UNIVERSITY OF CALIFORNIA, SAN DIEGO

# A Microfluidic Device for DNA Sequencing by Denaturation

A Thesis submitted in partial satisfaction of the requirements for the degree of Master of Science

in

Bioengineering

by

Matthew Thomas Walsh

Committee in change:

Professor Xiaohua Huang, Chair
Professor Michael Heller
Professor Kun Zhang

2011

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

---

---

---

Chair

University of California, San Diego

2011

Table of Contents

List of Abbreviations

A: Adenine
ATP: Adenosine triphosphate
B&W: Binding and washing
C: Cytosine
$C_0$: Initial concentration of double stranded DNA
CDF: Cumulative distribution function
dI $H_2O$: Deionized water
DMF: Dimethylformamide
DNA: Deoxyribonucleic acid
EDTA: Ethylenediaminetetraacetic acid
EMCCD: Electron-multiplying charged coupled device
*f*: Fraction of double stranded DNA
G: Guanine
GAPTES: Gamma-aminiopropyltriethoxysilane
NHS: N-hydroxysuccinimide ester
PCR: Polymerase chain reaction
PEEK: Polyether ether ketone
PEG: Polyethylene glycol
PID: Proportional, integral, derivative
PSI: Pounds per square inch
RCA: Rolling circle amplification
SBD: Sequencing by denaturation
SBS: Sequencing by synthesis
SDS: Sodium dodecyl sulfate
T: Thymine
TE: Thermoelectric
TEG: Triethylene glycol
Tm: Melting temperature
$\Delta H^0$: Standard free enthalpy
$\Delta S^0$: Standard free entropy

List of Figures

List of Tables

Acknowledgements

I would like to acknowledge Professor Xiaohua Huang for his guidance and support throughout this project. He has dedicated much time to teach me many valuable skills, from machining to chemistry. Without his vision and direction, this project would not have been feasible.

I would also like to acknowledge the entire Huang lab for their willing assistance whenever I needed it. In particular, I would like to thank Ying-Ja Chen for her development of the theoretical principle and experimental implementation of Sequencing by Denaturation. She provided me with the foundation for the project that allowed me to achieve the results I did. I would also like to express my appreciation to Eric Roller for programming the instrumentation used to image the denaturation reaction.

ABSTRACT OF THE THESIS

A Microfluidic Device for DNA Sequencing by Denaturation

by

Matthew Thomas Walsh

Master of Science in Bioengineering

University of California, San Diego, 2011

Professor Xiaohua Huang, Chair

The goal of this work is to develop an experimental platform to carry out controlled DNA denaturation reactions while monitoring the fractional hybridization using fluorescently labeled oligonucleotides. Because shorter oligonucleotides denature at a lower temperature, it is possible to decode the sequence based on their melting temperatures. This technique, known as Sequencing by Denaturation, is based on the sequential denaturation of DNA fragments generated by a Sanger dideoxy sequencing reaction on surface-bound DNA templates. To perform these experiments, a device was constructed consisting of nine microfluidic channels formed by an adhesive silicone gasket sandwiched between a coverslip and a stainless steel plate

that is encased in an aluminum block. Cooling and heating capabilities are provided by thermoelectric modules, which are in contact with another aluminum block that serves as the heatsink and microscope stage insert. The device satisfies the major criteria for performing this sequencing method. The microfluidic channels are uniformly heated to ensure that all the molecules experience the same temperature. Additionally, the device undergoes a consistent thermal expansion throughout the temperature ramp, allowing the imaging plane to be tracked by the microscope. We demonstrated the ability to determine the components of a denaturation profile created by the combination of four different oligonucleotide lengths. Based on simulation data, we also determined that it might be possible to resolve a denaturation curve using as few as fifty probes. Due to the speed, simplicity, and low cost of sequencing with SBD, it has potential to enable large-scale genome re-sequencing and genotyping.

# Chapter 1: Introduction

## 1.1 Genomics

The genome contains all information necessary for an organism to create and maintain itself, as well as pass along its hereditary information to the future generations. These genetic instructions are stored in the molecular form of deoxyribonucleic acid (DNA). The form of DNA is conserved among all forms of life and consists of a deoxyribose sugar attached to a phosphate group, and contains one of four possible bases: adenine (A), cytosine (C), guanine (G), or thymine (T). The order in which the DNA bases are arranged is of extreme importance for the actions carried out by the cell. The DNA sequence maintains a strict control over the proteins the organism synthesizes through the process of transcription and translation. The sequence essentially programs all cellular activity and therefore understanding of the sequence could prove invaluable in the biological and medical fields to better understand how cells and organisms function in healthy and diseased states. Before this can be accomplished, the daunting task of determining this sequence must be performed.

DNA sequencing strategies have greatly evolved from the initial Sanger's dideoxy biochemistry to encompass a large variety of different methods and sequencing chemistries, which all fall under the category of "second-generation" sequencing technologies (Shendure *et al.*, 2008). These second-generation approaches are all based on cyclic-array sequencing, in which reagents and enzymes are introduced in to a flow cell containing a dense array of DNA, in an iterative manner,

while imaging is performed after each step. While the chemistries vary greatly between the sequencing strategies, the general approach is to first create a library by amplifying the DNA, through methods such as emulsion PCR or bridge PCR (Bentley *et al.*, 2008, Diehl *et al.*, 2006, Fedorov *et al.*, 2006). The purpose is to generate clusters of clonally amplified DNA from a single template molecule that are spatially separated from each other. Once the DNA molecules are amplified and surface-bound, different enzymatic reactions can be performed to determine the sequence.

### 1.2 Sequencing by Synthesis

Many of the major sequencing chemistries rely on constructing the templates complementary strand, a technique called Sequencing by Synthesis (SBS). There are a variety of differing methods that fall within this scheme.

#### *Polymerase Driven Synthesis*

A common method utilizes polymerases to assemble the complementary DNA strand, in a manner similar to natural DNA replication. Pyrosequencing is currently employed by Roche/454, in which single nucleotides are introduced and washed away in a cyclic manner (Margulies *et al.*, 2005). When a polymerase incorporates the specific base type, the released pyrophosphate is converted to light via ATP sulfurylase and luciferase (Ronaghi *et al.*, 1998). Other platforms, such as that from Illumina Inc., employ a technique by which all four nucleotides are introduced together, but contain reversible terminators so that only one base can be added to each template at a time. The bases also contain a fluorophore that corresponds to their base type. The remaining unbound fluorophores are washed away so the DNA can be

effectively imaged. The fluorophores and reversible terminators are then cleaved before the next round begins.

***Ligase Driven Synthesis***

The sequence can also be deciphered through the process of flowing in oligonucleotides with random base sequences for all but one or two of the nucleotides, which are in a know position and fluorescently labeled. This method is employed by the Life Technologies' SOLiD platform, in which they use octamers with two-base encodings (Valouev *et al.*, 2008). By hybridizing and ligating these oligonucleotides to the template starting at different positions on the template, each of the bases is probed. Because two bases are know for each oligonucleotide, it provides an intrinsic error-correction scheme.

These methods of SBS have been designed to be high throughput and have the potential to produce a large amount of data from a single run. One of the major limitations of these approaches is the short read length that they are able to produce. This poses great difficulties for de novo sequencing, for which a reference genome is not available. These sequencing chemistries also consume large amounts of expensive reagents, as well as take a long time to complete, due to the cyclic nature of the reactions.

**1.3 Sequencing by Denaturation**

A novel sequencing technique has been developed based on the melting curve analysis of short oligonucleotides (Chen, 2008, Chen *et al.*, 2009). The method is based on the sequential denaturation of DNA fragments that are generated by a Sanger dideoxy sequencing reaction on surface-bound DNA templates. Because shorter

oligonucleotides denature at a lower temperature, it is possible to decode the sequence based on their melting temperatures. Thermodynamic simulations have proposed that it is possible to sequence up to 20 bases by comparing the denaturation profile of oligonucleotides from the thirteenth to thirty-second base (Chen *et al.*, 2009). Initial proof-of-concept experiments were also performed in which the sequencing of six bases was demonstrated (Chen *et al.*, 2010).

A basic thermodynamic model was developed, in which the equilibrium constant was used to relate fractional hybridization to free energy. The denaturation reaction can be expressed as $T^*P \leftrightharpoons T + P$, were T represents the surface-bound template DNA strand, and P represents the fluorescent labeled, complementary probe. Therefore the equilibrium constant is given by: $K_{eq} = [T][P]/[T^*P]$. By introducing a variable *f* as the fraction of double stranded DNA to initial concentration of double stranded DNA ($C_0$), $K_{eq}$ can be expressed as $K_{eq} = ((1-f)^2 * C_0 / f)$. Relating free energy to enthalpy and entropy leads to the equation $RT*\ln((1-f)^2 * C_0 / f) = \Delta H^0 - \Delta S^0 * T$, where R is the universal gas constant, T is the temperature, $\Delta H^0$ and $\Delta S^0$ are standard free enthalpy and entropy, respectively. Enthalpy and entropy of DNA hybridization have been thoroughly studied; these parameters can be predicted for a specific DNA sequence based on nearest-neighbor pairs (SantaLucia, 1998).

Using an initial double stranded DNA concentration of 0.1 pM, the theoretical denaturation profiles of oligonucleotides from 13- to 32-mer are displayed in Figure 1.

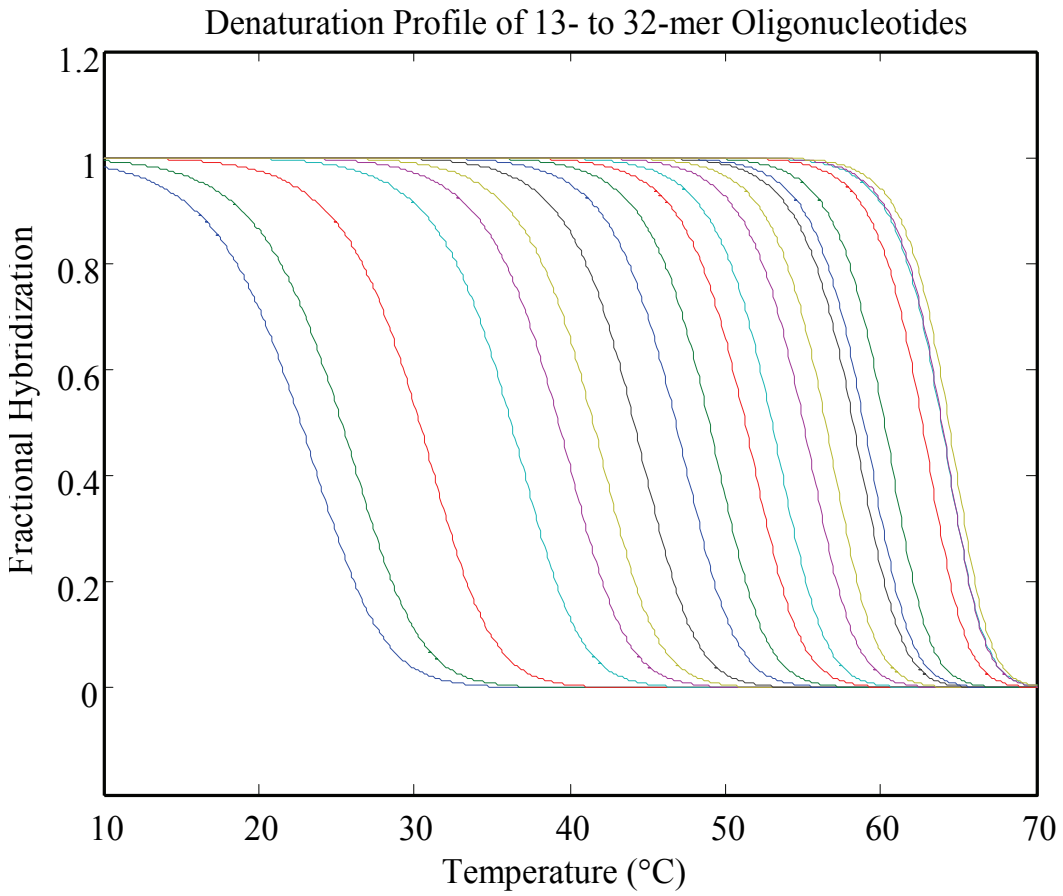



**Figure 1.1: Denaturation profile of 13- to 32-mer oligonucleotides.** The curves represent the fraction of hybridized probes for each length. Longer oligonucleotides denature at higher temperatures until they approach an upper limit at which differences in their melting temperatures cannot be resolved.

The range of oligonucleotides lengths that this method can be applied to is limited to within the thirteenth to thirty-second nucleotide because the difference in melting temperature of oligonucleotides longer than thirty-two bases become irresolvable and oligonucleotides less than thirteen bases are likely to denature at temperature lower than room temperature and would be difficult to assess experimentally.

# Chapter 2: Integrated Microfluidic Device

## 2.1 Introduction

To perform fluorescent microscopy, it is necessary to utilize a flow cell that is capable of maintaining the proper experimental conditions. In order to reduce the volume of reagents consumed during experiments, most flow cells are composed of microfluidic channels that efficiently compartmentalize the molecules of interest. To obtain the spatial-temporal resolution required by today's complex experiments, the target molecules are generally surface-bound through a variety of methods, such as biotin-stretavidin bonds or covalent methods. The fluidic device in which the molecules are contained has several important characteristics to ensure proper experimental results are obtained. One critical aspect of the flow cell is its ability to maintain a uniform temperature over the area of the channels. It is also imperative that the imaging surface remain level throughout the experiment. This is especially important if the experiment entails imaging during a temperature ramp. The device must undergo minimal or a predictable linear thermal expansion in order for the imaging system to track the desired focal plane for reliable imaging.

## 2.2 Flow Cell Design

Due to the high demand for reliable continuous image acquisition during SBD experiments in which precise temperature control is necessary, we came up with a design as illustrated in Figure 2.1. The flow cell is designed such that the fluidic channels are in direct thermal contact with the temperature controlling unit. This is accomplished by creating a fluidic chip that consists of nine microfluidic channels

formed by an adhesive gasket that bonds together a coverslip and stainless steel block. The stainless steel serves as a corrosion resistant backing that readily conducts heat to the fluid. The coverslip provides a clean surface for the molecules of interest to be bound and imaged.

Thermoelectric (TE) modules are incorporated to provide heating and cooling capabilities. A double stack of three TE modules in a row are used, providing up to 258 Watts of heating or cooling power. An aluminum block is used to encase the stainless steal chamber which serves two purposes: 1) its high heat conductivity reduces the temperature gradient in the less conductive steel block; 2) it allows the steel chamber to be easily removed, resulting in modular design. To allow for this easily modular feature, O-rings are utilized to create a compression seal with the steel block. A polyether ether ketone (PEEK) material is use to house the O-rings and as the interface to Lee Minstac tubing interconnects. A large aluminum block is fit on top of the TE modules to serve as a heatsink and stage insert. The water-cooled heatsink and stage insert help maintain a constant temperature, ensuring a predictable linear downward thermal expansion of the device from the aluminum encasement and steel blocks as the temperature is increased. Channels are machined into the heatsink block to allow for liquid cooling of the device with a chiller circulator. An exploded view of the device is depicted Figure 2.1. The dimensions of the individual components are specified in Figures 2.2 through 2.5.

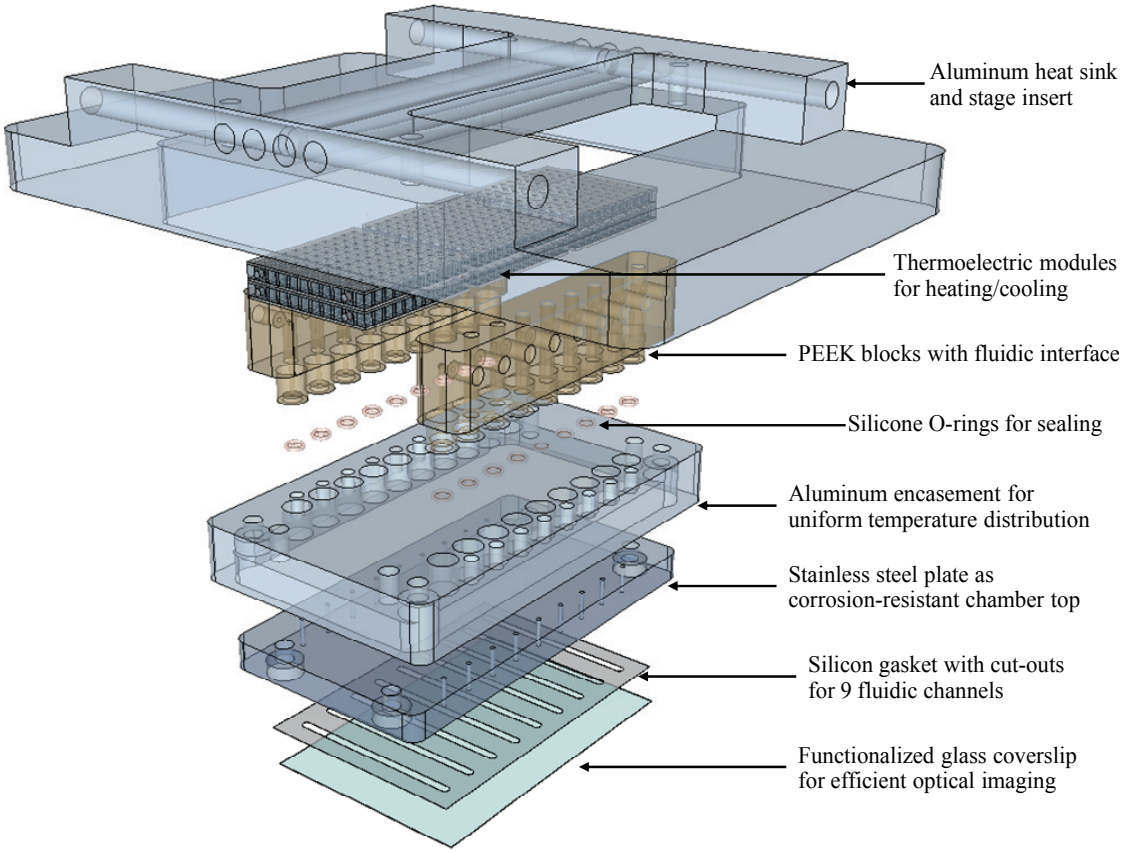
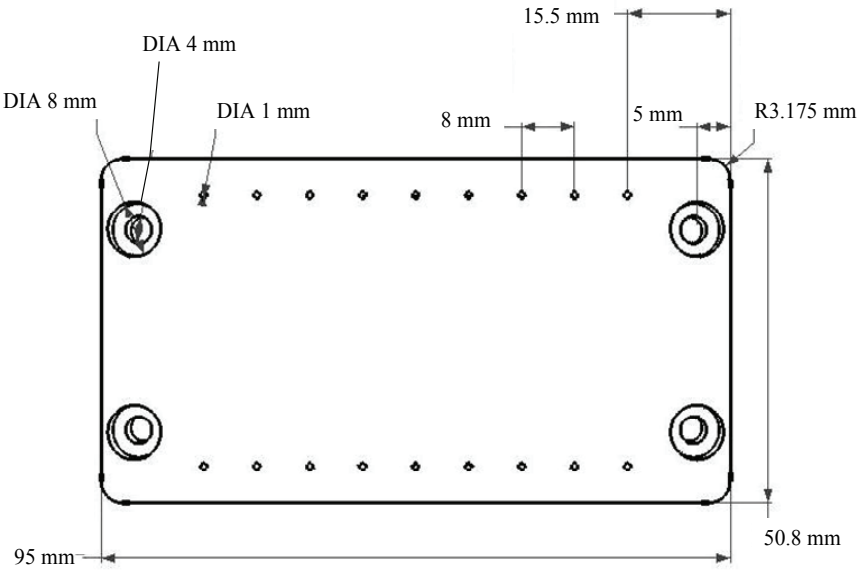



**Figure 2.1: Exploded view of the device.** The layers stack together and the top fits into the microscope stage. The bottom chamber is easily interchangeable to enable modularity.

**Stainless Steel Chamber**

**A**




**B**

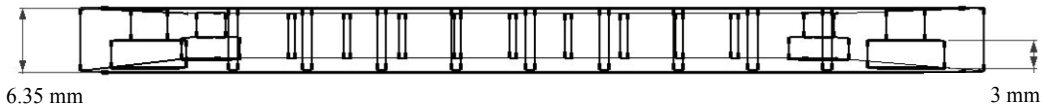



**Figure 2.2: Dimensions of stainless steel block.** Top view dimensions (A). Side view dimensions (B).

**Aluminum Encasement**

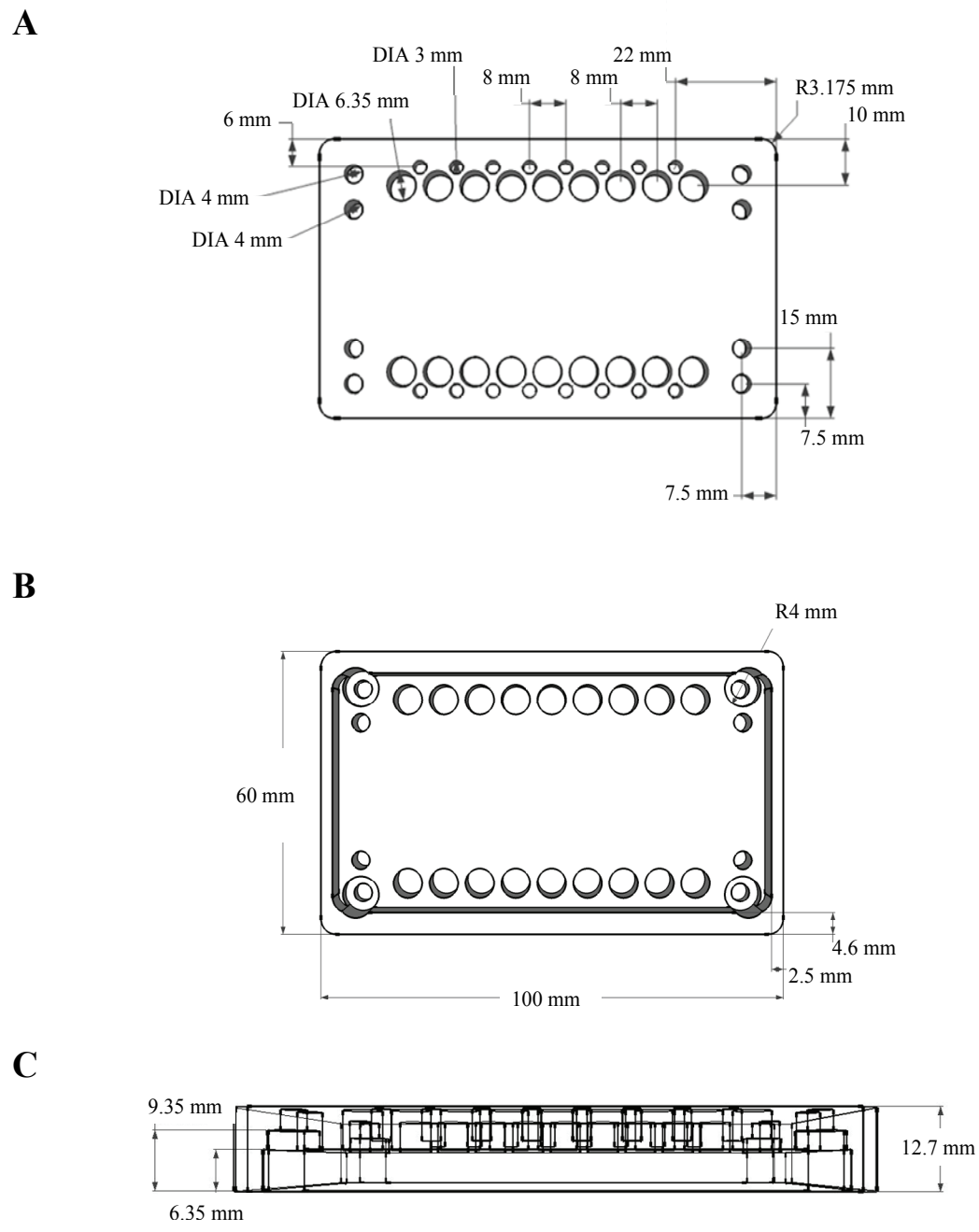



**Figure 2.3: Dimensions of aluminum encasement block.** Top view dimensions (A). Bottom view dimensions (B). Side view dimensions (C).

**PEEK Fluidic Interface**

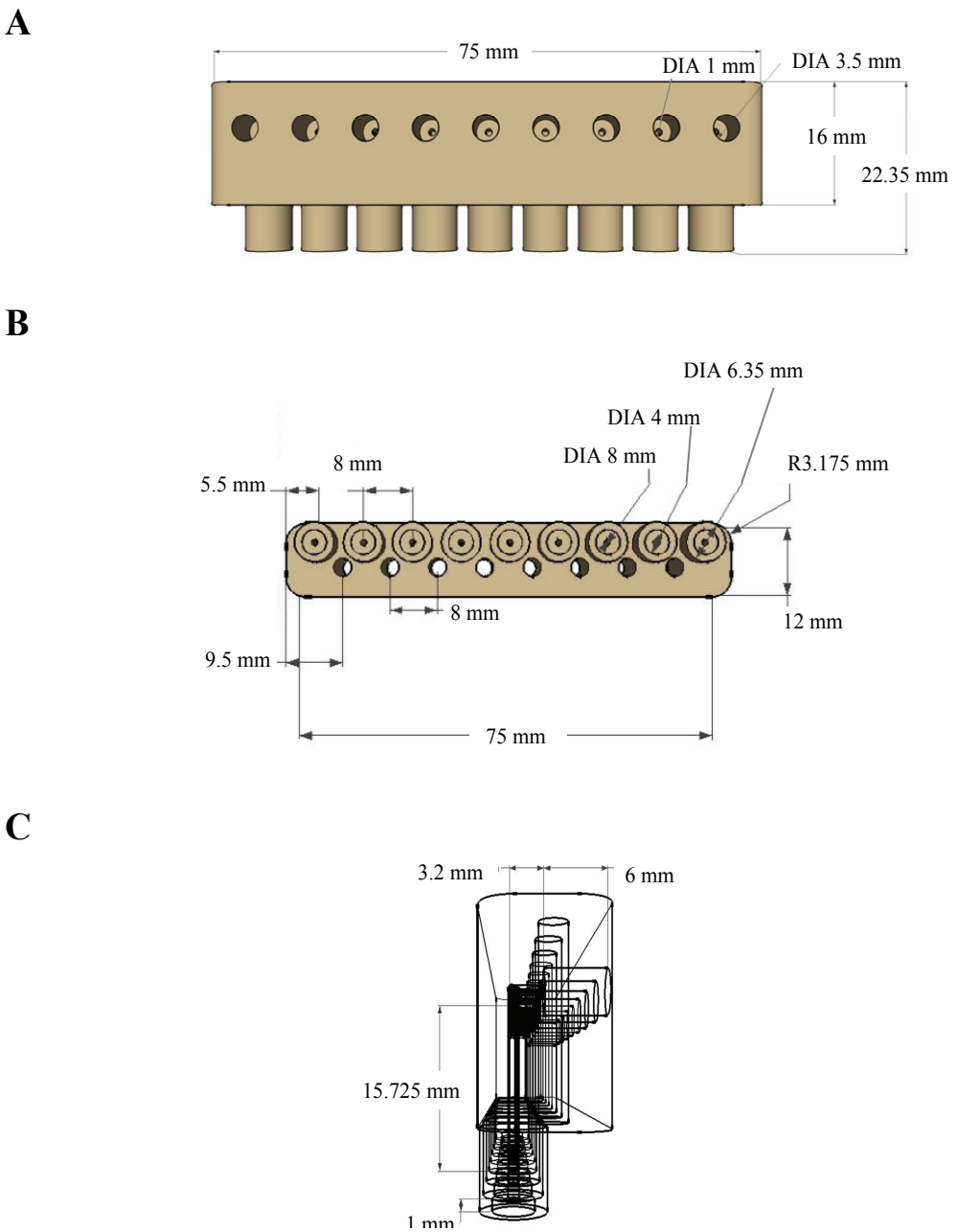



**Figure 2.4: Dimensions of PEEK fluidic interface.** Side view dimensions (A). Bottom view dimensions (B). Front view dimensions (C).

**Aluminum Heatsink/Stage Insert**

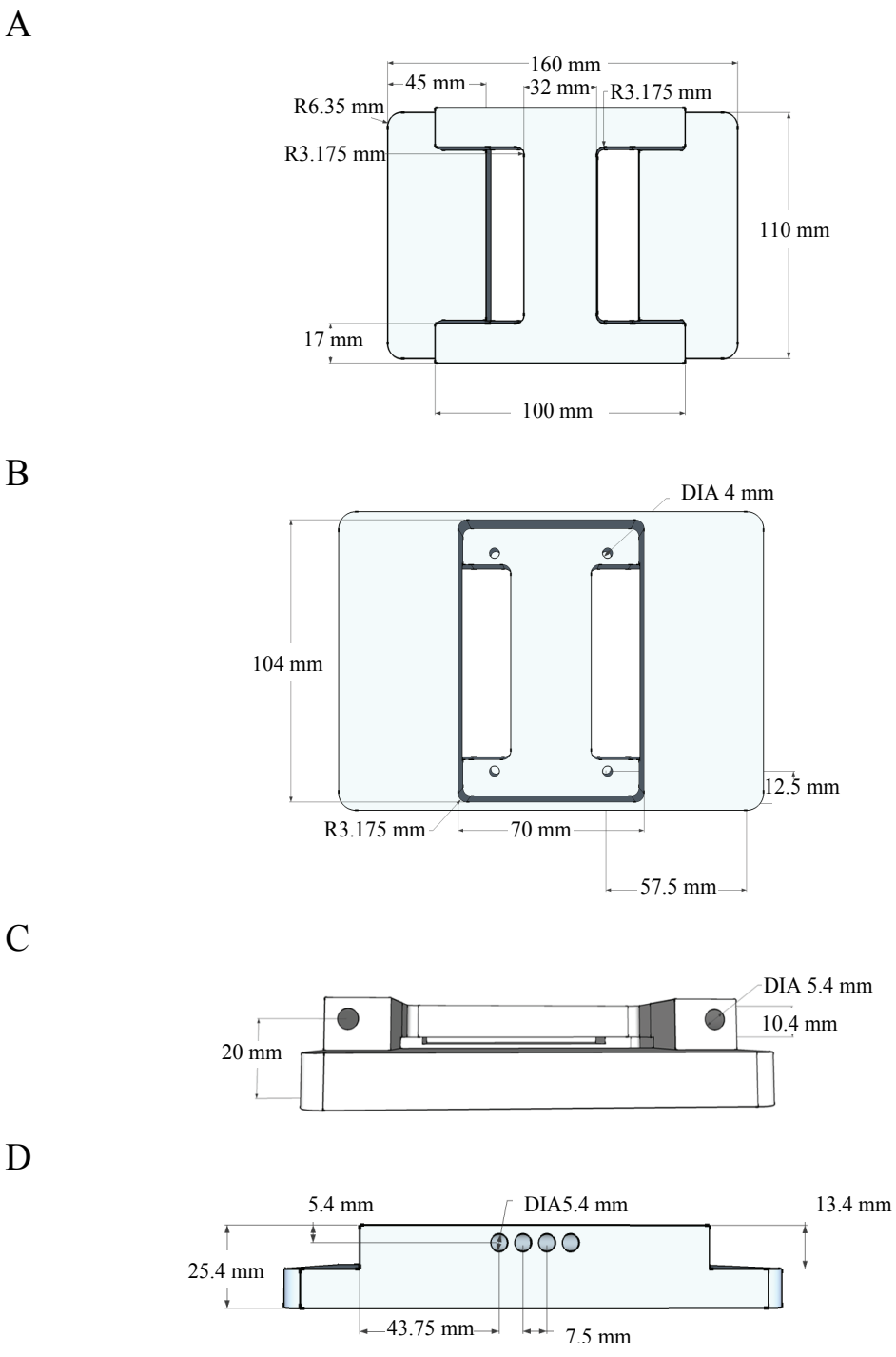



**Figure 2.5: Dimensions of aluminum heatsink/stage insert.** Top view dimensions (A). Bottom view dimensions (B). Side view dimensions (C). Front view dimensions (D).

### 2.3 Temperature Control

High temperature TE modules (VT-71-1.4-1.15, TE Technology) are chosen for their ability to operate at temperatures up to 150 °C. They are factory potted to prevent moisture from degrading their performance. Each module has the ability to pump 43 watts of heat and create a temperature different of 66 °C across it. Six of them are incorporated in the device. Two sets of three TE modules, connected in series, are stacked on top of each other and connected in parallel. A thermister (MP-2444, TE Technology) is placed near the outer surface of the stainless steel chamber to monitor the temperature and provide feedback to the system. A second thermister is set in the heatsink to observe its temperature. To control the temperature of the device, a controller was assembled for PID control of the temperature of the flow cell device. The controller consists of a temperature controller (TC-36-25-RS232, TE Technology), power supply (PS-24-20, TE Technology), and display (MP-2986, TE Technology). The temperature of the microfluidic chamber can be displayed in real time.

### 2.4 Heat Transfer Model

In order to verify the efficiency of the heat transfer within the device, a computer model was developed using the program COMSOL Multiphysics to determine the thermal gradients created in the device as it is heated. The geometry was created to match the cross section of the device with the TE module on top. The material parameters that were entered are shown in Table 2.1.

**Table 2.1: Material Thermal Properties**

| Material | Density ($kg/m^3$) | Heat Capacity (J/kg*K) | Thermal Conductivity (W/m*K)) |
|---|---|---|---|
| Aluminum | 2770 | 875 | 177 |
| Stainless Steel | 8238 | 468 | 13.4 |

The heat flux from the TE module was set to be 95.6 $kW/m^2$, which is the maximum heating capacity of the TE modules over the total contact area, and the initial temperature was set at 25°C. The following heat diffusion equation was used to define the physics of the problem in two spatial dimensions:

$$k\nabla T(x,y,t) = \rho c_p(\partial T(x,y,t)/\partial t),$$

where T is the temperature as a function of two spatial dimension, x and y, and time, t; k is the thermal conductivity; $\rho$ and $c_p$ are the density and heat capacity of the material, respectively. The heat flux from the TE modules is accounted for as a boundary condition at the interface of the aluminum encasement and TE modules.

The results from the numerical solver are depicted in Figure 2.6 as a time series from five to thirty seconds, in increments of five seconds. The solution demonstrates the effectiveness of uniformly spreading the heat by the aluminum block to the stainless steel piece. Because aluminum has such a greater thermal conductivity, it is able to more quickly distribute the heat throughout the block to ensure the even transfer of the heat to the entire stainless steel block, despite the low thermal conductance of the stainless steel.

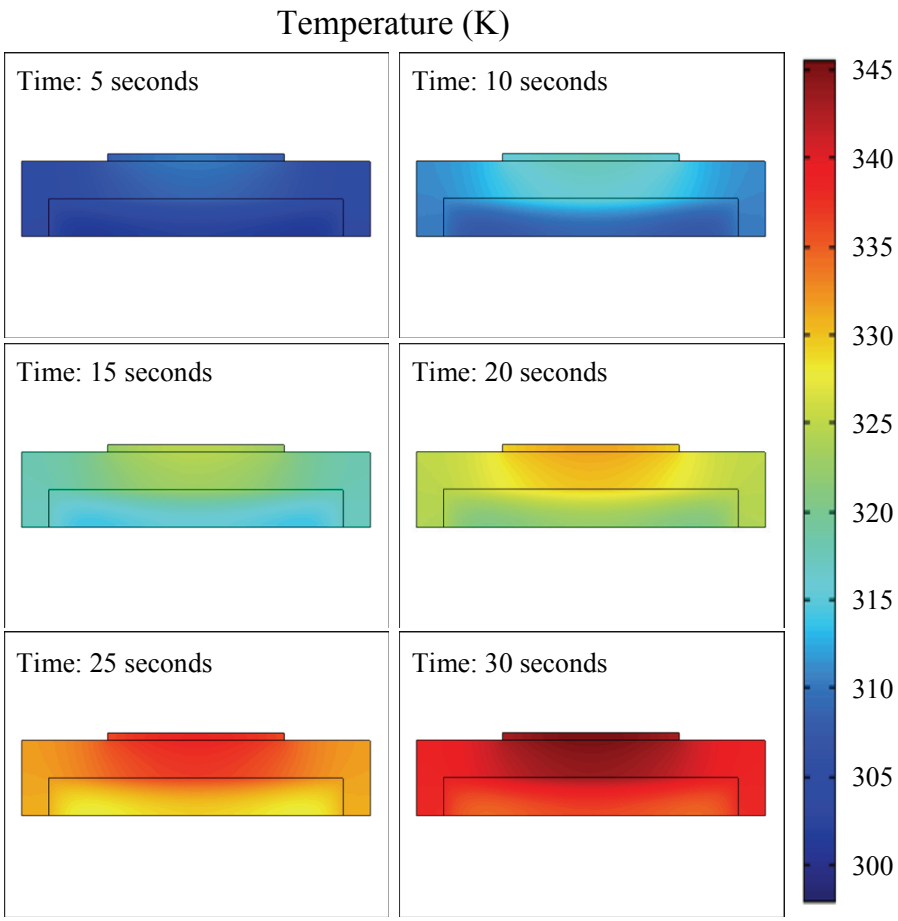



**Figure 2.6: Time series analysis from COMSOL Multiphysics.** The plots from 5 to 30 seconds demonstrate how the aluminum block helps the stainless steel heat more evenly by encasing it.

## 2.5 Device Characterization

In order for the temperature controller to regulate the output of the TE modules as desired, proper PID settings need to be established. The controller has an unconventional interpretation of the proportional gain (P) whereby if the error, calculated as the difference between the set point temperature and the current temperature, is greater than the P parameter, the power output will be at 100%. If the error is less than P, the power output will be a fraction of the total possible output

given by the ratio of the error divided by P. For this interpretation of the proportional gain, a large P will decrease the power output for a given error. The conventional manner in which proportional gain affects the output variable is by multiplying the error by P. Therefore, the larger P is, the greater the output variable is. Since the device has a large heat capacitance and there is an appreciable distance between the TE modules and the temperature sensor, there is a large time lag in the control system. This response lag can easily be seen by the large overshoot experienced by the system when using a P of 5, as shown in Figure 2.7. While the unit quickly reaches the desired temperature of 40 °C, the temperature continues the trajectory and overshoots by several degrees. It takes about one minute to reach a steady state at the set point value.

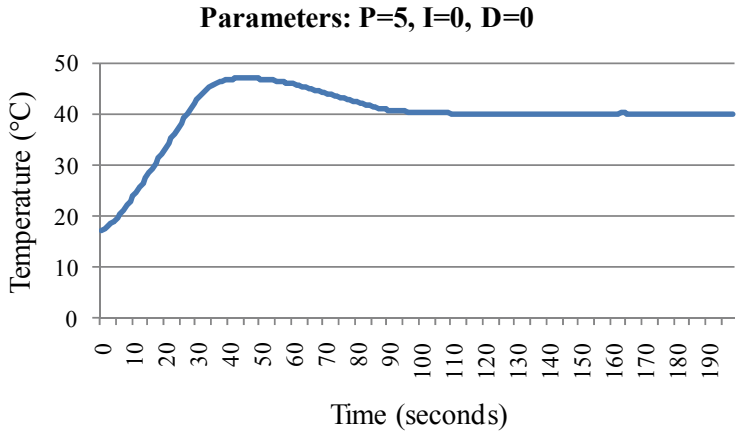



**Figure 2.7: Large overshoot in control system.** This small value for the proportional gain leads to a large overshoot due to the sluggish response by the physical system.

Using a larger P value allows the temperature to approach the desired temperature with little or negligible overshoot. As shown in the Figure 2.8, with a P of

75, it takes a little over one minute to reach 40 °C with no or very little overshoot.

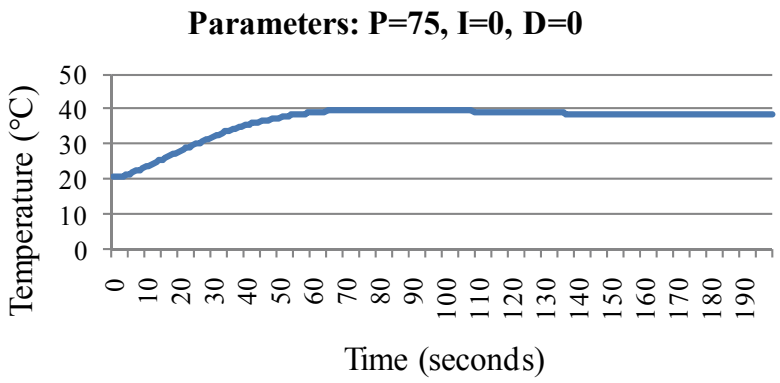



**Figure 2.8: Small offset in control system.** The large value for the proportional gain leads to a negligible overshoot but cannot correct for the steady-state offset.

However, at this large of a P value, the system is not capable of maintaining a steady-state value at the desired set point. In order to hold a steady-state value at the set point, an integral gain (I) component is necessary. The integral component takes into account the length of time for which the error has been occurring and multiplies it by the I parameter. Providing a small I parameter enables the system to correct for any offsets between the steady state value and set point, while only minimally affecting the approach to the set point, as seen in Figure 2.9.

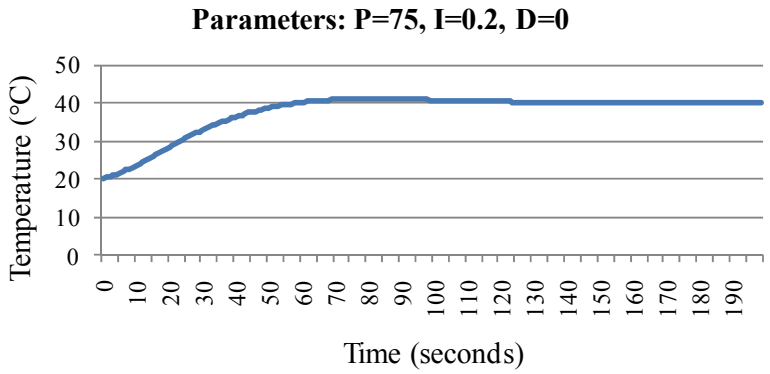



**Figure 2.9: Best controller PID parameters.** These PID values are effective at quickly reaching and holding the set point value.

The parameter values in the Figure 2.9 are chosen as the final PID parameters. With these setting, temperature of the system can be precisely controlled with minimal overshoots. A derivative gain (D) was not included because its affect on the system is complex and is not required to achieve the desired system response.

### 2.6 Focus Drift during Temperature Ramp

To fully observe the DNA denaturation, imaging is performed over a temperature ramp from 40 °C to about 75 °C. Due to thermal expansion, the imaging focal plane drifts significantly as the device is heated. A Zeiss “Definite Focus” system is utilized to maintain proper focusing during the temperature ramp. The system consists of an infrared distance measuring instrument that projects a grid of infrared light onto the desired imaging plane after an auto-focusing routine has been performed. The system can record the proper distance between the imaging plane and microscope objective. When the flow cell device expands downward as the

temperature is increased, the objective is adjusted via a motorized Z axis to maintain the proper focus.

By monitoring and recording the movement of the microscope objective, the focal plane of the glass substrate can be determined. The thermal expansion of the device can be characterized. Figure 2.10 shows a plot of the focus plane as a function of temperature. As can be observed, the surface of the glass substrate, which is the imaging focal plane, shifts downward as a linear function of temperature as the temperature of the device is increased, with a 20-micron shift over about 30°C.

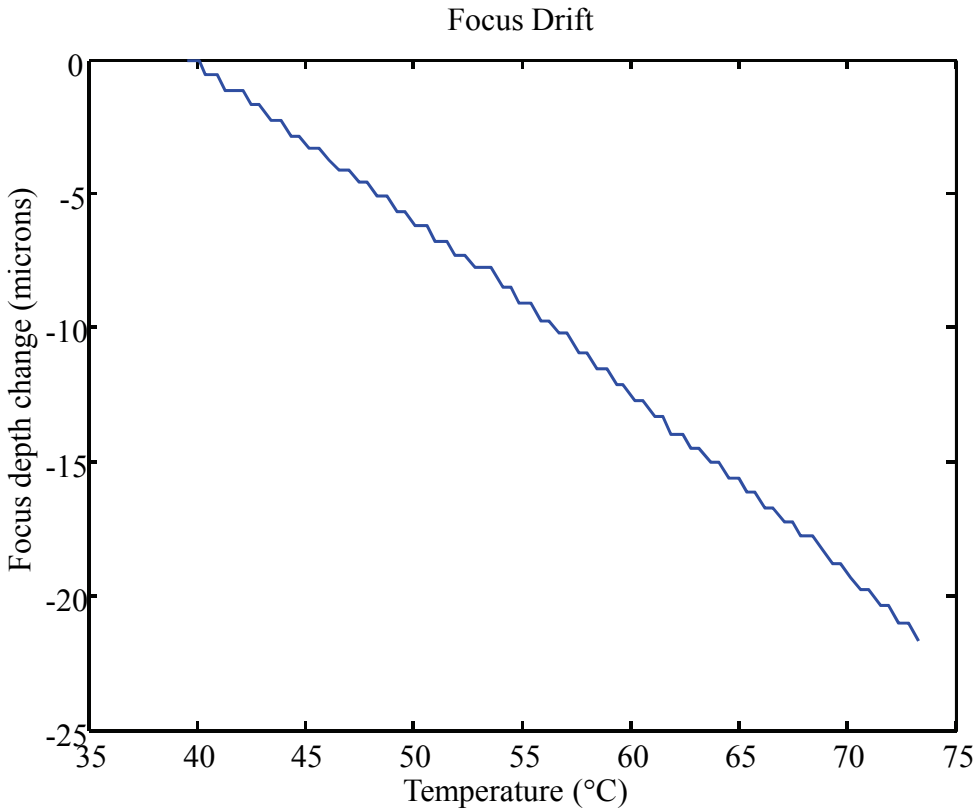



**Figure 2.10: Focus drift during temperature ramp.** As the temperature increases, the imaging plane of the device moves downward due to thermal expansion. Because the metal expands in a linear manner, the microscope in able to accurately track the change to ensure proper focus.

The step-wise curve seen in Figure 2.10 is due to limited resolution in the Z axis movement that can be achieved by the Definite Focus system. The linear thermal expansion coefficients for aluminum and stainless steel are $2.3*10^{-6}$/°C and $17.3*10^{-6}$/°C, respectively. Since both metal pieces are 6.35 mm thick, the total linear expansion can be calculated as:

$$6.35 \text{ mm} * ( 23*10^{-6} + 17.3*10^{-6} ) * 30 \text{ °C} = 7.68 \text{ microns.}$$

The observed 30 micron shift in the focal plane is much greater. The large discrepancy implies that the expansion of TE modules is much greater, accounting for most of the downward drift.

The large drift of focal plane due to heating is easily corrected for using the Definite Focus system. In order for the system to function properly, there must be an appreciable difference in refractive index of the aqueous solution in the channel and the top of the coverslip. The Definite Focus system sometimes has trouble in consistently detecting the correct focal plane with our device because the stainless steel top is highly reflective, returning a lot of the infrared light back to the sensor of the system. To overcome this issue, the stainless steel surface is made dull so that it does not reflect the light directly back to the sensor, but scatters the light randomly. To achieve the proper dull complexion, a lapping paste is used (McMaster Cat. #4428A19), with grain sizes from 30 to 60 microns. A cast iron lapping plate (McMaster Cat. #2352A11) is used for a flat and smooth surface against which the stainless steel block is ground until the steel became a uniformly dull gray. This lapping technique has proven quite successful in ensuring the proper performance of the Definite Focus system.

# Chapter 3: Measurements of DNA Denaturation Curves

Before a real sequencing experiment using Sanger's dideoxy reaction can be performed on an unknown template, initial tests must be run to test the ability of denaturation reactions to distinguish between different probe lengths. Being able to resolve different length oligonucleotides based on their differences in their denaturation profile is the basis of SBD. Once this has been established, the challenge arises of being able to identify which probes are present from a denaturation curve that has been created from the combination of multiple different oligonucleotide lengths.

## 3.1 Sample Preparation

To test the feasibility of SBD as a sequencing technology, synthetic probes with fluorophores attached were hybridized to a single population of template DNA. Each individual set of oligonucleotides was hybridized to the template separately to prevent competition between the oligonucleotides with different lengths. If the hybridization was performed with all the oligoculeotides in the mix, the hybridization of the longer oligonucleotides would be favored, resulting in uneven populations. The DNA template and probe in a one-to-two ratio are combined in a 1x Binding and Washing (B&W) buffer (5 mM Tris-HCl, 0.5 mM EDTA, 1 M NaCl, and 0.005% Triton X-100, pH7.5 ). The mix was heated to 70°C in a heat block, and then let to cool slowly to room temperature. One time excess of probe was used to ensure that every template would have a probe hybridized to it. A sixty-mer oligonucleotide was used for the template DNA. The template also contained two biotin moieties attached via a triethylene glycol (TEG) linker on the 5' end. The sequence of the template is:

5'-(Biotin-TEG)-(Biotin-TEG)-TAC AGA CTT AGT GGG GTA AAA CTA GCA TGA CTG ACT TCG TAC ATG ACT GAT GGT CGA TAC-3', where the underlined region represents that section where the probe hybridizes.

The DNA templates were bound to the Dynamicrobeads MyOne Streptavidin C1 magnetic microbeads from Invitrogen (Cat. 650) through the strong biotin-streptavidin non-covalent bond. Different mixtures of hybridized probe and template solutions were added to the microbeads. The microbead solution was diluted to $5.33*10^{-5}$ times the initial concentrations, resulting in a concentration of 3.73 to 5.33 $*10^{5}$ microbeads per milliliter. The microbeads were suspended in 1x B&W buffer. The desired combination of hybridized probe and template DNA was added to the microbead solution at a one-to-one ratio. The mixture of hybridized probe and template DNA was dependant on the desired experiment. By adding probes of different lengths but labeled with the same fluorescent dye, a much more complex denaturation curve is obtained compared to that of a single probe species. The solution was then incubated for thirty minutes in a thermal mixer at 350 rpm and 22 °C in microcentrifuge tubes. The microbeads were then washed 3 times by the addition of a wash buffer with 1X B&W buffer. The wash solution was removed by aspiration after the tube was placed against a magnet for two minutes to pull the microbeads to the sidewall of the tube, allowing for the solution to be aspirated without removing the microbeads.

Control microbeads were constructed in a similar manner. These microbeads contained single stranded DNA tethered to the microbead through two biotin groups on one end and included a fluorescent label on the other end. These microbeads acted

as a control used for the correction of photobleaching during the denaturation experiment. In each experiment, the control microbeads were mixed with the sample microbeads at a one-to-two ratio.

To bind the microbeads to the coverslip, the glass surface was functionalized with biotin molecules via a polyethylene glycol (PEG) linker. The general procedure as described by Barbee *et al.* was followed (Barbee *et al.*, 2008). Prior to the attachment of the biotin-PEG to the surface, the borosilicate coverslips were first cleaned in a two step process. First, the coverslips were placed in a 1:1:5 ratio of 30% ammonium hydroxide, 30% hydrogen peroxide, and water for one hour at 85 °C. The reaction mix was contained in a 2 L Pyrex container and the coverslips were held in a custom-made Teflon slide rack. This step is referred to as RCA-1, and is used to clean the glass of organic residues. The coverslips were then rinsed with deionized water (dI $H_2O$) before being placed in a 3:1 ratio of 98% sulfuric acid and 30% hydrogen peroxide for one hour at 85°C. This solution, known as Piranha solution, is extremely dangerous and should be handled with great care in a chemical hood. This step is referred to as RCA-2, and is used to remove all inorganic materials from the glass surface and leave the silicon with a pristine silicon dioxide surface. The coverslips were then rinsed thoroughly with dI $H_2O$ and placed in a different rack before rinsing with methanol. A new rack was used so that the rack that holds the coverslips during the cleaning reactions does not come in contact with organic solvents which could react violently then next time the rack is used to clean coverslips. This prepares the surface of the coverslip for covalent bond formation with the silicon group of gamma-aminiopropyltriethoxysilane (GAPTES). A solution of 2% GAPTES, 5% dI $H_2O$, and

93% ethanol was prepared and allowed undergo a hydrolysis/condensation reaction for 5 minutes before the rack of coverslips was placed in the solution. After 5 minutes the rack was removed and the coverslips were rinsed three times with acetone and cured in an oven at 110°C for 10 minutes. This provides a free amine group attached to the surface, capable of reacting with the N-hydroxysuccinimide ester (NHS) group that is attached to a PEG-biotin, with a total molecular weight of 5000 daltons. The NHS group forms a stable amide bond with the amine of the silane group to create a long and flexible covalent link attaching the biotin group to the coverslip surface. A solution of 1 mM NHS-PEG-biotin in DMF was prepared and 300 microliters was pipetted onto the surface of an individual coverslip. A second coverslip was then placed on first coverslip, using a number 1 coverslip in between as a spacer, so that the solution was sandwiched and evenly distributed between the coverslips. This reaction took place in a humidity controlled chamber so that the solution does not dry. The coverslips were then individually rinsed with acetone, a solution of 1% ammonium hydroxide with 0.1% SDS, and dI $H_2O$ before being cured 110°C for 8 minutes. The coverslips are stored under vacuum until they are needed.

The functionalized coverslips were then used to form microfluidic channels with the stainless steel block. A Roland CAMM-1 Servo GX-24 desktop cutter was used to cut channels in double-side adhesive silicone gasket (Scapa, 702). The tape was about 100 microns thick and was first set on the steel block before the coverslip was placed on to create the chambers. Pressure was then applied to the coverslip to produce a good seal between the steel and the glass coverslip. The total channel volume is approximately 15 microliters.

Before the microbeads were introduced to the channels, they were vortexed several times, spun down briefly, and sonicated for four minutes to break up any clumps of microbeads. The channel was first filled with a solution containing a small amount of Triton X-100 surfactant to wet the chamber, which helped reduce the non-specific binding of microbeads to the steel. The microbead solution was then flown into the channel and a magnet was quickly dragged across the coverslip to pull the microbeads down to the biotin surface. The channel was then left to incubate, with the coverslip facing downward, for 30 minutes to allow the biotin molecules on the PEG linkers to attach to the free streptavidin on the microbeads. After the incubation period, the magnet was again pulled across the coverslip to further assist the microbeads in binding to the surface. A depiction of the probe and control microbeads is shown in Figure 3.1.

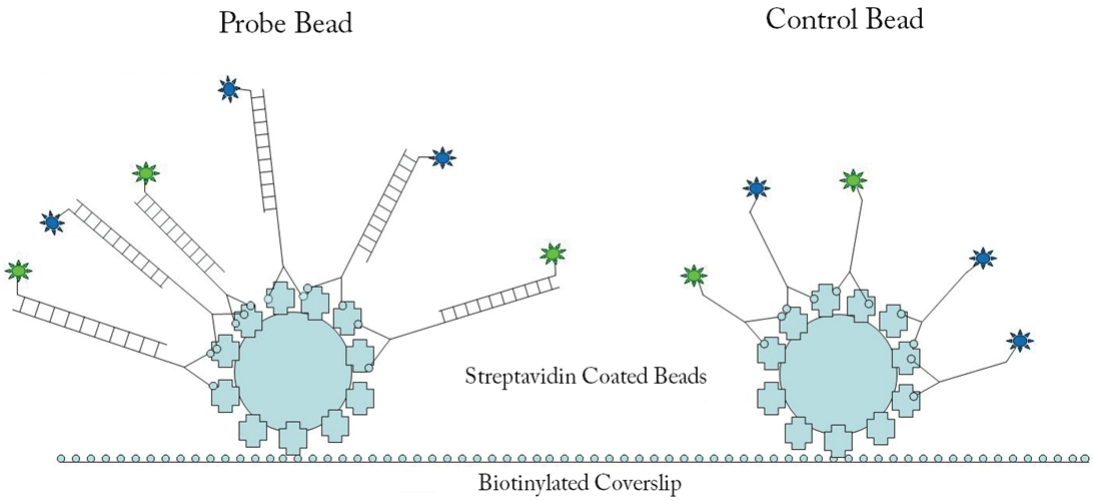



**Figure 3.1: Surface bound probe and control microbeads.** Streptavidin-coated magnetic microbeads were used to bind the fluorescently labeled DNA and hold them at the surface for efficient imaging.

**Table 3.1: List of Probe Sequences**

| Oligonucleotide | Sequence |
|---|---|
| 20mer Probe | 5'-Alexa488-CCATCAGTCATGTACGAAGT-3' |
| 21mer Probe | 5'-Alexa546-CCATCAGTCATGTACGAAGTC-3' |
| 24mer Probe | 5'-Alexa488-CCATCAGTCATGTACGAAGTCAGT-3' |
| 25mer Probe | 5'-Alexa546-CCATCAGTCATGTACGAAGTCAGTC-3' |
| 27mer Probe | 5'-Alexa488-CCATCAGTCATGTACGAAGTCAGTCAT-3' |
| 29mer Probe | 5'-Alexa546-CCATCAGTCATGTACGAAGTCAGTCATGC-3' |
| 30mer Probe | 5'-Alexa488-CCATCAGTCATGTACGAAGTCAGTCATGCT-3' |
| Alexa488 Control | 5'-Biotin-TEG-CCAT-(Biotin-TEG)-CAGTCATGTACGAAGTCAGT-Alexa488-3' |
| Alexa546 Control | 5'-Biotin-TEG-CCAT-(Biotin-TEG)-CAGTCATGTACGAAGTC-Alexa546-3' |

### 3.2 SBD Experimental Procedure

Once the sample had been prepared, the steel chamber was loaded into the aluminum encasement of the device and secured. The device was then mounted onto the microscope stage and screwed down. The wires to TE modules and thermistors on the device were connected to the controller. A Julabo F25-HE water chiller circulator was also connected to the heat sink to maintain the temperature at 20°C. Using a pressurized flow system, a steady flow of wash solution through the channel was maintained during imaging to remove the denatured probes (Chen *et al.*, 2010). The wash solution was a 33 mM Phosphate buffer which consists of 16.7 mM potassium phosphate monobasic, 16.3 mM potassium phosphate dibasic, and 0.05% Triton X-100, at pH 7.2. A pressure of 0.8 pounds per square inch (PSI) was applied to the solution to create a flow rate of 1 mL/min.

The automated imaging system consisted of a Ziess Axio Observer.Z1 microscope with a motorized stage from Ludl Electronics. A Plan-Apochromat 20x objective with a numerical aperture of 0.8 and working distance of 0.610 millimeters

was used for imaging. An iXon+ 885 EMCCD camera from Andor Technology PLC was used to capture the images. A Lambda DG-5 from Sutter Instruments was used as the light source for excitation. Autofocusing was performed with Zeiss Definite Focus system. FITC-3540B and Cy3-4040B cubes (Semrock Inc.) were used to select the proper excitation wavelength for the Alexa Fluor 488 and Alexa Fluor 546 dyes. The excitation and emission spectra of for both Alexa Fluor dyes are displayed in Figure 3.2.

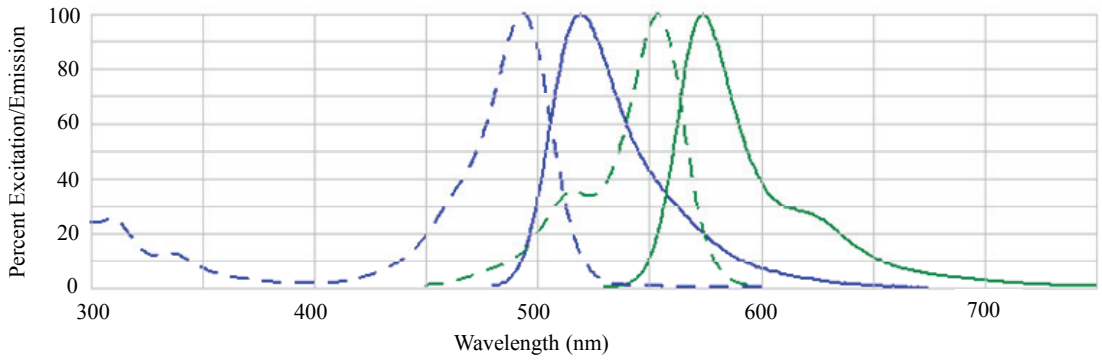



**Figure 3.2: Excitation and emission spectra for Alexa Fluor 488 and 546 dyes.** The green and blue curves correspond to Alexa Fluor 488 and 546, respectively. The dashed lines represent the excitation spectra, and the solid lines depict the emission spectra. These dyes have sufficient spectral separation so that the dyes can be individually excited and measured for emission using epifluorescent microscopy.

A custom software program was used for automated data acquisition. The software package was developed by Eric roller and Dr. Ying-Ja Chen (Chen *et al.*, 2010). The program was customized for the experiment. The optimized parameters for PID temperature control were used to gradually increase the channel temperature. The imaging parameters were determined by choosing an exposure time and gain so that

signal intensity was as large as possible without saturating the pixels of the camera. Once these parameters were established, an autofocusing routine was performed in which images were taken at discrete intervals over a small range. The images were analyzed for proper focus based on the signal intensity and the focal depth was determined. The definite focus system was then initialized to store the proper distance between the objective and coverslip. Lastly, the desired imaging spot was selected based on microbead density.

The imaging routine was then run to acquire a set of pictures of the microbeads every 15 seconds. If multiple fluorescent species were present in the sample, the motorized turret in the microscope was used to select the proper filter cube to image the different fluorophores in series. Before the acquisition of the fluorescence image, a function call to the Definite Focus system was performed to update the focal plane to ensure the microbeads were in focus. The temperature of the chamber was recorded when an image was captured. The entire experimental process lasted for 20 minutes. The temperature was linearly increased from 40°C to about 75°C to capture the full denaturation curve of the DNA probes.

The flow cell and imaging system were used to measure the denaturation profiles of seven different probe lengths. Oligonucleotides of lengths 21-, 25-, and 29-mer were labeled with an Alexa Fluor 546 dye. The 20-, 24-, 27-, and 30-mer probes were labeled with Alexa Fluor 488 (Refer to Table 3.1).

### 3.3 Data Processing

The SBD data were recorded as a series of files with the signal intensity of each pixel of the camera in each image. To extract useful information from this, the data set was processed by an ImageJ plugin written in Java. First, the background signal was removed using a rolling ball algorithm with a radius of 20 pixels. The first image was then thresholded according to the Otsu algorithm to detect the position of each microbead. Microbeads with a diameter from 5 to 25 pixels were selected and analyzed for the mean intensity within a 25 pixel diameter. This analysis was performed for each time point. The program was capable of tracking the movement of each microbead within the field of view, as it might have drifted during the experiment.

Once the images were processed, the data were saved as a matrix with each row corresponding to a specific microbead and each column corresponding to the data point. The temperature log file was used to correlate the data point to the temperature at which the image was taken. Both of these files were loaded into MATLAB and used to plot the fluorescent signal intensity change as a function of temperature. In the experiment to examine the denaturation profile containing 21-, 25-, and 29-mer probes, there were five different microbead populations present. There were four probe microbeads, one for each of the three individual probes separately, and one for the combination of all three probes. The other microbead type was the control microbead, with contains the bound fluorophore. Figure 3.3 shows the raw data from a single experiment, with a large amount of data obtained as the fluorescence intensity as a function of time. The data was immediately subjected to a filter in which the

numerical derivative of each time point was calculated and any microbead that has a derivative larger than a set value at any data point is removed.

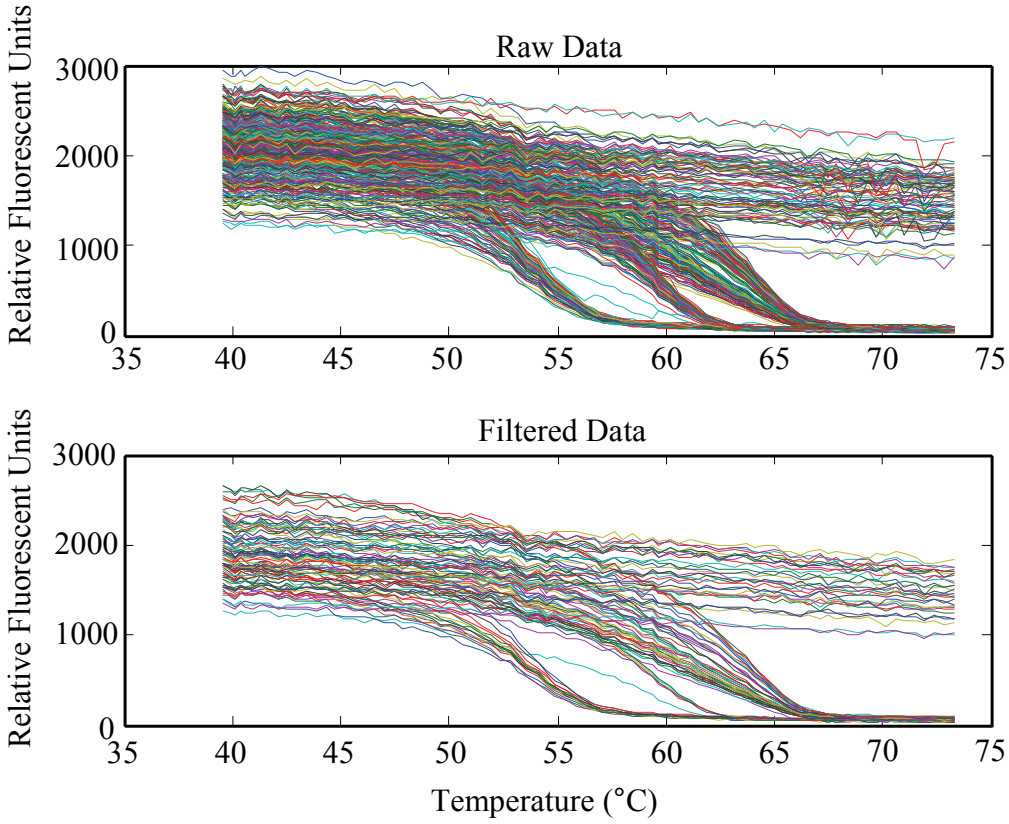



**Figure 3.3: Change in fluorescent intensities during temperature ramp.** The raw data is filtered to remove any beads with an abrupt change in intensity.

As can be seen, it is easy to distinguish between the probe and control microbeads based on the fluorescence intensity at the end of the experiment. The two groups were separated for subsequent analysis. The separated datasets are displayed in Figure 3.4.

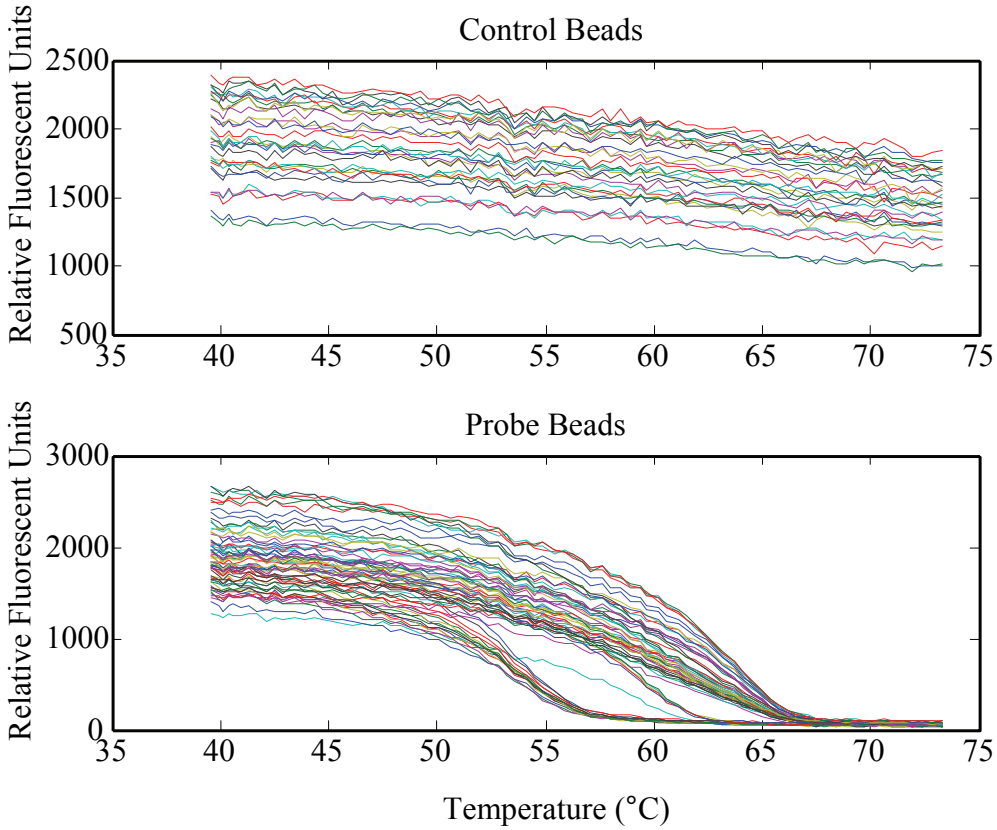



**Figure 3.4: Separation of control and probe beads.** These two sets of beads are easy to distinguish because of their greatly different behavior during the experiment.

The control microbead data were averaged and fit with a cubic regression line using a third order least squares algorithm. The effect of photobleaching was quantified by dividing the signal at each data point with the initial signal. These values were saved as a vector which was used to correct for the photobleaching effect on the probe signal decrease.

The probe microbeads intensity values were processed by first subtracting every data point by the intensity value of the last data point to reset the baseline. The curves were then normalized by dividing every data point with the intensity value of

the initial data point. Finally, the dataset was multiplied by the photobleaching correction vector. Because the data span the range of zero to one, it can be correlated to the fractional hybridization, as depicted in Figure 3.5.

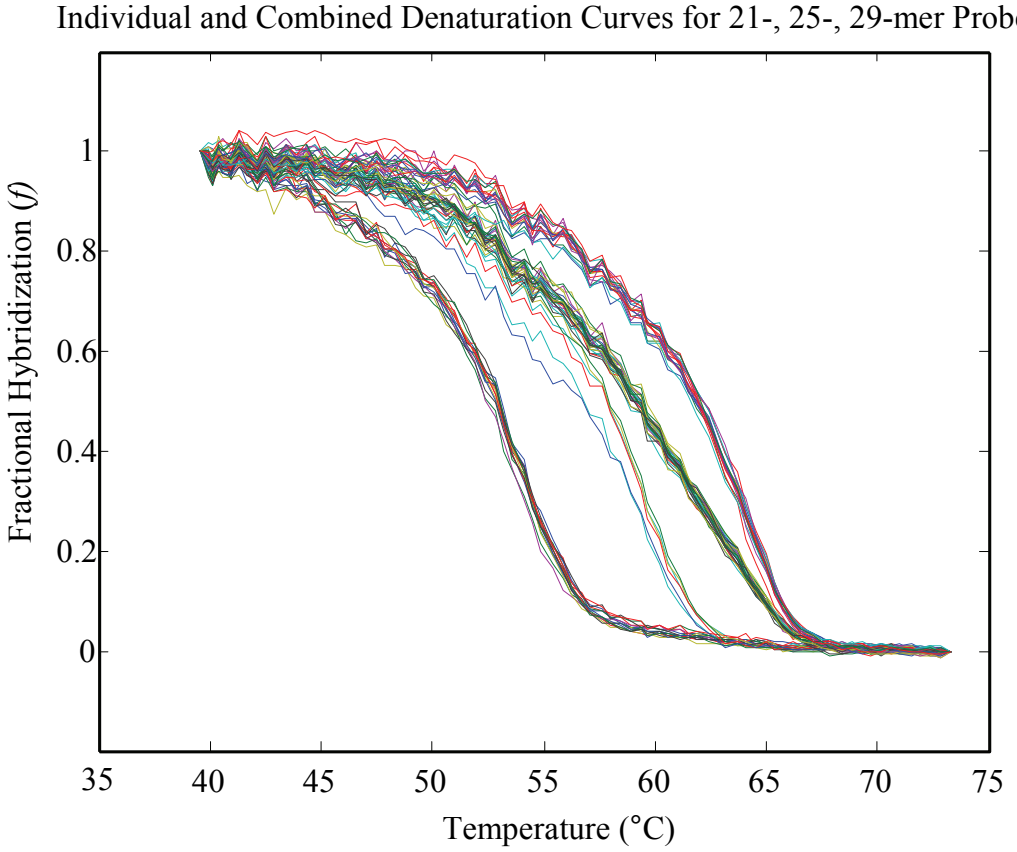



**Figure 3.5: Fractional hybridization of different bead types.** By correcting for photobleaching and normalizing by the initial intensity, the fractional hybridization is determined.

As shown in Figure 3.5, denaturation curves of the four microbead types are the clearly distinguishable. A common method of analyzing denaturation curves is to plot the negative first derivative of the fractional hybridization. Because the derivative will be taken numerically, it was necessary to smooth the curves first. A MATLAB

function was used to smooth the datasets. Smoothing was performed by local regression with weighted linear least squares and a second degree polynomial model with a span of 0.3 percent of the total data points. Figure 3.6 shows the smoothed curves of the four the microbead different types.

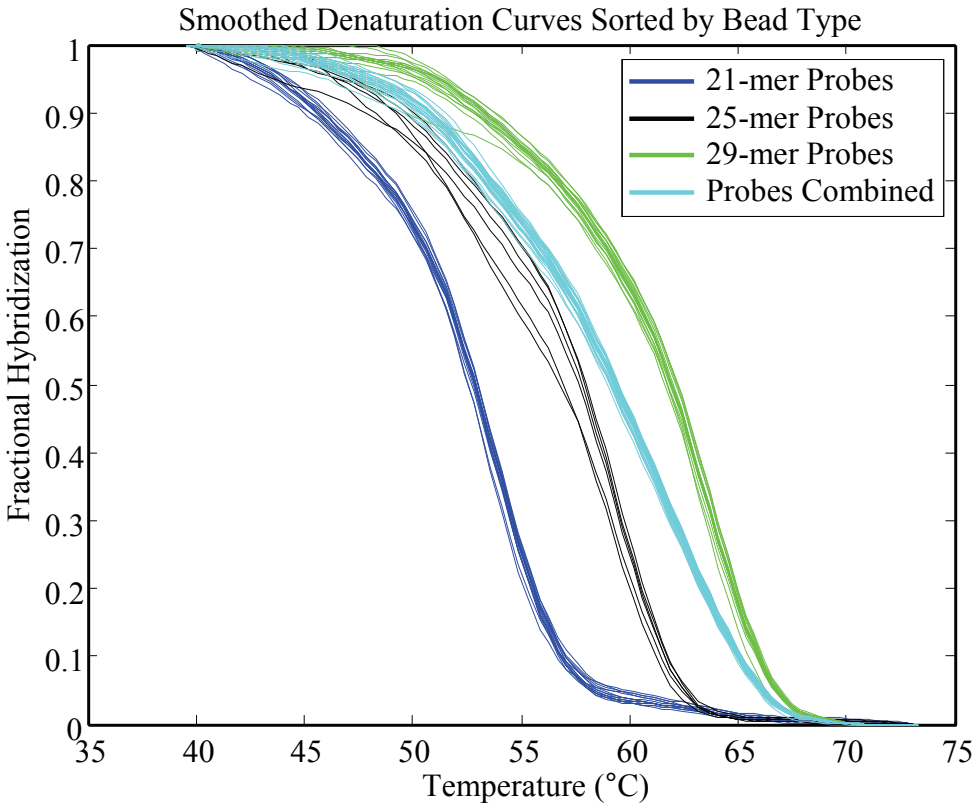



**Figure 3.6: Smoothed fractional hybridization.** A function in MATLAB was utilized to create smoothed datasets of the fractional hybridization so that the numerical derivative could be taken. The datasets are also separated into their respective bead types.

Taking the negative first derivative of the curves results in simple Gaussian-like curves for the microbeads with only one probe, while the microbeads with a combination of multiple probes will have a much more complex curve, which is the

superposition of the individual microbead's curves. The negative first derivative of the denaturation curves are shown in Figure 3.7.

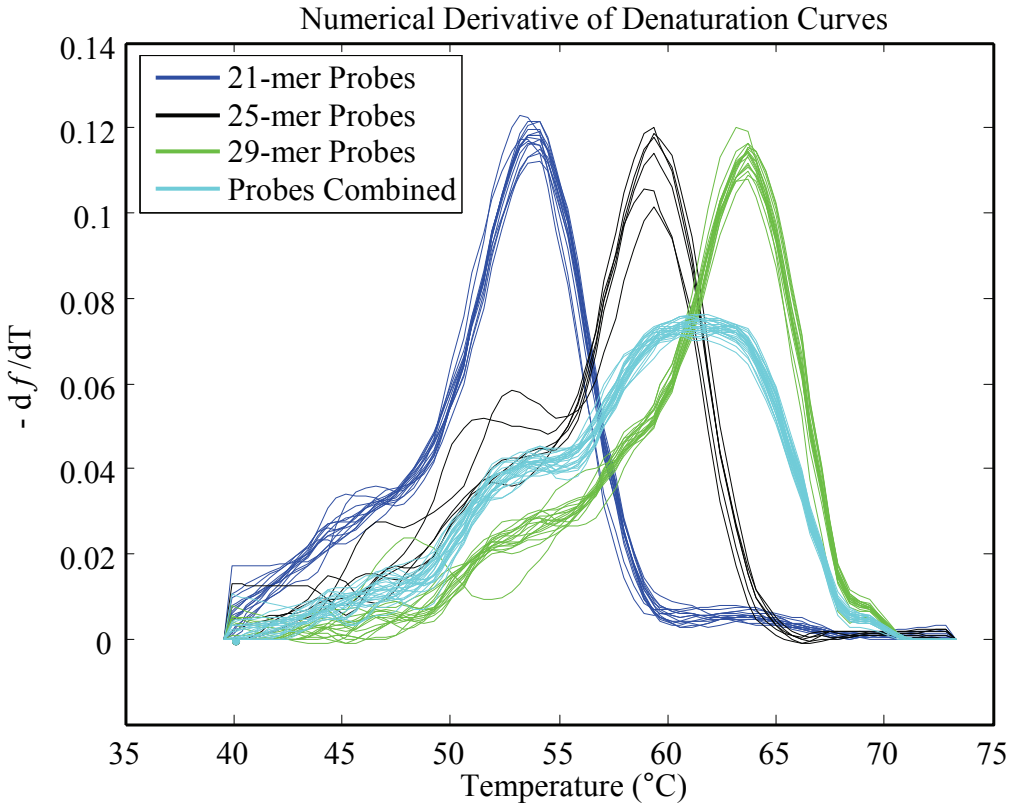



**Figure 3.7: Numerical derivative of 21-, 25-, 29-mer probes' denaturation curves.** These plots depict how the individual probe's Gaussian-like curves add to create the combined curve.

The denaturation profile with all four probes labeled combined was determined in the same manner as just described. The results are displayed in Figure 3.8. The curves for the four different individual and the combined probes are shown.

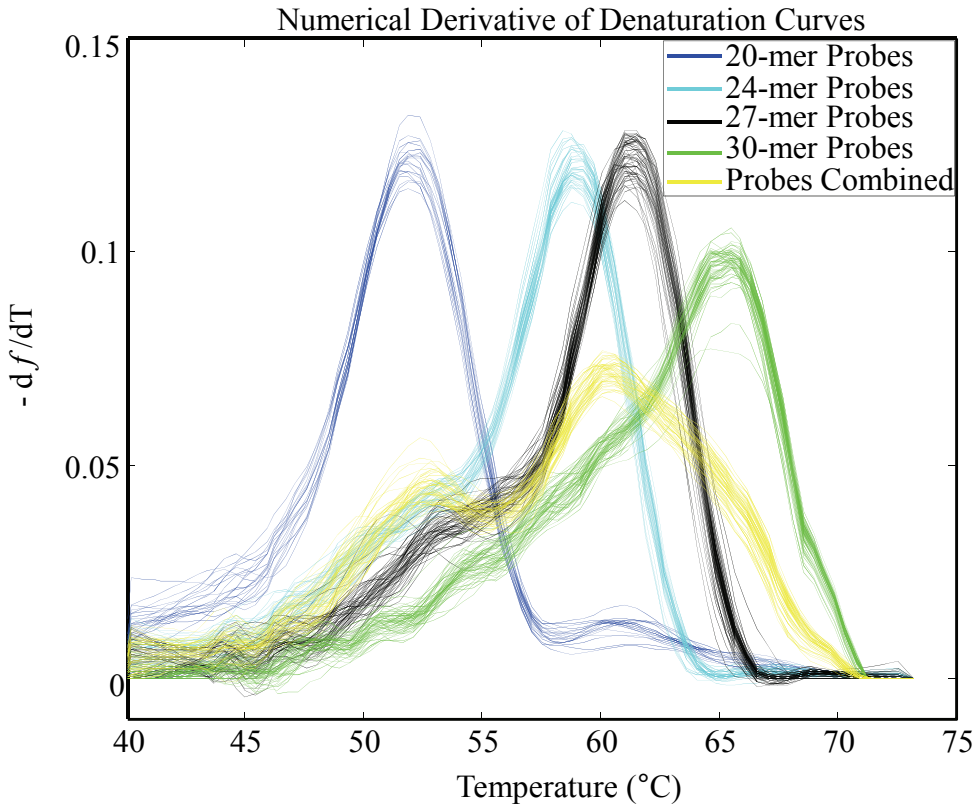



**Figure 3.8: Numerical derivative of 20-, 24-, 27-, 30-mer probes' denaturation curves.** This plot is similar to Figure 3.7, but it contains four different probe lengths which gives the curve of the probes combined a more complex shape.

## 3.4 Data Analysis

### *Modeling of SBD Data with a Sum of Gaussians*

To decode the base sequence from the processed data is much more challenging. One of the most basic approaches is to fit the derivative of the denaturation curve from the microbead containing multiple probes with a sum of Gaussian curves. For example, the denaturation curve of three combined probes can be fit with the following equation:

$$f(x) = a1 \times e^{-\left(\frac{x-b1}{c1}\right)^2} + a2 \times e^{-\left(\frac{x-b2}{c2}\right)^2} + a3 \times e^{-\left(\frac{x-b3}{c3}\right)^2},$$

which contains nine parameters. The a1- a3 parameters scale the height of each curve, thereby weighting the contribution each one has on the summated Gaussian. The b1-b3 parameters shift the positioning of the individual Gaussian along the x-axis. The c1-c3 parameters define the width of the Gaussian curve and are constrained between 3 and 4. The curve fitting function in MATLAB was used to compute the sum of Gaussian curve, and allowed the user to prescribe an initial value, as well as the range of possible values. In order for the fitting function to properly fit the curve, it was necessary to assign the 'b' parameters values close to the observed melting temperature. This severely hinders the use of this approach for deciphering an unknown sequence because the 'b' parameters are what need to be determined in the first place. However, this method does work well provided that the initial guess for the 'b' parameters are close to the observed values. As illustrated in the Figure 3.9, the experimental curve (blue dashes) was very closely fit by the sum of Gaussians curve (black line). The three individual component curves, which add together to create the curve fit, closely resemble the shape of the corresponding curves from the microbeads containing individual probes.

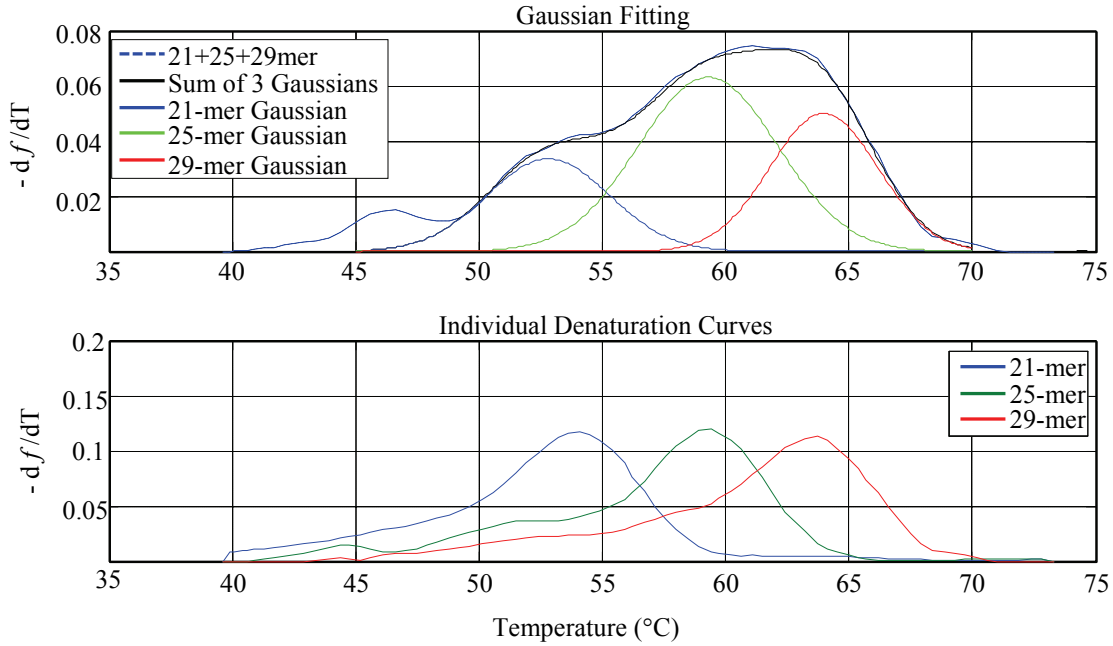



**Figure 3.9: Sum of 3 Gaussians for Alexa Fluor 546.** A curve fitting function in MATLAB is capable of accurately decomposing the total curve to individual curves, which closely match the observed individual curves, if tight constraints are applied. It is not feasible for decomposing an unknown curve.

The data from the four probes with the Alexa Fluor 488 was decomposed into a sum of four Gaussians in a similar manner, as shown in Figure 3.10.

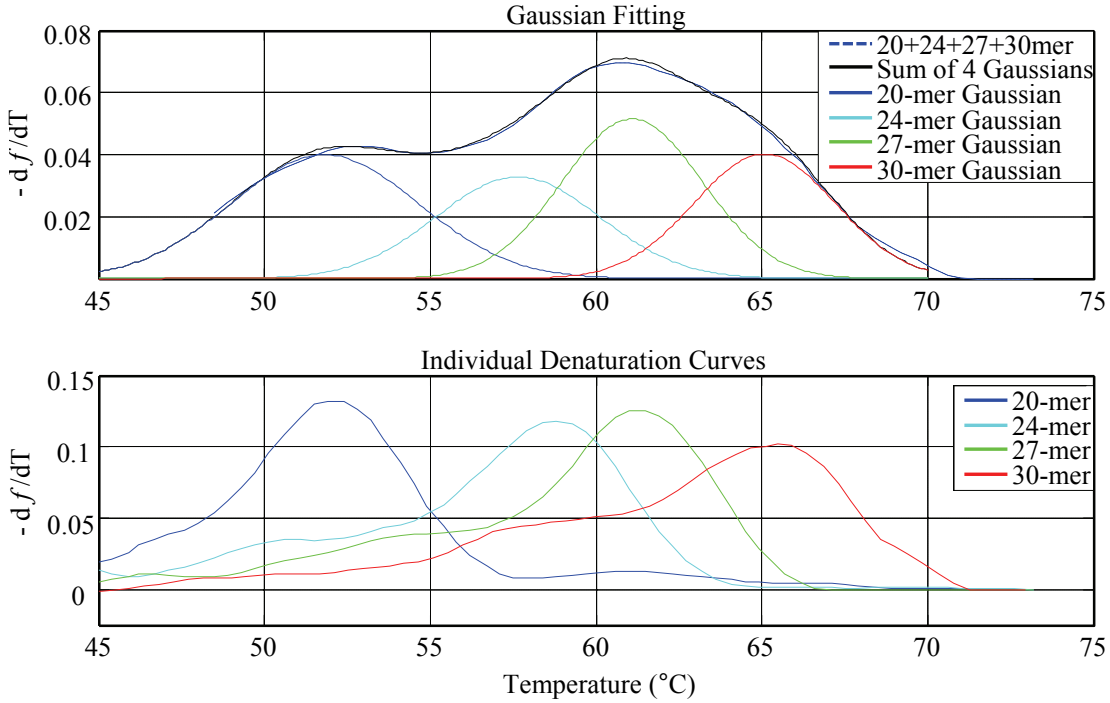



**Figure 3.10: Sum of 4 Gaussians for Alexa Fluor 488.** The curve fitting program is able to decompose the total curve into the individual components equally well for these four probes, but also has to have tight constraints applied.

The Gaussian curve fitting method proved to be limited in its ability to precisely identify the components that make up the curve. The individual denaturation curve of each probe has a distinct melting temperature and width of the transition from fully hybridized to fully denatured. This process is assumed to be governed by the thermodynamic equilibrium of the denaturation process. The assumption is a good approximation at best because in SBD the denaturation process may not reach equilibrium at each temperature. Therefore, the denaturation curve cannot be modeled precisely with a Gaussian curve.

***Curve fitting***

To compare the experimental data of the denaturation curves to the theoretical equation for fractional hybridization (*f*), a nonlinear least squares method was utilized. The theoretical equation for fractional hybridization as a function of temperature is given by the equation:

$$\mathbf{f(T) = 1 + \frac{1}{2C_0 \exp\left(-\frac{\Delta H^0 - T\Delta S^0}{RT}\right)} - \sqrt{\frac{4C_0 \exp\left(-\frac{\Delta H^0 - T\Delta S^0}{RT}\right) + 1}{4C_0^2 \exp\left(-\frac{2(\Delta H^0 - T\Delta S^0)}{RT}\right)}}}.$$

This method altered the parameters for initial concentration, $C_0$, and a scaling parameter to adjust the amplitude and position. An initial guess was applied to the system. A starting concentration of 0.1 picomolar and an amplitude of 1 were chosen. From the parameters, the fractional hybridization theoretical curve was calculated and the difference between the experimental data and theoretical curve was taken for each temperature data point. These values represent the error in the guess and were saved as an n X 1 vector, where n is equal to the length of the temperature points. Next, the derivatives of the fractional hybridization with respect to the concentration and amplitude were found and computed for each temperature point. These values were saved in a 2 X n matrix. By taking the transpose of this matrix and multiplying it by the error vector, matrix operations were used to solve for the reduced row echelon form, which gives the offset that should be applied to the original guess. Repeating this through an iterative process, the best fit was found by reducing the error vector to values less than $10^{-18}$ for the concentration, and less than 0.01 for the amplitude. In order for the solution to converge on the proper fit, an appropriate initial guess is required. Using the initial conditions that were previously described, the solution was

generally found within 10 iterations. An example of fitting with this algorithm is shown in Figure 3.11. This demonstrates the different curves which were chosen as the solution reaches the best fit, as well as the fits match with the negative derivative of the fractional hybridization.

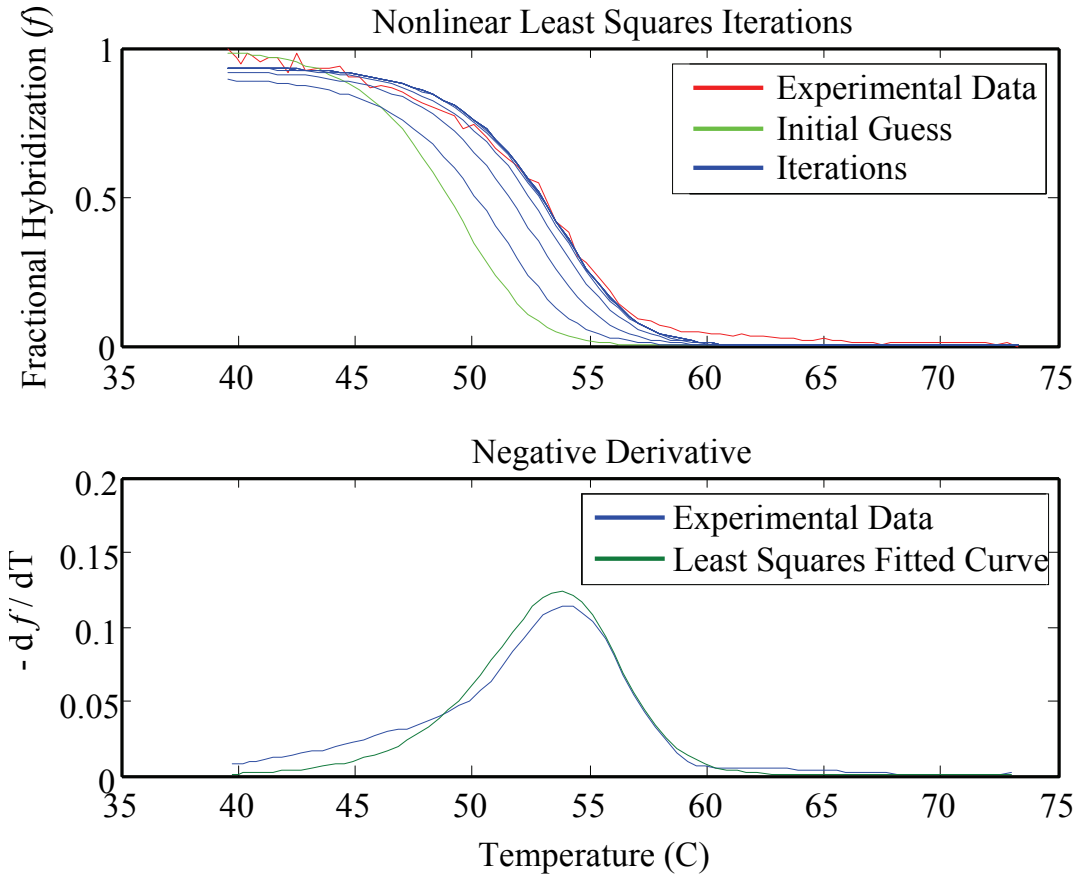



**Figure 3.11: Nonlinear least squares fitting of denaturation curve.** By altering the initial concentration and amplitude, the best fit of the experimental data is found after several iterations.

### *Fitting Data with Data*

Since the denaturation curve cannot be modeled precisely using a sigmoid function or other functions, we reasoned that it may be feasible to deconvolute the

experimental data by a sum of a known data set of individual oligonucleotides of different length, which can be determined experimentally. Using a linear least squares fitting method, the fractional hybridization curves for multiple length oligonucleotide probes were decomposed into the contribution from each individual probe length. This was performed for the three different probe lengths with the Alexa 488 dye selecting the experimental data from a single microbead in each of the four groups. There are curves for the 21-, 25-, and 29-mer microbeads, as well as the combination of all three. By running a standard least squares algorithm, the contribution that each of the individual probes has on the combined denaturation curve was determined. This algorithm chooses a coefficient to scale each individual curve in a manner that optimizes their fit when each curve is superimposed. Figure 3.12 depicts how the three component curves were scaled so that the sum of the individual curves create a close fit of the combined curve.

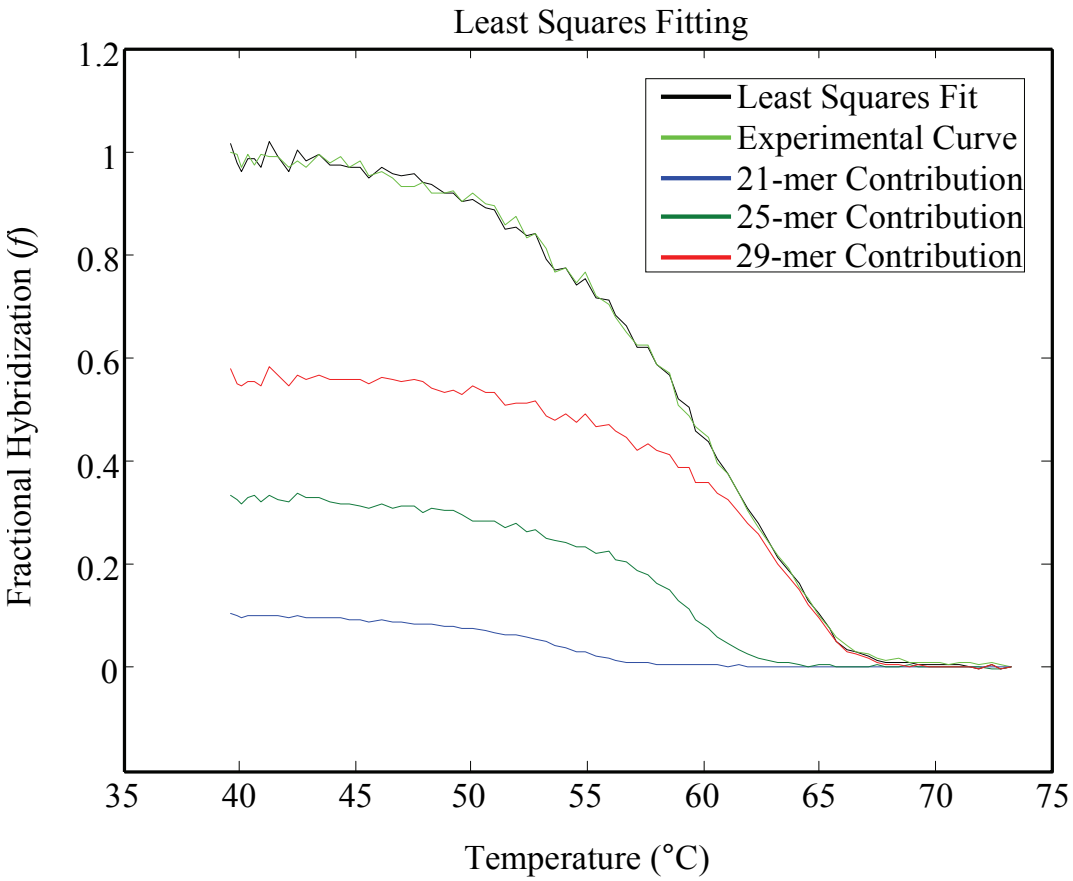



**Figure 3.12: Least squares fitting of data with data.** The denaturation curve of a bead containing 21-, 25-, and 29-mer probes was fit by superimposing the curves of the individual probes. A least squares algorithm was used to obtain the scaling factor for the curves to optimize the fit.

This strategy demonstrates some potential in the development of a base-calling algorithm for SBD. If the least squares curve fitting program that was previously described is modified so that the program incorporates data from a 20-mer microbead denaturation curve. In doing this, the algorithm can determine if the 20-mer probe contributes to the combined denaturation curve. This method proved to be successful in verifying that the 20-mer probe is not part of the combined curve. The results are

shown graphically in Figure 3.13. The least squares algorithm is capable of determining that the 20-mer probe was not part of the combined denaturation curve and actually assigns it a small, negative scaling factor. The curve fitting program recognized that the 20-mer microbead's denaturation curve falls before the combined curve and therefore it does not assign it a large contribution.

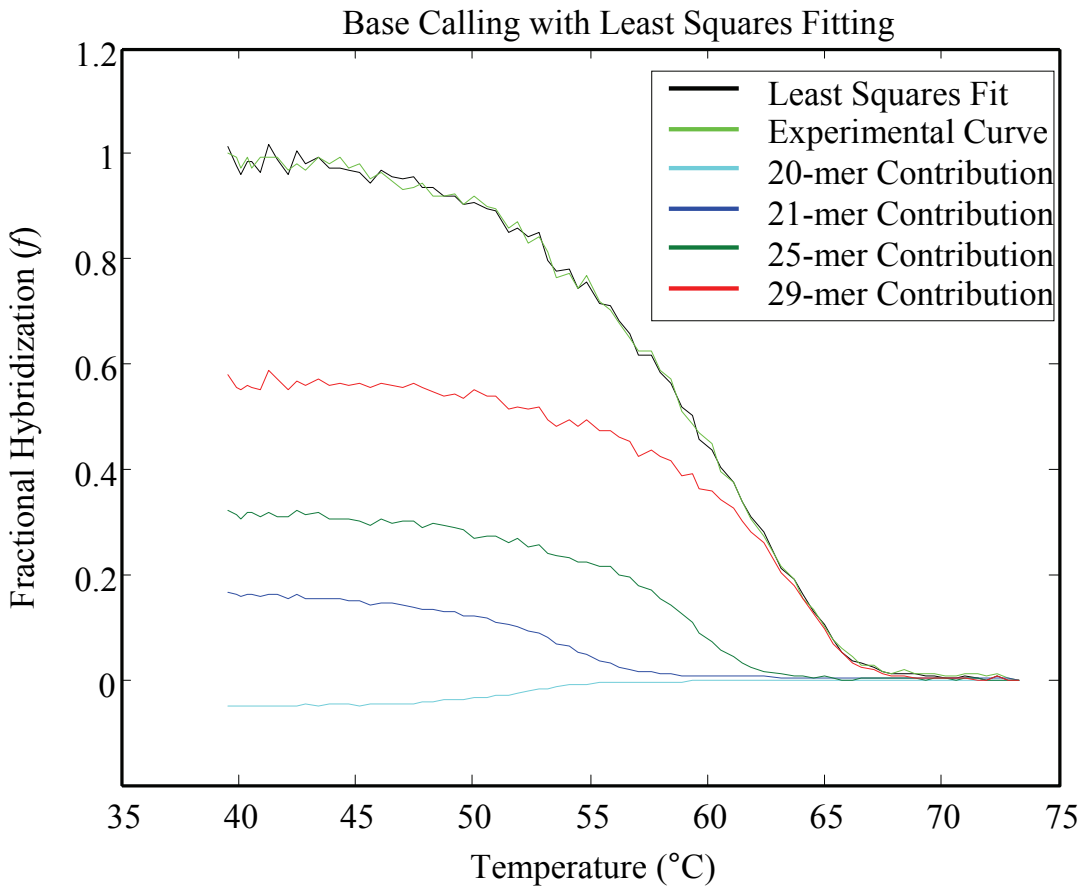



**Figure 3.13: Using least squares fitting to determine curve composition.** Because the fractional hybridization for the 20-mer probe falls before the experimental curve, the algorithm was able to determine that is does not contribute to the experimental curve.

This method is limited in that it only has the ability to reliably exclude a probe length that is less than the lengths that contribute to the experimental curve. For example, if the 24-mer probe replaced the 20-mer probe in the previous least squares algorithm, the program would incorporate the 24-mer probe into the composition in the process of finding the best fit. This severely limits the feasibility of using this process in the development of a base calling algorithm.

# Chapter 4: Implementation of SBD

## 4.1 Introduction

For a sequencing technology to be competitive, it needs to be easy to use and very robust. To accomplish these goals, the system should require minimal sample preparation and be capable of completing the sequencing process in a short time for minimal cost. Due to its simplicity in operation, SBD has the potential to be an inexpensive, rapid and low-cost sequencing method. Like many non-single molecule sequencing chemistries, each DNA fragment need to be amplified prior to sequencing. The genomic DNA library construction and DNA amplification process has turned out to be the rate-limiting step. The complexity and the cost associated with sample preparation process depend on the number of template copies needed to perform the sequencing chemistry.

In this chapter, Monte Carlo simulations are performed to determine how many copies of each DNA template are required to obtain an accurate denaturation curve. It is desirable to know the smallest number of probes necessary to accurately resolve the denaturation curve. This is highly relevant to the implementation of SBD as a sequencing method because this lower limit of probes needed constrains the amplification method used. If a large number of probes, in the range of one- to ten-thousand, are necessary for obtaining a clean denaturation curve, then magnetic microbeads must be used to hold the large number of bound DNA. For this strategy, emulsion PCR may be preformed in which microscopic water droplets in oil are used to compartmentalize the DNA amplification so that one microbead contains a clonal

template population (Diehl *et al.*, 2006). However, if only a hundred probes are needed, it might be possible to use rolling circle amplification to clone the templates (RCA). This method takes advantage of the ability for some polymerases to displace the strand of a circularized DNA to create multiple tandem repeats of the template sequence. RCA is a very efficient and simple method for linear amplification of circular DNA molecules (Huang, 2008, Lizardi *et al.*, 1998).

**4.2 Monte Carlo Simulation of SBD**

To obtain a prediction of the lower limit of probes needed for SBD, Monte Carlo simulations were performed to determine the stochastic denaturation process with various low copy number of the template molecules. The Monte Carlo approach is based on the random sampling of a cumulative distribution function (CDF). The CDF can be expressed as

$$F(x) = \int_{-\infty}^{x} f(t)dt ,$$

where F(x) represents the CDF and $f$(t) represents the probability distribution function. This equation demonstrates that the CDF is equal to the probability that a random variable, $x$, with a probability distribution function, $f$, will be found at a value less than $t$. Because the negative derivative of fractional hybridization with respect to temperature can be construed as the probability of a denaturation event happening at that temperature, the CDF can be interpreted as the negative of the equation for fractional hybridization.

For Monte Carlo simulations, a random number generator is used to replace the CDF equation, and then the independent variable is solved for. Utilizing this strategy,

an equation can be derived in which the fractional hybridization is replaced with a random number, and the temperature solved for is the melting temperature of that probe. This can be expressed as: $Tm=\Delta H^\circ/( R * \log((1 - r)^2 * C_0 / r ) + \Delta S^\circ )$, where r is the random variable, and Tm is the melting temperature. This approach takes the equation for fractional hybridization, which is based on the continuous denaturation of an infinite number of probes, and breaks it down into the digitized format in which the melting temperature of an individual probe can be considered.

To assemble the denaturation curve based on these individual melting temperatures, a binary system is used to assign each probe, a 1 if it is hybridized, and a 0 if denatured. By looping through the temperature range of interest, all the numbers are added and divided by the number of probes that are being sampled. This system allows the affect of sample size on the denaturation curve resolution of an individual microbead to be investigated. Figure 4.1 shows the simulation results for four different sample sizes of 10, 50, 100 and 1000. The $R^2$ values shown represent the average $R^2$ value of 1000 individual simulations.

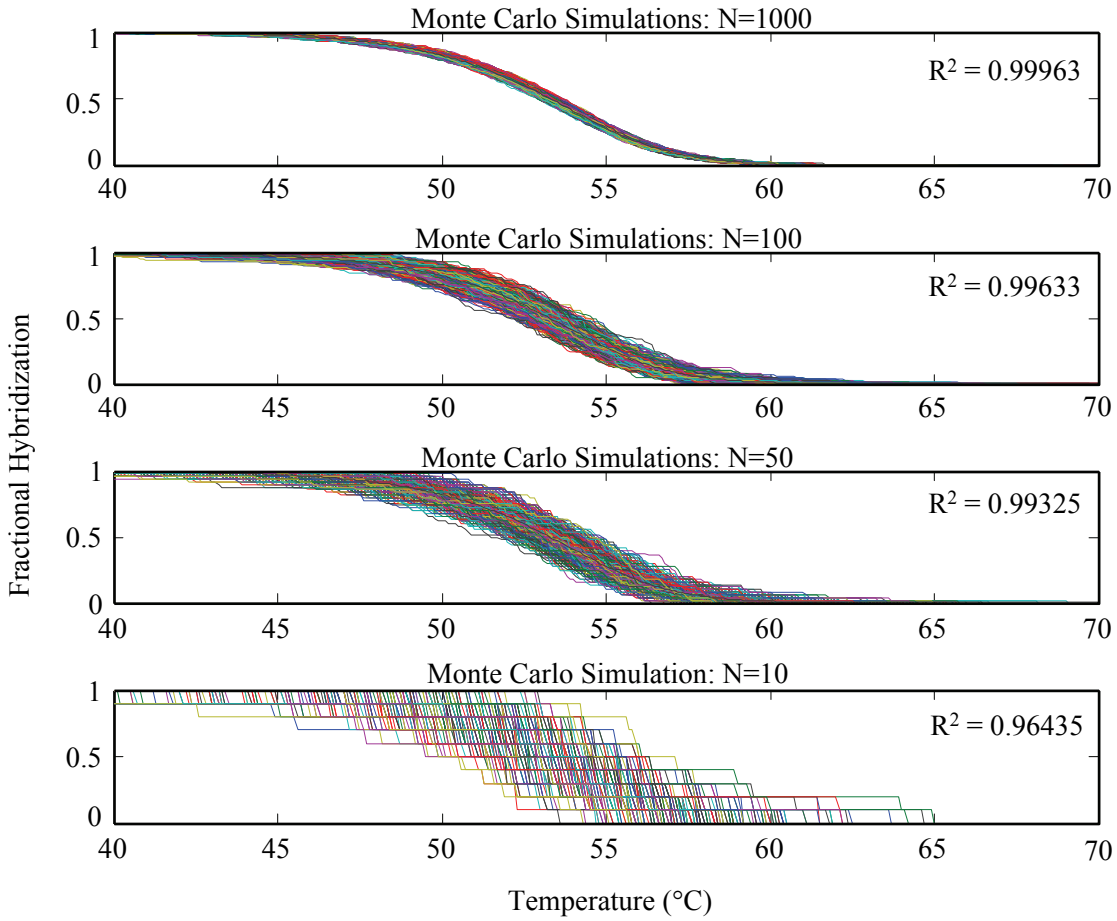



**Figure 4.1: Monte Carlo simulations.** The figure demonstrates the lost in resolution of the theoretical denaturation curve as fewer probes are used. This simulation suggests that the lower limit to be able to accurately determine the denaturation curve is about 50 probes.

### 4.3 Genome Sequencing by SBD Using Templates Prepared by RCA

From the Monte Carlo simulation results, it appears that a relatively accurate denaturation curve can be obtained using a small number of probes, perhaps as low as 50 copies. This has great implications, implying that the simple RCA method may be used to prepare genomic DNA libraries for sequencing by SBD. In fact, this method has been demonstrated as a viable approach for genome sequencing by the company

Complete Genomics, where circularized genomic DNA molecules are replicated by using *Phi*29 polymerase to create DNA nanoballs for sequencing (Drmanac *et al.*, 2009). Our simulations add to the appeal and feasibility of utilizing SBD in commercial genome sequencing applications.

# Chapter 5: Summary and Future Work

## 5.1 Summary

The work carried out so far in this thesis work has demonstrated the great potential that denaturation reactions have for determining a DNA sequence. An integrated microfluidic device has been designed and fabricated to achieve uniformly heated fluidic channels, as well as a steady thermal expansion throughout temperature ramping. The ability to rapidly change the temperature of the device in a controllable manner, while successfully tracking the focal plane by the microscope, has been proven. We have shown that different length probes display clearly distinguishable denaturation profiles, and that these individual profiles can superimpose to create a signature denaturation curve, representing a specific set of different length oligonucleotides. We displayed the capability to resolve the individual components of a denaturation profile formed by up to four different probe lengths, as well as the ability to closely fit the experimental data to the theoretical equation. We have also described how it might be possible to resolve the denaturation curve from only a small number of probes. Based on simulation data, we demonstrated how as few as fifty copies of a probe could be sufficient to accurately resolve the denaturation curve. These results give the technology commercial viability as a fast and cheap sequencing method for short read length applications.

## 5.2 Future work

Due to the simplicity of SBD, it has great potential for genome sequencing applications. Because the chemistry only requires a single Sanger's dideoxy

sequencing reaction, SBD requires much less reagents than other sequencing strategies that require cyclic loading and washing (Chen *et al.*, 2010). Additionally, the sequencing reaction only requires twenty minutes, which is much more rapid than other current methods. A cost analysis for sequencing a human genome using SBD with 10 fold coverage was estimated to be about $1000 (Chen, 2008). These advantages of SBD give it great potential as a commercial sequencing strategy.

Before this is possible however, several further principles must be demonstrated. The ability to obtain accurate denaturation curves of probes hybridized to RCA product must be shown. This can be done by performing an RCA reaction on a substrate that is capable of spatially separating single RCA molecules so that the fluorescence from each individual strand can be resolved. Additionally, the lower limit of probes needed to precisely resolve the denaturation curve must be experimentally determined.

If the denaturation curve is resolvable from probes hybridized on the tandem repeats generated by the RCA reaction, the number of probes required will determine the length of the RCA product needed. The final step would be to perform the reaction as it would be in a commercial application. Sanger's dideoxy reaction would be performed on the RCA product using fluorescently labeled dideoxy nucleotides, or other methods to image the probes, such as labeled primers. Once the probes have been synthesized on the tandem repeats of the RCA molecule, the standard denaturation experiment will be performed to obtain the denaturation curve profiles for all of the template DNA molecules. Methods for developing an accurate base-

calling algorithm must also be established in order to determine the sequence of an unknown template based on the denaturation curve profile.

# References


Barbee, K.D. and Huang, X. (2008) Magnetic assembly of high-density DNA arrays for genomic analyses, *Anal Chem*, **80**(6), 2149-54.

Bentley, D.R., Balasubramanian, S., Swerdlow, H.P., Smith, G.P., Milton, J., Brown, C.G.*, et al.* (2008) Accurate whole human genome sequencing using reversible terminator chemistry, *Nature*, **456**(7218), 53-9.

Chen, Y.J. (2008) *DNA Sequencing By Denaturation,* University of California, San Diego.

Chen, Y.J. and Huang, X. (2009) DNA sequencing by denaturation: principle and thermodynamic simulations, *Anal Biochem*, **384**(1), 170-9.

Chen, Y.J., Roller, E.E., and Huang, X. (2010) DNA sequencing by denaturation: experimental proof of concept with an integrated fluidic device, *Lab Chip*, **10**(9), 1153-9.

Diehl, F., Li, M., He, Y., Kinzler, K.W., Vogelstein, B., and Dressman, D. (2006) BEAMing: single-molecule PCR on microparticles in water-in-oil emulsions, *Nat Methods*, **3**(7), 551-9.

Drmanac, R., Sparks, A.B., Callow, M.J., Halpern, A.L., Burns, N.L., Kermani, B.G.*, et al.* (2009) Human genome sequencing using unchained base reads on self-assembling DNA nanoarrays, *Science*, **327**(5961), 78-81.

Fedorov, R., Witte, G., Urbanke, C., Manstein, D.J., and Curth, U. (2006) 3D structure of Thermus aquaticus single-stranded DNA-binding protein gives insight into the functioning of SSB proteins, *Nucleic Acids Res*, **34**(22), 6708-17.

Huang, X., *Genomes, Genome Technologies and Medicine*, in *An Introductory Text to Bioengineering*, Chien, S., Chen, P.C.Y., and Fung, Y.C., Editors. 2008, World Scientific Publishing Co. Pte. Ltd: Singapore. Chapter 27, pp 473-485.

Lizardi, P.M., Huang, X., Zhu, Z., Bray-Ward, P., Thomas, D.C., and Ward, D.C. (1998) Mutation detection and single-molecule counting using isothermal rolling-circle amplification, *Nat Genet*, **19**(3), 225-32.

Margulies, M., Egholm, M., Altman, W.E., Attiya, S., Bader, J.S., Bemben, L.A.*, et al.* (2005) Genome sequencing in microfabricated high-density picolitre reactors, *Nature*, **437**(7057), 376-80.

Ronaghi, M., Uhlen, M., and Nyren, P. (1998) A sequencing method based on real-time pyrophosphate, *Science*, **281**(5375), 363, 365.

SantaLucia, J., Jr. (1998) A unified view of polymer, dumbbell, and oligonucleotide DNA nearest-neighbor thermodynamics, *Proc Natl Acad Sci U S A*, **95**(4), 1460-5.

Shendure, J. and Ji, H. (2008) Next-generation DNA sequencing, *Nat Biotechnol*, **26**(10), 1135-45.

Valouev, A., Ichikawa, J., Tonthat, T., Stuart, J., Ranade, S., Peckham, H.*, et al.* (2008) A high-resolution, nucleosome position map of C. elegans reveals a lack of universal sequence-dictated positioning, *Genome Res*, **18**(7), 1051-63.