A Label-free, Fluorescence Based Assay for Microarray

Sanjun Niu

Dissertation submitted to the faculty of the
Virginia Polytechnic Institute and State University
In partial fulfillment of the requirements for the degree of

Doctor of Philosophy
In
Chemical Engineering

Ravi F. Saraf, Chair
David F. Cox
S. Ted Oyama
Eva Marand
Sam English

July 26, 2004
Blacksburg, Virginia

Keywords: DNA, sequencing, DNAchip, Biochip, microarray, fluorescence, label-free, SNP, Polystyrene

A Label-free, Fluorescence Based Assay for Microarray

Sanjun Niu

(Abstract)

DNA chip technology has drawn tremendous attention since it emerged in the mid 90's as a method that expedites gene sequencing by over 100-fold. DNA chip, also called DNA microarray, is a combinatorial technology in which different single-stranded DNA (ssDNA) molecules of known sequences are immobilized at specific spots. The immobilized ssDNA strands are called probes. In application, the chip is exposed to a solution containing ssDNA of unknown sequence, called targets, which are labeled with fluorescent dyes. Due to specific molecular recognition among the base pairs in the DNA, the binding or hybridization occurs only when the probe and target sequences are complementary. The nucleotide sequence of the target is determined by imaging the fluorescence from the spots.

The uncertainty of background in signal detection and statistical error in data analysis, primarily due to the error in the DNA amplification process and statistical distribution of the tags in the target DNA, have become the fundamental barriers in bringing the technology into application for clinical diagnostics. Furthermore, the dye and tagging process are expensive, making the cost of DNA chips inhibitive for clinical testing. These limitations and challenges make it difficult to implement DNA chip methods as a diagnostic tool in a pathology laboratory. The objective of this dissertation research is to provide an alternative approach that will address the above challenges..

In this research, a label-free assay is designed and studied. Polystyrene (PS), a commonly used polymeric material, serves as the fluorescence agent. Probe ssDNA is covalently immobilized on polystyrene thin film that is supported by a reflecting substrate. When this chip is exposed to excitation light, fluorescence light intensity from PS is detected as the signal. Since the optical constants and conformations of ssDNA and dsDNA (double stranded DNA) are different, the measured fluorescence from PS changes for the same intensity of excitation light. The fluorescence contrast is used to quantify the amount of probe-target hybridization. A mathematical model that considers multiple reflections and scattering is developed to explain the mechanism of the fluorescence contrast which depends on the thickness of the PS film. Scattering is the dominant factor that contributes to the contrast. The potential of this assay to detect single nucleotide polymorphism is also tested.

ACKNOWLEDGEMENTS

I thank my advisor Dr. Ravi F. Saraf. His patience, motivation, and enlightening guidance has been great support to me. In personal life, I treasure the friendship he and his family shared with my family. I acknowledge Dr. S. Ted Oyama, Dr. Eva Marand, Dr. David F. Cox and Dr. Sam English for serving on my dissertation committee.

I thank my friends in Blacksburg. Some of them are my laboratory mates and fellow students such as Gaurav, Vivek, Vikas, Ping, Wei, and Jean. We have spent so much time together sharing knowledge, discussing issues and of course having fun. I surely will miss my badminton and volleyball playmates such as Shuming and Lianying. I thank my friends from the Blacksburg Church of Christ, especially the Gillies, the Moores, the Woods, the Foys and the Roberts.

My special gratitude would go to my family. To my wife, Ying, thank you for your unselfish love and for keeping me motivated. To my son, Winter, you are the sweetest and the ultimate driving force for your dad. To my parents, my brother and sister, your caring and encouragement are always the strong support behind me.

Last but not least, I want to thank God for all the blessings I have enjoyed.

Table of Contents

List of Tables…… vi
List of Figures…… vii

1. Introduction and background…… 01
    1.1 DNA…….. 01
    1.2 DNA sequencing……06
    1.3 DNA Microarray…….. 08
        1.3.1 Solid impermeable supports……. 09
        1.3.2 Fabrication of microarray…… 10
            1.3.2.1 In situ synthesis of probes……... 10
            1.3.2.2 Arrays of presynthesized probes…… 11
        1.3.3 Processing……. 14
            1.3.3.1 Target and labeling…… 14
            1.3.3.2 Detection…… 15
        1.3.4 Applications of DNA arrays……. 16
            1.3.4.1 Gene expression and discovery……..…..16
            1.3.4.2 Detection of mutations and polymorphisms……...18
            1.3.4.3 Mapping genomic libraries…….19
            1.3.4.4 Drug discovery and development…….. 20
        1.3.5 Challenges of the current DNA chip technology……..21
        1.3.6 Label-free detection of DNA hybridization……. 27

2. A label-free assay for DNA chip…….. 30
    2.1 Design and concept of the label-free DNA chip assay……30
    2.2 Device fabrication…….33
        2.2.1 Substrate preparation ……33
        2.2.2 Characterization of modified PS surface……34
        2.2.3 Immobilization of ssDNA……. 36
        2.2.4 Hybridization and dehybridization…… 37
    2.3 Device performance evaluation ……37
        2.3.1 Fluorescence measurement……37
        2.3.2 Quantification of DNA coverage and extent of hybridization.……. 37
        2.3.3 Mismatch detection……40
    2.4 Result and discussion……40
        2.4.1 Signal contrast……40
        2.4.2 Mismatch detection……45
        2.4.3 DNA surface coverage quantification and assay response estimation…….. 50
        2.4.4 Sensitivity…….. 57

2.5 Conclusion…… 58
3. Mechanism of the label-free DNA chip assay……...60
3.1 Experimental determination of film thickness and optical property ……...61
3.2 Result and discussion……63
3.2.1 Thickness and optical constants for PS…… 63
3.2.2 Thickness and optical constants of ssDNA and dsDNA monolayers……66
3.3 The model and discussion……75
3.4 Conclusion…… 84

4. Conclusions and future work…… 85
4.1 Conclusions…….. 85
4.2 Some preparations and recommendations for future work……...85
4.2.1 Instrument development to detect signal contrast on micro-scale spot and automation to perform array analysis.. 85
4.2.2 Fabrication of a DNA macro-array…… 87
4.2.3 Nanoparticle/polymer composite for higher contrast and visible light probe…….. 87

References……...90
Appendix…….95
Vita……..99

List of Tables

Table 2.1 Detection limits of various methods……. 59

Table 3.1 Optical constants of immobilized 27 nt ssDNA and hybridized (25%) dsDNA on PS/Si measured by ellipsometry. A plot of the data is shown in Fig. 3.3……... 69

Table 3.2 Parameters in the Lorentz model for fitting optical constants and thickness of DNA…….. 70

Table 4.1 Sequences to be used in mismatch sensitivity test. Bold characters stand for mismatched nucleotides……..88

List of Figures

Figure 1.1 A schematic diagram of the structure of a nucleotide. From Ref. 1. 02

Figure 1.2 A skeletal model of double-helical DNA. The opposite strands are different colors, and the bases are seen on edge. From Ref. 1…….. 03

Figure 1.3 Ball-and–stick models of A-T and G-C base pairs. From Ref. 1. … 04

Figure 1.4 The T_m is a function of the G-C content of the DNA; the higher the G-C percentage, the greater the T_m. From Ref. 1…… 05

Figure 1.5 Number fraction of DNA synthesized in PCR with n nucleotide that may be tagged. As p increases the fraction of untagged nucleotide decreases monotonically. However the number fraction for r > 0 has a maximum.…… 24

Figure 1.6 The η increases with p since more tagged bases are available for insertion. When the number of sites (i.e., n) increases, the probability of DNA with multiple tags also increases. Thus η exhibits a maximum at n = 10. Maximum for n = 5 occurs at a larger p…… 26

Figure 1.7 The background due to multiple tagging increases monotonically as p increases……. 26

Figure 2.1 The basic element of the DNA chip device is shown. For the study described, the substrate is Si wafer and the fluorescent (FL) film is polystyrene..……..…. 31

Figure 2.2 Fluorescence of PS detected with the top layer of DNA changes from ssDNA to dsDNA then to dehybridized ssDNA. …….42

Figure 2.3 Contrast is dependent on the thickness of PS. At range of PS thickness from 30 nm to 35 nm, contrast reaches the maximum. Contrast drops to nominally zero when PS thickness is beyond 70 nm …… 44

Figure 2.4 Contrast responses to washing temperature for perfect match and mismatch. At all temperatures, contrast of perfect match is higher than that of mismatch. Detection of mismatch is possible…….. 46

Figure 2.5 Hybridization temperature effect on contrast for perfect match and mismatch. Unit of temperature is °C…….. 49

Figure 2.6 Calibration curves of Cy5 and Cy3 dye. The scanned area is $325 \times 325\ \mu m^2$ with the optical slice for Cy5: 9.8 μm, Cy3: 8.4 μm. Fitting R-square for Cy5 is 0.92 and for Cy3 is 0.97…… 50

Figure 2.7 Surface coverage estimation by using confocal microscope. Probe ssDNA was labeled with Cy5 and target was labeled with Cy3.(a) and (b) Raw confocal images scanned at the same position but for Cy5 and Cy3 respectively. (c) Fluorescence counts across the DNA coverage and non-cover region (background) for the images in (a) and (b). Counts used in coverage calculation are background subtracted……52

Figure 2.8(a) Probe and hybridized target coverage with regard to hybridization time. Probe coverage has small variation but remains at a constant level for all the time. Target coverage increases as hybridization time increases and saturated after 11 h…… 55

Figure 2.8(b) PS fluorescence contrast as a function of hybridization time. Before contrast is saturated, it responds linearly with regard to hybridization time……. 55

Figure 2.8(c) PS fluorescence contrast as a function of hybridization extent. It shows linear relationship. Extent of hybridization is obtained by checking the coverage of probe and target with corresponding calibration curves…….. 56

Figure 3.1 Spin curve of PS. Solvent is toluene. Thickness of PS on Si was measured by x-ray reflectivity. Error bars for all the data points are negligible……. 65

Figure 3.2 Optical constants of polystyrene…… 65

Figure 3.3 Optical constants of immobilized ssDNA and hybridized dsDNA(25%)……. 68

Figure 3.4 The refractive index change between ssDNA and dsDNA as a function of wavelength. The calculated contrast from simple increase in molecular weight is negative and insignificant. The contrast is nominally constant in the near UV and visible range…… 74

Figure 3.5 Polarizability per nucleotide for ssDNA and dsDNA is calculated from the data in Table.3.1 and Equations (3.21) and (3.2.2)……. 74

Figure 3.6 Scheme for reflection model. Only the first three rays are considered in the simulation…….. 75

Figure 3.7 Comparison of the result from reflection model and the experiment data……. 77

Figure 3.8 Scheme of the scattering model……. 79

Figure 3.9 Comparison of the result from scattering model and the experiment data. Also shown is the result from the reflection model……. 83

Figure 4.1 Block diagram showing the setup of micro-photometer…… 86

Figure 4.2 Macro-array for mismatch test. The upper is the array design and lower is a picture of a fabricated sample…….. 89

Chapter 1

Introduction and background

The primary building material of a living organism is protein. DNA stores information on all the protein that the organism is capable of synthesizing. The type of protein made depends on the chemical structure of DNA, i. e. the sequence of the four basic units that construct DNA. Therefore if the DNA sequence is known, it is possible to predict, for example, the disease a person will contract in near or far future (gene sequence), the response to a particular drug on a particular person (drug efficacy) and the optimum response to an antigen in a particular person (personal medicine). DNA chip technology is arguably the most promising method to impact above mentioned applications of molecular diagnostics by deciphering DNA at high speed and low cost. In this chapter, characters of DNA, salient classical DNA sequencing techniques and DNA chip technology will be described.

1.1 DNA (1)

DNA is the abbreviation of deoxyribonucleic acid. DNA is a linear polymer of nucleotides. It is the key molecule that stores information required to build the cells and tissues of an organism. The number of units in a DNA molecule is typically in the millions.

DNA chains are composed of four monomers called nucleotides. A nucleotide has three parts: a five-carbon sugar (called pentose), a phosphate group, and a nitrogen

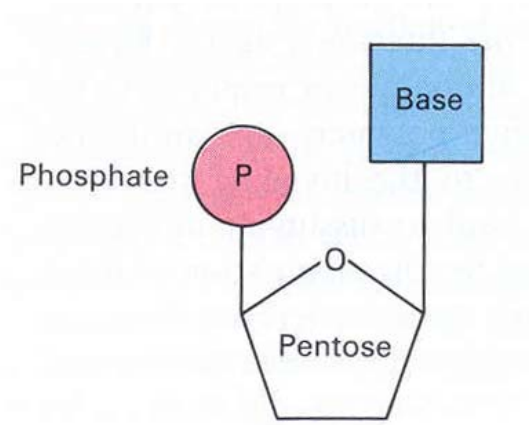


Fig.1.1. A schematic diagram of the structure of a nucleotide. From Ref. 1.

containing organic base as shown in Fig. 1.1. These bases can be either purines or pyrimidines. A pyrimidine has one ring while a purine has a pair of fused rings. The bases adenine, guanine, cytosine, and thymine are abbreviated A, G, C, and T. Since the four nucleotides differ in the base, these letters for the bases also represent the whole nucleotides containing that base. During the polymerization of nucleotides to form the nucleic acids, the hydroxyl group on the 3' carbon of a pentose of one nucleotide forms an ester bond with the phosphate of another nucleotide, eliminating a water molecule. Thus a nucleic acid molecule is a phosphate-pentose polymer (polyester) with purine and pyrimidine bases as side groups. A nucleic acid strand has a chemical orientation: the 3' end has a free hydroxyl group attached to the 3' carbon of a pentose, and the 5' end has a free hydroxyl group or phosphate group attached to the 5' carbon of a pentose.

In 1953, James D. Watson and Francis H. C. Crick first described the double-helical structure of DNA (2). DNA double helix consists of two phosphate-pentose strands that wind around each other; the base pairs are stacked in between the strands as shown in Fig. 1.2. The directions from 5' to 3' of the two strands are opposite, in other words, the orientation of the two strands is anti-parallel. The two strands are held together by hydrogen bonding and hydrophobic interactions. The hydrogen bonding between the bases on opposite strands are in strict match: A pairs with T by two hydrogen bonds; G pairs with C by three hydrogen bonds as shown in Fig 1.3. This base-pair

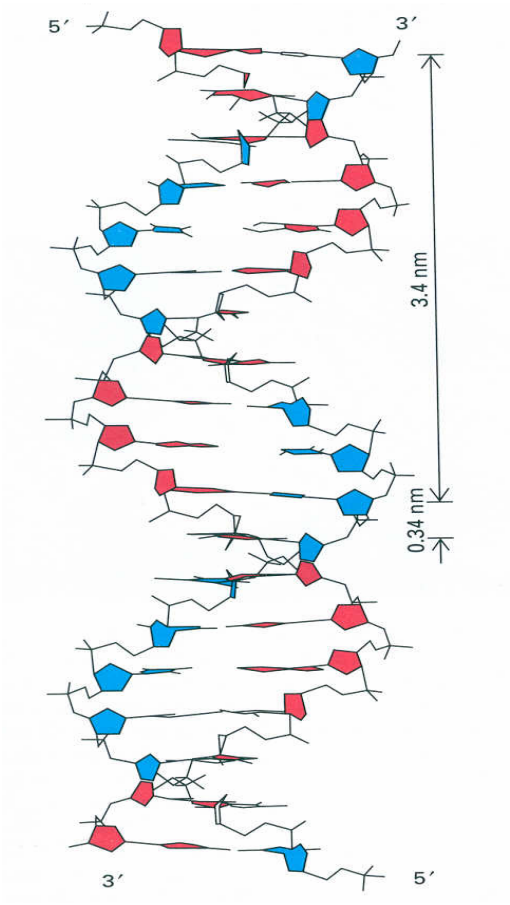


Fig.1.2. A skeletal model of double-helical DNA. The opposite strands are different colors, and the bases are seen on edge. From Ref. 1.

complementariness is determined by the structure, dimension, and chemical composition of the bases. Both the bases and the base-pairs are planar, and thus pairs stack on top of one another layer by layer. Hydrophobic and van der Waals interactions between adjacent base pairs in the stack contribute significantly to the overall stability of the double helix.

In double helixes, a purine pairs with a pyrimidine. However, this may be violated adding complexity to the structure. For example, a guanine residue (a purine) could theoretically form hydrogen bonds with a thymine (a pyrimidine), but this would cause a minor distortion in the helix. A similar distortion would allow the formation of a base pair between two pyrimidines, cytosine and thymine. Normally, G-T or C-T base pairs are not found in DNA. Native double stranded DNA is mostly right-handed helixes. The pitch of the helix is 3.4 nm, and there are about 10 pairs per turn.

If a solution of DNA is heated, the hydrogen bonds and other forces that stabilize the double helix will eventually be broken and the two strands will separate. This process is called denaturation or dehybridization of DNA or melt of DNA. The temperature at

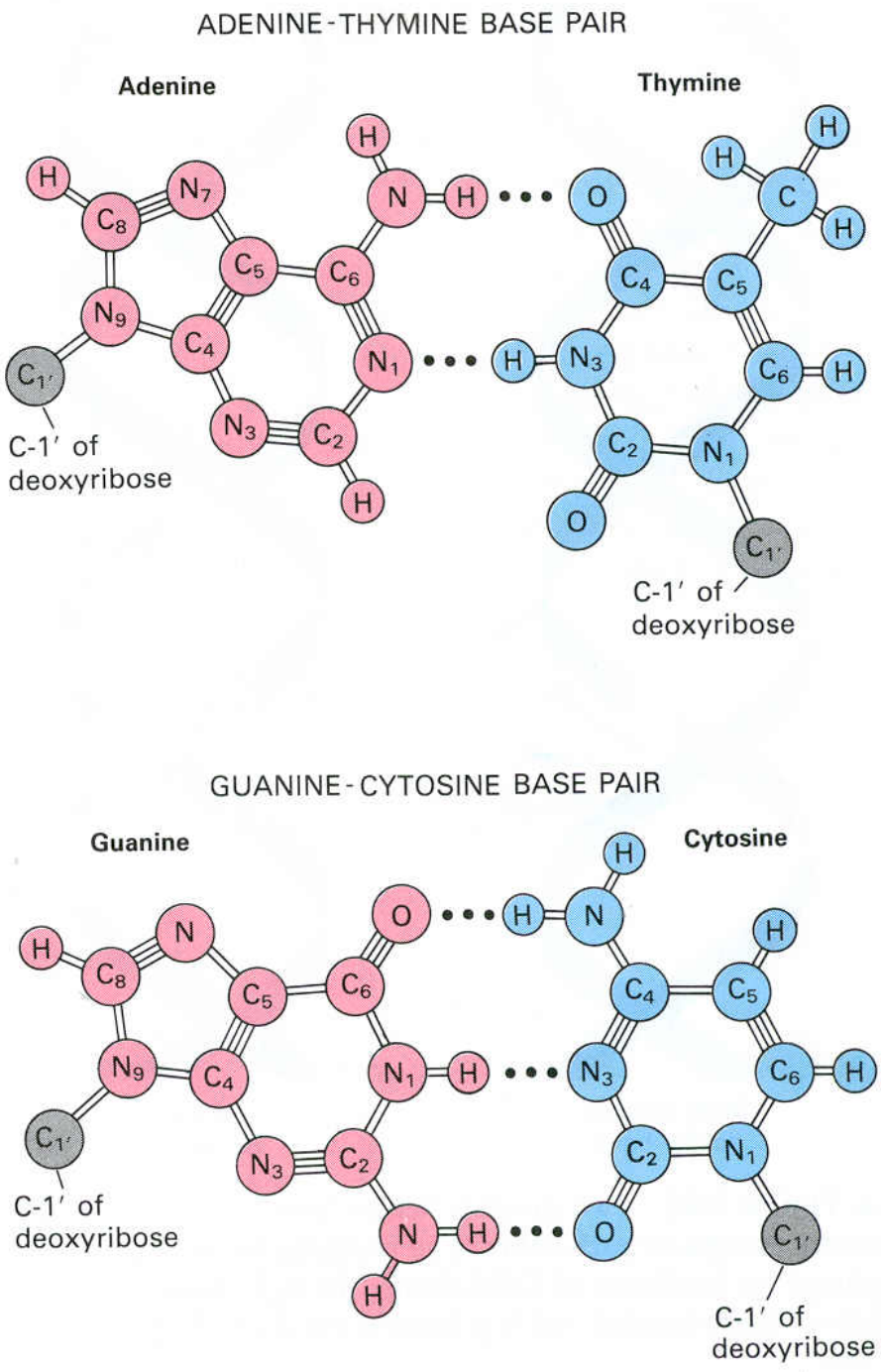


Fig. 1.3. Ball-and–stick models of A-T and G-C base pairs. From Ref. 1.

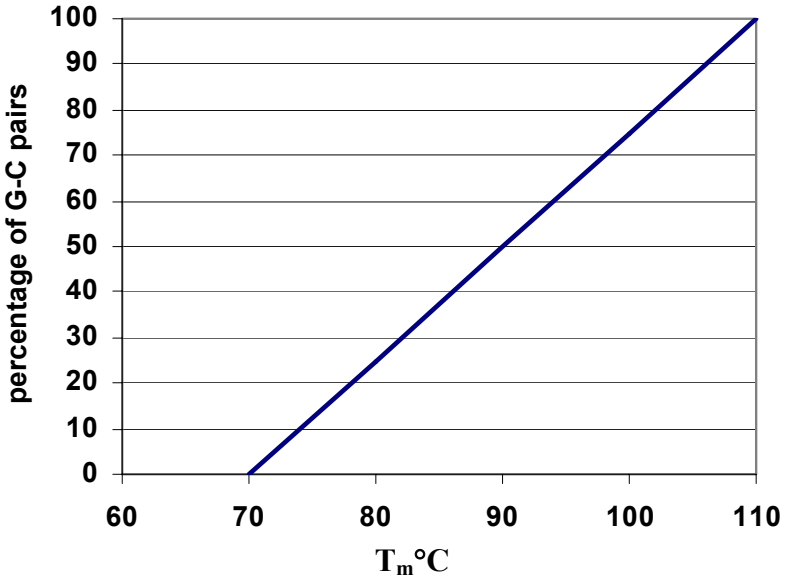


Fig. 1.4. The T_m is a function of the G-C content of the DNA; the higher the G-C percentage, the greater the T_m. From Ref. 1.

which the DNA denatures depends on the strength of the hydrogen bonds between the two strands. DNA that contains a higher number of G-C pairs require higher temperatures to denature because G-C pair has three hydrogen bonds compare to A-T pair which has only two (see Fig 1.4). A solution of double helixes with a low salt concentration tends to melt at a lower temperature than one with higher salt concentration. Exposure to some agents such as alkaline solutions and concentrated solutions of formamide or urea would also denature DNA by breaking hydrogen bonds. Using precise, tiny increments of temperature, A-T rich regions of a double stranded DNA can be melted while the entire strand of DNA is still held together by residual G-C rich regions.

The single stranded molecules that result from denaturation are stable, even if the temperature is lowered. They generally do not renature into the native double-stranded molecule, but form tangled balls. However, by adjusting the temperature and the salt concentration, the two complementary strands can be made to renature or hybridize and

form a double helix. This property is the basis of the powerful technique of DNA microarray.

1.2 DNA sequencing

Methods to determine the sequence of DNA were developed in the late 1970s and have revolutionized the science of molecular genetics. The DNA sequences of many different genes from diverse sources have been determined, and the information is stored in international databanks such as EMBL, GenBank, and DDBJ. Scientists now accept that sequence analysis will provide an increasingly useful approach to the characterization of biological systems. In the recent past, such large-scale sequencing projects were often dismissed as prohibitively expensive and of little short-term benefit. DNA sequencing was considered a repetitive and an unintellectual pursuit. However, this view is now changing and most scientists recognize the indispensable aspect of DNA sequencing. Recent technological advances, especially in the area of automated sequencing, have removed much of the drudgery that was associated with the technique. Innovative computer software has greatly simplified the analysis and manipulation of sequence data. Large-scale sequencing projects, such as the Human Genome Project, produce the DNA sequences of many unknown genes. Such data provide an impetus for molecular biologists to apply the techniques of reverse genetics to produce probes and antibodies that can be used to identify the developing cell (3).

The two original methods of DNA sequencing described in 1977 (4, 5) differ considerably in principle. The enzymatic (or dideoxy chain termination) method of Sanger (4) involves the synthesis of a DNA strand from a single-stranded template by a

DNA polymerase. The Maxam and Gilbert (or chemical degradation) method (5) involves chemical degradation of the original DNA. Both methods produce populations of radioactively labeled polynucleotides that begin from a fixed point and terminate at points dependent on the location of a particular base in the original DNA strand. The polynucleotides are separated by polyacrylamide gel electrophoresis, and the order of nucleotides in the original DNA can be read directly from an autoradiograph of the gel (6).

Although both techniques are still used today, there have been many changes and improvements to the original methods. While the chemical degradation method is still in use, the enzymatic chain termination method is by far the most popular and widely used technique for sequence determination. This process has been automated by utilizing fluorescent labeling instead of radioactive labeling, and the concepts of polymerase chain reaction (PCR) technology have been harnessed to enable the sequencing reaction to be 'cycled". Other recent innovations include solid phase sequencing, and the use of robotic work stations to automate sample preparation and sequencing reaction.

The method of sequencing by hybridization, in short SBH, was first disclosed in a patent application filed by Drmanac and Crkvenjakov (7) in 1987. This method is based on the specific molecular recognition in DNA hybridization. Well-known techniques such as Southern blot and reverse dot blot are examples of SBH (8). In SBH, a short ssDNA of known sequence is used for the probe. The probe is exposed to *target* that is of unknown sequence. The collection of sequences of targets is labeled with fluorescence dyes. Hybridization is recognized by detecting fluorescence from the target. SBH solved the repetitiveness problem of the previous DNA sequencing methods. For example if the

sequence of a gene is known, possibly only one experiment is needed to determine the existence of this gene. In 1979, Wallace et al. (9) demonstrated that oligonucleotide probes as short as 11 bases in length can be used to discriminate between perfect probe-target duplexes and those containing a single internal base mismatch. In 1989, Saiki et al. (10) described a "reverse dot blot" method of testing a patient's DNA samples for known mutations based on specificity of oligonucleotide hybridization. In this method, patient samples prepared by polymerase chain reaction (PCR) were tested for hybridization against an array of oligonucleotides designed to complement normal sequences and known mutations. In 1986, Breslauer et al. (11) demonstrated methods of predicting sequence-dependent oligonucleotide hybrid stability. In a 1988 patent application, Southern (12) proposed a method for combinatorial *in situ* synthesis of large oligonucleotide arrays for mutation detection or complete sequencing. Lysov et al. (13) proposed using an array of physically attached oligonucleotides for SBH analysis. The high accuracy of the SBH method was demonstrated in a blind experiment in which 330 bases of DNA in three samples were sequenced without a single error (14). These techniques are considered as the forerunner and macro version of the current powerful DNA microarray technology.

1.3 DNA Microarray

DNA microarray are arrays of probe ssDNA spots on a solid substrate, called the chip. Since each spot of probe interrogates a specific sequence, analysis of large number of DNA sequences can be simultaneously, i.e. combinatorial analysis. Each spot on the array is in micron scale so that as many as 100,000 spots, i.e. 100,000 different probes,

can be arrayed on a chip of the same size as a microscope slide. The principle of this technology is based on specific molecular recognition interactions between the arrayed probe ssDNA and the test DNA of interest, i.e. sequencing by hybridization as described in the last section. Oligonucleotide probes of known sequence are tested for complementary pairing (e.g. hybridization) against a DNA target. The process, which in some respects is similar to a keyword search of text by an Internet browser, is designed to identify matching sequence strings within the target DNA. In one scenario, software programs are then used to assemble the target sequence by ordering the set of overlapping, high-scoring probes.

1.3.1 Solid impermeable supports

All blotting procedures (15) unavoidably use porous supports, which have some advantages. For example, since the pores of the supports provide a larger total surface for attachment, it is possible to load quite large amounts of nucleic acids on a small area. Additionally, the nucleic acids can be applied in a relatively large volume as they soak into the pores of the membrane without excessive lateral spreading. However, the boundaries and shapes of the spots are poorly defined and the amount of oligonucleotide deposited is difficult to control accurately. The demands of genome projects brought the need for analysis on a much larger scale. Although it was possible to increase the area of dot blots, it was not possible to reduce the size of spots beyond certain limits, or to control their size and shape on a porous membrane. These factors become crucial for automated analysis of hybridization signals, when it is necessary to locate accurately the positions of the spots and to know in advance their precise shape and size. An additional

advantage of solid supports is their dimensional stability and rigidity. In contrast, permeable membranes swell in solvent and tend to shrink and distort when dried. Thus, it is not possible to locate spots with the high precision that can be achieved on a rigid substrate.

The introduction of impermeable substrates was a major departure that has several advantages over porous supports. Since the nucleic acids form a monolayer close to saturated concentration, the number density of probe per spot is constant over the array. Direct contact with the solution phase leads to faster reaction, because molecules do not have to diffuse into and out of the pores. All stages of the binding process benefit from high accessibility and low steric hindrance. Washing is also efficient since the excessive labeled material does not have to diffuse out of the pores of the membrane. All these aspects speed up the procedure, improves reproducibility, and reduces background. All these factors contribute to achieving reliable hybridization signals at a high level of accuracy needed to distinguish small differences in gene sequences.

1.3.2 Fabrication of microarray

1.3.2.1 *In situ* synthesis of probes

A further benefit of using impermeable supports is that it permits array fabrication by *in situ* synthesis of nucleic acids on the surface. *In situ* synthesis has a number of advantages over deposition of presynthesized probes. It combines the advantages of solid-phase synthesis (high coupling yields and high purity) with those of combinatorial chemistry (a large diversity of compounds can be made in a few steps) (16). Typically, the number of coupling steps is a small multiple of the length of probes made on the

array. For example, there are combinatorial methods for making all 4^8 octanucleotides with only eight coupling steps (17). This is significantly smaller than 8 × 65,536 = 524,288 steps if the probes are made individually. Two types of approach were developed to confine the synthesis to small, defined regions of the solid support.

The simpler approach adapted existing chemistry, delivering reagents to confined areas: e.g., using drop-on-demand ink-jet technology (18) or irrigating the surface through flow channels (17, 19, 20). A more specialized method adapted the photolithographic methods used in the semiconductor industry (21) and required the development of new photolabile protecting groups for nucleotide precursor. The ink-jet method is flexible and makes economical use of the most expensive reagents. The dimensions of spots made by ink-jet are about 100-150 μm in diameter, at 100-200 μm centers and flow channels generate cells with sides equal to the narrowest channel width. It is possible by micromachining to make flow channels < 100 μm wide. The more expensive photolithography method makes smaller arrays than the ink-jet and flow channel methods. Arrays fabricated by photolithography with 65,536 oligonucleotides in an area 1.28 × 1.28 cm are commercially available. A disadvantage of the photolithography method is that coupling yields (about 95%) (22) are lower than for conventional chemicals (> 99%). Thus, the yield of a 20-mer will be about 36% as compared with > 80%.

1.3.2.2 Arrays of presynthesized probes

The method to making arrays by spotting probes of cloned sequences, or PCR product, has been straightforward. The method of application is an adaptation of a

computer-controlled xyz stage with a head carrying a pin or pen device to pick up small drops of solution from the multi-well plates and carry them to the surface. For chemically synthesized probes, attachment is favored, and various methods for attachment to solid supports have been used (23). Quality control is becoming an important consideration, especially as nucleic acid arrays enter clinical diagnostic applications. Presynthesized probes seem to be advantageous since their quality can be checked before they are attached to the surface.

. A substantial body of work on immobilization of oligonucleotides to solid supports has been reported. Efficiency of coupling, the stability of attachment and the steric accessibility of oligonucleotides for hybridization of these technologies are important considerations. There are basically two types of immobilization techniques: (i) non-covalent adsorption and (ii) covalent attachment. Non-covalent methods depend on the ionic attraction between the negative charge on the DNA backbone and the positive charged substrate, for example, polylysine coated glass slides (24, 25). There are quite a few reports on non-covalent coupling techniques for different applications (26-30). However, most of them result in poorly defined strand orientation, low packing density, and limited mobility of the strands. The immobilization is susceptible to release from the surface when exposed to strong ionic strength solution and/or high temperature. The covalent attachment method overcame these drawbacks thus has been extensively explored and practiced. The chemistry of immobilization varies depending on the substrates. Most used substrates include glass, quartz, optical fibers, gold, and polymers.

The well-known and versatile silianization chemistry has been widely used to attach DNA onto glass slides (31-34). There are also methods to attach oligonucleotides

to underivatized glass. In one method (35), the 3'-propanolamine derivatized oligos are coupled directly to glass slides. The linkage is stable in hot water and attachment density is 10^{12}-10^{13} oligos/cm^3. Avidin can bind to glass by physical adsorption and glutaraldehyde treatment will form crosslinks between avidin and the glass thus stabilizing the connection. A biotinylated DNA then can bind tightly to the avidin-conjugated glass (36).

Immobilization of oligos to a gold surface are used primarily in surface plasmon resonance technology. Gold-sulfur bond chemistry is dominant in this category (37-40). A commonly used coupling agent is 11-mercaptoundecanoic acid, where the thiol group on one end binds to Au and the carboxylic group on the other end is versatile for further attachment to derivatized oligonucleotides.

Commonly available polymers such as polypropylene (PP) (41-43) and polystyrene (PS) (44) have been reported to serve as support for oligonucleotides attachment. Dry chemistry such as radio frequency plasma discharge or wet chemistry is used to functionalize the inert polymer surface with reactive groups including amino and carboxylic groups. There are several commercially available products. For example, Nalge Nunc International's "NucleoLink" microwells are amine modified polystyrene (45). Motorola's Surmodics surface activation technology applies to a few kinds of plastics including PS, PP, polymethylmathacrylate (PMMA) and polycarbonate (PC) (46). Polymers are inexpensive and easy to fabricate in mass production.

In summary, using presynthesized probes for microarray fabrication is straightforward. There are several points needed to be taken into consideration for the probe immobilization procedures. The attachment of oligo probes should be stable.

Density of the bound probes should fall in a suitable range for maximum hybridization efficiency. Chemical properties (hydrophobicity, charge) of linkers between the surface and the probe play important roles in hybridization kinetics, efficiency and specificity.

1.3.3 Processing

1.3.3.1 Target and labeling

The target nucleic acid to be analyzed can be RNA or DNA. For current DNA chip technology, targets are labeled so that the hybridization can be directly detected by fluorescence microscopy. PCR, which is commonly used, produces targets that are double stranded and unsuitable for hybridization to oligonucleotides. Asymmetric amplification makes enough single strands, but a better method is to destroy one strand by treatment with exonuclease (47, 48). Modifications to one of the PCR primers prevent access of the exonuclease to the strand that it primes. This method was found to be easy and reliable, and to produce targets that hybridize well with the probes. Alternatively, if an appropriate promoter is incorporated into the sequence of one of the PCR primers, a single-stranded transcript can be made readily by a bacterial polymerase, such as the T7 polymerase (19). This method has several advantages: there is substantial additional amplification as a result of the transcription, and the RNA can be labeled to a high specific activity by incorporating labeled precursors. However, RNA molecules fold as a result of intramolecular base pairing to form stable structures that interfere with the hybridization process—the corresponding structures in DNA are less stable. The problem with RNA can be partly relieved by degrading the transcripts to fragments of a size comparable with that of the oligonucleotide probes. The problem is less severe for arrays

of spotted cDNAs (complementary DNA) because hybridization can be carried out at higher temperatures, which melt the intramolecular base pairing.

Radioactivity provides sensitive detection, but it has a wide "shine." The limitation on spatial resolution is not a problem with membranes based blotting where the dimensions of the features are large. However for good spatial resolution for chip on solid substrates such as glass, fluorescent labels are suitable because imaging at high spatial resolution is possible by a method such as confocal microscopy. Unfortunately, fluorescent labels are expensive, unstable and generate statistical errors for data analysis, as will be described in section 1.3.5. The proposed method is adaptable to fluorescence read-out means but does not require any labeling of the targets. Therefore it can potentially overcome the difficulties related to the use of fluorescent dyes.

1.3.3.2 Detection

Radioactivity detection has many advantages. It has a wide dynamic range. Even with a single exposure, the range can be extended by varying the exposure time. Quantitation is very precise. It is easy to label targets to a high specific activity by a number of well-established methods. ^{32}P has a wide shine, but ^{33}P can be imaged by phosphorimaging to a resolution of about 200 μm. The resolution is limited by the grain structure of the phosphorimager screen.

Fluorescent labels have become more and more popular because of advantages such as longer shelf life, less environmental and health hazard compared to radioactive labels. They also enable double labeling (with two different dyes of one target) and high-resolution imaging. Confocal microscopy reduces noise by removing background due to

out-of-focus light. Several readers that apply the confocal principle to a large format have been developed and are now commercially available. However, as mentioned earlier, there are several problems associated with fluorescent labeling and will be reviewed in section 1.3.5.

1.3.4 Applications of DNA arrays

1.3.4.1 Gene expression and discovery

DNA chips have been used to measure expression levels of genes in plant (24, 49), yeast (50, 51) and human (52, 53) samples.

Schena et al. (52) monitored the expression of 1046 human cDNAs of unknown sequence using two-color differential expression analysis of heat shock or phorbol ester-regulated genes. The cDNAs of induced genes were sequenced and identified by comparison to known structures. Heat shock resulted in the induction of known heat shock genes for molecular chaperones and mediators of molecular degradation. Similarly, phorbol ester exposures resulted in the detection of genes characteristic of the phorbol ester signaling pathway. Additionally, three known genes that were expressed at low levels had not been previously assigned to this pathway. Four new genes were identified, each expressed at a low level. It is likely that previous conventional screens were not sufficiently sensitive for their detection. These experiments demonstrated the ability of DNA arrays to rapidly provide data for correlation of gene expression to biochemical pathways.

This technology was extended to the study of gene expression characteristic of the inflammatory diseases rheumatoid arthritis (RA) and inflammatory bowel disease (IBD)

(54). Probes were produced from RA tissue or IBD mucosa labeled with either Cy3 or Cy5 fluorophores and exposed to a microarray consisting of cDNA targets from genes known to be involved in the disease processes. Genes known to be expressed in inflammatory disease were observed. Some genes were expressed that had not previously been associated with inflammatory diseases. The newly assigned genes could become therapeutic targets. Certain genes were more strongly expressed in RA tissue, demonstrating the ability of DNA chips to rapidly provide information about the genetic basis of disease.

DeRisi et al. (55) investigated the genetic basis of tumorigenicity by the use of two color fluorescence dyes. The tumorigenic properties of human melanoma cell line UACC-903 can be reversed by insertion of human chromosome 6, and probes were made from both types of cells and labeled with two different fluorescence dyes. These probes were mixed and applied to a microarray of 1,161 cDNAs selected to study tumor suppression. The differential expression values provided valuable information about the molecular pathology of this tumor. For example, elevated expression of the WAF1 (p21) gene, which mediated *p53* tumor suppression, was observed only from a nontumorigenic probe. Likewise, elevated levels of human brown locus protein gene were observed only for tumorigenic cells.

Complementary DNA microarrays containing virtually all of the genes of *Saccharomyces cerevisiae* have been fabricated (51). This permitted the genome-wide study of the effects of the diauxic shift from anaerobic to aerobic metabolism under glucose limitation and the concomitant switch to ethanol as a carbon source. The significance of this work is that it mapped the changes in expression of genes with known

function to their metabolic pathways and vividly showed which metabolic pathways were reprogrammed by the shift. Additionally the expression patterns of many previously unknown genes were obtained.

Dramanac et al. (56) have used massive cDNA microarrays for large-scale gene discovery in infant brain tissue. Their protocols are used for large-scale, commercial, gene screening in which microarrays of 55,000 cDNAs screen 800,000 clones per month.

1.3.4.2 Detection of mutations and polymorphisms

Hacia et al. (57) used a DNA chip containing 96,600 oligonucleotide probes to detect all possible heterozygous mutations in the 3.45 kb exon 11 of the BRCA 1 breast and ovarian cancer gene. Controls and samples were differentiated by the two-color fluorescence protocol. Fifteen patient samples were analyzed with one false negative and eight single base pair polymorphisms (single nucleic polymorphisms, in short SNPs) were detected.

Lipshutz et al. (58) demonstrated that DNA chips could be used for screening of mutations in the reverse transcriptase and protease genes in the HIV-1 virus. Such mutations can cause resistance to antibiotics such as AZT. Kozal et al. (59) used a genechip to study the occurrence of polymorphisms in the HIV-1 clade protease inhibitors. Genechip results were checked with those obtained by Sanger sequencing (4). An agreement of 98% over 114 samples was obtained between the two methods and a large degree of polymorphism was observed.

Mutations in the cystic fibrosis transmembrane conductance regulator (CFTR) were studied using the Affymetrix chip (60). An array was designed to detect known

deletion, insertion or base substitution mutations in exons 10 and 11 of CFTR. Ten unknown patient samples were tested and the results were confirmed exactly by PCR product restriction fragment analysis performed by independent workers.

Chee et al. (61) fabricated a DNA chip containing 135,000 25-mer probes for the probing of the 16.6 kb human mitochondrial genome. Two-color fluorescence was used to compare mutated mitochondrial DNA (mtDNA) to control mtDNA. Mitochondrial genomes from 10 individuals were analyzed and 505 polymorphisms were detected. Each sample could be read in 12 min. It was estimated that during a working day 40 mtDNA genomes could be read compared to two by a modern gel sequencer.

Mutations at three positions in the β–globin gene from blood samples were detected by a 10-mer microarray (62). Contiguous stacking hybridization was then used to enhance detection of the IVS-1-1 (G→A) mutation by addition of a soluble pentamer probe.

1.3.4.3 Mapping genomic libraries

DNAchips have been used for mapping genomic libraries by determining the order of overlapping clones (63). *S. cerevisiae* cosmid DNA was prepared from 12 genomic clones. After PCR amplification, labeling with fluorescent marker and production of ssDNA, the product was hybridized to a 256-feature array. Fluorescence intensities were normalized and a correlation score was determined for each adjacent clone pair by statistical analysis. The 10 cosmids that gave the strongest signals were arranged as a continuous sequence, and in the correct order, by using correlation scores in a simulated annealing procedure. This approach demonstrated the utility of DNA chip

technique to map clones in a highly parallel fashion. Since all the chemical reactions for each clone were implemented in a single test tube, rapid parallel assays of many clones could be envisioned. Furthermore, all the assay steps were amenable to automation for increased throughput. It is estimated that a single operator could map several hundred clones per day.

1.3.4.4 Drug discovery and development

With an emphasis on functional genomics rather than sequencing, drug discovery programs are using custom chips to find lead compounds. It has already been shown that it is possible to treat cells with compounds and compare the resulting patterns of gene expression with patterns previously obtained when treating cells in known ways, thereby identifying which proteins or targets the compound is altering (64). Such *in vitro* target identification should greatly improve the inefficient conventional methods of developing drugs. Because animal testing of compounds is expensive, time consuming and has other negative aspects, DNA microarrays are likely to improve the efficiency of drug discovery by supplementing the information obtained by traditional animal testing (65).

Microarrays have the ability to rapidly survey and compare gene expression levels between reference and test samples, thus can highlight differences between normal and disease-related genes and document the effects of drugs on gene function. Pharmaceutical companies use microarray technology to predict and survey side effects of drugs at a genetic level (66). Gene expression profiling has been used to predict and understand the toxicity of drugs. Toxicogenomics, the evaluation of gene expression profiles in response

to toxic compounds by use of bioinformatics, has been facilitated by use of DNA microarrays (67).

1.3.5 Challenges of the current DNA chip technology

In current DNA chip technology, fluorescence labeling is the pervasive detection method. Fluorescence label, combined with an appropriate imaging instrument, is a sensitive and quantitative method that is widely used in molecular biology. Fluorescence detection offers a number of important advantages. The sensitivity of fluorescence applications is approaching that of radioisotope applications and is improved by the designing of better imaging instruments and invention of "brighter" dyes. Fluorescence molecules are low hazardous with minimal environment and health related problems. With respect to stability, if stored properly, fluorescence molecules are stable for more than six months while some radioisotopes used for study decay very fast. For example, ^{32}P decays significantly in about a week. Because of their long shelf-life, fluorescence labeled reagents can be prepared in large batches that can be standardized and used for extended periods, thus minimizing reagent variability between assays. Other advantages of fluorescence include that detection is a linear response with respect to the amount of fluorescence molecules over a wide range, and detection angle can avoid excitation light, thus minimizing background signal.

In spite of these advantages of the fluorescence labeling method, there are significant challenges, such as photobleaching of dyes and the interference with double helix formation due to incorporation of dyes on one of the oligo strands. Berlier et al. (68) tested the photobleaching of commonly used Cy5 and Cy3 dyes. After 95 s of

constant exposure, Cy3 dye retained 75% of its initial fluorescence and Cy5 only retained 55%. The experiment was performed on a E400 fluorescence microscope with light source of a 100W mercury arc lamp. Armersham Pharmacia (69) reported that Cy5 dye loses 50% after 24 h exposure to laboratory strip lighting and 1% loss per scan with red laser scans. Thus, photobleaching of dyes affects the robustness of DNA chip methods for potentially routine clinical applications. Another challenge for fluorescence labeling is that the incorporation of dye molecules interferes with the probe-target hybridization and stability of the double helix duplex. In an extreme case, due to the label interference, the probe-target hybridization between sequences with a mismatch occurs more efficiently than a perfectly matched pair (70). The labels affect the integrity of the structure as manifested by the reduction in the melting temperature (71, 72). The decrease in melting temperature is especially a concern because it is perhaps the most reliable method to determine whether the probe-target hybridization is perfectly matched or mismatched (73, 74). The apparent drop in the melting temperature due to the label in a perfectly matched system can therefore lead to an erroneous result that the pair has a mismatch. For most fluorescence dyes such as Cy3, Cy5, and FTIC, the difference in wavelength between excitation and emission light is < 25 nm, therefore requiring narrow band-width filters which attenuate the emission light signal, further limiting the sensitivity. Although there are some reports regarding loss of fluorescence on conjugation to proteins (68), none has been published on loss of fluorescence conjugating to DNA.

Another challenge of fluorescence labeling is the statistical distribution of dye over the target DNA molecules. Niu and Saraf (75) demonstrated this by a quantitative model. In current DNA chip applications, fluorescence labels are incorporated during the

Polymerase Chain Reaction (PCR) amplification process. In the PCR process, the DNA is replicated by polymerizing the nucleotides in accordance with the DNA fragment of interest. Since the nucleotides are composed of A, G, T and C bases, if one of the bases is tagged with a dye, then it will be incorporated in the synthesized DNA during PCR.

In a PCR, let the DNA fragment (typically ~25 bases) for replication has n G bases. Let p mole fractions of the G bases in the reaction mixture be tagged with a dye. Assuming no change in reactivity of the base due to tagging, the number fraction of ssDNA synthesized in the PCR that have r tags is,

$$P(n,r) = \frac{n!}{(n-r)!r!} p^r (1-p)^{(n-r)} \tag{1.1}$$

Tagged nucleotide bases are expensive. Thus, the mole fraction p is low, usually in 10% range. Fig.1.5 (a) and 1.5 (b) show the effect of p on number fraction of synthesized DNA with no tags (i.e., r = 0), 1 tag (i.e., r = 1) and two tags (i.e., r = 2) for n = 5 and 10. At p = 0, P (n, 0) = 1 as expected for both n = 5 and 10. As expected, P (n, 0) decreases monotonically to 0 as p increases to 1. In contrast, as expected, at low p, P (n, 1) and P (n, 2) increase monotonically as p increases. However, note that the number fraction is below 40% for the conditions chosen indicating a high level of inefficiency in the tagging process. Interestingly, P (n, 1) shows a maximum for both n = 5 and 10. This is because as p increases there are more DNA fragments with (undesired) multiple tags than a single tag. P (n, 2) exhibits similar behavior; however for the range shown, the maximum is visible only for n = 10, but not for n = 5

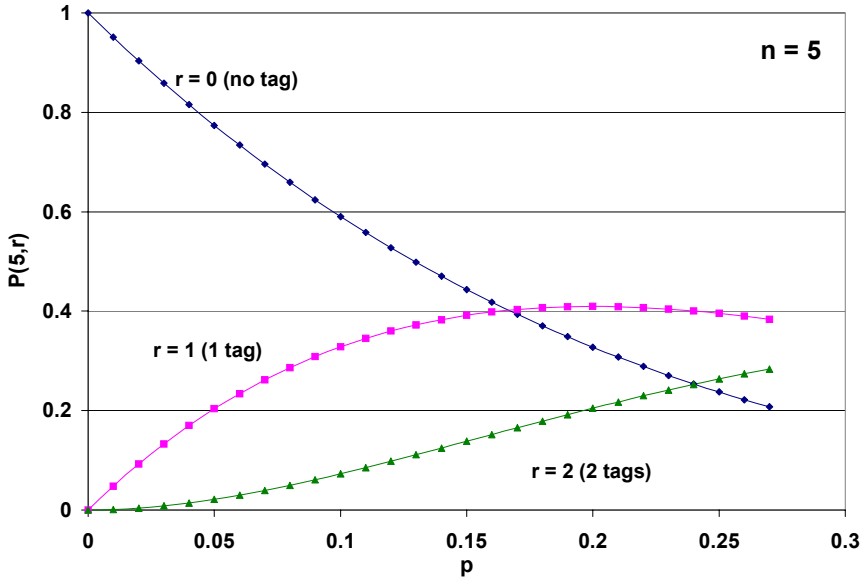


Fig. 1.5 (a)

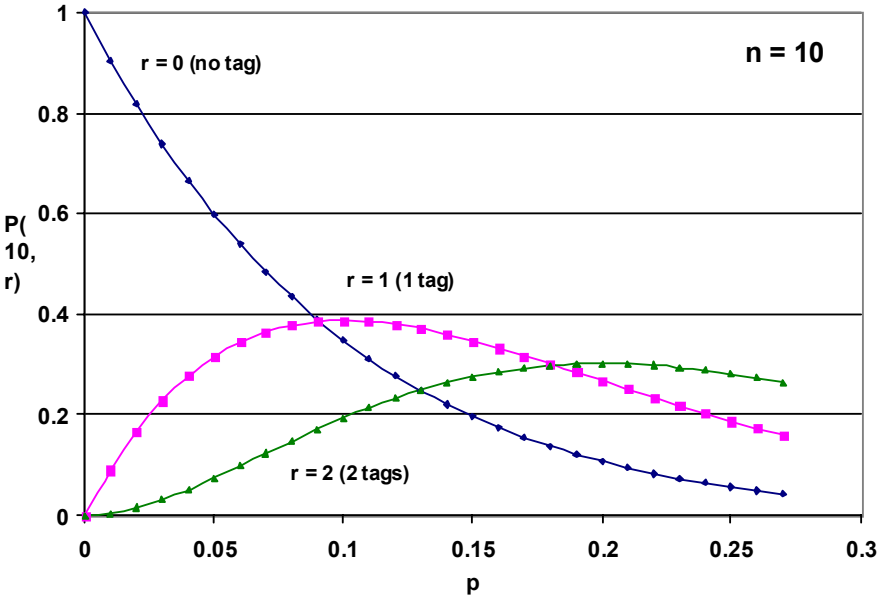


Fig. 1.5(b)

Fig. 1.5. Number fraction of DNA synthesized in PCR with n nucleotide that may be tagged. As p increases the fraction of untagged nucleotide decreases monotonically. However the number fraction for r > 0 has a maximum.

Since P (n, 0) > 0, it implies that there are significant amounts of DNA fragments synthesized in the PCR reactor that do not have a tag. We define a tag-efficiency-factor, $\eta = P(n, 1)/(1 + P(n, 0))$ such that $\eta = P(n, 1)$ (< 1) for p = 1. In an ideal case, if all the ssDNA synthesized have only one tag, P (n, 1) = 1, thus $\eta = 1$. Fig.1.6 shows the change in η as p increases. High η is expected as p increases. However for large values of p, η drops because most of the ssDNA synthesized will have more than one tag.

If we assume that p is small (usually due to cost reasons), then the number of ssDNA with more than two tags will be relatively small. For example, for n = 5, with p as high as 0.2, P (5,0) = 0.328, P (5,1) = 0.401, P (5,2) = 0.205 and P (5,3) = 0.051, i.e., DNA with three tags is ~8 and 4 times less than DNA with one and two tags, respectively. Thus, P (n, r) for r > 2 are ignored. Furthermore, P (n, 0) does not contribute to the PL signal. The background due to excess tagging is then defined as relative amount of ssDNA with two tags relative to one tag. In other words, the background is defined as $\phi = P(n, 2)/P(n, 1)$. Fig. 1.7 shows the calculated background at n = 5 and 10. At p > 0.2, ssDNA with two tags are more than single tag ssDNA for n =10 and half the amount for n = 5.

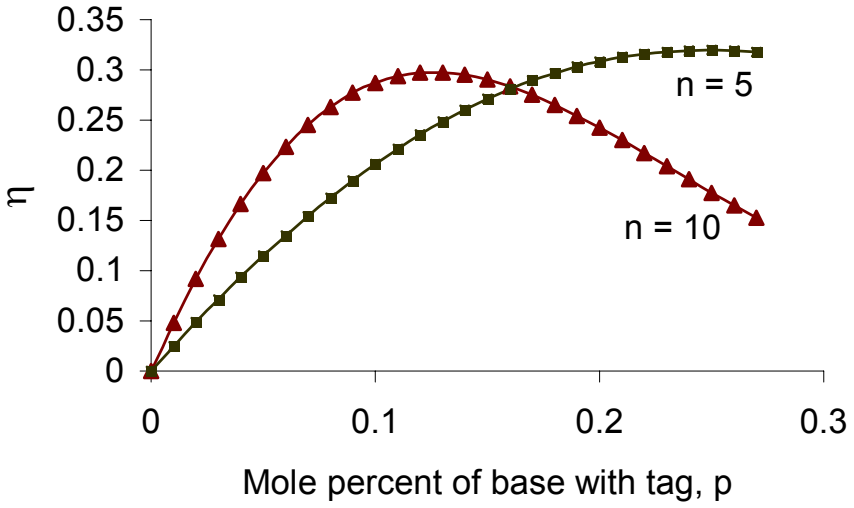


Fig. 1.6. The η increases with p since more tagged bases are available for insertion. When the number of sites (i.e., n) increases, the probability of DNA with multiple tags also increases. Thus η exhibits a maximum at n = 10. Maximum for n = 5 occurs at a larger p.

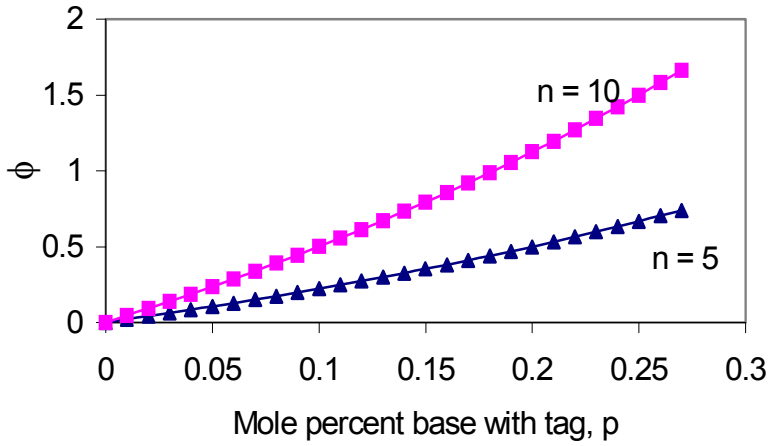


Fig. 1.7. The background due to multiple tagging increases monotonically as p increases.

One can see from the above reasoning that this incorporation process could bring statistical error during DNA chip data analysis. If one pixel has a stronger fluorescence signal than another, that does not guarantee more hybridized strands on the pixel than the other. Since PCR also preferentially amplifies certain sequences (76), it is very hard to set a standard to determine if one pixel "scored" or not, i.e. the result the DNA chip could be ambiguous. Comparisons between repeated experiments are difficult. Thus, the same spots need to be repeated to ensure good statistics for reliable conclusion, affecting the information density of the chip. Such statistical problems in data analysis due to PCR and dye tagging are been described in other reports (77, 78).

In conclusion, from the above discussion it is apparent that tag-free detection will significantly impact the reliability, performance and vitality of DNA chip method.

1.3.6 Label-free detection of DNA hybridization

Label-free methods have been developed in many laboratories and have attracted increasing attentions. Surface plasmon resonance (SPR) (79 - 83), electrochemical methods (84, 85), and mass based methods such as quartz crystal microbalance (QCM) (86) and silicon cantilever (87, 88) have been suggested and developed.

Surface plasmon resonance absorption is a phenomenon occurring at a clean metal surface when an incident light beam strikes the surface at a particular angle. The kinematics condition of SPR is sensitivity to changes in the refractive index of the medium next to the clean metal surface, which makes it possible to measure the attachment of molecules. In other words, the incident angle of the light for SPR absorption shifts due to variation in the dielectric environment. A commonly used oxide

free metal for SPR is Au. Since the refractive index of ssDNA and dsDNA monolayer is significantly different (89), a shift in SPR absorption angle is expected (90, 91). Due to low contrast, SPR is efficient for macroarrrays in the dimension of several (92, 93). Sensitivity of conventional SPR is around 10 nM. Many efforts have been made to increase the sensitivity. The most effective and popular method is to introduce higher refractive index tags onto the sample. Metallic nanoparticles are good choices because of large dielectric constant, high density and biocompatibility (92, 94, 95). Reported sensitivity reaches 10 pM (92).

Several electrochemical techniques have been developed to detect DNA hybridization. The first emerged and the most used method is to detect the changes in the intrinsic electroactivity of DNA (primarily the oxidation of guanine) before and after hybridization (96, 97). In this method, guanines in the probe ssDNA are replaced by inosine that pairs with cytosines. The hybridization is subsequently detected by the guanine signal from the complementary target. The detection limit for this method is 100 nM for a 19 base DNA containing 4 guanines. In another method, conductive polymer polypyrrole has been used as electrodes and sensing surface (98 - 101). Probe ssDNA was covalently immobilized on the polymer film at 5' end. In some cases, probe ssDNA was grafted onto the polymer chain during electropolymerization of polypyrrole. Hybridization is recognized by reading a transient current that increases due to the flow of small mobile cations into the polymer film to balance the negative charged target DNA. Sensitivity of this system is not competitive. Recently, Leiber's group (102, 103) reported nanowire sensors for DNA hybridization detection. In their method, probe ssDNA was grafted onto a Si nanowire. Conductance was measured before and after

hybridization as a hybridization indicator. Human cystic fibrosis mutation gene at one site is discriminated from the wild type. A femtomolar detection limit is reported.

After DNA is hybridized, its mass is doubled or in some cases becomes many times the size of the single stranded DNA. Caruso et al. (86) in 1997 first reported using quartz crystal microbalance (QCM) to detect DNA hybridization. Sensitivity of this method is in the range of 10^{-18} M. However, this method suffers high cost and lack of integration. Another technique, thin film bulk acoustic resonators on silicon substrates, was reported by Gabl et al.(104) recently. The detection principle of this method relies on a resonance frequency shift caused by mass loading of the acoustic resonator. Reported sensitivity is 25 ng/cm^2 for this method. Mckendry et al. (105, 106) reported a nanomechanical detection technique by using miniature silicon cantilevers. DNA probes were immobilized on the cantilever. The cantilever bends when hybridization occurs and the deflection of cantilever is detected by optical means. Femtomoles of DNA on the cantilever can be detected and the detection limit is 75 nM in solution. The methods are not demonstrated to detect mismatches.

Other label-free methods for DNA hybridization detection include molecular beacons (107), mechanical techniques (108, 109), and AFM nanoshaving approach (110).

Chapter 2

A label-free method for DNA chip

DNA chip technology has attracted enormous interest with its characteristics such as its combinatorial approach, possibility of complete automation, miniaturization, and fast output. Fluorescence detection in DNA chip technology offers high sensitivity, linearity and high thoroughput. The current read-out platforms such as confocal microscopy are well established. However, due to the dye-labeling requirement in the current DNA chip method, the robustness, accuracy/reliability and high cost of the assay are hindrances for this technology to be applied in clinical diagnostics. A label-free assay would potentially overcome these difficulties.

In this chapter, the concept and device design of a label-free method will be introduced. The fabrication process and the evaluation of the device performance will be described.

2.1 Design and concept of the label-free DNA chip method

The elements of the device structure are shown in Fig.2.1 (75, 110). The structure is composed of three layers. The top layer of single stranded DNA (ssDNA) is covalently immobilized onto a thin film of polystyrene (PS), a fluorescent polymeric material. The PS is deposited on a Si substrate by solution spinning cast. The substrate may be a fiber optic probe or a flat wafer. The signal that can be collected is reflection or transmission

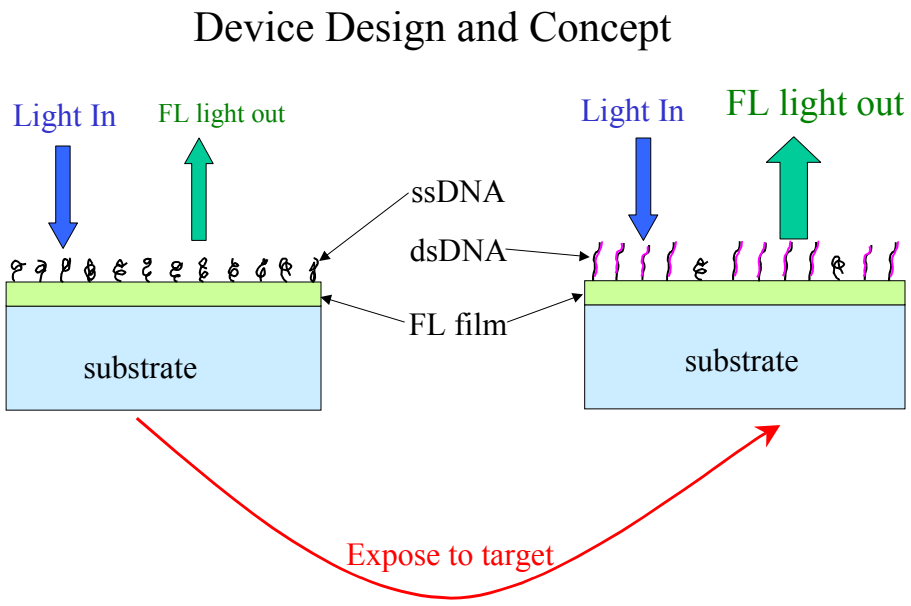


Fig. 2.1. The basic element of the DNA chip device is shown. For the study described, the substrate is Si wafer and the fluorescent (FL) film is polystyrene.

mode depending on the optical property of the substrate. For this study, the substrate is polished silicon wafer, i.e. the signal is collected in reflection mode.

As the incident light strikes the surface as shown in Fig. 2.1, some will be reflected, scattered, and absorbed by the DNA layer and the DNA/PS. The light travels into the PS film and is reflected at the DNA/PS and PS/Si interfaces. Thus, the light I_o that reaches the underlying polystyrene film that contributes to the overall fluorescence is,

$$I_O = I_{incident} - I_{scattering} - I_{reflection} - I_{absorption} \tag{2.1}$$

The resulting detectable fluorescence light I_{PL} produced by the PS layer is

$$I_{FL} = KI_O \tag{2.2}$$

where the constant K is proportional to the total fluorescent efficiency of PS and the sensitivity of the detector.

A significant change in optical properties of DNA is expected as it converts from ssDNA to dsDNA (89). This change is expected to be significant because ssDNA is a dielectric material and dsDNA is a semiconductor with a band gap of a few hundred meV (89). Thus, the surface refractive index will significantly change as the ssDNA layer transforms to dsDNA. It is this change that will cause the difference in I_{FL} between ssDNA and dsDNA. The fluorescence contrast depends on the change in refractive index and thickness of the top layer as it transforms from ssDNA to dsDNA. Since multiple reflections occur in the system, interference effect is not avoidable and the thickness of PS film will affect the contrast in I_{FL}.

2.2 Device fabrication

Chemicals and reagents are from Sigma-Aldrich unless otherwise mentioned. Water refers to deionized water unless otherwise mentioned. Water used in incubation and washing buffers is DNAse free water from Invitrogen.

2.2.1 Substrate preparation

Si substrates were prepared by dicing commercial Si wafers (Virginia Semiconductor) into 1.5 cm × 1.5 cm slides. The slides were soaked in ethanol (Fisher Scientific) overnight to preliminarily remove organic contamination. After rinsing well with water, the slides were washed in Piranha solution (98% sulfuric acid: 30% hydrogen peroxide = 3:1 v /v) for about 1 min. Piranha solution is a super-strong oxidizer. The Si wafers were considered clean and the surfaces of the Si wafers were hydrophilic after Piranha treatment. The wafers were rinsed well in water and stored in 70% ethanol solution.

Polystyrene (PS) solution (1%) was prepared by dissolving 250,000 delton molecular weight PS pellets (Scientific Polymer) in toluene. Si wafers were rinsed in water then dipped in 1:200 HF solution for 45 s. Si was treated with HF to make the Si surface hydrophobic for better PS adhesion. The wafers were then rinsed in water and dried in an oven at 98°C for 1 min, then cooled down to room temperature. A PS thin film was made by spinning cast 50 μl of 1% PS onto a Si slide at 3,000 rpm for 15 s. The film was then vacuum dried at 50°C for 1 h to remove residue solvent. The film was then ready for surface treatment.

Based on the phosphoramidate chemistry chosen to immobilize DNA onto PS surface, amine groups were expected on the PS surface. An integrated plasma discharge machine from March Plasma was used to modify the PS surface. Base pressure was 50 mtorr and the chamber was purged with NH_3 gas at 590 mtorr for 20 min. The pressure of NH_3 was adjusted to 465 mtorr and RF generator was turned on at a power of 62 W. The treatment lasted for 1 min. The chamber was then evacuated to 100 mtorr before purging with air. The PS film was taken out for the immobilization process.

2.2.2 Characterization of modified PS surface

X-ray Photoelectron Spectroscopy (XPS) was used to detect nitrogen element concentration on the modified PS surface. Plasma treated PS films were taken to the XPS for measurement within 1 h after the treatment. FTIR spectrum was used to determine the form, such as primary or secondary, of the nitrogen element.

PS is an inert polymer. In order to covalently immobilize DNA on to PS, the surface must be activated with the desired functional groups. In this case, amine groups are required for EDC chemistry as described in section 2.2.3. Several surface modification methods have been tried to create amines onto the PS surface, such as corona discharge, wet chemistry, and plasma discharge. Corona discharge is easy to operate and can readily activate the PS surface. However, in this method surface uniformity of functional groups was not achieved. Wet chemistry can bring dense and uniform distributed functional groups onto the PS surface, but the process is long (over 12 h) and uses harsh chemicals such as fuming sulfuric acid and nitric acid. Plasma discharge can uniformly modify the PS surface at a process cycle time of less than 2 min.

Although quite a few citations in the literature described using nitrogen or ammonia plasma discharge to bring amine groups on to polymers (112, 113), optimal operation parameters depend on the particular plasma systems. The most important parameters are pressure of gas, RF power input, and treatment duration. Both nitrogen gas plasma and ammonia gas plasma were tried and the assumption was that nitrogen element on the surface after treatment is in the form of either primary or secondary amine. Since both of these amines can react with 5' phosphorylated DNA in the presence of EDC and imidazole to form covalent bonds, maximum nitrogen element content on the surface is desirable.

Nitrogen and ammonia gas plasmas were tested. Under each of their optimized conditions, XPS showed nitrogen element concentration on the nitrogen plasma modified PS surface reached a maximum of 7.56% relative to carbon, oxygen and nitrogen. Ammonia gas plasma treated PS surface showed a maximum of 14.17% nitrogen concentration.

The assumption is that all the nitrogen on the PS surface after treatment is in the form of either secondary amine or primary amine. The chemical composition was tested by FTIR. Since the concentration of nitrogen element is low, reflection mode FTIR was employed. The spectrum indicates existence of the nitrogen on the surface of PS film. FTIR spectrum showed peaks at 1,603 nm, 1,026 nm, and 1,073 nm that compared to the stretching modes of primary amine groups (114).

2.2.3 Immobilization of ssDNA

The probe ssDNA is attached to the PS surface using a highly subscribed method (115). The probe has a phosphate-group at the 5'-end that can react with amine groups to form a phosphoramidate covalent bond with the presence of 1-[3-(Dimethylamino)propyl]-3-ethylcarbodiimide, better known as EDC (water soluble carbodiimide), and imidazole. Fresh-made 50 μl 0.1 M EDC in 0.1 M imidazole buffer with pH in the range of 7.2-7.5 and 75 μl of 7.5 μg/μl oligo solution were added onto the NH_3 plasma treated PS film surface. The film was kept in a humid environment at 50°C for 5 h. After incubation, the chip was washed in 0.2 M NaOH + 0.25% SDS solution for 2 min at 50°C and then rinsed thoroughly in water. The chip was blown dry by compressed air filtered through 0.1 μm Millipore filter. ssDNA (27-mer, 5'phosphate-CAA-AAT-AGA-CGC-TTA-CGC-AAC-GAA-AAC-3') was synthesized and supplied by IDT. The first three nucleotides CAA are used as spacers and do not participate in hybridization. The provision of such a spacer is important to achieve fast hybridization kinetics (121).

In this chemistry, EDC can rapidly react with phosphates to form an active complex that is able to couple with amine containing compounds. The carbodiimide activates and alkyl phosphate group to a highly reactive phosphodiester intermediate. The phosphodiester subsequently reacts with imidazole to form a second reactive intermediate, a phosphorimidazolide that rapidly couples to amine containing molecules to form a phosphoramidate linkage. The salient reason to have imidazole incorporated is to avoid hydrolysis of the phosphodiester.

2.2.4 Hybridization and dehybridization

Complementary ssDNA was supplied by IDT with sequence as (27-mer, 5'-ATA-GTT-TTC-GTT-GCG-TAA-GCG-TCT-ATT-3'). Hybridization buffer containing 0.1 M Na_2HPO_4, 2.5% SDS, 1 mM EDTA, and 0.1 M NaCl with pH = 7.4 and 1 μg/μl complementary ssDNA solution were dispensed 50 μl each onto the chip on which probe ssDNA was attached. The chip was then incubated in a humid chamber at 42°C for 13-14 h. After incubation the chip was washed twice for 10 min with 6 × SSC + 0.1% SDS at 60°C and then rinsed thoroughly with water. The chip was finally air dried. For dehybridization or removal of target ssDNA, the chip was immersed in 8.3 M urea solution at 42°C for 20 min followed by thoroughly washing in water.

2.3 Device performance evaluation

2.3.1 Fluorescence measurement

Fluorescence measurements were performed on a Hitachi fluorescence spectrophotometer F-4500. Incident UV light of 265 nm was projected on samples at an angle of 30°. The effective fluorescence area is about 0.8 cm × 0.3 cm. Instrument parameters were kept the same for all samples. Peak values of emitted light around 325 nm were used for signal contrast calculations.

2.3.2 Quantification of DNA coverage and extent of hybridization

Confocal microscopy (Zeiss, LSM 510) was used to determine surface coverage of the probe ssDNA and hybridized complementary ssDNA. A 40X oil immersion objective with numerical aperture 1.3 was used for all the applications. In these

experiments, the probe ssDNA was attached with Cy5 fluorescence dye at 3' end and the complementary ssDNA was attached with Cy3 at 5' end. After grafting the Cy5 tagged probes onto PS surface by the same chemistry as described in section 2.2.3, the chips were examined by confocal microscopy. Probes were only grafted in a confined area on the PS surface so that there was a significant interface boundary between the probe covered and untreated area. With this approach the fluorescence signal from background noise could be discriminated. Scanned image and fluorescence signal counts from the probe area were recorded for further quantification purposes. The samples then were hybridized with Cy3 tagged complementary ssDNA by the same method as described in section 2.2.4 and subsequently examined by confocal microscopy.

Calibration curves were generated to determine surface coverage of the probe and hybridization percentage. A series of Cy5 labeled ssDNA solutions were tested for fluorescence with the confocal microscopy under identical settings. From a concentration of 0.0192 μg/μl to 0.0012 μg/μl, total of nine solutions were tested. Each solution was dropped on the glass slides with 50 μl of volume. Focusing was performed by first focusing on the edge of the drop, then moving the sample so that focal point was near the drop center, and followed by moving the focal point into the solution. At least three images were taken for each solution for different sample to objective distance. Fluorescence intensities of an arbitrary unit were obtained by using the counting software. Counts of all the images taken from one solution were averaged as final intensity for that solution. Counts taken at different gain settings were normalized to one certain gain setting assuming linear relationship between counts and gain. The confocal volume scanned each time is 9×10^{-8} cm^3 with dimension of 0.03 cm × 0.03 cm × 0.0001

cm. By using this volume, concentrations of the solutions were converted into numbers of molecules. A counts versus number-of-molecules curve was then plotted as the calibration curve. A similar curve was generated also for Cy3 dye with which target DNA was attached. By checking with the calibration curve, the number of probe DNA molecules immobilized on PS and the number of hybridized targets was obtained.

An assay response curve was obtained as follows. Twenty-six pieces of PS film were prepared as described earlier. At the probe DNA grafting step, 13 of them were fully covered with grafting solution and the other 13 were only spotted in the middle of the film. Fluorescence of PS for the fully probe-covered samples was tested after grafting. During the hybridization step, one fully probe-covered sample was paired with one partially probe-covered sample. Hybridization durations for the 13 pairs were respectively 1 h to 13 h with 1h increments. After hybridization, the fully probe-covered samples were scanned by spectrophotometer for PS fluorescence and the partially covered samples were scanned by confocal microscopy for Cy5 and Cy3 fluorescence. The fluorescent signal from Cy5 and Cy3 was measured and the numbers of immobilized and hybridized DNA molecules were extracted from the respective calibration curves. Instrumental parameters for fluorescence measurements and calibration were kept the same for the same dyes. The experiments were repeated three to five times with multiple spots scanned on each wafer. The immobilization and hybridization data presented were the average of these repetitions and the error bars represent the variance observed on this average.

2.3.3 Mismatch detection

One important application of DNA chips is to detect point mutations and single nucleotide polymorphism (SNP) in genes. In other words, the assay should successfully decipher a perfectly matched and one nucleotide mismatched probe-target pair. The possible ability of the designed chip method to detect such a single nucleotide mismatch was tested as follows: The probe ssDNA sequence and perfect match ssDNA sequence are the same as in the section of immobilization and hybridization. The mismatch target sequence is 27-mer, 5'-ATAGTTTTCGTTG (G) GTAAGCGTCTATT-3' in which the G in parenthesis at position 14 from the 5' end is the mismatched nucleotide. In each experiment, two samples were processed in parallel as in the PS film preparation and probe immobilization. At the step of hybridization, one sample was hybridized with perfect match and the other with mismatch. The rest of the hybridization conditions were kept the same. After hybridization, the two samples were separately washed in 6× SSC + 0.1% SDS at a series of temperature of 50, 60, 70, and 80°C starting from the lowest temperature to the highest one sequentially. After each wash, the samples were rinsed thoroughly with water and fluorescence measurements were conducted.

2.4 Result and discussion

2.4.1 Signal contrast

As introduced in section 2.1, in the design of the DNA chip method, fluorescence signals from the PS are expected to be different when the DNA layer transforms from ssDNA to dsDNA. Of the three optical activities in equation 2.1, reflection and scattering are the dominating factors. Furthermore, scattering turns out to be the more critical of the

two. The contrast mechanism will be analyzed in the next chapter and a simple model will also be described. In this section, only the experiment results and corresponding discussions will be covered.

The change of fluorescence from PS as the top DNA layer in Fig.2.1 changes from ssDNA (probe) to dsDNA (probe-target duplex) is measured by Hitachi spectrophotometer at an excitation wavelength of 265 nm. This wavelength was chosen because it is at the maximum excitation of PS and it is only 5 nm apart from 260 nm where the reflective index difference between ssDNA and dsDNA is maximum (will be described in Chapter 3). Fig.2.2 shows the measured fluorescence of the PS when the top layer is (a) ssDNA, (b) 25% hybridization to dsDNA, and (c) reversed ssDNA after dehybridization. The contrast, defined as $(I_{ds} - I_{ss})/ I_{ss}$, is around 35% as an average value from dozens of samples. The contrast is highly amplified compared to the expected difference due to simple absorption difference between ssDNA and dsDNA. I_{ss} is lower than I_{ds} indicating that the fraction of the incident light sieving into the PS layer is larger when the top layer is dsDNA than when it is ssDNA. The major change is occurring due to change in reflectivity and scattering. This will be discussed in detail in the next chapter.

To prove that the contrast is due to the change of the top layer from ssDNA to dsDNA but not to other factors, a hybridization-dehybridization cycle experiment was designed. In this experiment, a sample went through a couple of hybridization-dehybridization cycles after grafting of the probe ssDNA. Fluorescence tests were performed after every step. Not surprisingly, after dehybridization, the fluorescence signal dropped back to the same level as ssDNA. And after re-hybridizations, the

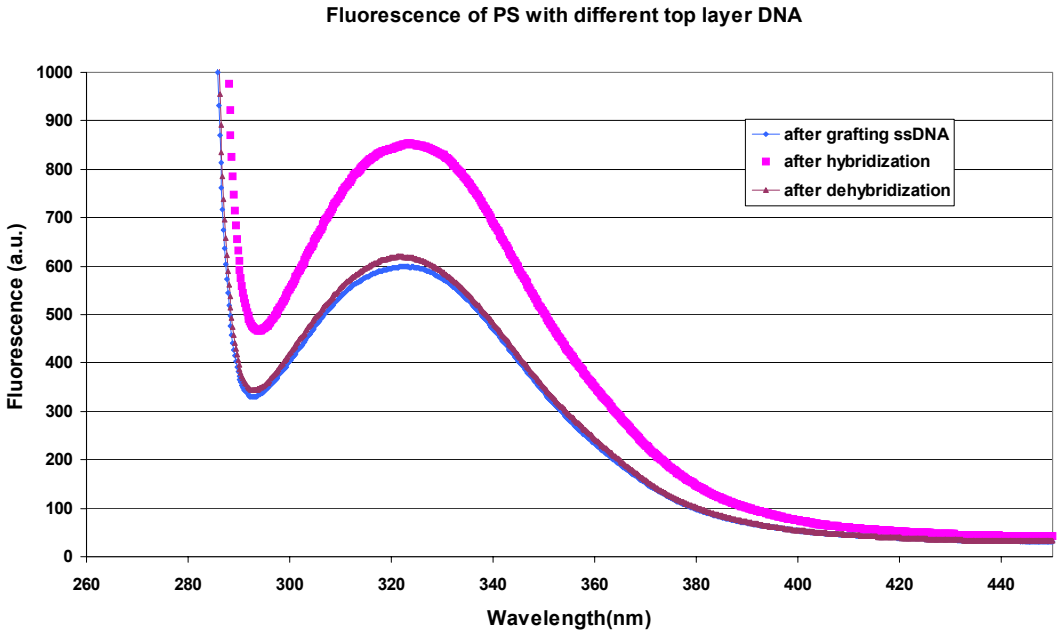


Fig. 2.2. Fluorescence of PS detected with the top layer of DNA changes from ssDNA to dsDNA then to dehybridized ssDNA.

fluorescence signal rose again to the hybridized level. Several cycles can be repeated until the possible loss and degradation of DNA and PS.

The contrast is dependent on the thickness of the PS film. Fig. 2.3 shows how the contrast varies as the thickness of the PS film changes. At the beginning of the study when the mechanism of the contrast was not at all understood, PS films of around 30 to 40 nm thick were used for all the experiments. These films were obtained by spin casting 1% PS/toluene solution at 3,000 rpm. Spin time for all the films was 15 s. These two parameters determined the thickness of the PS films and coincidentally led the contrast close to the maximum value. Fig. 2.3 also shows that when the thickness of PS is beyond 70 nm or so, contrast dropped close to zero. Thicker films were prepared by spin casting PS/Toluene solution of higher concentrations. It was found that film quality was not as good when the thickness increased. The surface became rough and the thickness across the film became uneven and in some cases there were visible cracks in the film. Since the contrast is sensitive to optical properties of the PS, probably the decrease in film quality causes the lowering of the contrast. Another possible reason could be the absorption of the excitation and fluorescence light by the PS film becomes significant, thus less fluorescence light comes out of the system.

The dependence of contrast on PS thickness was first found when the PS thickness was increased for the purpose of increasing the contrast or dynamic range, but the contrast was decreased. Intuitively, this thickness dependence is perhaps due to an interference effect from multiple reflections of the light in the PS film. Simulation from multiple reflections treating DNA as one of the layers in DNA/PS/Si multilayer structure did show a dependence of contrast on thickness of PS, and the thickness of PS at which

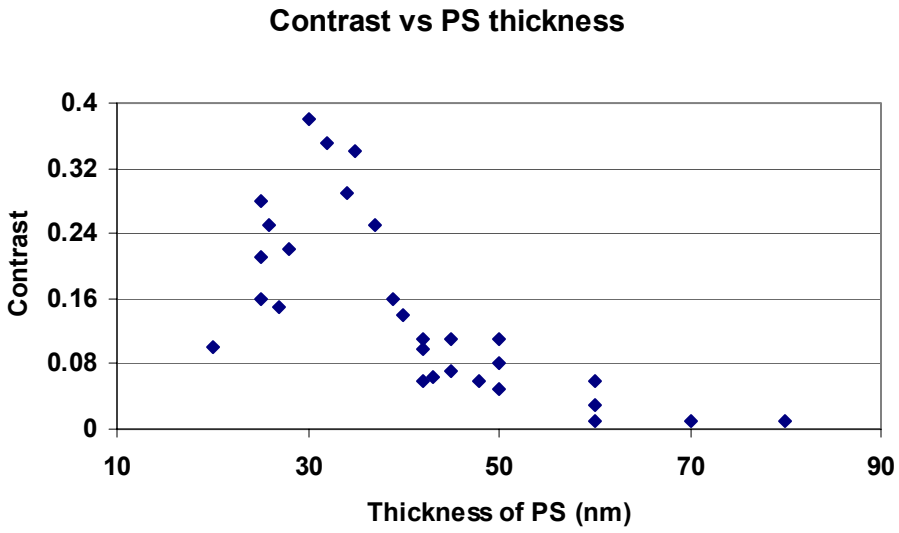


Fig. 2.3. Contrast is dependent on the thickness of PS. At range of PS thickness from 30 nm to 35 nm, contrast reaches the maximum. Contrast drops to nominally zero when PS thickness is beyond 70 nm.

the contrast reached the peak value matches well with the experimental result. However, the predicted contrast was significantly lower than the experimentally measured values. A model incorporating DNA molecules as individual scattering centers other than a continuous film was built and it matched fairly well with real data. The model and simulation will be covered in detail in the next chapter.

In biomedical applications, target DNA is usually a longer strand containing hundreds of nucleotides digested from natural DNA that could have millions of nucleotides. In order to test if the proposing DNA chip method has potential in real medical application, long complementary DNA strands were used in the hybridization and all the rest of the experiment conditions were kept the same as described before. The results were excellent. Contrasts from long complementary hybridization were higher than those from short complementary hybridization. From the data of five independent experiments, average contrast value for l00 nt complementary hybridization is 40% versus 33% for short complementary hybridization, with the same PS thickness.

2.4.2 Mismatch detection

If the target has one or two mismatched nucleotides to the probe, it can still bind to the probe. Since mismatched nucleotide does not form hydrogen bonds with the corresponding nucleotide in the probe, less energy is required to break the two strands apart. In other words, the melting temperature of a duplex containing mismatch is lower than that of a perfect double stranded duplex. Both theoretical calculations and experimental results from the literature (80, 116, 117) stated that for a 26 base-pair dsDNA, the melting temperature drops by ~5 °C and ~10 °C with respect to a perfect

match for one and two mismatch(es). The melting temperature for the perfect match sequence studied is ~68 °C calculated from the technical bulletin of IDT (118), thus the melting temperature for one mismatch is ~63 °C and for two mismatches is ~58 °C. The post-hybridization washing temperature for the perfect match was set to 60°C at which nonspecifically bound target will be washed away while the specifically bound target will remain.

The commonly used strategy for the mismatch determination is to wash the chip at temperatures below the melting temperature of the perfect match but above that of the mismatch. The expectation is that at a certain temperature range, mismatched targets are almost all washed away while perfect matched targets remain binding so that the fluorescence contrast from perfect matched sample is higher. Fig. 2.4 shows the contrast variation at different washing temperature. 50°C is below both mismatch and perfect match melting temperature. At this temperature, mismatch (MM) showed low contrast of ~ 8% while perfect match (PM) showed 22%. At this

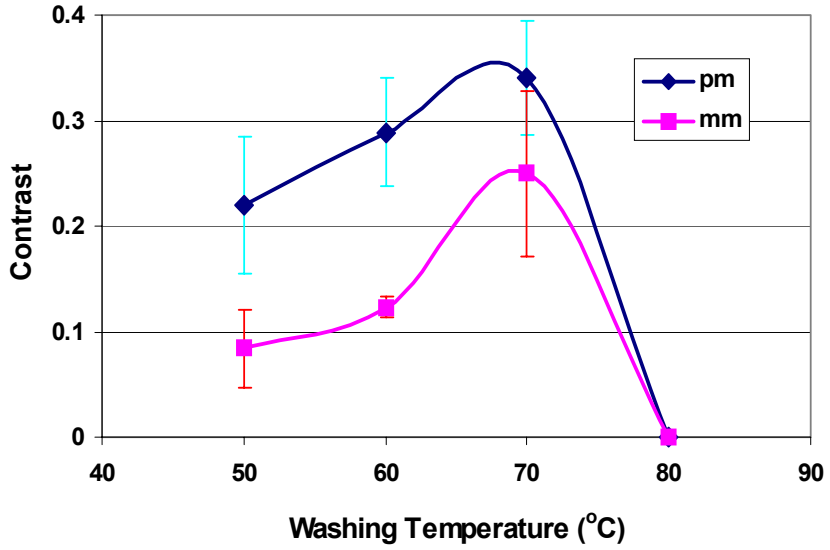


Fig. 2.4. Contrast responses to washing temperature for perfect match and mismatch. At all temperatures, contrast of perfect match is higher than that of mismatch. Detection of mismatch is possible.

temperature, some of the nonspecifically bound target is possibly retained on the chip and the effect is similar to additional ssDNA probe on the chip, making the contrast low. At 60°C, contrast for MM and PM increased to 13% and 29% respectively. This temperature is 3°C below the theoretical melting temperature of MM and 8°C below that of the PM. Most of the nonspecifically bound targets should have been washed away. The results at 70°C were surprising. This temperature is 2°C beyond PM's theoretical melting temperature and 7°C higher than MM's. Most of the hybridized MM and some of the PM were expected to be detached from the probes at this temperature. However, contrast for both cases increased, to 23% and 33% for MM and PM respectively. There is no detectable contrast for both MM and PM after 80°C washing since all the target ssDNA dehybridizes. At this point, there is no reasonable explanation for the higher contrast at 70°C, especially for the MM. Perhaps the melting temperatures for both PM and MM are higher than the theoretical ones. The contrasts from PM are higher than those from MM at washing temperatures of 50°C to 70°C indicating it is possible to differentiate PM from MM by comparing contrast values. The contrasts for both PM and MM are highest at 70°C washing temperature, but there are only 10 units in between and error bars overlapped. Probably the surface coverage of both PM and MM is not stable and lacks reproducibility. Washing temperatures at 50°C and 60°C are both applicable, but at 60°C the difference between the contrasts is slightly greater than that at 50°C. Additionally, reproducibility at 60°C seems better than at 50°C indicated by the smaller error bar at 60°C. Probably this is because there is less nonspecific binding at 60°C than 50°C.

Hybridization temperature used is 42°C which is 26°C below the theoretical melting temperature of the PM. Hybridization temperature at this level has little effect on discriminating PM and MM. By tuning to higher hybridization temperature, binding of the complementary targets to probes becomes difficult, however it affects MM binding more than PM thus enhances the discrimination in contrast between PM and MM (109, 119).

This strategy was tested on the assay and the result is shown in Fig. 2.5. Let the ratio in fluorescence from the PS for the top layer corresponding to ssDNA and dsDNA, respectively, be $R(x) = I_{ds}/ I_{ss}$, where x is the number of mismatched nucleotides between the probe ssDNA and the target ssDNA. Thus, for perfect match, $x = 0$ and for one mismatch, $x = 1$. As the hybridization temperature increases, the binding for both perfect match and one mismatch decreases. At lower temperatures, $R(0)$ and $R(1)$ are >1 indicating that both perfectly matched sample and one mismatch bind to the probe ssDNA. At a temperature higher than 48°C, the one mismatch does not bind, i.e., $R(1) = 1$, however for perfect match, $R(0) > 1$.

In real applications, the design of both hybridization temperature and post-hybridization washing temperature could be combined to obtain the best mismatch detection condition.

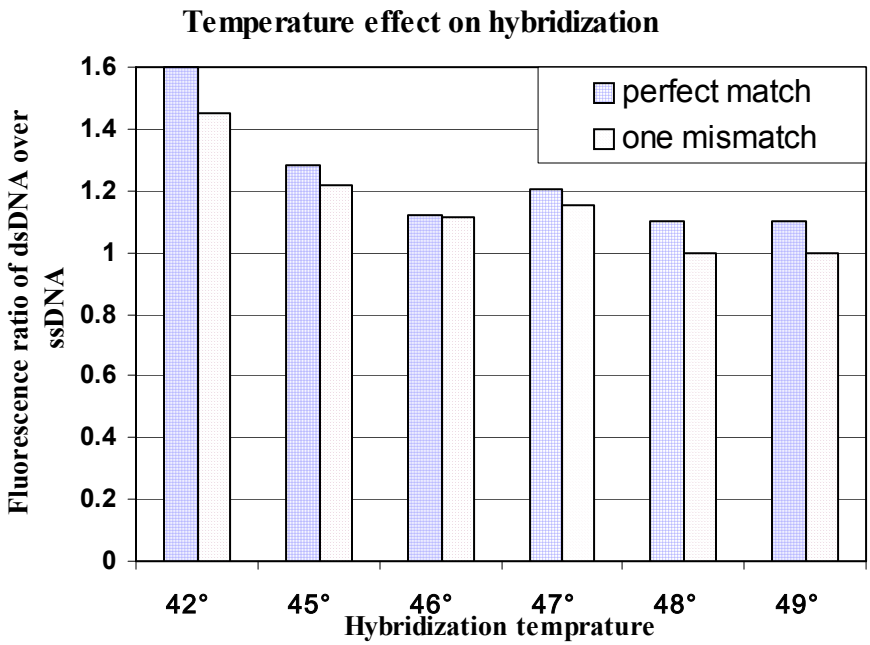


Fig. 2.5. Hybridization temperature effect on contrast for perfect match and mismatch. Unit of temperature is °C.

2.4.3 DNA surface coverage quantification and assay response estimation

Although the goal of this dissertation research is to develop a label-free assay, fluorescence labels were used to estimate surface coverage of probes and extent of hybridization. Probe ssDNA with Cy5 tagged at 3' end were tethered partially on PS surface and followed by hybridizing with 5'-Cy3 tagged complementary as described in the experimental section. Confocal scanning microscopy was used to quantify surface coverage. The calibration of the fluorescence from Cy5 and Cy3 tagged ssDNA of known concentration was conducted and the result is shown in Fig. 2.6. Fluorescence signals from both Cy5 and Cy3 showed linear relationship with regard to the number of fluorophores. As expected, both lines pass through the origin, consistent with no

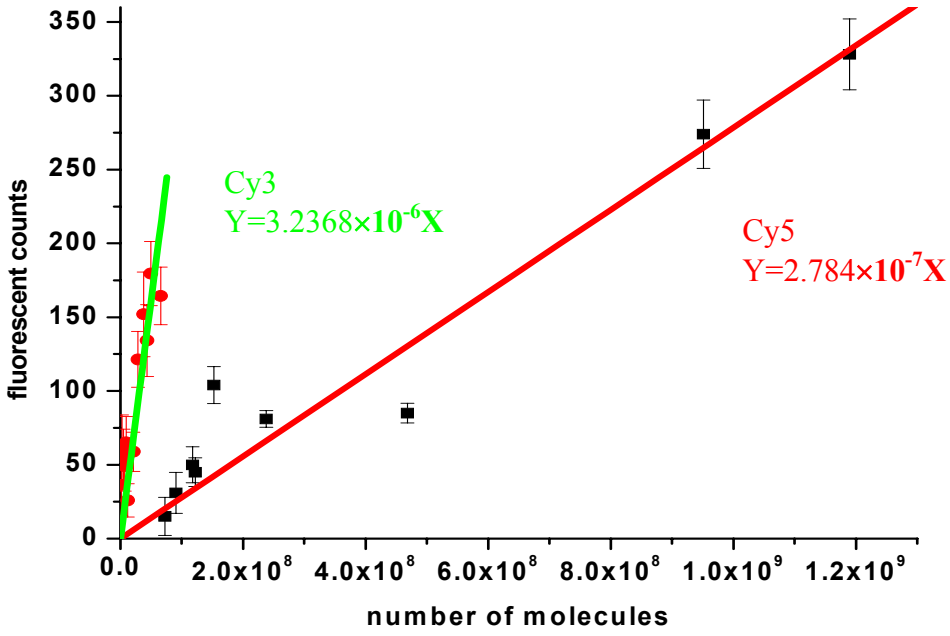
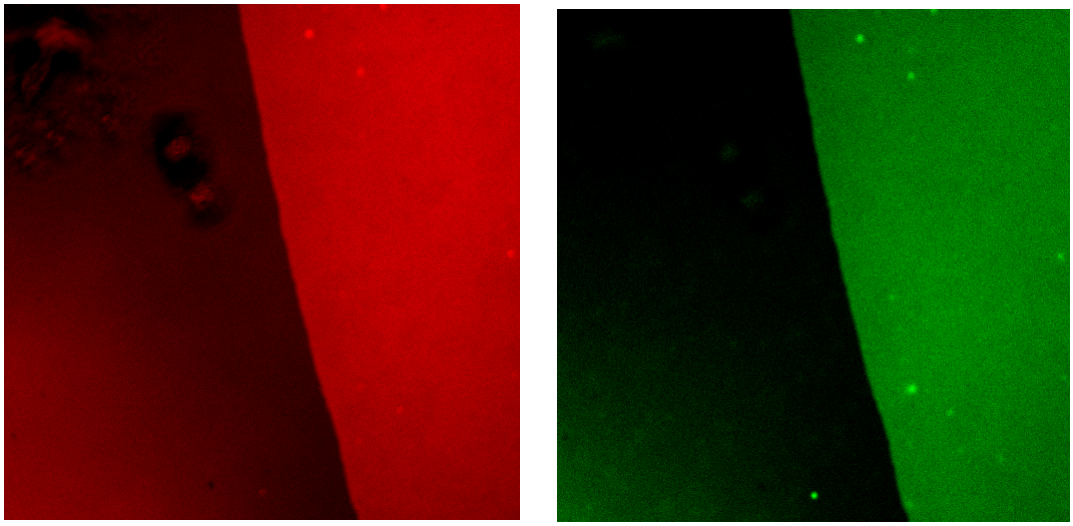


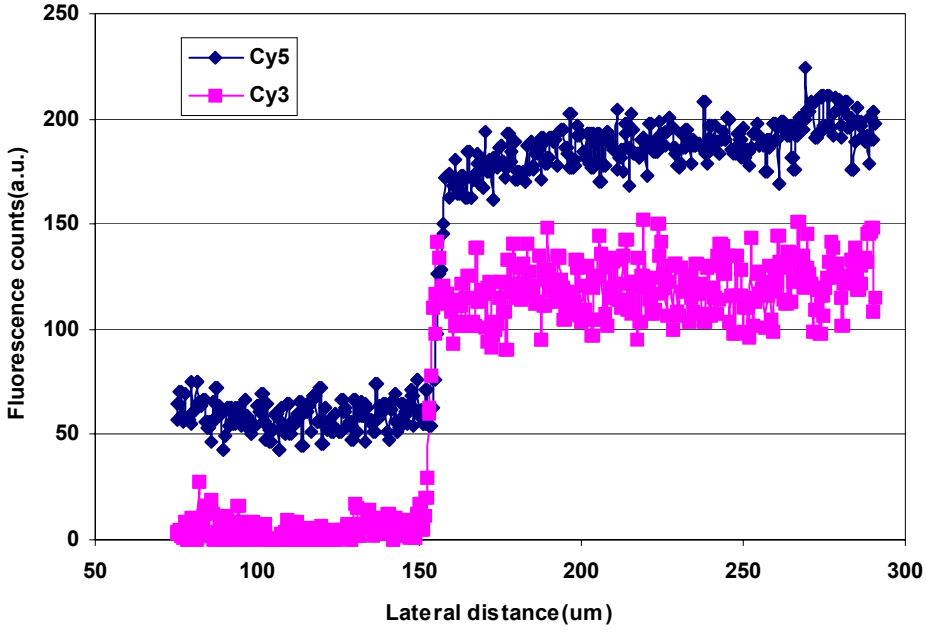
Fig. 2.6. Calibration curves of Cy5 and Cy3 dye. The scanned area is 325 × 325 μm^2 with the optical slice for Cy5: 9.8 μm, Cy3: 8.4 μm. Fitting R-square for Cy5 is 0.92 and for Cy3 is 0.97.

fluorescence at zero concentration (background for all the images were subtracted). The slopes of the two calibration lines are quite different, but there is no practical meaning to compare the two.

Fig.2.7 shows two confocal images of partially covered PS with immobilized 3'Cy5 probe ssDNA after hybridization with the 5'Cy3 target ssDNA. The fluorescence from Cy5 and Cy3 indicate the presence of both the probe and the target ssDNA. From the independently measured calibration curve, the averaged fluorescence intensities (34 samples for Cy5 and 9 fully hybridized samples for Cy3) from Cy5 and Cy3 correspond to a probe coverage of 6×10^{11} molecules/cm^2 and target coverage of 1.5×10^{11} molecules /cm^2. Assuming random tethering, the average distance between the anchored ends of probe ssDNA is around $[4 \times 10^{14}/ (\pi \times 6 \times 10^{11})]^{0.5} = 14.6$nm. The extent of hybridization defined as the fraction of probes that have hybridized to the targets is 25%. The probe coverage and extent of hybridization in this study are within good agreement with reported values (120, 121). Extent of hybridization more than 25% could not be achieved in this study, perhaps because the steric hindrance of the probes by other probes and the nonspecifically bound targets. The spacer nucleotides at the 5' end of a probe supply flexibility and decrease the interactions between DNA molecules and substrate surface and DNA molecules themselves so that higher extent of hybridization and fast hybridization kinetics are possible (122). In this study, only three nucleotides serve as the spacer and this could be one of the reasons that higher extent of hybridization could not be achieved. Another possible reason for low extent of hybridization could be the probe coverage. Peterson et al. (123) reported that with higher coverage of probes, hybridization kinetics and final level of binding decrease. This topic requires further



(a) (b)



(c)

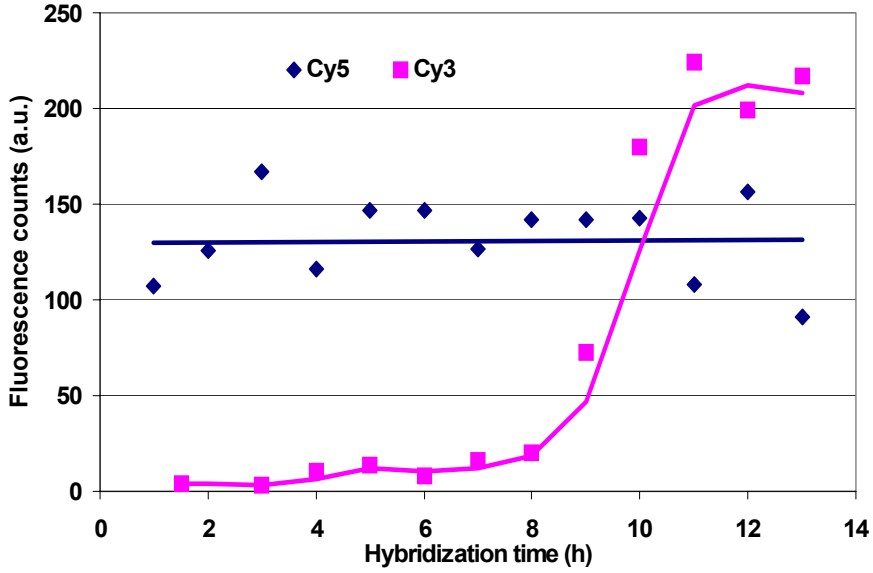
Fig. 2.7. Surface coverage estimation by using confocal microscope. Probe ssDNA was labeled with Cy5 and target was labeled with Cy3. (a) and (b) are raw confocal images scanned at the same position but for Cy5 and Cy3 respectively. (c) Fluorescence counts across the DNA coverage and non-cover region (background) for the images in (a) and (b). Counts used in coverage calculation are background subtracted.

study since it was also found that the contrast of PS fluorescence increases as the extent of hybridization increases as shown by the assay response curve in Fig. 2.8.

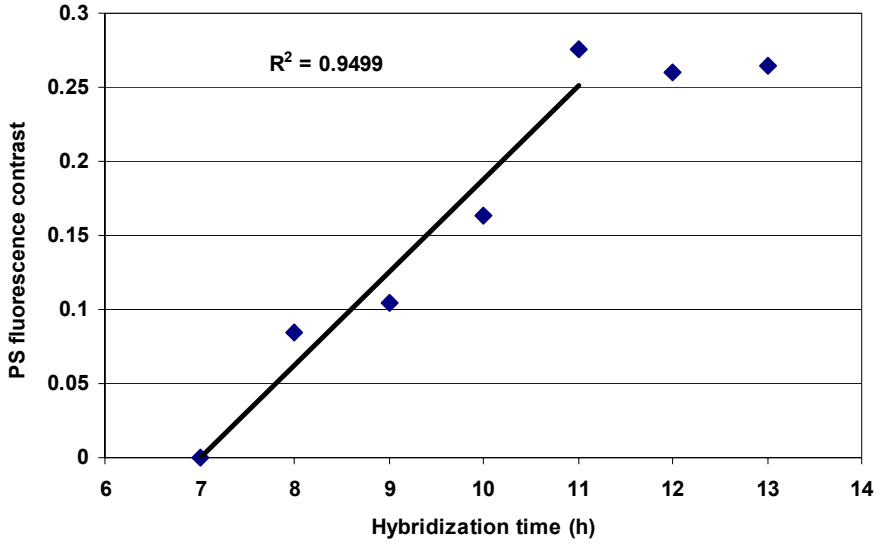
The assay response curve was generated as described in section 2.3.2. For each hybridization time, the fluorescence intensities of Cy5 and Cy3 were converted to molecular coverage of the probe and extent of hybridization using the calibration curves. Cy5 signals for all the samples showed a reasonably constant line indicating small variation in the immobilization process. Fluorescence signal from Cy3 was measurable for binding time above 1.5 h and did not deviate significantly for binding time below 7 h. After 7 h, Cy3 signal started increasing significantly until saturation at 11 h as shown in Fig. 2.8(a). From 7 to 11 h the increase of Cy3 showed linear response to hybridization time. The corresponding paired samples for PS fluorescence showed no contrast until 7 h hybridization. The contrast versus hybridization time after 7 h was plotted in Fig. 2.8 (b) also showed a linear relation that is consistent with the data from confocal fluorescence. Thus, as expected, the contrast versus extent of hybridization plot shows linear response as shown in Fig. 2.8(c). The reason behind the very slow rising plateau from 2 to 7 h is not understood. My conjecture is as follows. Since there was only a perfect-match target in the hybridization solution and the amount of target is in excess, diffusion of the target to the anchored probe should not take long. The first group of targets that make contact with the probes do not completely hybridize. Most of them bond nonspecifically with the probes and thus prevent further access of the probe to other targets and at the same time hindered the access of the neighboring probes. Almost all of the nonspecifically bound target would be removed during the post-hybridization washing process. The time taken for the perfect double helix to form under the described condition probably leads to the

induction time of 6 to 7 h. A small amount of hybridization took place with a faster speed thus contributing to the Cy3 signal for the binding time below 7 h. Another possibility is that nonspecifically bound targets form more severe steric hindrance than double helixes. That could partially explain the faster linear rising of the assay response curve after 7 h.

The response of the assay is highly linear with hybridization from 3% to 25%. The highly linear assay response allows for quantitative measurement of hybridization progress. Since the contrast of PS fluorescence occurs only due to the hybridization at the surface of PS, similar to surface plasmon resonance (78-82), the assay can potentially be extended to measure kinetics of hybridization.



2.8(a)



2.8(b)

Fig. 2.8(a). Probe and hybridized target coverage with regard to hybridization time. Probe coverage has small variation but remains at a constant level for all the time. Target coverage increases as hybridization time increases and is saturated after 11 h.

Fig. 2.8(b). PS fluorescence contrast as a function of hybridization time. Before contrast is saturated, it responds linearly with regard to hybridization time.

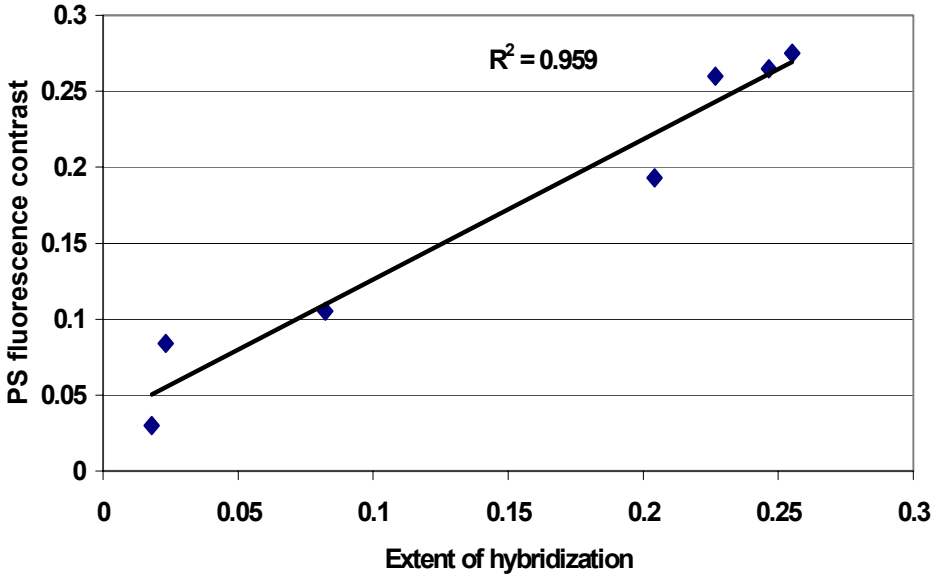


Fig. 2.8(c). PS fluorescence contrast as a function of extent of hybridization. It shows linear relationship. Extent of hybridization is obtained by checking the coverage of probe and target with corresponding calibration curves.

2.4.4 Sensitivity

Surface sensitivity of the assay with the current detection tool, spectrophotometer, was estimated. From the assay response curve, there is no detectable contrast when the hybridization time is shorter than 7 h. Cy3 signal from 7-hour hybridized sample give ~1.1×10^{10} molecules/cm^2 and this number corresponds to 2% hybridization. This can be considered as the sensitivity of the spectrophotometer to the current probe coverage. If the surface coverage of the probe increases, lower extent of hybridization probably can be detected. Sensitivity of the assay with regard to the target DNA concentration can be estimated as follows. Target concentration was diluted two-fold from the value of 100 nM. The rest of the hybridization condition and washing condition were kept unchanged. Hybridization time for all the samples was kept at 14 h. The lowest target concentration that showed stable contrast was 6 nM and a contrast of 13% was obtained. For 3 nM target, contrast varies from as high as 5% to as low as zero. The solution sensitivity is 3 nM therefore.

Note that the sensitivities mentioned above were evaluated based on the spectrophotometer. The sensitivity of the instrument is the bottleneck of the assay. Confocal microscopy shows a response at a lower extent of hybridization, although the detection subject is different. The dynamic range of the spectrophotometer data is 1 in 200 with signal-to-noise ratio of 200:1. If the signal-to-noise ratio is relaxed to 2:1, the corresponding dynamic range is 1 in 10^4 and sensitivity will be improved significantly. Better instrumentation design and possibly PS fluorescence enhancement deserves further study to obtain higher sensitivity.

Table 2.1 shows the collection of some recent reported detection limit for various methods. Note that the sensitivity of the studied method is very competitive in terms of both concentration and number of targets detected.

2.5 Conclusion

In this chapter, the design and evaluation of the label-free DNA detection method was reported. Fluorescence came out of underlying polystyrene supplied as the contrast agent for the changing of the top layer DNA from single strand to double strand. Experimental results have shown 40% signal change before and after hybridization with a dynamic range of 1 in 200 and a signal to noise ratio of 200:1. The contrast responds linearly with regard to extent of hybridization. This method has the potential to detect single mismatch and the sensitivity is competitive compared to other recently reported DNA detection techniques.

Detection Method	Detection limit (concentration of target)	Sample volume (μl)	Detection limit (no. of molecules)	Hybridization time (h)	Reference
Fluorescence	5 pM	10-50	1,000 per 100 x 100 μ m spot		(124, 125)
"Scanometric" (nanoparticel based)	50 fM				(124)
SPR (label-free)	10 nM		6×10^8 per 500 x 500 μm spot	1	(78)
SPR (Au-amplified)	10 pM		8×10^8 / cm^2	2	(90)
Dye-containing liposomes	220 pM		6×10^8	0.5	(126)
BARC sensor (magnetic beads)	100 fM			3	(127)
Microcantilever deflection	400 nM		10^{10}	1	(86)
Molecular beacons	100 pM	10		0.3	(128)
Electochemical	100 pM to 100 fM	500	10^8		(129)
Optical interference	10 fM	10-25		2	(130)
Nanopartical assisted electrochemical detection	500 fM				(131)

Table 2.1. Detection limits of various methods**.**

Chapter 3

Mechanism of the label-free DNA chip assay

In chapter 2, the design and principle of the label-free DNA chip method and experimental study to illustrate the concept were reported. The original design motivation was that more excitation light at 265 nm reaches the PS film when the top layer is dsDNA compared to single stranded DNA. As will be shown in this chapter, the dominating factors contributing to the contrast turn out to be reflectivity and scattering.

There are mainly two reasons for which PS was chosen as the fluorescence layer. First is that the maximum excitation wavelength for PS is at 265 nm that is only 5 nm longer than DNA's resonance absorption peak. At a wavelength of 260 nm, optical properties, namely refractive index and absorption coefficient of ssDNA and dsDNA, all reach their maximum values and the differences between these values also become the largest. Thus a maximum signal contrast between ssDNA and dsDNA is expected in the 265 nm region. The second reason is that PS has been proven to be biologically safe and its surface can be readily modified for bio-conjugation. Additional advantage of PS is that excitation and emission peaks at 265 nm and 325 nm respectively are 60 nm apart compared to 20-30 nm for commonly used dyes in biology. Thus, PS fluorescence imaging does not require sharp-cut optical filters as Cy3, Cy5 and other bio-dyes.

In order to understand the mechanism behind the observed phenomena, a quantitative model is developed with no fitting parameters. All the physical parameters required for the model, such as optical properties and physical dimensions of PS film,

ssDNA and dsDNA are independently measured. In this chapter, the model and the independent measurement of various physical parameters will be derived, reported and discussed.

3.1 Experimental determination of film thickness and optical property

A Woollam Variable Angle Scanning Ellipsometry (VASE) (132) was used to determine the optical properties and thickness of PS film on Si substrate, ssDNA and dsDNA monolayers attached to PS films on Si substrate.

The diameter of the ellipsometer beam is ~1 mm. Three spots on each sample were analyzed. Ellipsometry scans were performed in the UV-Vis range for incident wavelength, λ from 250 to 700 nm. Scans were performed at an angle of incidence with respect to the film-normal of 70°, 73°, 76°, which are close to the Brewster angle of the Si substrate at ~70^{O}. Theoretically, ellipticity (Ψ) and phase (Δ°) of the reflected "s" and "p" polarized light can be expressed by models containing thickness and optical constants of the films. The data then were fit to the measured ellipticity (Ψ) and phase (Δ°). Most commonly used models are Lorentz model and Cauchy model. The former is for metallic, conducting and semi-conducting materials and the latter is for dielectric materials.

First, samples of blank Si substrate were treated with only Piranha and HF as described in section 2.2.1 and were characterized by ellipsometry. Optical properties of Si and SiO_2 are well documented (133) and are in the database of ellipsometer. During the data fitting, Si was assumed infinitely thick and thickness of SiO_2 was fitted. Average thickness of 2.3 nm for native SiO_2 was measured. The Si/SiO_2 model with the fitted SiO_2 thickness was stored for further data analysis on other samples.

For PS, a spin curve was first generated by measuring PS film thickness by x-ray reflectivity. PS films spun from different concentration solutions at spin speed of 3,000 rpm were scanned by a home built x-ray reflectivity machine. A chart of film thickness versus solution concentration was drawn (see Fig. 3.1). Three of the films of different thicknesses were then scanned by ellipsometry. Cauchy model was used to fit for optical parameters of PS. Film thickness and MSE (mean square error) of the fitting were used to judge the fitness. The final parameters are average values of the physical properties measured by ellipsometry. The three films were then treated with ammonia plasma by the procedure described in 2.2.1 and scanned by ellipsometry again. It was found that thickness of a PS film drops approximately 3-4 nm due to NH_3 plasma treatment. Thus, the thickness and optical constant of the PS films in the DNA/PS/Si multilayers are determined.

ssDNA was grafted onto the film by the procedures described in section 2.2.3. After washing and drying, PS fluorescence measurement was conducted and then the film was scanned by ellipsometry. During the modeling, optical constants of Si, SiO_2 and PS, and thickness of Si and SiO_2 were fixed from the previous study. Thickness of PS was fitted. If the fitted PS thickness was in the range of ± 5 nm from the predicted thickness, the thickness was considered reasonable. Lorentz model was used for fitting optical constants of ssDNA layer. The dielectric constant $\varepsilon(\infty)$ and the oscillator position En were fixed at 2.1 and 4.76 eV respectively as determined by UV-Vis spectroscopy. The other optical parameters and thickness were all fitted.

The sample was then hybridized by the method described in section 2.2.4 and again PS fluorescence was measured and followed by ellipsometry scanning. Modeling

was very similar to that described for ssDNA except that the dielectric constant $\varepsilon(\infty)$ was fixed at 2.2 for dsDNA.

Samples with PS films of different thicknesses were tested. Some randomly selected samples were dehybridized and fluorescence and ellipsometry measurement were performed sequentially. The measurement of fluorescence was to verify that the hybridization occurred.

3.2 Result and discussion

3.2.1 Thickness and optical constants of PS

Thickness and optical properties of PS are important irrespective of the model. Two methods were available to measure thickness of a thin solid film, x-ray reflectivity and ellipsometry. X-ray reflectivity measures thickness of homogeneous and continuous films with no fitting procedures, i.e. thickness result is the average thickness value of the incident area of the x-ray beam. Drawback is that the area averaged is over $1cm^2$ that makes the technique inadequate to determine thickness non-uniformity. Ellipsometry does not have this drawback and at the same time can determine both optical constants and thickness. Nevertheless in ellipsometry, fitting the experiment data to a suitable model needs very careful and self-consistent analysis. There are six fitting parameters in both Cauchy and Lorentz model. Usually the mean square error (MSE) of a fitting is used to judge fitness with the experimental data. However, fitness alone can not differentiate a good fitting from a "fake good" fitting. The strategy taken was to eliminate as many fitting parameters as possible to lower the chance of getting "fake good" fittings. The

elimination is achieved by determining some of the parameters by other methods. For Cauchy model in PS, film thickness of a PS film can be measured by X-ray reflectivity and absorption peak position can be measured by UV-Vis. Then only the parameters determining optical constants are fitted. As described in section 3.1, a spin curve of the PS was first generated as shown in Fig.3.1. All the thicknesses on the spin curve were measured by x-ray reflectivity. Since the PS thickness is known from spin curve and thickness deduction due to plasma treatment was around 4 nm, the values are compared to the fitted value estimated by ellipsometry. Optical constants for PS films are expected to be the same even though the thicknesses of the films are different. By averaging fitted optical constants from five different films, the result is shown in Fig.3.2. MSEs for these fittings are all smaller than 4.0. Refractive indexes at a certain wavelength match very well with the values in literature (134).

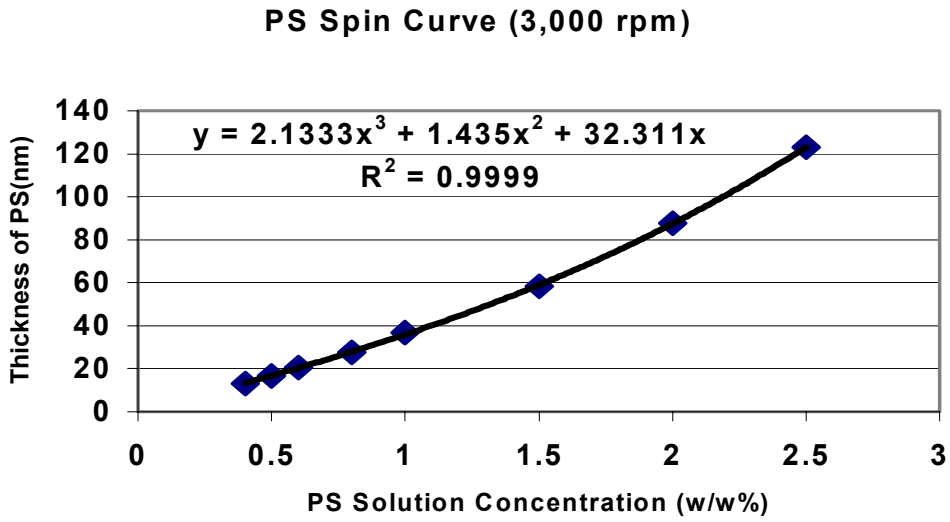


Fig. 3.1. Spin curve of PS. Solvent is toluene. Thickness of PS on Si was measured by x-ray reflectivity. Error bars for all the data points are negligible.

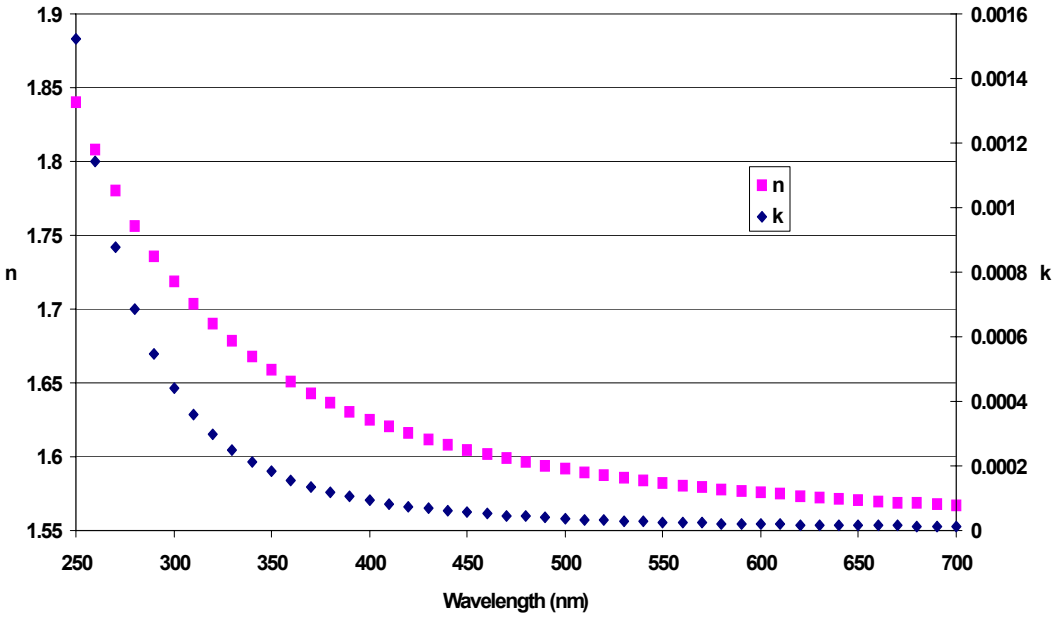


Fig. 3.2. Optical constants of polystyrene.

3.2.2 Thickness and optical constants of ssDNA and dsDNA monolayer

Average thicknesses measured from ellipsometry for ssDNA layer and dsDNA layer are 2 nm and 4.3 nm respectively. Since in ellipsometry modeling, the DNA layer was treated as a continuous film, the thickness result from ellipsometry is an averaged effective thickness of the DNA layer other than individual DNA molecule size.

In order to estimate the DNA molecular dimension, the conformation of the immobilized DNA needs to be considered first. Unperturbed immobilized ssDNA probe can be considered as a Gaussian coil of Kuhn step length b of ~ 2 nm that corresponds to a 4 nt segment (135, 136). For 27 nt ssDNA with contour length of 14.85 nm, the number of Kuhn steps, N is ~ 27/4 or 6.75 steps that corresponds to end-to-end distance of the chain, $R_{ss} = N^{0.5}b$ of ~ 5 nm. One ssDNA molecule can be considered as an anchored balloon with diameter of 5 nm. The balloon swings randomly because of thermal energy; thus the volume that is covered by a quickly swinging balloon is a hemisphere with radius of 5 nm. With coverage ϕ equal to 6×10^{11} molecules/cm^2 the effective height of the ssDNA layer is $[0.5 \times (4/3) \times \pi \times (5\text{nm})^3 \times \phi] = 1.57\text{nm}$. This value is consistent with the measured average thickness 2 nm of the ssDNA layer. For the coverage of 6×10^{11} molecules/cm^2, the average inter chain distance between the grafted chains is $[4 \times 10^{14} / (\pi \times 6 \times 10^{11})]^{0.5} = 14.6$ nm. Since 14.6 nm is much larger than 5 nm, the ssDNA molecular size, no significant inter-chain interaction is expected.

Confocal measurements indicated that the extent of hybridization is 25%, i.e. one in four ssDNA hybridized. The average thickness from ellipsometry for hybridized DNA layer is 4.3 nm. The dsDNA is a semi-flexible macromolecule that is modeled as a worm-

like chain with a persistence length of ~ 70-80 nm (136). Thus, the 24 bp dsDNA is essentially a rod of 8.16 nm in length (137). Three nucleotides serve as a spacer and do not hybridize. This makes the total length of the hybridized duplex 9.66 nm. Since only one in four ssDNA hybridized, it is reasonable to assume that dsDNA chains do not entangle with each other. Therefore interaction between the chains is insignificant. Similar to the anchored balloon model, one dsDNA chain can be seen as a tethered rod that can swing randomly. Thus it covers a hemispheric volume with radius of 9.7 nm. The effective thickness for this ssDNA and dsDNA mix can be expressed as $[0.5 \times (4/3) \times \pi \times (5\text{nm})^3 \times \phi \times 0.75] + [0.5 \times (4/3) \times \pi \times (9.66\text{nm})^3 \times \phi \times 0.25] = 4$ nm. This is consistent with the average value measured from ellipsometry.

The increase in measured DNA layer thickness can be attributed to double-helix formation upon hybridization. The increase in thickness also supports the above interpretation of coiled chain to helix conformation transformation of the DNA. If the immobilized ssDNA were to be fully stretched and form a brush-like conformation, (i) the thickness of ssDNA would be much more than 2 nm, and (ii) the thickness would decrease due to the formation of a double helix.

Fig.3.3 and Table 3.1 shows the optical properties of immobilized ssDNA and corresponding hybridized (25%) DNA from ellipsometry measurement. A complex index of refraction N is expressed as $n - ik$ where n is the real refractive index and k is absorbency coefficient or imaginary refractive index. Both n and k can be measured by ellipsometry. For ssDNA and dsDNA, n and k are represented as n_{ss} and k_{ss} and n_{ds} and k_{ds}, respectively. The properties are based on modeling the ellipsometry data with a four-layer structure. The structure from air to substrate is composed of (i) ssDNA covalently

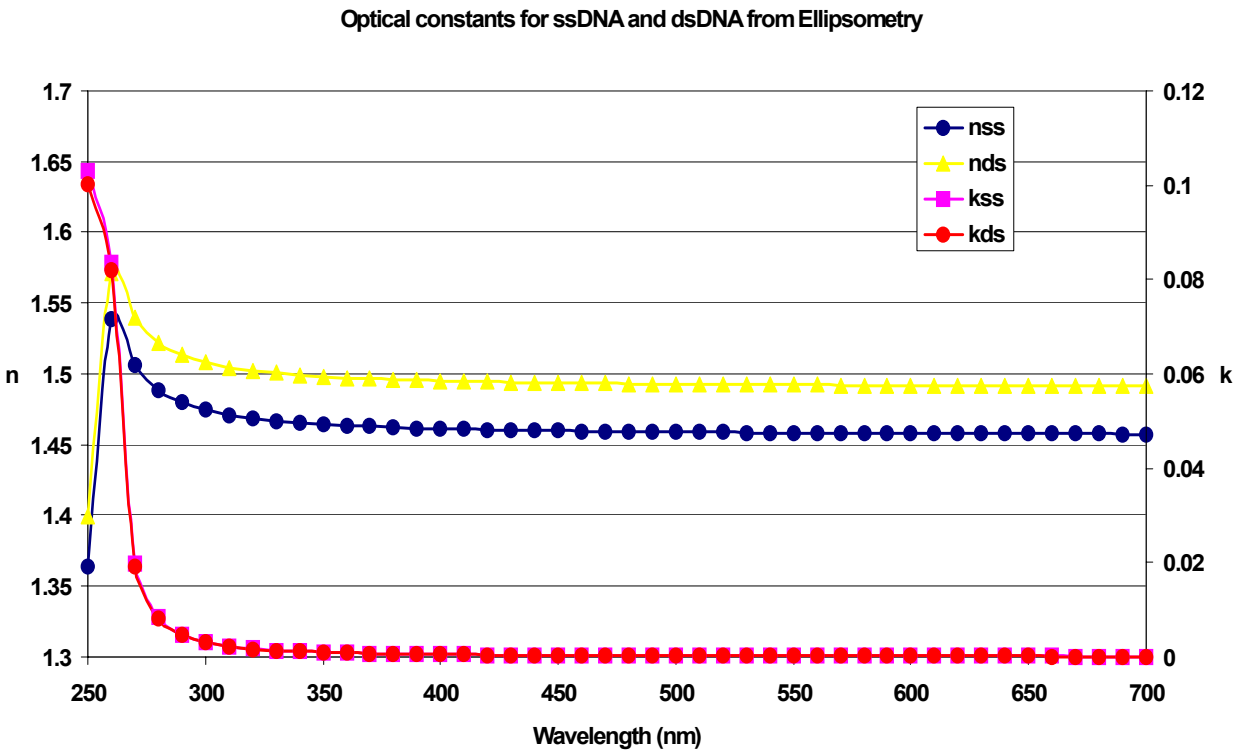


Fig. 3.3. Optical constants of immobilized ssDNA and hybridized dsDNA (25%).

Wavelength (nm)	n_{ss}	k_{ss}	n_{ds}	k_{ds}	n_{ds}-n_{ss}	k_{ds}-k_{ss}
260	1.5392	0.083595	1.5712	0.08189	0.032	-0.0017
270	1.5067	0.019706	1.5395	0.019285	0.0328	-0.0004
280	1.4888	0.008435	1.522	0.008251	0.0332	-0.0002
290	1.4798	0.004775	1.5132	0.004669	0.0334	-0.0001
300	1.4745	0.003137	1.5081	0.003068	0.0336	-7E-05
310	1.4711	0.002259	1.5047	0.002208	0.0336	-5E-05
320	1.4687	0.001729	1.5023	0.00169	0.0336	-4E-05
330	1.4669	0.001382	1.5006	0.001351	0.0337	-3E-05
340	1.4655	0.001141	1.4993	0.001116	0.0338	-3E-05
350	1.4644	0.000966	1.4982	0.000945	0.0338	-2E-05
360	1.4636	0.000835	1.4973	0.000816	0.0337	-2E-05
370	1.4629	0.000733	1.4966	0.000716	0.0337	-2E-05
380	1.4622	0.000652	1.496	0.000637	0.0338	-1E-05
390	1.4617	0.000586	1.4955	0.000573	0.0338	-1E-05
400	1.4613	0.000532	1.4951	0.00052	0.0338	-1E-05
410	1.4609	0.000487	1.4947	0.000476	0.0338	-1E-05
420	1.4606	0.000448	1.4944	0.000438	0.0338	-1E-05
430	1.4603	0.000416	1.4941	0.000406	0.0338	-9E-06
440	1.46	0.000387	1.4939	0.000379	0.0339	-9E-06
450	1.4598	0.000363	1.4936	0.000354	0.0338	-8E-06
460	1.4596	0.000341	1.4934	0.000333	0.0338	-8E-06
470	1.4594	0.000322	1.4933	0.000314	0.0339	-7E-06
480	1.4592	0.000304	1.4931	0.000298	0.0339	-7E-06
490	1.4591	0.000289	1.4929	0.000283	0.0338	-7E-06
500	1.4589	0.000275	1.4928	0.000269	0.0339	-6E-06
510	1.4588	0.000263	1.4927	0.000257	0.0339	-6E-06
520	1.4587	0.000251	1.4926	0.000246	0.0339	-6E-06
530	1.4586	0.000241	1.4924	0.000235	0.0338	-5E-06
540	1.4585	0.000231	1.4923	0.000226	0.0338	-5E-06
550	1.4584	0.000222	1.4923	0.000217	0.0339	-5E-06
560	1.4583	0.000214	1.4922	0.00021	0.0339	-5E-06
570	1.4582	0.000207	1.4921	0.000202	0.0339	-5E-06
580	1.4581	0.0002	1.492	0.000195	0.0339	-5E-06
590	1.458	0.000193	1.4919	0.000189	0.0339	-4E-06
600	1.458	0.000187	1.4919	0.000183	0.0339	-4E-06

Table 3.1. Optical constants of immobilized 27 nt ssDNA and hybridized (25%) dsDNA on PS/Si measured by ellipsometry. A plot of the data is shown in Fig. 3.3.

attached on (ii) PS film that is deposited on (iii) native SiO_2 layer on (iv) Si substrate. The optical parameter for Si and SiO_2 are built into the fitting program. The thickness of SiO_2, optical properties and thickness of PS film, were independently measured as described in the previous section. Note that, although the differences in Ψ and Δ were ~1^o between ssDNA and dsDNA, the precision and accuracy of the ellipsometric measurements is 0.02° and 0.05° for ψ and Δ, respectively (138), making the change significant and measurable. The Cauchy model is valid for dielectric materials and fits well for the PS film but does not model ssDNA and dsDNA layers. The Lorentz model that is more applicable for conductive materials fits the data well for DNA layers. The reason for the validity of the Lorentz model is the specific oscillator at E_n = 4.76 eV for both ssDNA and dsDNA (see Table 3.2) independently confirmed by absorption peak at 259 nm in UV-Vis spectrum. Additionally, the Lorentz dispersion equation for dielectrics is Kramer-Kronig consistent (139).

	ssDNA	dsDNA
Lorentz Model for ellipsometry data		
Oscillator energy, E_n (eV), fixed	4.76 eV	4.76 eV
Oscillator strength, A_m (eV^2), fitted	0.5	0.5
Oscillator bandwidth, B_r (eV),fitted	0.2	0.2
Dielectric constant, $\varepsilon(\infty)$, fixed	2.1	2.2
Film thickness (nm), fitted	2	4.3
k_{ss} and k_{ds} at 260 nm from ellipsometry, fitted	0.08359	0.08189
Experimentally measured parameters		
Oscillator energy from UV-vis spectrum	4.76 eV	4.76 eV
k_{ss} and k_{ds} at 260 nm from UV-vis spectrum	0.074	0.072

Table 3.2. Parameters in the Lorentz model for fitting optical constants and thickness of DNA.

Except for the significant dependence in the 260 nm region, the optical properties of DNA are nominally independent of λ, as can be seen in Fig. 3.3. This was expected because of the strong Lorentz oscillator at 4.76 eV. For maximum absorption at 260 nm, Table 3.2 compares the k for ssDNA and dsDNA directly measured by UV-Vis transmission and values obtained by the Lorentz fitting of the ellipsometry data. The excellent consistency between the experimentally obtained value (from UV-Vis absorption spectrum) and fitted value (from ellipsometry data) indicates that the oscillator parameters A_m and B_r for the Lorentz model are reasonable. Electron conduction in dsDNA increases compared to ssDNA, thus lead to higher k for dsDNA. However the difference in UV-Vis range is insignificant. The insensitivity to conductivity on k may be perhaps due to slow mobility of charge transport in dsDNA compared to the high frequency of optical light.

Another experimental evidence for the consistency of the fitting parameter for the Lorentz model is the high frequency dielectric constant, ε (∞), where ω is the frequency of the applied electric field. The optical frequency may be considered high frequency where optical dispersion (i.e., n and k) only depends on the electronic polarizability. Since k is negligible in the optical range, by Maxwell's relation, $\varepsilon = n^2$. From the n at the plateau values in Fig. 3.3 (where the oscillator effects at 260 nm are insignificant), the calculated dielectric constant ε (∞) for ssDNA and dsDNA is 2.12 and 2.23, respectively, that are consistent with the values obtained by Lorentz fitting as indicated in Table 3.2. The n_{ss} at the plateau region (λ in 600-700 nm) of 1.458 in Fig. 3.3 is consistent with 1.462 at 633 nm reported for 25 nt tethered ssDNA (140).

In principle, due to electron delocalization effects in dsDNA due to π-electron stacking of the bases, the polarizability of dsDNA should be higher. As noted earlier, such electron delocalization effect is reported to impart a semi-conducting property to the dsDNA chain. The change in k (that is proportional to electrical conductivity) also indicates that the delocalization effects at these frequencies are not insignificant. The density of dsDNA is higher than that of ssDNA because of the addition of hybridized target ssDNA chains. This could also cause the increase in n because of change in polarizability due to densification of the monolayer. One way to quantitatively estimate the relative amount of the two effects is by comparing the expected change in refractive index due to increase in density (from ssDNA to dsDNA) with the experimentally obtained change in Table 3.1.

Let, α_1 and α_2 be the average polarizability per nucleotide in ssDNA and dsDNA. For simplicity, assume that the average molecular weight of all the base units is the same. From Clausius-Mossotti equation (141),

$$\frac{n_{ss}^2 - 1}{n_{ss}^2 + 2} = \frac{x\phi}{d_{ss}}\alpha_1 \tag{3.2.1}$$

$$\frac{n_{ds}^2 - 1}{n_{ds}^2 + 2} = \frac{2x\phi\eta}{d_{ds}}\alpha_2 + \frac{x\phi(1-\eta)}{d_{ss}}\alpha_1 \tag{3.2.2}$$

where x =27 is the number of units per chain, d_{ss} = 2 nm, d_{ds} = 4.3 nm, η = 0.25, ϕ = 6×10^{11} molecules/cm^2, and n_{ss} and n_{ds} are given in Table 3.1.

Fig. 3.4 shows the expected change in refractive index due to the densification effect. The expected change due to densification is calculated by assuming $\alpha_1 = \alpha_2$. From n_{ss} in Table 3.1 and equation 3.2.1, α_1 is calculated and substituted in equation 3.2.2 to

obtain n_{ds}. The refractive index contrast due to the calculated densification effect is negative and the absolute value is insignificant when compared to the experimentally measured value, as shown in Fig 3.4. This suggests that the densification of dsDNA is not the main cause of the increase of n and perhaps $\alpha_1 < \alpha_2$. The average polarizability per nucleotide unit for ssDNA and dsDNA is estimated from the experimental n and k for ssDNA and dsDNA in Table 3.1 and equations 3.2.1 and 3.2.2, respectively. As expected, the estimated α_2 is higher than α_1 as shown in Fig. 3.5. Consistent with the refractive index in Table 3.1, the polarizability increases for lower λ, and is relatively constant in the near-UV and visible range.

Upon dehybridization, dsDNA samples reversibly change to 'original' immobilized ssDNA monolayers. The average thickness and optical constants measured by ellipsometry go back to the ranges of the ssDNA. The average fitting parameters are quoted in Table 3.2.

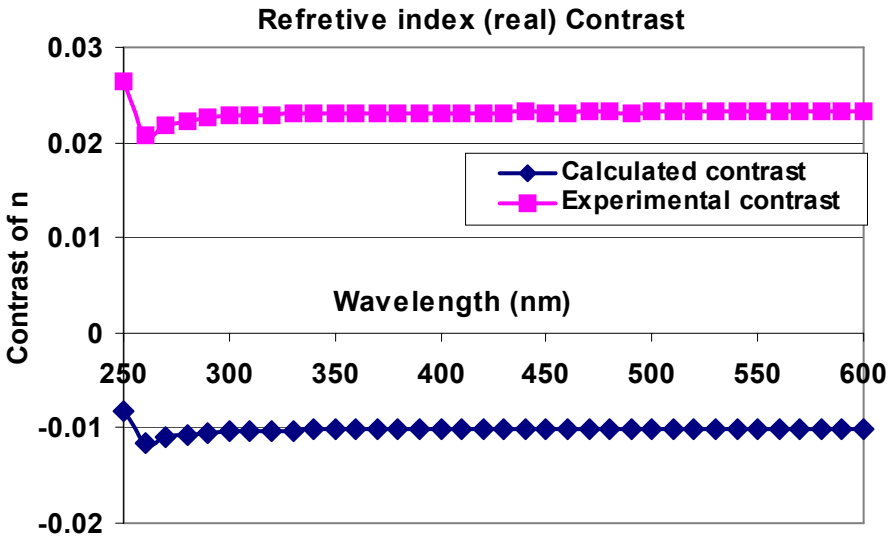


Fig. 3.4. The refractive index change between ssDNA and dsDNA as a function of wavelength. The calculated contrast from simple increase in molecular weight is negative and insignificant. The contrast is nominally constant in the near UV and visible range

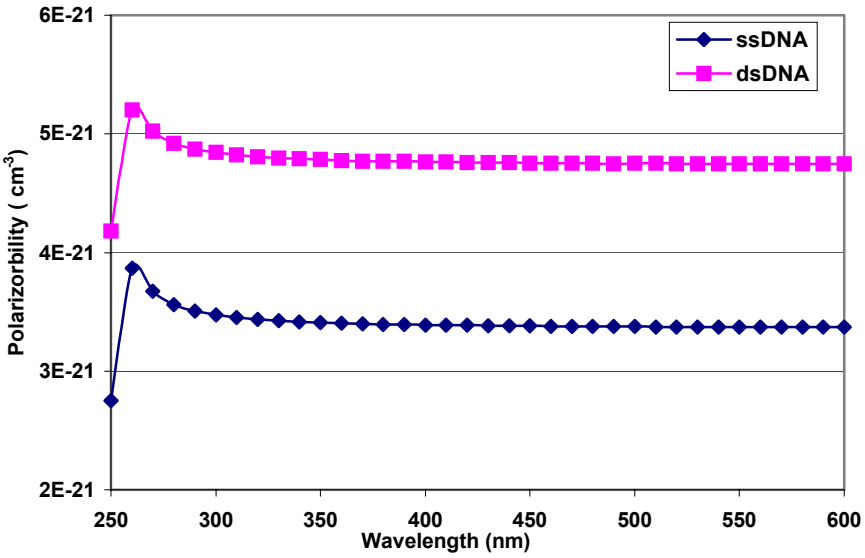


Fig. 3.5. Polarizability per nucleotide for ssDNA and dsDNA is calculated from the data in Tab.3.1 and Eqs. (3.21) and (3.2.2).

3.3 The model and discussion

First, absorbance alone is considered. If light passes through a material with thickness of y, and if at the incident interface the irradiance is I (0), the intensity passing through the material is given by Beer's Law as,

$$I(y) = I(0)\, e^{-\alpha y} \qquad (3.3.1)$$

where the absorbance coefficient $\alpha = 2\omega k/C$, C is the speed of light, ω is the frequency of the light wave and k is the imaginary refractive index of the material.

At wavelength 325 nm that is the fluorescence emission of PS, k of ssDNA and dsDNA is almost equal. Therefore no contrast is generated due to absorption of emitted light. Contrast due to absorption of the excitation light can be estimated by putting the k value from Table 3.1 into equation 3.3.1, and y is 2 nm and 4.3 nm for ssDNA and dsDNA respectively. The change in transmitted light [$I(4.3\text{ nm})_{ds} - I(2\text{ nm})_{ss}$] / $I(2\text{ nm})_{ss}$ equals -1.2%. Two points are noted: (a) the predicted change is very small compared to 40% obtained experimentally; (b) the contrast predicted is opposite to that from the experimental result. Furthermore, absorption alone

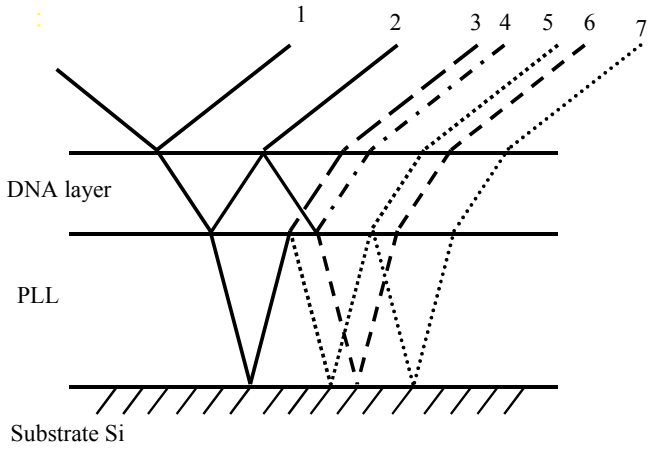


Fig. 3.6. Scheme for reflection model. Only the first three rays are considered in the simulation.

cannot explain the contrast dependence on PS film thickness.

The second model developed was to consider change in reflectivity due to hybridization, as shown in Fig. 3.6. Layers 0, 1, 2 and 3 represent respectively air, DNA, PS and Si substrate. The DNA layer was modeled as a continuous solid film. To simplify the model, only the first three reflections from the multilayer system were taken into consideration. The higher order reflections are ignored because the intensity levels become less than 1% compared to the first term. Total reflected amplitude R is,

$$R = r_{01} + t_{01}r_{12}t_{10}e^{-j2\beta_1} + t_{01}t_{12}r_{23}t_{21}t_{10}e^{-j2\beta_1}e^{-j2\beta_2} \qquad (3.3.2)$$

where r and t are Fresnel's coefficients for reflection and transmission at the interfaces denoted by the subscripts. For example, r_{01} is the reflection coefficient at the 0-1 interface and t_{10} is the transmission coefficient at the 1-0 interface. β is the phase change that the multiply-reflected wave inside the film experiences as it traverses the film once from one boundary to the other. β are given by

$$\beta_1 = 2\pi(d_1/\lambda)N_1\cos\phi_1 \qquad (3.3.3)$$

and $$\beta_2 = 2\pi(d_2/\lambda)N_2\cos\phi_2 \qquad (3.3.4)$$

where λ is the wavelength, d is the film thickness, N is the film complex refractive index and ϕ is the refraction angle. Subscripts indicate the corresponding layers.

Incident light is non-polarized light and can be considered as 50% p (parallel to the plane of incidence) polarized light and 50% s (perpendicular to the plane of incidence) polarized light. β for p and s polarization is the same. Terms in equation 3.3.2 are the corresponding Fresnel coefficients at that polarization (142). All Fresnel

coefficients are expressed in terms of complex refractive index N and the refraction angel ϕ of corresponding layers. Final reflectivity would be

$$R\% = (R_p^2 + R_s^2)/2 \qquad (3.3.5)$$

where R_p and R_s are the reflection amplitude of p and s polarization light respectively.

Given that the incident light wavelength is 265 nm, n and k for ssDNA and dsDNA are $n_{ss} = 1.523$, $k_{ss} = 0.05165$, $n_{ds} = 1.555$ and $k_{ds} = 0.0506$ as obtained by interpolating the data in Table 3.1. Contrast due to reflection from ssDNA to dsDNA versus PS film thickness was calculated and plotted in Fig. 3.7. The Matlab program used for reflection calculation is in Appendix B.

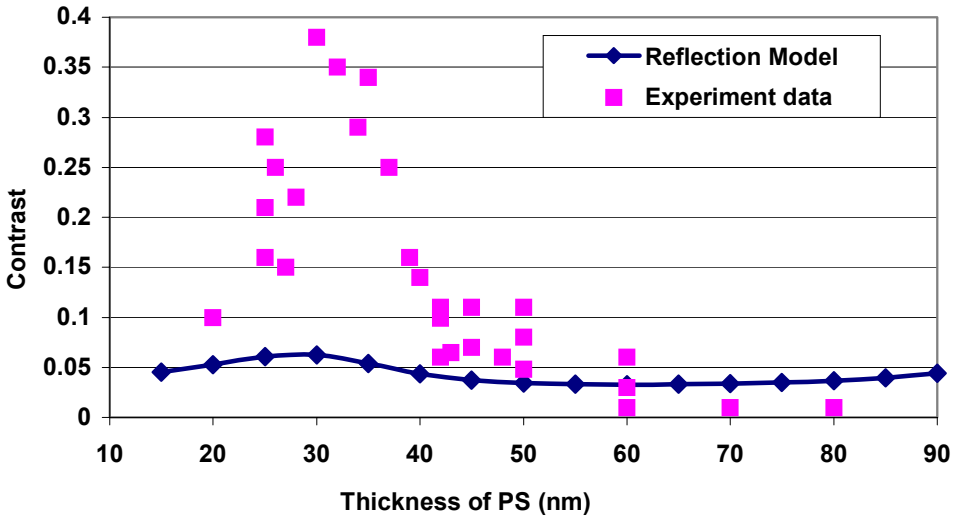
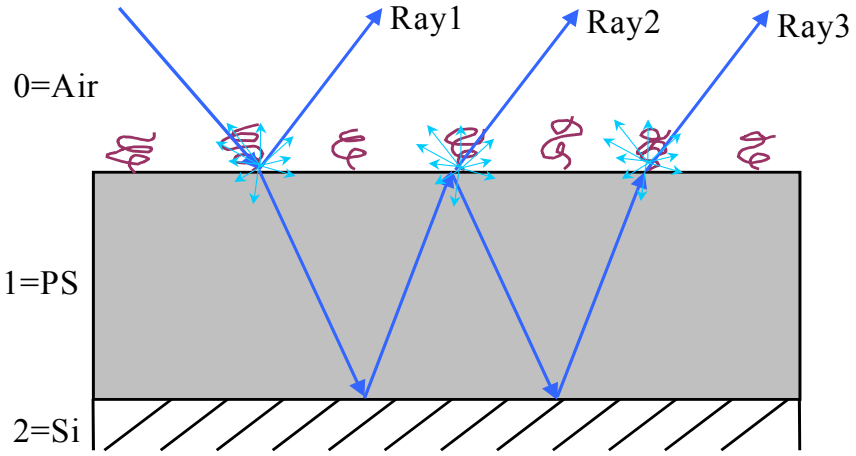


Fig. 3.7. Comparison of the result from reflection model and the experimental data.

The contrast magnitude of this simulation is larger than the prediction from absorbance but still small compared to the experiment value. However at PS thickness equals 30 nm, the simulated contrast reaches maximum. This matches well with the experiment peak position. The peak width also seems to match the experimental data. Simulation also shows that changes of refractive indexes of DNA in a reasonable range do not affect the contrast much and the optical constants for DNA can not vary significantly. The contrast seems more sensitive to thickness of DNA layers. An extreme case is to assume 1 nm and 9 nm for ssDNA and dsDNA respectively. Consequentially, the contrast increases with the maximum value to 17% versus 35% of experiment data. The peak position shifted from PS thickness at 30 nm to 70 nm. The conclusion that can be drawn is that multiple reflection contributes to the contrast and causes dependence of the contrast on PS thickness but does not completely explain the experimental observation. Simulation results of contrast are lower than experimental data even at extreme cases favoring contrast.

The above is a pure multilayer reflection model with absorption being considered by way of using complex refractive index. Although the thickness of the DNA layers is obtained from ellipsometry, thus it is the average effective thickness of the layers, in reality DNA molecules occupy only a small portion of the volume of the assumed “DNA film” and they are most probably distributed on the PS surface as isolated individuals, away from each other. Another reason could be that light activities also include scattering as described in Chapter 2, but in this model scattering was not considered and it could be the amplifying factor.



Ray1: $r01 \times (1+f_s)$

Ray2: $t01 \times r12 \times t10 \times (1-f_s) \times e^{-2j\beta}$

Ray3: $t01 \times r12^2 \times r10 \times (1+f_s) \times t10 \times (1-f_s) \times e^{-4j\beta}$

Fig. 3.8. Scheme of the scattering model.

A new model was built where the DNA molecules scatter light, as shown in Fig. 3.8. Since the coverage is low, DNA molecules are considered as individual scattering centers. Incident light first strikes on these scattering centers before reaching the PS film. The forward scattered light is then reflected at the air/PS interface and refracted into the PS film followed by reflection at the PS/Si interface. The light coming out of the PS film again is scattered by DNA molecules. This model was improved from the reflection model by incorporating scattering activity and treating the DNA probes as individual molecules, perhaps a better physical picture. Because the radius of gyration of DNA molecules 2 nm is much smaller than incident wavelength 265 nm, Rayleigh scattering mechanism is valid and given by

$$I_s = \frac{I_o}{r^2}\alpha^2\left(\frac{2\pi}{\lambda}\right)^4 \frac{1+\cos^2(\Theta)}{2} \tag{3.3.6}$$

where I_s is the intensity scattered by one molecule, I_0 is the incident intensity, r is the distance from the scattering center, λ is the wavelength of the light, Θ is the angle between the incident direction and scattering direction and α is the polarizability of DNA molecules and is given by the Lorentz-Lorentz formula,

$$\alpha = \frac{3}{4}\frac{1}{\pi}\frac{1}{N_s}\left(\frac{N^2-1}{N^2+2}\right) \tag{3.3.7}$$

where N_s is the number of molecules per unit volume and N is the complex refractive index.

For a monolayer system, the molecules per unit volume N_s can be calculated as ϕ/d where d is the thickness of the monolayer and ϕ is the surface coverage of the molecules in units of number of molecules per unit area. By multiplying I_s with the

sphere area of radius r, $4\pi r^2$, and integrating Θ term from 0 to π, total scattered irradiance by the molecules in the unit area on a surface is obtained as,

$$I_{st} / I_0 = 96\pi^3 d^2 \left(\frac{N^2 - 1}{N^2 + 2} \right)^2 \frac{1}{\lambda^4} \frac{1}{\phi} \tag{3.3.8}$$

This is the energy portion scattered from the incident light.

In order to incorporate scattering in the model together with the reflection, an amplitude factor is defined as,

$$f_s = \sqrt{I_{st} / I_0} = 54.5d \left(\frac{N^2 - 1}{N^2 + 2} \right) \frac{1}{\lambda^2} \frac{1}{\sqrt{\phi}} \tag{3.3.9}$$

where f_s is the ratio by which the amplitude of the scattered light wave to that of the incident light wave.

Since DNA molecules are treated as individuals, the optical constants measured by ellipsometry probably need some modification for dsDNA because only 25% of ssDNA were hybridized. Assuming a composite relationship, $n_{dsr} \times 0.25 + n_{ss} \times 0.75 = n_{ds}$, where n_{dsr} is the n corresponding to 100% hybridization and, n_{ss} and n_{ds} are the refractive index measured by ellipsometry. $n_{ss} = 1.523$ and $n_{ds} = 1.555$ gives $n_{dsr} = 1.651$. Estimating k the same way, with interpolated $k_{ss} = 0.05165$ and $k_{ds} = 0.0506$ at 265 nm, $k_{dsr} = 0.04745$ is obtained. From these optical constants and assuming 100% hybridization, f_s for ssDNA and dsDNA calculated by equation 3.3.9 is 0.61 and 1.57 respectively, which indicates 157% contrast. Again assuming composite relationship, 25% hybridized sample would have $f_s = f_{s,\,\mathrm{ssDNA}} \times 0.25 + f_{s,\,\mathrm{dsDNA}} \times 0.75 = 0.85$, which indicates 40% contrast. Notice that scattering contrast is very high; therefore it is possible

that scattering contributes strongly to the final contrast. Also note that the scattering effect is independent of the PS film thickness while experimental data show that contrast varies with the thickness of the PS film indicating interference activity from reflection. Thus, the model must incorporate both reflection and scattering.

If DNA molecules are not considered as a continuous reflective film but only as scattering centers shown in Fig. 3.8, reflection only occurs at air/PS and PS/Si interface. Incident light is scattered by DNA molecules before it is reflected at the PS surface. Half of the scattered light enters the PS film and the other half is scattered back into the air. So a scattering factor is added to Fresnel coefficients. Equation 3.3.2 would then be modified to,

$$R = r_{01}(1+f_s) + t_{01}\, r_{12}\, t_{10}(1-f_s)\, e^{-j2\beta_1} + t_{01}\, r_{10}\, r_{12}^2\, t_{10}(1+f_s)(1-f_s)\, e^{-j4\beta_1} \qquad (3.3.10)$$

where f_s is from equation 3.3.9, the subscripts in equation 3.3.10 are different from equation 3.3.2. Here 0, 1 and 2 represent air, PS film and Si respectively. Factors $(1-f_s)$ and $(1+f_s)$ are for attenuation and enhancement due to scattering. Since all the optical constants for dry ssDNA and dsDNA of the sequence used in this study is experimentally measured, ϕ is experimentally estimated, and d is estimated by both experimentally and theoretically, all the constants in equations 3.3.9 and 3.3.10 are known. Thus, the calculation of contrast does not require any fitting parameters. A Matlab program for contrast calculation is attached in Appendix B.

Fig. 3.9 shows the result of the calculated contrast from the scattering-reflection model versus PS thickness and experimental result for comparison. Also shown is the

calculated contrast from the reflection-only model. The result from the scattering-reflection model and the experiment curve match well. The mechanism is proposed as follows. (a) The dependence of the contrast on PS thickness is attributed to the multiple reflection; (b) The scattering of UV light of 265 nm by DNA is the "amplifier" to the contrast.

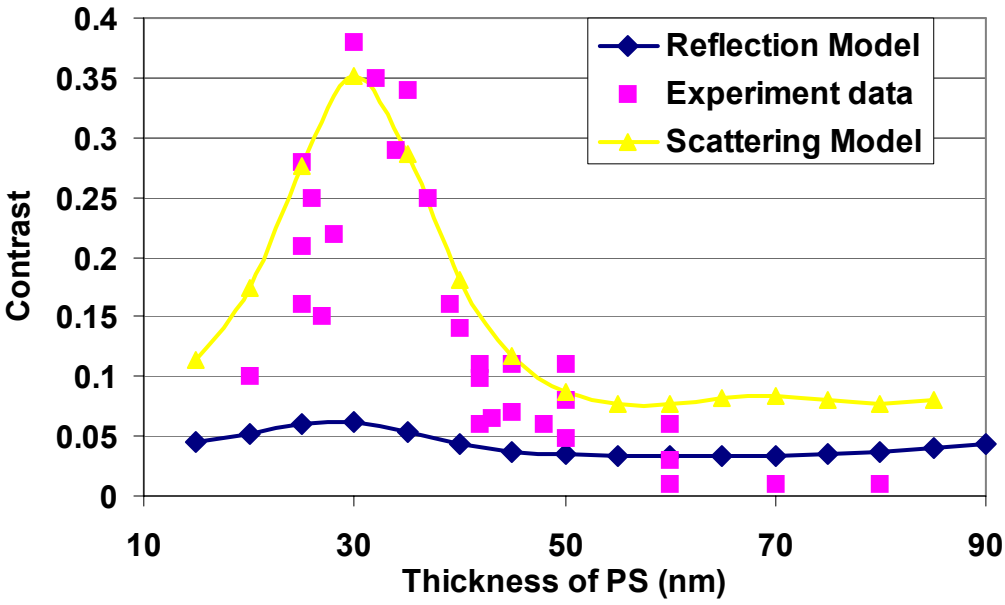


Fig. 3.9. Comparison of the result from scattering model and the experiment data. Also shown is the result from the reflection model.

3.4 Conclusion

The reason behind the contrast for the designed DNA-chip technique is the difference in refractive index and dimension between ssDNA and dsDNA. Refractive indexes and conformation change due to the transition from ssDNA to partially hybridized ssDNA of the top DNA layer of the device is amplified by the multiple reflections of the excitation light in PS films and the scattering from the DNA molecules.

Chapter 4

Conclusions and future work

4.1 Conclusions

This research has demonstrated a label-free fluorescence method that is compatible with the current DNA microarray read-out (confocal microscope) platform. The mechanism of the signal contrast generation is the change in scattering caused by large difference in refractive indexes and conformations between ssDNA and dsDNA. The scattering change is converted to the change in fluorescence from the underlining polystyrene layer that has excitation wavelength within 5 nm from the wavelength of maximum refractive index change between ssDNA and dsDNA. The characteristic curve of contrast as a function of degree of hybridization (i.e. binding) is a highly linear function over the entire percent hybridization range from 0 to 25%. The high linearity of the assay response allows for quantitative analysis of level of binding. The method is sensitive to single mismatch.

4.2 Some preparations and recommendations for future work

4.2.1 Instrument development to detect signal contrast on micro-scale spot and automation to perform array analysis

If the studied technique can be further developed to DNA microarray, the characteristic detection dimension should be in micrometer scale. In earlier research, detection was performed by using the Hitachi F-4500 fluorescence spectrophotometer.

Incident beam shape of this spectrophotometer is rectangular with an area of 0.24 cm^2. This is too large for a DNA chip. A micro-spectrophotometer has been designed as shown in Figure 4.1. The detection size as small as 300 μm is possible and the reduction of spot size will be compensated by increasing numerical aperture from the current ~ 0.1 to 0.5. A photon counter with an integrated low noise amplifier will be used for detector. Counting time of the photon counter is software adjustable and longer counting time will ensure a higher signal to noise ratio.

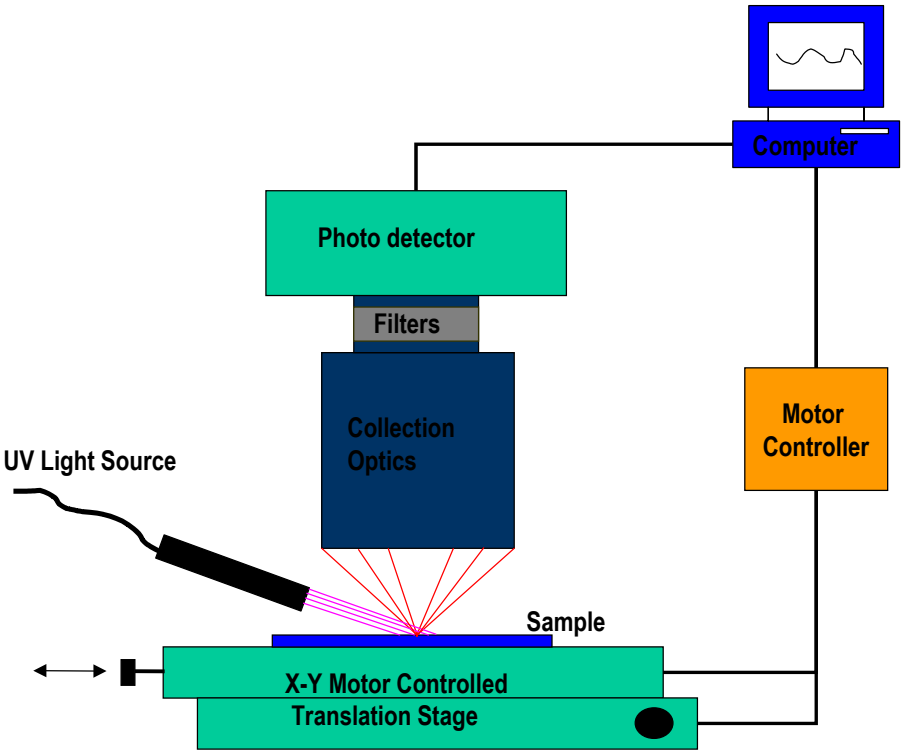


Fig. 4.1. Block diagram showing the setup of micro-photometer.

In order to perform array analysis, some automation work has been done. Instead of setting the sample on a stationary plate, a two-dimensional step motor driven translation stage (Newport) is assembled to enable automated scanning of the chip. In-house software has been developed to search for the same starting position on each chip and to recognize partitions on an array.

4.2.2 Fabrication of a DNA macro-array

A 3 × 3 array has been designed for testing the potential of the studied DNA chip method for SNP detection, as shown in Fig. 4.2.

Three different probes are put on the pixels. Each probe will be immobilized onto three separate pixels as show in Fig. 4.2. The dimension of each pixel is approximately 3.5 × 3.5 mm. The sequences of these probes and of the target are shown in Table 4.1. Each pixel holds the aqueous solution of ssDNA and grafting buffer and serves as an individual reactor. After the grafting process, the sample is scanned in the micro-photometer and fluorescence levels of each pixel are recorded. Then the sample is hybridized with the target that is also listed in Table 4.1. The target is a perfect match to probe PM and has one and two mismatches to probe MM1 and MM2. The sample is then washed at different temperatures starting from low to high. After each wash, the sample is scanned in the micro-photometer. The potential of sensing mismatches is tested in this way.

4.2.3 Nanoparticle/polymer composite for higher contrast and visible light probe

As described in Chapter 3, the fluorescence contrast when the top layer changes from ssDNA to dsDNA is proportional to the total fluorescence efficiency of the PS. While there is no way to improve the fluorescence efficiency of PS itself, it is possible to incorporate inorganic fluorescence nanoparticles such as ZnS in the PS film. Another reason to consider nanoparticles is that their fluorescence can be in the visible wavelength range. This can make signal collection and detection easier and more efficient. Although differences of refractive indexes between ssDNA and dsDNA are

very small in the visible wavelength range, the fluorescence efficiency improvement possibly can compensate. Many polymers are good support materials. Their surface is relatively easy for modification and is safe to biomaterials. Thus, a nanoparticle/polymer composite can be a good choice for the fluorescence layer. Novel nanoparticles chemistry is being explored to obtain high fluorescence with excitation in the 265 nm and visible range. Preliminary studies on doped II-VI compounds exhibit some promise.

Sequence Name	Sequence
Probe PM	5'-CAA-AAT-AGA-CGC-TTA-CGC-AAC-GAA-AAC-3'
Probe 1MM	5'-CAA-AAT-AGA-CGC-TTA-C**C**C-AAC-GAA-AAC-3'
Probe 2MM	5'-CAA-AAT-AG**T**-CGC-TTA-CGC-AAC-**T**AA-AAC-3'
Target	3'-TTA-TCT-GCG-AAT-GCG-TTG-CTT-TTG-ATA-5'

Table 4.1. Sequences to be used in mismatch sensitivity test. Bold characters stand for mismatched nucleotides.

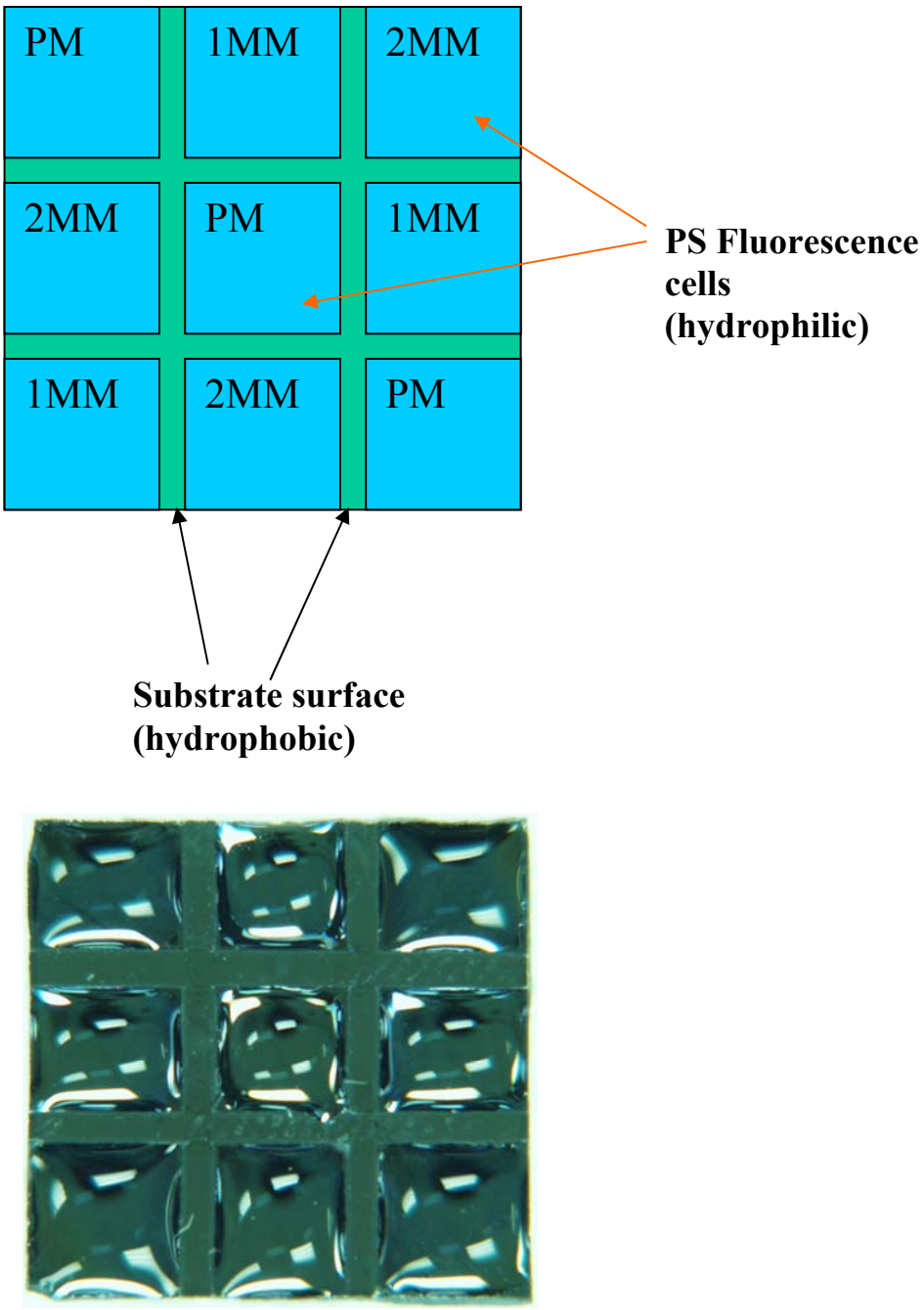


Fig. 4.2. Macro-array for mismatch test. The upper is the array design and lower is the fabricated sample.

Reference

1. Darnell, J.; Lodish, H.; Baltimore, D. *Molecular Cell Biology*, 2nd Ed., Scientific American Book: New York, 1990
2. Watson, J.D.; Crick, F. H. *Nature* **248**,737-738 (1953)
3. Barrell, B. *FASEB J.* **5**, 40-45 (1991)
4. Sanger, F.; Nicklen, S.; Coulson, A.R. *Proc. Natl. Acad. Sci. USA* **74**, 5463-5467 (1977)
5. Maxam, A. M.; Gilbert, W. *Proc. Natl. Acad. Sci. USA* **74**, 560-564 (1977)
6. Sambrook, J.; Fritisch, E.F.; Maniatis, T. *Molecular Cloning: A laboratory Manual*, 2nd ed., Cold Spring Harbor Laboratory Press: Could Spring Harbor, NY, 1989
7. Drmanac, R.; Crkvenjakov, R. US patent no. 5,202,231 (1993)
8. Southern, E. M., *J. Mol. Biol.* **98**, 503-517 (1975)
9. Wallace, R.B.; Shaffer, J.; Murphy, R.F. *Nucleic Acids Res.* **6**, 3543-3557 (1979)
10. Saiki, R.K.; Walsh, P.S.; Levenson, C.H.; Erlich, H.A. *Proc. Natl. Acad. Sci.* **86**, 6230-6234 (1989)
11. Breslauer, K. J.; Frank, R.; Blocker, H. et al. *Proc. Natl. Acad. Sci. USA* **83**, 3746-3750 (1986)
12. Southern E. M. International Patent Application PCT/GB 89/00460 (1988)
13. Lysov, Y. P. *Dokl. Akad. Nauk. SSSR* **303**, 1508-1511 (1988)
14. Drmanac, R. *Science* **260**, 1649-1652 (1993)
15. Southern, E. M. *J. Mol. Biol.* **98**, 503-517 (1975)
16. Maskos, U.; Southern, E. M. *Nucleic Acids Res.* **20**, 1675-1678 (1992)
17. Southern, E. M.; Maskos, U.; Elder, J.K. *Genomics* **13**, 1008-1017 (1992)
18. Blanchard, A. P. *Biosersors and Bioelectronics* **11**, 687-690 (1996)
19. Maskos, U.; Southern, E. M. *Nucleic Acids Res.* **21**, 2269-2270 (1993)
20. Southern, E. M.; Casegreen, S.C.; Elder, J.K. et al. *Nucleic Acids Res*. **22**, 1368-1373 (1994)
21. Fodor, S. P. A.; Read, J.L.; Pirrung, M.C. et al. *Science* **251**, 767-773 (1991)
22. Pirrung, M. C.; Fallon, L,; McGall, G. *J. Org. Chem.* **63**, 241-246 (1998)
23. Guo, Z.; Guilfoyle, R.A.; Thiel, A.J. et al. *Nucleic Acids Res.* **22**, 5456-5465 (1994)
24. Schena, M.; Shalon, D.; Davis, R.W. et al. *Science* **270**, 467-470 (1995)
25. Larsson, L.; Hougaard, D.M. *Histochemistry*, **101**, 325 (1994)
26. Xu, X.; Yang, H.C.; Mallouk, T.E. et al. *Amer. Chem. Soc*., **116** 8386 (1994)
27. Wang, J.; Nielsen, P.E.; Jiang, M. et al. *Anal. Chem*. **69**, 5200 (1997)
28. Su, H.; Yang, M.; Kallury, K.M.R. et al. *Analyst*, **118**, 309 (1993)
29. Pandey, P.C.; Weetall, H.H. *App. Biochem. Biotechnol*., **55**, 87 (1995)
30. Allemand, J.F.; Bensimon, D.; Jullien, L. et al. *Biophys. J.* **73**, 2064-2070 (1997)
31. Beattie, W.G.; Meng, L.; Turner, S.L. et al. *Molec. Biotech*. **4**, 213 (1995)
32. Rogers, Y.H.; Jiang-Baucom, P.; Huang, Z.J. et al. *Anal. Biochem*. **266**, 23 (1999)
33. Joos, B.; Kuster, H.; Cone, R. *Anal. Biochem*, **247**, 96 (1997)

34. Maskos, U.; Southern, E. M. *Nucleic Acids Res.* **20**, 1679-1684 (1992)
35. Doktycz, M.J.; Beattie, K.L. *Automation Technologies for Genome Characterization*; Edit by Beugelsdijk, T.J. John Wiley & Sons: N.Y. 1997
36. Tyagi, S.; Kramer, F.R. *Nat. Biotech*. **14**, 303 (1996)
37. Boncheva, M.; Scheibler, L.; Lincoln, P.et al. *Langmuir* **15**, 4317 (1999)
38. Herne, T.; Tarlov, M.J. *J. Am. Chem. Sco*. **119**, 8916 (1997)
39. Yang, M.; Yau, H.C.M.; Chan, H.L. *Langmuir* **14**, 6120 (1998)
40. Storhoff, J.J.; Elghanian, R.; Mucic, R.C.et al. *J. Am. Chem. Soc*. **120**, 1959 (1998)
41. Matson, R.S.; Rampal, J. B.; Coassin, P.J. *Anal. Biochem*. **217**, 306 (1994)
42. Weiler, J.; Hoheisel, J.D. *Anal. Biochem*. **243**, 218-227 (1996)
43. Beier, M. Hoheisel, J.D. *Nucleic Acids Res*. **27**(9) 1970-1977 (1999)
44. Zammatteo, N,; Girardeaux, C.; Delforge, D.; Pireaux, J.J.; Remacle, J. *Anal. Biochem*. **236**, 85-94 (1996)
45. www.nalgenunc.com TechNote Volume 3 No.17 NucleoLink versus Covalink Surfaces.
46. Liu, Y.; Rauch, C.B. *Anal. Biochem*. **317**, 76-84 (2003)
47. Shchepinov, M. S.; Udalova, I.A.; Bridgman, A.J. et al. *Nucleic Acids Res.* **25**,4447-4454 (1997)
48. Nikiforov, T. T.; Rendle, R.B.; Goelet, P. et al. *Nucleic Acids Res.* **22**, 4167-4175 (1994)
49. Schena, M. *BioEssays* **18**, 427-431 (1996)
50. Shalon, D.; Smith, S.J.; Brown, P.O. *Genome Res*. **6**, 639-645 (1996)
51. DeRisi, J. L. *Science* **278**, 680-686 (1997)
52. Schena, M.; Shalon, D.; Heller, R. et al. *Proc. Natl. Acad. Sci. USA* **93**, 10614-10619 (1996)
53. Iyer, V.; DeRisi, J.; Eisen, M. et al. *Fed. Am. Soc. Exp. Bio*. **11**, 1126 (1997)
54. R. N. Heller *Proc. Natl. Acad. Sci. USA* **94**, 2150-2155 (1997)
55. DeRisi, J.; Penland, L.; Brown, P.O. et al. *Nat. Genet.* **14**, 457-460 (1996)
56. Drmanac, S.; Stavropoulos, N.A.; Labat, I. et al. *Genomics* **37**, 29-40 (1996)
57. Hacia, J.G.; Brody, L.C.; Chee, M.S. et al. *Nat. Genet*. **14**(4), 441-447 (1996)
58. Lipshutz, R.J.; Morris, D.; Chee, M. et al. *Biotechniques* **19**, 442-447 (1995)
59. Kozal, M.J.; Shah, N.; Shen, N.P. et al. *Nature Medicine* **2**, 753-759 (1996)
60. Cronin, M.T.; Fucini, R.V.; Kim, S.M. et al. *Human Mut*. **7**, 244-255 (1996)
61. Chee, M.; Yang, R.; Hubbell, E. et al. *Science* **274**, 610-613 (1996)
62. Yershov, G.; Barsky, V.; Belgovskiy, A. et al. *Proc. Natl. Acad. Sci. USA* **93**(10), 4913-4918 (1996)
63. Sapolsky, R. J.; Lipshutz, R. J. *Genomics* **33**, 445-456 (1996)
64. Braxton. S.; Bedilion, T. *Curr. Opin. Biotechnol*. **9**, 643-649 (1998)
65. Friend, H. *Br. Med. J.* **319**, 1306 (1999)
66. Marton, M.J.; DeRisi, J.L.; Bennett, H.A. et al. *Nature Med*. **4**(11), 1293-1301 (1998)
67. Corton, J. C.; Anderson, S.P.; Stauber, A.J. et al. CIIT ACT **19**, 1-9 (1999)
68. Berlier, J.E.; Rothe, A.; Buller, G. et al. *J. Histochem Cytochem* **51** (12), 1699-1712 (2003)

69. www.amershambiosciences.com, Application Note. *Stability studies of Dyes in Microarray Applications.*
70. Naef, F.; Lim, D.A.; Patil, N.; Magnasco M. *Physical Review e* **65**, (2002).
71. Haralambidis, J.; Chai, M.; Tregear, G. W. *Nucleic Acids Research* **15**, 4857-4876 (1987).
72. Randolph, J. B.; Waggoner, A. S. *Nucleic Acids Research* **25**, 2923-2929 (1997)
73. Erickson, D.; Li, D. Q.; Krull, U. J. *Analytical Biochemistry* **317**, 186-200 (2003).
74. Hagihara, S.; Kumasawa, H.; Goto, Y. et al., *Nucleic Acids Research* **32**, 278-286 (2004).
75. Niu, S.; Saraf, R.F. *Smart Materials and Structures*, **11**, 778-782 (2002)
76. Rosenzweig, B.A.; Pine, P.S.; Domon, O.E. et al *Environmental Health Perspectives*, **112**(4), 480-487 (2004)
77. Schuchardt, J. *Nucleic Acids Res*. **28** E47 (2000)
78. Nadon, R.; Shoemaker, J. *Trends in Genetics*, **18**(5), 265-271 (2002)
79. Nelson, B.P.; Grimsrud, T.E.; Liles, M.R. et al. *Anal. Chem*. **73**, 1-7 (2001)
80. Peterson, A.W.; Heaton, R.J.; Georgiadis, R.M. *Nucleic Acids Res*. **29**, 5163-5168 (2001)
81. Peterson, A.W.; Wolf, L.K.; Georgiadis, R.M. *J. Am. Chem. Soc*. **124**, 14601-14606 (2002)
82. Song, F. Y.; Zhou, F.M.; Wang, J. et al. *Nucleic Acids Res*, **30**, e72 (2002)
83. Georgiadis R. Peterlinz, K.P.; Peterson, A.W. *J. Am. Chem. Soc*. **122**, 3166-3171 (2000)
84. Wang, J.; Kawde, A.N. *Analyst* **127**, 383-386 (2002)
85. Palecek, E.; Fojta, M. *Anal. Chem*., **73**, 75A (2001)
86. Caruso, F.; Rodda, E.; Furlong, D.F. et al. *Anal. Chem*. **69**, 2043-2049 (1997)
87. McKendry, R.; Zhang, J.Y.; Arntz, Y. et al. *Proc. Natl. Acad. Sci. USA* **99**, 9783-9788 (2002)
88. Reiteri, R.; Grattarola, M.; Butt, H.J. et al. *Sens. Actuators*, **79**, 115-126 (2001)
89. Rakitin, A. *Phys. Rev. Lett.,* **86**, 3607-3673 (2001).
90. Thiel, A.J.; Frutos, A.G.; Jordan, C.E. et al. *Anal. Chem*. **69**, 4948-4956 (1997)
91. Kukanskis, K.; Elkind, J.; Melendez, J. et al. *Anal. Biochem.* **274**, 7-17 (1999)
92. He, L.; Musick, M.D.; Nicewarner, S.R. et al. *J. Am. Chem. Soc*. **122**, 9071-9077 (2000)
93. Brockman, J.M.; Frutos, A.G.; Corn, R.M. *J.Am. Chem. Soc*. **121**, 8044-8051 (1999)
94. Lyon, L.A.; Musick, M.D.; Natan, M.J. *Anal. Chem*. **70**, 5177-5183 (1999)
95. Buckle, P.E.; Davies, R.J.; Kinning, T. et al. *Biosens. Bioelectron*. **8**, 355-363 (1993)
96. Johnston, D.H.; Glasgow, K.; Thorp, H.H. *J. Am. Chem. Soc*. **117**, 8933-8939 (1995)
97. Wang, J.; Kawde, A. Erdem, A.; Salazare, M. *Analyst*, **126**, 2020-2025 (2001)
98. Wang, J.; Jiang, M. *Langmuir* **16**, 2269-2274 (2000)
99. Korri-Youssoufi, H.; Garnier, P.; Srivastava, P. et al. A. *J. Am. Chem. Soc*. **119**, 7388-7389 (1997)
100. Wang, J. Jiang, M. Fortes, A. Mukherjee, B. *Anal. Chim. Acta* **402**, 7-12 (1999)

101. Garnier, F.; Korri-youssoufi, H.; Srivastava, P. et al. *Synth. Metals*, **100**, 89-93 (1999)
102. Cui, Y.; Wei, Q.; Park, H.; Lieber, C.M. *Science*, **293**, 1289-1295 (2001)
103. Hahm, J-I.; Lieber, C.M. *Nanoletters*, **4**(1), 51-54 (2004)
104. Gabl, R.; Feucht, H.-D.; Zeininger, H. et al. *Biosensors Bioelectronics*, **19**, 615-620 (2004)
105. Fritz, J.; Baller, M.K.; Lang, H.P. et al. *Science* **288**, 316-318 (2000)
106. Mckendry, R.; Zhang, J.; Arntz, Y. et al. *Proc. Natl. Acad. Sci. USA,* **99**, 9783-9788 (2002)
107. Bonnet, G.; Tyagi, S.; Libchaber, A. et al. *Proc. Natl. Acad. Sci. USA*, **96**, 6171-6176 (1999)
108. Singh-Zocchi, M.; Dixit, S.; Ivanov, V. et al. *Proc. Natl. Acad. Sci. USA*, **100**, 7605-7610 (2003)
109. Cho, Y-K.; Kim, S.; Lim, G. et al. *Langmuir*, **17**, 7732-7734 (2001)
110. Zhou, D.; Sinniah, K.; Abell, C. et al. *Angew. Chem. Int. Ed*. **42**, 4934-4937, (2003)
111. Saraf, R.F.; Niu, S. *US Patent 6,706,479 B2* (2004)
112. Wild, S.; Kesmodel, L.L. *J. Vac. Sci. Technol. A*, **19**(3) 856-860 (2001)
113. Idage, S.B.; Badrinarayanan, S. *Langmuir*, **14**, 2780-2785 (1998)
114. Alpert. N.L.; Keiser, W.E.; Szymanski, H.A. *IR Theory and Practice of Infrared Spectroscopy*. 2nd edition, Plenum Publishing Corporation. 1973
115. Rasmussen, S.R.; Larsen, M.R.; Rasmussen, S.E. *Anal. Biochem*. **198**, 138-142 (1991)
116. SantaLucia, J. *Proc. Natl. Acad. Sci. U.S.A.*, **95**, 1460-1465 **(**1998)
117. Peyret, N.; Seneviratne, P. A.; Allawi, H. T. et al. J. *Biochem.*, **38**, 3468-3477. **(**1999)
118. http://biotools.idtdna.com/Analyzer/oligocalc.asp.
119. Letowski, J.; Brousseau, R.; Masson, L. *J. Microbiological Meth*. **57**, 269-278 (2004)
120. Cavic, B. A.; McGovern, M. E.; Nisman, R. et al. *Analyst* **126**, 485-490 (2001)
121. Jin, L.; Horgan, A.; Levicky, R. *Langmuir* **19**, 6968-6975 (2003).
122. Shchepinov, M.S.; Case-Green, S.C.; Southern, E.M. *Nucleic Acids Res*. **25**(6), 1155-1161 (1997)
123. Peterson, A. W.; Heaton, R. J.; Georgiadis, R. M. *Nucleic Acids Research* **29**, 5163-5168 (2001).
124. Mirkin C.A.; Taton T.A.; Letsinger R.L. *Abstracts of Papers of the American Chemical Society*, **221**, U380 (2000)
125. Will G.; Kudryashov E.; Duggan E. et al. *Spectrochimica Acta Part A – Molecular and Biomolecular Spectroscopy* **55** (13), 2711-2717. (1999)
126. Rule G.S.; Montagna R.A.; Durst R.A. *Clinical Chemistry* **42** (8), 1206-1209 (1996)
127. Edelstein R.L.; Tamanaha C.R.; Sheehan P.E. et al. *Biosensors & Bioelectronics* **14**, 805-813 (2000)
128. Steemers F.J.; Ferguson J.A.; Walt D.R. *Nature Biotechnology* **18**, 91-94 (2000)

129. Terbrueggen R.H.; Vielmetter J.; Chen Y.P. et al. *Abstract of Papers of the American Chemical Society* **218**, U172 (1999).
130. Jenison R.; Yang S.; Haeberli A. et al. *Nature Biotechnology* **19** (1) 62-65 (2001).
131. Shi W.M.; Reed M.R.; Liu Y.B. et al. *Clinical Chemistry* **48** (11), 2090-2090 (2002).
132. Snyder, P.G.; Rost, M.C.; Bu-Abbud, G.H. et al. *Journal of Applied Physics.,* **60**(9), 3293-3302. (1986).
133. Palik, E.D. *Handbook of optical constants of solids*. 3rd Edition., Academic Press: San Diego. 1998.
134. Brandrup, J.; Immergut, E.H. *Polymer Handbook, 2nd edition*. John Wiley & Sons: New York. 1974.
135. Rivetti, C.; Walker, C.; Bustamante, C. *Journal of Molecular Biology.,* **280**, 41-59 (1998).
136. Tinland, B.; Pluen, A.; Sturm, J.; Weill, G. *Macromolecules.* **30**, 5763-5765 (1997).
137. Hagerman, P.J. *Annual Review of Biophysics and Biophysical Chemistry,* **17**, 265-286 (1988).
138. Herzinger, C.M. *Journal of Applied Physics* **77**, 4677-4687 (1995).
139. Hummel,R.E. *An Electronic Properties of Materials*., Springer-Verlag: Berlin. 1985.
140. Gray, D.E.; CaseGreen, S.C.; Fell, T.S. et al. *Langmuir,* **13**, 2833-2842 (1997).
141. Solymar, L.; Walsh, D. Lectures on the Electrical Properties of Materials. 4th Eds, Oxford Science Publications. 1988
142. Azzam, R.M.A.; Bashara, N.M. *Ellipsometry and Polarized Light*, North-Holland Personal Library. 1987

Appendix

A. Abbreviations and glossary

bp: Base pair
cDNA library :A pool of complementary DNA clones produced by cDNA cloning of total messenger RNA from a single source (cell type, tissue, embryo).
cytidine : The nucleoside containing cytosine as its base
DDBJ: DNA Data Bank of Japan
Denature: The separation of the two strands of a DNA double helix
DNA clone: A section of DNA that has been inserted into a vector molecule, such as a plasmid or a phage chromosome, and then replicated to form many identical copies.
Dot-blot: A technique for measuring the amount of one specific DNA or RNA in a complex mixture. The samples are spotted onto a hybridization membrane (such as nitrocellulose or activated nylon, etc.), fixed and hybridized with a radioactive probe. The extent of labeling (as determined by autoradiography and densitometry) is proportional to the concentration of the target molecule in the sample. Standards provide a means of calibrating the results
Electrophoresis: A technique for separating the components of a mixture of charged molecules (proteins, DNAs, or RNAs) in an electric field within a gel or other support. The movement of electrically charged molecules in an electric field often resulting in their separation
EMBL: European Molecular Biology Laboratory
Exon : A region of a gene that is present in the final functional transcript (mRNA) from that gene. Any non-intron section of the coding sequence of a gene; together the exons constitute the mRNA and are translated into protein.
Exonuclease: An enzyme which digests nucleic acids starting at one end. An example is Exonuclease III, which digests only double-stranded DNA starting from the 3' end
Expression: To "express" a gene is to cause it to function. A gene which encodes a protein will, when expressed, be transcribed and translated to produce that protein. A gene which encodes an RNA rather than a protein (for example, a rRNA gene) will produce that RNA when expressed.
Gene map: The DNA sequence of a gene annotated with sites of regulatory elements, introns, exons and mutations
Genotype: The specific allelic composition of a cell, either of the entire cell or more commonly for a certain gene or a set of genes. The genes that an organism possesses
Nuclease: One of the several classes of enzymes that degrade nucleic acid. An enzyme that can degrade DNA or RNA by breaking phosphodiester bonds.
nt: nucleotides
P53 : A 53KDa protein the product of a tumour suppressor gene.
Primer: A small oligonucleotide (anywhere from 6 to 50 nt long) used to prime DNA

synthesis. Primers are necessary for DNA sequencing and PCR

Renature: Spontaneous realignment of two single DNA strands to re-form a DNA double helix

Screening: To screen a library (see "Library") is to select and isolate individual clones out of the mixture of clones.

Wild-type: The genotype that is found in nature or in the standard laboratory stock for a given organism.

B. Matlab program for mechanism simulation.

```
N0=1;
N1=[PS complex N];
N2=[Si complex N];
N3=1;
d1(1)=15;
d2=500;
lamda=265;
theta0=6;
theta0rad=3.14/180*theta0;

for k=2:20
  d1(k)=d1(k-1)+5;
end

for j=1:20
      theta1=asin(N0/N1*sin(theta0rad));
      theta2=asin(N1/N2*sin(theta1));
      theta3=asin(N2/N3*sin(theta2));
      r01s=(N0*cos(theta0rad)-N1*cos(theta1))/(N0*cos(theta0rad)+N1*cos(theta1));
      r10s=-r01s;
      t01s=(2*N0*cos(theta0rad))/(N0*cos(theta0rad)+N1*cos(theta1));
      t10s=(1-r01s^2)/t01s;
      r01p=(N1*cos(theta0rad)-N0*cos(theta1))/(N0*cos(theta1)+
            N1*cos(theta0rad));
      r10p=-r01p;
      t01p=(2*N0*cos(theta0rad))/(N0*cos(theta1)+N1*cos(theta0rad));
      t10p=(1-r01p^2)/t01p;
      r12s=(N1*cos(theta1)-N2*cos(theta2))/(N1*cos(theta1)+N2*cos(theta2));
      r12p=(N2*cos(theta1)-N1*cos(theta2))/(N1*cos(theta2)+N2*cos(theta1));
      r23s=(N2*cos(theta2)-N3*cos(theta3))/(N2*cos(theta2)+N3*cos(theta3));
      r23p=(N3*cos(theta2)-N2*cos(theta3))/(N2*cos(theta3)+N3*cos(theta2));
      beta1(j)=2*3.14*(d1(j)/lamda)*N1*cos(theta1);
      beta2(j)=2*3.14*(d2/lamda)*N2*cos(theta2);
      Rp(j)=((r01p+r12p*exp(-2i*beta1))+(r01p*r12p+exp(-2i*beta1))*r23p*exp(-
            2i*beta2(j)))/((1+r01p*r12p*exp(-2i*beta1))+(r12p+r01p*exp(-
```

```
        2i*beta1))*r23p*exp(-2i*beta2(j)));
Rs(j)=((r01s+r12s*exp(-2i*beta1))+(r01s*r12s+exp(-2i*beta1))*r23s*exp(-
        2i*beta2(j)))/((1+r01s*r12s*exp(-2i*beta1))+(r12s+r01s*exp(-
        2i*beta1))*r23s*exp(-2i*beta2(j)));
    Rpreal(j)=abs(Rp(j));
    Rsreal(j)=abs(Rs(j));
    R(j)=(Rpreal(j)^2+Rsreal(j)^2)/2;
    R01sreal=abs(r01s);
    R01preal=abs(r01p);
    R01=(R01sreal^2+R01preal^2)/2;
    R12sreal=abs(r12s);
    R12preal=abs(r12p);
    R12=(R12sreal^2+R12preal^2)/2;

    Rs123(j)=r01s+(t01s-0.005)*r12s*t10s*exp(-2i*beta1(j))-0.005 +
             r10s*r12s*t10s*exp(-2i*beta1(j))*(t01s-0.005)*r12s*exp(-
             2i*beta1(j))-0.05;
    Rp123(j)=r01p+(t01p-0.005)*r12p*t10p*exp(-2i*beta1(j))-0.005+
              r10p*r12p*t10p*exp(-2i*beta1(j))*(t01p-0.005)*r12p*exp(-
              2i*beta1(j))-0.05;

    Rs123a(j)=r01s+t01s*(1-0.005)^2*r12s*t10s*exp(-2i*beta1(j))+
              r10s*r12s*exp(-2i*beta1(j))*t01s*(1-0.005)^2*r12s*t10s*exp(-
              2i*beta1(j));
    Rp123a(j)=r01p+t01p*(1-0.005)^2*r12p*t10p*exp(-2i*beta1(j))+
              r10p*r12p*exp(-2i*beta1(j))*t01p*(1-0.005)^2*r12p*t10p*exp(-
              2i*beta1(j));

    Rs123b(j)=r01s+t01s*r12s*t10s*exp(-2i*beta1(j))*(1-0.005^(1/2))^2+
               r10s*r12s*t10s*exp(-2i*beta1(j))*t01s*r12s*exp(-
              2i*beta1(j))*(1-0.005^(1/2))^2;
    Rp123b(j)=r01p+t01p*r12p*t10p*exp(-2i*beta1(j))*(1-0.005^(1/2))^2+
               r10p*r12p*t10p*exp(-2i*beta1(j))*t01p*r12p*exp(-
              2i*beta1(j))*(1-0.005^(1/2))^2;

    Rs123c(j)=r01s*(1+0.0099^(1/2))+t01s*r12s*t10s*exp(-2i*beta1(j))*(1-
              0.0099^(1/2))+r10s*r12s*t10s*exp(-2i*beta1(j))*t01s*r12s*exp(-
              2i*beta1(j))*(1-0.0099);
    Rp123c(j)=r01p*(1+0.0099^(1/2))+t01p*r12p*t10p*exp(-2i*beta1(j))*(1-
              0.0099^(1/2))+r10p*r12p*t10p*exp(-
              2i*beta1(j))*t01p*r12p*exp(-2i*beta1(j))*(1-0.0099);

    Rs12b(j)=r01s*(1-0.005^(1/2))+t01s*r12s*t10s*exp(-2i*beta1(j))*(1-
              0.005^(1/2))^2;
    Rp12b(j)=r01p*(1-0.005^(1/2))+t01p*r12p*t10p*exp(-2i*beta1(j))*(1-
              0.005^(1/2))^2;
```

```
        Rs123ps(j)=r01s+t01s*r12s*t10s*exp(-2i*beta1(j))+r10s*r12s*t01s*exp(-
                2i*beta1(j))*t01s*r12s*exp(-2i*beta1(j));
        Rp123ps(j)=r01p+t01p*r12p*t10p*exp(-2i*beta1(j))+r10p*r12p*t01p*exp(-
                2i*beta1(j))*t01p*r12p*exp(-2i*beta1(j));

        Rthird(j)=((abs(Rs123(j)))^2+(abs(Rp123(j)))^2)/2;

        Rthirda(j)=((abs(Rs123a(j)))^2+(abs(Rp123a(j)))^2)/2;

        Rthirdb(j)=((abs(Rs123b(j)))^2+(abs(Rp123b(j)))^2)/2;

        Rthirdc(j)=((abs(Rs123c(j)))^2+(abs(Rp123c(j)))^2)/2;

        R12b(j)=((abs(Rs12b(j)))^2+(abs(Rp12b(j)))^2)/2;

        Rthirdps(j)=((abs(Rs123ps(j)))^2+(abs(Rp123ps(j)))^2)/2;

end

plot(d1,Rthirdc)
dlmwrite('ss_L',transpose(Rthirdc)
```

Vita

Sanjun Niu was born to the Niu family in 1971 in town of ShangXian, Shannxi Province, China. He was raised there until 1989 when he went to Beijing for his college education. He graduated from Beijing University of Chemical Technology in 1993 with Bachelor of Science degree. He joined a manufacturing research institute right after graduation as a polymer processing consultant. In 1999, he came to graduate school at Virginia Tech pursuing the Ph.D. degree.