cooling to room temperature, the fibers were cleaned and etched by an 8 hour incubation at 50 °C in 30% (w/v) KOH in 1:1 mixture of water and methanol. Sonication was performed periodically to enhance the cleaning process and to aid in the removal of bubbles. After a thorough rinse with MilliQ water, the fibers were dried before being packed in a polyfluoroalkene (PFA) tube (Cole-Parmer, 5/32 inch inner diameter, ¼ inch outer diameter) to a length of 1 cm and compressed between two custom-machined Teflon stoppers each containing a stainless steel frit (IDEX Health & Science, 10 µm porosity) insert at the end to block the passage of any fibers. The stoppers have ridged edges to securely hold them in position and a hole through the center with a threaded port for connecting inlet and outlet lines. An assembled column is depicted in Figure 4.2A. To enable reliable and reproducible reagent delivery, an automated system consisting of a 9-port syringe pump connected to an additional nine-channel rotary valve was used. To provide thermal control, the column is encased within an aluminum block which is in direct contact to a thermoelectric module. The total system is shown in Figure 4.2B.

Using this fluidic setup, the fibers were chemically functionalized. First, silanol generation were performed by incubating 3.16 M nitric acid for at least one day. After washing with 9 ml of water, silanization was performed by first exchanging the water for

200 proof ethanol (Gold Shield) then flowing in 1 ml of a 2% APTES solution in ethanol. Following a 30-minute incubation, the fibers were washed with 9 ml of ethanol. To complete the silane binding, the fibers were dried by flowing through argon gas at 30 PSI for 5 minutes, baked at 100 °C and placed under vacuum for 20 minutes.

Short homobifunctional carboxymethyl poly(ethylene glycol) (Lysan Bio, 1000 g/mol) was conjugated to the amine silane using a carbodiimide coupling chemistry. A solution of 1 mM carboxymethyl PEG, 10 mM EDC and 5 mM TEA was flowed into the column and incubated for 2 hours at room temperature. Amine-reactive PEG was used to cap potential remaining amine silane groups by reacting with a solution of 5 mM N-hydroxysuccinimide PEG_4 (Pierce), 10 mM EDC, and 5 mM TEA in DMF for 1 hour. Subsequently, the column was washed with ethanol and exchanged with 100 mM MES buffer, pH 5 containing 0.02% Triton X-100. Poly(dT)50 were then coated to the carboylated PEG surface by incubating 20 µM NH2-poly(dT) with 16 mM EDC in MES buffer in the column for 2 hours.

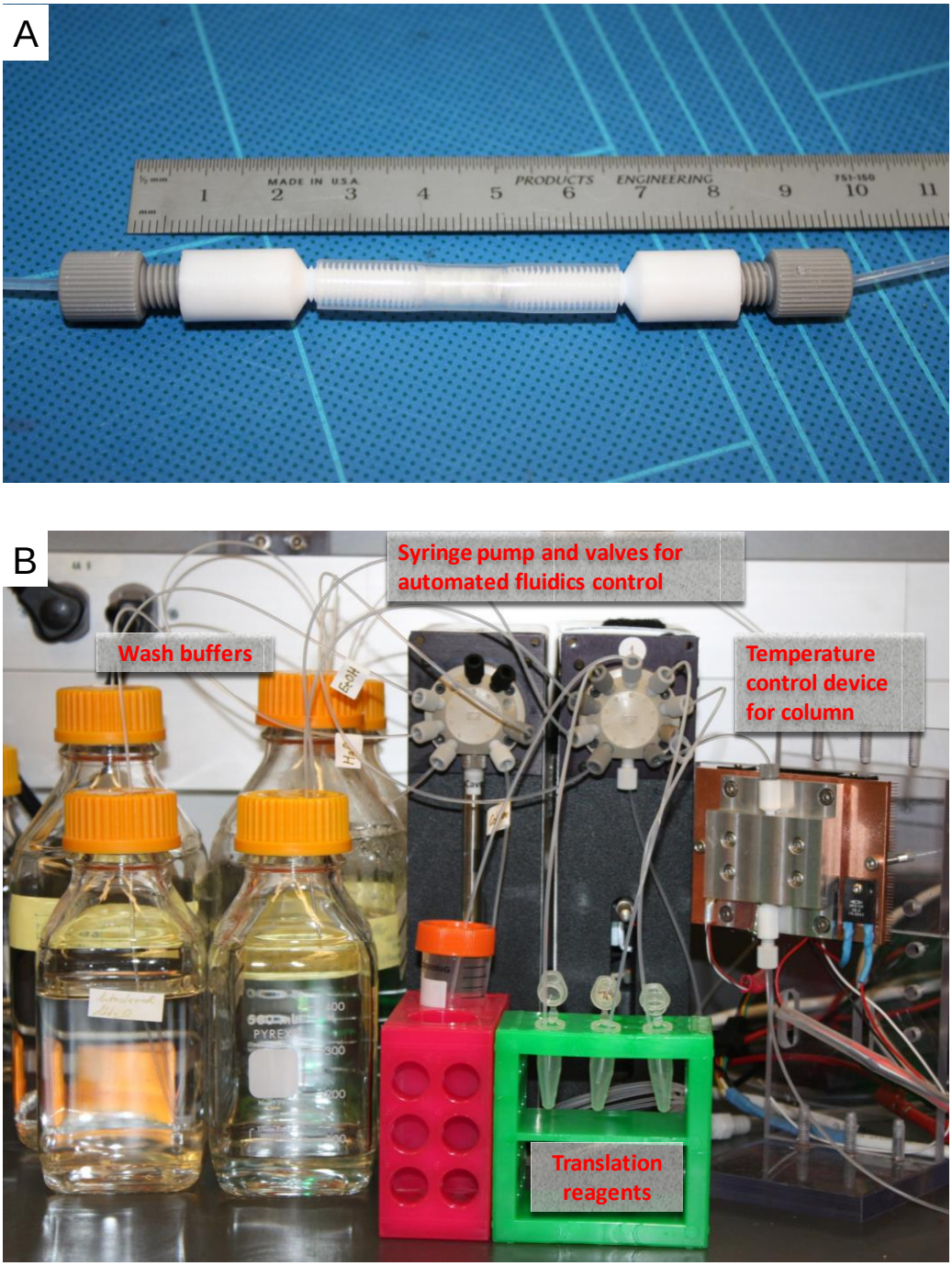


Figure 4.2. Experimental Setup for *in vitro* Translation. A. Quartz glass fibers are packed inside a custom-made column. B. The entire fluidic system including automated syringe-pumps and valve connected to the column. The column is mounted inside a temperature control device.

4.3.2 mRNA Hybridization in Column

Hybridization of mRNA to the surface was performed by passing 250 µl of 100 nM mRNA with a 150mer poly(A) tail in a hybridization buffer (200 mM NaCl, 15 mM Tris-Cl, 0.05% SDS, 5% PEG6000, pH 8) containing 0.13 units of Murine RNase inhibitor (NEB) through the column slowly over 15 minutes while holding the thermal block at 50 °C. The poly(A)-tailed mRNA was prepared by *in vitro* transcription from a plasmid containing a phi29 DNA polymerase gene. Subsequently, the temperature was linearly ramped to 25 °C over 20 minutes. To determine the mRNA binding capacity of the functionalized fiber-packed column, mRNA was quantified by UV absorption. The transcript is approximately 1900 bases long with a molecular weight of ~665 kDa. The extinction coefficient is estimated to be 16.6 μM^{-1} cm^{-1}. By measuring the absorbance and volume of the mRNA solution loaded into and removed from the column, the performance of the column and fluidic system can be quantitatively assessed.

The column was washed with 400 µl of translation wash buffer (20 mM NaCl, 50 mM KCl, 20 mM Tris-Cl, 0.05% Tween 20, pH 7.5), and the flow-through was collected in 100 µL aliquots. A low salt elution buffer (20 mM Tris, 0.05% SDS, pH 8) was then loaded. The column was heated to 55 °C for two minutes and more elution buffer was washed through the column and the flow-through was collected in 100 µL aliquots. All fractions were quantified by the absorbance at 260 nm measured using a spectrophotometer (NanoDrop ND-1000). Concentration was determined from the absorbance value for RNA and the molecular weight of the transcript. The results are listed in Table 4.1.

Table 4.1. Binding of RNA on Glass Fibers Packed in Column.

Sample	A260 (1cm path)	Conc. (nM)	Volume (µL)	Picomoles
mRNA loaded	1.768	107	250	26.6
Wash 1	1.359	81.9	92.1	7.54
Wash 2	0.693	41.7	102.6	4.28
Wash 3	0.184	11.1	102.6	1.14
Wash 4	0.016	0.964	101.9	0.0982
Elution 1	0.887	53.4	105.3	5.63
Elution 2	0.258	15.5	100.1	1.56
Elution 3	0.035	2.11	101.1	0.213
Elution 4	0.026	1.57	100.5	0.157

The total amount collected from the wash and elution is 13.1 picomoles and 7.55 picomoles, respectively. Thus, the total mRNA recovery is about 20.6 picomoles, which is 77% of the original mRNA. This verifies that the loss due to non-specific binding or residuals in the fluidic system is small. The column was packed with 0.104 g of fibers at a length of 10 mm inside a tubing of 4 mm inner diameter. Using the density of Pyrex glass of 2.2 g/cm^3, the void volume is approximated to be 79 µl. The binding capacity of the column is determined to be 73 pmol mRNA per gram of fibers. Taking into account of the volume taken up by the glass fibers, the effective concentration of the mRNA is about 96 nM.

4.3.3 Solid-Phase Cyclic in vitro Translation

A reaction mixture for translating the surface-bound mRNA into proteins was prepared by Dr. Zhixia Liu. The solution utilized the prokaryotic cell-free protein expression kit E. coli T7 S30 Extract System for Circular DNA (Promega), consisting of T7 S30 extract, S30 premix without amino acids, amino acid mixture minus cysteine, a custom prepared solution of cys-tRNA pre-charged with cysteine labeled with ATTO 488 in total tRNA, and RNase inhibitor. This reagent was introduced into the fiber column containing surface-hybridized mRNA and was incubated for 3 hours at 30 °C. Following the reaction, the sample was collected by exchanging the column with 150 µl of wash buffer (10 mM NaCl, 50 mM Tris-Cl, 0.1% SDS, pH 7.2). Three wash steps were performed with a 2-minute incubation in between each fraction. The fractions were analyzed by SDS-PAGE. A Bst DNA polymerase large fragment chemically labeled with ATTO 488 was added as a control. After electrophoresis, the gel was visualized by exciting the lanes with 488 nm light from a Lambda DG 5 (Sutter Instruments). The fluorescence emission was passed through a bandpass filtered designed for FITC imaging and collected on a digital DSLR camera (Canon Rebel XL). The resulting image is shown in Figure 4.3. The gel demonstrates the successful solid-phase translation and elusion of labeled Bst polymerase.

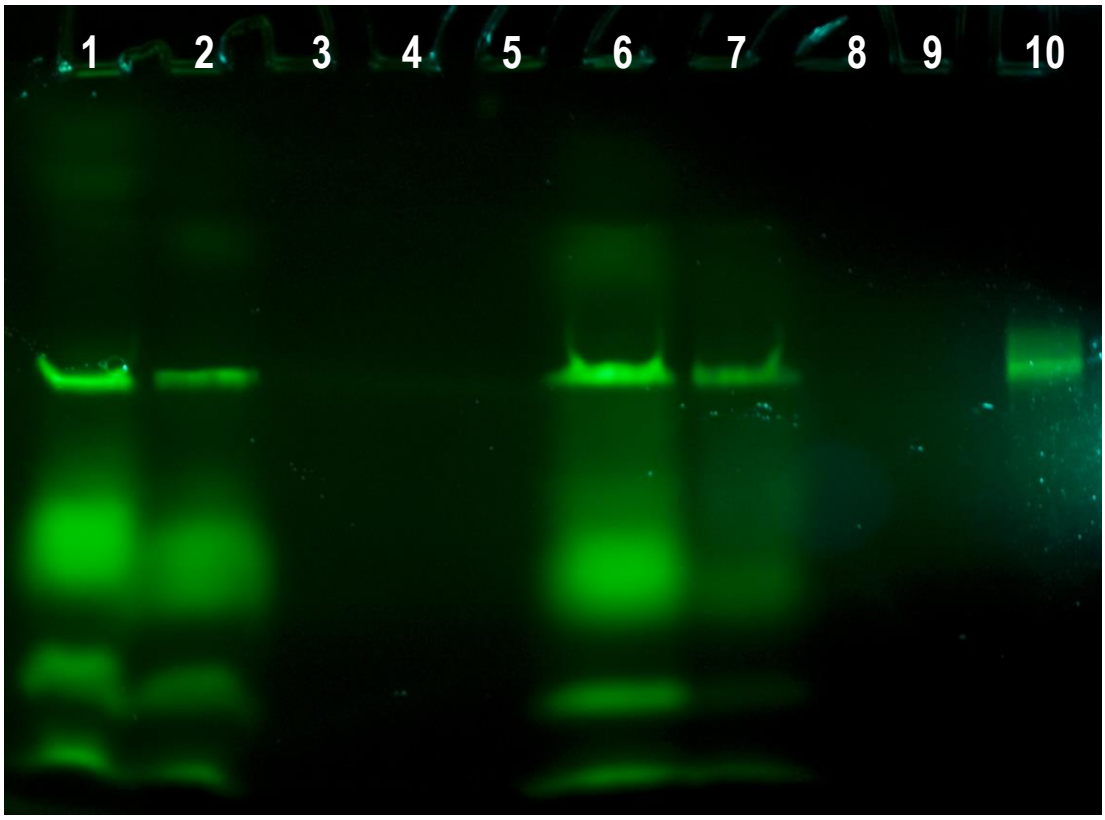


Figure 4.3. SDS-PAGE analysis of fluorescently labeled Bst polymerase translated on surface-bound mRNA. Lane 1: Reaction mix collection; lane 2: first wash; lane 3: second wash; lane 4: third wash; lane 5: loading buffer; lane 6: reaction mix collection concentrated; lane 7: first and second was combined and concentrated; lane 8 & 9: loading buffer; lane 10: chemically-labeled Bst polymerase-ATTO488.

4.3.4 Chemical Dual Labeling of Bst DNA Polymerase

Chemical labeling was assessed for the ability to evenly label the mutant Bst DNA polymerase. Chemical labeling of a protein with two exposed cysteine residues and two thiol-reactive fluorophores (dye 1 and dye 2) results in a mixture of six fluorescently distinct populations: unlabeled, singly-labeled with dye 1, singly-labeled with dye 2,

doubly-labeled with dye 1, doubly-labeled with dye 2, dually-labeled with both dye 1 and dye 2. Other permutations also exist depending on which of the cysteine residues the dyes are attached. Assuming an equal reactivity of both thiol-reactive fluorophores, an equal degree of labeling should be observed. To achieve such protein labeling, a reaction mixture was prepared containing 100 times excess of each dye over the protein concentration. Maleimide-activated ATTO488 and ATTO594 (ATTO-TEC) were diluted in dry DMF and the concentrations were determined using a NanoDrop Spectrophotometer.

The mutant Bst polymerase was expressed in and purified from E. coli by Dr. Zhixia Liu. To prepare the mutant Bst polymerase for labeling, 15 µl of 10 µM of the enzyme was washed with labeling buffer (100 mM KCl, 100 mM sodium phosphate, 0.1% Triton X-100, pH 7). Washing was performed by diluting and spinning down four times using an Amicon Ultra-0.5 10K (EMD MilliPore). Subsequently, reaction mixes consisting maleimide-ATTO488 and maleimide-ATTO594 was added to give final concentrations of ~500 µM for each dye. The reducing agent TCEP was also included at 250 µM to prevent the crosslinking of the polymerases cysteine residues, preventing aggregates and maximizing labeling. This reaction was allowed to proceed for one hour at room temperature before the solution was washed in the same manner. The mixture was finally diluted to 300 µl for measurement in a spectrophotometer to assess the degree of labeling. Based on absorbance values, it is determined that there is 1.12 times more ATTO 594 than ATTO 488, which is fairly close to desired 1:1 labeling. Due to a strong absorbance of Triton X-100 in the UV spectra, the amount of polymerase could not be quantified, preventing the calculation of the degree of labeling of the polymerase. The

raw and corrected absorbance spectra for the dual labeled Bst polymerase are shown in Figure 4.4. The concentrations for each fluorescent dye are calculated based on their extinction coefficient and shown in Table 4.2.

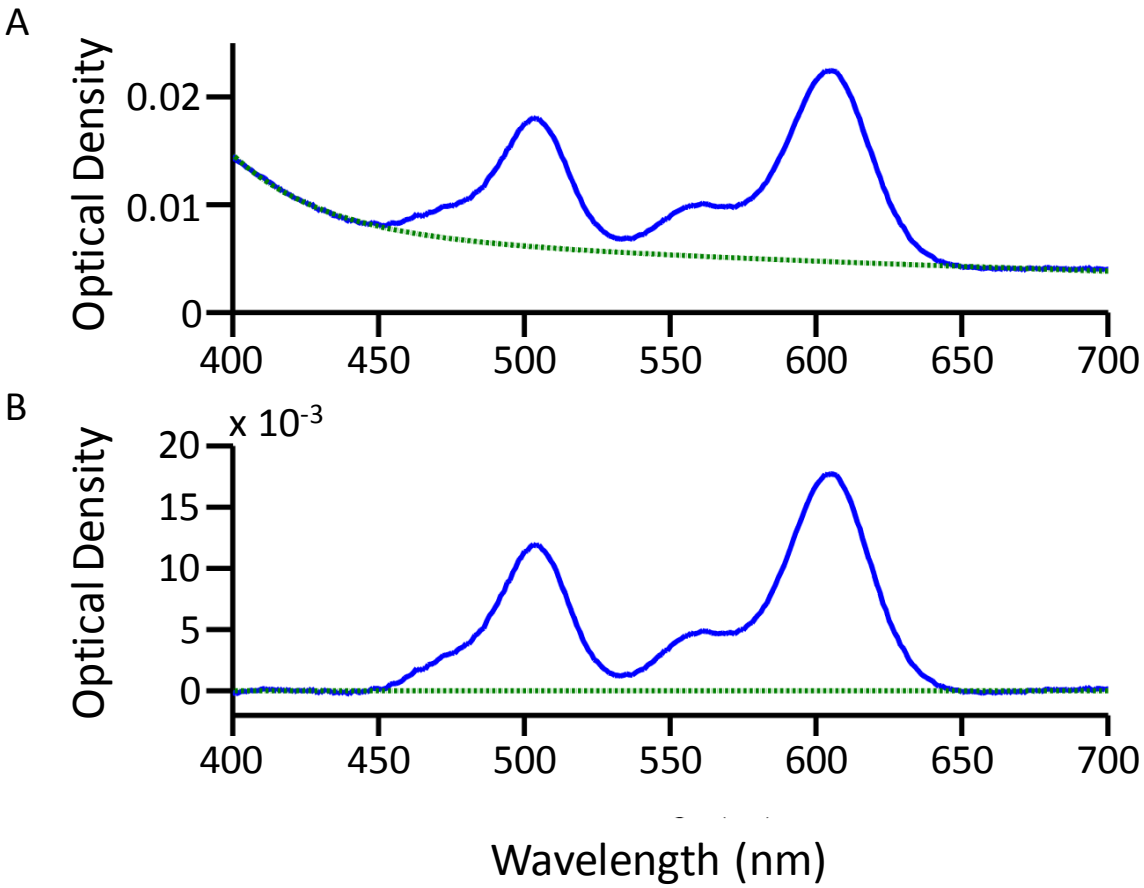


Figure 4.4. Baseline Correction of Absorbance of Dual-Labeled Fluorescent Bst DNA Polymerase. A. The raw spectrum of Bst DNA polymerase after dual chemically labeling. B. Baseline-corrected spectrum. To obtain accurate optical density values the data outside the wavelengths of interest was fit to an exponential curve and baseline was corrected.

Table 4.2. Concentration of Fluorescent Dyes Measured by Spectrophotometry.

Fluorophore	Extinction Coeff.: ϵ_{max} (M^{-1} cm^{-1})	Max Absorbance (1 cm path)	Conc. (nM)
ATTO 488	90,000	0.0114	126.7
ATTO 594	120,000	0.0171	142.5

This sample was further analyzed using single-molecule TIRF imaging and counting of immobilized polymerase. A 30 nM solution of polymerase was injected into a flowcell containing DNA on the surface. The DNA polymerase was allowed to bind to the primed DNA molecules on the surface for 5 minutes. The flowcell was washed briefly before imaging. A 488 nm laser was used for excitation for imaging the donor (ATTO 488), and the FRET between the donor (ATTO 488) and acceptor (ATTO 594). The acceptor was directly detected using 532 nm laser excitation. From a total of eight fields of views, the average numbers of donors, acceptor by FRET, and acceptors directly excited was 1118, 555 and 1156, respectively. The ratio of donors and directly excited acceptors is also close to one. This agrees well values determined by absorbance measurements.

4.3.5 Protein Engineering Summary

A mutated Bst DNA polymerase containing two introduced cysteine residues was labeled with fluorescent dyes by an *in vitro* translation method, as well as by traditional chemical labeling. The *in vitro* translation method relies on mRNA bound to a solid-support and a fluidic system to control the flow of custom-prepared translation reaction

mixes with the desired charged tRNAs and labeled amino acids. While we were unsuccessful at demonstrating the ability to incorporate specific fluorescently labeled amino acids at designated locations, the ability to produce full length enzymes with fluorescently labeled cysteine residues using surface-bound mRNA for translational templates was established. Maleimide-activated dyes were conjugated to a mutated Bst DNA polymerase expressed in E. coli in order to prepare polymerase samples for imaging experiments.

4.4 Single-Molecule FRET Measurements by TIRF Microscopy

Single-molecule FRET imaging is quite challenging due to the limited sensitivity and single-to-noise (SNR) that can be achieved. Sensitivity is determined by the optics and sensor of the imaging system, photophysical properties of the dye and detection noise. The maximum achievable sensitivity and SNR are primarily determined by photon collection efficiency of the imaging system which in turn is defined by the magnification and numerical aperture of the objective lens, intensity of the excitation and wavelength, bandwidth of the emission filter, and integration time and quantum efficiency of the detector. The fluorescence properties of the dyes can largely be characterized by the quantum yield and extinction coefficient, but also by other factors such as fluorescent lifetime and photostability. Noise stems from many intrinsic sources such as Raman and Raleigh scattering and extrinsic sources such as electronic noise in the detection system.

In high speed imaging, as the exposure time decreases, the signal intensity decreases. At certain point, the signal would be at the background level.

To monitor rapid enzyme kinetics and conformational dynamics, fast frame rate is essential to capture the molecular motions. For observing the real-time serial nucleotide incorporation by DNA polymerases which usually has a rate constant in the range of 10-100 nucleotides per second, the ideal integration time would be 10 ms or less so that sufficient number data points (e.g. 10 or greater) can be acquired for each nucleotide incorporation event. This sampling rate would allow the subtle and brief motions of DNA polymerase to be detected. The ability to detect the subtle but yet distinct conformational dynamics of the polymerase during the multi-step nucleotide incorporation process would potentially enable the real-time readout or sequencing of single DNA molecules. In this work, we aim to use high-speed single-molecule FRET imaging to dissect the molecular mechanism of DNA polymerases and to enable a new sequencing technology, which is called READS.

4.4.1 Simultaneous Imaging of FRET Donor and Acceptor Fluorescence

In order to accurately calculate the FRET efficiency, intensity measurements of the donor and acceptor dyes must be measured simultaneously. This is achieved by splitting the fluorescence emission according to wavelength. The imaging system is described in depth in Section 2.3.3 and shown in Figure 2.2. Briefly, an inverted fluorescence microscope (AxioObserver Z1, Carl Zeiss) is used to perform objective-based TIRF imaging. A TIRF slider (TIRF 3 Slider, Carl Zeiss) is used to introduce laser

excitation slightly beyond the TIR angle from sources at 405 nm, 488 nm, 532 nm, and 647 nm. A 100X oil objective with a numerical aperture of 1.46 (Alpha Plan Apochromat, Carl Zeiss) is used for laser excitation and fluorescence detection. A dual camera adapter (Carl Zeiss) with an image splitting dichroic mirror is used to spectrally separate the donor and acceptor fluorescence emission. The longer wavelength fluorescence emission from the acceptor is transmitted through the dichroic mirror to a horizontally mounted camera and the shorter wavelength fluorescence emission from the donor is reflected by the dichroic mirror a vertically mounted camera. iXon3 897 EMCCD cameras are used to achieve high sensitivity single-molecule fluorescence detection. Bandpass filters are placed after the dichroic mirror to reduce background fluorescence and cross-talk between the donor and acceptor fluorophores. The acceptor images must be computationally aligned to the donor images by rotation and translation.

For imaging FRET pairs in the Cy3/Cy5 channels, the 532 nm laser is used for excitation. An image splitting dichroic mirror with a cutoff wavelength of 662 nm is used for spectral separation. The dichroic mirror reflects the donor emission through a Cy3 emission bandpass filter (FF01-580/60-25-D, centered at 580 nm with a width of 60 nm). The acceptor emission is transmitted through the dichroic mirror through a Cy5 emission bandpass filter (FF01-675/67-25, centered at 675 nm with a width of 67 nm).

4.4.2 Single-molecule FRET Experiments Using DNA Scaffolds

To investigate the ability of the microscope system to detect donor and acceptor signals at various frame rates and FRET efficiencies, fluorescently labeled DNA probes were hybridized to a DNA template and immobilized on a coverslip surface. The use of

dsDNA as a molecular ruler for smFRET experiments provides an efficient method for creating a variety of fluorophore pairs and spacing. By utilizing three oligos labeled with different fluorescent dyes and three scaffolding oligos, nine different FRET conditions were created.

The probe oligos share a common sequence and contain a label at the 5' end. The scaffolding oligos were designed with a 24-base sequence to fully hybridize one probe at one end and a sequence that can be used to hybridize another probe. The separation between the donor and acceptor dyes can be spaced 12, 17, and 24 bp apart to give approximately 4, 6, and 8 nm distance, respectively. Table 4.3 shows the sequence of the oligos used. DNA constructs were formed by hybridizing the desired scaffolding oligo at 1 µM with two corresponding probes at 2 µM in 2x SSCt. The sample was heated at 65 °C for 5 minutes and cooled to 25 °C at 0.1 °C/s using an MJ Thermocycler. These hybridized oligos were then diluted (50 pM for the scaffolding oligo) and injected into a flowcell containing a poly$(dT)_{50}$ coated surface. The construct was hybridized to the surface via the poly(dA) region on the 5' of the scaffolding oligo for at least 5 minutes. The channel was washed and time-serial images was acquired to measure the single molecule FRET intensities.

Table 4.3. DNA Scaffolding Oligo Sequences. The underlined and bold sequences on the scaffold oligos represent the different parts of the template that are used to hybridize the probe oligos.

Oligo Name	Sequence
24mer_5Alx488	5'-Alexa488-CCA TCA GTC ATG TAC GAA GTC AGT-3'
25mer_5Alx546	5'-Alexa546-CCA TCA GTC ATG TAC GAA GTC AGT C-3'
26mer_5Alx647	5'Alexa647-CCA TCA GTC ATG TAC GAA GTC AGT CA-3'
4 nm Separation Scaffold	5'- AAA AAA AAA AAA AAA AAA AAA AAA ACT GAC TTC GTA CAT GAC TGA TGG **CAT GAC TGA TGG** -3'
6 nm Separation Scaffold	5'- AAA AAA AAA AAA AAA AAA AAA AAA ACT GAC TTC GTA CAT GAC TGA TGG **TCG TAC ATG ACT GAT GG** -3'
8 nm Separation Scaffold	5'- AAA AAA AAA AAA AAA AAA AAA AAA ACT GAC TTC GTA CAT GAC TGA TGG **ACT GAC TTC GTA CAT GAC TGA TGG** -3'

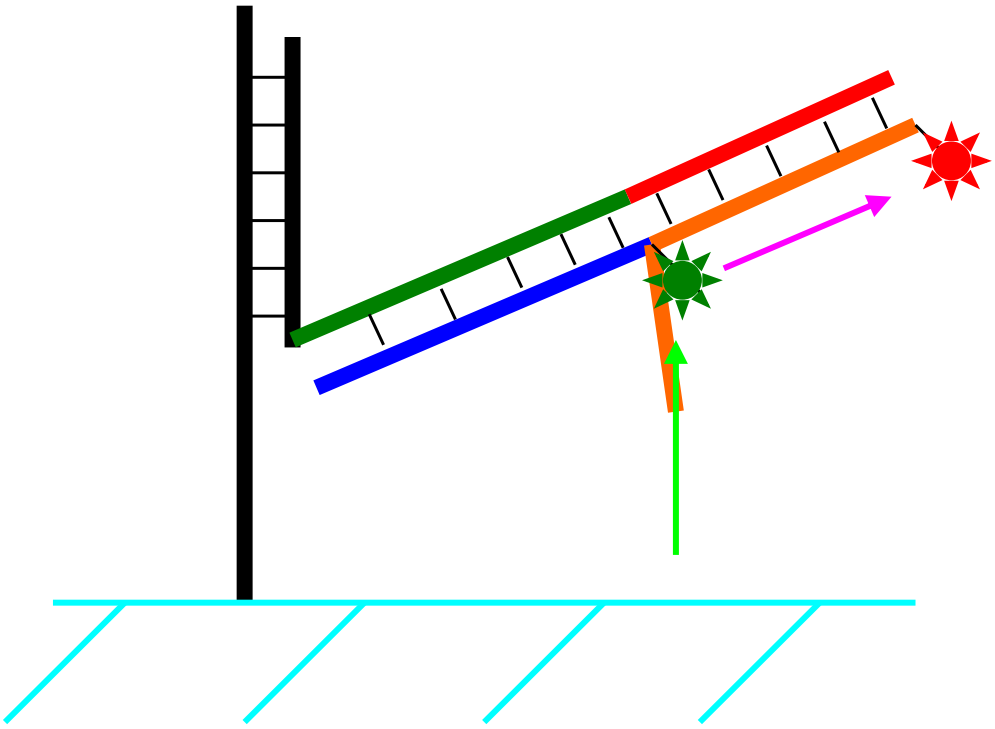


Figure 4.5. DNA Scaffold for smFRET Experiment. DNA scaffolding templates hybridized to two fluorescently-labeled oligos to create a structure with precise distance between the fluorophores along the double-stranded DNA. This design enables easy combination of FRET pairs and FRET distances by switching the labeled oligos and scaffolding templates, respectively. The templates contain a short poly(dA) sequence to hybridize to poly(dT) groups on a coverslip surface to enable TIRF imaging.

Table 4.4. Theoretical FRET Efficiencies for Different Donor-Acceptor Pairs and Distances.

FRET Pair	Förster Radius (nm)	FRET Efficiency at 4 nm Separation	FRET Efficiency at 6 nm Separation	FRET Efficiency at 8 nm Separation
Alx488-Alx546	6.37	94.2%	58.9%	20.3%
Alx488-Alx647	5.50	87.1%	37.1%	9.5%
Alx546-Alx647	7.02	96.7%	72.0%	31.3%

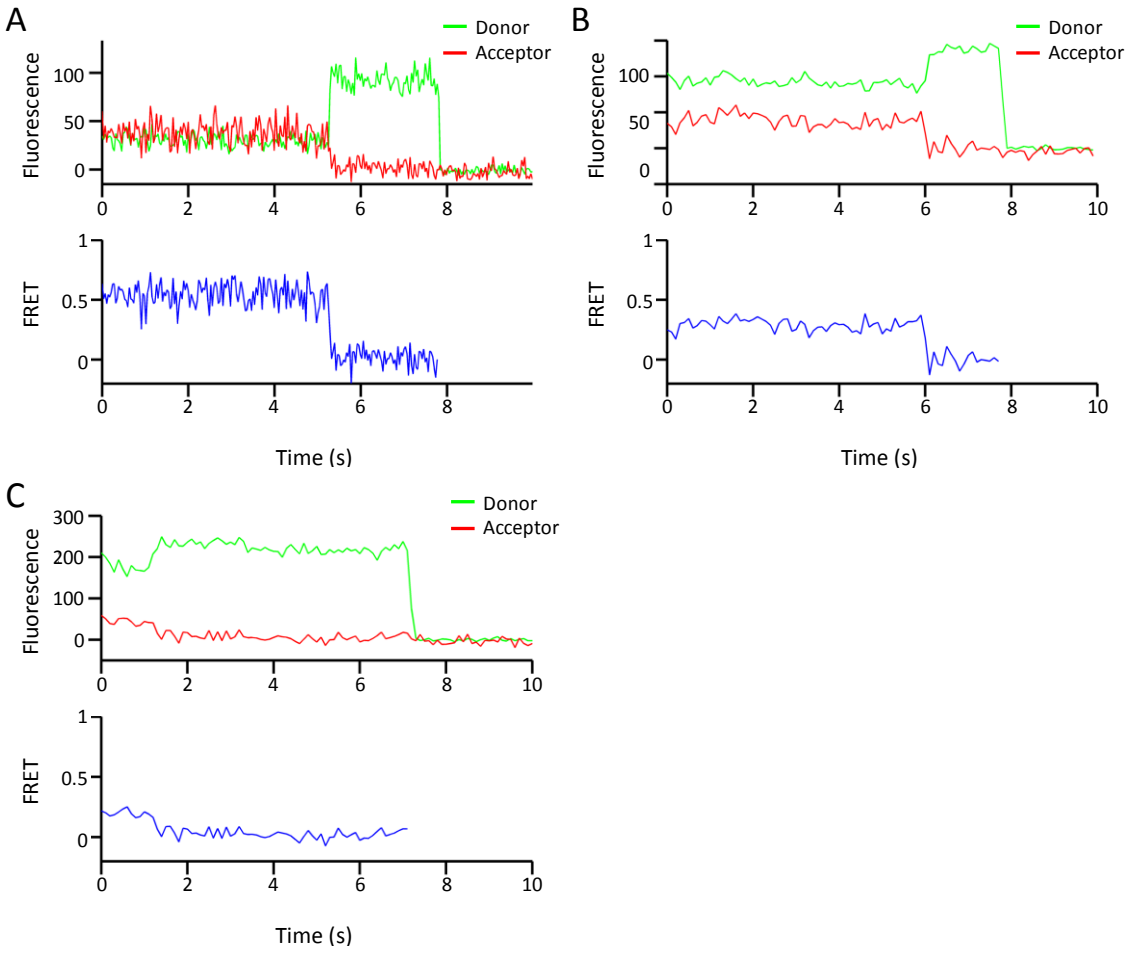


Figure 4.6 FRET Traces of Alexa 546 and Alexa 647 at different distances. The effect of the distance between dyes on the FRET efficiency is demonstrated. A. At an inter-dye separation of 4 nm, the FRET efficiency is measured to be 59.3% for this trace. B. For a spacing of 6 nm, the FRET efficiency is observed to decrease to 29.4%, almost half the value as shown in A. C. As the distance increases to 8 nm, the FRET efficiency decreases to 20.0%. Due to the low FRET efficiency for the 6 nm and 8 nm separation, longer exposure times were used to ensure suitable signal-to-noise values for analysis.

The ability of our microscopy system to detect changes in FRET efficiency at different separations between Alexa 546 and Alexa 647 were assessed. Using the DNA scaffold, inter-dye distances of approximately 4 nm, 6 nm and 8 nm were constructed. All imaging was performed using a 100x objective and 532 nm laser excitation. Due to the

high FRET efficiency of the 4 nm separated pair, a 35 ms exposure time was able to easily detect both the donor and acceptor fluorophores using an EM gain of 300 and laser intensity of 8.7 mW. For the FRET pairs a longer with 6 nm and 8 nm separation, 100 ms exposure time was necessary to effectively image the fluorescence of the acceptor. The EM gain of the EMCCD camera was set at 200. For the pair with 8 nm separation,17.5 mW of laser excitation was use, while only 8.7 mW excitation was used for the pair with 6 nm separation. The emitted fluorescence of the donor and acceptor dyes and the FRET traces are displayed in Figure 4.6.

The FRET efficiency was calculated to be 59.3%, 29.4%, and 20.0% for the FRET pair with 4 nm, 6 nm, and 8 nm separation, respectively. All measured FRET efficiency of all the pairs are lower than those predicted and listed in Table 4.4. It is likely that some of the discrepancy is attributed to extra distance in the dye linkers and the simple approximation of the 0.34 nm per base pair of helical structure of the double-stranded DNA scaffold. This inaccuracy would be more pronounced at shorter linear distance, which is seen in the data. This set of experiments highlights the difficulty in detecting the low signal from an acceptor with a low FRET efficiency. A long exposure time was required to achieve good SNR, even at relatively high laser intensity. To detect the rapid kinetics and conformational changes of DNA synthesis would require a much faster frame rate.

4.4.3 Analysis of Signal-to-Noise at Different Frame Rates

To explore our ability to detect fluorophores at a high frame rate, oligos labeled with Alexa 546 were hybridized to a surface and imaged at three different frame rates of

5 ms, 17 ms, and 35 ms per frame. All measurements were performed using a 100x objective, a 300 EM gain, and 14.5 mW of 532 nm laser excitation. To enable data acquisition at rates higher than 35 ms per frame (the maximum full frame rate of the camera), the camera was operated with 2 pixels x 2 pixels binning. The binning doubles the frame rate to 17 ms per frame. To decrease the exposure time further, the sensor size was cropped such that only the bottom 64 rows of the pixels were kept active for faster data transfer.

By collecting fluorescence emission for a long duration (60 seconds), the kinetics of photobleaching of multiple fluorophores was collected. Single molecules were identified and their fluorescence traces analyzed as previously described. Photobleaching events were identified and the corresponding time point was used to determine the signal of the fluorophore. The fluorescence emission signal was determined by calculating the average intensity before photobleaching subtracted by the average intensity of the fluorophore location after photobleaching had occurred. This background subtraction provides the true value of fluorescence signal from the fluorophore. The SNR values are calculated by dividing the true signal intensity of the fluorophore by the standard deviation of the fluorescence intensity of the fluorophore. Representative fluorescence traces with 5 ms, 17 ms, and 35 ms exposure are shown in Figure 4.7. Histograms displaying the distribution of calculated SNR values for each exposure time are shown in Figure 4.6. The mean SNR values are determined to be 2.58, 4.71, and 4.54. From these plots, it is apparent that the SNR decreases as the exposure time decreases. A SNR of 2 or greater would be required for the confident detection of a variation in intensity due to a

meaningful event rather than system noise. For a specific fluorophore, one way to increase the SNR at a given frame rate is to increase excitation power.

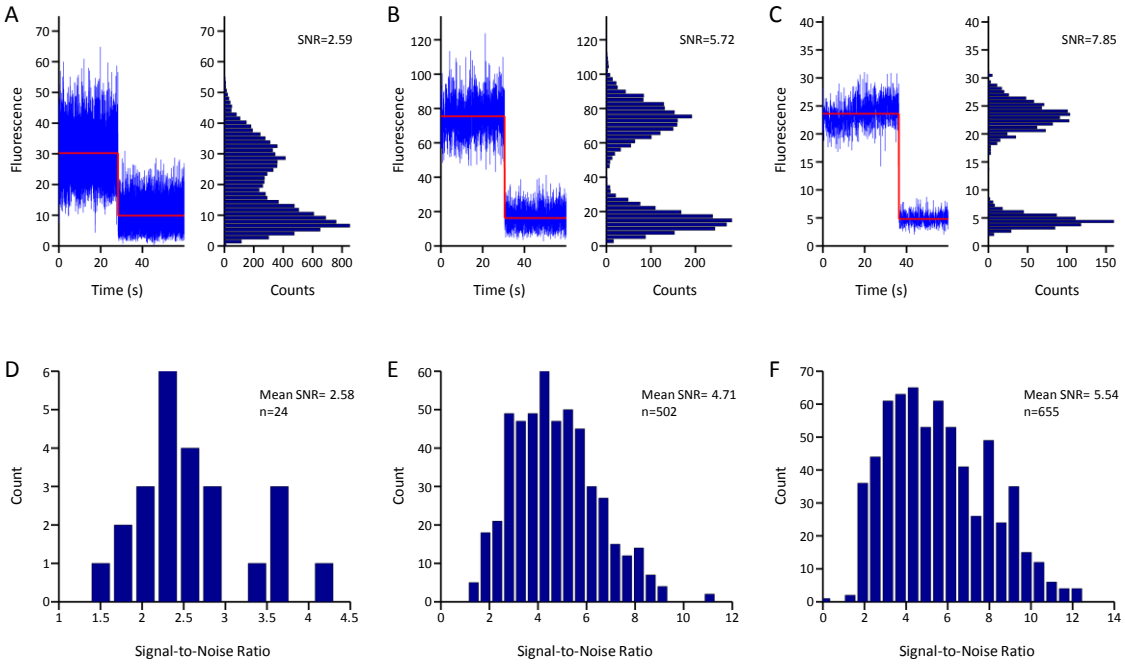


Figure 4.7. Signal-to-noise analysis of Alexa 546 imaged at different frame rates. Fluorescence traces for photobleaching events of Alexa 546 are shown. A. Data acquisition at 5 ms exposure. SNR = 2.59. B. Data acquisition at 17 ms exposure. SNR = 5.72. C. Data acquisition at 35 ms exposure. SNR = 7.85. Note that 2 x 2 binning was not used in this case. This results in lower fluorescence counts for both signal and background. D. The histogram of SNR values for the 5 ms exposure. Mean SNR = 2.58. Few molecules are detected due to the cropping of the sensor for the high frame rate and low SNR. Only spots a SNR > 1.5 were detected with the image processing algorithm. E. The SNR histogram for 17 ms exposure. A Gaussian distribution centered on the mean SNR of 4.71 is observed. F. The SNR histogram for 35 ms exposure. SNR values exhibit a broader distribution with a mean of 5.54.

4.5 Single-Molecule FRET Imaging of Labeled Bst DNA Polymerase

Our experiments with fluorescently-labeled DNA provide some insight into the potential capability and limitations of our current imaging system in detecting the single-

molecule FRET signal from dual-labeled polymerases. There are additional challenges working with proteins. One such challenge is the preparation of DNA-bound proteins at appropriate density for single-molecule imaging. Other challenges are related to fluorescence intensity fluctuation observed when fluorophores are conjugated to proteins.

Labeled Bst DNA polymerase was prepared by the chemical labeling method describing in Section 4.3.4 followed by purification by size exclusion chromatography using high-performance liquid chromatography (HPLC) to remove the free dye. Two samples were prepared. One contained Cy3 as the only fluorescent label. The second was created using a mixture of ATTO532 and ATTO647N to create a dual-labeled polymerase. The labeled enzyme solutions were analyzed for concentration and degree of labeling using a spectrophotometer (NanoDrop), diluted into storage buffer containing 50% glycerol, and stored at -30 °C.

4.5.1 Real-time Monitoring of DNA Polymerase Binding to Surface-Immobilized Primed DNA Template Using smFRET

Preparing an imaging surface with fluorescently labeled DNA polymerase bound immobilized primed DNA templates poses the challenge of reducing the non-specific binding of the polymerase while achieving a high enough amount of DNA-bound polymerase. A large amount of active enzymes are necessary in order to attain statistically relevant data and increase experimental through-put. Non-specifically bound polymerases reduce the imaging area available for the functioning polymerases and increase image processing complexity. To solve this problem, 10 nM of the polymerase

was incubated with 1 nM of primed DNA template containing a poly(dA) tail in 1x Bst buffer solution. The nanomolar range concentration of both species enables the binding equilibrium to be shifted towards a high degree of bound polymerase-DNA complexes. This mixture was immediately diluted by 10x before loading into a channel with a poly(dT) coated coverslip. The concentration of 100 pM DNA template provided a suitable density for single molecule imaging and the 1 nM polymerase has minor non-specific binding to the surface.

To visualize Bst DNA polymerase's proper binding to the immobilized DNA, a primer containing Alexa647 label on the 3' end was used with a Cy3-labeled polymerase. The oligonucleotide sequences used are shown in Table 4.5. The polymerase-DNA solution was incubated in the channel for 5 minutes to allow for the surface capture by hybridization of the poly(dA) tail to the poly(dT) surface. Subsequently, imaging buffer (0.15 μM PCD, 5 mM PCA, and 2 mM Trolox in 1x Bst buffer, pH 8) was introduced into the channel and Cy3/Cy5 smFRET imaging was initiated using the imaging system described in 4.2.1 with a 35 ms exposure time, EM gain of 300, and 532 nm laser excitation at 2 mW.

Table 4.5. Sequences of oligos for smFRET monitoring of Cy3-labeled Bst DNA polymerase and Cy5-labeled primed templates, Cy3-labeled Bst DNA polymerase and Cy5-labeled dCTP, and dual-labeled Bst DNA polymerase.

Template & Capture Oligo/3’-Labeled Primer for Polymerase-DNA Binding Monitoring
5′-GCTCACTCAACTACACGTAACCATATCACTAAGAGCATGACTGACTTCGTACATGACTGATGG-poly(dA)-3′ 3′-Alexa647-ACTGAAGCATGTACTGACTACC-5′
Spaced 3 dCTP Incorporation Template & Capture Oligo/5’-Labeled Primer for Labeled dCTP Incorporation
5′-GCTCACTCAACTACACGTAACCATATCACTAAGAGCATGACTGACTTCGTACATGACTGATGG-poly(dA)-3′ 3′-TCGTACTGACTGAAGCATGTACTGACTACC-Alexa488-5′
Homopolymer Incorporation Template & Capture Oligo/Primer for Dual Labeled DNA Polymerase Synthesis
5'-TTTCCCCCCCCCCTTTCCCCCCCCCCCGTATCGACCATCAGTCATGTACGAAGTCAGTCATGCTAGTTTTACCCCACTAAGTCTGTA-poly(dA)-3′ 3′-CATAGCTGGTAGTCAGTACATGCTTCAGTCAGTACGATCAAAATGGGGTGATTCAGACAT-Biotin-5′

Figure 4.8 depicts the different types of polymerase-DNA interactions that were observed. Interactions of the polymerase with the 3' end of the primer are distinguishable by the increase in FRET efficiency. Having the donor dye on the polymerase provides the unique ability to visualize when the enzyme is near the coverslip surface (presumably bound to the DNA), but may not be close to the 3' end of primer. This enabled the ability to monitor and detect 1-dimensional diffusion of the polymerase along the template, which is a known phenomenon for DNA binding proteins searching for their target position.[12] A wide variety of traces of Bst polymerase binding were observed, ranging from relatively stable binding at the priming site to sporadic localization near the priming site. Figure 4.8 A and B reveal the transient interactions with the 3' end of the primer, but a sustained presence of the polymerase on the template, indicative of 1-D diffusion. Figure 4.8 C and D exhibit a state of stable polymerase binding at the priming site with minimal and brief dissociations. The long dwell times of the polymerase on the template indicate a high affinity for the enzyme towards the DNA.

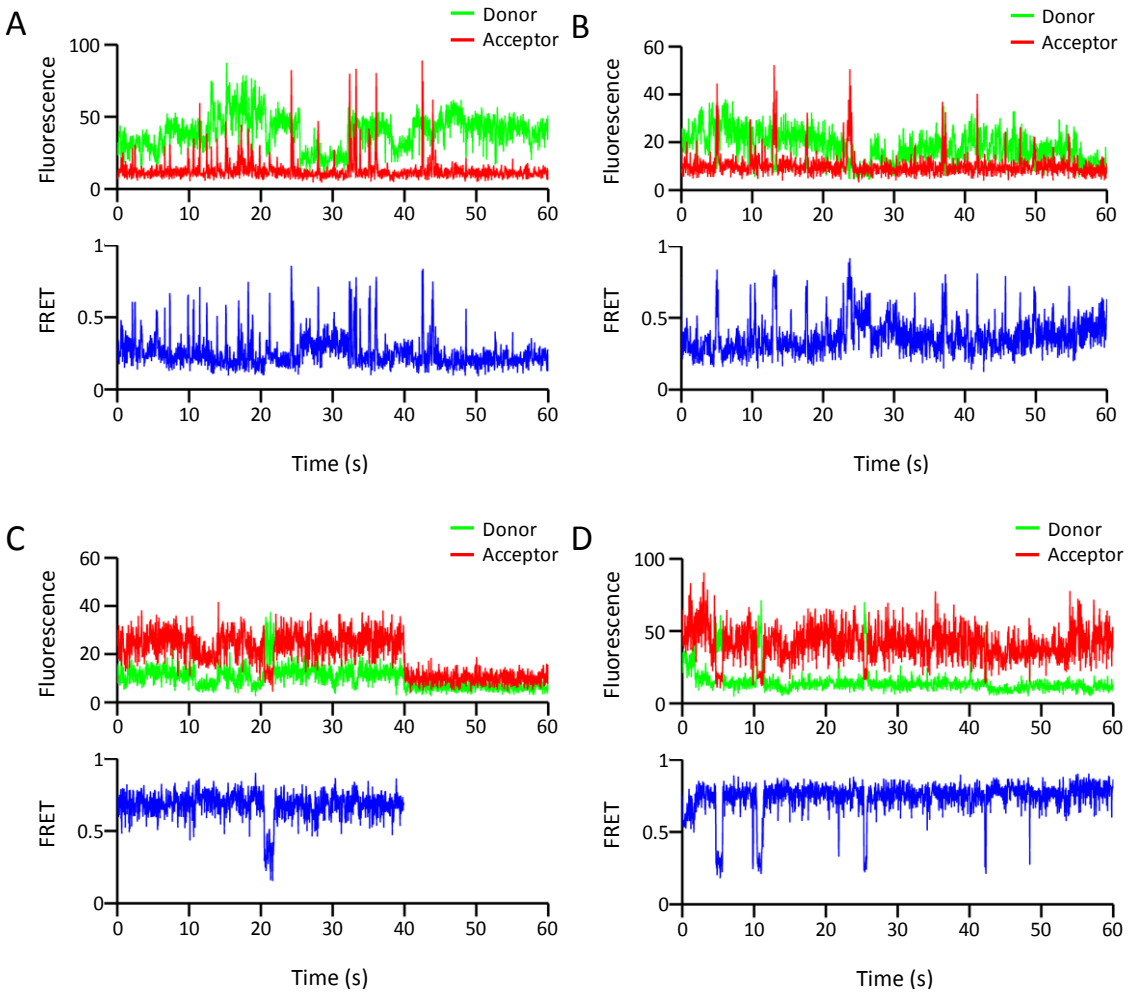


Figure 4.8. Fluorescence and FRET traces of Cy3-labeled Bst DNA polymerase and Cy5-labeled primer. These traces represent some examples of a variety of polymerase-DNA interactions that were observed. A and B. These FRET traces display a highly dynamic binding of a polymerase to DNA residing within close proximity of the 3' end of the primer for brief periods. The polymerase moves away from the acceptor fluorophore attached to the 3' end of the primer while remaining located in the same location of the template potentially due to 1-dimensional diffusion. C and D. These FRET traces represent stable complexes of polymerase bound to DNA at the primer site. The long binding events and brief duration indicate that the polymerase has a strong affinity for the primer site for some cases.

It is unclear what causes the polymerase to bind securely to some priming sites and to have highly dynamic interactions with others. To more thoroughly quantify these differences in binding activity, more experiments would be required, perhaps using

different primers with internal labeling at various positions and multiple concentrations of polymerase. Overall, this experiment served to demonstrate that the fluorescently labeled DNA polymerase mutant retained the ability to bind DNA at the priming site. It also indicated that the method of pre-incubating polymerase and DNA followed by the surface capture of the DNA by hybridization provides a suitable method of sample preparation for subsequent experiments involving nucleotide incorporation.

4.5.2 Real-Time Monitoring of Cy5-Labeled Nucleotide Incorporation by Cy3-Labeled DNA Polymerase

To monitor the incorporation of Cy5-labeled nucleotides, Cy3-labeled Bst DNA polymerase bound to Alexa488-labeled primed DNA template (sequences shown in Table 4.5) was loaded into the channel and captured to the surface by hybridization to the poly(T) on the surface as described earlier. After a brief wash to remove any unbound polymerase, the DNA synthesis reaction was conducted by flowing into the channel a solution containing 5 μM each of dATP, dGTP, dTTP and Cy5-labeled dCTP, 100 μg/ml BSA, 2.5 mM Trolox, mM PCA, 1 μM PCD in 1x Bst buffer at pH 8. Right before reaction mixture entered into the channel, time-lapse fluorescence imaging was initiated using 2 mW of 532 nm laser excitation. Imaging was performed with 35 ms exposure and EM gain of 300.

As with the polymerase binding experiments, a variety of traces were observed. There appeared to be distinguishable differences between what may be considered an incorporation event versus a transient binding event based on the duration of the events in which fluorescence intensity of the acceptor and FRET efficiency between the donor and

acceptor increase significantly. Example traces from individual DNA polymerase molecules are shown in Figure 4.9. Some traces exhibit events with long duration while other have more events with shorter duration. In the DNA synthesis process, three Cy5-labeled dNTPs are supposed to be incorporated at 10-base apart. As such, the fluorescence intensity of the acceptor and FRET efficiency between the donor and acceptor are supposed to go up, down, up, down and then up, and finally down when the polymerase dissociates from the template. Interestingly, this trend is observed in many traces, as shown in Figure 4.9 A and B. Therefore, it is very likely that the event with the longer duration is due to the active DNA synthesis by the polymerase. The events with shorter duration very likely are due to transient binding and sampling of nucleotides by the DNA polymerase or other activities of the DNA polymerase. Further work will be required to characterize the kinetics and the conformational changes associated with base-by-base DNA synthesis by the DNA polymerase. These experiments demonstrate that the labeled polymerase mutants retain nucleotide incorporation activity and that the rate of synthesis for 23 nucleotides takes around 10 seconds.

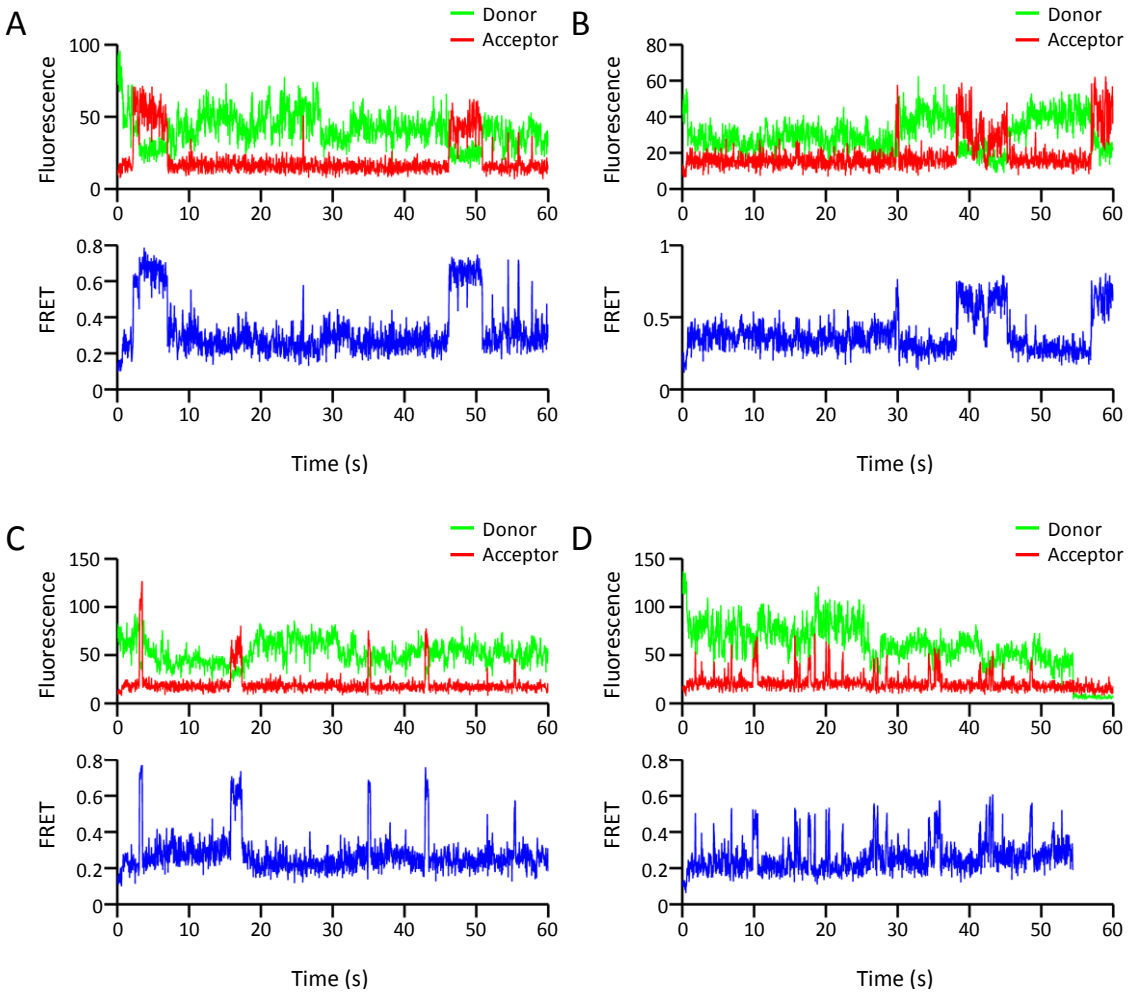


Figure 4.9. Fluorescence and FRET traces of Cy3-labeled Bst DNA polymerase during the incorporation of Cy5-labled dCTP. These traces represent some examples of a variety of nucleotide-polymerase interactions that were observed. A and B. Fluorescence and FRET traces of two DNA polymerase molecules with events having a long duration in which the FRET efficiency is significantly increased. The events could potentially be interpreted as the process where nucleotides are being incorporated by the DNA polymerase. C and D. Fluorescence and FRET traces of two DNA polymerase molecules with events having shorter duration in which the FRET efficiency is significantly increased. The shorter bursts of acceptor intensity reflect a transient interaction that might be more indicative of the sampling of the nucleotide by the DNA polymerase or other activities.

4.5.3 Single-Molecule FRET Imaging of Dual-Labeled Bst DNA Polymerase During DNA Synthesis

Once the mutant Bst DNA polymerase has been proven active for binding to DNA and incorporating nucleotides while monitoring using smFRET, the next step is to perform DNA synthesis using native nucleotides and a dual-labeled polymerase. Surfaces were prepared in a similar manner as described in Section 4.5.1, except that the channel temperature was held at 37 °C. The polymerase-DNA complex was incubated for 2 minutes to allow the complex to bind to the surface before the channel was washed with 100 µl of 1x Bst buffer. Immediately following the wash, an imaging buffer (0.15 µM PCD, 5 mM PCA, and 2 mM Trolox in 1x Bst buffer, pH 8) was flowed through the channel and a reaction mix (10 µM dNTPs, 0.1 µg/µl BSA, 0.15 µM PCD, 5 mM PCA, and 2 mM Trolox in 1x Bst buffer, pH 8) separated by a minimal air gap from the imaging buffer was injected into the channel. The flow was timed precisely such that right before the entrance of the reaction mix high-speed data acquisition was initiated using a 10 ms exposure, EM gain of 100, and 10 mW of 532 nm laser excitation. The template is designed to contain two stretches of a 10mer deoxycytidine homopolymer separated by a 3-base deoxythimidine homopolymer. We aim to demonstrate whether the incorporation of dGTP and dATP can be distinguished. In other experiments, a lower concentration of dATP was used to promote further variation in incorporation behavior.

While it remains difficult to interpret the incorporation kinetics from these FRET traces at this point, it appears that conformational change of the DNA polymerase upon the initiation of DNA synthesis is detectable. The raw and smoothed fluorescence data of

the donor and acceptor, and the calculated FRET efficiencies for two individual polymerase molecules are shown in Figure 4.10. Both traces display reasonably high SNR, but vary greatly in their FRET efficiencies. Reasons for the large differences in FRET efficiencies could be due to local environment of the individual dyes, which could influence the efficiency of energy transfer. Both enzymes show an abrupt increase in FRET efficiency (a decrease in donor and simultaneous increase in acceptor) at 0.48 seconds. This is the time point the reaction mix is expected to enter the fluidic channel. The change in FRET efficiency indicates a change in the polymerase conformation that has brought the finger and palm domains closer in the presence of dNTPs as expected .[9]

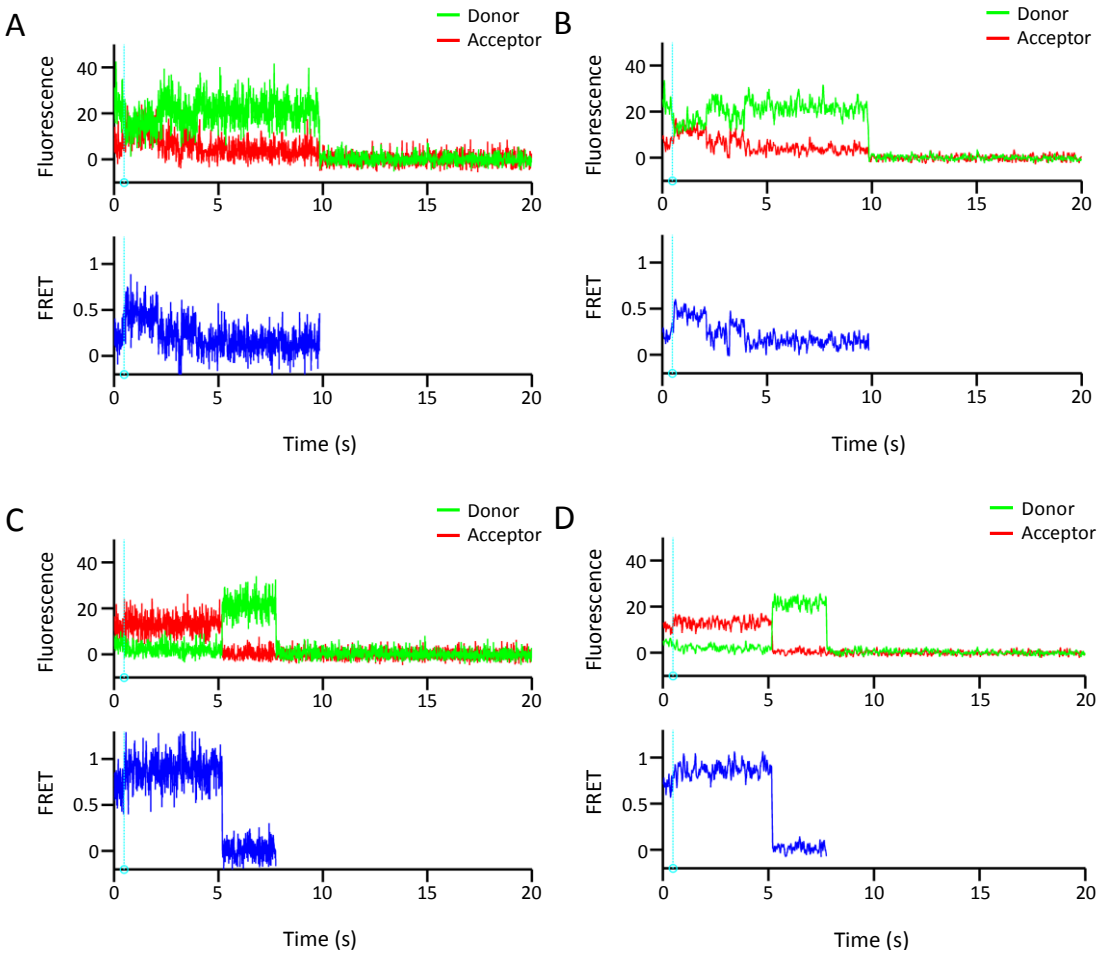


Figure 4.10. Raw and smoothed fluorescence and FRET traces from dual-labeled DNA polymerase during DNA synthesis. The fluorescence data from the donor and acceptor on Bst DNA polymerase displays an anti-correlated change upon the introduction of nucleotides. From the FRET traces, the increase in energy transfer efficiency indicates a decrease in the distance between the fluorophores as the result of a structural change in the enzyme. A. Raw fluorescence traces and the corresponding FRET traces. B. Smoothed fluorescence and FRET traces shown in A. C. Raw fluorescence traces and the corresponding FRET traces of two individual polymerase molecules. D. Smoothed fluorescence and FRET traces shown in C.

4.5.4 Variation of Fluorescence Signals Observed on Individual Dye-Labeled DNA Polymerases

A major challenge in the interpretation of smFRET signal is due to the variations in behavior of the fluorescence dye. Intrinsic factors such as dye rotation and blinking

can cause fluctuations in signal on a range of time scales. The flexible linkers used to attach dyes should allow a large degree of movement such that the signal will average out over the integration time. Blinking is caused by the transient formation of the triplet state, which can be effectively quenched by triplet-state oxygen. However, due to the propensity of singlet-state oxygen to cause photobleaching, dissolved oxygen is often depleted from the imaging solution using oxygen scavenging systems. Organic compounds capable of quenching triplet states, such as Trolox (6-hydroxy-2,5,7,8-tetramethylchroman-2–carboxylic acid), can also be used to prevent blinking.[13]

Other extrinsic factors, such as solvent environments, can also greatly influence the properties of fluorophores. Conditions where fluorophores are positioned very closely have been shown to generate multiple fluorescence intensity states for the acceptor, deviation from normal FRET behavior.[14] The interaction between the dye molecule and the amino acid residues with aromatic or other conjugated bond systems on the protein are known to alter the fluorescence properties of the dyes.[15] This behavior was observed for some polymerases labeled with ATTO532 and ATTO647N. A large degree of fluctuation in signal intensity was observed for some fluorescence traces of the DNA polymerase molecules labeled with the FRET pair in the absence of dNTPs. One example trace is shown in Figure 4.11A. The trace displays a transition between two unique states with different FRET efficiencies. The data could be interpreted as a change in the protein structure oscillating between two confirmations. However, based on observations of the direct excitation of the acceptor dye (ATTO647N) as shown in Figure 4.11B, it is evident that the fluorophore probably interacts with some amino acid side chains on the protein or other entity in the environment causing the dye to alternate between two different

fluorescence states. This effect can be overcome by using a different dye which may not interact in the same manner, or alternatively by either mutating the surrounding residues or relocating attachment position of the dye so that the dye is not in proximity to potentially problematic side chains. Other environmental influences on the fluorophores can also introduce a large degree of variability and noise, making the data interpretation very challenging.

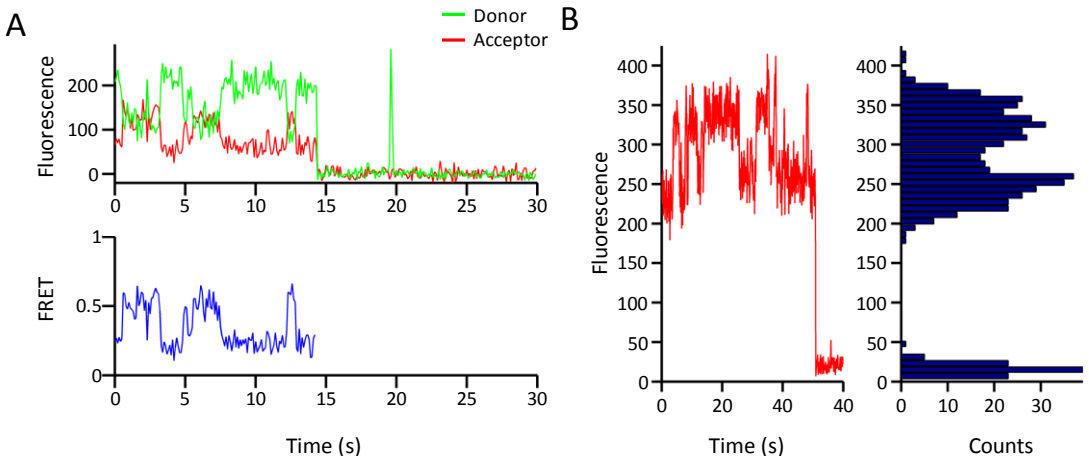


Figure 4.11. Variation of fluorescence signal observed on dye-labeled DNA polymerases. Fluctuations between two distinct fluorescence states can be due to environmental factors, such as the interaction of one of the dyes with local amino acid side chains. A. smFRET trace. This trace could be mistakenly interpreted as the oscillation of the protein between two unique structural conformations. B. Florescence trace of the acceptor dye with direct excitation. It is apparent that interactions of this fluorophore with its local environment cause it to fluctuate between two fluorescence states.

4.6 Conclusions

This work described in this chapter has provided a solid foundation to towards the development of technologies for high-speed, single-molecule imaging of engineered DNA polymerases and real-time single-molecule DNA sequencing. My work on the development of a novel method for engineering labeled proteins by cyclic *in vitro*

translation on solid supports was presented. Also described in detail is the large body of work on the characterization of an optical system for high-speed FRET imaging of single molecules to investigate the ability for real-time monitoring of conformational dynamics of DNA polymerase during DNA synthesis. While many challenges in the engineering of FRET-labeled DNA polymerase and single-molecule imaging remain and have been identified, we have demonstrated that nucleotide incorporation kinetics can be monitored in real-time using fluorescently-labeled DNA polymerase and high-speed single molecule FRET imaging. Further efforts will be required to demonstrate the feasibility of the READS technology.

4.7 References

1. Bentley, D.R., Balasubramanian, S., Swerdlow, H.P., Smith, G.P., Milton, J., Brown, C.G., Hall, K.P., Evers, D.J., Barnes, C.L., Bignell, H.R., Boutell, J.M., Bryant, J., Carter, R.J., Keira Cheetham, R., Cox, A.J., Ellis, D.J., Flatbush, M.R., Gormley, N.A., Humphray, S.J., Irving, L.J., Karbelashvili, M.S., Kirk, S.M., Li, H., Liu, X., Maisinger, K.S., Murray, L.J., Obradovic, B., Ost, T., Parkinson, M.L., Pratt, M.R., Rasolonjatovo, I.M., Reed, M.T., Rigatti, R., Rodighiero, C., Ross, M.T., Sabot, A., Sankar, S.V., Scally, A., Schroth, G.P., Smith, M.E., Smith, V.P., Spiridou, A., Torrance, P.E., Tzonev, S.S., Vermaas, E.H., Walter, K., Wu, X., Zhang, L., Alam, M.D., Anastasi, C., Aniebo, I.C., Bailey, D.M., Bancarz, I.R., Banerjee, S., Barbour, S.G., Baybayan, P.A., Benoit, V.A., Benson, K.F., Bevis, C., Black, P.J., Boodhun, A., Brennan, J.S., Bridgham, J.A., Brown, R.C., Brown, A.A., Buermann, D.H., Bundu, A.A., Burrows, J.C., Carter, N.P., Castillo, N., Chiara, E.C.M., Chang, S., Neil Cooley, R., Crake, N.R., Dada, O.O., Diakoumakos, K.D., Dominguez-Fernandez, B., Earnshaw, D.J., Egbujor, U.C., Elmore, D.W., Etchin, S.S., Ewan, M.R., Fedurco, M., Fraser, L.J., Fuentes Fajardo, K.V., Scott Furey, W., George, D., Gietzen, K.J., Goddard, C.P., Golda, G.S., Granieri, P.A., Green, D.E., Gustafson, D.L., Hansen, N.F., Harnish, K., Haudenschild, C.D., Heyer, N.I., Hims, M.M., Ho, J.T., Horgan, A.M., Hoschler, K., Hurwitz, S., Ivanov, D.V., Johnson, M.Q., James, T., Huw Jones, T.A., Kang, G.D., Kerelska, T.H., Kersey, A.D., Khrebtukova, I., Kindwall, A.P., Kingsbury, Z., Kokko-Gonzales, P.I., Kumar, A., Laurent, M.A., Lawley, C.T., Lee, S.E.,

Lee, X., Liao, A.K., Loch, J.A., Lok, M., Luo, S., Mammen, R.M., Martin, J.W., McCauley, P.G., McNitt, P., Mehta, P., Moon, K.W., Mullens, J.W., Newington, T., Ning, Z., Ling Ng, B., Novo, S.M., O'Neill, M.J., Osborne, M.A., Osnowski, A., Ostadan, O., Paraschos, L.L., Pickering, L., Pike, A.C., Chris Pinkard, D., Pliskin, D.P., Podhasky, J., Quijano, V.J., Raczy, C., Rae, V.H., Rawlings, S.R., Chiva Rodriguez, A., Roe, P.M., Rogers, J., Rogert Bacigalupo, M.C., Romanov, N., Romieu, A., Roth, R.K., Rourke, N.J., Ruediger, S.T., Rusman, E., Sanches-Kuiper, R.M., Schenker, M.R., Seoane, J.M., Shaw, R.J., Shiver, M.K., Short, S.W., Sizto, N.L., Sluis, J.P., Smith, M.A., Ernest Sohna Sohna, J., Spence, E.J., Stevens, K., Sutton, N., Szajkowski, L., Tregidgo, C.L., Turcatti, G., Vandevondele, S., Verhovsky, Y., Virk, S.M., Wakelin, S., Walcott, G.C., Wang, J., Worsley, G.J., Yan, J., Yau, L., Zuerlein, M., Mullikin, J.C., Hurles, M.E., McCooke, N.J., West, J.S., Oaks, F.L., Lundberg, P.L., Klenerman, D., Durbin, R. and Smith, A.J. (2008) Accurate whole human genome sequencing using reversible terminator chemistry. *Nature*, **456**, 53-59.

2. Frommer, M., McDonald, L.E., Millar, D.S., Collis, C.M., Watt, F., Grigg, G.W., Molloy, P.L. and Paul, C.L. (1992) A genomic sequencing protocol that yields a positive display of 5-methylcytosine residues in individual DNA strands. *Proc Natl Acad Sci U S A*, **89**, 1827-1831.

3. Laszlo, A.H., Derrington, I.M., Ross, B.C., Brinkerhoff, H., Adey, A., Nova, I.C., Craig, J.M., Langford, K.W., Samson, J.M., Daza, R., Doering, K., Shendure, J. and Gundlach, J.H. (2014) Decoding long nanopore sequencing reads of natural DNA. *Nat Biotechnol*, **32**, 829-833.

4. Korlach, J., Bibillo, A., Wegener, J., Peluso, P., Pham, T.T., Park, I., Clark, S., Otto, G.A. and Turner, S.W. (2008) Long, processive enzymatic DNA synthesis using 100% dye-labeled terminal phosphate-linked nucleotides. *Nucleosides Nucleotides Nucleic Acids*, **27**, 1072-1083.

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

6. Laszlo, A.H., Derrington, I.M., Brinkerhoff, H., Langford, K.W., Nova, I.C., Samson, J.M., Bartlett, J.J., Pavlenok, M. and Gundlach, J.H. (2013) Detection and mapping of 5-methylcytosine and 5-hydroxymethylcytosine with nanopore MspA. *Proc Natl Acad Sci U S A*, **110**, 18904-18909.

7. Eid, J., Fehr, A., Gray, J., Luong, K., Lyle, J., Otto, G., Peluso, P., Rank, D., Baybayan, P., Bettman, B., Bibillo, A., Bjornson, K., Chaudhuri, B., Christians, F., Cicero, R., Clark, S., Dalal, R., Dewinter, A., Dixon, J., Foquet, M., Gaertner, A., Hardenbol, P., Heiner, C., Hester, K., Holden, D., Kearns, G., Kong, X., Kuse, R., Lacroix, Y., Lin, S., Lundquist, P., Ma, C., Marks, P., Maxham, M., Murphy, D., Park, I., Pham, T., Phillips, M., Roy, J., Sebra, R., Shen, G., Sorenson, J.,

Tomaney, A., Travers, K., Trulson, M., Vieceli, J., Wegener, J., Wu, D., Yang, A., Zaccarin, D., Zhao, P., Zhong, F., Korlach, J. and Turner, S. (2009) Real-time DNA sequencing from single polymerase molecules. *Science*, **323**, 133-138.

8. Christian, T.D., Romano, L.J. and Rueda, D. (2009) Single-molecule measurements of synthesis by DNA polymerase with base-pair resolution. *Proc Natl Acad Sci U S A*, **106**, 21109-21114.

9. Berezhna, S.Y., Gill, J.P., Lamichhane, R. and Millar, D.P. (2012) Single-molecule Forster resonance energy transfer reveals an innate fidelity checkpoint in DNA polymerase I. *J Am Chem Soc*, **134**, 11261-11268.

10. Markiewicz, R.P., Vrtis, K.B., Rueda, D. and Romano, L.J. (2012) Single-molecule microscopy reveals new insights into nucleotide selection by DNA polymerase I. *Nucleic Acids Res*, **40**, 7975-7984.

11. Santoso, Y., Joyce, C.M., Potapova, O., Le Reste, L., Hohlbein, J., Torella, J.P., Grindley, N.D. and Kapanidis, A.N. (2010) Conformational transitions in DNA polymerase I revealed by single-molecule FRET. *Proc Natl Acad Sci U S A*, **107**, 715-720.

12. Shimamoto, N. (1999) One-dimensional diffusion of proteins along DNA. Its biological and chemical significance revealed by single-molecule measurements. *J Biol Chem*, **274**, 15293-15296.

13. Rasnik, I., McKinney, S.A. and Ha, T. (2006) Nonblinking and long-lasting single-molecule fluorescence imaging. *Nat Methods*, **3**, 891-893.

14. Di Fiori, N. and Meller, A. (2010) The Effect of dye-dye interactions on the spatial resolution of single-molecule FRET measurements in nucleic acids. *Biophys J*, **98**, 2265-2272.

15. Chen, H., Ahsan, S.S., Santiago-Berrios, M.B., Abruna, H.D. and Webb, W.W. (2010) Mechanisms of quenching of Alexa fluorophores by natural amino acids. *J Am Chem Soc*, **132**, 7244-7245.

5 SUMMARY AND FUTURE WORK

5.1 DNA Polymerase Incorporation Kinetic Measurements

We have demonstrated a method for measuring the single-base extension kinetics of fluorescently labeled nucleotides by DNA polymerase. To obtain adequate signal-to-noise, TIRF microscopy was used, which minimized the penetration depth of the laser to a few hundred nanometers above the imaging surface thereby reducing the excitation volume. Custom-synthesized coumarin-labeled nucleotides were incorporated by Bst DNA polymerase into a high-density of surface immobilized primed DNA templates. By monitoring the incorporation reaction kinetics at a range of concentrations for each nucleotide, a curve can be generated that details the dependence of reaction rate on nucleotide concentration. From this data the enzyme turn-over rate and nucleotide dissociation constant can be extracted.

This strategy of determining incorporation kinetics offers several advantages over the currently used methods, such as stopped-flow or quench-flow. A main advantage is the measurements are performed in real-time, eliminating the need to perform subsequent analysis by gel electrophoresis. Additionally, these measurements can be made by any lab with a TIRF microscope, which eliminates the need for additional equipment to perform the stopped-flow or quench-flow reactions. Limitations do exist, such as the requirement of using fluorescently labeled nucleotides and a concentration threshold above which the background noise of fluorophores in solution will overwhelm the increase in surface

intensity as nucleotides are incorporated. It is possible to extract the incorporation kinetics of native nucleotides, however, by performing these incorporation kinetics using a mixture of native and fluorescently labeled nucleotides. Using a specific ratio of native to labeled nucleotides and measuring the incorporation kinetics of multiple concentrations, the turn-over rate and dissociation constant of the native nucleotides can be extracted if those parameters are known for the labeled nucleotide. It is also feasible to determine the incorporation bias of DNA polymerase towards native nucleotides over fluorescently labeled nucleotides by performing single molecule incorporation experiments with labeled primers and counting the fraction of primers that have a fluorescent nucleotide incorporated.

5.2 Capture and Enumeration of mRNA in a Microfluidic Device

We demonstrate the use of a microfluidic platform that allows for the capture, isolation, and lysis of individual HeLa cells, followed by the on-chip capture and enumeration of mRNA. The poly(A) tail on the mRNA was used to rapidly capture the transcripts onto a surface functionalized with poly(dT) oligos, where they were enumerated by single-molecule imaging. This process was performed on 9 individual cells in two separate devices and the mRNA capture numbers were found to vary significantly, as is expected from single-cell experiments. The surface immobilization of single-cell mRNA serves as a useful method of preparing transcriptome libraries for on-

chip analysis. Preliminary work was performed to demonstrate the ability to conduct single-molecule sequencing using Illumina's reagent set, which would enable the direct digital whole transcriptome profiling of single cells within a single integrated microfluidic device. The implementation of a transcript probing or sequencing scheme with a microfluidic device would provide a powerful tool for single-cell transcriptomics. By allowing the on-chip digital analysis of mRNA, more accurate gene expression profile would be possible at lower cost.

5.3 Technologies to Enable Single-Molecule Enzymology and DNA Sequencing

Our approaches and steps taken towards the implementation of a novel DNA sequencing method involving the use of single-molecule FRET to monitor DNA synthesis by a mutated Bst DNA polymerase are described. Our progress towards developing protein engineering methods involving solid-phase cyclic *in vitro* translation is outlined. The *in vitro* translation method relies on mRNA bound to a solid-support and a fluidic system to control the flow of custom-prepared translation reaction mixes with the desired charged tRNAs and labeled amino acids. While we were unsuccessful at demonstrating the ability to incorporate specific fluorescently labeled amino acids at designated locations, the ability to produce full length enzymes with fluorescently labeled cysteine residues using surface-bound mRNA for translational templates was established.

Our microscopy system was characterized to determine the ability to the high-speed and high-sensitivity imaging required to perform fluorescence detection of the donor and acceptor dyes necessary for single-molecule FRET monitoring of DNA synthesis. Incorporation reactions were performed on surface-bound DNA using dual-labeled polymerases. We were unable to decipher the FRET trace of the dual-labeled polymerase to identify individual nucleotide incorporation, however, it appears that there was a distinguishable change in FRET efficiency upon the introduction of the nucleotides into the flow cell which could indicate the transition to a DNA synthesis state. To enable the real-time monitoring of DNA synthesis by dual-labeled polymerases, improvements to the imaging system, such as using scientific CMOS cameras which have higher frame rates, would assist in detecting rapid fluctuations in FRET efficiency. The choice of donor and acceptor dyes, as well as their spacing on the polymerase, is also critical to maximize the change in FRET efficiency with the structural movement of the enzyme. All methods could benefit from improved surfaces that provide a highly specific and rapid capture of the desired target, while reducing the non-specific binding of unwanted molecules.

5.4 Impact

This dissertation covers a range of applications and approaches for using fluorescence microscopy and microfluidic devices for performing cellular and enzymatic anlysis. By combining the fluidic control required to perform multi-step cell processing and biochemical reactions with molecular detection capabilities, new tools for biological

analysis were developed. Through further refinement of these and similar technologies, the scientific and medical community will be better able to study the molecular basis of diseases to continue improving healthcare.