Louisiana State University
LSU Digital Commons

LSU Doctoral Dissertations | Graduate School

2012

Single-molecule detection of molecular beacons generated from LDR on thermoplastic microfluidic device for bioanalysis

Zhiyong Peng
Louisiana State University and Agricultural and Mechanical College, zpeng2@tigers.lsu.edu

Follow this and additional works at: https://digitalcommons.lsu.edu/gradschool_dissertations

 Part of the Chemistry Commons

Recommended Citation
Peng, Zhiyong, "Single-molecule detection of molecular beacons generated from LDR on thermoplastic microfluidic device for bioanalysis" (2012). *LSU Doctoral Dissertations*. 843.
https://digitalcommons.lsu.edu/gradschool_dissertations/843

This Dissertation is brought to you for free and open access by the Graduate School at LSU Digital Commons. It has been accepted for inclusion in LSU Doctoral Dissertations by an authorized graduate school editor of LSU Digital Commons. For more information, please contactgradetd@lsu.edu.

SINGLE-MOLECULE DETECTION OF MOLECULAR BEACONS GENERATED FROM LDR ON THERMOPLASTIC MICROFLUIDIC DEVICE FOR BIOANALYSIS

A Dissertation
Submitted to the Graduate Faculty of the
Louisiana State University and
Agricultural and Mechanical College
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

in

The Department of Chemistry

by
Zhiyong Peng
B.S., Tianjin University, Tianjin, China, 1995
M.S., Louisiana State University, 2003
May, 2012

ACKNOWLEDGEMENTS

The author would like to thank his collaborators first. This work is finished through a multidisciplinary approach that leads to the final completion of the project. The Soper group members are always willing and ready to provide help anytime when needed. Hong Wang imparted me valuable knowledge and skills on PCR, LDR, and other molecular biology techniques. Jason Emory and Wonbae Lee shared with me thoughts in single-molecule detection techniques. In addition, Jason Guy was always enthusiastic and helpful in fabricating my mold insert. George Gascon did a terrific job in processing my machining submissions and made any effort to have my job done as soon as he can. Proyag Datta helped me making the COC microfluidic device using the hot-embossing facility in CAMD. Moreover, my collaborator, Maneesh Pingle in Medical College of Cornell University, provided me insights on designing the primers for pathogen detection work. Lloyd Davis provided the software for counting the photon burst events in single-molecule detection.

Most importantly I would like to express my deep gratitude to my advisor Dr. Soper. You spent a tremendous amount of time and energy mentoring me and helping me grow from a student to a professional and scientist. Thank you!

TABLE OF CONTENTS

ABSTRACT

Current clinical techniques for nucleic acid detection and analysis often involve PCR, lacking adequate specificity or sensitivity to meet the stringent requirements in certain applications. This research aims to develop an innovative molecular assay and the associated hardware to rapidly signal the presence of certain targets using reporter sequence found in their genome without requiring PCR. This assay coupled the sensitivity of single-pair fluorescence resonance energy transfer (spFRET) with the specificity of ligase detection reaction (LDR) to provide near real-time readout of target biomarkers. Heightened concerns on potential bioterrorism threats, such as rapid dissemination of pathogenic bacteria or viruses into water and/or food supplies, demand fast detection strategies. In this work, a pair of strain-specific primers was designed based on the 16S rRNA gene and were end-labeled with a donor (Cy5) and acceptor (Cy5.5) dyes. In the presence of the target bacterium, the primers were joined using LDR to form a reverse molecular beacon (rMB), thus bringing Cy5 and Cy5.5 into close proximity to allow FRET to occur. These rMBs were analyzed using single-molecule detection of the FRET pairs (spFRET). The LDR was performed in a Cyclic Olefin Copolymer (COC) microfluidic device equipped with 2 or 20 thermal cycles in a continuous flow format. Single-molecule photon bursts from the resulting rMBs were detected on-chip and registered using a laser-induced fluorescence (LIF) instrument. The presence of target pathogens could be reported in as little as 2.6 min using spFRET. In another development, a similar assay format was utilized to quantify mRNA expression levels of MMP-7 gene, which is highly relevant to invasion, metastasis and progression of a variety of tumors. HT-29 cells were found to express the highest levels of MMP-7 transcripts among the studied cell lines using LDR primers specific to MMP-7

gene. This observation is consistent with the results obtained with RT-qPCR. The LDR-spFRET assay was also used for stroke subtyping by designing primers specific to AMPH gene and using a microfluidic chip with tapered detection window to improve sampling efficiency. The detection could be completed in ≤15 min with extended readout time to glean low copy number transcripts.

CHAPTER 1 INTRODUCTION

1.1 Background

Advancement of genome sequencing technology has allowed for determining the nucleic acid sequence of any species. Deoxyribonucleic acid (DNA) is a linear biopolymer, most of which contain two single stranded oligonucleotides that are intertwined with each other to form a double helix structure in its natural state. The oligonucleotides strand are a combination of four different types of deoxyribonucleic acid monomers, each of which consists of a sugar-phosphate unit attached with one of the four bases (A, G, C, T). The permutation of the four types of monomers with different bases along the oligonucleotide strand defines the genetic information carried by most living organisms.

Ribonucleic acid (RNA), by contrast, is another important biomolecule and most RNA molecules are single-stranded and adopt complex three-dimensional structure.[1] Unlike DNA, whose backbone is composed of deoxyribose sugar units, RNA has ribose sugar units in its backbone and it contains a uracil (U) in its four nucleobases compared to DNA which contains a thymine (T) instead. Figure 1.1 makes a side-by-side comparison of the structures of DNA and RNA.

Recognition of certain nucleic acid sequences, which consists of determining the order of the monomer units (*i.e.*, primary structure), is one of the most established tools in modern biological studies and have found broad applications in many fields such as identification of suspects in forensic investigations, diagnosis/prognosis of cancer and drug resistance, recognition of infectious organisms in clinical diagnostics, and identification of microorganisms in food industry and environment monitoring.[2]

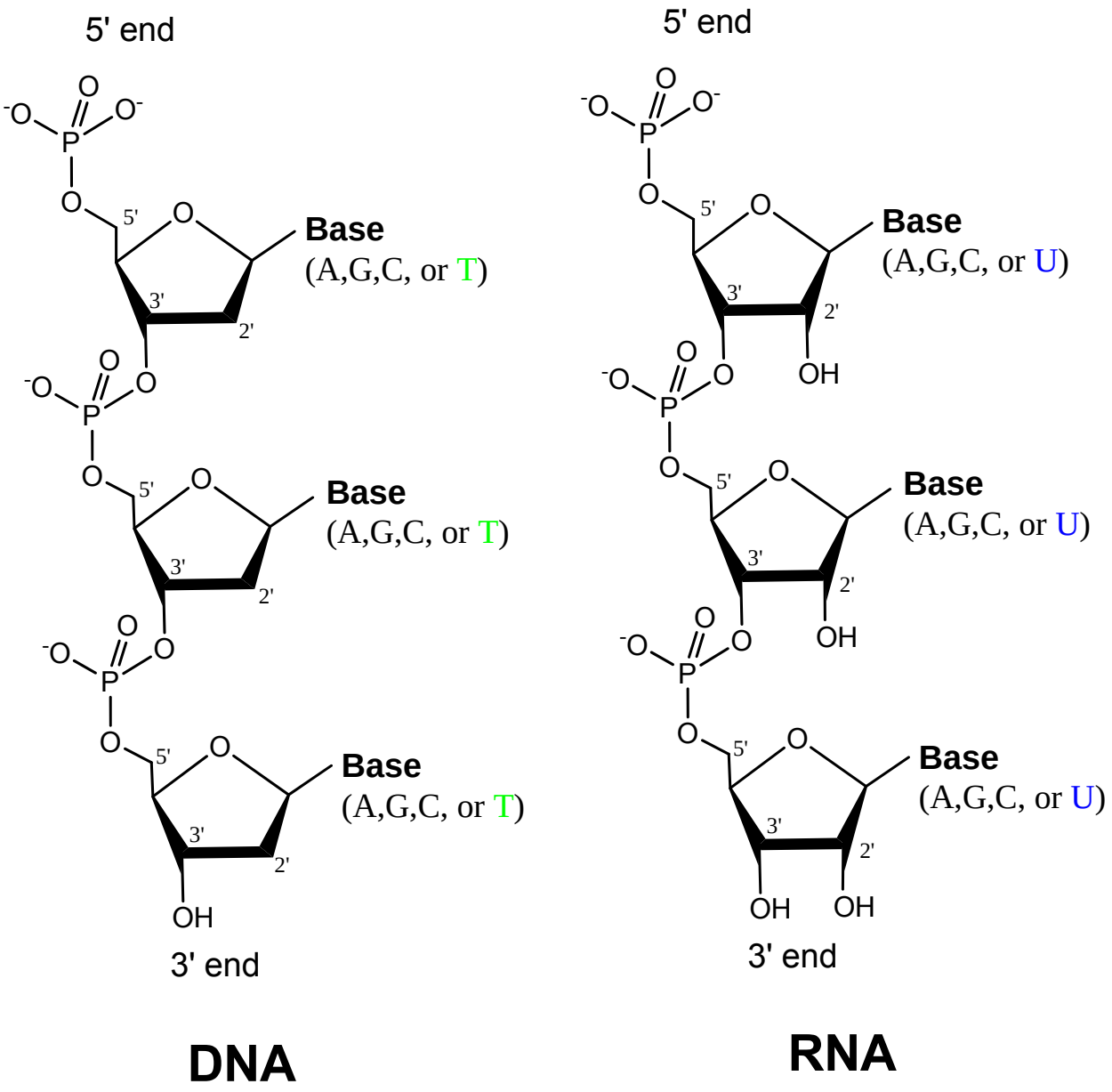


Figure 1.1 Comparison of DNA and RNA molecular structures

1.1.1 Human Genomics

Proteins are essential components of cells, and they are encoded from genes, the functional oligonucleotide sequence. Advances in DNA sequencing technology has practically allowed researchers to secure genetic information of any living organisms, and made it possible to identify and screen those protein-coding genes more rapidly.

With completion of the Human Genome Project (HGP) through a tremendous international effort for 13 years, people were able to obtain the whole genome sequence of the human being,[3, 4] which provided an unprecedented opportunity to better understand the role of genetic components in human health and disease, and facilitated further development of DNA-based biomedical diagnostics and therapeutics that highly relied on availability of accurate genome sequence. Toward the completion of HGP another multi-national effort, the International HapMap Project (IHP), was initiated to identify and catalog genetic similarities and differences between human beings and discover the causative genes for common human diseases and individual response to therapeutic medication and environmental factors.[5] Studies have shown that the human genome of any two individuals differ only by approximately 0.1%, and the most common genetic variations occur at single sites, the so-called single nucleotide polymorphism (SNP).[6, 7] Through phase I and phase II of the IHP, over 3.1 million SNPs in the human genome have been identified,[8] as well as numerous structural variants, including deleted, duplicated, or rearranged DNA segments,[9] which lays a solid foundation for genome-wide association studies to pinpoint those genes that are responsible for certain types of mutations related diseases, such as cancer.[10]

1.1.2 Bacteria Genomics

Bacteria are single-celled prokaryote microorganisms that inhabit ubiquitously on earth in soils, waters, organic environments like bodies of plants and animals to even extreme environments such as acidic hot springs and the deep crust of the earth. Due to their omnipresence, bacteria have found great impact on the health of human beings. It has been known that bacteria are much more diverse in genome structures and

nucleotide variations than their eukaryote counterpart.[11] Examining the sequence of the commonly encountered bacterium, *Escherichia coli*, demonstrated that the bacterial genomes possess enormous amount of diversity even in the same bacterial species,[12] therefore determining the sequence of bacteria genomes will be critical to quickly genotyping those harmful strains in diagnostics and enable preventative treatment in a timely manner. Since completion of the first bacterial genome sequencing in 1995,[13] hundreds of bacterial genomes have been successfully sequenced, which stimulated the interest in developing DNA-based diagnostic tools for pathogenic bacteria as well as more effective anti-biotic medicines and other types of treatment for bacterial infections.

1.1.3 Human Microbiomics

The human microbiome refers to all microorganisms, including prokaryotes, viruses, and microbial eukaryotes, that colonize in the human body. The term “Microbiome” was coined by Joshua Lederberg in 2001 to signify the crucial role that the microorganism community plays in the well being of mankind.[14] It was estimated that the total microbes residing in or on the human body outnumbers the cells in human body by 10 to 1, and they encode 100-fold more genes than our genome.[15] These microorganisms live widely in different body sites such as the gastrointestinal tract, urogenital/vaginal tract, oral cavity and skin, and the vast majority of the microbes reside in the gastrointestinal tract.[16, 17] The gut microbes have a great impact on human physiology and metabolism and their behavior is widely associated with many bowel diseases and chronic diseases like obesity and diabetes.[18-20] Therefore, the collective genes from all of these microbes residing in the human body are even regarded as the second genome of human beings.[21] The National Institute of Health (NIH) launched an ambitious Human

Microbiome Project (HMP) in 2007 following the accomplishment of the HGP, attempting to catalog and characterize the genes from a selected number of bacteria sampled from certain individuals and gain insights on their relationship with human health and disease.[14]

The overwhelming amount of genes that are involved in the human microbiome makes it extremely challenging in contrast to the human genome project. With the advancement of DNA sequencing technology and the invention of the next-generation sequencing platforms, which dropped the sequencing cost exponentially and increased the throughput by many-fold, people have successfully acquire an initial reference genome from 178 bacterial genomes from multiple body sites.[16]

1.1.4 Detection of Bacterial Pathogens

Pathogens, also called infectious agents, are a class of microorganisms including viruses and bacteria that form a parasitic association with their host. Bacterial pathogens are a major cause of human disease and death. Epidemics caused by pathogenic bacteria have taken place recurrently in human history and deprived lives from millions of people. Recent outbreaks of several well-known bacteria, such as *Salmonella* in peanut butter and *E. coli* O157:H7 in fresh spinach aroused extensive concerns on the risk of foodborne pathogens.[22, 23] Therefore, accurate and rapid identification of pathogen infection is of great importance to public health.

Early identification and characterization of bacterial pathogens relied on cultivation of bacteria followed by phenotypic recognition and biochemical tests, which is time-consuming, labor-intensive and lack of specificity. Many bacteria entities are not able to be identified by phenotypic approaches because there is no appropriate growth media

for their cultivation.[24] Genome sequencing technology has advanced drastically in the last decade and the entire genomic information of numerous bacteria species is now available for genotyping applications.

Pulsed-field gel electrophoresis (PFGE) is a popular technique for foodborne pathogen subtyping, where the genomic DNA from the interrogated pathogens was first isolated and cleaved by restriction enzymes and the resulting DNA fragments are subjected to PFGE with the band separation pattern compared to standards to find their featured fingerprint.[25]

PCR-based strategies utilize a pair of oligonucleotide sequence, called primers, to target a specific region in the bacterial genome to increase the copy number of low abundant DNA analytes. Coupled with DNA sequencing, this technique can find single nucleotide difference between similar strains and increase our understanding to those infectious diseases.[26]

Microarrays are a versatile tool to detect DNA, mRNA, or cellular samples in a parallel format on one single solid substrate. In DNA microarrays, millions of identical short oligonucleotide sequences are immobilized on individual microscopic spots on the array surface, and thousands of different probes can be assigned onto different spots, depending on the density of the microarray, to interrogate different target analytes simultaneously. The extraordinary high-throughput makes DNA microarray appealing in clinical microbiology and it has been employed extensively in various applications.[27]

1.1.5 Gene Expression Analysis of mRNA

With the human genome being completely sequenced and the genomes of more and more living organisms being deciphered, research interests have begun to migrate

to functional genomics, which aims to identify and characterize biological functions of each individual gene as well as gene networks at the cell, tissue and organ levels. The central dogma of molecular biology gave an accurate description on how the genetic information that was originally stored in the genome was first transferred to messenger RNA (mRNA), then to proteins, which are the final products of gene expression in the cell. This sequential flow of genetic information is illustrated in Figure 1.2.

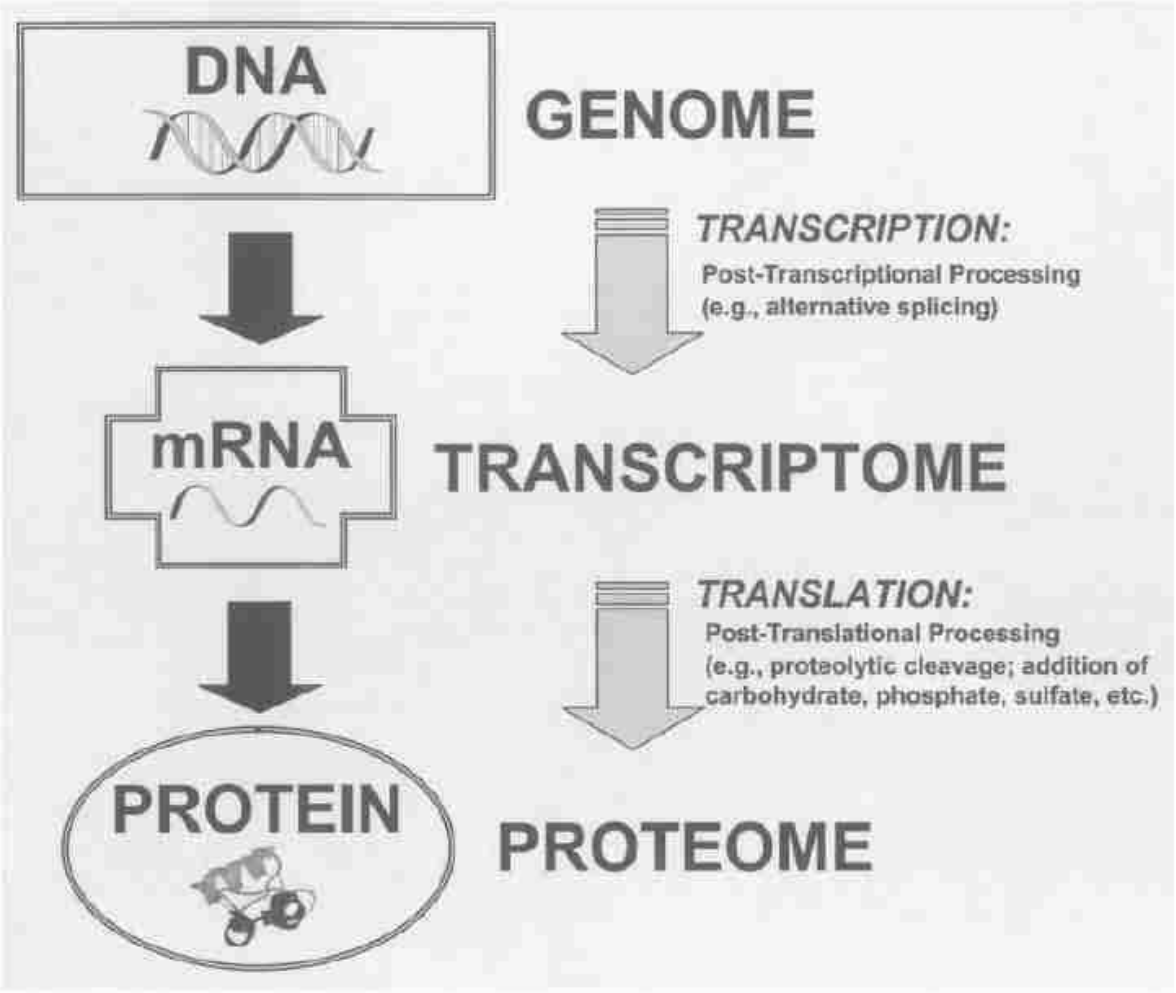


Figure 1.2 Illustration of sequential flow of genetic information from DNA to protein described in central dogma [28]

In the human genome 20,000 ~ 25,000 protein-encoding genes have been identified and their expression in the cells is precisely regulated by various intra- and extracellular

signals.[29] Although all the cells in the human body have the same genome, each individual cell essentially has its own expression pattern characterized by the abundance of certain transcripts (mRNAs). Misexpression of genes is usually indicators of a variety of human diseases such as cancer. Detecting the activity of these genes in the cells from different tissues or at development stages will provide both spatial and temporal information on transcription pattern, which could then be utilized to predict their relevance to certain cellular functions, biological processes or certain type of diseases.

Among the many techniques for gene expression analysis, northern blotting, RNase protection assay (RPA), in-situ hybridization (ISH), and real-time reverse transcription quantitative polymerase chain reaction (RT-qPCR) are the most commonly used ones for mRNA studies and quantification. Northern blotting was introduced in 1977 for detection of specific RNAs and often served as a gold standard in molecular biology labs for result confirmation.[30] In Northern blotting, the RNA sample is first separated via gel electrophoresis, and then transferred to a nylon membrane where the target mRNA can be hybridized to radiolabeled DNA or RNA probes and detected by X-ray films and quantified by densitometry.

RNase protection assay is a technique based on solution hybridization and is the method of choice when multiple targets need to be detected simultaneously. In this assay, an anti-sense RNA probe is used in the solution hybridization and the remaining unused probes and sample RNA are digested by nucleases. These nucleases are then deactivated and the probe-target hybrids are precipitated for further analysis through gel electrophoresis.[31]

In-situ hybridization is another attractive tool for gene expression analysis specifically useful for mRNA detection in cells or tissues. In this technique hybridization of probe with its target takes place inside the cell or tissue as evidenced by its name, which can provide unique spatial information of the mRNA within the sample.[32]

Real-time reverse transcription polymerase chain reaction is an extremely sensitive method for expression analysis and mRNA quantification and it has the capability of detecting mRNA levels from even single cells.[33] In RT-qPCR the mRNA is first converted into cDNA via reverse transcription, and then amplified through PCR to increase its copy number. The resulting PCR amplicons are quantified during the thermal cycling of PCR by mixing various fluorescent reporters into the reaction cocktail.[34]

1.2 Research Goals

Current techniques of pathogen detection and transcriptional expression analysis are time inefficient, taking from hours to days, which is not able to meet the need in most real-time applications. In this dissertation, my research is focused on developing a universal molecular assay in conjunction with single-molecule laser-induced fluorescence readout to rapidly report the presence of designated biomarkers on a thermoplastic microfluidic device with high sensitivity, reliability, and low limit-of-detection. In one development, the assay was deployed to detect bacterial pathogens using 16S ribosomal RNA as the biomarker to differentiate various strains of bacteria samples and provide a near real-time answer to their identities. In another development, a similar assay format was utilized to conduct expression analysis of messenger RNA of MMP-7 gene, which is highly correlated with invasion and progression of certain tumors.

The long-term goal of this research is to improve and expand this assay to a high throughput format, which can quickly and precisely perform the analysis of multiple biomarkers in parallel on a single microchip. Each microchannel on the chip is responsible for each individual strain specific pathogen and the concurrent signals from different channels can be read out by a multi-element detector, like a photodiode array or CCD. The whole system can be further integrated and miniaturized into a portable device, which has the flexibility to be used in applications under different circumstances.

1.3 Hybridization Based Genotyping Techniques

1.3.1 PCR

In most genetic analyses or diagnostic tests, PCR has been regarded as an indispensible step to identify target DNA with specific nucleic acid sequences. Briefly, PCR is a cyclic reaction mediated by a unique enzyme, called polymerase, to amplify single or double stranded oligonucleotides through repetitive thermal cycling. The specificity of PCR comes mainly from primers, which are short single-stranded oligonucleotides of ~20 base long and are designed to be complementary to specific sites on the target DNA. A typical PCR is composed of consecutive denaturing, annealing and extension steps that are performed at different temperatures. In the denaturing step, the oligonucleotide strand is heated to 90-97ºC to allow the double helix structure to be unraveled. When the temperature is lowered to 50-65ºC, which is selected to be in the neighborhood of the melting temperature (T_m) of the PCR primers, the pair of primers are annealed to the two strands of the template DNA through hydrogen bonding. In the extension step, four deoxyribonucleotide triphosphates (Adenine: dATP, Cytosine: dCTP, Thymine: dTTP, Guanine: dGTP) are added by the

polymerase to the 3' end of the primers one after another in a 5' to 3' direction according to Watson-Crick base pairing to form a long strand of oligonucleotide until the duplex is denatured again. These steps are carried out in a thermal cycler to control the required temperatures and repeated normally for 25-40 cycles. DNA copy number is nearly doubled in each thermal cycle and millions of copies of DNA are generated in an exponential fashion.

Early PCR experiments were performed using DNA polymerase extracted from *Escherichia coli*, which is a naturally occurring enzyme and susceptible to thermal inactivation at high temperatures. Thus, the amplification reaction had to be conducted at a temperature around 37ºC and a fresh aliquot of *Escherichia coli* DNA polymerase was added after each denaturation step, which is very inconvenient.[35] The low annealing temperature also led to nonspecific hybridization and it turned out that only about 1% of the PCR product was the targeted sequence.[36] This technical hurdle was not overcome until the discovery of thermally stable *Taq* DNA polymerase that was isolated successfully in 1988 from thermophilic bacterium, *Thermus aquaticus*, which can easily grow in geysers at a temperature of ≥110ºC.[37] The introduction of heat-resistant *Taq* polymerase actually revolutionized PCR technology from various aspects. It allowed primer annealing and extension to be carried out at higher temperatures, up to 94ºC, which greatly increased hybridization specificity by reducing mismatch hybridization to non-target sequence. It also increased reaction efficiency and yield, making longer amplicons available by minimizing secondary structure of genomic DNA at higher temperatures and simplified the automation of PCR by eliminating the tedious polymerase addition in each thermal cycle.[36-38]

Since its inception PCR has gained widespread popularity in the scientific community and has become a prevalent tool in clinical diagnosis,[39] microbiology examination,[40] infectious disease identification,[41] and gene expression analysis.[42] Despite its popularity, PCR does have limitations in some respects such as specificity, quantification, and turnaround time. For example, PCR is not able to provide the required specificity when distinguishing two similar alleles differing by only a single base pair such as AZT resistant HIV that featured multiple point mutations.[43]

Generally PCR proceeds through an exponential phase and a linear phase, and an endpoint measurement of the PCR product will not be able to adequately determine the initial copy number of the targeted DNA. Besides, PCR is prone to carryover contamination and primers design has to be exquisitely optimized to minimize mis-priming and false amplification. In addition, the long processing time due to repetitive thermal cycling makes it inappropriate for real-time applications. Thus, a molecular assay without PCR is needed to meet special requirements in certain applications.

1.3.2 LCR and LDR

Ligase chain reaction (LCR) is an alternative form of PCR for nucleic acid amplification that is specifically useful to detect a single base mutation.[44] Unlike PCR, which amplifies the target DNA by adding one nucleotide after another to the hybridized primer, LCR generates the amplicons by directly joining the two immediately adjacent oligonucleotide probes that are hybridized to their complementary target DNA. The ligation will occur only if the oligonucleotides are perfectly paired to the complements and have no gaps between them, and therefore a single-base substitution can be detected. The strategy of using DNA ligase for genetic analysis was first demonstrated

by Landegren for identifying specific alleles in human β globin gene that are differentiated by a single nucleotide substitution.[45] The experiment in this demonstration was conducted at 37ºC using T4 DNA ligase, which could possibly render ligation to mismatched hybridization. Barany and coworkers greatly advanced this technique by successfully isolating and cloning thermophilic DNA ligase from *Thermus aquaticus* and applied it to a Ligase Chain Reaction (LCR), the term that Barany coined, to distinguish human β globin alleles with pinpoint accuracy.[46] This thermostable ligase can survive repetitive denaturation at 94ºC and is highly active when performing at 65ºC for ligation making it unnecessary to add fresh ligase after each denaturation step when conducted in a thermal cycler. Compared to the mesophilic enzyme T4 ligase, it is much more specific and sensitive and can discriminate a single-base mismatch with a signal-to-noise of 75-500.[44]

The strategy of LCR is illustrated in Figure 1.3.[47] In LCR, four oligonucleotide probes are designed with two adjacent probes uniquely flanking one strand of the denatured target DNA and the other two probes flanking the complementary strand of target DNA. The two adjacent oligonucleotide probes are juxtaposed in a way that the 3' end of one probe is immediately adjacent to the 5' end of the other upon hybridizing to the target, therefore these two ends can be covalently linked by the thermostable DNA ligase to form a phosphodiester bond provided that the nucleotides at the junction are correctly base-paired with the target strand. Because in each reaction cycle the number of ligated oligonucleotides is doubled and the ligation product can serve as the DNA template for the next round of reaction, the target DNA is amplified in an exponential fashion, analogous to PCR. The mismatch at the ligation junction will not be amplified due to

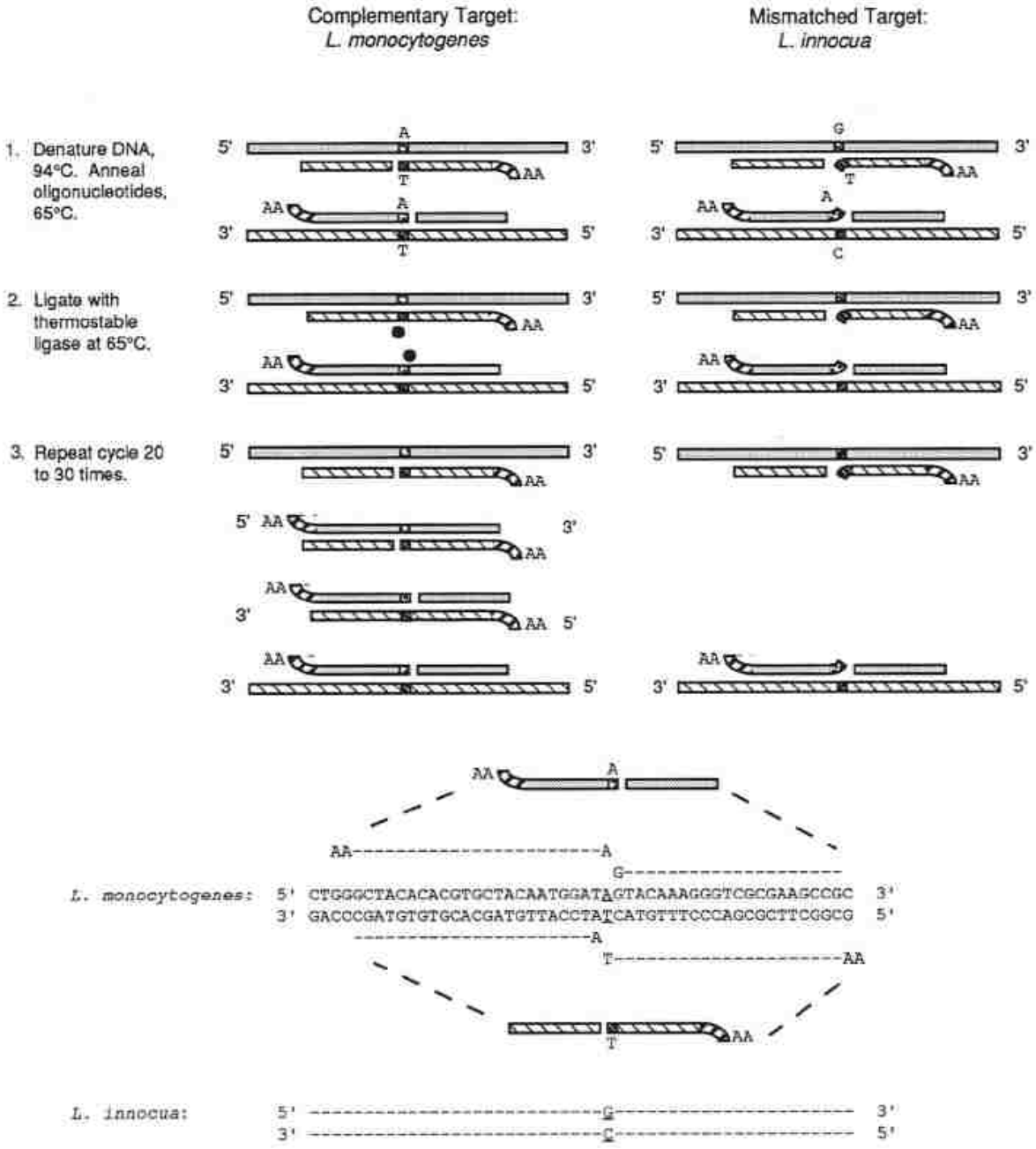


Figure 1.3 Schematic illustration of ligase chain reaction.[47] The matched target shown in this example is *L. monocytogenes* and mismatched target is *L. innocua*. They are distinguished from each other by a single base-pair difference in the 16S rDNA. *L. monocytonenes* has an A-T base-pair at nucleotide 1258, whereas *L. innocua* has a G-C base-pair at this position.

fidelity of DNA ligase and is thus distinguished. The same concept of using DNA ligase for amplification and discrimination of single-nucleotide mismatch was also adopted and

developed simultaneously in other research groups, and termed as oligonucleotide ligation assay (OLA) for a variety of applications.[48-50]

Ligase detection reaction (LDR) was derived from the concept of LCR and was shown in Figure 1.4.[44] In LDR, two oligonucleotide probes (termed common primer and discriminating primer, respectively) instead of four as in LCR are utilized and hybridized to only one strand of the denatured target DNA. When the two primers match exactly to the sequence of target DNA, a new oligonucleotide is formed by covalently linking the 3' end of discriminating primer and the 5' end of its adjacent common primer. This oligonucleotide differs from its parent primers by size and can be separated by running gel electrophoresis. Even if the discriminating primer has only one base mismatch to the target DNA, these two primers will no be linked successfully to form a longer oligonucleotide. Each cycle of LDR creates a new copy of template DNA and thus the target DNA is amplified in a linear fashion. In practice, LDR is usually preceded by a primary PCR to first generate enough DNA templates. Then, the LDR is subsequently carried out on these PCR products. The PCR-LDR assay has been used to detect low-abundant DNA point mutations in various diseased-related genes[51-53] and clinical microbiology tests.[54, 55] Another merit of LDR is its multiplexing capability when typing single-nucleotide polymorphisms (SNPs) at multiple loci of the gene at the same time.[56, 57] Either allele specific probes can be designed to have unique lengths, so that the wild type and mutant type ligation products can be separated by size, or they can be labeled with different fluorescent reporters, so that they can be discriminated by color.

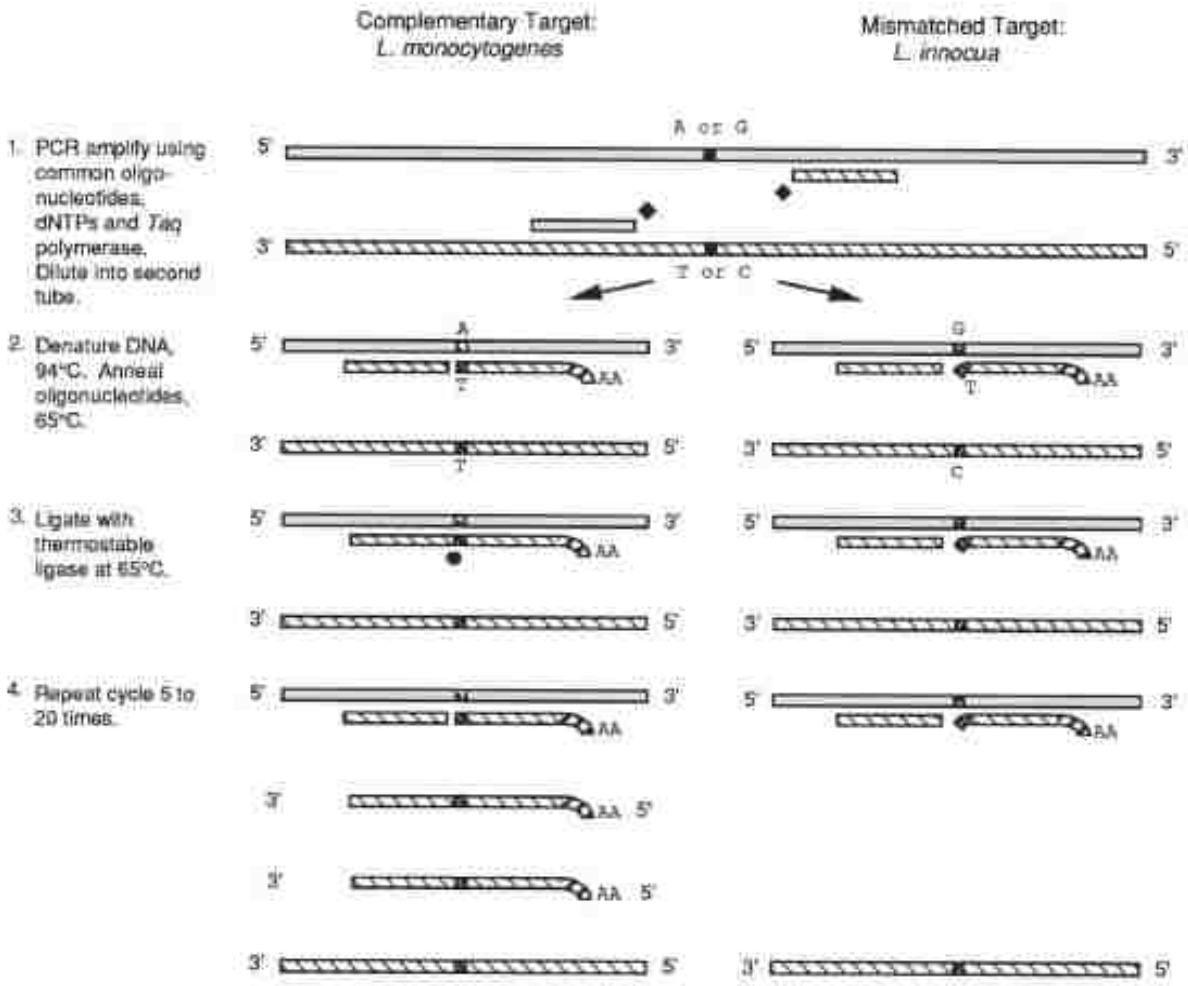


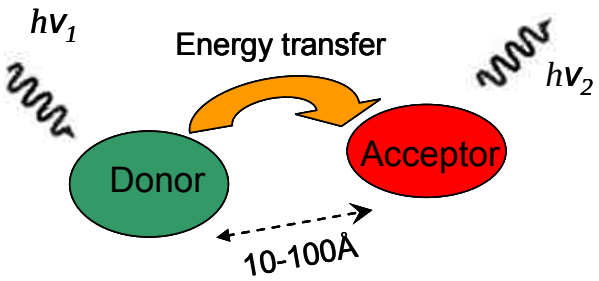
Figure 1.4 Schematic illustration of ligase detection reaction.[47] The matched target *L. monocytogenes* and mismatched target *L. innocua*. are distinguished from each other by a single base-pair difference in the 16S rDNA.

1.4 FRET

1.4.1 Principle of FRET

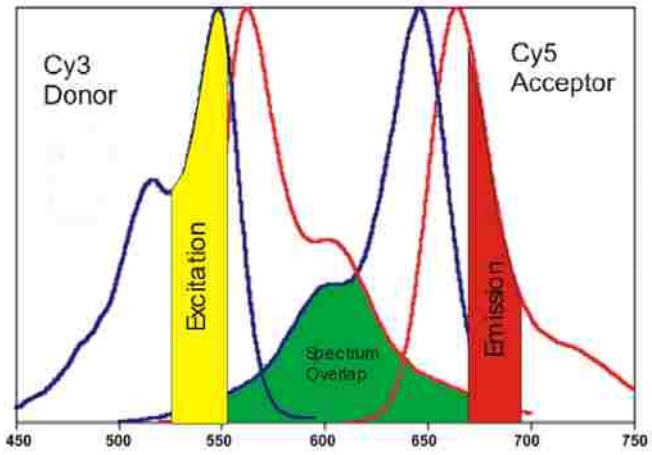
Fluorescence Resonance Energy Transfer (FRET) is a photophysical process first introduced by Stryer and Haugland as a spectroscopic ruler to measure distances in macromolecules[58] and has been revitalized in recent years and widely applied to biological studies attributed to development of new fluorescent dyes, and advancement in optics and instrumentation.[59] The efficiency of energy transfer in FRET strongly relies on the distance between the donor fluorophore and the acceptor fluorophore, which is

usually in the range of a few nm (~20-90Å), a length-scale not easily accessible through other biophysical techniques in solution. The principle of FRET is illustrated in Figure 1.5. In FRET, a donor fluorophore (Cy3 in this example) is excited by incident light and the energy in its excited state can be transmitted nonradiatively to an acceptor fluorophore (Cy5 in this example) via a long-range dipole-dipole interaction provided that they are in close proximity. This leads to a decrease in donor's emission intensity and excited state lifetime, and an increase in acceptor's emission intensity.



Donor: Cy3
Acceptor: Cy5

A: Fluorescence Energy Transfer from Donor to Acceptor



B: Spectral Overlap between Donor and Acceptor

Figure 1.5 Schematic representation of principle of FRET

Several prerequisites have to be met in order for FRET to occur: (a) The donor fluorophore and the acceptor molecules have to be in close proximity, typically 20~90 Å; (b) there is a large spectral overlap between the emission of the donor and the absorption of the acceptor, and the degree of spectral overlap is described by the integral $J(\lambda)$; (c) the transition dipole of the donor must be approximately parallel with that of the acceptor, and their relative orientation is described by κ^2. The efficiency of energy transfer relies on the distance between the donor and the acceptor molecules. Their relationship was first described by Förster in the following equation:[60]

$$E = \frac{R_O^6}{R_O^6 + r^y} \qquad (1.1)$$

where E is the efficiency of energy transfer, r is the actual distance between donor and acceptor, R_O is Förster distance, which is determined by factors such as κ^2 and $j(\lambda)$.

This equation clearly shows that the efficiency of resonance energy transfer in FRET is highly dependent on the distance *r* between the donor-acceptor pair, and is a reciprocal of the sixth power of this distance. For example, the transfer efficiency is 50% when $r = R_o$, and drops precipitately to only 1.6% when $r = 2R_o$.The predicted distance dependence of transfer efficiency has been experimentally demonstrated using poly-L-proline, which is an oligomer labeled with a naphthyl (donor) and a dansyl (acceptor) on its opposite ends.[58, 61] The distance between the donor and acceptor moieties is fixed because poly-L-proline forms a rigid helix structure with known atomic dimensions. By varying the numbers of proline residues, the transfer efficiency was confirmed to decrease as $1/r^6$. However, most samples in solution don't have a fixed donor-acceptor separation as in the poly-L-proline molecule, and the calculation of transfer efficiency is

complicated by their transient spatial separation that is constantly changed. Therefore, transfer efficiency is usually averaged in real applications over a number of potential spatial configurations.

1.4.2 Measurement of FRET Efficiency

Typically FRET efficiency is determined by comparing the donor fluorescence intensity of the sample to that of the donor only and FRET efficiency can be evaluated according to the following equation:

$$E = 1 - \frac{F_{DA}}{F_D} \quad (1.2)$$

where F_{DA} is the fluorescence intensity of the donor in the presence of the acceptor, and F_D is the fluorescence intensity of donor in the absence of acceptor.

This method depends on the concentration of the donor and acceptor in the sample, which has to be delicately matched to give an accurate measurement. Another popular method for determining transfer efficiency measures the lifetime of the donor. As we know, the fluorescence lifetime of the donor in its excited state will be shortened when FRET occurs and the energy is transferred to the acceptor. Thus, FRET efficiency can be determined by measuring the lifetime of the donor fluorophore in the presence and absence of the acceptor, respectively, and evaluated by the following equation:

$$E = 1 - \frac{\tau_{DA}}{\tau_D} \quad (1.3)$$

where τ_{DA} is the fluorescence lifetime of the donor in the presence of the acceptor, and τ_D is the fluorescence lifetime of the donor in the absence of the acceptor.

The advantage of this method is that lifetime of fluorophores does not dependent on concentration and it is possible to resolve lifetime of multiple fluorophores from the

same sample and determine the transfer efficiency for each donor-acceptor pair at the same time.

1.4.3 Applications of FRET

The Föster distance usually ranges from 20-90Å, which is comparable to the size of biological macromolecules, like proteins and the thickness of cell membranes. For a biomolecule that is fluorescently labeled on two different sites, a small change in distance between these two sites can cause a significant fluctuation in the energy transfer efficiency, which can be measured by the methods discussed above. This momentary change can be further used to investigate molecular interactions and dynamics like DNA interactions, protein folding, DNA structural dynamics, DNA-protein interactions.

An oligonucleotide can be used as a backbone to tag one or several fluorophores to form a DNA-FRET probe for many applications such as detection of hybridization, PCR, ligation, cleavage etc. For example, a double-stranded DNA probe was made by two complementary oligonucleotides to monitor a hybridization reaction.[62] One strand of the probe was 5' end labeled with fluorescein, and the other strand was 3' end labeled with a quencher of fluorescein. The target DNA sequence was detected by competitive hybridization of one of the oligonucleotides to form a probe duplex so that fluorescence from the donor was recovered and observed. Another example of FRET involves a dual-labeled double-stranded DNA probe to monitor the activity of PaeR7 endonuclease.[63] In this probe fluorescein was attached on its 5' end as a donor, and rhodamine was attached on its 3' end as an acceptor with a 6-mer spacer placed between them containing the PaeR7 recognition site. The extent of the cleavage reaction of this

double-stranded probe catalyzed by the PaeR7 endonuclease was then monitored in real time by observing the restoration of donor fluorescence and kinetic constant of this reaction can also be derived.

Another important application of FRET in biological research is to study DNA-protein interaction in situ. Cremazy et al. studied DNA binding to H2B histone, an abundant nuclear protein, and glucocorticoid receptor, a hormone-dependent transcription factor, in the cell nucleus using Fluorescence Lifetime Imaging Microscopy (FLIM).[64] The HeLa cells used in this study were genetically modified to express Green Fluorescent Protein (GFP), which was tagged to H2B histone and glucocorticoid receptor, respectively. Successful binding of the double-stranded DNA that was labeled with an acceptor fluorophore Sytox Orange, to H2B histone or glucocorticoid receptor caused the fluorescence lifetime of GFP to reduce, which was monitored by FLIM in real time.

Conformational changes in proteins, such as folding and unfolding, can modulate its functions, which is very important in many biological processes such as signaling pathways. The change of protein conformation is usually indicated by a change of specific atomic distances within the protein molecule, which can be reported dynamically by the change in FRET signal when these atoms are labeled with donor-acceptor pair. The challenge associated with this study is to successfully introduce two distinct fluorescent reporters to the desired sites within the same protein molecule using the same labeling chemistry. In some cases this hurdle can be overcome by reversibly blocking one labeling site when the other one is in use so that attachment of donor and acceptor can be conducted alternately.[65] In a study on conformational change of myosin, the commonest protein in muscle, single reactive cysteine residues have been

engineered to replace the native cysteines in *Dictyostelium* myosin-II so that they can be labeled with donor and acceptor selectively. By this approach, three conformational changes in myosin mediated by ATP-hydrolysis have been identified.[66]

1.5 Molecular Beacon

1.5.1 Principle of Molecular Beacon

Molecular Beacons (MB) are synthetic single-stranded oligonucleotide hybridization probes that can form a stem-loop structure.[67] The loop of the molecule, typically 10-40 bases long, is designed to be complementary to a predetermined region of the target DNA so that a hybridization reaction can take place whenever the target DNA is present in the sample solution. The stem of the molecule, contains two arm sequences, typically each 5-10 bases long, with equal length, which are complementary to each other, but not to the target DNA. At the end of each arm, a fluorescent molecule is attached to the end of one arm and a quenching molecule is attached to the end of the other arm. The working mechanism of molecular beacon is shown in Figure 1.6.

In its free state, the two arms of the probe anneal to each other by self-complementary sequences and the molecular beacon adopts a stem-loop structure. The fluorophore and quencher are brought in close proximity so that the fluorescence from the fluorophore is suppressed by the quencher due to fluorescence resonance energy transfer and direct energy transfer.[68] When the probe encounters the target oligonucleotides of interest, the stem of the beacon is denatured and the loop sequence is hybridized to its complementary sequence in the target oligonucleotide due to formation of more thermodynamically stable duplex. The fluorophore and the quencher are then spatially separated leading to restoration of fluorescence signal.

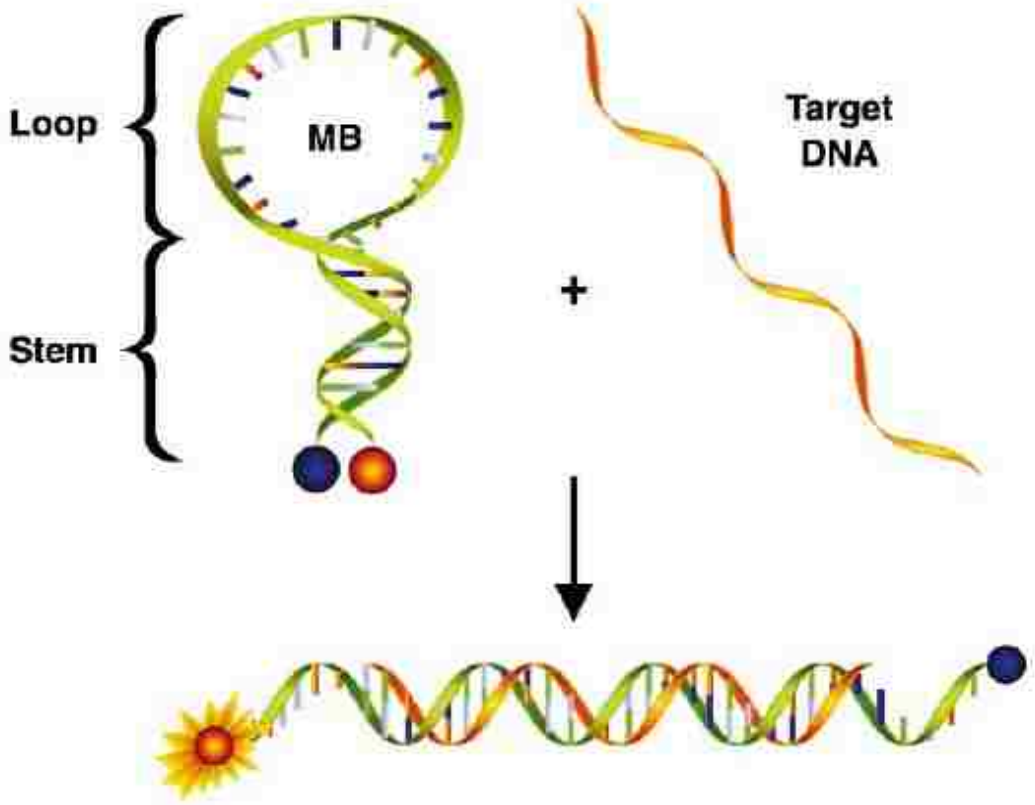


Figure 1.6 Conformational structure and working mechanism of molecular beacon probe [69]

1.5.2 Different Forms of Molecular Beacon

A modified version of a molecular beacon was developed termed wavelength-shift molecular beacon, and is shown in Figure 1.7.[70] This probe was designed with the same loop-stem scheme as the earlier version, but labeled with two fluorophores instead of one at the end of one arm and quencher at the end of the other arm. The first fluorophore, called harvest fluorophore, is directly attached to the 5' arm of the probe and has a wide absorbance spectrum and is excited by the incident light. The second fluorophore actes as a reporter, whose emission is monitored. This design allowed using the same monochromatic light as the excitation source, which was very convenient in my instrumentations.

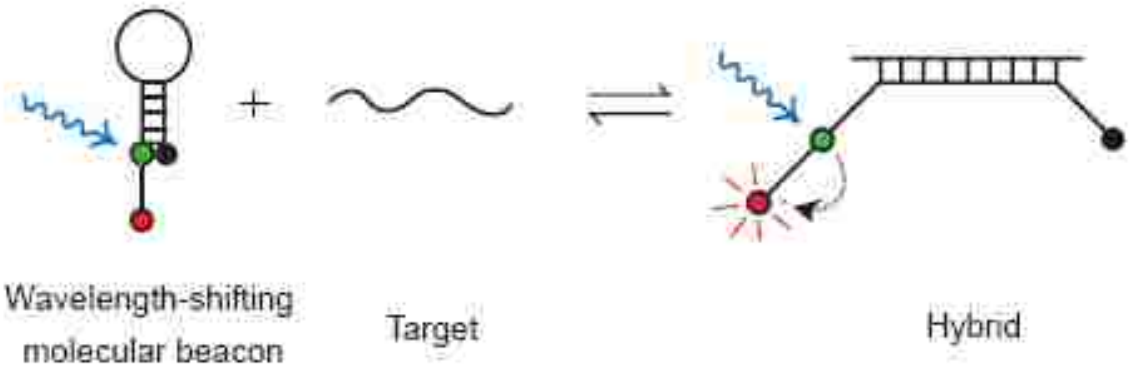


Figure 1.7 Illustration of wavelength-shifting molecular beacon [70]

When monitoring the ongoing nucleic acid events in living cells the conventional molecular beacon is prone to false positive signals due to interference from complex intracellular environment.[71, 72] For example, the molecular beacon could be degraded by nuclease or its hairpin structure could be opened by the nucleic acid binding proteins. A novel approach was developed by Tsourkas and Bao in order to overcome this problem by using two MB simultaneously.[73] The scheme is shown in Figure 1.8.

A pair of molecular beacons were designed each of which was labeled with a fluorophore and a quencher, separately. The fluorophore from one beacon served as a donor, and the fluorophore from the other beacon served as an acceptor in this dual FRET format. The sequences of the beacons were designed to be complementary to adjacent regions in a single nucleic acid target. Upon excitation of the donor fluorophore, FRET occurred only when both beacons were hybridized to the same oligonucleotide and the donor was brought into close proximity of the acceptor. Thus any fluorescence from nonspecific opening or degradation of the probes was excluded.

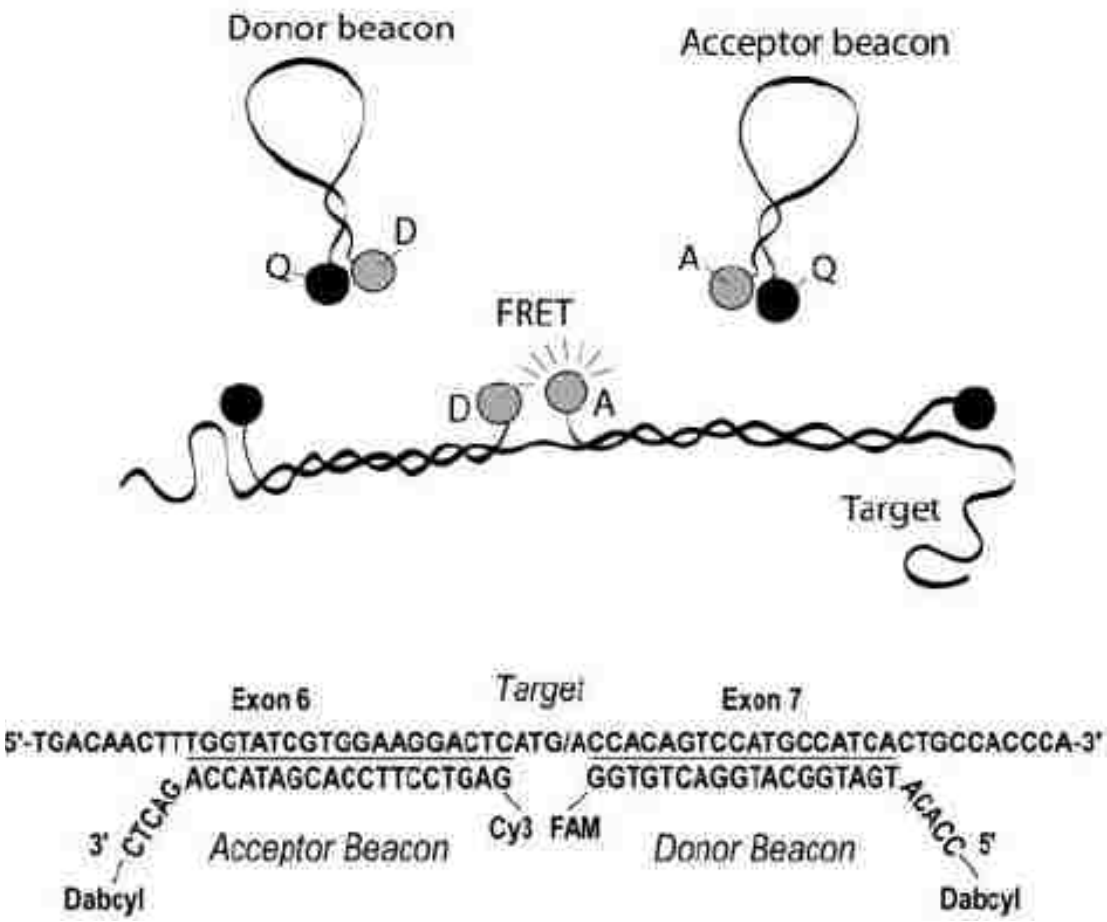


Figure 1.8 Schematic representation of dual FRET molecular beacons [73]

1.5.3 Advantages of Molecular Beacon

Molecular beacons are powerful molecular probes for nucleic acid recognition in homogeneous solution. The long sequence in the loop portion of the molecular beacon endows molecular beacon high selectivity in recognizing the target oligonucleotide, and the stem in the hairpin structure further destabilizes the formation of the duplex between the probe and mismatched target compared to linear oligonucleotide probes with equivalent length. This enhanced specificity due to formation of a stem-loop is a general feature of structurally constrained probes, and has been proven through thermodynamic

analysis using phase diagram.[74] Even a single nucleotide substitution in the target DNA sequence can be successfully distinguished due to its extraordinary selectivity.[74, 75]

Traditional hybridization assays involve a series of steps including: 1) labeling the hybridization probe, 2) immobilizing the probe onto a solid surface, 3) removing unhybridized probes from the mixture, 4) quantifying the remaining probes. This multi-step approach was not only time-consuming but also incapable of real-time monitoring hybridization events. Molecular beacons can measure the degree of hybridization without having to remove the unhybridized probes, because the unhybridized probes are in a closed loop and remain dark. This makes it possible to monitor in real-time nucleic acid hybridization events, which is especially important in living cells.[67]

Moreover, the design of molecular beacons makes it work in a way similar to an On-Off switch, and provide a high signal-to-background ratio and better detection sensitivity than other fluorescent probes. When the molecular beacon is hybridized to its target, the fluorescence intensity of the probe can be enhanced by 200-fold under optimal conditions.[75]

Another appealing property of molecular beacon probes is its multiplexing capability, by which a number of different samples can be detected in the same solution. The multiplexing capability of molecular beacons come from its unique stem-loop structure and signaling mechanism. On one hand, the loop sequence can be designed to contain oligonucleotides with either different lengths or different oligonucleotide sequences specific to each target DNA. On the other hand, each individual beacon can be assigned with a different fluorescent reporter, which can be quenched by a universal quencher like DABCYL. When multiple targets are present and found simultaneously an

array of colors will be emitted and observed accordingly. This approach has been used for multiplexed allele discrimination of different nucleic acid fragments.[75]

1.5.4 Applications of Molecular Beacon

The intriguing features of molecular beacons made it prevalent in the scientific community shortly after its invention in 1996 by Tyagi and Kramer, and its applications have been rapidly expanded into different fields in biomedical studies. For example, molecular beacon are often used to monitor the PCR progress to obtain quantitative information. Without gel electrophoresis, PCR products can be quantified by simply mixing the finishing reaction cocktail with molecular beacon probes in a microtiter well, where the probes are immobilized and the fluorescence intensity directly reflects the amount of PCR product.[76] But the more attractive feature for this application is that the ongoing PCR can be monitored in real-time by spiking the probes into the sealed PCR vial.[75] At the annealing step of each thermal cycle, the probes hybridize to their target amplicons and fluoresce. At the extension step when the temperature is increased, the probe-amplicon hybrid is dissociated and will not interfere with primer extension. The fluorescence level is registered in the last few seconds of each annealing step and directly indicates the number of amplicons produced as a result of the PCR process.

The rapid reporting of molecular beacon probes upon target hybridization make it a promising tool for rapid clinical diagnosis. In an example of identifying pathogenic retroviruses, a multiplex PCR was performed to amplify the retroviral DNA sequences in the same tube and four different molecular beacon probes were used as reporters. The abundance of retroviruses HIV-1, HIV-2, and human T-lymphotrophic virus type I and II were quantitatively detected.[77] In another example of molecular beacon coupled real-

time PCR assay, four different bacteria that could potentially be used as bioterrorism agents were successfully identified.[78]

Detecting, localizing and monitoring specific genes in living cells in real-time is a focus area in recent years and it provides a vivid image on the processing and transportation of specific mRNA in different physiological conditions. Current methods for *in vitro* gene expression analysis such as real-time PCR, in situ hybridization and Northern blotting are not applicable to *in vivo* analysis. The intrinsic capability of molecular beacons to identify targets without separation makes it exceptionally suitable for intracellular applications. It can be used to not only detect specific mRNA, but also visualize and track its sub-cellular location in real-time. Initial studies confirmed that molecular beacons are able to be used to hybridize specific gene transcripts in their native environment.[79, 80] Tyagi *et al.* demonstrated that molecular beacons can be used to visualize the distribution and transport of mRNA in living cells.[81] Cui *et al.* studied the virus-host cell interactions with the help of molecular beacons.[82] In this study, the viral nucleic acid, polio plus-strand RNA, was labeled with molecular beacon and imaged using a fluorescence microscope and its dynamic behavior and translocation in its host cells was patterned.

Sequence-specific DNA binding proteins are one of the most important class of proteins in living cells and plays a role in the regulation of many cellular processes, such as metabolism, immune response, and cell cycle. Abnormal function of these proteins often indicates development of certain diseases like cancer, which makes them become attractive diagnostic biomarkers that can potentially be used as targets for drug development.[83] Classical methods to detect DNA binding proteins such as gel shift

assays and ELISA involve multi-steps, which is time-consuming and not amenable to high throughput formats.[84] The straightforward binding-and-reporting fashion of molecular beacons allows them to capture any DNA binding protein in real time. Heyduk described a binary molecular beacon assay (shown in Figure 1.9) to detect CAP, a bacterial catabolite activator protein that binds DNA.[83] In this assay, two double-stranded oligonucleotide probes were made with each one containing about half of the sequence corresponding to the binding site. Both probes also contained a short single-stranded oligonucleotide overhang that were complementary to each other. One probe was tagged with a fluorophore and the other one was tagged with a quencher so that successful protein binding brought these probes adjacent and caused the fluorescence to be reduced.

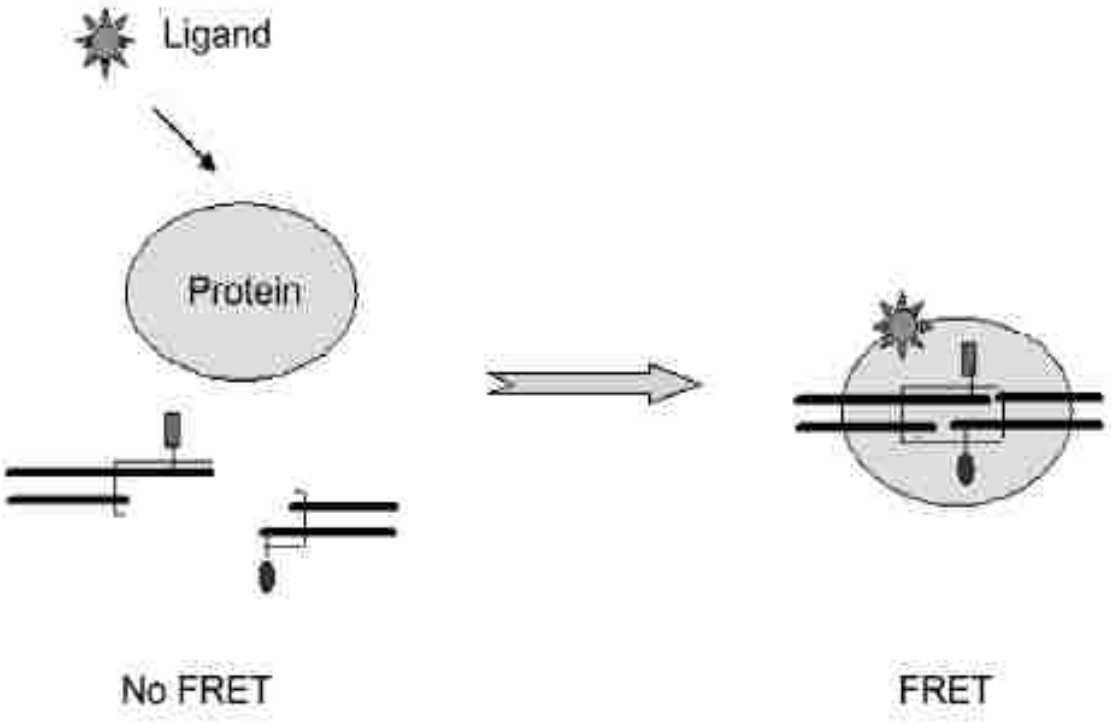


Figure 1.9 Illustration of molecular beacon for DNA-binding protein assay [84]

In addition to the applications mentioned above, molecular beacons were also used in SNP detection[85], real-time monitoring of ligation[86], phosphorylation[87]. Besides, molecular beacons can be used as the recognition element in the fabrication of biosensors, where they were immobilized onto a solid surface and arranged into an array format to allow parallel detection.[88, 89]

1.5.5 Limitations of Molecular Beacon

Although molecular beacons have excellent properties as a versatile tool for many challenging applications, the conventional format using a fluorophore-quencher pair does have limitations. Theoretically, the donor should be completely quenched when the pair stays in close proximity, with this distance defined by the Föster equation. In practice, however, there is always residual fluorescence from the donor due to incomplete quenching.[90] The magnitude of residual fluorescence varies greatly depending on factors including fluorophore/quencher pair selection, molecular beacon synthesis, and the way that the fluorophores are attached. This disadvantage impairs the sensitivity of molecular beacons and hinders their use in demanding situations, such as single-molecule detection where the SBS is critical. Therefore, molecular beacons with lower background fluorescence and better sensitivity need to be developed.

1.6 Single-Molecule Detection and Its Biological Applications

1.6.1 Overview of Single-Molecule Fluorescence Spectroscopy

In conventional biological or chemical analysis, the property that is measured comes from the ensemble average from the bulk analyte that is comprised of a plethora of molecules. This ensemble measurement, however, conceals the diversity of actual molecular behavior and keeps chemists from further understanding the characteristics of individual molecules. For example, the fluorescence lifetime of any given fluorophore

is not a fixed number and it is always distributed over a certain range; individual copies of an oligonucleotide or protein are featured by their heterogeneous behavior when they are at different configurations, different folded states, or different stages of an enzymatic cycle.[91] Therefore, being able to detect single molecules and characterize their individual properties will allow exploring the hidden heterogeneity of the seemingly identical molecules and interrogating their dynamic behavior on a one-by-one basis in complex local environments.

Studies of single-molecule have grown precipitously in the past years since its inception in the mid 1970s. The popularity of single-molecule studies can be reflected by the number of publications on this topic. A literature search of single-molecule methods from the PubMed database (www.pubmed.gov) with “single molecule” in the title yields nearly 3,000 publications, and the number of papers in each year over the past 30 years has increased significantly is shown in Figure 1.10. The search result from journals administrated by the American Chemical Society is represented as the blue bar and the result from the broader PubMed database is represented as a red bar. In this plot, we see an explosive growth of single-molecule studies after the year of 2000, which is probably due to its widespread acceptance by investigators in life sciences as an innovative experiment tool.

There are two major approaches to attain single-molecule sensitivity of target analytes: fluorescence and force manipulation. Force measurements on single-molecule basis include atomic force microscopy (AFM) and optical tweezers, which will not be discussed in this dissertation. Thus, the following discussion will center on fluorescence techniques for single-molecule experiments, especially laser-induced fluorescence (LIF).

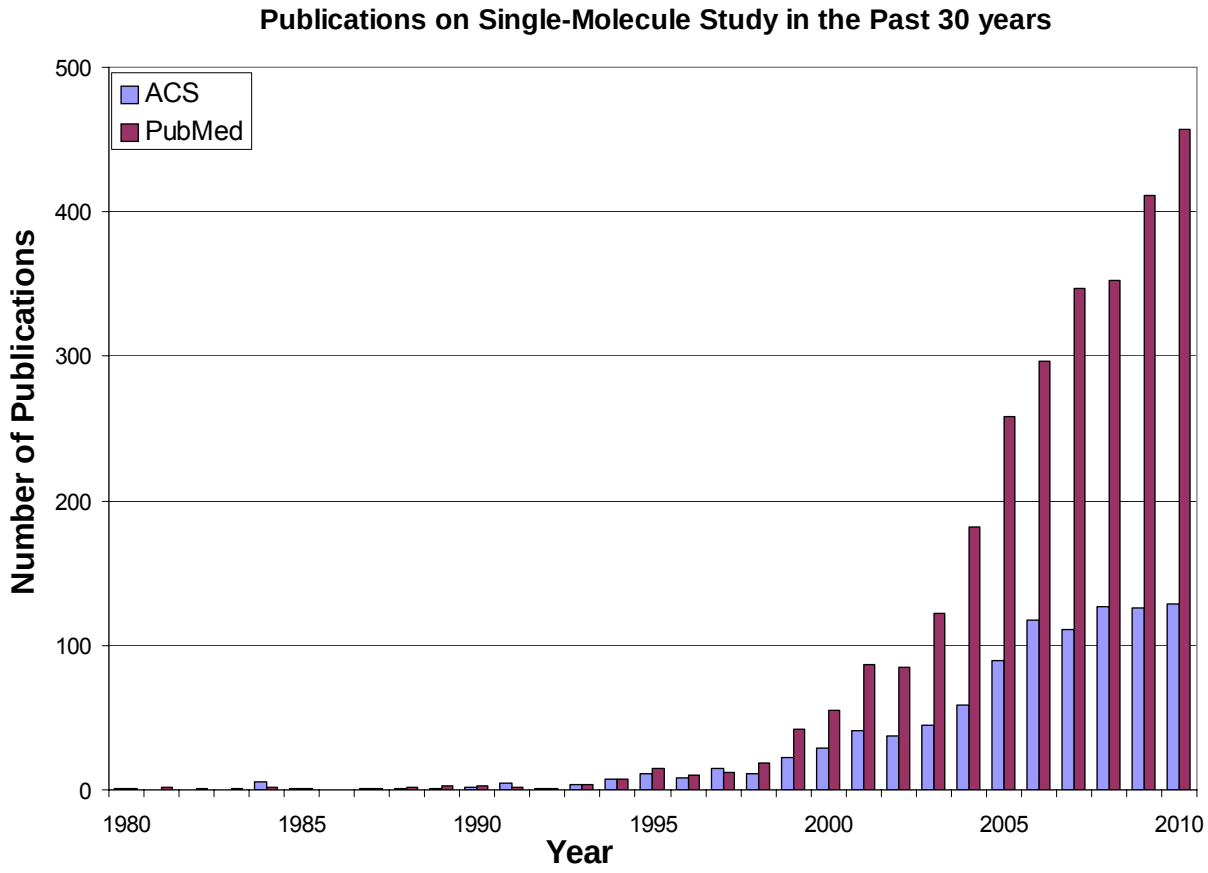


Figure 1.10 Publications of scientific papers using single-molecule detection methods

The first single-molecule experiment using an optical method was conducted by Hirschfeld, who successfully observed the fluorescence signature of a single antibody γ globulin, which was tagged with 80 to 100 fluoresceine isothiocyanate molecules.[92] His work demonstrated in principle the possibility of direct optical detection of a single fluorescent entity in liquid. The research on single molecule detection and spectroscopy (SMD and SMS) in solution at ambient conditions were then initiated and actively progressed in Keller's group at Los Alamos National Lab as well as in other research group in the 1980s and early 1990s. Early studies on SMD were exclusively focused on demonstrating the feasibility of detecting fluorescence signatures of single dye

molecules using fluorophores like Rhodamine 6G, and ultimate sensitivity of single molecule photon bursts was eventually achieved through incremental developments of many researchers.[93-101]

In the work during this time, a number of core issues in SMD including optical configurations, signal-to-noise ratio, photophysical properties of fluorophores, solvent effects, system optimization, autocorrelation criterion, etc. were thoroughly investigated, which laid the foundation towards further developments of SMD and establishment of its status as a widely accepted technique in other scientific fields.

1.6.2 Single-molecule Detection and FRET

The combination of single-molecule detection and FRET is a perfect marriage between these two exquisite techniques and offers benefits over themselves alone by two fold. In FRET the sixth power dependence of FRET efficiency, E_{FRET}, on donor-acceptor separation has been widely used to measure small distance change between different sites on large biological molecules like proteins. However, ensemble measurements cannot provide an accurate description of the dynamic conformational changes of these molecules in real-time because these conformational changes are difficult to be synchronized.[102] As a result, the detailed kinetics of structural changes of these molecules is obscured by the ensemble averaging. By contrast, single-pair FRET (spFRET), which detects the resonance energy transfer from a single donor to a single acceptor, potentially allows one to obtain the entire distribution of relevant properties as well as the time evolution of structural changes. Moreover, in classic SMD experiments one type of fluorophore is typically used in the sample. Most organic fluorescent molecules are characterized by their small red Stokes shift and the optimal emission

wavelength is usually only a few nanometers longer than the optimal excitation wavelength. Consequently, a portion of the excitation light can reach the detector by scattering and reflection, thus limiting detection sensitivity. Using two fluorophores that are optimally excited at two different wavelengths, on the other hand, increases the spectral gap between the optimal excitation and the optimal emission of the fluorophore-pair, thus allowing more efficient discrimination of excitation light and enhancing the sensitivity of target detection. This is particularly important when spFRET is deployed to detect probe-target hybridization in living cells because autofluorescence from cellular components will introduce extra background and interfere with faithful observations.

The first spFRET experiment was demonstrated by Ha who attempted to image a short DNA molecule that was labeled with a single donor and acceptor and was absorbed on a dry glass surface, with the FRET monitored using near-field scanning optical microscopy (NSOM).[103] Another early demonstration of spFRET was made using confocal microscopy equipped with two-color detection to differentiate and monitor subpopulations of molecules in solution.[104] These molecules were labeled with a donor and an acceptor at different sites, which gave different intramolecular distances or conformational states and could be separated by their distribution of FRET efficiencies.

In practice there are two primary ways of conducting single-molecule FRET measurements: in solution and on the surfaces. Currently most spFRET experiments are performed in solution because of the ease of implementation. But constant diffusion of analyte molecules makes these measurements inappropriate for slower processes (>10ms).[105] The alternative way is to tether the dye-labeled analyte molecules onto a solid surface, usually glass, to provide prolonged observation times. Then, detailed

reaction kinetics or unique signaling pathways on individual molecules can be comprehensively studied by tracking the time evolution of fluorescence. Surface spFRET measurements, however, are limited by photobleaching of fluorophores and complicated by artifacts caused by surface interactions on tagged analyte. Thus special care has to be taken when surface spFRET measurements are performed.

Traditional spFRET measurements were conducted between one donor and one acceptor. For multiple binding interactions in the same macromolecule, it can be beneficial to have more than one FRET pair to monitor correlated events. With improvement in fluorescence detection technology, spFRET experiments with three-color capabilities have been reported.[106-108] Hohng *et al.* developed a three-color FRET scheme with one donor (Cy3) and two acceptors (Cy5, Cy5.5) attached to three arms of a DNA four-way junction. Using this design, they were able to not only observe that the Cy5 and Cy5.5 fluorescence was emitted alternatively upon excitation, but also found that the Cy5 arm and the Cy5.5 arm actually moved towards the Cy3 arm at an asynchronous pace. These multi-color experiments were characterized by their color multiplexing capability, but in practice, they are difficult to implement because it is hard to find multiple fluorophores that own the following two properties: 1) Large spectral overlap for FRET; 2) sufficient emission spectral separation to allow detection in three or more colors. Kapanidis *et al.* designed a novel switchable FRET to overcome this technical barrier by using one donor and multiple identical acceptors, instead of multiple different acceptors.[109] In this method, FRET between each donor-acceptor pair proceeded alternately by photoswitching, in which the acceptor could be reversibly switched between a dark state (no spectral overlap between donor emission and

acceptor excitation) and a fluorescent state (overlap between donor emission and acceptor excitation). This type of time multiplexing circumvented many problems associated with multi-color approaches and could in principle be extended to distance measurements of more than three FRET pairs, which was not able to be achieved by multi-color measurements currently.

1.6.3 Single-molecule Absorption Detection at Room Temperature

Current single-molecule spectroscopy predominantly focuses on detecting the fluorescence emission, which is undoubtedly sensitive but limited to molecules that contain a fluorophore. Detecting a weak or nonfluorescent single molecule, however, overcomes this barrier and can be applied to a broader category of molecules that absorb photons though not able to fluoresce. The first single molecule optical absorption measurement in condensed matter was achieved in 1989 at low temperatures (1.6K).[110] Under cryogenic conditions, the absorption cross section of the sample molecule was extremely large, which boosted the absorption signal by several orders of magnitude.

Detecting single absorbing molecules at ambient conditions is more appealing but more challenging primarily because of abrupt decreases in the absorption cross section at higher temperatures. The absorption of single molecules at room temperature was first accomplished by Orrit and coworkers using photothermal microscopy.[111] In this experiment the molecule of interest was excited by a modulated laser and the surrounding glycerol solvent molecules were heated causing a change in the refractive index of the solvent molecules, which was detected through a probe laser. An alternative approach was applied by Xie *et al.*, whom used ground-state depletion microscopy equipped with two collinear laser beams at slightly different wavelengths

and both within the absorption band of the molecule.[112] The modulation of the pump beam caused small changes in the intensity of the transmitted probe beam, and variations in the laser intensity could discriminate the adsorption of single molecules.

1.6.4 Applications of Single-molecule Detection

Single-molecule approaches have gained enormous popularity in biological studies because it provides data on elementary biological processes which cannot be attained by conventional methods. Typically, reaction mechanisms are described on a single-molecule basis. Utilization of single-molecule techniques not only allows direct observation of these fundamental phenomena but also opens enormous opportunities for new discoveries.

1.6.4.1 DNA Fragment Sizing

Determination of DNA fragment size of minute amounts of nucleic acids is very important in many clinical diagnostics and forensic analyses. For example, DNA fragments after digestion by restriction enzymes at specific recognition sites can be used to fingerprint the genotype of patients. DNA fragments are usually separated through gel electrophoresis and visualized by staining, but tiny amounts of DNA fragments are not able to be observed directly due to sensitivity limitations of staining techniques. Fortunately, single-molecule fluorescence detection provides a good alternative approach when the DNA fragments of interest are rare. Sizing DNA fragments using single-molecule techniques was widely reported in the 1990s.[113-116] The principle of SMD approaches is that the dye molecules used to stain DNA fragments can be inserted into the dsDNA helix at a fixed ratio so that the amount of intercalating dyes bound to the fragment is proportional to the fragment length. When

each individual dye-labeled DNA molecule traverses a laser probe volume, the resulting photon burst can be recorded and the size of the photon burst is directly proportional to the length of the fragment being interrogated. Initially, the SMD fragment sizing was demonstrated in flow cytometer or capillary, and later on was migrated to microfluidic platforms with a variety of modifications and improvements like flow confinement to increase detection efficiency.[117] Fragments ranging from 2,000 to 200,000 base pairs can be successfully resolved using this device.

1.6.4.2 Protein Folding

Proteins are linear biopolymers that can spontaneously fold into a highly ordered 3-D structure from its initially random coiled state. The folding process is a self-assembly process and generally regarded to be pathway dependent. The protein molecule eventually finds its stable folded state after a stochastic search of conformations that intrinsically is a heterogeneous process.[118] Classical FRET measurements only yield ensemble averages, thus the detailed stochastic folding pathways of each protein are hidden. Single-molecule measurements allowed one to look at one protein molecule at a time and gain insight into the mechanism of protein folding. Single-pair FRET is the method of choice for protein folding investigations because of its unparalleled capability of measuring the intra- or inter-molecular distance in the nanometer range on the single-molecule level.

Protein folding observations can be carried out either in solution or on a surface. For example, the simplest single-molecule protein-folding experiment is to count the photon bursts of freely diffusing protein molecules at different denaturant concentrations.[119] By constructing a histogram of FRET efficiency distribution and counting the number of

peaks, the number of thermodynamic states is obtained. The limitation of this measurement is the short residence time of the protein in the probe volume, which makes slow folding processes unobservable. But the advantage is that this measurement is not subject to artifacts from surface interactions.

In order to directly observe slow protein folding events, like transitions between folded and unfolded states, the protein molecule has to be immobilized onto a solid surface. Jia *et al.* first reported monitoring the dynamics of single GCN-4 peptides attached on a glass surface using FRET confocal microscopy.[120] Practically, however, these experiments suffered from sporadic interactions between immobilized peptides and glass surfaces. Rhoades *et al.* used lipid vesicles to encapsulate the protein before being attached onto the surface, so that the interference arising from surface interactions was eliminated.[121] With this approach, they obtained time trajectories of target proteins and found that FRET from donor to acceptor of each encapsulated protein was not correlated. This provided evidence for the hypothesis that folding pathways of proteins are independent.

1.6.4.3 Single-molecule DNA Sequencing

During the current genomics revolution, the genomic sequences of a large number of living organisms have been fully deciphered. However, the emerging field of personal genomics, which aims to study genome variations in individuals, demands faster and more reliable sequencing technology. The National Institutes of Health has established a target of the "$1,000 genome" to address the increasing need for more efficient sequencing approaches, and a couple of technologies have been involved, including single-molecule sequencing.[122]

The concept of obtaining genomic sequence by single-molecule fluorescence detection was proposed and described ~20 years ago by Keller's group soon after successful detection of single dye molecules in solution.[123-125] In their method, a strand of DNA to be sequenced was first replicated in the presence of fluorescently labeled dNTPs, so its complementary strand was synthesized by successive addition of these fluorescent dNTPs according to its sequence. Then, the synthesized DNA strands made of dye-tagged dNTPs was suspended in a flow stream and a DNA exonuclease was attached to the end of this DNA strand to cleave the terminal nucleotide one at a time. Because the four types of nucleotides were tagged with four different fluorophores, sequential release of the nucleotides can be detected and differentiated downstream by single-molecule optics. A number of feasibility studies have been performed to address various issues in this technology, such as controlled movement of single-stranded DNA, exonuclease kinetics, complete detection of nucleotides in the order they are released, etc. Brakmann reported an elegant approach that combined several features for single-molecule sequencing in microstructures (7 x 10 μm).[126] A single dye-labelled DNA strand was first immobilized onto a microbead and trapped in the microchannel by optical tweezers. Then, released single nucleotides by exonuclease cleavage events were transported via electro-osmotic flow to a downstream detection window. The entire cross section of the microchannel at a given point was illuminated and an array of seven adjacent glass fiber tips were coupled to seven photon detectors to assure all nucleotides flowing through the channel could be observed. Werner *et al.* observed the progressive digestion of a single-labeled DNA fragments by a single exonuclease enzyme in flow cytometer, which was an essential component in this scheme.[127] But this

technique was still subject to high detection background probably due to fluorescent impurities, and the slow progressive rate of exonuclease.

Another DNA sequencing approach is sequencing-by-synthesis (SBS), in which the addition of each dye-labeled nucleotide is mediated by an enzyme and monitored by fluorescence signal at the single-molecule level. Quake *et al.* reported the repeated addition of single fluorescent nucleotides by a DNA polymerase into a single DNA strand on a microchip with single base resolution, and a sequence of up to 5 nucleotides was determined.[128] Harris *et al.* first reported successful sequencing of the genome of M13 virus using SBS with a read length of >25nt and 100% coverage.[129] The first commercial next-generation sequencer using single-molecule SBS was launched in 2008 by Helicos Bioscience.[130]

Both sequencing by cleavage and sequencing by synthesis entail action of an enzyme, which in some cases complicates the sequencing process. There is another appealing technique under development in recent years for DNA sequencing, which entails the use of a nanopore, where no enzyme is involved in the process. In this approach, a thin membrane was utilized which contained nanopores of 1.5-2 nm in diameter.[131] The target single-stranded DNA was placed on one side of the membrane and was stretched into the nanopore under the influence of an electric field. When the negatively charged DNA strands traversed through the nanopore, the channel was partially blocked and the electrical conductance of the membrane was changed, manifested by a small change in the measured current, usually in the range of pA. This current change depended on the DNA sequence rather than its length and can be used to decipher the oligonucleotide sequence. Currently this technology is still in the stage

of proof-of-concept, and one of the obstacles is that the elongated DNA strand travels through the nanopore in a speed that is too fast so that individual bases cannot be resolved.[130] In addition, the width of the nanopore encompasses several bases pairs and thus, single-base resolution cannot be achieved.

1.6.4.4 Other Applications of SMD

In addition to the applications mentioned above, SMD techniques have also been used for sensing DNA hybridization,[132] investigating biotin-avidin binding kinetics,[133] screening drug candidates,[134] etc, which will not be discussed in details here.

1.6.5 Perspective of SMD

The most promising application of SMD is to investigate biological processes by tracking individual biomolecules in three spatial dimensions in single living cells, which is also the ultimate goal of biological studies.[135] Chemical analysis and imaging of biomolecules in single cells *in vivo* have attracted great attention and seen strong advance by many research groups from their recent publications.[136] Measuring position and motion of individual fluorophores with high accuracy are the basis for recent breakthroughs in SMD in cellular environments. In SMD, the fluorophore attached to each single biomolecule is a local reporter of its physiological vicinity. The detailed single molecule biochemical process like protein folding or enzyme activity will be fully exposed by monitoring the movement of these single fluorescent reporters.

1.7 Microfluidic Device for Bioanalytical Applications

In the recent years, microfluidics has obtained widespread recognition across a wide variety of scientific disciplines as a promising tool to replace traditional ways for biological and chemical studies. Almost every issue currently involved in chemical or biological analysis has been addressed using a microfluidic platform, including nucleic

acid analysis with PCR,[137] protein analysis,[138] cell culturing and cell study,[139, 140] and tissue engineering.[141]

1.7.1 Advantages of Microfluidics

A microfluidic device is a miniaturized platform that is comprised of microchannels connected to liquid reservoirs, on which a variety of chemical or biological analysis can be carried out by manipulation of the fluid through these microchannels. The concept of microfluidics stems from the invention of early microanalytical methods such as gas phase chromatography, high-pressure liquid chromatography, and capillary electrophoresis, by which high sensitivity and high resolution are both achieved using only tiny amounts of sample when laser optical instruments are utilized for detection. Compared to conventional chemical or biological analysis on benchtop instruments, where fair amounts f sample are needed, analysis using microfluidic devices offers a number of advantages:

1) There is a minimal amount of sample and reagents required to perform an analysis, primarily due of the lateral dimensions of the microfluidic channels, which is usually in the range of μm or even sub-μm. For example, only a tiny drop of blood (a few μL) is needed to run a diagnostic test on a microfluidic chip. By contrast, traditional clinical labs usually requires a whole test tube of blood (a few mL) to be collected from the patient. This is critical when sample is rare and reagents are expensive, thus reducing the overall cost for the analysis.

2) Conventional bioanalytical methods are subject to inaccurate results caused by contaminants introduced when reagents are transferred between test tubes. A well designed microfluidic device integrates different functions including reaction, separation,

and detection on the same device, which allows the analysis to be conducted in a sample-in/answer-out fashion. Because the entire analysis is performed in an enclosed system, external contaminants will not be brought in, which otherwise could lead to false positive results. Use of disposable microfluidic chip eliminates instrument contaminations from previous runs.

3) Results can be obtained much faster than conventional methods, attributed to miniaturization of microfluidic devices or opportunities for selecting alternative processing strategies not amenable to benchtop processing. The small size of these devices allows only a small volume of fluid (a few μL) to be dispensed, which greatly facilitates mass transport and heat transfer during the whole process, making the analysis much faster.

4) Advancement of microfabrication techniques makes it possible to integrate different functionalities onto a single microfluidic unit. The architecture of microchannels could be designed to perform multiple experimental steps on one single device. This not only reduces the overall cost of analysis, but also facilitates system automation.

5) Multiple samples can be analyzed in parallel in the same microfluidic device. A typical example is capillary array electrophoresis microchannel plates fabricated on a glass wafer for DNA sequencing that allowed 96 samples to be processed simultaneously.[142] Increases in throughput greatly reduce the average time and cost needed for analyzing each sample.

1.7.2 Fabrication Techniques of Microfluidic Devices on Various Substrates

Early microfluidic devices were solely fabricated on silicon wafer or glass plates using the technology borrowed from the semiconductor industry, such as dry or wet

etching, photo-lithography, and electron beam lithography. In spite of mature fabrication techniques and their excellent chemical and physical properties, the high cost associated with the processing steps limits the use of these materials for disposable devices, which is required for clinical diagnostics.

Polymers, by contrast, are characterized by their low cost, ease to process, and suitability for mass production of disposable devices. The first polymer-based microfluidic device was prototyped by Whitesides and co-workers using PDMS, an elastomeric material, as the substrate.[143] In order to make a microfluidic chip using PDMS, a silicon master had to be first fabricated as a template from which the chips could be replicated. The silicon master was fabricated using soft lithography, where a thin layer of curable photoresist was spin coated onto a blank silicon wafer, followed by exposing the photoresist with UV light through a mask and dissolving the unexposed portion to form the desired pattern. Then a mixture of PDMS solution was casted onto the silicon master and cured to generate the microfluidic chip. The concept of making microfluidic devices by replication using a mold master is analogous to invention of printing technology, and has revolutionized microfluidics in the sense that theoretically it could tremendously reduce the cost and time required to make a device.

Since then microfluidic devices started to be made on a variety of polymeric materials such as polymethylmethacrylate (PMMA), polycarbonate (PC), cyclic olefin copolymer (COC), and polystyrene (PS). These polymers are thermoplastics and able to be reshaped when heated near their glass transition temperature (T_g). A mold master is typically fabricated on a metal using micromilling machine or LIGA process when fine structures were demanded. Then a piece of polymer sheet is placed against the mold

master under hydraulic pressure and heat so that the designed features stored on the mold master can be imprinted onto the polymer substrate. This process is called hot embossing and has been widely used in prototyping microfluidic devices of small runs. Alternatively, injection molding, which is commonly used in the plastic industry, can also be used to make microfluidic devices, especially in mass production due to its high throughput.[144]

1.7.3 Fluid Transportation in Microfluidic Device

There are basically two main methods to drive fluids through microchannels: electrokinetic-driven and pressure-driven flows. Flow by electrokinetics is commonly used in capillary electrophoresis, and is based on movement of molecules that are either electrostatically charged and driven by an electric field, or non-charged and carried by electroosmetic flow. This modality has been successfully migrated to microfluidic platforms and is among the earliest applications for driving fluids in microfluidic devices. Using an external electric field, the fluid inside the microchannel can be manipulated by switching the voltage on and off, which circumvented the need for valves. During electrokinetics an electroosmetic flow (EOF) was concurrent, which was caused by the bulk movement of the positive counterions that were tightly attracted to the neighborhood of the inner walls of a capillary or microchannel, which was usually negatively charged. The EOF can be suppressed or even reversed when the inner walls of a capillary of microchannel are coated with cationic polymers, like polybrene[145]

Another way to drive fluid through microfluidic devices is applying a pressure across the microchannel, usually through an external pump or vacuum. Pressure-driven flow is easy to implement and can provide constant flow rates, especially for long

microchannels. It can be used for a wide range of fluid compositions (conductive or non-conductive) and for microfluidic devices made from a wide range of substrates. However, the parabolic velocity profile of the laminar flow driven under pressure differences causes peak broadening due to longitudinal dispersion of analytes, which is not amenable to high-resolution separations.

1.7.4 Single-molecule Detection on Microfluidic Device

On-chip detection of analytes after separation or chemical reactions in microfluidic devices is challenging because only small volumes of sample are available to be interrogated due to the intrinsic feature of microfluidic devices. Both optical and electrical methods such as laser-induced fluorescence (LIF), chemiluminescence, conductivity, and Raman spectroscopy have been developed or incorporated onto microfluidic platforms. Among these detection techniques, LIF remains the most prevalent one because of the exquisite sensitivity, spectral selectivity, and low limit-of-detection it offers in microfluidic applications.[146]

Detecting the fluorescence signature of an analyte at the single-molecule level on microfluidic devices provides a powerful tool for bioanalytical research with extraordinary sensitivity and accuracy, which is unmatched by conventional ensemble detection methods. Single-molecule detection allows observing extremely low number of molecules without compromising the assay sensitivity, so that the probe volume and sample concentration can both be reduced, which is exactly the challenge in detection on microfluidic devices. Typical examples include detecting low-abundance nucleic acids without the need of amplification,[147] and quantifying mRNA gene expression levels extracted from single cells.[148]

1.8 Integrated Microfluidic System for Micro-Total-Analysis

The past two decades have seen a rapid growth of microfluidic technology, which enabled miniaturization of large analytical devices and allowed most benchtop analysis to be performed on microdevices, the so-called micro-total analysis system (μTAS) or lab-on-a-chip. The primary advantage of μTAS is the possibility to build a complete analytical microsystem with various functional modules integrated into a single device, so that any end user can directly use it to perform their test by simply loading the sample and waiting for their answer without any special training. A complete analysis usually involves sample preparation, analyte extraction, chemical reaction, analyte separation, and end-point detection, all of which can now be realized on microfluidic devices after tremendous amount of research. In the μTAS, microfluidic device is the core part and it delivers all the fluid through the microchannel network and dispenses all the chemicals to their designated sites, akin to the blood vessel in human body. Mechanical units such as pumps, valves, mixers, and electrical parts such as solenoids, actuators, are all necessary in order for the μTAS to operate properly. The micro-versions of these elements were successfully engineered attributed to advancements in microfabrication technology with functions comparable to their macro counterparts. A variety of detection techniques were also miniaturized by development and embedment of microelectrodes, microlenses, and waveguides onto microfluidic chips. This lays the foundation for all individual functions to be integrated.

System integration of microdfluidic devices is proceeding along two approaches. The first one adopts a monolithic approach and has been used by most existing systems. In this approach different function units such as chemical reactors, sensors, and actuators are all arranged on a single chip, usually on a single plane. In one

demonstration (shown in Figure 1.11), Liu and co-workers presented a self-contained, fully integrated biochip, which was able to detect pathogenic bacteria and single-nucleotide polymorphisms.[149] This biochip contained pumps, valves, channels, microfluidic mixers, chambers, heaters, and DNA microarray for nucleic acid analysis. Sample preparation, PCR, DNA hybridization and electrochemical detection were all performed on this single chip.

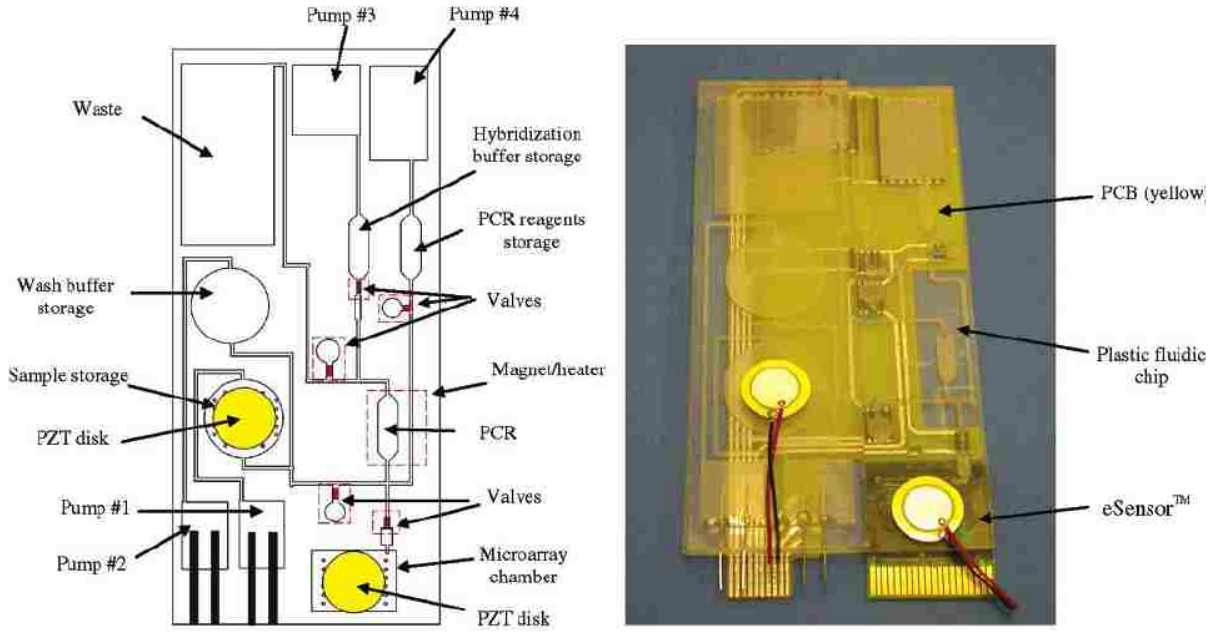


Figure 1.11 Illustration of self-contain, fully integrated biochip designed by Motorola Labs [149]

Landers *et al.* showed another integrated system using the similar concept also shown for microfluidic genetic analysis.[150] In this system, microchip electrophoresis separation was conducted after PCR and the sample was detected by LIF to complete the analysis in <30 min. This approach is straightforward, but it requires a microfabrication process common to all components, which sometimes may need to sacrifice some functionality to accommodate all components on the same device. In another approach for system integration, individual modules with clearly defined

functionalities are developed first, which are then assembled to construct the integrated system.

Quake first described a large-scale microfluidic channel system made of PDMS that contained two distinct layers for complex fluid manipulation.[151] In this system the flow layer contained the network of channels to be controlled, and the control layer was stacked on top of the fluid layer in a perpendicular way so that a microvalve was created at the junction of a flow channel and a control channel and can be deflected by hydraulic actuation to direct the flow traffic.

Liu *et al.* demonstrated an integrated microfluidic architecture (shown in Figure 1.12) using a modular construction scheme for biochemical analysis.[152] In this demonstration, the components were separated into categories and built on two different physical layers. The first layer contained all passive components such as microchannels and chemical reactors, and the second layer contained all active components such as valves, pump, mixers, and sensing elements. The second layer served as the microfluidic breadboard (FBB) and was mated to the first layer to perform various biochemical analyses. Compared to the monolithic approach, this method is more flexible and cost effective. The two layers can be made of different materials and customizable passive layer can be designed according to specific applications and shared the same FBB which has standardized configurations.

Current integration of μTAS is still in its incipient stage and most works continue to be dedicated to development and improvement of those stand-alone functionalities. With all these efforts, the promise of μTAS integration of all laboratory equipments onto

an integral microdevice, commercialization of a truly portable, disposable, analytical instrument can be expected in the near future.

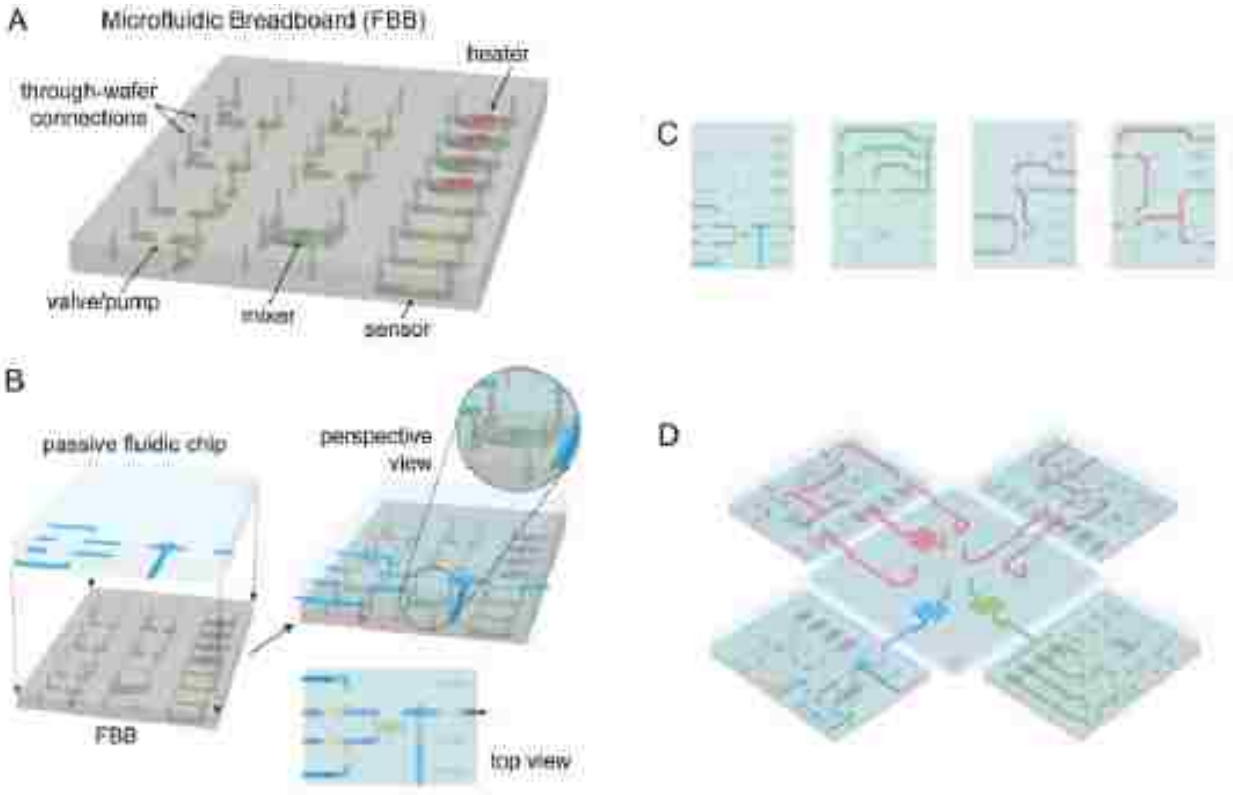


Figure 1.12 Schematic representation of the modular microfluidic architecture [152]

1.9 References

(1) Berg, J. M.; Tymoczko, J. L.; Stryer, L. *Biochemistry*, 6th ed.; W.H. Freeman: New York, 2007.

(2) Silverman, A. P.; Kool, E. T. *Chemical Reviews* 2006, *106*, 3775-3789.

(3) Collins, F. S.; Green, E. D.; Guttmacher, A. E.; Guyer, M. S. *Nature* 2003, *422*, 835-847.

(4) Collins, F. S.; Morgan, M.; Patrinos, A. *Science* 2003, *300*, 286-290.

(5) *Nature* 2003, *426*, 789-796.

(6) Li, W. H.; Sadler, L. A. *Genetics* 1991, *129*, 513-523.

(7) Wang, D., et al. *Science* 1998, *280*, 1077-1082.

(8) Frazer, K. A., et al. *Nature* 2007, *449*, 851-861.

(9) *Nature* 2005, *437*, 1299-1320.

(10) Chung, C. C.; Magalhaes, W. C.; Gonzalez-Bosquet, J.; Chanock, S. J. *Carcinogenesis*, *31*, 111-120.

(11) Casjens, S. *Annu Rev Genet* 1998, *32*, 339-377.

(12) Binnewies, T. T.; Motro, Y.; Hallin, P. F.; Lund, O.; Dunn, D.; La, T.; Hampson, D. J.; Bellgard, M.; Wassenaar, T. M.; Ussery, D. W. *Funct Integr Genomics* 2006, *6*, 165-185.

(13) Koonin, E. V.; Mushegian, A. R.; Rudd, K. E. *Curr Biol* 1996, *6*, 404-416.

(14) Peterson, J., et al. *Genome Res* 2009, *19*, 2317-2323.

(15) Ley, R. E.; Peterson, D. A.; Gordon, J. I. *Cell* 2006, *124*, 837-848.

(16) Nelson, K. E., et al. *Science*, *328*, 994-999.

(17) Qin, J., et al. *Nature*, *464*, 59-65.

(18) Zhang, H., et al. *Proc Natl Acad Sci U S A* 2009, *106*, 2365-2370.

(19) Cani, P. D., et al. *Diabetes* 2007, *56*, 1761-1772.

(20) Ley, R. E.; Turnbaugh, P. J.; Klein, S.; Gordon, J. I. *Nature* 2006, *444*, 1022-1023.

(21) Zhao, L. *Nature*, *465*, 879-880.

(22) *MMWR Morb Mortal Wkly Rep* 2006, *55*, 1045-1046.

(23) *MMWR Morb Mortal Wkly Rep* 2009, *58*, 85-90.

(24) Versalovic, J.; Lupski, J. R. *Trends Microbiol* 2002, *10*, S15-21.

(25) Gerner-Smidt, P.; Scheutz, F. *Foodborne Pathog Dis* 2006, *3*, 74-80.

(26) Holt, K. E.; Parkhill, J.; Mazzoni, C. J.; Roumagnac, P.; Weill, F. X.; Goodhead, I.; Rance, R.; Baker, S.; Maskell, D. J.; Wain, J.; Dolecek, C.; Achtman, M.; Dougan, G. *Nat Genet* 2008, *40*, 987-993.

(27) Miller, M. B.; Tang, Y. W. *Clin Microbiol Rev* 2009, *22*, 611-633.

(28) John M. Walker, R. R. *Medical BioMethods Handbook*, 2005.

(29) *Nature* 2004, *431*, 931-945.

(30) Alwine, J. C.; Kemp, D. J.; Stark, G. R. *Proc Natl Acad Sci U S A* 1977, *74*, 5350-5354.

(31) Chan, S. D.; Dill, K.; Blomdahl, J.; Wada, H. G. *Anal Biochem* 1996, *242*, 214-220.

(32) Gall, J. G.; Pardue, M. L. *Proc Natl Acad Sci U S A* 1969, *63*, 378-383.

(33) Bustin, S. A. *Journal of Molecular Endocrinology* 2000, *25*, 169-193.

(34) Gibson, U. E.; Heid, C. A.; Williams, P. M. *Genome Res* 1996, *6*, 995-1001.

(35) Mullis, K. B.; Faloona, F. A. *Methods Enzymol* 1987, *155*, 335-350.

(36) Erlich, H. A.; Gelfand, D.; Sninsky, J. J. *Science* 1991, *252*, 1643-1651.

(37) Saiki, R. K.; Gelfand, D. H.; Stoffel, S.; Scharf, S. J.; Higuchi, R.; Horn, G. T.; Mullis, K. B.; Erlich, H. A. *Science* 1988, *239*, 487-491.

(38) Gibbs, R. A. *Anal Chem* 1990, *62*, 1202-1214.

(39) Ronai, Z.; Yakubovskaya, M. *J Clin Lab Anal* 1995, *9*, 269-283.

(40) Mackay, I. M. *Clin Microbiol Infect* 2004, *10*, 190-212.

(41) Yang, S.; Rothman, R. E. *Lancet Infect Dis* 2004, *4*, 337-348.

(42) VanGuilder, H. D.; Vrana, K. E.; Freeman, W. M. *Biotechniques* 2008, *44*, 619-626.

(43) Larder, B. A.; Kemp, S. D. *Science* 1989, *246*, 1155-1158.

(44) Barany, F. *PCR Methods Appl* 1991, *1*, 5-16.

(45) Landegren, U.; Kaiser, R.; Sanders, J.; Hood, L. *Science* 1988, *241*, 1077-1080.

(46) Barany, F. *Proc Natl Acad Sci U S A* 1991, *88*, 189-193.

(47) Wiedmann, M.; Wilson, W. J.; Czajka, J.; Luo, J.; Barany, F.; Batt, C. A. *PCR Methods Appl* 1994, *3*, S51-64.

(48) Nickerson, D. A.; Kaiser, R.; Lappin, S.; Stewart, J.; Hood, L.; Landegren, U. *Proc Natl Acad Sci U S A* 1990, *87*, 8923-8927.

(49) Deng, J. Y.; Zhang, X. E.; Mang, Y.; Zhang, Z. P.; Zhou, Y. F.; Liu, Q.; Lu, H. B.; Fu, Z. J. *Biosens Bioelectron* 2004, *19*, 1277-1283.

(50) Mendy, M. E.; Kaye, S.; Le Roux, E.; Kirk, G. D.; Jeng-Barry, A.; McConkey, S.; Cotten, M.; Kuniholm, M. H.; Leligdowicz, A.; Hainaut, P.; Rowland-Jones, S.; Whittle, H. *J Clin Microbiol* 2008, *46*, 2723-2730.

(51) Sinville, R.; Coyne, J.; Meagher, R. J.; Cheng, Y. W.; Barany, F.; Barron, A.; Soper, S. A. *Electrophoresis* 2008, *29*, 4751-4760.

(52) Hashimoto, M.; Barany, F.; Soper, S. A. *Biosens Bioelectron* 2006, *21*, 1915-1923.

(53) Niederhauser, C.; Kaempf, L.; Heinzer, I. *Eur J Clin Microbiol Infect Dis* 2000, *19*, 477-480.

(54) Rondini, S., et al. *J Clin Microbiol* 2008, *46*, 2269-2279.

(55) Mikhailovich, V., et al. *J Clin Microbiol* 2001, *39*, 2531-2540.

(56) Eggerding, F. A. *PCR Methods Appl* 1995, *4*, 337-345.

(57) Tobe, V. O.; Taylor, S. L.; Nickerson, D. A. *Nucleic Acids Res* 1996, *24*, 3728-3732.

(58) Stryer, L.; Haugland, R. P. *Proc Natl Acad Sci U S A* 1967, *58*, 719-726.

(59) Selvin, P. R. *Nat Struct Biol* 2000, *7*, 730-734.

(60) Lakowicz, J. R. *Principles of fluorescence spectroscopy*, 3rd ed.; Springer: New York, 2006.

(61) Latt, S. A.; Cheung, H. T.; Blout, E. R. *J Am Chem Soc* 1965, *87*, 995-1003.

(62) Morrison, L. E.; Halder, T. C.; Stols, L. M. *Anal Biochem* 1989, *183*, 231-244.

(63) Ghosh, S. S.; Eis, P. S.; Blumeyer, K.; Fearon, K.; Millar, D. P. *Nucleic Acids Res* 1994, *22*, 3155-3159.

(64) Cremazy, F. G.; Manders, E. M.; Bastiaens, P. I.; Kramer, G.; Hager, G. L.; van Munster, E. B.; Verschure, P. J.; Gadella, T. J., Jr.; van Driel, R. *Exp Cell Res* 2005, *309*, 390-396.

(65) Heyduk, T. *Curr Opin Biotechnol* 2002, *13*, 292-296.

(66) Shih, W. M.; Gryczynski, Z.; Lakowicz, J. R.; Spudich, J. A. *Cell* 2000, *102*, 683-694.

(67) Tyagi, S.; Kramer, F. R. *Nat Biotechnol* 1996, *14*, 303-308.

(68) Fang, X.; Li, J. J.; Perlette, J.; Tan, W.; Wang, K. *Anal Chem* 2000, *72*, 747A-753A.

(69) Drake, T. J.; Tan, W. *Appl Spectrosc* 2004, *58*, 269A-280A.

(70) Tyagi, S.; Marras, S. A.; Kramer, F. R. *Nat Biotechnol* 2000, *18*, 1191-1196.

(71) Mitchell, P. *Nat Biotechnol* 2001, *19*, 1013-1017.

(72) Dirks, R. W.; Molenaar, C.; Tanke, H. J. *Histochem Cell Biol* 2001, *115*, 3-11.

(73) Tsourkas, A.; Behlke, M. A.; Xu, Y.; Bao, G. *Anal Chem* 2003, *75*, 3697-3703.

(74) Bonnet, G.; Tyagi, S.; Libchaber, A.; Kramer, F. R. *Proc Natl Acad Sci U S A* 1999, *96*, 6171-6176.

(75) Tyagi, S.; Bratu, D. P.; Kramer, F. R. *Nat Biotechnol* 1998, *16*, 49-53.

(76) Ortiz, E.; Estrada, G.; Lizardi, P. M. *Mol Cell Probes* 1998, *12*, 219-226.

(77) Vet, J. A.; Majithia, A. R.; Marras, S. A.; Tyagi, S.; Dube, S.; Poiesz, B. J.; Kramer, F. R. *Proc Natl Acad Sci U S A* 1999, *96*, 6394-6399.

(78) Varma-Basil, M.; El-Hajj, H.; Marras, S. A.; Hazbon, M. H.; Mann, J. M.; Connell, N. D.; Kramer, F. R.; Alland, D. *Clin Chem* 2004, *50*, 1060-1062.

(79) Sokol, D. L.; Zhang, X.; Lu, P.; Gewirtz, A. M. *Proc Natl Acad Sci U S A* 1998, *95*, 11538-11543.

(80) Perlette, J.; Tan, W. *Anal Chem* 2001, *73*, 5544-5550.

(81) Bratu, D. P.; Cha, B. J.; Mhlanga, M. M.; Kramer, F. R.; Tyagi, S. *Proc Natl Acad Sci U S A* 2003, *100*, 13308-13313.

(82) Cui, Z. Q.; Zhang, Z. P.; Zhang, X. E.; Wen, J. K.; Zhou, Y. F.; Xie, W. H. *Nucleic Acids Res* 2005, *33*, 3245-3252.

(83) Heyduk, T.; Heyduk, E. *Nat Biotechnol* 2002, *20*, 171-176.

(84) Dummitt, B.; Chang, Y. H. *Assay Drug Dev Technol* 2006, *4*, 343-349.

(85) Marras, S. A.; Kramer, F. R.; Tyagi, S. *Genet Anal* 1999, *14*, 151-156.

(86) Tang, Z.; Wang, K.; Tan, W.; Li, J.; Liu, L.; Guo, Q.; Meng, X.; Ma, C.; Huang, S. *Nucleic Acids Res* 2003, *31*, e148.

(87) Tang, Z.; Wang, K.; Tan, W.; Ma, C.; Li, J.; Liu, L.; Guo, Q.; Meng, X. *Nucleic Acids Res* 2005, *33*, e97.

(88) Epstein, J. R.; Leung, A. P.; Lee, K. H.; Walt, D. R. *Biosens Bioelectron* 2003, *18*, 541-546.

(89) Steemers, F. J.; Ferguson, J. A.; Walt, D. R. *Nat Biotechnol* 2000, *18*, 91-94.

(90) Zhang, P.; Beck, T.; Tan, W. *Angew Chem Int Ed Engl* 2001, *40*, 402-405.

(91) Moerner, W. E. *Journal of Physical Chemistry B* 2002, *106*, 910-927.

(92) Hirschfeld, T. *Appl Opt* 1976, *15*, 2965-2966.

(93) Dovichi, N. J.; Martin, J. C.; Jett, J. H.; Keller, R. A. *Science* 1983, *219*, 845-847.

(94) Dovichi, N. J.; Martin, J. C.; Jett, J. H.; Trkula, M.; Keller, R. A. *Analytical Chemistry* 1984, *56*, 348-354.

(95) Nguyen, D. C.; Keller, R. A.; Trkula, M. *Journal of the Optical Society of America B-Optical Physics* 1987, *4*, 138-143.

(96) Shera, E. B.; Seitzinger, N. K.; Davis, L. M.; Keller, R. A.; Soper, S. A. *Chemical Physics Letters* 1990, *174*, 553-557.

(97) Soper, S. A.; Shera, E. B.; Martin, J. C.; Jett, J. H.; Hahn, J. H.; Nutter, H. L.; Keller, R. A. *Analytical Chemistry* 1991, *63*, 432-437.

(98) Soper, S. A.; Davis, L. M.; Shera, E. B. *Journal of the Optical Society of America B-Optical Physics* 1992, *9*, 1761-1769.

(99) Mathies, R. A.; Peck, K.; Stryer, L. *Analytical Chemistry* 1990, *62*, 1786-1791.

(100) Peck, K.; Stryer, L.; Glazer, A. N.; Mathies, R. A. *Proceedings of the National Academy of Sciences of the United States of America* 1989, *86*, 4087-4091.

(101) Soper, S. A.; Mattingly, Q. L.; Vegunta, P. *Analytical Chemistry* 1993, *65*, 740-747.

(102) Ha, T. *Methods* 2001, *25*, 78-86.

(103) Ha, T.; Enderle, T.; Ogletree, D. F.; Chemla, D. S.; Selvin, P. R.; Weiss, S. *Proceedings of the National Academy of Sciences of the United States of America* 1996, *93*, 6264-6268.

(104) Deniz, A. A.; Dahan, M.; Grunwell, J. R.; Ha, T. J.; Faulhaber, A. E.; Chemla, D. S.; Weiss, S.; Schultz, P. G. *Proceedings of the National Academy of Sciences of the United States of America* 1999, *96*, 3670-3675.

(105) Rasnik, I.; Mckinney, S. A.; Ha, T. *Accounts of Chemical Research* 2005, *38*, 542-548.

(106) Hohng, S.; Joo, C.; Ha, T. *Biophysical Journal* 2004, *87*, 1328-1337.

(107) Clamme, J. P.; Deniz, A. *Biophysical Journal* 2004, *86*, 326a-326a.

(108) Hohng, S.; Joo, C.; Ha, T. *Biophysical Journal* 2004, *86*, 326a-326a.

(109) Uphoff, S.; Holden, S. J.; Le Reste, L.; Periz, J.; van de Linde, S.; Heilemann, M.; Kapanidis, A. N. *Nature Methods* 2010, *7*, 831-U890.

(110) Moerner, W. E.; Kador, L. *Phys Rev Lett* 1989, *62*, 2535-2538.

(111) Gaiduk, A.; Yorulmaz, M.; Ruijgrok, P. V.; Orrit, M. *Science*, *330*, 353-356.

(112) Shasha Chong, W. M., and X. Sunney Xie *The Journal of Physical Chemistry Letters* 2010, *1*, 7.

(113) Castro, A.; Fairfield, F. R.; Shera, E. B. *Analytical Chemistry* 1993, *65*, 849-852.

(114) Ambrose, W. P.; Goodwin, P. M.; Jett, J. H.; Johnson, M. E.; Martin, J. C.; Marrone, B. L.; Schecker, J. A.; Wilkerson, C. W.; Keller, R. A.; Haces, A.; Shih, P. J.; Harding, J. D. *Berichte Der Bunsen-Gesellschaft-Physical Chemistry Chemical Physics* 1993, *97*, 1535-1542.

(115) Goodwin, P. M.; Johnson, M. E.; Martin, J. C.; Ambrose, W. P.; Marrone, B. L.; Jett, J. H.; Keller, R. A. *Nucleic Acids Research* 1993, *21*, 803-806.

(116) Petty, J. T.; Johnson, M. E.; Goodwin, P. M.; Martin, J. C.; Jett, J. H.; Keller, R. A. *Analytical Chemistry* 1995, *67*, 1755-1761.

(117) Chou, H. P.; Spence, C.; Scherer, A.; Quake, S. *Proceedings of the National Academy of Sciences of the United States of America* 1999, *96*, 11-13.

(118) Dobson, C. M. *Nature* 2003, *426*, 884-890.

(119) Gopich, I. V.; Szabo, A. *Journal of Physical Chemistry B* 2007, *111*, 12925-12932.

(120) Jia, Y.; Talaga, D. S.; Wei, L.; Lu, H. S. M.; DeGrado, W. F.; Hochstrasser, R. M. *Biophysical Journal* 1999, *76*, A10-A10.

(121) Rhoades, E.; Gussakovsky, E.; Haran, G. *Proceedings of the National Academy of Sciences of the United States of America* 2003, *100*, 3197-3202.

(122) Service, R. F. *Science* 2006, *311*, 1544-1546.

(123) Jett, J. H.; Keller, R. A.; Martin, J. C.; Marrone, B. L.; Moyzis, R. K.; Ratliff, R. L.; Seitzinger, N. K.; Shera, E. B.; Stewart, C. C. *J Biomol Struct Dyn* 1989, *7*, 301-309.

(124) Davis, L. M.; Fairfield, F. R.; Harger, C. A.; Jett, J. H.; Keller, R. A.; Hahn, J. H.; Krakowski, L. A.; Marrone, B. L.; Martin, J. C.; Nutter, H. L.; Ratliff, R. L.; Shera, E. B.; Simpson, D. J.; Soper, S. A. *Genetic Analysis-Biomolecular Engineering* 1991, *8*, 1-7.

(125) Harding, J. D.; Keller, R. A. *Trends in Biotechnology* 1992, *10*, 55-57.

(126) Klaus Dorrey, S. B., Michael Brinkmeiery,; Kyung-Tae Hany, K. R., Petra Schwilley,; Jens Stephany, T. W., Markus Lapczynaz,; Michael Stukez, R. B., Michael Hinzx,; Hartmut Seligerx, J. H., Manfred Eigeny and; Rigler, R. *Bioimaging* 1997, *5*, 14.

(127) Werner, J. H.; Cai, H.; Jett, J. H.; Reha-Krantz, L.; Keller, R. A.; Goodwin, P. M. *Journal of Biotechnology* 2003, *102*, 1-14.

(128) Braslavsky, I.; Hebert, B.; Kartalov, E.; Quake, S. R. *Proceedings of the National Academy of Sciences of the United States of America* 2003, *100*, 3960-3964.

(129) 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.; Xie, Z. *Science* 2008, *320*, 106-109.

(130) Gupta, P. K. *Trends in Biotechnology* 2008, *26*, 602-611.

(131) Kasianowicz, J. J.; Brandin, E.; Branton, D.; Deamer, D. W. *Proceedings of the National Academy of Sciences of the United States of America* 1996, *93*, 13770-13773.

(132) Singh-Zocchi, M.; Dixit, S.; Ivanov, V.; Zocchi, G. *Proceedings of the National Academy of Sciences of the United States of America* 2003, *100*, 7605-7610.

(133) Wayment, J. R.; Harris, J. M. *Analytical Chemistry* 2009, *81*, 336-342.

(134) Skinner, G. M.; Visscher, K. *Assay and Drug Development Technologies* 2004, *2*, 397-405.

(135) Xie, X. S.; Yu, J.; Yang, W. Y. *Science* 2006, *312*, 228-230.

(136) Lord, S. J.; Lee, H. L. D.; Moerner, W. E. *Analytical Chemistry* 2010, *82*, 2192-2203.

(137) Auroux, P. A.; Koc, Y.; deMello, A.; Manz, A.; Day, P. J. *Lab Chip* 2004, *4*, 534-546.

(138) Lion, N., et al. *Electrophoresis* 2003, *24*, 3533-3562.

(139) El-Ali, J.; Sorger, P. K.; Jensen, K. F. *Nature* 2006, *442*, 403-411.

(140) Kim, S. M.; Lee, S. H.; Suh, K. Y. *Lab Chip* 2008, *8*, 1015-1023.

(141) Andersson, H.; van den Berg, A. *Lab Chip* 2004, *4*, 98-103.

(142) Paegel, B. M.; Emrich, C. A.; Wedemayer, G. J.; Scherer, J. R.; Mathies, R. A. *Proc Natl Acad Sci U S A* 2002, *99*, 574-579.

(143) David C. Duffy, J. C. M., Olivier J. A. Schueller, and George M. Whitesides *Analytical Chemistry* 1998, *72*, 11.

(144) Fiorini, G. S.; Chiu, D. T. *Biotechniques* 2005, *38*, 429-446.

(145) Liu, X. Z.; Erickson, D.; Li, D. Q.; Krull, U. J. *Analytica Chimica Acta* 2004, *507*, 55-62.

(146) de Mello, A. J. *Lab Chip* 2003, *3*, 29N-34N.

(147) Wang, T. H.; Peng, Y.; Zhang, C.; Wong, P. K.; Ho, C. M. *J Am Chem Soc* 2005, *127*, 5354-5359.

(148) Zhong, J. F.; Chen, Y.; Marcus, J. S.; Scherer, A.; Quake, S. R.; Taylor, C. R.; Weiner, L. P. *Lab Chip* 2008, *8*, 68-74.

(149) Liu, R. H.; Yang, J.; Lenigk, R.; Bonanno, J.; Grodzinski, P. *Anal Chem* 2004, *76*, 1824-1831.

(150) Easley, C. J.; Karlinsey, J. M.; Bienvenue, J. M.; Legendre, L. A.; Roper, M. G.; Feldman, S. H.; Hughes, M. A.; Hewlett, E. L.; Merkel, T. J.; Ferrance, J. P.; Landers, J. P. *Proc Natl Acad Sci U S A* 2006, *103*, 19272-19277.

(151) Thorsen, T.; Maerkl, S. J.; Quake, S. R. *Science* 2002, *298*, 580-584.

(152) Shaikh, K. A.; Ryu, K. S.; Goluch, E. D.; Nam, J. M.; Liu, J.; Thaxton, C. S.; Chiesl, T. N.; Barron, A. E.; Lu, Y.; Mirkin, C. A.; Liu, C. *Proc Natl Acad Sci U S A* 2005, *102*, 9745-9750.

CHAPTER 2 SINGLE MOLECULE DETECTION TECHNIQUES

2.1 Introduction

In chemical analysis, detecting single molecules represents the ultimate level of sensitivity and has been a longstanding goal of analytical chemistry. In the past three decades, significant progress has been made in the field of single-molecule optical detection since the first successful observation of a fluorescein-labeled single antibody using a microscope by Hirschfeld.[1] Parallel to the invention of scanning tunneling microscopy (STM) and Atomic Force Microscopy (AFM) in the early and mid-1980s,[2, 3] which were used to probe surface morphology with single atom or molecule resolution, studies on single-molecule fluorescence spectroscopy in liquids were initiated and actively pursued by various researchers using fluorophores in flowing conditions and eventually, a fluorescent signature from a single molecule was observed.[2-5] Moerner and co-workers also demonstrated the capability of detecting single molecules in solids by detection of single dopant molecules in a host crystal at cryogenic condition.[6]

The popularity of single-molecule techniques arose from its unmatched sensitivity of exploring the characteristics of each individual molecules other than merely measuring the average property of a collection of molecules. Conventional chemical quantitative analysis measures the ensemble average from the bulk in an analog manner. A calibration curve is usually established by first measuring a series of standard samples, and the concentration or amount of analyte was estimated from the amplitude of the measured signal through inspection of the calibration curve. Single-molecule detection, in principle, can quantify an analyte in a digital manner by direct counting the number of photon peaks that is proportional to the number of analyte molecules without the need

of calibration curves, which is favorable in many situations where either standard samples are difficult to obtain or the analyte is too rare to be detected by ensemble measurements. The rapid development of single-molecule detection techniques in recent years has made it a pervasive tool in many chemical analysis and biological studies, where exceptional sensitivity of detection is pursued to gain insight on information such as enzymatic reaction mechanisms and protein structure dynamics at the molecular level.

Among the different optical methods for single-molecule detection, laser-induced fluorescence still remains the most widely used because of its superior sensitivity, spectral selectivity, low background, and high signal-to-noise ratio which is ideal for probing minute amounts of analyte in small volumes non-invasively.[7] In practice, a single dye molecule can be attached to a host molecule under study to serve as a reporter so that the behavior of the host molecule becomes observable using optical detectors. Nowadays, fluorophore labeling has become a prevalent technology in every aspect of biological research including studies in nucleic acids, proteins and cells.

2.2 Principles of Laser-induced Fluorescence

Successful detection of a single fluorescent molecule through optical methods relies on the ability of the fluorophore to be cycled between its ground electronic state and an excited electronic state repeatedly to generate multiple photons before it gets fatigued, or photobleached. The fundamental principle of laser-induced fluorescence is depicted by a Jablonski diagram as shown in Figure 2.1.

At a certain temperature, fluorescent molecules are usually populated at different electronic states denoted by S_0, S_1 and S_2 and the distribution of the population is

determined by the Boltzmann equation. At each electronic energy level, the fluorophore can also exist at a number of vibrational energy levels denoted by 0, 1, 2, etc. Typically S_0 is the most populated electronic state, called the ground electronic state. When an incoming photon is in resonance with the energy gap of the fluorophore, it can absorb this photon and be promoted to a higher energy level such as S_1 or S_2, a process called excitation, which occurs in about 10^{-15} s.

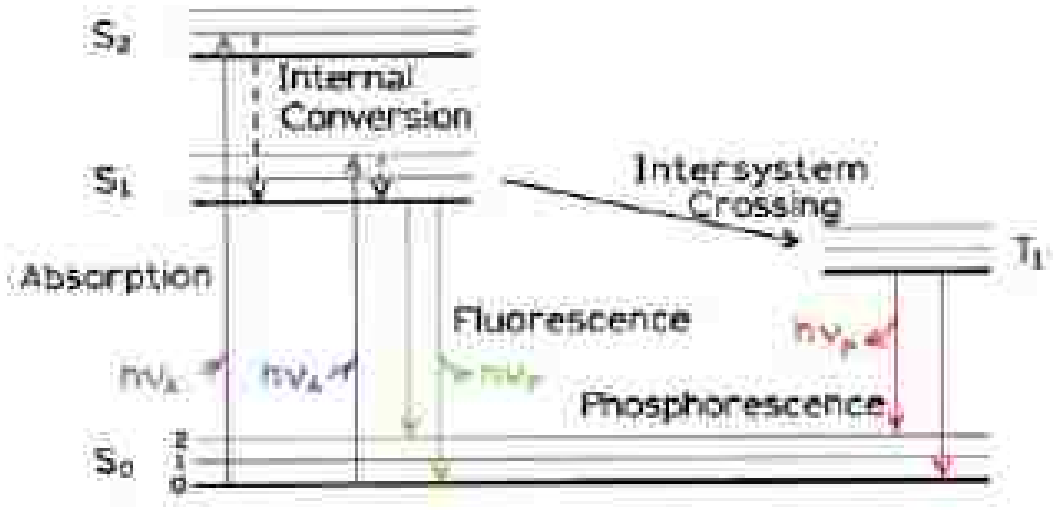


Figure 2.1 Schematic illustration of fluorescence using Jablonski Diagram [8]

A fluorophore can be excited to a higher vibrational level of S_1 or S_2, and quickly relaxes to the lowest vibrational level of S_1, which occurs in about 10^{-12} s. This process is called internal conversion and the energy is dissipated non-radiatively. Then, the fluorophore returns to its ground electronic state accompanied by emission of a photon, a process called fluorescence. Most fluorescent molecules have a lifetime of 10^{-9}-10^{-8} s, much slower than the relaxation process. Therefore the fluorescence photons are emitted from a thermal equilibrium energy level, which is the lowest vibrational level of S_1.

Alternatively, some fluorophores can also decay to the triplet state through spin conversion, a process called intersystem crossing, before it eventually returns to the ground state. The energy at the first triplet state can be dissipated in the form of heat non-radiatively, or by emitting a photon, a process called phosphorescence. Because the conversion from T_1 to S_0 is forbidden, the rate of phosphorescence is much slower than that of fluorescence.

The emitted photon in fluorescence typically has lower energy than the absorbed photon because of energy loss during the relaxation process. This phenomenon is called Stokes shift, which allows for spectral discrimination of fluorescence in the presence of high levels of scattering background.

Fluorophores can be energized by a variety of light sources. Compared to a regular light source, which has a broad spectral distribution, the energy of a laser is tightly distributed within a very narrow band of its central wavelengths, which makes it extremely efficient in exciting the fluorophores and the emitted fluorescence can be readily separated from scattering by spectral filtering. Therefore most fluorescence applications use laser as the source of excitation.

2.3 Theory of Single-Molecule Detection (SMD)

The reason that a single fluorescent molecule can be optically detected relies on its unique capability of being repetitively pumped to a higher energy level before it is photobleached. When a fluorescent molecule is brought into a tightly focused laser beam with a wavelength close to its adsorption maximum, it is promoted from the ground state to its first excited singlet state and gives off a photon during its return to the ground state. This process is cycled until it leaves the laser beam and a burst

containing multiple-photons is observed corresponding to the single molecule's fluorescence signature. In order to guarantee that only one fluorescent molecule resides within the laser excitation beam at any given time, the sample solution is first subjected to a series of dilutions so that statistically there is only one molecule is in the excitation volume. In a confocal laser-induced fluorescence (LIF) setup, the probe volume usually overlaps with the excitation volume, and the average occupancy of molecules in the probe volume is described by the following equation:[11]

$$K = C_a \cdot V_p \cdot N_A \qquad (2.1)$$

where K is the molecular occupancy, C_a is the concentration of analyte, V_p is the probe volume, and N_A is Avogadro's constant. For example, when a sample solution containing 1 pM of analyte flows through a probe volume of 10 fL, the average occupancy of molecules within the probe volume is 0.006, which means on average there are 0.006 sample molecules in that volume element at any given time.

The probability of finding any given number of molecules in the probe volume can be described by Poisson statistics, expressed by the following equation:[9, 10]

$$P(m) = \frac{K^m}{m!} e^{-K} \qquad (2.2)$$

where $P(m)$ is the probability of finding *m* molecules in the probe volume, and K is the average occupancy of molecules. The Poisson distribution of probability of finding *m* molecules in the probe volume is illustrated in Figure 2.1, and the values of probability are tabulated in Table 2.1. We can see from this figure that when the average molecular occupancy K is greater than 1, the probability of detecting multiple molecules $P(m)$ dominates. When K decreases to 0.1, the probability of detecting one molecule $P(1)$ is

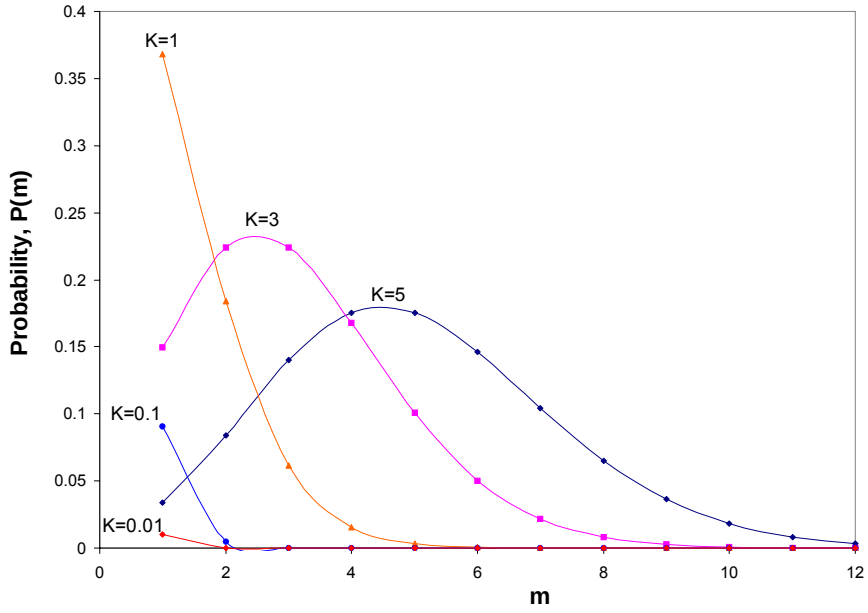


Figure 2.1 Poisson probability distribution of finding *m* molecules in the probe volume as a function of average molecular occupancy K

Table 2.1 Probability of detecting m molecules in probe volume

m	N=5	N=3	N=1	N=0.1	N=0.01
1	0.0337	0.1494	0.3679	0.0905	0.0099
2	0.0842	0.2240	0.1839	0.0045	4.95E-05
3	0.1404	0.2240	0.0613	0.0002	1.65E-07
4	0.1755	0.1680	0.0153	3.77E-06	4.13E-10
5	0.1755	0.1008	0.0031	7.54E-08	8.25E-13
6	0.1462	0.0504	0.0005	1.26E-09	1.38E-15

reduced to 0.09, while the probability of detecting two molecules $P(2)$ is reduced to 0.0045, giving a $P(1)$ to $P(2)$ ratio of 20, which suggests that among 20 detected photon bursts there is about one burst resulting from simultaneous emission of two fluorophores. When K further decreases to 0.01, $P(1)$ is reduced to 0.01 and $P(2)$ is reduced to 5 X 10^{-5}, giving a $P(1)$ to $P(2)$ ratio of 200, which suggests that >99% of the observed photon bursts arise from single fluorophores. Thus, a molecular occupancy of

0.01 is usually selected as the criteria when performing single-molecule detection experiments to make sure that observed signals indeed arise from single molecules.

2.4 Detectability of Single Molecules

The real challenge in single-molecule detection (SMD) is to detect the emission of a single fluorophore in the presence of billions of solvent molecules that surround it and scatter the excitation light, which can swamp the weak photon burst signal. This challenge can be vividly analogized to finding a needle in a haystack considering the dimensions of a single florophore (1 Å in diameter), probe volume (1 μm^3), needle (1 mm in diameter), and haystack (10mX10mX10m) [11]. The feasibility of detecting a single fluorescent molecule from billions of solvent molecules relies on two operational characteristics: 1) The fluorophore is able to withstand repetitive and uninterrupted excitation and emit hundreds to thousands of photons before it leaves the probe volume or photobleachs; 2) background photons generated from the solvent molecules can be significantly suppressed using a variety of techniques, such as time-gating, spectral filtering or reducing the probe volume.

The effect of reducing the probe volume on background fluorescence suppression can be illustrated by evaluating the cross-section of fluorophore and solvent molecules in different detection volumes.

The cross-section of a molecule reflects its capability of intercepting an incoming photon, and the larger the cross-section, the higher the probability that the molecule can be electronically excited by the incoming photon. The absorption cross-section σ of a fluorophore i estimated from its molar extinction coefficient ε using the following equation:[8]

$$\sigma = 2.303\varepsilon / N_A \qquad (2.3)$$

where N_A is the Avogadro's number. For example, the absorption cross-section of the commonly used fluorophore, Rhodamine 6G (R6G), is estimated to be ~4Å^2. By contrast, the Raman scattering cross-section of a single ethylene glycol (EG) molecule is about 3 X 10^{-12} Å^2. Consider a situation where one single R6G molecule is present in 1.3 X 10^{12} EG solvent molecules, which occupies a volume of 120 μm^3. The total Raman scattering cross-section is 4 Å^2, which is equivalent to the absorption cross-section of an R6G molecule. When the probe volume is reduced to 1.2 μm^3, which will contain 1.3 X 10^{10} EG solvent molecules, the total Raman scattering cross-section of EG molecules is proportionally reduced to 0.04 Å^2, which is only 1% of the absorption cross-section of an R6G molecule.

2.5 Signal and Background in SMD

In SMD experiments, successful detection of single-molecule events relies on both increasing the signal intensity and reducing background noise. Photon bursts from single-molecule events can be detected only if they significantly exceed the background fluorescence. Usually a threshold level is set above the average fluorescence background to minimize false positive results from an blank, and only those photon bursts above the threshold level can be counted as true single molecule events.

2.5.1 Photon Burst Signal

To obtain an intense photon burst signal, several conditions need to be sufficed: 1) Fluorophores with high quantum yield, large absorption cross-section, large Stokes shift, high photostability, and narrow triplet bottleneck; 2) optical components with high

collection efficiencies; and 3) detectors with high quantum efficiencies, low dark count rates, and single photon counting capabilities.

Dye molecules commonly used for biological fluorescent labeling that have aromatic structures, such as carbocyanine molecules, are featured by their favorable photostablility structures, large extinction coefficients, and high quantum yields.[12]

The amplitude of the photon burst is determined primarily by the number of photons that a fluorophore can give off, which is evaluated by the follow equation:

$$n = \frac{\Phi_f}{\Phi_d} \qquad (2.4)$$

where n is the number of photons that a fluorophore can emit, Φ_f is the quantum efficiency of the fluorophore, and Φ_d is the photodestruction quantum efficiency of the fluorophore.[13] At room temperature fluorescent molecules in aqueous solutions usually can emit up to ~10^6 photons before photobleaching.[14] In practice, only ~1% of these photons can eventually be registered on the photodetector even for the most sensitive optical system.[15] Thus, a total of 10,000 photons can be detected from a single fluorescent molecule.

The photon collection efficiency in an LIF system is determined by the objective that is used to gather the photons. The performance of an objective in terms of its collection efficiency can be evaluated by its numeric aperture (NA), which is defined as:

$$NA = n \cdot \sin(\frac{\varphi_{\max}}{2}) \qquad (2.5)$$

where n is the refractive index of the media between the objective and sample, and $\phi_{\max}$ is the full angle of the light that can be collected from the focal volume (see Figure 2.2).

Microscope objectives with higher magnification power generally have a higher NA, thus higher collection efficiency. In some applications, the objective is immersed in a liquid with high refractive index to increase its collection efficiency.

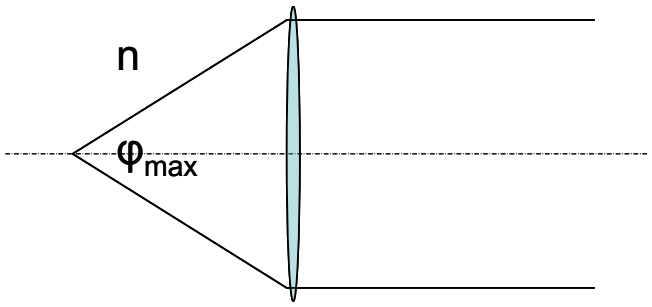


Figure 2.2 Illustration of the Numeric Aperture of a microscopy objective

The amplitude of the photon burst also depends on the emission rate of the fluorophore. For a fluorescent molecule that has a typical fluorescence lifetime of 4 ns, the maximum emission rate can reach 2.5 X 10^8 photons/s under optical saturation conditions if the quantum efficiency is assumed to be 100%. However this rate is often limited by alternative decay pathways of the excited fluorophore, such as intersystem crossing into long-live triplet state, where the fluorophore is not able to fluoresce for a relatively long period of time. These effects lead to premature saturation of the fluorophore and reduce the emission rate.

2.5.2 Fluorescence Background

Successful detection of fluorescence from single molecule events calls for a conspicuous photon burst above the background level or a large signal-to-background ratio. It is generally regarded among SMD practitioners that most single molecule techniques are background limited and improvements in the detectability of single-molecule events are more of a background issue than of a signal issue. Therefore,

suppressing the background level is the main challenge in the design of single molecule experiments.

2.5.2.1 Sources of Background

In single-molecule experiments, background refers to the detected photons not arising from the single molecule of interest. The background photons can arise from instrumental considerations, including dark counts from the detector, residual fluorescence from optical components such as colored glass filters and microscope objectives, and residual emission from the excitation laser. These background photons can be well controlled by scrupulously selecting premium components.

Background photons can come from the sample itself and are typically harder to control. Background photons arising from the sample contains two main sources: Rayleigh scattering and Raman scattering. Rayleigh scattering is elastic and appears at the same wavelength as the excitation source, which can be suppressed by using interference filters and ultraclean substrates. Raman scattering is inelastic and appears at a different wavelength than the excitation source and in many cases, can be in the spectral range as that of the fluorescence making it more difficult to remove. The magnitude of both scattering sources increases linearly with the sample volume. Autofluorescence from impurity molecules also contribute to fluorescence background and can be minimized by pre-photobleaching the solvent for a brief period of time before it reaches the detection window.[16]

2.5.2.2 Approaches to Reduce Background

2.5.2.2.1 Minimizing Probe Volume and Confocal Setup

In SMD experiments, the background photons from sample are proportional to the sampling volume, while the signal from a single fluorophore is independent of the probe

volume. This feature inspired single molecule experimentalist to investigate various schemes to scale down the detection volume to obtain a high signal-to-background ratio for detection. The excitation volume can be reduced by using a high numeric aperture microscope objective to tightly focus the laser beam into a small spot. The waist of a circular laser beam at the focal point is governed by the Rayleigh criterion:[17]

$$\omega = \frac{0.61\lambda}{NA} \qquad (2.6)$$

where ω is the beam waist, λ is the wavelength of the laser, and NA is the numeric aperture of the objective, assuming the spot size is diffraction limited. For example, for a laser working at 600 nm focused by a 60X objective with a NA of 0.85, the beam waist is 430 nm.

The probe volume can be effectively reduced by using a spatial filter like pinhole, which was first reported by Rigler and coworkers in their confocal microscope for single molecule sorting studies.[18] A typical confocal setup for LIF single-molecule detection is illustrated in Figure 2.3.

In this setup, the laser beam is tightly focused by a microscope objective with high numerical aperture into a small spot in the sample solution containing fluorescent molecules. The fluorescence photons from these fluorophores are collected by the same objective and spectrally separated from the excitation laser by a combination of interference filters placed in the optical path. A spatial filter, typically a pinhole, is placed in the back focal plane of the objective to reject out-of-focus light and allows only photons from the focal volume to reach the detector. Using this optical configuration individual fluorescent molecules could be detected with a S/B ratio ≥100.[19]

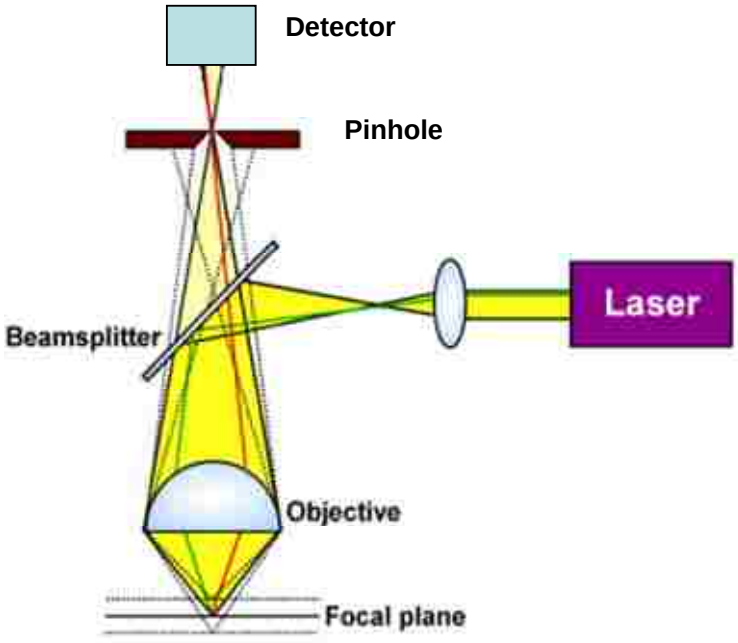


Figure 2.3 Illustration of confocal setup for LIF single-molecule detection. The laser beam is focused by a high NA objective into the fluorescent sample solution. The fluorescence photons from these fluorophores are collected by the same objective and spectrally separated from the excitation laser by a combination of interference filters. A pinhole is placed in the back focal plane of the objective to reject out-of-focus light and allows only photons from the probe volume to reach the detector.

2.5.2.2.2 Time-Gated Detection

Background photons due to Raman scattering that are difficult to remove by interference filters can be discriminated by setting a time gate in the detection channel because a fluorescence photon arriving at the detector is usually delayed by the average lifetime of the fluorophore, while photons arising from Raman scattering arrive at the detector almost instantaneously. In one of the early single-molecule experiments, Shera and co-workers used a pulsed laser that could generate a short laser pulse of 70 ps with an 82 MHz repetition rate to excite the R6G molecule. A time window of 4 ns, which is approximately the same as the lifetime of R6G, was set to analyze those photons appearing beyond that window. With the combination of pulsed-laser excitation

and time-gate discrimination, more than 99% of the scattered photons were removed from the single molecule trace data.[4]

2.5.2.2.3 Two-Photon Excitation

Instead of being excited by one photon, a fluorescent molecule can be excited by simultaneously absorbing two photons of the same wavelength.[20, 21] Figure 2.4 illustrates the differences between one-photon and two-photon excitation.

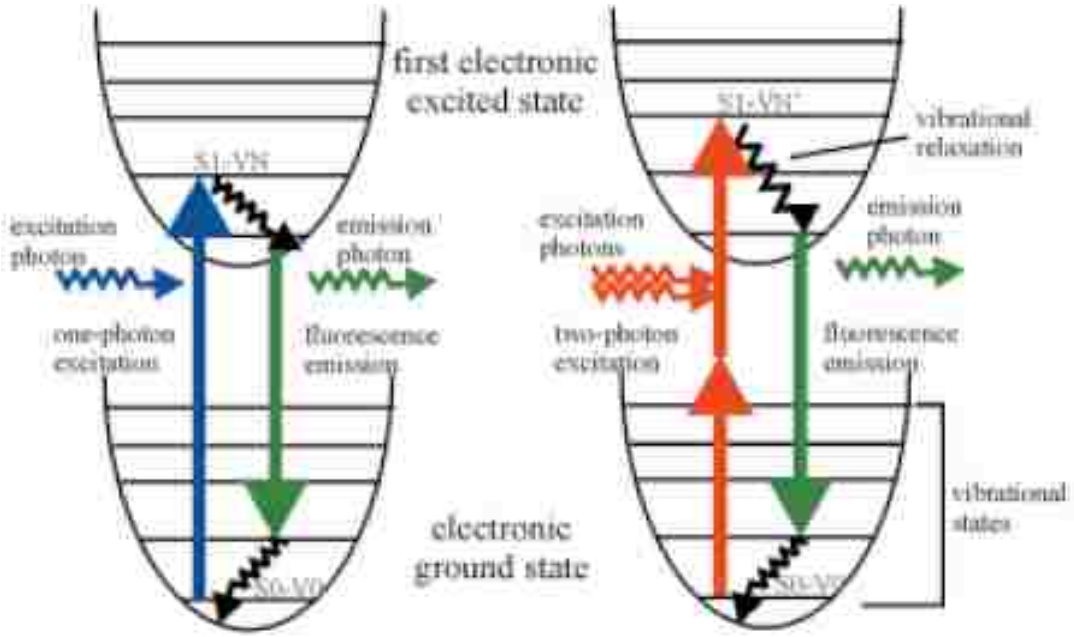


Figure 2.4 Jablonski Diagram for one-photon and two-photon excitation [22]

In contrast to one-photon excitation (OPE), two-photon excitation (TPE) produces extremely low scattering backgrounds due to its wide spectral separation. Because the energy of a photon is inversely proportional to its wavelength, two absorbed photons with twice the wavelength are required in TPE to produce the same electronic transition as that associated with OPE. For example, a fluorescent molecule that is normally excited by an ultraviolet photon at 300 nm can also be excited by two near-IR photons at 600 nm. As a result, the Rayleigh and Raman scattering can be readily filtered because their emission spectra are far away from the fluorescence emission spectrum.

The concept of TPE was actually proposed in as early as 1931, but was not experimentally realized until introduction of femto-second lasers, which have extremely high peak intensities to allow for significant two-photon excitation rates. The peak laser intensity is achieved by tightly focusing a mode-locked pulsed laser, which can generate extremely short laser pulses of 100 fs in width at a high repetition rate of 80 MHz (low duty cycle), yet the average laser power remains fairly low.[21] In TPE, the fluorophore is populated into the same singlet excited state as that in OPE, and emission process follows the same path as that in OPE.

2.6 Signal-to-Noise Ratio (SNR) and Signal-to-Background Ratio (SBR)

For an electronic measuring system, the quality of a measurement is usually evaluated by its signal-to-noise ratio (SNR), which determines the performance of the system. In SMD, the photon burst of single fluorescent molecules has to be detected above the background noise, which demands a high SNR.

The SNR for fluorescence detection of a single molecule is determined by the following equation:[14]

$$SNR = \frac{D\phi_F(\frac{\sigma_P}{A})(\frac{P_O}{h\nu})T}{\sqrt{(\frac{D\phi_F\sigma_P P_O T}{Ah\nu}) + C_b P_O T + N_d T}} \qquad (2.7)$$

where $P_O / h\nu$ is the number of incident photons per second, A is the beam area, T is the detector counting interval, ϕ_F is the quantum yield of the fluorophore, σ_P is the absorption cross section of the fluorophore, C_b is the background count rate per watt of excitation power, N_d is the dark count rate of the detector, and D is the instrument collection factor. D is an instrument dependent parameter, which is the product of a

combination of factors including the angular collection factor F_{coll} of the detection system, transmission factor F_{opt} of lenses, filter transmission factor F_{filter}, and quantum efficiency of detector η_Q, expressed by:

$$D = \eta_Q F_{coll} F_{opt} F_{filter} \qquad (2.8)$$

Equation 2.7 gives a comprehensive description of all factors that influence the SNR, but it is complicated and difficult to use because knowledge of many parameters is required *a priori*. In practice, the SNR of single-molecule detection can be evaluated directly from the data trace of photon burst by the following formula:[23]

$$SNR = \frac{n_f}{\sqrt{n_b}} \qquad (2.9)$$

where n_f represents the signal, and $\sqrt{n_b}$ represents the noise, or the mean fluctuation in the background.

Another used measure for the quality of SMD data is the signal-to-background ratio (SBR), which is defined as:

$$SBR = \frac{S - B}{B} \qquad (2.10)$$

where S is the signal strength of photon burst, and B is the background level of a blank sample, both of which are readily available from the data trace.

2.7 Photophysics and Photochemistry of Fluorophores

2.7.1 Anti-bunching

The photons emitted by a single fluorescent molecule are correlated by their arrival time at the detector. Measurement of the delay time of successive fluorescence photons should show anti-bunching, a phenomenon that was first observed for the pentacene

fluorophore in *p*-terphenyl host crystal in single-molecule emission at room temperature.[24] Briefly, emitted photons from a single fluorophore cannot follow each other arbitrarily close and have to be separated by a few nanoseconds, corresponding to the fluorescence lifetime of the fluorophore.[25] If two detected photons are within this time frame, they must have arisen from different emitting sources. Auti-bunching is a quantum signature of single-molecule spectroscopy and is purely due to the intrinsic quantum-mechanical effect of light.

2.7.2 Photobleaching

The repetitive cycling of fluorophore between the ground state and excited states eventually causes the fluorophore to be destroyed by photochemical reactions in the excited state, a process called photobleaching, where the chemical structure of the fluorescent molecule is irreversibly changed. Usually a fluorophore can continuously emit about 10^6 photons before photobleaching,[26] and it has been reported that a total of $1.7X10^6$ emitted photons were detected from R6G that was dissolved in ethanol.[13] It was generally regarded that oxygen molecules in the solution might be the culprit for photobleaching, because molecular oxygen can reacted with fluorophores in the triplet state and generate singlet oxygen, which further attacks the ground-state fluorophore and bleaches it.[25] Actually the maximum number of photons a fluorophore can emit is determined by the ratio of its fluorescence quantum yield to its photodestruction quantum yield.[27]

2.7.3 Fluorescence Intermittence

Stochastic behavior has been observed in some single-molecule experiments, which is characterized by fluctuating or blinking of fluorophores when they travel

through the laser excitation volume.[14] For example in an early single-molecule lifetime measurement using NSOM, some fluorescent molecules showed fluctuations in the detected emission even though a laser with fixed-wavelength was used.[28] This was caused by shifting of the absorption band in and out of the resonance energy of the fluorophore being probed due to phonon-driven fluctuations in the vicinity of the fluorophore being probed, which is termed “spectral diffusion”. In contrast to amplitude fluctuation of fluorescence, the emission of a fluorophore can be completely turned off momentarily and turned back on shortly after, which is termed “blinking”. Blinking has been shown for a single green fluorescent protein and its emission was switched on and off sporadically even though the laser was continuously shedding light onto the sample.[29] Intersystem crossing could be one of the causes of blinking because transition into the non-emitting triplet state interrupts the fluorescence continuity, and gives rise to on and off times of detected photons.[30]

2.8 Near-IR Fluorescence Dyes for SMD

The group of organic cyanine dyes such as Cy3, Cy5, Cy5.5 and Cy7, which emit in the near infra-red (IR) range, offer a number of advantages over dyes fluorescing in visible range making them ideal probes for SMD applications. As we know, Raman scattering from solvent molecules is a major contributor to background levels, which is difficult to be removed completely by interference filters due to spectral overlap with the single-molecule fluorescence. Because the amplitude of Raman scattering is inversely proportional to the 4^{th} power of the excitation wavelength,[12] it tends to diminish by a factor of $1/\lambda^4$ when the excitation wavelength approaches the near-IR region. For example, the Raman scattering level for detecting around 820 nm is reduced by greater

than six-fold compared to detection at 500 nm. Also, fluorescence from impurities is reduced or eliminated due to the fact that only a small fraction of compounds are known to demonstrate intrinsic fluorescence in near-IR region.[31] The reduction in overall background fluorescence results in enhancement of the signal-to-noise ratio for SMD. Moreover, semiconductor diode lasers working at different wavelengths in near-IR region have become commercially available, which are inexpensive and compact making them suitable for development of a portable instrument.

2.9 Excitation Modes and Optical Layout for SMD

2.9.1 Wide-Field Epi-fluorescence

The most straightforward approach to observe single molecule behavior is through wide-field epi-fluorescence microscopy. The optical configuration of this approach is illustrated in Figure 2.5.

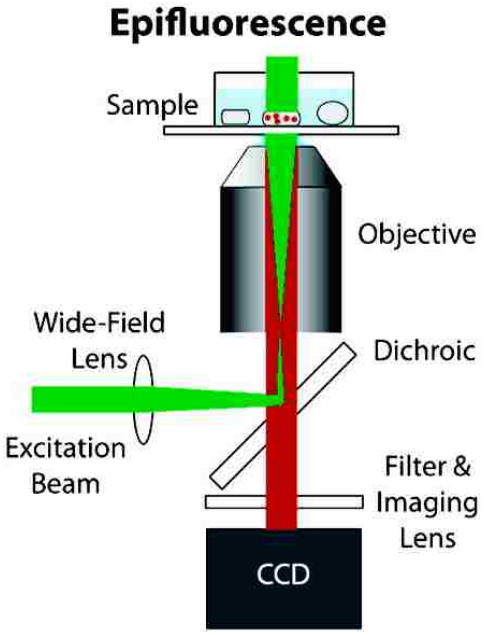


Figure 2.5 Optical layout of wide-field epi-fluorescence microscopy for SMD [32]

In this configuration, a collimated laser beam is first focused at the back focal plane of the microscope objective, then de-focused by the objective to form a collimated excitation beam impinging on the sample. Fluorescence emission from the sample is collected by the same microscope objective, filtered by a dichroic mirror and a series of interference filters before the sample is imaged onto a CCD camera. Because excitation light in the epi-illumination mode is not able to directly get to the detector, which is the case in the transmission illumination mode, background signal due to Rayleigh scattering is minimized. This approach has been used to study fluorophore labeled single biological molecules that were either attached on a surface[33] or freely diffusing in solution.[34] The advantage of wide-field is a large number of samples can be illuminated and observed simultaneously, where throughput of detection is a major concern. But it also suffers from reduced S/B ratio resulting from a large detection volume.

2.9.2 Far-field Confocal Microscopy

In wide-field epi-fluorescence microscopy, a large volume of sample is excited and the single-molecule signal in many cases is overwhelmed by a large background level. This problem can be overcome by employing a pinhole in the optical path to reject out-of-focus photons referred to as confocal microscopy (see Figure 2.6).

In confocal microscopy, a collimated laser beam slightly overfills the back aperture of the microscope objective and is directed and focused into a diffraction-limited spot at the sample. The fluorescence is collected by the same objective and processed by a dichroic mirror and interference filters before reaching a point detector like an avalanche photodiode (APD). Because a small pinhole (50 – 100 μm) is placed at the secondary

image plane of the objective, photons from out-of-focus regions are blocked resulting in high signal-to-noise ratio for fluorescence detection.

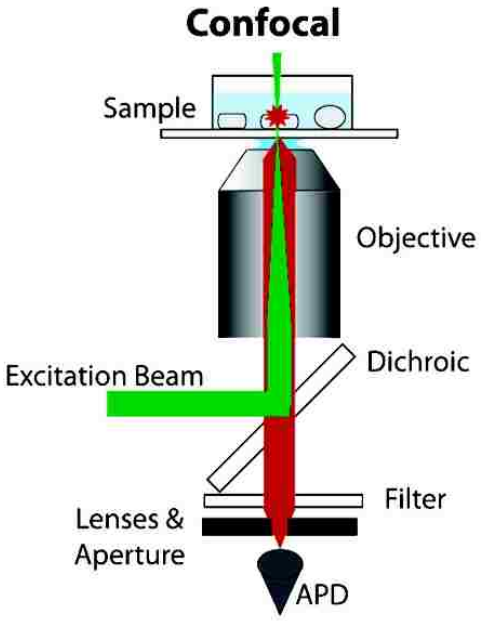


Figure 2.6 Optical layout of confocal microscopy for SMD [32]

The concept of using confocal arrangement for single-molecule detection was first proposed by Rigler *et al.* in their studies of fluorophore-tagged DNA molecules using fluorescence correlation spectroscopy.[18] It has been used for monitoring sensitive DNA hybridization dynamics and conformational transitions at the single-molecule level.[35, 36] The disadvantage of confocal microscopy is that only one point in the sample can be observed at one time. In order to obtain the total image of the sample of interest, a scanning stage is required to observe the sample point-by-point and combining the individual pixels to construct the image.

2.9.3 Total Internal Reflection

In addition to confocal microscopy, another frequently used way to reduce out-of-focus florescence in SMD is through evanescent wave excitation. As is known in

fundamental optics, when a collimated beam strikes the boundary between two optical media with different refractive indices at the critical angle with respect to the normal of the boundary surface, the beam will not be able to cross the boundary and an evanescent wave field is generated along the boundary interface, whose intensity decays exponentially into the low-index medium, according to equation:[37]

$$I(z) = I(0) \cdot e^{-\frac{z}{d}} \qquad (2.11)$$

where $I(0)$ is the intensity at the boundary interface, $I(z)$ is the intensity at a distance z into the low-index medium, and d is the characteristic decay distance, defined as:

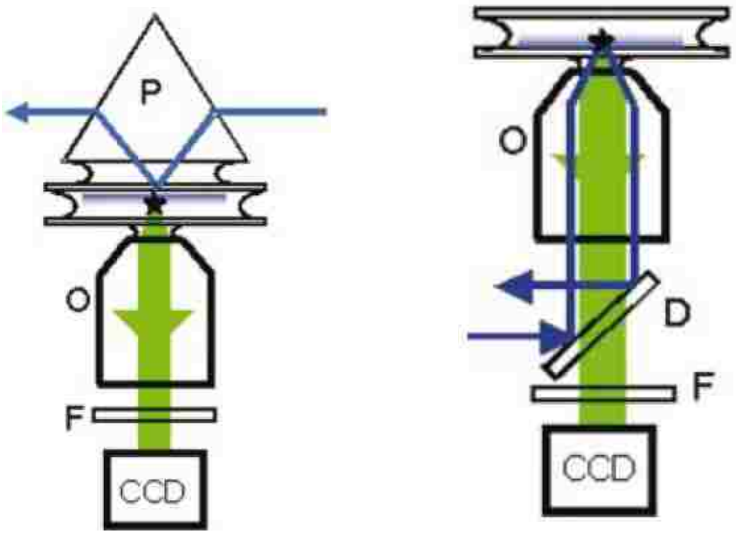
$$d = \frac{\lambda}{2\pi}(n_2^2 \sin^2\theta - n_1^2)^{-1/2} \qquad (2.12)$$

where n_1 is the refractive index of the low-index medium, n_2 is the refractive index of the high-index medium, θ is the critical angle, and λ is the wavelength of excitation beam.

Typically, the evanescent field is only a few hundred nanometers thick that is comparable to the lateral dimension of the laser spot waist in the confocal setup. For the evanescent excitation only sample sufficiently close to the interface can be excited, which greatly reduces the background levels. In practice, evanescent wave excitation is implemented by total internal reflection (TIR), which is illustrated in Figure 2.7.

In TIR, the excitation laser beam is brought to the glass/quartz coverslip through a prism with an index matching oil at an angle slightly above the critical angle (Figure 2.7 A), so that the excitation beam can be completely reflected back into the prism and the evanescent wave is created. Then, the fluorescent molecules that are either freely diffusing in solution or attached to the coverslip can be imaged by a CCD camera. TIR can also be generated by guiding the laser beam through the very edge of the

microscope objective with a high numerical aperture (shown in Figure 2.7B). The primary advantage of TIR is that only an extremely thin layer of the sample is probed thus the interfering background from the bulk sample is minimized. Yeung's group pioneered applications of TIR for single-molecule studies. Diffusion and photobleaching characteristics of single fluorescent molecule like R6G in solution were investigated quantitatively by single-molecule imaging.[38] Moreover, dynamics of fluorophore labeled individual λDNA molecules in free solution was also monitored and analyzed with high throughput of thousands of molecules per second by a TIR instrument.[39]



(A) (B)

Figure 2.7 Optical configuration of total internal reflection for SMD [14]

2.9.4 Breakthrough of Diffraction Limit

SMD using classical optical methods is limited by diffraction and the spatial resolution of conventional microscopy is about half of the wavelength of the excitation light.[26] To circumvent the diffraction limit, novel approaches such as NSOM and zero-mode waveguides have been used which can provide sub-wavelength resolution.

2.9.4.1 NSOM

Near-field scanning optical microscopy (NSOM) is an alternative method for single molecule imaging, which can provide sub-wavelength resolution. In NSOM the excitation light is transmitted to the sample surface by a modified optical fiber, which is pulled to a tapered tip of 50-100 nm in diameter, and coated with metal like aluminum of 100 nm thick on the sides of the tip to prevent light leakage.[40-42] The optical fiber is brought within 5-10 nm to the sample surface for the sample to be sufficiently illuminated. The fluorescence emission from the sample is detected in the far-field using an optical setup similar to confocal microscopy. The target sample surface is raster scanned point-by-point by a two-dimensional scanning stage, and each illuminated spot is imaged onto an APD. The integral image of the sample surface can be synthesized by recombining these individual image pixels in the order they are acquired. Because the diameter of the optical probe is much smaller than the wavelength of the excitation light, the resolution of the image is actually determined by the size of the aperture on probe, rather than by the light wavelength.

NSOM has been used to image single dye molecules, study molecular dynamics, and determine fluorophore's dipole orientation attributed to its extraordinary resolution and sensitivity.[40, 41] However, this technique has a few limitations that restrict its widespread applications. Firstly, NSOM suffers from low laser power throughput, which is only 1-50 nW for a 50-100 nm tip.[43] This results in weak signal strength in detection of single molecules. Secondly, the reproducibility of tip preparation is very poor and sample surface can be perturbed by the coated fiber tip.[28, 44]

2.9.4.2 Zero-mode Waveguide

Traditionally single-molecule analysis is especially applicable to very dilute samples, because single molecule occupancy will not be obtained until the sample is diluted in a low concentration range. For example, in fluorescence correlation spectroscopy (FCS) using confocal microscopy, the probe volume can be reduced to 0.2 fL, and the working concentration of fluorophores is usually in the nM range.[18] But many enzymatic reactions *in vivo* require much higher ligand concentrations, usually in the range of μM to mM, and low ligand concentrations will inevitably affect the enzyme kinetics.[45] In order to successfully study and measure enzyme kinetics arising from single molecule reactions in natural environments, the observation volume needs to be reduced by at least three orders of magnitude.

Zero-mode waveguide (ZMW) is an innovative approach developed by Craighead and co-workers using nano-fabrication technology to effectively obtain an array of observation volumes of zeptoliters for parallel analysis of multiple single molecule reactions.[46] The scheme of ZMW is illustrated in Figure 2.8.

The ZMW is fabricated with a thin metal film like aluminum deposited on a microscope coverslip and an array of small holes with diameters of about 50 nm are produced by electron-beam lithography. The holes act as the core of a waveguide and the surrounding metal film acts as its cladding. Because the diameter of these holes is much smaller than the wavelength of excitation light, no light can actually be guided inside the holes (zero-mode propagation) and light intensity decays exponentially along the light path at the entrance of the holes, which results in zeptoliter (10^{-21} L) detection volumes. In order to monitor the activity of individual enzymes, single enzyme

molecules are immobilized on the bottom of each hole, and solution containing the reaction mixture is placed on top of this coverslip to allow reactants to diffuse into the holes. A microscope objective of high numerical aperture is placed underneath the coverslip to focus excitation light into these holes and collect the fluorescence emission from single molecule reactions. ZMW has been used to monitor DNA polymerization in real-time,[47] and to study the dynamics of lipid diffusion in cell membranes.[48]

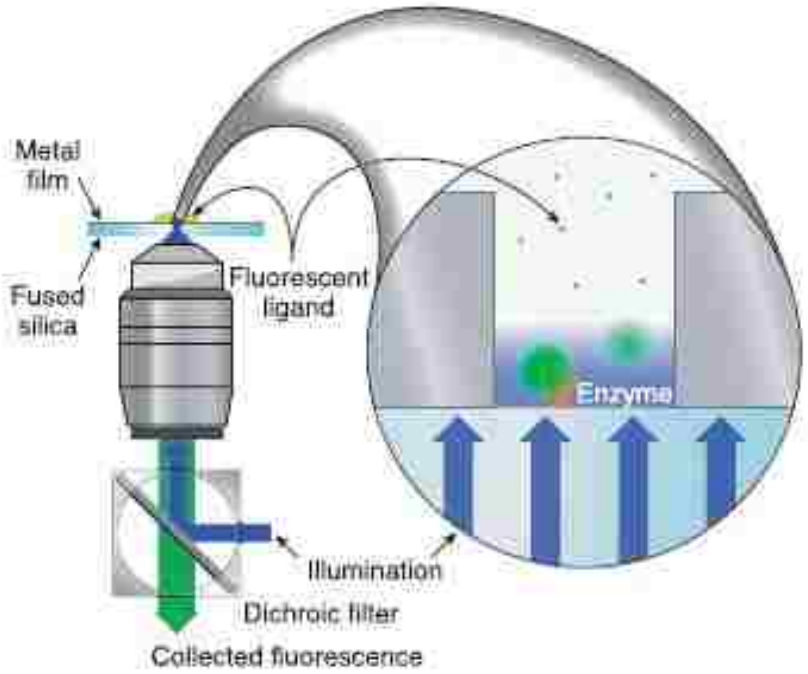


Figure 2.8 Illustration of zero-mode waveguide for single-molecule detection [46]

Finally, what needs to be specially pointed out in this discussion is that in both NSOM and ZMW, the light is propagated in the evanescence wave fashion, where the photons are confined into a sub-diffraction region at the cost of extremely low power throughput and extremely short propagation distance. These drawbacks inherent to these techniques have to be taken into account when their sensitivity and resolution are pursued in certain SMD applications.

2.10 Detectors for Single-molecule Detection

In order to capture rare photon flux emitted from a single fluorophore, the detector used in SMD has to have short response times, low dark counts, high quantum efficiency in the wavelength range of interest, and low noise levels. There are basically two types of detectors used in single-molecule studies, point detector and two-dimensional array detectors, both of which rely on the photoelectric effect on either photocathodes or semiconductors.

2.10.1 Point Detectors

Point detectors are used in confocal systems to count emitted photons from a small probe volume. There are two types of point detectors for SMD, photomultiplier tubes (PMTs and single-photon avalanche photodiodes (SPADs). In both detectors, each captured photon is converted into a photo-electron or electron-hole pair, whose number is then multiplied by several orders of magnitude in an avalanche fashion.

2.10.1.1 PMT

PMT is a vacuum device and consists of two main elements, photocathode for photon to electron conversion and dynodes for photon-electron multiplication.[8] When a photon impinges onto the photocathode that is deposited as a thin film on a glass window in a vacuum tube, a photon-electron is released with a finite probability (quantum efficiency), and accelerated to the first dynode under a positive potential held between the dynode and the photocathode. When this electron arrives at the first dynode, it gains enough momentum to further generate a number of secondary electrons, which are subsequently accelerated towards the second dynode because this dynode is held at a higher potential than the first dynode. This process is repeated

through a chain of dynodes, until all electrons finally arrive at the anode where they are detected. By this amplification fashion, a single incident photon is converted into millions of electrons that can be easily read by electronic devices.

2.10.1.2 SPAD

SPAD was introduced in 1980's as a cost-effective alternative to PMTs for photon counting applications. SPAD is a solid state semiconductor photo-detector with an on-board amplifier.[49] In SPAD, each incoming photon impinging on the diode creates an electron-hole pair at the p-n junction of the diode by promoting an electron from the valence to the conduction band leaving a vacancy (hole) in the valance band. Then, a divergent avalanche current pulse is triggered that is strong enough to be detected by electronic device. In order to generate the avalanche current, the diode has to be biased above its breakdown voltage to sustain the high electric field. Once a photon initiates an avalanche current, the voltage has to be reset below the breakdown voltage to dampen the current so that the subsequently arriving photon can be detected. This scheme results in a dead time of 30-40 ns, and limits its maximum counting rate to a few MHz. In earlier SPAD detector, passive quenching was applied by placing a large resistor R_Q (~ 100 kΩ) in series with the SPAD. The high avalanche current triggered by a photon together with the high resistance of R_Q drops the electric field across the SPAD below the breakdown voltage. Then, the detector is ready to detect the next photon. However, this large resistance results in long dead times. In active quenching, the biased voltage is quickly lowered by employing a quenching circuit, lowering the dead time.

2.10.2 Two-Dimensional Array Detectors

Charge-coupled device (CCD) is a semiconductor detector consisting of many thousands of photon sensing elements arranged in a two-dimensional array format.[50]

When an image is projected onto the photoactive region of a CCD through a lens, photoelectrons are generated on each sensing element (pixel) and the number of electrons is proportional to the intensity of incident light at that pixel. These sensing elements are physically segregated from each other by metal oxide barriers to prevent free migration of charges. A number of microelectrodes are attached to each element so that it acts as a charge capacitor and photoelectrons are trapped in this potential well when the appropriate voltage is applied. The stored charges in the sensing elements of each column can be transferred to the next adjacent column by alternating voltages applied to each individual element, and the charges in the last column are dumped into a shift register and read out after being amplified and converted into a voltage signal.

Based on its readout scheme, CCDs can be operated in a snapshot mode or time-delayed integration (TDI) mode.[51] In the snapshot mode, a shutter is used to control the exposure time of the detector to the incident light, and the CCD takes a snapshot of the image when the shutter is open. When the shutter is closed, the accumulated charges in the imaging area are transferred to its adjacent storage area, which contains an opaque mask. Then, the stored image is slowly read out during the time the shutter is re-opened and a new image is being exposed to the photoactive area. The long delays between consecutive snapshots however limit the use of this operation mode in tracking a moving objective continuously.

In TDI mode, the shutter is kept open constantly and the shifting rate of the columns is synchronized to the migration rate of the object that is being tracked. The photo-generated charges are integrated for the entire time the object stays in the observation zone before they eventually get to the shift register and are read out. This operation

mode greatly reduces the read time as well as the read noise of the CCD and allows detection of migrating objects with improved SNR.

2.10.3 Comparison of Point Detectors and Array Detectors

A detailed side-by-side comparison of commonly used photon counting devices in single-molecule experiments is summarized in Table 2.2.

Table 2.2 Comparison of different photon counting detectors [52]

	PMT	SPAD	ICCD *(Intensified CCD)*	EMCCD *(Electron Multiplied CCD)*
Photocathode	Multi Alkali	Silicon	Multi Alkali	Silicon
QE (at 600 nm)	6%	80%	6%	90%
Gain	10^6	10^6	10^6	10^3
Time Resolution	100 ps	100 ps	100 ms	100 ms
No of Pixels	1	1	~1024 x 1024	~1024 x 1024
Detection Area	1 cm x 1 cm	180 µm (*dia.*)	1 cm x 1 cm	1 cm x 1 cm
Count Rate	10 MHz	10 MHz	100 kHz	1 MHz
Readout Speed	1 MHz	1 MHz	20 Hz	20 Hz

Early experiments of single molecule fluorescence in flowing samples were accomplished using PMT,[2, 3] due to its adequate time resolution, and low dark counts level when electronically cooled. But PMT suffers from a low quantum efficiency (QE), which is usually <20% in the visible region, limiting its use in SMD. On the other hand, PMT possesses a large photo-active area on the order of 1 cm X 1 cm, which is convenient in some cases. SPAD, by contrast, has a much higher QE of 80% in the near-IR region due to the special band structure of silicon. Both PMT and SPAD have a dark count of <100 counts/s and time resolution on the order of 100 ps. This temporal resolution allows them to be used in demanding applications such as time-correlated single-photon counting (TCSPC) for fluorescence lifetime measurements. A limitation of

SPAD compared to PMT is its small photo-sensing area of ~200 μm in diameter. This small sensing area of SPAD can actually serve as an extra pinhole in confocal SMD setup to further block out-of-focus photons, which favors a higher S/B ratio. Point detectors are usually used to monitor photon burst arising from single molecules at a fixed spatial position.

Array detectors have a large sensing area containing thousands of sensing elements, which makes it suitable to image many single molecule events simultaneously and provide spatial resolution of these events. But CCD also suffers from readout noise during each analog-to-digital conversion. In addition, its low time resolution in milliseconds due to slow readout makes it impossible to resolve arrival time of each photon.

2.11 References

(1) Hirschfeld, T. *Appl Opt* 1976, *15*, 2965-2966.

(2) Dovichi, N. J.; Martin, J. C.; Jett, J. H.; Trkula, M.; Keller, R. A. *Analytical Chemistry* 1984, *56*, 348-354.

(3) Nguyen, D. C.; Keller, R. A.; Trkula, M. *Journal of the Optical Society of America B-Optical Physics* 1987, *4*, 138-143.

(4) Shera, E. B.; Seitzinger, N. K.; Davis, L. M.; Keller, R. A.; Soper, S. A. *Chemical Physics Letters* 1990, *174*, 553-557.

(5) Soper, S. A.; Davis, L. M.; Shera, E. B. *Journal of the Optical Society of America B-Optical Physics* 1992, *9*, 1761-1769.

(6) Moerner, W. E.; Kador, L. *Phys Rev Lett* 1989, *62*, 2535-2538.

(7) de Mello, A. J. *Lab on a Chip* 2003, *3*, 29n-34n.

(8) Lakowicz, J. R. *Principles of fluorescence spectroscopy*, 3rd ed.; Springer: New York, 2006.

(9) Culbertson, M. J.; Williams, J. T.; Cheng, W. W.; Stults, D. A.; Wiebracht, E. R.; Kasianowicz, J. J.; Burden, D. L. *Anal Chem* 2007, *79*, 4031-4039.

(10) Li, L. Q.; Davis, L. M. *Appl Opt* 1995, *34*, 3208-3217.

(11) Moerner, W. E.; Kador, L. *Analytical Chemistry* 1989, *61*, A1217-A1223.

(12) McWhorter, S.; Soper, S. A. *Electrophoresis* 2000, *21*, 1267-1280.

(13) Soper, S. A.; Shera, E. B.; Martin, J. C.; Jett, J. H.; Hahn, J. H.; Nutter, H. L.; Keller, R. A. *Analytical Chemistry* 1991, *63*, 432-437.

(14) Moerner, W. E.; Fromm, D. P. *Review of Scientific Instruments* 2003, *74*, 3597-3619.

(15) Ambrose, W. P.; Goodwin, P. M.; Jett, J. H.; Van Orden, A.; Werner, J. H.; Keller, R. A. *Chem Rev* 1999, *99*, 2929-2956.

(16) Affleck, R. L.; Ambrose, W. P.; Demas, J. N.; Goodwin, P. M.; Schecker, J. A.; Wu, M.; Keller, R. A. *Anal Chem* 1996, *68*, 2270-2276.

(17) Masters, B. R. *Confocal microscopy and multiphoton excitation microscopy : the genesis of live cell imaging*; SPIE Press: Bellingham, Wash., 2006.

(18) Eigen, M.; Rigler, R. *Proc Natl Acad Sci U S A* 1994, *91*, 5740-5747.

(19) Bernard Valeur, J.-C. B. *New Trends in Fluorescence Spectroscopy: Applications to Chemical and Life Sciences*, 2001.

(20) Mertz, J.; Xu, C.; Webb, W. W. *Opt Lett* 1995, *20*, 2532.

(21) Denk, W.; Strickler, J. H.; Webb, W. W. *Science* 1990, *248*, 73-76.

(22) So, P. T.; Dong, C. Y.; Masters, B. R.; Berland, K. M. *Annu Rev Biomed Eng* 2000, *2*, 399-429.

(23) Peck, K.; Stryer, L.; Glazer, A. N.; Mathies, R. A. *Proc Natl Acad Sci U S A* 1989, *86*, 4087-4091.

(24) Basche, T.; Moerner, W. E.; Orrit, M.; Talon, H. *Physical Review Letters* 1992, *69*, 1516-1519.

(25) Xie, X. S.; Trautman, J. K. *Annual Review of Physical Chemistry* 1998, *49*, 441-480.

(26) Moerner, W. E.; Orrit, M. *Science* 1999, *283*, 1670-+.

(27) Soper, S. A.; Mattingly, Q. L.; Vegunta, P. *Analytical Chemistry* 1993, *65*, 740-747.

(28) Ambrose, W. P.; Goodwin, P. M.; Martin, J. C.; Keller, R. A. *Science* 1994, *265*, 364-367.

(29) Dickson, R. M.; Cubitt, A. B.; Tsien, R. Y.; Moerner, W. E. *Nature* 1997, *388*, 355-358.

(30) Zumbusch, A.; Fleury, L.; Brown, R.; Bernard, J.; Orrit, M. *Phys Rev Lett* 1993, *70*, 3584-3587.

(31) Patonay, G.; Antoine, M. D. *Analytical Chemistry* 1991, *63*, A321-A326.

(32) Lord, S. J.; Lee, H. L.; Moerner, W. E. *Anal Chem* 2010, *82*, 2192-2203.

(33) Funatsu, T.; Harada, Y.; Tokunaga, M.; Saito, K.; Yanagida, T. *Nature* 1995, *374*, 555-559.

(34) Schmidt, T.; Schutz, G. J.; Baumgartner, W.; Gruber, H. J.; Schindler, H. *Proc Natl Acad Sci U S A* 1996, *93*, 2926-2929.

(35) Kinjo, M.; Rigler, R. *Nucleic Acids Res* 1995, *23*, 1795-1799.

(36) Edman, L.; Mets, U.; Rigler, R. *Proc Natl Acad Sci U S A* 1996, *93*, 6710-6715.

(37) M. F. Paige, E. J. B., and W. E. Moerner *Single Molecule* 2001, *2*, 11.

(38) Xu, X. H.; Yeung, E. S. *Science* 1997, *275*, 1106-1109.

(39) Ma, Y.; Shortreed, M. R.; Yeung, E. S. *Anal Chem* 2000, *72*, 4640-4645.

(40) Betzig, E.; Chichester, R. J. *Science* 1993, *262*, 1422-1425.

(41) Xie, X. S.; Dunn, R. C. *Science* 1994, *265*, 361-364.

(42) Nie, S. M.; Zare, R. N. *Annual Review of Biophysics and Biomolecular Structure* 1997, *26*, 567-596.

(43) Harris, T. D.; Grober, R. D.; Trautman, J. K.; Betzig, E. *Applied Spectroscopy* 1994, *48*, A14-A21.

(44) Valaskovic, G. A.; Holton, M.; Morrison, G. H. *Applied Optics* 1995, *34*, 1215-1228.

(45) Schomburg, D.; Schomburg, I.; Chang, A. *Springer handbook of enzymes*, 2nd ed.; Springer: Berlin ; New York, 2001.

(46) Levene, M. J.; Korlach, J.; Turner, S. W.; Foquet, M.; Craighead, H. G.; Webb, W. W. *Science* 2003, *299*, 682-686.

(47) Korlach, J.; Levene, M.; Turner, S. W.; Craighead, H. G.; Webb, W. W. *Biophysical Journal* 2002, *82*, 507a-507a.

(48) Edel, J. B.; Wu, M.; Baird, B.; Craighead, H. G. *Biophysical Journal* 2005, *88*, 195a-195a.

(49) Li, L. Q.; Davis, L. M. *Review of Scientific Instruments* 1993, *64*, 1524-1529.

(50) Hanley, Q. S.; Earle, C. W.; Pennebaker, F. M.; Madden, S. P.; Denton, M. B. *Analytical Chemistry* 1996, *68*, A661-A667.

(51) Sweedler, J. V.; Shear, J. B.; Fishman, H. A.; Zare, R. N.; Scheller, R. H. *Analytical Chemistry* 1991, *63*, 496-502.

(52) Tohru Ohnukia, X. M., Arun Tripathia, Shimon Weissb, Katsushi Arisaka *Proc. of SPIE* 2006, *6092*, 9.

CHAPTER 3 CONSTRUCTION OF A CONFOCAL LASER-INDUCED FLUORESCENCE SYSTEM FOR DETECTION OF SINGLE DNA MOLECULES IN A THERMOPLASTIC MICROCHIP

3.1 Single-molecule Detection in Microfluidic Chip

In recent years chemical analysis has migrated from benchtop devices to microfluidic platforms arising from its lower sample consumption, shorter analysis time, disposability, higher throughput, and potential for automation. However, ordinary detection techniques on microfluidic chips are challenged by the significantly reduced sample size associated with miniaturization of the analytical device. For LIF single-molecule detection, an ultra-small probe volume was created by the exquisite confocal setup and only a small fraction of analytes were interrogated, thus the detection is not subject to the sample size. This consequently provided an effective method for on-chip detection. The first example of single-molecule detection (SMD) in microfluidic chip was demonstrated by Effenhauser, where an electrophoresis microchip was fabricated in polydimethylsiloxane (PDMS) using soft lithography and single λ–DNA molecules loaded with intercalating dyes were observed when they migrated through the microfluidic channel.[1] In another example presented by Ramsey and co-workers, fluorescent molecules were electrodynamically driven through microfluidic channel that was patterned on glass substrate using photolithography and wet chemical etching.[2] Two types of chromophore molecules, Rhodamine 6G and Rhodamine B, were separated and fluorescence bursts from individual molecules were counted in this microchip. Haab et al. demonstrated that single DNA sizing ladders of 100-1000 bp long can be fluorescently resolved by performing capillary gel electrophoresis in a glass microfluidic chip to separate these ladders followed by single-molecule photon burst

counting.[3] Gösch studied the velocity profile of hydrodynamic flow within microchannels fabricated on a silicon wafer by scanning the microchannel in both the vertical and horizontal directions and detecting the photon burst from single tetramethylrhodamine molecules on different spatial locations in the channel.[4] Wabuyele et al. reported the detection of single double-stranded DNA (dsDNA) molecules that were electrokinetically pumped in microfluidic channels fabricated on poly(methylmethacrylate) (PMMA) and polycarbonate (PC) substrates.[5] In this implementation, near-IR intercalating dyes were used to stain the dsDNA molecules so that autofluorescence arising from substrate material was minimized in the laser-induced fluorescence detection.

In contrast to single-molecule fluorescence detection with the classic confocal setup, where a collimated laser beam was tightly focused into a tiny spot in the sample solution and fluorescence photon bursts from individual dye molecules were recorded sequentially by a single-element detector such as SPAD, wide-field imaging provides another approach in single-molecule measurement, where the behavior of many single molecules can be monitored simultaneously by a multi-element detector such as CCD. Kang and Yeung reported observation of single DNA molecules in a microchip fabricated on PDMS and glass.[6] In their study, the migration of individual λ–DNA molecules were directly observed by a differential interference contrast microscopy without fluorescence labeling. Emory and Soper demonstrated the detection of many individual dsDNA molecules in a high-throughput fashion on a PMMA microchip using a CCD camera operated in time-delayed integration mode.[7] These fluorescently labeled dsDNA molecules were driven through parallel microfluidic channels and a collimated laser beam was launched into the microchannels in an orthogonal format from the side

of the microchip so that these molecules can be illuminated and detected simultaneously when they passed through the laser beam. Okagbare reported an innovative optical setup for high-throughput single-molecule detection performed on a PMMA microchip consisting of 30 microchannels.[8] In his design, the DNA molecules in microchannels were excited in an epi-illumination format, where the collimated laser beam was pre-focused behind the input aperture of the microscope objective so that a large field-of-view illumination zone was formed to span all 30 microchannels and single DNA molecules flowing through different microchannels can be excited and observed simultaneously.

In this chapter, a laser-induced fluorescence single-molecule detection system will be constructed using confocal optical setup. Different substrate material will be compared to select the most appropriate one to fabricate the microfluidic chip for single-molecule measurement. Various aspects of single-molecule detection experiments will be investigated and discussed in great details so that optimal experimental conditions can be obtained to facilitate subsequent studies.

3.2 Experimental

3.2.1 Reagents and Materials

The dye labeled single-stranded DNA molecules used in this study were synthesized by Integrated DNA Technologies (IDT, Coralville, IA) and resuspended in 1X TE buffer. The sequence of the oligonucleotide strand was: 5'- CGCGCCGCCT-3', and the fluorescent molecule Cy5.5 was attached to the 5'-end of this oligonucleotide strand as a reporter. All buffers and nuclease-free water were purchased from Ambion (Carlsbad, California) and used as received. The substrate poly(methylmethacrylate),

PMMA, was obtained from Goodfellow (Oakdale, PA). Cyclic olefin copolymer, COC, was purchased from Topas Advanced Polymers (Florence, Kentucky). The polymer substrates were heated in an oven at 70ºC overnight before hot-embossing.

3.2.2 Confocal Setup of the LIF Instrument

The laser-induced fluorescence (LIF) system for single molecule detection was built in-house on an optical breadboard and is illustrated in Figure 3.1.

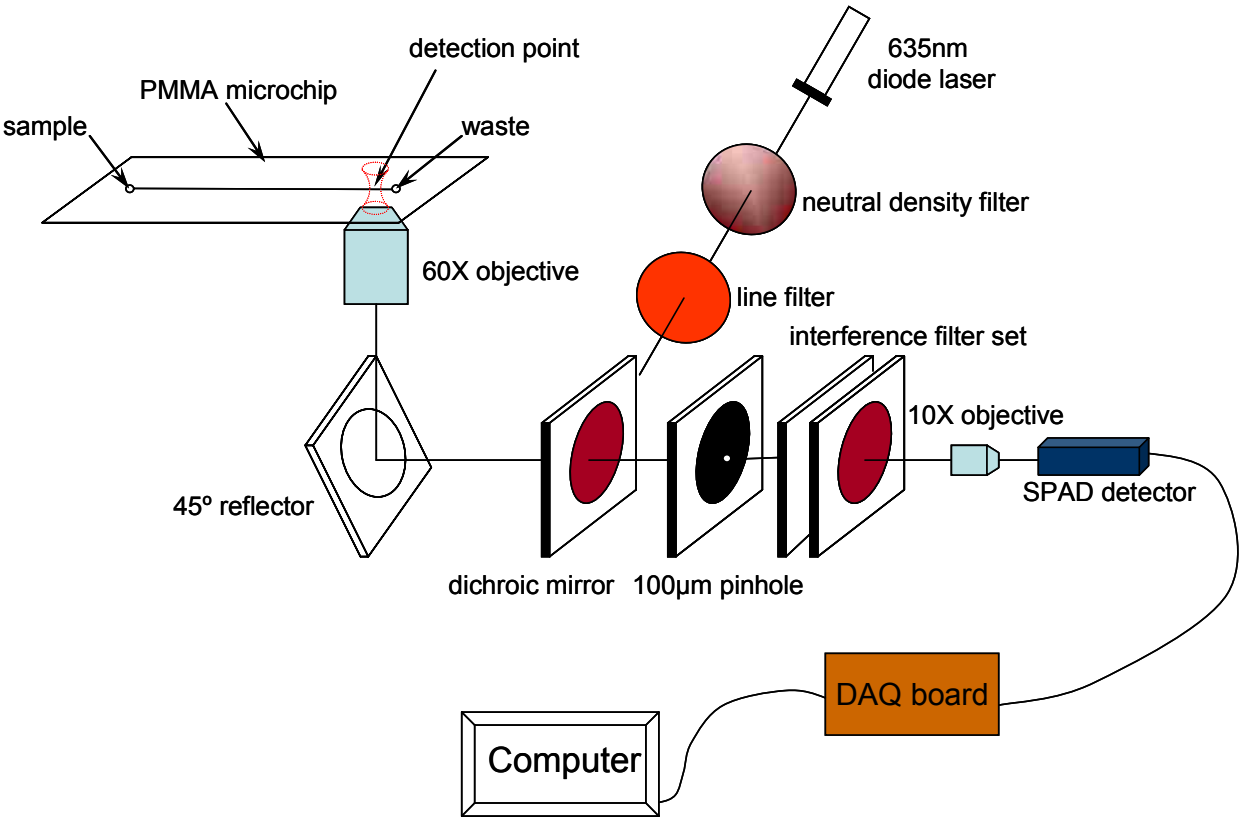


Figure 3.1 Illustration of confocal LIF instrument for single-molecule measurement

The excitation source consisted of a diode laser (Model CPS198, Thorlabs, Newton, NJ) lasing at 670 nm and was collimated by a collimation lens. The collimated laser beam was further conditioned by a laser line filter (670DF10, Omega, Brattleboro, VT). The laser beam was then directed into a microscope objective (M-60X, NA = 0.85,

Newport, Irvine, CA) by a dichroic mirror (Z670RDC, Chroma Technology, Rockingham, VT) and an optical reflector, both of which were positioned at a 45º angle of incidence.

The emitted fluorescence from flowing molecules inside the microchannel was collected by the same objective and transmitted through a dichroic into a pinhole (i.d. = 100 μm), which was positioned in the confocal plane of the microscope objective and served as a spatial filter to eliminate any out-of-focus light. The fluorescence light was further conditioned using a combination of interference filters, which contained a longpass filter (3RD690LP, Omega Optical) and a bandpass filter (3RD700-750, Omega Optical). The fluorescence light was focused by a 10X microscope objective onto the active area of a SPAD (SPCM-200, EG&G, Vandreuil, Canada). The pulses from the SPAD were transformed into TTL pulses and processed using a PCI-6602 digital counting board (National Instruments, Austin, TX) for subsequent data analysis.

3.2.3 Hydrodynamic Pumping of Single DNA Molecules

When single molecule measurements were performed on the microchip, the Cy5.5 labeled DNA molecules were diluted to appropriate concentrations in 1X TE buffer prior to use, and loaded into a glass syringe (SGE, Austin, TX), which was connected to the microchip through a capillary. Both the syringe and the microchip were thoroughly rinsed with isopropanol and ddH_2O before sample loading and fluorescence measurement. The sample was loaded into the glass syringe that hydrodynamically drove the DNA molecules through the microfluidic device using a syringe pump (PicoPlus, Harvard Apparatus, Holliston, MA).

3.3 Results and Discussion

3.3.1 Selection of Substrate Material for Microchip

The strategy in single-molecule detection is to make the photon burst of single-molecule events as bright as possible compared with fluctuations in the background. The intensity of photon burst is generally determined by the photophysical and photochemical properties of the dye molecules, optical setup of the LIF system, and performance of the photodetector. Thus, the background plays a crucial role in detecting single-molecule events, and the detection limit of single molecules is mainly limited by the count rate of the background in the measurement.

It has been widely reported that polymer substrates shows a wide range of autofluorescence levels in LIF measurement.[9] PMMA and COC are two commonly used substrates for microfluidic devices and LIF applications due to their exceptional optical properties. In order to evaluate the suitability of PMMA and COC in single-molecule measurement, microfluidic chips were hot embossed into PMMA and COC substrates. A straight channel of 100 μm wide and 100 μm deep was patterned on this microchip and the sample reservoir and waste reservoir were drilled, separately, on each end of the microchannel. 1X TE buffer was filled into the microchannel to examine the background fluorescence of the microchip.

Figure 3.2 gives a side-by-side comparison of autofluorescence background between PMMA and COC microchips. The autofluorescence level was found to be about 1,000 counts/s for COC substrate and 5,000 counts/s for PMMA substrate. To indentify real single-molecule events in sample solution, a threshold criterion is generally set to discriminate single molecule photon burst against the background fluorescence.[10] Choosing a low threshold value allows more single-molecule photon

bursts to be identified, whereas more false positive bursts will also be counted. Increasing the threshold level eliminates false positives, whereas a real single-molecule event may also be ignored (false negative). It was found empirically that three times the average background fluorescence was an appropriate threshold level to essentially eliminate the majority of background fluctuations so that any photon bursts that exceeded this threshold could be counted as real single-molecule events. By this criterion, the single-molecule threshold level was set at 3,000 counts/s for COC, and 15,000 counts/s for PMMA (see Figure 3.2).

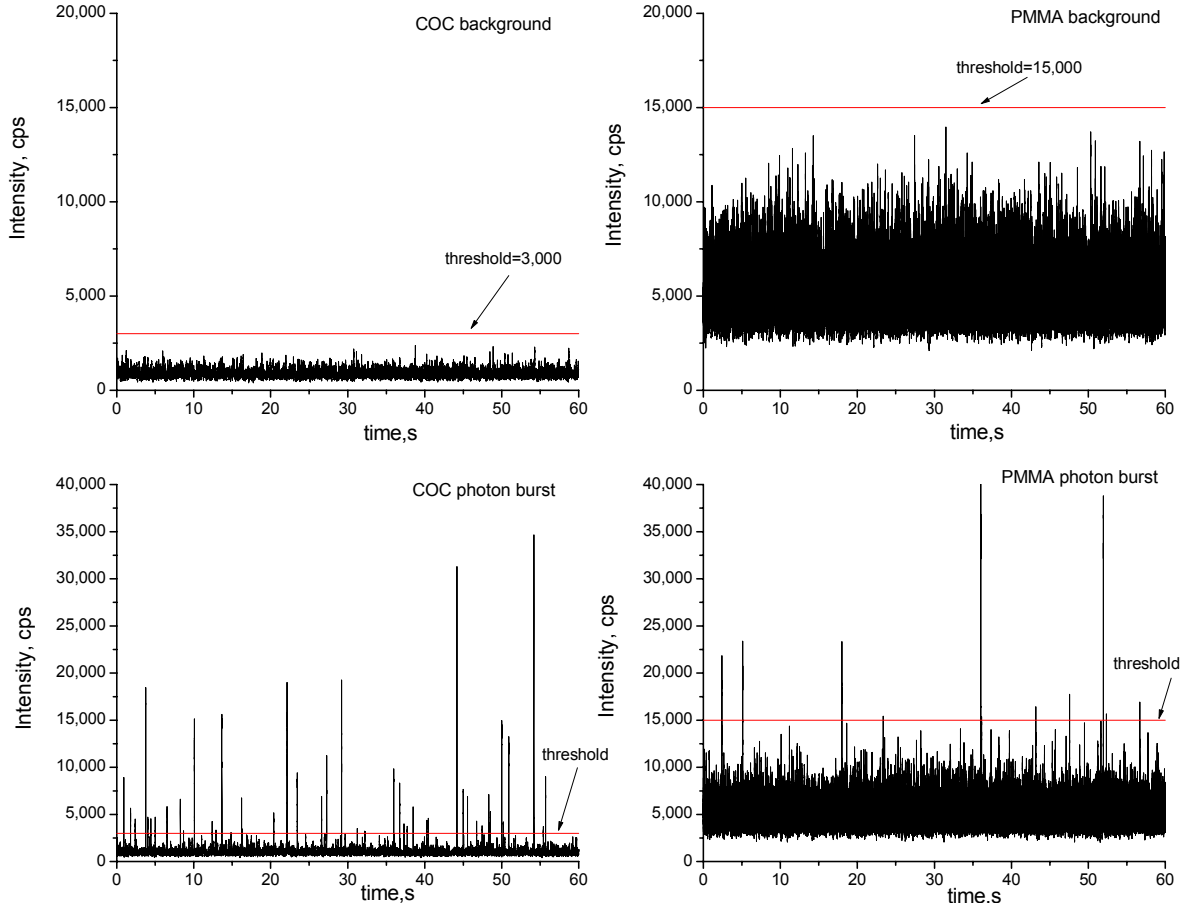


Figure 3.2 Comparison of PMMA and COC microchip for single-molecule measurement. Cy5.5 labeled oligonucleotides were diluted to 0.5 pM and driven at a flow rate of 0.3 μL/min in microchannel. The channels in both PMMA and COC chips are 100 μm wide and 100 μm deep. The output of the laser power was adjusted to 1 mW for excitation.

To evaluate the effect of background fluorescence on detectability of single molecule events, Cy5.5 labeled oligonucleotides were diluted to 0.5 pM and driven through the microchip at a volumetric flow rate of 0.3 μL/min. The single-molecule trace data are displayed in Figure 3.2. The photon burst events have comparable magnitude for PMMA and COC microchips, and there are a total of 43 single-molecule events that were identified above the threshold level of 3,000 counts/s on the COC microchip. This indicates that about 0.05% of the Cy5.5 labeled DNA molecules were detected since during the 60 s sampling time for photon burst registration there were 90,000 molecules flowing through the microchannel. However, when the same oligonucleotide sample was analyzed in the PMMA microchannel, only 10 events were identified, corresponding to a detection efficiency of about 0.01%. This is because in PMMA chip photon bursts with smaller peak height were easily overwhelmed by the increased background level. In order to achieve high single-molecule detection efficiency, COC is determined to be the substrate of choice for microchip fabrication in all subsequent studies attributed to its exceptionally low autofluorescence background.

3.3.2 Optimization of Excitation Laser Power

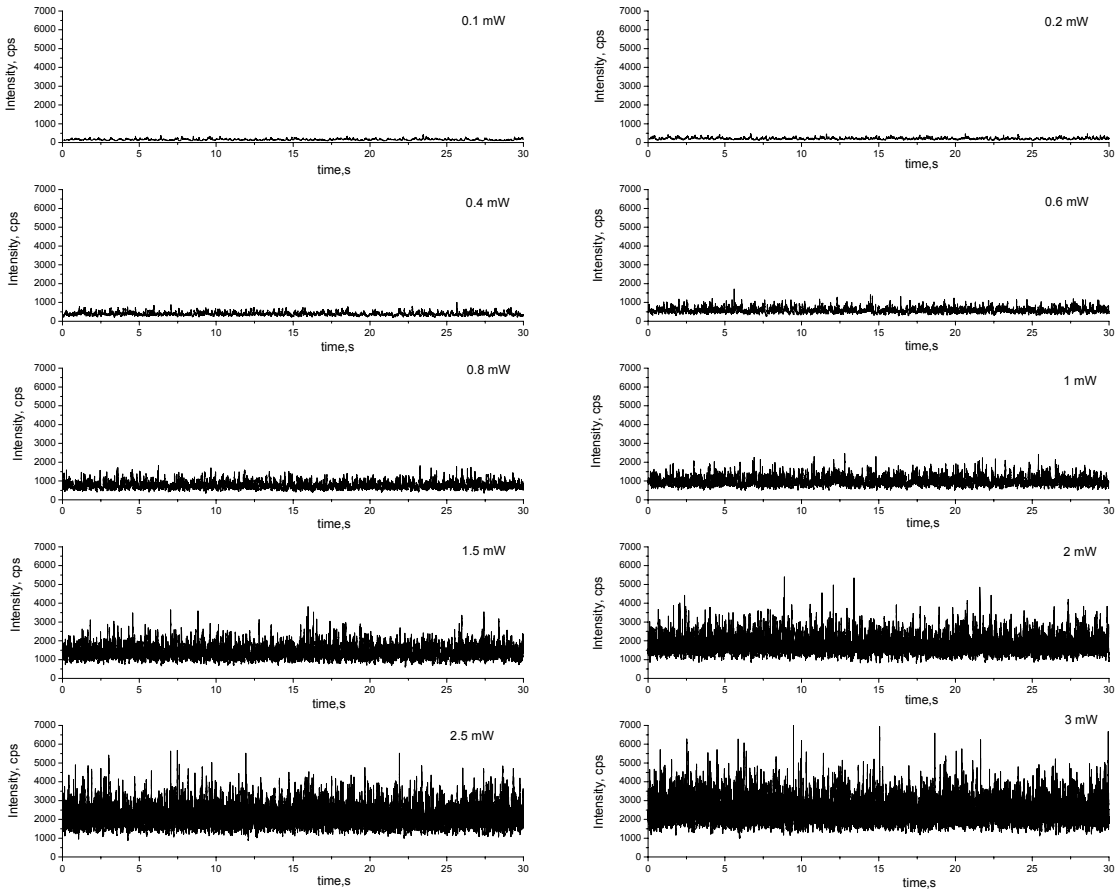
Higher SNR or SBR are always pursued in any analytical measurement, which, in this LIF single-molecule detection technique, strongly depends on the photon density supplied by the excitation laser. As illustrated in the Jablonski diagram, fluorescence of a single fluorophore mainly involves two steps: 1) excitation from the ground electronic state to an excited electronic state by absorption of a photon. 2) Radiative or non-radiative decay from the excited state to the ground state. Below optical saturation, the fluorescence increases almost linearly with laser power. Above optical saturation, the

vast majority of fluorophores are promoted to their excited states causing ground-state depletion, thus increasing the intensity of incident photon flux will not help the fluorophore absorb more photons and saturation occurs. In contrast, the background scattering increases linearly with laser intensity.

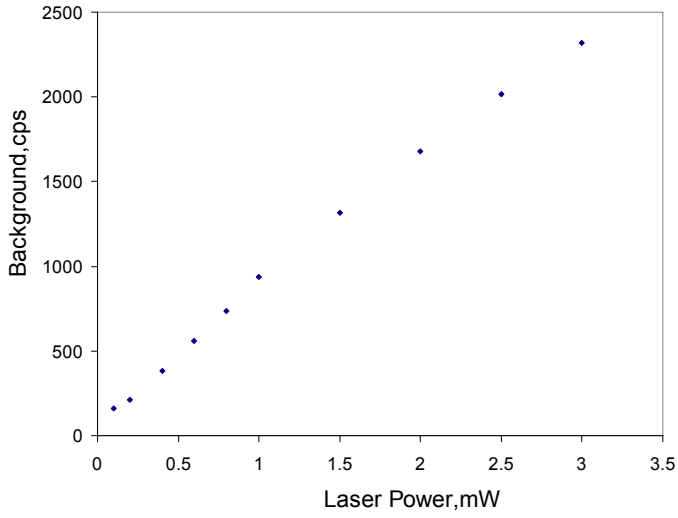
The COC microchip that was filled with 1X TE buffer and excited at different incident laser power levels with the results illustrated in Figure 3.3. It is clearly seen that as laser power increases, the background scattering increased linearly.

To evaluate the signal amplitude of single DNA molecules at different excitation intensities, the Cy5.5 labeled oligonucleotides were diluted to 1pM in 1X TE buffer and driven through the COC microchannel at a volumetric flow rate of 0.2 μL/min. The excitation laser intensity was set from 0.1 mW to 3 mW by tuning the neutral density filter in the LIF system; the photon burst data at each laser power are shown in Figure 3.4. Ideally, all photon bursts should have uniform amplitude if they travel exactly through the center of the focused laser beam. Practically, however, because the cross section of the microchannel (100 x 100 μm) is much larger than the focused laser spot (1 × 2 μm in diameter), a significant fraction of the DNA molecules flow through the microchannel without being excited. For those molecules that happen to pass through the laser spot, there still exists a large disparity in the magnitude of these photon bursts due to the Gaussian intensity of the focused laser spot.

The top 10 peaks with the highest burst intensity in each data trace were picked and their average peak intensity was utilized to represent the signal fluorescence at each laser power level. The SBR, defined as (S-B)/B, was obtained for each laser power level and is displayed in Figure 3.5. In this plot we find that the SBR increases rapidly until it



(A)



(B)

Figure 3.3 Background scattering of COC microchannel as a function of excitation laser power. The microchannel was filled with 1X TE buffer. (A) Trace data of background scattering. (B) Background scattering increased linearly with excitation laser power.

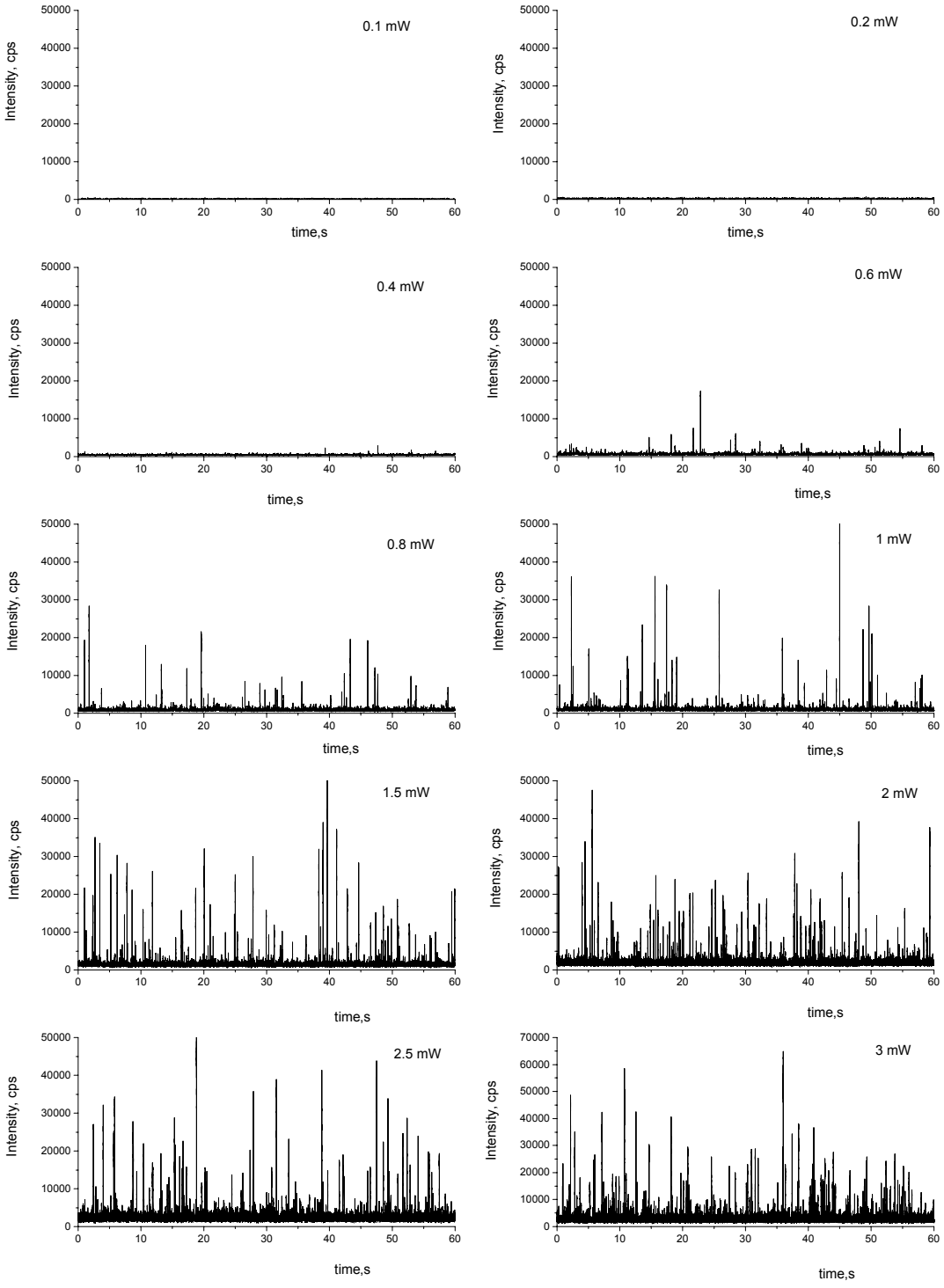


Figure 3.4 Data trace of single-molecule photon burst at different excitation laser power. Cy5.5 labeled oligonucleotides were diluted to 1pM in 1X TE buffer and driven through the COC microchannel at a volumetric flow rate of 0.2 μL/min. The excitation laser intensity was set from 0.1 mW to 3 mW.

reaches its maximum when the laser power increased from 0.1 mW to 1 mW. However, when the laser power was further increased, the SBR dropped, which is the result of saturation of the fluorophores at high illumination intensity. At 1 mW, the average photon burst intensity from the top 10 peaks was 29,158 counts/s against the background fluorescence of 939 counts/s, with a SBR of 30 that was found to be the optimal value.

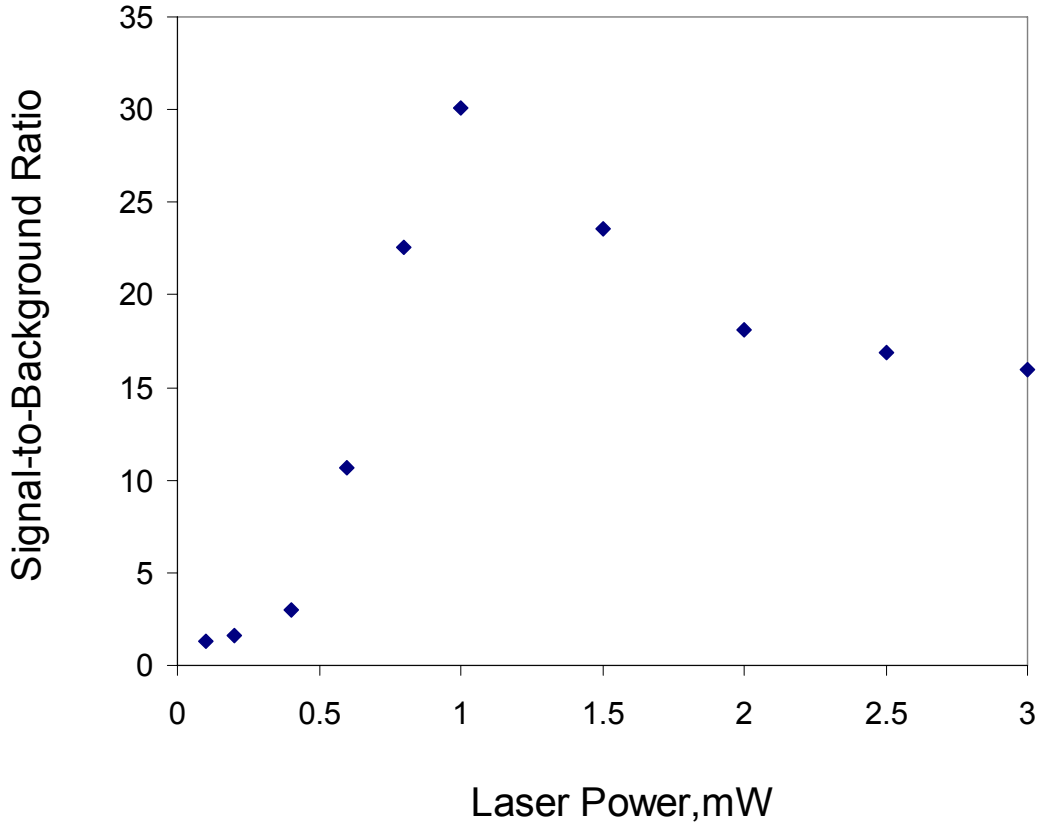


Figure 3.5 SBR of single-molecule measurement as a function of excitation laser power. In determining the signal intensity of photon burst, the top 10 peaks with the highest burst intensity in each data trace were picked and their average peak intensity was utilized to represent the signal fluorescence at each laser power level.

3.3.3 Autocorrelation Analysis and Transit Time of Single Molecules

In single molecule measurements, the photon bursts of single molecule events are a random process, which implies no fixed time interval between photon burst events. But

there still exists a non-random feature in this measurement because all detected single molecules have to pass through the laser spot at the same linear velocity. Autocorrelation analysis has been found to be a sensitive indicator of single molecule events when they travel through the laser beam.[11] The normalized autocorrelation function (ACF) in this application is defined as:[12]

$$g(\tau) = \frac{\langle F(t)F(t+\tau)\rangle}{\langle F(t)\rangle\langle F(t+\tau)\rangle} \qquad (3.1)$$

where $F(t)$, $F(t+\tau)$ are the photon intensity at t, and $t+\tau$, respectively, and

$$\langle F(t)F(t+\tau)\rangle = \underset{T\to\infty}{Lim}\frac{1}{2T}\int_{-T}^{T} F(t')F(t'+\tau)dt' \qquad (3.2)$$

Figure 3.6 shows the normalized autocorrelation function of buffer as well as a 2 pM dye labeled oligonucleotide solution that was driven at different volumetric flow rates. We can see that there is no non-random correlation in the ACF for the buffer only. However, for the dye labeled DNA solution, there is a strong non-random feature. This non-random feature in the autocorrelation function is due to correlated photon bursts from single dye molecules when they pass through the laser beam and fluoresce.

The width of the autocorrelation function also provides information about the average time a single molecule spends in the focused laser beam. By fitting the non-random feature of the ACF to a Gaussian function, the average transit time of the molecules can be obtained by reading at the $1/e^2$ height of the ACF peak. For example, the average molecular transit time at a flow rate of 0.8 μL/min was found to be 0.5 ms (Figure 3.6). Figure 3.7 graphs the relationship between the average transit time of single molecules and the reciprocal of each corresponding volumetric flow rate. The

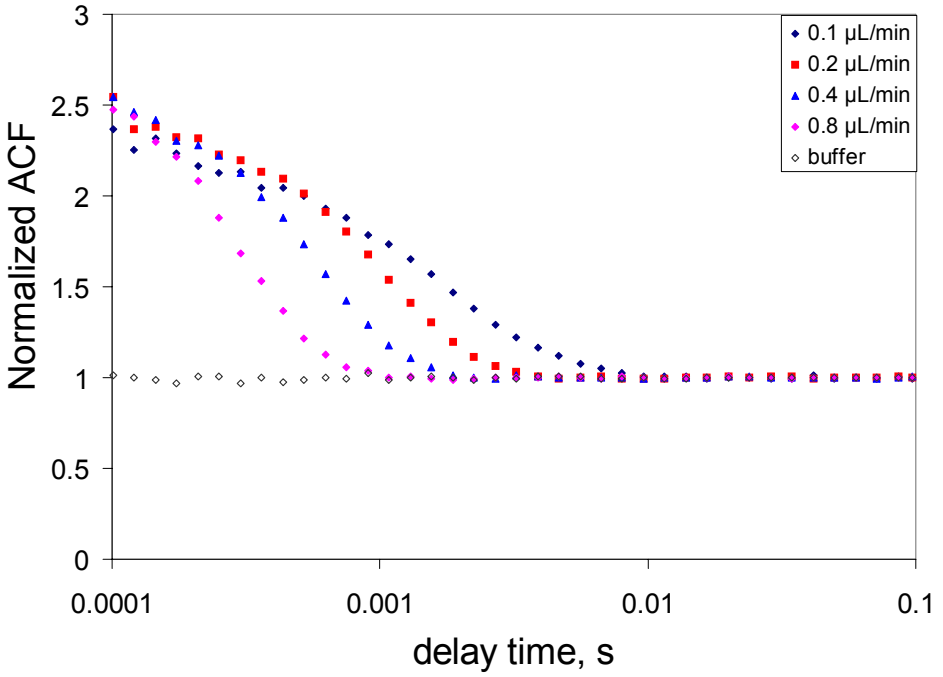


Figure 3.6 Effect of volumetric flow rate on normalized autocorrelation function. Cy5.5 labeled oligonucleotides were diluted to 2 pM and driven through the COC microchannel at a volumetric flow rate of 0.1-0.8 µL/min.

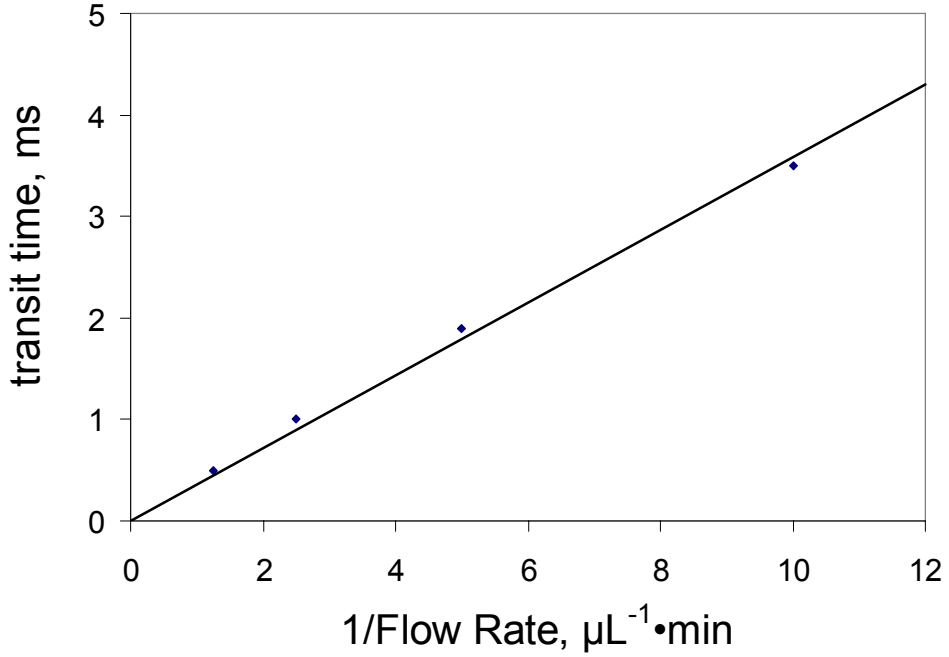


Figure 3.7 Plot of average molecular transit time with volumetric flow rate

faster these molecules traverse the laser beam, the shorter time they spend within the laser beam.

The average transit time of single molecules can also be estimated from the volumetric flow rate used in the experiment. For example, the average linear velocity of molecules at a volumetric flow rate of 0.8 μL/min was 1.33 mm/s, considering the cross sectional dimension of the microchannel (100 x 100 μm). The velocity profile of this laminar flow is parabolic, and the flow velocity in the center of the channel is twice the average flow velocity, which turns out to be 2.66 mm/s. The average transit time through the laser beam can be evaluated by:[11]

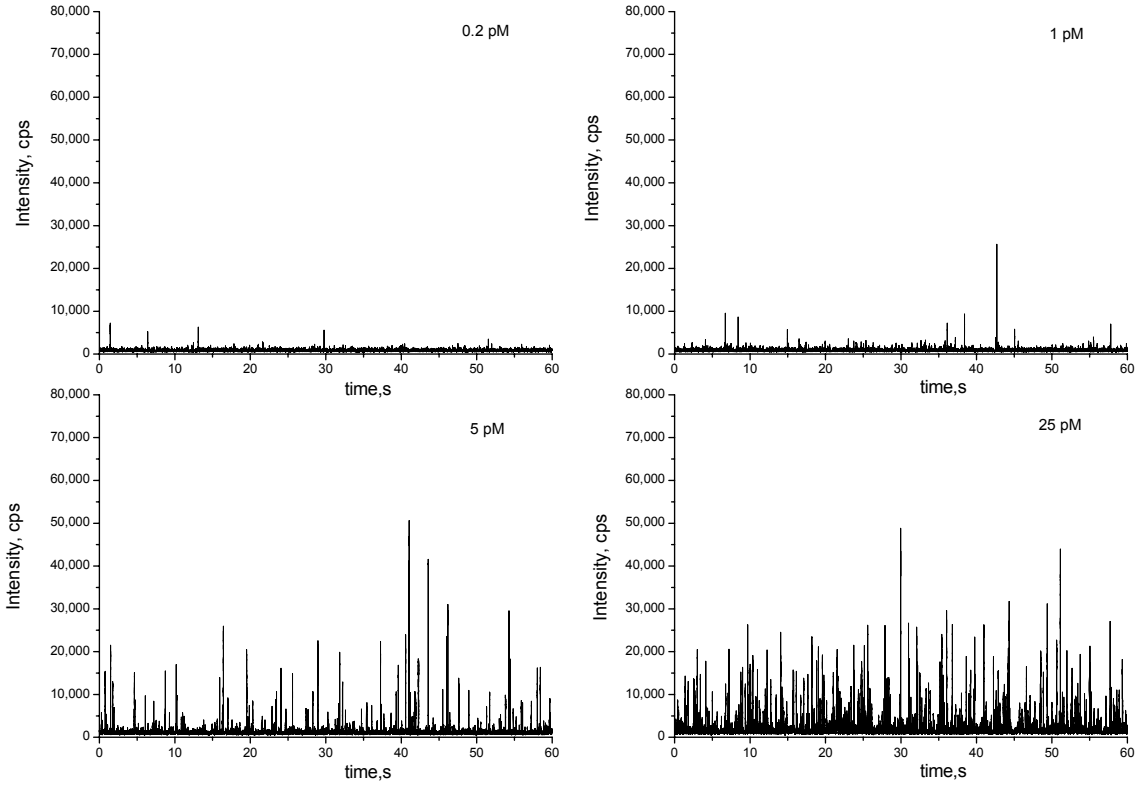
$$\tau = \frac{\pi\omega}{2\upsilon} \qquad (3.3)$$

where ω is the $1/e^2$ beam waist of the laser spot, which is 1 μm for this laser. The average transit time τ is determined to be 0.59 ms from this equation, which agrees well with the value obtained from the ACF.

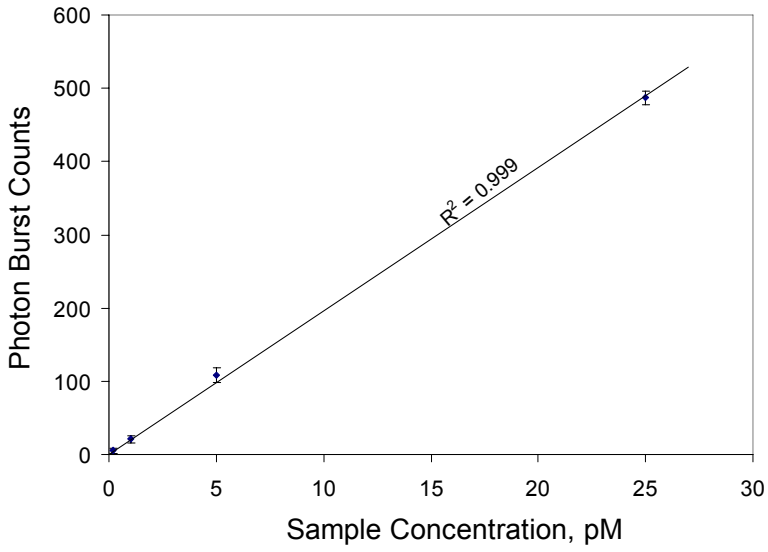
3.3.4 Single-Molecule Detection for Digital Molecule Counting

Conventional ensemble chemical analysis operates like an analog circuit, where the magnitude of the signal is proportional to the amount of input. In contrast to ensemble measurements, SMD operates like a digital circuit, where the magnitude of the TTL pulse is inconsequential and only the number of pulses matters, which is directly proportional to the input number.

The digital counting capability of SMD is demonstrated in Figure 3.8. Cy5.5 labeled DNA molecules were first diluted into a series of concentrations and driven through the COC microchannel at a volumetric flow rate of 0.1 μL/min. The photon burst spectra were collected for each sample and shown in Figure 3.8 (A). In this data, each photon



(A)



(B)

Figure 3.8 Counts of single-molecule photon bursts for Cy5.5 labeled oligonucleotides as a function of analyte concentration. The DNA molecules were diluted to a concentration of 0.2-25 pM in TE buffer. The sample solution was driven through the microchannel at a volumetric flow rate of 0.1 μL/min. (A) Trace data of photon burst. (B) Counts of photon burst vs. concentration of DNA molecules.

burst arises from individual Cy5.5 labeled DNA molecules and the counts of photon burst reveal the number of DNA molecules present. For the sample of 0.2 pM, there are only 5 photon bursts identified, whereas for the sample of 25 pM, there are a total of 529 photon bursts observed. The relationship between the counts of photon bursts and the sample input concentration is plot in Figure 3.8 (B). As seen, as the concentration of the sample solution increased, the observed numbers of photon bursts increased in a linear fashion with a R^2 value of 0.999 for the linear plot.

For the current digital counting attempts, only a very small fraction of molecules were interrogated by the photodetector due to the fact that the diameter of the focused laser beam (1 μm in diameter) was much smaller than the cross sectional dimension of the microchannel (100 × 100 μm). This technical bottleneck can be broken through the fabrication of a nanofluidic device, by making a nanochannel with a cross sectional dimension below 100 nm. This extremely thin channel would be overfilled by the focused laser spot and the sampling efficiency would approach unity.

3.3.5 Photon Burst Amplitude Distribution

In single-molecule measurements, we are also interested in understanding the distribution in the amplitudes of the photon bursts, which is affected by the illumination profile within the excitation volume. For an illumination profile with uniform photon intensity distribution, all photon bursts must have equally high amplitudes if these molecules travel at the same velocity. For a typical confocal LIF single-molecule detector, the collimated laser beam is tightly focused to a diffraction limited spot with the $1/e^2$ beam waist defining the excitation volume and the photon intensity within the excitation volume has a Gaussian distribution. When a single fluorescent molecule is

brought into this excitation volume, a photon burst is produced due to repetitive cycling of the fluorophore from the ground to excited state with subsequent relaxation back to the ground state accompanied by photon emission. The magnitude of the burst is directly proportional to the local photon density experienced by the molecule during its passage through the laser. The distribution pattern of the photon burst amplitude can thus serve as an indicator of the intensity profile of the excitation volume.

An experiment was run by driving Cy5.5 labeled DNA molecules through the microchannel at a flow rate of 0.4 μL/min. The DNA molecules were diluted to a concentration of 0.5 pM to keep the occupancy probability low in order to observe single molecule events. The photon burst data was plot and is shown in Figure 3.9 (A).

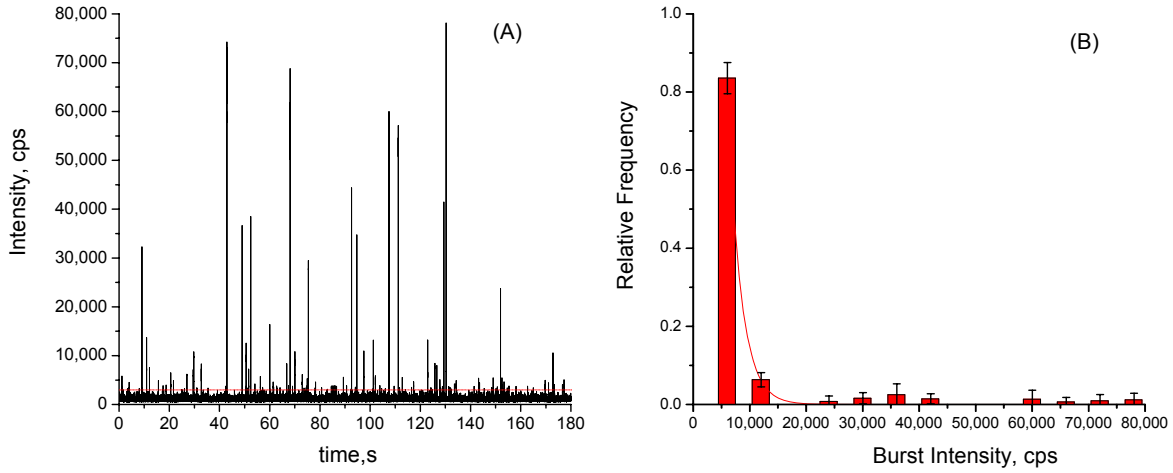


Figure 3.9 Amplitude histogram of photon burst from single-molecule measurement. Cy5.5 labeled oligonucleotides were diluted to a concentration of 0.5 pM and driven through the microchannel at a flow rate of 0.4 μL/min. (A) Trace data of photon burst. (B) Histogram plot. The threshold was set at 3,000 counts/s, and the photon burst magnitude was binned with an increment of 6,000 counts/s above the threshold level. The counts in each bin were normalized to the total number of photon bursts.

A histogram plot on the height distribution of the photon burst was constructed from this photon burst data trace and shown in Figure 3.9 (B). The data trace in Figure 3.9 (A)

was split into three consecutive 1 min long data traces so that their photon burst distribution could be averaged to get the mean and standard deviation to check the consistency of the distribution. In this histogram plot, the photon burst magnitude was binned with an increment of 6,000 counts/s above the threshold level, and the counts in each bin were normalized to the total number of photon bursts. It was found that the photon burst intensity was best fit to a single exponential, and the fitting curve using the exponential function was overlaid with the histogram in Figure 3.9 (B). This is not surprising, considering the Gaussian profile of light intensity within the focused laser spot. In this figure, we found the amplitude of photon bursts were dominantly distributed in the range between 3,000 to 6,000 counts/s with a relative frequency of 0.836 ±0.04. The reduced χ^2 value of the least square fit of this histogram was 1.48 × 10^{-4} and the R^2 value of the fit was 0.997.

3.3.6 Single-Molecule Detector for Analog Fluorescence Measurement

Figure 3.10 shows the fluorescence intensity of a series of dye labeled DNA solutions whose concentration ranged from 1 pM to 10 nM. The fluorescence intensity was averaged over the course of a 1 min sampling time in each run. When the concentration of the analyte was low, the detected fluorescence intensity increased almost linearly with concentration (see inset of Figure 3.10). This linearity was maintained until the concentration of the sample was set above 1 nM. Above this concentration, the fluorescence intensity leveled off, which suggested saturation.

There are two factors that explain fluorescence saturation. (1) The SPAD in this detection system is a digital counter, which functions by generating a pulse signal upon arrival of an incident photon. This detector has a dead time of 30-40 ns with a maximal

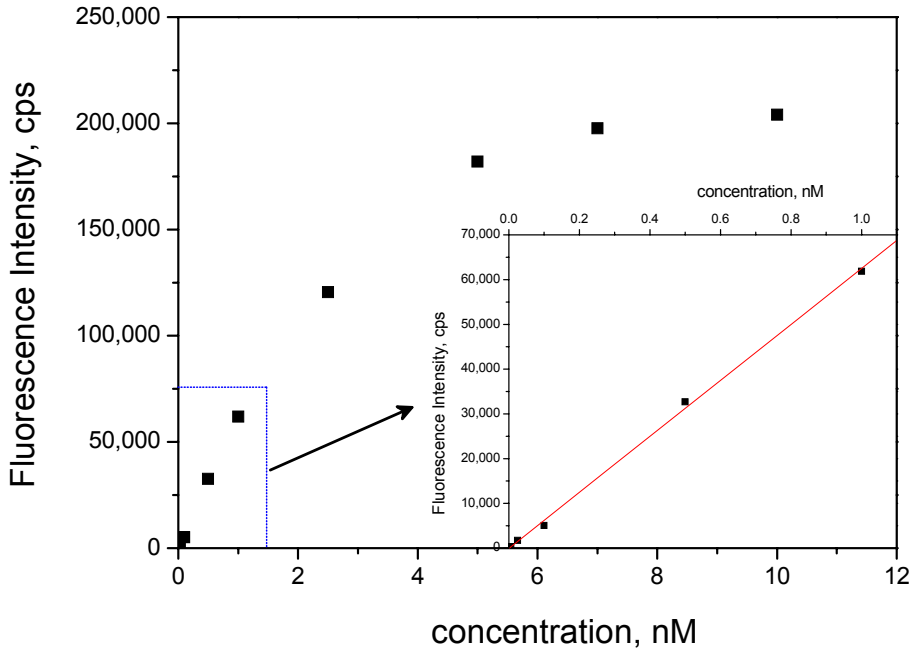


Figure 3.10 Intensity of fluorescence emission as a function of analyte concentration. The concentration of the Cy5.5 labeled oligonucleotides ranges from 1 pM to 10 nM. The output of the excitation laser power is 1 mW. The inset shows an expanded view of the measurement for sample concentration ranging from 1 pM to 1 nM.

counting rate of a few MHz. If the difference of arrival times between the leading photon and subsequent photons is shorter than the dead time of the SPAD, the detector cannot respond quick enough to trigger a cascade of discrete pulses corresponding to each arriving photon. The pulses pile up on the detector and only one pulse signal can be transmitted into the DAQ board for counting. Pulse pileup happens when the sample solution becomes highly concentrated because the photon flux will exceed the recovery time of the avalanche process. (2) Fluorescence saturation caused by an inner filter effect.[13] The influence of inner filter effect on fluorescence measurement of bulk sample is twofold: (i) The intensity of excitation light is not constant throughout the sample solution; and (ii) the photons that are emitted by a fluorphore could also be reabsorbed

by another fluorophore so that the emitted photon cannot reach the detector. Fortunately, the single-molecule experiment operates in digital mode and thus, does not suffer from inner filter effects because of the extremely dilute sample solution used.

3.4 Conclusions

In this chapter, a benchtop laser-induced fluorescence single-molecule detection system with confocal setup was established and optimized. A solution of Cy5.5 labeled oligonucleotides in the concentration range of sub pico-molar was hydrodynamically driven through a COC microfluidic device and detection of single DNA molecules has been demonstrated. The appropriate substrate among different polymeric materials for fabrication of the microfluidic device was selected to give the optimal SBR and maximal detection efficiency. This single-molecule detector has a broad dynamic range (up to 1 nM) and down to the "single molecule" level. The comprehensive study on different aspects of the single-molecule detection in this chapter sets up a solid foundation for designing and conducting single-molecule measurements in different bioanalytical applications.

3.5 References

(1) Effenhauser, C. S.; Bruin, G. J. M.; Paulus, A.; Ehrat, M. *Analytical Chemistry* 1997, *69*, 3451-3457.

(2) Fister, J. C.; Jacobson, S. C.; Davis, L. M.; Ramsey, J. M. *Analytical Chemistry* 1998, *70*, 431-437.

(3) Haab, B. B.; Mathies, R. A. *Analytical Chemistry* 1999, *71*, 5137-5145.

(4) Gosch, M.; Blom, H.; Holm, J.; Heino, T.; Rigler, R. *Analytical Chemistry* 2000, *72*, 3260-3265.

(5) Wabuyele, M. B.; Ford, S. M.; Stryjewski, W.; Barrow, J.; Soper, S. A. *Electrophoresis* 2001, *22*, 3939-3948.

(6) Kang, S. H.; Lee, S.; Yeung, E. S. *Analytical Chemistry* 2004, *76*, 4459-4464.

(7) Emory, J. M.; Soper, S. A. *Analytical Chemistry* 2008, *80*, 3897-3903.

(8) Okagbare, P. I.; Soper, S. A. *Analyst* 2009, *134*, 97-106.

(9) Shadpour, H.; Musyimi, H.; Chen, J. F.; Soper, S. A. *Journal of Chromatography A* 2006, *1111*, 238-251.

(10) Barnes, M. D.; Whitten, W. B.; Ramsey, J. M. *Analytical Chemistry* 1995, *67*, A418-A423.

(11) Soper, S. A.; Mattingly, Q. L.; Vegunta, P. *Analytical Chemistry* 1993, *65*, 740-747.

(12) Davis, L. M.; Williams, P. E.; Ball, D. A.; Swift, K. M.; Matayoshi, E. D. *Curr Pharm Biotechnol* 2003, *4*, 451-462.

(13) Lakowicz, J. R. *Principles of fluorescence spectroscopy*, 3rd ed.; Springer: New York, 2006.

CHAPTER 4 LDR GENERATION OF REVERSE MOLECULAR BEACONS FOR NEAR REAL-TIME ANALYSIS OF BACTERIAL PATHOGENS USING spFRET AND COC MICROFLUIDIC CHIP *

4.1 Introduction

Each year in the United States, the number of illnesses associated with bacterial contamination of consumer food products is estimated to be as high as 5 million cases causing more than 4,500 deaths according to the USDA.[1] Therefore, the rapid and highly specific identification of potential pathogenic contaminations are significant for maintaining public health and safety by minimizing the spread of the contamination.[2-5]

Conventional culture-based methods for pathogen detection are labor-intensive and time-consuming with a minimum of 2 days required for identification of the suspect bacteria and also, interpretation is prone to human error.[6] Immunoassays are an attractive alternative due to the highly selective antigen-antibody interactions they afford. In addition, immunoassays can be applied to complex biological matrices with little sample preparation,[7, 8] as well as the ability to perform parallel analyses.[9, 10] However, some target bacteria cannot be easily identified via an immunoassay due to difficulties associated with finding appropriate monoclonal antibodies to impart the necessary specificity for particular strains.[11]

Genome-specific identification utilizes unique reporter sequences within the genome of the target pathogen. In most cases, a polymerase chain reaction (PCR) is used to generate sufficient copies of the target sequence to aid in detection. The PCR amplicons can be subjected to an electrophoretic separation[12] or liquid chromatography[13] for identifying the bacterium. Due to the high sensitivity and

* The work reported in this chapter has been published in the Journal of Analytical Chemistry and has been reprinted with permission.

specificity of PCR, it has been widely used to detect trace amounts of microorganisms in many scenarios, such as monitoring water quality,[14] food contamination[4] and infectious biological agents.[5] By incorporating different primer pairs within the same reaction chamber, PCR has multiplexing capabilities as well to allow for the analysis of as many as 10 different pathogens.[13, 15, 16] Mathies's group recently reported a capillary electrophoresis-based microanalysis system to perform cell pre-concentration, purification and PCR for pathogen analysis.[17] This system could detect *E. coli* O157:H7 with a detection limit of 0.2 cfu/μL in a processing time >73 min.

Optical biosensors have also been reported for pathogen analysis using different transduction modalities.[18-25] Rider *et al.*[26] showed the recognition of biopathogenic species using a B lymphocyte cell line engineered to expresses both a bioluminescent protein and pathogen-specific membrane-bound antibodies. Low levels of certain pathogens were detected through binding to the antibodies of the B cells and thus, triggering the bioluminescent protein to emit light; 50 cfu of *Yersinia pestis* could be detected with a total processing time of ~3 min.

While the aforementioned techniques can be viewed as effective tools for monitoring the presence of certain bacterial species from a number of different sample inputs, they do possess limitations. For example, PCR-based schemes need several hours to obtain the required results. Even real-time PCR has turnaround times of 20-30 min due to the high number of thermal cycles employed, especially for cases where the bacterial copy number found in the sample can be low. In addition, PCR techniques are limited in terms of their specificity because different strains may possess single base variations in their genetic sequence, which is difficult to register via PCR. For example, *B. anthracis*

has been discovered to possess ~3,500 different single nucleotide variations among its eight strains[27] and application of PCR-based biosensing becomes impractical to identify the specific strain in a timely manner.

In contrast to PCR, the ligase detection reaction (LDR) can offer superior sequence specificity even for single base variations.[28] In the LDR assay, a successful ligation event between two designed primers (common and discriminating primer) can occur only if they are completely complementary to the target DNA, especially at the 3' end of the discriminating primer.[29] The reaction can distinguish specific sequence variations even in the presence of a majority of DNA that does not possess the variation.

In this study, single-pair Fluorescence Resonance Energy Transfer (spFRET) scheme was coupled to an LDR to provide near real-time readout of different bacterial pathogenic species with high strain-specificity. LDR thermal cycling and single-molecule readout were carried out directly on a thermoplastic microfluidic device to provide rapid assay results. The assay strategy (LDR-spFRET) for strain-specific identification of bacterial species is illustrated in Figure 4.1. A similar assay was adopted by Wabuyele *et al.* for detecting single nucleotide mutations in *KRAS* genes.[30] In this assay, a discriminating primer and a common primer were designed based on the sequence of a reporter region within the genome of the bacterial target. These primers also contained a 10-base arm with sequences that were complementary to each other and were covalently attached to a donor and acceptor fluorophore. Successful ligation of the primers will occur only if the complementary sequence is contained within the target generating a reverse molecular beacon (rMB), bringing the donor and acceptor dyes into close proximity producing a FRET response. Because the arm sequences of the

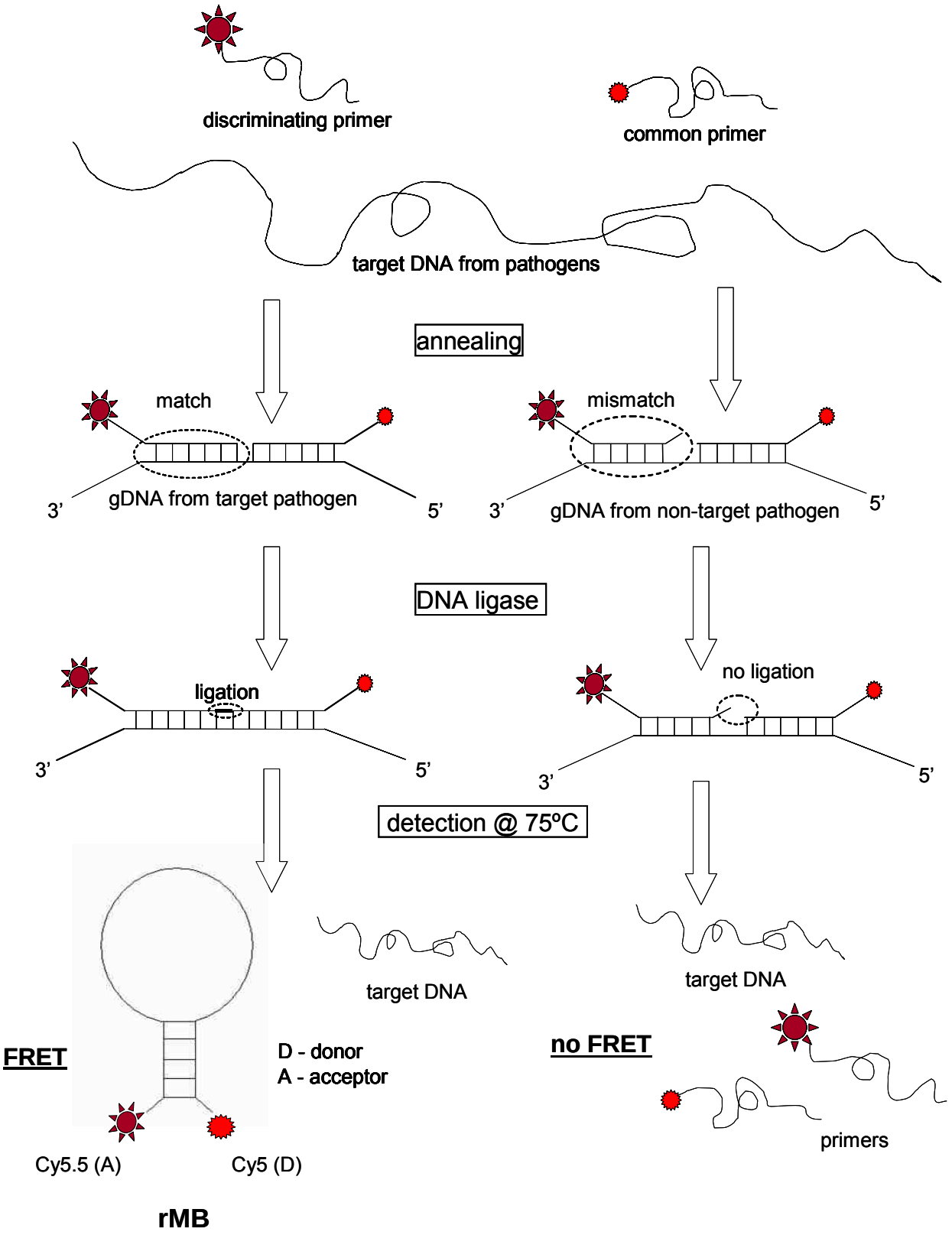


Figure 4.1 Illustration of the LDR-spFRET assay

rMB were designed to possess a higher melting temperature (T_m) and thus, thermodynamically more stable than the target-oligonucleotide duplex, the rMB reorganizes itself into a stable stem-loop conformation following ligation. The LDR was carried out directly on genomic DNA isolated from lysed bacterial cells with no PCR required and direct quantification accomplished using single-molecule counting. The lack of a primary PCR step significantly reduced the assay turnaround time, providing near real-time readout. In contrast to that previously reported using this assay for detecting single nucleotide variations in *K-ras* genes,[30] the genome of bacteria usually possesses sporadic sequence variations at multiple sites and not at a single location. Therefore, the present format of this LDR-spFRET assay did not depend on a single point mutation at the 3'-end of the discriminating, but multiple mismatches between the discriminating and common primers used for the LDR and the target genomic DNA isolated from the bacterial species. This should result in improved specificity due to differences in the T_m of the matched and mismatched duplexes (LDR primers and target genomic DNA) in addition to a possible mismatch at the ligation site.

Two Gram-positive (Gram(+)) pathogens, *S. aureus subsp. aureus*, *S. epidermidis RP-62A*, and one Gram-negative bacterium (Gram(-)), *E. coli K-12*, were employed in this work as models. *S. aureus* is an aggressive pathogen responsible for a range of acute and pyogenic infections and *S. epidermidis* is primarily associated with infections produced from such devices as implanted prosthetic joints or heart valves.[31, 32] *E. coli K-12* is a strain of bacteria with little harm to humans, but members of its family can be very harmful, such as *E. coli* O157:H7.

4.2 Experimental Section

4.2.1 Materials and Reagents

The oligonucleotide primers required for the LDR were synthesized by Integrated DNA Technologies (IDT, Coralville, IA), purified by RP-HPLC and suspended in 1X TE buffer. *Thermus aquaticus* (Taq) DNA ligase was purchased from New England Biolabs (Beverly, MA). AmpliTaq Gold polymerase was purchased from Applied Biosystems (Foster City, CA). Cyclic olefin copolymer, COC, was purchased from Topas Advanced Polymers (Florence, Kentucky).

4.2.2 Bacterial Samples

Three bacterial strains, *Staphylococcus aureus subsp. aureus* (ATCC 700699), *Staphylococcus epidermidis* (ATCC 35984) and *Escherichia Coli K-12* (ATCC 700926) served as models for this study and were acquired from ATCC (Manassas, VA). A series of dilutions to produce the desired cell densities were made in 1X TE buffer consisting of each bacterial genomic DNA. The concentration of the DNA was examined with a UV/vis spectrophotometer (Ultrospec 4000, Amersham Bioscience) using the 260 nm/280 nm absorption ratio. The bacterial genomic DNA was stored at -20ºC until used.

4.2.3 PCR and LDR

The PCR contained 1X PCR buffer II (Applied Biosystems), 2.5 mM $MgCl_2$, 200 μM dNTPs, 1 μM forward and reverse primers, 1.25 units of DNA polymerase (AmpliTaq Gold Polymerase, Applied Biosystems), ~10 ng of template DNA and enough nuclease-free H_2O to make a total reaction volume of 50 μL. The PCR was run in a thermal cycling machine (Eppendorf MasterCycler, Hamburg, Germany) with polymerase added under hot start conditions. The reaction cocktail was subjected to 35 thermal cycles at

94ºC for 15 s, 60ºC for 1 min, 72ºC for 1 min and a final extension at 72ºC for 7 min. The temperature was held at 99ºC to deactivate the polymerase enzyme prior to LDR.

Bench-top LDRs were carried out in 0.2 mL polypropylene microtubes using a bench-top thermal cycler (Eppendorf). The reaction cocktail consisted of 2 units/μL of thermostable DNA ligase, 20 mM Tris-HCl (pH 7.6), 25 mM potassium acetate, 10 mM magnesium acetate, 1 mM NAD^+ cofactor, 10 mM dithiothreitol, 0.1% Triton X-100, 10 nM of each LDR primer, the appropriate amount of the DNA target and nuclease-free H_2O to make a total reaction volume of 20 μL. Prior to thermal cycling, the reaction was first heated to 94ºC to denature the DNA followed by addition of the ligase enzyme. The reaction mixture was processed using the appropriate number of thermal cycles (linear amplification of LDR products), each of which was composed of a denaturation step at 94ºC for 30 s and an annealing/ligation step at 65ºC for 2 min. The reaction was stopped by quickly cooling to 4ºC and adding 0.5 μL of 0.5 M EDTA.

On-chip LDRs were similar to those noted for the bench-top reactions except that bovine serum albumin (BSA, 0.5 mg/mL) was included in the reaction mixture to minimize any potential non-specific adsorption artifacts of the ligase enzyme onto the thermal reactor surface.[33, 34] Kapton heaters were attached directly to the bottom of the microchip to provide the required temperatures for on-chip LDR using a continuous flow reactor format (see below). All PCR and LDR products were verified using slab gel and capillary gel electrophoresis. See the Supporting Information for discussion of the experimental conditions and results for the electrophoresis.

4.2.4 Primer and rMB Design for the LDRs

In designing the oligonucleotide primers for the LDR-spFRET assay, the 16S rRNA gene was selected as the biomarker for bacteria identification due to its highly

conserved sequence.[6, 32, 35] Moreover, the 16S rRNA gene appears at multiple locations within the genome of each bacterial cell increasing its copy number and thus, aiding detection. The sequence of the primers used for PCRs and LDRs are listed in Table 4.1. Two regions within the 16S rRNA gene were interrogated. The PCR primers denoted 'AMP1' and the LDR primers denoted as 'Gram+' were designed for one region of the 16S rRNA gene and used to differentiate Gram(+) from Gram(-) bacteria following LDR. When specific strain within the Gram(+) species needed to be identified, a second region of the 16S rRNA was examined with primer sets 'AMP2'. The LDR primer pair denoted as 'epid' were based on 16S rRNA gene used to identify *S. epidermidis*, and the LDR primer pair denoted as 'aureus' were designed to identify *S. aureus*.

Table 4.1 Oligonucleotide sequences used as the PCR and LDR primers for the strain-specific identification of bacterial species.

	Primers	Sequence (5'-3')*
PCR primers	AMP1 forward	ACTGAGACACGGTCCAGACTCCTAC
	AMP1 reverse	GTAGCGGTGAAATGCGCAGAGATA
	AMP2 forward	CAAACAGGATTAGATACCCTGGTAGTC
	AMP2 reverse	GAAGGTGGGGATGACGTCAAAT
LDR primers	Gram(+) disc	Cy5.5-C3-AGGCGGCGCGAGCGAAAGCCTGACGGAGCA
	Gram(+) com	p[a]ACGCCGCGTGAGTGATGAAGGTACGCGCCGCCT-C3-Cy5
	S. aureus disc	Cy5.5-C3-AGGCGGCGCGTTACCAAATCTTGACATCCTTTGACA
	S. aureus com	pACTCTAGAGATAGAGCCTTCCCCTTCGGCGCGCCGCCT-C3-Cy5
	S. epid disc	Cy5.5-C3-AGGCGGCGCGCGTAAAACTCTGTTATTAGGGAAGAACAA
	S. epid com	pATGTGTAAGTAACTATGCACGTCTTGACGCGCGCCGCCT-C3-Cy5

[a] p, phosphorylation.

* The underlined sequence consists of the stem of the rMB, which is formed following ligation. In all cases, a 3-carbon linker was used to attach the donor or acceptor to the oligonucleotides to maximize energy transfer efficiencies.

The design of the rMB was assisted by the DNA folding program from IDT to assure that the dominant conformation was indeed the closed hairpin form. This folding program uses minimum free energy of formation to predict the T_m of duplexed DNA as well as any possible secondary structure(s). The major input parameters to the analysis program included the sequence of the DNA, the salt concentration and the folding temperature. The T_m of the rMB's stem was computed by this program and possible secondary structures were generated and displayed as well. The loop sequence of the rMB was designed based on the known sequence of Gram(+) bacterial strains or to exactly match the corresponding sequence in *S. aureus* and *S. epidermidis* if strain-specificity was required. The two stem sequences were anchored to each end of the loop sequences and were designed to be complementary to each other, but not to the target. They possessed a high GC content resulting in a relatively high T_m. Therefore, the formation of the stem was thermodynamically favored over the loop-target hybrid at the concentrations employed for the LDR. At particular temperatures, the loop-target duplex and possible secondary structures associated with the rMB could coexist with the typical stem-loop hairpin structure and thus, reduce the observed FRET signal. However, at a single-molecule detection temperature of ~75ºC, the loop-target duplex was predominantly denatured and the possibilities of undesirable secondary structures of the rMB were significantly minimized. Thus, the predominant species was the rMB.

Figure 4.2 shows a secondary structure of a 63 base rMB at 75ºC in a solution buffered with 50 mM NaCl and 10 mM $MgCl_2$, which is obtained from the DNA folding program. The loop sequence (43 nt) contained the reporter sequence of the particular bacteria under investigation while the stem sequence was similar for all of the rMBs

used in these investigations. The loop sequence was selected to minimize any bulges formed due to secondary structure. The T_m of the resulting stem structure for this rMB was calculated to be 83.4ºC. This figure indicates that the rMB was the most thermodynamically favored conformation when folding was performed at 75°C using the primer pairs selected for spFRET.

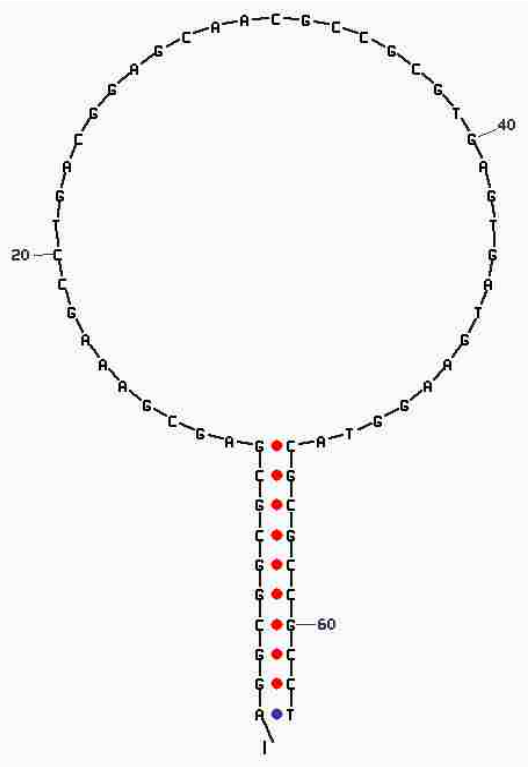


Figure 4.2 Secondary structure of a 63-base rMB folded at 75ºC.

For differentiation between Gram(+) and Gram(-) bacteria, the discriminating primer was composed of a 20 base loop sequence and a 10 base stem sequence with the 5'-end of the stem labeled with Cy5.5. The 33 base common primer, which contained a 23 base loop sequence and a 10 base stem sequence, was phosphorylated at its 5'- end and Cy5-labeled at its 3'-end. In the presence of target DNA, the common primer and the discriminating primer both can hybridize to the target and undergo ligation, but only

if the primers are completely complementary to the target DNA. At the detection temperature employed, the ligated primers formed the rMB that provided a spectroscopic signature of the presence of the target through spFRET.

4.2.5 Analysis of PCR and LDR Products

Five μL of each PCR product was pre-mixed with 2 μL of a loading dye, 5 μL of nuclease-free H_2O and loaded into a 3% ReadyAgarose Mini Gel (Bio-RAD, Hercules, CA). The electrophoresis was run in 1X TAE buffer under an electric field strength of 8 V/cm for 50 min. This was followed by staining the gel with ethidium bromide for 20 min and subsequently soaking in clean H_2O for another 20 min to remove excess staining dye. The gel was then sent to a gel imaging system (Gel Logic 200, Carestream Molecular Imaging, New Haven, CT) to confirm the presence of PCR products.

To examine the fidelity and yield of the designed LDRs, these products were analyzed via capillary gel electrophoresis with laser-induced fluorescence detection (Beckman Coulter CEQ8000, Fullerton, CA). Five μL of each LDR product was loaded into different wells in a 96-well titer plate and mixed with 35 μL of loading buffer. DNA size standards were also added into a separate well. A drop of mineral oil was applied into each sample well to prevent evaporation during the thermal denaturation step. The separation buffer was added into wells of another 96-well titer plate. The LDR products were first denatured at 94ºC for 2 min and then injected into the capillary gel column using a voltage of 2 kV for 15 s. The separation was carried out using 6 kV for 30 min. The sieving gel contained denaturing additives and the capillary was maintained at a temperature of 60ºC to ensure that the hairpin structure of the LDR product was fully opened during the electrophoretic separation.

4.2.6 Measurement of Energy Transfer

To evaluate the design of the rMBs, the LDR products were examined in a fluorometer (FluoroLog 3, HORIBA Jobin Yvon, Edison, NJ) to determine the extent and efficiency of energy transfer. A Peltier heating unit was equipped with the fluorometer to allow spectral interrogation at 75ºC, which was below the T_m of the stem associated with the rMB (83.4ºC, see Supporting Information), but above the T_m of the unligated or ligated primers hybridizing to their complementary sequences in the target DNA at the concentrations employed for the LDR/spFRET assays. The sample was excited at 635 nm and the emission collected between 650 – 750 nm.

4.2.7 Fabrication of the COC Microfluidic Device

The microfluidic chip was fabricated using procedures developed in our laboratories,[36, 37] which involved patterning microstructures on a brass molding tool through a micro-milling process followed by transferring the microstructures into a COC substrate (Topas Advanced Polymer, Florence, KY), which has a glass transition temperature (Tg) of 130ºC, using hot embossing and a brass molding tool. COC was selected as the substrate due to its high Tg to minimize any thermally-induced structure deformation at the temperatures employed for the LDR and also, its excellent optical properties.[38, 39] The layout of the microchip architecture is shown in Figure 4.3. Following embossing, the COC substrate was rinsed and sonicated in isopropanol for 10 min and then in ddH_2O for 20 min. A COC cover plate was thermally fusion bonded to the embossed microchip by sandwiching the cover plate/substrate assembly between two glass plates in a convection oven and maintaining the temperature at 134ºC for 20 min. After cooling, the assembled microchip was rinsed thoroughly with ddH_2O.

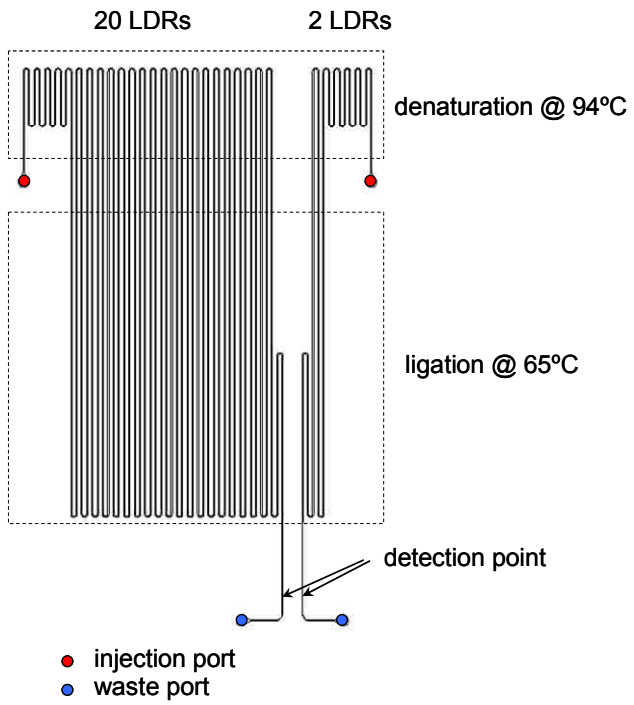


Figure 4.3 Schematic representation of the microfluidic device hot embossed from a COC substrate

The microchip was composed of two different devices, both of which performed LDRs in a continuous flow format.[37] One device contained a continuous flow thermal reactor to carry out 20 LDR thermal cycles (1,500 mm total length) while the other possessed only 2 LDR thermal cycles (total length = 204 mm). The channels used for the thermal cycling in both devices were 100 μm in width and 100 μm in depth. At the end of each continuous flow thermal cycler device was positioned a detection region, which allowed for on-chip single-molecule observation of spFRET signals generated from the rMBs.

4.2.8 Operation of the Microchip

A fused silica capillary (365 μm O.D.; 100 μm O.D., Polymicro Technologies, Phoenix, AZ) was inserted into the microchip and connected to a glass syringe (SGE, Australia) via a syringe-to-capillary adapter (InnovaQuartz, Phoenix, AZ). An LDR

cocktail was first run through a 0.2 μm filter to remove any large particulates. Then, the reaction mixture was loaded into a glass syringe and driven by a syringe pump (Pico Plus, Harvard Apparatus, Holliston, MA) through the microchannels of the chip. Thin-film Kapton heaters were placed at the appropriate positions on the chip to provide the desired temperatures for the LDR and spFRET detection. The denaturation and renaturation/ligation zones are shown in Figure 4.3. A 3.5 mm gap was situated between the denaturation and renaturation/ligation zones to minimize any thermal crosstalk. The volume flow rate was set between 0.1- 4 μL/min depending on the required ligation time and also, optimizing the signal-to-noise ratio for the single-molecule measurement.

4.2.9 Instrumentation for LIF Single-molecule Detection

LDR rMB products were measured using a laser-induced fluorescence (LIF) system. Briefly, a pre-collimated 635 nm diode laser (Model CPS196, Thorlabs, Newton, NJ) was used for excitation. The laser beam was conditioned using a laser line filter (640DF20, Omega, Brattleboro, VT) and directed by a dichroic mirror (690DRLP, Omega Optical) into a microscope objective (M-60X, NA = 0.85, Newport, Irvine, CA). The emitted fluorescence was then collected by the same objective and transmitted through a dichroic and pinhole (i.d. = 100 μm). The fluorescence light was spectrally filtered using interference filters, including a longpass filter (3RD690LP, Omega Optical) and a bandpass filter (HQ710/20m, Chroma Technology, Rockingham, VT). Finally, the fluorescence was focused onto the active area of a single photon avalanche diode (SPCM-200, EG&G, Vandreuil, Canada) using a 10X microscope objective. The signal

from the SPAD was transformed by a pulse converter into a TTL pulse and processed using a digital counting board (PCI-6602, National Instruments, Austin, TX).

Single-molecule fluorescence measurements were performed by measuring the time-of-arrival of single photon events. To accomplish this, a PCI-6602 data acquisition board provided 12.5 ns time resolution at its maximum clock frequency of 80 MHz. Each photon event was stamped with a time, which allowed binning the individual photon events into a user-defined bin.

4.3 Results and Discussions

4.3.1 Generating LDR Targets Using PCR Amplified gDNA for Assay Validation

To initially examine the design of the LDR primers and the rMB structure as well as the efficiency of the LDR, PCR was carried out on genomic DNA isolated from *S. aureus*, *S. epidermidis* and *E. coli* to provide sufficient targets so that the products could be examined by gel electrophoresis or a conventional fluorometer. However, the PCR step was eliminated for the on-chip single-molecule detection experiments, which used genomic DNA as the LDR targets.

After 35 PCR cycles, the products were validated by running agarose gel electrophoresis. Two different PCR amplicons were evaluated, which were defined by the primer sets ‘AMP1’ and ‘AMP2’. The primer set AMP1 was designed to produce a 388 bp amplicon from the 16S rRNA gene for each species, while primer set AMP2 was designed to produce a 423 bp amplicon from a different region of the 16S rRNA gene. Ten ng of genomic DNA from each bacterial species were used as the targets for the amplification.

Figure 4.4 shows the 388 bp amplicons (lanes B, C, D) and the 423 bp amplicons (lanes E, F, G) were clearly seen when primer set AMP1 or AMP2 were used, respectively. From this data, successful amplification of the target sequences was secured using all bacterial species. These regions of the genome were then interrogated via LDR to provide information on the Gram(+) or Gram(-) status of the target and/or determining its strain. A multiplexed PCR was also performed where both primer sets were added into the same reaction cocktail. Lanes H, I, J in Figure 4.4 shows the presence of two amplicons with different lengths for *S. aureus*, *S. epidermidis* and *E. coli*.

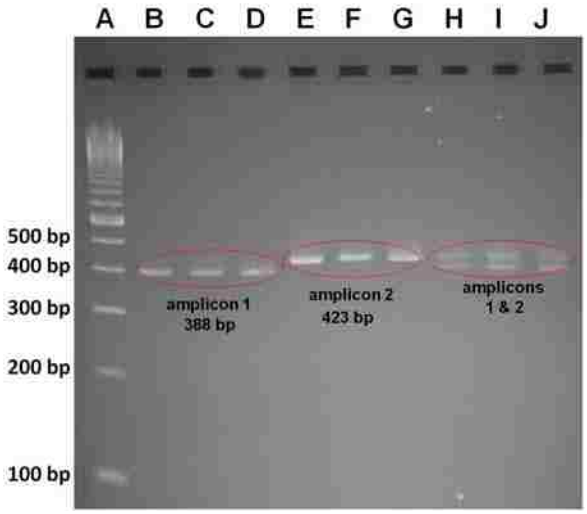


Figure 4.4 Gel electrophoresis images of dsDNA PCR products along with a sizing ladder. Lanes A = sizing ladder; B = *S. aureus* (AMP1); C = *S. epidermidis* (AMP1); D = *E. coli* (AMP1); E = *S. aureus* (AMP2); F = *S. epidermidis* (AMP2); G = *E. coli* (AMP2); H = *S. aureus* (AMP1 + AMP2); I = *S. epidermidis* (AMP1 + AMP2); J = *E. coli* (AMP1 + AMP2).

4.3.2 LDR for Determining Gram (+/-) or Bacterial Strain

To initially evaluate the efficiency of the spFRET assay using LDR from genomic bacterial DNA, we carefully controlled the copy number of DNA targets included in the

assay and used PCR amplicons as the input, which was generated from all bacterial species and the subsequent LDRs were carried out using these PCR amplicons.

PCR products from *S. aureus*, *S. epidermidis* and *E. coli* were quantified by UV absorption at 260 nm and diluted to the required concentration, which were then used as targets for LDR. In these experiments, the LDR was run using 20 thermal cycles and the LDR products were subjected to capillary gel electrophoresis to validate the appropriate LDR products were indeed generated.

Figure 4.5(A) shows a capillary gel electrophoretic analysis of LDR products generated following 20 thermal cycles with 10 nM of each primer and 1 nM of the PCR product secured from *S. aureus*. In the electropherogram, the peaks that eluted early in the electropherogram resulted from unligated primers. Size standards, which consisted of equal amounts of a 13 base and 88 base single-stranded oligonucleotides were co-electrophoresed with the LDR products and used for the size determination. From this figure, it can be seen that the two primers were successfully ligated to form a 63 base oligonucleotide, the designed length of the rMB. The magnitude of the peaks associated with the shorter (unligated primers) and longer (ligated primers) fragments also provided information as to the relative amounts of these two fragments from which the efficiency of the ligation reaction could be evaluated. In this example, the ratio of the peak height of the LDR products to the peak height of the primers was 3.96, indicating a concentration of ~8 nM for the LDR products and 2 nM for the unligated primers. Considering that 100% LDR efficiency for each thermal cycle would result in 1 nM of product, the average reaction efficiency was estimated to be 40% for each LDR cycle.

Figure 4.5(B) shows the results of a 20 cycle LDR run on PCR amplicons using primers based on the 16S rRNA gene geared for the identification of Gram (+) bacteria,

S. aureus and *S. epidermidis*. LDR products were successfully detected for both, but no LDR products were found when *E. coli* was used as the target, consistent with the Gram (+) and Gram (-) nature of these species.

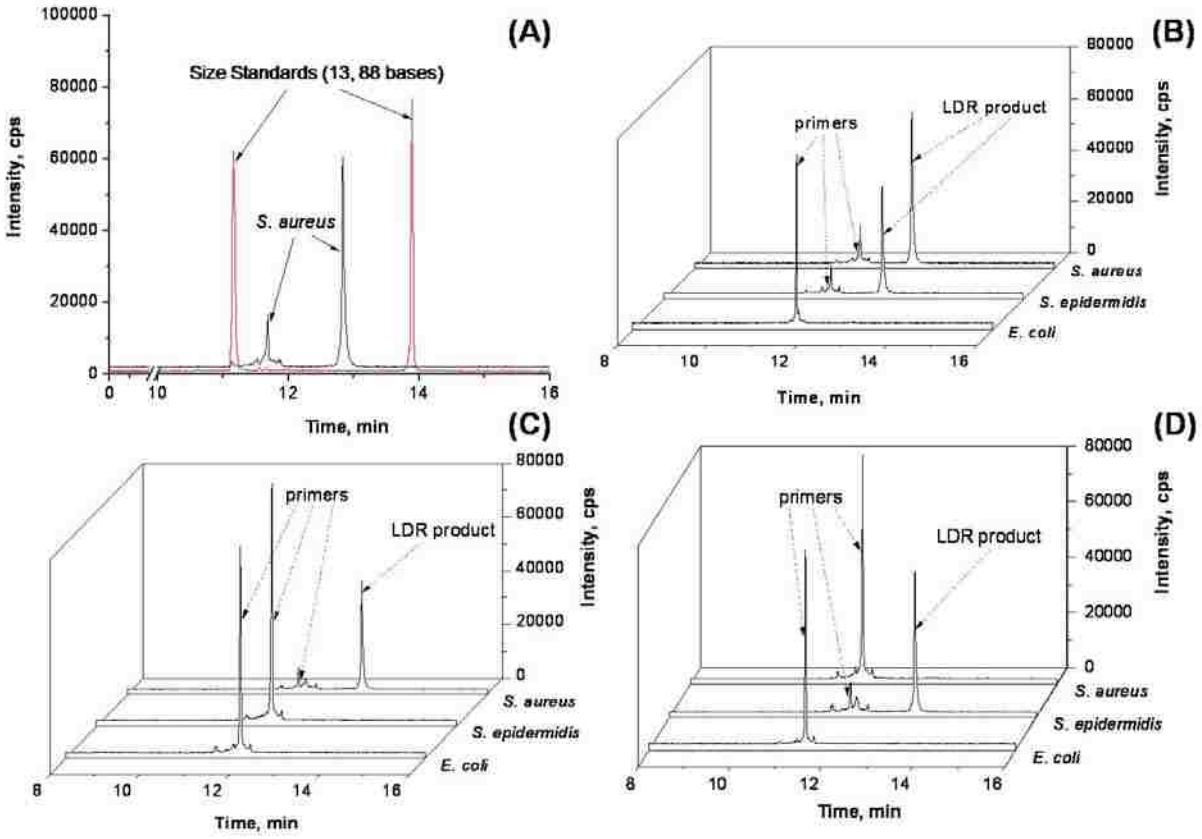


Figure 4.5 Electropherograms of LDR products from different bacterial samples. (A) LDR products from *S. aureus* amplicons (black) and the size standards (red). The size standards contained roughly equal amounts of a 13 nt and 88 nt single-stranded DNA fragment. (B) LDR products generated using the primer set specific for Gram(+) bacteria to allow differentiation between Gram(+) and Gram(-) strains. (C) LDR products generated using *S. aureus* specific LDR primers. (D) LDR products using *S. epidermidis* specific LDR primers. For all electropherograms, the LDR was run for 20 cycles at 94ºC for 30 s and 65ºC for 2 min using a bench-top thermal cycler. The LDR cocktail contained 10 nM of each primer and 1 nM of the PCR product from different bacterial samples. LDR products were analyzed via capillary gel electrophoresis equipped with laser-induced fluorescence detection. The sample was denatured at 94ºC for 2 min and then injected into the capillary using a voltage of 2 kV for 15 s. The separation was carried out using a voltage of 6 kV for 30 min.

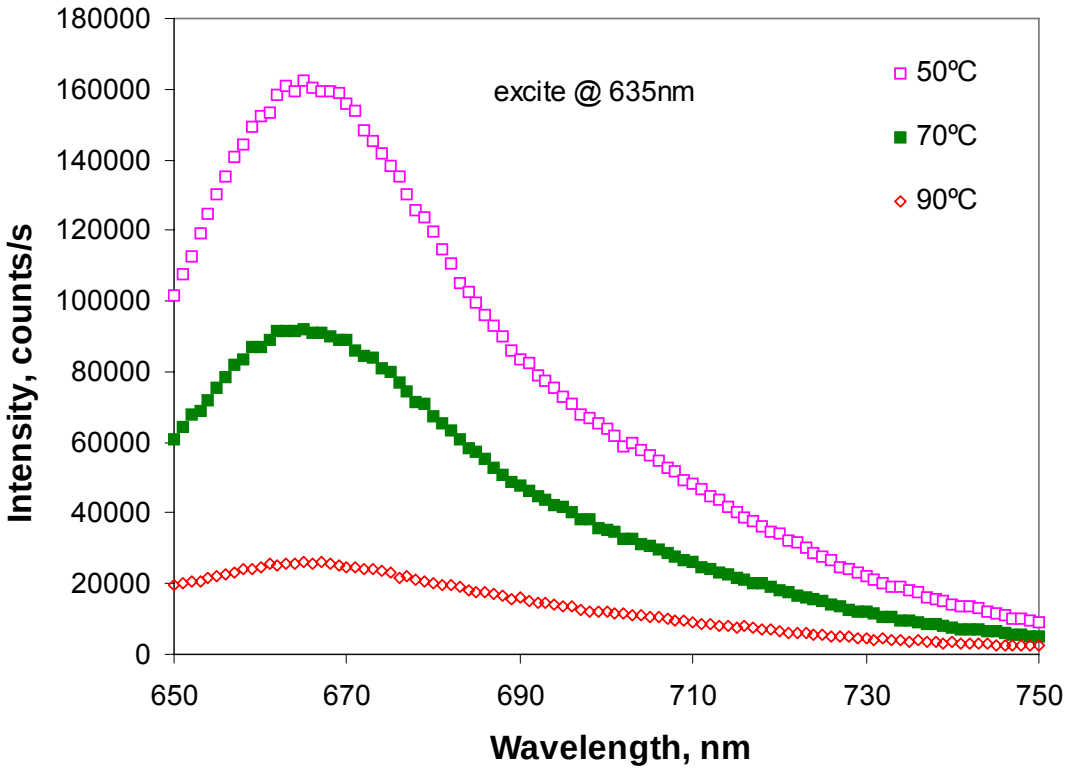
To distinguish between *S. aureus* and *S. epidermidis*, a strain-specific primer pair was used for the LDR that interrogated the section of the genome defined by PCR primer set AMP2. Figures 4.5(C) and 4.5(D) show the results of a 20 cycle LDR when the strain-specific primers were used for the LDR. As can be seen in Figure 4.5(C), when the primer pair specific for *S. aureus* was used, an LDR product was found for *S. aureus* only. Conversely, when the primer pair specific for *S. epidermidis* was used, an LDR product was found only from *S. epidermidis*. No LDR product was found for *E. coli* in both cases. This indicated that *S. aureus* and *S. epidermidis* could be differentiated using LDR and strain-specific primers.

4.3.3 Fluorescence Resonance Energy Transfer (FRET)

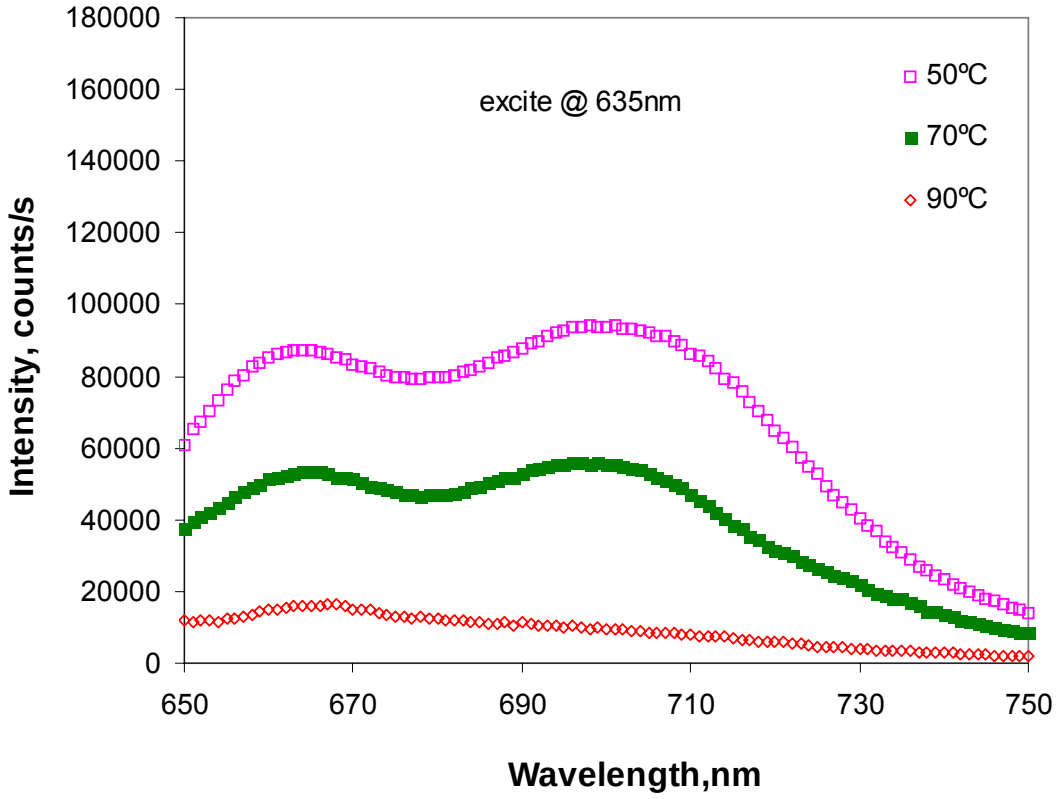
To determine the optimal configuration of the rMB to maximize the FRET efficiency, LDR products were evaluated using bulk fluorescence measurements. In one LDR cocktail, 10 nM of the Cy5-labeled common primer and 10 nM of the Cy5.5-labeled discriminating primer were mixed with 100 nM of DNA target and subjected to 20 LDR thermal cycles. The high concentration of the target as well as the utilization of 20 thermal cycles made the rMB the dominant component in the reaction with most of the primers consumed following thermal cycling. In another LDR cocktail, the PCR template was not introduced so that only the primers were present following thermal cycling.

Figure 4.6(A) shows the emission spectra of the primer mixture without target. These spectra consisted of two features, one with an emission maximum around 664 nm corresponding to the Cy5-labeled primer and another feature with an emission maximum at ~700 nm associated with the Cy5.5-labeled primer, which appears as a weak shoulder on the major band seen at 664 nm. The fluorescence in this case

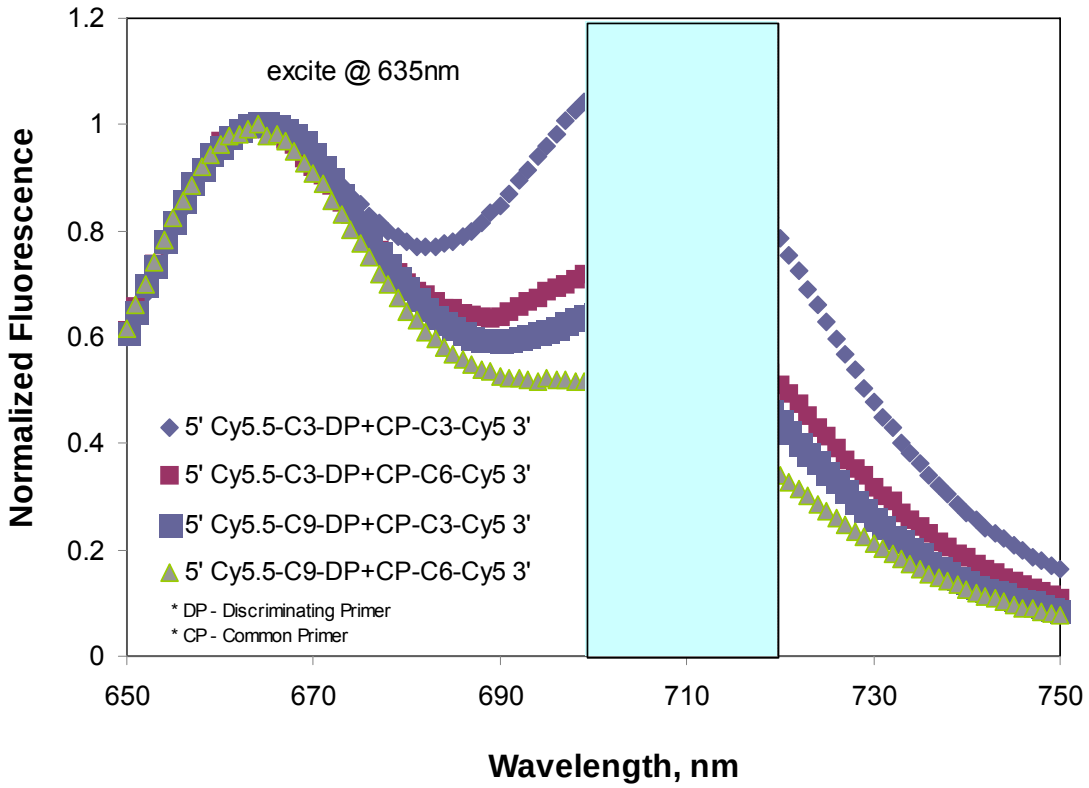
Figure 4.6 Measurements showing the emission spectra of the donor/acceptor dye pair of the rMB for FRET verification. (A) Emission spectra of a mixture of dye-labeled primers (10 nM each) after 20 LDR thermal cycles in the absence of the DNA target. (B) Emission spectra of LDR dye-labeled primers (10 nM of each primer) and 1 nM of the target DNA following 20-LDR thermal cycles. (C) Emission spectra of LDR products after 20-LDR thermal cycles using primers with different linkers used to attach the donor/acceptor dyes to the primer. The shaded area in (C) indicates the detection window used in the LIF detector for monitoring acceptor fluorescence.


180000
160000
140000
120000
100000
80000
60000
40000
20000
0
Intensity, counts/s
excite @ 635nm
50ºC
70ºC
90ºC
650
670
690
710
730
750
Wavelength, nm

(A)


180000
160000
140000
120000
100000
80000
60000
40000
20000
0
Intensity, counts/s
excite @ 635nm
50ºC
70ºC
90ºC
650
670
690
710
730
750
Wavelength,nm

(B)



(C)

originates primarily from direct excitation of the dye-labeled primers with the emission of Cy5 much higher than that of Cy5.5 because the excitation was set at Cy5's absorption maximum. When the temperature was increased, the fluorescence of both dyes was found to decrease due to the temperature dependent fluorescence quantum yields associated with many carbocyanine dyes.[40]

Figure 4.6(B) shows the emission spectra of the LDR cocktail containing the PCR template, which possessed a sequence complementary to the two primers. Compared to Figure 4.6(A), the donor fluorescence was significantly reduced while the acceptor fluorescence increased. This indicated energy transfer between the donor and acceptor due to formation of the rMB structure.

The efficiency of energy transfer was determined to be 42% at 70°C according to the equation: $E_{FRET} = 1 - I_{da} / I_d$,[41] where I_{da} is the fluorescence of donor in the presence of acceptor, and I_d is the fluorescence of donor when acceptor was absent. When the temperature was increased from 50ºC to 70ºC, the fluorescence from both dyes decreased similar to what was seen in Figure 4.6(A). However, at a temperature of 90ºC, the acceptor fluorescence was completely absent in the emission spectrum. This was due to the temperature being set above the T_m of the stem structure of the rMB (83.4ºC), melting the duplexed DNA and thus, spatially separating the donor from the acceptor dye significantly reducing the energy transfer efficiency.

4.3.4 Effects of Linker Structure on FRET Efficiency

In developing the spFRET assay, we also investigated different linker structures to assess their impact on the efficiency of energy transfer between the donor and acceptor. The different linker structures evaluated are shown in Table 4.2. The cyanine dye molecules, Cy5 and Cy5.5, were attached to either the 5' or 3' terminus of the oligonucleotides used for the LDR employing either a 3-carbon, 6-carbon, or 9-carbon linker structure, which provided different effective lengths between the donor/acceptor dyes based on molecular structure considerations.

Table 4.2 Different linker structures used for dye-labeling of the LDR primers.

Primers	Size	Sequence
discriminating	30	5' Cy5.5-C3-AGGCGGCGCGAGCGAAAGCCTGACGGAGCA 3'
discriminating	30	5' Cy5.5-C9-AGGCGGCGCGAGCGAAAGCCTGACGGAGCA 3'
common	33	5' pACGCCGCGTGAGTGATGAAGGTACGCGCCGCCT-C3-Cy5 3'
common	33	5' pACGCCGCGTGAGTGATGAAGGTACGCGCCGCCT-C6-Cy5 3'

Figure 4.6(C) shows a comparison of the FRET emission from LDR products formed using primers with different linker structures (carbon spacer was either 3, 6 or 9 units). The shaded area indicates the detection window for registering acceptor fluorescence that was defined by the bandpass filter used for the LIF detector. The results indicated that using more than 3 carbon spacers for the linkers for both the discriminating and common primers produced lower acceptor fluorescence, suggesting reduced energy transfer efficiency. In FRET, the energy transfer efficiency, E_{FRET}, depends on the time-averaged spatial separation (R) between the donor and acceptor according to the Förster theory, given by $E_{FRET} = \frac{1}{1+(R/R_O)^6}$, where R_O is the Förster distance. The additional carbons used in the linkers resulted in larger time-averaged distance between the donor/acceptor pair compared to shorter carbon linkers;[42] it has been reported that the projection fluctuation of a 5-carbon linker is 0.9 nm versus 0.25 nm for a 2-carbon linker.[43] However, when the donor and acceptor are brought extremely close, they may also undergo static quenching.[44] Therefore, we engineered the LDR primers used to form the rMBs to possess a 3-carbon spacer on both ends to mitigate any possible quenching. Based on the data shown in Figure 4.6(C), the optimal linker structure of those evaluated consisted of a 3 carbon spacer used for both the donor and acceptor attachment to the 5' and 3' ends of the stem structures.

4.3.5 Selection of Appropriate Polymer Substrate for the Microfluidic Chip

When performing the thermally cycled LDR on a thermoplastic chip and detect single molecule events on the same chip, we were also interested in deciding which polymer would be optimal for the fluidic chip. The requirements for this selection included; (i) the material must be easily molded into the desired structures using hot

embossing; (ii) chemically and thermally stable during the thermal cycling conditions required for the reaction;; and (iii) demonstrate low levels of autofluorescence to produce high signal-to-noise ratios for the single-molecule detection.. In single-molecule measurements, COC is very attractive because it has excellent optical properties and extremely low fluorescence in the near-IR region.[45] Table 4.3 provides a comparison of autofluorescence levels for several commonly used thermoplastic materials. As can be seen, COC produced the lowest background fluorescence of the three materials investigated. In addition, its high Tg provided good microstructure integrity when heated zones were applied to the chip for the LDR. Therefore, COC was used for all of the reported on-chip LDR and single-molecule measurements.

Table 4.3 Autofluorescence of commonly used thermoplastic materials used for microfluidic applications

Thermoplastic Material	Dimension of Microchannel	Fluorescence Background
Poly(methyl methacrylate) (PMMA)	100 x 100 μm	~ 5,000 counts/s
Polycarbonate (PC)	100 x 100 μm	~ 40,000 counts/s
Cyclic Olefin Copolymer (COC)	100 x 100 μm	~ 1,000 counts/s

*The measurement was carried using the laser-induced fluorescence system described in this manuscript. Each polymer was molded via hot embossing to form microchannels. All three chips were then thermally fusion bonded to a 100 μm thick cover plate made from the same material as the substrate. The fluidic channel was filled with 1X TBE buffer prior to the measurement.

4.3.6 Single-molecule Measurements of LDR-generated rMBs Directly from Genomic DNA

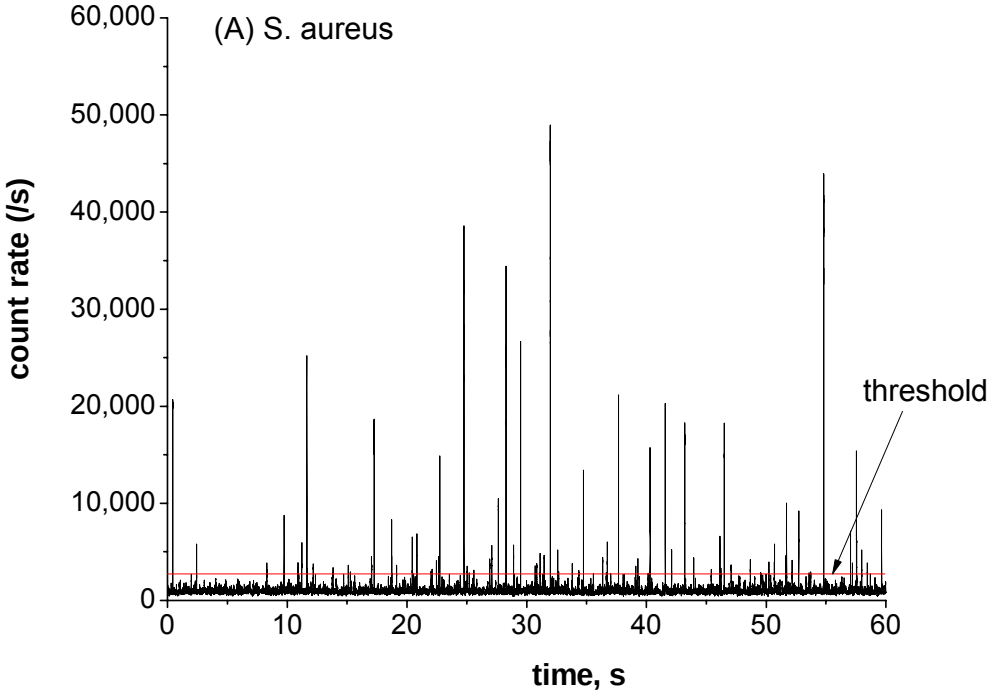
Figure 4.7 shows single-molecule photon bursts from rMBs generated from a 20-cycle LDR performed on a COC microfluidic chip. The reaction cocktail consisted of 10 pM of the common and discriminating primers in the presence of 6 copies/nL of genomic DNA from *S. aureus*, *S. epidermidis* and *E. coli*. The reaction mixture was driven at a volume flow rate of 0.78 μL/min, which was experimentally optimized to

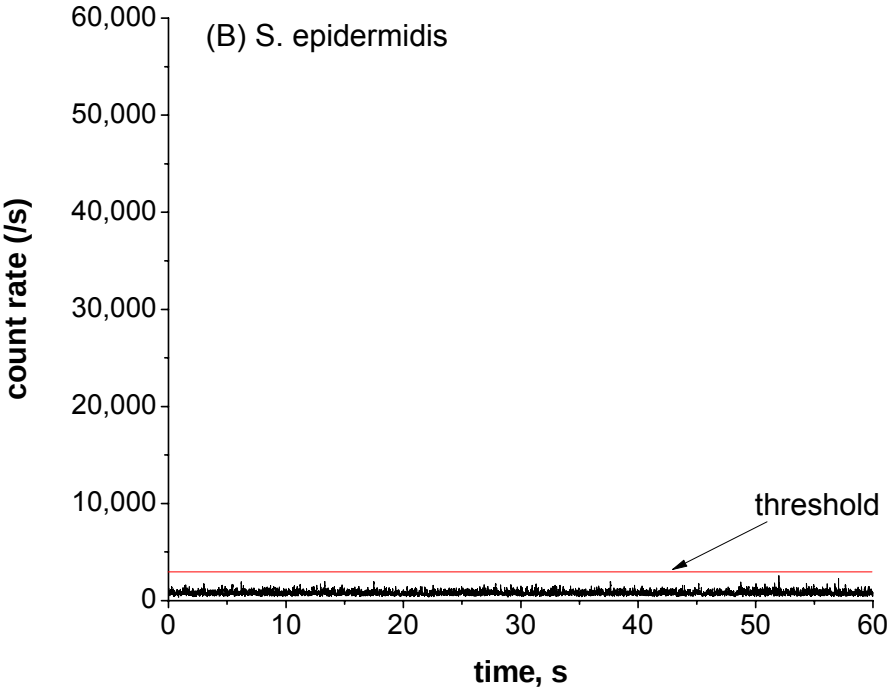
provide high single molecule signal-to-noise ratio and efficient LDR yields.[46] From Figure 4.7, *S. aureus* resulted in clear single molecule signatures in the data trace, but when *S. epidermidis* or *E. coli* were present, no photon bursts from single molecules were detected. Because the LDR cocktail included primers specific for *S. aureus*, these results are consistent with the primers used for the LDR. We note that although the LDR primers contain complementary sequences even in the absence of ligation, the thermodynamic nature of the T_m for inter-strand duplexes (*i.e*., non-ligated primers) at pM concentrations are well below the detection temperature and thus, no spFRET signal would be expected for unligated primers.[30] However, in the case of the ligated primers, which form intra-strand duplexes, the T_m is not concentration dependent.

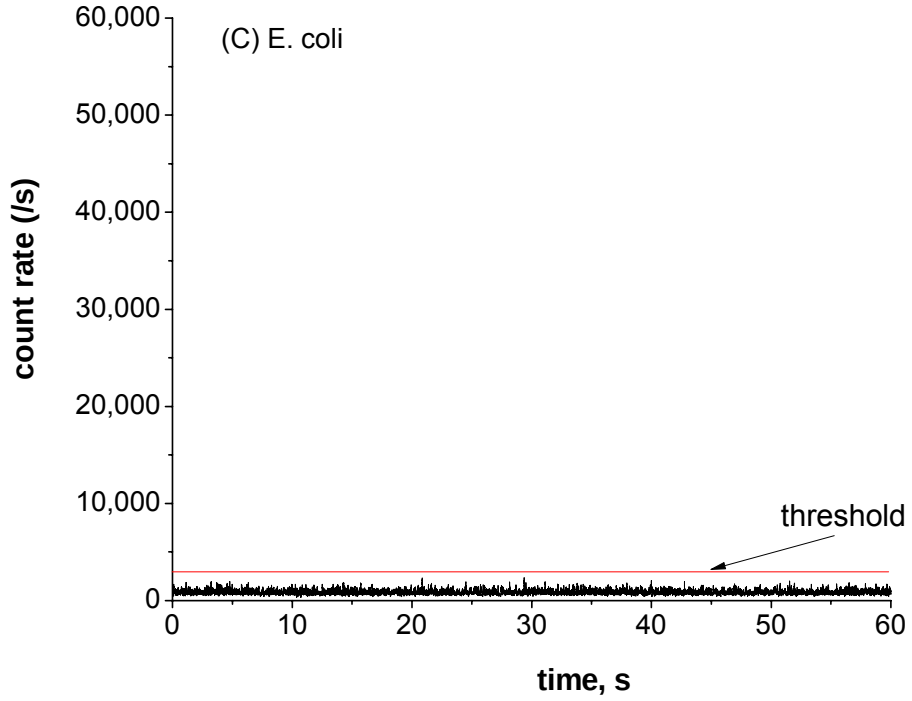
Single-molecule intensity histograms were analyzed for both background and samples containing the target bacterial species to set a threshold level to minimize errors due to false positive signals. Lower threshold levels reduce false negative errors, but typically at the expense of increasing false positive errors. For the single-molecule histograms (data not shown), the average background fluorescence was determined to be ~1,000 counts/s. By setting a threshold level at 3 times this average background level, there were no photon bursts detected from the blank producing a false positive rate of 0 per data stream. At this threshold level, a total of 83 photon burst events were found for the positive control.

For the bacterial genomes interrogated, the 16S rRNA gene appears in five locations. Considering the linear amplification associated with LDR and an average LDR efficiency of 40% per cycle, a 20-cycle LDR would generate 40 rMBs for each copy of gDNA. By driving the reaction mixture at 0.78 μL/min and collecting data for 1 min,

Figure 4.7 Single-molecule photon bursts of LDR-generated rMBs. LDR was performed on the COC microfluidic chip using a continuous flow process. (A) Photon bursts generated from *S. aureus* genomic DNA; (B) photon bursts from *S. epidermidis* genomic DNA; and (C) photon bursts from *E. coli* targets. The reaction cocktail consisted of 10 pM of the common and discriminating primers as well as 6 copies/nL of genomic DNA from *S. aureus*, *S. epidermidis* and *E. coli*. A 20-cycle LDR was run for all three bacteria on the COC microchip. When analyzing the data, a threshold of 3,000 counts/s was used to discriminate the single-molecule events from background fluorescence.


(A) S. aureus
count rate (/s)
60,000
50,000
40,000
30,000
20,000
10,000
0
0
10
20
30
40
50
60
time, s
threshold


(B) S. epidermidis
count rate (/s)
60,000
50,000
40,000
30,000
20,000
10,000
0
0
10
20
30
40
50
60
time, s
threshold



we would expect 187,200 rMBs flowing through the fluidic chip over this 1 min processing time. In the data trace of Figure 4.7(A), 83 photon burst events were observed, which means that 0.04% of the rMBs were analyzed. For the LIF setup used herein, the probe volume can be approximated by a cylinder with a $1/e^2$ beam waist of the elliptical laser measured to be 2 x 4 μm. The height of the cylinder was determined by the size of the pinhole and estimated to be ~2 μm. Thus, the probe volume was estimated to be 12.6 fL, which yielded a single molecule sampling efficiency of 0.06%, which was obtained by dividing the probe volume $(\pi \times 1\mu m \times 2\mu m \times 2\mu m)$ by the cross sectional volume of the fluidic channel $(100\mu m \times 100\mu m \times 2\mu m)$. The experimentally observed sampling efficiency (0.04%) agreed favorably with the percent of single rMBs calculated to travel through the probe volume with respect to the fluidic channel dimensions (0.06%).

Figure 4.8 shows a calibration plot of the number of detected photon burst events as a function of the input genomic DNA copy number. As can be seen, the calibration plot was linear (r = 0.95 for the 20-cycle LDR) over the copy numbers investigated. To examine if we could produce a shorter assay turnaround time, a 2-cycle LDR was also performed in the same COC chip with downstream single-molecule interrogation. The number of photon burst events was again linear with genomic DNA copy number with the slope of the calibration plot equal to 0.0032 photon bursts/genomic DNA copy, compared to 0.015 photon bursts/genomic DNA copy for the 20-cycle LDR. The slope describes the analytical sensitivity of the assay and the 20-cycle LDR was roughly 5 times more sensitive than the 2-cycle LDR. Clearly, the improved analytical sensitivity arose from increases in the number of cycles imposed on the LDR phase of the assay, which provides linear amplification of the number of rMB generated. However, the 20-cycle LDR reported results in 19.2 min, while the 2-cycle LDR results in 2.6 min assay turnaround time.

In our case, the readout phases of the assay are instantaneous as the output from the continuous flow thermal reactor is sent directly into the fluorescence detector. As can be seen from the results depicted in Figure 4.8, the 20-cycle LDR can provide better analytical sensitivity in terms of discerning differences in copy numbers compared to the 2-cycle LDR. However, the photon bursts detected do originate from a single genomic DNA molecule, but not as the input because the sampling efficiency is well below 100% and thus, many DNA molecules are not detected due to the fact that they do not travel through the probe volume. To realize single copy detection as the input, the sampling efficiency must be increased to near 100%, which can be generated by

reducing the geometrical dimensions of the fluidic channel where the laser-induced fluorescence detection occurs and/or increasing the probe volume.

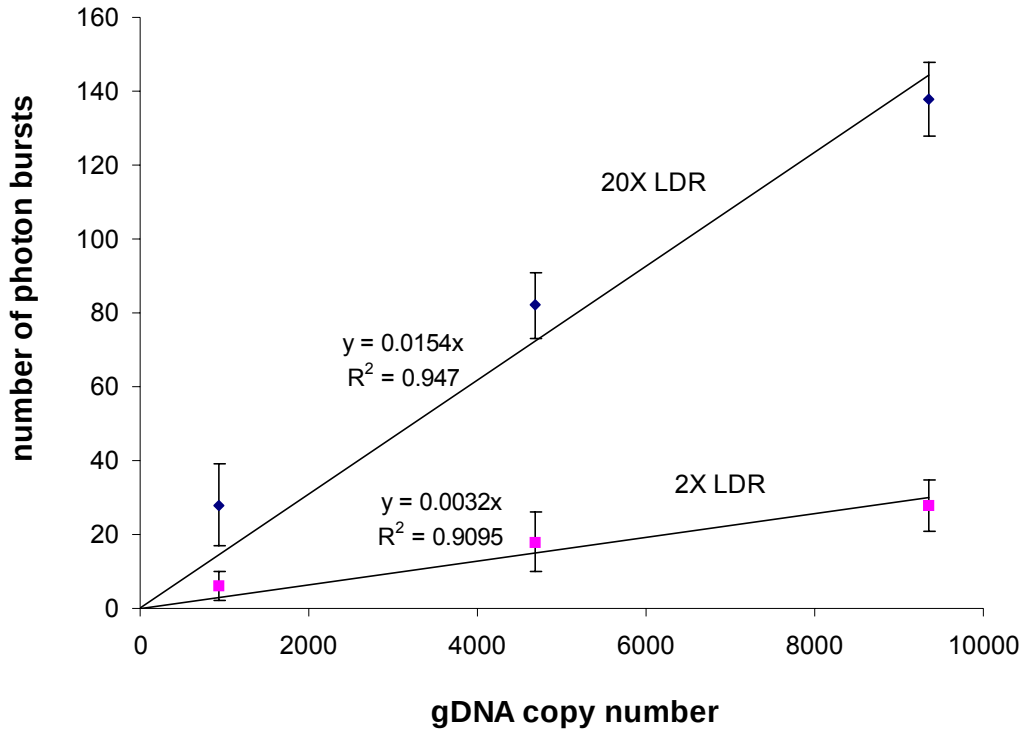


Figure 4.8 Plot of the number of detected photon burst events versus the input genomic DNA copy number. The upper curve shows the linear fit to a 20-cycle LDR with a correlation coefficient of 0.95 and the lower curve shows the fit to a 2-cycle LDR with a correlation coefficient of 0.91. The slope of the linear fitting function for the 2-cycle LDR was 0.0032, compared to 0.015 for the 20-cycle LDR. The slope of the fitting line described the sensitivity of the assay; the 20-cycle LDR was ~5 times more sensitive than the 2-cycle LDR.

4.4 Conclusions

Results have been obtained demonstrating the capability of LDR/spFRET to provide a low limit-of-detection and rapid reporting of bacterial pathogens with strain specificity as well as the ability to distinguish Gram(+) from Gram(-) species. These measurements were performed directly within a COC microchip and demonstrated the ability to process the input DNA sample using LDR without a PCR amplification step and detect the

products on-line in an automated fashion. With the measurements presented herein, the process time was found to be 2.6 min for a 2-cycle LDR and 19.2 min for a 20-cycle LDR. However, the larger number of cycles did improve the analytical sensitivity of the measurement. Significant improvements in assay sensitivity, even for the 2-cycle LDR, could be realized by simply increasing the sampling efficiency by reducing the channel size at the single-molecule detection zone of the chip and/or increasing the probe volume. For example, reducing the channel size to 1 μm (width) and 1 μm (depth) would provide a sampling efficiency of 100% for the laser beam size adopted herein.

The sampling rate, which is determined by the processing volume flow rate, must be balanced by optimizing the performance of the LDR[34] and the single-molecule detection efficiency.[46] It was determined to be 0.78 μl/min in the present case. This relatively low volumetric flow rate will make it difficult to process samples in which the bacterial copy number per unit volume is low, which would then require some type of bacterial target pre-concentration prior to the LDR/spFRET measurement. This can be envisioned by using an affinity pre-concentrator that can process large input volumes and select targets with high recovery. We have recently demonstrated the ability to use polyclonal antibodies to select certain pathogenic bacteria that are of low abundance from water samples.[47] Future work in our laboratory will integrate this rare cell selection device to LDR/spFRET to provide the ability to identify rare bacterial species from environmental samples in near real-time.

4.5 References

(1) Chuang, H.; Macuch, P.; Tabacco, M. B. *Analytical Chemistry* 2001, *73*, 462-466.

(2) McDowell, A.; Mahenthiralingam, E.; Moore, J. E.; Dunbar, K. E. A.; Webb, A. K.; Dodd, M. E.; Martin, S. L.; Millar, B. C.; Scott, C. J.; Crowe, M.; Elborn, J. S. *Journal of Clinical Microbiology* 2001, *39*, 4247-4255.

(3) Parkhurst, D. F.; Stern, D. A. *Environmental Science & Technology* 1998, *32*, 3424-3429.

(4) Alarcon, B.; Garcia-Canas, V.; Cifuentes, A.; Gonzalez, R.; Aznar, R. *Journal of Agricultural and Food Chemistry* 2004, *52*, 7180-7186.

(5) Belgrader, P.; Elkin, C. J.; Brown, S. B.; Nasarabadi, S. N.; Langlois, R. G.; Milanovich, F. P.; Colston, B. W.; Marshall, G. D. *Analytical Chemistry* 2003, *75*, 3446-3450.

(6) Tang, Y. W.; Ellis, N. M.; Hopkins, M. K.; Smith, D. H.; Dodge, D. E.; Persing, D. H. *Journal of Clinical Microbiology* 1998, *36*, 3674-3679.

(7) Rowe, C. A.; Scruggs, S. B.; Feldstein, M. J.; Golden, J. P.; Ligler, F. S. *Analytical Chemistry* 1999, *71*, 433-439.

(8) Rowe, C. A.; Tender, L. M.; Feldstein, M. J.; Golden, J. P.; Scruggs, S. B.; MacCraith, B. D.; Cras, J. J.; Ligler, F. S. *Analytical Chemistry* 1999, *71*, 3846-3852.

(9) Duburcq, X.; Olivier, C.; Malingue, F.; Desmet, R.; Bouzidi, A.; Zhou, F. L.; Auriault, C.; Gras-Masse, H.; Melnyk, O. *Bioconjugate Chemistry* 2004, *15*, 307-316.

(10) Delehanty, J. B.; Ligler, F. S. *Analytical Chemistry* 2002, *74*, 5681-5687.

(11) Kim, B. C.; Park, J. H.; Gu, M. B. *Analytical Chemistry* 2005, *77*, 2311-2317.

(12) Lagally, E. T.; Scherer, J. R.; Blazej, R. G.; Toriello, N. M.; Diep, B. A.; Ramchandani, M.; Sensabaugh, G. F.; Riley, L. W.; Mathies, R. A. *Analytical Chemistry* 2004, *76*, 3162-3170.

(13) Mayr, B. M.; Kobold, U.; Moczko, M.; Nyeki, A.; Koch, T.; Huber, C. G. *Analytical Chemistry* 2005, *77*, 4563-4570.

(14) Loge, F. N.; Thompson, D. E.; Call, D. R. *Environmental Science & Technology* 2002, *36*, 2754-2759.

(15) Brasher, C. W.; DePaola, A.; Jones, D. D.; Bej, A. K. *Current Microbiology* 1998, *37*, 101-107.

(16) Kong, R. Y. C.; Lee, S. K. Y.; Law, T. W. F.; Law, S. H. W.; Wu, R. S. S. *Water Research* 2002, *36*, 2802-2812.

(17) Beyor, N.; Yi, L. N.; Seo, T. S.; Mathies, R. A. *Analytical Chemistry* 2009, *81*, 3523-3528.

(18) Zourob, M.; Mohr, S.; Brown, B. J. T.; Fielden, P. R.; McDonnell, M. B.; Goddard, N. J. *Analytical Chemistry* 2005, *77*, 232-242.

(19) Zourob, M.; Hawkes, J. J.; Coakley, W. T.; Brown, B. J. T.; Fielden, P. R.; McDonnell, M. B.; Goddard, N. J. *Analytical Chemistry* 2005, *77*, 6163-6168.

(20) Premasiri, W. R.; Moir, D. T.; Klempner, M. S.; Krieger, N.; Jones, G.; Ziegler, L. D. *Journal of Physical Chemistry B* 2005, *109*, 312-320.

(21) Jarvis, R. M.; Goodacre, R. *Analytical Chemistry* 2004, *76*, 40-47.

(22) Tok, J. B. H.; Chuang, F. Y. S.; Kao, M. C.; Rose, K. A.; Pannu, S. S.; Sha, M. Y.; Chakarova, G.; Penn, S. G.; Dougherty, G. M. *Angewandte Chemie-International Edition* 2006, *45*, 6900-6904.

(23) Weeks, B. L.; Camarero, J.; Noy, A.; Miller, A. E.; Stanker, L.; De Yoreo, J. J. *Scanning* 2003, *25*, 297-299.

(24) Campbell, G. A.; Mutharasan, R. *Biosensors & Bioelectronics* 2005, *21*, 462-473.

(25) Gupta, A. K.; Nair, P. R.; Akin, D.; Ladisch, M. R.; Broyles, S.; Alam, M. A.; Bashir, R. *Proceedings of the National Academy of Sciences of the United States of America* 2006, *103*, 13362-13367.

(26) Rider, T. H.; Petrovick, M. S.; Nargi, F. E.; Harper, J. D.; Schwoebel, E. D.; Mathews, R. H.; Blanchard, D. J.; Bortolin, L. T.; Young, A. M.; Chen, J. Z.; Hollis, M. A. *Science* 2003, *301*, 213-215.

(27) Van Ert, M. N.; Easterday, W. R.; Simonson, T. S.; U'Ren, J. M.; Pearson, T.; Kenefic, L. J.; Busch, J. D.; Huynh, L. Y.; Dukerich, M.; Trim, C. B.; Beaudry, J.; Welty-Bernard, A.; Read, T.; Fraser, C. M.; Ravel, J.; Keim, P. *Journal of Clinical Microbiology* 2007, *45*, 47-53.

(28) Pingle, M. R.; Granger, K.; Feinberg, P.; Shatsky, R.; Sterling, B.; Rundell, M.; Spitzer, E.; Larone, D.; Golightly, L.; Barany, F. *Journal of Clinical Microbiology* 2007, *45*, 1927-1935.

(29) Barany, F. *Proc Natl Acad Sci U S A* 1991, *88*, 189-193.

(30) Wabuyele, M. B.; Farquar, H.; Stryjewski, W.; Hammer, R. P.; Soper, S. A.; Cheng, Y. W.; Barany, F. *Journal of the American Chemical Society* 2003, *125*, 6937-6945.

(31) Gill, S. R.; Fouts, D. E.; Archer, G. L.; Mongodin, E. F.; DeBoy, R. T.; Ravel, J.; Paulsen, I. T.; Kolonay, J. F.; Brinkac, L.; Beanan, M.; Dodson, R. J.; Daugherty, S. C.; Madupu, R.; Angiuoli, S. V.; Durkin, A. S.; Haft, D. H.; Vamathevan, J.; Khouri, H.; Utterback, T.; Lee, C.; Dimitrov, G.; Jiang, L. X.; Qin, H. Y.; Weidman, J.; Tran, K.; Kang, K.; Hance, I. R.; Nelson, K. E.; Fraser, C. M. *Journal of Bacteriology* 2005, *187*, 2426-2438.

(32) Krimmer, V.; Merkert, H.; von Eiff, C.; Frosch, M.; Eulert, J.; Lohr, J. F.; Hacker, J.; Ziebuhr, W. *Journal of Clinical Microbiology* 1999, *37*, 2667-2673.

(33) Chen, J. F.; Wabuyele, M.; Chen, H. W.; Patterson, D.; Hupert, M.; Shadpour, H.; Nikitopoulos, D.; Soper, S. A. *Analytical Chemistry* 2005, *77*, 658-666.

(34) Hashimoto, M.; Hupert, M. L.; Murphy, M. C.; Soper, S. A.; Cheng, Y. W.; Barany, F. *Analytical Chemistry* 2005, *77*, 3243-3255.

(35) Clarridge, J. E. *Clinical Microbiology Reviews* 2004, *17*, 840-+.

(36) Wang, H.; Chen, J. F.; Zhu, L.; Shadpour, H.; Hupert, M. L.; Soper, S. A. *Analytical Chemistry* 2006, *78*, 6223-6231.

(37) Hashimoto, M.; Barany, F.; Soper, S. A. *Biosensors & Bioelectronics* 2006, *21*, 1915-1923.

(38) Pu, Q. S.; Oyesanya, O.; Thompson, B.; Liu, S. T.; Alvarez, J. C. *Langmuir* 2007, *23*, 1577-1583.

(39) Steigert, J.; Haeberle, S.; Brenner, T.; Muller, C.; Steinert, C. P.; Koltay, P.; Gottschlich, N.; Reinecke, H.; Ruhe, J.; Zengerle, R.; Ducree, J. *Journal of Micromechanics and Microengineering* 2007, *17*, 333-341.

(40) Mali, K. S.; Dutt, G. B.; Mukherjee, T. *J Chem Phys* 2008, *128*, 124515.

(41) Chen, Y.; Mauldin, J. P.; Day, R. N.; Periasamy, A. *Journal of Microscopy-Oxford* 2007, *228*, 139-152.

(42) Ha, T.; Enderle, T.; Ogletree, D. F.; Chemla, D. S.; Selvin, P. R.; Weiss, S. *Proceedings of the National Academy of Sciences of the United States of America* 1996, *93*, 6264-6268.

(43) Best, R. B.; Merchant, K. A.; Gopich, I. V.; Schuler, B.; Bax, A.; Eaton, W. A. *Proceedings of the National Academy of Sciences of the United States of America* 2007, *104*, 18964-18969.

(44) Marras, S. A. E.; Kramer, F. R.; Tyagi, S. *Nucleic Acids Research* 2002, *30*, -.

(45) Okagbare, P. I.; Emory, J. M.; Datta, P.; Goettert, J.; Soper, S. A. *Lab on a Chip* 2010, *10*, 66-73.

(46) Mathies, R. A.; Peck, K.; Stryer, L. *Analytical Chemistry* 1990, *62*, 1786-1791.

(47) Dharmasiri, U.; Witek, M. A.; Adams, A. A.; Osiri, J. K.; Hupert, M. L.; Bianchi, T. S.; Roelke, D. L.; Soper, S. A. *Analytical Chemistry* 2010, *82*, 2844-2849.

CHAPTER 5 QUANTIFICATION OF mRNA TRANSCRIPTS FOR EXPRESSION ANALYSIS OF MMP-7 GENE WITH RT-PCR AND RT-qPCR

5.1 Introduction

With the epic completion of the human genome project and securing genomic sequence data for more and more organisms is being realized, research emphasis in the post-genomic era has been geared towards understanding the connections between the expression of individual genes or a group of genes and their unique biological functions. Determining the presence of specific genes and their expressional abundance in different cells, or tissues under different physiological conditions is playing a significant role in the elucidation of new signal transduction pathways,[1] discovery of new drug targets,[2] revelation of subtypes of diseases,[3] and the prediction of certain therapeutics to treat specific diseases (i.e., personalized medicine).[4, 5] Expression of genes can be evaluated either on the transcriptional level, in which the presence and quantity of some specific mRNA sequences transcribed from a particular gene contained within genomic DNA (gDNA) is examined, or on the translational level, in which the presence and quantity of certain proteins encoded from genes is investigated.

Proteins are the final product of gene expression and are generally regarded as the functional molecular machines of life. Profiling of protein expression levels among different cells or tissues has been routinely performed using western blotting, two-dimensional gel electrophoresis, liquid chromatography coupled with mass spectrometer and ELISA techniques.[6-8] However, methods to analyze protein expression levels are usually cumbersome, laborious and costly compared to the methods to measure mRNA transcript levels.

In the central dogma of molecular biology, the genetic information stored in gDNA is translated into proteins with mRNA as the intermediate agent, which suggests that there might be a direct relationship between mRNA and protein expressions. A number of studies have shown that there are significant general correlations between expression on transcriptional and translational levels.[9-12] Therefore, studies on mRNA expression levels can be used as a proxy to predict the expression abundance at the protein levels in some circumstances because differences in protein expression are reflected by the differences in mRNA expression as well.

5.1.1 Significance of MMP-7 Expression in Cancer Prognosis

Colorectal cancer is becoming a growing health problem around the world in the recent decade with over one million new cases of colorectal cancer diagnosed and 609,051 people dying from this disease in 2008 alone.[13] Thus, discovery of sensitive and reliable biomarkers has important clinical value in early diagnosis of this fatal disease. The matrix metalloproteinases (MMPs) are a family of secreted and membrane-bound endopeptidases that can degrade components associated with the extracellular matrix. Among the family of MMPs, Matrix Metalloproteinase-7 (MMP-7), also known as matrilysin, is the smallest member with a molecular weight of 28 kDa. MMP-7 has long been known as an important factor in early tumor growth and it has been extensively reported that expression of MMP-7 was associated with a wide variety of cancer types including colorectal cancer as well as carcinomas from the brain, lung, neck, stomach, prostate and adenocarcinoma from breast, colon and pancreas.[14-22] Although the exact mechanism of this association has not be unraveled yet, the capability of matrilysin to degrade the basement membrane and extracellular matrix is

believed to play a crucial role in invasiveness, metastasis and progression of various tumors, which has been demonstrated by experiments using MMP-7 transcript inhibitors to control metastasis.[23]

5.1.2 Quantitative Measurement of MMP-7 transcripts by RT-qPCR

5.1.2.1 Comparison of RT-qPCR and Conventional RT-PCR

RT-qPCR, reverse transcription quantitative PCR, is a popular method for gene expression analysis due to the many benefits it offers, such as extremely high sensitivity, large dynamic range of 7 to 8 log orders, no post-amplification manipulation, and ease of automation.[24] RT-PCR is typically performed in a closed tube and the quantity of PCR products is measured by gel electrophoresis with ethidium bromide staining at the end of the reaction (end-point PCR), which can be a time-consuming measurement. End-point PCR has a limited dynamic range for quantitative measurements and its results are based on comparing the intensity of an amplified band on a gel to standards of a known concentration, which is only semi-quantitative. In addition, end-point PCR is prone to high variability because the measurement is performed at the plateau phase of the PCR.[25] Figure 5.1 demonstrates the fluorescence data measured at the plateau stage of PCR from 96 identical reactions. We can see from this figure that the measured fluorescence data are spread over a certain range making precise quantitative analysis difficult using end-point PCR quantification.

By contrast, in RT-qPCR the amplification and fluorescence detection were combined into a single step, and the fluorescence data were collected during each PCR cycle to monitor the progress of the reaction in real-time, which is a great convenience.

Besides, the fluorescence data from the exponential stage was used for quantification, which is much less variable than the data from the plateau stage of PCR.

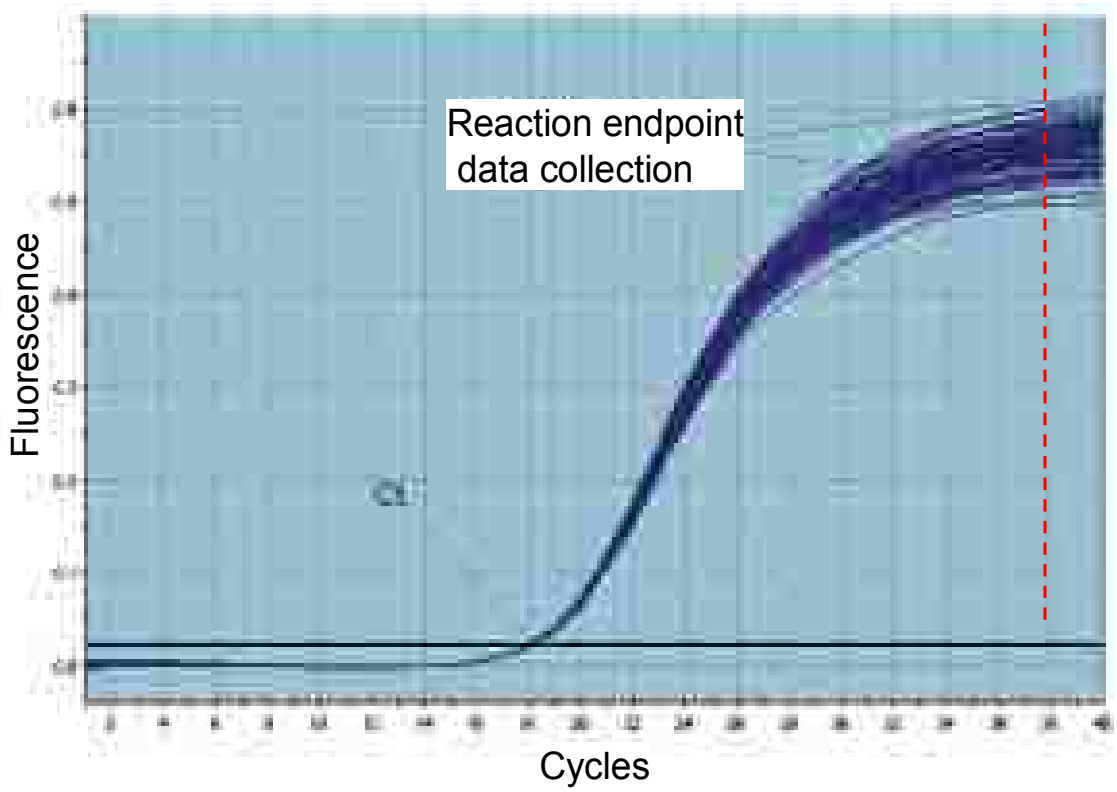


Figure 5.1 RT-qPCR run using 96 identical reactions.[26]

5.1.2.2 Principles of RT-qPCR

A typical PCR is composed of linear ground stage, early exponential stage, log-linear stage, and plateau stage.[24] During the linear ground stage, usually the first 10-15 thermal cycles, the fluorescence at each PCR cycle has not risen above the background fluorescence fluctuations and at this phase, the baseline is determined. At the early exponential stage, enough fluorescence is accumulated with increased amounts of amplicons to reach a threshold that is significantly above the background fluorescence (usually 10 times the standard deviation of the baseline).[24] The cycle number at this point is defined as C_T, which is closely related to the amount of initial templates in the

PCR mixture. The more templates present at the beginning of the reaction, the fewer number of cycles it takes to reach the threshold level, thus smaller C_T. During the log-linear stage the amplification of PCR reaches its best performance and PCR products are nearly doubled in every cycle. When the plateau stage is reached, reaction components are almost exhausted so the fluorescence intensity doesn't further increase with cycle number. Because C_T always occurs during the exponential stage of PCR, quantification is not affected by any reaction components becoming limited as in the plateau stage.

5.1.2.2.1 Real-Time Detection in RT-qPCR

In RT-qPCR the mRNA transcripts were first reversed transcribed into their complementary DNA (cDNA) followed by PCR amplifications. The PCR products are detected during the annealing or extension steps during each thermal cycle using fluorescence methods (see Figure 5.2). There are two fluorescence methods that are generally utilized in real-time monitoring: DNA intercalating dye or FRET probes.

Some fluorescent dyes such as SYBR Green I, show very little fluorescence when free in aqueous solutions. But when it is bound to double-stranded DNA, its fluorescence is enhanced by over 1000-fold. The detected fluorescence level is proportional to the dsDNA concentration, because the more dsDNA that is present, the more binding sites that are available for SYBR Green I to attach. Therefore, this dye can be used to monitor the accumulation of PCR products. As the target is amplified, the increasing concentration of double-stranded amplicons is deduced from the increasing fluorescence signals. The mechanism of SYBR Green I for the real-time monitoring of PCR progress is illustrated in Figure 5.2.

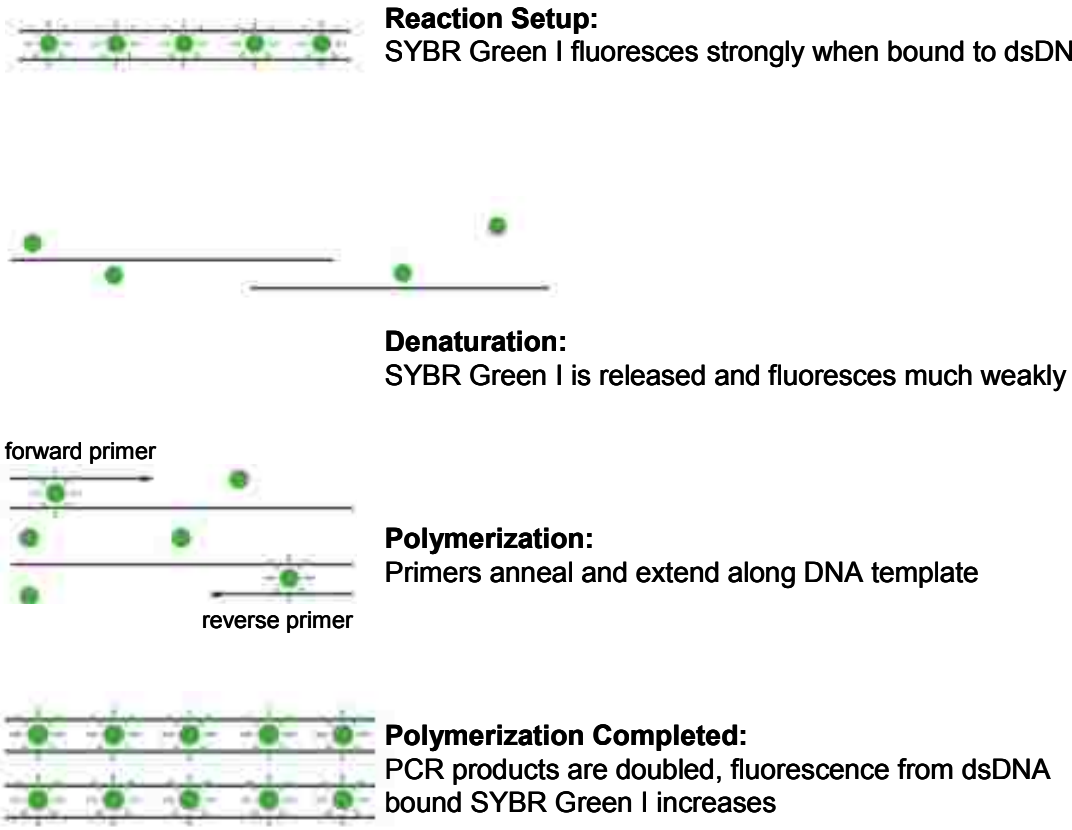


Figure 5.2 Mechanism of SYBR Green I in real-time monitoring of PCR progress

SYBR Green I can be added into each individual reaction tube, or into the PCR master mix, which is simple to use. It can be used to monitor the amplification of any dsDNA sequence without need to design probes for each specific target gene as is the case in probe-based detection chemistry. Therefore it provides the most economical solution in RT-qPCR applications. However, SYBR Green I is limited by its inherent non-specificity since it can bind to any dsDNA in the reaction, including primer-dimers and any non-specific dsDNA products. Therefore, it is susceptible to false positive signals, and the detection specificity is only ensured by the specificity of the primer design.

A higher level of detection specificity can be generated by using sequence-specific probes to detect the PCR products of interest. TaqMan, for instance, is a widely used oligonucleotide probe system for real-time qPCR applications and its detection specificity comes from the oligonucleotide sequence that is designed to specific PCR amplicons. The principle of the TaqMan probe system used in conjunction with qPCR is shown in Figure 5.3.

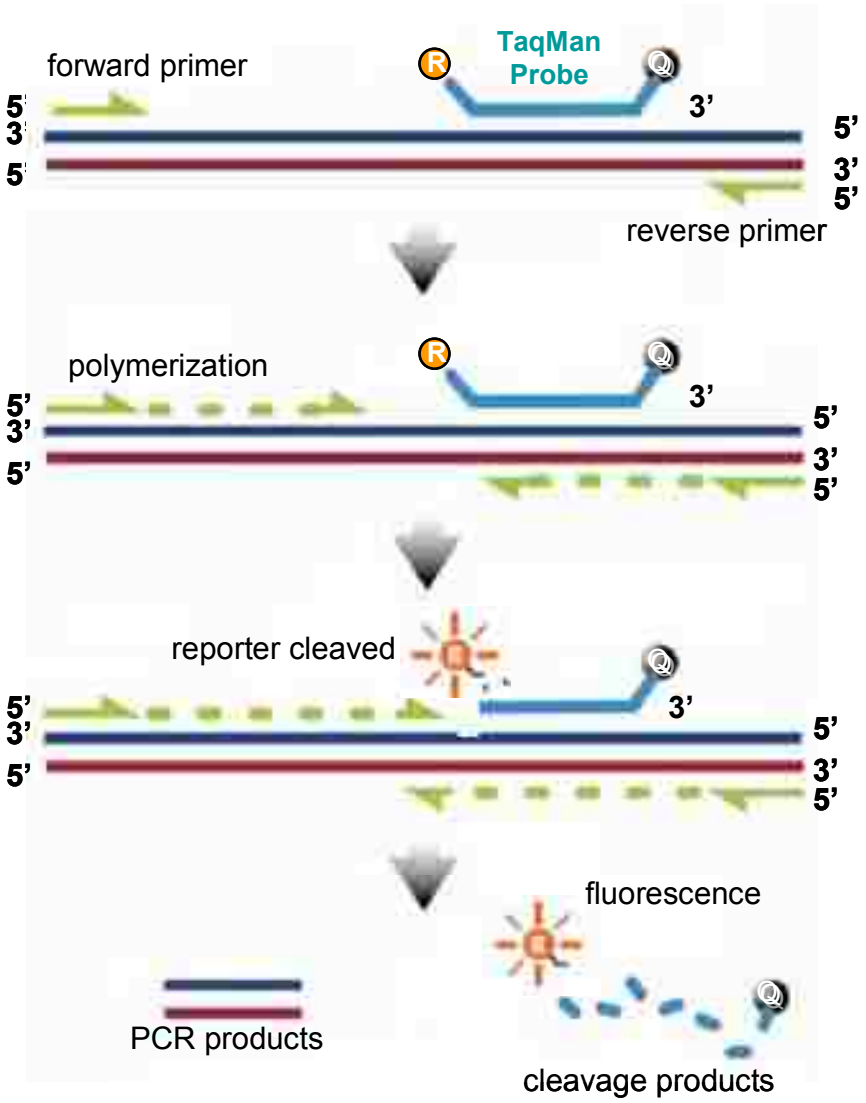


Figure 5.3 Principle of TaqMan probe in qPCR real-time monitoring.[26] This probe is a short oligonucleotide strand that is labeled with a fluorescent reporter on its 5' end and a quencher moiety on its 3' end.

Briefly, the TaqMan probe system consists of a dual-labeled oligonucleotide strand that is covalently attached to a fluorescent reporter on its 5' end and a quencher moiety on its 3' end. In its free state, the fluorescence from the reporter is turned off by the quencher through fluorescence resonant energy transfer (FRET). The probe is designed to anneal to one strand of the target sequence just slightly downstream of one of the primers. During PCR, when the polymerase extends the primer, it will encounter the 5' end of the probe. The Taq DNA polymerase has 5' exo-nuclease activity so it cleaves the 5' end of the probe and releases the reporter fluorophore into solution, restoring its fluorescence. The magnitude of fluorescence is directly proportional to the amount of PCR amplicons generated, which enables real-time monitoring of PCR progress. Compared to SYBR Green I, the TaqMan probe system is highly specific as noted above and detects amplification products only because the probe is designed to be complementary to the internal sequence of the amplicons. The probes can be designed on different PCR products and labeled with different, spectrally distinguishable reporter dyes, which allows for multiplexed detection of different PCR products in one reaction tube. The disadvantage of TaqMan probes is that a different probe has to be designed for each unique target sequence interrogated.

5.1.2.2.2 Standard Curve

In order to get the exact copy number of mRNA transcripts in the sample, a standard curve is usually constructed first. The mRNA standards are first subjected to a series of dilutions and mixed into the RT-qPCR cocktail. The dilution series is usually made from 3 - 10-fold concentration range to ensure that the reaction performs at equal efficiency for high and low concentrations of starting template copies. These individual

reactions are carried out in the same batch, and C_T from each reaction is plotted against the mRNA copy number spiked into each reaction. Supposedly, the C_T value is inversely proportional to the log of the initial mRNA copy number.[27] This curve is then used as a reference to extrapolate quantitative information of the target mRNA of unknown concentration in the real sample.

5.1.2.2.3 Reference Dye

It is a common practice in qPCR to include a reference dye in the reaction mixture. The reference dye will not bind to PCR products and is not related to any amplification variations. Therefore, the fluorescence from the reference dye should stay constant throughout the amplification reaction when placed in the PCR tube. Theoretically, the fluorescence for the reference dye should also be the same in each reaction tube. Practically, however, there always exists slight fluorescence differences caused by artifacts, such as differences in transparency and reflectivity of plastic microtubes, gas bubbles produced in the reaction mixture, and volume differences due to aliquotting errors. It has been shown that reaction volume is a key factor in PCR amplification efficiency, and volume differences in the master mix when using the same amount of starting template result in different amplification efficiencies.[28] A passive reference dye, such as ROX, is thus included in the master mix to normalize for these non-amplification related uncertainties during the fluorescence readout of the assay.

5.1.3 Housekeeping Genes

When gene expression levels in different samples are compared, it is crucial to consider experimental variations such as amount of starting template, RNA extraction quality and reverse transcription efficiency, which affected the quantification accuracy of

RT-qPCR.[29] Therefore, the RT-qPCR results are usually normalized to an internal standard, often referred to as a housekeeping gene or endogenous reference.[30]

The suitable housekeeping gene is required to be expressed adequately in the sample of interest, and more importantly, its expression level should show minimal sample-to-sample variability under the experimental conditions used. β–Actin is one of the earliest used housekeeping genes in transcriptional expression analysis and is expressed at moderately abundant levels in most cell types.[31] Glyceraldehyde-3-phosphate dehydrogenase (GAPDH) and beta-2-microglobulin (B2M) are another two frequently used reference genes that are ubiquitously expressed at moderately abundant levels.[32, 33] Therefore, these three commonly used reference genes are selected for further investigation on their expression stability in this study.

In this chapter, the expression level of MMP-7 transcripts from a selected group of cell lines will be investigated using conventional RT-PCR and RT-qPCR techniques, and their performance will be evaluated in quantitative measurement. Running conditions will be optimized to obtain the most efficient amplification in RT-PCR and RT-qPCR. The appropriate housekeeping gene in this application will also be selected through experiments to serve as a good internal control in future studies.

5.2 Experimental

5.2.1 Cell Cultures

HT-29, HeLa, SW480, SW620, LS180, MCF-7, and HEL299 cells were purchased from ATCC and cultured according to ATCC's protocol. All cells were cultured in Dulbecco's Modified Eagle's Medium supplemented with high glucose containing 1.5 g/L sodium bicarbonate ($NaHCO_3$), 15mM HEPES buffer, and 10% fetal bovine serum.

Cultures were maintained by replacing fresh medium twice a week, and the cell density was maintained between $2X10^5$ to $5X10^6$ cells/mL. A 0.25% trypsin solution was prepared in 150 mM PBS buffer to release the cultured cells when needed.

5.2.2 Total RNA Extraction

The harvested cells were first pelleted by centrifugation at 300 RPM for 5 min. The total RNA in the pelleted cells was extracted using a Total RNA Extraction kit (RTN70, Sigma-Aldrich, St. Louis, MO). After extraction the purified total RNA was incubated with Amplification Grade DNase to digest any potential genomic DNA contaminants that co-eluted from the spin column. The treated total RNA was divided into aliquots and stored at -80ºC until use. The concentration of the total RNA was determined by measuring the absorbance at 260 nm in 1X Tris·HCl buffer. The quality of the extracted RNA was examined by determining the ratio of absorbance between 260 nm to 280 nm (A_{260}/A_{280}).

5.2.3 RT-PCR

Reverse transcription PCR (RT-PCR) was performed using the one-step Access RT-PCR System (A1250, Promega, Madison, WI). The reaction mixture was prepared according to the manufacturer's protocol and the final cocktail contained 1X AMV/*Tfl* reaction buffer, 0.2 mM dNTP mix, 200 nM forward primer, 200 nM reverse primer, 1 mM $MgSO_4$, 0.1U/μL *Tfl* DNA polymerase. 0.1-1 μg extracted total RNA was added into each reaction well except for the no-template control. Enough nuclease-free H_2O was then added to make the total reaction volume 50 μL. AMV reverse transcriptase (RT) was added lastly into each reaction well except for the no-RT control. Each reaction mixture was overlaid with nuclease-free mineral oil to prevent evaporation during thermal cycling. The RT-PCR was run using a commercial thermal cycler (Eppendorf

MasterCycler, Hamburg, Germany). The reaction cocktail was first incubated at 45ºC for 45 mins to synthesize the first strand cDNA from the target mRNA. Then, the reaction mixture was heated to 94ºC for 2 min to deactivate the AMV RT, followed by 30-40 thermal cycles of amplification at 94ºC for 30 s, 50-65ºC for 1 min, and 68ºC for 2 min. The reaction was maintained at 68ºC for an additional 7 min to allow for complete extension. The PCR products were verified by agarose gel electrophoresis using ethidium bromide staining.

5.2.4 RT-qPCR

RT-qPCR was conducted using the one-step Brilliant II SYBR Green RT-qPCR kit (Stratagene, Santa Clara, CA). To minimize pipetting errors and to achieve better reproducibility, a master mix of the common components was first prepared and aliquotted into each PCR tube. Each reaction was carried out in a total volume of 25 µL, consisting of 1X Brilliant II SYBR Green qRT-PCR master mix, 200 nM upstream primer, 200 nM downstream primer, and 100 ng of experimental total RNA. One µL of RT/RNase block enzyme was added into each reaction tube. All reactions were prepared in triplicates to determine potential sample-to-sample variations. The PCR tube containing the reaction mixture was sealed with optically clear cap to minimize autofluorescence.

RT-qPCR was performed on the MX4000 Multiplex Quantitative PCR system (Stratagene, Santa Clara, CA) for real-time fluorescence monitoring. The reaction mixture was first incubated at 50ºC for 30 min to synthesize the first strand cDNA, then heated to 95 ºC for 10 min to deactivate the reverse transcriptase and activate the hot-start Taq DNA ligase. The amplification reaction was run for 40 thermal cycles at 95ºC

for 30 s, 50-65ºC for 1 min, 72ºC for 30 s. Fluorescence was measured at the end of the extension step to monitor the accumulated amplicons. At the conclusion of 40 cycles, a dissociation curve analysis was performed to check any non-specific amplification that may have occurred. The amplified products were incubated at 95ºC for 1 min, cooled to 55ºC at a rate of 0.2ºC/sec, followed by 81 cycles of incubation where the temperature was increased by 0.5ºC/cycle, beginning at 55ºC and ending at 95ºC with a duration step of 30 s for each cycle. Fluorescence was also measured at the end of each incubation cycle.

5.2.5 Design of RT-PCR Primers

When performing RT-qPCR for mRNA analysis, an important consideration is residual genomic DNA or unspliced RNA that tends to persist in many RNA preparations. These contaminants can also be amplified by PCR primers designed for specific cDNAs and lead to false positive results. The common solution to distinguish target cDNAs generated from their associated mRNAs from gDNA or unspliced RNA is to design the forward and reverse primers either from different exons or primers that bridge an exon-exon boundary in the target sequence. In the former case, amplicons from gDNA or unspliced RNA are bigger in size than the amplicons used for the target transcripts and can easily be differentiated based on gel electrophoresis. In the latter case, at normal annealing temperatures, the PCR primers are not able to hybridize with the gDNA or unspliced RNA sequence because there is a large intron between them. Thus, a positive result will exclude both gDNA contamination and unspliced transcripts. The RT-PCR primers used in this study to amplify MMP-7, GAPDH, β-Actin, and B2M transcripts are listed in Table 5.1.

Table 5.1 RT-PCR primers for MMP-7 and housekeeping gene transcripts

	primer direction	primer sequence (5'-3')	T_m, ºC	amplicon length, bp
MMP-7	forward	AAACTCCCGCGTCATAGAAAT	54.3	395
	reverse	TCCCTAGACTGCTACCATCCG	57.8	
GAPDH	forward	TCACCAGGGCTGCTTTTAAC	55.5	403
	reverse	CAGGAGGCATTGCTGATGAT	55.2	
β–Actin	forward	CACCACACCTTCTACAATGAGC	55.8	388
	reverse	GCTTCTCCTTAATGTCACGCAC	56.2	
B2M	forward	TGTCTTTCAGCAAGGACTGG	54.9	360
	reverse	TAGAGCTACCTGTGGAGCAA	55.2	

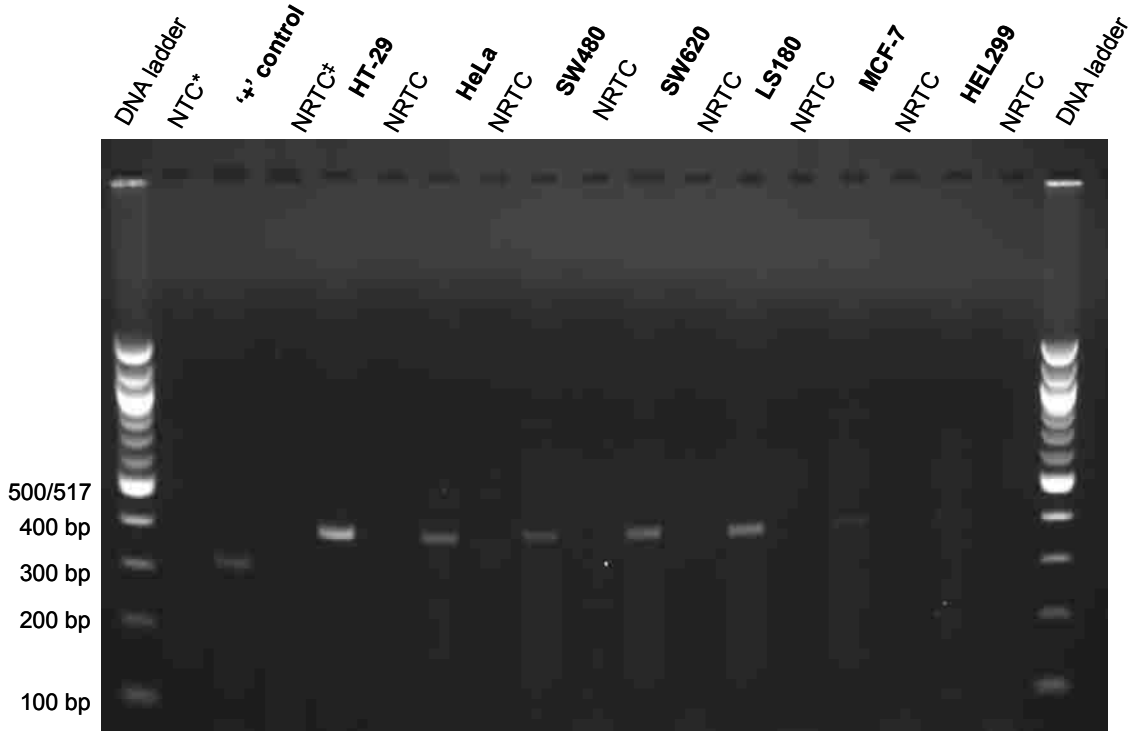
5.3 Results and Discussions

5.3.1 RT-PCR

5.3.1.1 Detection of MMP-7 mRNA in Different Cells

In this study, seven different types of cultured cells were used to detect their expression level of the MMP-7 gene by RT-PCR or RT-qPCR. MMP-7 has been reported to play an important role in metastasis and progression of colorectal and breast cancer.[21] Cell lines HT-29, HeLa, SW480, SW620, LS180 that originate from colorectal carcinomas, and MCF-7 from a breast carcinoma, were examined with regard to their MMP-7 expression level. HEL299 is from normal human embryonic lung tissue and was also examined as a negative control. In Figure 5.4, 0.1 μg of total RNA extracted from all seven cell lines were subjected to 35 thermal cycles of PCR following reverse transcription using primers specific to the MMP-7 transcript; the PCR products in all cases were 395 bp in length. The intensity of the band in the image provided an estimate of the relative amount of MMP-7 mRNA contained within the total RNA extract of each cell line. In this figure, HT-29 seemed to have the highest expression level of the MMP-7 gene based on its electrophoretic band intensity. All of the other four

colorectal cancer cell lines (HeLa, SW480, SW620, LS180) showed diverse levels of MMP-7 transcripts. The band intensity for MCF-7 cells in the gel image is much weaker compared to the five colorectal cancer cells, indicating a lower MMP-7 expression level. The band for HEL299 was nearly invisible in the image, indicating there is very little if any MMP-7 expression from HEL299. This observation is consistent with published results in the literature.



* NTC: no template control
‡ NRTC: no reverse transcriptase control

Figure 5.4 Gel electrophoresis image of RT-PCR to verify MMP-7 transcripts in different cell lines. 0.1 µg of total RNA was extracted from all seven cell lines and used as template for RT-PCR. The anticipated PCR products were 395 bp long. The reaction cocktail was incubated at 45ºC for 45 mins to synthesize the first strand of cDNA, followed by 35 thermal cycles at 94ºC for 30 s, 60ºC for 1 min, and 68ºC for 2 min for amplification. The PCR products were separated using a 3% agarose gel. The electrophoresis was run in 1X TBE buffer under an electric field of 8 V/cm for 50 min.

For expression analysis using RT-PCR or RT-qPCR, it is very important to include various controls to make the results meaningful.[26] A positive control was introduced into

the reaction cocktails to validate the RT-PCR reagents utility, including enzymes and buffer conditions, in case the desired RT-PCR products were not produced as expected. This positive control is provided in the RT-PCR kit for quality verification and has a target amplicon of 308 bp long. In Figure 5.4, a clear band was generated for the positive control corresponding to an amplicon of 308 bp long, suggesting that the quality of RT-PCR reagent was secured. RT-PCR assay should also include a no template control (NTC) to monitor contamination or non-specific amplifications. In Figure 5.4, there is no band observed for the NTC, indicating a clean sample mixture for RT-PCR, and there are no non-specific amplifications. As discussed before, RT-PCR is prone to genomic DNA contaminations, which can result in overestimation of the expression level of target transcripts using RT-PCR.[34] The no reverse transcriptase control (NRTC) is an effective method to identify contaminated gDNA because residual gDNA can be amplified by the polymerase in the RT-PCR kit and appear as a longer fragment in the gel image than the band for target transcript. NRTC was included for every cell line in this analysis to ensure the quality of the RNA extraction. In Figure 5.4, we can see that there are no bands observed for the NRTC case, which clearly suggested that the RNA samples extracted from all cell lines were prepared without gDNA contamination.

5.3.1.2 Quantitative Detection of MMP-7 Transcripts

The purpose of expression analysis is not only to get qualitative information concerning the presence of a target transcript, but also to know precisely the amount of the transcripts in the sample (*i.e.*, quantitative data). In order to see if RT-PCR can be used to quantify the mRNA expression level, total RNA extracted from HT-29 was subjected to a series of dilutions and spiked into various RT-PCR reaction cocktails.

The total RNA in each reaction cocktail ranged form 0.4 pg to 0.4 µg and they underwent 35 thermal cycles of amplification using identical conditions. Figure 5.5(A) shows the gel image of the RT-PCR with different total RNA concentrations. As expected, the band intensity increased significantly with increased initial RNA concentrations. To determine quantitatively the amount of amplicons displayed in the gel image, the intensity of each band was measured using Image Quant software and normalized to the 500/517 band of the DNA ladder, whose mass is 97 ng, The mass of each amplicon was plotted in Figure 5.5(B).

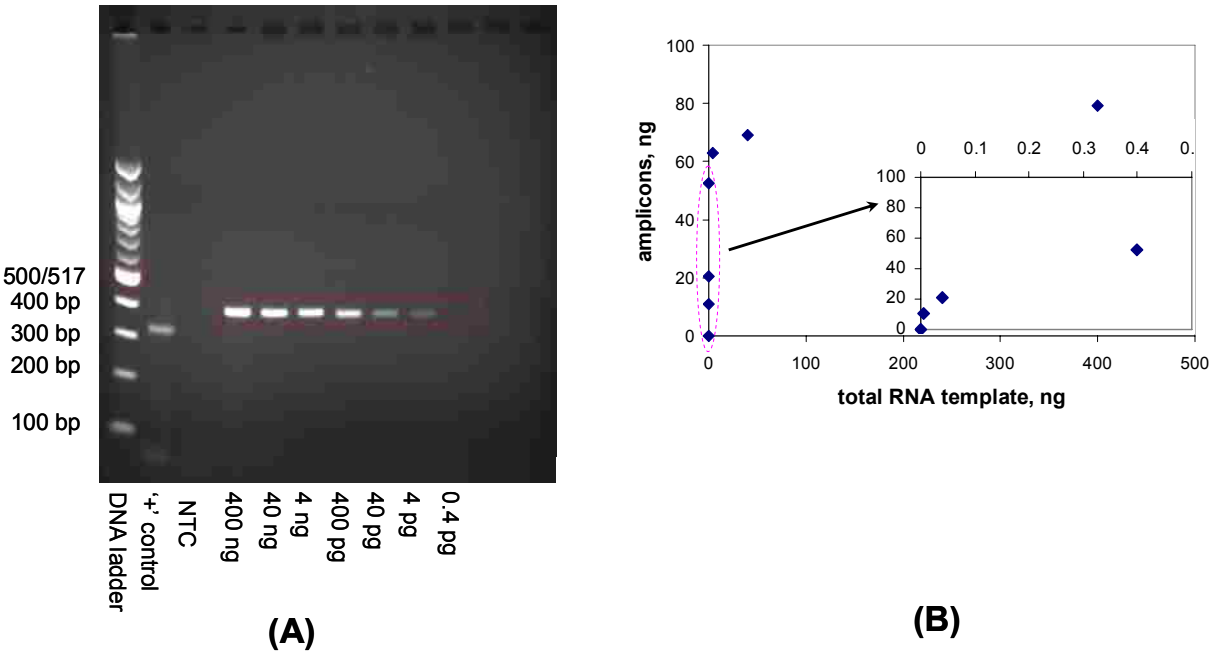


Figure 5.5 Quantification of MMP-7 transcripts using RT-PCR with end-point detection. Total RNA was extracted from HT-29, which was used as the template for RT-PCR. The running conditions of RT-PCR and gel electrophoresis in this experiment are the same as those used in Figure 5.4. (A) Gel image of RT-PCR. (B) Amount of PCR products vs. amount of total RNA template input to the assays.

In this figure, we found that the amount of amplicons increased rapidly as the amount of RNA templates increased from 0.4 pg to 400 pg. However, when the amount of input RNA was increased beyond 400 pg, the amount of amplicons generated did not

increase significantly. This is because after 35 thermal cycles the amplifications have already reached their plateau stages of the process. Therefore, RT-PCR with end-point detection can only provide semi-quantitative information, but only for cases where the input copy number is relatively low.

5.3.1.3. Expression of Housekeeping Genes in Different Cell Lines

Traditionally housekeeping genes with stable expression levels have been used in gene expression assays and serve an internal standard to normalize for experimental variations that may affect quantification accuracy. In practice, there is no universal housekeeping gene that exists for every RNA sample and even the most commonly used housekeeping genes such as GAPDH, β-actin have shown significant variations in their expression levels between different samples.[34] Thus, the expression stability of the selected housekeeping genes needs to be tested by each developed assay to determine its utility as an acceptable standard. The expression level of GAPDH, β-actin and B2M among all seven cell lines were examined and are shown in Figure 5.6.

0.2 μg of total RNA extracted from each cell line was subjected to 35 thermal cycles. The band intensity of each amplicon was normalized with respect to the corresponding 500/517 band, which consisted of a known mass of DNA; the results are shown in Figure 5.6(D). From this figure, we can see that after the same number of thermal cycles, all seven cell lines generated similar amounts of amplicons for GAPDH, β-actin and B2M, suggesting that the assay, including both the extraction and RT-PCR produce consistent results. However, this demonstration is only qualitative due to the semi-quantitative nature of end-point RT-PCR. The quantitative determination of the exact copy number of transcripts of each housekeeping gene is left to RT-qPCR to verify.

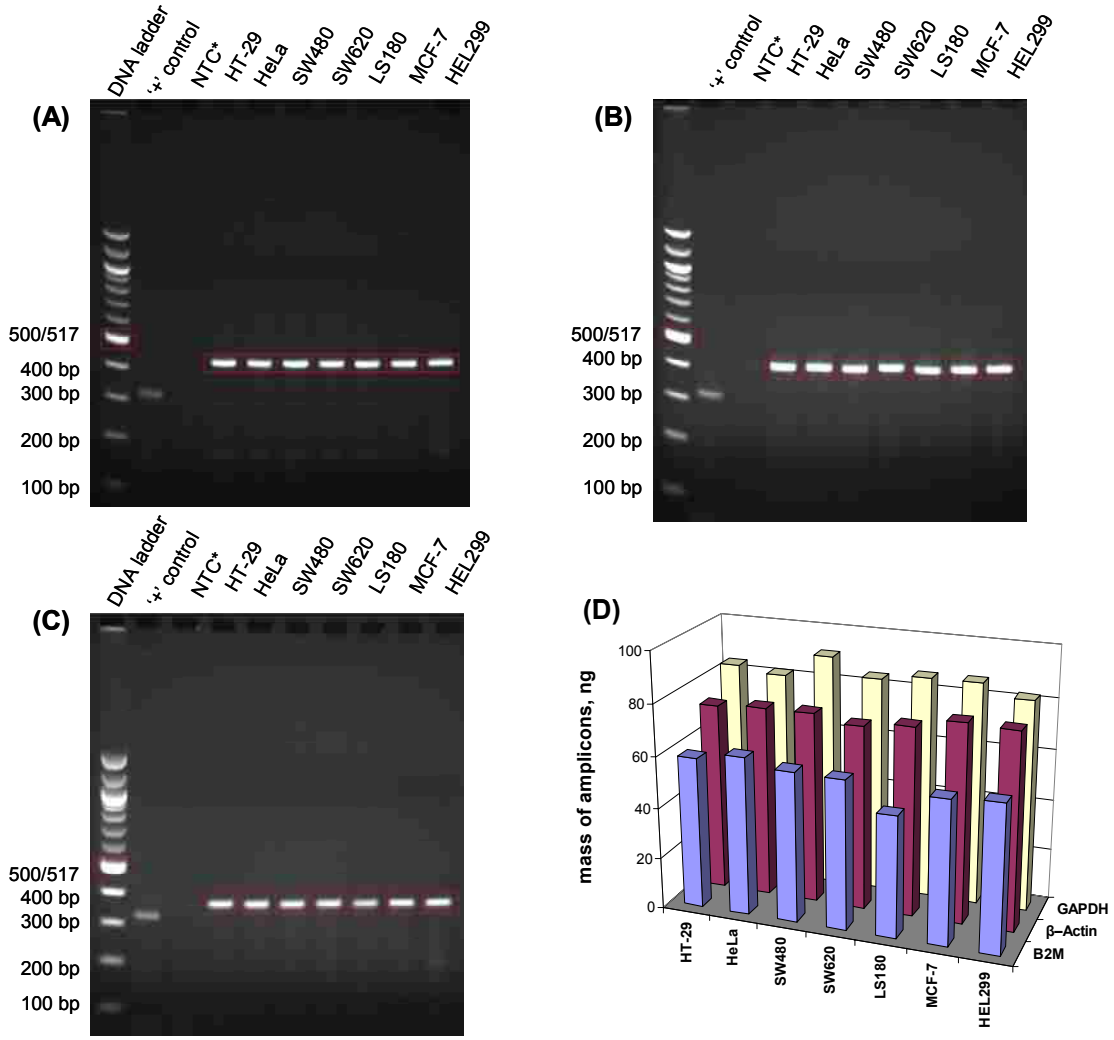


Figure 5.6 Expression levels of commonly used housekeeping genes among different cell lines. (A) GAPDH (B) β–Actin (C) B2M (D) Histogram of expression levels among different cell lines. 0.1 µg of total RNA extracted from all seven cell lines were used as template in RT-PCR. The PCR products are 403 bp, 388 bp, 360 bp long, for GAPDH, β–Actin, and B2M transcripts, respectively. The running conditions of RT-PCR and gel electrophoresis in this experiment are the same as used in Figure 5.4.

5.3.1.4 Optimization of Annealing Temperature for RT-PCR

In RT-PCR amplification, the PCR primers must first hybridize or anneal to their respective target sequence. Hybridization is a critical step in RT-PCR; fast and specific hybridization will lead to highly efficient amplification, thus high sensitivity for expression analysis. Hybridization usually takes place at temperatures in the neighborhood of the

T_m of the primers with respect to their respective complementary sequences in the target. In practice, the optimal annealing temperature for each primer combination needs to be obtained experimentally to ensure the most efficient PCR amplification.

In Figure 5.7, RT-PCR assays with gradient annealing temperatures were performed for MMP-7 and all three housekeeping genes. The band intensity of the amplicons at each annealing temperature was measured with Image Quant software and plotted in Figure 5.7(E). The optimal annealing temperature for each transcript was thus determined and is listed in Table 5.2. These annealing temperatures will be used in subsequent RT-qPCR experiments to achieve the best sensitivity of analysis.

Table 5.2 Optimal annealing temperature for various target transcripts

Transcript	Optimal Annealing Temperature, ºC
MMP-7	59.4
GAPDH	59.4
β-actin	57.3
B2M	61.5

5.3.2 Real-Time RT-PCR

Because end-point RT-PCR, as discussed above, is unable to provide accurate quantitative information on the mRNA level, especially for low abundance mRNA, quantitative RT-PCR (RT-qPCR) with real-time fluorescence detection was employed to quantify the copy number of target transcript due to its extreme sensitivity and broad dynamic range.

In RT-qPCR, a standard curve was first constructed by plotting the log of the initial template copy number against C_T for each sample containing a known template copy number. Quantification of initial copy number of target transcript in an unknown sample can then be accomplished by comparing its C_T value to that of the standard curve. Ideally, in transcriptional expression analysis of mRNA with known copy number should

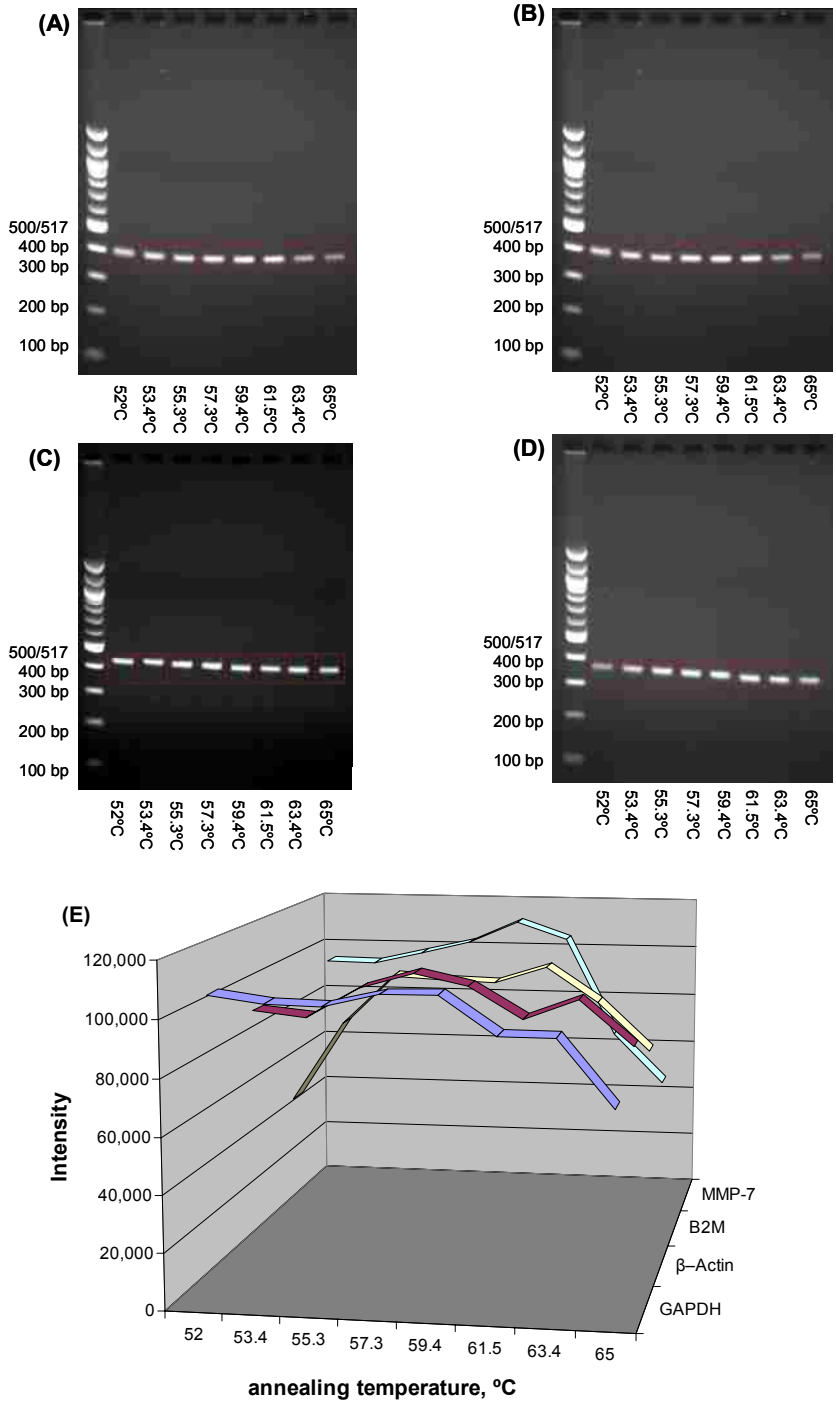


Figure 5.7 Amplification efficiency of RT-PCR with hybridization temperatures. 0.1 µg of total RNA extracted from HT-29 was used as template in RT-PCR. The reaction cocktail was incubated at 45ºC for 45 mins, followed by 30 thermal cycles at 94ºC for 30 s, 52-65ºC for 1 min, and 68ºC for 2 min for amplification. The gel electrophoresis was run in 1X TBE buffer under an electric field of 8 V/cm for 50 min. (A) Gel image of MMP-7 (B) GAPDH (C) β–Actin (D) B2M (E) Amplification efficiency v hybridization temperature. The band intensity was measured with Image Quant software from gel images.

be used as template to generate the standard curve. In practice, however, it is rare to directly use RNA as standards because they are susceptible to degradation and difficult to preserve for prolonged periods of time. It is more common to use plasmid DNA containing the cloned transcript of interest as the standard to generate a calibration curve. In this work, double-stranded PCR amplicons were used as standard templates to build the calibration curve. By knowing the efficiency of reverse transcription and assuming that the first strand of cDNA is fully reproduced into its double-stranded counterpart, the copy number of the target transcript can be derived using the calibration curve method that will be discussed subsequently.

5.3.2.1 RT Efficiency

In RT-PCR assay for mRNA analysis, the reverse transcription (RT) step is notorious for its high variability in terms of reaction efficiency.[35] Uncertainty in RT is caused by secondary and tertiary structure of the mRNA, and the properties of the reverse transcriptase, making it difficult to accurately generate cDNA.[36] RT efficiency ranging from 0.4% to 90% has been reported in the literature.[37] Therefore it is important to examine the efficiency of RT for accurate quantification of the copy number of the mRNA of interest. The reverse transcriptase used in this study was AffinityScript (Agilent Technologies, Santa Clara, CA), which is claimed to function at a relatively higher RT temperature (55ºC) with a higher RT efficiency compared to AMV reverse transcriptase.

The idea of deriving RT efficiency of reverse transcriptase is that if the RT efficiency is 100%, the mRNA standard containing a certain copy number will take only one more thermal cycle to reach the same fluorescence threshold than the dsDNA standard of the

same copy number. Any C_T difference that is >1 can be attributed to incomplete reverse transcription. In this method, $C_{T,lag}$ was defined as $C_{T,mRNA} - C_{T,dsDNA}$, where $C_{T,mRNA}$ was the threshold cycle using mRNA as input template, and $C_{T,dsDNA}$ was the threshold cycle using dsDNA of the same copy number as input template. RT efficiency can then be evaluated by the following equation that is derived based on the above reasoning.

$$RT\% = \frac{1}{2^{C_{T,lag}-1}} \qquad (5.1)$$

The mRNA standard originating from pAW109 plasmid DNA was obtained from Applied BioSystem(Foster City, CA), which contains 10^6 copies of mRNA per µL. In the first step of this approach, the mRNA standard, which was diluted to 10 – 10^6 copies in each reaction cocktail, was subjected to RT and 40 thermal cycles of PCR amplification and their C_T values were plotted against the input mRNA copy number with the results shown in Figure 5.8. In the second step, the pAW109 mRNA standard underwent conventional RT-PCR, and the dsDNA amplicons were purified and quantified by UV absorbance to get an accurate number of dsDNA templates. Then the high purity dsDNA was diluted to 10 – 10^6 copies in each reaction cocktail, and subjected to 40 thermal cycles of amplification using the same running conditions as those RT-qPCR reactions containing mRNA standard. The C_T values were also plotted against the input dsDNA copy number in Figure 5.8.

The C_T values of the RT-qPCR reactions with different copies of input mRNA are listed in Table 5.3, as well as the C_T values of the qPCR reactions with different copies of input dsDNA. Then the RT efficiency corresponding to each input mRNA copy number in the RT-qPCR is evaluated and listed in Table 5.3. We can see that it varies

from 45% to 93.3% in the range of the input mRNA copy number that were examined. To simplify further calculations, the RT efficiency of AffinityScript was taken as 76.18% and used throughout all studies in this work.

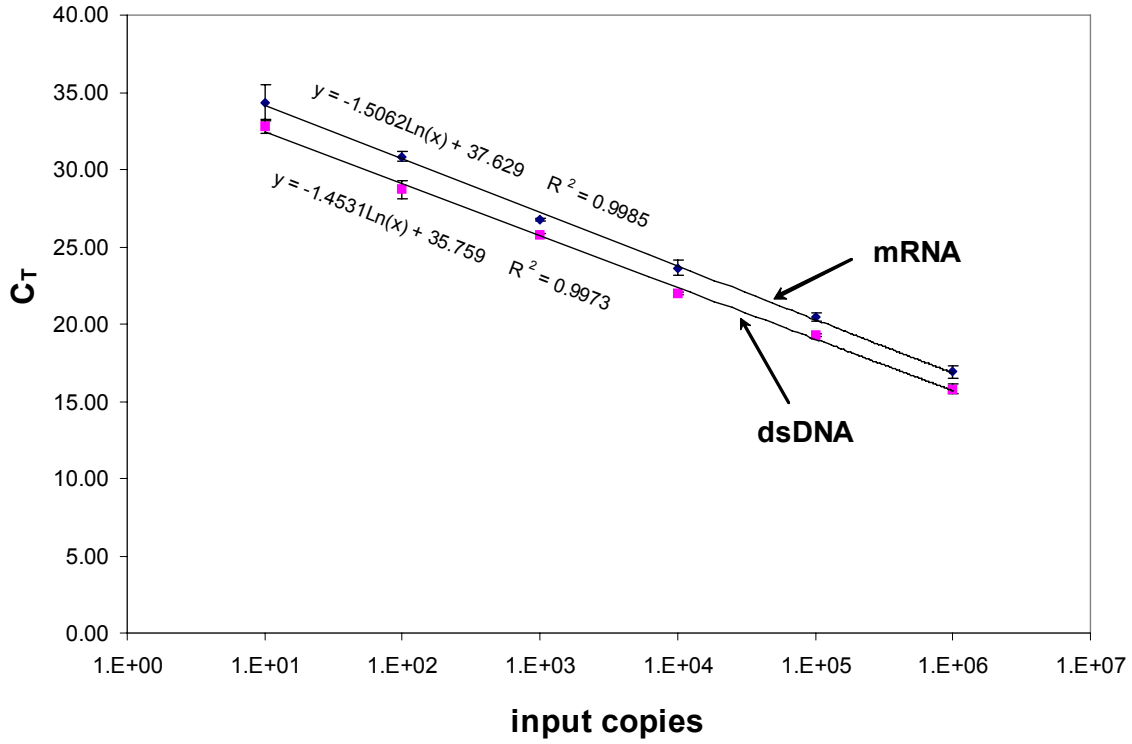


Figure 5.8 C_T of mRNA standard and dsDNA standard as a function of input copy number. Each RT-qPCR reaction tube contained 10 – 10^6 copies of mRNA or dsDNA standard. The reaction cocktail was incubated at 50ºC for 30 min, followed by 40 thermal cycles at 95ºC for 30 s, 60ºC for 1 min, 72ºC for 30 s.

Table 5.3 C_T values of RT-qPCR with different input copies of mRNA and dsDNA standard

input copies	C_T		RT efficiency, %
	mRNA standard	dsDNA standard	
10^6	16.91±0.38	15.81±0.30	93.30
10^5	20.45±0.28	19.26±0.11	87.26
10^4	23.64±0.47	21.98±0.06	63.00
10^3	26.76±0.10	25.73±0.14	98.40
10^2	30.85±0.28	28.70±0.56	44.96
10	34.33±1.13	32.81±0.47	70.14
	Average RT efficiency, %		76.18

5.3.2.2 Standard Curve for Quantification of MMP-7 Transcript

In order to quantify the MMP-7 mRNA expression level in different cell lines, a standard curve was first constructed using dsDNA standards obtained from conventional RT-PCR. The RT-PCR products were purified and quantified using UV absorbance to get the copy number of desired dsDNA. The reason dsDNA was used as standard is to avoid the cumbersome procedures involved in cloning approach as used by many expression analysis studies.[38] In qPCR and RT-qPCR assays, all experiments were run in replicates of three to verify their quality. In Figure 5.9 (A), we can see the real-time fluorescence signal is very reproducible for each input copy number, especially in the exponential stage where C_T is set. The fluorescence in the plateau stage has a much larger variation as discussed in an earlier section. The C_T values were obtained at the threshold fluorescence for each input copy. Then C_T is plotted against the log of initial template copy number in each diluted sample and fit to a linear function as shown in Figure 5.9 (B). The straight fitting line has a R^2 value of 0.997. From the slope of the fitting line, amplification efficiency was determined to be 106.7%, according to the following equation:[24]

$$PCR\% = 10^{(-1/SLOPE)} - 1 \quad (5.2)$$

Similar to end-point RT-PCR, negative controls were also used to ensure that the detected fluorescence from SYBR Green I was truly from the desired amplicons. In Figure 5.9 (A), we can see that the fluorescence from no-template control (NTC) and no-RT control (NRTC) were very weak, not exceeding the threshold fluorescence. This indicated that there is no gDNA contamination or other non-specific amplifications.

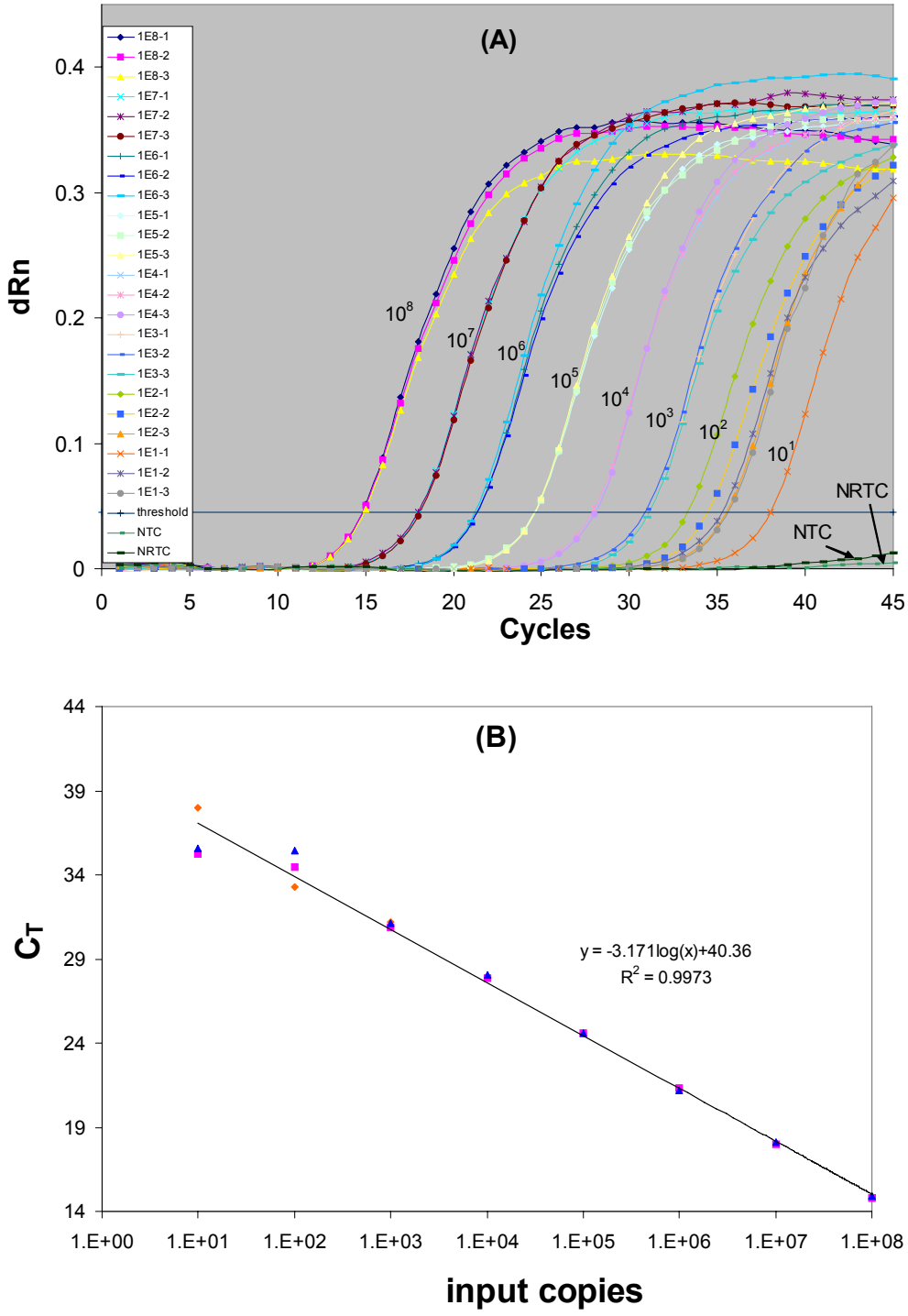


Figure 5.9 Standard curve for quantification of MMP-7 transcripts. (A) Raw fluorescence data. (B) Relationship of C_T with input copy number. 10 – 10^8 copies of dsDNA standard arising from HT-29 transcripts were used as the qPCR template. The reaction was incubated at 50ºC for 30 min, followed by 40 thermal cycles at 95ºC for 30 s, 59.4ºC for 1 min, 72ºC for 30 s. Each experiment was run in replicates of three.

In RT-qPCR assay, a dissociation curve or melting curve is usually performed to check for non-specific amplifications. Figure 5.10 shows the melting curve of the amplicons generated from the above experiment. The peak at 84ºC corresponded to the 395 bp long amplicons. In this plot, there were no other peaks identified. This suggested that the primers designed to amplify the MMP-7 transcript is highly specific to the target sequence and there were no non-specific amplification occurring.

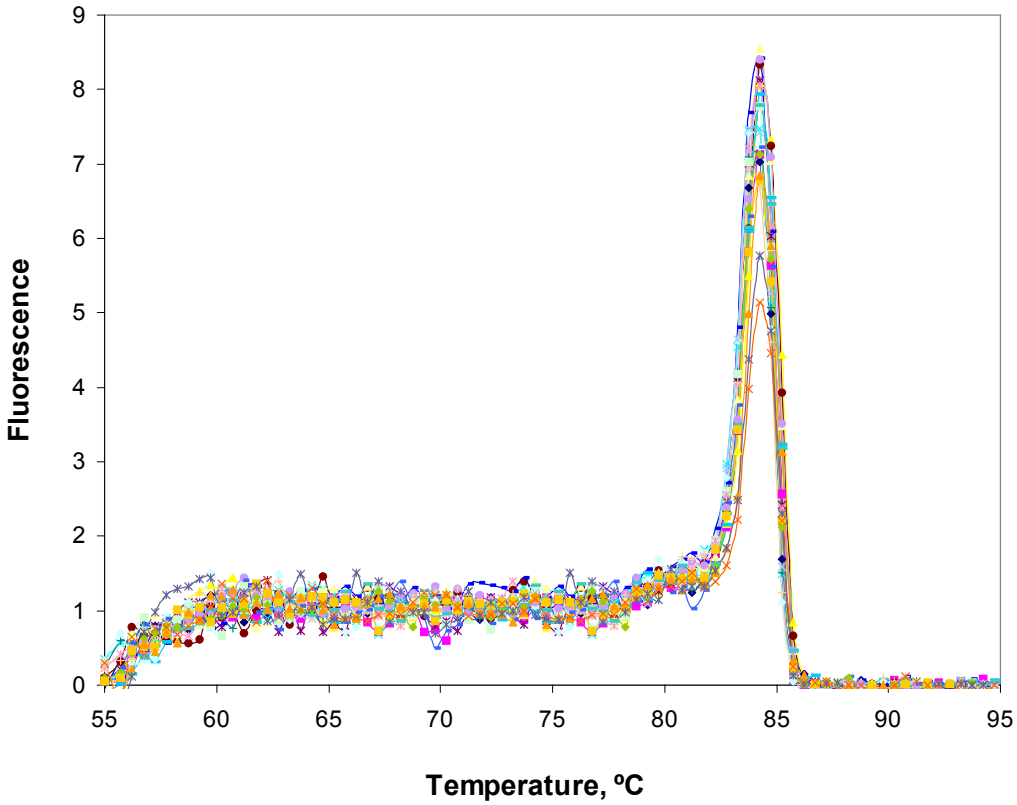


Figure 5.10 Dissociation curves for real-time RT-PCR using MMP-7 standard. The amplified products were incubated at 95ºC for 1 min, cooled down to 55ºC at a rate of 0.2ºC/sec, followed by 81 cycles of incubation where the temperature was increased by 0.5ºC/cycle, beginning at 55ºC and ending at 95ºC with a duration of 30 s for each cycle.

5.3.2.3 Quantification of mRNA Transcript Copy Number with Standard Curve

In RT-qPCR, the first thermal cycle after RT is always used to convert the reverse transcribed cDNA into the first dsDNA molecule. Because the standard curves

presented above were all constructed using a dsDNA standard, it is necessary to subtract 1 from the C_T value for each input template to compensate for this single thermal cycle conversion. This reduced C_T was then put back into the fitting function of the calibration curves to allow for the determination of the input copy number of cDNA. After correction for the RT efficiency, the initial mRNA copy number in the original sample could then be obtained.

The mRNA copy numbers of MMP-7 transcripts from different cell lines are displayed in Figure 5.11. One µg of total RNA extracted from all seven cell lines were used as template for the RT-qPCR. In panel (A), the RNA sample extracted from three consecutive batches of harvested cells was used for the examination. In panel (B), RNA extracted from the same cell line was split into three reaction mixtures to generate an intra-sample triplicate to monitor assay reproducibility.

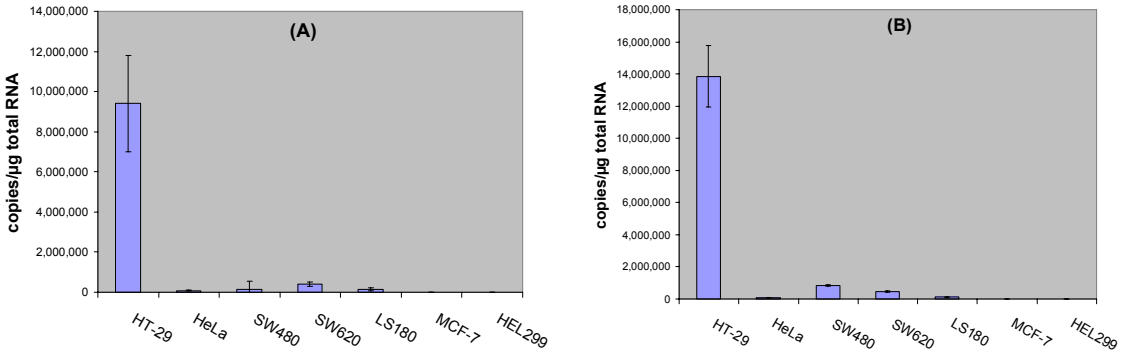


Figure 5.11 Copy numbers of MMP-7 mRNA in different cell lines. (A) Replicate runs of RT-qPCR using three groups of RNA samples extracted from cells harvested at different times. (B) Triplicate runs of RT-qPCR using the same RNA sample. One µg of total RNA was extracted from all seven cell lines and used as the input for RT-qPCR.

In this figure, it can be seen that HT-29 had a significantly higher MMP-7 expression level than the other cells used in this study. The other four colorectal cancer cells also showed different levels of MMP-7 expression, which were also seen from this figure. In contrast, MCF-7 and HEL299 have only 365 and 516 copies of MMP-7 transcript per µg

of total RNA, respectively. From comparisons between panel (A) and panel (B), it was also noted that for each cell line, the transcript copy number had a larger standard variation using RNA samples extracted from cells harvested at different times than that using the same RNA sample. The latter case stands for the artifact in RT-qPCR practice, such as sample preparation. The former case indicated that variations in cell culturing conditions can have an influence on expression levels of specific genes. To better compared the differences in MMP-7 expression level among different cell lines, the data in Figure 5.11 was further translated into transcript copy number per cell based on the total number of cells used for RNA extraction and the amount of RNA obtained from each cell line. The results are tabulated in Table 5.4. In this table, it is found that all cell lines associated with colorectal cancer (HT-29, HeLa, SW480, SW620, LS180) in this study have more than one copy of MMP-7 transcripts per cell, while HT-29 shows the highest expression level with 178 copies of MMP-7 transcripts per cell. This suggested that in these colorectal cancer cells, MMP-7 gene expression is up regulated, which agrees with that reported in the literature.[19, 21] By comparison, the average copy numbers of MMP-7 transcripts per cell in MCF-7 and HEL299 are much smaller than that found in the colorectal cancer cell lines.

Table 5.4 Copy number of MMP-7 transcripts per cell for different cell lines

Cell Line	average expression level copies/cell	standard deviation copies/cell
HT-29	178.397	24.553
HeLa	1.856	0.096
SW480	4.575	0.233
SW620	3.181	0.380
LS180	1.573	0.311
MCF-7	0.005	0.003
HEL299	0.009	0.002

The mRNA expression level of the housekeeping genes, described by copy number per µg of total RNA, were also quantified for each cell line using the standard curve method discussed above, and are shown in Figure 5.12. In this figure, we found that among all seven cell lines, the B2M gene showed the highest expression level among the three selected housekeeping genes, followed by GAPDH, and β–actin. In this figure, it is also noticed that B2M has a larger variability among the seven studied cell lines than the other two housekeeping genes.

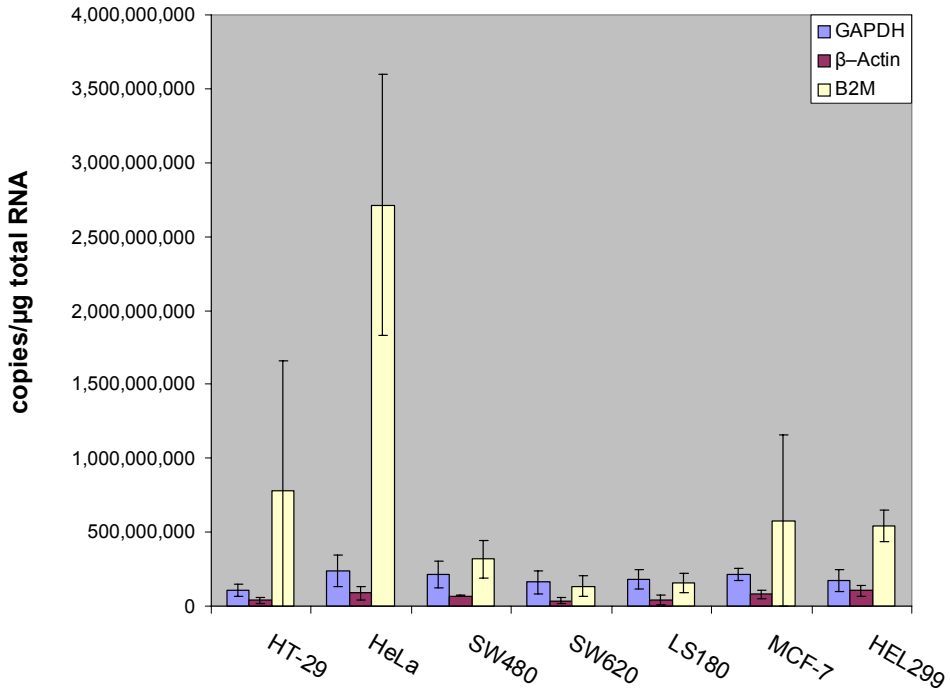


Figure 5.12 Expression levels of housekeeping genes in different cell lines. One µg of total RNA extracted from all seven cell lines were used as template for RT-qPCR.

Similar to the approach used for the MMP-7 quantification, the copy number for each housekeeping gene transcript based on one µg of total RNA was translated into a per cell basis and is shown in Table 5.5. In this table, the GAPDH transcript showed the smallest variability of expression among the seven cell lines as determined by its lower relative standard deviation (RSD) compared to the other genes, while B2M showed the

highest expression level variability. Thus GAPDH was selected to be most stable housekeeping gene among the studied cell lines and used for normalization throughout these studies.

Table 5.5 mRNA copy number of housekeeping genes per cell in different cell lines

Cell	mRNA expression level, copies/cell		
	GAPDH	β-actin	B2M
HT-29	1373±529	501±263	10094±11273
HeLa	4637±2074	1669±938	52324±17076
SW480	1178±480	370±25	1751±698
SW620	1096±533	247±156	913±490
LS180	2571±928	592±459	2188±878
MCF-7	3089±604	1167±406	8436±8475
HEL299	3077±1256	1854±700	9597±1873
average	2432	914	12186
standard deviatioin	1305	650	18130
RSD	53.65%	71.06%	148.77%

5.4 Conclusions

In this chapter, the expression level of the MMP-7 transcript, as well as several commonly used housekeeping genes were investigated using conventional end-point RT-PCR and RT-qPCR with real-time readout. In contrast to end-point RT-PCR, which only provided semi-quantitative information about the expression abundance of specific transcripts, RT-qPCR can accurately measure the copy number of the target transcript. The annealing temperatures for RT-PCR were optimized for each transcript and primer pairs to ensure that the highest efficiency of amplification and sensitivity of analysis are achieved during the real-time RT-qPCR measurement. Efficiency of reverse transcription was also evaluated using an innovative experimental design. In this study, HT-29 showed the highest level of MMP-7 expression of those cell lines studied, indicating up-regulation of MMP-7 gene activity with respect to the other cell lines. Of all

the three selected housekeeping gene candidates evaluated, GAPDH showed the most stable expression level among all seven used cell lines and can serve as an internal control to normalize for artifacts in experimental variations.

5.5 References

(1) Liu, Y. C.; Ringner, M. *Genome Biology* 2007, *8*.

(2) Evans, W. E.; Guy, R. K. *Nature Genetics* 2004, *36*, 214-215.

(3) Sorlie, T.; Tibshirani, R.; Parker, J.; Hastie, T.; Marron, J. S.; Nobel, A.; Deng, S.; Johnsen, H.; Pesich, R.; Geisler, S.; Demeter, J.; Perou, C. M.; Lonning, P. E.; Brown, P. O.; Borresen-Dale, A. L.; Botstein, D. *Proceedings of the National Academy of Sciences of the United States of America* 2003, *100*, 8418-8423.

(4) Barrier, A.; Lemoine, A.; Brault, D.; Houry, S.; Flahault, A.; Dudoit, S. *Clinical Cancer Research* 2005, *11*, 8999s-8999s.

(5) Barrier, A.; Lemoine, A.; Boelle, P. Y.; Tse, C.; Brault, D.; Chiappini, F.; Breittschneider, J.; Lacaine, F.; Houry, S.; Huguier, M.; Van der Laan, M. J.; Speed, T.; Debuire, B.; Flahault, A.; Dudoit, S. *Oncogene* 2005, *24*, 6155-6164.

(6) Gygi, S. P.; Corthals, G. L.; Zhang, Y.; Rochon, Y.; Aebersold, R. *Proceedings of the National Academy of Sciences of the United States of America* 2000, *97*, 9390-9395.

(7) Chelius, D.; Bondarenko, P. V. *Journal of Proteome Research* 2002, *1*, 317-323.

(8) Sydor, J. R.; Nock, S. *Proteome Sci* 2003, *1*, 3.

(9) Fu, N.; Drinnenberg, I.; Kelso, J.; Wu, J. R.; Paabo, S.; Zeng, R.; Khaitovich, P. *Plos One* 2007, *2*.

(10) Greenbaum, D.; Colangelo, C.; Williams, K.; Gerstein, M. *Genome Biology* 2003, *4*.

(11) Gygi, S. P.; Rochon, Y.; Franza, B. R.; Aebersold, R. *Molecular and Cellular Biology* 1999, *19*, 1720-1730.

(12) Lu, P.; Vogel, C.; Wang, R.; Yao, X.; Marcotte, E. M. *Nature Biotechnology* 2007, *25*, 117-124.

(13) Ferlay, J.; Shin, H. R.; Bray, F.; Forman, D.; Mathers, C.; Parkin, D. M. *International Journal of Cancer* 2010, *127*, 2893-2917.

(14) Mimori, K.; Yamashita, K.; Ohta, M.; Yoshinaga, K.; Ishikawa, K.; Ishii, H.; Utsunomiya, T.; Barnard, G. F.; Inoue, H.; Mori, M. *Clinical Cancer Research* 2004, *10*, 8243-8249.

(15) Rome, C.; Arsaut, J.; Taris, C.; Couillaud, F.; Loiseau, H. *Molecular Carcinogenesis* 2007, *46*, 446-452.

(16) Muller, D.; Wolf, C.; Abecassis, J.; Millon, R.; Engelmann, A.; Bronner, G.; Rouyer, N.; Rio, M. C.; Eber, M.; Methlin, G.; Chambon, P.; Basset, P. *Cancer Research* 1993, *53*, 165-169.

(17) Muller, D.; Breathnach, R.; Engelmann, A.; Millon, R.; Bronner, G.; Flesch, H.; Dumont, P.; Eber, M.; Abecassis, J. *International Journal of Cancer* 1991, *48*, 550-556.

(18) Mcdonnell, S.; Navre, M.; Coffey, R. J.; Matrisian, L. M. *Molecular Carcinogenesis* 1991, *4*, 527-533.

(19) Pajouh, M. S.; Nagle, R. B.; Breathnach, R.; Finch, J. S.; Brawer, M. K.; Bowden, G. T. *Journal of Cancer Research and Clinical Oncology* 1991, *117*, 144-150.

(20) Fukushima, H.; Yamamoto, H.; Itoh, F.; Nakamura, H.; Min, Y. F.; Horiuchi, S.; Iku, S.; Sasaki, S.; Imai, K. *Carcinogenesis* 2001, *22*, 1049-1052.

(21) Luo, H. Z.; Zhou, Z. G.; Yang, L.; Yu, Y. Y.; Tian, C.; Zhou, B.; Zheng, X. L.; Xia, Q. J.; Li, Y.; Wang, R. *Japanese Journal of Clinical Oncology* 2005, *35*, 739-744.

(22) Basset, P.; Bellocq, J. P.; Wolf, C.; Stoll, I.; Hutin, P.; Limacher, J. M.; Podhajcer, O. L.; Chenard, M. P.; Rio, M. C.; Chambon, P. *Nature* 1990, *348*, 699-704.

(23) Witty, J. P.; Mcdonnell, S.; Newell, K. J.; Cannon, P.; Navre, M.; Tressler, R. J.; Matrisian, L. M. *Cancer Research* 1994, *54*, 4805-4812.

(24) Wong, M. L.; Medrano, J. F. *Biotechniques* 2005, *39*, 75-85.

(25) VanGuilder, H. D.; Vrana, K. E.; Freeman, W. M. *Biotechniques* 2008, *44*, 619-626.

(26) Stratagene *Introduction to Quantitative PCR: Methods and Applilcations Guide*, 2007.

(27) Higuchi, R.; Fockler, C.; Dollinger, G.; Watson, R. *Bio-Technology* 1993, *11*, 1026-1030.

(28) Liu, W. H.; Saint, D. A. *Analytical Biochemistry* 2002, *302*, 52-59.

(29) Dheda, K.; Huggett, J. F.; Bustin, S. A.; Johnson, M. A.; Rook, G.; Zumla, A. *Biotechniques* 2004, *37*, 112-114, 116, 118-119.

(30) Barber, R. D.; Harmer, D. W.; Coleman, R. A.; Clark, B. J. *Physiological Genomics* 2005, *21*, 389-395.

(31) Kreuzer, K. A.; Lass, U.; Landt, O.; Nitsche, A.; Laser, J.; Ellerbrok, H.; Pauli, G.; Huhn, D.; Schmidt, C. A. *Clin Chem* 1999, *45*, 297-300.

(32) Winer, J.; Jung, C. K.; Shackel, I.; Williams, P. M. *Anal Biochem* 1999, *270*, 41-49.

(33) Dydensborg, A. B.; Herring, E.; Auclair, J.; Tremblay, E.; Beaulieu, J. F. *Am J Physiol Gastrointest Liver Physiol* 2006, *290*, G1067-1074.

(34) Bustin, S. A. *Journal of Molecular Endocrinology* 2000, *25*, 169-193.

(35) Mannhalter, C.; Koizar, D.; Mitterbauer, G. *Clinical Chemistry and Laboratory Medicine* 2000, *38*, 171-177.

(36) Stahlberg, A.; Hakansson, J.; Xian, X. J.; Semb, H.; Kubista, M. *Clinical Chemistry* 2004, *50*, 509-515.

(37) Stahlberg, A.; Kubista, M.; Pfaffl, M. *Clinical Chemistry* 2004, *50*, 1678-1680.

(38) Bartlett, J. M. *Methods Mol Med* 2001, *39*, 399-409.

CHAPTER 6 SINGLE-MOLECULE QUANTIFICATION OF mRNA TRANSCRIPTS ON A CONTINUOUS FLOW COC MICROFLUIDIC REACTOR FOR GENE EXPRESSION PROFILING

6.1 Introduction

Quantification of mRNA transcripts is an important approach for gene expression analysis and has far-reaching implications for diagnosis of various disease conditions.[1, 2] Conventionally, mRNA molecules are analyzed and quantified using Northern blotting, ribonuclease protection assays (RPA), in-situ hybridization (ISH), reverse transcriptase PCR (RT-PCR), and cDNA microarray techniques with different readout modalities such as fluorescence, electrochemical, radioactivity or direct visualization.

In many molecular biology labs, Northern blotting still remains as a standard method and is often used as a confirmation in many applications.[3] It can be used to determine the size of the transcript and allow comparison of mRNA abundance from different samples on a single membrane. However, Northern analysis generally requires significant amounts of mRNA transcripts due to its low sensitivity. RPA is a solution-based hybridization technique for expression analysis and can be used to detect mRNAs that are expressed at low levels. ISH is a powerful method especially useful for localizing and detection the expression of mRNA sequences in specific regions of cells or within morphologically preserved tissue sections by hybridizing a complementary oligonucleotide strand to the mRNA sequence of interest.[4] These methods, however, involves many laborious steps to get the desired result, which makes the entire assays time-consuming and not amenable to examination of many different transcripts in a high-throughput manner.[5] In addition, they require fair amounts of mRNA for the analysis.

cDNA microarrays are featured by the ability to simultaneously analyze a large number of gene transcripts in parallel due to its high spatial resolution and the capability of automation, making cDNA microarrays the method of choice for expression profiling on a global scale, which is essential in unraveling complex cellular signal pathways to explore the underlying mechanism of certain diseases.[6] But, this technique is limited by its low sensitivity and accuracy when quantitative measurements of mRNA expression are desired.[7, 8]

Real-time RT-PCR is a much more efficient method for expression analysis and is characterized by a aster turn-around time and low limit of detection compared to the aforementioned techniques. It has been regarded as the most sensitive method for mRNA expression analysis with a theoretical limit of detection of one copy of the target transcript.[9] However, the extreme sensitivity of real-time RT-PCR primarily due to its exponential amplification of target sequence also subjects it to experimental variations because even minute errors introduced at the early stages of the experiment will be magnified with increasing PCR cycles. Additionally the accuracy of the results is greatly affected by sample preparation, variation in reagents and the skills of the operator.[10] More importantly, in practice, real-time RT-PCR requires construction of calibration curves with known concentrations of pure mRNA each time a new expression transcript needs to be assayed, which is both time and reagent consuming.

In addition to the techniques for mRNA expression analysis mentioned above, there are various PCR-free techniques that have been developed to emulate the sensitivity of real-time RT-PCR, while avoiding its susceptible amplification steps to errors. Medley *et al.* utilized dual-labeled molecular beacon probes to monitor the expression of multiple

gene transcripts in living breast carcinoma cells.[11] Chen *et al.* reported an electrical sensor array combined with a two-step hybridization for mRNA analysis.[12] In their scheme, the sensor array was fabricated by overlaying two layers of Au electrodes on a silicon wafer using standard photolithographic techniques and chemical vapor deposition with the gap between the electrodes controlled within 10 to 100 nm. A pair of hybridization probes specific to the target mRNA transcripts were used to capture the target mRNA and form a bridge between the two layers of electrods. This oligonucleotide bridge was then coated with silver to make it electrically conductive. Noticeable conductance changes were observed for as little as 0.1 fM of mRNA. Lin and co-workers adopted another PCR-free scheme using a branched-DNA assay (bDNA) for mRNA quantification.[13] In contrast to multiplication of target mRNA sequences as used in real-time RT-PCR, which inevitably biased the quantity of target gene in the original sample, the bDNA assay made detectable minute quantities of target mRNA by simply enhancing the signal of detection using multiply labeled oligonucleotide probes. This technique has been reported to detect 1 fM target mRNA in p185 BCR-ABL leukemia fusion transcripts.

Aside from the various techniques described above, single-molecule detection techniques have also been used for monitoring expression levels of mRNAs and can offer extraordinary sensitivity without requiring PCR. Neely *et al.* reported a laser-induced fluorescence SMD system for the quantification of microRNA (miRNA) expression, which is used for regulating fundamental cellular processes.[14] In this approach, a four-color LIF instrument was used to count the coincident events arising from fluorescently labeled LNA-DNA probes hybridized to target miRNA molecules in

solution. The event numbers were directly correlated to the number of miRNA molecules in complex biological samples and as low as 500 fM of miRNA were successfully detected. Rigler and co-workers developed a hybridization assay containing two oligonucleotide probes for gene expression level measurements.[15] Simultaneous hybridization of the dye labeled probes to the target gene caused enhancement of cross-correlation features, which were measured using a two-color fluorescence correlation spectroscopy (FCS) single-molecule detection system. The magnitude of the correlation function was associated with the concentration of target genes. It was noted, however, that in these two single-molecule detection systems mentioned above, dual-color excitation and detection methods were needed, which is challenged by complicated instrumental setups and optical alignment. In addition, these techniques require hybridization-based recognition, which can be challenging when single-base variations must be monitored due to the required stringent hybridization conditions including tight control on temperature and salt conditions. Finally, due to the thermodynamic nature of hybridization, high excess copy numbers of interfering sequences can make detection of the target problematic when a minority is in a mixed population. These observations can give rise to extensive amounts of false positive results.

Herein, we report a unique single-molecule detection scheme combined with reverse transcription and the ligase detection reaction (RT-LDR) to effectively count and quantify mRNA transcripts using digital techniques (*i.e.* molecular counting). The strategy of this assay is illustrated in Figure 6.1. A similar scheme has been adopted for detecting single nucleotide mutations in K-ras gene and strain specific differentiation of

bacterial pathogens, where the primer pairs were designed to directly identify the genomic DNA of the target species using LDR and spFRET readout.[16, 17]

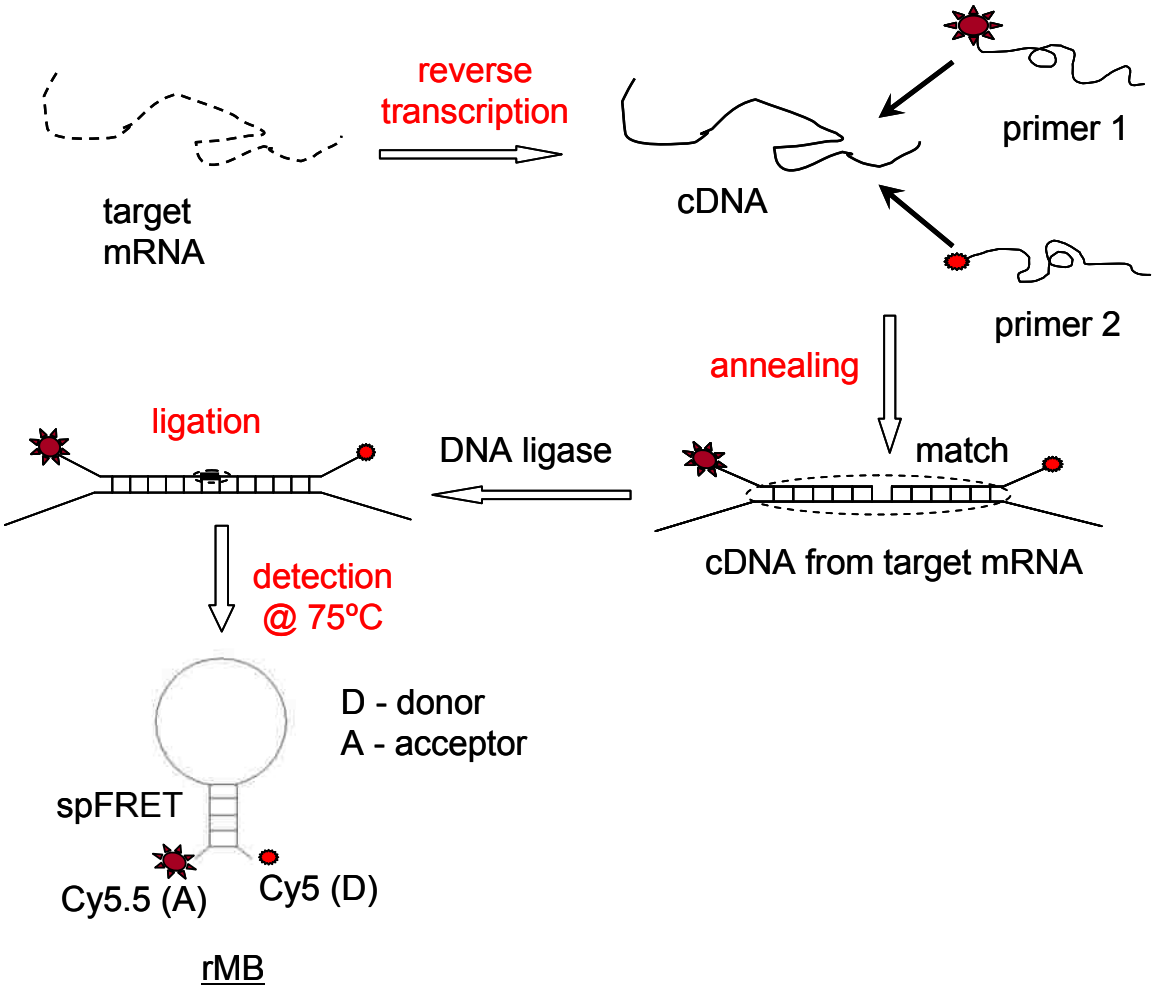


Figure 6.1 Illustration of RT-LDR scheme coupled to spFRET readout format for gene expression analysis of mRNA transcripts

In this assay, the mRNA molecules of interest are first converted into their complementary DNA sequences by reverse transcription. A pair of hybridization probes is designed to flank two adjacent fragment sequences located in a reporter region of this transcript, which consists of a span of its sequence that does not appear in other mRNA transcripts. Each of these probes also contains a 10-base arm sequence that is complementary to each other while not complementary to the target sequence. The end

of each arm sequence is attached to a donor dye and an acceptor dye for fluorescence readout using fluorescence resonance energy transfer (FRET) only when the arm sequence are hybridized. Successful ligation of the LDR probes will occur only if the target sequence is complementary to both LDR probes, resulting in the formation of a molecular beacon (MB). Thus, the donor and acceptor dyes are brought into close proximity resulting in a FRET signal. The arm sequences of the rMB are designed to possess a higher melting temperature (T_m) by incorporating a GC-rich sequence and thus, thermodynamically more stable than the target-oligonucleotide duplex. Therefore, the rMB adopts a stable stem-loop conformation following ligation. The FRET signal from individual rMBs is reported by a LIF single-molecule detection instrument. LDR thermal cycling and single-molecule readout are performed directly on a thermoplastic microfluidic device to provide near real time assay results.

The exceptional fidelity of LDR in recognizing matched targets even in the presence of great excesses of mismatched substrates can provide high specificity of the assay. This fidelity is derived from: 1) LDR requires two concurrent hybridization events. Therefore, primers can be designed that target a unique reporter sequence within the target mRNA transcript, negating the need for single-base discrimination. 2) DNA ligase is intolerant of a 3'-end mismatch. The outstanding specificity of LDR allows it to discriminate a single base mismatch.[18] This specificity is especially valuable for differentiation of miRNAs because some human miRNA members differ by only one or two nucleotide bases.[19] In addition, there is only one color that is required in this approach for the detection of fluorescently labeled probes, which greatly facilitates simplification of the instrumentation. Also, because there is no PCR involved, the copy

number of mRNA transcripts can be directly counted in the measurement without any upfront calibration, as required in PCR.

In recent years colorectal cancer has come to public awareness as a serious health problem in developed countries and early diagnosis of this disease is playing a crucial role in improving patient survival by allowing timely treatment.[20] MMP-7 has been identified to be an important biomarker of colorectal cancer, and it has been extensively reported that over-expression of this gene has strong implications with invasion, progression, and metastasis.[21-23] In this study, the mRNA molecules of MMP-7 will be extracted from cells originating from colorectal carcinomas with the number of MMP-7 transcripts accurately quantified using RT-LDR assay. The results will be compared to the MMP-7 transcriptional level in human normal cells. In this application demonstrating the utility of RT-LDR, quantification of low-copy number transcripts will be demonstrated with comparisons to conventional RT-PCR.

Another promising application of RT-LDR is focused on expression profiling of stroke biomarkers. Stroke has become the third leading cause of death and disability in the United States with ~795,000 new and recurrent cases reported each year in the US alone.[24] For proper treatment, stroke patients have to be given therapeutic treatment within the very first few hours of onset of a stroke event. Clinically, stroke is classified into two forms: (1) ischemic stroke, which is due to occlusion of arterial vascular network and accounts for 83% of the total number of stroke cases; and (2) hemorrhagic stroke, which is due to vascular rupture and accounts for 17% of the total stroke cases.[25, 26] Currently, stroke patients need to be hospitalized to determine the type of stroke with imaging techniques such as computed tomography (CT) or magnetic

resonance imaging (MRI) for proper diagnosis. This is because the efficacious treatment of ischemic stroke through intravenous injection of blood clot dissolving medicine, such as tissue plasminogen activator, is contraindicated in hemorrhagic stroke.[27] Thus, fast diagnosis is critical to allow prompt treatment after stroke has occurred.

Monitoring changes in expression levels of functional genes has been used for the diagnosis of stroke as well as other brain disorders. Moore *et al.* conducted a pilot-scale gene expression profiling using oligonucleotide microarrays of the peripheral blood mononuclear cells (PBMCs) from 20 ischemic stroke patients and investigated the expression level of a panel of 22 genes.[28] Recently, Baird discovered two genes, amphiphysin (AMPH) and IL1R2, which were expressed differentially among hemorrhagic and ischemic stroke patients.[29] Unfortunately, the extended assay turn-around-time and sophisticated nature of both RT-PCR and cDNA microarrays make these assay strategies prohibitive for field-monitoring of mRNA transcripts for diagnostics. Therefore, the use of RT-LDR will be demonstrated as a viable assay strategy for the diagnosis of stroke with short assay turn-around-time (*i.e.*, time-sensitive assay) to allow for the proper management of stroke patients.

6.2 Experimental

6.2.1 Materials and Reagents

The oligonucleotide primers for LDR and RT-PCR were synthesized by Integrated DNA Technologies (Coralville, IA), purified by RP-HPLC and resuspended in 1X TE buffer. AmpliTaq Gold polymerase was purchased from Applied Biosystems (Foster City, CA). Taq DNA ligase was purchased from New England BioLabs (Ipswich, MA).

Cyclic olefin co-polymer, COC, was purchased from Topas Advanced Polymers (Florence, Kentucky).

6.2.2 On-chip LDR

Before making single-molecule measurements of MBs generated from cDNA of the target transcript, the LDR needed to be evaluated with respect to its amenability to being carried out using COC microchip in a continuous flow format to allow for thermal cycling. For these evaluative experiments, a primary RT-PCR was conducted to provide enough templates material for the on-chip LDR in order to provide sufficient products so that this assay step could be verified using capillary gel electrophoresis. But, in the single-molecule measurements of the transcript quantification, no RT-PCR was used.

The LDR cocktail consisted of 2 units/μL of thermostable DNA ligase, 20 mM Tris-HCl (pH 7.6), 25 mM potassium acetate, 10 mM magnesium acetate, 1 mM NAD^+ cofactor, 10 mM dithiothreitol, 0.1% Triton X-100, 25 nM of each LDR primer, 0.5 μL of DNA template and nuclease-free H_2O. Ultrapure bovine serum albumin (BSA, 0.5 mg/mL) was included in the reaction mixture to minimize non-specific adsorption of the ligase enzyme onto the thermal reactor surfaces.[30, 31] Kapton heating tape was attached directly to the bottom of the microchip to provide the required temperatures for on-chip LDR. The reaction mixture was loaded into a glass syringe, and driven continuously through the serpentine microchannel by a syringe pump at a flow rate of 0.1-8 μL/min. In the micro-reactor, the LDR mixture was pre-heated to 94ºC, then subjected to 2-20 thermal cycles by alternately flowing through a 94ºC zones for denaturation and 65ºC zone for ligation. The resulting LDR products were collected into a PCR microtube for capillary gel electrophoresis analysis.

6.2.3 LDR Primer Design

Considerations in designing LDR primers for expression analysis of MMP-7 and AMPH transcripts are similar to those in designing primers for RT-PCR discussed in Chapter 5. Briefly, residual genomic DNA or unspliced RNA tends to persist in many RNA preparations. These contaminants can also serve as templates for LDR after reverse transcription and lead to false positive results. Therefore the LDR primer pairs should be derived from two adjacent exons or at least one of them has to span an exon-exon junction in the target sequence. In either case, ligation cannot occur on gDNA or unspliced RNA reverse transcripts because the intron between adjacent exons will set the two LDR primers apart and prevent ligation. Thus, a successful ligation event indicated by a spFRET signal will exclude both gDNA contamination and unspliced transcripts with this design strategy invoked. The LDR primers designed are listed in Table 6.1.

Table 6.1 LDR primers for MMP-7 and AMPH transcripts

LDR primers		sequence (5'-3')*
MMP-7	upstream	Cy5.5-C3-AGGCGGCGCGTCCAAAGTGGTCACCTACAGGAT
	downstream	p[a]CGTATCATATACTCGAGACTTACCCGCGCCGCCT-C3-Cy5
AMPH	upstream	Cy5.5-C3-AGGCGGCGCGCAGTGTGACAACACCTTCCCAG
	downstream	pAATGAAGTCCCTGAGGTGAAGAAACGCGCCGCCT-C3-Cy5

[a] p, phosphorylation.
* The underlined sequence consists of the stem of the rMB, which is formed following ligation. A 3-carbon linker was used to attach the donor or acceptor to the oligonucleotides.

6.2.4 Fabrication of Continuous Flow Microfluidic Reactor

The designed microstructures were first patterned on a brass molding tool with a high-precision micromilling machine (Kern MMP 2522, Murnau, Germany). The layout of the microfluidic reactor is shown in Figure 6.2.

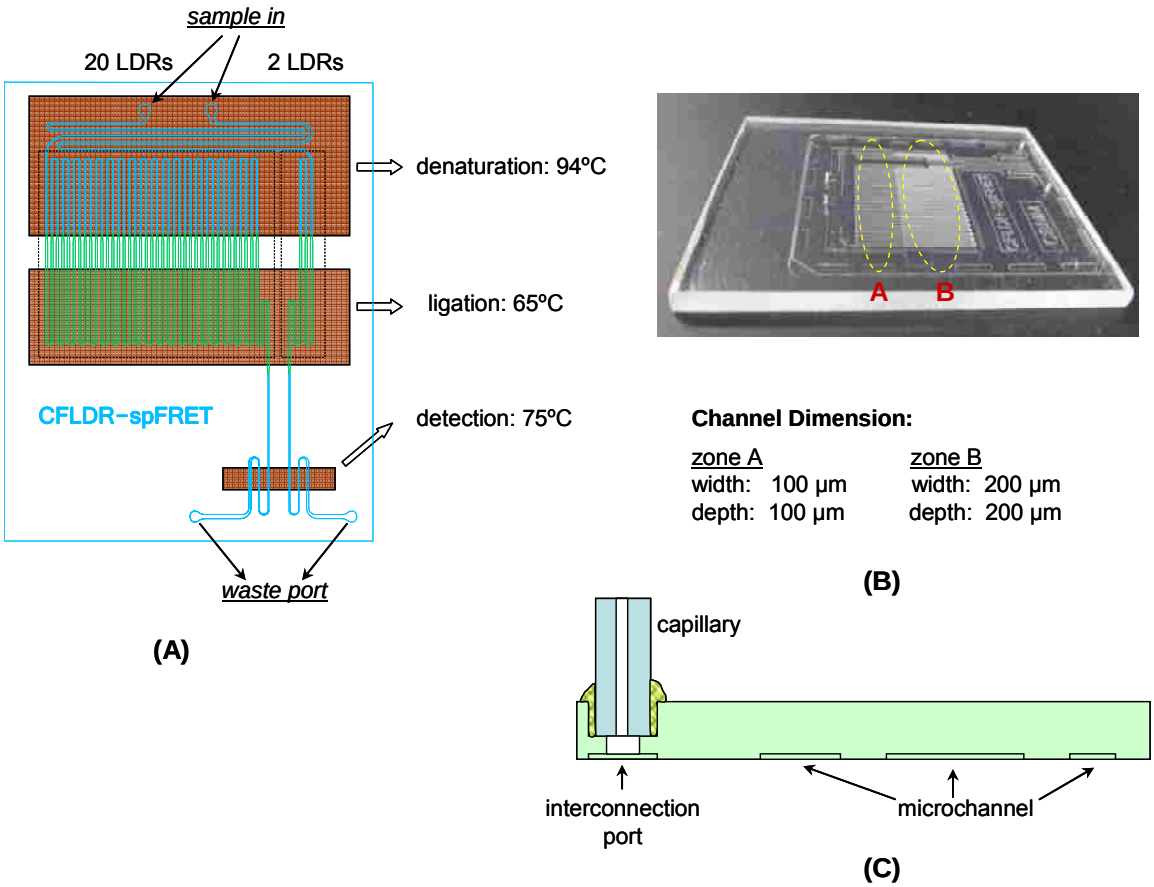


Figure 6.2 Layout of the microchannels on the hot-embossed COC chip. (A) Schematic diagram of the microfluidic chip. (B) Photographic image of the chip. (C) Illustration for the attachment of a capillary to the chip.

To make the plastic chip, the molding tool was used to replicate structures into a polymer substrate using a hot-embossing machine (PHI Precision Press TS-21-H-C, City of Industry, CA). COC (5013L-10, Tg = 130ºC, Topas Advanced Polymer, Florence, KY) was chosen to be the substrate for the microfluidic reactor due to its capability to sustain high operating temperatures (95ºC) during the denaturation step for LDR thermal cycling and its exceptionally low autofluorescence, which is critical in single-molecule measurements. After hot-embossing, holes were drilled into the molded chip to facilitate connection to external tubing. The finished COC substrate was then sonicated in ddH_2O for 20 min to remove contaminants. A COC cover plate was

thermally fusion bonded to the embossed chip by sandwiching the cover plate-substrate assembly between two glass plates in a convection oven and maintaining the temperature at 132ºC for 20 min. After cooling, the assembled chip was rinsed thoroughly with ddH_2O.

The chip contained two different devices both of which performed LDRs in a continuous flow format.[32] One device contained a continuous flow thermal reactor to carry out 20 LDR thermal cycles while the other one consisted of 2 LDR thermal cycles. In both reactors, the cross section of the microchannels was 100 x 100 μm at the 94ºC zone, and 200 x 200 μm at the 65ºC zone. The larger cross section in the 65ºC zone allowed extended residence time in this zone to accommodate the kinetics of the ligation process. The detection window was located at the end of each continuous flow thermal cycler for single-molecule observation of the spFRET signals generated from the MBs.

To allow connection of the COC chip to an external syringe pump, a through hole (300 μm O.D.) was drilled from the patterned side of the substrate, and another larger hole (460 μm O.D.) was drilled half-way through the substrate from the opposite side. Thus, a fused silica capillary (365 μm O.D.; 100 μm O.D., Polymicro Technologies, Phenix, AZ) could fit tightly into the larger hole and stopped by the shoulder of the smaller hole creating minimal unswept volumes (Figure 6.2C). After insertion of the capillary into the guide holes, the capillary was sealed to the chip using epoxy glue.

Another concern in making the microfluidic device for single-molecule measurements is selection of the appropriate cover plate for the chip. In the LIF single-molecule detection instrument, a microscope objective with high magnification power

was chosen to obtain higher photon collection efficiency. This objective, however, possesses a very short working distance (*e.g.* the working distance of a 60X objective (M-60X, Newport, Irvine, CA) is 0.3 mm). Thus, the thickness of the cover plate must be less than the objective's effective working distance. On the other hand, the cover plate needs to be thick enough to withstand the hydraulic pressure in the microfluidic channels when the sample was driven by a syringe pump. A 125 μm cover plate was tested and found to be adequate for the above mentioned operating conditions.

6.2.5 Operation of the Chip

The capillary attached to the chip was connected to a glass syringe (SGE, Australia) via a syringe-to-capillary adapter (InnovaQuartz, Phoenix, AZ). The LDR cocktail was first run through a 0.2 μm filter to remove any large particulates. Then, the reaction mixture was loaded into a glass syringe and was driven by a syringe pump (Pico Plus, Harvard Apparatus, Holliston, MA) through the thermal reactor's microchannels. Thin-film Kapton heaters were placed at the appropriate zones of the chip to provide the desired temperatures for the LDR and spFRET measurement. The denaturation and renaturation/ligation zones are shown in Figure 6.2A. A 3.5 mm gap was situated between the denaturation and renaturation/ligation zones to minimize any thermal crosstalk. The volume flow rate was set between 0.1- 8 μL/min depending on the required ligation time.

6.2.6 LIF Single-Molecule Detection Instrumentation

The instrument used for counting MBs arising from LDR has been described elsewhere.[17] A picture of the LIF single-molecule detection system is illustrated in Figure 6.3.

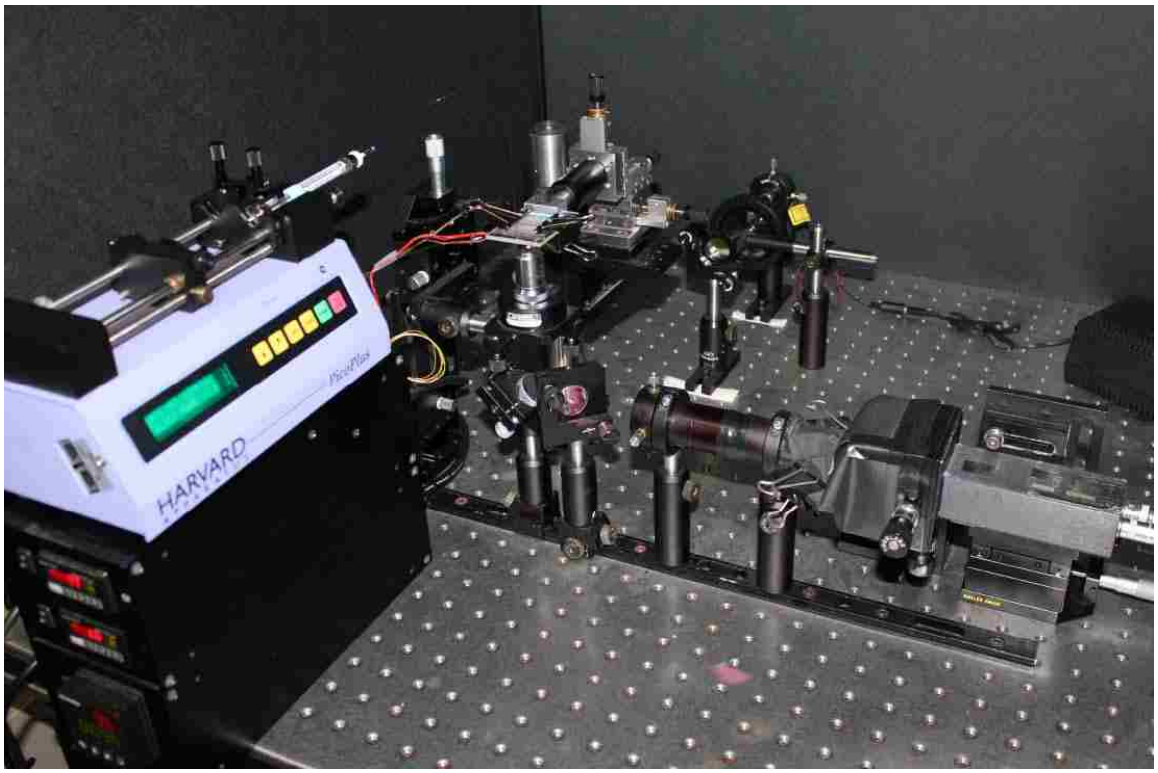


Figure 6.3 Layout of the instrument for LIF single-molecule detection

In this system the excitation source consists of a 635 nm diode laser (Model CPS196, Thorlabs, Newton, NJ), a laser line filter (640DF20, Omega, Brattleboro, VT), and a neutral density filter. The laser diode outputs an elliptical laser beam (2.45X0.54 mm), which was first collimated by its self-configured head lens and then conditioned by the laser line filter. The power impinging on the chip was set by a neutral density filter and fixed at 1 mW, which has been optimized in Chapter 3 to give the best signal-to-background ratio for single-molecule measurements. The laser beam was then directed into a microscope objective (M-60X, NA = 0.85, Newport, Irvine, CA) by a dichroic mirror (690DRLP, Omega Optical). The collimated laser was focused by this objective into the microfluidic channel positioned on the top of the objective with the shorter axis of the elliptical beam parallel to the width of the microchannel to produce a bigger illumination

cross section within the channel. When the fluorescently labeled MBs flowed through the channel, they were excited and the emitted fluorescence was then collected by the same objective and transmitted through a dichroic and pinhole (i.d. = 100 μm), which served as a spatial filter to remove out-of-focus fluorescence. A combination of interference filters, including a longpass filter (3RD690LP, Omega Optical) and a bandpass filter (HQ710/20m, Chroma Technology, Rockingham, VT) were placed after the pinhole to spectrally filter the fluorescence light. The fluorescence was finally focused onto the active area of a single photon avalanche diode (SPCM-200, EG&G, Vandreuil, Canada) using a 10X microscope objective. The signal from the SPAD was transformed into a TTL pulse and processed using a digital counting board (PCI-6602, National Instruments, Austin, TX).

Single-molecule fluorescence measurements were performed by registering the arrival time of each incoming photon to the SPAD. Single molecule fluorescence measurements were usually performed by recording the number of photons collected in a predefined interval of time. This method, however, carry less information than the photon arrival time approach, especially when smaller time intervals are needed after data acquisition. The time-of-arrival photon registration approach is an exquisite technique and provides the ultimate solution for photon counting application. A PCI-6602 data acquisition board provided 12.5 ns time resolution at its maximum clock frequency of 80 MHz, which can differentiate any two photons that arrive at the SPAD beyond this interval and alleviate pulse pileup. Also, it provided great flexibility to bin the photon counts into any user defined intervals, after data acquisition was completed.

6.2.7 Contact Angle Measurement

Sessile water contact angles were acquired using a VC 2000 contact angle system equipped with a CCD camera (VCA, Billerica, MA). For each measurement, 2 μL of ddH_2O (18MΩ·cm) was dispensed onto the polymer surface by a precision syringe equipped with the instrument. The left and right contact angles of the water droplet were measured immediately using software provided by the manufacturer. An average of at least five measurements on separate positions was reported for each given substrate.

6.3 Results and Discussions

6.3.1 Surface Passivation of the COC Substrate

The efficiency of LDR performed on-chip often suffers from enzyme deactivation caused by non-specific interactions between the inner surface of the microchannels and biomolecules.[31] These interactions are directly related to the hydrophobic properties of the surface. To reduce these unfavorable interactions, the surface of the microchannel needs to be converted to a more hydrophilic state.[33] Contact angles can measure the hydrophobicity/hydrophilicity of a polymer surface, which were measured and shown in Figure 6.4 with the results were summarized in Table 6.2. Among the polymer substrates tested, pristine COC showed the highest water contact angle of 90.1º, indicating it is the most hydrophobic amongst the three evaluated polymers. Poly(methylmetharcylate), PMMA, showed a relatively lower water contact angle of 77.2 º, while polycarbonate (PC) showed an intermediate water contact angle (84.5º).

The hydrophobicity of pristine polymer can be altered by adsorptive coverage of a passivation material, such as a protein like BSA. BSA has been reported to adsorb with a high propensity to many hydrophobic surfaces.[34] A passivation scheme was thus

adopted to coat the pristine COC substrate with a layer of BSA. Thin films of PMMA, PC, COC of 250 μm thick were incubated in a solution containing 0.5 mg/mL BSA for 20 min at room temperature. Following Passivation, their contact angles were measured, and the results of which are shown in Figure 6.4 and Table 6.2.

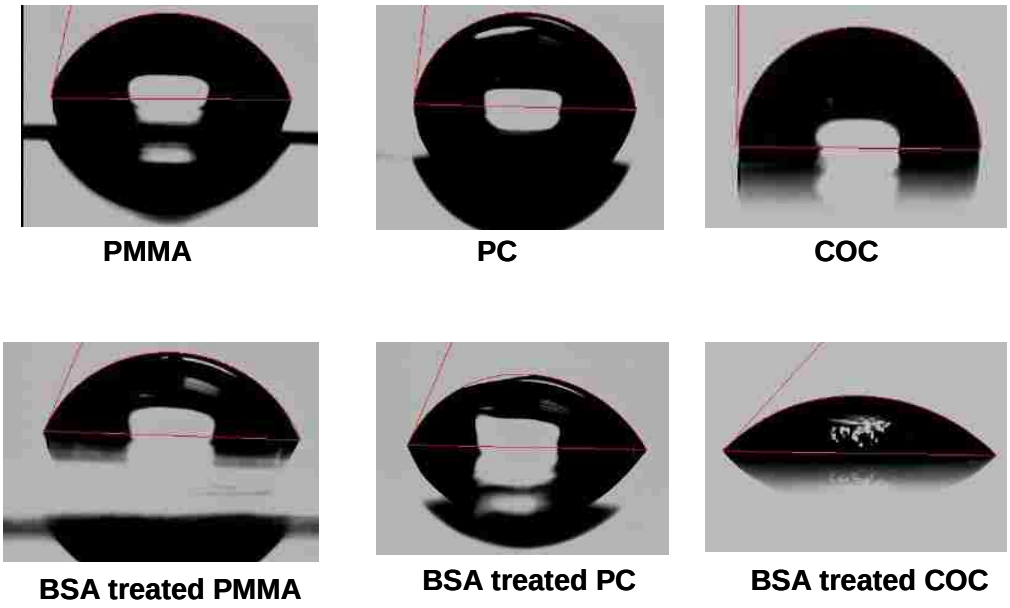


Figure 6.4 Water contact angle measured from various polymer substrates. The three images in the upper panel are from pristine substrates. The three images in the lower panel are from substrates that were BSA treated.

Table 6.2 Contact angle of various polymer substrates

	PMMA	PC	COC
pristine substrate	77.2±1.89	84.5±1.37	90.1±0.5
BSA treated substrate	66.5±2.56	69.2±1.25	48.0±0.44

For all three polymer substrates, treatment with BSA solution increased their hydrophilicity as indicated by reductions in their measured water contact angles. In these polymers, BSA-treated COC showed the most significant increase in its hydrophilic character compared to BSA-treated PC and PMMA; its contact angle was reduced by 47% from 90º to 48º after BSA treatment. Interestingly, PMMA and PC

showed similar contact angles after BSA treatment with values higher than that of COC, consistent with the observation that the adsorption of BSA to COC was more favorable.

6.3.2 LDR in Continuous Flow COC Microfluidic Chip

Thermal cycling using in a continuous flow reactor has been found to be more thermally efficient compared to benchtop thermal cyclers due to better thermal management capabilities and the lack of heating/cooling large thermal masses.[31, 32, 35, 36] This is because the continuous flow microreactor is characterized by a large surface area, and the reaction mixture is fully exposed to the heating elements and brought to thermal equilibrium instantaneously. In the benchtop device, however, the reagents are usually contained in a microtube sitting in a heating block, thus longer times are required for the entire reaction mass to reach thermal equilibrium. Moreover, in the benchtop device, the entire heating block has to be heated and cooled during thermal cycling. By contrast, the heating elements for the continuous flow reactor are maintained at constant temperatures, while the reaction mixture is moved sequentially through different isothermal zones to realize thermal cycling. In spite of the higher thermal efficiency of continuous flow microreactor compared to benchtop devices, their high surface-to-volume ratio must invoke special precautions to avoid potential adsorption artifacts of enzymes to the reactor walls.

The performance of the ligase-mediated reaction on a COC chip was verified to make sure appropriate products were generated in the continuous flow thermal reactor. The LDR cocktail contained 25 nM of primers, 0.5 μL of RT-PCR products generated from a MMP-7 transcript, and 2 U/μL of DNA ligase in 1X LDR buffer. The reaction was carried out in the reactor containing 20 thermal cycles. Following thermal cycling, the

LDR products were analyzed using capillary gel electrophoresis, the results of which are shown in Figure 6.5.

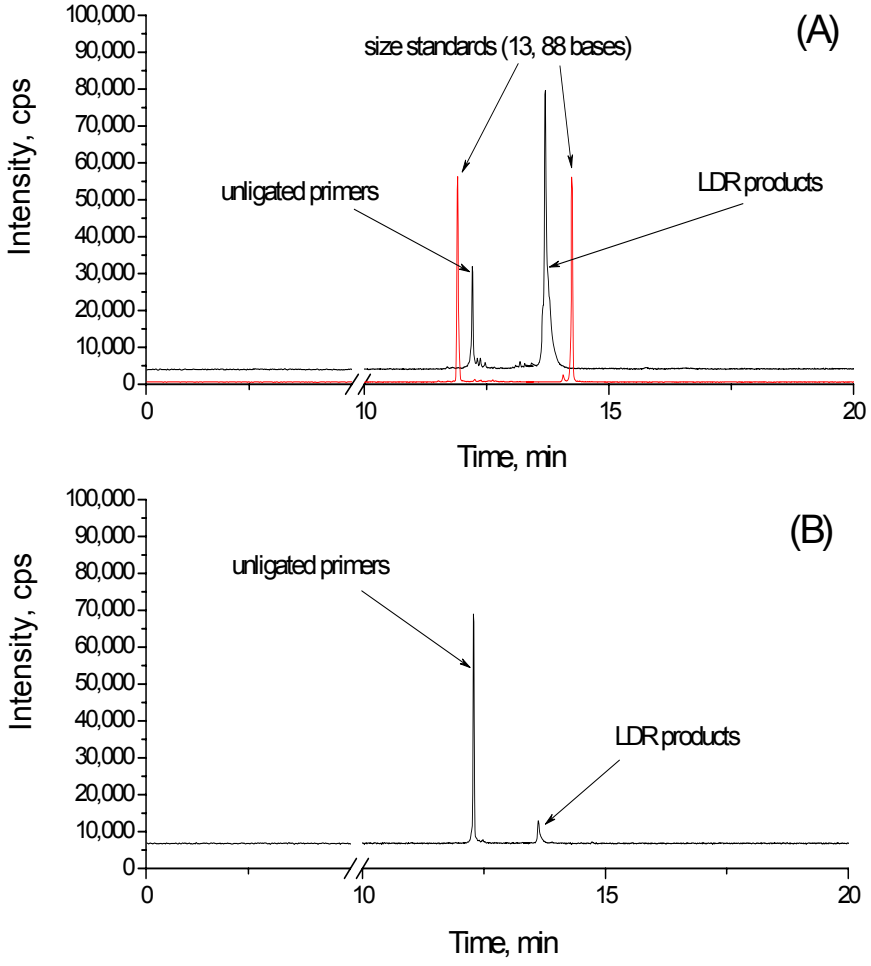


Figure 6.5 Electropherograms of LDR products obtained from the COC continuous flow thermal reactor. (A) LDR products from the COC chip with BSA treatment. (B) LDR products from a pristine COC chip. The templates used for the LDR were RT-PCR amplicons produced from MMP-7 transcripts. The reaction was run for 20 thermal cycles. The reaction mixture was driven at a flow rate of 1 µL/min. The sample was denatured at 94ºC for 2 min and injected into the capillary using a voltage of 2 KV for 20 s. The separation voltage was 6 KV.

In Figure 6.5 (B), it is seen that there was very little LDR product formed when the LDR was carried out in a pristine COC microfluidic reactor. The low yield of LDR was primarily due to non-specific adsorption of DNA ligase onto the walls of the pristine COC

thermal reactor. It is well known that proteins tend to adsorb to solid surfaces and undergo conformational changes.[37, 38] Usually, the extent of protein adsorption increases with increasing surface hydrophobicity.[39] Because pristine COC substrates are highly hydrophobic as revealed by its high contact angle, the DNA ligase molecules were severely denatured and lost their enzymatic activity when they were adsorbed onto the COC surface. The surface of pristine polymers can be passively coated by PEG, PVP, or BSA moieties to modify its surface hydrophobicity.[40] BSA is a thoroughly studied model protein and has been widely used in many chip applications as a surface modifier. When adsorbed onto hydrophobic surfaces, it tends to unfold and spread over the attached surface to form a full monolayer over the surface. The 0.5 mg/mL BSA solution contains $4.46X10^{12}$ BSA molecules per μL, considering its molecular weight of 68 kD. The ligase concentration in LDR mixture is 2 U/uL, which corresponds to $0.18X10^{12}$ ligase molecules per μL. Since binding of BSA onto the pristine polymer surface is a dynamic process through weak molecular interactions, and it is still possible that a ligase could be adsorbed onto the COC surface. The high BSA-to-ligase ratio of 25:1 in the LDR working solution make BSA preferentially binds to COC surface and effectively inhibits the non-specific binding of ligase, leaving the majority of the ligase molecules in solution intact. Figure 6.5 (A) shows the gel electrophoresis result of the LDR products obtained from reaction in COC microchip that was BSA treated. The LDR was run using the same thermal cycling conditions and volumetric flow rate as in the pristine COC microchip. In this figure, it was clearly seen that significantly higher amount of LDR product was observed, suggesting that treatment with BSA is effective on preventing enzyme deactivation due to non-specific adsorption of DNA ligase.

6.3.3 Single-molecule Detection of MMP-7 mRNA on COC Microchip

The aim of this study was to quantitatively measure the expression level of specific mRNA transcripts of clinical significance using LDR performed in a continuous flow polymer chip with the resultant LDR products being interrogated by spFRET readout on the same chip. As RNA cannot typically serve as a template for LDR due to low yield, the mRNA transcript of target gene needs to be first reverse transcribed to its complementary DNA (cDNA). In this study, HT-29 and SW620 metastatic colorectal cancer cell lines were used as models for expression profiling of MMP-7 transcripts due to their various expression levels of this gene demonstrated in Chapter 5. HEL299 was used as a negative control due to its extremely low expression level of MMP-7. In this set of experiments, the sensitivity and quantification capabilities of spFRET for expression profiling were to be evaluated.

Two hundred ng of total RNA was extracted from each of these three cell lines and were converted into cDNA by AffinityScript reverse transcriptase, which exhibited high RT efficiency and has been investigated thoroughly in the RT-PCR experiments discussed in Chapter 5. The produced cDNA was then included into the LDR cocktail for subsequent analysis. A LDR mixture of 40 μL was prepared containing 10 pM of upstream and downstream primers, 2 U/μL of DNA ligase, 0.5 mg/mL of BSA, and the cDNA template generated from the preceding reverse transcription. The mixture was loaded into a 50 μL glass syringe, affixed to the COC chip through a capillary tube. The LDR was carried out using the BSA-treated COC micro-reactor containing 20 thermal cycles and the reactants were driven at a flow rate of 1 μL/min hydrodynamically. The generated photon bursts collected continuously for 1 min are shown in Figure 6.6.

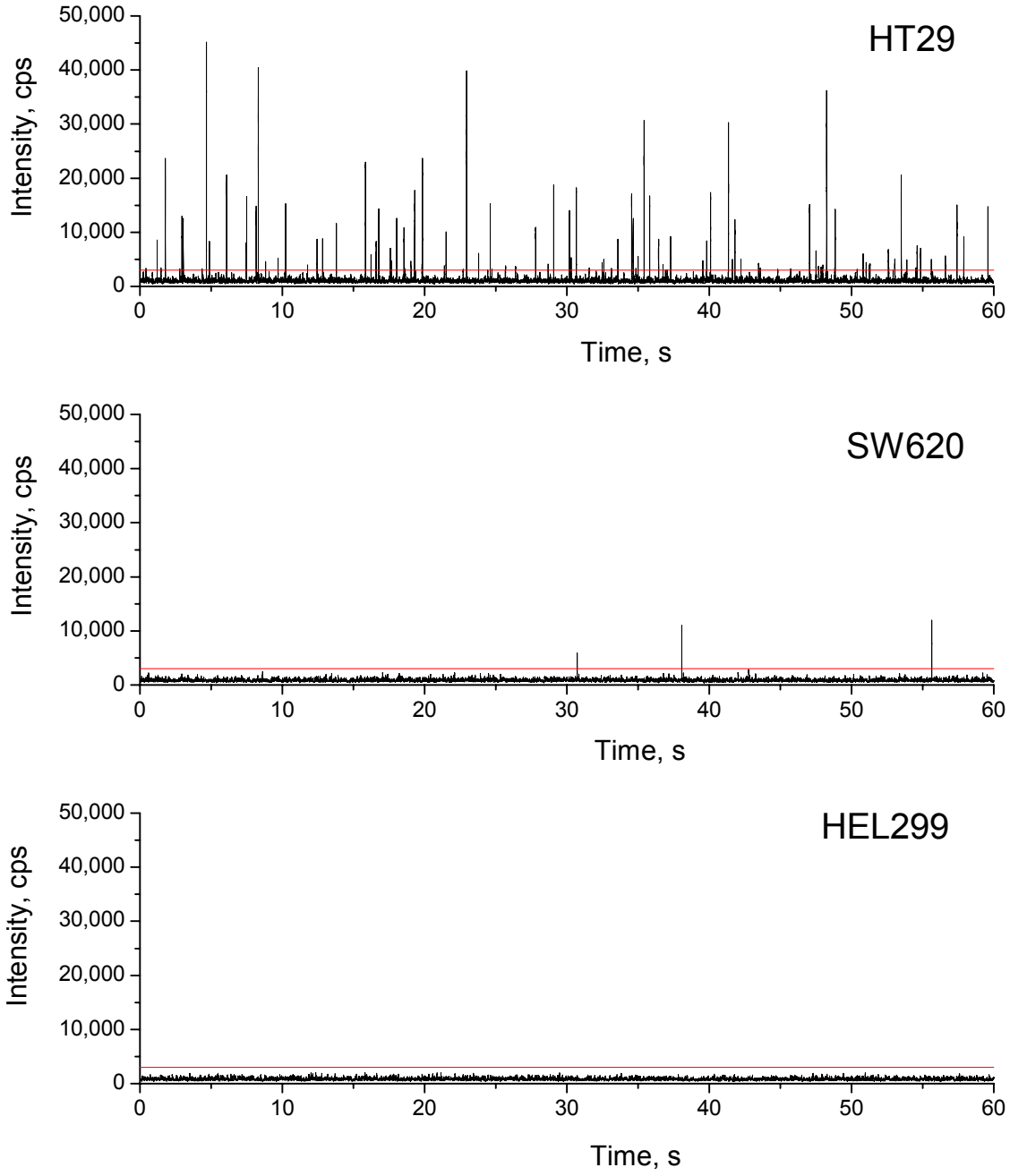


Figure 6.6 Photon burst generated from LDR for expression profiling of MMP-7 transcript. The LDR mixture contained 10 pM of upstream and downstream primers, 2 U/μL of DNA ligase, 0.5 mg/mL of BSA and cDNA template from HT29, SW620 and HEL299. The LDR was subjected to 20 thermal cycles in the continuous-flow microreactor that was BSA-treated and the reactants were driven at a flow rate of 1 μL/min. The produced photon bursts were collected for 1 min. A threshold of 3,000 counts/s was set to count the eligible photon burst.

To exclude false positives generated by autofluorescence fluctuations, a threshold of 3,000 counts/s was set as a discriminator threshold, which was shown to be appropriate for COC substrates through earlier studies.[17] In Figure 6.6 a total of 97 photon bursts were identified using reversed transcribed cDNA from HT-29 cells in a 1 min sampling time. This number corresponds to $7.98X10^6$ copies/μg total RNA, or 103 copies/cell of MMP-7 transcripts in the original total RNA sample after being corrected for RT efficiency of 0.76, LDR efficiency of 0.4, the number of LDR amplification cycles (20), and a sampling efficiency of 0.04%, which is an instrument specific parameter of the LIF single-molecule detection system, according to the following formula:

$$\text{Transcription Level}\left(\frac{\text{mRNA copies}}{\mu\text{g of total RNA}}\right) = \frac{\#\text{ of photon burst / }\mu\text{g of total RNA}}{\eta_{RT} \times \kappa_{LDR} \times \eta_{LDR} \times \eta_{sampling}} \quad (6.1)$$

where η_{RT} is the efficiency of reverse transcription, κ_{LDR} Is the number of LDR thermal cycles, η_{LDR} is the average efficiency of LDR performed in the COC microreactor, and $\eta_{sampling}$ is the sampling efficiency of the LIF single-molecule detection system. This value is fairly consistent with the result of 178 copies/cell obtained by the RT-qPCR approach described in Chapter 5. Similarly, there are 3 photon bursts identified from the LDR using cDNA from the SW620 cell line in the 1 min sampling time. This value accounts for 0.247 X 10^6 copies of MMP-7 transcripts per μg of total RNA or 1.68 copies/cell in the original total RNA sample, which agrees well with the result of 3.18 copies/cell obtained by the RT-qPCR approach. There were no photon bursts identified from the LDR using cDNA from HEL299. This is not surprising taking into account the low abundance of MMP-7 gene expressed in HEL299 cells, which is about 27,000 times lower than that in HT-29 from the RT-qPCR analysis.

To illustrate the effect of sampling time on the sensitivity of spFRET measurements, the reactions used in Figure 6.6 were carried out under the same running conditions and the photon bursts from the resultant LDR products were collected with an extended 10 min sampling time. The results are shown in Figure 6.7.

For LDR performed to quantify the expression level of MMP-7 transcripts from the HT-29 cell line, there were a total of 920 photon burst events gathered from the data trace in Figure 6.7 during the 10 min sampling time. This readout is almost 10 times more sensitive than the measurement using a 1 min sampling time. Similarly, for LDR performed using SW620 cell there were 27 photon bursts recognized, which, again, is nearly 10 times as sensitive as in the 1 min sampling time. For LDR run using HEL299, there were still no photon bursts identified because of the extremely low abundance of MMP-7 transcripts expressed in this cell line. Therefore, we can see that the sensitivity of spFRET quantification after LDR amplification did increase with sample readout time. The sensitivity multiplication is especially valuable for analysis of transcripts of low copy number, because the number of photon bursts generated from LDR is typically low for low abundance samples, as seen in the case of SW620, due to the linear amplification associated with LDR.

6.3.4 Effects of Flow Rate on spFRET Measurement using Continuous Flow LDR

When performing LDR in the continuous-flow microfluidic reactor for expression analysis, a shorter turnaround time is always favored. The overall analysis time of the assay is comprised of both the time to conduct thermal cycling for on-chip LDR, and the time to collect photon bursts during the phase of spFRET readout, which has been discussed earlier. The time spent for LDR thermal cycling is determined by the flow rate

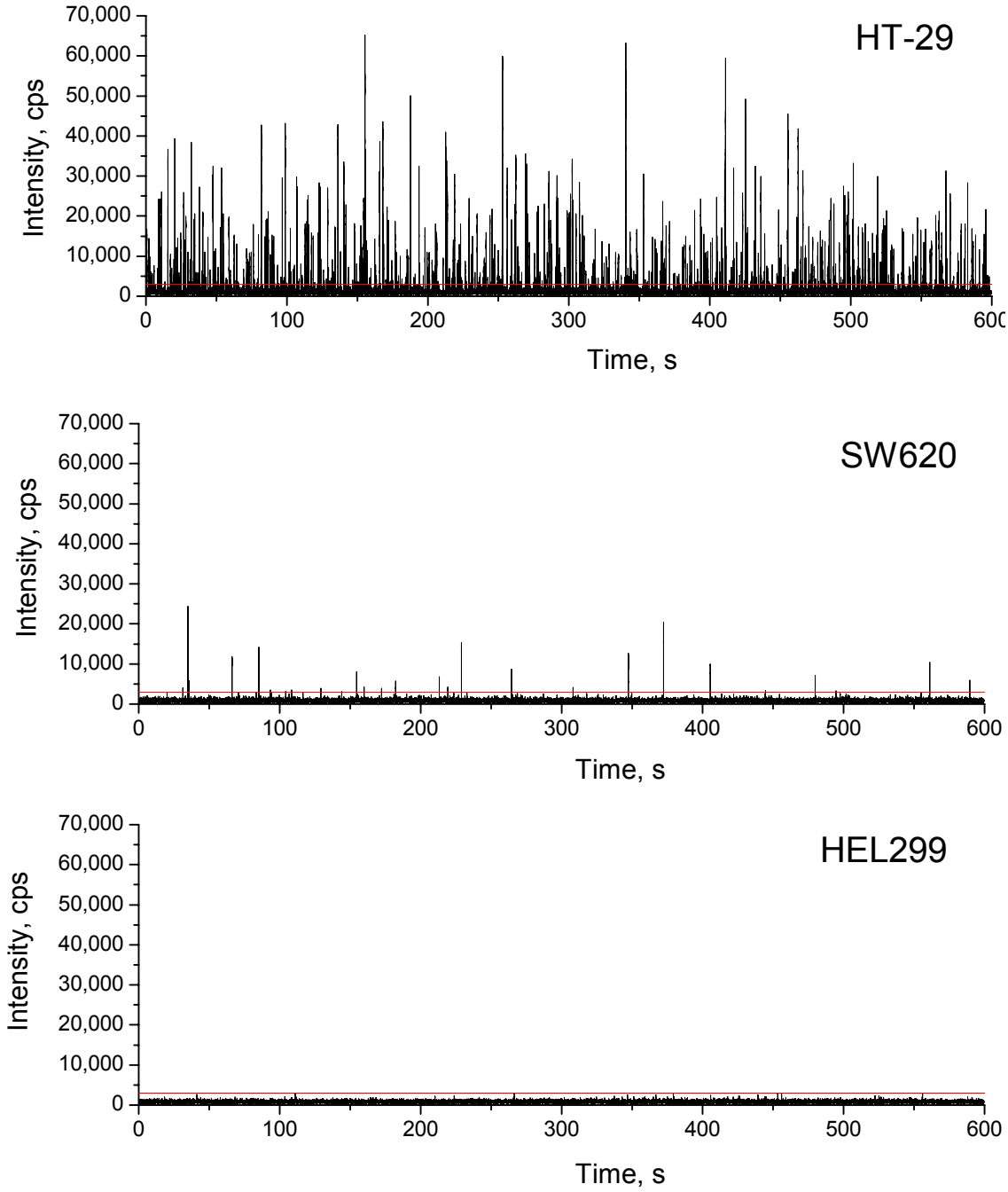


Figure 6.7 Photon burst generated from LDR for expression profiling of MMP-7 transcript with extended sampling time. HT-29, SW620, HEL299 were used to generate cDNA served as templates for LDR. The LDR contained exactly the same compositions and run under the same conditions as in Figure 6.6. The produced photon bursts were collected continuously for 10 min.

of reaction mixture contained in the microfluidic channel. A complete LDR assay consists of multiple cycles of denaturation, hybridization, and ligation, which occurred sequentially at the 95ºC and 65ºC temperature zones in the chip, respectively. Previous studies have shown that denaturation and hybridization take place in as short as 1 s.[41] Hence, the total time of LDR is set by the time needed for sufficient ligation, which is limited by kinetics of the ligase enzyme.

In making quantitative measurements with spFRET, the sensitivity of the measurement was evaluated by the number of photon bursts that were successfully counted from the preceding LDR. Figure 6.8 shows photon bursts of LDRs performed at different flow rate in the COC continuous-flow microreactor to profile the expression level of MMP-7 transcripts from the HT-29 cell line. A series of volumetric flow rate of 0.35, 0.52, 0.78, 1.56, 3.12, 6.24 μL/min were investigated, which corresponded to a dwell time of 90, 60, 40, 20, 10, 5 s within the 65ºC ligation zone. At each flow rate the number of photon bursts was counted, and the average peak height of the photon bursts were evaluated, which were plotted against sample flow rate in Figure 6.9.

In photon burst measurement, a higher peak height of photon burst is always favored attributed to its high signal-to-background ratio. Figure 6.9 (A) shows the average peak height of photon bursts as a function of sample flow rate. As can be seen, when flow rate of the LDR mixture increases, the average peak height of detected photon bursts drops significantly, indicating decreased signal-to-background ratio of the measurement. This can be understood from the fundamentals of LIF single-molecule detection. A photon burst was typically generated when a fluorescent molecule travelled through a focused laser spot and experienced repetitive excitation from its ground state.

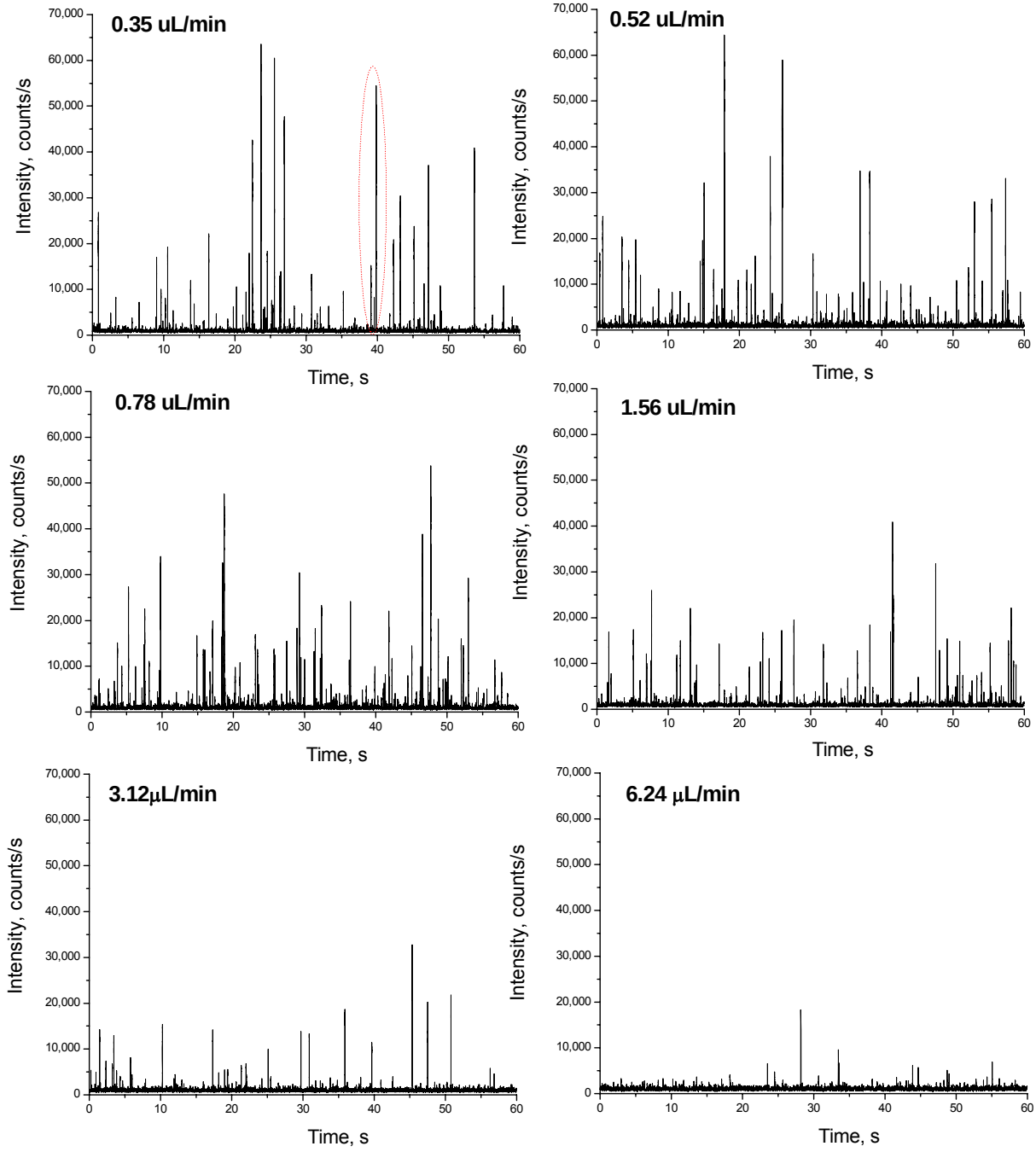


Figure 6.8 Photon burst generated from LDR run at different sample flow rate for expression profiling of MMP-7 transcript. The LDR mixture contained 10 pM of upstream and downstream primers, 2 U/μL of DNA ligase, 0.5 mg/mL of BSA and cDNA template from HT29. 20 thermal cycles were carried out on the LDR mixture in the continuous-flow microreactor. The produced photon bursts were collected for 1 min. A threshold of 3,000 counts/s was set to count the eligible photon burst.

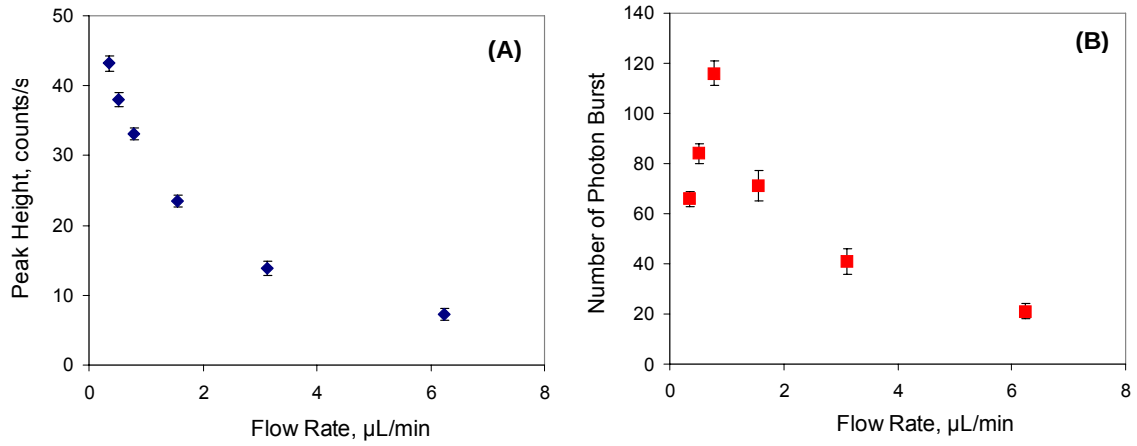


Figure 6.9 Effect of sample flow rate on spFRET measurement. (A) Average peak height of photon burst as a function of flow rate (B) Number of detected photon burst as a function of flow rate

The longer the fluorophore resided within the laser spot, the more excitations it will experience, thus a higher photon burst peak will be produced. When the LDR mixture flowed through the microchannels slowly, the produced molecular beacons were exposed to the laser spot for extended period of time, resulting in higher peaks of photon bursts. As LDR flow rate increased, the produced molecular beacons were not able to experience sufficient excitations within the laser spot. Hence, reduced peak heights of photon bursts were observed.

In making quantitative measurement with spFRET, the sensitivity of the measurements relies on the number of photon burst events that can be successfully captured from the fluorescent molecules, and a more sensitive measurement arises from more countable photon bursts. In spFRET measurement of the molecular beacons generated from LDR, since the photon bursts were collected directly from the microchannel where the preceding LDR was carried out and the flow rate of the LDR mixture through the microreactor is the same as the produced molecular beacons

through the single-molecule detection window, the effect of sample flow rate on the photon burst counting is two folds. On one hand, the higher the sample flow rate, the more photon bursts that can be collected in unit sampling time from molecular beacons that flow through the detection window. On the other hand, the yield of LDR is dependent upon the kinetics of the ligase. The higher the sample flow rate, the shorter the dwell time of the LDR mixture within the 65ºC ligation temperature zone, thus the lower LDR yield. The effect of ligation time on the yield of LDR performed in polymer continuous-flow microfluidic reactor has been investigated in previous studies.[32] Therefore the net effect of sample flow rate on the countable photon bursts resulted from preceding LDR is a combination of these two effects discussed above. Figure 6.9 (B) shows the counts of photon burst as a function of sample flow rate in the LDR/spFRET assay for MMP-7 transcript profiling. At lower flow rates, the number of photon bursts increased almost linearly with sample rate because theoretically the number of photon burst events is proportional to the sample flow rate regardless of ligase kinetics. At higher flow rates, however, the LDR mixture does not have sufficient dwell time within the ligation zone to ensure 100% yield of LDR, and ligase kinetics dominates in this regime. Thus, the number of photon bursts drops with further increases in sample flow rate because the number of detectable molecular beacons is limited by the decreased rate of ligation reaction. It was found that at an optimal flow rate of 0.78 μL/min the maximum number of 116 photon bursts was detected in 1 min sampling time. At this condition, a turnaround time of 22.3 min is required to make a quantitative measurement with the LDR/spFRET assay.

6.3.5 Photostability of Molecular Beacons

In spFRET measurement of LDR products, the detection of a photon burst is not only dependent on a successful ligation reaction, where a molecular beacons is formed, but also dependent on the photostability of the chosen donor/acceptor dye pairs within the beacon structure. In laser-induced fluorescence single-molecule experiments, the dye molecule could be cycled up to 10^9 times/s between the ground and first excited states with the exact number dependent on the fluorescence lifetime of the dye molecule. Photo-labile dye molecules are prone to photobleaching and thus, permanently lose their fluorescing capability.

The peak shape of single molecule photon bursts is supposed to be symmetrical adopting a Gaussian shape due to the Gaussian profile of the excitation beam. However, when a molecule is photobleached, it suddenly stops fluorescing and the peak shape is characterized by a sudden drop to background. Figure 6.10 shows an expanded view of a few typical single molecule events from one panel in Figure 6.8, where the LDR was performed at the lowest flow rate of 0.35 μL/min, and the generated MBs are most likely to be photobleached due to longer exposure time within the focused laser spot.

It can be seen in this figure that the shape of these photon bursts is nearly symmetrical and characterized by Gaussian profile. Careful inspection of the photon bursts from LDR perfomed on different running conditions indicated that these events are basically symmetrical, showing no abrupt cessation of photon emission, which is indicative of irreversible photobleaching during the travelling of the MB through the excitation laser beam.

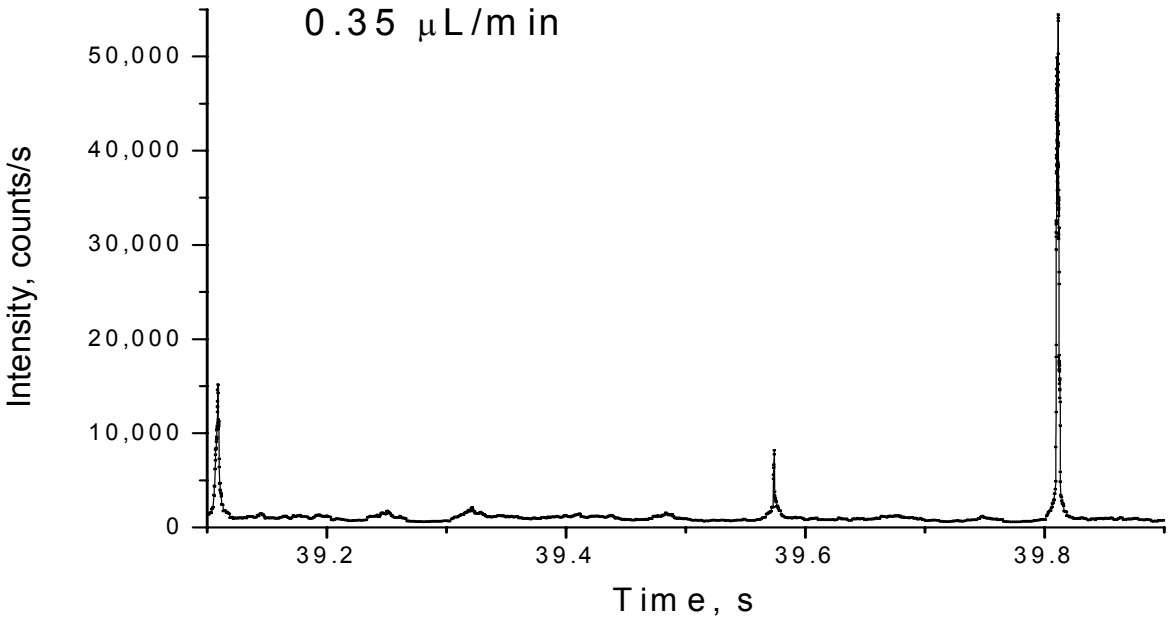


Figure 6.10 Expanded view of a few typical single-molecule photon bursts. The shape of these photon burst peaks is symmetrical adopting a Gaussian shape. The LDR was run at a flow rate of 0.35 μL/min.

6.3.6 Improvement of Sampling Efficiency in Single-Molecule Detection

In photon burst detection of single molecules using laser-induced fluorescence, it is desirable to create a diffraction-limited probe volume to improve the SBR in the measurement.[42] For quantitative analysis in flowing samples, however, diffraction-limited sampling volumes are challenged by poor sampling efficiency, where only a small portion of the sample molecules are actually observed, leaving the vast majority of the sample molecules in the solution undetected.

Sampling efficiency in SMD can be improved by confining the sample molecules to travel through the center of the fluid channel through hydrodynamic focusing, electrokinetic focusing, or channel dimension downsizing, so that a higher percentage of sample molecules will travel through the probe volume and be registered by the

photodetector. For example, Dovichi and co-workers demonstrated hydrodynamic focusing in single particle counting using a capillary column by introducing two lateral sheath flows flanking the sample stream to narrow the flowing sample particles through the center of the probe volume.[43] de Mello and co-workers demonstrated hydrodynamic focusing in microfluidic channels for detection of fluorescent particles with increased detection efficiency.[44] However, the challenge of hydrodynamic focusing in planar microfluidic channels is that it is difficult to implement in three dimensional focusing. Mathies and co-workers developed a microfluidic device with electrokinetic focusing for electrophoresis separation and single-molecule detection of dye-labeled DNA.[45] This device featured a cross channel to electrokinetically focus the molecular streams to the center of the separation channel to allow detecting a larger proportion of the DNA molecules, and the detection efficiency increased to 1.1%, which is two orders of magnitude improvement over the conventional capillary system. Wang *et al.* directly patterned microelectrodes onto the walls of a microfluidic channel to precisely confine the flowing molecules into a tiny detection region to facilitate highly efficient counting of sample molecules.[46] Another simpler and more straightforward approach involves fabrication of microchannels with a physically narrowed region to directly confine the sample molecules to the center of the fluid channel. Nie and co-workers studied the flow dynamics of single fluorescent molecules in ultra-thin capillaries by pulling glass capillaries to create a tapered region with inner diameter of 500 – 600 nm so that most molecules could be interrogated.[47] Foquet *et al.* reported fabrication of sub-micron fluidic channels on a silicon wafer using a sacrificial layer method for single-molecule detection.[48] Analyte molecules were delivered through the diffraction-limited probe

volume under electrokinetic force and were exhaustively detected. Although a nearly perfect sampling efficiency could be theoretically achieved by flowing single molecules through a fluidic channel with sub-micron dimensions, this approach is subjected to challenges such as fabrication complexity, filling of sample, elevated fluorescence background, and increased interactions between analyte molecules and the channel walls.

In this study, a microfluidic channel with a tapered detection region of 25 X 25 μm in cross section was fabricated on a COC substrate to improve sampling efficiency of the molecular beacons. The structure of the chip is illustrated in Figure 6.11.

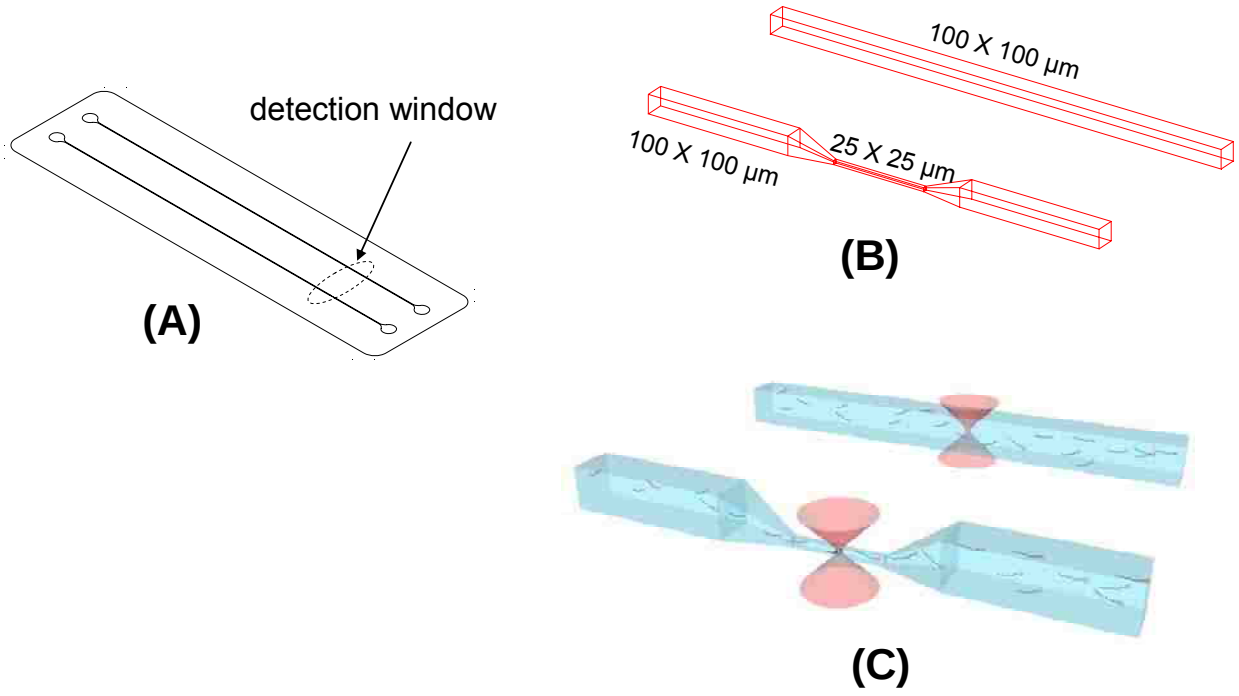


Figure 6.11 Illustration of the microfluidic chip with tapered detection window for more efficient collection of photon burst. (A) Schematics of the microfluidic channel with varied dimensions in detection window (B) Expanded view of the tapered detection window. One channel has a uniform cross section of 100 X 100 μm, while the other channel has a tapered cross section of 25 X 25 μm (C) Illustration of flowing of fluorescent molecules through the tapered detection window

In this chip, one section of the channel had a cross section of 100 X 100 μm, while the other section of the channel had a reduced cross section to confine more analyte molecules to flow through the laser-defined probe volume. This cross section was 25 X 25 μm, and its length was only 0.5 mm to minimize increased back pressure due to restricted flow.

Fluorescent microspheres (Molecular Probes, T-8878, Eugene, OR) were used to investigate the differences in photon burst numbers and intensity from microchannels with different dimensions. These microbeads had an average diameter of 0.1 μm, and were preloaded with multiple FRET dye pairs (excite @ 633 nm, emit @ 716 nm). They were diluted to $3X10^{-14}$ mol/L in 1X TE buffer and driven through the microchannel at a flow rate of 1 μL/min. Figure 6.12 (A) shows the photon bursts from single microspheres flowing through the straight channel with a uniform cross section of 100 X 100 μm. The data were collected for 2 min with a total of 36,000 microbeads that passed through the channel. The background fluorescence of this microchannel was about 1,000 counts/s, which was obtained by running TE buffer only through the channel. A threshold level of 3,000 counts/s was set so that the false positive rate was nearly zero in the blank. The photon burst events arising from the microspheres are shown on the bottom panel and 32 bursts were revealed above the threshold level, corresponding to a sampling efficiency of 0.09%. The background fluorescence of the microchannel with a tapered detection window is shown on the top of Figure 6.12 (B), which is about 2,000 counts/s. A threshold level of 6,000 counts/s was set to suppress any false positive signals from the blank as noted above. At this level, a total of 312 events were observed, corresponding to a sampling efficiency of 0.9%. The confinement of the analytes to the

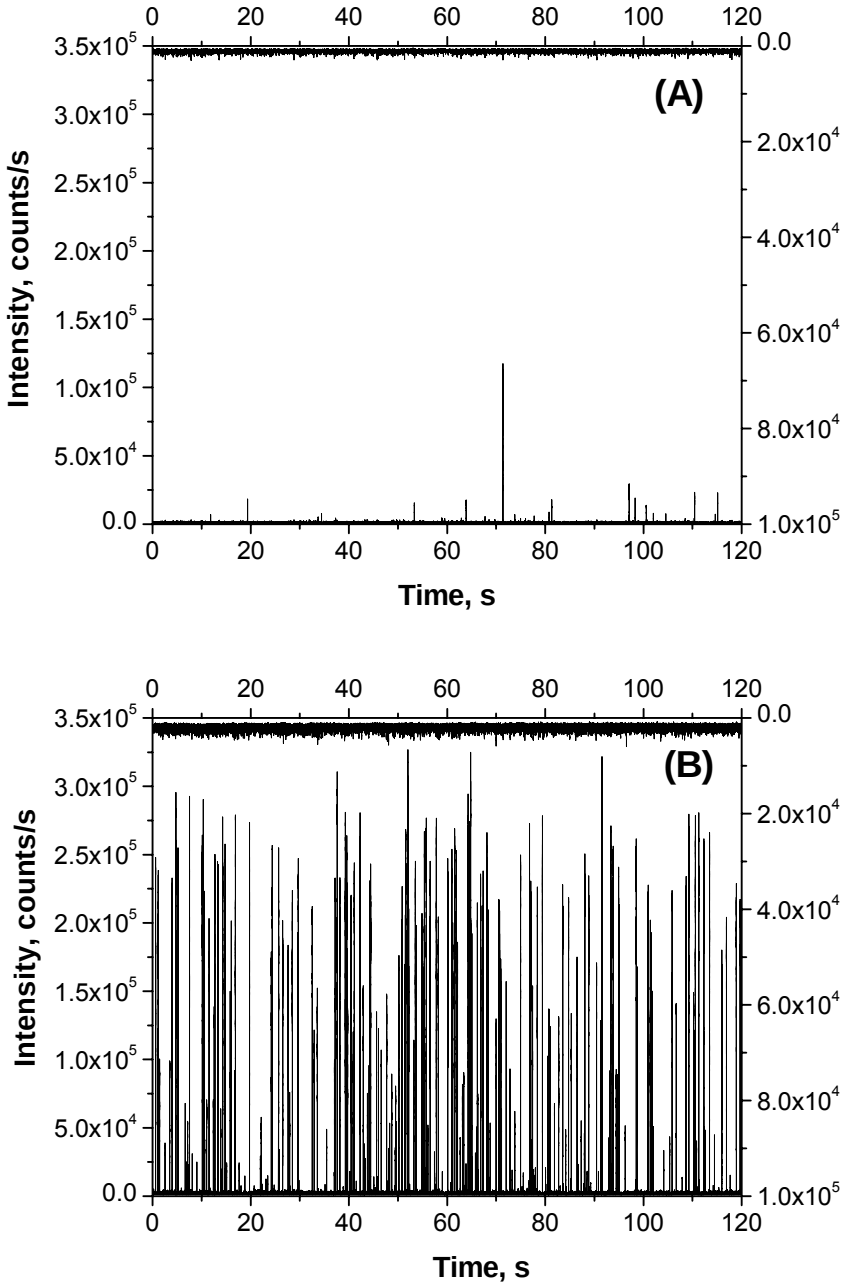


Figure 6.12 Photon bursts collected from single fluorescent microspheres on microfluidic chip with detection window of different dimensions. These microbeads were preloaded with multiple FRET dye pairs, which were excited at 633 nm and emitted 716 nm. They were diluted to $3X10^{-14}$ mol/L in 1X TE buffer and driven through the microchannel at a flow rate of 1 μL/min. The produced photon bursts were collected for 2 min. Threshold levels of 3,000 and 6,000 counts/s were set for the 100 X 100 μm channel and 25 X 25 μm channel, respectively, to count the eligible photon bursts from flowing microspheres.

probe volume is obvious by comparing the number of photon bursts from panels (A) and (B) in Figure 6.12. It was also noted that the average peak height in panel (B) is significantly higher than that in panel (A). The peak height enhancement in tapered microchannel is another evidence of molecular confinement due to the fact that in microchannel with reduced size the analyte molecules are more likely to pass through the center of the channel, where the focused laser beam is situated, thus producing a strong photon burst than in microchannel with larger cross section. These results suggested that reduction in the detection window led to increase in both the number and peak height of photon bursts that can be detected from the microfluidic channel.

6.3.7 Quantitative Measurements of MMP-7 Transcripts on a COC Microchip

In performing quantitative analysis of mRNA expression levels, the number of detected photon bursts is dependent upon the amount of input transcripts. Figure 6.13 shows a calibration plot of the counts of detected photon bursts as a function of the input copy number of MMP-7 transcripts.

The LDR cocktail contained 10 pM of upstream and downstream primers, 2 U/μL of DNA ligase, 0.5 mg/mL of BSA, and cDNAs that were reverse transcribed from MMP-7 transcripts with input copy numbers ranging from 300 to 30,000. The reaction mixture was driven at a flow rate of 0.78 μL/min and subjected to 20 thermal cycles through the 65ºC ligation zone and 94ºC denaturation zone. The photon bursts were directly collected from the COC chip where LDR was performed. In Figure 6.13, we can see that for the 100 X 100 μm detection channel the number of photon burst events was linear with the input copy number of MMP-7 mRNA, and the slope of the calibration plot was equal to 0.0022 photon bursts per MMP-7 mRNA. For the 25 X 25 μm detection channel,

the calibration plot was also linear with the slope of the plot equal to 0.022 photon bursts per MMP-7 mRNA. The slope of the calibration plot represents the analytical sensitivity of the assay and using the 25 X 25 μm channel provides nearly 10 times more sensitivity than the 100 X 100 μm channel due to improvements in the sampling efficiency in the tapered detection channel.

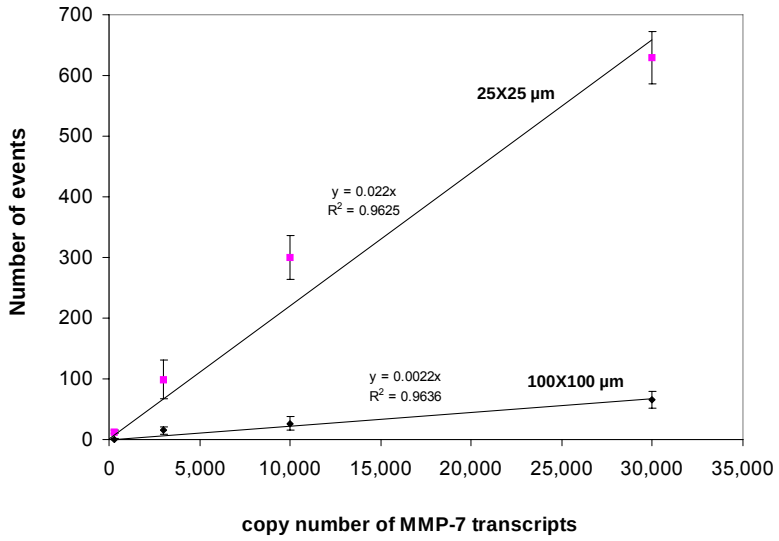


Figure 6.13 Calibration plot of photon bursts as a function of input copy number of MMP-7 transcripts. The LDR was run for 20 thermal cycles and driven through the microchannel at a flow rate of 0.78 μL/min. For 100 X 100 μm detection channel photon bursts were directly collected on the same chip where LDR was performed. For 25 X 25 μm detection channel LDR was performed on the primary chip, cleaned with a 0.2 μm filter, then loaded into the secondary chip for photon bursts collection.

6.3.8 Fast Detection of Stroke Biomarker with LDR/spFRET

Currently stroke is becoming one of the major causes of death and disability in the US. Stroke can be categorized into ischemic stroke and hemorrhagic stroke, and each one is treated distinctively. Survival and rehabilitation of stroke patients rely on proper treatment and medicine during the first few hours of disease. Due to the stringent requirement for identification of stroke types, it is important to obtain an assessment on the clinical state of the patient: stroke versus no stroke and if stroke is detected, is it

ischemic or hemorrhagic stroke. It has been reported that the expression level of AMPH gene is upregulated in hemorrhagic stroke but not in ischemic stroke.[29] By profiling AMPH transcripts from patient's blood, it would be possible to subtype the stroke of a patient and give timely and proper treatment.

In this study, the LDR cocktail contained 10 pM of upstream and downstream primers, 2 U/μL of DNA ligase, 0.5 mg/mL of BSA, and cDNAs that were reverse transcribed from AMPH transcripts with varied input copy numbers. The reaction mixture was driven at a flow rate of 0.78 μL/min through the microreactor with 2 thermal cycles to reduce the turnaround time of the assay. The generated molecular beacons were directly observed downstream of the LDR microreactor where a 25 X 25 μm detection window was patterned. The photon bursts were collected for an extended period of 10 min to improve the sensitivity of the measurement. Figure 6.14 displays the calibration plot of the detected photon bursts as a function of input transcript copy number of AMPH.

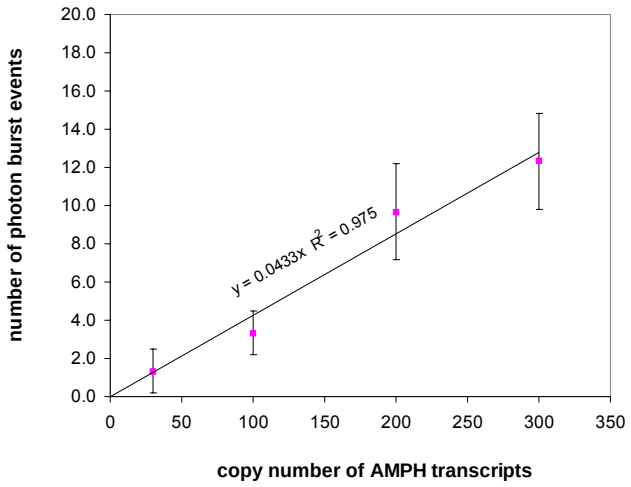


Figure 6.14 Calibration plot of photon bursts as a function of input copy number of AMPH transcripts. The LDR was run for 2 thermal cycles and the generated photon bursts were directly collected from the chip downstream of the thermal cycling region where a 25 X 25 μm detection channel was patterned. The photon bursts were collected for 10 min.

In this figure, the number of photon burst events was linear with the input copy number of AMPH mRNA, and the slope of the calibration plot equals to 0.0433 photon bursts per AMPH mRNA. As low as 30 copies of AMPH transcripts were detected with an average of 1.3 photon bursts observed during the extended 10 min sampling time. This analysis needs a total of 15 min to read the copy number of the target transcripts present in the LDR cocktail using only 2 thermal cycles. However, increasing the cycle numbers would increase the yield, but also increase the assay turnaround time.

6.4 Conclusions

In this chapter, a COC fluidic chip was used to accommodate the LDR assay for expression profiling. The original mRNA transcripts were linearly amplified and thus, no bias introduced into the quantification as would be the case for RT-qPCR that employs exponential amplification. The pristine COC microchip was treated with BSA and the hydrophilicity/hydrophobicity properties were studied using water contact angle measurements. BSA treatment had a significant effect on preventing non-specific adsorption of the ligase enzyme, and thus ensured a high yield of the LDR when performed using the COC chip. Significant improvements in assay sensitivity, even for the 2-cycle LDR, could be realized by simply increasing the sampling efficiency through reducing the channel size at the single-molecule detection zone of the chip and/or increasing the probe volume. In this application the channel dimension was tapered to 25 X 25 μm in cross section and the resulting sampling efficiency of generated rMBs was increased by a factor of 10. The processing flow rate of LDR must be balanced by optimizing the performance of the LDR and the single-molecule detection efficiency to maximize the detection sensitivity of analysis. It was determined to be 0.78 μl/min in the

present case. Expression analysis of the MMP-7 gene was performed to analyze the transcriptional level from different cell lines. The results are comparable to those obtained using conventional RT-qPCR. In the spFRET readout, the Cy5/Cy5.5 dye pair was fairly photostable, showing no photobleaching artifacts as evident from the photon burst trace data. The flow rate of LDR was optimized to maximize the detection sensitivity of analysis. Expression of AMPH transcripts was also analyzed using this assay and gave a result in as short as 15 min.

6.5 References

(1) Cohen, C. D.; Frach, K.; Schlondorff, D.; Kretzler, M. *Kidney International* 2002, *61*, 133-140.

(2) Tenedini, E.; Roncaglia, E.; Ferrari, F.; Orlandi, C.; Bianchi, E.; Bicciato, S.; Tagliafico, E.; Ferrari, S. *Cell Death & Disease* 2010, *1*.

(3) Rapley, R.; Walker, J. M. *Molecular biomethods handbook*, 2nd ed.; Humana Press: Totowa, NJ, 2008.

(4) Jin, L.; Lloyd, R. V. *J Clin Lab Anal* 1997, *11*, 2-9.

(5) Yamaguchi, Y.; Yatsushiro, S.; Yamamura, S.; Abe, H.; Abe, K.; Watanabe, M.; Kajimoto, K.; Shinohara, Y.; Baba, Y.; Kataoka, M. *Analyst* 2011, *136*, 2247-2251.

(6) Schulze, A.; Downward, J. *Nature Cell Biology* 2001, *3*, E190-E195.

(7) Wang, E.; Miller, L. D.; Ohnmacht, G. A.; Liu, E. T.; Marincola, F. M. *Nature Biotechnology* 2000, *18*, 457-459.

(8) Nelson, P. T.; Baldwin, D. A.; Scearce, L. M.; Oberholtzer, J. C.; Tobias, J. W.; Mourelatos, Z. *Nature Methods* 2004, *1*, 155-161.

(9) Palmer, S.; Wiegand, A. P.; Maldarelli, F.; Bazmi, H.; Mican, J. M.; Polis, M.; Dewar, R. L.; Planta, A.; Liu, S.; Metcalf, J. A.; Mellors, J. W.; Coffin, J. M. *J Clin Microbiol* 2003, *41*, 4531-4536.

(10) Wong, M. L.; Medrano, J. F. *Biotechniques* 2005, *39*, 75-85.

(11) Medley, C. D.; Drake, T. J.; Tomasini, J. M.; Rogers, R. J.; Tan, W. H. *Analytical Chemistry* 2005, *77*, 4713-4718.

(12) Chen, X. J.; Roy, S.; Peng, Y. F.; Gao, Z. Q. *Analytical Chemistry* 2010, *82*, 5958-5964.

(13) Lee, A. C.; Dai, Z. Y.; Chen, B. W.; Wu, H.; Wang, J.; Zhang, A. G.; Zhang, L. R.; Lim, T. M.; Lin, Y. H. *Analytical Chemistry* 2008, *80*, 9402-9410.

(14) Neely, L. A.; Patel, S.; Garver, J.; Gallo, M.; Hackett, M.; McLaughlin, S.; Nadel, M.; Harris, J.; Gullans, S.; Rooke, J. *Nature Methods* 2006, *3*, 41-46.

(15) Korn, K.; Gardellin, P.; Liao, B.; Amacker, M.; Bergstrom, A.; Bjorkman, H.; Camacho, A.; Dorhofer, S.; Dorre, K.; Enstrom, J.; Ericson, T.; Favez, T.; Gosch, M.; Honegger, A.; Jaccoud, S.; Lapczyna, M.; Litborn, E.; Thyberg, P.; Winter, H.; Rigler, R. *Nucleic Acids Research* 2003, *31*.

(16) Wabuyele, M. B.; Farquar, H.; Stryjewski, W.; Hammer, R. P.; Soper, S. A.; Cheng, Y. W.; Barany, F. *Journal of the American Chemical Society* 2003, *125*, 6937-6945.

(17) Peng, Z. Y.; Soper, S. A.; Pingle, M. R.; Barany, F.; Davis, L. M. *Analytical Chemistry* 2010, *82*, 9727-9735.

(18) Barany, F. *Proc Natl Acad Sci U S A* 1991, *88*, 189-193.

(19) Miska, E. A.; Alvarez-Saavedra, E.; Townsend, M.; Yoshii, A.; Sestan, N.; Rakic, P.; Constantine-Paton, M.; Horvitz, H. R. *Genome Biology* 2004, *5*.

(20) Levin, B.; Lieberman, D. A.; McFarland, B.; Andrews, K. S.; Brooks, D.; Bond, J.; Dash, C.; Giardiello, F. M.; Glick, S.; Johnson, D.; Johnson, C. D.; Levin, T. R.; Pickhardt, P. J.; Rex, D. K.; Smith, R. A.; Thorson, A.; Winawer, S. J.; Ad, A. C. S. C. C.; Canc, A. C. R. C. *Gastroenterology* 2008, *134*, 1570-1595.

(21) Barrier, A.; Lemoine, A.; Brault, D.; Houry, S.; Flahault, A.; Dudoit, S. *Clinical Cancer Research* 2005, *11*, 8999s-8999s.

(22) Luo, H. Z.; Zhou, Z. G.; Yang, L.; Yu, Y. Y.; Tian, C.; Zhou, B.; Zheng, X. L.; Xia, Q. J.; Li, Y.; Wang, R. *Japanese Journal of Clinical Oncology* 2005, *35*, 739-744.

(23) Zeng, Z. S.; Shu, W. P.; Cohen, A. M.; Guillem, J. G. *Clinical Cancer Research* 2002, *8*, 144-148.

(24) Lloyd-Jones, D.; Adams, R. J.; Brown, T. M.; Carnethon, M.; Dai, S.; De Simone, G.; Ferguson, T. B.; Ford, E.; Furie, K.; Gillespie, C.; Go, A.; Greenlund, K.; Haase, N.; Hailpern, S.; Ho, P. M.; Howard, V.; Kissela, B.; Kittner, S.; Lackland,

D.; Lisabeth, L.; Marelli, A.; McDermott, M. M.; Meigs, J.; Mozaffarian, D.; Mussolino, M.; Nichol, G.; Roger, V. L.; Rosamond, W.; Sacco, R.; Sorlie, P.; Thom, T.; Wasserthiel-Smoller, S.; Wong, N. D.; Wylie-Rosett, J. *Circulation* 2010, *121*, e46-e215.

(25) Rosamond, W. D.; Folsom, A. R.; Chambless, L. E.; Wang, C. H.; McGovern, P. G.; Howard, G.; Copper, L. S.; Shahar, E. *Stroke* 1999, *30*, 736-743.

(26) Rothwell, P. M.; Coull, A. J.; Silver, L. E.; Fairhead, J. F.; Giles, M. F.; Lovelock, C. E.; Redgrave, J. N. E.; Bull, L. M.; Welch, S. J. V.; Cuthbertson, F. C.; Binney, L. E.; Gutnikov, S. A.; Anslow, P.; Banning, A. P.; Mant, D.; Mehta, Z.; Study, O. V. *Lancet* 2005, *366*, 1773-1783.

(27) Levy, D. E.; Brott, T. G.; Haley, E. C.; Marler, J. R.; Sheppard, G. L.; Barsan, W.; Broderick, J. P. *Stroke* 1994, *25*, 291-297.

(28) Moore, D. F.; Li, H.; Jeffries, N.; Wright, V.; Cooper, R. A.; Elkahloun, A.; Gelderman, M. P.; Zudaire, E.; Blevins, G.; Yu, H.; Goldin, E.; Baird, A. E. *Circulation* 2005, *111*, 212-221.

(29) Alison E. Baird, D. F. M., Ehud Goldin, Kory Johnson: US, 2010, pp 143.

(30) Chen, J. F.; Wabuyele, M.; Chen, H. W.; Patterson, D.; Hupert, M.; Shadpour, H.; Nikitopoulos, D.; Soper, S. A. *Analytical Chemistry* 2005, *77*, 658-666.

(31) Hashimoto, M.; Hupert, M. L.; Murphy, M. C.; Soper, S. A.; Cheng, Y. W.; Barany, F. *Analytical Chemistry* 2005, *77*, 3243-3255.

(32) Hashimoto, M.; Barany, F.; Soper, S. A. *Biosensors & Bioelectronics* 2006, *21*, 1915-1923.

(33) Dolnik, V. *Electrophoresis* 2004, *25*, 3589-3601.

(34) Smith, J. R.; Cicerone, M. T.; Meuse, C. W. *Langmuir* 2009, *25*, 4571-4578.

(35) Hashimoto, M.; Chen, P. C.; Mitchell, M. W.; Nikitopoulos, D. E.; Soper, S. A.; Murphy, M. C. *Lab on a Chip* 2004, *4*, 638-645.

(36) Hashimoto, M.; Barany, F.; Xu, F.; Soper, S. A. *Analyst* 2007, *132*, 913-921.

(37) Welzel, P. B. *Thermochimica Acta* 2002, *382*, 175-188.

(38) Horbett, T. A. In *Surfactant Science Series*, 2003, pp 393-413.

(39) Gray, J. J. *Current Opinion in Structural Biology* 2004, *14*, 110-115.

(40) Kricka, L. J.; Wilding, P. *Analytical and Bioanalytical Chemistry* 2003, *377*, 820-825.

(41) Wittwer, C. T.; Fillmore, G. C.; Garling, D. J. *Analytical Biochemistry* 1990, *186*, 328-331.

(42) Eigen, M.; Rigler, R. *Proceedings of the National Academy of Sciences of the United States of America* 1994, *91*, 5740-5747.

(43) Zarrin, F.; Dovichi, N. J. *Analytical Chemistry* 1985, *57*, 1826-1829.

(44) de Mello, A. J.; Edel, J. B. *Journal of Applied Physics* 2007, *101*.

(45) Haab, B. B.; Mathies, R. A. *Analytical Chemistry* 1999, *71*, 5137-5145.

(46) Wang, T. H.; Peng, Y. H.; Zhang, C. Y.; Wong, P. K.; Ho, C. M. *Journal of the American Chemical Society* 2005, *127*, 5354-5359.

(47) Lyon, W. A.; Nie, S. M. *Analytical Chemistry* 1997, *69*, 3400-3405.

(48) Foquet, M.; Korlach, J.; Zipfel, W. R.; Webb, W. W.; Craighead, H. G. *Analytical Chemistry* 2004, *76*, 1618-1626.

CHAPTER 7 CONCLUSIONS AND FUTURE WORKS

7.1 Conclusions

The past three decades has witnessed the evolvement of single-molecule fluorescence detection techniques from a purely proof-of-concept demonstration to an increasingly refined methodology that is becoming popular in many scientific disciplines. It has been adopted as an indispensible tool by practitioners to address some of the most challenging issues in biochemical studies, such as conformational dynamics of proteins, DNA sequencing, single cell gene expression profiling, etc. Combined with the advancement of microfluidic technology in recent years, it allows to integrating ultrasensitive single-molecule detection with different functional units such as reaction and separation into a single miniaturized device, which contains all features of a complete lab for bioanalysis with minimal sample consumption. This dissertation presents the efforts to build such a versatile platform on polymer substrate to perform enzymatic reaction, hydrodynamic sample delivery, confocal fluorescence detection as well as integrated thermal management all in the same device.

In Chapter 1, the role of pathogen detection and expression analysis of gene transcripts is signified. Various conventional techniques for genetic analysis like PCR and LDR were extensively reviewed. The fundamentals of FRET and molecular beacons were reviewed as well as their applications in bioanalytical applications. The principles and applications of single-molecule detection, spFRET, microfluidics and micro-total-analysis (μTAS) were thoroughly discussed. Various aspects of laser-induced fluorescence single-molecule detection with confocal configuration were specifically reviewed in more details in Chapter 2.

In Chapter 3, a benchtop laser-induced fluorescence single-molecule detection system with confocal setup was established and optimized. A solution of Cy5.5 labeled oligonucleotides in the concentration range of sub pico-molar was hydrodynamically driven through a COC microfluidic device and detection of single DNA molecules was demonstrated. The appropriate substrate was selected among different polymeric materials for fabrication of the microfluidic device to give the optimal SBR and maximal detection efficiency. This single-molecule detector has a broad dynamic range (up to 1 nM) and down to the “single molecule” level. The comprehensive study on different aspects of the single-molecule detection in this chapter sets up a solid foundation for designing and conducting single-molecule measurements in different bioanalytical applications.

In Chapter 4, the capability of LDR/spFRET has been demonstrated to provide a low limit-of-detection and rapid reporting of bacterial pathogens with strain specificity as well as the ability to distinguish Gram(+) from Gram(-) species. These measurements were performed directly within a COC microchip and demonstrated the ability to process the input DNA sample using LDR without a PCR amplification step and detect the products on-line in an automated fashion. With the measurements presented herein, the process time was found to be 2.6 min for a 2-cycle LDR and 19.2 min for a 20-cycle LDR. However, the larger number of cycles did improve the analytical sensitivity of the measurement. The relatively low volumetric flow rate of 0.78 µl/min used in this application will make it difficult to process samples in which the bacterial copy number per unit volume is low, which would then require some type of bacterial target pre-concentration prior to the LDR/spFRET measurement. This can be envisioned by using

an affinity pre-concentrator that can process large input volumes and select targets with high recovery. We have recently demonstrated the ability to use polyclonal antibodies to select certain pathogenic bacteria that are of low abundance from water samples.[1] Future work in our laboratory will integrate this rare cell selection device to LDR/spFRET to provide the ability to identify rare bacterial species from environmental samples in near real-time.

In Chapter 5, the expression level of the MMP-7 transcript, as well as several commonly used housekeeping genes were investigated using conventional end-point RT-PCR and RT-qPCR with real-time readout. In contrast to end-point RT-PCR, which only provided semi-quantitative information about the expression abundance of specific transcripts, RT-qPCR can accurately measure the copy number of the target transcript. The annealing temperatures for RT-PCR were optimized for each transcript and primer pairs to ensure that the highest efficiency of amplification and sensitivity of analysis are achieved during the real-time RT-qPCR measurement. Efficiency of reverse transcription was also evaluated using an innovative experimental design. In this study, HT-29 showed the highest level of MMP-7 expression of those cell lines studied, indicating up-regulation of MMP-7 gene activity with respect to the other cell lines. Of all the three selected housekeeping gene candidates evaluated, GAPDH showed the most stable expression level among all seven used cell lines and can serve as an internal control to normalize for artifacts in experimental variations.

In Chapter 6, COC microfluidic chip was used to accommodate LDR assay for expression profiling. The original mRNA transcript were linear amplified thus there is no bias introduced like RT-qPCR. The pristine COC microchip was treated with BSA

dynamic coating, and the hydrophilicity/hydrophobicity property was studied using water contact angle measurement. BSA treatment has a significant effect on preventing non-specific adsorption of ligase, thus ensure the yield of LDR assay on COC microchip. Significant improvements in assay sensitivity, even for the 2-cycle LDR, could be realized by simply increasing the sampling efficiency by reducing the channel size at the single-molecule detection zone of the chip and/or increasing the probe volume. In this application the channel dimension was tapered to 25 X 25 μm in cross section and the resulting sampling efficiency of generated rMBs was increased by a factor of 10. The sampling rate, which is determined by the processing volume flow rate, must be balanced by optimizing the performance of the LDR and the single-molecule detection efficiency. It was determined to be 0.78 μl/min in the present case. Expression analysis of MMP-7 gene was performed to analyze the transcriptional level from different cell lines. The results are comparable to the ones obtained using conventional RT-qPCR assay. In the spFRET readout, the Cy5/Cy5.5 dye pair are fairly photostable Iin this measurement, showing no bleaching from the photon burst trace data. Expression of AMPH transcript for stroke subtyping was also analyzed using this assay and can give a result in as short as 15 min.

7.2 Future Works

Expression analysis of specific mRNA transcript plays an important role in cancer prognosis and identifying the conditions of certain diseases. In chapter 6, expression analysis using microfluidic device combined with single-molecule readout format has been demonstrated for quantifying MMP-7 mRNA transcript in different cell lines. In this demonstration, the total RNA was extracted using benchtop devices including spin

columns and centrifuge, which is not only time consuming, but also incapable of analyzing rare samples. Hence it is profitable to incorporate RNA isolation into the microfluidic device in order to efficiently detect the expression level of sample with low abundance of transcripts in some applications. Marcus et al reported a microfluidic device on PDMS substrate to isolate mRNA from even a single cell, which demonstrated the capability of microfluidic device in obtaining mRNA transcript from rare biological sample.[2] Witek et al. successfully purified and concentrated genomic DNA from whole cell lysates using microfluidic chip fabricated from polycarbonate substrate.[3] Compared to genomic DNA, mRNA molecules are extremely susceptible to degradation, and special care has to be taken when RNA samples are handled on microfluidic device. The biocompatibility of mRNA with selected polymeric substrates also needs to be verified during RNA extraction implementation.

In mRNA transcript analysis it' is a common practice to for mRNA to be first reverse transcribed into cDNA, producing a mixture of DNA fragments served as template for subsequent PCR amplification. RT is known for its high variability, which compromises the accuracy of mRNA quantification. It has been reported in recent literature that a ligation reaction can actually take place on a RNA template using T4 DNA ligase.[4, 5] The reported ligation reaction performed using RNA template can discriminate a single-base mutation, which was only demonstrated using DNA template in previous studies. It is a great benefit to directly conduct the ligation detection reaction on RNA template in this LDR-spFRET assay. Challenges associated with uncertainties in the RT step will thus be eliminated.

In Chapter 6, the effect of cross-sectional dimension of microchannels on the sampling efficiency in photon burst detection has been illustrated, and reduction in size of microchannel improved the number of photon bursts that can be detected in fixed sampling time. However, the sampling efficiency is far from satisfactory because the cross-sectional dimension of the current design (25 X 25 μm) is still much larger than the beam waist of the focused laser spot. In the future, microchannel with cross-sectional dimension down to 1 X 1 μm will be fabricated to eventually detection all the LDR generated rMBs flowing through the detection window.

Current microfluidic chip patterned on COC substrate is merely a prototype of a miniaturized device to perform mRNA expression analysis with on-chip LDR and integrated single-molecule readout. In this proof-of-principle demonstration one single sample with one single gene was analyzed. Clinical applications such as drug candidate screening often requires that many samples or many genes need to be interrogated at the same time in a highly parallel fashion. Therefore, in the long run this assay will be able to perform high-throughput analysis by containing parallell fluidic channels, each of which corresponds to a single target transcript that needs to be examined.

The layout of the microfluidic chip for parallel expression analysis is shown schematically in Figure 7.1, which allows for simultaneously detection of four different cell samples or four different transcripts within the same cell. In this assay, the cells from blood draw are introduced into the device by a vacuum pump. They are lysed with their RNA molecules isolated and purified in the solid-phase extraction (SPE) bed. Then the LDR cocktail is introduced and mixed with the isolated mRNA molecules. The LDR mixture is subjected to a certain number of thermal cycles for ligation reaction, and the

produced molecular beacons are eluted through the detection window for photon burst registration and analysis. Since there are totally four different samples to be analyzed at the same time, four microfluidic channels are patterned, each of which is responsible for an individual transcript. A multi-element photo-detector is equipped and placed below the detection zone that has microchannels with reduced cross-sectional dimensions. Currently a 4-SPAD array is commercially available from PerkinElmer or PicoQuant and could be used as the detector of choice in this application to simultaneously count photon bursts from four different microchannels, which are directly related to the abundance of each transcript.

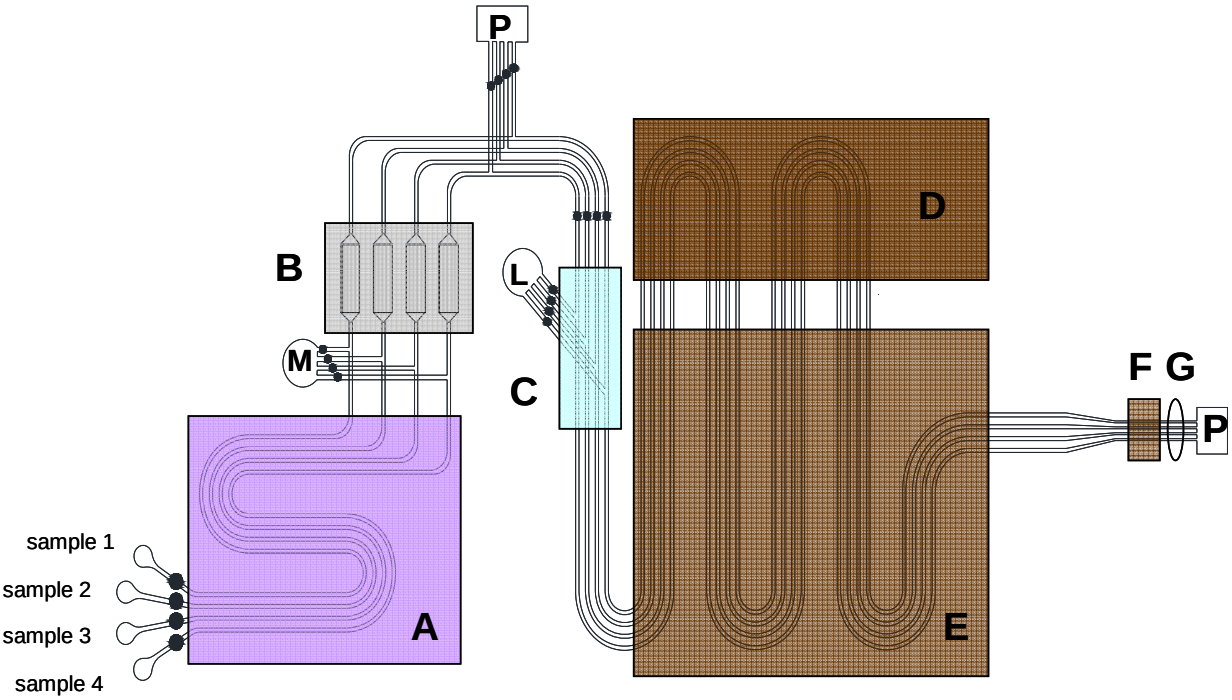


Figure 7.1 Layout of the COC microfluidic device for parallel expression analysis of mRNA. This device contains 6 functional sections: (A) cell lysis (B) solid-phase extraction (C) LDR mixing by a high aspect-ratio passive mixer (D) denaturation at 95ºC (E) ligation at 65ºC (F) detection of molecular beacons at 75ºC. The detection window (G) has reduced channel size for sampling efficiency improvement. Four different samples can be interrogated in this device simultaneously. The sample reservoir contains blood sample and lysis buffer. (M) – elution buffer, ethanol (L) – LDR cocktail (P) vacuum pump. Sample solution is driven hydrodanymically by a vacuum pump. Valves (•) are patterned in the chip to control flow of different reagents.

7.3 References

(1) Dharmasiri, U.; Witek, M. A.; Adams, A. A.; Osiri, J. K.; Hupert, M. L.; Bianchi, T. S.; Roelke, D. L.; Soper, S. A. *Analytical Chemistry* 2010, *82*, 2844-2849.

(2) Marcus, J. S.; Anderson, W. F.; Quake, S. R. *Analytical Chemistry* 2006, *78*, 3084-3089.

(3) Witek, M. A.; Llopis, S. D.; Wheatley, A.; McCarley, R. L.; Soper, S. A. *Nucleic Acids Research* 2006, *34*, -.

(4) Abe, H.; Kondo, Y.; Jinmei, H.; Abe, N.; Furukawa, K.; Uchiyama, A.; Tsuneda, S.; Aikawa, K.; Matsumoto, I.; Ito, Y. *Bioconjug Chem* 2008, *19*, 327-333.

(5) Tang, H.; Yang, X.; Wang, K.; Tan, W.; Li, H.; He, L.; Liu, B. *Talanta* 2008, *75*, 1388-1393.

APPENDIX: PERMISSIONS

PERMISSION/LICENSE IS GRANTED FOR YOUR ORDER AT NO CHARGE

This type of permission/license, instead of the standard Terms & Conditions, is sent to you because no fee is being charged for your order. Please note the following:

- Permission is granted for your request in both print and electronic formats.
- If figures and/or tables were requested, they may be adapted or used in part.
- Please print this page for your records and send a copy of it to your publisher/graduate school.
- Appropriate credit for the requested material should be given as follows: "Reprinted (adapted) with permission from (COMPLETE REFERENCE CITATION). Copyright (YEAR) American Chemical Society." Insert appropriate information in place of the capitalized words.
- One-time permission is granted only for the use specified in your request. No additional uses are granted (such as derivative works or other editions). For any other uses, please submit a new request.

VITA

Zhiyong Peng was born in Tianjin, China, to Yu-Rui Peng and Jin-Ying Gong, in 1972. He obtained his bachelor's degree from Tianjin University in 1995. In fall 2000, he attended Louisiana State University and obtained a degree of Master of Science, majoring in chemical engineering, which was conferred in 2003.

In the fall of 2003, Zhiyong Peng was enrolled in the doctoral program of Louisiana State University in the Chemistry Department. His work was performed under the direction of Professor Steven A. Soper. Zhiyong's research was focused on the development of a novel ligase detection reaction molecular assay coupled with online single-pair fluorescence resonance energy transfer readout that is carried out on thermoplastic microfluidic device for biological analysis. The degree of Doctor of Philosophy will be conferred at the May 2012 commencement ceremony.